

A global analysis of viral mRNA biogenesis during infection with Adenovirus type 5

DISSERTATION

with the aim of achieving a doctoral degree at the
Faculty of Mathematics, Informatics and Natural Sciences,
Department of Biology,
University of Hamburg

submitted

by

Margarita Valdés Alemán

May 2018 in Hamburg

Day of disputation: 10.10.2018

Disputation committee:

Prof. Dr. T. Dobner (dissertation reviewer)

Prof. Dr. N. Fischer (dissertation reviewer)

Prof. Dr. Z. Ignatova (disputation examiner)

Prof. Dr. J. Kehr (disputation chair)

Clayton Southerly

4000 Wilson Boulevard, #809
Arlington, VA 22203
United States of America
(703) 969-7574
Clayton.southerly@gmail.com

24th February 2018

The Faculty of Mathematics, Informatics, and Natural Sciences

Department of Biology
Universität Hamburg, MIN-Dekanat
Rothenbaumchaussee 19
20148 Hamburg

To whom it may concern,

This letter serves as verification that the English language used in the PhD thesis entitled "A global analysis of viral mRNA biogenesis during infection with Adenovirus type 5" by Margarita Valdés Alemán reads fluently and is well-written. I give my support that the thesis is correctly articulated in English to general standards.

Regards,

A handwritten signature in black ink, appearing to read 'Clayton S.', with a long horizontal flourish extending to the right.

Clayton Southerly
Product Development Analyst
Bulletin Intelligence
B.A. International Relations
The College of William & Mary

Statutory Declaration

I declare that this thesis is a result of my personal work and that no other than the indicated aids have been used for its completion. Furthermore, I assure that all quotations and statements that have been inferred literally or in a general manner from published or unpublished writings are marked as such. Beyond this I assure that the work has not been used, neither completely nor in parts, to achieve an academic grading or is being published elsewhere.

21.02.2018

Cernavaca, Mexico.

Signature

A handwritten signature in blue ink, consisting of a large, stylized 'S' shape with a vertical line through it, and a small 'S' to the right.

Abstract

During the late phase of infection with adenovirus, cellular protein synthesis is shut off, due to a translational block of host cell mRNAs. It has been documented that cellular mRNAs fail to accumulate in the cytoplasm despite continued nuclear synthesis and processing. In contrast, the viral late mRNAs are selectively exported to the cytoplasm via the cellular export receptor TAP/NXF1. Interestingly, the activity of an E3-ubiquitin ligase complex composed of the viral E1B-55K and E4 Orf6 proteins and cellular factors is required, indicating that ubiquitin-dependent proteasomal degradation of one or more proteins could contribute directly or indirectly to the regulation of mRNA export during the late phase of infection. So far, the identity of such substrates is not known, nor has the mechanism by which viral mRNA species are distinguished from their cellular counterparts for export to the cytoplasm been identified; although several different models have been proposed. In these study, we search for possible degradation candidates for the viral-formed ubiquitin ligase. Also, attempting to identify additional parameters that control the differential export of viral and cellular transcripts, we performed global transcriptome analyses (RNA-Seq) to monitor changes in the cytoplasmic accumulation of RNAs expressed from cellular and viral genes as a function of time after infection of A549 cells.

In this study, none of the tested cellular proteins were found to be degraded during infection. Thus, the role of the viral E1B/Orf6/E3 ubiquitin ligase in the selective export of viral late mRNAs remains unanswered. During this evaluation, hnRNP M was found to conjugated by SUMO-1 and SUMO-2. Higher forms of SUMO-2-modified hnRNP M are formed at late time points of infection, depending on the presence of the viral E1B-55K protein. Our results suggest that phosphorylated E1B-55K enhances production of this higher forms of SUMO-2 conjugated hnRNP M. It was also found that hnRNP M accumulates around the viral RCs at late time points of infection. In our transcriptome analysis, we observed an unexpectedly large fraction of cellular transcripts was refractory to the export block and thus escaped the virus-induced export inhibition. Importantly, export ratios of this cluster were increased by a factor of two or greater during the late phase, when compared to the early phase. Gene functional classification analysis shows that mRNA species of this cluster are linked to RNA metabolism and translation. A time comparison between the total amount of cellular vs. viral mRNAs shows that even though cellular mRNAs are equally abundant as in early times, the viral mRNA population is significantly more abundant than cellular transcripts at late times of infection. In sum, these findings show that the impact of HAdV-5 infection on nucleo-cytoplasmic RNA transport is greater than appreciated previously, and suggest that cellular export blockage is

not absolute as we found a subset of host cell mRNA transcripts that are preferentially exported to the cytoplasm in the late phase of infection. Also, the high amounts of viral mRNAs that are detected at late times suggest that previous findings, where mostly viral mRNAs were found in the cytoplasm, could be more likely explained by the extremely successful expression of the viral mRNAs than by the blockage of cellular mRNAs export.

Zusammenfassung

Es wird weitgehend angenommen, dass die Synthese zellulärer Proteine in der späten Phase einer adenoviralen Infektion inhibiert wird. Die Ursache dieser Inhibierung ist begründet durch die ausbleibende Akkumulierung zellulärer mRNAs im Zytoplasma, die während der Infektion unverändert im Zellkern kontinuierlich synthetisiert und prozessiert wird. Im Gegensatz dazu werden die späten viralen mRNAs selektiv über den zellulären Export-Rezeptor TAP/NXF1 ins Zytoplasma transportiert. Erforderlich dafür ist die Aktivität eines E3-Ubiquitin Ligase Komplexes, bestehend aus den viralen Proteinen E1B-55K und E4 Orf6 Proteinen und weiteren zellulären Faktoren. Dieser Zusammenhang deutet darauf hin, dass der Ubiquitin-abhängige proteasomale Abbau von einem oder mehreren Proteinen, die direkt oder indirekt an der Regulation des mRNA Exportes während der späten Phase der Infektion beteiligt ist, erfolgen könnte. In den vergangenen Jahren wurde weder dieser Faktor noch der Mechanismus für den selektiven Export viraler mRNAs identifiziert. In der vorliegenden Arbeit wurden putative zelluläre Zielgene, die durch den E3-Ubiquitin Ligase Komplex abgebaut werden, untersucht. Des Weiteren wurde eine globale Transkriptom-Analyse (RNA-Seq) in A549-Zellen durchgeführt, mit Hilfe derer die Veränderungen der zytoplasmatischen Akkumulierung der exprimierten RNAs von zellulären und viralen Genen im zeitlichen Verlauf der Infektion untersucht wurden. Damit sollen weitere bisher unbekannte Faktoren, die den differentiellen Export viraler und zellulärer Transkripte induzieren, identifiziert werden.

Zunächst konnte keiner der putativen zellulären Proteine als Substrat des E3-Ubiquitin Ligase Komplexes identifiziert werden, womit die Rolle der viralen E1B/Orf6/E3 Ubiquitin Ligase im selektiven Export der viralen späten mRNAs weiterhin unbekannt bleibt. Im Rahmen dieser Arbeit wurde jedoch gezeigt, dass das zelluläre Protein hnRNP M sowohl SUMO-1 als auch SUMO-2 konjugiert wird, was zum späten Zeitpunkt einer adenoviralen Infektion in Abhängigkeit des viralen E1B-55K Proteins in zunehmend längeren SUMO-2-Ketten mündet. Zudem wurde gezeigt, dass phosphoryliertes E1B-55K die Ausbildung dieser höher SUMO-2-konjugierten hnRNP M-Formen fördert. Außerdem akkumuliert hnRNP M zum späten Zeitpunkt der Infektion im Bereich der viralen Replikationszentren (RCs). Die Transkriptom-Analyse identifizierte eine große Fraktion zellulärer Transkripte im Zytoplasma, die nicht wie zuvor angenommen durch die virale Infektion im Nukleus zurückgehalten wurde. Zudem ist das Export Verhältnis dieser Fraktion in der späten Phase im Vergleich zur frühen Phase um einen Faktor von zwei oder mehr erhöht. Funktionelle Genklassifikationsanalysen zeigen, dass die Produkte der mRNAs dieser Gruppe am RNA-Metabolismus und an der RNA-Translation beteiligt sind. Ein zeitlicher Vergleich der Gesamtmengen an zellulärer und viraler mRNAs zeigt

bei gleichbleibender Menge an zellulärer mRNA Transkripten einen signifikanten Anstieg der viralen mRNA-Transkripte in der späten Phase der Infektion.

Die große Menge viraler mRNAs zum späten Zeitpunkt der Infektion im Zytoplasma im Vergleich zur gleichbleibender Menge zellulärer mRNAs könnte zu Fehlinterpretationen bisheriger Forschungsergebnisse geführt haben, bei denen ein Ausschalten des Transports zellulärer mRNA bei viraler Infektion postuliert wurde.

Zusammengefasst zeigt diese Arbeit, dass der RNA-Transport während einer HAdV-5 Infektion einem umfassenderen Mechanismus als zunächst angenommen unterliegt. Durch Transkriptomanalysen konnte gezeigt werden, dass es zu keiner absoluten Blockierung des zellulären mRNA-Exportes kommt, da eine Gruppe an Wirtszell-mRNA-Transkripten gefunden wurde, die bevorzugt in der späten Phase der Infektion ins Zytoplasma exportiert wird.

Table of Contents

1	Introduction.....	1
1.1	Adenoviruses.....	1
1.1.1	Generalities.....	1
1.1.2	Classification.....	2
1.1.3	Virion Structure.....	2
1.1.4	Genome Organization.....	4
1.2	Single-cell reproductive cycle.....	6
1.2.1	Early Phase: Interaction with host cell and transcriptional activation.....	6
1.2.2	Late Phase: Viral genome replication and selective viral late gene expression.....	9
1.3	Early genes.....	13
1.3.1	Transcriptional activation by E1A.....	13
1.4	Inhibition of host cell defenses.....	15
1.4.1	Apoptosis inhibition.....	15
1.4.2	Inhibition of host cell response to dsDNA breakage.....	16
1.4.3	Inhibition of interferon-induced antiviral response.....	17
1.5	Late genes.....	18
1.5.1	Transcriptional activation of viral late genes.....	18
1.5.2	Post-transcriptional processing.....	19
1.5.3	Selective expression of viral late mRNAs.....	21
1.5.3.1	Selective Export of viral late mRNAs.....	22
1.5.3.2	Selective Translation of viral late mRNAs.....	24
1.6	Adenovirus Replication Centers.....	25
1.7	mRNA processing and mRNA export of viral late mRNAs are linked to Replication Centers 27	
1.8	SUMOylation during HAdV infection.....	29
2	Project Aims.....	32
3	Materials and Methods.....	33
3.1	Materials.....	33
3.1.1	Cells and Viruses.....	33
3.1.2	Antibodies.....	34
3.1.3	Standards and markers.....	35
3.1.4	Commercial systems.....	36
3.1.5	Chemicals, enzymes, reagents and equipment.....	36
3.1.6	Software and databases.....	36
3.2	Methods.....	37

3.2.1	Tissue culture	37
3.2.2	Infection with adenovirus	39
3.2.3	SUMO pull-down assay	40
3.2.4	Cycloheximide assay	42
3.2.5	Preparation of protein cell lysates.....	42
3.2.6	SDS polyacrylamide gel electrophoresis (SDS-PAGE)	43
3.2.7	Western Blot	44
3.2.8	Immunofluorescence.....	45
3.2.9	Cytoplasmic and nucleoplasmic fractionation for RNA extraction.....	46
3.2.10	Next Generation Sequencing (NGS)	47
3.2.11	NGS data analysis	47
4	Results.....	49
4.1	Analysis of cellular proteins involved in mRNA processing and export pathways during infection	49
4.1.1	Time-course analysis of cellular proteins involved in mRNA processing and mRNA export	49
4.1.2	Stability analysis of cellular proteins involved in mRNA processing and mRNA export	52
4.1.3	hnRNP M SUMOylation analysis.....	54
4.1.4	Analysis of E1B-55K dependence for changes in hnRNP M SUMOylation patterns during infection.....	56
4.1.5	Analysis of E1B-55K phosphorylation status dependence for changes in hnRNP M SUMOylation patterns during infection	57
4.1.6	hnRNP M intracellular distribution analysis.....	58
4.2	Analysis of viral and host mRNA accumulation and export rates during infection .	60
4.2.1	Analysis of viral mRNA accumulation during infection	61
4.2.1.1	Time-course alignment of sequenced reads to the viral transcript annotations .	61
4.2.1.2	Quantification of viral transcripts at different time points of infections	64
4.2.1.2.1	Analysis of viral early mRNAs.....	67
4.2.1.2.2	Analysis of viral late mRNAs	72
4.2.2	Analysis of viral mRNA cytoplasmic to nuclear rates during infection.....	76
4.2.3	Analysis of cellular mRNA cytoplasmic to nuclear rates during infection	79
4.2.3.1	Functional analysis: Networks and Classification	81
4.2.4	Comparison between viral and cellular expression during the late phase of infection	84
5	Discussion.....	90

5.1	Contribution of the E1B/Orf6/E3 Ubiquitin ligase in degradation of proteins involved in mRNA biogenesis.	90
5.2	hnRNP M is hyper-SUMOylated at late times of infection with HAdV-5	91
5.2.1	Hyper-SUMOylation of hnRNP M depends on the presence of E1B-55K	92
5.3	hnRNP M is localized towards the viral RCs during infection with HAdV-5	93
5.4	Global analysis of viral and cellular mRNAs during infection with HAdV-5	94
5.4.1	Time-course analysis of viral mRNAs	95
5.4.1.1	Viral early mRNAs	95
5.4.1.2	Viral late mRNAs	99
5.4.2	Effects of HAdV-5 infection on cytoplasmic mRNA accumulation	100
5.4.2.1	Viral mRNAs	100
5.4.2.2	Cellular mRNAs	101
6	Literature	104
7	Tables	120
8	Publications	128
	Scientific Meetings	128
	Scientific Workshops	128
9	Acknowledgements	129

aa	amino acids
Ad pol	adenoviral polymerase
ATM	Serine/threonine kinase
ATR	Serine/threonine protein kinase
BCL-2	B-cell lymphoma 2
bp	base pair
BSA	bovine serum albumin
CAR	Coxsackie-Adenovirus Receptor
CBP	CREB-binding protein
CDK	Cyclin-dependent kinase
CDS	coding sequence
CR	conserved regions
CRM1	Chromosomal Maintenance 1 or Exportin 1
cyto	cytoplasmic
DAPI	4',6'-diamidino-2-phenylindole
Daxx	Death domain-associated protein
dCMP	Deoxycytidine monophosphate
DMEM	Dulbecco's Modified Eagles Medium
DMSO	dimethyl sulfoxide
DNA	Deoxyribonucleic Acid
DSBR	double-strand break repair
dsDNA	double-stranded DNA
E1B	E1B-55K
eIF2a	Eukaryotic translation initiation factor 2 subunit 1
eIF4E	Eukaryotic translation initiation factor 4E
eIF4F	Eukaryotic initiation factor 4F
EJC	exon junction complex
FCS	fetal calf serum
ffu	foci formation units
HAdV	human adenovirus
HAdV-2	human adenovirus type 2
HAdV-5	human adenovirus type 5
hnRNA	heterogeneous nuclear RNA
hpi	hours post-infection
INF	Interferon
ITR	inverted terminal repeat
kbp	kilo base pair
kDa	kilo Dalton
MAPK	Mitogen-activated protein kinase
MDM2	Mouse double minute 2 homolog
MED23	Mediator complex subunit 23
ML	Major Late
MLP	Major Late promoter
MOI	multiplicity of infection

MRN	Mre11/ Rad50/ Nbs1 complex
mRNA	messenger RNA
mRNP	messenger ribonucleoprotein
MTOC	Microtubule-organization center
NaN	not a number/ unrepresentable number
Nbs1	Nijmegen breakage syndrome 1 (nibrin)
ND10	Nuclear Domain 10
NES	nuclear export sequences
NHEJ	non-homologus end joining
NLS	nuclear localization sequences
NPC	Nuclear Pore Complex
nt	nucleotides
nuc	nucleoplasmic
Orf3	E4 Orf3
Orf6	E4 Orf6
PBS	Phosphate Buffered Saline
PKR	Protein kinase RNA-activated
PML	Promyelocytic leukemia protein
poly(A)	polyadenylation
pRB	Retinoblastoma protein
PRZ	peripheral replication zone
RCs	Replication Centers
RNA	Ribonucleic Acid
RNA-seq	RNA sequencing
RNApol	RNA polymerase
RPKM	Reads Per Kilobase of transcript per Million mapped reads
rRNA	ribosomal RNA
RT	room temperature
S phase	synthesis phase
snRNPs	small nuclear ribonucleoproteins
ssDNA	single-stranded DNA
SUMO	Small Ubiquitin-like Modifier
TBP	TATA-binding protein
TNF α	Tumor Necrosis Factor α
TP	Terminal Protein
TREX	Transcription-Export Complex
Ub	Ubiquitin
VA	Virus-Associated
WB	Western Blot
wt	wild type

1 Introduction

1.1 Adenoviruses

1.1.1 Generalities

Adenoviruses were initially isolated in 1953, from adenoids tissue, in patients with acute respiratory infections (Rowe *et al.*, 1953; Hilleman and Werner, 1954). They are an important cause of hospitalization in children (Sun *et al.*, 2014), however, they are only responsible for a small percentage of acute respiratory morbidity. Infection with adenoviruses in immunocompetent patients usually only causes mild acute infections which are self-limiting. Outbreaks have been reported in places where there is a high concentration of people, such as in military training centers (Kunz and Ottolini, 2010). Human adenovirus (HAdV) can also cause other diseases such as conjunctivitis and infantile gastroenteritis (Albert, 1986; Aoki *et al.*, 2011). Particularly, they are dangerous for neonates and immunosuppressed people, such as patients with AIDS or hematopoietic stem cell transplantation recipients (HSCT), in which they can cause fulminant fatal pneumonia, hepatitis and encephalitis (Lion *et al.*, 2010; Tebruegge and Curtis, 2010). Infection of HAdV is commonly diagnosed by polymerase chain reaction (PCR) assays (Lion *et al.*, 2003; Leruez-Ville *et al.*, 2004; Kehl and Kumar, 2009).

Adenoviruses were the first known human viruses to promote oncogenesis when studies showed that they are capable of inducing tumors in newborn hamsters (Trentin, Yabe and Taylor, 1962). This discovery increased interest in HAdV as an experimental system for studying oncogenesis and therefore, several viral mutants have been developed allowing researchers to decipher their oncogene functions (Speiseder, Nevels and Dobner, 2014). However, no epidemiologic association of human cancer caused by adenovirus infection has been reported (Trentin *et al.*, 1962; Mackey, Rigden and Green, 1976; Endter and Dobner, 2004). Further, HAdVs are used as a model for studying processes such as cellular transcription and translation, DNA replication, cell-cycle control and cellular growth regulation. Messenger RNA (mRNA) splicing was first discovered in HAdV (Berget, Moore and Sharp, 1977; Chow *et al.*, 1977; Joseph R. Nevins, 1979). Currently, HAdV are widely studied as a part of increased interest in their potential as vectors for gene therapy, vaccination, and cancer gene therapy (Wold and Toth, 2013).

1.1.2 Classification

Adenoviruses belong to the *Adenoviridae* family, which divides into five genera: *Mastadenovirus*, isolated from mammals; *Aviadenovirus*, isolated from birds; *Atadenovirus*, isolated from mammals, reptiles, birds and a marsupial (and are named after their unusual high content of AT in their genomes); *Siadenovirus*, are isolated from reptiles and birds; and *Ichtadenovirus*, isolated from fish (Davison, Benko and Harrach, 2003).

They are classified into seven species (A-G) and there are 57 different known types of human adenoviruses (HAdV) to date. Initially, they were classified to “serotypes” because they were categorized by their degree of neutralization against serums and their hemagglutination (Rosen, 1960) as well as their oncogenicity in rodents and transformation of cultured primary cells. However, this terminology has changed into “types”, since they are now classified by their amino acid sequence (mainly of the major capsid protein hexon) (Allard, Albinsson and Wadell, 2001; Robinson *et al.*, 2011) (Figure 1).

Species	Hemagglutination Groups	Types	Oncogenic potential		% GC	Associated disease
			Tumors in animals	Transformation in cell culture		
HAdV-A	IV (little or none)	12, 18, 31	High	Positive	46–47	Cryptic enteric infection
HAdV-B	I (complete for monkey erythrocytes)	3, 7, 11, 14, 16, 21, 34, 35, 50	Moderate	Positive	49–51	Conjunctivitis Acute respiratory disease Hemorrhagic cystitis Central nervous system
HAdV-C	II (partial for rat erythrocytes)	1, 2, 5, 6	Low or none	Positive	55	Endemic infection Respiratory symptoms
HAdV-D	III (complete for rat erythrocytes)	8, 9, 10, 13, 15, 17, 19, 20, 22–30, 32, 33, 36–39, 42–49, 51, 53, 54	Low or none (mammary tumors)	Positive	55–57	Keratoconjunctivitis in immunocompromised and AIDS patients
HAdV-E	III	4	Low or none	Positive	58	Conjunctivitis Acute respiratory disease
HAdV-F	III	40, 41	Unknown	Negative	51	Infantile diarrhea
HAdV-G	Unknown	52	Unknown	Unknown	55	Gastroenteritis

Figure 1. Classification of human adenoviruses. Human adenoviruses grouped according to their hemagglutination group and their type, showing their degree of oncogenic potential, their GC content percentage and to what diseases they associate (image from Fields Virology 6th Edition, Chapter 55 *Adenoviridae*).

1.1.3 Virion Structure

Adenoviruses are particles with a diameter of 70-100 nm. They have no lipidic envelope and their capsid has an icosahedral shape with fibers along the 12 vertices. The capsid surrounds

the core of the virion, where the viral genome is associated and folded with viral proteins. The capsid is formed by nine different structural proteins, mainly hexon and penton. There are 240 capsomers of hexon, each one is a trimer of protein II. There are 12 hexon trimers at each triangular facet of the icosahedron. Penton capsomers are at each of the icosahedral vertices. They are formed by a penton base and a fiber that projects from the surface of each vertex (Figure 2, panel A). The base of the penton capsomer is formed by a pentamer of protein III. Each fiber shaft terminates in a distal knob and it is composed of a trimer of protein IV. Fiber knobs interact with cellular receptor proteins and the length of the fiber shaft varies between HAdV types. Polypeptides IIIa, VIII and IX are stabilizing the interactions between the hexon capsomers and allow different non-equivalent arrangements in the capsid (Nemerow *et al.*, 2009; Liu *et al.*, 2010).

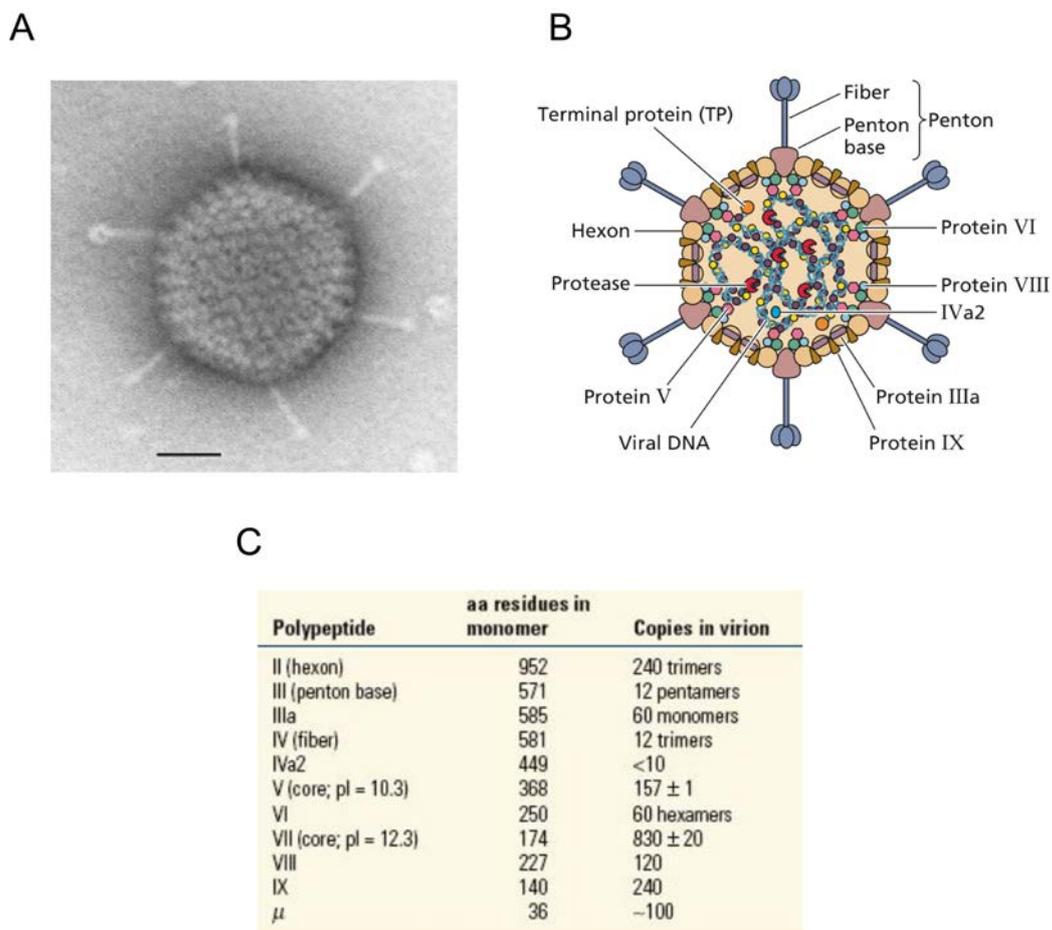


Figure 2. Adenovirus structure and protein composition. A) Virion structure. Negatively stained human adenovirus type 5 particle visualized by electron microscopy. B) Representation of virion structure and viral protein interactions within the capsid and core (panels A and B, from Principles of Virology, 4th edition). C) Adenovirus structural polypeptides and their copies per virion (from Fields Virology, 6th edition, chapter 55 *Adenoviridae*, adapted from Liu *et al.*, 2010).

The core of the virion is constituted by the viral genome and seven viral proteins. The most abundant is protein VII, with more than 800 copies per virion (van Oostrum and Burnett, 1985) (Figure 2, panel C). The viral DNA is in contact with polypeptides V, VII and μ , which are arginine-rich proteins and compact the genome in the core (Russell, Laver and Sanderson, 1968; Chatterjee, Vayda and Flint, 1986; Anderson, Young and Flint, 1989). The terminal protein (TP) is covalently attached to both 5'-ends of the viral DNA genome by a phosphodiester link between its serine 562 and the 5'-hydroxyl of the terminal deoxycytidine (Smart and Stillman, 1982). Therefore, there are only two TP per core. Protein VI associates with hexon trimers from the inner surface of the capsid and with protein V they tether the highly-ordered capsid with the less-ordered DNA-protein core (Saban *et al.*, 2006; Pérez-Berná *et al.*, 2009) (Figure 2, panel B). It has been observed that DNA is organized in spherical structural elements, called "adenosomes", inside the core; it is proposed that there are 12 adenosomes each directed towards one of the vertices of the icosahedron (Newcomb, Boring and Brown, 1984). Finally, the protease is present in ~10 copies inside the core. It is important during assembly and maturation of the virus particles as it cleaves precursors of several virion proteins and it also plays a role during disassembly and escape from the endosomes during viral entrance (Greber *et al.*, 1996).

1.1.4 Genome Organization

The HAdV-5 genome is a linear double-stranded DNA (dsDNA) chain that measures ~36 kilo base pairs (kbp). It has two identical inverted terminal repeat (ITR) sequences in each end, which function as DNA replication origins (Wides *et al.*, 1987). As described before, a TP is covalently linked at each 5'-phosphate end of each strand. Near the left terminal repeat, it contains several repeats of cis-acting packaging sequences, which are required for proper packaging of viral DNA into virion particles (Hearing *et al.*, 1987). Both DNA strands are transcribed and there is an overlap between coding regions.

The viral genome is organized into transcriptional units, which are transcribed by RNA polymerase (RNAPol) II. There are five early transcriptional units (E1A, E1B, E2, E3 and E4), four intermediate units, which are transcribed at the onset of viral DNA replication (IX, IVa2, L4 intermediate, and E2 late) and one late transcriptional unit, under control of the Major Late (ML) promoter that is divided into five mRNA families (L1-L5). Viral late mRNA families are produced by alternative splicing and differential polyadenylation (poly(A)) site selection. All viral late mRNAs share three small introns at their 5'-end, which together are called the "tripartite leader" (described in more detail in section Late gene). Also, there are two genes

transcribed by RNAPol III that code for two small RNAs named Virus Associated (VA) RNAs: VA RNA-1 and VA RNA-2 (Roberts, O'Neill and Yen, 1984; Flint, 1999) (Figure 3).

RNAPol II transcribed units (except for IVa2 and IX) code for more than one product by using alternative splicing (Berget *et al.*, 1977; Chow *et al.*, 1977). Major late, E2 and E3 transcription units also use alternative poly(A) sites. These post-transcriptional regulations of the viral mRNAs generate more than 40 different transcripts. Consequently, proteins coded in a same transcriptional unit can share sequences and domains that participate in common functions. The E1A unit codes for two proteins that activate transcription and induce entry into S phase of the cell cycle; E1B codes for two proteins that block apoptosis activation in the cell; E2, codes for three proteins that participate in viral genome replication; the E3 unit codes for proteins that regulate the host response to infection; and the viral late mRNA families code for structural proteins and proteins involved in capsid assembly. The E4 transcriptional unit is the only one that codes for proteins with widely different functions. E4 products regulate processes like viral gene transcription, mRNA transport, viral genome replication, and apoptosis of the host cell.

Organization into transcriptional units allows to control, in a single transcriptional element, the expression of multiple genes, which are required for a specific function, such as transcriptional activation (E1A) or replication of the viral genome (E2). It is thought that the position of each transcriptional unit inside the genome might give them a temporal regulation over their expression (Sambrook *et al.*, 1980). Thereby, E1 and E4 are the first transcribed units in the early phase, as they are located at the ends of the genome (Nevins *et al.*, 1979).

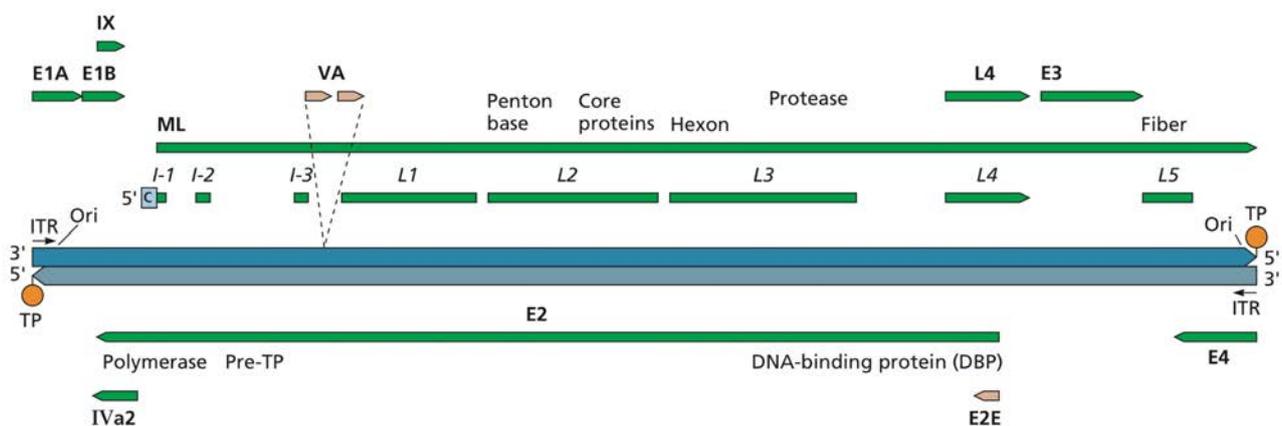


Figure 3. Organization of HAAdV-5 genome and transcriptional units. By convention, the genome is represented showing the E1A gene towards the left side. The blue strands represent the viral dsDNA genome with TP protein attached to the 5' ends. The green arrows represent each of the transcriptional units transcribed by RNAPol II. They are pointing in the direction and in the side of the genome that they are coded. The ML unit is represented by a long arrow and below, it is segmented to show where each late family starts and ends (as

well as the three introns of the tripartite leader). The two small yellow arrows represent the small VA RNAs that are transcribed by RNAPol III (image from Principles of Virology, 4th edition).

1.2 Single-cell reproductive cycle

The adenovirus reproductive cycle has been studied mainly in types 2 and 5 (HAdV-2 and HAdV-5) from subgroup C. Therefore, they are considered prototypes for adenovirus. By convention, the replication cycle of viruses with DNA genomes is divided mainly into two phases, an early and a late phase, separated by the onset of the viral genome replication.

1.2.1 Early Phase: Interaction with host cell and transcriptional activation

The immediate early events start with the interaction of the virus with the permissive host cell. Attachment takes place first, between the fiber knob domain and Coxsackie-Adenovirus Receptor (CAR) at the surface of the cell (Bergelson *et al.*, 1997; Bewley, 1999). The second interaction occurs between the penton base and integrins $\alpha_v\beta_3$ and $\alpha_v\beta_5$, which promotes the clathrin-mediated endocytosis of the virion (Nemerow and Stewart, 1999; Meier and Greber, 2003). This interaction also leads to the detachment of the fibers, which occurs before endocytosis (Nakano *et al.*, 2000). In the endosome, the penton bases and peripentoneal hexon trimers dissociate due to the low pH (Puntener *et al.*, 2011). This acidification reactivates the viral protease resulting in a cut on protein VI, which tethers the core and the capsid; this is required for final disassembly of internal capsid proteins IIIa, VI, and VIII (Greber *et al.*, 1993). Before fusion of the endosome with the lysosome, the virion escapes into the cytoplasm by a process mediated by a membrane-lytic domain of protein VI that remains associated to the inner surface of hexon trimers of the intact virion (Wiethoff *et al.*, 2005). Once in the cytoplasm, the virion is transported to the nucleus via dynein-mediated transport through microtubules arriving at the juxtannuclear microtubule-organization center (MTOC) (Dales and Chardonnet, 1973; Leopold *et al.*, 2000; Bailey, Crystal and Leopold, 2003). There, the partially disassembled particles associate with the nuclear pore complexes (NPCs) via interaction between hexon trimers and filament proteins from the nuclear pore complex CAN/Nup214 (Trotman *et al.*, 2001). Viral uncoating occurs at the nuclear pore as the capsid binds kinesin-1 light chains that pull the capsid away from the nucleus (Fay and Panté, 2015). Finally, only the viral DNA associated with core protein VII is imported into the nucleus by association with transportin, histone H1, and importin-7/importin- β heterodimer (Trotman *et al.*, 2001; Hindley, Lawrence and Matthews, 2007) (Figure 4). Once in the nucleus, protein VII-viral DNA complex associates with the nuclear matrix by interaction with TP (Schaack *et al.*, 1990). Protein VII remains associated with the viral DNA until it is dissociated by transcription. However, cellular histones

are known to replace some of the protein (Chen, Morral and Engel, 2007). Association of protein VII with the viral DNA prevents recognition during entry and therefore, activation of the cellular dsDNA break response (DSBR) (Karen and Hearing, 2011) (Figure 5, steps 1-5).

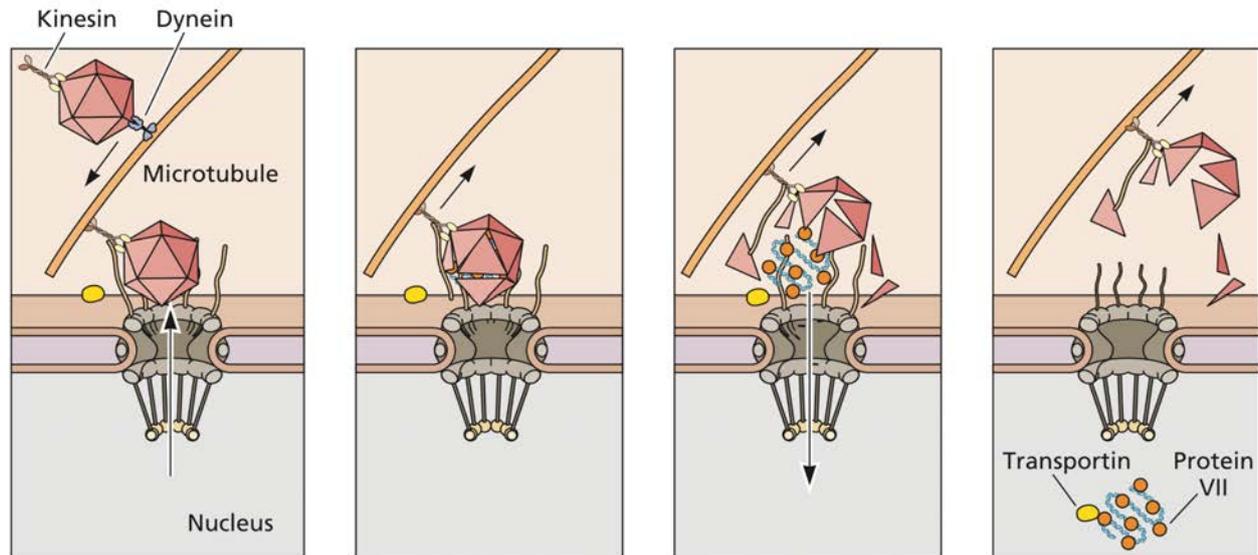


Figure 4. Uncoating of adenovirus at the nuclear pore complex. Partly disassembled capsids arrive at the nucleus via dynein-mediated microtubule transport. Particles interact with Nup24 in the NPC and the capsid binds kinesin-1 light chains and moves away from the nucleus. The protein VII-viral DNA complex enters the nucleus by an association of protein VII to transportin and other import proteins (image from Principles of Virology, 4th edition).

The immediate early E1A gene is the first to be transcribed (Nevins *et al.*, 1979). E1A proteins, 12S and 13S, regulate transcriptional activation of the early transcriptional units (Montell *et al.*, 1984). They also promote as the direct and indirect activation of cellular genes responsible for inducing the cell into S phase and creating a favorable environment required for viral replication (Ben-Israel and Kleinberger, 2002; Ghosh and Harter, 2003). Activation of the viral early genes results in accumulation of products from the transcriptional unit E2: the single-stranded DNA binding protein (DBP); the precursor of terminal protein (Pre-TP); and the viral DNA polymerase (Ad pol); which are all required for replication of the viral genome. The viral early mRNAs are transcribed in the nucleus and exported to the cytoplasm where they are translated. The recently synthesized early proteins are imported into the nucleus where they interact with cellular transcription factors (Oct 1 and NF1), inducing synthesis of the viral DNA (Bosher, Robinson and Hay, 1990; Mul, Verrijzer and van der Vliet, 1990) (Figure 5, steps 6-14).

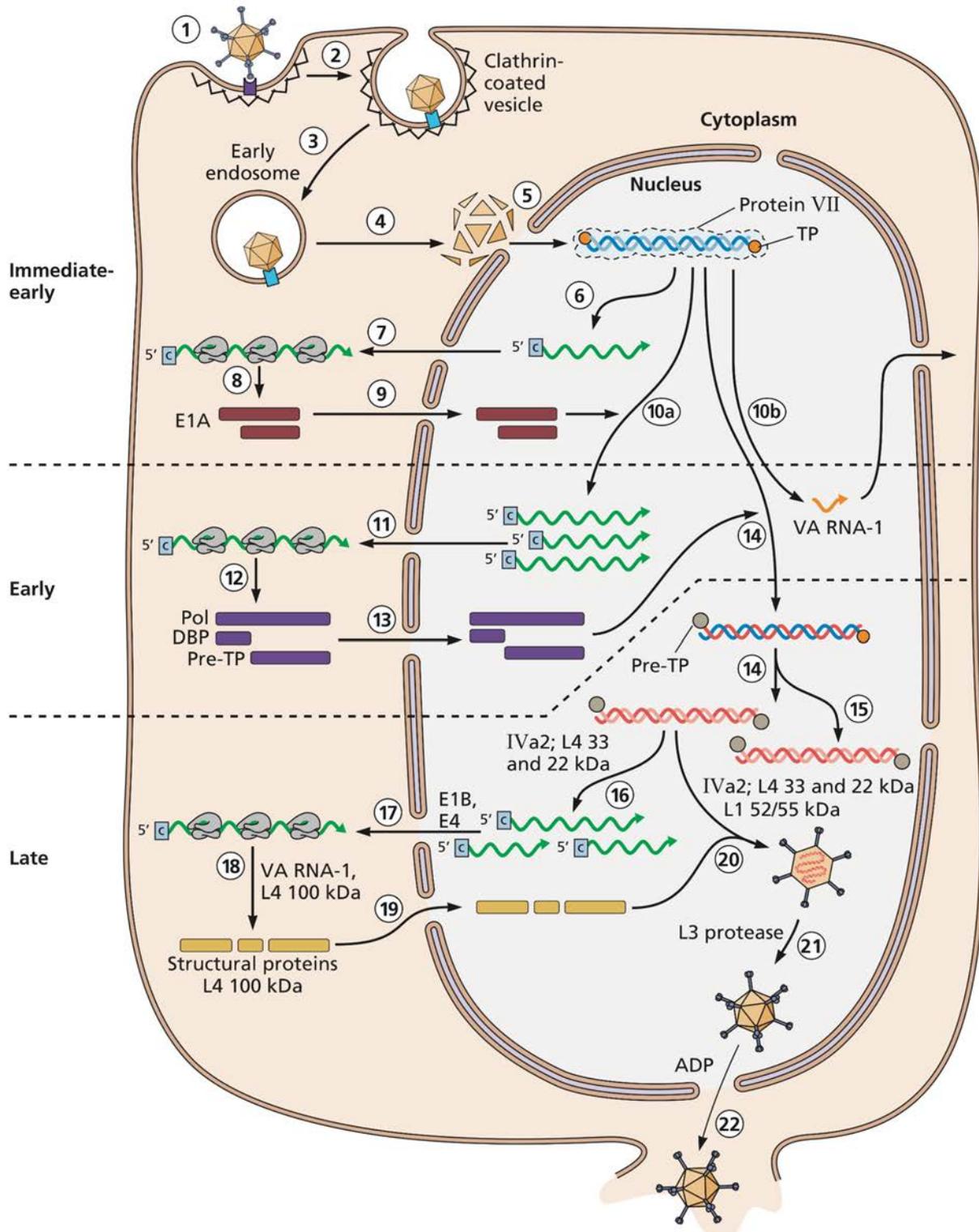


Figure 5. Single-cell reproductive cycle of adenovirus. Immediate early phase: The adenovirus particle attaches and enters the host cell, via clathrin-mediated endocytosis. Virion disassembles during the transit towards the nucleus. Only the viral genome attached to protein VII enters the nucleus. E1A is the first viral gene to be transcribed (steps 1-13). Early phase: the E1A products activate transcription of the other early genes and induce the host cell into S phase of the cell cycle. Mechanisms for inhibition of host cell defenses are established. Early proteins required for viral DNA synthesis accumulate in the nucleus (steps 1-13). Late phase: Replication of the viral DNA starts, which activates the transcription of the viral late mRNAs. Structural proteins are synthesized in the cytoplasm, involving regulating mechanisms by late proteins IVa2, L4-22K, L4-33K and L4-

100K. Hexon and penton capsomers form in the cytoplasm and enter the nucleus. Assembly of capsids and DNA packaging takes place in the nucleus. Protease activity matures the virion particles. Viral progeny is liberated by cell lysis (steps 14-22) (image from Principles of Virology, 4th edition).

1.2.2 Late Phase: Viral genome replication and selective viral late gene expression

Once the E2 proteins start to accumulate in the nucleus, replication of the viral DNA starts, commencing the late phase of the single-cell reproduction cycle. The inverted terminal repeats at each side of the viral chromosome, serve as replication origins for the viral polymerase (Ad pol). The Pre-TP is covalently linked to the 5'-ends of the viral DNA by Ad pol. A dCMP is covalently linked to serine 580, of the Pre-TP, providing a free 3'-OH that works as a primer for DNA synthesis (Smart and Stillman, 1982) and converting into the first nucleotide of the 5'-end of the DNA chain (Challberg, Desiderio and Kelly, 1980; Lichy, Horwitz and Hurwitz, 1981). Viral DNA synthesis occurs by chain displacement mechanism in two steps (Figure 6): 1) Replication starts at any of the chromosome's replication origins by recruitment of the preinitiation complex (Pre-TP, Ad pol, and DBP). Chain elongation continues (requiring Ad pol, DBP and cellular topoisomerase I) (Nagata, Guggenheimer and Hurwitz, 1983), until the other end producing a dsDNA product and a displaced single-stranded DNA (ssDNA) chain. 2) The ssDNA template circularizes by annealing of its complementary inverted terminal repeats forming a duplex "panhandle" that can be recognized as dsDNA substrate by the Ad pol (Lechner and Kelly, 1977). The newly synthesized dsDNA molecules can be used as templates to continue replication or for transcription of the viral late genes coded in the late transcriptional unit ML (Figure 5, steps 14-15).

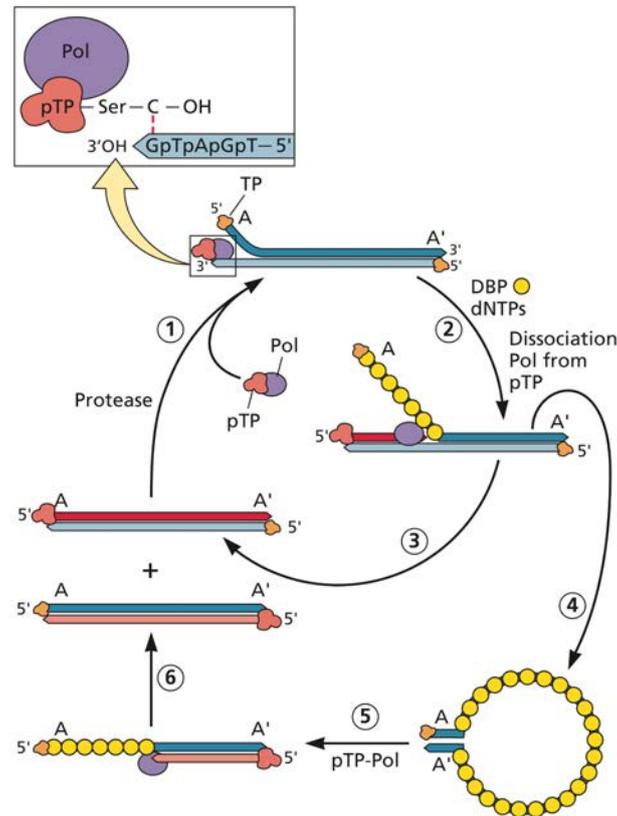
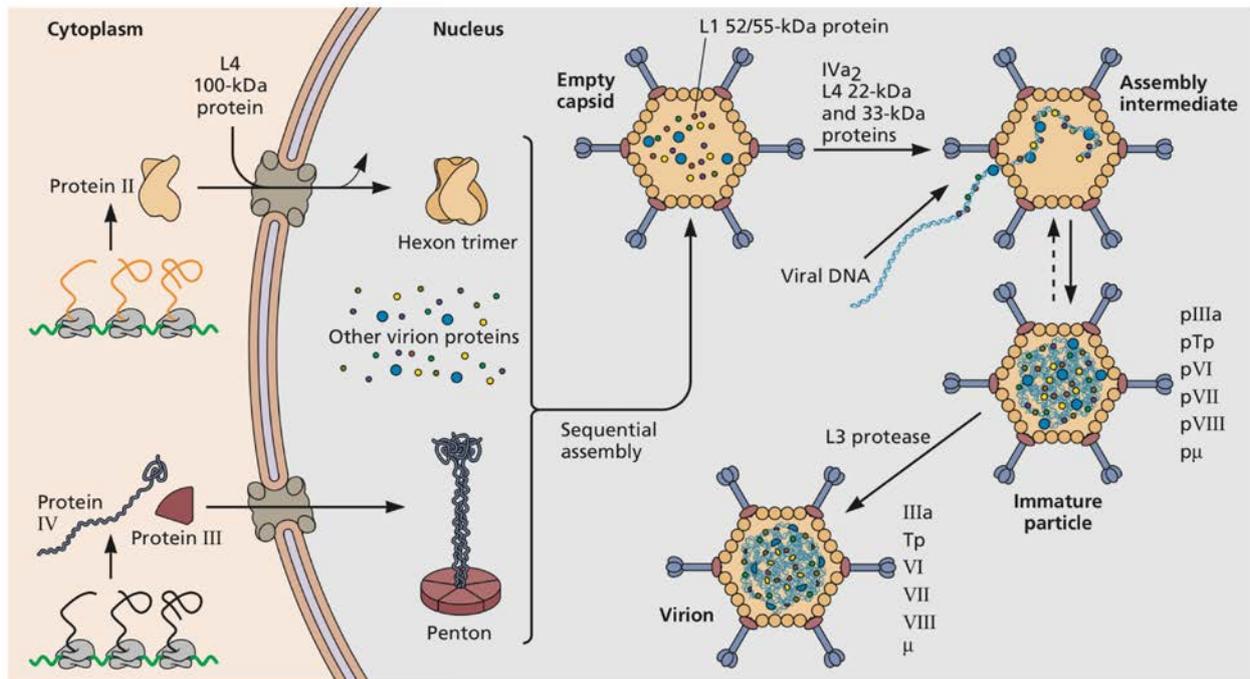


Figure 6. Replication of adenoviral DNA. Replication occurs in two stages: synthesis of the entire length of one strand first (steps 1-3). The replication of the complementary strand is synthesized second, requiring circularization of the displaced ssDNA to generate a dsDNA region ("panhandle" structure) to start replication (steps 4-6). As described in the text, initiation requires assembly of the pTP and Ad pol into a preinitiation complex in the terminal origin of replication. Ad pol covalently links the pTP with a dCMP. Replication is primed by the free 3'-OH group of the pTP-dCMP (image from Principles of Virology, 4th edition).

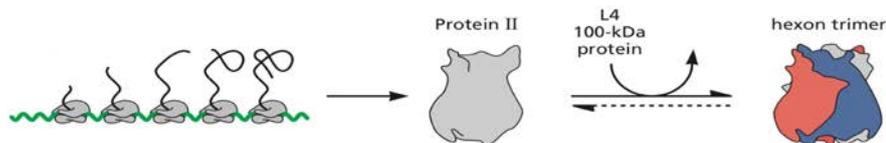
Transcription of viral late mRNAs is regulated by a series of mechanisms (involving late proteins IVa2, L4-22K, and L4-33K) that activate with the start of viral DNA replication. All late transcripts are transcribed by cellular RNAPol II and by differential poly(A) site selection. Thereby, five families of messengers are generated (L1-L5). During the late phase, several mechanisms are established to favor selection of viral late mRNAs over cellular transcripts for the use of cellular machinery required for gene expression (further described in section 1.5.3). The viral late mRNAs are exported through the NPC, requiring viral proteins E1B-55K and E4 Orf6 for their efficient accumulation in the cytoplasm. As explained further, viral genome replication and synthesis of the viral late mRNA take place in viral-induced nuclear structures named Replication Centers (RCs) (see section 1.6). Translation of viral late mRNAs occurs by ribosome shunting mechanism, which is promoted by viral L4-100K and VA-RNA I (Figure 5, steps 16-18). Viral late gene expression is described in detail in section 1.5.

After synthesis of large quantities of viral genomes and structural proteins, viral particles are assembled. L4-100K functions as a chaperone for hexon folding and as a scaffold for hexon trimers formation (Hong *et al.*, 2005) (Figure 7, panel B). Both hexon and penton capsomers are assembled in the cytoplasm and imported into the nucleus where virion assembly takes place (Horowitz GA, Scharff MD, 1969; Velicer and Ginsberg, 1970) (Figure 7, panel A). Encapsidation of the viral genome is a polarized process that requires a packing sequence of ~200 bp near the left end (Figure 7, panel A and C). Mutants of IVa2 and L1-52/55K have defects in virion packaging, suggesting that they are involved in this process (Hasson *et al.*, 1989; Gustin and Imperiale, 1998; Zhang and Imperiale, 2003). The precursor of protein VII (pVII) associates with the viral DNA during the late phase. The IVa2 and L1-52/55K interact with pVII and the viral DNA-pVII as a complex is packaged with core proteins V and precursor of μ (Weber and Philipson, 1984; Zhang and Arcos, 2005). The mature virions are formed when the precursor capsid proteins (VI, VII, VII, μ , and TP) are cleaved by the viral protease (L3-23K) (Figure 7, panel A). The protease requires DNA and pVI (or a cleaved peptide of pVI) as cofactors, which prevent the protease from activating before the protease and precursors are incorporated into virion particles (Mangel *et al.*, 1993; Tihanyi *et al.*, 1993; Webster, Hay and Kemp, 1993; Gupta *et al.*, 2004). The adenovirus death protein (ADP) (E3-11.6K) accumulates in the nuclear envelope during the late phase and promotes virus release from the nucleus by an unknown mechanism (Tollefson *et al.*, 1992, 1996). Finally, cleavage of cytokeratin K18, a cytoplasmic intermediate filament, by the protease compromises the structural integrity of the host cell and allows liberation of viral progeny by cellular lysis (Chen, Ornelles and Shenk, 1993) (Figure 5, steps 19-22).

A



B



C

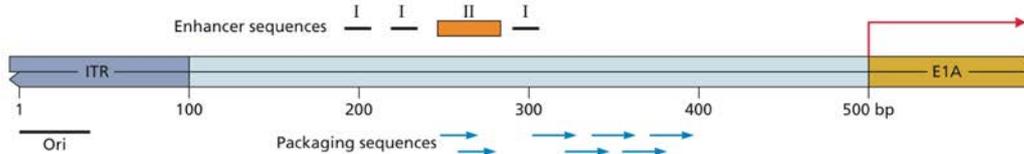


Figure 7. Adenovirus assembly and viral DNA-packaging. A) Capsid assembly starts with synthesis and assembly of hexon and penton capsomers in the cytoplasm and are then imported to the nucleus and empty capsids are assembled. The viral DNA is packaged into empty capsids requiring a packaging signal and viral proteins IVa2 and L1-52/55K, forming an assembly intermediate. Once the viral genome is completely packaged, the immature particle is formed and L3 protease activity is required to form mature virions. B) L4-100K is required as chaperone and scaffold protein to form hexon trimers. C) Viral DNA packaging signals are located near the left end of the viral genome, between the ITR and the E1A transcriptional unit (image modified from Principles of Virology, 4th edition).

1.3 Early genes

The viral early proteins establish three important events during the early phase that are required for a successful viral replication: 1) The S phase of the cell cycle is induced in the host cell to establish an adequate environment that allows replication of the viral genome. 2) Inhibition of the cellular defenses (apoptosis, interferon antiviral response, and double-strand break repair) that are activated during viral infection to protect the cell. 3) Synthesis of viral proteins that are necessary for viral genome replication. These three events depend on activation of the early transcriptional units at the start of the replication cycle, which are activated by the E1A coded proteins.

1.3.1 Transcriptional activation by E1A

In the immediate early phase, after the viral DNA enters the nucleus, the E1A unit is the first to be transcribed due to a very strong enhancer, making its promoter constitutively active (Nevins *et al.*, 1979; Hearing and Shenk, 1983). E1A generates two principal products, E1A 12S and E1A 13S, which result from differential splicing processing (Perricaudet *et al.*, 1979) (Figure 8, panel A). They only differ in 46 amino acids (aa) from the center region of 13S. They have five conserved regions (CR): The N-terminal, CR1, CR2, CR3 (only included in 13S) and CR4. They function as *trans*-activators and activate the viral transcriptional units as well as cellular genes. However, the E1A proteins do not have DNA binding domains nor depend on a specific promoter sequence, they activate transcription by binding to regulatory proteins and cellular transcriptional factors.

The E1A 13S activates the early transcriptional units. The activation domain is localized in the CR3 region of the E1A 13S, which binds to the MED23 subunit of the Mediator Complex and is essential in the regulation of transcription by RNAPol II (Boyer *et al.*, 1999; Stevens *et al.*, 2002) (Figure 8, panel A). This interaction favors assembly of the preinitiation complexes on promoters and stimulates transcription elongation (Cantin, Stevens and Berk, 2003; Wang *et al.*, 2005). Therefore, the CR3 domain of E1A 13S is responsible for stimulating the activity of early promoters E1B, E2E (E2 early promoter), E3, and E4 (Winberg and Shenk, 1984).

E1A proteins also regulate transcription of E2F-regulated promoters. E2F was discovered in studies with HAdV because it regulates the promoter of the E2 transcriptional unit. The viral E2 transcriptional unit encodes proteins required for viral DNA replication. In uninfected cells, retinoblastoma protein (pRB) binds to E2F preventing the activation of E2F regulated genes. E1A 12S and 13S bind to pRB, via CR1 and CR2, releasing E2F repression (Whyte *et al.*, 1988;

Bagchi, Raychaudhuri and Nevins, 1990) and therefore, activating transcription of the E2 and cellular genes required for progression into the S phase of the cellular cycle (such as DNA pol α , cyclin A and E, cdc2 and c-myc) (Egan, Bayley and Branton, 1989) (Figure 8, panel B). Therefore, regulation of E2F by E1A ensures progression into the late phase of the infectious cycle. During infection, the E1A proteins are regulated through several phosphorylation processes by cellular kinases CDK2, CDK4, and CDC2. E1A phosphorylation increases its association with pRB, thereby disrupting E2F-pRB more efficiently (Mal, Piotrkowski and Harter, 1996).

E1A proteins also regulate histone acetyltransferases, p300 and CBP, via their N-terminal domain. E1A interaction with these proteins results in epigenetic reprogramming of the host cell, stimulating cell cycle progression, inhibiting antiviral responses and cell differentiation (Ferrari *et al.*, 2008) (Figure 8, panel C).

The E1A transcription unit also codes for three alternatively spliced mRNAs: 9S, 10S, and 11S. These accumulate at later times of the infectious cycle, however, their functions remain unknown (Svensson, Petterson and Akusjärvi, 1983; Stephens and Harlow, 1987).

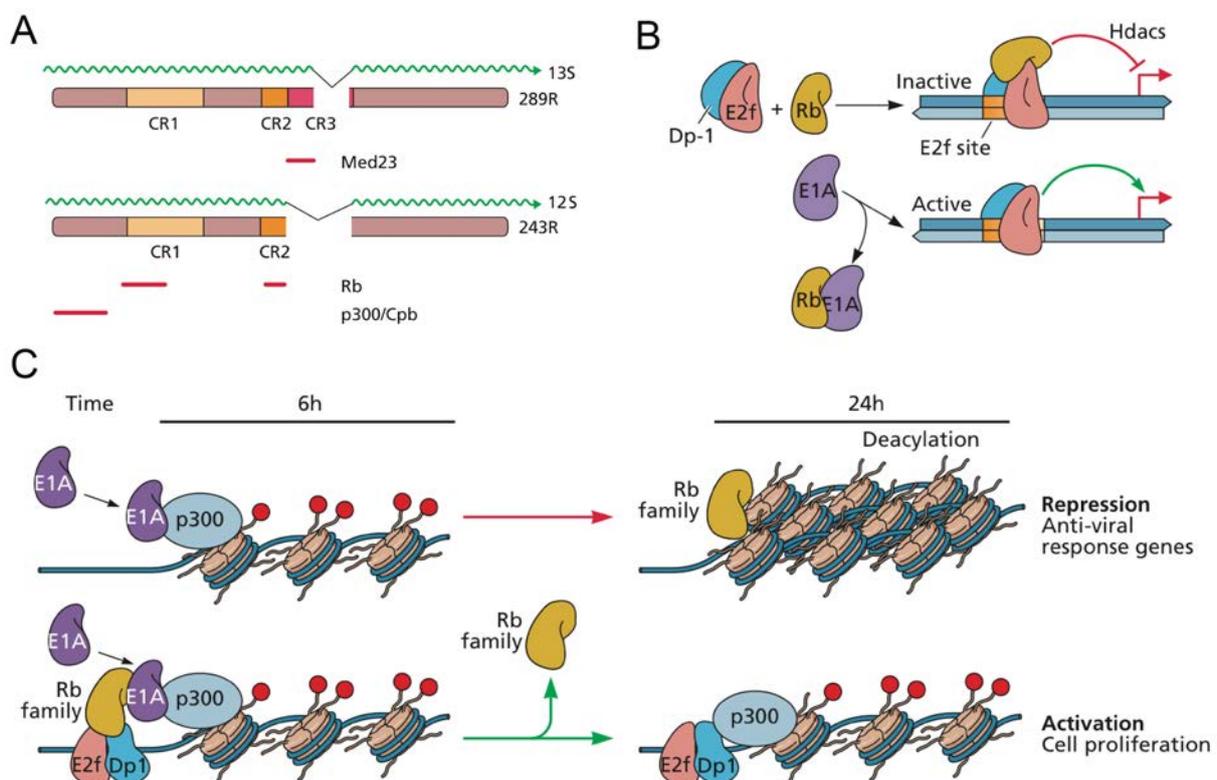


Figure 8. Transcriptional activation by E1A. A) E1A conserved regions that interact with cellular transcriptional factors (interactions are indicated with red lines). B) E1A 13S releases repression of E2F by binding pRB, activating E2F regulated genes. C) Epigenetic reprogramming of host cell by E1A. Interaction of E1A with p300 and Rb family proteins results in repression of antiviral response genes and activation of cell proliferation genes (image modified from Principles of Virology, 4th edition).

1.4 Inhibition of host cell defenses

1.4.1 Apoptosis inhibition

The early E1A-mediated transcriptional activation and induction of the cell cycle into S phase turns on p53-dependent and -independent apoptosis. The p53-dependent activation occurs by induction of the tumor suppressor p19^{ARF} by E1A (De Stanchina *et al.*, 1998). This tumor suppressor binds to MDM2, a protein required for ubiquitination and later degradation of p53, resulting in an increase of p53 concentration (Lowe and Earl Ruley, 1993). The p53-independent activation is induced by the viral protein E4 Orf4, which is also activated by E1A (Lavoie *et al.*, 1998). E4 Orf4 binds to the tumor suppressor phosphatase 2A (PP2A) and inhibits its activity (Kleinberger and Shenk, 1993; Marcellus *et al.*, 2000).

During infection, viral proteins E1B-55K (E1B), E1B-19K, and E4 Orf6 (Orf6) are responsible for inhibiting apoptosis. The E1B-19K is a homolog of BCL-2 and associates to pro-apoptotic proteins BAK and BAX inhibiting their oligomerization and formation of pores in the mitochondrial membrane (Debbas and White, 1993; Lowe and Earl Ruley, 1993; Cuconati and White, 2002). E1B binds to the N-terminal domain of p53, inhibiting transcriptional activation of p53-regulated genes (Yew and Berk, 1992; Renee Yew, Liu and Berk, 1994), as the p53-E1B interaction blocks the p53 interaction site with MDM2 (Sarnow *et al.*, 1982; Kao, Yew and Berk, 1990). However, p53 degradation is induced by the viral complex E1B/Orf6 which associates with cellular complex E3 ubiquitin ligase, integrated by Cullin 5, Elongin B and C, and Rbx1. The E1B/Orf6 complex binds to p53 directing it to the E3 ubiquitin ligase (E1B/Orf6/E3 ubiquitin ligase) for ubiquitination and further degradation via proteasome (Figure 9) (Querido *et al.*, 2001; Harada *et al.*, 2002). Orf6 alone can also associate with the C-terminal end of p53, also blocking its transcriptional activation activities (Dobner *et al.*, 1996).

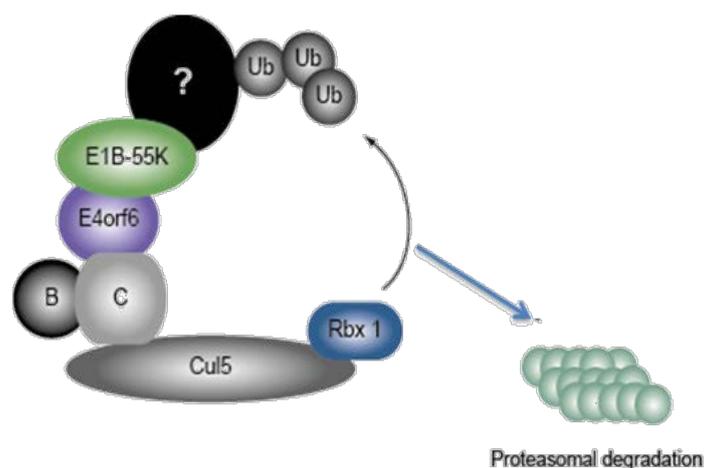


Figure 9. Diagram of the E1B/Orf6/E3 ubiquitin ligase. The E3 ubiquitin complex is formed by Elongin B and C (represented as B and C), they interact with the N-terminal domain of Cullin 5 (Cul5) which interacts, through its N-terminal domain, with the Rbx1. The E1B/Orf6 complex interacts with the E3 ubiquitin complex through binding of Orf6 with Elongin C. The E1B/Orf6 complex can bind to different cellular proteins and direct them to the E3 ubiquitin ligase complex. Polyubiquitination of selected proteins results in their degradation via proteasome.

Viral proteins coded in the E3 transcriptional unit also inhibit antiviral host responses by blocking the induction of apoptosis by Cytotoxic T Lymphocytes (CTLs) and TNF α (Andersson *et al.*, 1985; Burgert and Kvist, 1985; Li *et al.*, 1999).

1.4.2 Inhibition of host cell response to dsDNA breakage

Assays with adenoviral E4 null mutants have shown that during infection the dsDNA break response (DSBR) is activated during infection and concatemers with multiple copies of the viral DNA are formed, which are ligated by the non-homologous end joining (NHEJ) mechanism. Recognition of the viral genome termini is inhibited by different functions of the early proteins E4 Orf3 (Orf3), Orf6 and E1B (Stracker, Carson and Weitzman, 2002). Mutants in which E4 is not expressed show activation of the DNA damage response kinases ATM and ATR. These kinases activate cellular proteins involved in DNA damage response, like the MRN complex (composed by Mre11, Rad50, and Nbs1). E1B and Orf3 counteract this response by sequestering MRN complexes towards the Nuclear Domain 10 (ND10) and subsequently exporting them to aggresomes at the MTOC (Araujo *et al.*, 2005; Evans and Hearing, 2005; Liu *et al.*, 2005; Blanchette *et al.*, 2013). Relocalizing the MRN complex prevents activation of ATR protein kinase; thereby, the activation of the DSBR during infection is averted (Carson *et al.*, 2009). Additionally, the E1B/Orf6/E3 ubiquitin ligase complex, which was described previously (Figure 9), can ubiquitinate and send all components of the MRN complex for degradation (Stracker *et al.*, 2002), DNA ligase IV, and Bloom helicase, all involved in the DSBR (Baker *et al.*, 2007; N. I. Orazio *et al.*, 2011).

1.4.3 Inhibition of interferon-induced antiviral response

The Orf3 protein accumulates in the early phase of infection in the ND10 (Nuclear Domain 10), which are also known as PML-bodies (Carvalho *et al.*, 1995). In non-infected cells, the ND10 are round punctuated structures inside the nucleus. They contain several isoforms of the PML (Promyelocytic leukemia protein) protein, which are responsive to interferon (INF) and function as a scaffold for other ND10 components (Maul *et al.*, 2000). Interaction of Orf3 with PML isoform II induces a reorganization of these structures from rounded to track-like structures called PML tracks (Figure 10) (Hoppe *et al.*, 2006; Leppard *et al.*, 2009). Other ND10 components such as Daxx, which inhibits antiviral response induced by INF, are also redistributed by interaction with E1B (Zhao *et al.*, 2003). In the late phase, E1B marks Daxx for degradation via proteasome, by association with the E3 ubiquitin ligase, but it occurs independently of Orf6 (in contrast with p53 and Mre11 degradation) (Schreiner *et al.*, 2010). As described further on, the reorganization of the ND10 is accompanied by the formation of nuclear structures where the viral genome is replicated and transcribed (section Adenovirus Replication Centers).

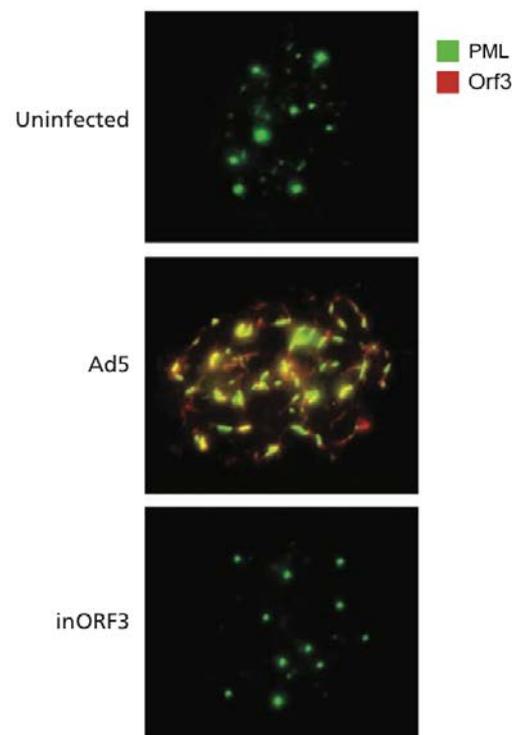


Figure 10. Reorganization of ND10 by Orf3. Immunofluorescence of infected monkey cells with wt HAdV-5 (Ad5) and null-mutant of Orf3 (inORF3). Images are a close-up to the nucleus. Orf3 is stained in red and PML in green (image taken from Principles of Virology, 4th edition, adapted from (Ullman, Reich and Hearing, 2007)).

1.5 Late genes

The viral late genes code for structural and catalytic proteins that are required for viral progeny production. Productive formation of new viral particles depends on the efficiency of the viral genome replication and the synthesis of the structural proteins. Importantly, expression of viral late genes requires the cellular machinery for mRNA transcription, processing, nuclear export, and translation. However, during infection cellular mRNA synthesis is not blocked, therefore, cellular and viral mRNAs compete for cellular machinery at each point of the synthesis pathway. Nevertheless, at late time points of the single-cell reproduction cycle, almost all of the synthesized proteins are viral proteins (Beltz and Flint, 1979). This is the result of different selective mechanisms, established by viral proteins during the late phase. Viral late mRNA selection mechanisms start at transcription, with signals that highly amplify the activation of the ML promoter (MLP) at the start of viral DNA replication. Followed by alterations in the cellular post-translational processing machinery that favors the adequate processing of the different viral late mRNA families, which require splicing of very large introns. Finally, two mechanisms are known to confer an advantage to viral late mRNAs over the host cell transcripts: 1) the preferential selection of viral late mRNAs for their cytoplasmic export, and 2) preferential selection of viral late mRNAs by the translation machinery. Establishment of these mechanisms involves interaction and/or interference of viral proteins with cellular machinery components, compromising their activity with the synthesis of viral late genes and resulting in a very efficient viral progeny production. As it is described further (section 1.6), several of these steps, such as replication of the viral genome, synthesis of viral late mRNAs, and possibly post-transcriptional processing, take place in nuclear structures induced by HAdV infection named Replication Centers (RCs).

1.5.1 Transcriptional activation of viral late genes

The five families that code for the viral late mRNAs (L1-L5) are organized in one unique transcriptional unit under the regulation of the MLP. This promoter is activated at very low levels, by E1A, since the early phase of the single-cell reproductive cycle. However, its transcription is very low and abortive, and only the genes closest to the promoter are synthesized (L1-52/55K) (Akusjarvi and Persson, 1981a; Nevins and Wilson, 1981a). It is not until the viral DNA synthesis starts, that the MLP activity increases several hundred-folds (Shaw and Ziff, 1980). The delay in the activity increase of this promoter is due to its positive regulation mechanisms by *cis* and *trans* elements, which ensure its maximum activity until the late phase.

The *cis*-activation of ML depends on the start of the viral DNA replication (Thomas and Mathews, 1980). The reason for the requirement of viral DNA replication is not understood, but it is thought that this mechanism requires access into a binding site in the viral genome, which is only exposed on newly synthesized genomes. Besides E1A, the IVa2 and viral late proteins L4-22K and L4-33K (L4-22/33K) are necessary for complete *trans*-activation of late unit synthesis.

IVa2 is known as a delayed early protein because it is transcribed after the rest of the early genes in the transition into the late phase, due to its transcriptional repression, which is regulated by the replication of the viral genome (Binger and Flint, 1984). Once IVa2 is synthesized, it binds to the MLP, acting as a transcriptional activation factor (Pardo-Mateos and Young, 2004). However, the late transcriptional unit measures 28 kbp and there is not enough transcriptional progression to generate the mRNAs of the most distant families from the ML transcriptional unit (such as L3, L4, and L5). Increase in transcriptional progression requires viral late proteins L4-22/33K, which are under the regulation of a promoter inside the ML transcriptional unit (within the L4-100K coding region), independent from the MLP. Activation of this internal promoter depends on proteins E1A, Orf3 and IVa2 (Morris, Scott and Leppard, 2010). Together, IVa2 and L4-22/33K, successfully stimulate expression of all five viral late families (Backström, Kaufmann, Lan, & Akusjärvi, 2010; Wu, Guimet, & Hearing, 2013). The L4-22K can also stimulate expression of IVa2 resulting in a positive regulatory loop ensuring activation of other MLP in newly synthesized viral genome copies (Morris *et al.*, 2010). Altogether, these regulating mechanisms work as a timer that activates the synthesis of viral late genes until the viral genome replication starts and the number of viral genome copies increases.

1.5.2 Post-transcriptional processing

All the viral transcriptional units that are transcribed by the RNAPol II, except for IVa2 and IX, give place to more than one product by alternative splicing. During the late phase, changes in the post-transcriptional machinery, which favor the cleavage of bigger introns, are induced both in early genes and in late genes. This change is especially important to favor the progression during transcription of the late genes coded in the ML transcriptional unit.

The ML transcriptional unit has a length of ~28 kb and at least 15 different late mRNAs are produced from this sequence (Nevins and Darnell, 1978). All viral late mRNAs share a segment of three small exons in the 5'-end named tripartite leader (Berget *et al.*, 1977; Chow *et al.*, 1977), but the synthesis of each late mRNA family depends on the transcription progression, the 3'-intron acceptor selected to bind the tripartite leader during post-transcriptional

processing and the poly(A) site selection (Nevins and Darnell, 1978; Ziff and Fraser, 1978). Synthesis of the mRNA families that are further from the promoter requires bigger intron fragments to be processed each time (Figure 11).

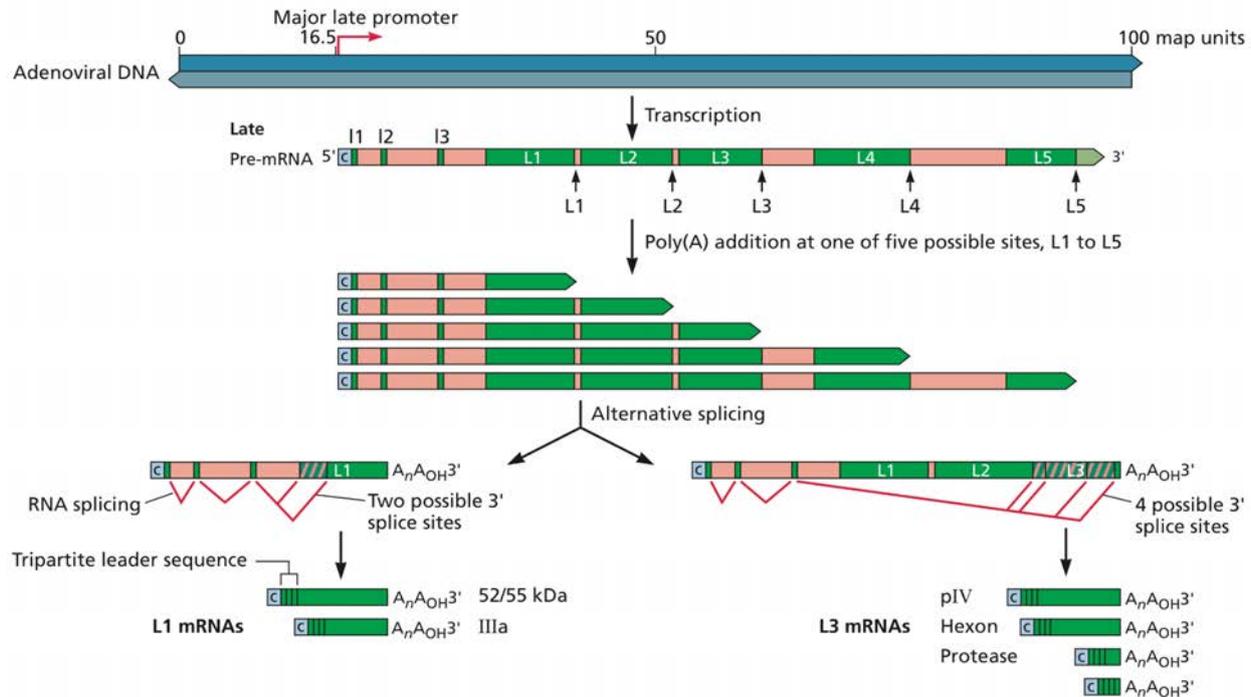


Figure 11. Alternative polyadenylation and splicing of viral Major Late transcripts. All viral late mRNAs are coded in a single transcriptional unit under the regulation of the Major Late promoter. A single long pre-mRNA is coded; however, there are five possible poly(A) sites that can be selected. Each poly(A) site gives place to a late mRNA family (L1-L5), and each family can be processed by alternative splicing to produce more than one mRNA. All mRNAs share three short exons (I1, I2, and I3) at the 5'-end. The L1 mRNA has two possible splice sites that can be linked with the tripartite leader and will produce two different proteins, L1-52/55K and L1-IIIa. The L3 mRNA has four possible splice sites that generate pIV, hexon and the protease. These mechanisms allow the adenovirus to synthesize at least 15 different mRNAs from a single transcriptional unit (image from Principles of Virology, 4th edition).

During the late phase, interactions of the viral E4 Orf4 protein with the cellular phosphatase PP2A induces phosphorylation of SR proteins (serine- and arginine-rich proteins, involved in splicing processing), which results in a change in the available SR proteins altering the patterns of hnRNA processing (A Kanopka *et al.*, 1998; Brestovitsky *et al.*, 2011; Zhang *et al.*, 2011). Synthesis of late protein L4-22/33K induces splicing processing in specific sites (Biasiotto *et al.*, 2015). At the early phase, the L1-52/55K mRNA is synthesized in low quantities, and in the late phase, the alternative spliced mRNA that codes for the IIIa protein is produced (Akusjarvi and Persson, 1981a; Nevins and Wilson, 1981a; Larsson, Svensson and Akusjärvi, 1992). The splice site for L1-52/55K mRNA is similar to the mammalian canonical splicing sites, and it is selected during the early phase. However, in the late phase, the L4-33K protein binds upstream the IIIa 3' splice site and acts as a splicing enhancer favoring its selection (Figure 12) (Törmänen *et al.*, 2006). L4-33K can also stimulate splicing of other suboptimal 3' splice sites in late pre-mRNAs

such as L2 to produce protein V and pVII. Since phosphorylation of the SR proteins is necessary for them to bind to their consensus splicing signals, it is possible that the dephosphorylation induced by E4 Orf4 affects the host cell gene expression (A. Kanopka *et al.*, 1998). Also, during infection, the cellular processing machinery in the nucleus is redistributed towards the sites of viral late mRNAs synthesis, being more accessible for viral late mRNAs (further described in section 1.6).

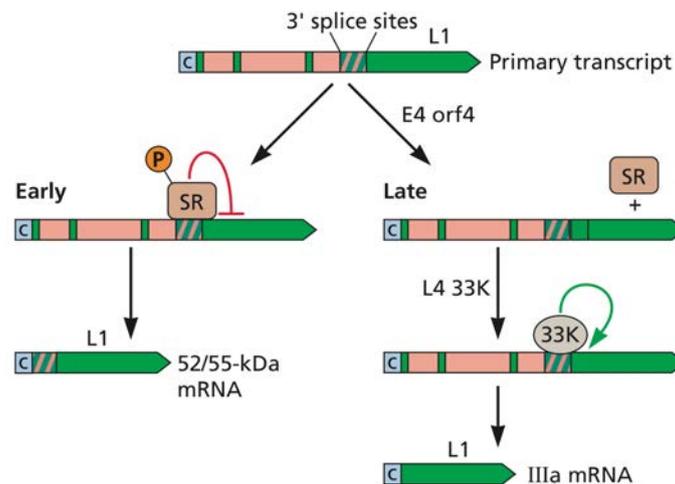


Figure 12. Regulation of alternative splicing. The L1 pre-mRNA has two alternative 3' splice sites. One produces the L1-52/55K protein and the other the IIIa protein. In the early phase, only the L1-52/55K mRNA is produced because SR proteins bind to the pre-mRNA and block recognition of the IIIa alternative splice site. In the late phase, E4 Orf4 induces dephosphorylation of the SR proteins by binding to PP2A. Dephosphorylated SR proteins do not bind the pre-mRNA allowing access to the alternative splice site for IIIa. Efficient splicing of this site requires binding of late protein L4-33K that acts as a splicing enhancer (image from Principles of Virology, 4th edition).

1.5.3 Selective expression of viral late mRNAs

As described before, during infection there is no inhibition of cellular mRNA expression, however most of the proteins synthesis in the late phase is viral (Beltz and Flint, 1979). One of adenovirus strategies is to synthesize a great quantity of viral mRNAs during the late phase, however, they still compete with cellular mRNAs for the use of the processing, export, and translation machinery, which would diminish the efficiency to form viral progeny. Two events have been described to favor synthesis of viral proteins, over cellular proteins: selective export and selective translation of viral late mRNAs. As discussed previously, the proteins E1B, Orf3, and Orf6 counteract cellular defenses allowing to establish optimum conditions for late gene synthesis; however, these proteins are also important during expression of viral late genes.

1.5.3.1 Selective Export of viral late mRNAs

While studying the inhibition of host cell protein synthesis, the overall metabolism of the host cell mRNA was analyzed during the late phase, when protein synthesis is drastically reduced by more than 90%. It was found that transcription of cellular mRNAs was not affected, but the transport of newly synthesized cellular mRNAs to the cytoplasm was decreased (Beltz and Flint, 1979; Babich *et al.*, 1983). It was also shown that this effect could not account completely for the whole inhibition in cellular protein synthesis (see section 1.5.3.2 Selective Translation). This suggested that the viral mRNAs were preferably selected for transport to the cytoplasm while cellular mRNA transport was blocked and they accumulated in the nucleus. Cytoplasmic export of small cellular RNA species (rRNA and tRNAs) was not found to be affected during infection (Castiglia and Flint, 1983; Smiley *et al.*, 1995; Rabino *et al.*, 2000). Therefore, the blockage on cytoplasmic export was discovered to be restricted to cellular mRNAs. In the absence of E1B, synthesis of viral mRNAs was not affected, but delivery to the cytoplasm was reduced and cellular mRNA synthesis and transport were unaffected (Babiss, Ginsberg and Darnell, 1985). And thus, the concept of selective export was originated, referring to effects shown in the late phase where viral late mRNAs being selected for cytoplasmic export while cellular mRNAs remain inside the nucleus.

Selective viral late mRNA export has been shown to depend on the presence of the E1B protein (Babich *et al.*, 1983; Pilder *et al.*, 1986; Williams *et al.*, 1986) and the complex it forms with Orf6 (Bridge and Ketner, 1990). Both proteins contain nuclear export sequences (NES) and nuclear localization sequences (NLS), so they can shuttle between the nucleus and cytoplasm (Dobbelstein *et al.*, 1997; Krätzer *et al.*, 2000; Dosch *et al.*, 2001). Initially it was proposed that E1B/Orf6 could act as a nucleo-cytoplasmic shuttle complex for viral late mRNAs, and therefore inhibiting cellular mRNAs export during the late phase. The E1B has an RNA-binding domain and has been shown to bind to mRNA *in vitro*; however, binding is non-specific to viral mRNAs (Horridge and Leppard, 1998). Also, inhibiting E1B and Orf6 CRM1-dependent nuclear export did not impact the selective export during the late phase (Schmid, R. Gonzalez and Dobner, 2012). Also, neither E1B or Orf6 are closely associated with the NPCs (Smiley, Young and Flint, 1990; Ornelles and Shenk, 1991). As described before, all viral late mRNAs share a tripartite sequence (see section 1.5.2), which has shown to increase the export efficiency of a mRNA to even five times (Flint, 1998). However, this is not found to be the factor to which selective export is related to E1B, since mRNAs without this tripartite leader (such as influenza mRNAs and cellular mRNAs), when expressed within adenovirus chromosomes, can escape export blockage and accumulate in the cytoplasm during the late phase (Gaynor, Hillman and Berk, 1984; Hearing and Shenk, 1985; Moore *et al.*, 1987). The early delayed proteins IVa2

and IX do not have the tripartite leader sequence and are also selectively exported during the late phase. The only requirement for cellular mRNA to be exported is to be transcribed in the sites of viral mRNA transcription (Yang, Huang and Flint, 1996).

Until now, the molecular mechanism behind selective export of viral late mRNAs remains unknown. Two different mechanisms (which are not mutually exclusive) have been proposed and are based on the activities known for the E1B/Orf6 complex: 1) the relocalization of cellular components to organize nuclear domains known as Replication Centers (RCs), where the virus late transcription takes place (see section 1.6), and 2) the association with the E3 ubiquitin ligase and targeting cellular components for degradation.

The first model suggests that sites of viral transcription occupy discrete nuclear microenvironments that are functionally specialized to promote mRNA export. An important observation supporting this model is that the requirement for an mRNA (cellular or viral) to be exported is to be transcribed in the sites of viral mRNA transcription (Yang *et al.*, 1996). The RCs are the sites where the viral genome is replicated and where the viral late mRNAs are synthesized and initially processed (Mul *et al.*, 1990; Pombo *et al.*, 1994; Puvion-Dutilleul *et al.*, 1994; Bridge and Pettersson, 1996; Aspegren, Rabino and Bridge, 1998). Several components of the cellular splicing machinery have been observed to relocalize towards the RCs. Furthermore, fractions enriched with RCs can support transcription and splicing of pre-mRNAs *in vitro* (Hidalgo *et al.*, 2016). It has also been shown by fractionation assays that, in the absence of E1B, viral mRNAs fail to move away from the nuclear matrix (Leppard, 1993). Suggesting that E1B function is upstream of the NPC translocation. The E1B protein also accumulates in the RCs. Mutants of the virus where the E1B/Orf6 interaction is interrupted, not only fail to accumulate late mRNAs in the cytoplasm but E1B also no longer locates towards the RCs (Ornelles and Shenk, 1991; Gonzalez and Flint, 2002). E1B also binds with cellular protein E1B-AP5 (AP5) which contains RNA-binding domains (Gabler *et al.*, 1998) and can interact with TAP-mediated export machinery. AP5 is also localized in the RCs during infection (Blackford, Bruton, Dirlík, S. Stewart, *et al.*, 2008). It has been found that viral early mRNAs export is dependent on CRM1 (Schmid, R. Gonzalez, *et al.*, 2012) but the viral late mRNA export is TAP mediated (Yatherajam, Huang and Flint, 2011). TAP-mediated mRNA export is known to be coupled to posttranscriptional machinery and only mature processed mRNAs are selected to be exported (Rougemaille *et al.*, 2008; Katahira, 2015). This supports the idea that in the late phase viral mRNA export requires that transcription takes place in specialized domains enriched with cellular mRNA processing machinery. Thereby, favoring viral late mRNA export by coupling mRNA transcription with processing and export machinery.

The second more recent model proposes that the E1B/Orf6/E3 ubiquitin ligase (see Figure 9) targets cellular proteins, involved in post-transcriptional processes and/ or mRNA export, for degradation resulting in a defect in cellular mRNA export but favors viral late mRNA export during the late phase. This model is proposed based on studies in which the viral ubiquitin complex is no longer formed, due to the absence of Cullin 5 (part of the ubiquitin ligase complex). Further, a mutation in Orf6, which inhibits assembly with the E3 ubiquitin ligase, showed a defect in late mRNA export (Woo and Berk, 2007; Paola Blanchette *et al.*, 2008). Inhibition of an active proteasome during infection is also reported to affect viral late gene expression (Corbin-Lickfett and Bridge, 2003). However, none of the known substrates for the E1B/Orf6/E3 ubiquitin ligase (p53, MRN complex, DNA ligase IV, integrin α 3, and Bloom helicase) are involved in mRNA biogenesis (Querido *et al.*, 2001; Stracker *et al.*, 2002; Baker *et al.*, 2007; Dallaire, Blanchette and Branton, 2009; Nicole I Orazio *et al.*, 2011). Also, when identifying TAP as the export pathway required for viral late mRNA export, no degradation of any of the principal components of this export pathway was observed (Yatherajam *et al.*, 2011). Therefore, no evidence for degradation of proteins involved in mRNA processing or export pathways have been found so far.

In this present work, we further investigate the phenomenon of selective export by using a global approach to analyze viral and cellular mRNAs during infection, and we search for possible candidates for the E1B/Orf6/E3 ubiquitin ligase involved in mRNA biogenesis (further described in Project Aims).

1.5.3.2 Selective Translation of viral late mRNAs

Once in the cytoplasm, viral mRNAs still compete with cellular mRNAs over the cellular translation machinery. However, transition into the late phase results in inhibition of host mRNA translation (Tal, Craig and Raskas, 1975; Cuesta, Xi and Schneider, 2001) and preferential selection of viral mRNAs by translational machinery (Babich *et al.*, 1983). This has been described to occur by two mechanisms: 1) Inhibition of Cap-dependent translation of cellular mRNAs. The subunit eIF4E of the Cap-dependent RNA helicase eIF4F becomes dephosphorylated and therefore inactivated in the late phase (Huang and Schneider, 1991). Dephosphorylation of eIF4E occurs as the result of the displacement of kinase Mnk1 from the translation initiation complex by L4-100K interaction with eIF4G (Cuesta, Xi and Schneider, 2000, 2004). Viral late mRNAs also contain a 5'-Cap. However, they can overcome the blockage of Cap-dependent translation because they possess the tripartite leader (described in section 1.5.2). This sequence is 201 nucleotides long and promotes translation by ribosome shunting mechanism, which allows the ribosome to scan the non-translated 5' region until the

start codon. This is due to the complementarity of the tripartite leader with the 18S rRNA (Yueh and Schneider, 1996). This mechanism is also stimulated by viral protein L4-100K, which binds both the tripartite leader and the eIF4G subunit of the eIF4F (Hayes *et al.*, 1990; Xi, Cuesta and Schneider, 2004, 2005). 2) Inhibition of PKR to avoid phosphorylation of the initiation factor eIF2 α . Phosphorylation of eIF2 α inhibits translation initiation globally. Adenovirus avoids activation of PKR by synthesizing RNA-VA I, which accumulate up to 10⁹ copies per cell. These small RNAs fold to produce a dsRNA and bind to PKR blocking its activation (Kitajewski *et al.*, 1986; Katze *et al.*, 1987).

Some evidence suggests that E1B could also be involved in selective translation: 1) Cells infected with an E1B-null mutant besides showing a deficient selective export, there is bigger defect in production of viral late proteins, which does not correlate with the defect on accumulating viral late mRNAs in the cytoplasm (Babiss *et al.*, 1985; Harada and Berk, 1999). 2) During infection with E1B-null mutants, the eIF4E factor remains phosphorylated suggesting E1B has a direct effect on translation (Zhang, Feigenblum and Schneider, 1994). 3) Proteomic analyses have shown an association of E1B with the translation machinery (Harada *et al.*, 2002).

1.6 Adenovirus Replication Centers

The HAdV RCs are structures formed during infection inside the host nucleus where the viral genome is replicated, viral late mRNAs are synthesized, and initially post-transcriptionally processed (Pombo *et al.* 1994). These structures are formed during the late phase of infection and are the result of reorganization and recruitment of both viral and cellular proteins (Schmid *et al.*, 2014). Molecular processes that occur in the RCs require a great quantity of cellular machinery. Part of this machinery is localized in nuclear domains which are reorganized during infection with HAdV (Figure 13).

Fluorescence microscopy analyses have shown that the RCs are characterized by the presence of the viral ssDNA binding protein DBP. This protein participates in viral genome replication, its affinity towards single DNA strands is so high that its interaction forces the aperture of the double helix, promoting replication (Stuiver and van der Vliet, 1990; van Breukelen *et al.*, 2003). At the beginning of the formation of RCs, there is an accumulation of DBP and E1B in structures with a dot-like shape and form adjacent to the PML tracks. The size of these structures increases by the accumulation of DBP, E1B, and other components from the ND10 such as SP100, p53, Daxx, ATR, ATM and MRN (Figure 13, panel II) (reorganization of ND10

is described in section 1.4.3). As the infection progresses, DBP distribution starts to change into ring-like structures (Figure 13, panel I).

Components of Interchromatin granules (ICG) and Cajal Bodies (CBs) are reorganized since the early phase of infection (Bridge *et al.*, 1993; Rebelo *et al.*, 1996). ICG and CBs are nuclear domains where the cellular splicing machinery, such as snRNPs, SR proteins and exon-binding complexes among other proteins are localized (Lamond and Sleeman, 2003; Lamond and Spector, 2003). Early after infection, snRNPs present in ICS and CBs redistribute and as infection progresses they accumulate in growing punctuated structures that surround the RCs (Figure 13, panel III).

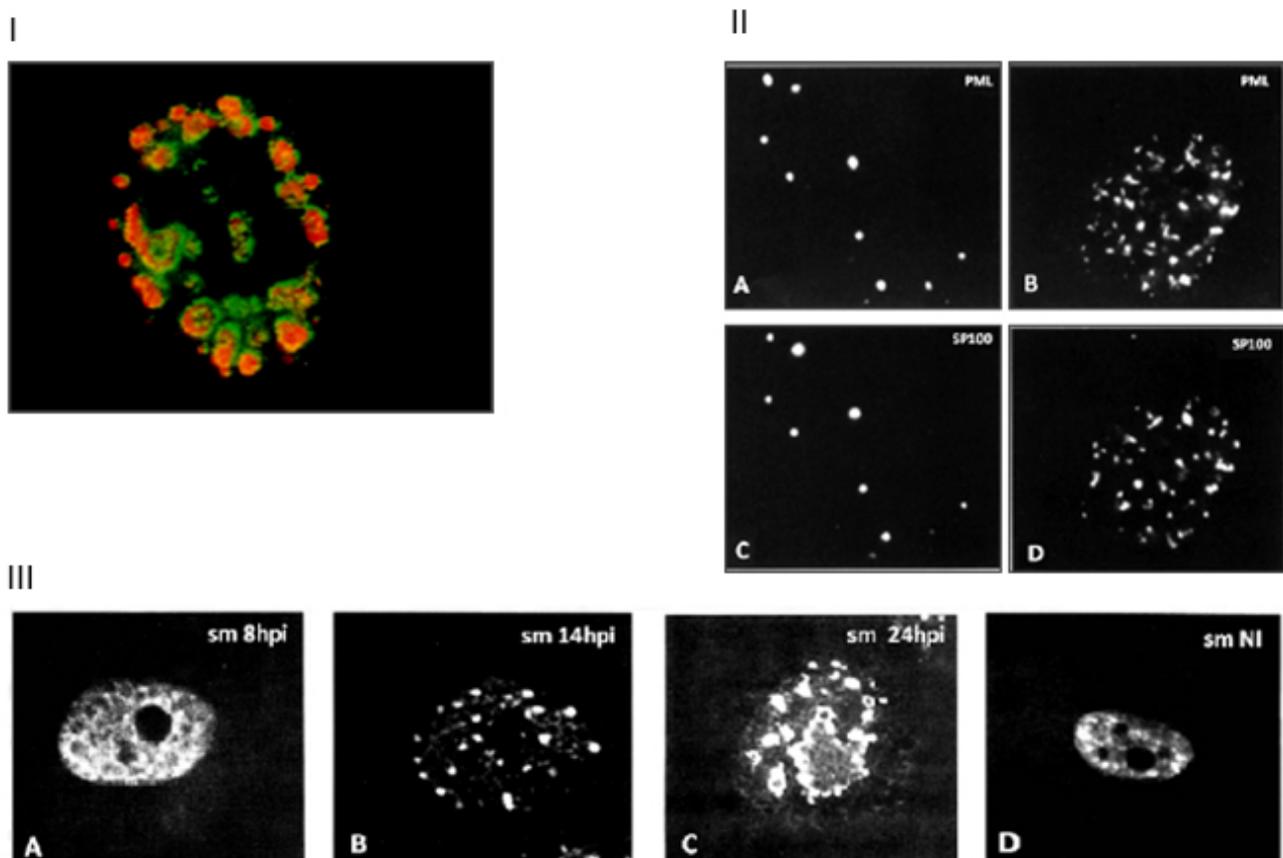


Figure 13. Replication Centers formation and nuclear structures reorganization. I) Replication Centers. Immunofluorescence showing a close-up to a HeLa cell nucleus infected with HAdV-2 at 20 hpi. Structures inside the nucleus are the viral-induced RCs. DBP is shown in red staining and viral RNA synthesis sites were marked by biotin-dUTP and are stained in green (adapted from Pombo *et al.*, 1994). II) Reorganization of ND10 domains. Immunofluorescences showing a nucleus close-up of non-infected (A and C) and infected (B and D) HeLa cells at 8 hpi with wt HAdV-5. Reorganization of ND10 abundant components, SP100 and PML, is shown in B and D (adapted from Carvalho, *et al.*, 1995). III) Reorganization of Interchromatin granules and Cajal Bodies. Immunofluorescences showing a nucleus close-up of non-infected (D) and infected HeLa cells at different time-points (A-C). Reorganization of ICG and CBs during infection is shown by staining with anti-Sm a domain present in snRNP proteins which are concentrated in both ICG and CBs (adapted from Bridge, *et al.*, 1993).

Electronic microscopy analyses have allowed observing the organization and thereby functions associated to RCs with higher resolution (Figure 14). The center of these sites contains

very dense structures where viral dsDNA that is not being transcribed accumulates. Surrounding these structures, there is a fibro-granular net named peripheral replication zone (PRZ) where sites in which ssDNA accumulates are embedded. This PRZ is the site where replication and transcription take place (Puvion-Dutilleul, Roussev and Puvion, 1992). It has been shown that transcription and replication occur in separate, but adjacent, sites within the RCs (A. Pombo *et al.*, 1994). Surrounding the PRZ there is an accumulation of components from the ICG of the host cell (Figure 13, panel II and Figure 14). Poly(A)-mRNAs are also found in these areas (Puvion-Dutilleul *et al.*, 1994). Therefore, it is suggested that the sites where clusters of ICG are recruited are sites where at least the start of viral late mRNAs post-transcriptional processing occurs.

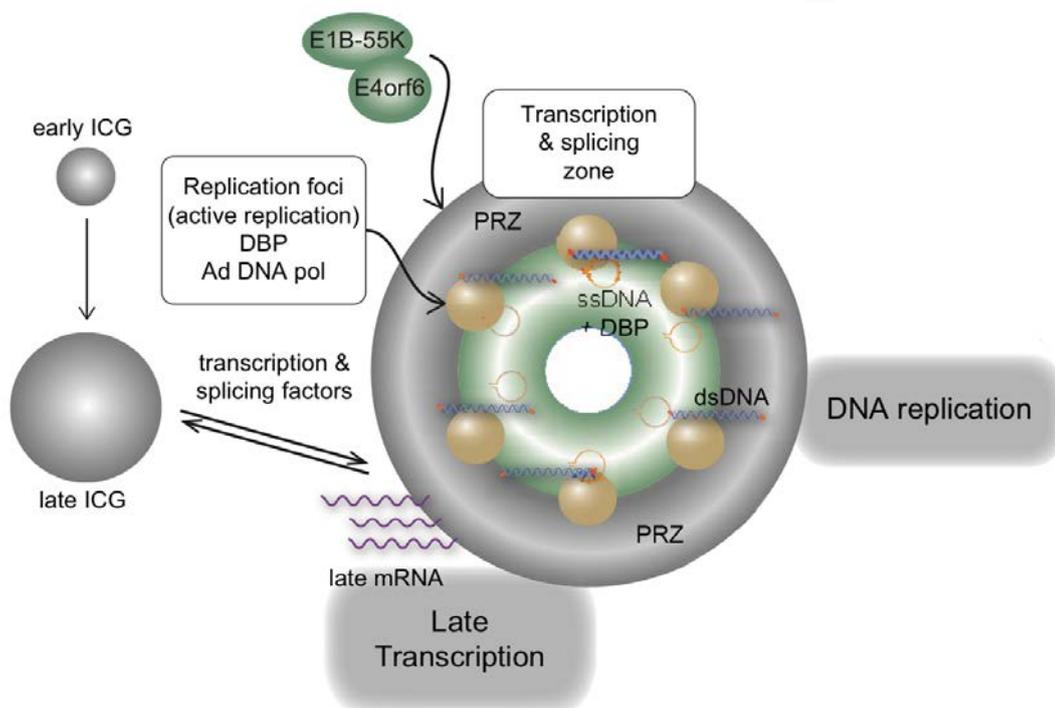


Figure 14. Replication Centers and associated functions. Scheme showing the distribution of the different functions associated to RCs, as well as localization of viral proteins that are recruited towards these sites (adapted from Schmid *et al.*, 2014).

1.7 mRNA processing and mRNA export of viral late mRNAs are linked to Replication Centers

In mammalian cells the mRNA splicing and export machinery are physically coupled to the mRNA during its splicing. Therefore, selection of the cellular mRNAs for export occurs during synthesis and processing events, resulting in the formation of a mRNPs (mRNA associated with proteins of the posttranscriptional machinery), which are selected by the export machinery (Rougemaille *et al.*, 2008). Normal bulk mRNA export occurs via the Aly/TAP pathway (Figure

15). During mRNA splicing, the TREX complex (formed by Aly, UAP56, and Hpr1) is assembled into the mRNAs. The adaptor protein Aly recruits the export receptor TAP (also known as Nxf1), which is part of the heterodimer TAP-p15, towards the processed mRNA (Katahira, 2015). SR proteins and components of the exon junction complex (EJC) also interact with TAP-p15 (Huang *et al.*, 2003). The Aly/TAP-p15 complex associates directly with the NPC and mediates the export of mature and fully processed mRNPs.

TAP-p15 has been reported to be the pathway used for export of adenoviral late mRNAs (Yatherajam *et al.*, 2011). As described previously, all viral late mRNAs require splicing processing machinery, much of which is distributed towards the RCs (see section 1.6, Figure 13 panel II). Also, selective export depends on the E1B/Orf6 complex but does not depend on any consensus sequence. The only factor associated with the selectivity for export during the late phase is that transcription takes place during this time (see section 1.5.3.1). These suggest that active transcriptional units during this phase occupy different nuclear microenvironments, specialized in facilitating the transport of mature mRNAs to the cytoplasm.

Crm1, also known as exportin-1, is involved in the export of rRNAs, snRNAs, and some specific cellular mRNAs (Figure 15). It is well studied that in infection with HIV-1, Crm1 can bind to the viral adaptor protein REV and export the unspliced viral mRNAs (Cullen, 2003). Crm1 itself is not an RNA binding protein, so it requires protein adaptors to transport mRNAs such as REV. It has been found that viral early mRNAs export is dependent on CRM1 (Schmid, R. Gonzalez, *et al.*, 2012).

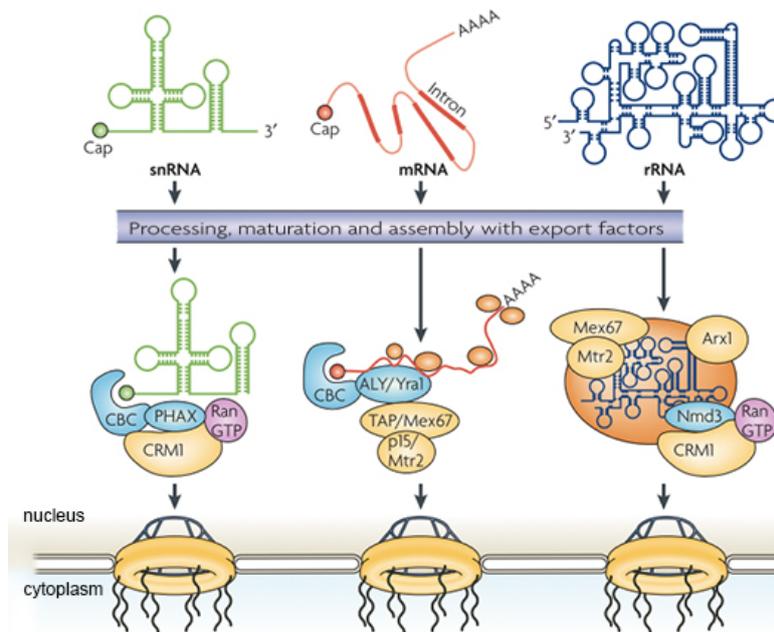


Figure 15. Cellular RNA export pathways. Representation of Tap and CRM1 mediated pathways. In each case, the primary RNA transcript is shown, as well as the transport-competent RNA after it has undergone processing, maturation and assembly with export factors (export adaptors are shown in blue, export receptors are shown in yellow) CBC, cap binding protein (adapted from Köhler & Hurt 2007).

The Heterogeneous Ribonucleoproteins are a group of around 25 different proteins named hnRNPs A-U. They are mainly nuclear proteins with a diffuse nucleoplasmic distribution. hnRNPs are grouped as a family of proteins not due to the similarity in sequence or domain composition, but because they were first isolated from nascent transcripts which are synthesized by RNA pol II. The hnRNP C protein has an N-terminal RNA-binding motif, and it is involved in splicing and packaging of nascent transcripts. The hnRNP L can play a role in alternative splicing, intron retention, suppression of multiple exons and alternative poly(A) site selection (Han, Tang and Smith, 2010). The hnRNP M is also involved in alternative splicing (Hovhannisyanyan & Carstens 2007). It is the only hnRNP in the nuclear pore complex (NPC), and it accompanies mRNAs to the nuclear envelope. AP5, which associates with E1B, is also an hnRNP protein named hnRNP U-like 1 (HNRPU1) and is known to interact with TAP (Bachi *et al.*, 2000; Barral *et al.*, 2005). The interaction between E1B and AP5 favors viral late mRNA export (Gabler *et al.*, 1998).

1.8 SUMOylation during HAdV infection

SUMOylation is a post-translational modification where small ubiquitin-like modifiers (SUMOs) are attached to lysine residues on proteins. Five mammalian SUMO genes (SUMO-1 to SUMO-5) have been found. SUMO-1, SUMO-2, and SUMO-3 (the last two are referred as SUMO-2/3

because of their very high homology) have been the most studied and SUMO-2/3 can form polymeric chains. SUMOylation requires an E1 enzyme which brings the SUMO proteins, and it is then passed to an E2 which is a conjugating enzyme (Ubc9). Finally, one of a small number of cellular E3 ligating proteins attaches it to the protein. SUMOs are mainly localized in the nucleus and SUMOylation is implicated in regulating several cellular processes and particularly nuclear events.

During infection with HAdV, the host SUMOylation system is regulated by some of the early proteins (E1A, E1B, and E4 Orf3) (Sohn and Hearing, 2016). The E1A protein can bind to Ubc9 and inhibit SUMOylation of pRb (Figure 16, panel A) (Ledl, Schmidt and Müller, 2005). E4 Orf3 induces transient SUMOylation of cellular proteins involved in DNA damage response, such as Mre11, and Nbs1, before their E1B/Orf6-mediated proteasome degradation (Figure 16, panel C) (Sohn and Hearing, 2012). However, the E1B is the only viral protein known to be a substrate for SUMOylation, which takes place in the lysine 104 (Endter *et al.*, 2001). It is known that phosphorylation of E1B at its C-terminus enhances its own SUMOylation (Wimmer *et al.*, 2013). This modification affects the E1B interaction with other proteins, its subcellular localization, its transformation ability, and its ubiquitin E3 ligase activity (described in section 1.4, Figure 9).

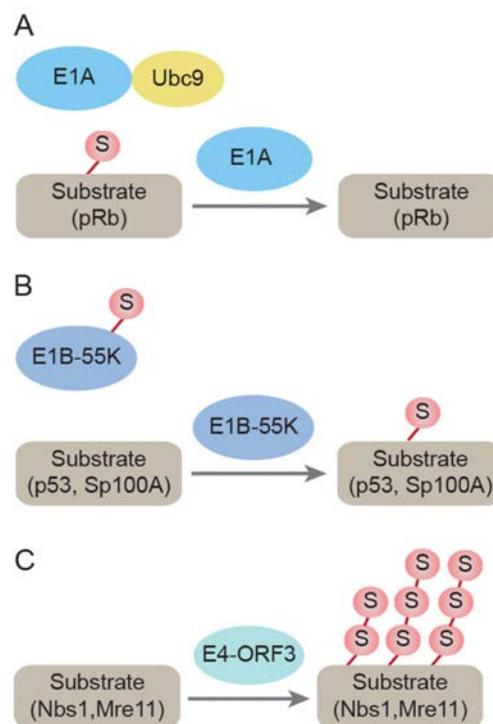


Figure 16. HAdV proteins that regulate SUMOylation during infection. Representation of the interplay between HAdV early proteins and the host SUMOylation system. A) E1A binds to Ubc9 inhibiting SUMO conjugation of pRB. B) E1B is both a SUMO substrate and a E3 ligase for p53 and Sp100A. C) Orf3 induces SUMOylation of Nbs1 and Mre11 (adapted from Sohn & Hearing 2016).

The SUMOylation of E1B is required to interact with components, such as PML-V and Sp100A, and to target Daxx for degradation (described in section 1.4.3) (Wimmer *et al.*, 2010; Schreiner *et al.*, 2011; Berscheminski *et al.*, 2016). Therefore, E1B-SUMO appears to be crucial for E4 Orf6-independent degradation. SUMOylation of E1B is also important for its subcellular localization. During infection E1B has a diffuse nuclear distribution that overlays with the ND10s and small perinuclear aggregates. As virus replication progresses into the late phase E1B localizes towards the RCs and accumulates at perinuclear aggresomes at the MTOC. However, when the lysine 104 of the E1B is mutated, inhibiting SUMOylation, the E1B is mainly accumulated in perinuclear aggregates (Kinds Müller *et al.*, 2007). E1B is not only a substrate for SUMOylation, but is also a SUMO E3 ligase and interacts with Ubc9 (Wimmer *et al.*, 2013). The best characterized SUMOylation target of E1B is p53 and is required for E1B-mediated inhibition of p53 transcriptional activation (Muller and Dobner, 2008; Pennella *et al.*, 2010). However, E1B-SUMO is also reported to stimulate SUMO-2 conjugation of Sp100A (Berscheminski *et al.*, 2016).

2 Project Aims

In this work, we attempt to search for possible cellular proteins involved in mRNA biogenesis, that could be degraded by the E1B/Orf6/E3 ubiquitin ligase, thereby, regulating viral late mRNA selective export. As well, we address the concept of selective export in a global approach by using Next Generation Sequencing technologies. We study both viral and cellular mRNA changes of cyto/nuc ratios and total expression values during infection.

3 Materials and Methods

3.1 Materials

3.1.1 Cells and Viruses

Cells:

Cell line	Genotype
A549	Human lung carcinoma cell line expressing wild-type p53 (Giard <i>et al.</i> , 1973).
HEK-293	Established HAdV-C5-transformed, human embryonic kidney cell line stably expressing the adenoviral E1A and E1B oncoproteins (Graham <i>et al.</i> , 1977).
H1299	Human lung carcinoma cell line, p53 negative (Mitsudomi T <i>et al.</i> , 1992).

Viruses:

Adenovirus	Characteristics
H5pg4100	Wt Ad5 containing an 1863 bp deletion (nt 28602-30465) in the E3 region (Kindsmüller <i>et al.</i> , 2007).
H5pm4101	E1B-55K mutant carrying three aa exchanges (L83/87/91A) within the NES region of E1B-55K (Kindsmüller, 2006).
H5pm4102	Ad5 E1B-55K mutant (K104R) containing an aa exchange in the SUMOylation site in position 104 (Kindsmüller, 2006).
H5pm4103	Ad5 E1B-55K mutant containing four aa changes (L83A/L87A/L91A/K104R) in the NES region and SUMOylation site (Kindsmüller, 2006).
H5pm4149	Ad5 E1B-55K null mutant containing four stop codons at the aa positions 3, 8, 86 and 88 of the E1B-55K sequence (Kindsmüller <i>et al.</i> , 2007).
H5pm4174	Ad5 E1B-55K mutant containing changes three aa changes (S490-491A/T495A) in phosphorylation sites 490, 491 and 495.
H5pm4219	Ad5 E1B-55K mutant containing three aa changes (A490-491D/T495D) a deletion in nt 2334/2335 of viral polymerase and a deletion in nt 760 of viral polymerase.

3.1.2 Antibodies

Primary antibodies:

Name	Properties	Source
2A6	Monoclonal mouse Ab; against N-terminus of HAdV-C5 E1B-55K (Sarnow <i>et al.</i> , 1982).	Group database
6B10	Monoclonal rat Ab; against HAdV-C5 L4-100K.	Group database
6His	Monoclonal mouse Ab; against 6xHis-tag.	Clontech
Aly/REF (11G5)	Monoclonal mouse Ab; raised against Aly/REF.	Santa Cruz
β -actin (AC-15)	Monoclonal mouse Ab; against β -actin.	Sigma Aldrich
B6-8	Monoclonal mouse Ab; against HAdV-C5 E2A protein (Reich <i>et al.</i> , 1983).	Group database
Crm1/Exp-1	Mouse Ab; against human hCRM1 aa 2-122.	BD Transduction
DO-1	Monoclonal mouse Ab; against the N-terminal aa 11-25 of human p53.	Santa Cruz
E1B-AP5	Monoclonal rat Ab; against human HNRPIL1.	NA
E2A	Polyclonal rabbit Ab; against HAdV-C5 E2A-72 kDa protein.	kindly provided by R. T. Hay
eIF4E	Monoclonal mouse Ab; raised against rabbit eIF-4E aa. 1-217.	BD Transduction
hnRNP L (4D11)	Monoclonal mouse Ab; raised against human hnRNP proteins.	Santa Cruz
hnRNP M3/4 (2A6)	Monoclonal mouse Ab; raised against fusion protein consisting of entire M4 protein sequence.	Santa Cruz
Mre11	Polyclonal rabbit Ab; against human Mre11.	Abcam/Novus
TAP	Polyclonal goat antibody raised against a N-terminal peptide of TAP.	Santa Cruz

Secondary antibodies:**Antibodies for Western Blotting**

Product	Properties	Company
HRP-Anti-Mouse IgG	HRP (horseradish peroxidase)-coupled; raised in sheep.	Jackson
HRP-Anti-Rabbit IgG	HRP (horseradish peroxidase)-coupled; raised in sheep.	Jackson
HRP-Anti-Rat IgG	HRP (horseradish peroxidase)-coupled; raised in sheep.	Jackson
HRP-Anti-Goat IgG	HRP (horseradish peroxidase)-coupled; antibody raised in sheep.	Jackson
HRP-Anti-Mouse IgG light chain specific	HRP (horseradish peroxidase)-coupled; raised in sheep.	Jackson

Antibodies for immunofluorescence staining

Product	Properties	Company
Alexa 488 Anti-Mouse IgG	Alexa 488 antibody raised in goat (H + L; F(ab') ₂ Fragment).	Invitrogen
Cy3-Anti-Rabbit IgG	Affinity purified, Cy3-coupled; raised in donkey; (H + L).	Dianova
FITC-Anti-Mouse IgG	Fluorescein-isothiocyanate (FITC)-coupled antibody raised against mouse IgGs in donkey; (H + L).	Dianova

3.1.3 Standards and markers

Product	Company
PageRuler Plus Prestained Protein Ladder	Pierce

3.1.4 Commercial systems

Product	Company
ScriptSeq™ v2 RNA-Seq Library Preparation Kit	Epicentre
High Sensitivity DNA Analysis Kit	Agilent
NEBNext® Poly(A) mRNA Magnetic Isolation Module	New England Biolabs
RNeasy Mini Kit	Qiagen
RNA Nano Kit	Agilent
Trizol® Reagent	Invitrogen

3.1.5 Chemicals, enzymes, reagents and equipment

Chemicals, enzymes, and reagents used in this study were purchased from Agilent, AppliChem, Biomol, Epicentre, Illumina, Merck, New England Biolabs, Quiagen, Roche and Sigma Aldrich. Cell culture materials as well as other plastic material and equipment were obtained from BioRad, Biozym, Brand, Engelbrecht, Eppendorf GmbH, Falcon, Gibco BRL, Greiner, Hartenstein, Hellma, Nunc, Pan, Peqlab, Sarstedt, Schleicher & Schuell, VWR and Whatman.

3.1.6 Software and databases

Software	Purpose	Source
Acrobat Pro DC	PDF data processing	Adobe
CLC Genomics Workbench 9.0	Bioinformatics and data processing	CLC bio
DAVID Bioinformatics Resources 6.8	Gene functional annotations	Open database (by LHR Leidos Biomedical Research, Inc.) 1
Excel 2016	Data analysis and graphics	Microsoft

Filemaker Pro 11	Database management	FileMaker, Inc.
Mendeley Desktop	Reference management	Free software (by Elsevier)
NetWalker 2	Gene clusters	Free software (by Vanderbilt University)
Illustrator CC 2018	Layout processing	Adobe
PANTHER 13.0	Gene Ontology	Open protein classification system (by GO Consortium)
Photoshop CC 2018	Image processing	Adobe
PubMed	Literature database, Adenoviral annotation reference	Open database (by NCBI)
R Studio	Data analysis and graphics	Open-source software
Word 2016	Text processing	Microsoft

3.2 Methods

3.2.1 Tissue culture

Maintenance of cell lines. A549 cells were cultured as monolayers in polystyrene cell culture dishes with *Dulbecco's Modified Eagles Medium* (DMEM; Sigma) containing 0.11 g/l sodium pyruvate, 10% FCS (Pan) and 1% of penicillin/ streptomycin solution (1000 U/ ml penicillin and 10 mg/ ml streptomycin in 0.9 % NaCl; Pan). Cells were cultured at 37°C in a CO₂ incubator (Heraeus) in 5% CO₂ atmosphere.

Passage of cell lines. All tissue culture techniques were performed under sterile conditions inside specialized flow hoods. To split or maintain cell lines, the existing medium was removed and the cell monolayer was washed twice with phosphate buffered saline (PBS). Cells were detached from the culture dish by adding enough trypsin/ EDTA (Pan) to cover the monolayer and this was incubated for 3-5 min at 37 °C. Subsequently, the trypsin activity was inactivated by adding standard culture medium, using the same volume used of trypsin (1:1 ratio). The detached cells were transferred into a 50 ml tube and centrifuged at 2,000 rpm for 3 min (*Multifuge 3S-R*, Heraeus). The remaining pellet of cells was resuspended in the desired volume of culture medium to either split or maintain the cell culture. If cells were split to propagate

them, then they were divided into the required number of cell dishes, always staying in the limits of ratios 1:5-1:20. If cells were used for an experiment, they were counted before seeding to achieve the desired concentration of cells. When cells were frozen for preservation the procedure below was performed (see Cryopreservation of cell lines).

PBS	140 mM	NaCl
	3 mM	KCl
	4 mM	Na ₂ HPO ₄
	1.5 mM	KH ₂ PO ₄
		(pH 7.0-7.7; autoclaved)

Counting and seeding cells. Cells in suspension were counted using a *Neubauer cell counter* (C. Roth). A mixture of suspended cells with trypan blue solution was prepared in a 1:1 ratio. To count the cells, 10 µl of the cell mixture was pipetted into the chamber. The cells were counted using a light microscope (Leica DM IL) and the final number was multiplied by a factor of 10⁴ to obtain the number of cells in 1 ml suspension. The desired dilution was calculated to seed the number of cells needed for each experiment (which is specified in each figure).

$$\text{cell number/ml} = \text{counted cell} \times 2 \text{ (dilution factor)} \times 10^4$$

Trypan blue solution	0.15 %	Trypan Blue
	0.85 %	NaCl

Cryopreservation of cell lines. After trypsination, cells were pelleted by centrifugation at 2,000 rpm for 3 min. The pellet of cells was then resuspended in pure FCS, supplemented with 10% dimethyl sulfoxide (DMSO) and aliquoted into *CryoTubes*TM (Nunc). The aliquoted cells were gradually cooled using a *Mr. Frosty* (Nalgene) to -80 °C overnight before transferring them into liquid nitrogen storage. For re-cultivation, cells were thawed to 37 °C using a water bath and afterwards they were pelleted to remove the DMSO. The cells were resuspended with 1 ml of fresh culture medium and seeded into cell culture dishes to which rest of the appropriate volume of culture medium was added.

Harvest of mammalian cells. For protein experiments a standard harvesting procedure was performed: The monolayer of cells was detached using scrapers (Sarstedt) keeping the

medium and collected into 15 or 50 ml conical tubes. Cells were centrifuged at 2,000 rpm for 3 min at RT (*Multifuge 35-R*; Heraeus). The supernatant was discarded and the pellet was washed once with PBS and then stored at $-20\text{ }^{\circ}\text{C}$ for following experiments. When cells were needed intact for cell fractionation protocols, they were harvested by trypsination: The medium was discarded and cells were washed with PBS. Enough trypsin/ EDTA (Pan) was added to cover the monolayer and this was incubated for 3-5 min at $37\text{ }^{\circ}\text{C}$. Subsequently, the trypsin activity was inactivated by adding standard culture medium, using the same volume used of trypsin (ratio 1:1). Cells were collected into 15 or 50 ml conical tubes and washed three times with cold PBS by centrifugation (2,000 rpm, 3 min, $4\text{ }^{\circ}\text{C}$, *Multifuge 35-R*; Heraeus). Pelleted cells were kept on ice until next procedure.

3.2.2 Infection with adenovirus

Infection of A549 cells with adenovirus. A549 cells were seeded into the desired culture plate size 24 h before infection, so they could reach a confluency of $\sim 80\%$ at the time of infection. Before infection, the medium was removed and cells were washed once with PBS to remove traces of serum. The virus dilutions were prepared using DMEM (a small volume enough to cover the monolayer surface) without supplements and this was added into the culture. The volume of virus stock added into the cells was calculated according to this formula:

$$\text{volume virus stock solution } (\mu\text{l}) = \frac{\text{multiplicity of infection (MOI)} \times \text{total cell number}}{\text{virus titer (focus forming units (ffu)/}\mu\text{l)}}$$

The virus dilution was left incubating on cells at $37\text{ }^{\circ}\text{C}$ for a period of 2 h. Afterwards, the virus was removed and fresh standard culture medium was added. Infected cells were left incubating at $37\text{ }^{\circ}\text{C}$ for the stipulated time of each experiment.

Propagation and storage of high-titer virus stock. Cell culture dishes (150 mm) with HEK-293 cells at 60% confluency were infected with established laboratory virus stocks at an MOI of 20 ffu/ cell, as described above. The infected cells were harvested 3-5 days after infection using cell scrapers (Sarstedt) and were collected into 15 or 50 ml conical tubes. Suspended cells were centrifuged at 2,000 rpm for 5 min at room temperature (RT) (*Multifuge 35-R*; Heraeus). The supernatant was discarded and the pellet was washed once with PBS and resuspended in DMEM without supplements, using $\sim 1\text{ ml/ } 150\text{ mm}$ dish. The viral particles were released by repeated cycles of freezing and thawing using liquid nitrogen and a water bath at $37\text{ }^{\circ}\text{C}$ (3-5 cycles). The cell debris was pelleted by centrifuging at 4,500 rpm for 10 min (*Multifuge 35-R*; Heraeus). The supernatant containing the virus was then mixed with an 87% sterile glycerol

solution to achieve a final concentration of 10% glycerol. Virus stocks were preserved at -80 °C until usage.

Titration of virus stocks. To determine the titer of new virus stocks, immunofluorescence assays using antibodies to stain against adenoviral protein E2A 72K (DBP) were performed (Reich *et al.*, 1983). From the virus stock, 1 ml aliquot dilutions were made from 10^2 - 10^6 . HEK-293 cells were seeded into a 6-well dish (5×10^5 cells/well) and infected with 1 ml of the virus stock dilutions. After 24 h of infection, cells were fixed with 1 ml ice-cold methanol and incubated during 15 min at -20 °C. The methanol was removed and cells were air-dried at RT and then incubated with 1 ml of PBS-Triton during 15 min. Afterwards, the PBS-Triton was removed and cells were blocked with 1 ml TBS-BG during 1 h at RT. The blocking solution was removed and then the primary antibody (B6-8) diluted 1:10 in TBS-G was incubated during 2 h at RT. The primary antibody was removed and cells were washed three times for 15 min with TBS-BG. The Alexa Fluor®488-coupled secondary antibody (Invitrogen), in a 1:500 dilution with TBS-BG, was incubated during 2 h at RT. The secondary antibody was removed and cells were washed three times with TBS-BG during 15 min. Stained cells were counted using a fluorescence microscope (Leica). The total number of infectious particles was determined from the infected cell number, the virus dilution, and the microscope magnification used to count, which resulted in fluorescence forming units (ffu).

TBS-BG	20 mM	Tris-HCl (pH 7.6)
	137 mM	NaCl
	3 mM	KCl
	1.5 mM	MgCl ₂
	0.05 % (v/v)	Tween20
	0.05 % (w/v)	Sodium azide (NaN ₃)
	5 % (w/v)	Glycine
	5 % (w/v)	BSA

3.2.3 SUMO pull-down assay

For this assay, H1299 cells that were transiently transfected with p6His-SUMO-1 or p6His-SUMO-2 were infected with wt of different HAdV-5 mutants. For every experimental time-point, two culture dishes with 4×10^6 cells were utilized. Cells were harvested by scrapping after the desired post-infection time point as described above (section 3.2.1). From the total harvested cells, 20% of the sample was saved to use as an input control and was lysed normally with RIPA buffer (described in section 3.2.5). The rest of the sample was resuspended in 5 ml of

guanidinium lysis buffer and then lysed by sonication. Separately, Ni-NTA agarose beads (Qiagen) were washed three times with 5 ml of guanidinium lysis buffer. The lysates were then incubated with 25 μ l/ sample of the prewashed beads over night at 4 °C. Samples were centrifuged during 10 min at 4,000 rpm and the supernatant was discarded carefully to not touch the beads, which stayed loose. The beads were washed twice with 1 ml of wash buffer pH 8.0 and three times with wash buffer pH 6.3. The His-SUMO conjugates were eluted from the agarose beads with 30 μ l of elution buffer. Subsequently, samples were boiled at 95 °C for 5 min. Finally, eluted proteins were separated by SDS-PAGE (see 3.2.6) and visualized by immunoblotting (see 3.2.7).

Guanidinium lysis buffer	6 M	Guanidinium-HCl
	0.1 M	Na ₂ HPO ₄
	0.1 M	NaH ₂ PO ₄
	10 mM	Tris-HCl (pH 8.0)
	20 mM	Imidazole
	5 mM	β -Mercaptoethanol
Wash buffer pH 8.0	8 M	Urea
	0.1 M	Na ₂ HPO ₄
	0.1 M	NaH ₂ PO ₄
	10 mM	Tris-HCl (pH 8.0)
	20 mM	Imidazole
	5 mM	β -Mercaptoethanol Protease inhibitors
Wash buffer pH 6.3	8 M	Urea
	0.1 M	Na ₂ HPO ₄
	0.1 M	NaH ₂ PO ₄
	10 mM	Tris-HCl (pH 6.3)
	20 mM	Imidazole
	5 mM	β -Mercaptoethanol Protease inhibitors
Elution buffer	200 mM	Imidazole
	0.1 % (w/v)	SDS
	150 mM	Tris-HCl (pH 6.8)
	30 % (v/v)	Glycerol
	720 mM	β -Mercaptoethanol
	0.01 % (w/v)	Bromphenol blue

3.2.4 Cycloheximide assay

Per each post-infection time point required, 10 culture dishes were seeded with A549 cells (2.5×10^6). Afterwards, cells were infected as described above. After the post-infection time had passed, the cycloheximide treatment was performed. If several time-points were used, cells were synced so that the cycloheximide treatment could be performed at the same time.

Cycloheximide treatment. Medium was removed from dishes and a mixture of 5 ml of cycloheximide (at a final concentration of 100 $\mu\text{g}/\text{ml}$) with 5 ml of standard culture medium was added. Cells were returned to the incubator and immediately proceeded to harvest cells by scrapping (section 3.2.1, Harvest of mammalian cells), at 0 (without treatment), 5 min, 10 min, 20 min, 30 min, 1 h, 4 h, 6 h, 8 h, and 24 h after adding the cycloheximide. Samples were kept on ice while collecting all samples and then proceeded to prepare protein cell lysates as described below.

3.2.5 Preparation of protein cell lysates

After harvesting (section 3.2.1, Harvest of mammalian cells), cells were pelleted and frozen or immediately proceeded to prepare protein cell-lysates. The cell pellets were resuspended in $\sim 100\text{-}300\ \mu\text{l}$ of RIPA lysis buffer with freshly added 0.2 mM PMSF, 0.001 mM DTT, 1 mg/ml pepstatin A, 5 mg/ml aprotinin, and 20 mg/ml leupeptin. Samples were incubated on ice for 30 min and vortexed every 10 min in between. To completely lyse cells and shred the genomic DNA, samples were sonicated (3 times, 30 secs, output 0.60; 0.8 Impulse/s; *Branson Sonifier 450*). The remaining debris and insoluble material was pelleted by centrifugation (11,000 rpm, 5 min, 4 °C; *Eppendorf 5417R*). The soluble sample was transferred into a new tube and protein concentration was determined by spectrophotometry, as described below (Quantitative determination of protein concentration). The protein lysates were stored at $-20\ \text{°C}$ until further use.

RIPA lysis buffer

50 mM	Tris-HCl (pH 8.0)
150 mM	NaCl
5 mM	EDTA
1 % (v/v)	Nonidet P-40
0.1 % (w/v)	SDS
0.5 % (w/v)	Sodium Desoxycholate

Quantitative determination of protein concentration. The *Protein-Assays* (BioRad), according to Bradford, were used to measure the protein concentration of the samples (Bradford, 1976).

From each protein lysate, 1 μ l was taken and mixed with 800 μ l of ddH₂O and 200 μ l of *Bradford Reagent* (BioRad). The mixture was incubated for 5 min at RT and measured in a *SmartSpec Plus spectrophotometer* (BioRad) at a wave length of 595 nm against a blank. A standard curve performed with increasing BSA concentrations (1-16 μ g/ μ l; New England Biolabs) was used to determine the final protein concentrations by interpolation.

3.2.6 SDS polyacrylamide gel electrophoresis (SDS-PAGE)

Protein samples from cell lysates of time course infections (see 3.2.2) and SUMO pull-down assays (see 3.2.3) were separated according to their molecular weights by SDS-PAGE (Biometra). The polyacrylamide gels were made using 30% acrylamide/bisacrylamide solution (37.5:1 *Rotiphorese Gel 30*; Roth) diluted to the final concentration of 10-15% with ddH₂O. APS ($f_{\text{final}}=0.1\%$) and TEMED ($f_{\text{final}}=0.01\%$) were added to accelerate the acrylamide polymerization. The polyacrylamide gels were prepared using the *Multigel* SDS-PAGE system (Biometra) according to the manufacturer's instructions. A low pH and low acrylamide concentration stacking gel (to which protein samples were loaded) was added, on top of the separating gel, to allow stacking of proteins by pH before entering the separating gel. Before loading, samples were adjusted to the same concentration and prepared by adding 2 x or 5 x SDS sample buffer, to achieve a final concentration of 1 x (Sambrook, Fritsch and Maniatis., 1989). Subsequently, the samples were boiled at 95 °C during 3 min using a thermoblock (*Thermomixer Comfort*; Eppendorf) and the same amount of protein from each sample was loaded into the gels. To determine the proteins' weights in the gel, the *Page Ruler™ Prestained Protein Ladder Plus* (Fermentas) was also loaded. Gels were ran in TGS-buffer at 10 mA/ gel through the stacking gel and at 20 mA/gel through the separating gel.

5 x SDS sample buffer	100 mM	Tris-HCl (pH 6.8)
	10 % (w/v)	SDS
	200 mM	DTT
	0.2 % (w/v)	Bromophenol blue
2 x SDS sample buffer	100 mM	Tris-HCl (pH 6.8)
	4 % (w/v)	SDS
	200 mM	DTT
	0.2 % (w/v)	Bromophenol blue
	20 %	Glycerol

5 % stacking gel	17 % (v/v)	Acrylamide solution (30 %)
	120 mM	Tris-HCl (pH 6.8)
	0.1 % (w/v)	SDS
	0.1 % (w/v)	APS
	0.1 % (v/v)	TEMED
10 % separating gel	34 % (v/v)	Acrylamide solution (30 %)
	250 mM	Tris-HCl (pH 8.8)
	0.1 % (w/v)	SDS
	0.1 % (w/v)	APS
	0.6 % (v/v)	TEMED
12 % separating gel	40 % (v/v)	Acrylamide solution (30 %)
	250 mM	Tris-HCl (pH 8.8)
	0.1 % (w/v)	SDS
	0.1 % (w/v)	APS
	0.6 % (v/v)	TEMED
15 % separating gel	50 % (v/v)	Acrylamide solution (30 %)
	250 mM	Tris-HCl (pH 8.8)
	0.1 % (w/v)	SDS
	0.1 % (w/v)	APS
	0.6 % (v/v)	TEMED
TGS buffer	25 mM	Tris
	200 mM	Glycine
	0.1 % (w/v)	SDS

3.2.7 Western Blot

Protein transfer. SDS-PAGE were transferred onto nitrocellulose (*Whatman*) using the *Trans-Blot Electrophoretic Transfer Cell System (BioRad)* in Towbin buffer. In the case of proteins smaller than 20 kDa, polyvinylidene fluoride (PVDF) membranes were used instead. The PVDF membranes were activated by soaking them in methanol before using. Afterwards, gels and membranes were soaked in Towbin buffer and placed between two blotting papers (*Whatman*) on both sides of the transfer sandwich. This was placed between two blotting pads in a plastic grid. The electrophoretic transfer was performed in “full wet” mode in a blotting tank filled with Towbin buffer at 400 mA during 90 min.

Towbin buffer	25 mM	Tris-HCl (pH 8.3)
	200 mM	Glycine
	0.05 % (w/v)	SDS
	20 % (v/v)	Methanol
PBS-Tween	0.1 % (v/v)	Tween20 in 1x PBS

Antibodies incubation. Before incubating with antibodies, membranes with transferred proteins were blocked to avoid unspecific antibody binding on the nitrocellulose membrane. Therefore, they were incubated with PBS-Tween 5% non-fat milk powder (*Frema*) for at least 2 hours at RT or overnight at 4 °C on an *orbital shaker* (GFL). Afterwards, the blocking solution was discarded and membranes were washed briefly with PBS-Tween to remove the excess blocking solution. The primary antibodies were diluted in PBS-Tween and incubated during 2 h at RT on an *orbital shaker* (GFL). The dilutions of primary antibodies as well as the amount of non-fat powder milk (*Frema*) added were determined individually for each antibody. After the incubation period, the primary antibody dilution was removed and the nitrocellulose membranes were washed three times for 10 min with PBS-Tween. The membranes were then incubated with the HRP-coupled secondary antibody, diluted with PBS-Tween (1:10,000; *Amersham*) and with 3 % non-fat powder milk (*Frema*), during 2 h at RT. The secondary antibody dilution was removed and the nitrocellulose membranes were washed three times for 10 min with PBS-Tween. Protein bands were visualized by enhanced chemiluminescence using *SuperSignal West Pico Chemiluminescent Substrate* (*Pierce*) according to the manufacturer's instructions. Bands were detected using X-ray films (*RP New Medical X-Ray Film*; CEA) and a *GBX Developer* (*Kodak*). The X-ray films were scanned and figures were prepared using *Illustrator CC 2018* (*Adobe*).

3.2.8 Immunofluorescence

For indirect immunofluorescence analysis, A549 cells were grown in 12-well dishes on glass coverslips (1×10^5 cells/ well) and infected as previously described (see 3.2.2). When several infection time points were required, infections were synchronized to collect all samples at the same time. After the period of infection, cells were washed once with PBS with and fixed with 3.7% PFA during 20 min at RT. Afterwards, cells were permeabilized with PBS/ 0.5 % (v/v) Triton X-100 for 5 min at RT with slow shaking. Detergent was washed three times with PBS during 5 min at RT and slow shaking. The primary antibody diluted in PBS, was added on top of the coverslips and incubated during 4 h in a wet chamber at 4 °C. The primary antibody was washed three times with PBS during 5 min at RT and slow shaking. The corresponding secondary antibodies were diluted in PBS and incubated during 2 h at 4 °C. Afterwards, coverslips were washed three times with PBS, and incubated with DAPI (diluted 1:10,000 in PBS) during 5 min at RT. DAPI was washed two times with PBS during 5 min at RT and slow shaking. Finally, coverslips were mounted into glass slides with *Glow Mounting Media* (*EnerGene*). Slides were stored at -20 °C until they were analyzed. Digital images were acquired with a *DM6000 fluorescence microscope* (*Leica*) using a charge-coupled device camera (*Leica*) and final figures were assembled using *Illustrator CC 2018* (*Adobe*).

3.2.9 Cytoplasmic and nucleoplasmic fractionation for RNA extraction

Cytoplasmic and nucleoplasmic isolation. Mock and adenovirus infected A549 cells (2.5×10^6) were harvested by trypsination (see section 3.2.1). Cells transferred into a reaction tube (*Eppendorf*) and centrifuged at 4,000 rpm for 3 min at 4 °C (*Eppendorf 5417R*). Pelleted cells were resuspended in 100 μ l of cold NSP40 buffer, incubated on ice for 2 min and centrifuged (470 g, 5 min, 4 °C; *Eppendorf 5417R*). The supernatant (~80 μ l) was transferred into a new tube with 600 μ l of *Trizol*® (Invitrogen) and saved as the cytoplasmic fraction. The remaining nuclear pellets are washed with 100 μ l of cold NP40 buffer and centrifuged (470 g, 5 min, 4 °C; *Eppendorf 5417R*). The supernatant was discarded and the pellet containing the nuclei was resuspended in 600 μ l of *Trizol*® (Invitrogen) and saved as the nucleoplasmic fraction.

NP40 buffer	10mM	Hepes (pH 7.8)
	10 mM	KCl
	20% (v/v)	glycerol
	0.25% (v/v)	NP40
	1 mM	DTT (before using)

RNA extraction. Samples were incubated in *Trizol*® (Invitrogen) during 5 min at RT. Afterwards, 120 μ l of chloroform were added, tubes were shaken together during 15 secs and incubated during 2-3 min at RT. Tubes were centrifuged at 12,000 g during 15 min at 4 °C (*Eppendorf 5417R*). Three phases resolve and the upper aqueous phase (~320 μ l) was transferred into a new tube. The RNA was precipitated by adding 1.5 times the volume (~480 μ l) of 100% ethanol and pipetting up and down to mix the sample. From this mixture, 700 μ l were transferred to an *RNeasy Mini spin column* (*RNeasy Mini Kit*, *Quiagen*) and the sample was centrifuged 15 secs (8,000 g, 4 °C; *Eppendorf 5417R*) and the flow was discarded. The same step was repeated with the remaining sample. Then, 350 μ l of Buffer RW1 (*RNeasy Mini Kit*, *Quiagen*) was added to the column and centrifuged (8,000 g, 15 secs, 4 °C; *Eppendorf 5417R*). The DNA was digested using RNase-free DNase I (1 U/ μ L, *Quiagen*); 80 μ l of the DNase I mixture (10 μ l of the DNase I was mixed with 70 μ l of the BDD buffer) was used per sample and incubated for 30 min at RT. Afterwards, 350 μ l of Buffer RW1 (*RNeasy Mini Kit*, *Quiagen*) was added to the column and incubated for 5 min before centrifuging and discarding the flow-through (8,000 g, 15 secs, 4 °C; *Eppendorf 5417R*). Then, 500 μ l of Buffer RPE RW1 (*RNeasy Mini Kit*, *Quiagen*) was added to the column and centrifuged (8,000 g, 15 secs, 4 °C; *Eppendorf 5417R*). Again, 500 μ l of Buffer RPE RW1 (*RNeasy Mini Kit*, *Quiagen*) was added to the column and centrifuged (8,000 g, 1 min, 4 °C; *Eppendorf 5417R*). The columns were transferred into a reaction tube (*Eppendorf*) and 60 μ l of RNase-free water was carefully added directly onto the column's membrane and incubated for 1 min at RT. Afterwards, tubes were centrifuged at maximum speed to elute the

RNA (1 min, 4 °C; *Eppendorf 5417R*). The final RNA concentration was measured with the *NanoDrop* spectrophotometer (*PEQLAB*; Erlangen) and the RNA samples were stored at -80 °C until further use.

3.2.10 Next Generation Sequencing (NGS)

Library preparation. The RNA isolated from both cytoplasmic and nucleoplasmic (see 3.2.9) were sent to the HPI's Technology Platform. There, the RNA quality was evaluated using a *Bioanalyzer* with the *RNA nano Chip* (Agilent Technologies). Afterwards, poly(A) mRNAs were purified with the *NEBnext® poly(A) mRNA Magnetic Isolation Module* (NEB sequencing). The cDNA libraries were generated with the *ScriptSeq v2 RNA Seq Kit* (Epicenter). The size and quality of the libraries were visualized with a *Bioanalyzer* and a High Sensitivity DNA Chip (High Sensitivity DNA Kit, Agilent Technologies).

Sequencing. Diluted libraries (2 mM) were multiplex-sequenced on the *HiSeq 2500* (Illumina), performing a paired end run (2x100 bp) with a depth of around 50 million reads per sample.

3.2.11 NGS data analysis

RNA-Seq analysis. The RNA-Seq and quality analysis were performed using the *CLC Genomics Workbench 9.0* software (CLCbio). To start, the human reference genome (hg19) was downloaded to the software (from PubMed) and the HAdV-5 reference genome was annotated manually (see Table 1). Afterwards, the gene and mRNA tracks of both genomes were generated. The fastq files, containing the raw sequencing data, were imported into software. As multiple fastq files were generated for a single biological sample from the sequencing (3.2.10), the RNA-Seq analysis for each sample was ran as a batch. Also, in the settings the strand specificity was set as "forward". The rest of the settings for the mapping were left as default and the option to calculate Reads Per Kilobase of transcript per Million mapped (RPKM) was selected. The resulting original expression values were subjected to a quantile normalization. The quality of the sequences was assessed with a principal component analysis, using both the original and normalized data. For every time-point an independent experiment between nucleoplasmic and cytoplasmic samples was set up (group 1= nucleoplasmic, group 2= cytoplasmic) and a proportions test (Baggerley) was performed to obtain the ratio between the cytoplasmic and nucleoplasmic data for each transcript. For this test, the normalized expression values were used and the NaN values were filtered. From this test, the Bonferroni and FDR corrected p-values were generated. Finally, the xls files for each time point of infection, with the cytoplasmic and nucleoplasmic data, was exported. Each file

contained the assigned reads to each mRNA of the reference genome that was used (human or viral), including the normalized values and the proportion test analysis. Data from these files was used for all further comparison analyses using the RPKMs as cytoplasmic and nucleoplasmic expression values and the ratios resulting from the Baggerley test as cytoplasmic to nucleoplasmic ratios. The total expression values for each time point were derived from the sum of cytoplasmic and nucleoplasmic expression values. Plots were done using Excel 2016 (Microsoft) and the R package ggplot2 (Computing, 2011). Images of viral transcripts (Figure 23), alignment of reads through the viral transcriptome (Figure 24) and representation of HAdV-5 transcript expression (Figure 31) were also generated from the CLC Genomics Workbench 9.0 software (CLCbio). Final figures were assembled and adapted using *Illustrator CC 2018 (Adobe)*.

Clustering and heat map of cyto/nuc ratios. The cytoplasmic to nucleoplasmic ratios were used to generate clusters and the values were represented as a heat map using the R package gplots (Computing, 2011). Values were z-score transformed and clustered by row (by transcripts, not time points). The hierarchical clustering method used was complete agglomeration and the metric used was Euclidean distance. Final figures were assembled and adapted using *Illustrator CC 2018 (Adobe)*.

Functional networks and classification. The functional networks were obtained from the gene names associated to each of the selected clusters using the NetWalker2 free software. The functions associated to the genes of each selected clusters were assessed using the DAVID online platform (Huang, Sherman and Lempicki, 2009a, 2009b). The list of genes was browsed for "gene functional classification" and the identifier was set as "OFFICIAL_GENE_SYMBOL". The list was compared to Homo Sapiens and the final classification stringency was set as low as the list of genes was short. The final figures were assembled and adapted using *Illustrator CC 2018 (Adobe)*.

4 Results

4.1 Analysis of cellular proteins involved in mRNA processing and export pathways during infection

As described in the introduction, in the absence of an active E1B/Orf6/E3 ubiquitin ligase, viral late mRNA export and viral late protein synthesis are less efficient (Woo and Berk, 2007). Therefore, it has been proposed that the E1B/Orf6/E3 ubiquitin ligase targets cellular components for degradation, required for cellular mRNA export and linked to mRNA maturation. Altering these cellular pathways could reduce the efficiency of production, processing or export of cellular mRNAs and thereby favor viral late mRNA processing and export during infection. Degradation of these components most likely would occur during late times of infection, since the absence of an active E3 ubiquitin ligase showed no effect on viral early mRNA export (Woo and Berk, 2007). However, it has been suggested that during the early phase viral mRNA export could take place through a different export pathway (Schmid, R. A. Gonzalez and Dobner, 2012). So far, none of the discovered E1B/Orf6/E3 ubiquitin ligase degradation targets have been found to participate in processes involving mRNA maturation or export, so the mechanism explaining how the viral complex impacts the viral late mRNA export efficiency remains unclear. Therefore, we chose to assess whether some of the key cellular components of mRNA splicing, maturation or export pathways are degraded during infection with adenovirus type 5.

4.1.1 Time-course analysis of cellular proteins involved in mRNA processing and mRNA export

To evaluate if the E1B/Orf6/E3 ubiquitin ligase degrades cellular mRNA biogenesis components, key cellular components involved in this process were selected, such as the RNA transporters TAP and Crm1, the adaptor protein Aly, and heterogeneous ribonucleoproteins that interact with TAP or are involved in mRNA splicing (hnRNP M, hnRNP L, and AP5), described in section 1.7. As described in the introduction, Crm1 is suggested to participate in viral early mRNA export and TAP/Aly in viral late mRNA export.

The first objective was to assess whether the steady state concentrations of any of the previously described proteins were affected during adenovirus infection. To answer this question, a time-course experiment using the human lung carcinoma cell-line A549, infected with a wt HAdV-5 was performed. Infected cells were harvested at early (6 and 12 hpi) and

late (24-72 hpi) time points post-infection (Figure 17). Samples were used to perform a Western Blot (WB) analysis to search for the presence of the previously described cellular proteins. Antibodies for viral proteins were also included in these experiments as bona fide markers of the transition from the early to the late phase of the infection cycle (E2-72K; DBP), and of the late phase and the start of viral late mRNA production (L4-100K). The presence or absence of these viral late proteins allows us to define whether the effects observed in our experiments correlate to the early or late phase of infection. As a positive control for viral induced degradation, Mre11 was also analyzed, since it is known to be a target of E1B/Orf6/E3 ubiquitin ligase mediated degradation.

Time-course experiments revealed that the signal for Aly decreased at late times of infection, but when normalized with the actin control this effect is very small. The WB signal for eIF4E, Tap and Crm1 remained constant throughout the infection (Figure 17; panel A, 72 hpi). However, at late time points a series of higher and lower migrating bands appeared for Tap as the infection progressed (Figure 17; panel A, 24-72 hpi). In the case of Crm1, it was observed that the band migrated with a slightly reduced molecular weight and lower migrating bands appeared at late time points (Figure 17; panel B, 24-72 hpi). This effect could be caused by proteolysis or degradation at late times of infection.

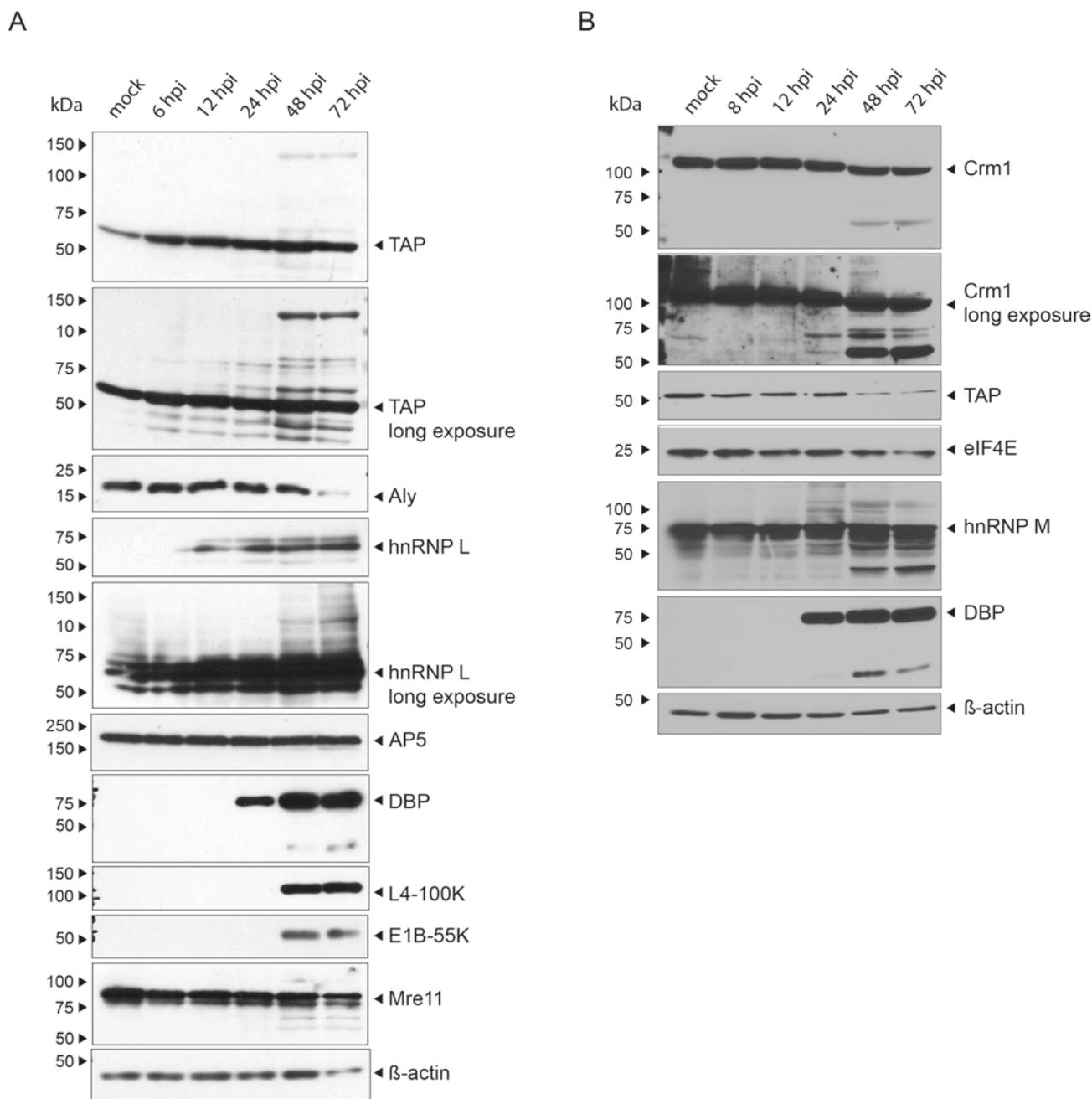


Figure 17. Time-course analysis of cellular proteins that participate in mRNA processing and export activities. WB of A549 cells (5×10^6) infected with wt HAAdV-5 (H5pg4100) at an MOI of 20 ffu/cell. Cells were harvested at 6, 12, 24, 48 and 72 hpi and whole cell protein lysates were prepared. Samples were resolved by 10% SDS-PAGE and cellular TAP, Aly, hnRNP L, hnRNP M, AP5, Crm1 and eIF4E proteins were visualized by immunoblotting. Viral proteins DBP, L4-100K, and E1B-55K were used as a control of infection. Cellular protein Mre11 was used as a control of degradation and β -actin as a protein concentration control.

The migrating patterns of the heterogeneous nuclear ribonucleoproteins M, L and AP5 in the time-course showed to be different during infection (Figure 17; panel A). AP5 remained constant at all time points. In contrast, hnRNP M levels increased at late time points (Figure 17; panel B, 48 and 72 hpi), also, lower and higher migrating bands appeared at 24 hpi. Similarly, hnRNP L levels increased during infection and some higher bands appeared at late time points (Figure 17; panel A 48 and 72 hpi).

In summary, in the steady state time-course analysis only Aly showed a small decrease in signal at very late time points of infection. In contrast, hnRNPs L and M had an increased signal after infection. However, TAP, Crm1 and hnRNP M showed lower migrating bands that appear at late time points of infection. In this analysis, none of the analyzed cellular proteins involved in biogenesis had an obvious effect of degradation which could be linked to the viral E1B/Orf6/E3 ubiquitin ligase complex. However, as the degradation control (Mre11) did not show an obvious effect in this steady state analysis, further stability assays were required.

4.1.2 Stability analysis of cellular proteins involved in mRNA processing and mRNA export

To analyze whether cellular proteins Aly, TAP, Crm1 and hnRNP M are degraded at late time points of infection, experiments designed to determine the stability of these proteins during infection were conducted (Figure 18). Cells were infected with wt HAdV-5 and treated with cycloheximide, during a period from 5 min to 24 h, after an early (12 hpi) and a late (48 hpi) time point post-infection. Cycloheximide is an inhibitor of translation in eukaryotes; therefore, this drug can be used to measure the half-life of a given protein and assess whether its stability changes during infection. As a control for degradation we also stained for cellular p53, which is known to be targeted for degradation by the viral E1B/Orf6/E3 ubiquitin ligase during the late phase of infection (Howley, Maki and Huibregtse, 1996; Querido *et al.*, 2001).

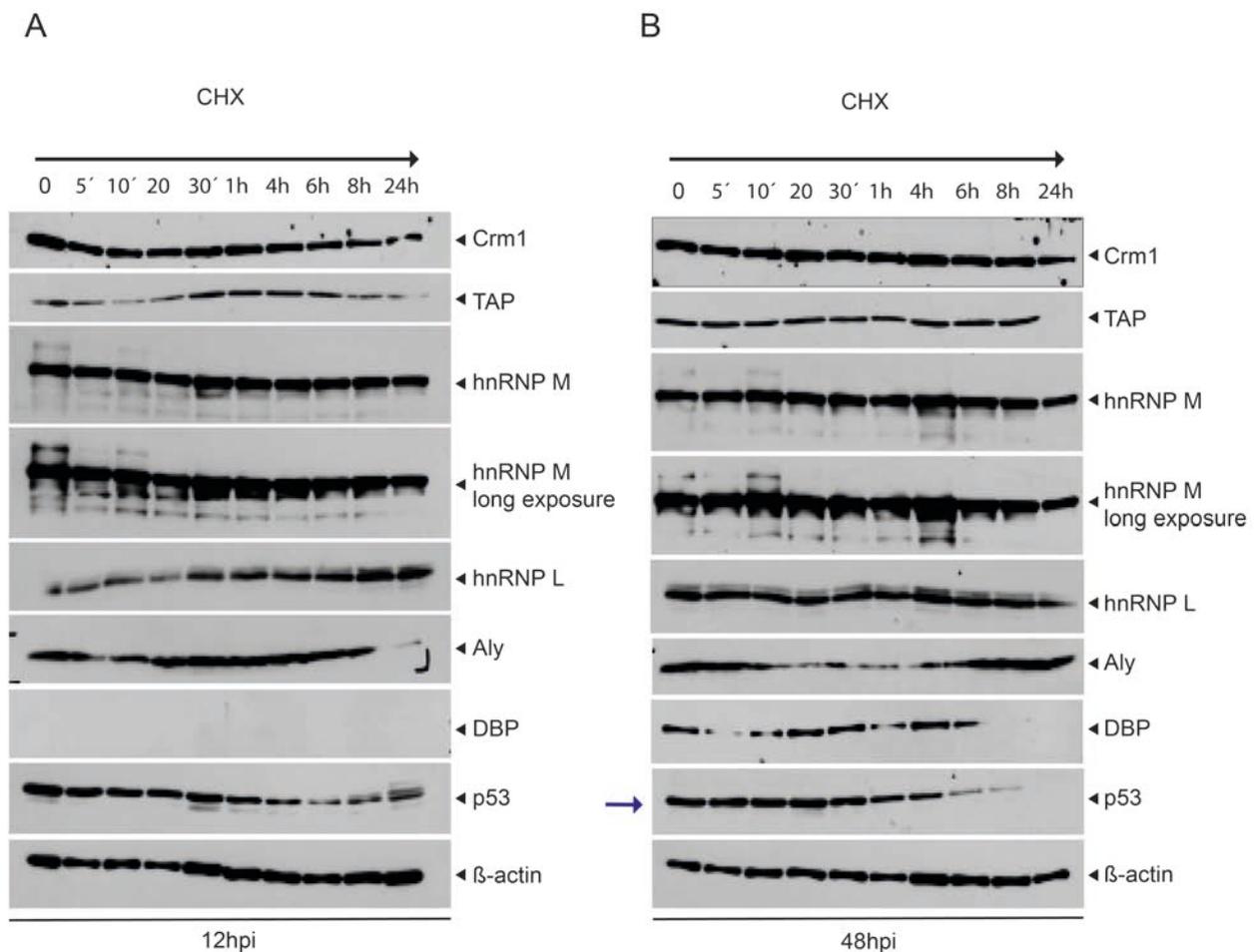


Figure 18. Protein stability analysis of cellular proteins that participate in biogenesis. Subconfluent A549 cells (2.5×10^6) were infected with wt HAAdV-5 (H5pg4100) at an MOI of 20 ffu/cell. At 12 (panel A) and 48 hpi (panel B), cells were treated with cycloheximide during indicated time points. After treatment, whole cell protein lysates were prepared. Samples were resolved by 10% SDS-PAGE and cellular Crm1, TAP, hnRNP M, hnRNP L and Aly proteins were visualized by Western Blot. Viral protein DBP was used as a control of infection. Cellular protein p53 is used as a control of degradation (shown with a blue arrow) and β -actin as a protein concentration control.

During the early phase of infection, p53 levels were constant even at 24h of treatment with cycloheximide (Figure 18; panel A, 12hpi), showing that at this point it is not targeted for degradation. Levels for Crm1, TAP, hnRNP M, hnRNP L also remained constant. At 12 hpi, Aly showed less abundance at 24h post-treatment, but it is still present at 48 hpi. Therefore, this degradation is probably not induced by the viral E1B/Orf6/E3 ubiquitin ligase, since this effect occurs mostly during the late phase of the replication cycle (as shown with p53 at 48 hpi).

In these conditions, no changes were observed in the half-life of any of the analyzed cellular proteins. Even though, in the late phase, the half-life of our positive control, p53 is decreased. This is shown by a decay in its signal after 4h of treatment with cycloheximide. As mentioned before, this effect has been previously reported to be caused by the viral E1B/Orf6/E3 ubiquitin ligase. Therefore, we concluded that the stability assay worked but the analyzed proteins

involved in mRNA biogenesis are not targeted for degradation during the late phase of infection.

4.1.3 hnRNP M SUMOylation analysis

None of the analyzed cellular mRNA biogenesis components showed evidence of degradation by the viral E1B/Orf6/E3 ubiquitin ligase during the late phase of infection. However, different post-translational modifications are known to regulate protein functions besides degradation that can alter the molecular weight of the protein enough to show higher forms of the protein band in a WB (e.g. phosphorylation, ubiquitination, SUMOylation, etc.). In our time-course experiments, several proteins showed higher bands in the late phase (Figure 17). Therefore, we decided to analyze if they were subjected to a post-translational modification. We suspected of SUMOylation as a possibility due to the size of the upper bands that appear at late time points. Each SUMO protein has a molecular weight of ~12 kDa and poly-SUMOylation can occur with SUMO-2 whereas SUMO-1 terminates SUMO-2 chains or mono-SUMOylates proteins. Therefore, one or several slower migrating bands can be generated because of SUMOylation modifications.

To evaluate if SUMOylation could be detected, a SUMO pull-down assay was performed in which HeLa cells overexpressing SUMO-1 or SUMO-2 proteins, containing a polyhistidine-tag, were infected (Figure 19). Cells were collected and tested for SUMOylation at early (6 hpi) and late (24 and 48 hpi) time points of infection. The SUMO pull-down samples were tested by WB for the presence of hnRNP M, TAP and AP5. SUMOylated TAP or AP5 could not be detected in the samples after the SUMO pull-down, neither in infected nor mock-infected samples (not-shown). However, SUMOylation of hnRNP M, which is known to be modified by SUMO-1, could be detected in mock cells and during infection (panel A, long exposure, marked with a white arrow). The signal for hnRNP M in WBs, has been described as a doublet ~68 kDa and its SUMO-1 modification as a triplet ~90 kDa (Vassileva and Matunis, 2004).

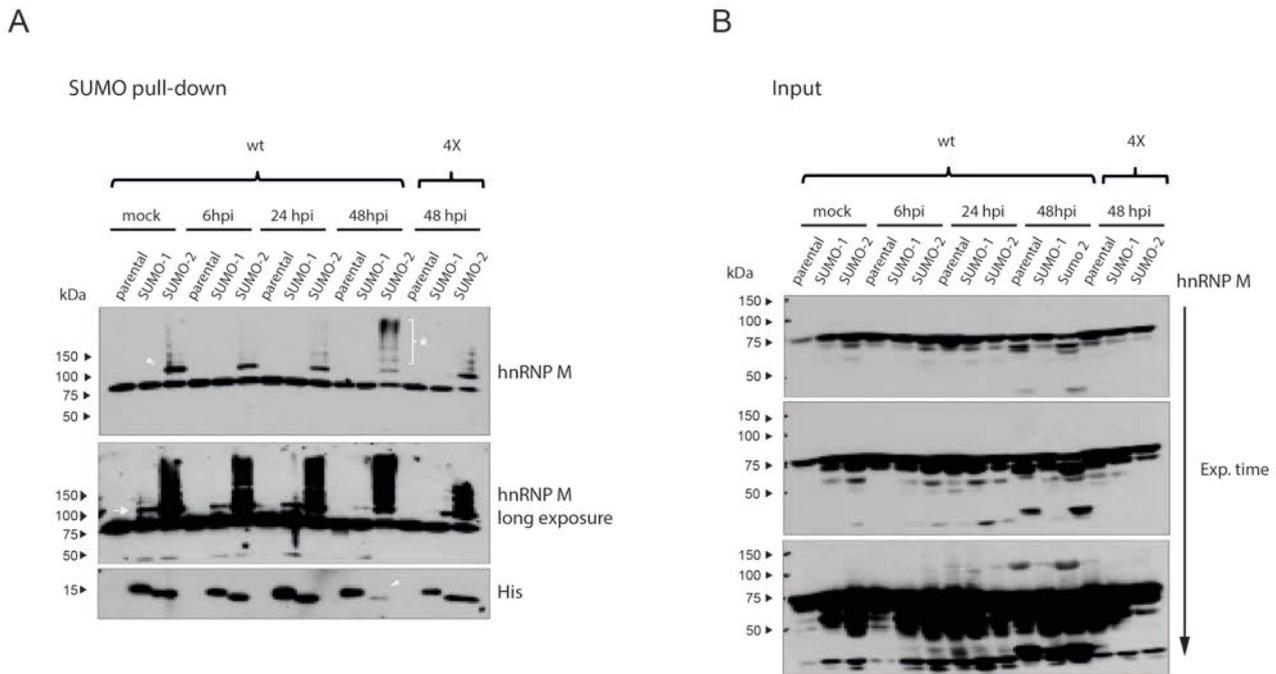


Figure 19. SUMOylation analysis of hnRNP M during infection. HeLa cells (4×10^6), constitutively expressing p6His-SUMO-1 or p6His-SUMO-2, were infected with the wt HAdV-5 (H5pg4100) and the E1B-55K null-mutant virus, 4X (H5pm4149), at an MOI of 20 ffu/cell. Cells were harvested at 6, 24 and 48 hpi. Whole cell protein lysates were prepared, and SUMO precipitation was performed using nickel beads to pull-down SUMOylated proteins. Eluted samples and input samples were resolved by 10% SDS-PAGE and cellular hnRNP M protein was visualized by immunoblotting. The SUMO pull-down experiment is shown in panel A, and the input samples in panel B. Histidine was also visualized by immunoblotting as a control for the purification of tagged SUMO proteins (panel A). The white arrow shows the triplet of SUMO-1-hnRNP M. The SUMO-2 modified hnRNP is shown with an arrow head and the higher forms of SUMO-2-hnRNP M are marked with an (*). At 48 hpi with wt, the white arrow heads show the decrease of the free form of Histidine tagged SUMO-2.

In our experimental set-up, SUMO-2 modification was also detected in both, infected and uninfected cells as a band at slightly below the SUMO-1-hnRNP M triplet. However, in the latest time-point of infection (48 hpi), higher SUMOylated forms of hnRNP M were observed above the SUMO-2 band in early time-points (Figure 19, panel A, 48 hpi, marked with an asterisk). In the long exposure image, these upper SUMOylated bands were also detected under all conditions. However, at 48 hpi less free SUMO-2 was detected, showing that at this time point there is more SUMO-2 attached to proteins (Figure 19 panel A, His blot indicated with an arrowhead). In the long-exposed image of hnRNP M, SUMO-1 modification decreases at 48 hpi, when poly-SUMO-2 modification becomes more abundant (Figure 19, panel A, 48 hpi).

In summary, we confirmed that hnRNP M is post-translationally modified by SUMO-1, as well as by SUMO-2, in uninfected and infected conditions. Also, an increase on heavier forms of SUMO-2-conjugated hnRNP M was observed at late times of infection. The presence of heavy forms of SUMO-2-conjugated hnRNP M was associated with a decrease in SUMO-1-conjugated hnRNP M.

4.1.4 Analysis of E1B-55K dependence for changes in hnRNP M SUMOylation patterns during infection

During infection, the E1B-55K (E1B) protein is a substrate for SUMOylation (described in section 1.8). This post-translational modification on E1B results in changes in its interaction partners, its subcellular localization, and its ubiquitin-ligase activity. E1B itself is a SUMO E3 ligase targeting cellular proteins p53 and Sp100A for SUMOylation (Muller and Dobner, 2008; Berscheminski *et al.*, 2014). Thus, we wanted to know if the higher SUMOylated forms of hnRNP M at late time points depended on the presence of E1B.

To test dependence on E1B, the SUMO pull-down was performed using cells infected with the E1B-null mutant H5pm4149 (4X) and collected them at 48 hpi to compare with wt (Figure 19, panel A, 48 hpi). At this time-point, the higher SUMOylation forms of hnRNP M could not be detected. It is known that the E1B-null mutant has a delayed effect in the entry towards the late phase of infection. Therefore, our observation time points were extended and tested for SUMOylation at 72 hpi (Figure 20). Even at this very late time-point no higher forms of SUMOylated hnRNP M were detected in the E1B-null mutant compared to wt (Figure 20, SUMO-2, wt 48hpi vs 4x 48-72 hpi).

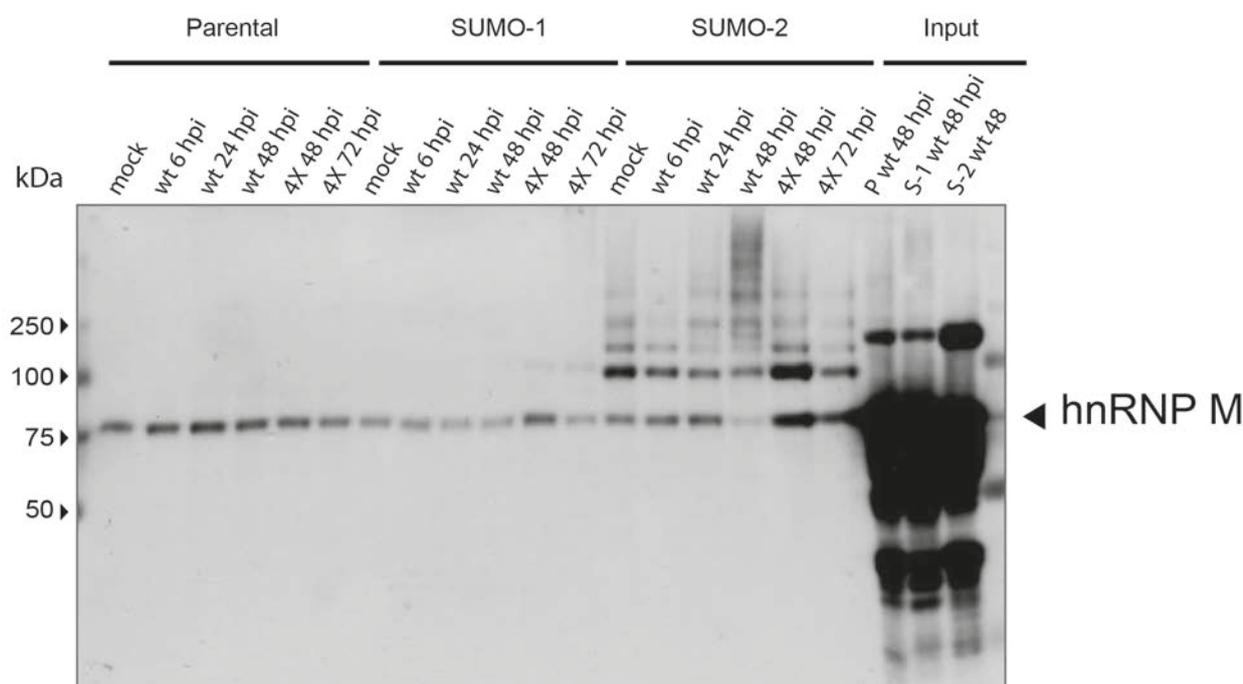


Figure 20. SUMOylation analysis of hnRNP M in the absence of the wt E1B-55K viral protein. HeLa cells (4×10^6), constitutively expressing p6His-SUMO-1 or p6His-SUMO-2, were infected with the wt HAdV-5 (H5pg4100) and the E1B-55K null-mutant virus, 4X (H5pm4149), at an MOI of 20. Cells were harvested at 6, 24, 48 and 72 hpi. Whole cell protein lysates were prepared, and SUMO precipitation was performed using nickel beads. The SUMO pull-down samples and input samples were resolved by 10% SDS-PAGE and cellular hnRNP M protein was visualized by immunoblotting.

E1B dependence for hnRNP M SUMO-2 modification, was also observed by mass spectrometry experiments performed by colleague Dr. Julia Berscheminski analysis (data not yet published). Her data further confirms our results, which indicate that during infection E1B is required, directly or indirectly, to observe higher forms of SUMO-2-conjugated hnRNP M.

4.1.5 Analysis of E1B-55K phosphorylation status dependence for changes in hnRNP M SUMOylation patterns during infection

During infection, E1B is post-translationally modified by phosphorylation and SUMOylation. Phosphorylation of the E1B protein is known to favor its own SUMOylation (Wimmer *et al.*, 2013). To further analyze the dependence of E1B for the SUMOylation of hnRNP M, we repeated the SUMO pull-down using different E1B mutants for posttranslational modifications (Figure 21). Thereby, the SUMO and the DelP mutants are unable to be SUMOylated and phosphorylated, respectively. The pMimic mutant is mimicking a continuously phosphorylated state of the protein (contrary to the DelP mutant) on three phosphorylation sites at the C-terminal end of the protein. We also included the NES mutant, which has a deletion in the NES sequence, inhibiting its CRM1-dependent nuclear export. Thus, deletion of the NES sequence results in an E1B which is permanently in the nucleus and in the viral RCs. This nuclear retention leads to higher SUMOylation of E1B. The SUMO/NES is the double mutant blocking the SUMOylation site and the nucleus export signal of E1B. This setup of mutants allowed us to ask if any of the known characteristics of the E1B, associated to each mutant, is important for the hyper-SUMOylation of hnRNP M.

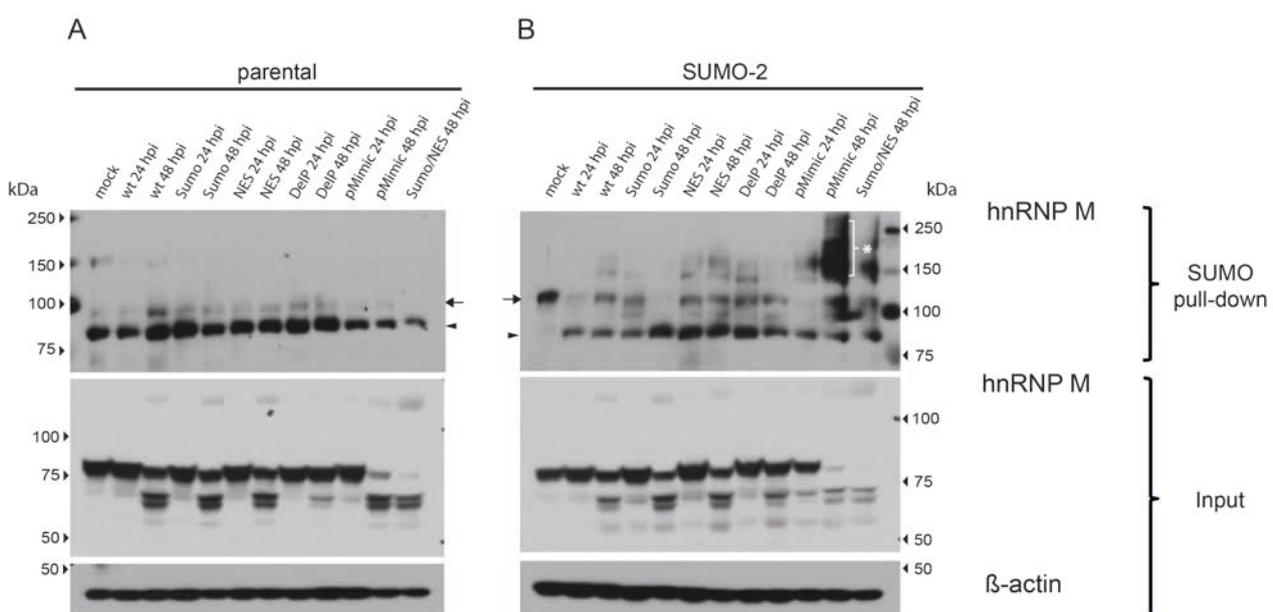


Figure 21. SUMOylation analysis of hnRNP M with mutants of the E1B-55K viral protein. HeLa cells (4×10^6), constitutively expressing p6His-SUMO-2, were infected with the wt HAdV-5 (H5pg4100) and the E1B mutant viruses: Sumo mutant (H5pm4102), NES mutant (H5pm4101), DelP mutant (H5pm4174), pMimic mutant

(H5pm4219) and the double Sumo/NES mutant (H5pm4103). Cells were infected at an MOI of 20 and harvested at 24 and 48 hpi. Whole cell protein lysates were prepared, and SUMO pull-down was performed using nickel beads. SUMO pull-down samples and input samples were resolved by 10% SDS-PAGE and cellular hnRNP M protein was visualized by immunoblotting. The black arrows show the single SUMO-2 modification of hnRNP M. The black arrow heads show unmodified hnRNP M. Higher forms of SUMO-2-hnRNP M are marked with a white (*).

The NES and DelP mutants showed higher SUMOylated bands since 24 hpi compared to wt, in which they appear until 48 hpi. On the contrary, the SUMO mutant showed normal SUMOylation bands at 24 hpi and then they disappear at 48 hpi. Interestingly, the pMimic mutant had a very intense signal for the higher SUMOylation bands of hnRNP M (Figure 21, SUMO-2, wt vs pMimic at 48 hpi). In a lesser extent, the double mutant SUMO/NES showed also an intense signal in the higher SUMOylation bands of hnRNP M. These experiments are preliminary, so they should be confirmed since they were only performed once. However, they suggest that when the E1B is in a constant state of phosphorylation and in the nucleus (as is the pMimic mutant), hyper-SUMOylation of the cellular hnRNP M is promoted even more than in the E1B wt.

Altogether, it was found that hnRNP M SUMOylation is increased during the late phase of infection, resulting in heavier SUMOylated forms of SUMO-2-hnRNP M. We could also confirm that the presence of the viral E1B is required, directly or indirectly, to observe this effect. And finally, preliminary experiments suggest that phosphorylation of the E1B could be responsible for this effect.

4.1.6 hnRNP M intracellular distribution analysis

After demonstrating that infection with HAdV-5 results in hnRNP M being modified by the addition of SUMO-2 chains, the next question was if its intracellular distribution is altered too. Therefore, an immunofluorescence analysis was performed to analyze changes of hnRNP M localization at different times of infection (Figure 22). For this analysis, the viral DBP protein was used as a marker of infection and to track progression of the late phase.

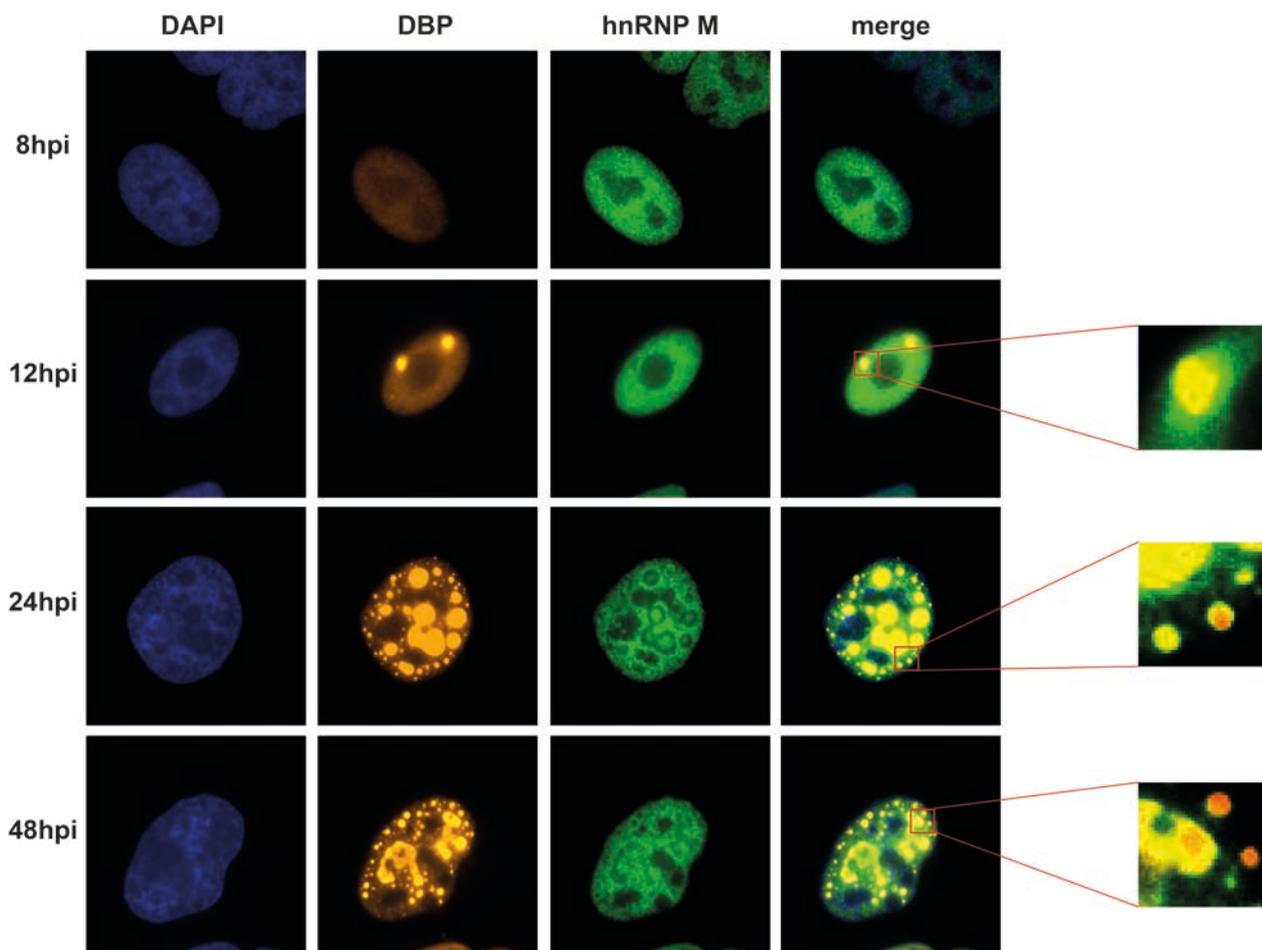


Figure 22. Subcellular distribution of hnRNP M during infection with HAdV-5. Immunofluorescence images, showing a nucleus close-up of representative distributions for each condition. A549 cells were infected with the wild-type virus H5pg4100 at an MOI of 30 ffu per cell and fixed at 8, 12, 24 and 48 hpi. Cells were labeled with anti-DBP (E2A) and anti-hnRNP M3/4 (2A6). Secondary antibody incubation was done using Cy3 (orange) for DBP, and FITC (green) for hnRNP M. The nucleus (DNA) is labeled with DAPI (blue). The overlays of all labels are shown in the merge distributions column. The panels at the right, are 5X zoomed images of the marked areas from the nucleus showing the distributions around the RCs.

At early time points (8 hpi), the distribution of the hnRNP M is the same as in non-infected cells, nuclear, diffuse, and excluded from nucleoli (Figure 22; 8 hpi, hnRNP M, compare nuclei with and without DBP). As the late phase progresses (12-48 hpi), DBP starts to assemble into the viral RCs, forming first round structures that accumulate and start fusing at very late time points (Figure 22, 12-48 hpi, DBP). During this period, the hnRNP M remains diffuse in the nuclei, however, it accumulates more intensely around DBP structures (Figure 22; RC zoomed images).

As described before, the mechanism of how the E1B/Orf6/E3 ubiquitin ligase impacts viral late mRNA export remains unknown. It has been hypothesized that it could degrade substrates involved in mRNA processing or export, so far, none of the known substrates for the complex are involved in mRNA biogenesis. In summary, we evaluated known key cellular components, with functions at different stages of mRNA maturation and export. None of the tested cellular proteins were found to be degraded during infection. Thus, it remains unanswered whether the

viral E1B/Orf6/E3 ubiquitin ligase is involved at targeting cellular proteins participating in mRNA biogenesis and if this is a mechanism used by adenovirus to selectively export viral late mRNAs during the late phase. During our evaluation, it was found that during infection, hnRNP M is conjugated by SUMO-1 and SUMO-2. This post-translational modification occurs in non-infected cells too, however, higher forms of SUMO-2 modified hnRNP M are formed at late time points of infection. It was also found that this effect on hnRNP M SUMOylation depends on the presence of the viral E1B protein. Preliminary results suggest that phosphorylated E1B enhances production of this higher forms of SUMO-2-conjugated hnRNP M. These results link the viral E1B protein with mRNA processing and export pathways. Finally, it was also discovered that hnRNP M accumulates around the viral RCs at late time points of infection. Interestingly, this effect correlates with the same time points where hnRNP M is highly SUMOylated, although, it is unclear if the hnRNP M that localizes at the RC is the highly SUMOylated hnRNP M.

4.2 Analysis of viral and host mRNA accumulation and export rates during infection

As described in the introduction, the E1B/Orf6 complex is required for efficient accumulation of viral mRNAs in the cytoplasm. However, the role they play in selective export is yet not understood. They are not functioning as a mRNA shuttle complex and so far, their role within the E1B/Orf6/E3 ubiquitin ligase is not clear on how it impacts mRNA export. In our experiments, no cellular components were found to be targeted for degradation during infection.

It has been published that viral late mRNAs export is via the Aly/Tap pathway (Yatherajam *et al.*, 2011). This pathway is the normal export pathway used by bulk cellular mRNAs in non-infected conditions. If viral late mRNAs require the same export pathway as cellular mRNAs, it is unclear if there is a complete cellular export shut off. The concept of selective viral late mRNA export arose since first publications could not detect cellular mRNAs in the cytoplasm (see section 1.5.3.1). Methods used in those first publications did not have enough specificity to measure specific mRNAs. Afterwards, several publications analyzing viral late mRNAs versus a handful of cellular mRNAs have been analyzed to test for selective export during infection, and how it is or not affected under certain conditions. However, these publications only test the most abundant viral late mRNAs and it is unclear if viral late selective export occurs. Therefore, we decided to perform a comprehensive transcriptome analysis to get a global idea of how infection alters viral and cellular mRNAs and readdress the concept of selective mRNA export.

4.2.1 Analysis of viral mRNA accumulation during infection

To investigate how the relative abundance of the viral and cellular mRNAs behaves during infection, we decided to do a global analysis by deep sequencing using the same experimental setup applied for previous time-course experiments. RNA samples were obtained from nucleus and cytoplasmic fractions at each time-point, to analyze the cytoplasmic to nucleus ratios (cyto/nuc). The RNA samples were obtained and processed according to the protocol described in the methods (section 3.2.9) and sent to the HPI's Technology Platform, where the RNA samples were first tested for quality. Further, the mRNAs were purified by poly(A) selection (to exclude rRNAs and other miRNAs) and the cDNA library was prepared with insert fragments of 100bp (see section 3.2.10). Samples were multiplex-sequenced (by synthesis method) on the Illumina HiSeq 2500 (2 x 100bp paired end run) at a depth of about fifty million reads per sample. Each sample was sequenced in duplicate. The obtained sequenced reads were further subjected to both viral and cellular differential expression analyses.

4.2.1.1 Time-course alignment of sequenced reads to the viral transcript annotations

Initially, we decided to analyze the relative abundance of the viral mRNA sequences during the course of infection to get a full notion of how viral mRNAs expression behaves under our experimental conditions. To do this, the viral transcriptome annotations were required as a reference to align the sequenced reads (Figure 23). Since only the HAdV-2 is currently fully annotated in PubMed (AC_000007.1), annotations of the HAdV-5 transcriptome were compiled from different sources for this analysis. The H5pg4100 full sequence was used as a reference and new annotations were added by comparing the CDS between HAdV-2 and HAdV-5, since both genomes have a very high homology (Chroboczek, Bieber and Jacrot, 1992). In addition, new alternative splicing forms of certain mRNAs from the E4 region have been detected in the early and late phases of infection (Dix and Leppard, 1993). These are included in our annotation indicated with suffix -early and -late (Figure 23; E4 transcripts with suffix -early and -late). Also, a more recent deep sequencing study of HAdV-2 infected cells showed that there are several alternative splicing sites not previously published (Zhao, Chen and Pettersson, 2014). These newly predicted transcripts were also incorporated into our transcriptome (Figure 23; transcripts with suffix -new). In total, 62 transcripts were annotated for this analysis (Table 1).



Figure 24. Mapping of the sequenced reads across the HAdV-5 transcriptome. The reads were assembled and mapped to the adenoviral transcriptome using the CLC Genomics Workbench software (see methods, section 3.2.11). The annotated transcripts are represented in green at the top of the figure. The blue histograms below show the frequency and distribution of the reads that were mapped with the viral transcript annotations. A histogram is shown for each of the analyzed time points. Numbers at the left side of each histogram indicate the total number of reads that mapped against the viral transcriptome at each time point.

Interestingly, the histogram at 24 hpi looks completely different from previous time points. Most reads localized in the middle section of the genome and in three small introns that belong to the tripartite leader, which is shared by all viral late mRNAs. Reads distributed on early regions are still present but are not as abundant as in previous time points (Figure 24; 24 hpi). This histogram clearly shows that the late phase of infection is already established at 24 hpi. At 48 hpi, reads are distributed very similarly to 24 hpi, showing almost no reads in the early regions (Figure 24; 48 hpi).

Importantly, the total number of reads that aligned to viral mRNAs increased more than 21 thousand times from 6 hpi to 48 hpi (Figure 24; numbers in the left side of the histograms). Samples were sequenced at a depth of around 50 million reads and at 6 hpi, only 234 of those

reads were viral. However, at 48 hpi, around 49 million reads belonged to viral mRNAs. This shows that at this late time-point most of the mRNAs in the samples were viral.

Therefore, this initial analysis showed that the reads that align with the viral transcriptome change, in pattern and number, along the infection according to what we expected to observe at early and late time points. This reassured us that our experimental setup would allow us to see the effects of the transition from the early into the late phase of infection, and it encouraged us to further quantify and analyze the viral mRNAs by RNA-Seq methods.

4.2.1.2 Quantification of viral transcripts at different time points of infections

The RNA-Seq analysis was performed using the CLC Genomics Workbench and sequenced reads were assigned to the matching mRNAs. Expression values were calculated taking the total number of reads per transcript and normalizing it to the length of the transcript, resulting in RPKM (Reads Per Kilobase of transcript per Million mapped reads) as an expression unit (see 3.2.1.1). Expression values were calculated for samples of nuclear and cytoplasmic fractions of each time-point to determine export ratios. In addition, total expression values were calculated from the sum of nuclear and cytoplasmic expression values (in Table 2 and plotted in Figure 25). Inspecting the total expression values of the viral mRNAs allowed us to have a better understanding of the variations on the relative abundance of early vs late mRNAs species along the replication cycle.

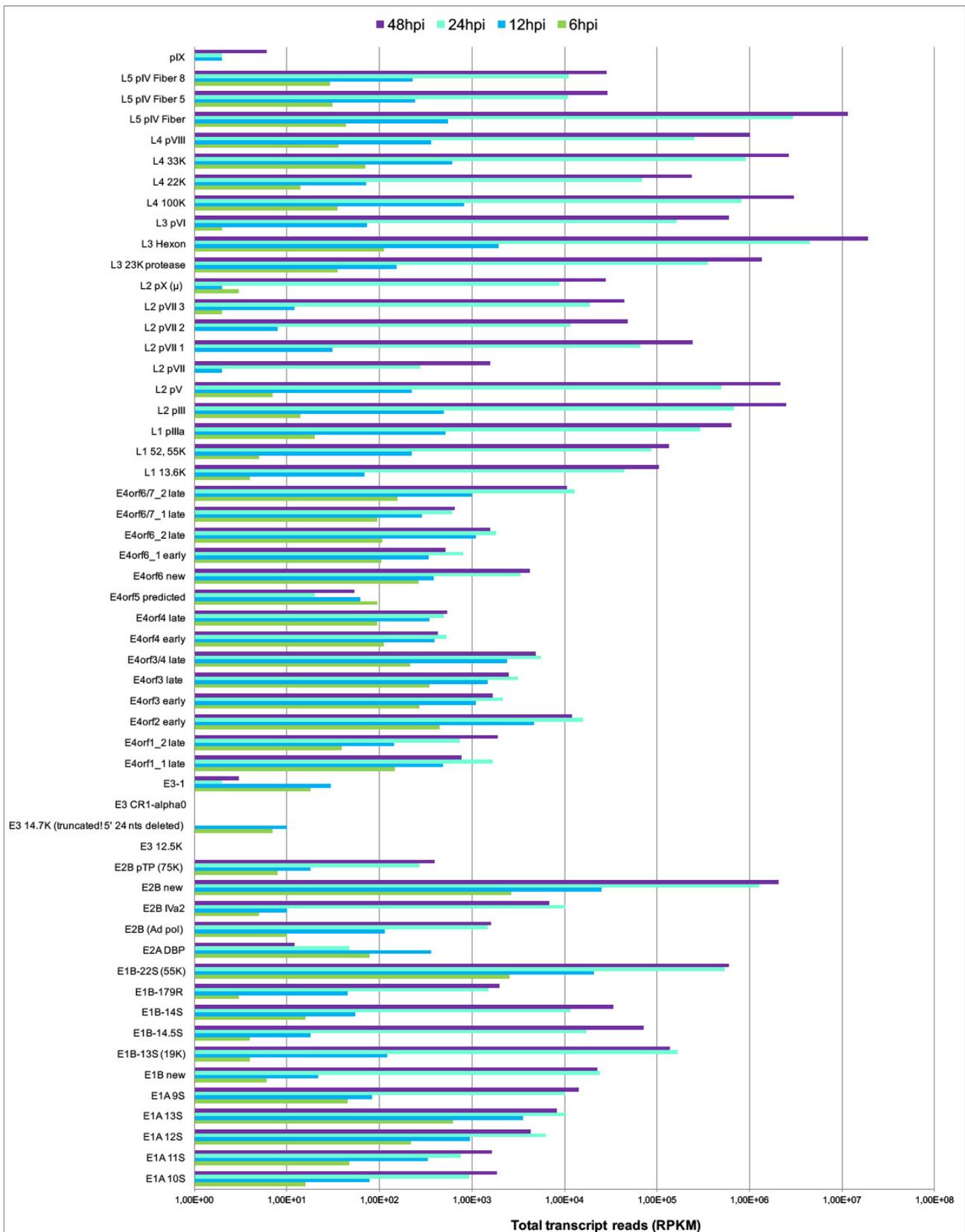


Figure 25. Total viral mRNAs expression values at different time points. The above bar chart shows the total expression values calculated from the sum of nucleoplasmic and cytoplasmic total values shown in Table 2. The total expression values for each time-point is represented in a different color (shown at the top of the chart) side by side. The total expression values are represented in RPKM (Reads Per Kilobase of transcript per Million mapped reads) units.

Using the total expression values, a general analysis of all viral gene families was conducted to track changes as the early phase transitions into the late phase (Figure 26). The total amount of both early and viral late mRNAs detected at 6 hpi was above 9,000 RPKM (Figure 26 panel B, 6 hpi). Though, early mRNAs were more abundant than viral late mRNAs, which represented only 5% of the total viral mRNAs (Figure 26 panel A, 6 hpi). At this time-point, the most abundant mRNAs were E1B-22S (55K) and a larger version of the E2B (E2B-new) followed by E1A-13S and E4 Orf2 early (Figure 25, 6 hpi).

At 12 hpi, the total amount of viral mRNAs measured was almost eight times higher than at 6 hpi with more than 70,000 RPKM (Figure 26 panel B, 12 hpi). Early viral mRNAs were still more abundant than viral late mRNAs, which slowly increased to 9% (Figure 26 panel A, 12 hpi). At this time-point the most abundant mRNAs are the same as at 6 hpi, however, E4orf3/4 late mRNA also became one of the most detected mRNAs (Figure 25, 12 hpi).

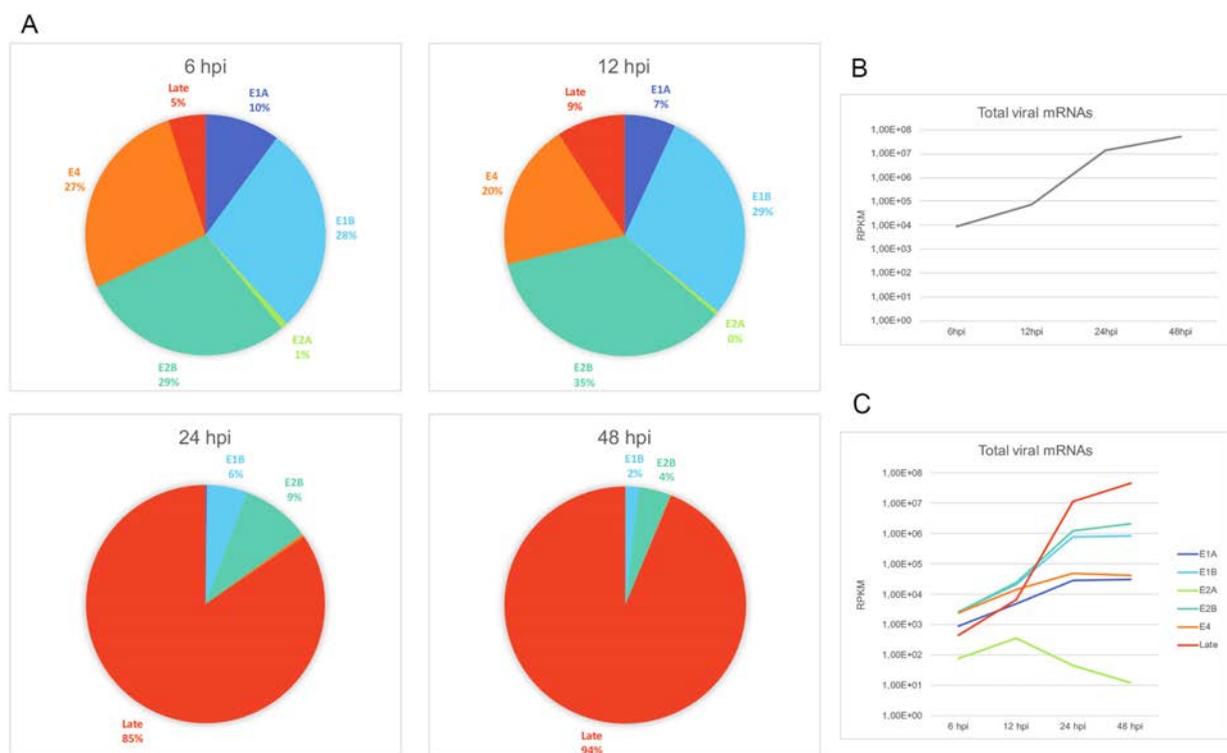


Figure 26. Expression of HAdV-5 gene families at different infection time points. Total viral mRNA expression values were used to analyze the changes and contribution of each viral gene at different time point. A) Pie-charts represent the percentage of RPKMs of each adenoviral gene at indicated time points. Line plot represents B) the total amount of viral RPKMs measured during the time-course of infection, and C) the kinetics and contribution of each viral gene at different time points of infection. The total expression values are represented in RPKM units, which are Reads Per Kilobase of transcript per Million mapped reads.

At 24 hpi, it is obvious that the transition to the late phase has been established, since most viral mRNAs present (85%) were late mRNAs (Figure 26 panel A, 24 hpi). In total, all viral mRNAs at this time-point reached almost 14 million RPKM, which is 190 times more than what was

observed at 12 hpi (Figure 26 panel B, 24 hpi). E2B-new was still one of the most abundant mRNAs, however, hexon and fiber mRNAs surpassed its levels, and from this point forward were the most abundant mRNAs during the late phase. The viral late mRNAs L4-100K and L4-33K were also highly abundant during this period (Figure 25, 24 hpi). Interestingly, most of the early mRNAs reached their highest expression levels at this time-point (10^4 - 10^6 RPKMs). Some then plateau and others decrease (Figure 26 panel C). Nevertheless, the viral late mRNAs were the most abundant of all.

At 48 hpi, almost all viral mRNAs present (94%) were viral late mRNAs (Figure 26 panel A, 48 hpi). At this time-point, there were more than 48 million RPKM, amounting to 3.5 times more than were measured at 24 hpi (Figure 26 panel B, 24 hpi). Viral late mRNAs were still being highly synthesized but the biggest increment in synthesis was from 12 to 24 hpi (Figure 26 panel C). Apart from hexon, fiber, L4-100K and L4-33K, which were the most abundant of all viral mRNAs, there was also an increase of L2 pII, L2 pV and L3.23K (protease) (Figure 25, 48 hpi).

As expected, this general evaluation of the total expression values showed that the early genes are predominant during the early phase and the late genes expression activates in the late phase. Also, the most expressed viral mRNAs at each time-point were identified and correlate with what is known (see introduction; sections 1.1.4, 1.2, 1.3.1, and 1.5.1). We observed that most of the early genes are not shut down as the infection progressed into the late phase, however, viral late mRNAs become more abundant by several orders of magnitude.

4.2.1.2.1 Analysis of viral early mRNAs

Next, a more detailed examination into each viral gene family was made to better understand the kinetics of the early and late mRNAs and how the transition into the late phase of infection is established. As shown previously, viral early mRNAs were the predominant viral mRNAs during the early phase and expression of viral late mRNAs was very low (6 and 12 hpi). The expression of the viral early mRNA continued during the infection cycle and most of them reached their highest levels at 24 hpi (Figure 26 panel C). However, the kinetics of viral early mRNA expression varied between genes (Figure 27). Most early mRNAs displayed similar kinetic tendencies to other members of their gene family, except for the E2 mRNAs, which all behaved differently (Figure 27, E2). In general, mRNAs of E1A and E4 genes showed similar kinetics, just as mRNAs of E1B and E2B genes behave similarly (Figure 26, panel C; and Figure 27).

Most of E1A and E4 mRNAs increase gradually during infection (Figure 27, E1A and E4). In contrast, E1B mRNAs as well as E1A 9S, E2B-new and E2B IVa2 mRNAs have a drastic increase in expression, from 12 to 24 hpi, of two or more orders of magnitude (Figure 27). Therefore, in the transition to the late phase, the percentage of E1A and E4 mRNAs severely diminishes (Figure 26, panel A), since it correlates with the drastic increase of viral late mRNAs (Figure 26, panel C). Only E2A DBP and the predicted E4orf5 mRNAs have very different kinetics from the rest. E2A DBP mRNA decreases during the late phase, reaching its highest levels of detection at 12 hpi, and dropping afterwards (Figure 27, E2). The predicted form of E4orf5 mRNA peaks at 6 hpi, then decreases until 24 hpi to increase again at 48 hpi (Figure 27, E4).

In general, expression levels of early viral mRNAs do not reach more than 6 orders of magnitude during the whole viral infection cycle (Figure 26, panel C). The E1B mRNAs, as a family, are the most abundant early mRNAs kind. Specifically, the E1B mRNAs, 22S and 13S, reach more than 10^5 RPKMs (Figure 27, E1B). However, the E2B new mRNA reaches more than 10^6 RPKMs, being the most abundant early mRNA; even through the late phase it is more abundant than several late mRNAs (Figure 25, E2B new and Figure 27, E2). The E1A, E2 (with exception of the E2B new mRNA) and E4 mRNAs stay around 10^4 RPKMs at their highest levels (Figure 27, E2 and E4).

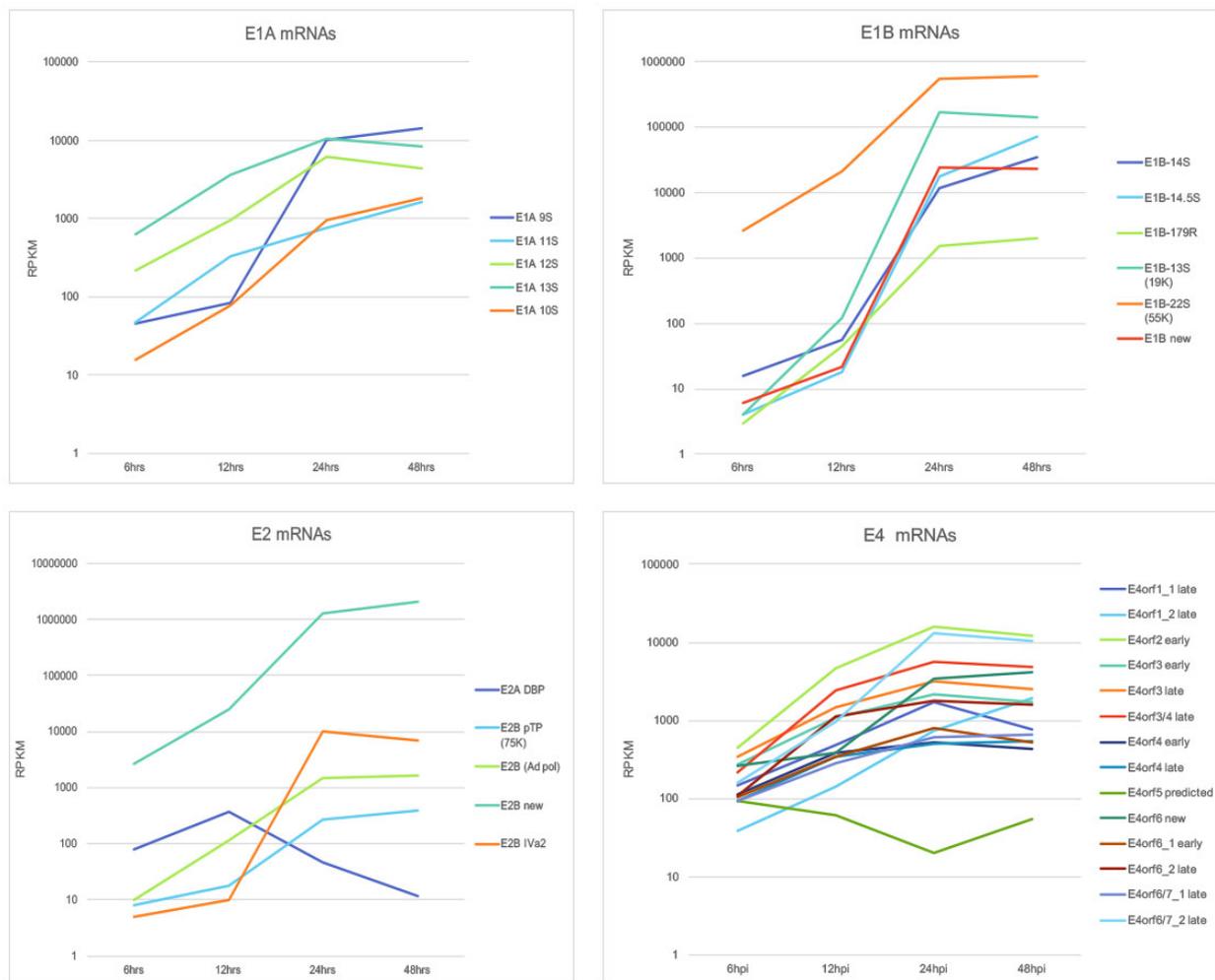


Figure 27. Kinetics of early viral mRNAs. Total viral mRNA expression values were used to analyze the kinetics of each viral early gene at different time points. Each panel shows the kinetics of every viral early mRNA for each early viral gene: E1A, E1B, E2 (E2A and E2B) and E4. The total expression values are represented in RPKM units, which are Reads Per Kilobase of transcript per Million mapped reads.

To analyze the relative abundance between the different mRNAs within each family we calculated the proportions at each time-point (Figure 28). Even though it was shown before that most viral early mRNAs increased constantly throughout infection, the proportions of each mRNA within the early genes families changed from the early to late phase. In the case of the E1A gene, during the early phase, 13S and 12S mRNAs were the most abundant and 9S was very low (Figure 28, E1A). At 48 hpi, proportions shifted as the 9S mRNA levels increased and it was more abundant than 12S and 13S mRNA (Figure 27 and Figure 28, E1A). Compared to the rest of E1A mRNAs, the 10S mRNA remained at low levels during infection, but it increased almost 3-fold at 48 hpi compared with the early phase. The proportions of the 12S and 11S mRNAs fluctuated less during infection, displaying very similar accumulation kinetics (Figure 27 and Figure 28, E1A).

The E1B mRNAs showed similar kinetics, but their proportions still change during infection cycle. The 22S mRNA was the most abundant during the whole infection cycle (Figure 27, E1B). During the early phase, it represented almost 100% of the E1B mRNAs (Figure 28, E1B). In the late phase 22S proportions decreased as the E1B new, 13S, 14.5S and 14S mRNA proportions increased due to their rapid accumulation from 12-24 hpi, which was greater than the accumulation of E1B 22S (Figure 27 and Figure 28, E1B). After transitioning into the late phase, 13S was the second most abundant mRNA in the E1B family; and at 48 hpi, 13S proportions decreased slightly as the 14.5S mRNAs accumulated faster (Figure 28, E1B).

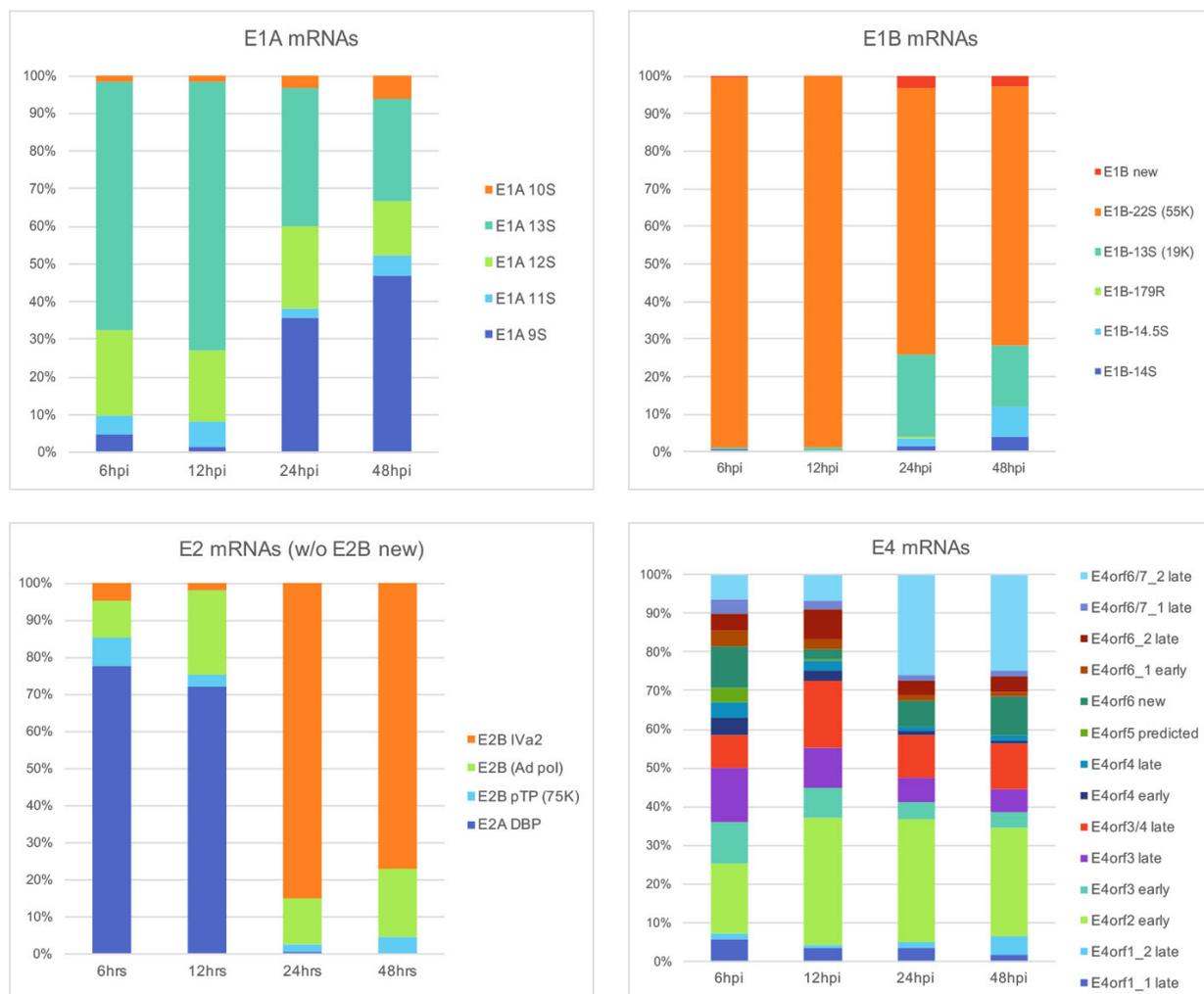


Figure 28. Proportions of early viral mRNAs. Total viral mRNA expression values were used to analyze the proportions of each viral early gene at different time points. Each panel shows the proportions at a given time point of every viral early mRNA for each early viral gene: E1A, E1B, E2 (E2A and E2B) and E4. Each mRNA is represented as a percentage relative to the total amount of viral mRNAs of each gene at each time point.

Changes in the E4 mRNAs proportions were not as drastic as the other viral genes, however, there were a few differences from the early to late phase (Figure 28, E4). At 6 hpi, proportions between all E4 mRNAs did not differ as much, being orf2 early and orf3 late mRNAs, slightly more abundant than the rest (Figure 28, E4). However, the orf6/7_2 late and the orf2 early

mRNAs increased in the late phase becoming the most abundant E4 mRNAs (Figure 27 and Figure 28, E4). The proportions of orf2 early increased since 12 hpi and the orf6/7_2 late increased at 24 hpi (Figure 28, E4). The orf3 early, orf4 early and orf1 early mRNAs decreased their proportions since 12 hpi and in the late phase they were half of their proportions at 6 hpi (Figure 28, E4).

After analyzing the proportions within the E2 genes (E2A and E2B), it was obvious that the E2B new mRNA was the most abundant during the whole infectious cycle. However, nothing has been described about this mRNA or its translation product, so we continued analyzing the rest of the E2 mRNAs without this new E2B transcript (Figure 28, E2). The proportions between E2A DBP and E2B IVa2 mRNAs shifted from the early to the late phase. DBP mRNA was the most abundant during the early phase and IVa2 mRNAs were present in very low levels. In the late phase, there was a shift after DBP mRNA levels dropped (Figure 27, E2) and IVa2 became the most abundant mRNA in the late phase. This change in expression could be important, since DBP acts as a viral DNA helicase, and it is essential for viral DNA replication. Thus, it is rational that it would accumulate in the early phase before the start of the late phase when DNA replication occurs. Expression of IVa2 is essential for the full activation of the Major Late transcriptional unit (see introduction 1.5.1). Therefore, a shift in expression from DBP to IVa2 could be crucial to determine the start of the late phase and transcriptional activation of viral late mRNAs. The E2B polymerase mRNA proportions were higher at 12 hpi (Figure 28, E2), but its levels continued to increase during the late phase (Figure 27, E2). Accumulation of E2B pTP mRNAs switched with DBP mRNAs, as they were very low at early time points and then increased after the drop of DBP mRNAs at 24 hpi (Figure 27, E2). This is also important, since pTP is needed once the viral replication has started, in the late phase. However, its proportions during infection remained very low within the E2 mRNAs (Figure 28, E2).

The E4 mRNAs proportions did not change as drastically as the other early viral genes, however, there were some fluctuations from the early to the late phase (Figure 28, E4). At 6 hpi, proportions between all E4 mRNAs differed only slightly, with orf2 early and orf3 late mRNAs, slightly more abundant than the rest (Figure 28, E4). However, the orf6/7_2 late and the orf2 early mRNAs increased in the late phase becoming the most abundant E4 mRNAs (Figure 27 and Figure 28, E4). The proportions of orf2 early increased from 12 hpi and the orf6/7_2 late increased by 24 hpi (Figure 28, E4). The orf3 early, orf4 early and orf1 early mRNAs decreased in proportion from 12 hpi and in the late phase they were half of their proportions at 6 hpi (Figure 28, E4).

Overall, the kinetic analysis of the early genes showed that most mRNAs behave similarly within each family, except for the E2 family. Analysis of proportions between early genes showed that there is clear a shift in expression patterns from the early to the late phase for E1A and E2 families. The E4 and E1B gene families have changes in proportions as well, but they are less drastic. This is interesting because, as described in the introduction, early gene families have several roles in activating transcription of the viral genome as well as the host cell, inhibiting the immune innate response of the host cell, regulating splicing of both viral and host mRNAs, and replication of the viral genome. Many of the early genes code for multifunctional proteins that change function during infection as they are modified post-translationally. These results show that in the transition into the late phase not only viral late mRNAs are affected, but also early viral mRNAs undergo a change in expression patterns, even though they are less abundant than the viral late mRNAs.

4.2.1.2.2 Analysis of viral late mRNAs

As previously shown, viral late mRNAs were present since the early phase, although in a very small percentage (Figure 26, panel A, 6-12 hpi). Their expression exponentially increased in the late phase (Figure 26, panel C), which resulted in them being the most abundant viral mRNAs in the late phase (Figure 26; panel A, 24-48 hpi) and had expression values from 10^4 to 10^7 orders of magnitude (Figure 30, panel A). Even though, all the viral late mRNAs families (L1-L5) had the same kinetics of expression and continued to increase during infection (Figure 29, panel B), there were some differences in the percentage of expression of each family during infection (Figure 29, panel A).

In the early phase, the most expressed families were the L3 and L4, with over 30% of the viral late mRNA's RPKMs each (Figure 29, panel A, 6 and 12hpi). At 6 hpi, the L1 and L2 families were the least represented, at this time-point. However, L1 and L2 proportions doubled at 12 hpi, as their expression levels rose slightly faster than the L5 mRNAs.

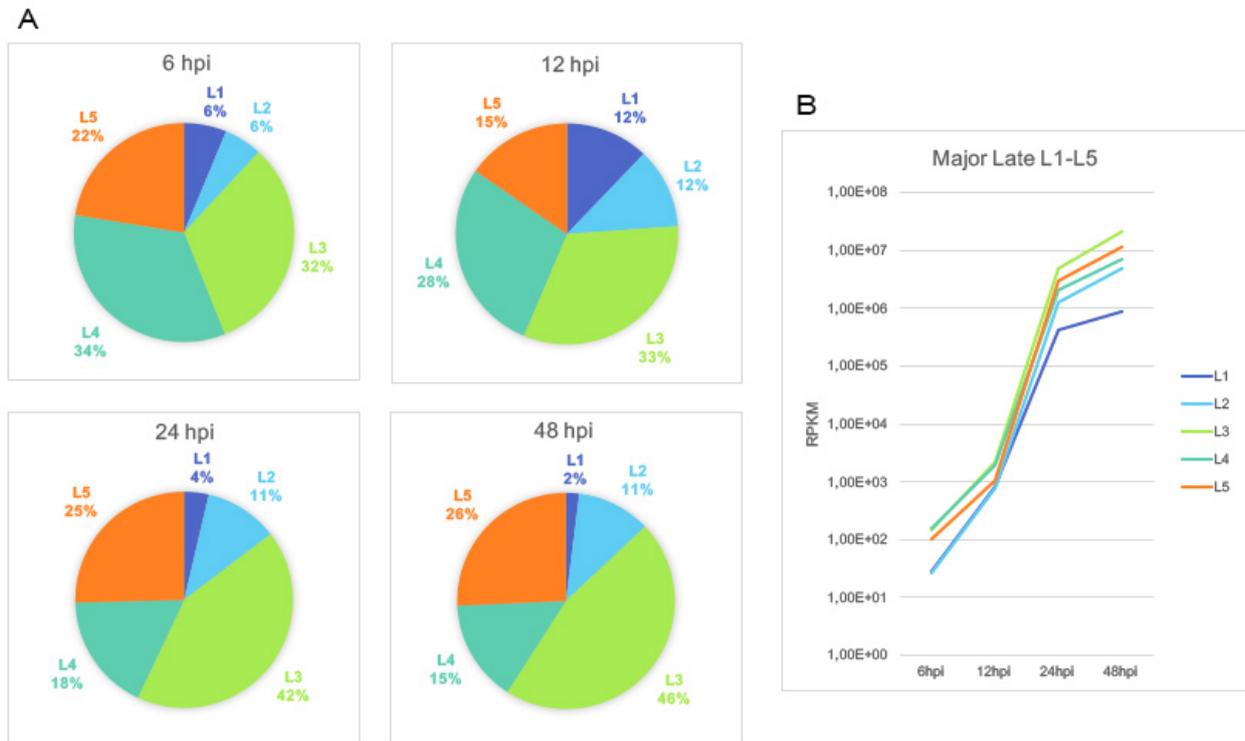


Figure 29. Analysis of Major Late families. Behavior of the Major Late families' expression at different infection time points. Total viral mRNA expression values for each family were used to analyze the changes and contribution of each viral late family at different time points. A) Pie-charts represent the percentage of RPKMs of each of the Major Late families at different time points. B) Line plot showing the kinetics and contribution of each of the Major Late families at different time points of infection. The total expression values are represented in RPKM units which are Reads Per Kilobase of transcript per Million mapped reads.

Once the late phase was established, proportions between the late genes did not change. The L3 family was the most abundant of all late mRNA families with almost half of the viral late mRNAs' RPKMs (Figure 29, panel A, 12 and 24 hpi). After L3, the L5 family was the second most represented at late time points. This was mainly due to L3 hexon and L5 pV fiber mRNAs being the most abundant mRNAs in the late phase, with expression levels that reached $\sim 10^7$ RPKMs (Figure 30, panel A).

The viral late mRNAs proportions, in contrast to the early genes, remained mostly the same during infection, specially once the late phase was established at 24 and 48 hpi (Figure 30, panel B). However, their changes in expression values went from three to five orders of magnitude ($\sim 10^2$ to $\sim 10^7$ RPKMs) due to their exponential kinetics (Figure 30, panel A). Besides the drastic change in expression levels, viral late mRNAs showed a few, but important, changes in their proportions during the transition from the early to late phase, as the L5 pV fiber and L3 hexon mRNAs accumulate a lot faster than the rest (Figure 30, panel A).

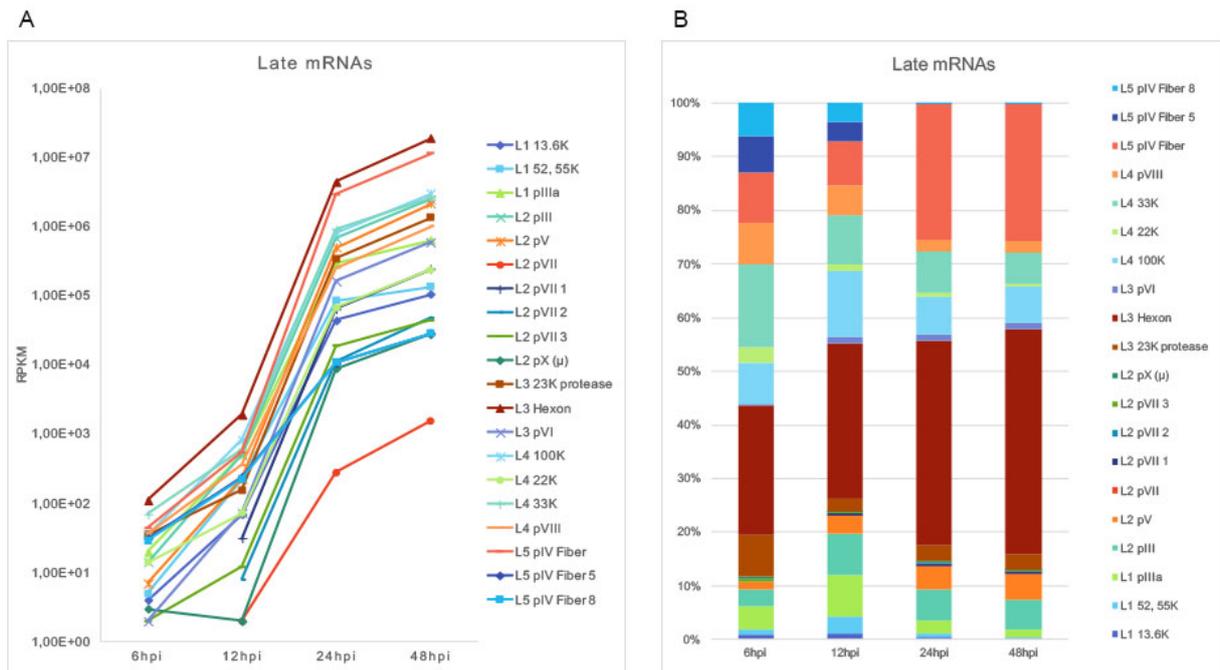


Figure 30. Kinetics and proportions of viral late mRNAs. Total viral mRNA expression values were used to analyze the kinetics and proportions of each viral late mRNA at different time points. A) shows the kinetics of every viral late mRNA for each viral late family (L1-L5). The total expression values are represented in RPKM units which are Reads Per Kilobase of transcript per Million mapped reads. B) Bar plot shows the proportions, at a given time point, of every viral late mRNA. Each viral late mRNA is represented as a percentage relative to the total amount of viral late mRNAs at each time point.

The L3 hexon mRNA was the most abundant viral late mRNA at all time points. Its expression levels increased five orders of magnitude (10^2 to 10^7 RPKMs) from 6 to 48 hpi (Figure 30, panel A) and its proportions increased by approximately one third from the early to late phase (Figure 30, panel B). The L5 pIV fiber mRNA, at 6 hpi has almost the same proportions as its alternative spliced counterparts, fiber 8 and 5 (Figure 30, panel B, 6 hpi). At late time points, as fiber 8 and 5 mRNAs decreased in their proportions, the canonically spliced fiber was the most prevalent form of fiber mRNA. Its expression levels increased also around 5 orders of magnitude, from 10^2 to almost 10^7 RPKMs, (Figure 30, panel A) and was the second most abundant viral late mRNA (Figure 30, panel B). This does not mean the splice variants, fiber 5 and 8, were not present anymore in the late phase; however, they were more than three orders of magnitude below the canonical splice form (10^4 vs 10^7 RPKMs). This clearly showed that even though all viral late mRNAs expression was exponential in the late phase, some were significantly more represented in the mRNA population than others, due to small differences in their kinetics.

The most abundant viral late mRNAs in the late phase code for the late capsid proteins hexon and fiber ($\sim 10^7$ RPKMs), the penton base pIII protein ($\sim 10^6$ RPKMs), the core protein pV ($\sim 10^6$ RPKMs) and the protease 23K ($\sim 10^6$ RPKMs). These highly expressed viral late mRNAs code for the most abundant proteins of the viral capsid (see introduction, section 1.1.3). Therefore, it

was expected that their corresponding mRNAs would be highly expressed to efficiently produce viral progeny. The viral late mRNAs coding for regulatory proteins L4-100K and L4-33K (~10⁶ RPKMs) are also highly expressed. Although they are not needed in the same quantities as hexon or fiber, these proteins participate in highly demanded processes in the late phase regulating the viral late proteins production, such as selective mRNA export, selective mRNA translation, and processes related to capsid formation such as DNA packaging and assembly of the hexon trimers (see introduction; section 1.2.2, Figure 5 and Figure 7).

As shown previously, most viral mRNAs are already present at 6 hpi (even in small quantities, as in the case of the viral late mRNAs), and continue to accumulate during the progression of the infection. However, due to variations in the kinetics of each mRNA, the proportions of viral mRNA change during the infection cycle. A summary of all the changes in proportions during the infection cycle of all viral mRNAs are represented in Figure 31.

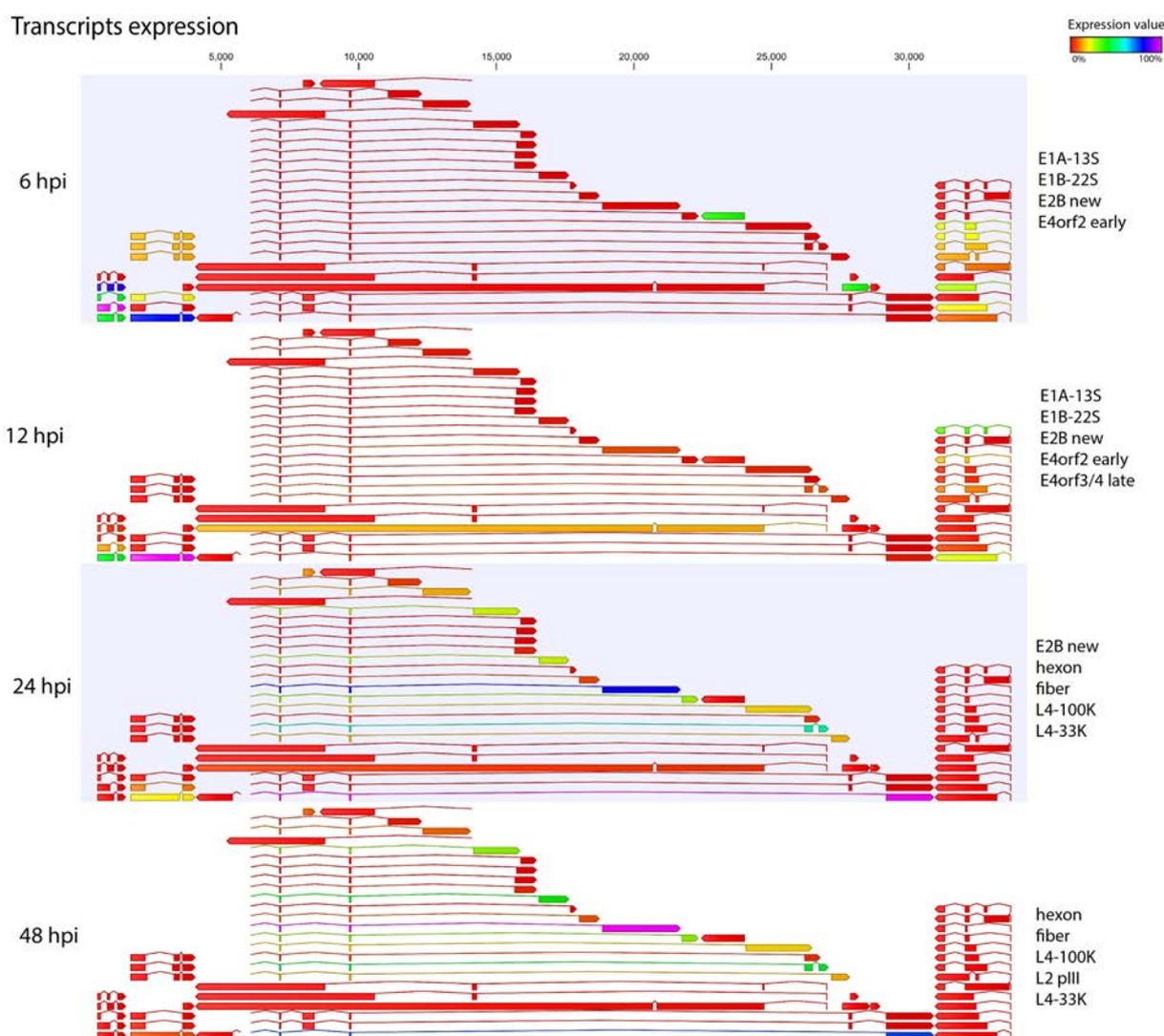


Figure 31. Representation of HAdV-5 transcript expression. Color-code representation of viral mRNA expression at different time points of infection. This image was rendered using the CLC Genomics Workbench software,

based on the expression values of a nuclear fraction representative of the changes in viral mRNA abundance. Annotated transcripts are color coded corresponding to the percentage of their abundance at each time point. The mRNAs at the right side of each diagram indicate the most abundant viral mRNAs at each time-point post infection. These transcripts were selected from the calculated total expression values in Table 2.

The main change in the viral mRNA profile clearly occurs after the entry into the late phase of infection. The viral late mRNAs exponential accumulation from 12 to 24 hpi results in their overpassing the viral early mRNAs by several orders of magnitude, to the point where most of the viral mRNA present are late mRNAs (94%), as shown in Figure 26. This change in kinetics is reflected in the profile of the most expressed viral mRNAs at different time points (Figure 31).

4.2.2 Analysis of viral mRNA cytoplasmic to nuclear rates during infection

After evaluating the kinetics and relative abundance of the total viral mRNAs, we confirmed that the RNA-Seq analysis showed the expected pattern for a HAdV-5 infection, with all the changes in expression after the transition from the early into the late phase. This reassured us that our experimental set up was suitable to further conduct a global examination showing how the transition into the late phase of infection influences the flow of viral mRNAs from the nucleus to cytoplasm. Therefore, the individual mRNA export ratios were calculated for each time-point, from the nucleoplasmic and cytoplasmic expression values. This was done by performing a Baggerley test (a proportion-based test) for which the expression values of each cellular fraction were used to compare with a reference group, in this case the nucleus fractions. Whenever the cytoplasmic values are higher than nuclear values, the ratio will be positive. In the opposite case, when the nuclear expression values are higher than those of the cytoplasm, the calculated ratio is negative. These values allowed us to know if a viral mRNA was more abundant in the cytoplasm or in the nucleus and to study if there were changes in these ratios during the time-course of infection.

To visualize how viral mRNAs' cytoplasmic to nucleus ratios (cyto/nuc) change during the time-course, ratios were clustered (see section 3.2.11) and a heat map was created (Figure 32). The main variations on the ratios occurred in the first three time points (6, 12 and 24 hpi), with 24 hpi and 48 hpi being similar. We decided to render the heat map showing only 6, 12 and 24 hpi to simplify the clustering process and interpretation.

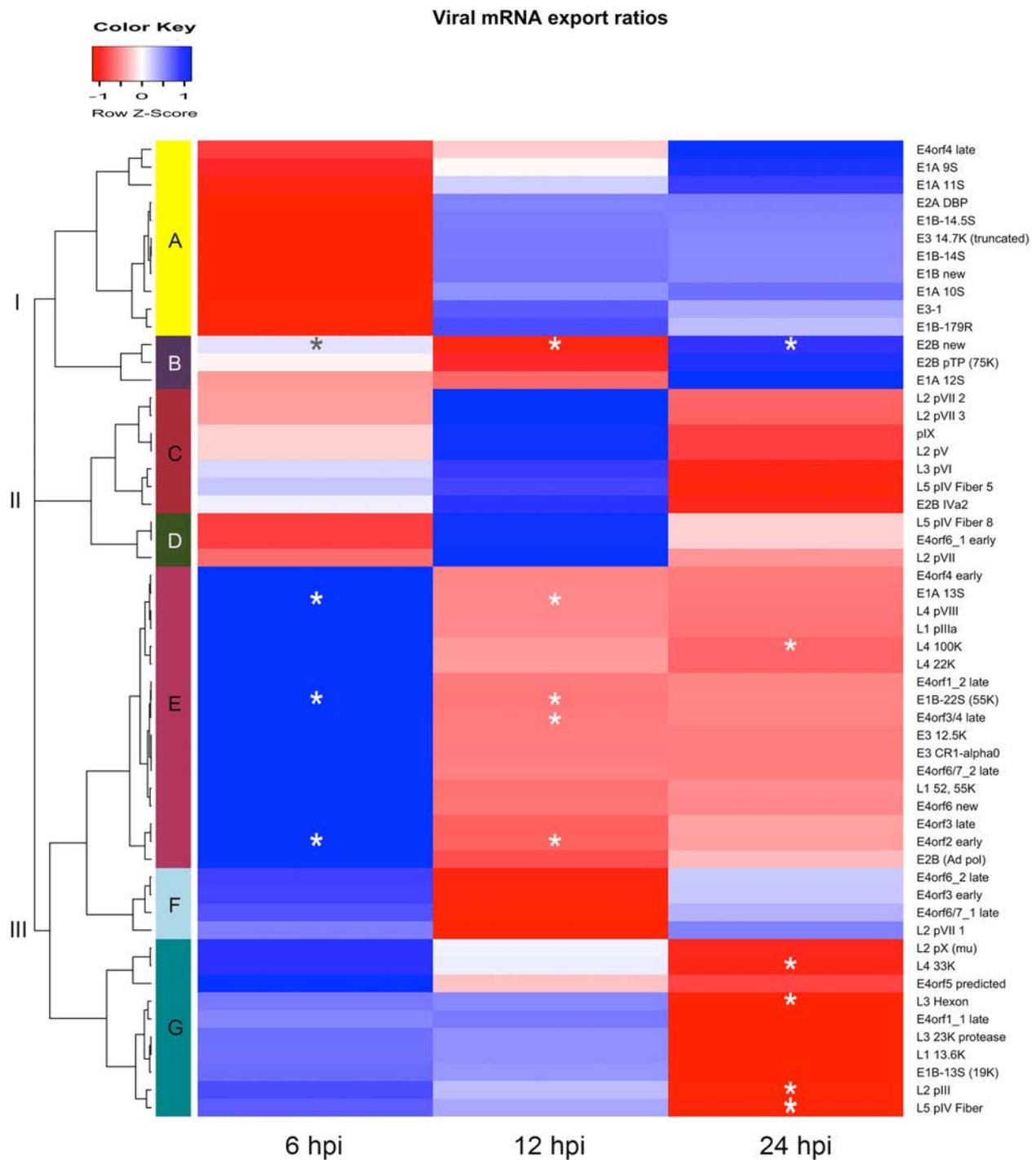


Figure 32. Viral mRNA cytoplasmic/nucleoplasmic ratios analysis. To perform the heat map above, the calculated cytoplasmic to nuclear ratios of viral mRNAs (Table 3) were z-score transformed and clustered by row using the Euclidean distance method. The resulting tree from the clustering is shown on the left side of the heat map. The three principal branches of the tree are marked with roman numerals (I-III). The seven obtained clusters are shown in different colors and marked with letters (A-G). In relation to the Color Key, mRNAs that are more abundant in the cytoplasm than in the nucleoplasm are represented towards the red side of the column z-score and, vice versa, mRNAs that are more abundant in the nucleoplasm are represented on the blue spectrum of the z-score. Variation in color represents changes, but not absolute values, of the ratios. The top most abundant mRNAs of each time-point of infection are marked with an (*).

The resulting tree from the clustering analysis has three principal tree branches (I-III), showing the main patterns of variations of the ratios within different time points. From these main branches, we obtained seven different clusters (A-G). In branch I, all the mRNAs clustered are viral early mRNAs, mainly from the E1 and E2 regions. The viral late mRNAs are all scattered between branches II and III, along with other viral early mRNAs, mainly from the E4 region (Figure 32).

Colors in the heat map represent relative values, not absolute values. Therefore, changes in color across time points can be interpreted as shifts in ratios. Interestingly, the color pattern from groups I and III at 6 hpi was the opposite of those at 24 hpi. This showed that between early and late time points of infection there was a transition in the cyto/nuc ratios pattern that started since 12 hpi. Branch I started with lower z-score values at 6 hpi and then changed to higher values at 12 and 24 hpi. Branch III showed the opposite. It started with lower z-score values and then clusters E and F changed to more positive values at 12 hpi, and cluster G also but at 24 hpi. Different to branches I and III, branch II started more towards the center of the z-score and at 12 hpi shifted to more positive values and reverted to lower values at 48 hpi. Interestingly, clusters A and E had the opposite z-score pattern.

The top most abundant mRNAs at each time-point (marked with an asterisk) were in clusters B, E and G. At 6 hpi, the most abundant mRNAs were towards the positive side of the z-score, which would suggest that the transcripts in the nucleus transition faster into the cytoplasm at that time point than at 12 hpi and 24 hpi. The viral late mRNAs at 24 hpi had a negative z-score, which means that more transcripts are detected in nucleus than at previous time points. Since, the cyto/nuc ratios were not calculated with newly synthesized mRNAs, but rather with the steady state mRNA levels, one possible explanation is that early mRNAs are expressed at levels where they are produced slightly slower than they are exported and are detected mainly in the cytoplasm. However, in the late phase (the transition from 12 to 24 hpi), expression of viral mRNAs is so high that the export machinery cannot export them as fast as they are synthesized and therefore they are detected more often in the nucleus than at earlier time-points. This would shift their cyto/nuc ratios towards more negative values. This seemed the case for some of the most abundant viral late mRNAs at 24 hpi (L3 hexon, L5 fiber and L4-33K), which were in cluster G.

In general, by analyzing the total expression values of viral mRNAs we observed that early and late mRNAs had similar kinetics within each transcriptional unit (except for the E2 gene). The viral late mRNAs had an exponential accumulation from 12 to 24 hpi resulting in them overpassing the early mRNAs by several orders of magnitude and becoming the most

abundant viral mRNAs, even when viral mRNAs continued to accumulate. When analyzing the relative proportions of viral mRNAs, the early genes showed that there is a shift in expression patterns from the early to the late phase. Proportions of viral late mRNAs changed slightly from the early to late phase, remaining the same during their exponential accumulation in the late phase. We analyzed the cyto/nuc ratios of the viral mRNAs steady states between the early and the late phase of infection. We did see changes in the cyto/nuc ratios from the early to late phase. The changes in ratios started since 12 hpi, and the pattern was fully reversed at 24 hpi and it remained the same at 48 hpi (not shown). Since our analysis did not consist of newly synthesized mRNAs we did not find the cyto/nuc ratios to be all positive in the late phase. Taking in count all our data, the transition into the late phase highly influences the viral late mRNAs' accumulation rates and the cyto/nuc ratios.

4.2.3 Analysis of cellular mRNA cytoplasmic to nuclear rates during infection

Earlier studies have described a shut-off of cellular mRNA export as a consequence of viral-induced selective export. Experiments initially used to show cellular export shut-off did not have enough resolution to measure specific mRNAs. Since then, experiments used to assess effects of viral proteins over the selective export have only focused on a very small set of cellular mRNAs and a few of viral early and late mRNAs. Therefore, we decided to use our experimental set-up to have a global evaluation of HAAdV-5 influence on cellular mRNAs.

To analyze if there is global impact of HAAdV-5 infection on cellular mRNA export, we ran an RNA-Seq analysis with our time-course experiment samples, but this time using the human genome transcriptome (hg19), as a reference to pair the sequenced reads using the CLC Genomics Workbench software. As before, nuclear and cytoplasmic samples of each time-point, and their experimental repeats, were ran in the same RNA-Seq analysis and analyzed independently but compared for further analysis. The mapping parameters used to align each read to the transcripts are described in the methods (section 3.2.11)

The mRNA export ratios were calculated for each time-point, from the nucleoplasmic and cytoplasmic expression values. This, as before, was done by performing a Baggerley test (a proportion-based test) where the expression values of each cellular fraction were used to compare with the nuclear fractions (section 3.2.11). As before, whenever the cytoplasmic values are higher than nucleus the ratio will be positive. In the opposite case, when the nuclear expression values are higher than the cytoplasmic, the calculated ratio is negative. To evaluate the impact of infection on the cyto/nuc ratios, infected samples were normalized with a none-infected sample (mock). Afterwards, we narrowed down our analysis by keeping

only the significant data, filtering with an FDR-corrected p-value ≤ 0.05 , and we obtained a total of 737 transcripts. To analyze the ratios that were changing due to infection, we filtered the data to keep the ratios that had at least a two-fold change (≥ 2). We finally obtained a set of 339 cellular mRNAs that were significantly changing during infection. The cyto/nuc ratios of these cellular mRNAs were clustered (section 3.2.11) and a heat map was generated to visualize them (Figure 33).

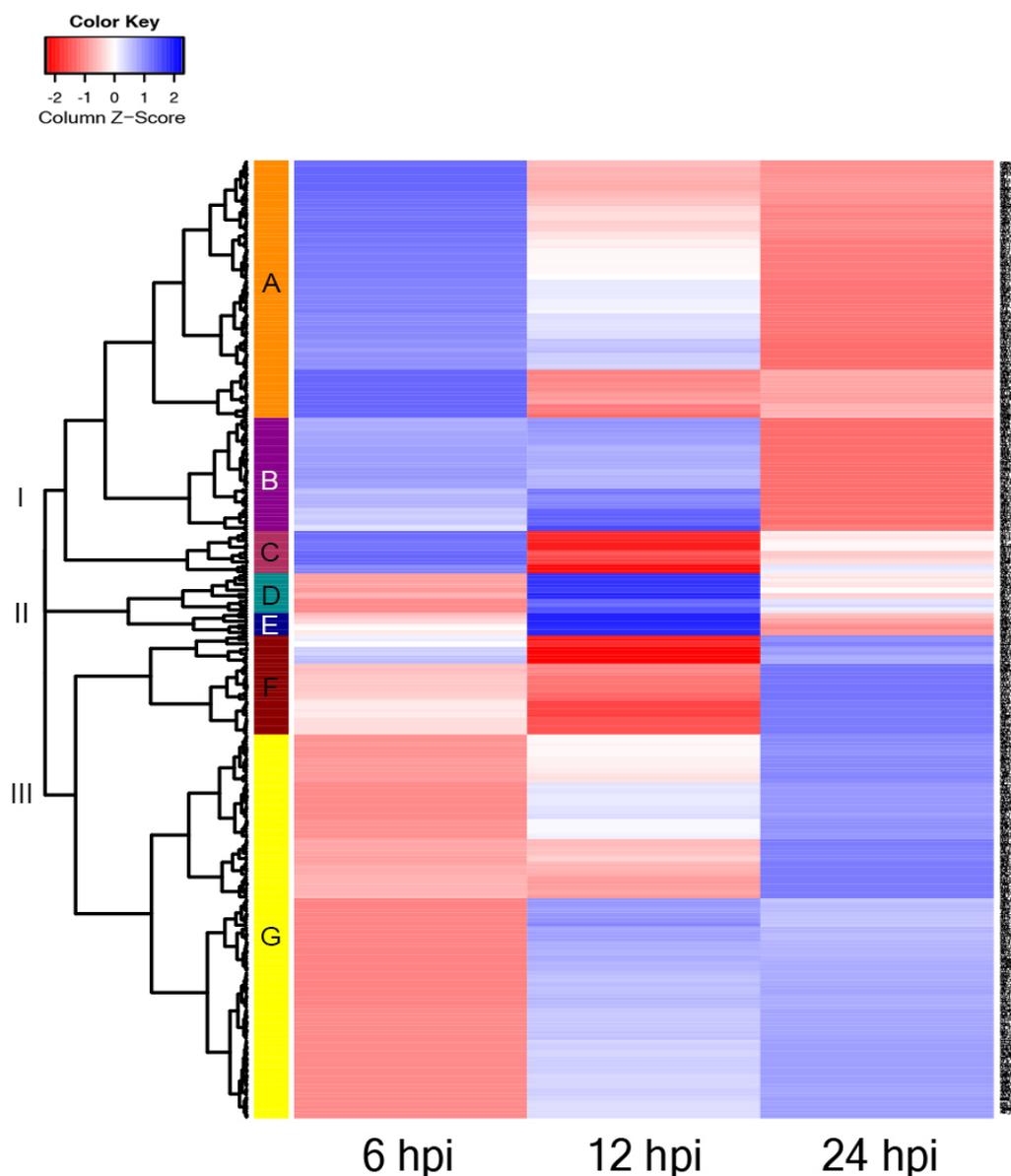


Figure 33. Cellular mRNA cytoplasmic/nucleoplasmic ratios analysis during infection. Cellular transcripts were first filtered to have an FDR corrected p-value ≤ 0.05 and a fold-change ≥ 2 . The weighted proportions fold-change of each transcript was then clustered. In relation to the Color Key, the calculated cytoplasmic to nuclear ratios of viral mRNAs where z-score transformed and clustered by row using the Euclidean distance method. The resulting tree from the clustering is shown in the left side of the heat map. The three principal branches of the tree are marked with roman numbers (I-III) and the seven obtained clusters are shown in different colors and marked with letters (A-G). The mRNAs that are more abundant in the cytoplasm than in the nucleoplasm are represented towards the red gradient side of the column z-score and vice versa, mRNAs that are more abundant in the nucleoplasm are represented on the blue spectrum of the z-score.

As a result, from the clustering analysis a tree was obtained with three principal branches (I-III) from which we obtained seven different clusters (A-G) showing the main patterns of variations of the ratios within different time points. Colors in the heat map represent relative values, not absolute values. Therefore, changes in color across time points can be interpreted as change in ratios. In general, we found that there are several changes in the cyto/nuc ratios as infection progressed into the late phase. Interestingly, the ratios in the early phase changed completely to the opposite in the late phase, and 12 hpi seems as a transition phase. In accordance to published reports describing a cellular export shut-off, we did find a group of cellular mRNAs (Figure 33, clusters A and B) that accumulate more efficiently in the cytoplasm during the early phase (positive z-score values) than after transition into the late phase, when cyto/nuc ratios were more negative. Cluster A's cyto/nuc ratios were negatively regulated since 12 hpi and cluster B's ratios until 24 hpi. A small group of mRNA ratios (cluster C) were very negatively regulated at 12 hpi and this was no longer the case at 24 hpi. However, contrary to what has been published, we also found a large set of cellular mRNAs cyto/nuc ratios that were positively regulated (clusters F and G), accumulating more efficiently in the cytoplasm during the late phase. Some of the mRNAs ratios in cluster G started to be positively regulated since 12 hpi. The mRNAs ratios in Cluster F are negatively regulated at 12 hpi but turned to be positively regulated at 24 hpi. A small group of cellular mRNAs ratios (clusters D and E) are very positively regulated at 12 hpi. Also, from the 737 mRNAs ratios that were analyzed, only 339 (46%) were changed more than two-fold during infection. These findings are interesting because they show that accumulation in the cytoplasm of cellular mRNAs is not completely blocked during the late phase, as it is believed.

4.2.3.1 Functional analysis: Networks and Classification

To further analyze which functions were associated with the cellular mRNAs from the down-regulated (clusters A and B) and up-regulated (clusters F and G) ratios showed before (Figure 33), we performed a functional analysis using the mRNAs from those clusters (Table 4). The functional analysis was done using the DAVID bioinformatics database. The significant terms obtained for each set of clusters (Figure 34 and Figure 35, panels B) showed that both set of clusters have functions associated to RNA splicing, RNA binding, translation, protein maturation, ribonucleoprotein, transmembrane region proteins, glycoproteins, phosphoproteins, ATP-binding, and post-translational modifications such as acetylation. Additionally, we performed a network analysis for each set of clusters using NetWalker2 free software (Figure 34 and Figure 35, panels A). We could show that proteins coded in the

analyzed genes of each set of clusters are highly interconnected in several protein-protein interactions, metabolic, and Reactome pathways.

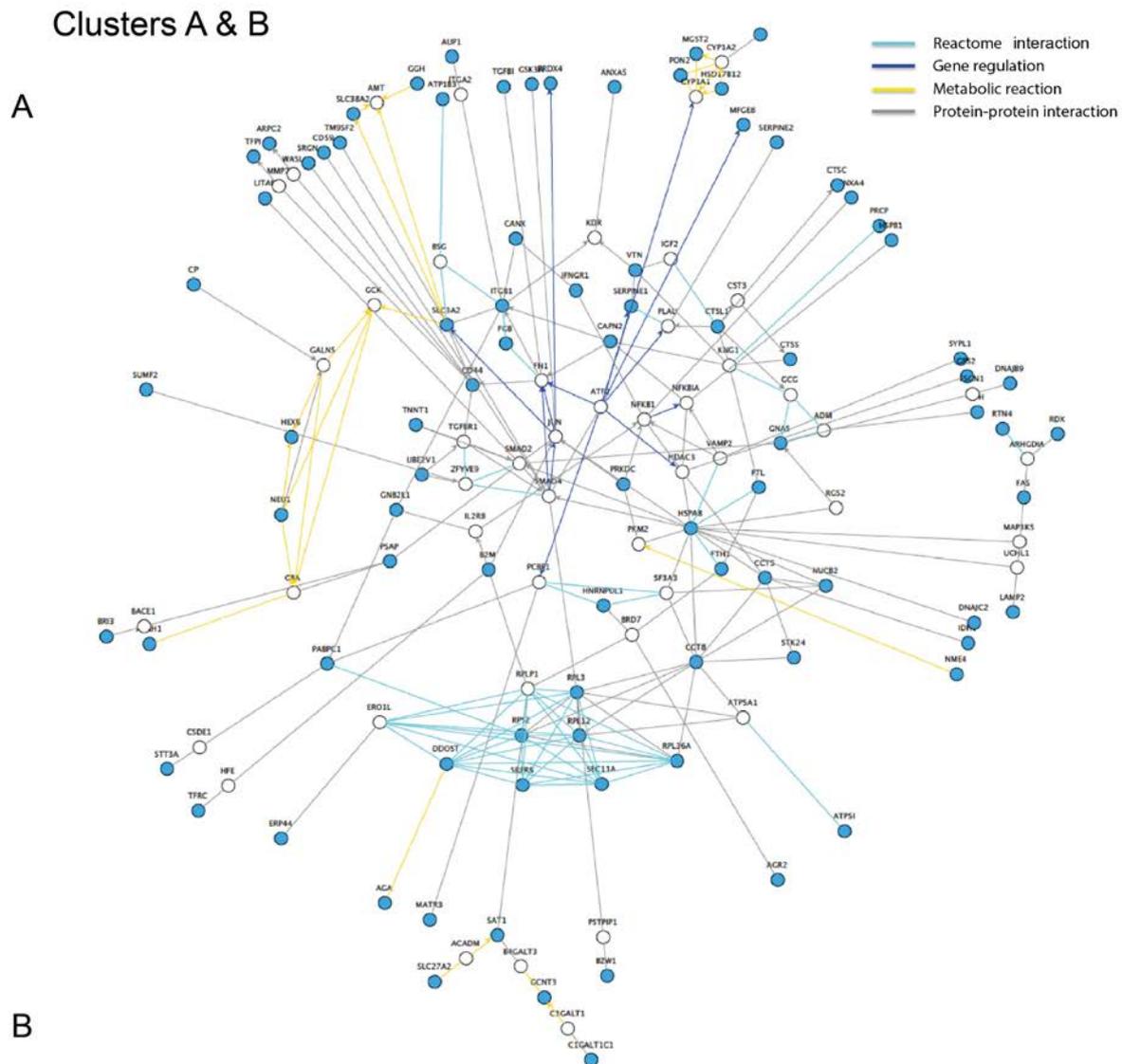


Figure 34. Clustering and functional analysis of differentially regulated mRNAs in clusters A and B. A) Cellular mRNAs from clusters A and B (Table 4) were used to perform a functional network of genes, using NetWalker2 software. The clustered genes are represented by the blue nodes and the lines connecting the gene nodes in the network are color coded depending on the type of interaction. B) Cellular mRNAs from clusters A and B (Table 4) were used to perform a functional analysis using the DAVID data base. The table shows a summary of the associated functions and their enrichment scores.

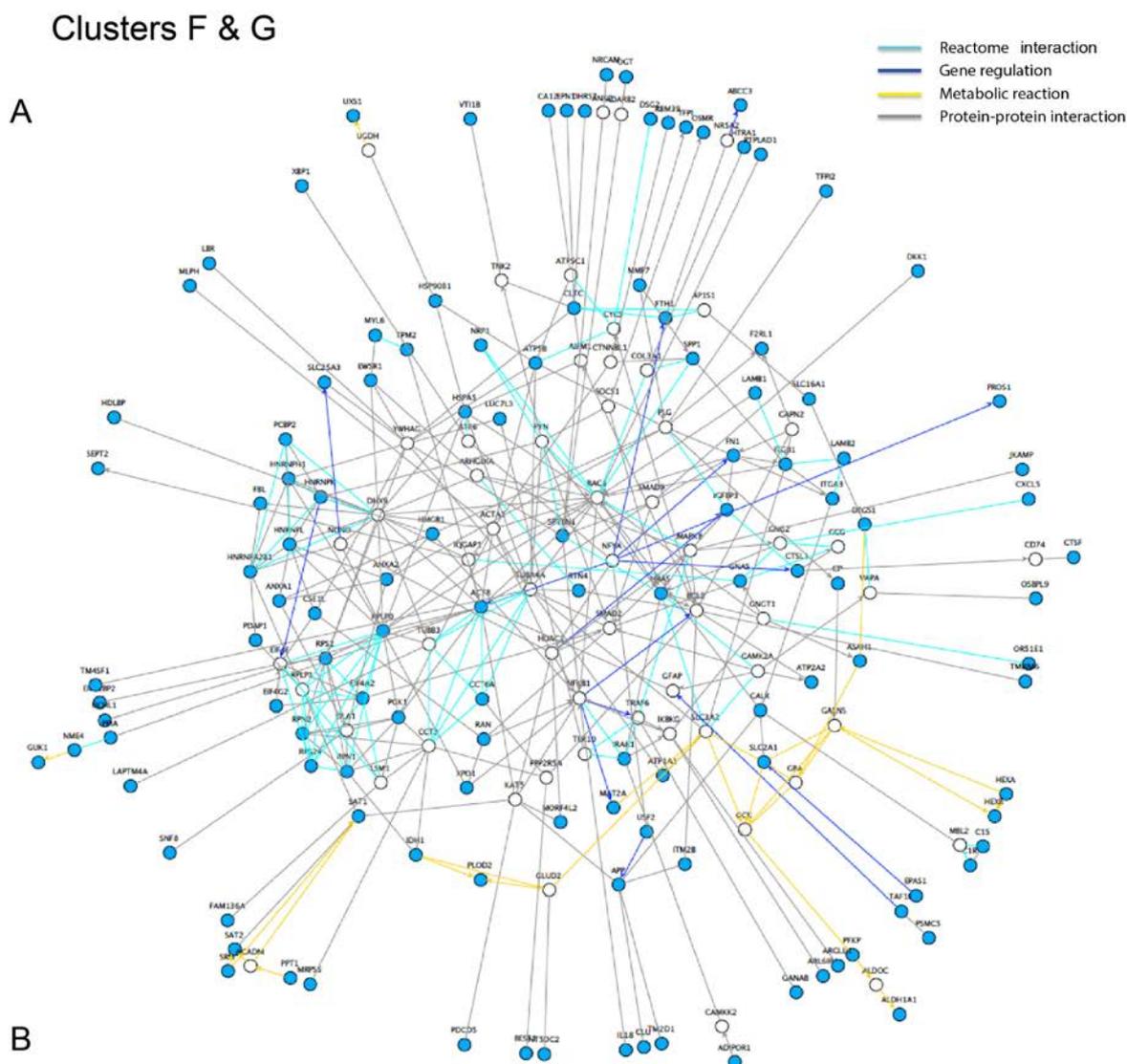


Figure 35. Clustering and functional analysis of differentially regulated mRNAs in clusters F and G. A) Cellular mRNAs from clusters F and G (Table 4) were used to perform a functional network of genes, using NetWalker2 software. The clustered genes are represented by the blue nodes and the lines connecting the gene nodes in the network are color coded depending on the type of interaction. B) Cellular mRNAs from clusters F and G (Table 4) were used to perform a functional analysis using the DAVID data base. The table shows a summary of the associated functions and their enrichment scores.

Altogether, this data shows that towards the late phase of the infection, similar cellular functions are regulated at an mRNA export level, which are part of common pathways mainly involved in RNA splicing and protein biogenesis.

4.2.4 Comparison between viral and cellular expression during the late phase of infection

As previously shown, the entry into the late phase resulted in a very high production of viral late mRNAs, accumulating in orders of magnitude $\sim 10^7$ RPKMs (Figure 30). We also showed that not all cellular mRNAs are blocked in the late phase. In fact, we could detect several cellular mRNAs that had an increase in their cyto/nuc ratios in the late phase (Figure 26). As it has been previously reported that very few cellular mRNAs accumulate in the cytoplasm during the late phase, our data suggested that this could be a consequence of a mass effect due to the high synthesis rates of viral mRNAs during the late phase. To address this, a comparison between the total reads of nuclear and cytoplasmic contributions from cellular and viral mRNAs was made (Figure 36).

The total reads assigned to viral transcripts increased in the late phase as expected and in accordance to our previous data. However, this analysis showed that the number of cytoplasmic total reads is greater than the nucleoplasmic and this effect increases during the late phase (Figure 36, panel A). At 24 hpi, 57% of the viral total reads were cytoplasmic and at 48 hpi it increased to 72%. This would indicate that viral transcripts are effectively accumulating in the cytoplasm during the late phase as previously reported. To analyze the total cellular reads, all cellular transcripts were filtered for a p-value ≤ 0.05 and the final total reads were counted at each time-point. This showed that in non-infected cells, they were $\sim 3 \times 10^7$ total reads (Figure 36, panel B). At 6 hpi, the cellular total reads significantly diminished to 1.4×10^7 and they increased again at 12 hpi to 2.4×10^7 . The increase in total reads from 6 to 12 hpi is probably an effect of the transcriptional activation by E1A (see section 1.3.1). At late time points, 24 and 48 hpi, cellular total reads diminished again to 1.8×10^7 and 1.3×10^7 respectively. This represents an 1.8 times decrease on cellular total reads from 12 hpi to 48 hpi. When analyzing the contribution of nucleoplasmic to cytoplasmic reads we found that non-infected cells had $\sim 57\%$ of the total cellular reads in the cytoplasmic fraction. This diminished to 46% and 42% at 24 and 48 hpi, respectively. Therefore, there is 15% less of the total cellular reads in the cytoplasmic fraction at late time points. This would indicate that both, the cellular total reads and the percentage of cellular transcripts in the cytoplasmic fraction, are decreasing during the late phase. However, the decrease of the total cellular reads has a higher contribution. When evaluating the contribution of both viral and cellular total reads at

each time point it was found that the number of total reads increased at late time points (Figure 36, panel C). At 24 hpi the total reads were slightly above the number of total reads in non-infected cells; however, at 48 hpi the total reads were almost twice the amount. During the early phase (6 and 12 hpi), the cellular reads comprised almost all the analyzed total reads. In contrast, in the late phase, almost half of the total reads were viral at 24 hpi; and at 48 hpi viral reads were almost five times the number of total cellular reads. This again showed that the decrease of the cellular total reads combined with the increase in viral total reads has a larger contribution than the decrease in cytoplasmic percentage of cellular reads. The percentage of cytoplasmic and nucleoplasmic total reads, of both cellular and viral contributions was calculated (Figure 36, panel D). The analysis showed that in non-infected cells and the early phase of infection, more than 50% of the cellular total reads are cytoplasmic. However, at late time points 70% of the total reads are cytoplasmic, but 60% are viral and only 10% are cellular.

To further evaluate how HAdV-5 infection affected specific cellular mRNAs, we chose the top 35 most abundant mRNAs at each time-point of infection (Table 5) and plotted them together to compare (Figure 37). As it would be expected, the top cellular mRNAs changed during infection in position of expression levels, and some disappeared or appeared in the list as infection progressed (Table 5). However, the important effect found was that the overall abundance of the most expressed cellular mRNAs diminished during the late phase (Figure 37, 24 and 48 hpi). The expression levels of the selected cellular mRNAs diminished ~8.7 times from 12 hpi to 48 hpi. Therefore, when competing with viral late mRNAs for splicing, export and translation machinery they would be expected to be outnumbered.

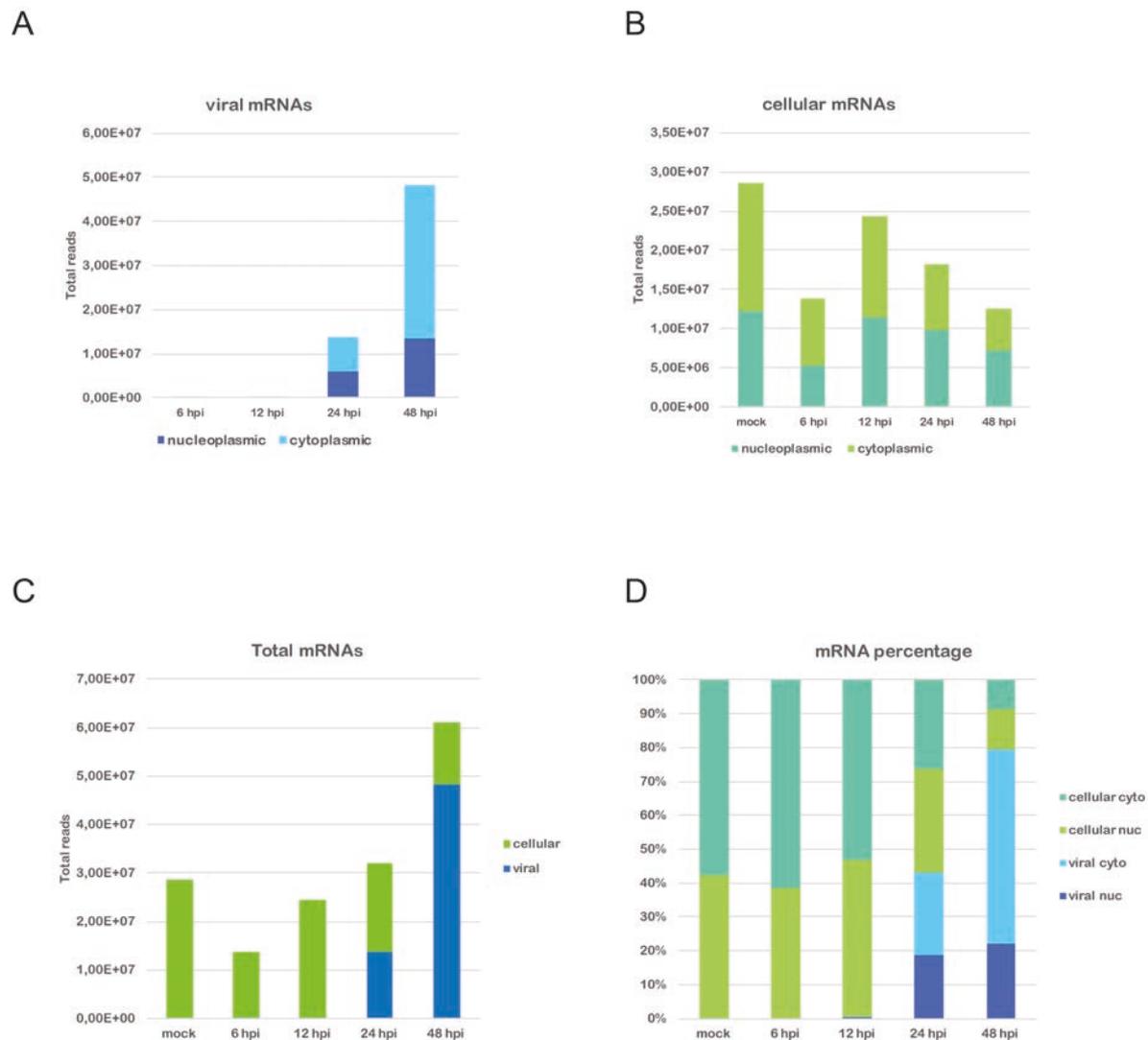


Figure 36. Comparison between cellular and viral total reads and their nuclear and cytoplasmic percentages at different times post infection. The reads assigned to viral (panel A) and cellular transcripts (panel B) are represented as bar plots showing the proportions of nucleoplasmic and cytoplasmic reads at each time point. The total number of reads detected at each time point are displayed as a bar plot showing the contribution of viral and cellular reads (panel C). A comparison between nucleoplasmic (nuc) and cytoplasmic (cyto) reads of cellular and viral reads at each time point are plotted and represented as percentages (panel D).

To evaluate if the most expressed cellular mRNAs are at comparable expression levels with the viral late mRNAs during the late phase, we plotted together their total transcript expression values at late time points of infection (Figure 38). Interestingly, we found that at 24 hpi, when total cellular mRNAs were ten times more abundant than viral mRNAs (Figure 36, 24 hpi), most viral late mRNAs had comparable expression levels (10^5 - 10^6 RPKMs) with the most expressed cellular mRNAs (Figure 38, panel A). At 48 hpi, when total viral mRNAs were almost the same as total cellular mRNAs (Figure 36, 48 hpi), most viral late mRNAs (taking in count the different isoforms or alternative splicing products) were more abundant than the cellular mRNAs, having expression levels from 10^6 - 10^7 RPKMs (Figure 38, panel B). The most expressed viral late mRNAs also correlated with the most abundant capsid proteins (see Figure 2, panel C), such as L3 Hexon, L5 Fiber (all isoforms), L2 pVII (all isoforms), L1 pIIIa, L2 pV, L2 pIII, and L4 pVIII. Viral

proteins involved in translation, protein folding, and capsid assembly, such as L4 100K, L4 33K and L4 22K, were also more abundant than cellular mRNAs (Figure 38, panel B).

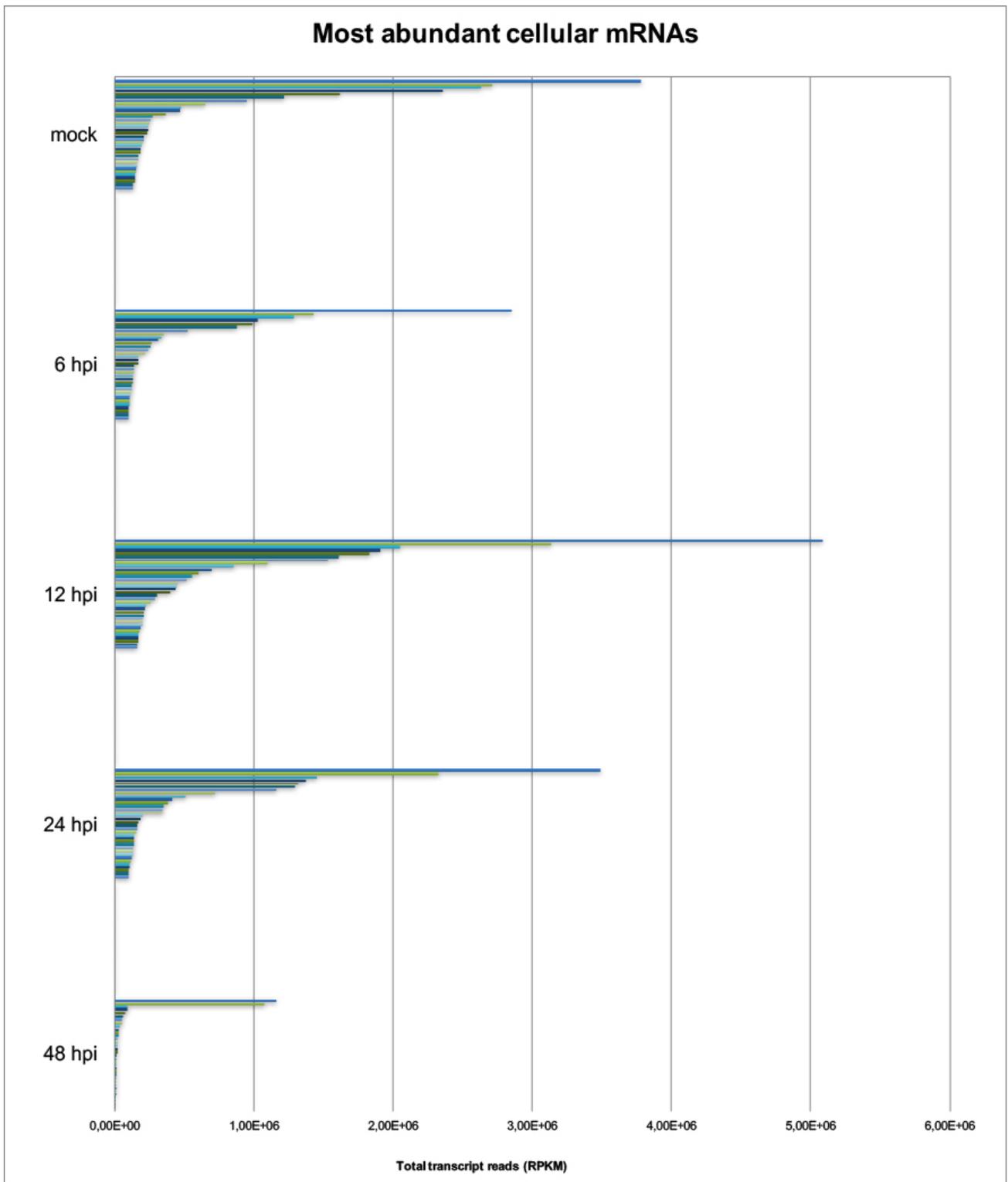


Figure 37. Top most abundant cellular mRNAs expression during infection. Bar plot showing in parallel the total transcript reads (RPKM) of the top 35 most abundant cellular mRNAs at each time of infection (Table 5).

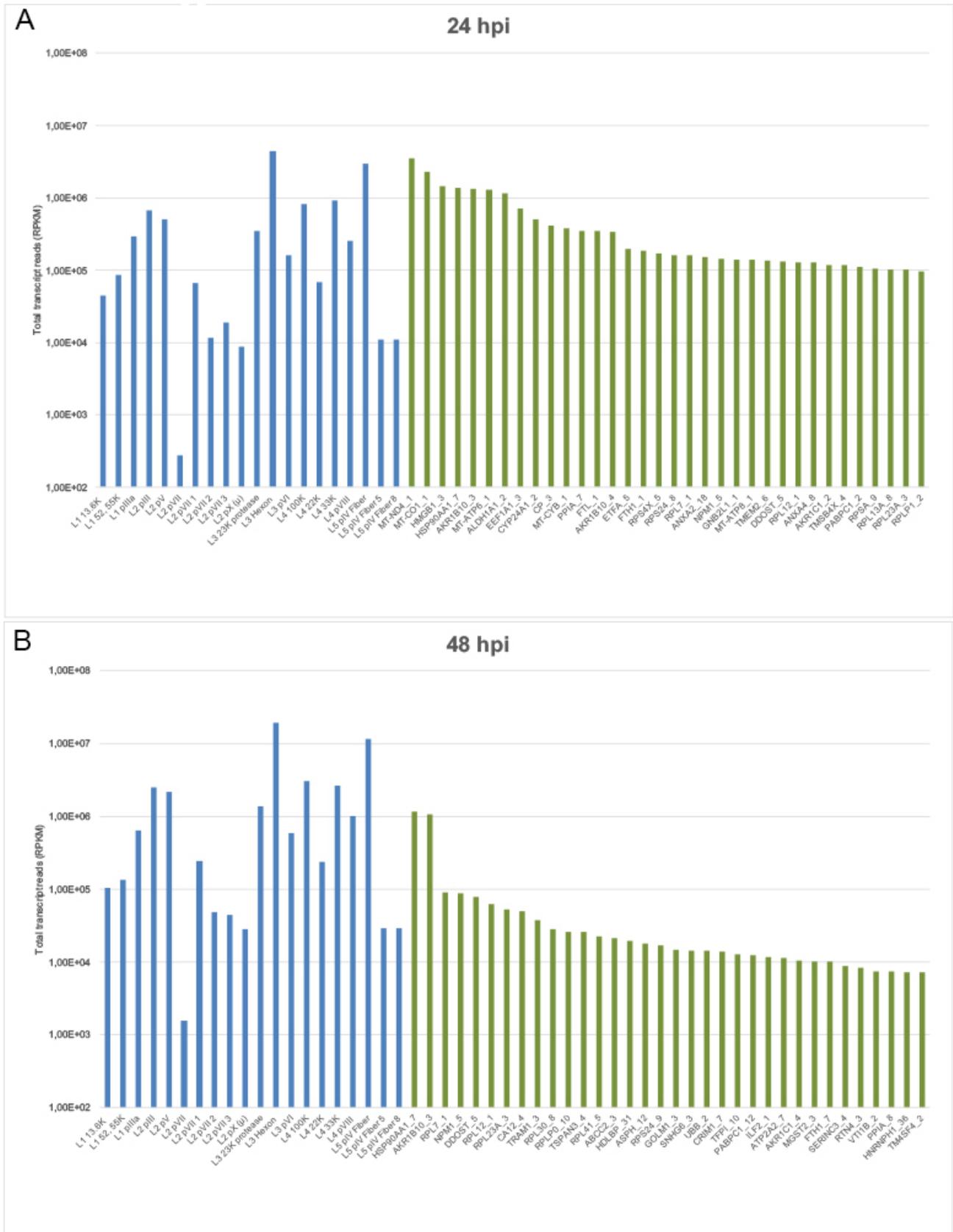


Figure 38. Comparison between viral late mRNAs and most abundant cellular mRNAs. Bar plots showing the total expression levels of all viral late mRNAs (blue) and the most abundant cellular mRNAs (green) at 24 hpi (panel A) and at 48 hpi (panel B).

Interestingly, two cellular mRNAs were very abundant at 48 hpi, in levels comparable to the viral late mRNAs. One was the mRNA for HSP90A, a well-known protein chaperone. The second one was the mRNA for AKR1B10 and Aldo-keto reductase. More of the top expressed cellular mRNAs had functions in translation and protein folding (Table 6 and Table 7). The mRNA NPM1 coding for the nucleolar protein nucleophosmin (B23) was also found in these highly expressed cellular mRNAs. Nucleophosmin has been reported to be relocalized from nucleolus to the nucleoplasm by protein V. It has also been shown to be associated with and to be a stimulatory factor for viral DNA replication (Matthews, 2001; Okuwaki *et al.*, 2001).

Taken together, these results showed that accumulation of cellular mRNAs in the cytoplasm is highly reduced during the late phase (Figure 36). At these time points, most of the cytoplasmic reads are associated to viral transcripts. We observed both a reduction of cellular total reads and of the percentage of cellular cytoplasmic reads. Therefore, less cellular mRNAs accumulate in the cytoplasm during the late phase. We suggest that this is mostly due to the global reduction of the total reads as the reduction in the percentage of cytoplasmic reads was only 15%. Cellular mRNA cytoplasmic accumulation was also highly impacted by the exponential accumulation of viral late mRNAs from 12 to 24 hpi and by the increase in cytoplasmic viral total reads (which is also around 15%) at the same time points. Due to all the above, viral late mRNAs are more abundant in the cytoplasm than cellular mRNAs at 48 hpi (Figure 36). The global reduction of cellular mRNAs is further confirmed when analyzing the most expressed cellular mRNAs (Figure 37). When individual mRNAs were analyzed, most viral late mRNAs were more abundant than cellular mRNAs at 48 hpi (Figure 38, panel B).

5 Discussion

5.1 Contribution of the E1B/Orf6/E3 Ubiquitin ligase in degradation of proteins involved in mRNA biogenesis.

Selective export of viral mRNAs during the late phase was first reported in the early 1980s when it was found that cellular mRNAs failed to accumulate in the cytoplasm during infection with HAdV-2 (Babich *et al.*, 1983). However, the molecular mechanism behind this viral effect remains unknown to this date. Studies showed that the E1B/Orf6/E3 ubiquitin ligase contributed to this effect when using Cullin 5 dominant-negative cells and Orf6 mutants, that could not assemble the viral ligase, were defective in accumulating viral late mRNAs L3 and L5 in the cytoplasm (Woo and Berk, 2007; P. Blanchette *et al.*, 2008). Therefore, it has been proposed that this complex sends cellular proteins involved in mRNA export for degradation, resulting in viral late mRNAs being exported to the cytoplasm more efficiently than cellular mRNAs. So far, none of the known targets of the viral ubiquitin ligase participate in processes that could directly affect mRNA export.

In this work, we attempt to search for possible degradation candidates for the E1B/Orf6/E3 ubiquitin ligase, thereby selected proteins known to be involved in viral late mRNA export such as Aly and Tap (Yatherajam *et al.*, 2011), as well as CRM1, which has been reported to contribute during viral early mRNA export (Schmid, R. Gonzalez, *et al.*, 2012), were investigated. We included as well RNA-binding proteins that participate in mRNA processing and alternative splicing, such as hnRNP L and hnRNP M, which are previous steps that contribute to mRNAs being selected for cytoplasmic export; and AP5 (HNRPUL1), which interacts with TAP and has been suggested to play a role in RNA transport (Gabler *et al.*, 1998). We also included the translation initiation factor eIF4E, which has been found in the nucleus where it is involved in mRNA export and promotes cap-dependent translation (Culjkovic *et al.*, 2005, 2006). In our experiments, none of the analyzed cellular targets were degraded. The steady state time-course analyses showed either a constant expression during infection or, in the case of hnRNP L and hnRNP M, an increased signal in late time-points of infection (Figure 17). We could detect lower migrating bands for TAP, Crm1, and hnRNP M at late time points. These lower migrating bands could appear due to degradation or proteolysis. However, stability assays showed no degradation for any of the selected cellular proteins (Figure 18). Therefore, whether the E1B/Orf6/E3 ubiquitin ligase has a role degrading cellular proteins

involved in mRNA biogenesis, resulting in selective export of viral late mRNAs, remains unknown.

It has been shown that there is a dependence of the E1B/Orf6/E3 ubiquitin ligase for selective export by showing a decrease in cytoplasmic viral late mRNA levels with stable total mRNA levels (Woo and Berk, 2007). Therefore, it is inconclusive whether defects in the viral ligase affect the overall expression levels of viral late mRNAs or only the mRNA export. It is possible that cellular proteins involved in different steps of mRNA biogenesis that have not yet been analyzed could be selected for degradation. Other possibilities also exist, such as targets of this viral ligase not being directly involved in mRNA export but still influencing it during the late phase. Several of the reported targets of the viral ligase are involved in inhibiting the host-cell response to viral infection (such as p53, MRN complex, DNA ligase IV, and integrin α 3) (see section 1.4). Therefore, it is possible that defects on the E1B/Orf6/E3 ubiquitin ligase overall affect the efficient inhibition of the host's innate immune response and thereby progression into the late phase and efficient production of viral late mRNAs.

5.2 hnRNP M is hyper-SUMOylated at late times of infection with HAdV-5

In our quest to find targets for the E1B/Orf6/E3 ubiquitin ligase, we observed higher bands at late infection time-points of some of the analyzed proteins (Crm1, TAP, hnRNP L, and hnRNP M) by WB (Figure 17). Due to the size of these bands, we suspected of SUMOylation. Each SUMO protein weighs ~12 kDa, which are conjugated to lysine residues on proteins. Thereby, several SUMOs can be added to one protein, and poly-SUMO chains can be added into a specific site. Furthermore, a protein can be conjugated via a lysine residue onto another protein's SUMO chain resulting in formation of very complex molecules (Bergink and Jentsch, 2009). SUMO modifications are reversible and are known to be implicated in regulating nuclear events, making them ideal for regulatory purposes.

In this work, we tested if the detected higher bands were SUMO modifications of the respective proteins using a SUMO pull-down assay. In our experiments, only hnRNP M showed to be modified by SUMO. However, TAP and hnRNP L have previously been identified in SUMO proteomic studies (Hendriks *et al.*, 2015; Richard, Vethantham and Manley, 2017). hnRNP M had previously been reported to be modified by SUMO-1, forming a triplet of ~90 kDa corresponding to addition of SUMO into a single lysine residue. Further, this SUMOylated form of hnRNP M copurifies with isolated nuclear pore complexes (NPCs) (Vassileva and Matunis, 2004). We could detect the SUMOylated triplet in non-infected cells as well as during infection (Figure 19). In our studies, we also found that hnRNP M is modified by SUMO-2, appearing as a

band migrating at a slightly smaller molecular weight than the SUMO-1-hnRNP M triplet, therefore we infer it is also a single SUMO-2 addition. Interestingly, we found that at late time-points of infection the single SUMOylation band shifted into poly-SUMO modification of SUMO-2 monomers. We also found that at late time-points of infection, when higher SUMOylated forms of hnRNP M are detected, the SUMO-1 band decreases. In vitro reactions with hnRNP M and SUMO-1 have also shown the poly-SUMO chains and reduction of the single SUMO-1 band (Vassileva and Matunis, 2004).

5.2.1 Hyper-SUMOylation of hnRNP M depends on the presence of E1B-55K

The E1B protein is known to act as a SUMO E3 ligase that targets cellular proteins for SUMOylation (described in section 1.8). Therefore, we were interested to know if E1B influenced the discovered hnRNP M poly-SUMO-2 modification. Our experiments show that this was the case since we could only detect higher forms of SUMO-2-hnRNP M during infection with a wt HAdV-5 and not when using an E1B null-mutant virus (Figure 20). Since this mutant is known to have a delayed entry into the late phase, we extended the infection period to 72 hpi. Nevertheless, higher SUMO-2 forms of hnRNP M could not be detected in the absence of E1B, even after 72 hpi. Therefore, we could confirm that higher forms of hnRNP M-SUMO-2 depend, directly or indirectly, on the presence of the E1B. These results are in congruence with an unpublished SUMO proteomic analysis from our laboratory, performed by Dr. Julia Berscheminski, where hnRNP M is present in a list of proteins that are highly SUMOylated in the wt infection, compared to the E1B null-mutant virus (wt virus/-E1B virus). This has also been confirmed in our group by colleague Tina Meyer in several WB analyses.

E1B is known to be phosphorylated and SUMOylated during infection; and these post-translational modifications can alter its subcellular localization, interaction with proteins, and its different known functions (Endter *et al.*, 2001; Ching, Dobner and Koyuncu, 2012) Therefore, we were interested to know if any of the post-translational modifications of E1B affected our previous observation of E1B influencing SUMOylation of hnRNP M. In this work, we found that a phosphorylated E1B favors the hnRNP M poly-SUMO-2 modification (Figure 21). The E1B is both a substrate for SUMOylation and can act as a SUMO E3 ligase (described in section 1.8); and phosphorylation of E1B C-terminus has been reported to enhance its own SUMOylation (Wimmer *et al.*, 2013). Interestingly, we also observed that in infection with an E1B mutant virus that could not be SUMOylated, higher forms of SUMO-2-hnRNP M could not be observed. These observations suggest that during infection, an E1B that is phosphorylated, and therefore SUMOylated, favors the hyper-SUMOylation of hnRNP M by SUMO-2 in the late phase. Whether the E1B itself is directly responsible for targeting hnRNP M for SUMOylation remains unanswered.

Co-immunoprecipitation assays between E1B and hnRNP M were tested; however, our data were inconclusive (data not shown).

The role of hnRNP M in the context of HAdV infection is not yet understood. However, the hnRNP M is known to be a splicing regulatory protein that can enhance or silence splicing of alternatively spliced exons (known as exon skipping and exon inclusion) (Hovhannisyan and Carstens, 2007). This function is highly required during the late phase, as the synthesis of viral late mRNAs requires alternative splicing to produce the different mRNAs coded in each of the late mRNA families (described in section 1.5.2). However, the SUMO-modified version of hnRNP M has been reported to be associated with the NPC; and it has been proposed that it plays a role helping mRNPs complexes organize at the NPC and facilitating their nuclear export (Vassileva and Matunis, 2004). This is a very interesting function proposed for hnRNP M but has not yet been confirmed. However, as it has been described in the introduction (section 1.5), viral late mRNAs are very abundant during the late phase and must be efficiently exported to the cytoplasm. Therefore, both functions associated with the hnRNP M protein would favor viral late mRNA production and export, and it would be logical that such functions would be exploited during infection. Further experiments are required to assess whether hnRNP M is associated with NPC during infection with HAdV-5.

5.3 hnRNP M is localized towards the viral RCs during infection with HAdV-5

SUMOylation is associated with redistribution of nuclear proteins. Therefore, in our work, we decided to evaluate whether the cellular distribution of hnRNP M changed during infection using immunofluorescence analyses (Figure 22). Our results showed that in non-infected conditions, hnRNP M has a diffused nucleoplasmic distribution and it is excluded from the nucleoli as previously reported (Gattoni *et al.*, 1996). However, during infection, and as RCs are formed in the infected nuclei, a fraction of the nucleoplasmic hnRNP M was found to accumulate in the surrounding areas of DBP distribution. These areas are named peripheral replication zone (PRZ) and are known to be the sites of viral late mRNA transcription (described in section 1.6). Interestingly, AP5, another hnRNP, is known to completely relocalize towards these sites during infection (Blackford, Bruton, Dirlík, G. S. Stewart, *et al.*, 2008). As the PRZ are the sites of viral late translation, it would be expected for hnRNP M to localize in these sites if it would participate in alternative splicing processes of viral late mRNAs. The hnRNP M has previously been reported to associate with nuclear speckles in non-infected cells (Marko *et al.*, 2010). These nuclear structures are the sites where cellular splicing machinery accumulates during infection, and they have been found to redistribute and relocalize towards the RCs (described in section 1.6) (Bridge *et al.*, 1993). Interestingly, hnRNP M distribution changes

correlate with time-points in which it is modified by SUMOylation. The distribution of SUMO-2/3 during infection with HAdV-5 in non-transformed cells has been reported to colocalize with DBP (Castillo-Villanueva *et al.*, 2014). However, further studies must be performed to test whether the hnRNP M surrounding the RCs is the SUMO-modified hnRNP M.

In this work, we could not detect any signs of degradation in the analyzed hnRNPs. However, we suggest that they could play a significant role during infection since we could detect that hnRNP M is modified by SUMOylation in an E1B-dependent manner. Further, we could also find that hnRNP M relocates towards the sites of viral late mRNA transcription, at the periphery of the RCs. Also, in our NGS experiments, several hnRNPs, such as hnRNP K, hnRNP DL, hnRNP U-L1 (AP5), hnRNP A-2B1, hnRNP H1, and hnRNP L, were found during the cellular cyto/nuc clustering analysis (Table 4).

5.4 Global analysis of viral and cellular mRNAs during infection with HAdV-5

The first part of this work centers on the contribution of the E1B/Orf6/E3 ubiquitin ligase in the viral late mRNA selective export by degrading key cellular components involved in mRNA biogenesis. The second part of this study focuses completely on the phenomenon of selective export. As previously described in the introduction (section 1.5.3.1), selective export was proposed after finding that cellular protein synthesis is reduced by more than 90% during the late phase. However, cellular mRNA expression is not inhibited and cellular mRNA accumulation in the cytoplasm decreased (Babich *et al.*, 1983). Most of the early data collected to report on this, is based on RNA/DNA and RNA/RNA hybridization experiments. Initially, unspecific probes were used to measure cytoplasmic presence of viral and cellular mRNAs (e.g. cellular cDNA clones and fragments of HAdV genome). However, these types of experiments do not have enough specificity and detection sensibility. Most of the later experiments measured individual mRNAs by RT-PCR which improves in sensibility, but it does not allow to measure enough cellular and viral probes to get a full understanding of how mRNAs behave during infection. Therefore, only a small number of cellular mRNAs (e.g. Hsp70, β -actin, tubulin, 6-16, and Mx-A) have been used to compare with the most abundant late viral mRNAs (e.g. L2, L3 and L5). The objective of this thesis was to analyze selective export in a global manner, therefore, for our experiments, we decided to use Next Generation Sequencing, which allowed us to analyze simultaneously all viral and cellular mRNAs in a time-course infection with HAdV-5. The approach was to analyze the total mRNA expression values as well as the cyto/nuc ratios.

5.4.1 Time-course analysis of viral mRNAs

Initial analysis of viral mRNAs showed that during the early phase, at 6 hpi, all reads mapped at the ends of the viral genome in early regions E1, E2, E3 (remaining regions) and E4. This is in accordance to previous reports using HeLa cells infected with HAdV-2 and different experimental approaches for mapping (such as R-loop mapping and RNA-DNA hybridizations), in which they demonstrate that region E1A and E4 are the first to be activated, followed by E3 and E1B, and finally E2 (Neuwald *et al.*, 1977; Nevins *et al.*, 1979). Our experiments show the same for exception of E3, which is deleted in our wt HAdV-5. In the late phase (12 and 24 hpi), almost all the reads mapped to the Major Late transcription unit and the tripartite leader (Figure 24). The change in patterns in our histograms during the time-course infection with HAdV-5 are very similar to R-loops patterns generated by early and late HAdV-2 RNA in a previous report (Neuwald *et al.*, 1977).

Further, we found that in the early phase, at 6 hpi, 5% of the total mRNAs was viral mRNA and at 12 hpi, this incremented to 9%. However, in the late phase at 24 hpi, 85% of the total mRNAs was viral mRNA, and at 48 hpi, 94% of the total was viral mRNA. These results are similar to previous reports using KB cells infected with HAdV-2, from which poly(A)-RNAs are isolated from polyribosomes. Thus, they find that in the early infection 16% is viral RNA, and in the late phase more than 90% of polyribosomal poly(A)-RNA in the infected cells was viral (Lindberg, Persson and Philipson, 1972). Similarly, a second study, in which they analyzed total poly(A)-RNA, showed that at early time points 5% of the total poly(A)-RNA is viral, and at late times it is more than 80% (Tal *et al.*, 1975). Therefore, our results support previous publications, in which the early-to-late transcriptional switch, has been described in HAdV infection (Nevins, 1987).

5.4.1.1 Viral early mRNAs

Quantification by RNA-Seq allowed us to analyze the kinetics and proportions of individual viral mRNAs along the time-course infection. In this work, the RNA-Seq data was normalized to Reads Per Kilobase of transcript per Million mapped reads (RPKM), however, we are aware that better normalization methods have now been proposed for Next Generation Sequencing data (Dillies *et al.*, 2013). Our experiments showed most early mRNAs have similar kinetics as other members of their gene family, except the E2 mRNAs. The E1A and E4 mRNAs, increase gradually during infection while E1B mRNAs and E2B IVa2 have a drastic increase of more than two orders of magnitude. Comparison of our data with other individual viral early RNA profiles is complicated since very different experimental conditions are used (e.g. cell-lines; time-points; type of RNA measured, such as total or cytoplasmic; type of measurement, such as

synthesis rates or steady states; and method used for mapping, such as DNA-RNA hybridizations or S1 nuclease mapping). However, other RNA profiles have also shown that the early mRNAs have different kinetics (Nevins *et al.*, 1979; Glenn and Ricciardi, 1988). Methods used previously for early mRNAs mappings do not allow for quantitative comparisons between them. However, our data showed that expression levels of mRNAs vary within each early family and between them. In general, expression levels of early viral mRNAs do not reach more than six orders of magnitude.

Our analysis of the kinetics and proportions of the E1A mRNAs showed that E1A 12S and 13S are more abundant in the early phase. In the late phase, expression of the E1A 9S is increased and becomes the most abundant E1A mRNA (Figure 27 and Figure 28). These results are in accordance with a previous publication in which they measure the appearance of cytoplasmic RNA in HAdV-2 infection, and show that E1A 9S appears only after viral DNA replication begins while E1A 12S and 13S are produced at both early and late times post-infection (Spector, McGrogan and Raskas, 1978). Similarly, our analysis of the E1B proportions and kinetics showed a change in expression from early to late phase of infection cycle. In the early phase, the E1B 55K is the most abundant mRNA, and after transition into the late phase the E1B 19K levels increase. Previous reports have also shown this effect in different transformed cells infected with HAdV-2 (Berk and Sharp, 1978; Spector *et al.*, 1978; Wilson *et al.*, 1978). Different mRNA stabilities have also been reported in such E1B mRNAs during the late phase, accounting for the increase of E1B 19K mRNA (Wilson and Darnell, 1981).

The proportion analysis within the E2 mRNAs showed a shift between E2A DBP and E2B IVa2 mRNAs from the early to the late phase. The DBP mRNA was the most abundant E2 mRNA during the early phase when the IVa2 mRNAs were present in very low levels. However, in the late phase, DBP mRNA levels dropped and IVa2 became the most abundant mRNA in the late phase (Figure 27). IVa2 is known to be expressed later in infection after viral DNA synthesis begins, and it is not produced when viral DNA synthesis is blocked (Binger and Flint, 1984). As discussed before, this change in expression patterns could be important, for the early-to-late switch mechanism. DBP has been reported to affect early mRNAs stabilities and proposed to influence the rapid turnover of early viral mRNAs (Babich and Nevins, 1981). On the other hand, the IVa2 is required to fully activate ML transcription unit expression during the late phase (described in section 1.5.3). Therefore, the timing of the expression of these two proteins could participate in the transition into the late phase. A concern in our analysis was the E2A DBP mRNAs expression levels, since they are very low in comparison to other early mRNAs and the kinetics show a decrease after 12 hpi (Figure 27). During replication, DBP is required in multiple copies per genome. As viral DNA is exponentially replicating, the amount of DBP

required should also increase. Therefore, we would expect the E2A DBP mRNA to be a lot more abundant in comparison to E2B pTP as an example, for which the number of copies needed per genome is only two. Further the steady states levels of DBP protein can be detected abundantly during the late phase (Figure 17). Low measurements of E2A DBP in our RNA-Seq analysis is surprising, since we can detect several reads mapping in the E2 region comparable to the other early regions (Figure 24). Thereby, we considered a possible problem could be in the assignment of reads towards the E2A DBP transcript during the RNA-Seq analysis. Therefore, other E2A annotations from HAdV-2, containing extra introns that could be missing in the HAdV-5 annotations were considered during the RNA-Seq analysis. Nevertheless, reads were only assigned to the HAdV-5 transcript with RPKMs being below three orders of magnitude. However, we found that synthesis rates comparing the relative maximal rates of expression between early transcription units showed that the E2A region had the lowest transcription rates (Nevins *et al.*, 1979). Further, Northern Blots showing expression levels of E2A mRNA in A549 cells at 6, 12, and 24 hpi, show very low amounts of E2A only at 12 hpi (Fessler and Young, 1998). These, is similar to the E2A DBP mRNA steady states kinetics in our analysis in which the peak of expression is at 12 hpi (Figure 24). The same report shows the presence of DBP proteins (steady state levels) from 12 hpi to 24 hpi. However, the synthesis rate of DBP protein has been reported to decline during the late phase of productive adenovirus infection (Beltz and Flint, 1979). Our observations of DBP decreasing during the late phase would support previous findings, even when the steady states of the protein can still be detected during the late phase. In previously reported mRNA single-end and RNA paired-end sequencing, which were set up to detect different splice sites in HAdV-2 infected human primary lung fibroblasts, new alternative splicing forms of E2A and E2B were detected (Zhao *et al.*, 2014). Thus, we included them in our transcriptome annotations (Figure 23, labeled as new). In our experiments, the E2A new transcripts were not detected, however, very high amounts of E2B new were measured. In our analysis, several reads aligned in this novel splice site, however, few reads aligned along the whole E2B new transcript (Figure 24). The authors also reported high levels of this mRNA splice site, in comparison to other E2B mRNAs. This transcript has a very small predicted open reading frame (from position 20723 to 19552 of HAdV-2 genome), however, no functions have been associated to it. The authors of this report discuss that it is unlikely that all of the novel splice sites detected in their experiments generate functional mRNAs.

In contrast to the other early mRNAs, the E4 mRNAs in our analysis did not show such a drastic change in proportions in the transition from early to late phase. Partly because the E4 region has a wide set of mRNAs, due complicated splicing patterns. In general, the E4 orf2 early mRNAs was the most abundant at early and late time-points and E4 orf6/7_2 late mRNAs highly increased in the late phase (Figure 28). Around 24 different E4 mRNAs have been

previously reported in different experimental settings (Herissé and Galibert, 1981; Freyer and Roberts, 1984; Rigolet and Galibert, 1984; Tigges and Raskas, 1984; Virtanen *et al.*, 1984). All E4 mRNAs share the 5' and 3' ends but have different internal splicing patterns. Studies have shown that there is some degree of temporal regulation on the splicing pattern. Fourteen different splice variants of E4 were measured in HeLa cells by quantitative RNase protection assays showing their cytoplasmic levels at different infection time-points with HAdV-5 (Dix and Leppard, 1993). A relative abundance comparison between these mRNAs was performed and most of them were classified as early or late, depending on their temporal expression. The mRNAs classified as early were first detected in the cytoplasm at early times post-infection and then remained constant or declined slightly during the late phase of infection. The mRNAs classified as late were barely detectable in the early phase and their levels increased dramatically thereafter. In our analysis, we included 12 of these E4 mRNAs splice variants to our transcriptome adding the suffix -early and -late depending on the reported temporal expression (Dix and Leppard, 1993). We also included a new detected splice form of E4 orf6 and a predicted form of E4 orf5 (Zhao *et al.*, 2014). Comparison of our sequencing analysis with previously reported quantitative RNase protection did not match all the temporal observations. Our measuring approach has higher sensitivity; therefore, we could detect all mRNAs since early time points, even in low quantities (Table 2). When analyzing the kinetics of the total expression values, most E4 mRNAs showed a gradual increase during the time-course, more in accordance to what was previously described for mRNAs from the early class. However, some of the mRNAs classified as late did show a drastic increase in the late phase, such as E4 orf1_1 late, E4 orf1_2 late, E4 orf6/7_2 late, and E4 orf6 new (Figure 27). Differences could be due to the use of different cell-lines and time-points, as well as the different techniques used to measure mRNAs. The biggest down-fall in our technique is the assignment of reads between mRNAs sharing most of the same sequence as only the intron-intron regions give specific read counts and the unspecific reads are distributed in proportion to the amounts of specific reads of each mRNA. Therefore, reads that align at the 5', 3' or common middle region are sorted between the E4 mRNAs and could affect the kinetics, especially at early time-points.

Synthesis rates of early mRNAs have been reported to stop during the late phase (Fessler and Young, 1998). However, the kinetics of the steady state early mRNAs showed that most of the early genes do not decrease as infection progresses into the late phase. Therefore, they are still present at late times, but viral late mRNAs are several orders of magnitude more abundant and early mRNAs are only 6% of viral mRNAs at 48 hpi (Figure 26).

5.4.1.2 Viral late mRNAs

The analysis of viral late mRNAs during the early phase showed that they were present in very low quantities, 5% and 9% at 6 hpi and 12 hpi, respectively. Previous reports have shown that in the early phase transcription in MLTU proceeds across the L1, L2, and L3 poly (A) sites, however, only mRNA from the L1 region can be detected to accumulate in the cytoplasm (Thomas and Mathews, 1980; Akusjarvi and Persson, 1981a; Nevins and Wilson, 1981b; Iwamoto *et al.*, 1986). In an intermediate state, immediately after DNA synthesis but before full activation of MLTU transcription, L1 and L4 products (L1 52/55K and L4 33K mRNAs) are detected in the cytoplasm (Larsson *et al.*, 1992; Farley, Brown and Leppard, 2004; Wu, Guimet and Hearing, 2013b). In this analysis, we do not detect L1 as the most abundant viral late mRNA in the early phase and only few reads align around the L1 region at 12 hpi, but not more than in other late regions (Figure 24). However, we learned that the L1 52/55K detected during the early phase contains an extra intron (i-leader), between the leaders 2 and 3 of the tripartite leader, which is excluded in the late phase (Farley *et al.*, 2004). The presence of the i-leader in the L1 52/55K mRNA does not translate into the L1 52/55K protein, but into a ~14K product coded inside the i-leader (Akusjarvi and Persson, 1981a). In our analysis, the i-leader was not added to the L1 52/55K mRNA annotations, therefore, it is possible that we do not detect this early form of L1 52/55K for that reason. However, in the alignment of reads to the transcriptome, a small peak is present in the i-leader region at 12 hpi (Figure 24). In our experiments the most expressed late families during the early phase were the L3 and L4 (Figure 29).

In the late phase, expression of the viral late mRNAs increased exponentially, resulting in them being the most abundant mRNAs during the late phase, 85% and 94% of viral mRNAs at 24 hpi and 48 hpi, respectively. The steady states kinetics between the late families were similar, showing a general activation of the ML promoter after the transition into the late phase as previously reported (Farley *et al.*, 2004). No other quantitative study with a complete viral late mRNA set was found to compare with our data. However, there are previous reports showing the L1-L5 families by Northern Blots analysis at late time points of infection to which we could compare the relative abundance between mRNAs of each family (Akusjarvi and Persson, 1981a; Larsson *et al.*, 1992; Wu *et al.*, 2013a). Although not all the annotated mRNAs in this analysis could be compared, we did see that the proportions of the main viral late mRNAs are very similar to previous reports. An exception was the L2 pVII mRNA, which in our analysis is less abundant than the rest of the L2 mRNAs. Also, L3 hexon and L5 pVI fiber viral late mRNAs were the most abundant (~10⁷ RPKMs), and most reads in the late phase also aligned in those transcripts. This has also been shown in Northern Blots of cytoplasmic poly(A) viral mRNAs of HAAdV-5 infected HeLa cells at late time points of infection (Halbert, Cutt and Shenk, 1985). An

interesting observation is that the amount of reads aligned in the tripartite leader regions are not all equal, as different rates of usage have been reported (Zhao *et al.*, 2014). It has been previously shown that L5 pVI fiber mRNA 5'-end has several alternative spliced products, incorporating not only the tripartite leader but a combinations of leaders (i-, x-, y-, and z-leader) (Akusjärvi and Pettersson, 1979; Zhao *et al.*, 2014; Hidalgo *et al.*, 2016). For this study, we considered the i- and y-leaders. The y-leader has been reported to be the most incorporated during splicing processes (Zhao *et al.*, 2014). Also, the deletion of the E3 region in the wt HAdV-5 used in for these experiments does not contain the z-leader. In this analysis, the L5 fiber mRNA with the i- and y-leaders was detected during the early phase, however in the late phase the fully spliced L5 pVI fiber mRNA (without the leaders), was the most abundant (Figure 30). This is in accordance to measurements of the spliced and unspliced L5 fiber mRNA by RT PCR, from RCs fractions of HAdV-5 infected HFF cells at late time point of infection (Hidalgo *et al.*, 2016). In general, the use of this technique allowed us to get a full understanding of how the steady states of all viral mRNAs behave during the time course of infection in A549 cells in a single experimental condition.

5.4.2 Effects of HAdV-5 infection on cytoplasmic mRNA accumulation

In this study, we also approached the question of how the cytoplasmic and nucleoplasmic mRNA levels are affected during infection with HAdV-5. We decided to use this technique which allowed us to asses this question in a global fashion, calculating cytoplasmic to nuclear mRNA ratios for all viral mRNAs during the time course infection. Previous reports have focused on a very small group of viral and cellular mRNAs; therefore, conclusions could be misleading. In this experiment, the measurement of cyto/nuc ratios was performed with steady state mRNA concentrations not with newly synthesized mRNAs. Therefore, what we measured is not the mRNA export rates but a measurement of cytoplasmic mRNA accumulation.

5.4.2.1 Viral mRNAs

The sum of all the data shows that viral mRNAs are present in very low quantities during the early phase. Only 0.07% and 0.3% of the global reads are viral at 6 hpi and 12 hpi, respectively. However, as the late phase commences and exponential accumulation of viral mRNAs is induced, reads associated to viral mRNAs occupy 43% and 79% of the global total reads at 24 hpi and 48 hpi, respectively (Figure 36, panel C). In this study, we found that the cytoplasmic accumulation of viral mRNAs also varies during infection. At 6 hpi, very few viral mRNAs are detected but 99% of them are measured in the cytoplasm. In contrast, at 12 hpi, as early viral transcription is activated, there is an eight-fold increase on viral mRNA reads but only 22% of

them were detected in the cytoplasm. In the late phase, at 24 hpi, there is an exponential increase (190 times) of total viral reads. The cytoplasmic viral reads measured were 57% of the total viral reads, representing almost 25% of the global reads and 48% of the total cytoplasmic reads. Later, at 48 hpi, the viral reads accumulated in the cytoplasm increased to 72%, representing 62% of the global reads and 87% of the total cytoplasmic reads (Figure 36, panel A and D). Together, these data show that in the late phase, viral mRNAs become more abundant than cellular mRNAs globally and in the cytoplasm. Therefore, in this study we do see an increase in the cytoplasmic accumulation of viral mRNAs in the late phase, accordingly to previously published data.

The analysis of the changes in cyto/nuc ratios of individual viral mRNAs along infection reflected some of the previous observations. We could observe a transition in the cyto/nuc ratio patterns between the early and late time point. As late in infection, the pattern looks completely opposite to the early phase, which is probably in part a reflection of the increase in transcription during the late phase. However, we could also observe that viral mRNAs transition faster from the nucleus to the cytoplasm at 6 hpi, when they are expressed in very low quantities, than at 12 and 24 hpi when they accumulate faster. As the cyto/nuc ratios of the most abundant viral mRNAs became more negative at late time points (Figure 32). This does not indicate that viral mRNAs are selectively exported in the late phase. However, it does suggest that they are being synthesized faster than they can be exported by cellular export pathways. Altogether, our data shows that even though the cyto/nuc rates diminish at late time points, they efficiently accumulate in the cytoplasm.

5.4.2.2 Cellular mRNAs

For the analysis of the cellular mRNA cyto/nuc ratios we selected only the ones that had a significant fold change of at least two, which was 46% of the total sample. This showed that more than half of our sample did not have a significant change in the cyto/nuc ratios during infection. This contradicts a previous report stating that cellular mRNA export is blocked during the late phase of adenovirus infection (Babich *et al.*, 1983). However, this report only analyzed four cellular cDNA clones to arrive to that conclusion. Further, when analyzing the cellular mRNA cyto/nuc ratios we found that not all cellular mRNAs are negatively regulated during the late phase, as a set of cellular mRNAs was found to be positively regulated at that time point (Figure 33). Previously, cellular mRNAs such as Hsp70, β -actin, Tubulin, 6-16, and Mx-A, were reported to escape cellular export blockage and to efficiently accumulate in the cytoplasm, suggesting that not all cellular mRNAs are subjected to nuclear export blockage (Yang *et al.*, 1996). As in the viral analysis, the cellular mRNAs cyto/nuc ratios pattern was

completely changed from the early to the late phase and 12 hpi seems to be a transition phase. This is expected as the transition into the late phase has an overall impact in both the mRNA transcription and mRNA cytoplasmic accumulation, as previously described (section 1.5). Additionally, the functional analysis showed that mRNAs which were regulated on a cyto/nuc ratio level, are associated to cellular functions involved in RNA splicing and protein biogenesis (section 4.2.3.1). This is interesting, but not surprising, as both cellular functions are crucial for the high levels of viral mRNA and protein synthesis required during the late phase. A previous transcriptomic report has also shown the enrichment of mRNAs involved in these same functions, during the late phase of the single-cell replication cycle of HAdV-5 infected cells (Miller *et al.*, 2009). Also, an increase in cellular proteins required for mRNA processing during the late phase has been reported earlier (Bridge *et al.*, 1993).

Cellular transcription has been shown not to be blocked during infection with HAdV, as cellular mRNA transcription rates are still detected during the late phase when assessing incorporation of H³ labeled uridine in hnRNA sequences (Beltz and Flint, 1979). In this study, we show that there is a global decrease of 1.8 times on the steady states levels of cellular mRNAs from 12 to 48 hpi (Figure 36, panel B). Also, when observing the top 35 most expressed cellular mRNAs, we also found that their overall abundance was diminished ~8.7 times from 12 to 48 hpi, showing a greater effect in the decrease of individual mRNAs' steady states levels (Figure 37). Therefore, our data show that even when cellular transcription is still ongoing during infection, the global accumulation of cellular mRNAs is reduced in the late phase of infection. Also, when comparing the top most expressed cellular mRNAs with viral mRNAs we found that cellular mRNAs were below the most expressed viral mRNA levels in the late phase (Figure 38).

To further comprehend these data, the cytoplasmic accumulation of cellular mRNAs was also analyzed. Cellular cytoplasmic reads in non-infected cells (and at early time-points of infection) represented more than half of the total cellular reads (Figure 36, panel B). However, in the late phase, we could observe a decrease of 15%, making the nuclear reads more abundant than the cytoplasmic reads. Therefore, we observed both a decrease in the total and the cytoplasmic steady states accumulation of cellular reads. However, the decrease in the percentage of cytoplasmic reads is not as substantial as the decrease of the total cytoplasmic reads. This alone would suggest that there is little effect on the cellular mRNA cytoplasmic accumulation due to HAdV-5 infection. However, when compared with viral mRNA reads it is obvious that most of the cytoplasmic reads at late time points are associated to viral mRNAs as they constitute 85.6% of the total cytoplasmic reads (Figure 36, panel D). This observation agrees with previous studies showing that fewer cellular poly(A)-mRNAs are detected in the cytoplasm at late time points when using hybridization assays (Babich *et al.*,

1983). However, our results suggest that the decrease in cytoplasmic cellular mRNAs is not due to a block in mRNA export, but to a mass effect. In the late phase, the total number of cellular mRNAs decreases while viral mRNAs increase, resulting in viral mRNAs being five times more abundant globally (Figure 36, panel C). As described before, the same export pathway is reported to be used by viral and cellular mRNAs. Also, no degradation of cellular components involved in this pathway, that could alter cellular mRNA export, has been reported so far (Yatherajam *et al.*, 2011). Therefore, the decrease of cytoplasmic accumulation of cellular mRNAs can be explained by effects other than cytoplasmic export blockage. Nevertheless, when observing the cytoplasmic accumulation of viral mRNAs during the late phase, there is a notorious increase in the percentage of viral cytoplasmic reads during the late phase, from 22% at 12 hpi to 72% at 48 hpi (Figure 36, panel A). Since this study was performed using steady-state mRNAs, we cannot discriminate whether this is due to an increase in mRNA export rates or due to higher mRNA stabilities, or both. However, the stability of newly synthesized viral and cellular mRNAs has been previously analyzed in HeLa cells infected with HAdV-2 at late time-points and no differences were found (Beltz and Flint, 1979). Our observations support a number of previous reports showing that viral mRNAs are preferably selected for cytoplasmic export (Flint and Gonzalez, 2003). Several studies have shown that the E1B and Orf6 proteins are required for this viral mRNA selective export (described in section 1.5.3.1). In this study, we do not assess the role of the E1B and Orf6 proteins; however, efforts to elucidate their contribution have been made and are still in progress.

Altogether, our data supports the hypothesis that nuclear microenvironments induced during infection, which are specialized in the transcription, processing and therefore priming for nuclear export of viral mRNAs, favor the accumulation of viral mRNAs in the cytoplasm at late time points, even when cellular mRNAs transcription is not inhibited. The first part of this work also contributes to this hypothesis, showing the relocalization of a cellular protein, involved in mRNA processing and which also associates to the NPC, to these viral induced structures. In this study, we do not see a complete block in cytoplasmic accumulation for cellular mRNAs, nor could we find degradation of the analyzed key cellular components involved in mRNA processing and export pathways. Other transcriptome analyses have been performed in the context of HAdV infection before this one, however, this is the first time this method has been used to assess questions regarding mRNA biogenesis and export during HAdV infection. Finally, these results show that mRNA regulation during infection with HAdV is more complex than previously thought.

6 Literature

- Akusjarvi, G. and Persson, H.** (1981a) 'Controls of RNA splicing and termination in the major late adenovirus transcription unit', *Nature*, 292(5822), pp. 420–426.
- Akusjärvi, G. and Pettersson, U.** (1979) 'Sequence analysis of adenovirus DNA: Complete nucleotide sequence of the spliced 5' noncoding region of adenovirus 2 hexon messenger RNA', *Cell*, 16(4), pp. 841–850.
- Albert, M. J.** (1986) 'Enteric adenoviruses', *Archives of Virology*, pp. 1–17.
- Allard, A., Albinsson, B. and Wadell, G.** (2001) 'Rapid typing of human adenoviruses by a general PCR combined with restriction endonuclease analysis', *Journal of Clinical Microbiology*, 39(2), pp. 498–505.
- Anderson, C. W., Young, M. E. and Flint, S. J.** (1989) 'Characterization of the adenovirus 2 virion protein, Mu', *Virology*, 172(2), pp. 506–512.
- Andersson, M. et al.** (1985) 'Impaired intracellular transport of class I MHC antigens as a possible means for adenoviruses to evade immune surveillance', *Cell*, 43(1), pp. 215–222.
- Aoki, K. et al.** (2011) 'Clinical features of adenoviral conjunctivitis at the early stage of infection', *Japanese Journal of Ophthalmology*, 55(1), pp. 11–15.
- Araujo, F. D. et al.** (2005) 'Adenovirus type 5 E4orf3 protein targets the Mre11 complex to cytoplasmic aggresomes.', *Journal of virology*, 79(17), pp. 11382–91.
- Aspegren, A., Rabino, C. and Bridge, E.** (1998) 'Organization of splicing factors in adenovirus-infected cells reflects changes in gene expression during the early to late phase transition', *Exp Cell Res*, 245(1), pp. 203–213.
- Babich, A. et al.** (1983) 'Effect of adenovirus on metabolism of specific host mRNAs : transport control and specific translational Effect of Adenovirus on Metabolism of Specific Host mRNAs : Transport Control and Specific Translational Discrimination'.
- Babich, A. and Nevins, J. R.** (1981) 'The stability of early adenovirus mRNA is controlled by the viral 72 kd DNA-binding protein', *Cell*, 26(3 PART 1), pp. 371–379.
- Babiss, L. E., Ginsberg, H. S. and Darnell, J. E.** (1985) 'Adenovirus E1B proteins are required for accumulation of late viral mRNA and for effects on cellular mRNA translation and transport.', *Molecular and cellular biology*, 5(10), pp. 2552–8.
- Bachi, A. et al.** (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), pp. 136–158.
- Backström, E. et al.** (2010) 'Adenovirus L4-22K stimulates major late transcription by a mechanism requiring the intragenic late-specific transcription factor-binding site', *Virus Research*, 151(2), pp. 220–228.
- Bagchi, S., Raychaudhuri, P. and Nevins, J. R.** (1990) 'Adenovirus E1A proteins can dissociate heteromeric complexes involving the E2F transcription factor: A novel mechanism for E1A trans-activation', *Cell*, 62(4), pp. 659–669.
- Bailey, C. J., Crystal, R. G. and Leopold, P. L.** (2003) 'Association of adenovirus with the microtubule organizing center.', *Journal of virology*, 77(24), pp. 13275–13287.

- Baker, A. et al.** (2007) 'Adenovirus E4 34k and E1b 55k Oncoproteins Target Host DNA Ligase IV for Proteasomal Degradation', *Society*, 81(13), pp. 7034–7040.
- Barral, P. M. et al.** (2005) 'The interaction of the hnRNP family member E1B-AP5 with p53.', *FEBS letters*, 579(13), pp. 2752–8.
- Beltz, G. a and Flint, S. J.** (1979) 'Inhibition of HeLa cell protein synthesis during adenovirus infection. Restriction of cellular messenger RNA sequences to the nucleus.', *Journal of molecular biology*, 131(2), pp. 353–73.
- Ben-Israel, H. and Kleinberger, T.** (2002) 'Adenovirus and cell cycle control.', *Frontiers in bioscience : a journal and virtual library*, 7(February), pp. d1369–d1395.
- Bergelson, J. M. et al.** (1997) 'Isolation of a common receptor for coxsackie B viruses and adenoviruses 2 and 5', *Science*, 275(5304), pp. 1320–1323.
- Berget, S. M., Moore, C. and Sharp, P. A.** (1977) 'Spliced segments at the 5' terminus of adenovirus 2 late mRNA', *Proceedings of the National Academy of Sciences*, 74(8), pp. 3171–3175.
- Bergink, S. and Jentsch, S.** (2009) 'Principles of ubiquitin and SUMO modifications in DNA repair.', *Nature*, 458(7237), pp. 461–7.
- Berk, A. J. and Sharp, P. A.** (1978) 'Structure of the adenovirus 2 early mRNAs', *Cell*, 14(3), pp. 695–711.
- Berscheminski, J. et al.** (2014) 'Sp100 isoform-specific regulation of human adenovirus 5 gene expression', *J Virol*, 88(11), pp. 6076–6092.
- Berscheminski, J. et al.** (2016) 'Sp100A is a tumor suppressor that activates p53-dependent transcription and counteracts E1A/E1B-55K-mediated transformation', *Oncogene*, 35(24), pp. 3178–3189.
- Bewley, M. C.** (1999) 'Structural Analysis of the Mechanism of Adenovirus Binding to Its Human Cellular Receptor, CAR', *Science*, 286(5444), pp. 1579–1583.
- Biasiotto, R. et al.** (2015) 'Regulation of human adenovirus alternative RNA splicing by the adenoviral L4-33K and L4-22K proteins', *International Journal of Molecular Sciences*, 16(2), pp. 2893–2912.
- Binger, M. H. and Flint, S. J.** (1984) 'Accumulation of early and intermediate mRNA species during subgroup C adenovirus productive infections', *Virology*, 136(2), pp. 387–403.
- Blackford, A. N., Bruton, R. K., Dirlik, O., Stewart, G. S., et al.** (2008) 'A role for E1B-AP5 in ATR signaling pathways during adenovirus infection.', *Journal of virology*, 82(15), pp. 7640–52.
- Blackford, A. N., Bruton, R. K., Dirlik, O., Stewart, S., et al.** (2008) 'A Role for E1B-AP5 in ATR Signaling Pathways during Adenovirus Infection A Role for E1B-AP5 in ATR Signaling Pathways during Adenovirus Infection □'.
- Blanchette, P. et al.** (2008) 'Control of mRNA Export by Adenovirus E4orf6 and E1B55K Proteins During Productive Infection Requires E4orf6 Ubiquitin Ligase Activity.', *Journal of virology*, 82(6), pp. 2642–51.
- Blanchette, P. et al.** (2008) 'Control of mRNA Export by Adenovirus E4orf6 and E1B55K Proteins during Productive Infection Requires E4orf6 Ubiquitin Ligase Activity Control of mRNA Export by Adenovirus E4orf6 and E1B55K Proteins during Productive Infection Requires E4orf6 Ubiquitin Lig'.
- Blanchette, P. et al.** (2013) 'Aggresome formation by the adenoviral protein E1B55K is not

conserved among adenovirus species and is not required for efficient degradation of nuclear substrates.', *Journal of virology*, 87(9).

Bosher, J., Robinson, E. C. and Hay, R. T. (1990) 'Interactions between the adenovirus type 2 DNA polymerase and the DNA binding domain of nuclear factor I', *New Biol*, 2(12), pp. 1083–1090.

Boyer, T. G. et al. (1999) 'Mammalian Srb/mediator complex is targeted by adenovirus E1A protein', *Nature*, 399(6733), pp. 276–279.

Bradford, L. (1976) 'Problems of Ethics and Behavior in the Forensic Sciences', *Jorunal of Forensic Sciences*, 21(4), pp. 763–768.

Brestovitsky, A. et al. (2011) 'The adenovirus E4orf4 protein targets PP2A to the ACF chromatin-remodeling factor and induces cell death through regulation of SNF2h-containing complexes', *Nucleic Acids Research*, 39(15), pp. 6414–6427.

van Breukelen, B. et al. (2003) 'Adenovirus type 5 DNA binding protein stimulates binding of DNA polymerase to the replication origin.', *Journal of virology*, 77(2), pp. 915–22.

Bridge, E. et al. (1993) 'Nuclear organization of splicing small nuclear ribonucleoproteins in adenovirus-infected cells', *J Virol*, 67(10), pp. 5792–5802.

Bridge, E. and Ketner, G. (1990) 'Interaction of adenoviral E4 and E1b products in late gene expression', *Virology*, 174(2), pp. 345–353.

Bridge, E. and Pettersson, U. (1996) 'Nuclear organization of adenovirus RNA biogenesis', *Experimental Cell Research*, 229(2), pp. 233–239.

Burgert, H. G. and Kvist, S. (1985) 'An adenovirus type 2 glycoprotein blocks cell surface expression of human histocompatibility class I antigens.', *Cell*, 41(3), pp. 987–97.

Cantin, G. T., Stevens, J. L. and Berk, A. J. (2003) 'Activation domain-mediator interactions promote transcription preinitiation complex assembly on promoter DNA', *Proceedings of the National Academy of Sciences*, 100(21), pp. 12003–12008.

Carson, C. T. et al. (2009) 'Mislocalization of the MRN complex prevents ATR signaling during adenovirus infection', *EMBO Journal*, 28(6), pp. 652–662.

Carvalho, T. et al. (1995) 'Targeting of adenovirus E1A and E4-ORF3 proteins to nuclear matrix-associated PML bodies', *Journal of Cell Biology*, 131(1), pp. 45–56.

Castiglia, C. L. and Flint, S. J. (1983) 'Effects of adenovirus infection on rRNA synthesis and maturation in HeLa cells.', *Molecular and cellular biology*, 3(4), pp. 662–71.

Castillo-Villanueva, E. et al. (2014) 'The Mre11 Cellular Protein Is Modified by Conjugation of Both SUMO-1 and SUMO-2/3 during Adenovirus Infection', *ISRN Virology*, 2014(52777), pp. 1–14.

Challberg, M. D., Desiderio, S. V and Kelly, T. J. (1980) 'Adenovirus DNA replication in vitro: characterization of a protein covalently linked to nascent DNA strands.', *Proceedings of the National Academy of Sciences of the United States of America*, 77(9), pp. 5105–9.

Chatterjee, P. K., Vayda, M. E. and Flint, S. J. (1986) 'Identification of proteins and protein domains that contact DNA within adenovirus nucleoprotein cores by ultraviolet light crosslinking of oligonucleotides 32P-labelled in vivo', *Journal of Molecular Biology*, 188(1), pp. 23–37.

Chen, J., Morral, N. and Engel, D. A. (2007) 'Transcription releases protein VII from adenovirus chromatin', *Virology*, 369(2), pp. 411–422.

- Chen, P. H., Ornelles, D. A. and Shenk, T.** (1993) 'The adenovirus L3 23-kilodalton proteinase cleaves the amino-terminal head domain from cytokeratin 18 and disrupts the cytokeratin network of HeLa cells.', *Journal of virology*, 67(6), pp. 3507–14.
- Ching, W., Dobner, T. and Koyuncu, E.** (2012) 'The human adenovirus type 5 E1B 55-kilodalton protein is phosphorylated by protein kinase CK2.', *Journal of virology*, 86(5), pp. 2400–15.
- Chow, L. T. et al.** (1977) 'An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA', *Cell*, 12(1), pp. 1–8.
- Chroboczek, J., Bieber, F. and Jacrot, B.** (1992) 'The sequence of the genome of adenovirus type 5 and its comparison with the genome of adenovirus type 2.', *Virology*, 186(1), pp. 280–5.
- Computing, R. F. for S.** (2011) 'R Development Core Team', *R: A Language and Environment for Statistical Computing*, 55, pp. 275–286.
- Corbin-Lickfett, K. A. and Bridge, E.** (2003) 'Adenovirus E4-34kDa requires active proteasomes to promote late gene expression', *Virology*, 315(1), pp. 234–244.
- Cuconati, A. and White, E.** (2002) 'Viral homologs of BCL-2: Role of apoptosis in the regulation of virus infection', *Genes and Development*, pp. 2465–2478.
- Cuesta, R., Xi, Q. and Schneider, R. J.** (2000) 'Adenovirus-specific translation by displacement of kinase Mnk1 from cap-initiation complex eIF4F.', *The EMBO journal*, 19(13), pp. 3465–74.
- Cuesta, R., Xi, Q. and Schneider, R. J.** (2001) 'Preferential translation of adenovirus mRNAs in infected cells', in *Cold Spring Harbor Symposia on Quantitative Biology*, pp. 259–267.
- Cuesta, R., Xi, Q. and Schneider, R. J.** (2004) 'Structural basis for competitive inhibition of eIF4G-Mnk1 interaction by the adenovirus 100-kilodalton protein.', *Journal of virology*, 78(14), pp. 7707–16.
- Culjkovic, B. et al.** (2005) 'eIF4E promotes nuclear export of cyclin D1 mRNAs via an element in the 3'UTR', *Journal of Cell Biology*, 169(2), pp. 245–256.
- Culjkovic, B. et al.** (2006) 'eIF4E is a central node of an RNA regulon that governs cellular proliferation', *Journal of Cell Biology*, 175(3), pp. 415–426.
- Cullen, B. R.** (2003) 'Nuclear mRNA export: Insights from virology', *Trends in Biochemical Sciences*, 28(8), pp. 419–424.
- Dales, S. and Chardonnet, Y.** (1973) 'Early events in the interaction of adenoviruses with HeLa cells. IV. Association with microtubules and the nuclear pore complex during vectorial movement of the inoculum', *Virology*, 56(2), pp. 465–483.
- Dallaire, F., Blanchette, P. and Branton, P. E.** (2009) 'A proteomic approach to identify candidate substrates of human adenovirus E4orf6-E1B55K and other viral cullin-based E3 ubiquitin ligases.', *Journal of virology*, 83(23), pp. 12172–84.
- Davison, A. J., Benko, M. and Harrach, B.** (2003) 'Genetic content and evolution of adenoviruses', *Journal of General Virology*, 84(11), pp. 2895–2908.
- Debbas, M. and White, E.** (1993) 'Wild-type p53 mediates apoptosis by E1A, which is inhibited by E1B', *Genes and Development*, 7(4), pp. 546–554.
- Dillies, M. A. et al.** (2013) 'A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis', *Briefings in Bioinformatics*, 14(6), pp. 671–683.
- Dix, I. and Leppard, K. N.** (1993) 'Regulated splicing of adenovirus type 5 E4 transcripts and

regulated cytoplasmic accumulation of E4 mRNA.', *Journal of virology*, 67(6), pp. 3226–31.

Dobbelstein, M. et al. (1997) 'Nuclear export of the E1B 55-kDa and E4 34-kDa adenoviral oncoproteins mediated by a rev-like signal sequence', *EMBO Journal*, 16(14), pp. 4276–4284.

Dobner, T. et al. (1996) 'Blockage by adenovirus E4orf6 of transcriptional activation by the p53 tumor suppressor', *Science*, 272(5267), pp. 1470–1473.

Dosch, T. et al. (2001) 'The adenovirus type 5 E1B-55K oncoprotein actively shuttles in virus-infected cells, whereas transport of E4orf6 is mediated by a CRM1-independent mechanism', *Journal of virology*, 75(12), pp. 5677–5683.

Egan, C., Bayley, S. T. and Branton, P. E. (1989) 'Binding of the Rb1 protein to E1A products is required for adenovirus transformation.', *Oncogene*, 4(3), pp. 383–388.

Endter, C. et al. (2001) 'SUMO-1 modification required for transformation by adenovirus type 5 early region 1B 55-kDa oncoprotein.', *Proceedings of the National Academy of Sciences of the United States of America*, 98(20), pp. 11312–7.

Endter, C. and Dobner, T. (2004) 'Cell transformation by human adenoviruses.', *Current topics in microbiology and immunology*, 273, pp. 163–214.

Evans, J. D. and Hearing, P. (2005) 'Relocalization of the Mre11-Rad50-Nbs1 complex by the adenovirus E4 ORF3 protein is required for viral replication.', *Journal of virology*, 79(10), pp. 6207–15.

Farley, D. C., Brown, J. L. and Leppard, K. N. (2004) 'Activation of the Early-Late Switch in Adenovirus Type 5 Major Late Transcription Unit Expression by L4 Gene Products', *Journal of Virology*, 78(4), pp. 1782–1791.

Fay, N. and Panté, N. (2015) 'Old foes, new understandings: Nuclear entry of small non-enveloped DNA viruses', *Current Opinion in Virology*, pp. 59–65.

Ferrari, R. et al. (2008) 'Epigenetic reprogramming by adenovirus e1a.', *Science (New York, N.Y.)*, 321(5892), pp. 1086–8.

Fessler, S. P. and Young, C. S. (1998) 'Control of adenovirus early gene expression during the late phase of infection.', *Journal of virology*, 72(5), pp. 4049–56.

Flint, S. J. (1998) 'The Tripartite Leader Sequence of Subgroup C Adenovirus Major Late mRNAs Can Increase the Efficiency of mRNA Export', *Microbiology*, 72(1), pp. 225–235.

Flint, S. J. (1999) 'Organization of the Adenoviral Genome', in *Adenoviruses: Basic Biology to Gene Therapy*, pp. 17–30.

Flint, S. J. and Gonzalez, R. A. (2003) 'Regulation of mRNA production by the adenoviral E1B 55-kDa and E4 Orf6 proteins.', *Current topics in microbiology and immunology*, 272, pp. 287–330.

Freyer, G. A. and Roberts, R. J. (1984) 'Characterization of the major mRNAs from adenovirus 2 early region 4 by cDNA cloning and sequencing', *Nucleic Acids Research*, 12(8), pp. 3504–3519.

Gabler, S. et al. (1998) 'E1B 55-kilodalton-associated protein: a cellular protein with RNA-binding activity implicated in nucleocytoplasmic transport of adenovirus and cellular mRNAs.', *Journal of virology*, 72(10), pp. 7960–71.

Gattoni, R. et al. (1996) 'The human hnRNP-M proteins: Structure and relation with early heat shock-induced splicing arrest and chromosome mapping', *Nucleic Acids Research*, 24(13), pp. 2535–2542.

Gaynor, R. B., Hillman, D. and Berk, A. J. (1984) 'Adenovirus early region 1A protein activates

transcription of a nonviral gene introduced into mammalian cells by infection or transfection', *Proc Natl Acad Sci U S A*, 81(4), pp. 1193–1197.

Ghosh, M. K. and Harter, M. L. (2003) 'A viral mechanism for remodeling chromatin structure in G0 cells', *Molecular Cell*, 12(1), pp. 255–260.

Giard, D. J. et al. (1973) 'In vitro cultivation of human tumors: establishment of cell lines derived from a series of solid tumors', *J Natl Cancer Inst*, 51(5), pp. 1417–1423.

Glenn, G. M. and Ricciardi, R. P. (1988) 'Detailed kinetics of adenovirus type-5 steady-state transcripts during early infection', *Virus Research*, 9(1), pp. 73–91.

Gonzalez, R. A. and Flint, S. J. (2002) 'Effects of Mutations in the Adenoviral E1B 55-Kilodalton Protein Coding Sequence on Viral Late mRNA Metabolism', *Journal of virology*, 76(9), pp. 4507–4519.

Graham, F. L. et al. (1977) 'Characteristics of a human cell line transformed by DNA from human adenovirus type 5.', *The Journal of general virology*, 36(1), pp. 59–74.

Greber, U. F. et al. (1993) 'Stepwise dismantling of adenovirus 2 during entry into cells', *Cell*, 75(3), pp. 477–486.

Greber, U. F. et al. (1996) 'The role of the adenovirus protease on virus entry into cells.', *The EMBO journal*, 15(8), pp. 1766–1777.

Gupta, S. et al. (2004) 'DNA binding provides a molecular strap activating the adenovirus proteinase.', *Molecular & cellular proteomics : MCP*, 3(10), pp. 950–9.

Gustin, K. E. and Imperiale, M. J. (1998) 'Encapsidation of viral DNA requires the adenovirus L1 52/55-kilodalton protein', *J Virol*, 72(10), pp. 7860–7870.

Halbert, D. N., Cuff, J. R. and Shenk, T. (1985) 'Adenovirus early region 4 encodes functions required for efficient DNA replication, late gene expression, and host cell shutoff.', *Journal of virology*, 56(1), pp. 250–7.

Han, S. P., Tang, Y. H. and Smith, R. (2010) 'Functional diversity of the hnRNPs: past, present and perspectives.', *The Biochemical journal*, 430(3), pp. 379–92.

Harada, J. N. et al. (2002) 'Analysis of the Adenovirus E1B-55K-Anchored Proteome Reveals Its Link to Ubiquitination Machinery', *Society*, 76(18), pp. 9194–9206.

Harada, J. N. and Berk, A. J. (1999) 'p53-Independent and -dependent requirements for E1B-55K in adenovirus type 5 replication.', *Journal of virology*, 73(7), pp. 5333–44.

Hasson, T. B. et al. (1989) 'Adenovirus L1 52- and 55-kilodalton proteins are required for assembly of virions', *J Virol*, 63(9), pp. 3612–3621.

Hayes, B. W. et al. (1990) 'The adenovirus L4 100-kilodalton protein is necessary for efficient translation of viral late mRNA species.', *Journal of virology*, 64(6), pp. 2732–42.

Hearing, P. et al. (1987) 'Identification of a repeated sequence element required for efficient encapsidation of the adenovirus type 5 chromosome.', *Journal of virology*, 61(8), pp. 2555–8.

Hearing, P. and Shenk, T. (1983) 'The adenovirus type 5 E1A transcriptional control region contains a duplicated enhancer element', *Cell*, 33(3), pp. 695–703.

Hearing, P. and Shenk, T. (1985) 'Sequence-independent autoregulation of the adenovirus type 5 E1A transcription unit', *Mol Cell Biol*, 5(11), pp. 3214–3221.

- Hendriks, I. A. et al.** (2015) 'Uncovering Global SUMOylation Signaling Networks in a Site-Specific Manner', *Nat Struct Mol Biol.*, 21(10), pp. 927–936.
- Herissé, J. and Galibert, F.** (1981) 'Nucleotide sequence of the EcoRI E fragment of adenovirus 2 genome', *Nucleic Acids Research*, 9(5), pp. 1229–1240.
- Hidalgo, P. et al.** (2016) 'Morphological, biochemical and functional study of viral replication compartments isolated from adenovirus-infected cells.', *Journal of Virology*, (January), p. JVI.00033-16.
- Hilleman, M. R. and Werner, J. H.** (1954) 'Recovery of New Agent from Patients with Acute Respiratory Illness', *Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, N.Y.)*, 85(1), pp. 183–8.
- Hindley, C. E., Lawrence, F. J. and Matthews, D. A.** (2007) 'A role for transportin in the nuclear import of adenovirus core proteins and DNA', *Traffic*, 8(10), pp. 1313–1322.
- Hong, S. S. et al.** (2005) 'The 100K-chaperone protein from adenovirus serotype 2 (subgroup C) assists in trimerization and nuclear localization of hexons from subgroups C and B adenoviruses', *Journal of Molecular Biology*, 352(1), pp. 125–138.
- Hoppe, A. et al.** (2006) 'Interaction of the adenovirus type 5 E4 Orf3 protein with promyelocytic leukemia protein isoform II is required for ND10 disruption.', *Journal of virology*, 80(6), pp. 3042–3049.
- Horowitz GA, Scharff MD, M. J. J.** (1969) 'Synthesis and assembly of adenovirus 2. I. Polypeptide synthesis, assembly of capsomers, and morphogenesis of the virion.', *Virology*, (39), pp. 682–694.
- Horridge, J. J. and Leppard, K. N.** (1998) 'RNA-binding activity of the E1B 55-kilodalton protein from human adenovirus type 5', *J. Virol.*, 72(11), pp. 9374–9379.
- Hovhannisyan, R. H. and Carstens, R. P.** (2007) 'Heterogeneous ribonucleoprotein M is a splicing regulatory protein that can enhance or silence splicing of alternatively spliced exons', *Journal of Biological Chemistry*, 282(50), pp. 36265–36274.
- Howley, P. M., Maki, C. G. and Huibregtse, J. M.** (1996) 'In Vivo Ubiquitination and Proteasome-mediated Degradation of p531', 1436(29), pp. 2649–2654.
- Huang, D. W., Sherman, B. T. and Lempicki, R. A.** (2009a) 'Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists', *Nucleic Acids Research*, 37(1), pp. 1–13.
- Huang, D. W., Sherman, B. T. and Lempicki, R. A.** (2009b) 'Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources', *Nature Protocols*, 4(1), pp. 44–57.
- Huang, J. and Schneider, R. J.** (1991) 'Adenovirus inhibition of cellular protein synthesis involves inactivation of cap-binding protein', *Cell*, 65(2), pp. 271–280.
- Huang, Y. et al.** (2003) 'SR splicing factors serve as adapter proteins for TAP-dependent mRNA export', *Molecular Cell*, 11(3), pp. 837–843.
- Iwamoto, S. et al.** (1986) 'Transcription unit mapping in adenovirus: regions of termination', *J Virol*, 59(1), pp. 112–119.
- Josheph R. Nevins** (1979) 'Processing of late adenovirus nuclear RNA to mRNA. Kinetics of formation of intermediates and demonstration that all events are nuclear.', *J. Mol. Biol*, 130(4), pp. 493–506.
- Kanopka, A. et al.** (1998) 'Regulation of adenovirus alternative RNA splicing by dephosphorylation of SR proteins', *Nature*, 393(6681), pp. 185–187.

- Kanopka, A. et al.** (1998) 'Regulation of adenovirus alternative RNA splicing by dephosphorylation of SR proteins.', *Nature*, 393(6681), pp. 185–187.
- Kao, C. C., Yew, P. R. and Berk, A. J.** (1990) 'Domains required for in vitro association between the cellular p53 and the adenovirus 2 E1B 55k proteins', *Virology*, 179(2), pp. 806–814.
- Karen, K. a and Hearing, P.** (2011) 'Adenovirus Core Protein VII Protects the Viral Genome from a DNA Damage Response at Early Times after Infection', *Journal of virology*, 85(9), pp. 4135–4142.
- Katahira, J.** (2015) 'Nuclear export of messenger RNA', *Genes*, 6(2), pp. 163–184.
- Katze, M. G. et al.** (1987) 'Adenovirus VAI RNA complexes with the 68 000 Mr protein kinase to regulate its autophosphorylation and activity.', *The EMBO journal*, 6(3), pp. 689–97.
- Kehl, S. C. and Kumar, S.** (2009) 'Utilization of Nucleic Acid Amplification Assays for the Detection of Respiratory Viruses', *Clinics in Laboratory Medicine*, pp. 661–671.
- Kindsmüller, K.** (2006) *Analysen zur Funktion des E1B-55K-Proteins von Adenovirus Typ 5 im lytischen Replikationszyklus*. Universität Regensburg.
- Kindsmüller, K. et al.** (2007) 'Intranuclear targeting and nuclear export of the adenovirus E1B-55K protein are regulated by SUMO1 conjugation.', *Proceedings of the National Academy of Sciences of the United States of America*, 104(16), pp. 6684–9.
- Kitajewski, J. et al.** (1986) 'Adenovirus VAI RNA antagonizes the antiviral action of interferon by preventing activation of the interferon-induced eIF-2 α kinase', *Cell*, 45(2), pp. 195–200.
- Kleinberger, T. and Shenk, T.** (1993) 'Adenovirus E4orf4 protein binds to protein phosphatase 2A, and the complex down regulates E1A-enhanced junB transcription.', *Journal of virology*, 67(12), pp. 7556–60.
- Krätzer, F. et al.** (2000) 'The adenovirus type 5 E1B-55K oncoprotein is a highly active shuttle protein and shuttling is independent of E4orf6, p53 and Mdm2.', *Oncogene*, 19(7), pp. 850–7.
- Kunz, A. N. and Ottolini, M.** (2010) 'The role of adenovirus in respiratory tract infections', *Current Infectious Disease Reports*, pp. 81–87.
- Lamond, A. I. and Sleeman, J. E.** (2003) 'Nuclear substructure and dynamics', *Current Biology*, 13(21), pp. R825–R828.
- Lamond, A. I. and Spector, D. L.** (2003) 'Nuclear Speckles: A Model for Nuclear Organelles', *Molecular Cell Biology*, 4(Agust), pp. 605–612.
- Larsson, S., Svensson, C. and Akusjärvi, G.** (1992) 'Control of adenovirus major late gene expression at multiple levels', *Journal of Molecular Biology*, 225(2), pp. 287–298.
- Lavoie, J. N. et al.** (1998) 'E4orf4, a novel adenovirus death factor that induces p53-independent apoptosis by a pathway that is not inhibited by zVAD-fmk', *Journal of Cell Biology*, 140(3), pp. 637–645.
- Lechner, R. L. and Kelly, T. J.** (1977) 'The structure of replicating adenovirus 2 DNA molecules', *Cell*, 12(4), pp. 1007–1020.
- Ledl, A., Schmidt, D. and Müller, S.** (2005) 'Viral oncoproteins E1A and E7 and cellular LxCxE proteins repress SUMO modification of the retinoblastoma tumor suppressor', *Oncogene*, 24(23), pp. 3810–3818.
- Leopold, P. et al.** (2000) 'Dynein- and microtubule-mediated translocation of adenovirus serotype 5

occurs after endosomal lysis.', *Human gene therapy*, 11, pp. 151–165.

Leppard, K. N. (1993) 'Selective effects on adenovirus late gene expression of deleting the E1b 55K protein.', *The Journal of general virology*, 74 (Pt 4)(4), pp. 575–82.

Leppard, K. N. et al. (2009) 'Adenovirus type 5 E4 Orf3 protein targets promyelocytic leukaemia (PML) protein nuclear domains for disruption via a sequence in PML isoform II that is predicted as a protein interaction site by bioinformatic analysis Printed in Great Britain', *Journal of General Virology*, pp. 95–104.

Leruez-Ville, M. et al. (2004) 'Real-time blood plasma polymerase chain reaction for management of disseminated adenovirus infection.', *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, 38(1), pp. 45–52.

Li, Y. et al. (1999) 'Identification of a cell protein (FIP-3) as a modulator of NF-kappaB activity and as a target of an adenovirus inhibitor of tumor necrosis factor alpha-induced apoptosis.', *Proceedings of the National Academy of Sciences of the United States of America*, 96(3), pp. 1042–1047.

Lichy, J. H., Horwitz, M. S. and Hurwitz, J. (1981) 'Formation of a covalent complex between the 80,000-dalton adenovirus terminal protein and 5'-dCMP in vitro', *Proceedings of the National Academy of Sciences*, 78(5), pp. 2678–2682.

Lindberg, U., Persson, T. and Philipson, L. (1972) 'Isolation and characterization of adenovirus messenger ribonucleic acid in productive infection.', *Journal of virology*, 10(5), pp. 909–919.

Lion, T. et al. (2003) 'Molecular Monitoring of Adenovirus in Peripheral Blood After', *Blood*, 102(c), pp. 1–29.

Lion, T. et al. (2010) 'Monitoring of adenovirus load in stool by real-time PCR permits early detection of impending invasive infection in patients after allogeneic stem cell transplantation', *Leukemia*, 24(4), pp. 706–714.

Liu, H. et al. (2010) 'Atomic structure of human adenovirus by cryo-EM reveals interactions among protein networks.', *Science (New York, N.Y.)*, 329(5995), pp. 1038–43.

Liu, Y. et al. (2005) 'Adenovirus Exploits the Cellular Aggresome Response To Accelerate Inactivation of the MRN Complex', *Journal of virology*, 79(22), pp. 14004–14016.

Lowe, S. W. and Earl Ruley, H. (1993) 'Stabilization of the p53 tumor suppressor is induced by adenovirus 5 E1A and accompanies apoptosis', *Genes and Development*, 7(4), pp. 535–545.

Mackey, J. K., Rigden, P. M. and Green, M. (1976) 'Do highly oncogenic group A human adenoviruses cause human cancer? Analysis of human tumors for adenovirus 12 transforming DNA sequences.', *Proceedings of the National Academy of Sciences of the United States of America*, 73(12), pp. 4657–61.

Mal, A., Piotrkowski, A. and Harter, M. L. (1996) 'Cyclin-dependent kinases phosphorylate the adenovirus E1A protein, enhancing its ability to bind pRb and disrupt pRb-E2F complexes.', *Journal of virology*, 70(5), pp. 2911–21.

Mangel, W. F. et al. (1993) 'Viral DNA and a viral peptide can act as cofactors of adenovirus virion proteinase activity', *Nature*, 361(6409), pp. 274–275.

Marcellus, R. C. et al. (2000) 'Induction of p53-independent apoptosis by the adenovirus E4orf4 protein requires binding to the Balpha subunit of protein phosphatase 2A.', *Journal of virology*, 74(17), pp. 7869–7877.

- Marko, M. et al.** (2010) 'hnRNP M interacts with PSF and p54nrb and co-localizes within defined nuclear structures', *Experimental Cell Research*, 316(3), pp. 390–400.
- Matthews, D. A.** (2001) 'Adenovirus Protein V Induces Redistribution of Nucleolin and B23 from Nucleolus to Cytoplasm', *Journal of virology*, i(2), pp. 1031–1038.
- Maul, G. G. et al.** (2000) 'Review: Properties and assembly mechanisms of ND10, PML2 bodies, or PODs', *Journal of Structural Biology*, pp. 278–287.
- Meier, O. and Greber, U. F.** (2003) 'Adenovirus endocytosis', *Journal of Gene Medicine*, pp. 451–462.
- Miller, D. L. et al.** (2009) 'The Adenoviral E1B 55-Kilodalton Protein Controls Expression of Immune Response Genes but Not p53-Dependent Transcription \square ', *Journal of virology*, 83(8), pp. 3591–3603.
- Mitsudomi T et al.** (1992) 'p53 gene mutations in non-small-cell lung cancer cell lines and their correlation with the presence of ras mutations and clinical features.', *Oncogene*, 7(171–180).
- Montell, C. et al.** (1984) 'Complete transformation by adenovirus 2 requires both E1A proteins', *Cell*, 36(4), pp. 951–961.
- Moore, M. et al.** (1987) 'Induced heat shock mRNAs escape the nucleocytoplasmic transport block in adenovirus-infected HeLa cells.', *Molecular and cellular biology*, 7(12), pp. 4505–12.
- Morris, S. J., Scott, G. E. and Leppard, K. N.** (2010) 'Adenovirus late-phase infection is controlled by a novel L4 promoter.', *Journal of virology*, 84(14), pp. 7096–104.
- Mul, Y. M., Verrijzer, C. P. and van der Vliet, P. C.** (1990) 'Transcription factors NF1 and NF110/oct-1 function independently, employing different mechanisms to enhance adenovirus DNA replication.', *Journal of virology*, 64(11), pp. 5510–8.
- Muller, S. and Dobner, T.** (2008) 'The adenovirus E1B-55K oncoprotein induces SUMO modification of p53', *Cell Cycle*, (March), pp. 754–758.
- Nagata, K., Guggenheimer, R. A. and Hurwitz, J.** (1983) 'Adenovirus DNA replication in vitro: synthesis of full-length DNA with purified proteins.', *Proceedings of the National Academy of Sciences of the United States of America*, 80(14), pp. 4266–70.
- Nakano, M. Y. et al.** (2000) 'The first step of adenovirus type 2 disassembly occurs at the cell surface, independently of endocytosis and escape to the cytosol.', *Journal of virology*, 74(15), pp. 7085–95.
- Nemerow, G. R. et al.** (2009) 'Insights into adenovirus host cell interactions from structural studies', *Virology*. Elsevier Inc., 384(2), pp. 380–388.
- Nemerow, G. R. and Stewart, P. L.** (1999) 'Role of alpha(v) integrins in adenovirus cell entry and gene delivery.', *Microbiology and molecular biology reviews: MMBR*, 63(3), pp. 725–734.
- Neuwald, P. D. et al.** (1977) 'Early gene expression of adenovirus type 2: R-loop mapping of mRNA and time course of viral DNA, mRNA, and protein synthesis.', *Journal of virology*, 21(3), pp. 1019–30.
- Nevins, J. R. et al.** (1979) 'Regulation of the primary expression of the early adenovirus transcription units.', *Journal of virology*, 32(3), pp. 727–33.
- Nevins, J. R.** (1987) 'Regulation of early adenovirus gene expression.', *Microbiological reviews*, 51(4), pp. 419–30.
- Nevins, J. R. and Darnell, J. E.** (1978) 'Groups of adenovirus type 2 mRNA's derived from a large

primary transcript: probable nuclear origin and possible common 3' ends', *J Virol*, 25(3), pp. 811–823.

Nevins, J. R. and Wilson, M. C. (1981a) 'Regulation of adenovirus-2 gene expression at the level of transcriptional termination and RNA processing', *Nature*, 290(5802), pp. 113–118.

Nevins, J. R. and Wilson, M. C. (1981b) 'Regulation of adenovirus-2 gene expression at the level of transcriptional termination and RNA processing', *Nature*, 290, pp. 113–118.

Newcomb, W. W., Boring, J. W. and Brown, J. C. (1984) 'Ion etching of human adenovirus 2: structure of the core.', *Journal of virology*, 51(1), pp. 52–56.

Okuwaki, M. et al. (2001) 'Identification of nucleophosmin/B23, an acidic nucleolar protein, as a stimulatory factor for in vitro replication of adenovirus DNA complexed with viral basic core proteins', *Journal of Molecular Biology*, 311(1), pp. 41–55.

van Oostrum, J. and Burnett, R. M. (1985) 'Molecular composition of the adenovirus type 2 virion.', *Journal of virology*, 56(2), pp. 439–48.

Orazio, N. I. et al. (2011) 'The Adenovirus E1b55K/E4orf6 Complex Induces Degradation of the Bloom Helicase during Infection', *Journal of Virology*, 85(4), pp. 1887–1892.

Orazio, N. I. et al. (2011) 'The adenovirus E1b55K/E4orf6 complex induces degradation of the Bloom helicase during infection.', *Journal of virology*, 85(4), pp. 1887–92.

Ornelles, D. A. and Shenk, T. (1991) 'Localization of the adenovirus early region 1B 55-kilodalton protein during lytic infection: association with nuclear viral inclusions requires the early region 4 34-kilodalton protein', *Journal of virology*, 65(1), pp. 424–429.

Pardo-Mateos, A. and Young, C. S. H. (2004) 'Adenovirus IVa2 protein plays an important role in transcription from the major late promoter in vivo', *Virology*, 327(1), pp. 50–59.

Pennella, M. A. et al. (2010) 'Adenovirus E1B 55-Kilodalton Protein Is a p53-SUMO1 E3 Ligase That Represses p53 and Stimulates Its Nuclear Export through Interactions with Promyelocytic Leukemia Nuclear Bodies', *Journal of Virology*, 84(23), pp. 12210–12225.

Pérez-Berná, A. J. et al. (2009) 'Structure and Uncoating of Immature Adenovirus', *Journal of Molecular Biology*, 392(2), pp. 547–557.

Perricaudet, M. et al. (1979) 'Structure of two spliced mRNAs from the transforming region of human subgroup C adenoviruses [23]', *Nature*, pp. 694–696.

Pilder, S. et al. (1986) 'The Adenovirus E1B-55K Transforming Polypeptide Modulates Transport or', 6(2), pp. 470–476.

Pombo, A. et al. (1994) 'Adenovirus replication and transcription sites are spatially separated in the nucleus of infected cells.', *The EMBO journal*, 13(21), pp. 5075–85.

Pombo, a et al. (1994) 'Adenovirus replication and transcription sites are spatially separated in the nucleus of infected cells.', *The EMBO journal*, 13(21), pp. 5075–85.

Puntener, D. et al. (2011) 'Stepwise Loss of Fluorescent Core Protein V from Human Adenovirus during Entry into Cells †', *Society*, 85(1), pp. 481–496.

Puvion-Dutilleul, F. et al. (1994) 'Rearrangements of intranuclear structures involved in RNA processing in response to adenovirus infection.', *Journal of cell science*, 107 (Pt 6, pp. 1457–68.

Puvion-Dutilleul, F., Roussev, R. and Puvion, E. (1992) 'Distribution of viral RNA molecules during the

- adenovirus type 5 infectious cycle in HeLa cells.', *Journal of structural biology*, 108(3), pp. 209–20.
- Querido, E. et al.** (2001) 'Degradation of p53 by adenovirus E4orf6 and E1B55K proteins occurs via a novel mechanism involving a Cullin-containing complex', *Genes and Development*, 15(23), pp. 3104–3117.
- Rabino, C. et al.** (2000) 'Adenovirus late gene expression does not require a Rev-like nuclear RNA export pathway.', *Journal of virology*, 74(14), pp. 6684–6688.
- Rebelo, L. et al.** (1996) 'The Dynamics of Coiled Bodies in the Nucleus of Adenovirus-infected Cells', *Molecular Biology of the Cell*, 7(July), pp. 1137–1151.
- Reich, N. C. et al.** (1983) 'Monoclonal antibodies which recognize native and denatured forms of the adenovirus DNA-binding protein', *Virology*, 128(2), pp. 480–484.
- Renee Yew, P., Liu, X. and Berk, A. J.** (1994) 'Adenovims E1B oncoprotein tethers a transcriptional repression domain to p53', *Genes and Development*, 8(2), pp. 190–202.
- Richard, P., Vethantham, V. and Manley, J. L.** (2017) 'SUMO Regulation of Cellular Processes', *Adv Exp Med Biol*, 963, pp. 15–33.
- Rigolet, M. and Galibert, F.** (1984) 'Organization and expression of the E4 region of adenovirus 2', *Nucleic Acids Research*, 12(20), pp. 7649–7661.
- Roberts, R. J., O'Neill, K. E. and Yen, C. T.** (1984) 'DNA sequences from the adenovirus 2 genome', *Journal of Biological Chemistry*, 259(22), pp. 13968–13975.
- Robinson, C. M. et al.** (2011) 'Computational analysis and identification of an emergent human adenovirus pathogen implicated in a respiratory fatality', *Virology*, 409(2), pp. 141–147.
- Rosen, L.** (1960) 'A hemmagglutination-inhibition technique for typing adenoviruses', *American Journal of Hygiene*, (71), pp. 120–128.
- Rougemaille, M. et al.** (2008) 'mRNA journey to the cytoplasm: attire required.', *Biology of the cell / under the auspices of the European Cell Biology Organization*, 100(6), pp. 327–42.
- Rowe, W. P. et al.** (1953) 'Isolation of a Cytopathogenic Agent from Human Adenoids Undergoing Spontaneous Degeneration in Tissue Culture', *Experimental Biology and Medicine*, 84(3), pp. 570–573.
- Russell, W. C., Laver, W. G. and Sanderson, P. J.** (1968) 'Internal components of adenovirus', *Nature*, 219(5159), pp. 1127–1130.
- Saban, S. D. et al.** (2006) 'Visualization of α -Helices in a 6-Angstrom Resolution Cryoelectron Microscopy Structure of Adenovirus Allows Refinement of Capsid Protein Assignments', *Journal of Virology*, 80(24), pp. 12049–12059.
- Sambrook, J. et al.** (1980) 'THE EVOLUTION OF THE ADENOVIRAL GENOME', *Annals of the New York Academy of Sciences*, 354(1), pp. 426–452.
- Sambrook, J., Fritsch, E. F. and Maniatis., T.** (1989) 'Molecular cloning: A laboratory manual.', in. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.
- Sarnow, P. et al.** (1982) 'Adenovirus E1b-58kd tumor antigen and SV40 large tumor antigen are physically associated with the same 54 kd cellular protein in transformed cells', *Cell*, 28(2), pp. 387–394.
- Schaack, J. et al.** (1990) 'Adenovirus terminal protein mediates both nuclear matrix association and

efficient transcription of adenovirus DNA', *Genes and Development*, 4(7), pp. 1197–1208.

Schmid, M. et al. (2014) 'DNA virus replication compartments.', *Journal of virology*, 88(3), pp. 1404–20.

Schmid, M., Gonzalez, R. A. and Dobner, T. (2012) 'CRM1-Dependent Transport Supports Cytoplasmic Accumulation of', 86(4).

Schmid, M., Gonzalez, R. and Dobner, T. (2012) 'CRM1-dependent transport supports cytoplasmic accumulation of adenoviral early transcripts.', *Journal of virology*, 86(4), pp. 2282–92.

Schreiner, S. et al. (2010) 'Proteasome-dependent degradation of Daxx by the viral E1B-55K protein in human adenovirus-infected cells.', *Journal of virology*, 84(14), pp. 7029–38.

Schreiner, S. et al. (2011) 'Adenovirus type 5 early region 1B 55K oncoprotein-dependent degradation of cellular factor Daxx is required for efficient transformation of primary rodent cells.', *Journal of virology*, 85(17), pp. 8752–65.

Shaw, A. R. and Ziff, E. B. (1980) 'Transcripts from the adenovirus-2 major late promoter yield a single early family of 3' coterminal mRNAs and five late families', *Cell*, 22(3), pp. 905–916.

Smart, J. E. and Stillman, B. W. (1982) 'Adenovirus terminal protein precursor. Partial amino acid sequence and the site of covalent linkage to virus DNA', *Journal of Biological Chemistry*, 257(22), pp. 13499–13506.

Smiley, J. K. et al. (1995) 'The metabolism of small cellular RNA species during productive subgroup C adenovirus infection', *Virology*, 206(1), pp. 100–107.

Smiley, J. K., Young, M. a and Flint, S. J. (1990) 'Intranuclear location of the adenovirus type 5 E1B 55-kilodalton protein.', *Journal of virology*, 64(9), pp. 4558–4564.

Sohn, S.-Y. and Hearing, P. (2012) 'Adenovirus regulates sumoylation of Mre11-Rad50-Nbs1 components through a paralog-specific mechanism.', *Journal of virology*, 86(18), pp. 9656–65.

Sohn, S. Y. and Hearing, P. (2016) 'Adenovirus early proteins and host sumoylation', *mBio*, 7(5), pp. 1–7.

Spector, D. J., McGrogan, M. and Raskas, H. J. (1978) 'Regulation of the appearance of cytoplasmic RNAs from region 1 of the adenovirus 2 genome', *Journal of Molecular Biology*, 126(3), pp. 395–414.

Speiseder, T., Nevels, M. and Dobner, T. (2014) 'Determination of the transforming activities of adenovirus oncogenes', *Methods in Molecular Biology*, 1089, pp. 105–115.

De Stanchina, E. et al. (1998) 'E1A signaling to p53 involves the p19(ARF) tumor suppressor', *Genes and Development*, 12(15), pp. 2434–2442.

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), pp. 2027–2035.

Stevens, J. L. et al. (2002) 'Transcription control by E1A and MAP kinase pathway via Sur2 Mediator subunit', *Science*, 296(5568), pp. 755–758.

Stracker, T. H., Carson, C. T. and Weitzman, M. D. (2002) 'Adenovirus oncoproteins inactivate the Mre11 Rad50 NBS1 DNA repair complex', *Nature*, 418(6895), p. 348.

Stuiver, M. H. and van der Vliet, P. C. (1990) 'Adenovirus DNA-binding protein forms a multimeric protein complex with double-stranded DNA and enhances binding of nuclear factor I.', *Journal of*

virology, 64(1), pp. 379–86.

Sun, Q. et al. (2014) 'Epidemiology and clinical features of respiratory adenoviral infections in children', *European Journal of Pediatrics*, 173(4), pp. 441–444.

Svensson, C., Pettersson, U. and Akusjärvi, G. (1983) 'Splicing of adenovirus 2 early region 1A mRNAs is non-sequential', *Journal of Molecular Biology*, 165(3), pp. 475–495.

Tal, J., Craig, E. and Raskas, H. (1975) 'Sequence relationships between adenovirus 2 early RNA and viral RNA size classes synthesized at 18 hours after infection.', *J Virol*, 15(1), pp. 137–144.

Tebruegge, M. and Curtis, N. (2010) 'Adenovirus infection in the immunocompromised host', *Advances in Experimental Medicine and Biology*, 659, pp. 153–174.

Thomas, G. P. and Mathews, M. B. (1980) 'DNA replication and the early to late transition in adenovirus infection.', *Cell*, 22(2 Pt 2), pp. 523–33.

Tigges, M. A. and Raskas, H. J. (1984) 'Splice junctions in adenovirus 2 early region 4 mRNAs: multiple splice sites produce 18 to 24 RNAs', *J. Virol.*, 50(1), pp. 106–117.

Tihanyi, K. et al. (1993) 'Isolation and properties of adenovirus type 2 proteinase', *J Biol Chem*, 268(3), pp. 1780–1785.

Tollefson, A. E. et al. (1992) 'The 11,600-MW protein encoded by region E3 of adenovirus is expressed early but is greatly amplified at late stages of infection.', *Journal of virology*, 66(6), pp. 3633–42.

Tollefson, A. E. et al. (1996) 'The adenovirus death protein (E3-11.6K) is required at very late stages of infection for efficient cell lysis and release of adenovirus from infected cells.', *Journal of virology*, 70(4), pp. 2296–306.

Törmänen, H. et al. (2006) 'L4-33K, an adenovirus-encoded alternative RNA splicing factor', *Journal of Biological Chemistry*, 281(48), pp. 36510–36517.

Törmänen Persson, H. et al. (2012) 'Two cellular protein kinases, DNA-PK and PKA, phosphorylate the adenoviral L4-33K protein and have opposite effects on L1 alternative RNA splicing.', *PloS one*, 7(2), p. e31871.

Trentin, J. J., Yabe, Y. and Taylor, G. (1962) 'The quest for human cancer viruses.', *Science (New York, N.Y.)*, 137(3533), pp. 835–41.

Trotman, L. C. et al. (2001) 'Import of adenovirus DNA involves the nuclear pore complex receptor CAN/Nup214 and histone H1', *Nature Cell Biology*, 3(12), pp. 1092–1100.

Ullman, A. J., Reich, N. C. and Hearing, P. (2007) 'Adenovirus E4 ORF3 Protein Inhibits the Interferon-Mediated Antiviral Response', *Journal of Virology*, 81(9), pp. 4744–4752.

Vassileva, M. T. and Matunis, M. J. (2004) 'SUMO Modification of Heterogeneous Nuclear Ribonucleoproteins', *Molecular and cellular biology*, 24(9), pp. 3623–3632.

Velicer, L. F. and Ginsberg, H. S. (1970) 'Synthesis, transport, and morphogenesis of type adenovirus capsid proteins.', *Journal of virology*, 5(3), pp. 338–352.

Virtanen, A. et al. (1984) 'mRNAs from human adenovirus 2 early region 4', *J Virol*, 51(3), pp. 822–831.

Wang, G. et al. (2005) 'Mediator requirement for both recruitment and postrecruitment steps in transcription initiation', *Molecular Cell*, 17(5), pp. 683–694.

- Weber, J. and Philipson, L.** (1984) 'Protein composition of adenovirus nucleoprotein complexes extracted from infected cells', *Virology*, 136(2), pp. 321–327.
- Webster, A., Hay, R. T. and Kemp, G.** (1993) 'The adenovirus protease is activated by a virus-coded disulphide-linked peptide', *Cell*, 72(1), pp. 97–104.
- Whyte, P. et al.** (1988) 'Association between an oncogene and an anti-oncogene: the adenovirus E1A proteins bind to the retinoblastoma gene product', *Nature*, 334(6178), pp. 124–129.
- Wides, R. J. et al.** (1987) 'Adenovirus origin of DNA replication: sequence requirements for replication in vitro.', *Molecular and cellular biology*, 7(2), pp. 864–74.
- Wiethoff, C. M. et al.** (2005) 'Adenovirus Protein VI Mediates Membrane Disruption following Capsid Disassembly', *Journal of Virology*, 79(4), pp. 1992–2000.
- Williams, J. et al.** (1986) 'The adenovirus E1B 495R protein plays a role regulating transport and stability of the viral late messages.', *Cancer Cells*, 4, pp. 275–284.
- Wilson, M. C. et al.** (1978) 'Adenovirus Type 2 mRNA in Transformed Cells : Map Positions and Difference in Transport Time', 25(1), pp. 97–103.
- Wilson, M. C. and Darnell, J. E.** (1981) 'Control of Messenger RNA Concentration by Differential Cytoplasmic Half-life', *J. Mol. Biol*, 148, pp. 231–251.
- Wimmer, P. et al.** (2010) 'SUMO modification of E1B-55K oncoprotein regulates isoform-specific binding to the tumour suppressor protein PML', *Oncogene*. Nature Publishing Group, 29(40), pp. 5511–5522.
- Wimmer, P. et al.** (2013) 'Cross-talk between phosphorylation and SUMOylation regulates transforming activities of an adenoviral oncoprotein', *Oncogene*. Nature Publishing Group, 32(13), pp. 1626–1637.
- Winberg, G. and Shenk, T.** (1984) 'Dissection of overlapping functions within the adenovirus type 5 E1A gene.', *The EMBO journal*, 3(8), pp. 1907–12.
- Wold, W. S. M. and Toth, K.** (2013) 'Adenovirus vectors for gene therapy, vaccination and cancer gene therapy.', *Current gene therapy*, 13(6), pp. 421–33.
- Woo, J. L. and Berk, A. J.** (2007) 'Adenovirus ubiquitin-protein ligase stimulates viral late mRNA nuclear export.', *Journal of virology*, 81(2), pp. 575–87.
- Wu, K., Guimet, D. and Hearing, P.** (2013a) 'The adenovirus L4-33K protein regulates both late gene expression patterns and viral DNA packaging.', *Journal of virology*, 87(April), pp. 6739–47.
- Wu, K., Guimet, D. and Hearing, P.** (2013b) 'The adenovirus L4-33K protein regulates both late gene expression patterns and viral DNA packaging.', *Journal of virology*, 87(12), pp. 6739–47.
- Xi, Q., Cuesta, R. and Schneider, R. J.** (2004) 'Tethering of eIF4G to adenoviral mRNAs by viral 100k protein drives ribosome shunting', *Genes and Development*, 18(16), pp. 1997–2009.
- Xi, Q., Cuesta, R. and Schneider, R. J.** (2005) 'Regulation of Translation by Ribosome Shunting through Phosphotyrosine-Dependent Coupling of Adenovirus Protein 100k to Viral mRNAs', *Society*, 79(9), pp. 5676–5683.
- Yang, U., Huang, W. and Flint, S. J.** (1996) 'mRNA Export Correlates with Activation of Transcription in Human Subgroup C Adenovirus-Infected Cells', *Journal of virology*, 70(6), pp. 4071–4080.
- Yatherajam, G., Huang, W. and Flint, S. J.** (2011) 'Export of adenoviral late mRNA from the nucleus

requires the Nxf1/Tap export receptor.', *Journal of virology*, 85(4), pp. 1429–38.

Yew, P. R. and Berk, A. J. (1992) 'Inhibition of p53 transactivation required for transformation by adenovirus early 1B protein', *Nature*, 357(6373), pp. 82–85.

Yueh, A. and Schneider, R. J. (1996) 'Selective translation initiation by ribosome jumping in adenovirus- infected and heat-shocked cells', *Genes and Development*, 10(12), pp. 1557–1567.

Zhang, W. and Arcos, R. (2005) 'Interaction of the adenovirus major core protein precursor, pVII, with the viral DNA packaging machinery', *Virology*, 334(2), pp. 194–202.

Zhang, W. and Imperiale, M. J. (2003) 'Requirement of the Adenovirus IVa2 Protein for Virus Assembly', *Journal of virology*, 77(6), pp. 3586–3594.

Zhang, Y., Feigenblum, D. and Schneider, R. J. (1994) 'A late adenovirus factor induces eIF-4E dephosphorylation and inhibition of cell protein synthesis', *J Virol*, 68(11), pp. 7040–7050.

Zhang, Z. et al. (2011) 'Genetic Analysis of B55 /Cdc55 Protein Phosphatase 2A Subunits: Association with the Adenovirus E4orf4 Protein', *Journal of Virology*, 85(1), pp. 286–295.

Zhao, H., Chen, M. and Pettersson, U. (2014) 'A new look at adenovirus splicing.', *Virology*. Elsevier, 456–457, pp. 329–41.

Zhao, L. Y. et al. (2003) 'Adenovirus E1B 55-kilodalton oncoprotein binds to Daxx and eliminates enhancement of p53-dependent transcription by Daxx.', *Journal of virology*, 77(21), pp. 11809–21.

Ziff, E. and Fraser, N. (1978) 'Adenovirus type 2 late mRNA's: structural evidence for 3'-coterminal species.', *J Virol*, 25(3), pp. 897–906.

7 Tables

Table 1. Annotations used for HAdV-5 transcriptome

Transcript_id	Region	Note
E1A 10	join(499..637,854..974,1229..1545)	*
E1A 11S	join(499..637,854..1112,1229..1545)	*
E1A 12S	join(499..974,1229..1545)	
E1A 13S	join(499..1112,1229..1545)	
E1A 9S	join(499..637,1229..1545)	*
E1B new	join(1702..2255,3598..4070)	*
E1B-13S (19K)	join(1702..2255,3595..4070)	
E1B-14.5S	join(1702..2255,3218..3510,3595..4070)	
E1B-14S	join(1702..2255,3276..3510,3595..4070)	
E1B-179R	join(1702..2324,3276..3510,3595..4070)	
E1B-22S (55K)	join(1702..3510,3595..4070)	
E2A DBP	complement(22443..24032)	
E2A	complement(join(22374..24032,24667..24745,26977..27045))	*
E2A new1	complement(join(22374..24042,24146..24745,26977..27045))	*
E2A new2	complement(join(22374..24331,24667..24745,26977..27045))	*
E2A-L	complement(join(22374..24032,24667..24745,25838..25909))	*
	complement(join(4060..8785,14111..14299,24667..24745,26977..27045))	*
E2B (Ad pol)	complement(join(5197..8784,14112..14120))	
E2B 120K DNA polymerase	complement(join(5197..8784,14112..14120))	
E2B IVa2	complement(join(4091..5427,5706..5718))	
E2B new	complement(join(4060..20695,20823..24745,26977..27045))	*
E2B pTP (75K)	complement(join(4060..10590,14111..14299,26977..27045))	*
E2B pTP 75K	complement(join(8583..10589,14112..14120))	
E3 12.5K	27852..28175	
E3 14.7K (truncated! 5' 24 nts deleted)	28593..28955	
E3 CR1-alpha0	28541..28592	
E3-1	27562..28592	
E4orf1_1 late	complement(join(30932..31308,32020..33642,33658..33724))	**
	complement(join(30932..31308,32020..32198,32719..33642,33658..33724))	**
E4orf1_2 late	complement(join(30932..33208,33658..33724))	**
E4orf2 early	complement(join(30932..32848,33658..33724))	**
E4orf3 early	complement(join(30932..31308,32020..32848,33658..33724))	**
E4orf3 late	complement(join(30932..31308,32020..32198,32719..32848,33658..33724))	**
E4orf3/4 late	complement(join(30932..32548,33658..33724))	**
E4orf4 early	complement(join(30932..31308,32020..32548,33658..33724))	**
E4orf4 late	complement(join(30932..32358,33658..33724))	*
E4orf6 new	complement(join(30932..32445,33658..33724))	**
E4orf6_1 early	complement(join(30932..32198,32404..32548,33658..33724))	**
E4orf6_2 late	complement(join(30932..31308,32020..32445,33658..33724))	**
E4orf6/7_1 late	complement(join(30932..31308,32020..32198,33658..33724))	**
E4orf6/7_2 late	complement(join(30932..31308,32020..32198,33658..33724))	**
L1 13.6K	7978..8427	***
L1 52. 55K	join(6049..6090,7110..7183,9643..9734,11050..12297)	
L1 pIIIa	join(6049..6090,7110..7183,9643..9734,12318..14075)	
L2 pIII (penton base)	join(6049..6090,7110..7183,9643..9734,14157..15872)	
L2 pV	join(6049..6090,7110..7183,9643..9734,16545..17651)	
L2 pVII	join(6049..6090,7110..7183,9643..9734,15879..16475)	
L2 pVII mRNA1	join(6049..6090,7110..7183,9643..9734,15659..16475)	*
L2 pVII mRNA2	join(6049..6090,7110..7183,9643..9734,15665..16475)	*
L2 pVII mRNA3	join(6049..6090,7110..7183,9643..9734,15729..16475)	*
L2 pX (μ)	join(6049..6090,7110..7183,9643..9734,17679..17921)	
L3 23K protease	join(6049..6090,7110..7183,9643..9734,21733..22347)	
L3 Hexon	join(6049..6090,7110..7183,9643..9734,18842..21700)	
L3 pVI	join(6049..6090,7110..7183,9643..9734,18004..18756)	
L4 100K	join(6049..6090,7110..7183,9643..9734,24061..26484)	
L4 22K	join(6049..6090,7110..7183,9643..9734,26195..26779)	
L4 33K	join(6049..6090,7110..7183,9643..9734,26195..26510,26713..27080)	
L4 pVIII	join(6049..6090,7110..7183,9643..9734,27168..27851)	
L5 pIV Fiber	join(6049..6090,7110..7183,9643..9734,29158..30903)	
	join(6049..6090,7110..7183,7951..8392,9643..9734,27800..27934,29158..30903)	*
L5 pIV Fiber mRNA5	join(6049..6090,7110..7183,7951..8392,9643..9734,27800..27934,29158..30903)	*
L5 pIV Fiber mRNA8	3609..4031	
pIX	complement(join(30932..31308,32031..32123,33658..33724))	
predicted E4orf5	complement(join(22374..24042,24667..24745,28984..29147))	*
UXP		

* predicted according to (Zhao et al. 2014) according to Ad2 (NCBI:NC_001405)

** reported in (Dix and Leppard, 1993)

*** putative i-leader, (NCBI: AFS0547.1)

Table 2. Total viral mRNAs expression values.

Transcript ID	6 hpi	12 hpi	24 hpi	48 hpi
E1A 10S	16	79	942	1856
E1A 11S	47	333	751	1645
E1A 12S	217	961	6243	4344
E1A 13S	625	3597	10492	8223
E1A 9S	45	84	10186	14197
E1B new	6	22	23905	22784
E1B 13S (19K)	4	121	165963	137588
E1B 14.5S	4	18	17317	71150
E1B 14S	16	55	11499	33699
E1B 179R	3	45	1508	1968
E1B 22S (55K)	2558	20707	539057	599301
E2A DBP	79	363	47	12
E2B (Ad pol)	10	114	1496	1619
E2B 120K DNA polymerase	0	1	1	2
E2B IVa2	5	10	10203	6783
E2B new	2641	24998	1279368	2061783
E2B pTP (75K)	8	18	268	394
E2B pTP 75K	0	1	6	2
E3 12.5K	1	0	0	0
E3 14.7K (truncated! 5' 24 nts deleted)	7	10	0	0
E3 CR1-alpha0	0	0	0	0
E3-1	18	30	2	3
E4orf1_1 late	146	489	1694	769
E4orf1_2 late	39	144	741	1904
E4orf2 early	446	4648	15885	11968
E4orf3 early	270	1094	2150	1695
E4orf3 late	350	1490	3120	2514
E4orf3/4 late	215	2422	5566	4907
E4orf4 early	112	391	532	426
E4orf4 late	95	347	495	540
E4orf5 (predicted)	95	62	20	54
E4orf6 new	267	387	3376	4223
E4orf6_1 early	105	344	807	515
E4orf6_2 late	108	1109	1818	1563
E4orf6/7_1 late	95	288	616	647
E4orf6/7_2 late	157	984	12988	10560
L1 13.6K	4	69	44785	104914
L1 52, 55K	5	222	86093	135745
L1 pIIIa	20	514	295731	638883
L2 pIII	14	501	676752	2476344
L2 pV	7	226	499012	2145840
L2 pVII	0	2	279	1566
L2 pVII mRNA1	0	31	65753	242660
L2 pVII mRNA2	0	8	11538	48703
L2 pVII mRNA3	2	12	18942	44487
L2 pX (μ)	3	2	8857	28047
L3 23K protease	35	155	350896	1355877
L3 Hexon	112	1929	4441986	18975066
L3 pVI	2	74	162650	595347
L4 100K	35	831	818077	3027630
L4 22K	14	72	68574	238117
L4 33K	71	614	914460	2631104
L4 pVIII	36	364	254933	1009438
L5 pIV Fiber	44	550	2947663	11597781
L5 pIV Fiber mRNA5	31	243	10857	29061
L5 pIV Fiber mRNA8	29	227	11004	28766
pIX mRNA	0	2	2	6
TOTAL	9274	72414	13817906	48365020

Total viral mRNA expression values were calculated by the sum of total nucleus and cytoplasmic values of each mRNA. The most abundant mRNAs at each time point are shadowed in gray.

Table 3. Nuclear / Cytoplasmic ratios.

Transcript_ID	6 hpi	12 hpi	24 hpi
E1A 9S	-3.320786	-1.261488	1.084952
E1A 11S	-6.6429	-1.41039	1.249813
E1A 12S	-1.399266	-1.417937	-1.263204
E1A 13S	1.864385	-1.598076	-1.745581
E1A 10S	0	1.149568	1.257156
E1B-14S	-6.410934	1.439114	1.161946
E1B-14.5S	-9.19956	1.384986	1.271613
E1B-179R	-10.099007	1.547351	-1.473924
E1B-13S (19K)	2.01893	1.93489	1.236597
E1B-22S (55K)	1.049779	-1.379934	-1.30159
E1B new	-7.962768	1.46486	1.026466
E2A DBP	-5.698974	1.25742	1.305485
E2B pTP (75K)	0	-1.10695	1.305485
E2B (Ad pol)	0	-1.458925	-1.097496
E2B new	0	-1.91519	1.288929
E2B IVa2	3.330275	4.580445	1.740494
E3-1	-8.653351	1.948673	0
E3 CR1-alpha0	1.546465	1	1
E3 14.7K (truncated)	-8.740425	1.368872	1
E3 12.5K	1.885684	0	0
E4orf3/4 late	10.000649	-1.329579	-1.079062
E4orf6_1 early	-2.830808	1.059772	-1.536149
E4orf4 early	2.08327	-1.00159	-1.093165
E4orf3 early	2.161171	-1.29785	1.109539
E4orf2 early	1.850476	-1.831962	-1.296335
E4orf6 new	9.098388	-1.579623	-1.091682
E4orf6/7_1 late	1.868148	-1.149607	1.199588
E4orf5 predicted	7.606923	1.00013	-1.66196
E4orf6/7_2 late	8.051263	-1.404631	-1.405668
E4orf1_2 late	2.603057	-1.821588	-1.689645
E4orf4 late	-2.12643	-1.007739	1.295126
E4orf3 late	3.508231	1.029558	1.410235
E4orf6_2 late	2.597734	-1.131439	1.410847
E4orf1_1 late	1.163432	1.234598	-2.197476
L1 13.6K	1.586948	1.262974	-2.371829
L2 pV	1.309116	1.819601	1.051379
L2 pVII	1.094919	4.355591	1.427284
L1 pIIIa	1.548796	-1.309067	-1.48205
L1 52, 55K	1.323484	-1.584524	-1.39165
L5 pIV Fiber 8	-3.050837	2.676616	-1.145028
L5 pIV Fiber 5	1.722956	3.05992	-1.353074
L2 pVII 3	1.547206	2.506049	1.378708
L2 pVII 2	1.54681	2.778863	1.35573
L2 pVII 1	1.58311	-1.124403	1.568758
L3 pVI	1.780582	1.827635	1.697807
L5 pIV Fiber	2.88452	2.559913	1.05817
L2 pIII	2.995721	1.936532	-1.175242
L4 pVIII	3.144626	1.32531	1.230518
L4 33K	4.045763	1.839037	-1
L4 22K	2.999286	1.602877	1.37745
L4 100K	2.827386	1.304946	1.090096
L3 23K protease	2.833319	2.504846	-1
L3 Hexon	2.947994	2.787393	-1.063273
L2 pX (mu)	2.725647	1.03195	-1.043854
pIX	1.842884	5.509265	0

Ratios were obtained using the Baggerly's proportions test and then were subjected to a log2 transformation.

Table 4. Cellular mRNAs clustered by cyto/nuc ratios

A	B	C	D	E	F	G
AGA_2	AHNAK2_4	ATL2_13	AKR1B1_14	ACLY_2	ANXA1_3	ABCC3_9
AGR2_2	ANXA4_5	CD46_19	ASAH1_8	ARRB2_10	APP_2	ACTB_10
ARPC2_4	ANXA5_4	CLU_12	ATP6AP1_3	ELOVL5_6	ARGLU1_4	ADIPOR1_1
ASAH1_12	ATP1B3_11	F5_2	CALU_7	ETFA_5	ASAH1_9	ALDH1A1_2
ATP5L_5	BZW1_3	FN1_9	FADS2_3	HNRNPK_5	ATP1A1_8	ANXA1_7
AUP1_2	CALU_4	FTH1_4	FLNA_12	LAMB1_8	C1R_11	ANXA2_18
B2M_9	CANX_11	GJA1_1	NCL_3	MSMO1_2	CALR_1	ARL6IP4_11
BRI3_3	CAPN2_8	HECTD1_17	P4HB_16	TGFBL14	CALR_5	ATP2A2_7
C1GALT1C1_2	CCT5_3	HSP90B1_2	RPS26_2		CCPG1_13	ATP5B_1
C5orf15_2	CCT8_9	NIT2_7	RPSA_9		CCT6A_10	ATP6AP2_7
CALU_8	CD44_17	PTGS2_1	SAT1_3		CLTC_4	ATRAID_7
CD59_3	CFH_4	RNF149_3	SDC1_2		CSE1L_4	BEST1_8
CLGN_3	CTSL_4	SDF2_6	TPM3_22		EIF4A2_6	C1S_19
CP_8	DDOST_5	SEP15_3	VCAN_10		HEXB_3	CA12_4
CTSC_4	DDOST_6	SPP1_9			HNRNPK_10	CCT6A_3
CTSS_1	ERP44_1				HSP90B1_3	CHMP3_9
CYP24A1_2	GNB2L1_32				HSP90B1_9	CLU_1
DNAJB9_1	HM13_24				IL18_5	CP_10
DNAJC2_4	IDH1_7				JAG1_1_3	CTD-2139B15.2_1
EPT1_1	ITFG1_9				MRPS5_1	CTD-2354A18.1_1
FAS_5	ITGB1_3				NRCAM_14	CTD-2354A18.1_3
FGB_5	MATR3_39				NRP1_5	CTDNEP1_12
FTH1_10	MFGE8_15				OGT_3	CTDNEP1_6
FTH1_2	NUCB2_17				PPIA_8	CTSF_3
FTH1_3	PON2_2				RPN2_8	CTSL_3
FTH1_5	POR_9				RPS24_8	CTSL_5
FTH1_6	PRKDC_9				SAT2_11	CXCL5_1
FTH1_7	PSAP_1				SLC16A1_1	DBNDD1_4
FTL_1	RDX_18				SLC25A3_10	DEGS1_1
GCNT3_10	RPL12_5				SLC2A1_4	DHRS7_7
GGH_2	RTN4_3				TM2D1_8	DKK1_2
GINM1_2	SERPINE1_2				TMBIM6_4	DMTN_2
GNAS_4	SERPINE2_7				VMP1_6	DSG2_1
GPS2_4	SLC27A2_2				VTI1B_5	EIF4EBP2_1
GSK3A_3	STT3A_8				XPO1_3	EIF4G2_28
HEXB_12	TFRC_2					EMC3_1
HNRNPDL_5	TGFBL7					EPAS1_2
HNRNPUL1_13	TM9SF2_1					EPN1_4
HSD17B12_8	UBE2V1_20					EWSR1_17
HSPA8_9	VTN_2					F2RL1_2
HSPB1_2						FAM136A_4
IDS_1_6						FBL1
IFNGR1_3						FLNB_1
LAMP2_2						FN1_16
LAMP2_4						FN1_20
LIPA_6						FTH1_1
LITAF_14						FTH1_8
MAGT1_2						GANAB_9
MGST2_3						GNAS_44
MTUS1_30						GPX8_7
NEU1_2						GUK1_31
NME4_6						HDLBP_31
NPTN_4						HEXA_9
OCIA1_12						HLA-B_1
PABPC1_13						HMGB1_3
PON2_11						HNRNPA2B1_7
PRCP_3						HNRNPH1_36
PRDX4_2						HNRNPL_13
PTRH2_5						HRAS_7
RDH11_11						HSPA5_1
						HTRA1_1
RP11-396K3.1_2						
RPL36A_4						IDH1_9
RPL3_17						IGFBP3_8
RPS2_5						IRAK1_11
SAT1_8						ITGA3_6
SEC11A_7						ITGB1_1
SLC35F5_2						ITM2B_4
SLC38A2_6						JKAMP_11
SLC3A2_5						LAMB1_13
SRGN_2						LAMB2_11
SRPRB_5						LAPTM4A_2
SRSF7_2						LBR_8
SRSF7_6						LUC7L3_10
STK24_7						MAT2A_3
SUMF2_18						MAT2A_6
SYPL1_2						MLPH_6
TFPL10						MMP24-AS1_5
TFPL9						MMP7_2
TFRC_6						MORF4L2_4

TM4SF4_2
TMCO1_2
TMCO1_3
TMEM106B_4
TMEM165_12
TMEM9_5
TNNT1_5
TXNDC15_4
UFL1_2
VMP1_15

ZMPSTE24_1
ZNF358_1

MYL6_3
NME4_5
NOP56_10
NPTN_2
NT5DC2_6
NTAN1_4
OR51E1_2
OSBPL9_15
OSMR_2
PCBP2_2
PDAP1_2

PDCD5_8
PFKP_4
PGK1_2
PLOD2_1
PPP1R14B_1
PPT1_5
PROS1_1
PSMC5_2
PTPLAD1_2
QPCT_6
RAN_2
RBM39_32
RPLP0_16
RPN1_3
RPS24_1
RPS2_4
RPS2_6
RTN4_10
SAT1_6
SEPT2_4
SLC25A23_2
SLC25A39_17
SLC2A4RG_1
SLC35A4_3
SNF8_11
SPP1_8
SPTBN1_6
SRM1
SRSF5_18
TAF10_5
TFPI2_1
TFPI2_2
TFPI_12
TM4SF1_3
TMCO1_5
TMED7_1
TMEM2_6
TMEM59_9
TPM2_3
TSR3_2
U2SURP_3
UCHL1_2
USF2_14
UXS1_6
VMP1_12
XBP1_2

Table 5. Top most abundant cellular mRNAs during infection with HAdV-5.

Transcript_ID	mock	Transcript_ID	6 hpi	Transcript_ID	12 hpi	Transcript_ID	24 hpi	Transcript_ID	48 hpi
MT-CO1_1	3783091	MT-ND4_1	2853740	MT-ND4_1	5086110	MT-ND4_1	3492953	HSP90AA1_7	1162716
HMGB1_3	2709171	MT-CO1_1	1429703	MT-CO1_1	3135777	MT-CO1_1	2322469	AKR1B10_3	1071133
HSP90AA1_7	2632634	HSP90AA1_7	1286921	HMGB1_3	2052746	HMGB1_3	1455053	RPL7_1	90750
AKR1B10_3	2359903	HMGB1_3	1030090	HSP90AA1_7	1905244	HSP90AA1_7	1376563	NPM1_5	87889
ALDH1A1_2	1613839	ALDH1A1_2	987985	AKR1B10_3	1826921	AKR1B10_3	1321628	DDOST_5	78161
MT-ND5_1	1212127	AKR1B10_3	876498	MT-CO2_1	1607606	MT-ATP6_1	1293645	RPL12_1	63271
CYP24A1_2	946902	EEF1A1_3	524778	ALDH1A1_2	1530898	ALDH1A1_2	1163818	RPL23A_3	52208
CP_3	652381	AKR1B10_4	351460	MT-ND5_1	1099835	EEF1A1_3	720382	CA12_4	49915
FTL_1	473592	FTL_1	336371	EEF1A1_3	852207	CYP24A1_2	505124	TRAM1_3	37472
AKR1B10_4	468894	CYP24A1_2	311950	CYP24A1_2	695051	CP_3	416924	RPL30_8	28349
ETFA_5	362561	TMSB10_1	261989	CP_3	603476	MT-CYB_1	383921	RPLP0_10	25756
PPIA_7	275709	MT-ND3_1	256347	MT-CYB_1	558434	PPIA_7	353599	TSPAN3_4	25622
FTH1_1	257240	CP_3	243847	AKR1B10_4	518872	FTL_1	345480	RPL41_5	22247
AKR1C3_1	247496	ETFA_5	219428	FTL_1	456397	AKR1B10_4	342488	ABCC2_3	21139
KRT81_1	242428	PPIA_7	172001	MT-ND3_1	445715	ETFA_5	198129	HDLBP_31	19638
RPS24_8	240398	AKR1C3_1	168839	PPIA_7	439806	FTH1_1	187437	ASPH_12	17899
TXNRD1_24	234210	DDOST_5	166239	TMSB10_1	399158	RPS4X_5	171596	RPS24_9	17035
ENO1_6	209240	MT-ATP8_1	137611	GAPDH_8	301388	RPS24_8	163121	GOLM1_3	14849
DDOST_5	208654	RPS24_8	136698	ETFA_5	291433	RPL7_1	162721	SNHG6_3	14220
GNB2L1_1	202204	TMSB4X_4	136690	FTH1_1	252684	ANXA2_18	150557	UBB_2	14123
SCD_1	196924	ANXA2_18	133822	TMEM2_6	228284	NPM1_5	141840	CRIM1_7	13921
TMEM2_6	188097	FTH1_1	131823	AKR1C3_1	215072	GNB2L1_1	141667	TFPI_10	12761
KRT18_6	182847	RPS6_3	129120	ENO1_6	207384	MT-ATP8_1	138917	PABPC1_12	12510
RPL12_1	173224	TMEM2_6	126198	RPL19_6	205679	TMEM2_6	135353	ILF2_1	11846
PGD_6	168449	RPL19_6	121686	RPS4X_5	200293	DDOST_5	132089	ATP2A2_7	11237
ANXA2_18	163387	RPL7_1	118005	RPL7_1	198785	RPL12_1	128908	AKR1C1_4	10425
PRDX1_4	162455	RPL12_1	111066	RPS24_8	198116	ANXA4_8	127843	MGST2_3	10153
KRT7_13	156506	ENO1_6	108359	MT-ATP8_1	185722	AKR1C1_2	118629	FTH1_7	10042
RPSA_9	151473	GNB2L1_1	107579	DDOST_5	174550	TMSB4X_4	116660	SERINC3_4	8944
TRAM1_3	148467	TRAM1_3	104789	RPS6_3	170413	PABPC1_2	110069	RTN4_3	8277
TPT1_2	145183	DSTN_2	100431	GNB2L1_1	169998	RPSA_9	105900	VTI1B_2	7529
EEF1G_3	144877	KRT18_6	100246	RPL12_1	165750	RPL13A_8	102499	PPIA_8	7360
PKM_23	133321	AKR1C1_2	97576	TMSB4X_4	164476	RPL23A_3	101409	HNRNPH1_36	7282
RPL23A_3	128191	TPT1_2	95950	HSP90AB1_3	163377	RPLP1_2	97361	TM4SF4_2	7180

Table 6. Top most abundant cellular mRNAs at 24 hpi.

Gene Symbol	Gene Name	Protein Class
AKR1B10	Aldo-keto reductase family 1 member B10;AKR1B10;ortholog	reductase(PC00176)
AKR1C1	Aldo-keto reductase family 1 member C1;AKR1C1;ortholog	reductase(PC00176)
ALDH1A1	Retinal dehydrogenase 1;ALDH1A1;ortholog	dehydrogenase(PC00176)
ANXA2	Annexin A2;ANXA2;ortholog	
ANXA4	Annexin A4;ANXA4;ortholog	
CP	Ceruloplasmin;CP;ortholog	transporter(PC00227);apolipoprotein(PC00219);membrane-bound signaling molecule(PC00052);receptor(PC00207);metalloprotease(PC00152);serine protease(PC00197);oxidase(PC00190);metalloprotease(PC00153);serine protease(PC00203);extracellular matrix protein(PC00176);enzyme modulator(PC00175);cell adhesion molecule(PC00121)
CYP24A1	1,25-dihydroxyvitamin D(3) 24-hydroxylase, mitochondrial;CYP24A1;ortholog	oxygenase(PC00176)
DDOST	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 48 kDa subunit;DDOST;ortholog	glycosyltransferase(PC00220)
EEF1A1	Elongation factor 1-alpha 1;EEF1A1;ortholog	translation elongation factor(PC00171);translation initiation factor(PC00031);hydrolase(PC00223);G-protein(PC00222)
ETFA	Electron transfer flavoprotein subunit alpha, mitochondrial;ETFA;ortholog	transferase(PC00220);dehydrogenase(PC00176);oxidase(PC00092)
FTH1	Ferritin heavy chain;FTH1;ortholog	storage protein(PC00210)
FTL	Ferritin light chain;FTL;ortholog	storage protein(PC00210)
GNB2L1	Guanine nucleotide-binding protein subunit beta-2-like 1;GNB2L1;ortholog	
HMGB1	High mobility group protein B1;HMGB1;ortholog	HMG box transcription factor(PC00218);signaling molecule(PC00024);chromatin/chromatin-binding protein(PC00207)
HSP90AA1	Heat shock protein HSP 90-alpha;HSP90AA1;ortholog	Hsp90 family chaperone(PC00072)
MT-ATP6	ATP synthase subunit a;MT-ATP6;ortholog	
MT-ATP8	ATP synthase protein 8;MT-ATP8;ortholog	ATP synthase(PC00227);hydrolase(PC00068)
MT-CO1	Cytochrome c oxidase subunit 1;MT-CO1;ortholog	oxidase(PC00176)
MT-CYB	Cytochrome b;MT-CYB;ortholog	
MT-ND4	NADH-ubiquinone oxidoreductase chain 4;MT-ND4;ortholog	dehydrogenase(PC00176);reductase(PC00092)
NPM1	Nucleophosmin;NPM1;ortholog	chaperone(PC00072)
PABPC1	Polyadenylate-binding protein 1;PABPC1;ortholog	
PPIA	Peptidyl-prolyl cis-trans isomerase A;PPIA;ortholog	isomerase(PC00135)
RPL12	60S ribosomal protein L12;RPL12;ortholog	ribosomal protein(PC00171)
RPL13A	60S ribosomal protein L13a;RPL13A;ortholog	ribosomal protein(PC00171)
RPL23A	60S ribosomal protein L23a;RPL23A;ortholog	ribosomal protein(PC00171)
RPL7	60S ribosomal protein L7;RPL7;ortholog	ribosomal protein(PC00171)
RPLP1	60S acidic ribosomal protein P1;RPLP1;ortholog	ribosomal protein(PC00171)
RPS24	40S ribosomal protein S24;RPS24;ortholog	ribosomal protein(PC00171)
RPS4X	40S ribosomal protein S4, X isoform;RPS4X;ortholog	ribosomal protein(PC00171)
RPSA	40S ribosomal protein SA;RPSA;ortholog	ribosomal protein(PC00171)
TMEM2	Transmembrane protein 2;TMEM2;ortholog	chromatin/chromatin-binding protein(PC00171)
TMSB4X	Thymosin beta-4;TMSB4X;ortholog	

The top most abundant cellular mRNAs at 24 hpi were submitted to PANTHER Gene Ontology web data base (<http://www.pantherdb.org/>) to search for their associate protein names.

Table 7. Top most abundant cellular mRNAs at 24 hpi.

Gene Symbol	Gene Name	Protein Class
ABCC2	Canalicular multispecific organic anion transporter 1;ABCC2;ortholog	ATP-binding cassette (ABC) transporter(PC00227)
AKR1B10	Aldo-keto reductase family 1 member B10;AKR1B10;ortholog	reductase(PC00176)
AKR1C1	Aldo-keto reductase family 1 member C1;AKR1C1;ortholog	reductase(PC00176)
ASPH	Aspartyl/asparaginyl beta-hydroxylase;ASPH;ortholog	hydroxylase(PC00176)
ATP2A2	Sarcoplasmic/endoplasmic reticulum calcium ATPase 2;ATP2A2;ortholog	cation transporter(PC00227);ion channel(PC00068);hydrolase(PC00133)
CA12	Carbonic anhydrase 12;CA12;ortholog	dehydratase(PC00144)
CRIM1	Cysteine-rich motor neuron 1 protein;CRIM1;ortholog	extracellular matrix glycoprotein(PC00102);cell adhesion molecule(PC00100)
DDOST	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 48 kDa subunit;DDOST;ortholog	glycosyltransferase(PC00220)
FTH1	Ferritin heavy chain;FTH1;ortholog	storage protein(PC00210)
GOLM1	Golgi membrane protein 1;GOLM1;ortholog	
HDLBP	Vigilin;HDLBP;ortholog	
HNRNPH1	Heterogeneous nuclear ribonucleoprotein H;HNRNPH1;ortholog	ribosomal protein(PC00171)
HSP90AA1	Heat shock protein HSP 90-alpha;HSP90AA1;ortholog	Hsp90 family chaperone(PC00072)
ILF2	Interleukin enhancer-binding factor 2;ILF2;ortholog	transcription cofactor(PC00218)
MGST2	Microsomal glutathione S-transferase 2;MGST2;ortholog	transferase(PC00220)
NPM1	Nucleophosmin;NPM1;ortholog	chaperone(PC00072)
PABPC1	Polyadenylate-binding protein 1;PABPC1;ortholog	
PPIA	Peptidyl-prolyl cis-trans isomerase A;PPIA;ortholog	isomerase(PC00135)
RPL12	60S ribosomal protein L12;RPL12;ortholog	ribosomal protein(PC00171)
RPL23A	60S ribosomal protein L23a;RPL23A;ortholog	ribosomal protein(PC00171)
RPL30	60S ribosomal protein L30;RPL30;ortholog	ribosomal protein(PC00171)
RPL41	60S ribosomal protein L41;RPL41;ortholog	
RPL7	60S ribosomal protein L7;RPL7;ortholog	ribosomal protein(PC00171)
RPLP0	60S acidic ribosomal protein P0;RPLP0;ortholog	ribosomal protein(PC00171)
RPS24	40S ribosomal protein S24;RPS24;ortholog	ribosomal protein(PC00171)
RTN4	Reticulon-4;RTN4;ortholog	membrane traffic protein(PC00150)
SERINC3	Serine incorporator 3;SERINC3;ortholog	transmembrane receptor regulatory/adaptor protein(PC00226)
TFPI	Tissue factor pathway inhibitor;TFPI;ortholog	serine protease inhibitor(PC00095)
TM4SF4	Transmembrane 4 L6 family member 4;TM4SF4;ortholog	
TRAM1	Translocating chain-associated membrane protein 1;TRAM1;ortholog	
TSPAN3	Tetraspanin-3;TSPAN3;ortholog	membrane-bound signaling molecule(PC00207);receptor(PC00152);cell adhesion molecule(PC00197)
UBB	Polyubiquitin-B;UBB;ortholog	ribosomal protein(PC00171)
VTI1B	Vesicle transport through interaction with t-SNAREs homolog 1B;VTI1B;ortholog	SNARE protein(PC00150)

The top most abundant cellular mRNAs at 48 hpi were submitted to PANTHER Gene Ontology web data base (<http://www.pantherdb.org/>) to search for their associate protein names

8 Publications

P. Hidalgo, L. Anzures, A. Hernández-Mendoza, A. Guerrero, C. D. Wood, **M. Valdés**, T. Dobner, and R. A. Gonzalez, "Morphological, biochemical and functional study of viral replication compartments isolated from adenovirus-infected cells.," *J. Virol.*, vol. 90, no. January, p. JVI.00033-16, 2016.

Scientific Meetings

DNA Tumor Virus Meeting (2015). Trieste, Italy. Oral presentation.

IX Congreso Nacional de Virología / IX National Virology Congress (2015). Morelos, Mexico. Poster presentation.

12th International Adenovirus Meeting (2016). Barsinghausen, Germany. Poster presentation.

Scientific Workshops

Soft skill course: Effective Scientific Writing (2012). Heinrich-Pette Institute. Hamburg, Germany.

Soft skill course: Application Training (2013). Heinrich-Pette Institute. Hamburg, Germany.

Soft skill course: Data Analysis (2014). Heinrich-Pette Institute. Hamburg, Germany.

Semester course: Introducción a la programación en R y Bioconductor / Introduction to R programming and Bioconductor (Aug-Dec 2017). Instituto de Biotecnología UNAM. Morelos, Mexico.

9 Acknowledgements

First, I would like to thank my supervisor Prof. Dr. Thomas Dobner for his mentoring, for giving me the opportunity to work in his laboratory and for allowing me to work on a project I feel passionate about.

I would also like to thank all the members of the advisory committee: Prof. Dr. Nicole Fischer and Prof. Dr. Zoya Ignatova for evaluating my thesis; and Prof. Dr. Julia Kehr for chairing the oral thesis defense.

Thanks to my lab mates Carolin Bürck, Nora Freudenberger, Jana Kondrajew, Tina Meyer, Sarah Müncheberg, Wengi Hang Ip and Remi Vaillant. They made every day special and fun to work in the lab. I will always treasure the long talks I had with each one. I also wish to thank my former lab members Sabrina Schreiner, Julia Berscheminski, Elena Lam, Shankar Kapnoor, Thomas Speiseder; and current lab members Estefania Gonzalez, Willi Ching, Marie Fiedler, Helga Hofmann-Sieber, Viktoria Kolbe, Lisa Kieweg, Michael Melling, Steewen Speck, and Nilgün Tekin-Bubenheim.

I am particularly grateful for the assistance given by Gabi Dobner and Britta Wilkens and the administrative assistance of Edda Renz who supported me during all these years.

Special thanks to Dr. Thomas Speiseder for his contribution in the annotation of the HAdV-5 transcriptome; and to Dr. Wengi Hang Ip, Dr. Ramón González, and Clayton Southerly who helped me by thoroughly reading and correcting my thesis manuscript.