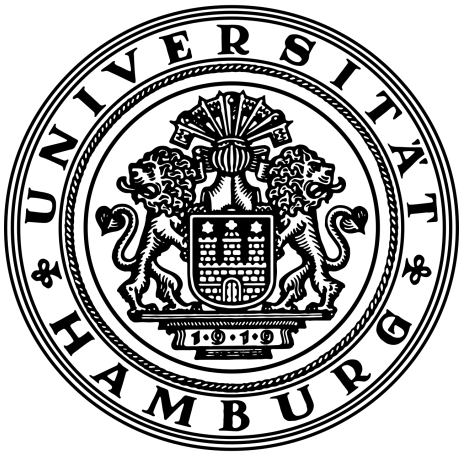
Adenovirus large E1B proteins regulate transcription through interaction with mammalian transcription factors

DISSERTATION



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

submitted by
Konstantin von Stromberg
November, 2022 in Hamburg

Disputation date: 27.01.2023

First Examiner: Prof. Thomas Dobner

Second Examiner: Prof. Nicole Fischer

Contents

Abbreviations

A | **B** | **C** | **D** | **E** | **F** | **G** | **H** | **I** | **K** | **L** | **M** | **N** | **O** | **P** | **Q** | **R** | **S** | **T** | **V** | **Y**

A

aa	amino acid
acetyl-CoA	acetyl coenzyme A
ADP	adenoviral death protein
AP-1	activator protein-1
APS	ammonium persulfate
ATL	adult T-cell leukemia-lymphoma
ATRX	alpha thalassemia/mental retardation syndrome X-linked

B

BAK	Bcl-2 homologous antagonist killer
bam	binary alignment map
BAX	Bcl-2-associated X
Bcl-2	B-cell lymphoma 2
bed	browser extensible data
BFP	blue fluorescent protein
BNIP-2	BCL2/adenovirus E1B-19K-interacting protein
BOZF-1	BTB/POZ and zinc finger domains factor on chromosome 1
BRK	baby rat kidney
bw	bigwig
bZIP	basic leucine zipper domain
BZLF1	BamHI Z fragment leftward open reading frame 1

C

C1orf43	chromosome 1 open reading frame 43
CAV	canine adenovirus
CBP	CREB-binding protein
CBX	chromobox
CDC	Centers for Disease Control and Prevention
CDKN1A	cyclin-dependent kinase inhibitor 1A
CDKN2A	cyclin-dependent kinase inhibitor 2A
ChIP	chromatin immunoprecipitation
COF	transcriptional cofactors
COMPASS	complex proteins associated with SET1
CR	conserved region
CRE	cAMP-response element
CRM1	chromosomal maintenance 1

D

DAPI	4,6-diamidino-2-phenylindole
Daxx	death-associated protein 6
dd	double distilled
DDR	DNA damage response
DSG	di(N-succinimidyl) glutarate
DTT	dithiothreitol

E

E1A	early Region 1A
E2F	E2 factor
EAdV	equine adenovirus
EBNA	Epstein-Barr virus nuclear antigen
EBV	Epstein–Barr virus
EDS-76	egg drop syndrome-1976 virus
eIF4E	eukaryotic translation initiation factor 4E
ELK-1	E26 transformation-specific - like 1
EMT	epithelial-to-mesenchymal transition
eRNA	enhancer RNA
EZH	enhancer of zeste homolog

F

FACS	fluorescence-activated cell sorting
Fas	FS-7-associated surface antigen
FBS	fetal bovine serum

G

GCN5	general control non-depressible 5
GPI	glucose-6-phosphate isomerase

H

H3K18ac	histone H3 lysine 18 acetylation
H3K4me3	histone H3 lysine 4 trimethylation
HAdV	human adenovirus
HAT	histone acetyltransferase
HBV	hepatitis B virus
HBZ	HTLV-1 bZIP factor
HCF-1	host cell factor 1
HCV	hepatitis C virus
HDAC	histone deactelyase
HFFs	normal human fibroblasts
hMSC	human mesenchymal stromal cells
HOX	homeobox
HP1	heterochromatin protein 1
HPV	human papillomavirus
HSV	herpes simplex virus
HTLV	human T-lymphotropic virus
HUSH	human silencing hub

I

ICP infected cell protein
IRF interferon regulatory factor
ITR inverted terminal repeats

K

KAP1 KRAB-associated protein-1
kbp kilobase pairs
kDa kilodalton
KDM5A lysine-specific demethylase 5A
KSHV kaposi sarcoma-associated herpesvirus

L

LANA latency associated nuclear antigen
LATS large tumor suppressor kinase
LB lysogeny broth
LeGO lentiviral ”gene ontology”

M

macs model-based analysis for ChIP-seq
MALAT1 metastasis associated lung adenocarcinoma transcript 1
MAPQ mapping quality
MCPyV Merkel cell polyomavirus
MDM2 mouse double minute 2 homolog
MLL mixed-lineage leukemia
MLP major late promoter
MLTU major late transcription unit
Mnk1 MAP kinase-interacting serine/threonine-protein kinase 1
MPP8 M-phase phosphoprotein 8
MSA multiple sequence alignments

N

NDR nucleosome depleted region
NES nuclear export signal
NGS next-generation sequencing
NLS nuclear localization signal
NPC nuclear pore complex

O

orf open reading frame

P

p14ARF p14 alternative reading frame
PAGE polyacrylamide gel electrophoresis
PCA principal component analysis
PCAF p300/CBP associated factor
PcG polycomb-group
PCR polymerase chain reaction
PEI polyethylenimine
PFA paraformaldehyde

PIC preinitiation complex
PML promyelocytic leukemia
PMSF phenylmethylsulfonyl fluoride
pRb retinoblastoma protein
PRC polycomb-repressive complex
PTM post-translational modifications

Q

qPCR real-time polymerase chain reaction

R

RAS rat sarcoma
RPGC reads per genomic content
RSF1 remodeling and spacing factor 1
RT room temperature

S

S serine
sam sequence alignment map
SAM S-adenosyl methionine
SCM SUMO conjugation motif
SDS sodium dodecyl sulfate
SET su(var)3-9/enhancer of zeste/trithorax
SET1C SET domain containing 1C
Sp specificity protein
SUMO small ubiquitin-related modifier
SUV39H suppressor of variegation 3-9 homolog
SV40 simian virus 40

T

TAD topologically associating domains
TAF TATA-box binding associated factors
TASOR transcription activation suppressor
TAZ WW domain-containing transcription regulator protein 1
TBP TATA-box binding protein
TEAD TEA domain family member
TEMED tetramethylethylenediamine
TF transcription factor
TFBS transcription factor binding site
TFIID transcription factor II D
THAP thanatos-associated proteins
TNF tumor necrosis factor
TRE TPA-responsive element
TrxG trithorax group
TSS transcriptional starting site
TTRAP TRAF and TNF receptor-associated protein

V

VGLL vestigial-like

VLPs virus-like particles
vTR viral transcriptional regulator

Y

YAP yes-associated protein

Zusammenfassung

Es wird vermutet, dass humanpathogene DNA-Tumorviren (engl. small DNA tumor viruses) die Ursache für etwa 12% bis 20% aller Krebsfälle weltweit sind. Die Grundlagenforschung in diesem Themenfeld ist jedoch komplex, da der Zeitraum zwischen der primären Infektion und dem Auftreten einer unkontrollierten Vermehrung von Tumorzellen von vielen Variablen beeinflusst wird und mehrere Jahrzehnte umfassen kann. Eines der Kennzeichen der viral-induzierten Tumorgenese ist die Inhibition der Apoptose und die exzessive Replikation von Krebszellen, die sonst durch den Tumorsuppressor p53 verhindert werden würde. Eine Manipulation dieses wichtigen regulatorischen Proteins wurde bei vielen DNA-Tumorviren beobachtet. Hierfür kodieren Adenoviren die frühe 1B Genregion, die unter anderem, das Protein E1B-55K enthält. Seine Hauptfunktion liegt in der Interaktion und Hemmung des oben erwähnten Tumorsuppressorproteins p53, wofür es mehrere synergistische Strategien entwickelt hat. Die derzeit am wenigsten untersuchte Wirkungsweise ist seine postulierte Funktion als transkriptioneller Repressor von p53-abhängigen Genen. In der vorliegenden Arbeit konnte ich durch verschiedene ChIP-seq-Experimente in primären BRK Zellen eine weitreichende und indirekte Bindung von E1B-55K über das gesamte Wirtsgenom *in vivo* nachweisen. Die Folgen dieser Interaktion wurden analysiert, indem die mit der Bindung assoziierten Gene identifiziert und ihre Expressionsänderung mittels mRNA-seq ausgewertet wurde. Diese Experimente liefern konkrete Hinweise dafür, dass E1B-55K als viraler Transkriptionsrepressor fungieren kann, indem es mit DNA-gebundenen Wirtstranskriptionsfaktoren interagiert und deren Aktivität hemmt. Diese Interaktion und Deregulation begünstigt ein pro-virales Milieu. Diese Aktivität ist unabhängig von dem SUMOylierungsstatus von E1B-55K, sofern das Protein im Zellkern lokalisieren kann, um mit der Transkriptionsmaschinerie der Wirtszelle zu interagieren. Die Analyse der assoziierten biologischen Signalwege ergab einen signifikanten Einfluss auf die Apoptose, die Wachstumshemmung und die Stressreaktion der Wirtszelle. Dieser repressive Mechanismus ist auch bei den großen E1B-Proteinen von anderen adenoviralen Spezies konserviert und korreliert mit deren Onkogenität im Tiermodell. Die Auswirkungen der Expression von E1B-55K auf die epigenetische Landschaft (engl. epigenetic lands-

cape) von primären humanen mesenchymalen Stromazellen lieferten Erkentnisse über den möglichen Mechanismus der oben beschriebenen transkriptionellen Inhibition über die Hemmung von aktivierenden Histonmodifikationen.

Abstract

Small human DNA tumor viruses are thought to be the etiological agent for approximately 12% to 20% cancer incidents around the world. The study of this development is quite arduous, as the time frame between primary infection and the emergence of unregulated proliferation of cancer cells is influenced by countless variables and can span multiple decades. One of the hallmarks of human viral tumorigenesis is the evasion of apoptosis and uncontrolled cellular replication, counteracted in part by the tumor suppressor p53 in healthy cells. Circumvention of this important regulatory protein has been observed in many DNA tumor viruses. In order to accomplish this, adenoviruses encode the early region 1B gene region, which contains the E1B-55K protein. It's main function upon expression lies in the interaction and inhibition of the aforementioned tumor suppressor protein, for which it has evolved multiple synergistic strategies. The currently least understood course of action is its postulated function as a transcriptional repressor of p53-responsive genes, initially hypothesized two decades ago. In this work, I could show an expansive and indirect *in vivo* occupancy across the host genome of primary baby rat kidney cells through various ChIP-seq experiments. The consequences of this interaction were analyzed by identifying the binding-associated genes and evaluating their respective mRNA expression level via mRNA-seq. These experiments provide tangible evidence that E1B-55K can act as a viral transcriptional repressor by interacting with DNA-bound host transcription factors beyond p53, inhibiting their activity, thereby modulating various downstream responses conducive to a pro-viral environment. I found that this activity is independent from its SUMOylation status, as long as the protein has a means to localize in the nucleus to interact with the host cellular transcriptional machinery. Multiple biological pathway analyses consequently revealed interference with host cell apoptosis, growth inhibition and stress response pathways, among others. This repressive mechanism is highly conserved among large E1B proteins from different adenoviral species and correlates with the observed oncogenicity of the respective viral species. Analysis of the epigenetic landscape of primary human mesenchymal stromal cells that stably expressed E1B-55K provided a potential mechanism of this transcriptional abrogation, as establishment of activating histone PTMs was found to be inhibited at several genomic loci.

Chapter 1

Introduction

1.1 Adenoviruses

1.1.1 Principles of *Adenoviridae*

The *Adenoviridae* family of pathogenic viruses was first discovered in 1953 by ROWE *et al.* (1953) as a cytopathogenic agent and subsequently described as a potential viral pathogen in 1956 by ENDERS *et al.* (1956). Up to the present day this family of icosahedral non-enveloped DNA-Viruses has grown to encompass over 6 genera with 87 species and has been found to infect a wide variety of vertebral hosts (WALKER *et al.*, 2020). The adenoviral genera are separated by modern high-resolution molecular phylogenetic analysis and named accordingly to their respective preferred host range. They consist of, in no specific order, the *Atadenovirus* genus infecting reptilian and ruminant hosts, the *Siadenovirus* genus infecting amphibian hosts, the *Aviadenovirus* genus infecting avian hosts, the *Ichtadenovirus* genus infecting fish, the *Testadenovirus* which infect testudines and the *Mastadenovirus* genus infecting mammalian hosts (Figure 1). The latter contains families that can be found in several different mammal hosts and includes the family of human adenovirus (HAdV). The HAdVs are subgrouped into seven different species (named A-G) containing 113 types, and were historically divided by their serum neutralization and hemagglutination capabilities (WADELL, 1984; EBNER *et al.*, 2005), a demarcation criterion which is currently being replaced by genomic criteria since the discovery of HAdV-52 (JONES *et al.*, 2007). Interestingly, ape mastadenoviruses are classified into species that belong to the human isolates due to their high degree of phylogenetic relatedness and are thought to have originated in apes and switched to humans during hominine evolution (ROY *et al.*, 2009; HOPPE *et al.*, 2015). While some human Adenoviruses are able to permissively infect rodent and ruminant cells, others display varying degrees of oncogenicity in newborn hamsters, ranging from low in the species B to high in the species A (TRENTIN

et al., 1962), which was the first piece of evidence that viruses can be the causative agent of cancer and led to the classification of Adenovirus as a DNA tumor virus.

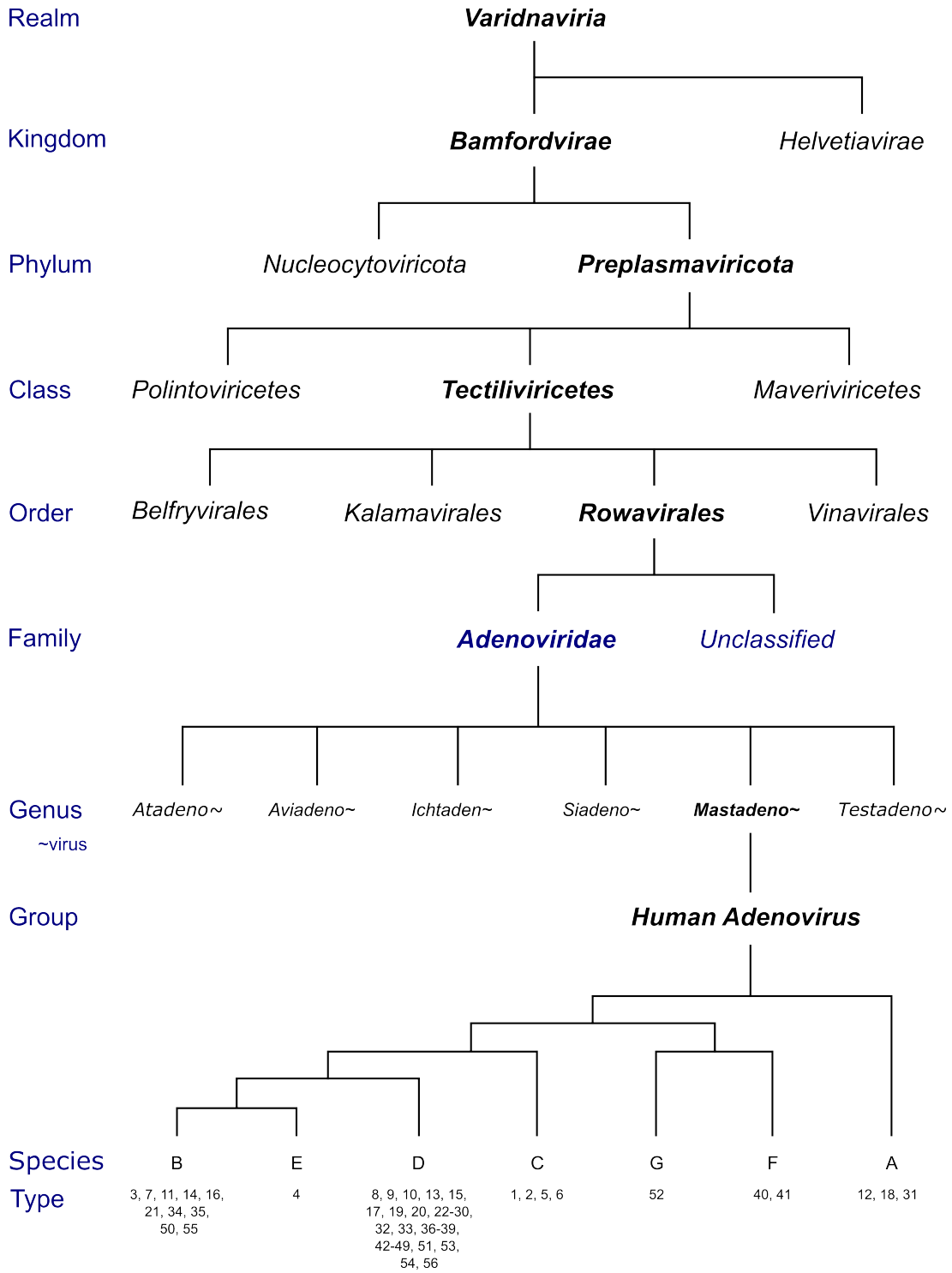


Figure 1: Classification of *Adenoviridae*. Hierarchical representation of the *Adenoviridae* family in the realm of *Varidnaviria* according to the International Committee of the Taxonomy of Viruses (ICTV) classification from the rank of family to type, focusing on human adenoviruses, accessed in June, 2022.

However, later studies revealed that HAdV DNA was not detectable in the investigated primary human gastrointestinal or lung tumors (MACKEY *et al.*, 1976). As yet, no causal relationship between Adenovirus infection and human carcinogenesis could be established. Undeterred by these findings, the study of Adenoviruses has proven to be vital in the field of molecular biology, leading to several groundbreaking discoveries through viral interactions and disturbances within its various host systems. From the discovery of RNA splicing (CHOW *et al.*, 1977; BERGET *et al.*, 1977) to cell cycle regulation (YANG *et al.*, 1996) and cellular transformation (CHINNADURAI, 1983; DEBBAS and WHITE, 1993), these works have undoubtedly contributed to advancing the knowledge on and deepening the understanding of basic eukaryotic cellular mechanisms. Today, Adenoviruses are widely used as the vector backbone of vaccine systems due to their ease of production, low risk of insertional mutagenesis and stable transient gene expression (HARVEY *et al.*, 2002; TATSIS and ERTL, 2004; EWER *et al.*, 2016). They are used extensively as a model system owing to their capability to reliably infect a broad range of different cell types and are applied in oncolytic cancer treatment (PESONEN *et al.*, 2011) and gene transfer systems (KAPLAN, 2005; SU, 2011).

1.1.2 Virion structure and genomic organization

Adenovirus is a non-enveloped and icosahedral eukaryotic virus, about 950 Å in diameter that contains a single linear double-stranded DNA (dsDNA) molecule which varies in size between species, ranging from 26 kilobase pairs (kbp) to 48 kbp (DAVISON *et al.*, 2000; DOSZPOLY *et al.*, 2019) (Figure 2). The viral particle is made up of minor and major capsid proteins, core proteins, a terminal protein, the packaging machinery and a maturation protease (reviewed in GALLARDO *et al.* (2021)). The icosahedron is created by facet arrangements of 720 individual units of the major coat hexon protein arranged as 240 trimers with 12 individual capsomers, consisting of a penton fiber and base, protruding from the vertices. The over 900 amino acid (aa) large highly conserved hexon protein accounts for 60% of the total mass of the viral particle. It contains hypervariable regions that create serotype-specific differences between the species (CRAWFORD-MIKSZA and SCHNURR, 1996) and is the protein responsible for the pseudo-T = 25 triangulation number, due to its complex oligomeric arrangement by four independent hexons placed in four different environments per facet (BURNETT *et al.*, 1985; BURNETT, 1985; VAN OOSTRUM *et al.*, 1987). The icosahedral facet structure of the viral particle leaves a gap in the center of five peripentonal hexons, which is filled with the pentamers of the penton polyprotein

III. This protein is highly conserved across all human adenoviral species and contains a hypervariable loop region (ZUBIETA *et al.*, 2005). Protruding fiber proteins binds to the outer surface of the penton pentamer in a non-covalent way and are the attachment protein of the virus. These trimeric proteins consist of a head-, a shaft-, and a tail region. The head region, which forms a globular knob, is responsible for binding to the host cell surface proteins as a first step to gain entry into the cell. The structure of the fiber protein varies widely between species, thereby creating enough diversity to explain the broad tropism of Adenovirus. For example, due to the small size of ß-loops, Atadenovirus has the smallest fiber while the fiber head of Siadenovirus is more similar to the fiber of Reovirus and is able to bind to its neighboring monomer (SINGH *et al.*, 2014, 2015). The flexibility and length of the fiber protein is important for viral entry into the cell, by influencing both primary interaction with the receptor and secondary interactions with the cell surface integrins $\alpha v\beta 1$, $\alpha v\beta 3$ and $\alpha v\beta 5$ (WICKHAM *et al.*, 1993; STEWART and NEMEROW, 2007). HAdV species F40 and F41 are particularly fascinating, as they have two fiber genes, differing in length, and with only one of them attached to a vertex at random (KIDD *et al.*, 1993). Interestingly, some HAdV are able to form dodecahedron-shaped virus-like particles (VLPs) with penton base and fiber proteins alone, without any hexon present. These VLPs are applied in gene and vaccine delivery and receptor identification (NORRBY and WADELL (1967), reviewed in BESSON *et al.* (2020)). Minor coat proteins are predominantly located on the inner surface of the viral particle. Five copies of the structurally diverse protein IIIa are underlying beneath the penton vertex. This protein is responsible for connecting the peripentonal hexons to the penton base, acting as a kind of glue for the viral particle (KUNDHAVAI NATCHIAR *et al.*, 2018; PÉREZ-ILLANA *et al.*, 2021; MARABINI *et al.*, 2021). Protein VIII acts as a further stabilizing factor between the interactions of protein IIIa and the peripentonal hexons (LIU *et al.*, 2010; MARABINI *et al.*, 2021). The only minor coat protein present on the outside is protein IX, often described as ”cement” of the viral particle and is unique to Mastadenovirus. Although being dispensable for assembly, the protein alone is able to confer capsid thermostability (VELLINGA *et al.*, 2005). The atadenoviral specific LH3 protein plays an identical role in the capsid of the virus (PANTELIC *et al.*, 2008) but displays high sequence similarity towards both Mastadenovirus protein IX and the E1B-55K protein, located directly upwards of the gene coding for IX. Recent combinations of structural and sequence analyses hint towards an evolutionary explanation of this conundrum. It is thought that an ancestor of Atadenovirus and Mastadenovirus has acquired an LH3-like gene from a bacterium or bacteriophage, a

gene that has subsequently undergone a duplication event in Mastadenovirus with one copy retaining the structural role of IX while the other drastically changed its C-terminus and lost its ability to bind the capsid, giving rise to the E1B-55K protein (MARABINI *et al.*, 2021).

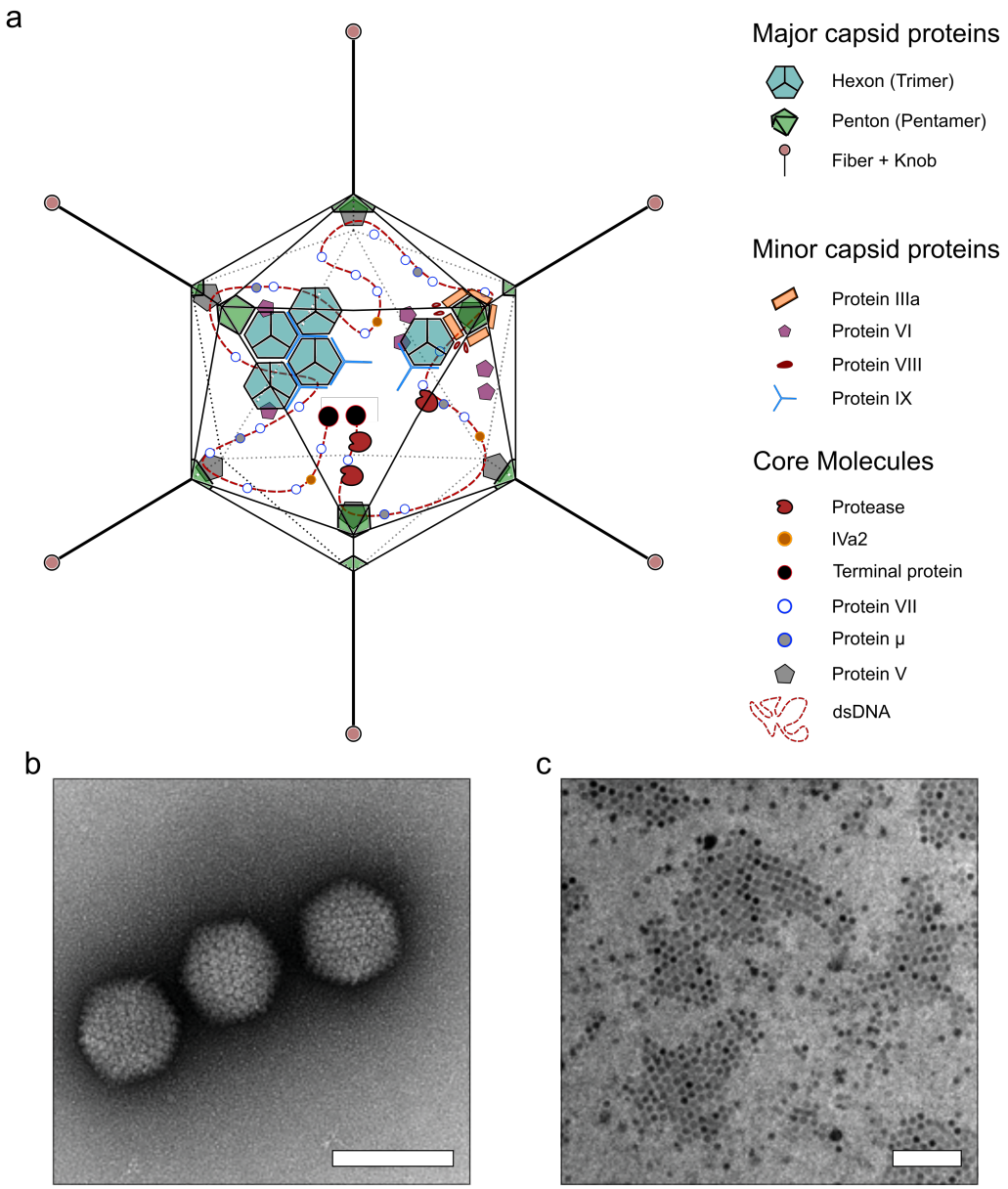


Figure 2: Schematic illustration and TEM micrographs of human Mastadenovirus. **a** Virus organization of a HAdV-C5 particle with annotated major-, minor- and core molecules. **b** Negative stain transmission electron micrographs of purified virus particles. The scale bar indicates 100 nm. **c** HAdV-C5 particles in infected A549 cells. The scale bar indicates 5 μm. Electron micrographs were adapted from PFITZNER *et al.* (2020).

The basic core proteins V, VII and μ of the virus are thought to act as condensing agents of the DNA (PÉREZ-BERNÁ *et al.*, 2015; MARION *et al.*, 2017), although structural data is still missing yet. Protein VII is highly associated with viral DNA, creating a ”bead-on-a-

string" structure reminiscent of cell chromatin (VAYDA *et al.*, 1983). While the protein does not seem to be required to form viral particles, it is indispensable for viral infection as particles lacking the protein are not able to expose protein VI to the endosome and are therefore unable to exit this stage of infection (OSTAPCHUK *et al.*, 2017; HERNANDO-PÉREZ *et al.*, 2020). Polyprotein VI also acts as a cofactor for the protease (MANGEL and SAN MARTÍN, 2014; HERNANDO-PÉREZ *et al.*, 2020).

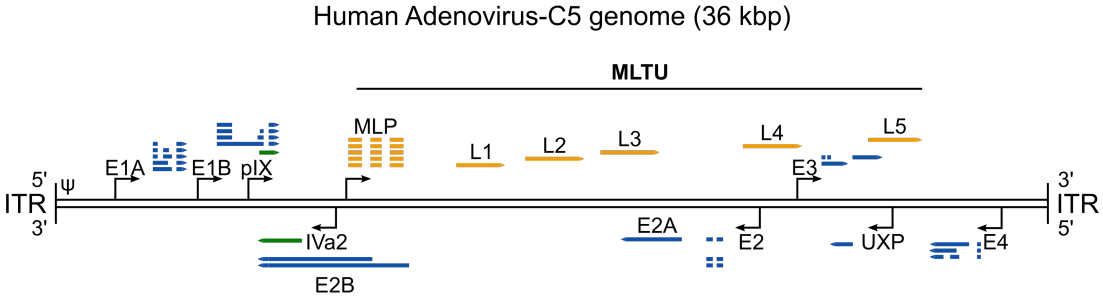


Figure 3: Overview of the HAdV-C5 genome. Visualization of the double stranded linear genome organization of HAdV-C5, showing the most important canonical transcripts of the virus. Genes are colored according to their expression in the early (blue), intermediate (green) and late (yellow) phase of infection. ITR: Inverted terminal repeat, E: Early, L: Late, MLP: Major late promoter, MLTU: Major late transcription unit, ψ: Packaging sequence, UXP: U exon protein

All adenoviral genomes across various species encode the major early and late proteins required to initiate viral DNA replication and assembly of viral particles. A visualization of the ∼36 kbp HAdV-C5 genome which encodes for around 40 proteins is displayed in Figure 3, showing all the major early (E1A, E1B, E2, E3, E4), intermediate (IVa2, pIX) and late (L1-5, U exon protein UXP) transcription units of the virus. Inverted terminal repeats (ITR) can be found at both ends of the linear and double-stranded DNA genome. Transcription is mediated by host RNA polymerase II, transcribing from both strands of the genome. All viral RNAs are capped and polyadenylated during processing. Many viruses, due to their compact genome structure, are making extensive use of overlapping reading frames and RNA-splicing with adenoviruses having an especially complex transcription and splicing system. Novel studies have shown that the predicted adenoviral transcriptome is more diverse and abundant than expected, with many different mRNAs effectively coding for the same protein. Its thought that this can be used to further fine-tune protein expression and allows for more flexibility in RNA processing, yielding a potential evolutionary fitness benefit (DONOVAN-BANFIELD *et al.*, 2020).

1.1.3 Reproductive life cycle

Adenoviruses are obligate intracellular pathogens and have to extensively modulate the infected host cell and co-opt its processes to produce viral progeny at the end of a replication cycle. The adenoviral life cycle is distinguished by two major phases which are separated by the onset of viral DNA replication. In short, the early phase starts with surface binding followed by clathrin-mediated endocytosis and particle disassembly and ends with the expression of the early viral regulatory proteins, while the late phase consists of expression of viral structural proteins, intranuclear assembly of the virion and cellular egress. Various cell types are permissive for adenoviral infection but the preferred tissue tropism varies between species. Differences in targeting cells can depend on the surface protein composition of the virion, with most adenoviral species interacting with the coxsackievirus and adenovirus receptor CAR, a cellular surface protein that is associated with tight junctions and assists in a barrier mechanism against paracellular flow of solutes and macromolecules (BERGELSON *et al.*, 1997; ROELVINK *et al.*, 1998; COYNE and BERGELSON, 2005). This Adenovirus-CAR interaction is attributed to the viral surface fiber protein and studies have shown that a mutation in this fiber can confer a change in cellular specificity of the virus (HUANG *et al.*, 1999; ISMAIL *et al.*, 2016). Members of species B are unique in their usage of CD46 as an entry receptor (GAGGAR *et al.*, 2003). Initial contact is established between the viral particle and the CAR receptor, which is directly followed by integrin clustering and binding to a single penton base pentamer (LI *et al.*, 1998; VEESLER *et al.*, 2014; NESTIĆ *et al.*, 2019, 2021). This process leads to the internalization of the viral particle into a vesicle that will fuse with an endosome shortly afterwards. Acidification of the cellular endosome to pH 6 will expose the endosome lysis protein VI, releasing the partially disassembled viral particle into the cytoplasm. This entry event may trigger macropinocytosis, a pathway of receptor-independent endocytosis and fluid-phase uptake that plays an important role in adenoviral exit from the endosome into the cytosol (MEIER *et al.*, 2002). The virus initiates its transport to the nucleus through interaction of the viral capsid hexon and the cytoplasmic dynein by a mechanism that is distinct from that of physiological dynein cargo (BREMNER *et al.*, 2009). The viral particle is transported bidirectionally along microtubules to the nucleus, where interaction with the nuclear export factor chromosome region maintenance 1 (CRM1) releases the particles from the microtubules to the nuclear pore complex (NPC) and promotes both capsid disassembly and envelope translocation (STRUNZE *et al.*, 2005). As soon as the viral DNA enters the nucleus, the cellular transcription machinery initiates transcription of the immediate early

Region 1A (E1A). E1A proteins interact with a variety of cellular regulators to reprogram gene expression in the infected cell and induce transcription of the other early gene E1B, E2, E3 and E4 region products (MORAN *et al.*, 1986; AVVAKUMOV *et al.*, 2002). The goal of the early transcription unit products is to alter host cell processes to be advantageous for efficient viral replication. Furthermore, they are involved in cellular transformation (reviewed in B.GREENBERG (2013)).

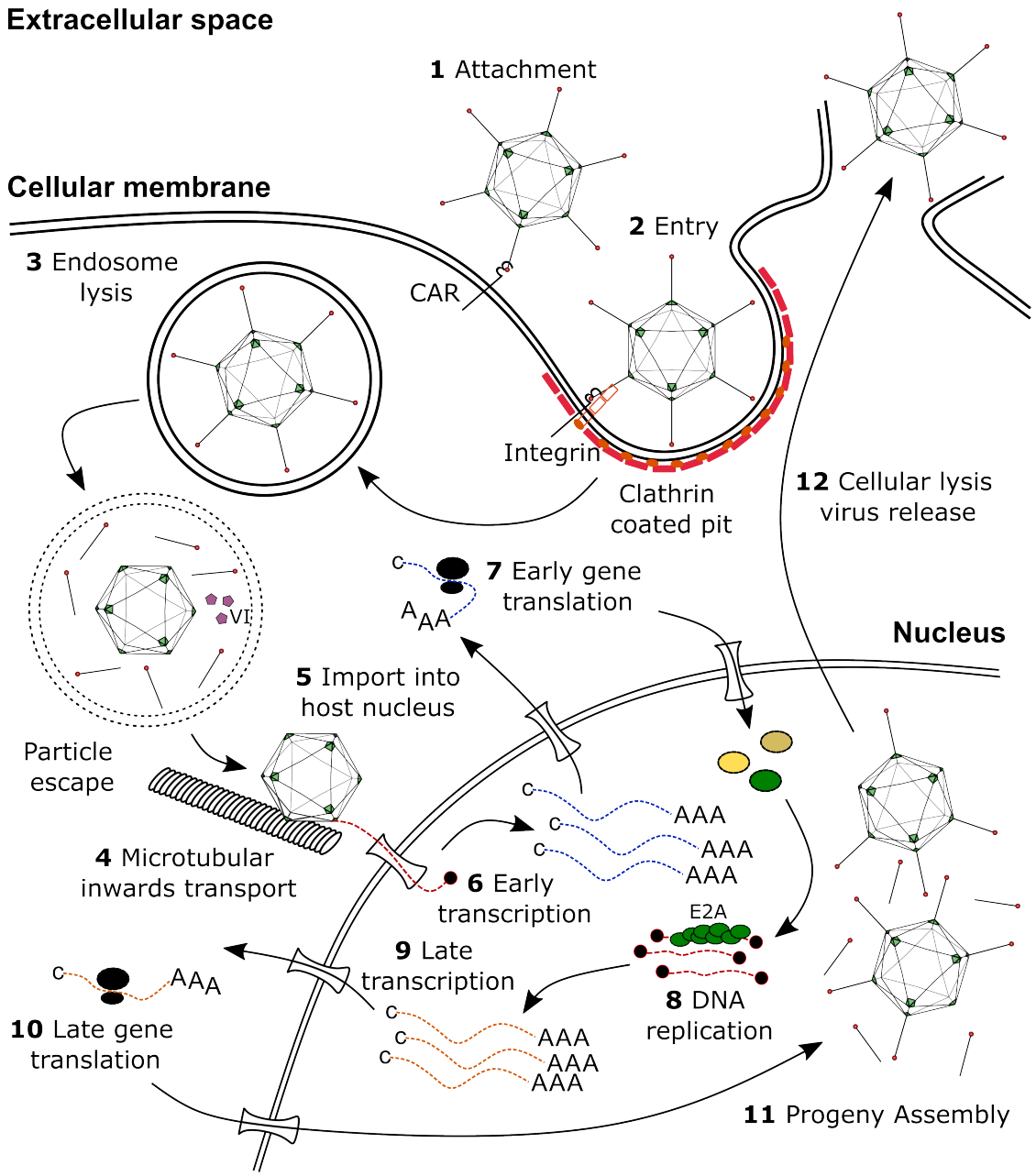


Figure 4: Overview of the HAdV-C5 life cycle. Simplified illustration of the general HAdV life cycle in an infected cell. (1) Viral attachment and (2) entry are mediated by interactions with the CAR receptor and integrins. After (3) protein VI-mediated endosomal lysis, (4) transport of the particle to the nucleus and (5) import of the viral DNA into the nucleus, the (6+7) early gene expression is initiated. (8) DNA replication of the viral genome is then followed by (9+10) late gene expression and (11) progeny assembly and the cycle is completed with (12) cellular lysis and viral escape.

Proteomic studies have revealed extensive remodelling of the host cell proteome during the early phase of infection, especially through E4 region proteins-mediated degradation of inhibitive cellular factors that would otherwise interact with viral DNA (Dybas *et al.*, 2021). At these early stages of infection, promyelocytic leukemia (PML) proteins are found adjacent to the viral DNA as part of an antiviral defense mechanism (Ishov and

MAUL, 1996). The emergence of numerous nuclear microdomains that function as hub areas and early viral replication sites starts after expression of the early viral proteins (reviewed in HIDALGO *et al.* (2019)). Different viral proteins localize into these structures: The DNA-binding protein E2A, the DNA polymerase Pol, the terminal protein pTP and other late viral components like the E1B-55K, E4orf6, L4-22K and -33K proteins. Additionally, numerous cellular proteins are known to concentrate into these viral structures as well, like the ubiquitin specific protease 7 (USP7) (CHING *et al.*, 2013), the speckled 100 kDa (Sp100) (DOUCAS *et al.*, 1996) or the CCCTC-binding factor (CTCF) (KOMATSU *et al.*, 2013). The late phase of infection starts with onset of DNA replication by the E2 region proteins of the virus and after formation of replication centers in the nucleus (FLINT and GONZALEZ, 2003). Activation of the major late promoter (MLP) leads to expression of the different structural, core and capsid proteins encoded in the major late transcription unit (MLTU), named L1 to L5, accordingly (SHAW and ZIFF, 1980). The late phase of viral infection additionally leads to a host translation shutoff by blocking the interaction between the eukaryotic translation initiation factor 4E (eIF4E) and the MAP kinase-interacting serine/threonine-protein kinase 1 (Mnk1), while extensive use of ribosome shunting on viral late mRNA allows for continuous expression of viral RNA (CUESTA *et al.*, 2004). The replication cycle ends after 24-36 h with assembly and lytic release of up to 10^4 viral particles per cell through action of the adenoviral death protein (ADP) (TOLLEFSON *et al.*, 1996a,b). Novel studies have shown that the mechanism of viral nuclear egress is more complex than previously assumed, as release of viral particles can be observed even in the absence of ADP, presumably through permeability of the nuclear membrane and multiple nuclear membrane destabilization events during the course of virus accumulation in the nucleus (PFITZNER *et al.*, 2021). An exemplary and simplified HAdV-C5 life cycle is illustrated in Figure 4.

1.1.4 Pathogenesis

Adenoviruses are common pathogens in vertebrates, from fish over birds to humans. Several members of distinct genera are very capable of causing potentially fatal disease. Canine adenovirus (CAV)-type 1 and CAV-type 2 can cause infectious canine hepatitis or enteric and respiratory disease in dogs (PRIER, 1962; WRIGHT *et al.*, 1972), while aviadenovirus egg drop syndrome-1976 virus (EDS-76) can cause an economically critical disease resulting in falls in egg production (MCFERRAN *et al.*, 1978). Equine adenovirus (EAdV)-type 1 and EAdV-type 2 predominantly lead to upper respiratory tract infection

or gastrointestinal tract infection in horses, respectively (CAVANAGH *et al.*, 2012). As already mentioned above, ape adenoviruses are human adenoviruses, establishing the inherent underlying zoonotic potential of adenovirus, with novel species continuously being identified, many of which readily cross interspecies host barriers (reviewed in BORKENHAGEN *et al.* (2019) and HARRACH *et al.* (2019)). Through the impact of several different pandemics caused by influenza viruses and with the emergence of SARS-CoV-2 at the beginning of 2020, we experienced firsthand how close human-animal contact can lead to recurrent zoonosis and anthroponosis. Adenoviruses feature distinct attributes that make it a definitive risk in terms of pandemic potential, from efficient intra- and interspecies transmission, persistent or latent infection and their ability to steadily recombine to escape herd immunity (reviewed in KREMER (2021)). According to the Centers for Disease Control and Prevention (CDC), HAdVs can emerge unpredictably and can cause contained short-term outbreaks. While infection in immunocompetent hosts usually causes conjunctivitis, respiratory illness or diarrhea with symptoms ranging from common cold to pneumonia and bronchitis, disease manifestation heavily depends on the type of virus and immune status of the infected person. For example, humans exposed to HAdV species B or E (types 3, 4, 7 and 14) are usually experiencing acute respiratory illness, species D (types 8, 19, 37, 53 and 54) can cause epidemic keratoconjunctivitis while species F (types 40 and 41) cause gastroenteritis (JAMES *et al.*, 2007; BIGGS *et al.*, 2018; KILLERBY *et al.*, 2019; LEE *et al.*, 2020; SHIEH, 2022; CENTERS FOR DISEASE CONTROL AND PREVENTION, 2022). For the immunocompromised host, opportunistic adenoviral infection or reactivation can pose a serious risk for severe health complications. Reactivation can occur in hosts that harbor persistent HAdV presence in different tissues, including tonsillar and adenoidal T lymphocytes or T cells isolated from the colon (GARNETT *et al.*, 2002, 2009; ROY *et al.*, 2011; KOSULIN *et al.*, 2007) and is associated with high morbidity and mortality in paediatric haematopoietic stem cell transplant recipients (LION *et al.*, 2010; FEGHOUL *et al.*, 2015). It is thought that intestinal lymphocytes represent a potential reservoir for persistence, whereas the intestinal epithelium is the main place of viral replication prior to viral spread in an immunocompromised host (KOSULIN *et al.*, 2016). Diagnosis of an adenovirus infection is mostly confirmed via real-time polymerase chain reaction (qPCR), but has an inherent need to cover the entire known HAdV spectrum to attribute for all potential different species and types (reviewed in LION (2014)). Anti-adenoviral treatment options are contingent on the individual risk situation of the host. The cytidine nucleotide analogue Cidofovir is the most used general antiviral therapy, but no specific drug against

adenovirus has been approved yet (SAFRIN *et al.*, 1997; MATTHES-MARTIN *et al.*, 2012). A recent review by HIWARKAR *et al.* (2018) has recommended to initiate anti-adenoviral treatment in patients with rapidly rising stool HAdV levels, as intestinal reactivation and expansion is extremely common prior to systemic adenoviral infection.

1.1.5 Viral oncogenes and cellular transformation

Today, we know of seven different viruses that can cause aberrant and uncontrolled cell growth during their infectious cycle (reviewed in KRUMP and YOU (2018)). These viruses are, in no specific order: Epstein–Barr virus (EBV), human T-lymphotropic virus (HTLV)-1, kaposi sarcoma-associated herpesvirus (KSHV), Merkel cell polyomavirus (MCPyV), hepatitis B virus (HBV), hepatitis C virus (HCV) and human papillomavirus (HPV). Although harboring various differences regarding their genome structure, cellular tropism, disease prevalence and cancer pathologies, they share features and hallmarks that can eventually lead to cancer growth in humans. One of these similarities is the targeting of various tumor suppressor pathways, thereby abrogating cellular apoptosis, senescence or cell cycle arrest. Two of the key regulators are the retinoblastoma protein (pRb) and p53, both of which are dysregulated, degraded or inhibited by almost all oncogenic viruses through action of their oncoproteins (reviewed in HOWLEY and LIVINGSTON (2009); LEVINE (2009)). Although being classified as a human DNA tumor virus, adenoviruses are not causally implicated in human pathogenic oncogenesis. Yet, they encode a set of regulatory proteins, namely the E1 and E4 region proteins, that are heavily involved in deregulating cellular pathways to the benefit of the virus (IP and DOBNER, 2020). Continuous expression of these proteins can lead to cellular transformation in different rodent or human primary cell culture systems (GRAHAM *et al.*, 1977; VAN DEN HEUVEL *et al.*, 1992; SCHIEDNER *et al.*, 2000; SPEISEDER *et al.*, 2017). This process requires accidental integration of either E1 or E4 regions into the host genome during abortive viral infection in rodents, while it has not yet been shown to occur in human cells. Control over the aforementioned pathways is exerted by the proteins contained in the E1 region, namely the E1A, E1B-19K and E1B-55K proteins, which are more extensively described in detail in the next sections.

Early region 1A protein

The adenoviral E1A protein is well researched and understood regarding its indispensable ability to convert the host cell into an amenable state for viral replication (JONES and SHENK, 1979). As the first adenoviral protein to be expressed during infection with

its main functions including activation of the viral gene transcription, induction of the cell cycle in otherwise quiescent cells and suppression of the inflammatory response induced by viral infection (NEVINS *et al.*, 1979; WHYTE *et al.*, 1988; SCHAACK *et al.*, 2004). It is thought to be largely disordered (reviewed in PELKA *et al.* (2008)) like its HPV E7 counterpart (OHLENSCHLÄGER *et al.*, 2006) and retains function even after denaturation through boiling (KRIPPL *et al.*, 1984). The recently published 3D-structure prediction tool Alphafold2 can be used to illustrate the disordered and unstructured conformation of E1A, shown in Figure 5**a** (JUMPER *et al.*, 2021; MIRDITA *et al.*, 2019, 2022). Splicing of HAdV-C5 E1A-mRNA leads to the expression of five distinct proteins: two large early isoforms named 13S/12S of 289 aa/243 aa, respectively, and three small late and less understood isoforms named 11S, 10S and 9S of 217 aa/171 aa/55 aa, respectively (PERRICAUDET *et al.*, 1979; ROWE *et al.*, 1983). These proteins are highly conserved across the different HAdV (AVVAKUMOV *et al.*, 2004) (Figure 5**b**). The largest isoform contains four different conserved region (CR)s, termed CR1-CR4 (Figure 5**c**), the first two of which bare striking similarity with the HPV E7 and the simian virus 40 (SV40) large-T antigens (PHELPS *et al.*, 1988; FIGGE *et al.*, 1988). The 13S isoform is predicted to be phosphorylated at three serine (S) sites: S89, S219 and S231 (TREMBLAY *et al.*, 1988) (Figure 5**c**). Earlier studies on E1A and viral replication in general have led to vital understanding of cellular replication pathways and identification of the involved key host factors, like the E2 factor (E2F) (CHALLBERG and KELLY, 1979; KOVESDI *et al.*, 1986). By itself, the E1A protein is able to immortalize primary rodent cells but requires cooperation with another oncogene, like HAdV E1B or activated rat sarcoma (RAS) protein to fully transform the host cell (LOWE *et al.* (1994) and reviewed in FRISCH and MYMRYK (2002)). Albeit containing a zinc finger motif in its CR3, E1A does not bind to DNA directly (AVVAKUMOV *et al.*, 2002), it rather functions through its vast array of interaction partners. These factors include the pRb, which is the first described example of interaction between a viral oncogene and a cellular tumor suppressor (WHYTE *et al.*, 1988). This interplay between E1A and pRb supposedly frees the transcription factor E2F and therefore allows the cell to enter the S-phase cell cycle by activating expression of different cyclins, cyclin-dependent kinases and other replication proteins (FATTAEY *et al.* (1993); LIU and MARMORSTEIN (2007) and reviewed in LONGWORTH and DYSON (2010)). Interestingly, this phenotype mirrors the functional inactivation observed in non-virally induced cancers and its resulting continuous proliferation signaling and evasion of growth suppression (KNUDSEN *et al.*, 2020). The lysine acetyltransferase p300 protein is another target of E1A, an interaction that readily

leads to cellular transformation through complex mechanisms that are to be fully elucidated (HOWE *et al.*, 1990; ECKNER *et al.*, 1994; TURNELL and MYMRYK, 2006). Through these interactions, E1A also exerts tremendous influence on the cellular epigenome by relocalizing histone H3 lysine 18 acetylation (H3K18ac) to a set of genes, activating their expression, which will lead to stimulation of cell cycling, inhibition of antiviral responses and cellular differentiation (FERRARI *et al.*, 2008; HORWITZ *et al.*, 2008; FERRARI *et al.*, 2014).

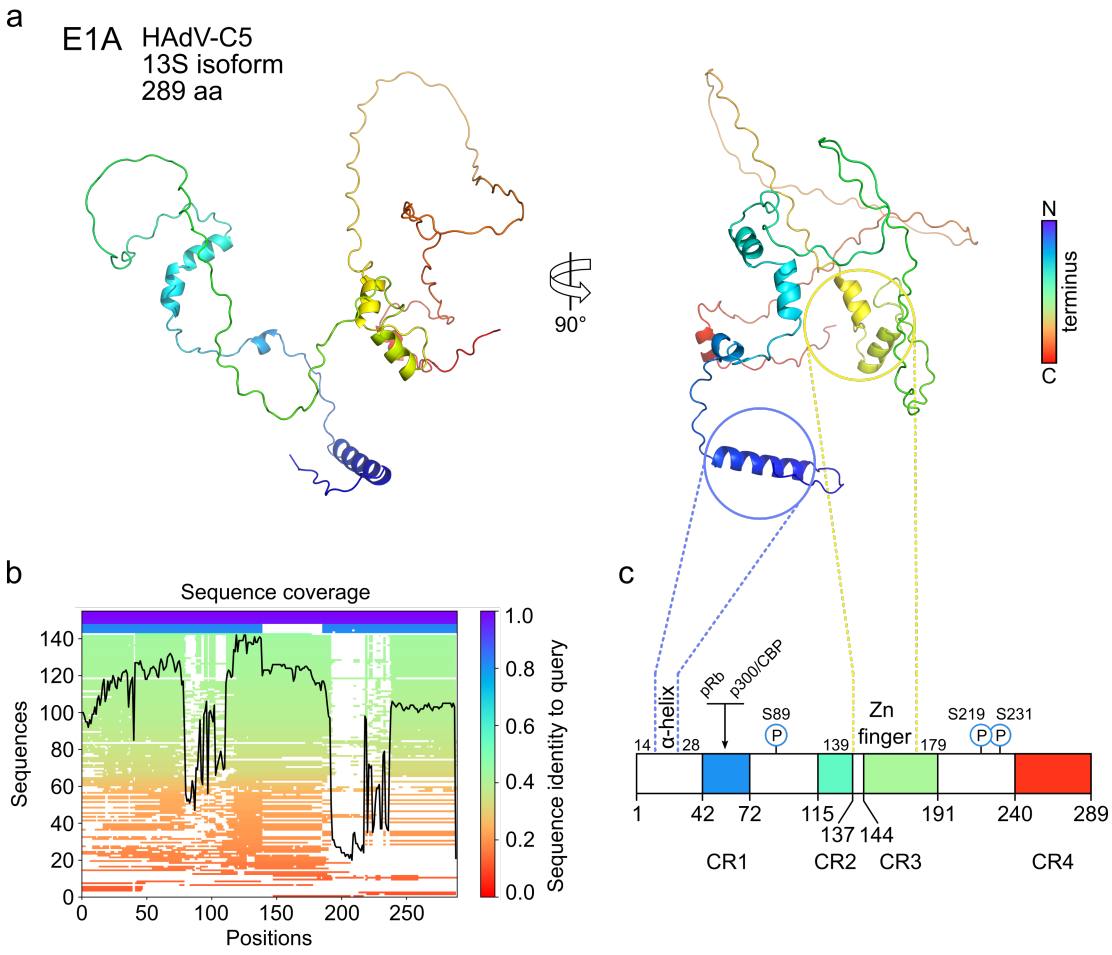


Figure 5: Alphafold2 prediction of the HAdV-C5 E1A structure. **a** Rendition of HAdV-C5 E1A 13S protein folding prediction by Alphafold2, colored from N-terminus (blue) to C-terminus (red) (JUMPER *et al.*, 2021; MIRDITA *et al.*, 2022), visualized by PyMOL (SCHRÖDINGER and DELANO, 2021). **b** Prediction quality sequence coverage plot visualizing the multiple sequence alignments (MSA) depth generated by MMseqs2, which searches the query amino acid sequence against the consensus sequences of the UniRef30 database (SUZEK *et al.*, 2015; STEINEGGER and SÖDING, 2017; MIRDITA *et al.*, 2019). **c** Overview of the full HAdV-C5 E1A 13S sequence structure with relevant domains and modification sites. CR: Conserved region, P: Phosphorylation site, S: Serine, Zn: Zinc.

E1A acts on many more cellular factors, like p300/CBP associated factor (PCAF), p400, general control non-depressible 5 (GCN5) and the TRAF and TNF receptor-associated protein (TTRAP) protein, among others (REID *et al.*, 1998; DELEU *et al.*, 2001; RASTI *et al.*,

2005; ABLACK *et al.*, 2012; ZHAO *et al.*, 2017). In conclusion, adenoviral E1A is able to compromise many key eukaryotic pathways and manipulates various cellular host systems to its advantage. It has a vast primary and secondary interaction network, intruding and shaping the host as a result of a continuous evolutionary arms race (extensively reviewed in KING *et al.* (2018) and TESSIER *et al.* (2021)).

Early region 1B proteins

These previously introduced actions of E1A will lead to preemptive p53-dependent apoptosis, a circumstance that HAdV avoids through actions of the E1B-region. This locus codes for five different gene products, with the prominant 176 aa E1B-19K and the 496 aa E1B-55K being the most closely examined. The three other E1B polypeptides, sorted by size, are termed 156R, 93R and 84R and are expressed later in the viral life cycle through differential splicing of the E1B-55K transcript (SPECTOR *et al.*, 1978; WILSON and DARNELL, 1981) and largely share a functional subset with the E1B-55K product (SIEBER and DOBNER, 2007). Large E1B proteins act in concordance with E1A to allow substantive cellular changes without apoptotic induction. It is interesting to note that sole expression of E1B proteins is usually not sufficient for complete cellular transformation (ELSEN *et al.*, 1983), but either one of the gene products is capable of producing this outcome when E1A is present, albeit with differing efficiency (MCLORIE *et al.*, 1991).

Early region 1B-19K

The E1B-19K protein, named after its approximate size of 19 kilodalton (kDa), is a functional B-cell lymphoma 2 (Bcl-2) homolog, interfering with tumor necrosis factor (TNF)-α and FS-7-associated surface antigen (Fas) cell death ligand pathways (WHITE *et al.*, 1992; HAN *et al.*, 1996; LOMONOSOVA *et al.*, 2005). Cellular Bcl-2 promotes cell survival by blocking apoptosis through inhibition of the Bcl-2-associated X (BAX) and Bcl-2 homologous antagonist killer (BAK) proteins, located at the mitochondrial membranes (VAUX *et al.*, 1988; TSUJIMOTO, 1989). Unhindered action of these proteins inadvertently leads to disruption of the mitochondrial membrane and subsequent release of cytochrome c (COSULICH *et al.*, 1997). Even though both E1B-19K and Bcl-2 only share a low level of amino acid sequence homology, either one of them can functionally substitute the other, as overexpression of the latter is able to induce full cellular transformation with E1A alone (CHIOU *et al.*, 1994; BOYD *et al.*, 1994). Research on E1B-19K was vital to our understanding of cellular apoptotic processes (reviewed in CUCONATI and WHITE (2002)) and invest-

igation of its interaction partners led to the discovery of host factors like BCL2/adenovirus E1B-19K-interacting protein (BNIP-2), a protein that plays a role in cancer cell migration and evasion, amongst other things (PAN *et al.*, 2020).

Early region 1B-55K

The large E1B-55K protein, named after its approximate size of about 55 kDa, is a multifunctional viral regulation protein, involved in many processes ranging from inhibition of p53-mediated apoptosis, influence on cell cycle control, degradation of antiviral cellular proteins with adenoviral E4 open reading frame (orf)-6 (E4orf6), to promotion of viral DNA replication and mRNA transport (reviewed in BLACKFORD and GRAND (2009); HIDALGO *et al.* (2019)). Species-specific sequence homology is low in the N-terminal region but comparatively high in the rest of the protein (HIDALGO *et al.*, 2019). Experimental evidence shows that the protein forms nonglobular dimers or tetramers (MARTIN and BERK, 1998), but to date, no 3D-structure could be proven by crystallographic means. Like with E1A, the 3D-structure prediction tool Alphafold2 can be used to illustrate the disordered N-terminus, the β-barrel structure of the central region and the relatively structured C-terminus of HAdV-C5 E1B-55K, shown in Figure 6**a** (JUMPER *et al.*, 2021; MIRDITA *et al.*, 2019, 2022). Since singular point mutations in the central region can interfere with either p53 or E4orf6 binding, it is thought that compared to the largely unstructured E1A protein, E1B-55K adopts and depends on a delicate 3D structural conformation. This central region is predicted to harbor a hydrophobic core, containing a ribonucleoprotein motif that is implicated in RNA binding (HORRIDGE and LEPPARD, 1998; GONZALEZ and FLINT, 2002; TEJERA *et al.*, 2019). The intracellular localization pattern of E1B-55K differs between viral infection and cellular transformation and depends on its interaction partners. In the former, localization also depends on the state of the viral replication cycle (ZANTEMA *et al.* (1985); CHING *et al.* (2012) and reviewed in HIDALGO *et al.* (2019)). HAdV-C5 E1B-55K shuttles between nucleus and cytoplasm, with export being mediated by chromosomal maintenance 1 (CRM1) through its nuclear export signal (NES) and import through its nuclear localization signal (NLS) that is predicted to be in the C-terminal region of the protein (KRÄTZER *et al.*, 2000; KINDSMÜLLER *et al.*, 2007) (Figure 6**c**). Furthermore, localization and function of E1B-55K is influenced by post-translational modifications (PTM) with the small ubiquitin-related modifier (SUMO), a process called SUMOylation.

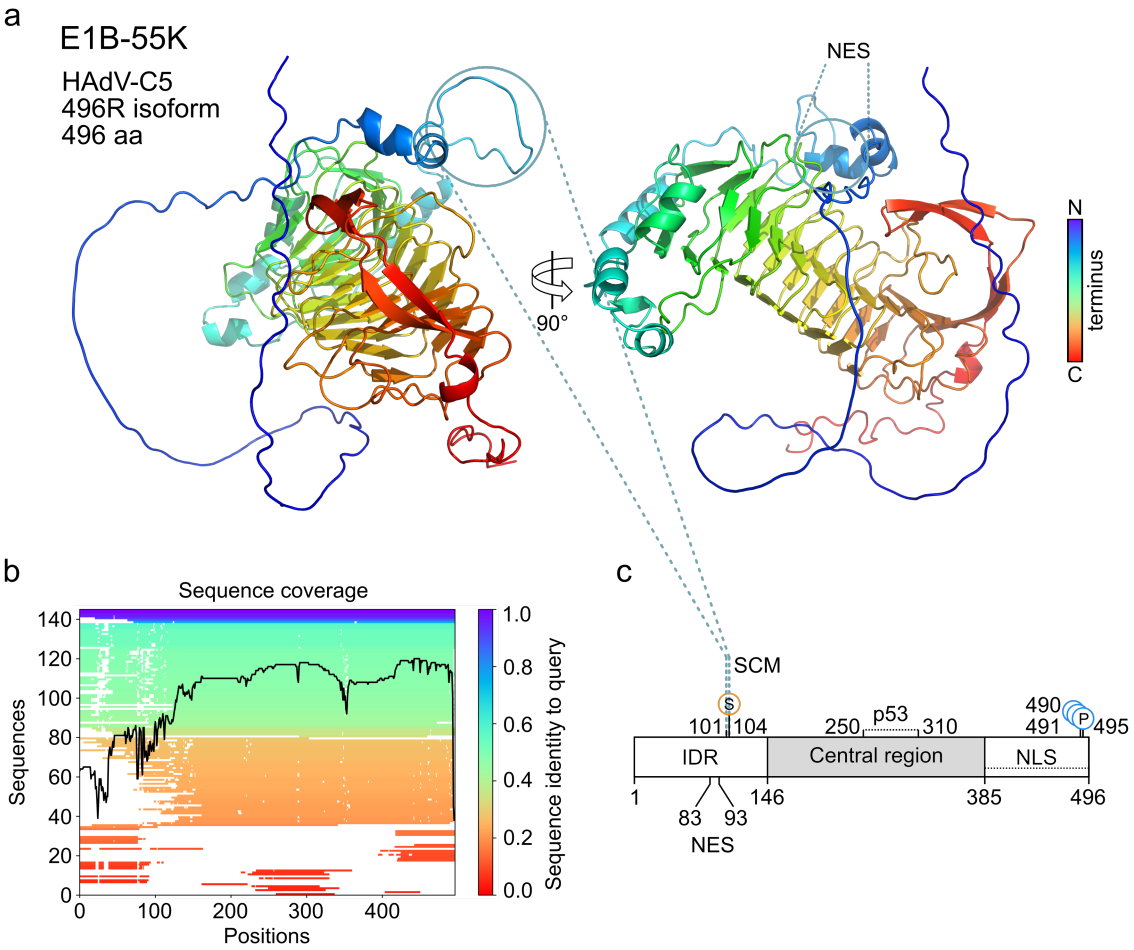


Figure 6: Alphafold2 prediction of the HAdV-C5 E1B-55K structure. **a** Rendition of HAdV-C5 E1B-55K 496R protein folding prediction by Alphafold2, colored from N-terminus (blue) to C-terminus (red) (JUMPER *et al.*, 2021; MIRDITA *et al.*, 2022), visualized by PyMOL (SCHRÖDINGER and DELANO, 2021). **b** Prediction quality sequence coverage plot visualizing the MSA depth generated by MMseqs2, which searches the query amino acid sequence against the consensus sequences of the UniRef30 database (SUZEK *et al.*, 2015; STEINEGGER and SÖDING, 2017; MIRDITA *et al.*, 2019). **c** Overview of the full HAdV-C5 E1B-55K 496R sequence structure with relevant domains and modification sites. IDR: Intrinsically disordered region, NES: Nuclear export signal, SCM: SUMO conjugation motif, S: SUMOylation site, NLS: Nuclear localization signal, P: Phosphorylation site.

Hence, SUMO is N-terminally conjugated to the lysine residue K104 within the SUMO conjugation motif (SCM) of E1B-55K, further described below. SUMOylation is an important and essential PTM involved in many key eukaryotic and viral mechanisms (reviewed in CELEN and SAHIN (2020)). Its influence ranges from cellular protein localization, structure, stability and general activity to nuclear integrity and transport, epigenetics and differentiation. A classical SCM is made up by a ΨKxE/D motif, (Ψ is a large hydrophobic amino acid, x is any amino acid and K, E and D are lysine, glutamic acid and aspartic acid, respectively), which, while being highly conserved across different adenoviral species (KOLBE *et al.*, 2022), is not present in HAdV-A12, -F40 and -G52. Unique in HAdV-C5 is a lysine at position 101 (K101), which specifically regulates the SUMOylation status of the

protein (KOLBE *et al.*, 2022). Beside SUMOylation, E1B-55K is posttranslationally modified and functionally regulated by phosphorylation at the C-terminal end at serines 490 and 491 and at threonine 495 (TEODORO *et al.*, 1994; ENDTER *et al.*, 2001; KOLBE *et al.*, 2022). Interestingly, phosphorylation and SUMOylation have been shown to influence each other, as a phosphorylation-null mutant of E1B-55K is abrogated in its SUMOylation status with significant consequences regarding specific protein interactions, like with p53 (SCHWARTZ *et al.*, 2008; WIMMER *et al.*, 2013). Due to the presence of a unique NES on HAdV-C5 E1B-55K that readily shuttles the protein out of the nucleus, SUMOylation becomes essential for some of its intra-nuclear activities, as a SUMO-null mutant E1B-55K (termed K104R, according to the substitution of its lysine to an arginine at aa position 104) is heavily impaired in most of its functional capabilites (KINDSMÜLLER *et al.*, 2007; KOLBE *et al.*, 2022). A sequence alignment between a selected subset of large E1B proteins can be found in Figure 7. The proteins of these species were more closely examined in this thesis. One of the main functions of E1B-55K is the inhibition of the E1A-activated p53-mediated apoptosis network, which is regulated through multiple different mechanisms: (i) The SUMOylation and relocalization of p53 into PML-nuclear bodies (MULLER and DOBNER, 2008; PENNELLA *et al.*, 2010) and subsequent nuclear export into cytoplasmic aggresomes (LIU *et al.*, 2005), (ii) the E1B-55K mediated tethering of a repressor domain to p53 to inhibit transcriptional activation. This function was first described *in vitro* by the group of Arnold Berk in YEW *et al.* (1994) and MARTIN and BERK (1998). (iii) The formation of a E3 ubiquitin ligase complex in conjunction with E4orf6 during viral infection, which targets a large portion of activated p53 to proteasomal degradation (QUERIDO *et al.*, 2001). Other cellular targets include the alpha thalassemia/mental retardation syndrome X-linked (ATRX) protein, members of the MRE11-RAD50-NBS1 complex, or death-associated protein 6 (Daxx) (QUERIDO *et al.*, 2001; STRACKER *et al.*, 2002; SCHWARTZ *et al.*, 2008; SCHREINER *et al.*, 2010, 2013). Unhindered, these proteins would establish transcriptional repressive complexes on the viral genome, inhibiting the replication cycle. Through its diverse set of cellular and viral interaction partners, E1B-55K can exert control over various cellular mechanisms and is heavily implicated in the process of cellular transformation. Yet, to this day, many functions of E1B-55K remain unknown.

Figure 7: Sequence alignment overlap between selected E1B-55K species. Sequence overlap and alignment between HAdV E1B-55K from types C5, A12, F40 and G52. Individual amino acids are colored in their respective RasMol color. Sequence conservation is shown by a color gradient below the alignment, ranging from red (no conservation) over white (semi-conserved) to blue (highly conserved), while sequence gaps are shown by a red dash. The consensus sequence logo is also shown, with the respective size of the letter indicating the grade of conservation. The NES and the SCM, identified in the HAdV-C5 E1B-55K, are annotated at their aa positions 83-93 and 103-106, respectively. This alignment was created with the CLC Main workbench software (QIAGEN, 2017).

1.2 Mammalian genome regulation

All domains of life follow the similar basic mechanism of gene expression in one way or another, also termed the *central dogma of molecular biology* by Nobel laureate Francis Crick (Crick, 1970). This can be described by the simplistic flow of genetic informa-

tion: First, specific gene regions of DNA get transferred to RNA replica through a process called transcription. Following this, the mRNA subset of this RNA will be adapted to amino acid polypeptides chains called proteins through ribosomal-mediated translation. Today, this view is understood as a basic assumption of how biological systems generally operate, with several exceptions: The process of reverse transcription employed by retroviruses (reviewed in JOHNSON (2019)), the presence of retrotransposons in eukaryotic genomes (reviewed in WICKER *et al.* (2007); HUA-VAN *et al.* (2011)) and the existence of the infectious proteinaceous agents knows as prions (reviewed in SCHECKEL and AGUZZI (2018)). Eukaryotes and prokaryotes do differ in several ways, as the existence of a membrane-separated nucleus in eukaryotic cells alone dramatically complicates this process. In prokaryotes, transcription and translation occur in a similar cellular environment, while eukaryotic cells have to transfer their RNA to the cytoplasm, where ribosomes can translate them into proteins. The comparatively sizable eukaryotic genome is also tightly associated with histones into structural units called nucleosomes, and forms higher order chromatin structures that have to be dynamically regulated to allow transcription to occur at all.

1.2.1 Epigenetics

Epigenetics are defined as the study of heritable and stable phenotypic changes that are not directly induced by nucleotide differences of the DNA, but by alterations to the chromosomal superstructure, thus controlling access to the DNA (BERGER *et al.*, 2009). This is mediated through the action of four mechanisms, in short: (i) modification of cytosine residues of DNA by addition of a methyl group to form 5-methylcytosine, (ii) positioning of nucleosomes, which are composed of DNA wrapped around a histone octamer and their PTMs and (iii) noncoding RNA (reviewed in LU *et al.* (2020)). In my thesis, I focus on various histone PTMs, as they provide an elegant opportunity to investigate both dynamic and long-term activating and repressing signals within a singular system. The eukaryotic genome is wrapped around histones, thereby forming the most basic unit of DNA packaging called nucleosomes. Between 145 and 147 bp are organized in a superhelix around an histone octamer (FLAUS *et al.*, 1996; LUGER *et al.*, 1997). The interaction between these molecules compacts the DNA and inhibits its accessibility by regulatory proteins. The histone octamer is a protein complex consisting of a tetramer of two copies of H3 and H4 each, combined with two dimers of H2A and H2B (Figure 8**a**) (RICHMOND *et al.*, 1984). The individual histones have extruding 'tails', which provide a domain structure

for PTM. Several different PTMs can be dynamically added and removed, predominantly acetylation (ac), mono- di- and tri-methylation (me1-3) and phosphorylation (reviewed in TAYLOR and YOUNG (2021)). Different proteins can interact with (by *readers*), add (by *writers*) or remove (by *erasers*) these histone modifications to create a 'histone code', originally proposed by STRAHL and ALLIS (2000), allowing either further compaction into transcriptionally (mostly) silent heterochromatin or transcriptionally active euchromatin.

Euchromatin

The lightly condensed euchromatin state allows transcriptional expression of selected genes by permitting interaction of the different parts of the transcriptional machinery to the unfolded DNA structure (reviewed in (MORRISON and THAKUR, 2021)). Regions commonly associated with euchromatin are promoters and enhancers. The first region is typically found upstream of the 5' end of both protein-coding and non-coding genes. As described by SMALE and KADONAGA (2003), the eukaryotic RNA polymerase II core promoter consists of special nucleotide motifs, like the TATA box (LIFTON *et al.*, 1977) and other motifs commonly required to initiate transcription through recruitment of the preinitiation complex (PIC) (8**b**). Directly adjacent to the transcriptional starting site (TSS) is the $\sim$150 bp wide nucleosome depleted region (NDR) which can contain different transcription factor binding site (TFBS), as shown in the case of the pioneer factor OCT4 (YOU *et al.*, 2011). Recent data suggest that NDR can contain highly dynamic and unstable nucleosomes consisting of H3.3 and H2A.Z histone variants in human cells (JIN *et al.*, 2009). On the other hand, NDRs are flanked by strictly positioned nucleosomes (STRUHL and SEGAL, 2013), a critical component for gene expression. Equally important for active transcription are special and potentially spatially distant regions called enhancers. These regions are primarily associated with activating histone marks and are actively transcribed to produce enhancer-associated ncRNA, described earlier as eRNA (CORE *et al.*, 2014; ALLEN *et al.*, 2014). TFBS are densely clustered around enhancers, bound with cell type-specific TFs, chromatin modifiers, transcriptional cofactors (COF) and architectural proteins like cohesin, condensin and CTCF (reviewed in PANIGRAHI and O'MALLEY (2021)). Through assembly of these proteins around the region, looping of the enhancer is mediated to get into physical contact with the target promoter region to aid activation of transcription. The histone PTMs that are associated with euchromatin are generally considered to be H3K4me3, H3K27ac and H3K36me.

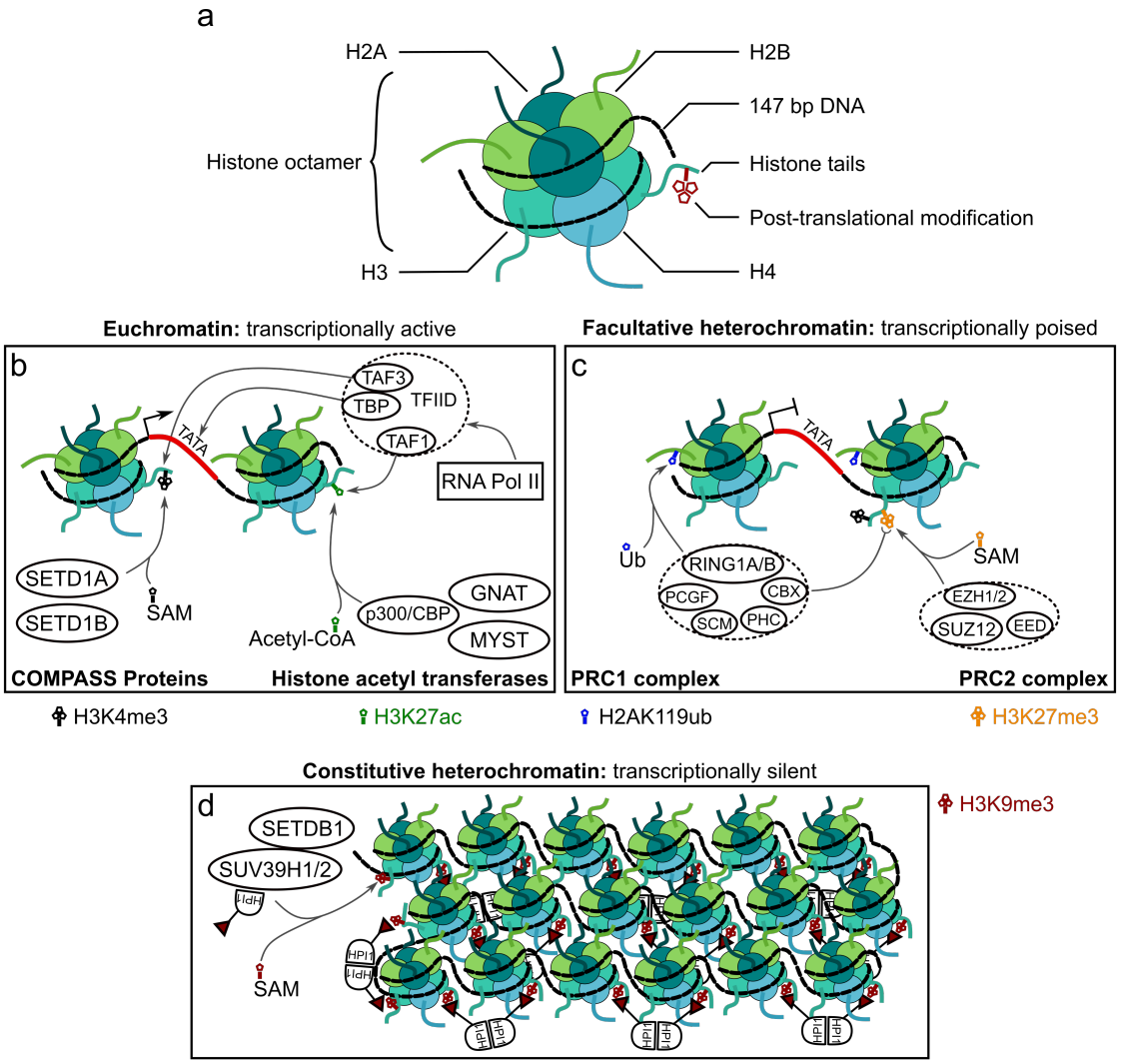


Figure 8: Histone PTMs help define chromatin states. **a** General nucleosome structure of DNA compacted around a histone octamer. **b** Euchromatic chromatin, a lightly condensed form of DNA around histones. Mostly found at genes under active transcription and usually enriched with H3K4me3 and H3K27ac, catalyzed by either COMPASS or different histone acetyltransferases, respectively. **c** Facultative heterochromatin, a mostly silent form of DNA compaction, regulated primarily by members of the polycomb-group (PcG) and associated with H2AK119Ub and H3K27me3. **d** Constitutive heterochromatin, commonly found at transcriptionally completely silent DNA. Represented by densely compacted DNA and associated with H3K9me3, catalyzed by histone lysine methyltransferases SUVH39H1/2 or SETDB1 and bound with dimerized HP1. H: histone, K: lysine, me: methylation, ac: acetylation, ub: ubiquitination. SAM: s-adenosyl methionine, acetyl-CoA: acetyl coenzyme A.

The methylation of histone 3 lysine 4, named H3K4me3, is a comparatively well understood modification and found especially enriched at promoter regions around the TSS (SANTOS-ROSA *et al.*, 2002). The catalyzation (*writing*) of H3K4me3 is mediated via transfer of methyl groups from S-adenosyl methionine (SAM) by histone lysine methyltransferases of the trithorax group (TrxG) proteins, colloquially termed the complex proteins associated with SET1 (COMPASS)-family (MILLER *et al.*, 2001). In mammals, these family consists of the su(var)3-9/enhancer of zeste/trithorax (SET)-domain containing protein

1A (SETD1A), SETD1B and the mixed-lineage leukemia (MLL)1-4 proteins. These can independently form multi-subunit COMPASS complexes with specific functions dependent on the exact composition (reviewed in SHILATIFARD (2012) and CENIK and SHILATIFARD (2021)). Multiple proteins are known to interact with H3K4me3 (*readers*), like the TATA-box binding associated factors (TAF)3, 4 and 9 proteins which are vital members of the transcription factor II D (TFIID)-containing PIC, visualized in Figure 8**a** (BEACON *et al.*, 2021). Removal of this PTM (*eraser*) is mediated by the lysine-specific demethylase 5A (KDM5A), which was found to specifically remove di- and trimethylated H3K4 by the group of CHRISTENSEN *et al.* (2007). The acteylation of histone 3 lysine 4, named H3K27ac, is another mark of transcriptionally active genes and found, similarly to H3K4me3, enriched at promoters but also understood to be one of the dominant marks present at active enhancers (CREYGHTON *et al.*, 2010). Catalyzation of this PTM is performed by a set of different histone acetyltransferase (HAT) proteins containing specific domains called bromodomains that recognize acetylated lysine residues (TAMKUN *et al.*, 1992). Exemplary members of these HATs are: the p300/CBP coactivator family and members of the MYST- or GNAT-family. HATs use acetyl coenzyme A (acetyl-CoA) to transfer the acetyl group to the N-terminal tails of several histones (WELLEN *et al.*, 2009; CAI *et al.*, 2011). Acetylation of histones is thought to (i) neutralize the positive charge of histones, conferred through their numerous arginine and lysine amino acids, thereby weakening the interaction between negatively-charged DNA and nucleosomes and (ii) serves as as a potential interaction site for bromodomain-containing proteins (reviewed in BANNISTER and KOUZARIDES (2011)). Removal of histone acetylation is conferred through action of one of 18 different human histone deactelyase (HDAC) proteins, grouped into four different classes (reviewed in PARK and KIM (2020)).

Heterochromatin

The antagonist of euchromatin is the structurally compacted form of DNA, called heterochromatin, which is predominantly associated with transcriptionally silent gene regulation. Heterochromatic regions play essential roles in the establishment of cell-type specific transcription patterns, in centromere functions and often contain transposable elements and satellite repeats, preventing them from spreading throughout the genome (extensively reviewed in ALLSHIRE and MADHANI (2018) and PENAGOS-PUIG and FURLAN-MAGARIL (2020)). Histone PTMs that are associated with this 'closed' state of the DNA are (i) H3K27me3 (Figure 8**c**), which is generally associated with facultative heterchromatin

through polycomb-mediated repression and (ii) H3K9me3 (Figure 8**d**), which marks the transcriptionally completely silent constitutive heterochromatin. DNA 5-Cytosine methylation plays a similarly integral role in heterochromatin establishment and maintenance, but will not be discussed further in this thesis. The trimethylation of histone 3 lysine 27 (H3K27me3), in addition to mono-ubi-quitination of histone H2A lysine 119 (H2AK119ub) are integral components of polycomb-mediated silencing. These two modifications are established through the cooperation between two protein formations: the polycomb-repressive complex (PRC)1 and PRC2 (CAO *et al.*, 2002; BOYER *et al.*, 2006). The smaller PRC2 complex contains a catalytic core via the activity of the enhancer of zeste homolog (EZH)1/2 subunits, which deposits mono-, di-, and trimethylation at H3K27 (MARGUERON *et al.*, 2008). This mark is antagonistic to the previously described H3K27ac, preventing the establishment of the activating histone PTM (TIE *et al.*, 2009) and is thought to recruit the PRC1 complex through its chromobox (CBX) component to establish H2AK119ub PTM. There exist multiple distinct canonical and non-canonical PCR1 complexes that catalyze H2AK119ub through action of their core E3 ubiquitin ligase RING1A/B with variied associated components (GAO *et al.*, 2012). This modification was found to be associated with gene silencing across multiple studies by inhibition of RNA polymerase II functionality (STOCK *et al.*, 2007; ZHOU *et al.*, 2008) and general chromatin compaction (FRANCIS *et al.*, 2004; GRAU *et al.*, 2011). Polycomb-repressed gene regions are characterized by both histone modifications, especially around developmental regulators - a state reversible by the induction of differentiation (BOYER *et al.*, 2006). A subset of promoter regions contain both H3K4me3 and H3K27me3, which are thought to be transcriptionally poised and tightly linked to a pluripotent state of stem cells in mammals and tissue-specific gene regulation (BLANCO *et al.*, 2020). Removal of H3K27me3, mediated by histone lysine demethylase KDM6B (AGGER *et al.*, 2007), would lead to activation of gene transcription from the poised state. The methylation of histone 3 lysine 9 (H3K9me3) is associated with the constitutive compacted heterochromatin and silenced genes. This modification is established through the histone methyltransferases suppressor of variegation 3-9 homolog (SUV39H)1/2 (REA *et al.*, 2000; BARSKI *et al.*, 2007), which participates with the histone methyltransferase SETDB1 in a multimeric complex (FRITSCH *et al.*, 2010). SUV39H1 also directly interacts with heterochromatin protein 1 (HP1) through its chromodomain (YAMAMOTO and SONODA, 2003), which is known to recruit more SUV39H1 activity to H3K9me3 PTM through a positive feedback loop. HP1 is known to also interact directly with H3K9me3 through its chromodomain (LACHNER *et al.*, 2001), whereas it fa-

cilitates homodimerization through its chromoshadow domain (COWIESON *et al.*, 2000) and is able to effectively spread the H3K9me3 mark to adjacent regions (CANZIO *et al.*, 2013). Maintenance of heterochromatin regions additionally depends on the human silencing hub (HUSH) complex, containing the transcription activation suppressor (TASOR), the chromodomain-containing M-phase phosphoprotein 8 (MPP8) and periphilin proteins (TCHASOVNIKAROVA *et al.*, 2015), which are able to interact and recruit SETDB1 and establish repressive H3K9me3. The KRAB-associated protein-1 (KAP1) protein was also identified to interact with SETDB1, recruiting it to specific sites and thereby contributes to HP1-mediated gene silencing (SCHULTZ *et al.*, 2002).

1.2.2 Transcription factors

Activation of transcription is a complex stepwise process, from creating an open chromatin environment, assembly of a preinitiation complex to productive elongation. In some cases, these steps are preceded by binding of so-called pioneer factors, which are transcription factors that are able to bind to closed heterochromatin (reviewed in BALSALOBRE and DROUIN (2022)). Transcription factor (TF)s are proteins that play an integral role in the interpretation of the genome by (i) directly interacting with specific nucleotide sequences, referred to as motifs or TFBS, through their DNA-binding domains and thereby (ii) regulating the expression of associated genes (extensively reviewed in LAMBERT *et al.* (2018)). These motifs are usually located in proximal or distal key regulatory regions, called promoters, enhancers or silencers and any typical human gene will have multiple different motifs in its regulatory regions (WEIRAUCH and HUGHES, 2014). Interestingly, the overwhelming majority of predicted TFBS are non-functional - a conundrum termed the 'futility theorem' by WASSERMAN and SANDELIN (2004). Additionally, follow-up studies have indicated that only a small subset of transcription-factor-bound TFBS are actually associated with differential expression of the adjacent gene (CUSANOVICH *et al.*, 2014). In the same work, the researchers found that functional TF binding, as in binding events leading to differential expression of the associated gene, is predicted by (a) the amount of different TFBS in the regulatory region, by (b) binding affinity to the canonical TF motif, the (c) number of binding sites of the specific factor in the vicinity of the gene and (d) lastly the designation of the bound region as an "active enhancer". It is also well understood that cooperation between different TFs is a vital determinant of gene expression (reviewed in REITER *et al.* (2017)). The mechanisms by which TFs can regulate gene expression are remarkably complex and depend on many factors, like the availability of COFs or are

contingent to the TF itself. Some TFs act by steric means, blocking other factors from binding to the regulatory site (AKERBLOM *et al.*, 1988), while others initiate binding of the RNA polymerase II as members of the transcription PIC, like the TATA-box binding protein (TBP) (HOFFMANN *et al.*, 1990). The predominant mode of transcriptional regulation is thought to be mediated through the recruitment of COFs, co-activators or co-repressors (REITER *et al.*, 2017). These COFs can contain domains involved in DNA-binding, modification of histone PTMs or remodeling of nucleosomes (FRIETZE and FARNHAM, 2011). The ATP-dependent SWI/SNF complex is a well-understood example of a COF in eukaryotes that acts through destabilization of histone-DNA interactions, leading to nucleosome rearrangement, providing easier access to chromatin (first described in STERN *et al.* (1984)). In the following sections, the p53, TEAD and AP-1 transcription factors are explained in more detail, as their investigation played a key part in this thesis.

Tumor suppressor p53

Commonly called the ”guardian of the genome” (LANE, 1992) and the most studied human gene to this day, the tumor suppressor and transcription factor p53 plays a vital role in the response to cellular stress involving DNA damage, oncogene activation, telomere erosion or hypoxia (reviewed in JOERGER and FERSHT (2016); HAFNER *et al.* (2019)). Its activity is predominantly regulated via the mouse double minute 2 homolog (MDM2) protein, an E3 ubiquitin ligase that continuously keeps the steady-state levels of p53 at a minimum (HAUPT *et al.*, 1997) and directly inhibits its transcriptional activity (OLINER *et al.*, 1993). This degradation and inhibition is reversible through the stress-induced DNA-PK-mediated phosphorylation of p53, rescuing the protein from MDM2 (SHIEH *et al.*, 1997). A different way to relinquish MDM2-mediated repression is expression of the p14 alternative reading frame (p14ARF), encoded in the cyclin-dependent kinase inhibitor 2A (CDKN2A) locus, a tumor suppressor that inactivates the MDM2 protein and therefore stabilizes the cellular tumor antigen p53 (DE STANCHINA *et al.* (1998); ZHANG *et al.* (1998), reviewed in GALLIMORE and TURNELL (2001)). Antagonization of MDM2 leads to myriad of consequences, most prominently (i) cell-cycle arrest, (ii) DNA repair, (iii) apoptosis, (iv) deregulation of homeostasis and (v) induction of autophagy, among others. Upon activation, p53 forms a tetrameric complex (WANG *et al.*, 1995) with each dimer binding to the half-site consensus sequence RRRCWWGYYY (R = A/G; W = A/T; Y = C/T) (EL-DEIRY *et al.*, 1992). Single-nucleotide polymorphisms within this core-motif results in reduced transactivation of a reporter gene (TOMSO *et al.*, 2005). It was shown that binding of p53 tetrameters partly

depends on the chromatin regulator remodeling and spacing factor 1 (RSF1), as knock-out experiments reduced the formation of p53/p300 complexes. Interestingly, p53 is sufficient in activation of gene expression without association of other transcription factors at a specific set of p53-responsive enhancers (VERFAILLIE *et al.*, 2016). Though independent from other transcription factors in some cases, activation of p53-dependent enhancer activity hinges on the premeditated p53-independent production of enhancer RNA (eRNA) (MELO *et al.*, 2013), which in turn stimulates this production even further through its interaction with these regions (ALLEN *et al.*, 2014). It is thought that transcriptional activation after the previously described enhancer- or promoter binding events of p53 is mediated through the recruitment of the acetyltransferase p300, an action required for the subsequent expression of e.g. cyclin-dependent kinase inhibitor 1A (CDKN1A), among others (BARLEV *et al.*, 2001) by induction of H3 acetylation. The interaction between p53 and p300 is additionally thought to induce recruitment of the methyltransferase SET domain containing 1C (SET1C), leading to accumulation of gene-activating histone H3 lysine 4 trimethylation (H3K4me3) (TANG *et al.*, 2013). To this day, p53 remains a major scientific challenge, with sometimes contradictory bifunctional activities and puzzeling outcomes (as argued by AYLON and OREN (2016)), and further research in exactly how and why viruses have evolved to deregulate its functions is warranted.

AP-1 family transcription factors

The activator protein-1 (AP-1) proteins were first discovered and described as viral oncogenes expressed by avian sarcoma- and murine osteosarcoma virus (CURRAN *et al.*, 1982; MAKI *et al.*, 1987). After identification of cellular homologs, the c-Jun and c-Fos proteins (BOHMANN *et al.*, 1987), the family has grown to further encompass the Maf- and Atf subfamilies (NISHIZAWA *et al.*, 1989). These four subfamilies can form different combinations of homo- and heterodimers (CHINENOV and KERPPOLA, 2001) that subsequently binds to DNA by formation of a leucine zipper structure (GLOVER and HARRISON, 1995) through their basic leucine zipper domain (bZIP). These dimers preferentially bind to the TPA-responsive element (TRE), the cAMP-response element (CRE), among others, although different combinations display higher affinity for specific sites (RAUSCHER *et al.*, 1988; NAKABEPPU *et al.*, 1988). The TRE motif, represented as the nucleotide sequence TGASTCA, is thought to be preferentially bound by heterodimers of c-Jun and c-Fos, while the CRE motif, represented as the nucleotide sequence TGASSTCA, is more associated with heterodimers of c-Jun and ATF. The AP-1 family of transcription factors is implicated in a

wide breadth of biological pathways, ranging from (i) cell cycle and cellular proliferation, to (ii) apoptosis, (iii) autophagy, (iv) immune cell development and (v) development and maintenance of cancer (reviewed in (WU *et al.*, 2021)). The AP-1 motif is adjacent to many other TFBS (KHERADPOUR and KELLIS, 2014), suggesting high degrees of cooperation with other TFs, is enriched in active enhancer regions (NGUYEN *et al.*, 2016) and is strongly associated with the transcriptional response of the Hippo signaling pathway (ZANCONATO *et al.*, 2015), among others. They are also implicated in the TGF-β-induced epithelial-to-mesenchymal transition (EMT) in A549 cells (CHANG *et al.*, 2016).

TEAD family transcription factors

The TEA domain family member (TEAD) proteins encompass a comparatively small family of four transcription factors (TEAD1-4) that are directly implicated as key regulators of cellular development (reviewed in CURREY *et al.* (2021)). All members of this group have a TEA-motif binding domain that recognizes the nucleotide sequence CATTCCW (FARRANCE *et al.*, 1992; BÜRGLIN, 1991). TEAD are the effectors of the Hippo signaling pathway, regulating (i) cellular proliferation, (ii) cell stemness and (iii) cellular differentiation (reviewed in MA *et al.* (2019)). In contrast to p53, the TEAD proteins require binding of a COF, primarily the yes-associated protein (YAP)1 and WW domain-containing transcription regulator protein 1 (TAZ) proteins. In the ”off”-state of the pathway, dephosphorylated YAP and TAZ are translocated into the nucleus, where they interact with TEAD family proteins to activate transcription of pathways predominantly associated with cell density, morphology, metabolic stress and others (reviewed in LOW *et al.* (2014)) and cross-talk with a variety of other signaling pathways, such as the TGF-β and Wnt/β-catenin pathways (HUH *et al.*, 2019). Upon specific cues, like cell contact inhibition, YAP and TAZ are inactivated and remain in the cytoplasm (ZHAO *et al.*, 2007a). It could be shown that a soft matrix surface would lead to a similar outcome, while a stiff matrix surface would activate them, re-establishing localization of YAP and TAZ into the nucleus (DUPONT *et al.*, 2011; ARAGONA *et al.*, 2013). In the same work, the authors could show that this regulation required Rho GTPase activity and tension mediated through the F-actin cytoskeleton. Inactivation of YAP and TAZ is conveyed through multiple phosphorylation events mediated by the large tumor suppressor kinase (LATS)1 and 2 proteins, which phosphorylate YAP and TAZ, inhibiting their nuclear localization, ultimately leading to their proteasomal degradation (ZHAO *et al.*, 2010). The other major interaction partners of the TEAD transcription factor that could be identified are members of the vestigial-like (VGLL) family (VAUDIN

et al., 1999). Multiple studies have shown that the TEAD4-VGLL complex targets different genes than the TEAD-YAP/TAZ complex, especially when the Hippo signaling pathways is in its ”on” state (POBBATI *et al.*, 2012; CAGLIERO *et al.*, 2013). Current evidence points towards multiple roles in tumor development and progression through its interaction with TEAD (reviewed in POBBATI and HONG (2013); YAMAGUCHI (2020)).

Chapter 2

Materials

2.1 Cells

2.1.1 Bacterial strains

Strain	Genotype
DH5α	supE44, ΔlacU169, (Δ80dlacZΔM15), hsdR17, recA1, endA1, gyrA96, thi-1, relA1 (HANAHAN, 1983)

2.1.2 Mammalian cells

Strain	Genotype
H1299	Human lung epithelial carcinoma cell line, homozygous partial deletion of the TP53 gene (MITSUDOMI *et al.*, 1992)
HCT116	Human colon carcinoma cell line, RAS-mutation in codon 13 (BRATTAIN *et al.*, 1981)
HEK293T	Human embryonic kidney (HEK) 293 cells, containing HAdV-C5 DNA (nucleotides 1 to 4344), early region 1 integrated into chromosome 19 and stable expression of large T antigen from SV40 (PEAR *et al.*, 1993)
hMSC	Human mesenchymal stromal cells, bone marrow-derived (LANGE *et al.*, 2007)
pBRK	Primary baby rat kidney cells, freshly isolated and propagated from 3 to 5 days old Sprague-Dawley rats (Janvier, France)
MCF-7	Human breast carcinoma cell line (S. LEUNG *et al.*, 1973)

2.2 Nucleic acids

2.2.1 Vector plasmids

All vector plasmids were numbered according to the internal group *Filemaker Pro* database.

#	Name	Characteristics
136	pcDNA3	Mammalian expression vector under control of a CMV promoter
152	pCMX3b-Flag	Mammalian expression vector under control of a CMV promoter, contains FLAG-tag
232	LeGO-iVLN2	LeGO vector containing a venus-neomycin resistance fusion protein under the control of a SFFV promoter (WEBER *et al.*, 2008)
234	LeGO-iBLB2	LeGO vector containing a BFP-blasticidin resistance fusion protein under the control of a SFFV promoter (WEBER *et al.*, 2008)

2.2.2 Recombinant plasmids

All vector plasmids were numbered according to the internal group *Filemaker Pro* database. Marked (*) plasmids have been created during this work.

#	Name	Characteristics
1968	pCMV-VSV-G	3rd generation lentiviral packaging plasmid, contains VSVG envelope protein (DULL *et al.*, 1998).
1969	pRSV Rev	3rd generation lentiviral packaging plasmid, contains HIV-1 rev under control of RSV U3 promoter (DULL *et al.*, 1998)
1970	pMDLg/pRRE	3rd generation lentiviral packaging plasmid, contains HIV-1 gag and pol and the RRE binding site (DULL *et al.*, 1998)
1319	pcDNA3-E1B-55K	pcDNA3 vector containing the HAdV-C5 E1B-55K gene under the control of the CMV promoter
2215	Ad5 HA55K pcDNA3	pcDNA3 vector containing the HAdV-C5 E1B-55K gene (with an HA-tag at its N-terminus) under the control of the CMV promoter
2217	Ad12 HA55K pcDNA3.1(+)	pcDNA3.1(+) vector containing the HAdV-A12 large E1B gene (with an HA-tag at its N-terminus) under the control of the CMV promoter
3713	pCMX3b-Flag-TEAD4	pCMX3b vector containing the human TEAD4 gene (with a Flag-tag at its N-terminus) under the control of the CMV promoter
3706*	LeGO-iVLN2 p53 wt	LeGO vector containing the human wildtype p53 gene region and a venus-neomycin resistance fusion protein under the control of a SFFV promoter

#	Name	Characteristics
2103	LeGO-iVLN2-E1Agenomisch	LeGO vector containing the HAdV-C5 E1A gene region and a venus-neomycin resistance fusion protein under the control of a SFFV promoter
2402	LeGO-iBLB2-E1Blang	LeGO vector containing the HAdV-C5 E1B gene region and a BFP-blasticidin resistance fusion protein under the control of a SFFV promoter
3455*	LeGO-iBLB2-E1Blang 55K 4xStopp	LeGO vector containing the HAdV-C5 E1B gene region (with four stop codons to stop expression of E1B-55K) and a BFP-blasticidin resistance fusion protein under the control of a SFFV promoter
3503*	LeGO-iBLB2 E1Blang HA-55K 4xStopp	LeGO vector containing the HAdV-C5 E1B gene region (with an HA-tag at N-terminus of E1B-55K and four stop codons to stop expression of E1B-55K) and a BFP-blasticidin resistance fusion protein under the control of a SFFV promoter
3667	LeGO-iBLB2-E1Blang HA-55K	LeGO vector containing the HAdV-C5 E1B gene region (with an HA-tag at N-terminus of E1B-55K) and a BFP-blasticidin resistance fusion protein under the control of a SFFV promoter
3665	LeGO-iBLB2-E1Blang HA-55K K101R	LeGO vector containing the HAdV-C5 E1B gene region (with an HA-tag at N-terminus of E1B-55K K101R) and a BFP-blasticidin resistance fusion protein under the control of a SFFV promoter (KOLBE *et al.*, 2022)
3657	LeGO-iBLB2-E1Blang HA-55K K104R	LeGO vector containing the HAdV-C5 E1B gene region (with an HA-tag at N-terminus of E1B-55K K104R) and a BFP-blasticidin resistance fusion protein under the control of a SFFV promoter (ENDTER *et al.*, 2001)
3666	LeGO-iBLB2-E1Blang HA-55K NES	LeGO vector containing the HAdV-C5 E1B gene region (with an HA-tag at N-terminus of E1B-55K NES) and a BFP-blasticidin resistance fusion protein under the control of a SFFV promoter (KRÄTZER *et al.*, 2000; KINDSMÜLLER *et al.*, 2007)
3668	LeGO-iBLB2-E1Blang HA-55K NES/K104R	LeGO vector containing the HAdV-C5 E1B gene region (with an HA-tag at N-terminus of E1B-55K NES/K104R) and a BFP-blasticidin resistance fusion protein under the control of a SFFV promoter (KRÄTZER *et al.*, 2000; KINDSMÜLLER *et al.*, 2007; KOLBE *et al.*, 2022)

#	Name	Characteristics
3154	Ad12 E1B-55K LeGO-iBLB2+	LeGO vector containing the HAdV-A12 large E1B gene region (with an HA-tag at N-terminus of large E1B) and a BFP-blasticidin resistance fusion protein under the control of a SFFV promoter (WEBER *et al.*, 2008)
3141	Ad40 E1B-55K LeGO-iBLB2+	LeGO vector containing the HAdV-F40 large E1B gene region (with an HA-tag at N-terminus of large E1B) and a BFP-blasticidin resistance fusion protein under the control of a SFFV promoter (WEBER *et al.*, 2008)
3637	Ad52 E1B-55K LeGO-iBLB2+	LeGO vector containing the HAdV-G52 large E1B gene region (with an HA-tag at N-terminus of large E1B) and a BFP-blasticidin resistance fusion protein under the control of a SFFV promoter (WEBER *et al.*, 2008)

2.2.3 Oligonucleotides

All nucleotides were ordered from Metabion and numbered according to the internal group *Filemaker Pro* database.

#	Name	Sequence (5' → 3')	Purpose
3554	GPI_promoter_fwd	CCTGCATCTACGGTCTGTGA	qPCR H3K4me3/K27ac
3555	GPI_promoter_rev	GTGGTGGTGTCTGTCACCTG	qPCR H3K4me3/K27ac
3558	C1orf43_promoter_fwd	AGTGGGTGGAGAATGCAGAC	qPCR H3K4me3/K27ac
3559	C1orf43_promoter_rev	GAGATTACCCCACCCCATTC	qPCR H3K4me3/K27ac
3564	HOXC12_proximal_fwd	AGTAGTTCGCCCCCAGATTT	qPCR H3K27me3
3565	HOXC12_proximal_rev	GCGGAAGGGAGGTAGAGAAT	qPCR H3K27me3
3568	HOXC13_fwd	GAGCCCGAGATTCACTCAAC	qPCR H3K9/K27me3
3569	HOXC13_rev	TTATGCCCAGTTTTGGGGTA	qPCR H3K9/27me3
3678	Malat1genebody1_fw	GCCTAGATGCAGAGAAAACAGC	qPCR MALAT1
3679	Malat1genebody1_rv	TCAGGGTTATGCTTATTCCCCA	qPCR MALAT1
3680	Malat1genebody2_fw	CAGTACTGTTCTGATCCCGCT	qPCR MALAT1
3681	Malat1genebody2_rv	CTCCAAGCATTGGGGAACACA	qPCR MALAT1
3682	Malat1upstream_fw	AGGGCCGTAGATGACGGAT	qPCR MALAT1
3683	Malat1upstream_rv	CCAGAGAATGCCAGAATGCCA	qPCR MALAT1
3684	Malat1downstream_fw	TGATCTAATGGGGGATGGGGT	qPCR MALAT1
3685	Malat1downstream_rv	ACTTAGGAGAGCACCTAGCTCA	qPCR MALAT1

2.3 Antibodies

2.3.1 Primary antibodies

All primary antibodies were numbered according to the internal group *Filemaker Pro* database.

#	Name	Properties	Purpose
1	2A6	Monoclonal mouse antibody against the N-terminus of HAdV-C5 E1B-55K (SARNOW *et al.*, 1982)	ChIP, Western blot
54	α-p53 (FL393)	Polyclonal rabbit antibody against full-length p53 (Santa Cruz, sc-6243)	Western blot
88	α-β-actin (clone AC-15)	Monoclonal mouse antibody against the N-terminus of the β-isoform of actin (Sigma Aldrich, A5441)	Western blot
98	pAB421	Monoclonal mouse antibody against p53 (HALAZONETIS *et al.*, 1993)	Western blot
131	M73	Monoclonal mouse antibody against HAdV-C5 E1A (HARLOW *et al.*, 1985)	Western blot
272	XPH9	Monoclonal mouse antibody against HAdV-A12 large E1B (MERRICK *et al.*, 1991)	ChIP, qPCR
490	α-E1B-19K	Polyclonal rabbit antibody against HAdV-C5 E1B-19K (LOMONOSOVA *et al.*, 2005)	Western blot
531	α-IgG mouse	Polyclonal mouse antibody against IgG (Santa Cruz, sc-2025)	ChIP
629	α-HA	Monoclonal rat antibody against HA-tag (Merck, 3F10, 12158167001)	IP, Western blot
634	α-H3K4me3	Monoclonal rabbit antibody against H3K4me3, ChIP-graded (Merck, 04-745)	ChIP, Western blot, qPCR
635	α-H3K27me3	Monoclonal rabbit antibody against H3K4me3, ChIP-graded (Merck, 04-449)	ChIP, Western blot, qPCR
636	α-H3ac	Monoclonal rabbit antibody against H3ac, ChIP-graded (Merck, 06-559)	ChIP, Western blot, qPCR
637	α-H3K9me3	Monoclonal rabbit antibody against H3K9me3, ChIP-graded (Active Motif, 39161)	ChIP, Western blot, qPCR

#	Name	Properties	Purpose
643	α-HA (ChIP)	Polyclonal rabbit antibody against HA-tag, ChIP-graded (Abcam, ab9110)	ChIP, Immunofluorescence, Western blot
646	α-p21	Monoclonal mouse antibody against full-length mouse p21 (Santa Cruz, sc-6246)	Immunofluorescence, Western blot
653	α-H3K27ac	Monoclonal rabbit antibody against H3K27ac, ChIP-graded (Abcam, ab4729)	ChIP, qPCR
697	α-IgG rabbit	Polyclonal rabbit antibody against IgG (Cell Signaling, 2729)	ChIP
698	α-p53 (DO-7)	Monoclonal mouse antibody against p53, ChIP-graded (Cell Signaling, DO-7, 48818)	ChIP, qPCR
707	α-TEAD4	Monoclonal mouse antibody against TEAD4/TEF-3 (Santa Cruz, sc-101184)	IP, Western blot

2.3.2 Secondary antibodies

All secondary antibodies were numbered according to the internal group *Filemaker Pro* database.

#	Name	Properties	Purpose
156	HRP-α-mouse	Polyclonal horseradish peroxidase-conjugated goat antibody against mouse IgG (Jackson, 115-036-003)	Western blot
685	HRP-α-rabbit	Polyclonal horseradish peroxidase-conjugated goat antibody against rabbit IgG (Jackson, 111-036-003)	Western blot
-	HRP-α-rat	Polyclonal horseradish peroxidase-conjugated goat antibody against rat IgG (Jackson, 112-036-003)	Western blot

2.3.3 Immunofluorescence antibodies

All immunofluorescence antibodies were numbered according to the internal group *Filemaker Pro* database.

#	Name	Properties	Purpose
-	Alexa Fluor™ 488-α-mouse	Polyclonal Alexa™488-conjugated goat antibody against mouse IgG (Invitrogen, A-11001)	Immunofluorescence
-	Alexa Fluor™ 488-α-rabbit	Polyclonal Alexa™488-conjugated goat antibody against rabbit IgG (Invitrogen, A-11008)	Immunofluorescence
-	Alexa Fluor™ 555-α-mouse	Polyclonal Alexa™555-conjugated goat antibody against mouse IgG (Invitrogen, A-21427)	Immunofluorescence
-	Alexa Fluor™ 555-α-rabbit	Polyclonal Alexa™555-conjugated goat antibody against rabbit IgG (Invitrogen, A-32732)	Immunofluorescence

2.4 Standards and markers

Product name	Company
1 kb or 100 bp DNA ladder	New England Biolabs
PageRuler™ Plus Prestained Protein Ladder, 10 to 250 kDa	Thermo Scientific

2.5 Buffers, media and solvents

Buffers for agarose gel electrophoresis, SDS-PAGE and western blot

Name	Composition	Purpose
6x DNA loading dye	0.25% Bromphenol blue (w/v) 0.25% Xylene cyanol (w/v) 50% Glycerol (v/v) 2% 50x Tris-acetate (v/v)	Agarose gel electrophoresis

Name	Composition	Purpose
5x TBE buffer (pH 7.8)	450 mM Tris/HCl (pH 8.0) 450 mM Boric acid 10 mM Ethylenediaminetetraacetic acid	Agarose gel electrophoresis
RIPA lysis buffer	50 mM Tris/HCl (pH 8.0) 150 mM Sodium chloride 5 mM Ethylenediaminetetraacetic acid 1% Nonidet P-40 (v/v) 0.1% Sodium dodecyl sulfate (w/v) 0.5% Sodium deoxycholate (v/v)	SDS-PAGE cellular lysis buffer for protein extraction
High salt lysis buffer	50 mM Tris/HCl (pH 8.0) 300 mM Sodium chloride 0.5% Triton X-100 (v/v) 0.1% Sodium dodecyl sulfate (w/v)	SDS-PAGE cellular lysis buffer for histone protein extraction
5x Laemmli buffer	100 mM Tris/HCl (pH 6.8) 200 mM Dithiothreitol 10% Sodium dodecyl sulfate (w/v) 0.2% Bromphenol blue (w/v)	SDS-PAGE protein sample buffer
Acrylamide stock (30%)	29% Acrylamide N (w/v) 1% N'Methylenbisacrylamide (w/v)	Organic compound used for SDS-PAGE protein gels
SDS-PAGE stacking gel (5%)	17% Acrylamide stock (30%) (v/v) 120 mM Tris/HCl (pH 6.8) 0.1% Sodium dodecyl sulfate (w/v) 0.1% Ammonium persulfate (w/v) 0.1% Tetramethylethylenediamine (v/v) Bromphenol blue	SDS-PAGE stacking gel for protein gathering
SDS-PAGE resolving gel (10%)	34% Acrylamide stock (30%) (v/v) 250 mM Tris/HCl (pH 8.8) 0.1% Sodium dodecyl sulfate (w/v) 0.1% Ammonium persulfate (w/v) 0.04% Tetramethylethylenediamine (v/v)	SDS-PAGE resolving gel for protein separation
SDS-PAGE resolving gel (15%)	50% Acrylamide stock (30%) (v/v) 250 mM Tris/HCl (pH 8.8) 0.1% Sodium dodecyl sulfate (w/v) 0.1% Ammonium persulfate (w/v) 0.04% Tetramethylethylenediamine (v/v)	SDS-PAGE resolving gel for protein separation

Name	Composition	Purpose
TGS buffer	25 mM Tris 200 mM Glycine 0.1% Sodium dodecyl sulfate (w/v)	SDS-PAGE buffer for protein separation
Towbin buffer	25 mM Tris/HCl (pH 8.3) 200 mM Glycine 0.05% Sodium dodecyl sulfate (w/v) 20% Methanol (v/v)	Buffer for continuous western blotting
PBS-Tween	0.1% Tween20 (v/v) in 1x Phosphate-buffered saline	Membrane washing buffer

Buffers for indirect immunofluorescence

Name	Composition	Purpose
Permeabilization solution	0.5% Triton X-100 (v/v) in 1x Phosphate-buffered saline	Cellular membrane permeabilization buffer
TBS-BG	20 mM Tris-HCl (pH 8.0) 137 mM Sodium chloride 3 mM Potassium chloride 1.5 mM Magnesium chloride 0.05% Tween-20 (v/v) 0.05% Sodium azide (w/v) 5% Glycine (w/v) 5% Bovine serum albumin (w/v)	Blocking buffer to mask potential unspecific antibody binding sites

Buffers for cross-linking chromatin immunoprecipitation (ChIP)

Name	Composition	Purpose
1M Tris-HCl (0.22 μm sterile filtered, pH 8.0)	12.04% Tris (w/v) in double-distilled H_20 Hydrochloric acid (HCl) for pH	Stock solution for ChIP buffers
1M Hepes-KOH (0.22 μm sterile filtered, pH 7.5)	23.8% Hepes (w/v) in double-distilled H_20 Potassium hydroxide (KOH) for pH	Stock solution for ChIP buffers
1.25M Glycine (0.22 μm sterile filtered)	9.4% Glycine (w/v) in double-distilled H_20	Quenching solution for formaldehyde fixation

Name	Composition	Purpose
5M Sodium chloride	29.22% Sodium chloride (w/v) in double-distilled H_20	Stock solution for various ChIP buffers
0.5M EDTA (0.22 μm sterile filtered, pH 8.0)	18.6% Ethylenediaminetetraacetic acid in double-distilled H_20 Sodium hydroxide (NaOH) for pH	Stock solution for various ChIP buffers
0.5M EGTA (0.22 μm sterile filtered, pH 8.0)	19.33% Triethyleneglycoldiamine-tetraacetic acid in double-distilled H_20 Sodium hydroxide (NaOH) for pH	Stock solution for various ChIP buffers
ChIP buffer I (0.22 μm sterile filtered)	50 mM Hepes-KOH (pH 8.0) 140 mM Sodium chloride 1 mM Ethylenediaminetetraacetic acid 10% Glycerol (v/v) 0.5% Nonidet P-40 0.25% Triton X-100 in double-distilled H_20	ChIP cellular membrane lysis for nuclei extraction
ChIP buffer II (0.22 μm sterile filtered)	10 mM Tris-HCl (pH 8.0) 200 mM Sodium chloride 1 mM Ethylenediaminetetraacetic acid 0.5% mM Triethyleneglycoldiamine-tetraacetic acid in double-distilled H_20	ChIP nuclei washing buffer
ChIP buffer III (0.22 μm sterile filtered)	50 mM Tris-HCl (pH 8.0) 10 mM Ethylenediaminetetraacetic acid 1% Sodium dodecyl sulfate (w/v) in double-distilled H_20	ChIP nuclear membrane lysis buffer
Dilution buffer (0.22 μm sterile filtered)	16.7 mM Tris-HCl (pH 8.0) 167 mM Sodium chloride 10 mM Ethylenediaminetetraacetic acid 1.1% Triton X-100 0.01% Sodium dodecyl sulfate (w/v) in double-distilled H_20	ChIP buffer for antigen precipitation

Name	Composition	Purpose
Low-salt buffer (0.22 μm sterile filtered)	20 mM Tris-HCl (pH 8.0) 150 mM Sodium chloride 2 mM Ethylenediaminetetraacetic acid 1% Triton X-100 0.1% Sodium dodecyl sulfate (w/v) in double-distilled H_20	ChIP buffer for removal of non-specific complexes
High-salt buffer (0.22 μm sterile filtered)	20 mM Tris-HCl (pH 8.0) 500 mM Sodium chloride 2 mM Ethylenediaminetetraacetic acid 1% Triton X-100 0.1% Sodium dodecyl sulfate (w/v) in double-distilled H_20	ChIP buffer for removal of non-specific complexes
LiCl wash buffer (0.22 μm sterile filtered)	10 mM Tris-HCl (pH 8.0) 250 mM Lithium chloride 1 mM Ethylenediaminetetraacetic acid 1% Nonidet P-40 1% Sodium deoxycholate (v/v) in double-distilled H_20	ChIP buffer for removal of non-specific complexes
TE buffer (0.22 μm sterile filtered)	10 mM Tris-HCl (pH 8.0) 1 mM Ethylenediaminetetraacetic acid in double-distilled H_20	ChIP washing buffer
Elution buffer (0.22 μm sterile filtered)	50 mM Tris-HCl (pH 8.0) 10 mM Ethylenediaminetetraacetic acid 1% Sodium dodecyl sulfate in double-distilled H_20	ChIP decrosslinking buffer
$CaCl_2$ buffer (0.22 μm sterile filtered)	10 mM Tris-HCl (pH 8.0) 300 mM Calcium chloride in double-distilled H_20	Ingredient for enzyme reactions
EB buffer (0.22 μm sterile filtered)	10 mM Tris-HCl (pH 8.0) 1% Sodium dodecyl sulfate in double-distilled H_20	ChIP DNA elution buffer

Buffers for native chromatin immunoprecipitation (native MNase-ChIP)

Name	Composition	Purpose
Nuclear isolation buffer (0.22 μm sterile filtered)	50 mM Hepes-KOH (pH 8.0) 140 mM Sodium chloride 10% Glycerol (v/v) 0.5% Nonidet P-40 0.25% Triton X-100 in double-distilled H_20	Native ChIP buffer for cellular membrane lysis
Immunoprecipitation buffer (0.22 μm sterile filtered)	20 mM Tris-HCl (pH 8.0) 2 mM Ethylenediaminetetraacetic acid 150 mM Sodium chloride 0.1% Triton X-100 in double-distilled H_20	Native ChIP buffer for antigen precipitation
Triton-sodium-deoxycholate solution	1% Triton X-100 (v/v) 1% Sodium deoxycholate (w/v) in double-distilled H_20	Ingredient for native ChIP

Buffers for cell counting and transfection experiments

Name	Composition	Purpose
Phosphate-buffered saline (PBS, autoclaved)	140 mM Sodium chloride 3 mM Potassium chloride 4 mM Disodium phosphate 1.5 mM Potassium phosphate in double-distilled H_20	Buffer used for various experiments
Trypan blue solution	0.15% Trypan blue (w/v) 0.85% Sodium chloride (w/v) in double-distilled H_20	Solution for cell counting
PEI solution (0.22 μm sterile filtered, pH 7.2)	0.1% Polyethylenimine (w/v) in double-distilled H_20 Hydrochloric acid (HCl) for pH	Transfection reagent
2x HBS precipitation buffer	50 mM Hepes 280 mM Sodium chloride in double-distilled H_20	Transfection reagent
LB-media (autoclaved)	0.5% Yeast extract (w/v) 0.5% Sodium chloride (w/v) in double-distilled H_20	Media for bacteria culturing

2.6 Enzymes, reagents and consumables

If not specifically annotated, chemicals and reagents were obtained from AppliChem, Hartenstein, Invitrogen, New England Biolabs, Merck, Roche, Sigma and Cell signaling. Cell culture material was obtained from Falcon and Gibco BRL. Plastic equipment was purchased from Falcon, Sarstedt, Whatman, Nunc, Biorad, Eppendorf GmbH, Biozym and ibidi.

2.7 Commercial systems

Name and version	**Source**	**Purpose**
CLC Main Workbench 7.9.1	QIAGEN Aarhus A/S	General nucleotide data processing
FileMaker Pro 19.0.1.116	Claris International	Database management
Inkscape 1.1	Inkscape	Figure creation
Acrobat X Pro 10.1.13	Adobe	PDF data processing
GraphPad Prism 9.4	GraphPad Software	Data graphing and statistical analysis
Integrative Genomics Viewer 2.14.1	ROBINSON *et al.* (2011)	Visualization tool for sequencing data
EaSeq 1.111	LERDRUP *et al.* (2016)	Visualization tool for sequencing data
Microsoft Excel 2016	Microsoft	Table processing and statistical analysis
Fiji (ImageJ) 1.51j	SCHINDELIN *et al.* (2012)	Immunofluorescence image processing
Overleaf (2022)	Digital Science UK Limited	LaTeX markup software
Metascape (2022)	ZHOU *et al.* (2019)	Biological pathway enrichment
Rstudio (2022.07.1, 554)	RSTUDIO TEAM (2020)	Integrated development environment for R
R language 4.2.1	R CORE TEAM (2022)	Programming language for statistical computing and graphics

2.8 Bioinformatic software

2.8.1 Linux command line tools

Command line tools used in the bioinformatic analysis

Name and version	Source	Purpose
Bioconda 4.11.0	DALE *et al.* (2018)	Collection and distribution of bioinformatic software
fastqc v0.11.9	ANDREWS (2019)	Quality control of sequencing data
multiqc 1.11	EWELS *et al.* (2016)	Quality control of sequencing data
samtools 1.7	DANECEK *et al.* (2021)	Interaction with sequencing data
bedtools v2.30.0	QUINLAN and HALL (2010)	Interaction with sequencing data
bowtie2 2.4.4	LANGMEAD and SALZBERG (2012)	Sequencing data read alignment
macs2 2.2.7.1	ZHANG *et al.* (2008)	Peak calling of sequencing data
epic2 0.0.52	STOVNER and SÆTROM (2019)	Diffuse peak calling of sequencing data
salmon 1.6	PATRO *et al.* (2017)	RNA transcript quantification
deeptools 3.5.1	RAMÍREZ *et al.* (2016)	Suite of python tools for interaction with sequencing data
diffReps 1.55.6	SHEN *et al.* (2013)	Detecting of differential chromatin modification sites
HOMER v4.11	HEINZ *et al.* (2010)	Suite of tools for motif discovery in sequencing data

2.8.2 R software packages

R packages used in the bioinformatic analysis

Name and version or year of use	**Source**
BiocManager (1.30.18)	MORGEN (2022)
TxDb.Rnorvegicus.UCSC.rn7.refGene (2022)	BIOCONDUCTOR CORE TEAM AND BIOCONDUCTOR PACKAGE MAINTAINER (2022)
TxDb.Rnorvegicus.UCSC.hg38.refGene (2022)	BIOCONDUCTOR CORE TEAM AND BIOCONDUCTOR PACKAGE MAINTAINER (2022)
org.Rn.eg.db (2022)	CARLSON (2022b)
org.Hs.eg.db (2022)	CARLSON (2022a)
biomaRt (2022)	DURINCK *et al.* (2009)
ggupset (0.3.0)	CONSTANTIN AHLMANN-ELTZE (2020)
ggimage (0.3.1)	GUANGCHUANG YU (2022)
ReactomePA (2022)	YU and HE (2016)
dplyr (1.0.9)	WICKHAM *et al.* (2022)
ChIPseeker (2022)	YU *et al.* (2015)
forcats (0.5.2)	WICKHAM (2022)
GenomicFeatures (2022)	LAWRENCE *et al.* (2013)
DESeq2 (2022)	SONESON *et al.* (2016)
TximportData (1.24.0)	LOVE (2022)
Tximeta (1.24.0)	LOVE *et al.* (2020)
ggplot2 (2022)	GUANGCHUANG YU (2022)
genefilter (1.78.0)	GENTLEMAN *et al.* (2022)
ggrepel (0.9.1)	SLOWIKOWSKI (2021)
AnnotationDbi (1.58.0)	PAGES *et al.* (2022)
pheatmap (1.0.12)	KOLDE (2019)
vsn (2022)	HUBER *et al.* (2002)
apeglm (2022)	ZHU *et al.* (2019)
RColorBrewer (1.1-3)	NEUWIRTH (2022)
ashr (2.2-54)	STEPHENS *et al.* (2022)
tidyr (1.2.0)	WICKHAM and GIRLICH MAXIMILIAN (2022)
EnhancedVolcano (1.14.0)	BLIGHE *et al.* (2022)
scales (1.2.1)	WICKHAM and SEIDEL (2022)
viridis (0.6.2)	GARNIER (2021)
shinyCircos (2021)	YU *et al.* (2018)

Chapter 3

Methods

3.1 Mammalian tissue culture techniques

3.1.1 Culture and passaging of mammalian cell lines

Tissue culture was performed in flow hoods (HerasafeTM KS1, Heraeus) under sterile conditions. Adherent human tumor cell lines and baby rat kidney (BRK) cells were cultured in *Dulbecco's Modified Eagle Medium* (DMEM; 4.5 g/l D-glucose, L-glutamine, pyruvate; Gibco®, Thermo Fisher Scientific), supplemented with 10% fetal bovine serum (FBS) (v/v) and 1% penicillin/streptomycin (v/v, 10,000 U/ml penicillin, 10 mg/ml streptomycin, Pan Biotech). Adherent human mesenchymal stromal cells (hMSC) were grown in *Alpha Minimum Essential Medium* (MEM; 1 g/l D-glucose, 2.2 g/l sodium bicarbonate, Merck), supplemented with 5% human platelet lysate supernatant (v/v, received from Claudia Lange, LANGE *et al.* (2007)), 0.04 % Heparin (200,000 I.E. Heparin-Natrium, Braun) and 2 mM L-alanyl-L-glutamine dipeptide (100X GlutaMAX™ Supplement, Thermo Fisher Scientific). All cells were incubated at 37°C with 95% relative humidity saturated with 5% CO_2 in cell culture incubators (*CO2 incubator BBD6220*, Heraeus). For cell passaging, medium was removed and cells were washed once with sterile PBS. Adherent cells were detached by incubation in an appropriate amount of 0.05% trypsin/EDTA (Pan) at 37°C. Primary BRK cells were detached with 0.25% trypsin/EDTA (Pan). Trypsin was subsequently inactivated by adding the appropriate fully supplemented culture medium. The cell suspension was transferred into a suitable reaction tube and centrifuged at 2,000 rpm for 3 min (*Multifuge* 3S-R, Heraeus). Subsequently, the supernatant was aspirated, pelleted cells were washed once (human tumor cell lines and BRK) or twice (hMSC) in PBS and subsequently resuspended in culture medium, diluted as desired for further propagation and reseeded on an adequate dish. Passaging was performed every 3-4 days, depending on the cell type, to avoid reaching maximal confluency.

3.1.2 Determination of cell numbers

Detached cells destined for specific downstream experiments were diluted 1:10 (v/v) in a trypan blue solution and counted with a hemocytometer (*Neubauer* cell counter, Carl Roth) using a light microscope (*Leica* DM IL). The total number of viable cells from four quadrants was calculated by applying the following formula:

$$\frac{cells}{ml} = \frac{counted\,cells}{4} \times dilution\ factor \times 10^4$$

3.1.3 Transfection with polyethylenimine

Pre-determined amount of cells were seeded 16 h prior to transfection on 6-well plates or 100 mm culture dishes. Plasmid DNA was mixed with 600 μl pre-warmed DMEM without additives and PEI was added in a 1:10 ratio to DMEM without supplements (v/v). The solution was vortexed shortly and incubated for 10 min at room temperature (RT). During this time, cell media was changed to DMEM without supplements. Subsequently, the transfection mixture was added dropwise to the cells and plates were swirled gently. After 6 h, the transfection solution was aspirated and replaced with standard fully supplemented culture medium. Cells were incubated at 37°C with 95% relative humidity and 5% CO_2 (*CO2 incubator BBD6220*, Heraeus) and subsequently harvested 24-72 h post transfection.

3.1.4 Transfection with calcium phosphate

This method was used for the creation of clonal BRK cell lines expressing adenovirus gene regions or for the production of lentiviral particles from 293T cells. Plasmid DNA was diluted with double distilled (dd) H_2O to a volume of 437.5 μl and mixed with 62.5 μl 2M $CaCl_2$ solution. 500 μl of 2x HBS (HEPES buffered saline) precipitation buffer was transferred to a 15 ml reaction tube and the DNA-$CaCl_2$ was added dropwise while blowing air into the precipitation buffer using a pasteur pipette attached to a pipet boy. The solution was incubated for 15 min at RT. The old cell media was swapped during this incubation step with 8 ml of standard culture medium supplemented with 25 μM chloroquine (Sigma Aldrich). The transfection mix was subsequently added dropwise to the cells and the plates were swirled gently. After 6-8 h the medium was aspirated and 6-8 ml of fresh fully supplemented DMEM or MEM (containing 20 mM HEPES in case of lentivirus production) was added. Virus supernatants were harvested after 48 h and sterile filtered through a 0.22 μm *MILLEX-GP* sterile filter (Merck Millipore). Virus supernatants were aliquoted and stored at -80°C or used directly. In case of the transformation assays

with primary baby rat kidney cells, the dishes were incubated for 4-6 weeks and media was changed every 3-4 days. Non-transfected cells generally died, whereas cells transfected with the E1-region infrequently formed multi-layered cell colonies, so called foci, which were subsequently isolated and propagated into clonal cell lines in 96 well dishes.

3.1.5 Transduction with lentiviruses

Cells to be transduced were seeded in the appropriate fully supplemented medium on 6-well plates to reach 70-80% confluency prior to transduction. The medium was subsequently aspirated and replaced with 1 ml of fresh fully supplemented culture medium containing 20 mM HEPES and 8 μg/ml Polybrene (Merck Millipore) per well. In the case of hMSCs, no heparin was added to avoid inactivation of virus particles (GUIBINGA *et al.*, 2002). Next, 500 μl of each lentivirus containing supernatant was added dropwise and plates were swirled. Subsequently, plates were centrifuged at 2,000 rpm and RT for 45 min and incubated for 24 h at 37°C. This transduction procedure was repeated after 24 h. For fluorescence-activated cell sorting (FACS), culture medium was removed and cells were washed twice with sterile PBS. Next, cells were harvested as described and resuspended in 500 μl PBS containing 1% FBS and transferred into a sterile 5 ml *Falcon*® Round-Bottom Polystyrene Tube (Stemcell) through the cell strainer cap. Samples were sorted at the in-house LIV FACS facility on a *BD FACS Aria Fusion* (BD Biosciences) flow cytometer based on their venus- and BFP-expression. Finally, sorted cells were pelleted and washed after cytometry before they were seeded on the appropriate culture plate. Cultivation medium was changed every day for the next two weeks.

3.1.6 Cryopreservation and re-cultivation

For long-term storage, cells were harvested as described above and resuspended in FBS (human tumor cell lines and BRK) or human serum (for hMSC, Albiomin 5%, Biotest) containing 10% DMSO. The cell suspension was transferred to *CryoPure* tubes (Sarstedt) and slowly frozen to -80°C utilizing a freezing container filled with 100% isopropanol (*Mr. Frosty*, Nalgene). Frozen cells were either stored short-term at -80°C or long term at -170°C in liquid nitrogen tanks. For re-cultivation of BRK cells or human tumor cell lines, frozen cells were thawed shortly in a water bath at 37°C, immediately added to pre-warmed fully supplemented cultivation medium in a 15 ml reaction tube, centrifuged at 2,000 rpm for 3 min (*Multifuge* 3S-R, Heraeus) before the medium was aspirated. Cells were then resuspended in an appropriate amount of culture medium and seeded on a

suitable dish. In the case of hMSC, cells were not centrifuged after thawing to avoid unnecessary stress, but added directly into 100 ml of fully supplemented MEM in a T175 flask to dilute the DMSO.

3.2 Bacteria culture techniques

3.2.1 Culture and storage

Liquid *Escherichia coli* DH5α cultures were grown in sterile lysogeny broth (LB)-media supplemented with ampicillin (100 μg/ml) in an *Inova 4000 Incubator* (New Brunswick) at 30°C (for bacteria containing E1B-55K constructs) or 37°C, by shaking at 200-400 rpm until the required density was achieved. Likewise, transformed bacteria or bacteria from glycerol stocks were grown by plating on LB agar plates containing 0.015% agar (w/v) and ampicillin (100 μg/ml), incubated in *Heraeus incubators* at 30°C or 37°C for 16 up to 48 h. Liquid or solid cultures were stored at 4°C for short time frames. Glycerol stocks from single colony bacteria were generated by pelleting 1 ml liquid culture for 3 min at 8,000 rpm and 4°C (*5417R* Centrifuge, Eppendorf). For long term storage, the bacterial pellet was resuspended in 1 ml LB-media, mixed 1:1 (v/v) with sterile glycerol (87%) and subsequently transferred into *CryoPure tubes* (Sarstedt) at -80°C.

3.2.2 Heat-shock chemical transformation

For all bacterial transformation experiments, aliquots of chemically competent *Escherichia coli* DH5α were thawed on ice. The suspension was transferred into a pre-cooled polypropylene culture tube on ice before adding the appropriate amount of the ligation mixture, PCR product or plasmid DNA. This mixture was incubated on ice for 30 min prior to heat shock at 42°C for 45 s in a water bath. Afterwards, the tube was rested on ice for 2 min. Subsequently, 1 ml of 37°C antibiotics-free LB-media was added and the suspension was transferred into a clean 1.5 ml reaction tube. Recovery was performed at 30°C or 37°C and 500 rpm for 1.5 h in a *Thermomixer comfort* (Eppendorf). Finally, 200 μl of the suspension was plated on LB agar plates containing ampicillin (100 μg/ml). The remaining mixture was centrifuged at 8,000 rpm for 3 min and the pellet was resuspended in 200 μl supernatant, plated on an additional LB agar plate containing ampicillin (100 μg/ml) and incubated overnight at 30°C or 37°C.

3.3 Nucleic acid techniques

3.3.1 Isolation of plasmid DNA from *E. coli*

For the Isolation of plasmid DNA from transformed DH5α, 500 ml of LB-medium was supplemented with the appropriate antibiotic, inoculated with a single bacteria colony and incubated for 16-24 h at 30°C or 37°C and 200 rpm (*Inova 4000 Incubator*, New Brunswick). Bacteria were centrifuged for 10 min at 6,000 rpm and 4°C (*Avanti-JE* centrifuge, Beckmann Coulter). The supernatant was discarded and the plasmid DNA was isolated using the *QIAGEN Plasmid Maxi* Kit. All subsequent centrifugation steps were performed in a Megafuge 1.0 (Heraeus) at 4°C. First, the pellet was resuspended in 10 ml of buffer P1, which was supplemented with RNAse A and LyseBlue reagent and transferred to a 50 ml reaction tube. Next, 10 ml of buffer P2 was added to lyse bacteria and the sample was mixed thoroughly by inverting multiple times. After 5 min, 10 ml of buffer P3 was added and the sample was mixed and incubated for 20 min at 4°C. Cellular debris was pelleted at 4,000 rpm and 4°C for 60 min and the supernatant was added onto a column that was previously equilibrated with 10 ml of buffer QBT. Afterwards, two washing steps were performed by applying 30 ml of wash buffer QC twice. The plasmid DNA was subsequently eluted with 15 ml of buffer QF into a 50 ml reaction tube. DNA was precipitated with 10.5 ml 100% isopropanol and centrifuged at 4,000 rpm and 4°C for 30 min (*Avanti-JE* centrifuge, Beckmann Coulter). The supernatant was discarded, the DNA pellet was washed with 5 ml 75% ethanol and centrifuged at 4,000 rpm and 4°C for 15 min (*Avanti-JE* centrifuge, Beckmann Coulter). Lastly, the ethanol was decanted and the DNA was air-dried before it was dissolved in a suitable volume of ddH_2O. For analytical purposes, the plasmid preparation was performed as described above but DNA-binding columns were omitted. Here, up to 2 ml of a liquid culture were pelleted at 8,000 rpm for 3 min (*5417R* centrifuge, Eppendorf). Briefly, the bacterial pellet was resuspended in 300 μl buffer P1. Next, 300 μl of buffer P2 was added and the samples were mixed well. After 5 min incubation at RT, 300 μl of buffer P3 was added to the sample and the mixture was incubated for min at RT before it was centrifuged at 13,000 rpm for 15 min (*5417R* centrifuge, Eppendorf). 750 μl of the supernatant was immediately transferred to a new 1.5 ml reaction tube with 600 μl isopropanol. The sample was inverted several times and the DNA was precipitated at 13,000 rpm for 15 min (*5417R* centrifuge, Eppendorf). Afterwards, the isopropanol was discarded and the pellet was washed with 500 μl of 75% ethanol and centrifuged as described before. Lastly, the ethanol was decanted, the pellet dried at 37°C and the DNA

was eluted in 30-50 μl ddH_2O. DNA concentration was measured on a *NanoDrop 1000* (Thermo Fisher Scientific) spectrometer.

3.3.2 Polymerase chain reaction

The *PfuUltra II Fusion* high-fidelity DNA Polymerase (Agilent) was used in all polymerase chain reaction (PCR) reactions. This method was used to (i) amplify a DNA sequence from a specific vector, (ii) insert additional restriction sites to a plasmid or (iii) for site-directed mutagenesis. In all cases, a 50 μl PCR reaction volume was prepared by mixing 5-30 ng of DNA template or vector DNA, 2 μl of forward and reverse primers (final concentration of 0.4 μM), 1-2 μl of a dNTP mix (containing 1 mM of each nucleotide), 5 μl of *PfuUltra II* 10x polymerase buffer and 1 μl of the *PfuUltra II Fusion* polymerase in a 0.2 mm thin-walled PCR tube (Biozym) with ddH_2O. Cycling was performed on a *MJ MiniTM* gradient thermal cycler (BioRad). For site-directed mutagenesis, 5-10 μl PCR reaction were analyzed on an agarose gel and the residual reaction was incubated with 1 μl DpnI (NEB) for 2 h at 37°C to digest methylated input template plasmid DNA. 10 μl of the digest were used for transformation of chemical competent DH5α. The PCR cycling protocol is depicted in the following table:

Cycling	**Temperature**	**Step**
2 min	95°C	Preheating
1 min	95°C	**DNA denaturation**
10 s - 1 min	55-70°C	**Annealing**
1 min / 1,000 bp	72°C	**Elongation**
10 min	68°C	Final extension
Endless	4°C	Storage

Marked(**bold**) steps were performed for 30 cycles

PCR was purified with the help of the *QIAquick PCR Purification* kit (QIAGEN) according to the manufacturer's protocol to remove PCR reagents from the product. Therefore, the sample was diluted with five volumes of buffer PB, applied to the spin column placed on a 2 ml collection tube and centrifuged at 13,000 rpm and RT for 1 min (*5417R* centrifuge, Eppendorf). The column-bound DNA was washed with 750 μl buffer PE and subsequently centrifuged again. The flow-through was discarded and the spin column was centrifuged again to remove residual buffer PE. In the last step, 30 μl ddH_2O was added to the center of the membrane, incubated for 4 min and DNA was collected into a 1.5 ml reaction tube through a final centrifugation step.

3.3.3 DNA agarose gel electrophoresis

Analytical, as well as preparative, agarose gels were prepared with agarose powder (*Seakem LE Agarose*, Biozym), diluted in 1x TBE buffer to a concentration of 1% (w/v) by bringing the mixture to a boil utilizing a microwave. The agarose solution was cooled to about 50-60°C in cold water and 100 ml were poured into a gel tray equipped with a well comb. Ethidium bromide was added to the liquid agarose in a final concentration of 0.2-0.5 μg/ml. The agarose gel solidified at RT for about 30 min before it was placed into an electrophoresis chamber filled with 1x TBE. DNA was mixed 5:1 (v/v) with 6x Purple DNA Loading Dye (NEB) before it was loaded into the wells. The DNA was separated at 5-10 V/cm gel length. 1 kb and 100 bp DNA ladders (NEB) were used as references. Analytical agarose gels were documented using the *G:Box system UV-Transilluminator* (SynGene) at a wavelength of 312 nm with the *GeneTools* software (SynGene). DNA fragments of interest were separated in preparative agarose gels, visualized at 365 nm on an UV table (*GelVue UV Transilluminator*, Syngene) and isolated with sterile scalpels. DNA was subsequently extracted from agarose using the *QIAquick* gel extraction kit (QIAGEN) following the manufacturers' protocol. The gel slice was weighed and dissolved in three volumes of buffer QG to one volume gel (v/w) and incubated at 50°C and 800 rpm for 10 min (*Thermomixer comfort*, Eppendorf). One gel volume of isopropanol (v/w) was added, the sample was mixed and applied to a *QIAquick* spin column in a 2 ml collection tube. The spin column was centrifuged at 13,000 rpm for 1 min (*5417R* centrifuge, Eppendorf) and the flow-through was discarded. Next, 750 μl of buffer PE was applied to the column and centrifugation followed as previously described (under 3.3.2). The flow-through was discarded and the spin column was centrifuged once again to remove residual wash buffer. Finally, the column was transferred to a clean 1.5 ml reaction tube, 30 μl ddH_2O were applied directly onto the membrane followed by a 5 min incubation period and a last centrifugation step at 13,000 rpm for 1 min (*5417R* centrifuge, Eppendorf) to collect the DNA.

3.3.4 Restriction-site mutagenesis

Restriction endonucleases were utilized to cleave plasmid DNA or DNA fragments designated either for (i) analytical testing or (ii) cloning and subsequent ligation. To investigate (i) DNA from plasmid preparation, 500-1000 ng of DNA were incubated with 5 U of the appropriate enzyme and corresponding buffer (both acquired from NEB) in a total volume of 20 μl. To examine the band pattern of restricted DNA, samples were mixed with load-

ing dye and agarose gel electrophoresis was performed afterwards. In case of (ii), a 50 μl reaction volume was prepared by mixing 5-10 U of the desired enzyme per μg DNA with the corresponding buffer and ddH_2O. The digest occurred for 1-2 hours or overnight at 37°C in a thermal block (*Thermomixer comfort*, Eppendorf). Restriction digest was then followed by preparative gel electrophoresis (see 3.3.3).

3.3.5 Plasmid ligation

For ligation, 50 ng of vector DNA were usually added to 150 ng of insert. Generally, the insert was combined at a 3:1 molar excess over vector DNA, however other ratios were tested in case the recommended ratio did not work initially. The vector DNA could additionally be dephosphorylated (according to manufacturer protocol using the *Rapid DNA Ligation* Kit from Roche) to decrease the amount of re-ligation of empty vector in the following steps. The calculated amounts of vector and insert DNA were mixed with 2 μl 5x DNA Dilution buffer to a total volume of 10 μl. The solution was mixed shortly before 10 μl T4 Ligation Buffer and 1 μl T4 DNA-Ligase were added. Incubation occurred at RT for 5-30 min. The mixture was either used directly for transformation of competent DH5α or could be stored without heat inactivation at 4°C for later transformation experiments (see 3.2.2.).

3.3.6 RNA extraction

RNA was isolated from harvested mammalian cells on ice under a fume hood, with all surfaces, gloves and pipettes previously being thoroughly cleaned using *RNAse ZAP* (Sigma Aldrich). Cell pellets were resuspended in 1 ml *TRIzol* Reagent (Ambion by Life Technologies), mixed with 200 μl chloroform, incubated for 2-3 min on ice and centrifuged at 12,000 g and 4°C for 15 min (*5417R* centrifuge, Eppendorf). The aqueous phase containing the extracted RNA was removed carefully and transferred to a fresh 1.5 ml reaction tube, mixed with 600 μl isopropanol (optional addition of 3 μl of 5 mg/ml glycogen to increase yield) and incubated for 5-10 min of ice to precipitate the RNA. After a second centrifugation step at 12,000 rpm and 4°C for 15 min (*5417R* centrifuge, Eppendorf), the supernatant was discarded and the visible RNA pellet was washed with 500 μl 75% ethanol. Subsequently, the sample was centrifuged at 7,500 g and 4°C for 15 min (*5417R* centrifuge, Eppendorf). The supernatant was discarded again, residual ethanol was removed and the RNA was air-dried for 5-10 min. RNA was finally eluted in an appropriate amount of nuclease-free water (QIAGEN) depending on the pellet size.

The eluted RNA was stored long-term at -80°C to avoid degradation.

3.3.7 Cross-linking chromatin immunoprecipitation (ChIP)

The cross-linking chromatin immunoprecipitation (abbreviated as X-ChIP) method was utilized to enrich DNA bound by specific transcription factors or histones. In case of direct interaction between target protein (e.g. p53 or TEAD4) and DNA, only formaldehyde was used as a fixation agent. When E1B-55K was the target protein of the experiment, the homobifunctional cross-linking agent di(N-succinimidyl) glutarate (DSG) which contains amine-reactive N-hydroxysuccinimide ester groups was used before the formaldehyde fixation (abbreviated as dual X-ChIP, method published in TIAN *et al.* (2012)).

Cross-linking

Freshly harvested cells were washed three times with autoclaved 1x PBS, supplemented with phenylmethylsulfonyl fluoride (PMSF), a serine protease inhibitor, to a final concentration of 1% (v/v, Sigma Aldrich). The cell number was determined before fixation steps and cells were transferred into 15 ml falcons. For dual cross-linking, half of the cells were resuspended in 1x PBS containing a final concentration of 2 mM DSG, dissolved in DMSO directly before the experiment, and subsequently incubated by rotation for 30 min at 15 rpm and RT. Afterwards, cells were pelleted at 2,000 rpm and RT for 3 min (*Multifuge* 3S-R, Heraeus), fixation solution was aspirated and cells were washed twice with RT 1x PBS 1% PMSF (v/v) before they were subjected to formaldehyde fixation as described below. For formaldehyde fixation, cells were pelleted and resuspended in 10 ml PBS supplemented with 270 μl of 37% formaldehyde to a final concentration of 1% (v/v). Reaction tubes were rotated at 15 rpm and RT for exactly 10 min. Cross-linking was quenched immediately by addition of 1 ml of 1.25 M glycine solution, followed by rotation for 5 min at 15 rpm and RT. Afterwards, the fixed cells were pelleted at 2,000 rpm and 4°C for 3 min and the supernatant was aspirated. Cells were washed twice with ice-cold 1x PBS supplemented with 1% PMSF (v/v) before they were resuspended in 500 μl TE buffer and transferred into a fresh 1.5 ml reaction tube. Fixed cells were stored in TE buffer at -20°C until chromatin isolation.

Chromatin isolation and shearing

Previously fixed cells were thawed on ice and pelleted at 500 g and 4°C for 5 min to remove the TE buffer (*5417R* centrifuge, Eppendorf). ChIP buffer I to III were supplemented

with protease inhibitors (final concentration: 1 mM PMSF, 1 μg/ml leupeptin, 1 μg/ml pepstatin and 10 U/ml aprotinin). The cellular membranes were lysed by resuspension in 1 ml of 4°C buffer I, which contains the mild, non-ionic detergents NP40 and Triton X-100, by rotation for 10 min at 4°C. Nuclei were pelleted by centrifuging for 5 min at 2,000 g and 4°C (*5417R* centrifuge, Eppendorf). The supernatant was discarded and the nuclei were washed in 1 ml of 4°C buffer II. Samples were rotated and centrifuged as described and nuclei were vigorously resuspended in 1 ml of RT Buffer III for lysis of the nuclear membrane. In the last step, isolated chromatin was transferred to 15 ml conical reaction tubes and subjected to shearing using a *Bioruptor* (Diagenode) and sonicated for 15 cycles (20 cycles for dual X-ChIP) in a refrigerated sonication bath (30 second intervals on/off at highest setting). The Bioruptor sonication system uses ultrasound to create focused mechanical stress to shear chromatin. Ultrasound waves generated by the ultrasound transducer below the rotating samples pass through the samples resulting in expansion and contraction of the liquid. During this expansion, bubbles are formed between molecules in a process called cavitation. These bubbles implode and thereby create focused shearing forces that break apart biomolecules such as nucleic acids. After this process, the sonicated sample was transferred to a 1.5 ml reaction tube, 100 μl 10% Triton X-100 was added and samples were centrifuged at 20,000 g and 4°C for 10 min and the chromatin containing supernatant was transferred to a fresh 1.5 ml reaction tube. At this point, fragmented chromatin from 2.5 x 10^5 cells was taken from each sample and stored at -20°C which would later be used as the input control.

Preclearing and preparation of magnetic beads

50 μl *Protein A/G magnetic beads* (Life Technologies) are required for preclearing of each individual chromatin and for each individual subsequent immunoprecipitation reaction. First, the storage solution of the beads was discarded by utilizing a magnetic stand. The beads were washed three times with 1 ml dilution buffer, supplemented with protease inhibitors (final concentration: 1 mM PMSF, 1 μg/ml leupeptin, 1 μg/ml pepstatin and 10 U/ml aprotinin) before they were blocked overnight rotating in dilution buffer and 50 μl of 20 mg/ml BSA at 4°C and washed an additional time pre-use before resuspension in the original volume of dilution buffer. For preclearing, 50 μl of washed magnetic beads were added to each sheared chromatin sample and rotated for 30 min at 4°C. Afterwards, the samples were placed on the magnetic rack and the supernatant was transferred to a fresh 1.5 ml reaction tube.

Immunoprecipitation

For immunoprecipitation, pre-cleared chromatin of 1 x 10^6 cells was mixed 1:10 in dilution buffer (v/v), supplemented with all four protease inhibitors and 4 μg of the target-specific antibody or the corresponding mouse or rabbit IgG control and subsequently incubated by overnight rotation at 4°C. In the next step, 50 μl of the blocked *Protein A/G magnetic beads* were added for 1 h at 4°C while rotating to precipitate bead-antibody-chromatin complexes. Afterwards, the samples were placed on a magnetic stand and the supernatant containing unbound DNA was carefully aspirated. The beads were washed with 1 ml of low-salt buffer and transferred into a fresh 1.5 ml reaction tube. This is followed by two subsequent washing steps using high-salt- and $LiCl_2$ wash buffers supplemented with the above mentioned protease inhibitors. Afterwards, beads were washed twice with TE buffer and transferred to a fresh 1.5 ml reaction tube. To elute the chromatin, beads were incubated in 210 μl elution buffer containing 1% sodium dodecyl sulfate (SDS) (w/v) on a thermal block at 65°C and 1,000 rpm for 30 min. The chromatin containing supernatant was removed by placing the reaction tube in a magnetic stand and transferred into a fresh 1.5 ml reaction tube with 8 μl of 5 M NaCl.

De-cross-linking and DNA purification

The previously frozen input samples were thawed at this step and diluted with elution buffer to a final volume of 210 μl and 8 μl of 5 M NaCl were added afterwards. The chromatin input as well as immunoprecipitation samples were incubated overnight at 65°C and 1,000 rpm in a thermal mixer for de-cross-linking of DNA and protein complexes. Next, potential RNA contaminants were degraded by adding 200 μl TE buffer containing 8.8 μl of (9 mg/ml RNAse A and subsequent incubation at 37°C for 2 h in a thermal shaker (*Thermomixer comfort*, Eppendorf). In the following step, 7 μl of $CaCl_2$ buffer and 8 μl of 10 mg/ml Proteinase K were added and samples were incubated at 55°C and 1,000 rpm in a thermal shaker. *Heavy Gel Phase-Lock* tubes (Quantabio) were prepared by initial centrifugation at 18.000 g and 4°C for 1 min. The samples were combined with 400 μl of *UltraPure* phenol:chloroform:isoamylalcohol (PCI, 25:24:1; v/v; Carl Roth), mixed by vigorously vortexing for 30 seconds and centrifuged for 4 min at 18.000 g at 4°C. Another 400 μl of PCI solution was added to the same phase-lock tube, mixed and centrifuged again. Next, 400 μl of chloroform was added, the sample was vigorously vortexed and the centrifugation step repeated. In the final step, the aqueous phase above the gel layer was carefully transferred into a new 1.5 ml reaction tube without disturbing the gel layer.

DNA was precipitated by adding 24 μl of 5 M NaCl, 3 μl of 5 mg/ml glycogen and 1055 μl of -20°C 100% ethanol. The DNA was precipitated overnight at -80°C before DNA was pelleted the next day by centrifuging for 15 min at 20,000 g and 4°C. The supernatant was discarded and the pellet was washed with 500 μl of 70% ethanol by gently shaking the tube until the pellet was detached from the tube wall. The samples were centrifuged like in the step before, the supernatant was aspirated and the pellet was dried shortly in a heat block at 37°C. The DNA was collected in 55 μl of EB buffer and dissolved at 37°C for 1-2 h. Samples were stored at -20°C for several months or until next-generation sequencing (NGS). To verify the chromatin shearing, 10 μl of input samples was applied to a 1% agarose gel to validate fragmentation of the chromatin.

3.3.8 Native MNase ChIP

The native MNase ChIP method (described in BRIND'AMOUR *et al.* (2015)) was utilized to enrich non-cross-linked and unmasked DNA bound by histones, targeting specific PTM of the proteins, like histone H3 lysine 4 trimethylation (H3K4me3), associated with transcriptionally active gene regions.

Chromatin isolation and MNase digestion

All steps of the following protocol were performed in 1.5 ml DNA/RNA *lobind tubes* (Eppendorf). Cells used for native ChIP were harvested as described in the precious section and the number of viable cells was determined by trypan blue staining. A total of 1 x 10^6 cells were transferred to a 1.5 ml reaction tube, washed in 1 ml of autoclaved 1x PBS containing protease inhibitors (final concentration: 1 mM PMSF, 10 U/ml aprotinin, 1 μg/ml leupeptin, 1 μg/ml pepstatin and 5 mM sodium butyrate) and pelleted at 500 g and 4°C for 5 min. The supernatant was discarded and the cell pellet was flash frozen in liquid nitrogen and stored at -80°C or directly used for DNA extraction. These cell pellets were thawed on ice and resuspended in 50 μl of complete nuclear isolation buffer (containing all previously stated protease and HDAC inhibitors) by pipetting up and down for 15-20 times and transferred on a RT reaction tube rack. The MNAse Master Mix was prepared on ice with 6 μl 10x MNase Buffer (NEB), 0.44 μl 200 mM dithiothreitol (DTT), 0.6 μl MNase (final concentration: 2 U/μl; NEB) and 2.96 μl nuclease-free water (QIAGEN). 10 μl of the master mix was added to each sample and incubated exactly for 7.5 min at 37°C. The enzymatic reaction was stopped by adding 6.6 μl 100 mM EDTA, followed by addition of 6.6 μl Triton-Sodium-deoxycholate-solution. Samples were subsequently incubated on

ice for 15 min before they were vortexed on medium settings for 30 s. Complete immunoprecipitation buffer was added to a final volume of 200 μl and the sample was rotated for 1 h at 4°C and then vortexed for 30 s. Next, 10% of the sample was transferred to a 1.5 ml reaction tube and used as input control. 10% volume of 10% SDS solution were added, the sample was mixed and elution buffer was added to a final volume of 100 μl. Input samples were stored at -20°C for later usage.

Preclearing and preparation of magnetic beads

10 μl *Protein A/G magnetic beads* (Life Technologies) are required for preclearing of each individual chromatin and for each individual subsequent immunoprecipitation reaction. First, the storage solution of the beads was discarded by utilizing a magnetic stand. The beads were washed three times with 1 ml complete immunoprecipitation buffer. For preclearing, 10 μl of washed magnetic beads were added to each digested chromatin sample corresponding of 1 x 10^5 cells and rotated for 6 h at 4°C. Afterwards, the samples were placed on the magnetic rack and the supernatant was transferred to a fresh 1.5 ml reaction tube.

Immunoprecipitation

At the same time, antibody-bead complexes were prepared by adding 100 μl of complete immunoprecipitation buffer buffer, 10 μl washed beads and 1 μg of the corresponding α-histone PTM antibody to a 1.5 ml reaction tube and rotated for 6 h at 4°C. In the next step, antibody-bead complexes were placed on a magnetic rack and the supernatant was discarded. The pre-cleared chromatin was also placed in a magnetic rack and chromatin-containing supernatant was transferred to the antibody-bead complexes. The samples were rotated overnight at 4°C before the reaction tubes were placed on a magnetic rack and the supernatant was discarded. The beads were resuspended in 200 μl complete immunoprecipitation buffer, transferred to a fresh 1.5 ml reaction tube and washed twice with 200 μl low-salt, followed by 200 μl high-salt wash buffer. Afterwards, beads were transferred to a clean 1.5 ml reaction tube and pulse-spinned shortly to completely remove all residual wash buffer by pipetting.

DNA purification

30 μl of freshly prepared ChIP Elution buffer were added and DNA was eluted for 1.5 h at 65°C and 1,000 rpm on a thermal mixer (*Thermomixer comfort*, Eppendorf). Samples were

placed on a magnetic stand and the chromatin-containing supernatant was transferred to a clean 1.5 ml reaction tube. Lastly, beads were washed once with 70 μl of elution buffer and combined with the eluted chromatin. *Heavy Gel Phase-Lock* tubes (Quantabio) were prepared by initial centrifugation at 18,000 g and 4°C for 1 min. The samples were combined with 100 μl of *UltraPure* phenol:chloroform:isoamylalcohol (PCI, 25:24:1; v/v; Carl Roth), mixed by vigorously vortexing for 30 seconds and centrifuged for 4 min at 13,000 g at RT. In the final step, the aqueous phase above the gel layer was carefully transferred into a new 1.5 ml reaction tube without disturbing the gel layer. DNA was precipitated by adding 6 μl of 5 M NaCl, 1 μl of 5 mg/ml glycogen and 275 μl of -20°C 100% ethanol. The DNA was precipitated overnight at -80°C before DNA was pelleted the next day by centrifuging for 15 min at 20,000 g and 4°C. The supernatant was discarded and the pellet was washed with 500 μl of 70% ethanol by gently shaking the tube until the pellet was detached from the tube wall. The samples were centrifuged like in the step before, the supernatant was aspirated and the pellet was dried shortly in a heat block at 37°C. The DNA was collected in 30 μl of EB buffer and dissolved at 37°C for 1-2 h. Samples were stored at -20°C for several months or until NGS. To verify the chromatin digestion, 5 μl of input samples was applied to a 1% agarose gel.

3.3.9 Real-time polymerase chain reaction

To validate the successful enrichment of DNA from fixation or native ChIP assays, real-time polymerase chain reaction (qPCR) was performed with purified DNA and oligonucleotides that amplify known target transcription factor binding sites or regions that are expected to harbour a certain histone PTM. Enrichment was quantified by comparison of the ChIP DNA to the corresponding input sample, thereby creating the %-input value. Additionally, negative control sites, which are not bound by the target factor, were amplified to determine the specificity of the antibodies and the background. In these experiments, 5 μl of 2x *SYBR Green* (Meridian Bioscience) was mixed with 0.6 μl primer mix (final concentration: 10 pmol of 1:1 (v/v) forward and reverse primer mix), 1.5 μl of the specific ChIP experiment and 2.9 μl nuclease-free water (QIAGEN). The 10 μl reaction volume was transferred to 0.1 ml Strip Tube (LTF Labortechnik) and qPCR was performed on the *Rotor Gene 6000* (Corbett Life Science). The following settings were used for cycling:

Cycling	Temperature	Step
10 s	95°C	**DNA denaturation**
40 s	57°C	**Annealing**
10 s	72°C	**Elongation**

Marked(**bold**) steps were performed for 40 cycles

3.3.10 ChIP-seq library preparation

The library preparation with isolated ChIP-DNA was performed by the in-house sequencing facility using the *NEXTflex ChIP-seq kit* (PerkinElmer Applied Genomics) according to manufacturer's instructions.

3.3.11 RNA-seq library preparation

The library preparation with extracted RNA was performed by the in-house sequencing facility using the *NEXTflex Rapid Directional qRNA-seq kit* (PerkinElmer Applied Genomics) according to manufacturer's instructions.

3.4 Protein techniques

3.4.1 Lysate preparation for protein extraction

Cells were harvested as described above and cell pellets were resuspended in an appropriate amount of ice-cold fully supplemented RIPA buffer (final concentration: 1 mM PMSF, 10 U/ml aprotinin, 1 μg/ml leupeptin and 1 μg/ml pepstatin). Resuspended cells were transferred into pre-cooled 1.5 ml reaction tubes and incubated on ice for 30 min with 30 s of vortexing every 10 min. Sonication was performed for 30 s at 4°C (30 pulses, output 0.8, 0.8 pulses per second) using a *Branson Sonifier 450*. To remove insoluble proteins and cellular debris, the samples were centrifuged at 11,000 rpm and 4°C for 3 min (*5417R* Centrifuge, Eppendorf). The supernatant was transferred to a fresh 1.5 ml reaction tube and stored on ice. Afterwards, the protein concentration was determined by spectrophotometry using a Bradford protein assay described in the next section. In the final step, all lysates were adjusted to the same concentration by adding ddH_2O and were denaturated by adding 5:1 5x Laemmli buffer (v/v) and boiling at 95°C for 10 min. These RIPA-extracts were stored at -20°C for subsequent western blotting. In case of protein extraction targeting specifically histones, a high-salt lysis buffer was used instead of the

RIPA buffer (adapted from SHECHTER *et al.* (2007)).

3.4.2 Quantification of protein concentration

Protein concentrations of lysates were determined using the bradford reagent based protein assay. This colorimetric assay for measuring soluble protein concentrations is based on the bradford dye-binding method. Samples are mixed with *Coomassie Brilliant Blue G-250* dye (BioRad), which leads to a measurable shift of the absorbance maximum of this dye from 465 to 595 nm. For these experiments, 800 μl of ddH_2O are mixed with 1 μl of protein extract and supplemented with 200 μl *Bradford Reagent* (BioRad) in a semi-micro cuvette (Sarstedt). Samples were prepared in duplicates and incubated for 5 min at RT prior to measurement on a *SmartSpec Plus spectrophotometer* (BioRad). Protein concentrations were determined by interpolation from a freshly prepared and recorded standard curve. For this purpose, five standards ranging from 1- 16 μg/μl were created using bovine serum albumin (BSA, NEB), mixed with ddH_2O.

3.4.3 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

Denatured protein samples were separated by SDS-polyacrylamide gel electrophoresis (PAGE), where SDS is providing a negative net charge to the denaturated proteins, allowing a separation solely by molecular mass instead of charge. SDS gels are composed of a 5% stacking gel with a pH of 6.8 pH and a resolving gel with a pH of 8.8. The resolving gel density depended on the molecular weight of the proteins of interest and usually ranged from 10-15% in this work. SDS gels were prepared with a 30% acrylamide stock solution (37.5:1 *Rotiphorese Gel 30*, Carl Roth), diluted to a final concentration of 10-15% with ddH_2O and 1.5 M Tris/HCl buffer. Addition of tetramethylethylenediamine (TEMED) and ammonium persulfate (APS) will catalyze the polymerization of the acrylamide gel. SDS-PAGE was performed in TGS-buffer at 10-20 mA/gel using the *Multigel* SDS-PAGE system (Biometra). The *Page RulerTM Plus Prestained Protein ladder* (Pierce), ranging from 10-200 kDa was used as a reference for protein sizes.

3.4.4 Western blot

Proteins separated by SDS-PAGE were subsequently transferred onto nitrocellulose membranes (pore size of 0.45 μm; Amersham) using the *Trans-Blot Electrophoretic Transfer Cell System* (BioRad). The electrophoretic transfer of proteins was performed in a *Trans-Blot Electrophoretic Transfer* tank filled with towbin buffer by applying 400 mA for 90 min.

After the transfer was completed, the membranes were transferred to 1x PBS supplemented with 5% non-fat dry milk (w/v; Frema) for 2 h or overnight at 4°C. Membranes were subsequently washed three times with PBS-Tween for 10 min prior to the incubation with primary antibody dilutions at 4°C for 2h. Primary antibodies were diluted individually and recycled after each incubation. Afterwards, membranes were washed again and incubated with the corresponding horseradish peroxidase-conjugated secondary antibodies for 2 h at 4°C. Secondary antibodies were diluted 1:10,000 in PBS-Tween (v/v) supplemented with 3% non-fat dry milk (w/v). In the next step, secondary antibody dilutions were discarded and membranes were washed three final times before they were incubated with the *SuperSignalTM West Pico Plus Chemiluminescent Substrate* (Thermo Fischer Scientific). Here, the horseradish peroxidase catalyzes the oxidation of luminol in the presence of a hydrogen peroxide which acts as an oxidizing agent, producing chemiluminescence. The chemiluminescence was visualized on *blue-sensitive X-ray films* (Super RX-N, Fujifilm) and developed using a *GBX Developer* (Kodak). X-ray films were scanned and figures were prepared with *Photoshop CS5* (Adobe) and *Inkscape*.

3.5 Microscopy

3.5.1 Immunofluorescence

Cells were seeded onto 18x18 mm glass coverslips in 6-well plates or onto *µ-Dish* (35 mm, high, ibidi). After 16-24 h, cells at 80% confluency were fixed with 4% paraformaldehyde (PFA) solution for 15 min at RT. PFA was subsequently carefully aspirated and cells were washed three times with 1x PBS for 5 min. In the next step, cells were permeabilized using a 0.5% Triton X-100 solution in PBS (v/v) for 10 min before they were washed again three times with 1x PBS for 5 min and blocked by incubation with TBS-BG solution for 60 min at 4°C, followed by a final three times wash step with 1x PBS. Primary antibodies were diluted in sterile PBS and coverslips were incubated with the solution at RT for 1 h. The samples were washed again as described before and incubated for 1 h at RT and in the dark with the corresponding *Alexa-Fluor* (Thermo fischer scientifc) secondary antibody dilution, which was created by diluting antibodies 1:400 in 1x PBS. Here, 4,6-diamidino-2-phenylindole (DAPI) was additionally diluted 1:10,000 to visualize dsDNA. In the final step, coverslips were washed and mounted onto clean object slides (Menzel-Gläser) with 10 µl *fluorescence mounting medium* (Agilent Dako). Object slides were dried overnight at 4°C in the dark. All samples were stored at 4°C after preparation. Immunoflourescence

samples were imaged using the Nikon A1 (confocal microscope with a Nikon Ti2 frame) and subsequently analyzed using the Fiji software (SCHINDELIN *et al.*, 2012).

3.6 Bioinformatic analysis

Any of the following tools were used in an Ubuntu LTS 20.04 operating system through bash shell commands. The softwares were obtained from the *bioconda* (DALE *et al.*, 2018) software distribution tool (version 4.11.0) and invoked from within different *bioconda* environments.

3.6.1 Fastqc and multiqc high-throughput sequencing data quality control

The *fastqc* software (version 0.11.9) was used to perform quality control on all ChIP-seq and mRNA-seq raw high-throughput sequencing .fastq files after demultiplexing. The software was invoked with standard settings and followed by the *multiqc* software (version 1.11) to create a .html file to compare the individual files and identify potential outliers.

3.6.2 Bowtie2 read alignment to reference genomes

The *Bowtie2* tool (version 2.4.4) was used to align all ChIP-seq sequencing data after quality control to its corresponding genome (ensembl version 105), the mRatBN7.2 toplevel assembly for *rattus norvegicus* and GRCh38 primary assembly for *homo sapiens*. In the first step, a bowtie2 index was created with:

```
$ conda run bowtie2-build --threads 10 GENOME.fa
```

In the next step, the individual .fastq files were aligned against their respective genome index while also removing the first three adapter nucleotides from the 5' end of the fragments. This process creates sequence alignment map (sam) files as its output.

```
$ conda run bowtie2 --trim5 3 -p 10 -x GENOME [.fastq]
```

3.6.3 Samtools interaction and conversion of data

The *samtools* software (version 1.7) was used to perform lossless conversion of the .sam file format into its sorted and compressed binary alignment map (bam) file format by:

```
$ conda run samtools sort -@ 12 -O bam [.fastq]
```

In the next step, all PCR duplicates were marked and removed by:

```
$ conda run samtools markdup -r -s -@ 12 [.bam]
```

Followed by removing all reads with a mapping quality (MAPQ) score below 30 by:

```
$ conda run samtools view -@ 12 -q 30 [.bam]
```

In the final step, an index was created, which is needed by some downstream software tools.

```
$ conda run samtools index [.bam]
```

3.6.4 Macs2 narrow peak calling

The model-based analysis for ChIP-seq (macs) software (2.2.7.1) was used for any peak-finding experiment targeting transcription factors or narrow histone PTM marks, like H3K4me3 or H3K27ac, with the following command:

```
$ conda run macs2 callpeak -t [.bam] -c [input.bam] -f BAM
    -g [genome size] -q 1e-2 --call-summits
```

3.6.5 SICER/epic2 broad peak calling

The epic2 implementation (version 0.0.52) of the SICER software was used to identify diffuse and broad differential histone PTM regions, e.g. heterochromatin markers H3K9me3 and H3K27me3.

```
$ conda run epic2 --treatment [.bam] --control [input.bam] --genome
    hg38 --false-discovery-rate 0.001 --chromsizes hg38.chrom.sizes
    --fragment-size 380
```

3.6.6 DiffReps histone PTM differential region identification

The *diffReps* software (version 1.55.6) was used to identify any differential region by between treatment and control of all histone PTM with both peak and broad settings, using replicates if applicable:

```
$ diffReps.pl --treatment [treatment.bam] --control [control.bam]
    --btr [treatment.input.bed] --bco [control.input.bed] --report
    --chrlen hg38.chrom.sizes --meth nb --nproc 6 --pval 0.0001
    --frag 380 --mode [p / block]
```

3.6.7 Bedtools interaction with genome coordinates

The *bedtools* software (version 2.30.0) was used in various settings whenever selection for specific gene regions or merging of peak files was necessary. *Bedtools* was extensively used when interacting with histone PTM datasets in conjunction with the DiffReps software by intersecting macs2-identified narrowPeak files with diffReps-identified genome regions, where an overlap fraction (-f) value of either 0.5 for differential peaks or 0.1 for differential blocks was required:

```
$ bedtools intersect -f [narrow: 0.5; broad: 0.1] -a [.narrowPeak /
    diffusepeaks.bed] -b [differentialregion.bed]
```

Additionally, browser extensible data (bed) files were used in the EaSeq software. Here, .bam files were transformed into .bed files with the following command:

```
$ bedools bamtobed -i [.bam]
```

3.6.8 DeepTools creation of heatmaps

The *deepTools* software suite (version 3.5.1) was employed to first convert .bam files into bigwig (bw) format prior to plotting them in heatmaps. The first step is to convert .bam into .bw files, applying reads per genomic content (RPGC) normalization and extending the read size. For DNA obtained from human cells, a blacklist was used here to remove unstructured, anomalous or general high signal in next-generation sequencing experiments (AMEMIYA *et al.*, 2019). When DNA was obtained from rat cells, removal of high-signal regions was performed at a later step in the pipeline.

```
$ conda run bamCoverage --effectiveGenomeSize [2.2e9 for rat, 2.7e9
    for human, -bs 5- normalizeUsing RPGC --centerReads --e 150 -p12
    --scaleFactor [Scale factor] --blackListFileName [hg38.blacklist]
```

This is followed by creation of an intermediate file from the scores per selected (e.g. peak) genome regions in a given window around the peak:

```
$ conda run computeMatrix reference-point --referencePoint center -S
    [.bw files] -R [peak region] -a 2000 -b 2000 -p 4 -bs 5 --
    skipZeros
```

The final step is the creation of the heatmap:

```
$ conda run plotHeatmap --refPointLabel Peak --plotFileFormat svg --
    heatmapHeight 20 --colorMap [] --missingDataColor black
    --interpolationMethod gaussian --plotType Lines
```

3.6.9 HOMER motif identification

The motif identification software *HOMER* (version 4.11) was used to identify transcription factor motifs from E1B-55K binding sites across various experiments with standard settings. This was generally preceded by *bedtools*-mediated removal of any non-canonical chromosomal regions and masked genomic regions:

```
$ findMotifsGenome.pl [.narrowPeak] rn7 -size 200 -p 4 -mis 2
```

3.6.10 Salmon RNA transcript quantification

The *salmon* transcript quantification software (version 1.6) was used as the first step in the mRNA-seq pipeline in its pseudo-mapping based mode. The initial step was to create a decoy-aware transcriptome for both human- and rat-based analysis by concatenating the genome to the end of the transcriptome. A decoy.txt file was prepared with the individual chromosome names from the primary assembly of the genome and followed by the command:

```
$ salmon index -t [transcriptome.genome.fa] -d decoys.txt -p 12
```

This decoy-aware transcriptome was then used in any of the subsequent transcript quantification steps:

```
$ salmon quant -i [decoy.aware.transcriptome] --validateMappings -l A
    -p 8 -r [.fastq] --seqBias --gcBias --fldMean [mean of fragment.
    length] --fldSD [standard deviation of fragment.length] --dumpEQ
```

3.6.11 Data availability

The data created in this work, e.g. tables, can be obtained from the Zenodo digital library with the DOI 10.5281/zenodo.8195779 by contacting the corresponding author.

Chapter 4

Results

4.1 Assessing HAdV5 E1B-55K DNA-binding potential to elucidate putative transcriptional repressor capabilities

HAdV-C5 E1B-55K is a multi-functional viral oncogene and regulatory protein, involved in diverse cellular and viral processes (reviewed in BLACKFORD and GRAND (2009); HIDALGO *et al.* (2019)). One of the main functions of E1B-55K is the inhibition of the p53-mediated apoptosis network, first described by the group of Arnold Berk in YEW *et al.* (1994). By performing *in vitro* transcription assays, they could show that E1B-55K acts as a transcriptional repressor by directly interfering with DNA-bound p53 (MARTIN and BERK, 1998). Scientific knowledge on whether and how E1B-55K interacts with the host genome *in vivo* and if there are epigenomic or transcriptomic consequences to this putative binding are still scarce. This part (4.1) of the thesis is therefore dedicated to answering this question and laying the groundwork for subsequent in-depth experiments. For the sake of clarity, the short species numerical will be used whenever the large E1B protein is mentioned (for example: Ad5 instead of HAdV-C5).

4.1.1 Transfection in HCT116 cells confirms no direct E1B-55K-DNA interaction and suggests a dose-dependent effect on the epigenetic landscape

The following experimental setup was initially established to verify the above-mentioned hypothesis in an *in vivo* cell culture system using the HCT116 tumor cell line and to evaluate the specific conditions necessary for future chromatin immunoprecipitation (ChIP)-seq of epigenetic histone marks, E1B-55K and cellular transcription factors. As illustrated in Figure 9**a**, the initial experiment consists of polyethylenimine (PEI) transfection of HCT116 cells with either pcDNA3 empty vector or an Ad5 E1B-55K containing plasmid, subsequent formaldehyde cross-linked ChIP-seq, NGS and qPCR validation of specific tar-

gets. After the ChIP experiment, the amount of precipitated DNA was compared to an input control from the same experiment by qPCR to validate the efficiency of the antibodies on a set of selected gene regions.

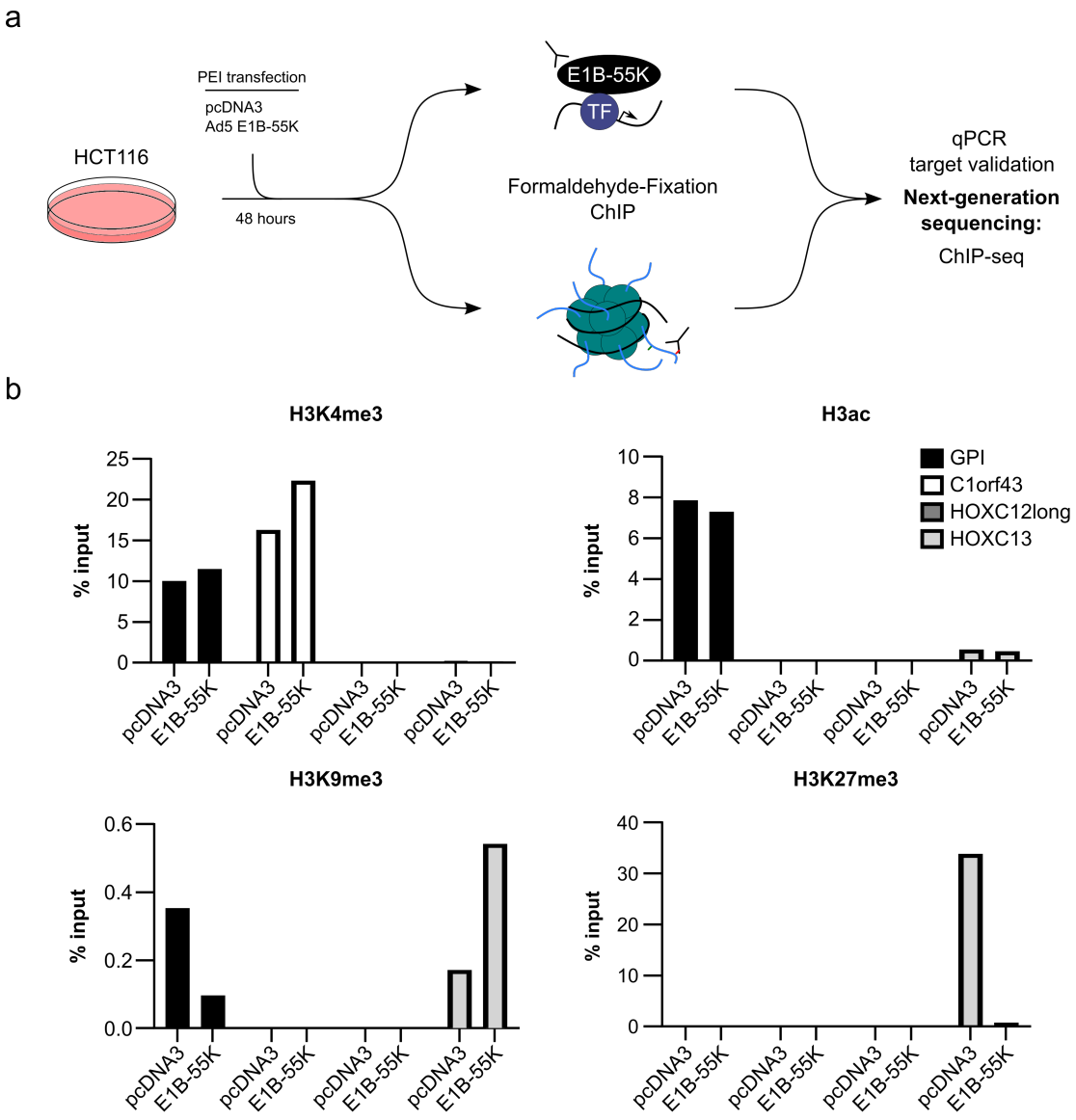


Figure 9: ChIP-seq in HCT116 cells verifies the specificity of antibodies targeting histone PTMs. a Schematic workflow of the PEI-transfection experiment with subsequent formaldehyde ChIP-seq in pcDNA3- (#136) or E1B-55K (#1319) expressing HCT116 cells for either E1B-55K (2A6 antibody), activating histone marks H3K4me3 (#634), H3K9/14ac (#636) and repressive histone marks H3K9me3 (#637) or H3K27me3 (#635). **b** qPCR analysis of selected epigenetic marks to verify the specificity of antibodies for subsequent experiments, performed prior to sequencing. The y-axis (%-input) visualizes the ratio of immunoprecipitated DNA to input DNA. PEI: Polyethylenimine, TF: Transcription factor

These consist of the promoter regions of the ubiquitously expressed glucose-6-phosphate isomerase (GPI) and chromosome 1 open reading frame 43 (C1orf43) housekeeping genes that should either harbor one or both activating histone H3K4me3 or H3K9/14 acetylation PTMs (EISENBERG and LEVANON, 2013). The controls for the repressive histone marks

H3K9me3 and H3K27me3 consist of two different homeobox (HOX)-genes: HOXC12 for the former and HOXC13 for the latter PTM. HOX genes are known to be associated with polycomb-mediated repression and are therefore highly enriched with H3K27me3 and, in somes cases, with the constitutive heterochromatin marker H3K9me3 (VIEUX-ROCHAS *et al.*, 2015; ROPA *et al.*, 2018).

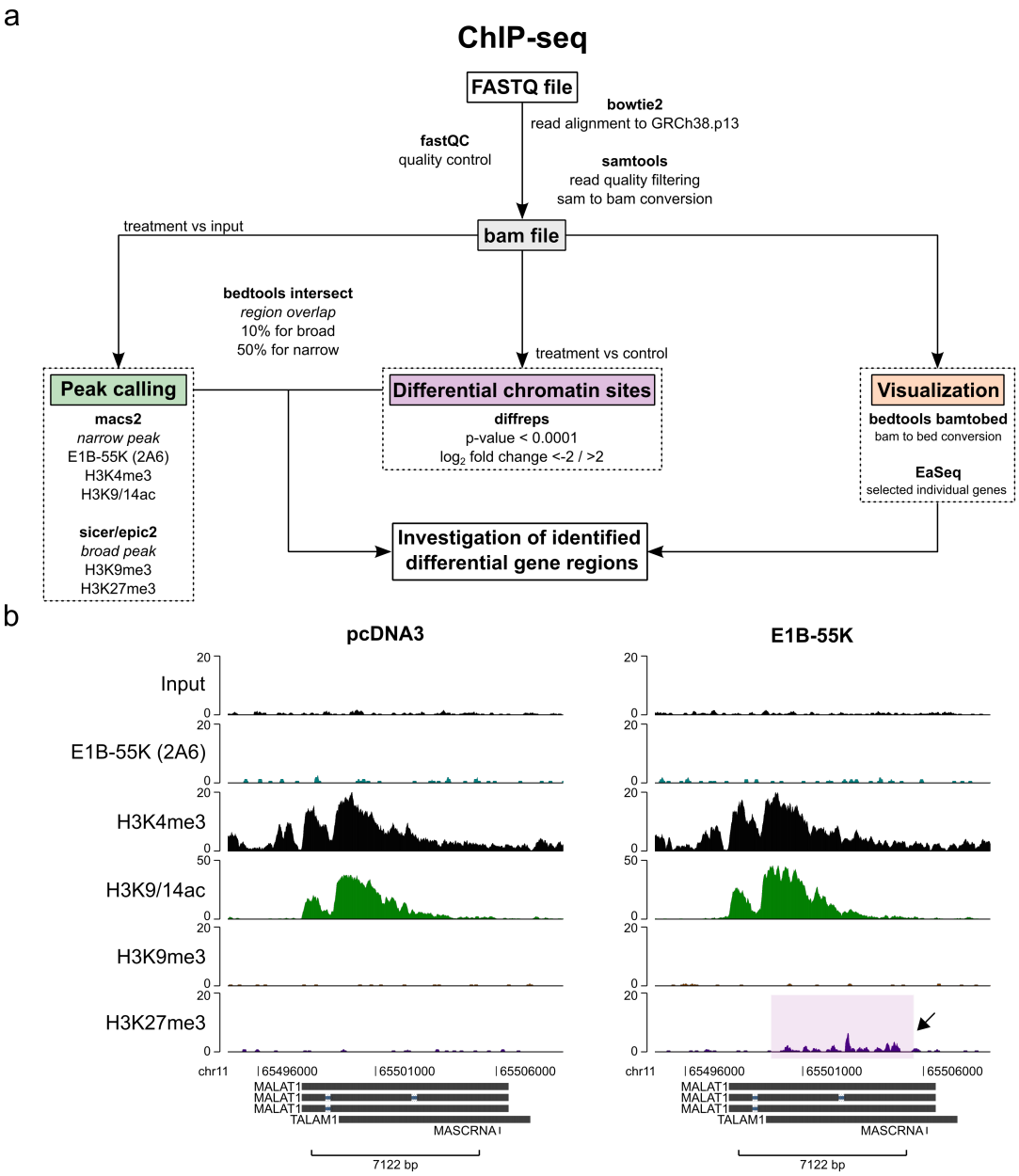


Figure 10: Identification of differential H3K27me3 PTM in the gene body of MALAT1 in E1B-55K-expressing HCT116 cells. a Flowchart diagram displaying the utilized bioinformatic software for DNA sequence alignment, narrow- and broad peak calling, identification of differential histone chromatin regions and visualization. **b** Distribution of the investigated histone PTMs in the sole differential region of potential E1B-55K-dependent H3K27me3 gain in the gene body of MALAT1. The arrow indicates the highlighted significant region. The y-axis indicates the respective ChIP-seq signal. Here, the input was used as a negative control for peak calling.

As shown in Figure 9**b**, the investigated regions mostly display the expected PTM status with significant enrichment in the activating histone marks over 5%-, while the H3K9me3-targetting ChIP enriched around 0.5%- and the H3K27me3-targetting ChIP accumulated around 35% of the input DNA. These results indicated a successful ChIP experiment and the samples were subsequently sequenced to a depth of approximately 10-20*10^6 reads. After obtaining the raw and demultiplexed FASTQ data from the *Illumina NextSeq 500*, the bioinformatic analysis according to the pipeline outlined in Figure 10**a** was performed. In short, this consists of aligning the FASTQ sequence files to the human reference genome assembly GRCh38.p13, setting a read quality filter and removing duplicated reads, performing either narrow (macs2 for 2A6, H3K4me3 and H3K9/14ac) or broad (sicer/epic for H3K9me3 and H3K27me3) peak calling and overlapping these regions with identified differential regions (diffReps). A single differential region could be identified regarding the H3K27me3 signal in the gene body of the metastasis associated lung adenocarcinoma transcript 1 (MALAT1) gene and subsequently visualized using the EaSeq-software in Figure 10**b**. To verify the biological relevance of this signal gain, the experiment was repeated and MALAT1-specific primers within- and around the gene body were designed. After ChIP, this region was tested for H3K27me3 signal strength via qPCR, visualized in Figure 11, but no significant difference between E1B-55K and the pcDNA3 control could be identified, hinting towards the initial finding being a sequencing or experimental artifact.

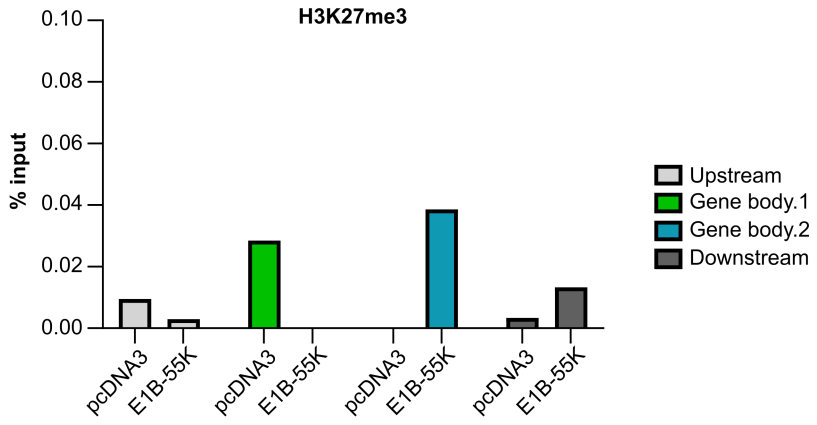


Figure 11: Significant change of the histone H3K27me3 signal in MALAT1 cannot be reproduced in subsequent ChIP-qPCR replicate. The transfection and ChIP experiment was repeated in HCT116 cells and the H3K27me3 signal in the gene body of MALAT1 was compared after 48 h between pcDNA3- and E1B-55K-transfected cells with specifically designed primers that target regions directly up- and downstream or directly within the gene body at two different sites. The y-axis (%-input) visualizes the ratio of immunoprecipitated DNA in comparison to input DNA.

Additionally, no significantly enriched region could be found by targeting untagged E1B-55K, leading to the assumption that formaldehyde as the sole fixation agent would not be sufficient to enrich potential indirect transcription factors and that E1B-55K does not directly bind to DNA by itself. An additional confounding factor could lie in the utilized hybridoma-derived supernatant 2A6 antibody with unclear concentration, which could be alleviated by usage of a HA-tagged E1B-55K. The experimental setup was also not ideal, as I did not specifically enrich for E1B-55K-positive cells beforehand.

4.1.2 Dual-fixation ChIP in BRK cells validates secondary interaction of E1B-55K with p53-responsive genes on host chromatin

Considering the idea of E1B-55K as a potential indirect transcriptional regulator, a modified ChIP protocol was adopted from TIAN *et al.* (2012), which proposed using a secondary fixation agent, like the homo-bifunctional N-hydroxysuccinimide ester DSG. This modified protocol allows enrichment of transcriptional coactivator proteins through creation of approximately 7 Å long covalent bridges between protein complexes. A HA-tag was also added to the N-terminus of E1B-55K, allowing subsequent usage of a validated antibody target for ChIP-seq with specific concentrations (α-HA-tag ChIP graded antibody #643). Additionally, to create a uniform cell line that allows stable and pronounced expression of adenoviral oncogenes, fresh BRK cells were transfected with the complete E1-region, containing both E1A and E1B proteins of Ad5 (HOUWELING *et al.*, 1980). After a four- to six week incubation period, transforming foci were isolated and propagated into clonal cell lines, as illustrated in 12**a**. The surviving clonal cell lines (Figure 12**b**, lanes 4-5; Clone2 and Clone4) were investigated regarding their oncogene expression via western blotting and compared to cell lines (Figure 12**b**, lanes 1-3; AB7, BRK2#377 and AB1#1008) that have been transformed by adenoviral oncogenes in prior experiments (NEVELS *et al.*, 1999). The Clone4 expressed more E1A than Clone2 (Figure 12**b**, lane 4), although the steady-state level of E1B-19K and E1B-55K were considerably more pronounced in the latter (Figure 12**b**, lane 5). Both newly established cell lines were used in subsequent ChIP-sequencing experiments. In Figure 12**c**, an agarose gel shows the effect of sonication employed in the protocol to verify the optimal degree of DNA shearing prior to ChIP-seq, where the bulk of DNA should be visible between 100 bp and 500 bp according to KIDDER *et al.* (2011). Albeit similar initial numbers of cells used in the protocol in theory, strong differences were visible regarding the absolute amount of DNA that was used in the subsequent antibody pull-down, which may impact the strength of the ChIP,

as does the expression level of E1B-55K (Figure 12**c**, lane 4 and 5).

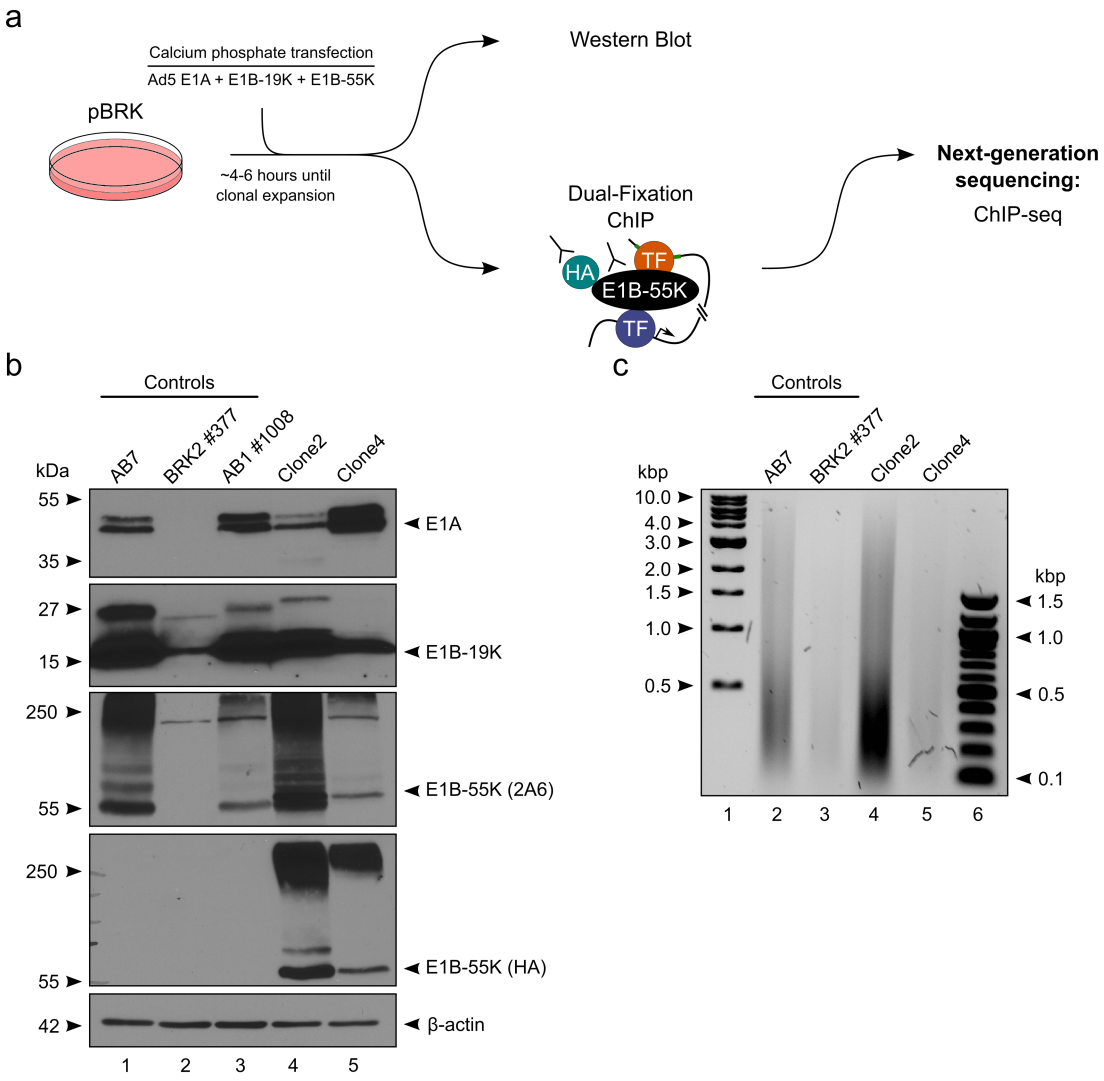


Figure 12: Establishment of HA-E1B-55K expressing BRK cells that are transformed with the Ad5 E1-region. **a** Schematic workflow of calcium phosphate mediated transfection and transformation of BRK cells with the Ad5 E1-region genes with subsequent formaldehyde- and DSG dual-fixation ChIP-seq. **b** Western blot of control cell lines AB7, BRK2#377 and AB1#1008 in lanes 1-3, and the two newly created HA-E1B-55K clonally expanded cell lines Clone2 and Clone4 in lanes 4-5. After total protein extraction, expression levels of key proteins were detected using M73 (α-E1A), 490 (α-Ad5 E1B-19K), 2A6 (α-Ad5 E1B-55K), ChIP-HA (α-HA-E1B-55K) and A5441 (α-β-actin). **c** Control agarose gel of approximately $1{,}25*10^5$ cells per sample from AB7, BRK2#377 and Clone2 and Clone4 sheared input DNA used in subsequent ChIP-seq experiment. The DNA was sheared using a *Diagenode Bioruptor* to an average length of 100–500 bp.

Nevertheless, the complete ChIP-seq procedure utilizing 2A6 or α-HA antibodies targeting HA-E1B-55K, with a rabbit IgG antibody as the appropriate control, was employed and DNA was subjected to NGS. After obtaining the raw, demultiplexed FASTQ files, the computational workflow illustrated in Figure 13**a** was employed to detect the indirect presence of E1B-55K on the DNA.

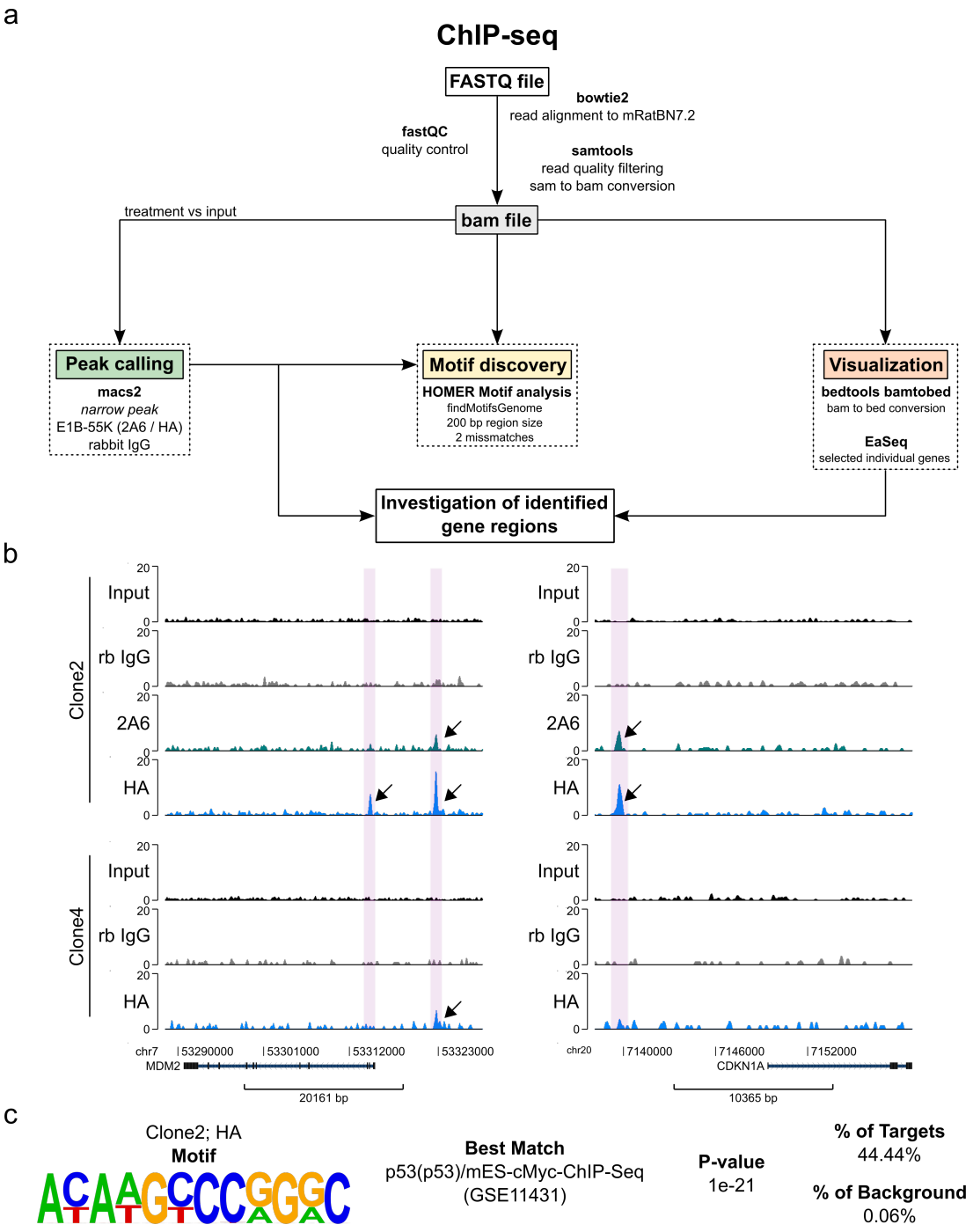


Figure 13: Dual ChIP-seq of HA-E1B-55K reveals indirect binding to p53-responsive genes MDM2 and CDKN1A. a Flowchart diagram displaying the utilized bioinformatic software for DNA sequence alignment, narrow peak calling, HOMER *de novo* motif discovery and visualization. **b** Illustration of the two most significant peaks in regulatory elements of p53-responsive gene regions MDM2 and CDKN1A. Arrows and purple regions indicate the presence of a significant enrichment of the annotated antigen. The y-axes indicate the respective ChIP-seq signal. **c** HOMER *de novo* motif analysis of 200 bp around the peak summit region of peaks with p53 as the best-annotated match.

A total of 19 peaks could be identified in the Clone2 ChIP-seq experiment, which were subsequently investigated regarding transcription factor motifs present around 200 bp of the peak summit using the *de novo* HOMER motif discovery software with default settings. The single statistically significant motif that could be identified best matched towards a p53 ChIP-seq profile with a p-value of $1*10^{-21}$, present in 44,44% of total peaks. The three most significant peaks were found in the promoter and putative enhancer regions of the p53 target genes MDM2 and CDKN1A, as can be seen in Figure 13**b**, albeit with clear differences between the two clones, which could be attributed to the above-mentioned extensive differences in input DNA. It is also apparent that the 2A6-mediated pull-down of E1B-55K lead to considerably less signal compared to the α-HA ChIP, which is why the latter was used in all subsequent experiments. Nonetheless, these results validate and potentiate the notion that E1B-55K indirectly binds to the regulatory regions of p53-targeted genes.

4.1.3 Lentiviral transduction of E1-region genes in BRK cells allows stable expression of wildtype E1B-55K and SUMOylation variants

The functionality of Ad5 E1B-55K is partly regulated via SUMOylation at lysine residue K104 (ENDTER *et al.*, 2001; KOLBE *et al.*, 2022). This PTM can influence the localization of the viral oncoprotein and thereby has the potential to alter its proposed indirect DNA-binding capability, warranting research into specific SUMOylation E1B-55K mutants. By changing either lysine K104 or K101 to an arginine, SUMOylation can be either completely abrogated (K104R, SUMO-null) or amplified tremendously (K101R, hyper-SUMOylation) (KINDSMÜLLER *et al.*, 2007; KOLBE *et al.*, 2022). The NES mutation of E1B-55K removes its ability to shuttle between the cytoplasm and nucleus, trapping the majority of the protein in the latter compartment and resulting in a higher SUMOylation-level compared to the wildtype. Introduction of an arginine at position K104 of the NES mutant creates a nuclear protein that is devoid of SUMOylation, allowing potential insight into how SUMOylation may influence the interaction with DNA-bound transcription factors. Considering the total amount of time investment that transformation via transfection and clonal expansion requires, a potentially faster alternative approach using a lentiviral vector system was investigated. The Ad5 E1A and E1B gene regions, containing all major E1A, E1B-19K and HA-tagged E1B-55K (either wildtype or any of the SUMOylation mutants K101R and K104R, nuclear export mutant NES or the NES-K104R double-mutant isoforms) proteins, were cloned into lentiviral ”gene ontology” (LeGO) vector systems (WEBER *et al.*, 2008). After production of lentiviruses in 293T helper cells, freshly harvested BRK cells were transduced and subsequently sorted into an E1A and E1B double-positive subset using FACS by virtue of venus and blue fluorescent protein (BFP) fluorophore expression. These cells were expanded into three replicates and investigated via western blotting, immunofluorescence, ChIP-seq and mRNA-seq. The general workflow is illustrated in Figure 14**a**, while initial expression levels of selected proteins are visualized in Figure 14**b**. For the sake of readability of this and further parts, ”wildtype”, ”K101R”, ”K104R”, ”NES and ”NES/K104R” will be used whenever the corresponding E1B-55K variant or E1B-55K expressing cell line is mentioned. Expression of E1A is relatively constant across wildtype and SUMOylation mutants, with the exception being two of the three replicates of the previously mentioned SUMO-negative K104R populations (Figure 14**b**, top right, lanes 1-2). HA-E1B-55K expression was investigated with both 2A6 and α-HA antibodies, with only weak quantitative differences visible between the replicates, while analysis between the variants revealed mediocre variance of E1B-55K levels. Here again, the exceptions are

found in the K104R replicates, with major differences especially in E1B-19K- and p53 protein expression, even when taking actin normalization into account (Figure 14**b**, top right, lanes 1-3). As E1B-55K was the major factor of interest in these experiments, subsequent analysis was focused on the second- and third replicate (Figure 14**b**, top right, lanes 2-3). E1B-19K expression was largely comparable between all replicates and mutants, with the exception of NES/K104R, in which substantially more E1B-19K was detectable (Figure 14**b**, bottom, lanes 4-6). It could also be observed that all NES/K104R replicates were expressing slightly less p53 than the other variants.

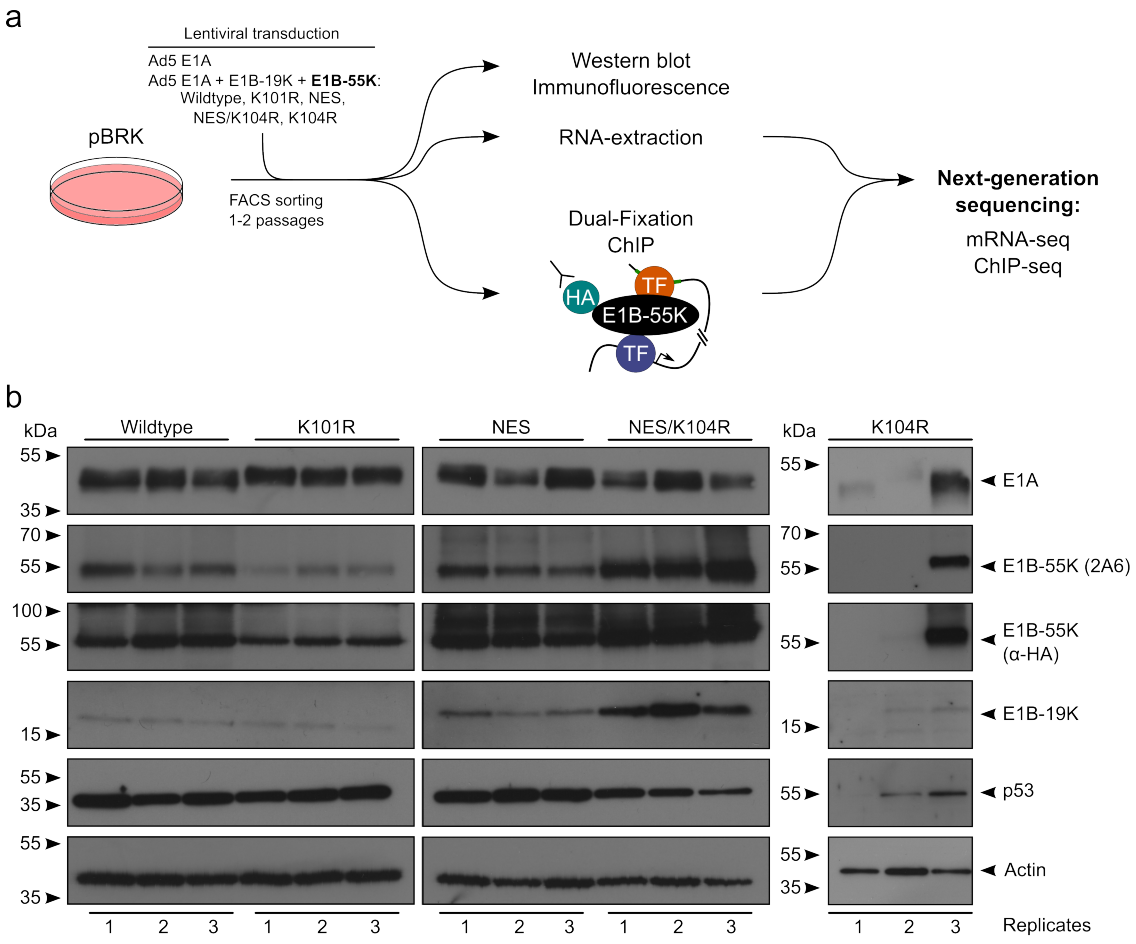


Figure 14: BRK cell transduction with lentiviruses harboring the Ad5 E1-region genes reliably creates HA-E1B-55K expressing cell lines. a Schematic workflow of LeGO-mediated transduction of BRK cells with the Ad5 E1-region genes with subsequent protein expression and localization verification experiments, as well as mRNA-seq and formaldehyde- and DSG dual-fixation ChIP-seq. **b** Western blot of E1A and E1B expressing replicates. After total protein extraction, expression levels of key proteins were detected using M73 (α-E1A), 490 (α-Ad5 E1B-19K), 2A6 (α-Ad5 E1B-55K), ChIP-HA (α-HA-E1B-55K), pAB421 (α-p53) and A5441 (α-β-actin).

After confirmation of adequate protein expression, analysis of intracellular E1B-55K and p53 distribution was quantitatively assessed via immunofluorescence (Figure 15) in order to better contextualize and interpret downstream ChIP- and mRNA-seq results.

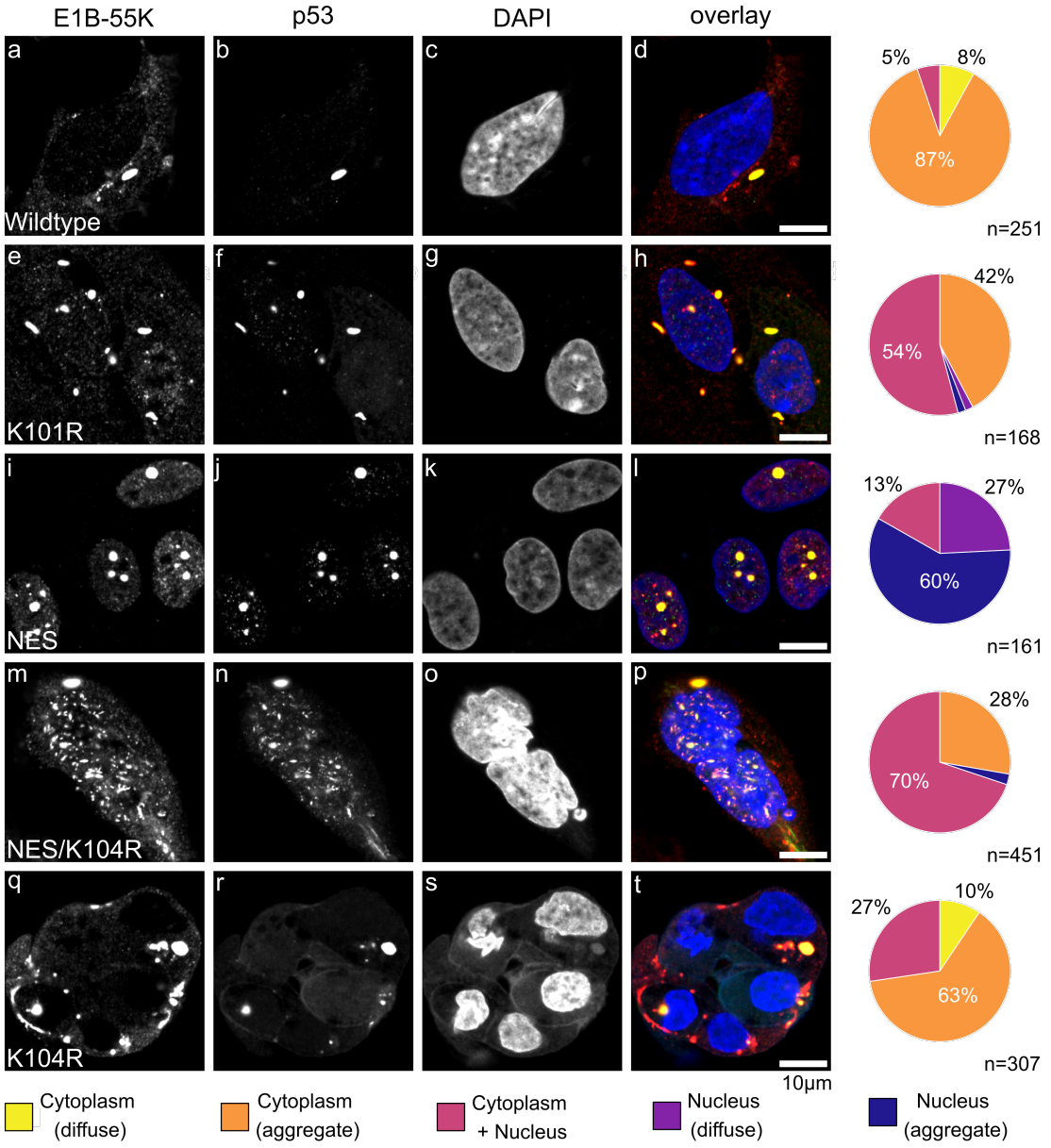


Figure 15: SUMOylation status and presence of a functional NES influence E1B-55K localization. BRK cell replicates expressing different variants of E1B-55K were analyzed regarding their E1B-55K (ChIP-HA) and p53 (pAB421) localization. The dsDNA is visualized by DAPI staining. Imaging was performed with the *Nikon A1R HD25* microscope at 60x magnification with oil-immersion. Multiple images were fused by using the Nikon image stitch tool, extracted via the Fiji software and counted by virtue of the cell counter plugin. BRK cells were grouped according to E1B-55K localization ranging from diffusely in the cytoplasm (in yellow) to aggregates in the nucleus (in dark blue). Images shown here are examples of the observed major distribution patterns. The scale bar corresponds to 10 μm.

E1B-55K was found to be largely localized as expected (KINDSMÜLLER *et al.*, 2007; KOLBE *et al.*, 2022), with wildtype E1B-55K being predominantly (in 87% of cells) found in cytoplasmic perinuclear aggregates with cellular p53 (Figure 15, a-d). hyper-SUMOylation, conferred by the K101R mutation, altered the presence of E1B-55K to both (in 54% of cells) cytoplasmic and nuclear localization in aggregates (Figure 15, e-h). Mutation of the nuclear export signal had the most pronounced effect on E1B-55K, with 60% of cells only displaying nuclear aggregates and no exclusive cytoplasmic presence at all (Figure 15, i-l). The SUMO-null and nuclear export mutant NES/K104R was majorly (in 70% of cells) found in both cytoplasm and nucleus, mostly as aggregates, which confirms the theory that SUMOylation influences the localization of proteins (Figure 15, m-p). Interestingly, more cells that displayed both nuclear and cytoplasmic K104R were found compared to wildtype E1B-55K (Figure 15, q-t). It is important to note that K104R expressing BRK cells were predominantly clumped together, often with a shared cytoplasm, while wildtype E1B-55K expressing BRK cells were more often found in singlets that were easy to distinguish from their respective neighboring cells. This observation makes the exact localization of E1B-55K hard to determine in cluster-structured cellular clumps and may explain the difference between the two cell populations. Taken together, both the presence of a functional nuclear export signal and modification with SUMO mandate the intracellular localization of Ad5 E1B-55K. The viral protein is likely using both signals to actively shuttle between the nucleus and the cytoplasm and to fine-tune its localization.

4.1.4 Expression of wildtype or mutant E1B-55K has global consequences on the BRK cell transcriptome

As soon as the required confluency was reached, the expression of viral proteins and their respective intracellular nuclear localization analyzed, both RNA-extraction and dual-fixation ChIP-seq were performed. As initial observations provided evidence for indirect interaction with the host DNA via p53, I was particularly interested in expanding the interaction repertoire of E1B-55K, while also investigating the influence of localization and SUMOylation on this functionality. In general, ChIP-seq was performed with two of the three technical replicates, while mRNA-seq was performed with all three replicates. The exception here is the K104R mutant, in which mRNA-seq was performed once on the second replicate and twice on different passages of the third replicate (Figure 14**b**, top right, lanes 2 and 3), as the first replicate had no detectable levels of E1B-55K.

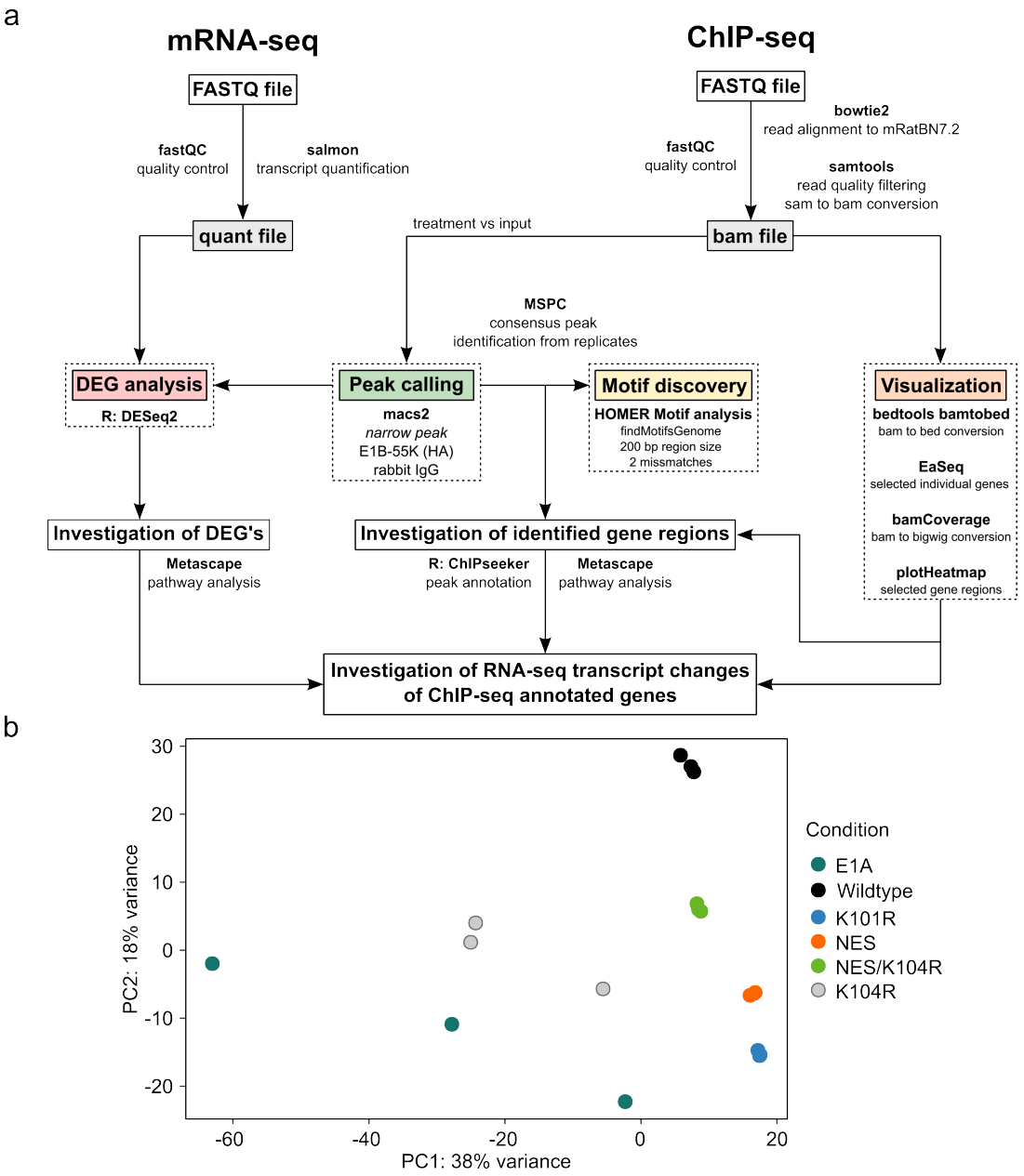


Figure 16: Bioinformatic workflow and mRNA-seq variance between E1B-55K variants. a Flowchart diagram displaying the utilized bioinformatic software for DNA sequence alignment, narrow peak calling, HOMER *de novo* motif discovery, visualization and R packages used for functional assessment, as well as software used for mRNA transcript quantification and subsequent differential gene expression analyses. **b** DESeq2-computed PCA that visualizes the transcriptomic variance between the groups. The negative control E1A is represented by early passages of BRK cells that are transduced by the E1A region of Ad5. The other groups are the different E1B-55K variants that have been used in prior experiments. The RNA was gathered either at the earliest possible time point or at the time of the ChIP-seq experiment. PC: Principal component.

The first part of this analysis aims to capture the global effects on mRNA expression induced by either wildtype E1B-55K or its mutants in comparison to an E1A-only expressing BRK control cell line. Whole cell RNA extraction was followed by poly(A) enrichment, library creation and sequencing to a depth of around 10-20*10^6 reads. Bioinformatic analysis was performed through usage of a pipeline that includes read quality control with

fastQC, transcript quantification using salmon, differential gene expression analysis with the R package DESeq2 and final biological contextualization through usage of Metascape (Figure 16**a**). A principal component analysis (PCA) plot is illustrated in Figure 16**b**, which is a basic 2D representation of the first two principal components of a given data set obtained through a process of dimensionality reduction to visualize the experimental characteristics, variance and batch effects of the mRNA-seq data and functions as an exploratory first view into the data.

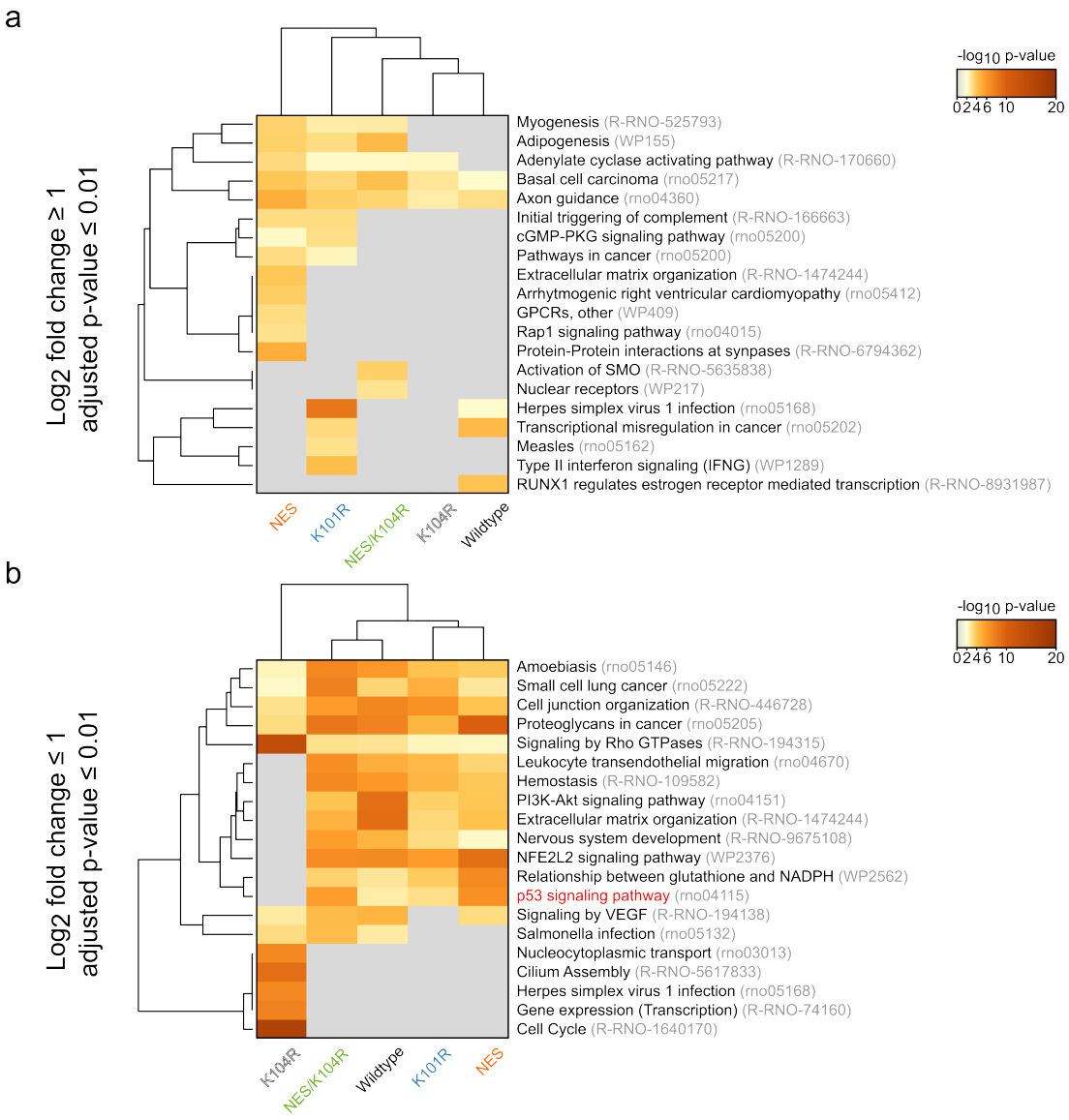


Figure 17: Complete overview of effects on the BRK cell transcriptome induced by expression of E1B-55K variants. Metascape was applied to genes that are deregulated in E1A and E1B-55K variant expressing BRK cells in comparison to the E1A-only control cell line. Pathway and process enrichment analyses were carried out with the KEGG Pathway, Reactome Gene Sets and WikiPathways ontology sources. Pathways involving p53 have been marked in red for the purpose of clarity. Significant pathways are colored based on their -$\log_{10}$ p-value from grey (0, not significant) to dark orange (20, very significant). As a cut-off, only genes with a $\log_2$ fold change of **a** ≥ 1 (upregulated) or **b** $\leq$ -1 (downregulated) and an adjusted p-value of ≤ 0.01 were allowed in this analysis.

The technical replicates of most E1B-55K variants group closely together with the exception of both the E1A and K104R sets, which display comparatively strong intragroup variance. The K104R replicates display differences in protein expression which may explain this finding (14**b**). At the same time, the E1A-only expressing BRK cells were rapidly undergoing apoptosis and extracted RNA of one replicate was of exceptionally low quality. Nonetheless, as this sample was grouping in the vicinity of the other two, it was kept in the control group for the analysis to have a complete triplicate. The basic R DESeq2 workflow (LOVE *et al.*, 2014) was utilized to compute gene $\log_2$ fold changes between E1B-55K variant triplicates and the E1A-only control. To gain insight of potential direct- or indirect intrusion into biological pathways by E1B-55K, a metascape pathway and process enrichment analyses was performed and compiled in Figure 17 (ZHOU *et al.*, 2019). In a bid to increase the significance of this examination, only genes with a $\log_2$ fold change of $\leq$ -1 or $\geq$ 1 and an adjusted p-value of $\leq$ to 0.01 were used. Metascape subjects the gene lists to an enrichment clustering algorithm, where the input is compared to various gene sets defined by their involvement in certain biological pathways or processes. Overrepresentation will lead to integration into the output, colored according to a specific $-\log_{10}$ p-value. Through the integration of multi-gene lists, metascape allows a better interpretation of individual and shared characteristics of the effect of E1B-55K variants on their respective transcriptomes. The chosen categories were KEGG pathways, Reactome gene sets and WikiPathways annotated in brackets, as they offer a wide spectrum of well-curated data repositories (reviewed in CHOWDHURY and SARKAR (2015)). Pathways that are regulated by p53 in the individual ontologies are colored in red, as data presented here (Figure 13**c**) and elsewhere heavily implicate E1B-55K interference with its activity (YEW *et al.*, 1994). The pathway analysis performed with upregulated gene sets (Figure 17**a**) is less dense and more diverse between the variants in comparison to the pathway analysis with the downregulated gene sets (Figure 17**b**). The SUMOylation status seemingly contributes to the distinction between the variants when it comes to potential E1B-55K direct or indirect upregulation of pathways, indicated by the close grouping of both the K101R and NES variants (Figure 17**a**). Many pathways enriched here are related to various differentiation processes (”Myogenesis”, ”Adipogenesis”, ”Axon guidance”) or general transformation- and cancer associated pathways (”Pathways in cancer”, ”Basal cell carcinoma”). Interestingly, only K101R seemingly triggers immune response- and virus infection-related pathways (”Type II interferon signaling”, ”HSV1 infection”, ”Measles”), with NES and K101R both enriching an innate immune system pathway (”Initial triggering

of complement"). On the other hand, when it comes to the downregulated pathways (Figure 17**b**), the only outlier is the K104R variant, being the only investigated sample that lacks in repression of a set of pathways, ranging from "p53 signaling pathway" to "Leukocyte transendothelial migration", while being enriched with an unique arrangement of infection- and cell cycle-related set of ontologies (e.g. "Cell Cycle", "HSV1 infection"). Several cancer- and differentiation- associated pathways are found to be downregulated by all E1B-55K expressing cell lines, regardless of SUMOylation status.

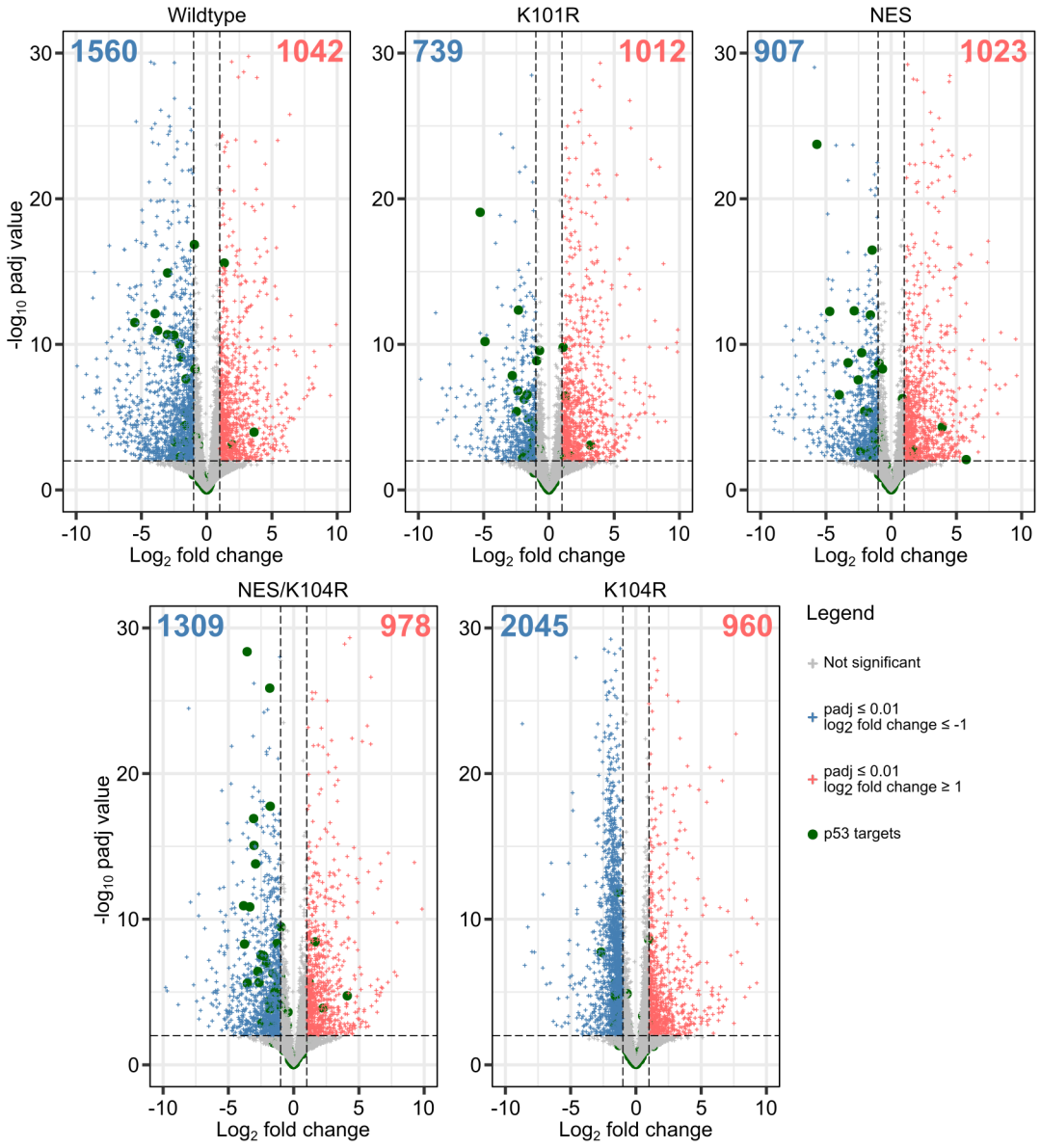


Figure 18: Volcano plots of differential gene expression analysis. DESeq2-computed -$\log_{10}$ adjusted p-value and $\log_2$ fold changes of comparisons between E1B-55K variants and E1A-only expressing control BRK cells visualized by volcano plots (BLIGHE *et al.*, 2022). Blue or red plus signs represent individual genes with an adjusted p-value $\leq$ 0.01 and a $\log_2$ fold change of $\leq$ -1 or $\geq$ 1, respectively with absolute numbers of genes in the top corners. Genes that are enriched in 16 high-throughput p53-targeted ChIP-seq data sets (FISCHER, 2017) are highlighted in green while non-significant genes are colored in grey. padj: adjusted p-value.

A volcano plot in Figure 18 shows the statistical significance represented by the -$\log_{10}$ adjusted p-value (cut-off at 0.01) versus the magnitude of $\log_2$ fold changes (excluding any value between -1 and 1) to visualize the effects and results from the differential gene expression analysis. Several p53 target genes that are curated through combination of 16 published high-throughput ChIP-seq data sets by FISCHER (2017) are annotated here as green circles. K104R is severely hindered regarding deregulation of these p53 target genes, although co-localization could be shown (Figure 15, q-t), while the other E1B-55K variants predominantly and significantly downregulate p53 target genes. The absolute numbers of significant up- and downregulated genes are visible in the top corners of the respective plots. Both wildtype and NES/K104R display equivalent numbers of up- versus downregulated genes and, especially in comparison to the K101R variant, show a wider distribution and magnitude of more significantly downregulated genes. The K101R (1012 up versus 739 down) and NES variants (1023 up versus 907 down), both have more upregulated- versus downregulated genes in total. The K104R variant displays a 1:2 ratio of up/down genes, the latter having a unique pattern of being both very significant while having weak $\log_2$ fold changes. This may be explained with one replicate having comparatively faint expression of E1A (Figure 14**b**, top right, lane 2), as E1A-related functions generally lead to transcriptional activation of a diverse set of genes. These genes may belong to the mRNA-seq enriched pathways related to ”Cell Cycle”, ”Gene expression” and ”Herpes simplex virus 1 infection” (Figure 17**b**).

4.1.5 E1B-55K specifically targets and inhibits regulatory sites of p53-responsive genes, promoting their transcriptional repression

The relative number of transcripts of the previously described curated set of p53-associated genes was investigated with the help of k-means clustering, separating these genes into two clusters based on the distribution of its respective row Z-score and illustrating them using a heatmap (Figure 19). The large cluster 1 comprises of genes that show a clear qualitative difference between E1A and K104R on the one hand and the remaining variants on the other, while the small cluster 2 shows a more divergent and complex transcription pattern of p53-related genes. This illustrates that the presence of the functional K104 SUMOylation site is important to repress transcription of p53-associated genes as BRK cells expressing K104R have roughly equivalent and elevated levels of p53-associated genes compared to the E1A-only control. However, if the nuclear export is additionally abrogated, as seen in the NES/K104R variant, the inhibition of p53-associated genes is

rescued to a level that is comparable to the wildtype. To investigate, whether transcriptional repression of p53-responsive genes correlates with presence of E1B-55K in the TSS of these genes, the ChIP-seq signal enrichment was compared independently from peaks and visualized in Figure 20.

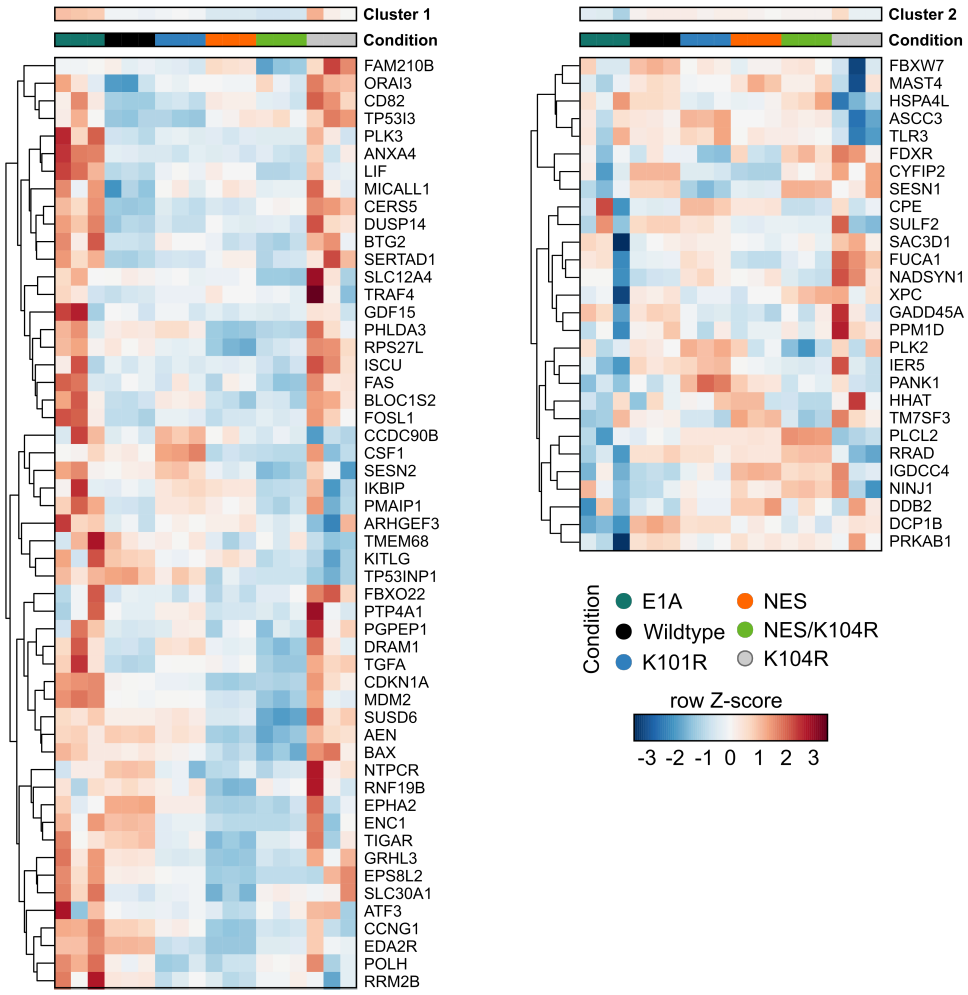


Figure 19: Clustering of p53-associated genes based on their transcript row Z-score distribution. DESeq2-calculated row Z-score of p53-associated genes (FISCHER, 2017) and distribution into two groups by the k-means clustering algorithm. The gene row Z-score for each sample is calculated by subtracting the individual gene transcript number from the transcript number mean of all samples and subsequent division with the standard deviation of all samples. Negative values (in blue) show a comparatively lesser number of transcripts compared to the mean while the inverse is true for positive values (in red). A cut-off of at least 50 transcripts across all conditions was put in place to avoid showing genes with only faint transcription in any of the conditions.

After initial DNA sequence alignment of the ChIP-seq FASTQ sequence files to the rat reference genome assembly mRatBN7.2, TSS coordinates of this specific set of p53 target genes were extracted through usage of the Ensembl (version 105) BioMart tool (CUNNINGHAM et al., 2022) in which I subsequently analyzed and compared the E1B-55K variants indirect occupancy on the DNA. The above mentioned k-means clusters from Figure 19 were used in this context. Neither unspecific rabbit antibody (short: rb IgG),

nor the input sample show any enrichment in these regions, the same holds true for the K104R ChIP-seq sample. A marked difference is visible in the average ChIP-seq signal between cluster 1 and cluster 2, the former showing more enrichment in all variants besides K104R, predominantly apparent in the wildtype. This shows that regulatory regions of p53-associated genes are more strongly occupied with E1B-55K and more likely to be transcriptionally repressed than p53-associated genes that lack binding of E1B-55K.

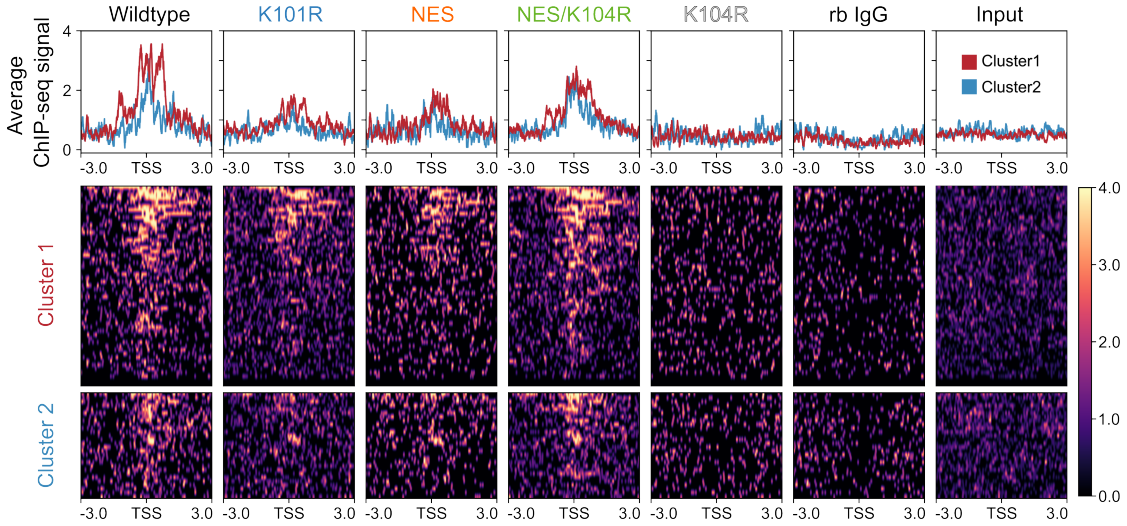


Figure 20: Presence of E1B-55K variants on TSS of p53-responsive genes correlates with transcriptional repression. The deepTools suite (RAMÍREZ *et al.*, 2016) was used with plotHeatmap to visualize the presence of the E1B-55K on the TSS of p53-associated genes that were subdivided earlier into two groups by k-means clustering. Cluster 1 (red) and cluster 2 (blue) are therefore identical to the clusters shown in Figure 19. Visualized here is a range of 6 kb around the TSS of each gene in a heatmap (lower box), sorted by signal strength and ranging from black (weak) to bright yellow (strong). The average ChIP-seq signal (upper box) of the respective Ad5 E1B-55K variant is shown above. The exact genome coordinates were obtained through the usage of the Ensembl (version 105) BioMart tool (CUNNINGHAM et al., 2022). Here, ChIP of an unspecific rabbit IgG antibody (rb IgG) and an input sample were used as controls to verify the specificity of the utilized α-HA antibody.

4.1.6 E1B-55K indirectly binds to many and diverse global genomic regions with differing efficiencies, depending on its SUMOylation status

To gather and analyze the full breadth of E1B-55K binding sites on the host DNA, peak calling was performed and selected exemplary peaks were illustrated using the EaSeq software in Figure 21**a**. This initial visualization marks a significant contrast between wildtype E1B-55K and its variants. At these p53-regulated gene regions, moderately SUMOylated wildtype E1B-55K presence was more pronounced when compared to its variants, with both hyper-SUMOylated K101R and SUMO-negative K104R displaying no signal at all. These regions were specifically chosen as examples, as they (i) contain peaks with the highest significance in the whole data set and (ii) consist of genes that are key regulators

across many ChIP-seq data sets of p53 (FISCHER, 2017). This illustrates that the SUMOylation status of E1B-55K and consequently its cellular localization may generally influence the effectiveness of ChIP-seq.

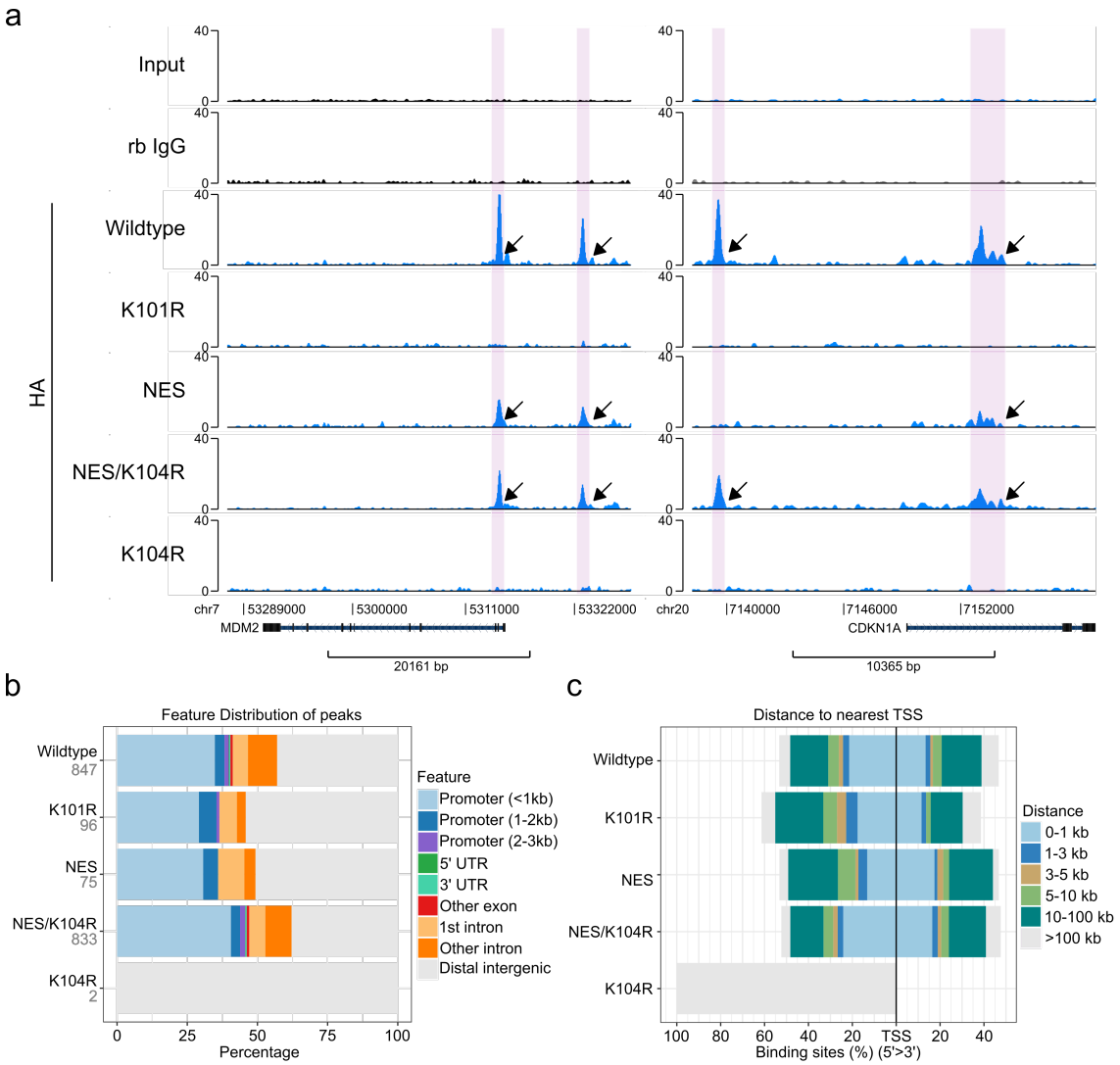


Figure 21: HA-E1B-55K ChIP-seq reveals binding sites directly adjacent to genes as well as regulatory distal regions. **a** Visualization of the identified peaks in regulatory elements of p53-responsive gene regions MDM2 and CDKN1A of all investigated E1B-55K variants. Arrows indicate the presence of a significant enrichment of the annotated antigen. The y-axis indicates the respective ChIP-seq signal. **b** Feature classification of significant peaks in either gene-associated or distal intergenic regions. **c** Percent distribution of binding events located either up- or downstream of the nearest TSS, ranging from directly adjacent (0-1 kb) to distant (> 100 kb).

The MSPC software was employed to improve both sensitivity and specificity by creating a consensus-peak set from individual replicates by removing weak background peaks from the data set (JALILI *et al.*, 2015, 2016). In total, 847 consensus peaks were identified in wildtype, 96 in K101R, 75 in NES, 833 in NES/K104R and only 2 in K104R. Visualized in Figure 21**b** and **c** is the peak feature distribution, based on the nearest up- or downstream gene and approximate distance to the nearest TSS. This analysis was performed

with the R package ChIPseeker (YU *et al.*, 2015), which allows functional assessment of the indirect E1B-55K binding sites identified by ChIP-seq. With the exception of K104R, where no enrichment beyond two binding events could be observed, the majority of enrichment was located directly in the promoter regions (< 1 to 3 kb), with another approximately 10% of peaks being generally gene-associated, distributed into the 5' and 3' untranslated regions, introns or exons (Figure 21**b**). This distribution is also visible in the respective TSS-distance plot of Figure 21**c**, where a small upstream binding preference is visible, with the majority of peaks being located less than 10 kb away from the TSS.

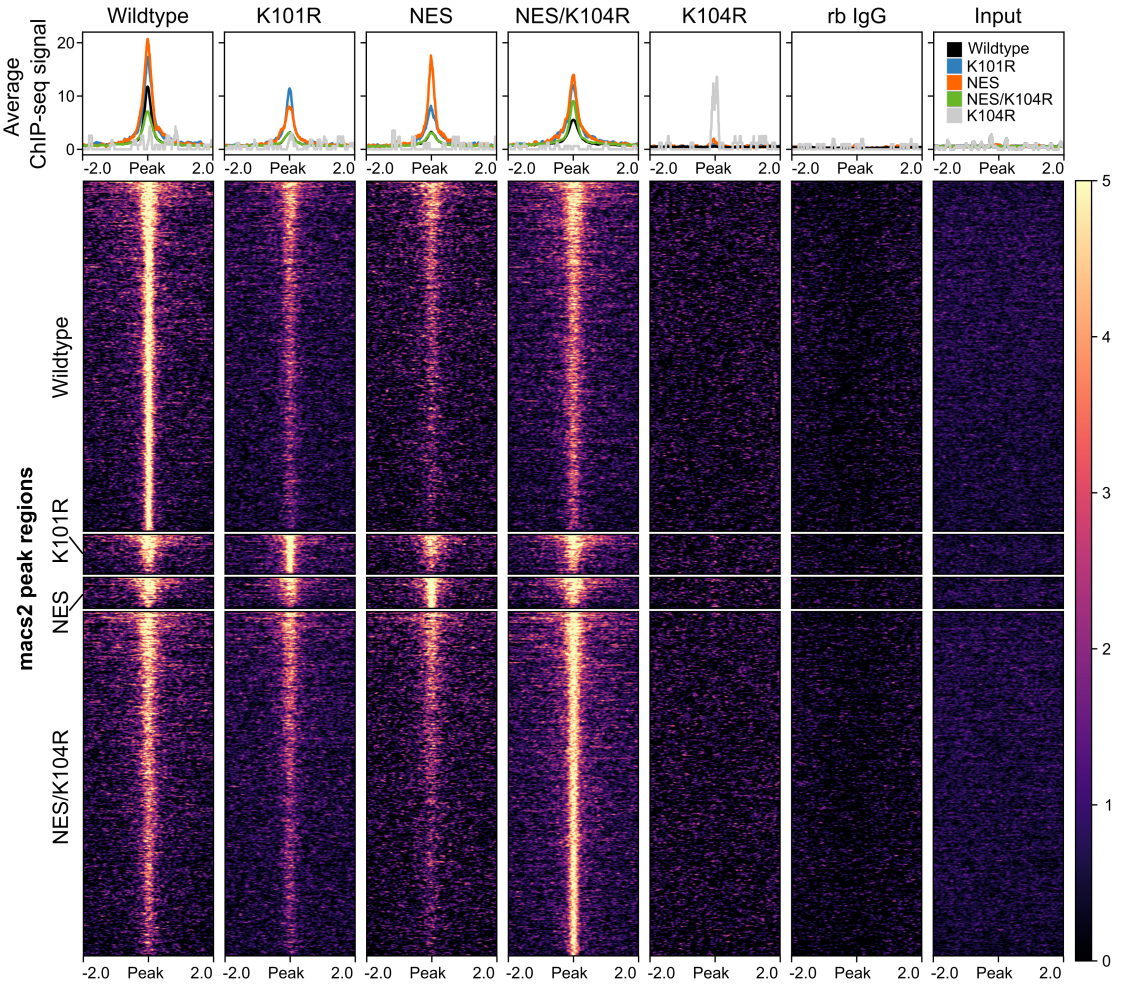


Figure 22: E1B-55K binds to numerous different DNA binding sites on the host genome. Average ChIP-seq signal illustration (upper box) of the investigated Ad5 E1B-55K ChIP-seq in a 4 kb region around significant peak regions (lower box) with signal strength ranging from weak (dark blue) to strong (yellow). Here, ChIP-seq of an unspecific rabbit IgG antibody (rb IgG) and an input sample were used as controls to verify the specificity of the utilized α-HA antibody.

Next, I used the plotHeatmap tool of the deepTools suite (RAMÍREZ *et al.*, 2016) to visualize all identified consensus peaks on the host genome. Even though the absolute number of significant peaks differs drastically, a distinct overlap between the binding patterns of

the Ad5 E1B-55K variants is apparent (Figure 22), where all significant peak regions of either one of the variants is occupied by at least one of the others to a different degree. This demonstrates that both SUMOylation, as a post-translational modification, and the presence of a functional NES directly influence the steady-state levels of E1B-55K present on the DNA and thereby predict the efficiency of the ChIP procedure, but do not fundamentally change the DNA-interaction repertoire. It is interesting to note that the K101R and the NES mutants, both showing significantly stronger SUMOylation patterns (KOLBE *et al.*, 2022), have comparable steady-state levels (Figure 14**b**) and display a predominantly intra-nuclear localization pattern (Figure 15, a-l) compared to the wildtype, were notably more difficult to enrich via ChIP. At the same time, the SUMO-negative K104R, while generally sharing the major cytoplasmic localization pattern with the wildtype protein in immunofluorescence experiments (Figure 15, a-d and q-t), was completely unfeasible to localize on the DNA over repeated ChIP experiments (n = 5, data not shown). It can therefore be concluded that immunofluorescence experiments are a poor predictor of ChIP success with E1B-55K. Concurrently, the SUMO-null, intra-nuclear NES/K104R double mutant protein was directly comparable to the wildtype regarding the outcome of the ChIP-seq experiments, with both proteins having an almost similar number of significant peaks.

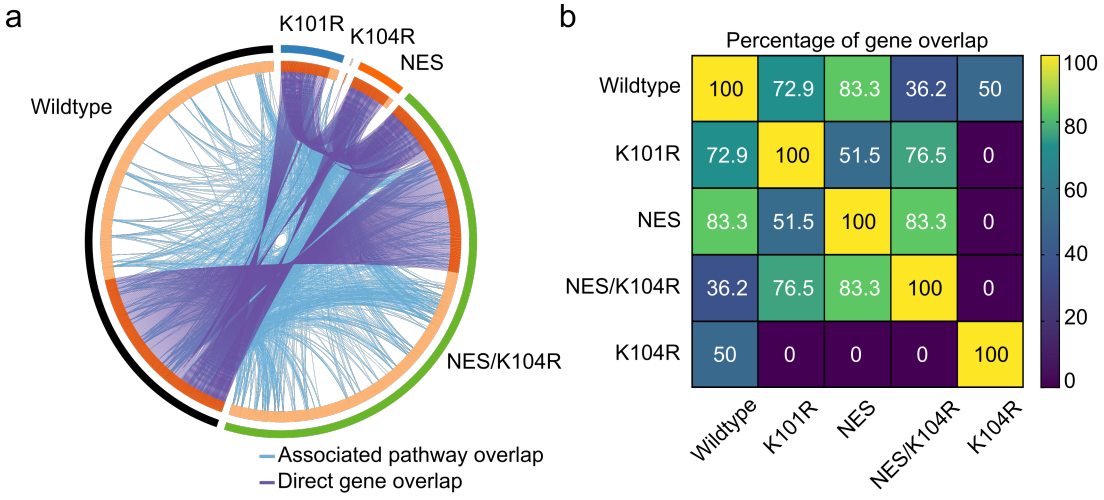


Figure 23: Peak to nearest gene association reveal distinct differences between wildtype and NES/K104R. a Metascape circos plot showing the direct gene overlap (purple) and gene pathway overlap (light blue) between the E1B-55K variants. **b** R heatmap plot with gene percentage overlap between E1B-55K variants. Numerical values describe the ratio of the smaller group divided by the larger group, multiplied by 100.

When annotating the respective peaks with their nearest gene and comparing the overlap between the variants, significant differences between wildtype and NES/K104R become apparent, visualized by a circos plot in Figure 23**a**. In Figure 23**b**, the magnitude of gene overlap is written-out in a heatmap plot (BARTER and YU, 2018).

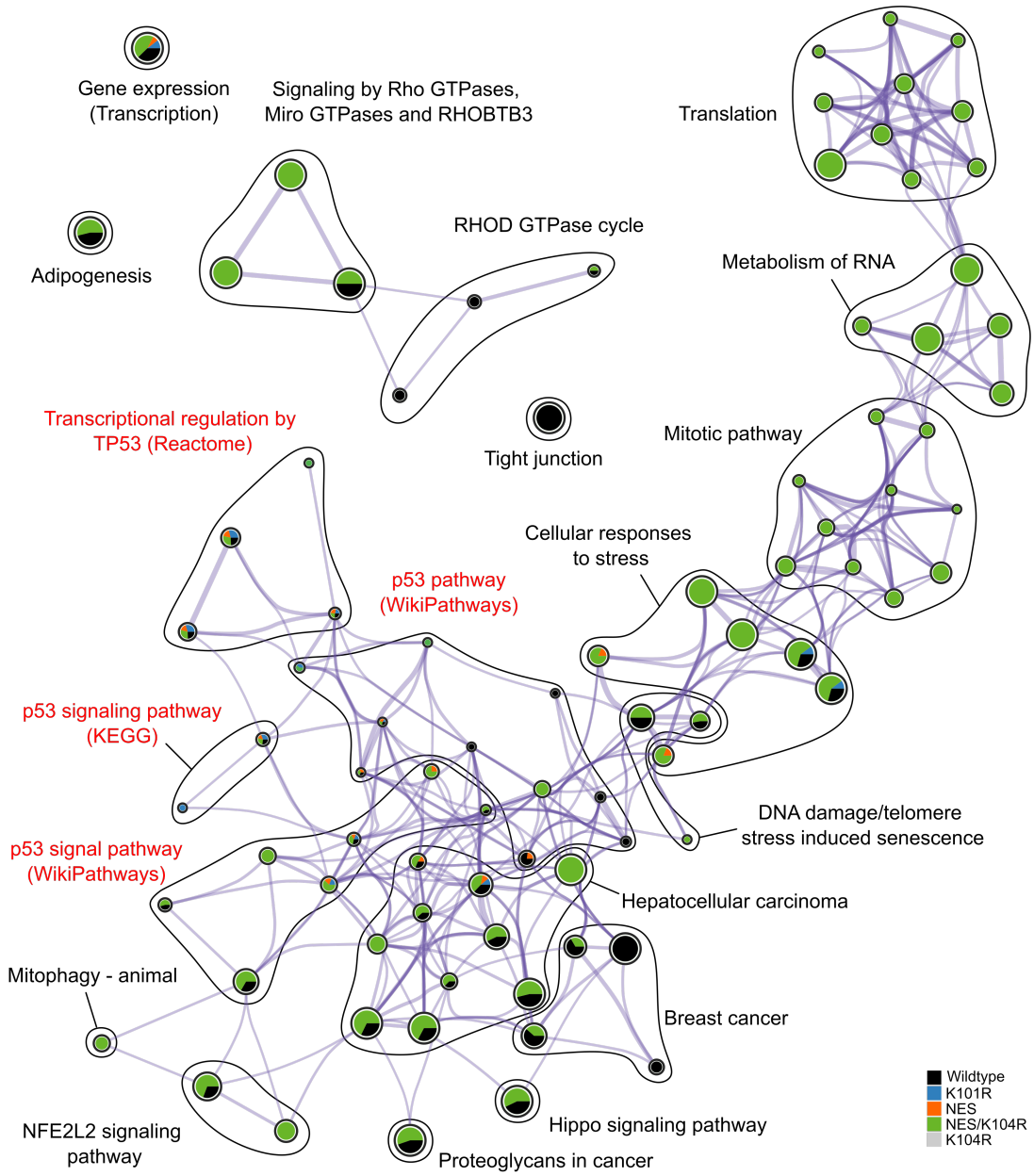


Figure 24: Biological network enrichment analyses reveals E1B-55K intrusion into pathways beyond p53 signaling. Cytoscape network plot visualizing the degree of similarities and differences in overrepresented pathways between individual Ad5 E1B-55K ChIP-seq peak sets, calculated by the metascape analysis tool. Nodes that have a kappa similarity of at least 0.3 are connected by purple lines. The proximity between the nodes is an indicator of redundancy between the gene sets of the nodes, and redundant clusters have been termed by their respective most statistically significant member node. Pathways related to p53 have been marked in red to highlight involvement. Participation of different E1B-55K variants is colored by their respective proportion. A color code for the nodes is provided in the lower right corner.

This is directly contrasted by the earlier observation that both variants do extensively overlap regarding their respective ChIP-seq signal (Figure 22). Since macs2 peak calling works through specific p-value cut-offs of ChIP enrichment in comparison to its local surrounding background signal and to the input control, a weak signal might not be sufficient to be identified as a noteworthy region. This difference in annotated genes might be explained by a slightly changed interaction portfolio of moderately SUMOylated wildtype E1B-55K compared to the SUMO-negative NES/K104R variant, thereby slightly shifting the strength of individual respective binding sites over or under the p-value threshold. As SUMOylation is a very dynamic process, wildtype and all individual variants were therefore investigated comparatively as reflections of a complete mosaic, thereby allowing a more biologically relevant interpretation than either of the variants by itself. To integrate these global E1B-55K-enriched gene sets into an analytical pipeline to produce biologically contextualized results, the metascape enrichment tool was used. Metascape also allows meta-analysis of multi-gene lists, allowing a better interpretation of the individual and shared characteristics of the Ad5 E1B-55K ChIP-seq datasets. As this severely complicates the analysis, only the major modern up-to-date ontologies KEGG pathways, Reactome gene sets and WikiPathways were chosen again as target gene lists for enrichment clustering (KANEHISA and GOTO, 2000; FABREGAT *et al.*, 2018; MARTENS *et al.*, 2021). Here, the individual gene lists were subjected to an enrichment clustering algorithm, where the input is compared to various gene sets defined by their involvement in certain biological pathways or processes. If a specific gene set is significantly overrepresented in the input, it will be integrated into the results with a specific p-value. Since analysis on individual resulting pathways does not take into account any inter-cluster similarities and intra-cluster redundancies, an enrichment network illustration was chosen and presented in Figure 24, depicted with the cytoscape software package (REIMAND *et al.*, 2019). In this network, any significant term is represented by an individual node, which is colored according to the cluster membership of the respective used input gene list. Redundancy in biological pathways and processes drives the closeness of a cluster of nodes which can thereby be categorized under a single descriptive term, usually represented by the most significant node in the cluster. Individual nodes are connected by a purple line, representing a kappa similarity of at least 0.3 - a statistical measure of pairwise similarity first presented in COHEN (1960). Pathways that are regulated by p53 in the individual ontologies are colored in red, as intrusion of E1B-55K in p53 signaling is expected based on prior experiments presented here and highlights the reliability of the ChIP-seq procedure.

In addition, the ”Transcriptional regulation by TP53” (Reactome), the ”Gene expression” (Figure 24, top left), and parts of the ”Hepatocellular carcinoma” cluster are the only pathways that are significantly enriched in all E1B-55K variants. This illustration implicates that Ad5 E1B-55K transcriptionally interferes in a more diverse set of biological pathways beyond p53 signaling and suggests interactions with host transcription factors that have not yet been described in this context. This includes pathways that regulate cellular growth (”Hippo signaling pathway”), cellular transformation (”Breast cancer”, ”Hepatocellular carcinoma”, ”Proteoglycans in cancer”) and stress signaling (”Cellular responses to stress”, ”DNA damage/telomere stress induced senescence”), pathways that are generally targeted in viral infection. Interestingly, a distinct set of pathways could be specifically enriched by ChIP-seq of the NES/K104R mutant (Figure 24, top right), consisting of ”Translation”, ”Metabolism of RNA” and ”Mitotic pathway”.

4.1.7 E1B-55K interacts with a diverse set of transcription factors identified via *de novo* motif analysis

Most transcription factors preferably interact with specific and unique nucleotide sequences to regulate gene expression in any kind of way. To identify E1B-55K-dependent, DSG-fixed transcription factor complexes enriched by ChIP, various amounts of transcription factor binding motifs around the summit region of the peak were identified. Analyzing these motifs across the peak sets allows unbiased insight into potential E1B-55K transcription factor targets. Therefore, the HOMER motif discovery algorithm is the method of choice, as it provides a straightforward combination of the *de novo* and known motif discovery strategies (HEINZ *et al.*, 2010). The difference between these approaches lies in how exactly the software performs motif enrichment, with *de novo* enrichment discovering any sequence that is differentially enriched in the target gene set compared to a set of background sequences, while the known motif algorithm simply shows which of the already known motifs are overrepresented. While the former approach is generally more trustworthy, the latter is advantageous when only a limited set of regions are used in the analysis. As predicted by the metascape pathway ontology (Figure 24), all different E1B-55K variants, with the exception of K104R, revealed p53 as their most significant potential interaction partner with varying degrees of significance (Figure 25**b**). The p53 binding motif was the only one that could be enriched significantly with K101R and NES despite their low (less than 100) peak number (Figure 25**b**).

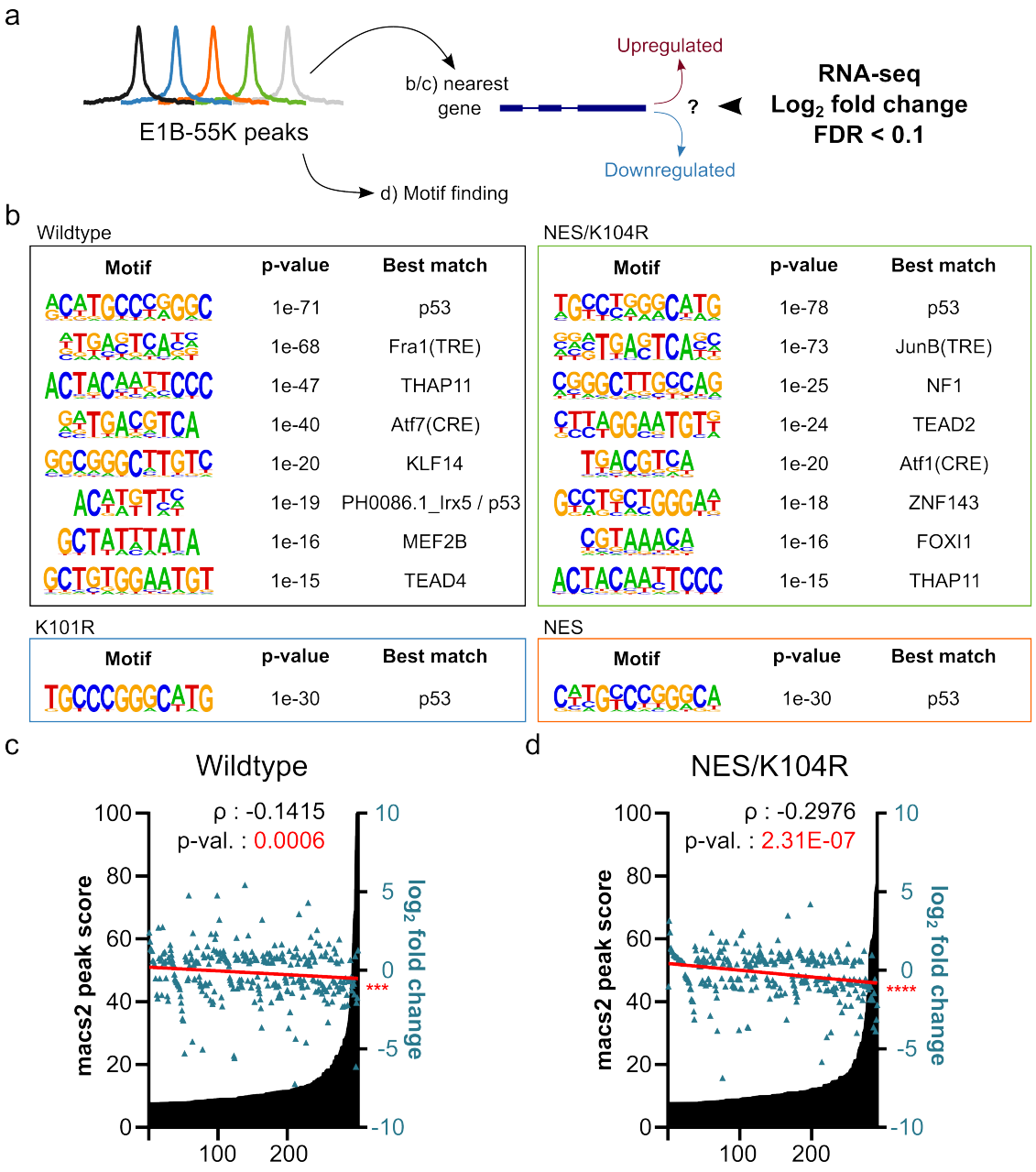


Figure 25: E1B-55K binds different transcription factors on the rat genome with repressive transcriptional consequences. **a** Scheme of combining ChIP-seq significant peaks with mRNA-seq derived mRNA $\log_2$ fold change of genes to examine consequences of E1B-55K binding. **b** HOMER *de novo* motif analysis identifies multiple potential transcription factor motifs bound by E1B-55K, sorted by their individual significance (represented by their p-value). **c** Wildtype- and **d** NES/K104R E1B-55K macs2-calculated peak score and their relationship with the transcriptional status (blue triangles) of the nearest gene in comparison to the control. Peaks are ranked according to the increasing macs2 peak score of that associated gene. Only genes with an adjusted p-value (FDR) of ≤ 0.1 were allowed in this analysis. Linear regression analysis (red line) of $\log_2$ fold changes is included in figure to visualize correlation. A Pearson correlation coefficient (ρ) was calculated to quantify the relationship between E1B-55K binding events and transcriptional consequences of associated genes. A negative ρ-value corresponds to an inverse relationship between peak strength and $\log_2$ fold change, while the p-value presents the calculated significance of ρ with the help of a two-tailed t-test (***$p \leq 0.001$, ****$p \leq 0.0001$).

According to RILEY *et al.* (2008), the human p53 motif is usually depicted as a motif comprising of two decanucleotide repeats of the form RRRCWWGYYY (R = A, G; W = A, T; Y = C, T). In the rat model system, this usually manifested as up to two repeats with the sequence pattern CATGCCCGGGCATG. The second most significant motif in both wildtype and NES/K104R was identified as the TGASTCA (S = G,C) TRE-motif, with the remarkably similar TGASSTCA CRE-motif being slightly less significant. Both motifs are preferentially bound by members of the AP-1 complex, a dimeric complex comprising of members of the Atf-, Jun- or Fos protein families that regulates a diverse set of cellular mechanisms, ranging from differentiation and proliferation to survival and apoptosis (reviewed in GAZON *et al.* (2018)). Other transcription factors of note were TEAD and thanatos-associated proteins (THAP), albeit with differing significance between wildtype and NES/K104R. The former is implicated in the Hippo signaling pathway that regulates cell growth and proliferation (reviewed in HUH *et al.* (2019)), while the latter is part of a complex with the host cell factor 1 (HCF-1) and implicated in a diverse array of functions (PARKER *et al.* (2012)). To investigate the functional consequences of indirect E1B-55K binding to the identified motifs, I combined genes that had a significant transcriptional change (FDR ≤ 0.1) compared to the E1A-only control with motifs enriched in the vicinity of these genes. I was specifically interested in the relationship between the strength of the macs2-calculated peak score and the $\log_2$ fold change deviation of mRNA expression from the nearest respective gene (Figure 25**a**). I ranked the peaks by decreasing macs2-calculated scores and performed a linear correlation analysis by using the Pearson coefficient (ρ) to calculate whether there was a positive or negative relationship between the peak score and the $\log_2$ fold change of the nearest genes (Figure 25**c** and **d**). The NES/K104R mutant displayed a stronger repressive association between binding events and transcriptional regulation as the wildtype, but both E1B-55K variants were found to be weak- and significantly associated with transcriptional repression by virtue of this analysis, with a ρ of -0.1415 and ρ of -0.2976 and a p-value of 0.0006 (***) and 2.31 x 10^{-7} (****), respectively. The negative (positive) linear regression slope corresponds to an association between magnitude of E1B-55K presence on the DNA and transcriptionally repressed (activated) genes. Non-significant results, due to low total peak numbers, of K101R, K104R and NES are not shown here. To investigate, whether a certain motif is associated with genes that are more- or less downregulated compared to the average and whether there is a difference between the occurrence of the motif in a promoter or enhancer region, significant motifs that occurred in at least 30 individual peaks were extracted

using the findMotifsGenome.pl script of the HOMER suite, separated into promoter- and enhancer-associated peaks and finally associated with RNA-seq data to investigate possible transcriptional modulation (Figure 26).

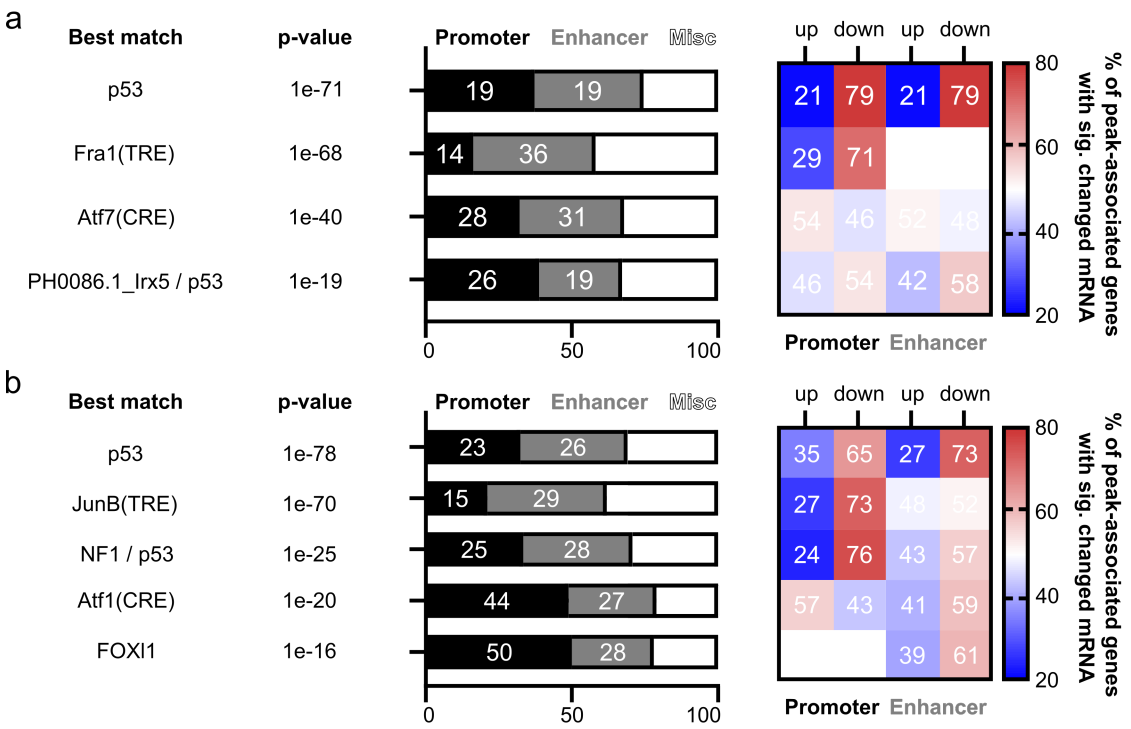


Figure 26: Genes with significant adjacent E1B-55K-bound transcription factor motifs are more likely to be transcriptionally repressed than the average. After **a** wildtype and **b** NES/K104R E1B-55K HOMER *de novo* motif identification and classification, motifs were separated according to their gene-respective position into either promoter- or enhancer-associated. The "Misc" genes are intron-, exon-, or UTR-associated. Only motifs with $\geq$ 30 individual peaks were chosen for this analysis, individually matched with the mRNA-seq data of their respective nearest gene and separated into either up- or downregulated genes. The color gradient indicates the proportion of genes that occur in either set, ranging from blue (low relative amount) over white (equal amount) to red (high relative amount). The numbers in the bars represent the total number of motifs in either promoter or enhancer peaks. The individual motifs are sorted by their individual significance (represented by their p-value).

Here, motifs that occurred around 3 kb of a gene TSS were classified as "promoter", while motifs that were initially described as "distal intergenic" were further selected through overlap with a combined set of predicted enhancer regions to exclude any region that could not readily be described as "enhancer"-associated, which resulted in an approximate loss of around 10% of peaks. I created this special set of enhancers by overlapping predicted enhancers of several different *Rattus norvegicus* cell lines from the EnhancerAtlas 2.0 database (GAO and QIAN, 2020) and using the UCSC liftOver tool to transform these from rn5 to the modern rn7 genome assembly (HINRICHS *et al.*, 2006). The "Misc"-set combines peaks that are annotated as either intron-, exon- , or UTR-associated and are not investigated in detail here. It is immediately apparent that the proportion of downreg-

ulated genes correlates with the significance of the investigated motif present in promoter regions. Besides, this correlation resulted more often in negative transcriptional modulation compared to enhancer-associated localization. E1B-55K presence on p53-associated motifs was accompanied with transcriptional downregulation, which was slightly more pronounced in the wildtype (79% in both promoter and enhancer, Figure 26**a**) than in NES/K104R (65% in promoter and 73% in enhancer, Figure 26**b**), compared to an average of all peaks of 49.67% or 49.14%, respectively. The absolute number of genes associated with these motifs is not sufficient to conclude a functional difference between the variants. Furthermore, TRE-associated motifs in the promoter region of the associated genes were also more often than not found to be transcriptionally repressed (71% in wildtype and 73% in NES/K104R). Unexpectedly, and for the first time to our knowledge, this data could attribute a non-p53 transcription factor to E1B-55K-mediated transcriptional deregulation through indirect interaction on the host cell chromatin, as TRE-associated motifs were also more often than not found to be transcriptionally repressed (71% in wildtype and 73% in NES/K104R), but only if the motif occurred in the promoter region of the associated gene. The motif annotated with either NF1 (enriched by NES/K104R) or KLF14 (enriched by wildtype) as its best matches might be derived from a miss-attribution of a reverse-strand oriented p53-motif, which, according to the HOMER *de novo* motif analysis, is scored slightly worse by the algorithm (0.57 for p53 compared to 0.67 for NF1). Likewise, the motif that was best attributed to PH0086.1_Irx5 in the wildtype E1B-55K ChIP-seq analysis is noticeably similar to the ”CATG” motif that is also present in the p53- responsive element, resulting in a false positive match. It is interesting to note that the CRE-motif that is bound by E1B-55K wildtype and NES/K104R via the AP-1 complex was not significantly attributed to transcriptional deregulation, as opposed to the above mentioned TRE element, suggesting a functional preference for Jun-Fos heterodimer- over Jun-Jun homodimer complexes. Conclusions drawn from enhancer-associated motifs are generally less rigid than promoter-associated motifs, as the sample size was small and associaton with the regulated gene was more complex, as described in PENNACCHIO *et al.* (2013). In this context, the nearest up- or downstream gene was always chosen, resulting in the possibility of an increase in the false-positive error rate. Next, in order to investigate whether co-occurrence of motifs in peaks would explain the downregulation of associated genes, an overlap analysis between the different significant motifs was performed. Here, instances of the four primary motifs that are present in at least 30 individual peaks were analyzed regarding the transcriptomic consequences of respective isolation- and overlapped motif

sets, based on their unique peak region ID (Figure 27**a** and **b**)

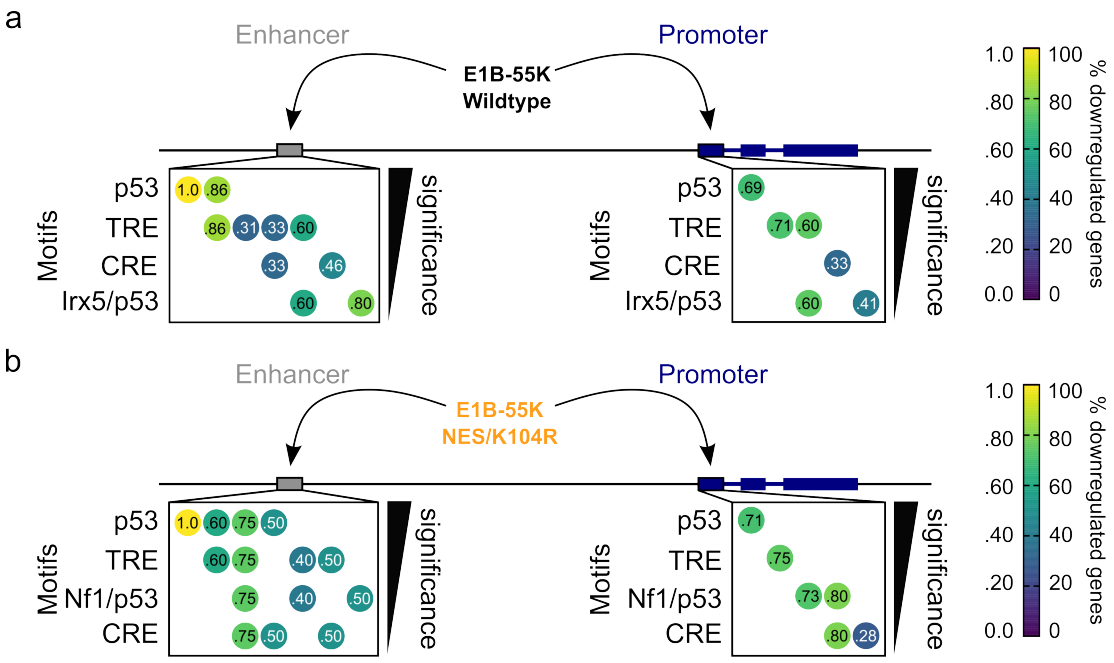


Figure 27: Simplified overview of motif enrichment through E1B-55K ChIP-seq on promoter- or enhancer regions. a Wildtype- and **b** NES/K104R E1B-55K associated co-occurrence motif separated into promoter- and enhancer sets. A circle represents at least five unique motif occurrences in peaks with overlaps in at least five cases being represented by multiple, vertically aligned, circles at the respective motifs. The color code ranges from dark blue (mostly upregulated) to bright yellow (mostly downregulated).

Thus, any two or more motifs are required to be present in a 200 bp region around the summit of a peak to count as overlapping motifs. In Figure 27**a**, the overlap of wildtype E1B-55K motifs are annotated by multiple vertically aligned circles at the present motifs, which illustrates a generally low amount of motif co-occurrence in the promoter region. No promoter-associated overlaps between the canonical p53- and TRE motifs were found, while a small part of the Irx5/p53 subset overlaps with TRE-associated motifs. Interestingly, genes that have either p53- or TRE-associated motifs in their promoter region are considerably more likely to be downregulated (69.23% and 71.43%), while genes with CRE-associated motifs in their promoter or enhancers are less likely to be downregulated (only 33.33% and 46.15%, respectively). A striking difference becomes apparent when E1B-55K was found to be associated with enhancer-associated TRE motifs with no p53 motif present, where only 30.77% of the 13 genes were transcriptionally downregulated. The amounts of downregulated genes increase to 85.71% if a p53 motif was also found in the vicinity, while a sole isolated p53 motif in the enhancer region lead to 100 % transcriptional repression. This pattern is comparable to the NES/K104R-associated motifs (Figure 27**b**). In any case, interaction of E1B-55K wildtype and NES/K104R on either p53- or TRE

motifs in promoter regions was coinciding with transcriptional repression in more than 69% of cases, which suggests that E1B-55K represses TRE-dependent transcription independent from p53. Fundamentally, both wildtype- and the NES/K104R variant of E1B-55K were associated with transcriptional repression of the nearest gene when they are bound to a p53-associated motif, irrespective of exact location. TRE-associated motifs shared this repressive association when found in promoters, whereas isolated enhancer-localized motifs were not related to gene repression. When it comes to isolated CRE-associated motifs, adjacent genes had the tendency to be upregulated. In essence, it could be repeatedly shown here and in the previous parts that, regarding the Ad5 E1B-55K wildtype protein, SUMOylation is absolutely essential for the functional capacity of the viral oncogene, conceivably due to its confinement to the cellular cytoplasm, if abrogated. This is the case with K104R, as the protein was found to be completely devoid of any interaction with the DNA (Figure 22) and a marked loss of transcriptional changes was observed. The wildtype phenotype could be rescued by the mutation of the NES, allowing the protein to remain in the nucleus to establish contact with DNA-bound cellular transcription factors, among its other functions.

4.1.8 Ad5 E1B-55K interacts with TEAD4 without apparent deregulation of associated genes

Identifying transcription factor motifs across my datasets intrigued me to take a more in-depth view into potential E1B-enriched interaction factors. I subsequently tested various different transcription factor targets by ChIP-seq and co-immunoprecipitation in the *Rattus norvegicus* system. Unfortunately, ChIP as a procedure is extremely dependent on the quality of the antibodies, and most commercially obtained antibodies are usually targeted towards human orthologs. Antibodies that function well in the *homo sapiens* system are not expected to behave similarly in the *Rattus norvegicus* system, even when taking high grades of target protein amino acid sequence conservation across species into consideration. This conundrum resulted in my inability to create any functional ChIP-seq data from either p53 and c-Jun, as a major constituent of the AP-1 family, across several experiments and using different antibodies. Nonetheless, successful DNA enrichment could be achieved targeting TEAD4, the protein binding to a significant motif identified by HOMER in both wildtype and NES/K104R ChIP-seq (Figure 25**b**). TEAD4 was targeted as representative of the TEA-motif targeting proteins, but the high grade of conservation across the TEAD family potentially allows enrichment of either one of the protein paralogs (HOLDEN and

CUNNINGHAM, 2018). Here, the ChIP was performed once on formaldehyde-fixed cells of equal passage to the DSG-fixed cells that were used in E1B-55K ChIP experiments.

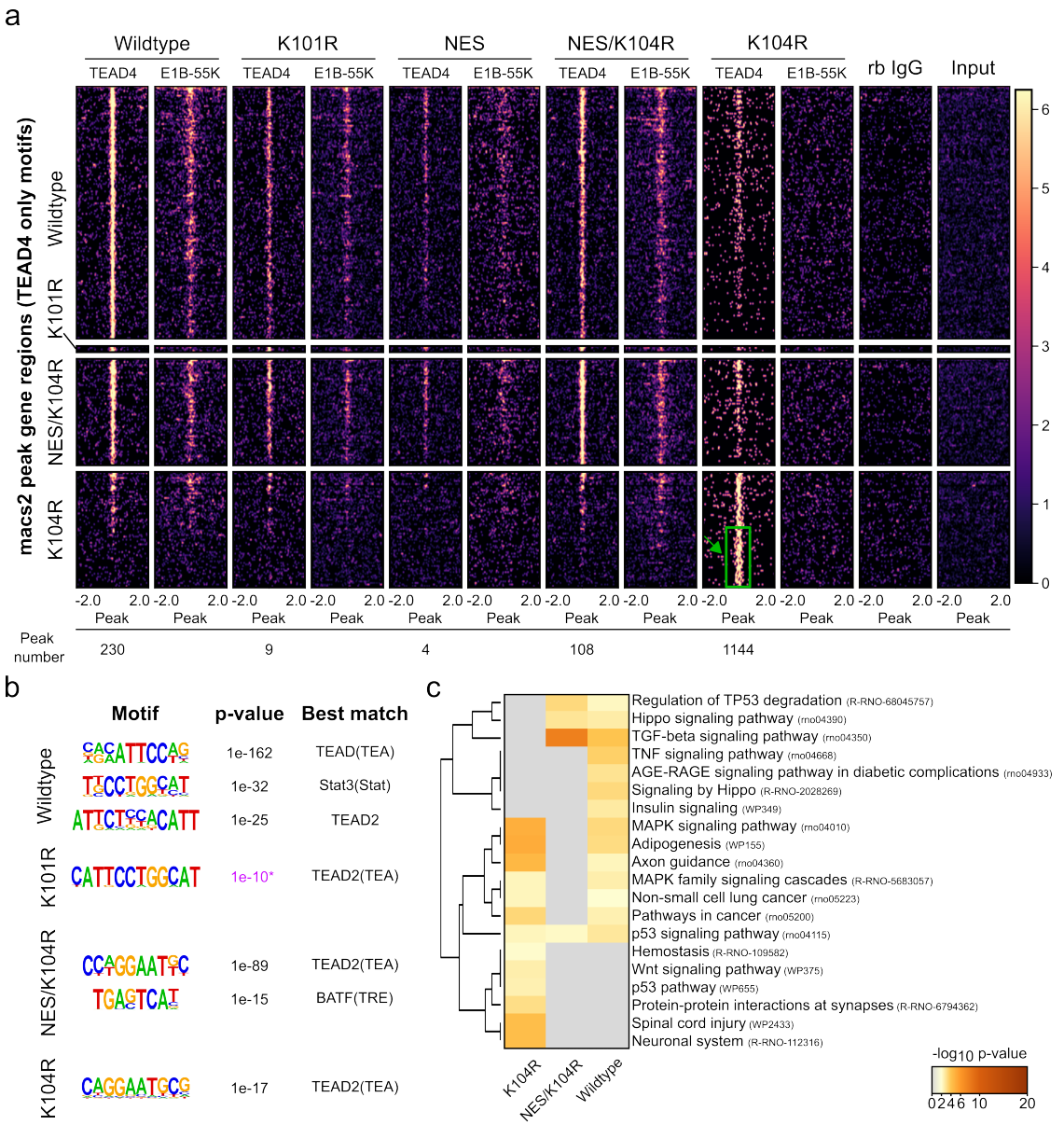


Figure 28: Ad5 E1B-55K interacts with endogenous TEAD4 in ChIP and co-immunoprecipitation assays. a Illustration of TEAD4-ChIP-seq in Ad5 E1A and E1B transformed BRK cells in a 4 kb region around significant peak regions with signal strength ranging from weak (dark blue) to strong (bright yellow). Here, ChIP-seq of unspecific rabbit IgG antibody (rb IgG) and an input sample were used as controls to verify the specificity of the utilized α-TEAD antibody. The green box and arrow highlight the TEAD binding regions that are unique in K104R. The absolute number of peaks found in the TEAD4-ChIP-seq is annotated below the heatmaps. **b** HOMER *de novo* motif identification of TEAD ChIP-seq peak regions with respective p-value. Only motifs with a p-values of at least 1e-15 were shown with the exception of K101R (in violet). **c** Metascape pathway and process enrichment analyses were carried out with the KEGG Pathway, Reactome Gene Sets and WikiPathways ontology sources on the TEAD4-derived ChIP-seq annotated peak sets. The significant pathways are colored based on their -$\log_{10}$ p-value from grey (0, not significant) to dark orange (20, very significant).

A comparison of the different TEAD-peaks and general binding patterns between the Ad5 E1B-55K variants is displayed in Figure 28**a**. The astonishing number of peaks obtained in the K104R replicates was contrasted by the comparatively weak statistical significance of the TEAD motif, with a p-value of only $1*10^{-17}$ (Figure 28**b**). I decided to visualize only the verified peaks containing a TEA-motif, thereby tremendously reducing the amount of peaks in the K104R TEAD4-ChIP-seq from 1144 to 76 peaks. Surprisingly, the general efficiency of TEAD4-targeted ChIP-seq seemingly correlated with the E1B-55K ChIP-seq quality and an overlap between protein-DNA binding patterns is visible in all tested variants. When no E1B-55K K104R presence on DNA was detectable, this pattern was slightly altered, with a proportion of unique peaks that are not found in the other replicates (Figure 28**a**, green box and arrow), albeit presenting a TEA-motif. This opens the possibility that E1B-55K can interact and relocalize TEAD, but no significant effect on the transcriptomic profile of affected genes could be observed (data not shown). A metascape pathway and enrichment analysis was performed with the complete peak-adjacent gene set obtained from TEAD4-ChIP-seq. The ”Hippo signaling” and ”TGF-beta signaling” gene sets were the most significantly enriched pathways in both wildtype and NES/K104R expressing cells, both of which were not present in the K104R ChIP-seq dataset, which also lacked any direct connection to TEAD-controlled pathways. The large amount of K104R expressing TEAD4 peaks contributed to enrichment of distant, weakly significant pathways that may be close to background noise and difficult to interpret. Another interesting finding was the prevalence of p53-regulated pathways in the analysis of TEAD4 peaks, regardless of presence or absence of available E1B-55K on the host chromatin. The next step was to provide additional evidence that E1B-55K can interact with TEAD4 besides the ChIP-seq data. This was investigated by using the co-immunoprecipitation method in H1299- and MCF7 cells. The latter breast cancer cell line was chosen as it is often used in assays investigating the Hippo signaling pathway and has high level of basal TEAD protein expression. We found that both endogenous and co-transfected TEAD4 interacts with all tested Ad5 E1B-55K variants (Figure 29). No significant difference between the variants could be detected in these experiments. In summation, I found that E1B-55K does readily colocalize with TEAD4 regarding its presence on the host genome, but with no significant effect on transcriptional outcomes of TEAD4-adjacent genes. We also found that E1B-55K and TEAD proteins are directly interacting with each other by virtue of co-immunoprecipitation.

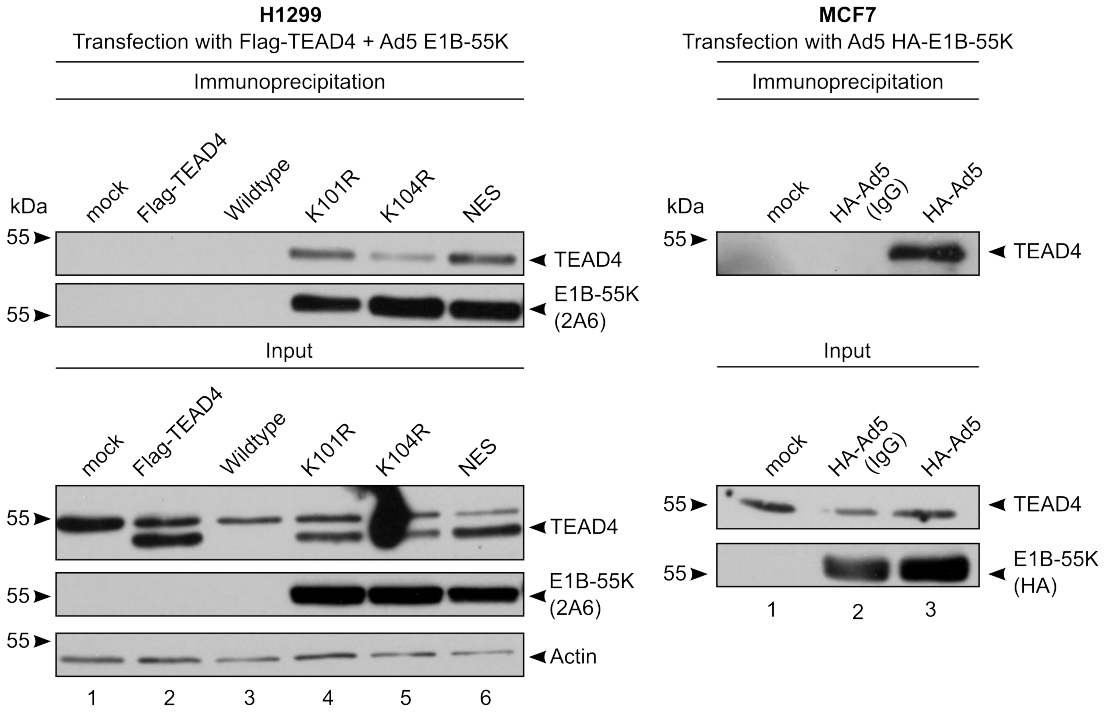


Figure 29: Co-immunoprecipitation of TEAD4 and Ad5 E1B-55K variants in MCF7- and H1299 cells. **a** Western blot of protein co-immunoprecipitation in H1299 cells, which were transfected with TEAD4 (#3713) and different Ad5 E1B-55K variants (#2217). Cells were harvested 48 hours after transfection and co-immunoprecipitation was performed with either the 2A6 antibody (α-Ad5 E1B-55K) or a Flag antibody (M1) conjugated to sepharose beads (α-Flag TEAD4). After 10% SDS-PAGE, proteins were immunoblotted with the aforementioned antibodies and A5441 (α-β-actin). **b** Western blot of co-immunoprecipitated TEAD4 with transfected Ad5 E1B-55K variant (#2215) in MCF7 cells. Cells were harvested 48 hours after transfection and co-immunoprecipitation was performed with either HA antibody (α-E1B-55K) or TEF-3 (α-TEAD4). After 10% SDS-PAGE, proteins were immunoblotted with the aforementioned antibodies. These transfection, co-immunoprecipitation and western blotting experiments were performed by and published in Eileen Dudda's bachelor's thesis.

4.2 Exploring the large E1B indirect DNA-binding potential across different adenoviral species

As in most other adenovirus-related research, the brunt of knowledge on E1B-55K and its multi-faceted functions is focused on just Ad2 and Ad5 from species C. It has been shown that the large E1B proteins from different human adenovirus species members can vary considerably in their amino acid sequences and functions (see Figure 7). For example, a functional NES, represented by several hydrophobic residues with a defined spacing in the N-terminal region of the protein, allowing the protein to actively shuttle between cytoplasm and nucleus, is conserved in members from species B, C and D. Members from species A, F and G do not contain a verifiable NES (Figure 7, BLANCHETTE *et al.* (2013)). Thus, the next part of this work aims to elucidate the general differences and similarities

between the large E1B protein species regarding their potential interaction capacity with DNA-bound transcription factor complexes and resulting putative transcriptional changes.

4.2.1 Indirect DNA-binding capabilities of adenovirus large E1B proteins from species A, F and G are conserved

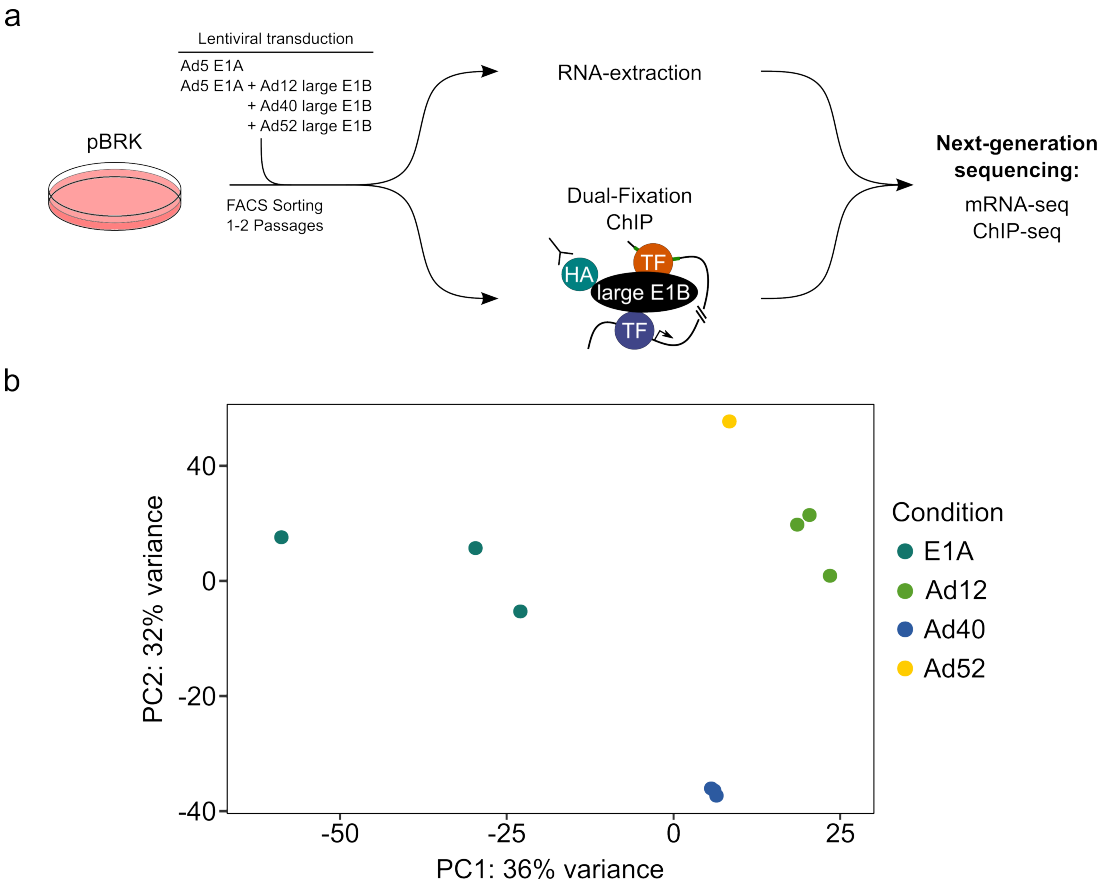


Figure 30: BRK cells transduced with lentiviruses harboring the Ad5 E1A- and large E1B gene regions from different species show substantial differences. **a** Schematic workflow of LeGO-mediated transduction of BRK cells with Ad5 E1A- and different adenoviral species E1B-55K gene regions with mRNA-seq and formaldehyde- and DSG dual-fixation ChIP-seq. The transformation and ChIP-seq experiments were performed by several bachelor's and master's students and published in their respective theses: Max Hüppner (Ad12, B.Sc.), Sophie Weinert (Ad12, M.Sc.), Jennifer Rothe (Ad40, B.Sc.) and Leah Kobza (Ad52, B.Sc.). **b** DESeq2-computed PCA visualizes the variance between the groups. The negative control E1A is represented by early passages of BRK cells that are transduced by the E1A region of Ad5. The other groups are represented by the different E1B species. The RNA was gathered either at the earliest possible time point or at the time of the ChIP-seq experiment. PC: Principal component.

The general workflow of transduction and subsequent transformation of BRK cells through a combination of Ad5 E1A and the large E1B proteins from different species is largely similar to the preceding part 4.1.3, illustrated in Figure 30**a**. The same is true for the bioinformatic approach visualized in Figure 16**a**. The transformed cell lines were created, protein expression was confirmed (data not shown) and ChIP-seq was performed either

once (F40 and G52) or in biological triplicates (A12).

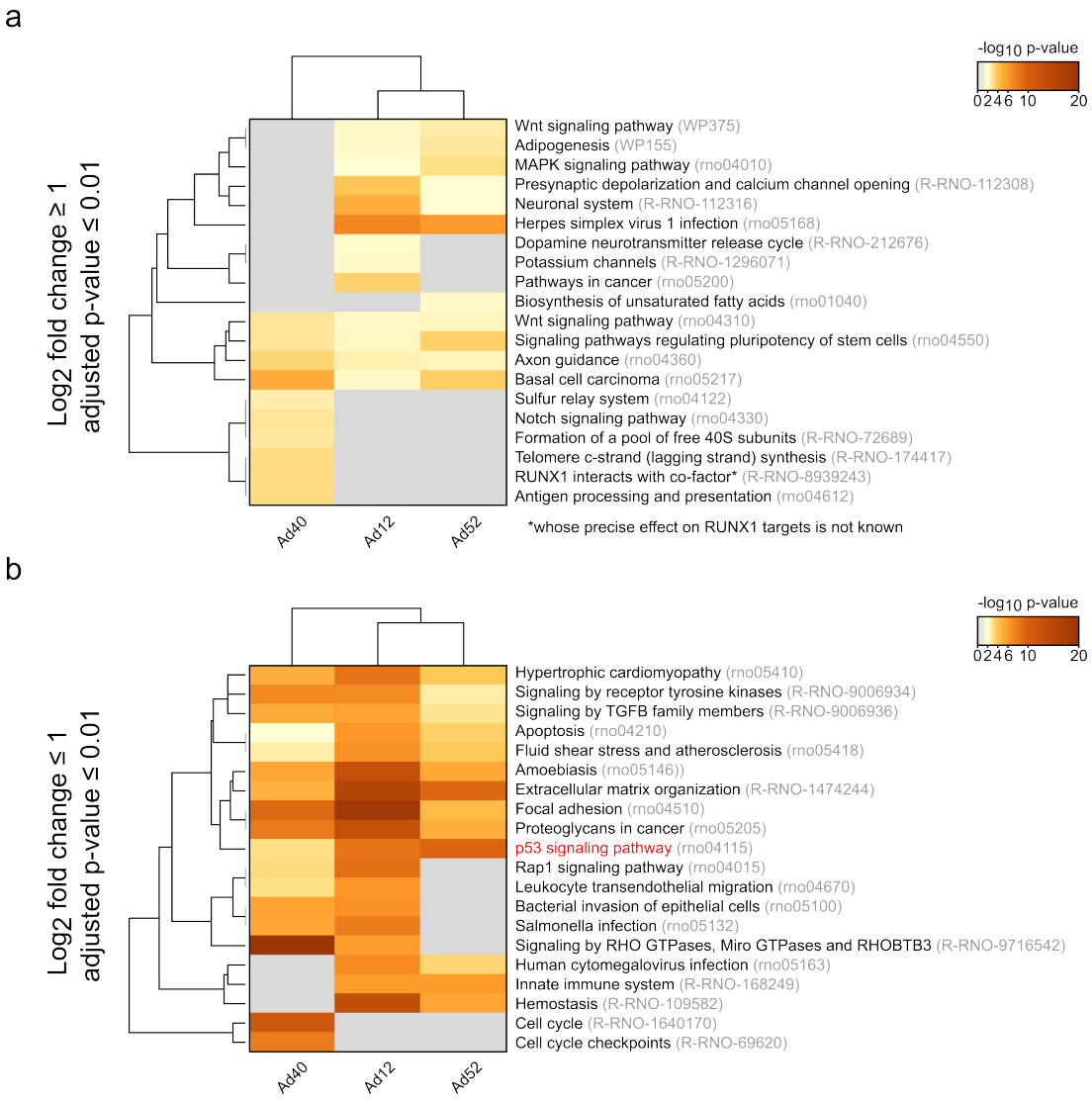


Figure 31: Complete overview of effects on the BRK cell transcriptome induced by expression of large E1B species. Metascape was applied to genes that are deregulated in Ad5 E1A and large E1B species expressing BRK cells in comparison to the E1A-only control cell line. Pathway and process enrichment analyses were carried out with the KEGG Pathway, Reactome Gene Sets and WikiPathways ontology sources. Pathways involving p53 have been marked in red for the purpose of clarity. Significant pathways are colored based on their -$\log_{10}$ p-value from grey (0, not significant) to dark orange (20, very significant). As a cut-off, only genes with a $\log_2$ fold change of **a** $\geq$ 1 (upregulated) or **b** $\leq$ -1 (downregulated) and an adjusted p-value of $\leq$ 0.01 were allowed in this analysis.

The experimental setup was applied to a wider set of adenoviruses species, but successful and stable transformation could only be achieved in the species listed above, as other transduced BRK cells went into apoptosis either shortly after transduction or FACS procedures. Similarly to the procedure presented in 4.1.4, a global transcriptomic analysis is performed to complement the preceding ChIP experiment. A PCA plot is illustrated in Figure 30**b**. While mRNA from Ad12 and Ad40 was analyzed through three biological

replicates, it is important to note that mRNA-seq of Ad52 was only performed once with a single sample which was copied two times to obtain the required three samples for the DESeq2 analysis. This method is only used in an exploratory context and any conclusion drawn from Ad52 is strictly preliminary. This procedure is easily observable in Figure 30**b**, where the Ad52 pseudo-triplicate is focused into a singular dot. To capture the potential influence of the E1B species on biological pathways, metascape pathway and process enrichment analyses were performed and compiled in Figure 31. In an attempt to increase the significance of this approach, only genes with $\log_2$ values of $\leq$ -1 or $\geq$ 1 and an adjusted p-value of $\leq$ 0.01 were considered. The pathway analysis performed with upregulated gene sets (Figure 31**a**) resulted in generally less significant pathways between the species in comparison to the pathway analysis with the downregulated gene sets (Figure 31**b**). Here, ”Axon guidance”, ”Herpes simplex virus 1 infection” and ”Basal cell carcinoma” were equally upregulated in the earlier global transcriptomic analysis performed on the Ad5 E1B-55K variants (Figure 17**a**). Many similar downregulated pathways can be found in this analysis with high significance compared to Figure 17**b**, insinuating a general cross-species conservation of transcriptionally repressive functions of adenoviral large E1B proteins. All species were consistently able to readily repress p53 signaling. Repression of the ”Signaling by TGFB family members” pathway was found in all three species, but was not found to be repressed by Ad5 E1B-55K, while ”Signaling by RHO GTPases, Miro GTPases and RHOBTB3” was found to be expressed by Ad12 and Ad40, but not by Ad52 (Figure 17**b**). In general, BRK cells expressing Ad12 and 52 large E1B species were more akin to each other compared to the Ad40 transduced cell replicates, visible via both PCA (Figure 30**b**) and biological pathway analysis (Figure 31).

4.2.2 Adenovirus large E1B proteins share common DNA-bound transcription factors with little divergence between species

To reveal, whether different large E1B proteins from other species than C regulate transcription by indirect interaction with the host genome, binding sites of were investigated via, ChIP-seq, NGS, and subsequent peak calling was performed while selected exemplary peaks were illustrated using the EaSeq software in Figure 32**a**.

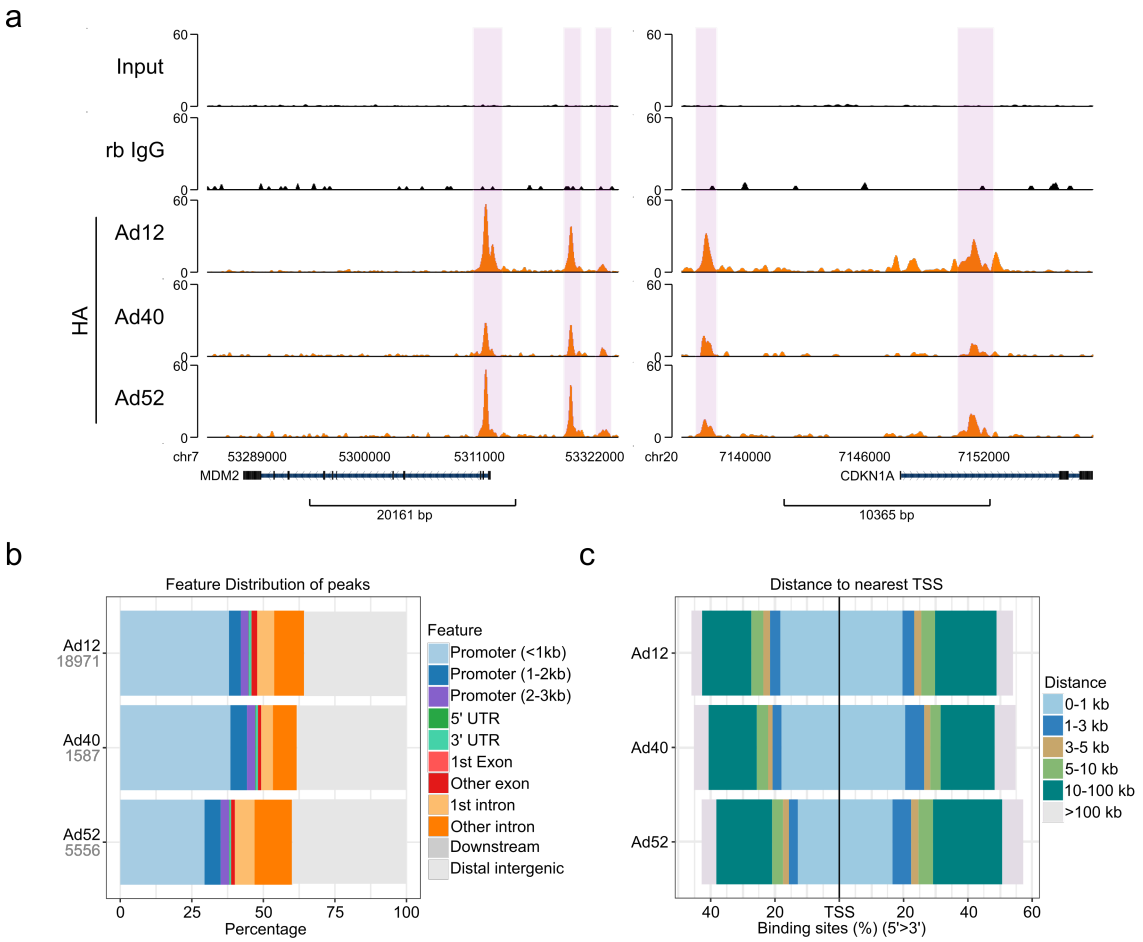


Figure 32: Peak to nearest gene association is largely similar between large E1B species. a Visualization of the identified peaks in regulatory elements of p53-responsive gene regions MDM2 and CDKN1A of all investigated large E1B species. The y-axis indicates the respective ChIP-seq signal. **b** Feature classification of significant peaks in either gene-associated or distal intergenic regions. **c** Percent distribution of binding events located either up- or downstream of the nearest TSS, ranging from directly adjacent (0-1 kb) to distant (> 100 kb).

Consistent to the results obtained from mRNA-seq, any of the tested E1B species were found to directly interact with the regulatory regions of the exemplary p53-regulated gene regions. The absolute number of peaks varied widely between the species, with Ad12 ChIP-seq being enriched at 18791 sites (3 biological replicates, verified by the MSPC software), Ad52 enriched at 5556 sites and Ad40 at 1587 sites. Any of these peak numbers eclipsed the ChIP-seq data gathered from Ad5 HA-E1B-55K by a large margin, in which only 847 sites were significantly enriched by the wildtype protein (Figure 21**b**). Visualized in Figure 32**b** and **c** is the peak feature distribution based on the nearest up- or downstream gene and approximate distance to the nearest TSS. Large E1Bs from different species bind preferably in promoter regions (< 1 to 3 kb), with another approximately 15% of peaks being generally gene-associated, distributed into the 5' and 3' untranslated regions, introns or exons (Figure 32**b**).

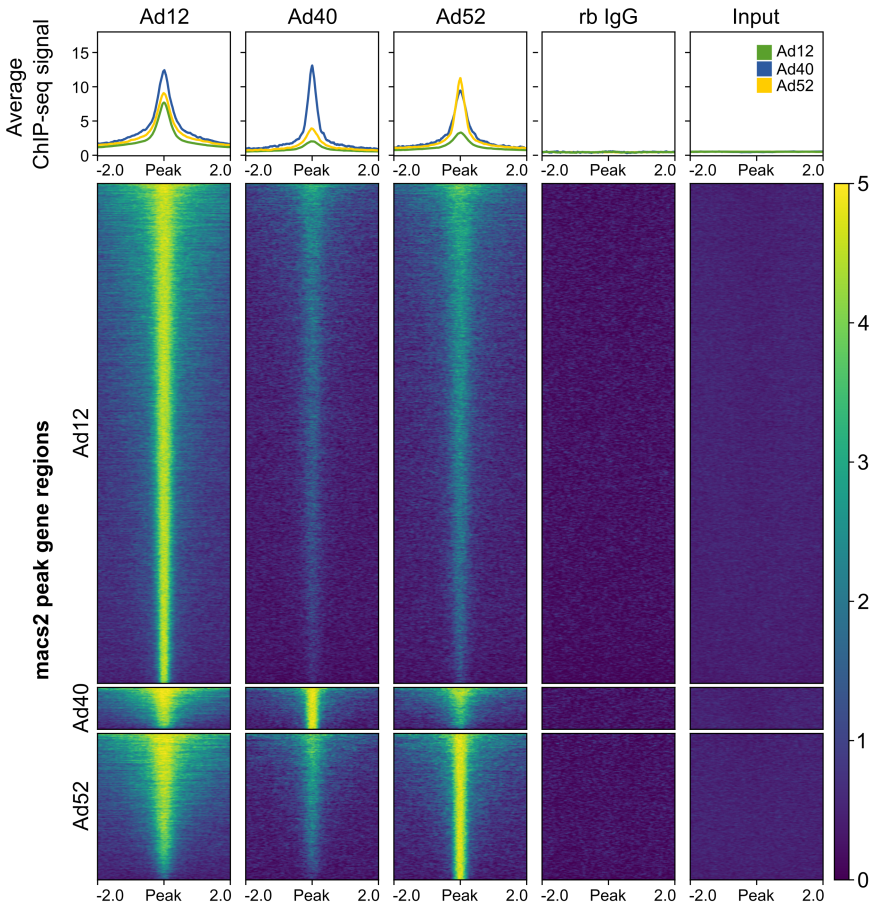


Figure 33: Large E1B species overlap in their ChIP-seq binding profile. Average ChIP-seq signal illustration (upper box) of the investigated E1B species ChIP-seq in a 4 kb region around significant peak regions (lower box) with signal strength ranging from weak (dark blue) to strong (yellow). Here, ChIP-seq of an unspecific rabbit IgG antibody (rb IgG) and an input sample were used as controls to verify the specificity of the utilized α-HA antibody.

This distribution is also visible in the respective TSS-distance plot of Figure 32**c**, where a small upstream binding preference of the proteins is visible, with the majority of peaks being located less than 10 kb away from the TSS. This pattern was found to be largely similar between the species and akin to Ad5 E1B-55K (Figure 21**b**). Next, I used the plotHeatmap tool of the deepTools suite to visualize all identified consensus peaks on the host genome. Even though the absolute number of significant peaks differs drastically, a distinct overlap between the binding patterns of the large E1B species is apparent (Figure 33), where all significant peak regions of either one of the species is occupied by at least one of the others. Here, visible signals on any identified significant peak region also showed enrichment by all the other E1B species, which were absent in both rabbit IgG control and input. These results suggest that the indirect DNA-binding capability of large E1Bs from different species is a general and conserved mechanism.

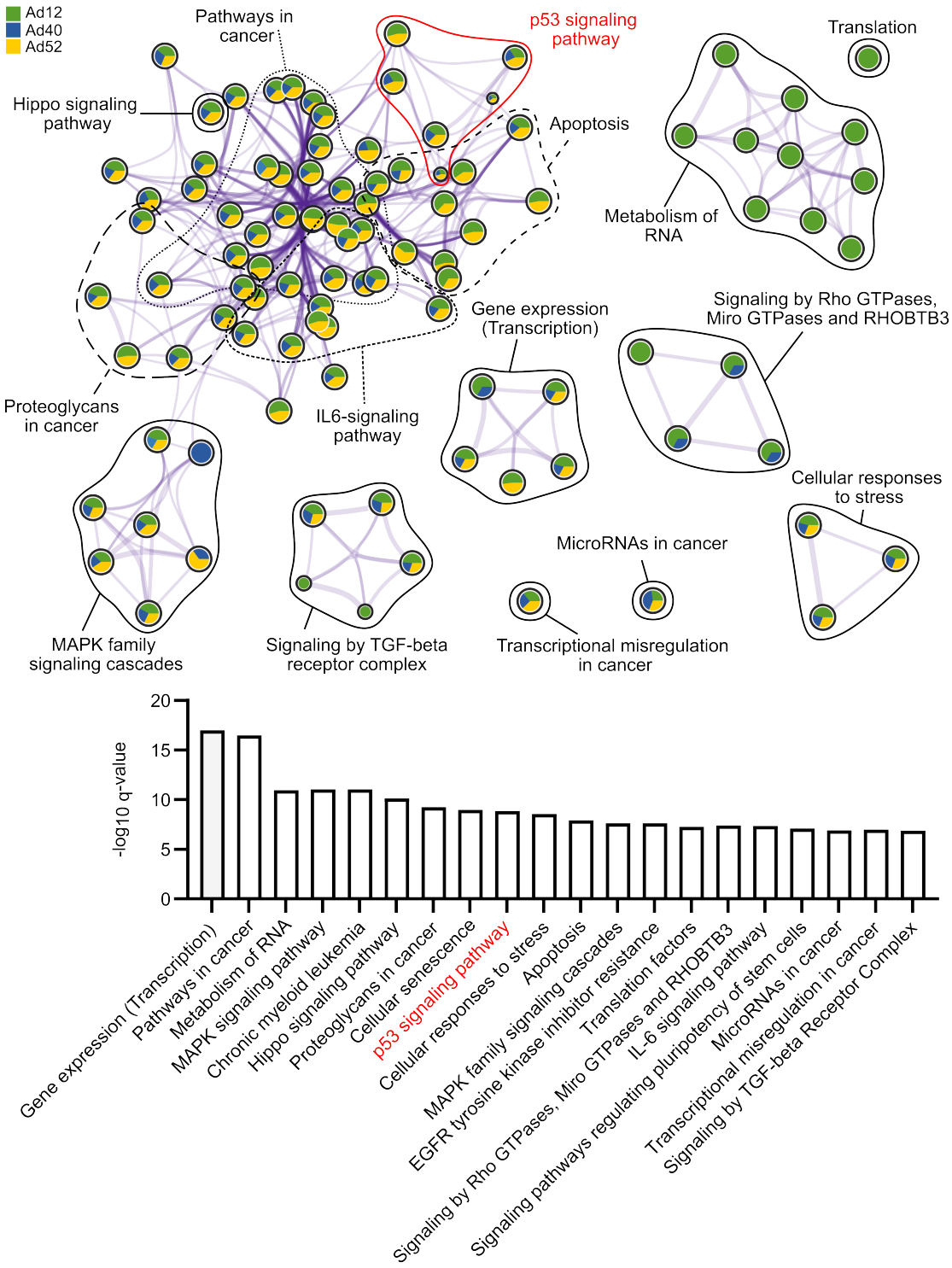


Figure 34: Conserved pathway regulation between E1B-species upon DNA-interaction. Cytoscape network plot visualizing the degree of similarities and differences in overrepresented pathways between individual large E1B species ChIP-seq peak sets, calculated by the metascape analysis tool. Nodes that have a kappa similarity of at least 0.3 are connected by purple lines. The proximity between the nodes is an indicator of redundancy between the gene sets of the nodes, and redundant clusters have been termed by their respective most statistically significant member node. Pathways related to p53 have been marked in red to highlight their involvement. Detected proportions of large E1B species are indicated in the nodes by their respective color. A color code for the nodes is provided in the top left corner. The y-axis represents the -$\log_{10}$ q-value of each pathway.

Biological network analysis allows a more in-depth insight into potential large E1B-influenced pathways through interaction with different transcription factors resulting in indirect binding to DNA. Therefore, the top 3000 significantly deregulated genes were analyzed by the metascape enrichment clustering algorithm and were visualized through a cytoscape plot in Figure 34, showing the -$\log_{10}$ q-values of the individual identified biological KEGG pathways, Reactome gene sets and WikiPathways at the bottom. Thus, a large cluster of closely related pathways comprising ”Pathways in cancer”, ”Apoptosis”, ”Proteoglycans in cancer”, ”IL6-signaling pathway” and ”Hippo signaling pathway” are enriched in any large E1B species ChIP-seq to various degrees. Intriguingly, the ”Metabolism of RNA” and ”Translation” clusters that were enriched solely by Ad12 were also found by Ad5 E1B-55K NES/K104R, illustrated in Figure 24. Additionally, ChIP-seq significantly revealed enriched gene regions for Ad12 and Ad40 but not Ad52 that can be assigned to the ”Signaling by RHO GTPases, Miro GTPases and RHOBTB3” pathway which in line leads to transcriptional repression of genes involved in this pathway (Figure 31**b**). This further establishes a link between DNA-binding and gene repression.

4.2.3 Ad12 large E1B is the most potent DNA-binding transcriptional modulator of all tested E1B-55K species

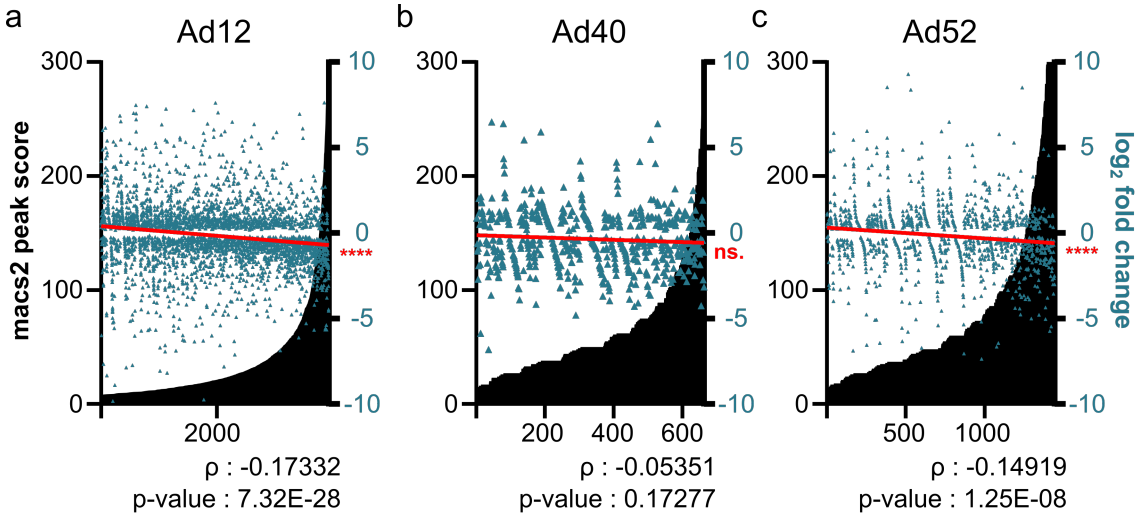


Figure 35: Global relationship between large E1B ChIP-seq peak strength and transcriptional deregulation. **a** Ad12-, **b** Ad40- and **c** Ad52 large E1B macs2-calculated peak score and its relationship with the transcriptional status (blue triangles) of the nearest gene in comparison to the control. Peaks are ranked according to the increasing macs2 peak score of that gene. Only genes with an adjusted p-value (FDR) of ≤ 0.1 were included in this analysis. Linear regression analysis (red line) of $\log_2$ fold changes visualizes correlation. A negative (positive) slope value corresponds to an association between the magnitude of large E1B presence on the DNA and transcriptionally repressed (activated) genes. A Pearson correlation coefficient (ρ) was calculated to quantify the relationship between large E1B binding events and transcriptional consequences of associated genes. A negative ρ-value corresponds to an inverse relationship between peak strength and $\log_2$ fold change, while the p-value presents the calculated significance of ρ with a two-tailed t-test (****$p \leq 0.0001$, ns: not significant).

To further investigate the consequences of indirect DNA-binding by E1B proteins from Ad12, Ad40 and A52 to the DNA, ChIP-seq-associated genes were combined with their respective differential mRNA transcript level. The Ad12 large E1B protein was found to bind directly or adjacent to 49.89% of significant genes (RNA-seq FDR ≤ 0.1) of which 54.84% were downregulated (data not shown), whereas of all differentially expressed genes only 46.73% were transcriptionally repressed. A similar but less pronounced effect was found in the other species: Ad40 with 59.57% of associated genes versus 53.59% of total genes downregulated and Ad52 with 53.68% of associated genes versus 47.97% of total genes downregulated. Global correlation analysis of ChIP-seq and transcriptional deregulation showed that only Ad12 and Ad52 were significantly correlated with negative transcriptional outcomes for associated genes with Pearson correlation coefficients (ρ) of -0.1733 and -0.1492, respectively (Figure 35**a** and **b**), while no significant relationship has been revealed for Ad40 (Figure 35**b**). This correlation was exceptionally significant in the Ad12 large E1B data set with a p-value of 7.32×10^{-28}, followed by Ad52 with a

p-value of 1.25 x 10^{-8}. Interestingly, slight differences between the species of DNA-bound transcription factors became apparent after *de novo* HOMER motif analysis. A collection of the respective top six motifs with their associated best match, p-value, absolute number of promoter- or enhancer-associated peaks and their percentile proportion of downregulated genes is presented in Figure 36.

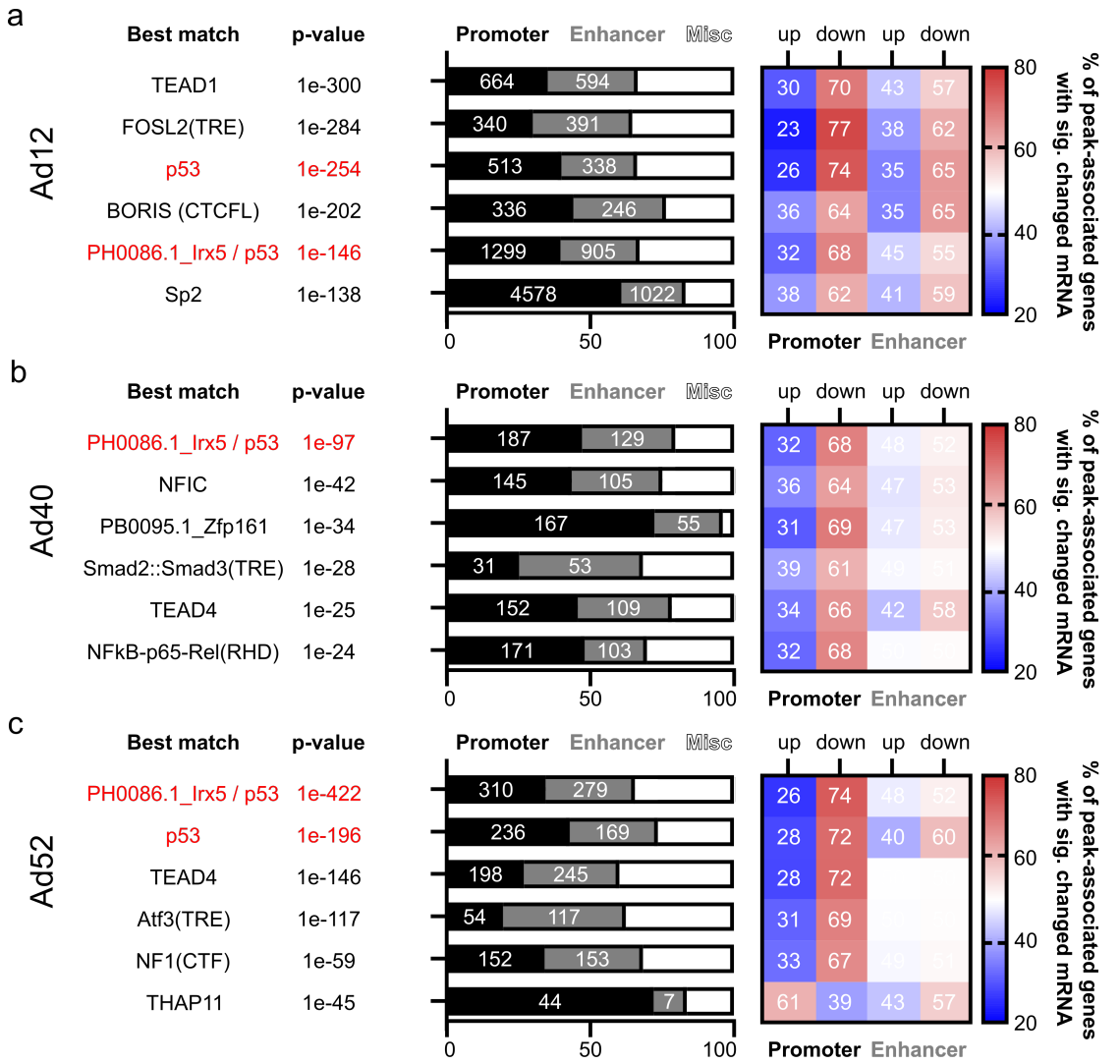


Figure 36: Ad12 and G52 binding to transcription factor motifs lead to significant repression of adjacent genes. After **a** Ad12, **b** Ad40 and **c** Ad52 large E1B motif identification and classification, motifs were separated according to their gene-respective position into either promoter- or enhancer-associated. The ”Misc” genes are intron-, exon-, or UTR-associated. Only motifs with at least 30 individual peaks were chosen for this analysis, individually matched with the mRNA-seq data of their respective nearest gene and separated into either up- or downregulated genes. The color gradient indicates the proportion of genes that occur in either set, ranging from blue (low relative amount) over white (equal amount) to red (high relative amount). The numbers in the bars represent the total number of motifs in either promoter or enhancer peaks.

All three species significantly enriched p53-associated motifs with different p-values, with a preference for promoter-, rather than enhancer regions and related to transcriptional repression in both cases (Figure 36). In contrast to the previous experiment, in

which the TRE-associated motif was only slightly less transcriptionally repressed compared to p53-associated genes with Ad5 E1B-55K (Figure 26**a**), Ad12 displayed the overwhelming proportion of TRE motif-adjacent genes to be transcriptionally repressed, even more so than the genes associated with p53-motifs (Figure 36**a**). All peak sets additionally identified TEAD-associated motifs that are distinctively less significant in Ad5 E1B-55K (Figure 25**b**), and a general correlation between p-value of the identified motif with the degree of transcriptional repression could be observed here, which could not previously be established in Ad5 (Figure 26**b**). Unique to Ad12 was the identification of a motif best matched to the CCCTC-binding factor CTCF and CTCFL (BORIS) transcription factors, the latter scoring slightly higher. CTCF is a multifunctional zinc-finger protein implicated in genome regulation and architecture (reviewed in KIM *et al.* (2015)).

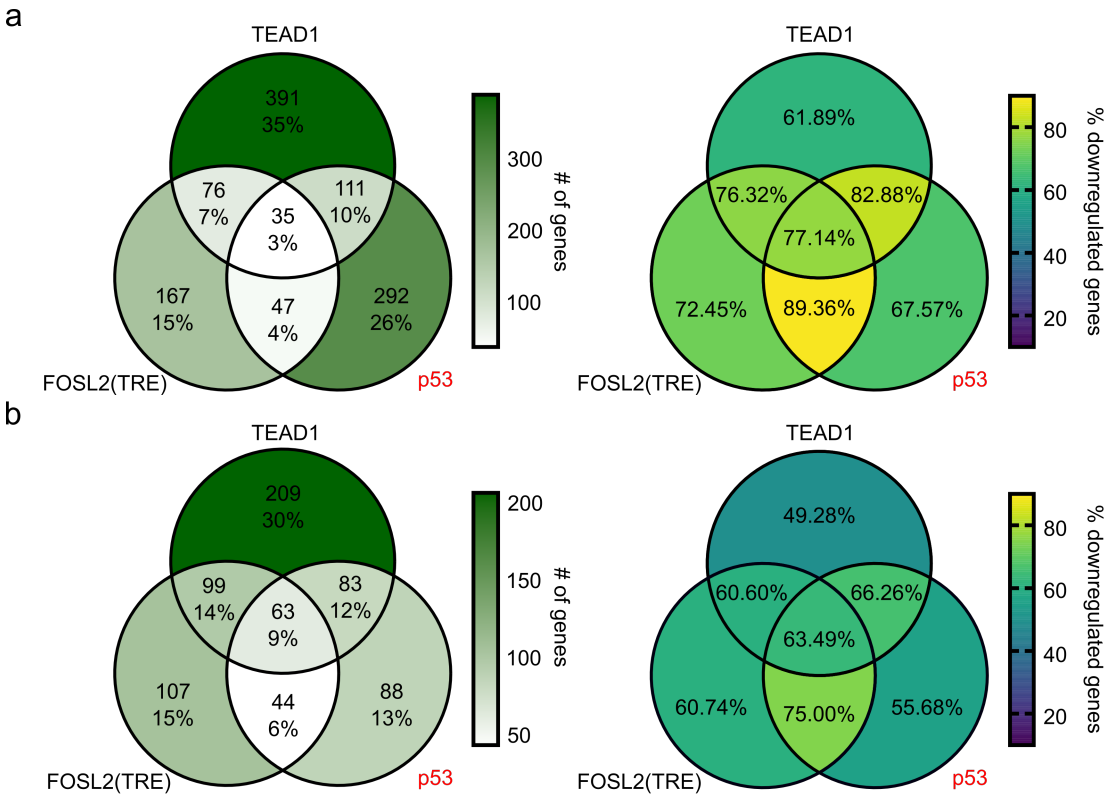


Figure 37: Overlap between motif occurrences in Ad12 large E1B-enriched peaks is strongly associated with transcriptional repression. **a** Promoter- and **b** enhancer-associated co-occurrence motif overlap. On the left side, specific instances of significant motifs were extracted by running the findMotifsGenome.pl script of HOMER and subsequently investigated regarding their overlap with the other motifs via their respective unique peak region ID. mRNA-seq data of the respective nearest gene was combined with the overlap data and illustrated on the right side. The color code ranges from violet (mostly upregulated) to bright yellow (mostly downregulated).

The protein acts as one of the core architectural proteins that helps to establish the organization of the DNA with up to 60.000 potential binding sites identified in mammalian genomes. As the association with Ad12 large E1B was the strongest predictor of tran-

scriptional repression across the largest amount of peaks, I was specifically interested in how extensively its top TEAD-, TRE- and p53 motifs were overlapping in regards to their relationship with the expression of adjacent genes. This overlap was visualized using Venn diagrams in Figure 37. The numerically most prevalent and statistically significant motif, a sequence mostly found on the reverse strand with a p-value of 1 x 10^{-300}, was matched to the TEAD transcription factor. Analysis of the promoter region for TEAD-dependent genes revealed downregulation in 61.89% of cases when isolated. This value increased to 82.88% when found in conjunction with a p53-associated motif. This pattern is present in any case of overlap and most notable in the case of a co-occurance of p53- with a TRE motif, which led to transcriptional repression in 89.36% of the 47 associated nearest genes. The strongest repressive correlation was found in genes with an isolated TRE motif, with downregulation in 72.45% of 167 cases. The difference between promoter- and enhancer-associated motifs that I described earlier in Figure 26, was mirrored for Ad12 with a slight reduction of association between binding and repression in the latter. These findings provide solid proof within a comparatively large data set, that large E1B proteins interact and repress diverse transcription factors.

4.2.4 Ad12 large E1B proteins interact with TEAD4

I performed a ChIP assay by precipitating TEAD4 bound DNA in conjunction with Ad12 and found strong and specific Ad12 signal in the vicinity of the 129 significant TEAD4 peaks, presented in Figure 38**a**.

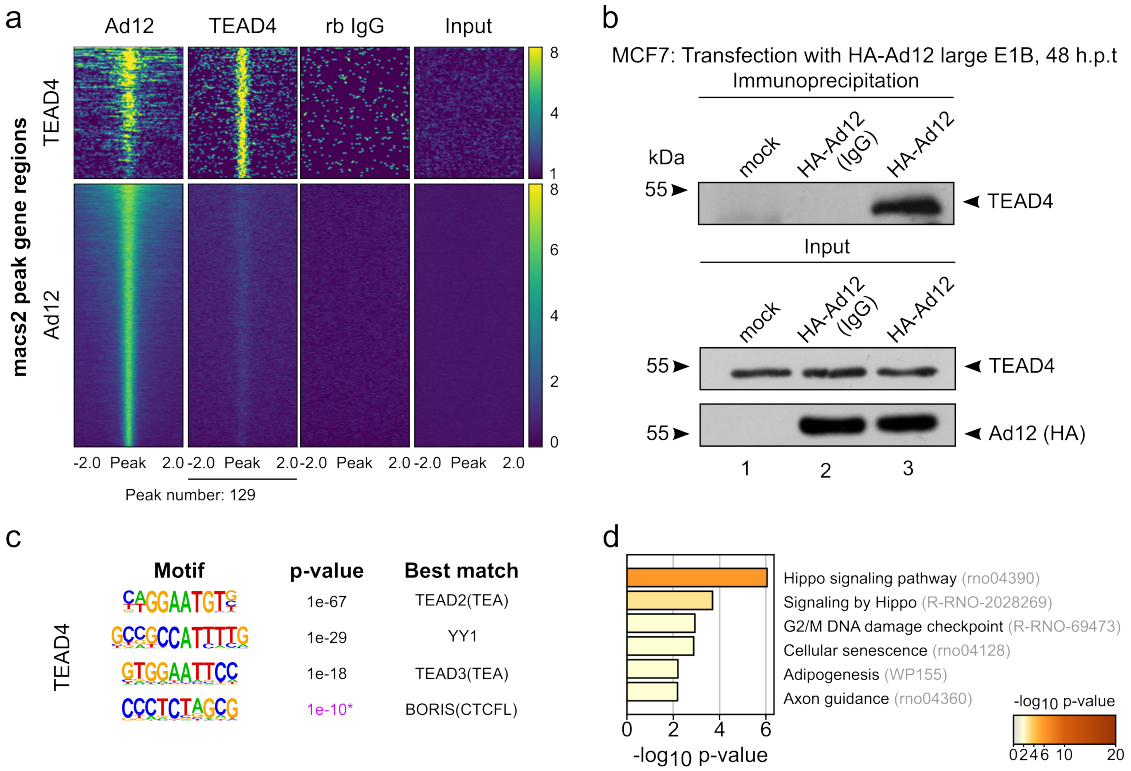


Figure 38: Ad12 large E1B interacts with endogenous TEAD4 in ChIP and co-immunoprecipitation assays. **a** Illustration of TEAD4-ChIP-seq in Ad5 E1A and Ad12 large E1B transformed BRK cells in a 4 kb region around significant peak regions with signal strength ranging from weak (dark blue) to strong (bright yellow). Here, ChIP-seq of an unspecific rabbit IgG antibody (rb IgG) and an input sample were used as controls to verify the specificity of the utilized α-HA or α-TEAD4 antibodies. **b** Western blot of co-immunoprecipitated TEAD4 with transfected Ad12 large E1B (#2217). Cells were harvested 48 hours after transfection and co-immunoprecipitation was performed with the HA antibody (α-large E1B). After 10% SDS-PAGE, proteins were immunoblotted with the aforementioned antibodies. These transfection, co-immunoprecipitation and western blotting experiments were performed by in Eileen Dudda's bachelor's thesis. **c** HOMER *de novo* motif identification of TEAD4-ChIP-seq peak regions with respective p-values. Non-significant motifs are shown in purple. **d** Metascape pathway and process enrichment analyses were carried out with the KEGG Pathway, Reactome Gene Sets and WikiPathways ontology sources on the TEAD4-derived ChIP-seq annotated peak sets. The significant pathways are colored based on their -$\log_{10}$ p-value from grey (0, not significant) to dark orange (20, very significant).

TEAD4-ChIP-seq was successful, as shown by motif enrichment (Figure 38**c**) and biological pathway enrichment (Figure 38**d**), in which the "Hippo signaling pathway" was the most significant hit. In line with the Ad12 ChIP-seq, I found unique enrichment of the BORIS(CTCFL) motif in the ChIP-seq of TEAD4, which is the third most significant motif found in the Ad12 ChIP-seq (Figure 36**a**). Furthermore, transcriptional repression of TEA-

motif associated genes could also be verified as 48 of 81 (59.26%) genes were downregulated with a FDR ≤ 0.1, whereas 46.73% of total significant genes were transcriptionally repressed. Ad12 association with endogenous TEAD4 has also been confirmed by co-immunoprecipitation experiments (Figure 38**b**).

4.3 Analysis of Ad5 E1B-55K in p53-negative H1299 cells reveals p53-independent binding to DNA

Previous experimental setups addressed E1B-55K interaction with DNA-bound transcription factors in the presence of an E1A-activated p53 network. As the interaction between E1B-55K and p53 is remarkably substantial, I was interested in tracking the behaviour of the adenoviral oncogene in the absence of the cellular tumor suppressor. I chose the H1299 cell line for this reason, as it harbors a homozygous partial deletion of p53, completely abolishing expression of the protein (LIN and CHANG, 1996). After transduction and FACS-sorting of fluorophore expressing cells, the positive population was expanded and the expression level of E1B-55K was validated by western blotting (Figure 39**a**). As expected, expression of p53 could not be detected while substantial protein levels of E1B-55K were present in the transduced H1299 cells. I moved on to enrich E1B-55K via its N-terminal HA-tag in two technical replicates from two different passages of the cells (HA-ChIP#1 and -#2). The negative controls for this experimental series consisted of a (i) p53-ChIP to examine potential signals of closely-related paralogs that could bind to E1B-55K and (ii) a second H1299 cell line that was transduced with a E1B-55K plasmid harboring multiple stop codons to prevent protein expression in order to track the background signal of the antibody I used in the human system. After initial quality control and alignment of the FASTQ sequence files to the human reference genome assembly GRCh38.p13, I performed macs2 peak calling and subsequent identification of consensus regions between the replicates with the help of MSPC and validated 1760 binding events. The created peak sets were subsequently annotated by ChIPseeker to determine the relative position and the nearest genes of the individual peaks (Figure 39**b**). It is apparent that the proportion of promoter-associated peaks drastically increased to over 80% in the absence of cellular p53 and adenoviral E1A, in contrast to earlier findings where the values ranged between 20% and 45% (Figure 21**b**). As no p53 is present in this system, E1B-55K is free to bind to any other transcription factors, possibly leading to this observed increase. HOMER *de novo* motif enrichment revealed a predominant presence of THAP11-, TRE- and ZNF143- nuc-

leotide target sequences in the E1B-55K bound peaks (Figure 39**c**). THAP11 and ZNF143 are known to be found in the direct vicinity of each other (VINCKEVICIUS *et al.*, 2015), due to overlap of the ending sequence of the first motif ([...]TTCCC) and the starting sequence of the second motif (TCCC[...]). Intriguingly, no TEAD-associated motifs could be enriched at all in this context.

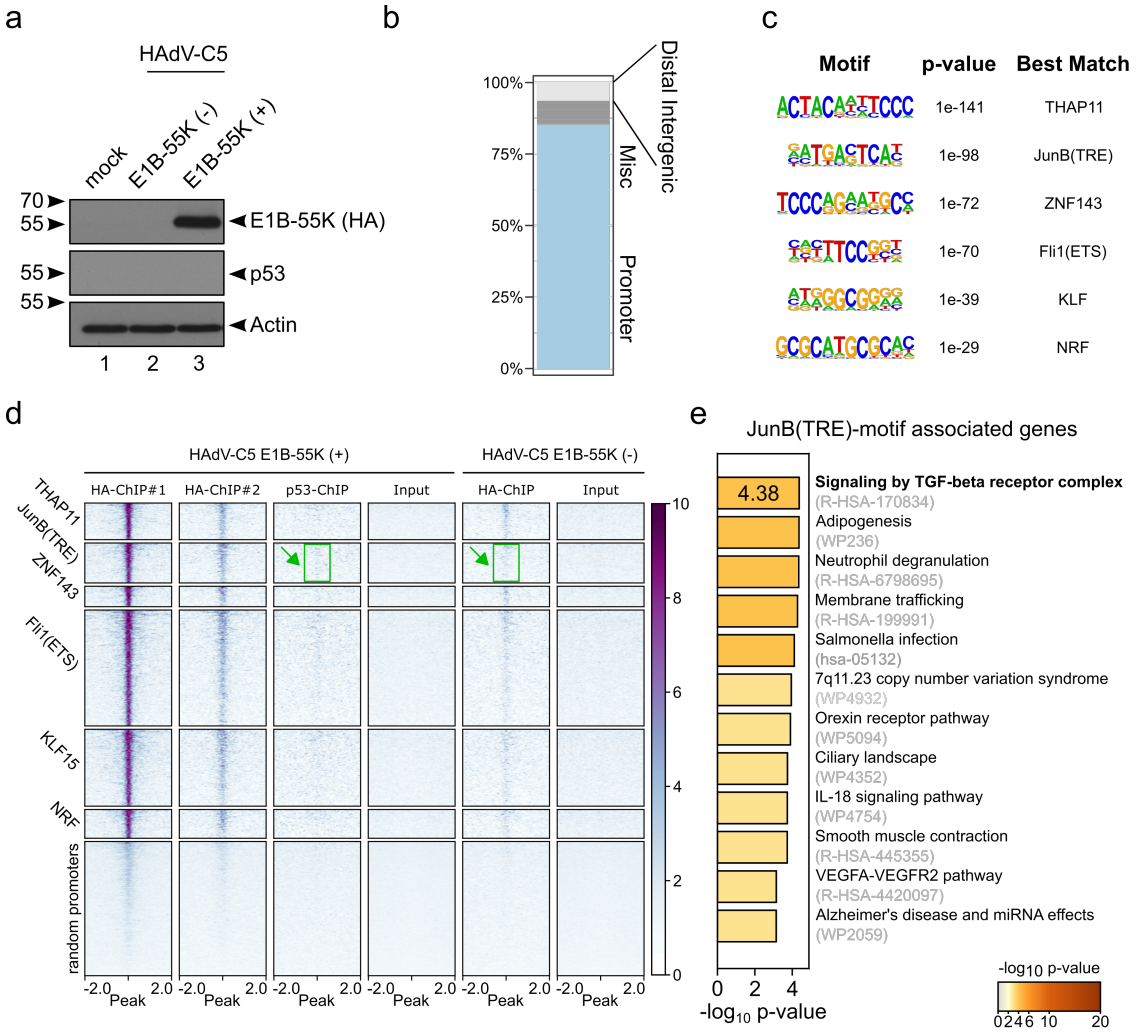


Figure 39: Successful E1B-55K ChIP-seq in the p53-negative H1299 tumor cell line. a Western blot of steady-state levels of Ad5 E1B-55K and p53 in Ad5 E1B-55K transduced H1299 cells. Cells were harvested and immunoblotting was performed with either ChIP-HA (α-Ad5 E1B-55K), DO-7 (α-p53) and A5441 (α-β-actin) antibodies. **b** R ChIPseeker-mediated annotation of verified E1B-55K peaks. **c** Top six motifs of HOMER *de novo* motif analysis of verified E1B-55K peaks. **d** Deeptools plotHeatmap of the top six motif regions and a random human promoter set in the presence and absence of E1B-55K. The green box and arrow annotate a region of suspected weak background signal. Visualized is a range of 4 kb around peaks, sorted by signal strength and ranging from white (weak) to purple (strong). Here, ChIPs of respective input samples were used as controls to verify the specificity of utilized α-HA and α-TEAD4 antibodies. **e** Metascape pathway and process enrichment analyses were carried out with the KEGG Pathway, Reactome Gene Sets and WikiPathways ontology sources on the isolated TRE motif-specific gene set. The significant pathways are colored based on their -$\log_{10}$ p-value from grey (0, not significant) to dark orange (20, very significant).

The obstacle of general THAP11-, AP-1- and ETS- overrepresentation (HUNT and WASSERMAN, 2014) in different ChIP-seq datasets warranted further inquiry into these results. Weak background signals could be found in the vicinity of E1B-55K identified peaks compared to the respective input samples, especially pronounced in the HA-, but also apparent in the p53-ChIP when neither of the target proteins were expressed during the ChIP-seq experiment (Figure 39**c**). This leads to the assumption that there is a natural propensity of these antibodies to bind to similar regions in this context. I suspect unspecific binding of the used antibodies to promoters as regions of high transcriptional activity, an obstacle that is not present in the *Rattus norvegicus* system. As speculated in JAIN *et al.* (2015), many components of the transcription machinery contain unstructured protein surfaces creating interaction-prone regions that could be an unspecific target of antibodies, especially once their native assembly configuration may be disrupted through sonication. I created six different subsets of E1B-55K peaks that predominantly contain the annotated motifs, and found that the background signal was weakly present in all motifs with the exception of TRE motifs, where the background signal was comparatively fainter (Figure 39**c**, green box and arrow). I also generated a set of 5000 random human promoter regions (Figure 39**c**, bottom segment) and identified non-specific signals in a small part of the investigated samples in about equal proportions as found in the separate motif regions, lending credibility of general weak overrepresentation of transcription-active regions in all H1299 datasets. Regardless, in the presence of E1B-55K, this signal was potentiated by multiple orders of magnitude, visible in the first two HA-ChIP-seq experiments. The TRE motif dataset was the least likely to be background-associated, which is why I performed a metascape pathway analysis on this isolated set (Figure 39**d**). Here, the most significant pathway was identified to be ”Signaling by TGF-beta receptor complex” that is closely connected to the AP-1 network (SUNDQVIST *et al.*, 2020) with a $-\log_{10}$ p-value of 4.38, followed by pathways related to the immune system, general presence of infectious agents and others. In summation, the DNA-binding potential of E1B-55K could be readily established in the absence of p53, even when strictly accounting for non-specific antibody background signal. In this system, the most promising interaction target of E1B-55K are members of the AP-1 complex targeting the TRE motif, similarly to previous observations in the *Rattus norvegicus* system.

4.4 Effect of Ad5 E1B-55K expression on the epigenetic landscape of human primary mesenchymal stromal cells

During my doctoral work, I was primarily interested in investigating the PTM of histone tails, as prior work presented here suggested a direct influence of the adenoviral E1B-55K oncogene on the host genomic DNA. Following the work of SPEISEDER *et al.* (2017), who could show efficient transformation of hMSC through continuous expression of the adenoviral E1A- and E1B-55K oncogenes, my goal was to explore if and how the latter protein might sway the epigenetic landscape of a primary human cell system. I applied our established strategy of stable oncogene expression via lentivirus-mediated transduction, similar to the previous BRK cell transformation assays.

4.4.1 Transduction with adenoviral E1A and E1B oncogenes in combination with FACS-based selection allows efficient transformation of hMSC

In general, hMSCs were transduced with Ad5 oncogenes. Cells that showed expression of both fluorescent markers were reseeded and expanded after FACS selection. HMSC from three male donors of different ages, representing biological replicates, were thereby transduced with differing combinations of oncogenes: Either empty vectors (short: mock), E1A only, E1A and E1B-19K (short: 55K-, containing multiple stop codons to prevent expression of E1B-55K) or full-length E1A and E1B gene regions (short: 55K). As expected, the first two conditions were exceedingly difficult to maintain in cell culture, as they either stopped replicating or rapidly underwent apoptosis after FACS. Even though multiple attempts were undertaken, I was only able to readily and repeatedly create transformed cultures of E1A + E1B-19K and full-length E1A + E1B of all three replicates, while expression of E1A resulted in partial immortalization and sustained expansion of up to 30 passages in only a single specimen. Mock-transduced cells were undergoing senescence immediately after transduction, hindering any further passaging or expansion. My main interest was rooted in analyzing E1B-55K-dependent specific epigenetic changes, explicitly regarding its effect on histone PTM and associated gene mRNA expression. In order to potentially enhance the specificity of α-histone PTM antibodies, I chose to apply an adapted native MNase-ChIP protocol as described by BRIND'AMOUR *et al.* (2015). This method allowed me to work with smaller amounts of cellular material, antibodies and reagents compared to the conventional formaldehyde-fixed X-ChIP protocol that was used in 4.1.1.

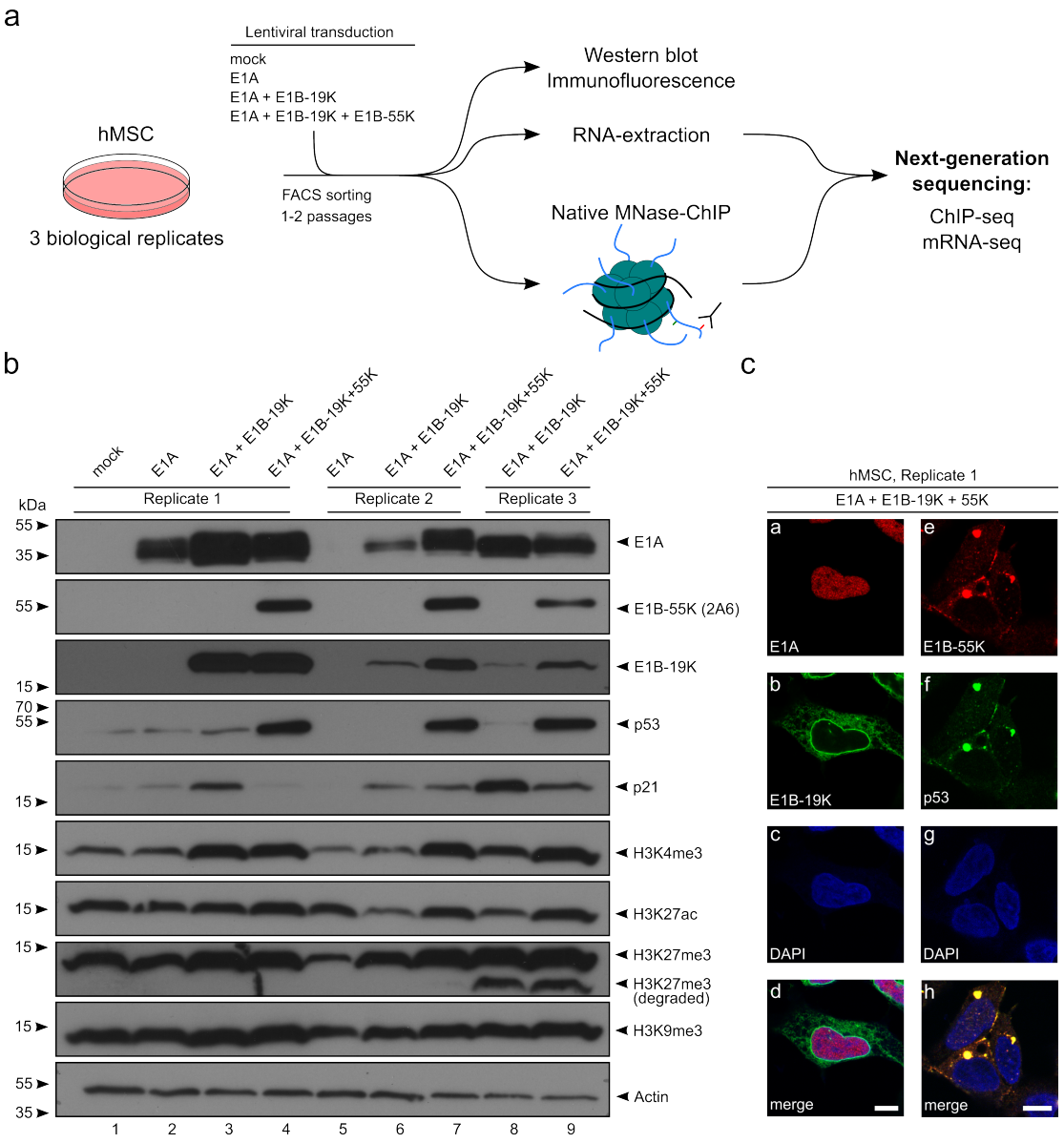


Figure 40: Overview of the hMSC workflow and verification of protein expression. a Schematic workflow of the lentiviral transduction of hMSC with subsequent native MNase-based ChIP-seq using different α-histone PTM antibodies. **b** Western blot showing steady-state levels of several early passage hMSC. Cells were harvested and immunoblotting was performed with either 2A6 (α-Ad5 E1B-55K), 490 (α-Ad5 E1B-19K), DO-7 (α-p53) and A5441 (α-β-actin) antibodies. Specific histone PTM steady-state levels were investigated by α-H3K4me3 (#634), α-H3K27ac (#653), α-H3K9me3 (#637) or α-H3K27me3 (#635) antibodies. **c** Exemplary immunofluorescence plot visualizing the expression patterns of selected viral and cellular targets of E1A and full-length E1B-transduced hMSC from an advanced passage. Antibodies used here are similar to the ones used in the western blot. The scale bars correspond to 10 μm.

A scheme of the workflow is presented in Figure 40**a** and initial protein expression of the earliest passages is visualized using western blotting in Figure 40**b**. Slight alterations in cellular protein p53 and p21 expression are expected regarding the inherent bias introduced by using hMSC from three different biological donors, while differences in viral proteins can vary as these replicates were produced via three independent transduction

experiments. Here, p21 (also known as CDKN1A) was chosen as a primary target for p53-mediated transcriptional activation. Interestingly, only replicate 1 and 3 display a strong reduction of p21 expression in the presence versus absence of E1B-55K (Figure 40**b**, lane 3-4 and 8-9) while replicate 2 shows no discernible difference, albeit with amplified levels of p53 steady-states when E1B-55K is present (Figure 40**b**, lane 6-7). Repression of the p53 pathway is still expected in all three replicates, indicated by their generally elevated p53 steady-state levels. This is a potential side-effect introduced by E1B-55K-mediated abrogated expression of MDM2, an E3 ubiquitin ligase that targets and degrades p53. I chose to only perform MNase-ChIP on the first and third replicate, as I expected the repressive effect of E1B-55K to be the strongest in these cells, based on this western blot, as p21 was visibly less expressed in contrast to E1B-55K-negative cells. Global tested histone PTM levels were found to be largely comparable between the different conditions. Notably, in the third replicate, H3K27me3 showed a secondary band slightly below the main band, which potentially shows slight amounts of degradation. Immunofluorescence was used to identify whether localization of adenoviral proteins differs from the expected patterns observed in literature. I found that all proteins expressed from the E1A and E1B gene regions were localized in hMSC as expected with E1A being diffusely distributed in the cellular nucleus (Figure 40**c**, a), E1B-19K found to be associated with cytoplasmic and nuclear membranes (Figure 40**c**, b) and both E1B-55K and cellular p53 found either in a perinuclear dot or distributed sparsely in the nucleus, identically to other established cell lines (Figure 40**c**, e and f) (PERSSON *et al.*, 1982; LYONS *et al.*, 1987; CHIOU *et al.*, 1994; KRÄTZER *et al.*, 2000; KINDSMÜLLER *et al.*, 2007).

4.4.2 Identification of a significant E1B-55K-dependent influence on activating H3K4me3 and H3K27ac histone PTMs

To analyze, whether expression of the postulated transcriptional repressor E1B-55K would lead to specific changes in the canonically activating histone PTMs, I analyzed activating histone H3 lysine 4 trimethylation (H3K4me3) and histone H3 lysine 27 acetylation (H3K27ac). Both modifications are predominantly found in promoter regions of actively transcribed genes and marked nucleosomes are detected as relatively narrow and sharp peaks in the vicinity of TSS. After transduction, FACS and MNase-based native ChIP from two of the three biological replicates, I verified the success of the experiment via qPCR (see 4.1.1) and sent the samples to the in-house NGS facility where deep sequencing was performed with the *Illumina NextSeq 500* to a depth of approximately $20*10^6$ reads. In

the next step, a bioinformatic pipeline consisting of aligning the FASTQ sequence files to the human reference genome assembly GRCh38.p13, setting a read quality filter and removing duplicated reads was undertaken.

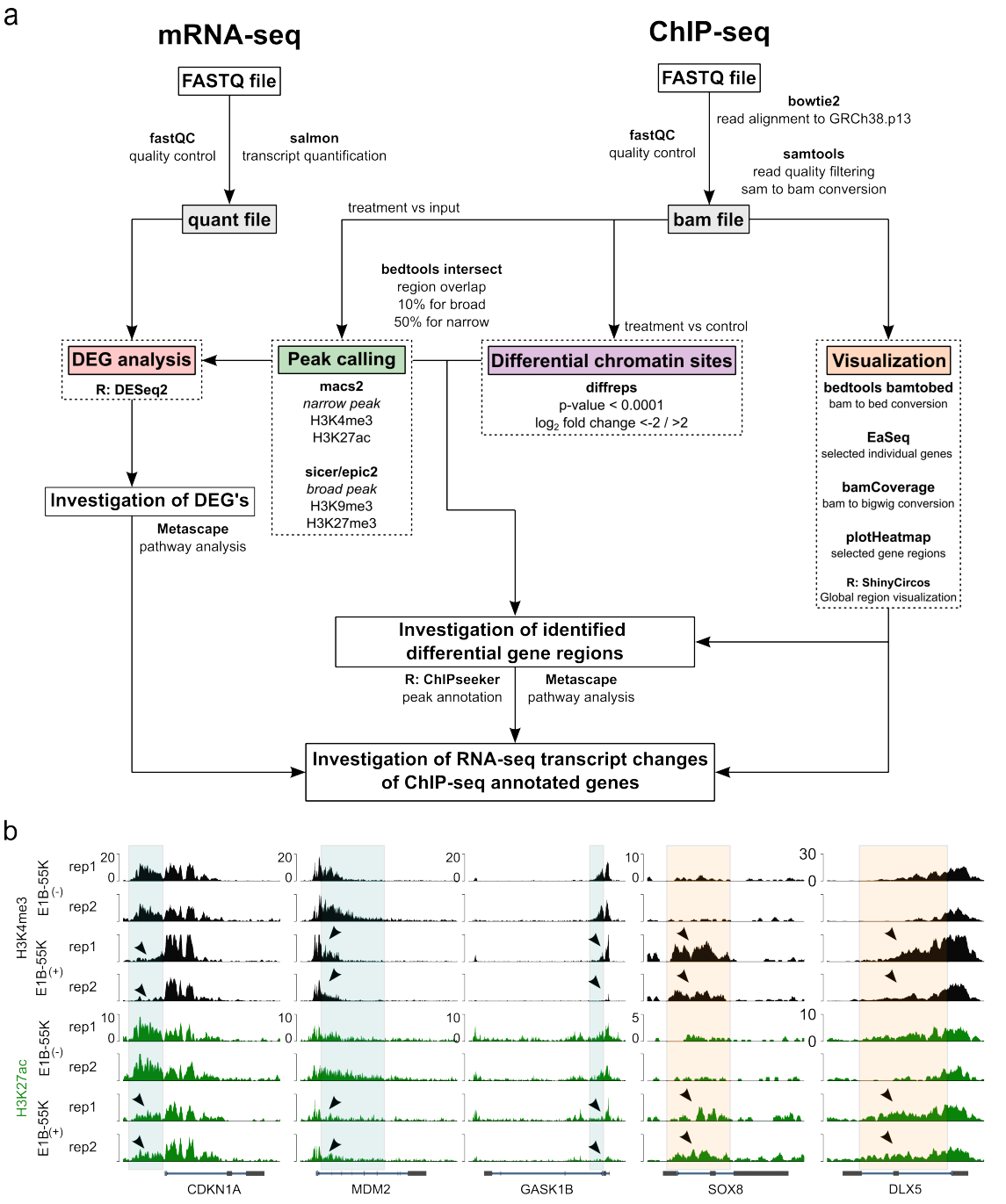


Figure 41: Initial observations of substantial E1B-55K-dependent gain and loss of activating histone PTM signal. a Flowchart diagram displaying the utilized bioinformatic software for DNA sequence alignment, narrow peak calling, visualization and R packages used for functional assessment, as well as software used for differential gene expression analyses. **b** Visualization of five identified selected differential regions, three of which have lost both H3K4me3 and H3K27ac signal (green boxes, arrowheads) and two of which have gained signal (orange boxes, arrowheads). The y-axis indicates the respective ChIP-seq signal.

Afterwards, narrow peak calling, peak-verification with MSPC and overlap-analysis of

these regions with the identified differential region signal strength (diffReps) between 55K+ and 55K- transduced hMSC was performed (Figure 41**a**). In total, 250 promoter regions that gained H3K4me3 signal in the presence of E1B-55K could be identified, while 441 regions displayed lost signal. Regarding H3K27ac signal strength, 55 regions were found that have gained-, while 99 have lost signal. A more detailed analysis of selected significant differential regions is visualized in Figure 41**b**. As examples, the CDKN1A, MDM2 and GASK1B gene regions are shown regarding their respective H3K4me3 (black) and H3K27ac (green) signals. In either one of them, a significant drop in signal strength is visible in the presence of E1B-55K, indicated with arrows and highlighted with colored boxes. Besides, gained signal strength is visible in the SOX8 and DLX5 regions, also indicated by arrows. This observation indicates E1B-55K-dependent direct or indirect regulation of histone PTM. A global distribution of the differential histone PTM regions is illustrated using a circos plot in Figure 42, showing the wide dispersion of the regions across the human chromosomes. Here, each dot indicates a significantly changed region annotated with the respective $\log_2$ fold change of the histone PTM signal in the presence versus the absence of E1B-55K into the orange- (upregulated) or green sector (downregulated).

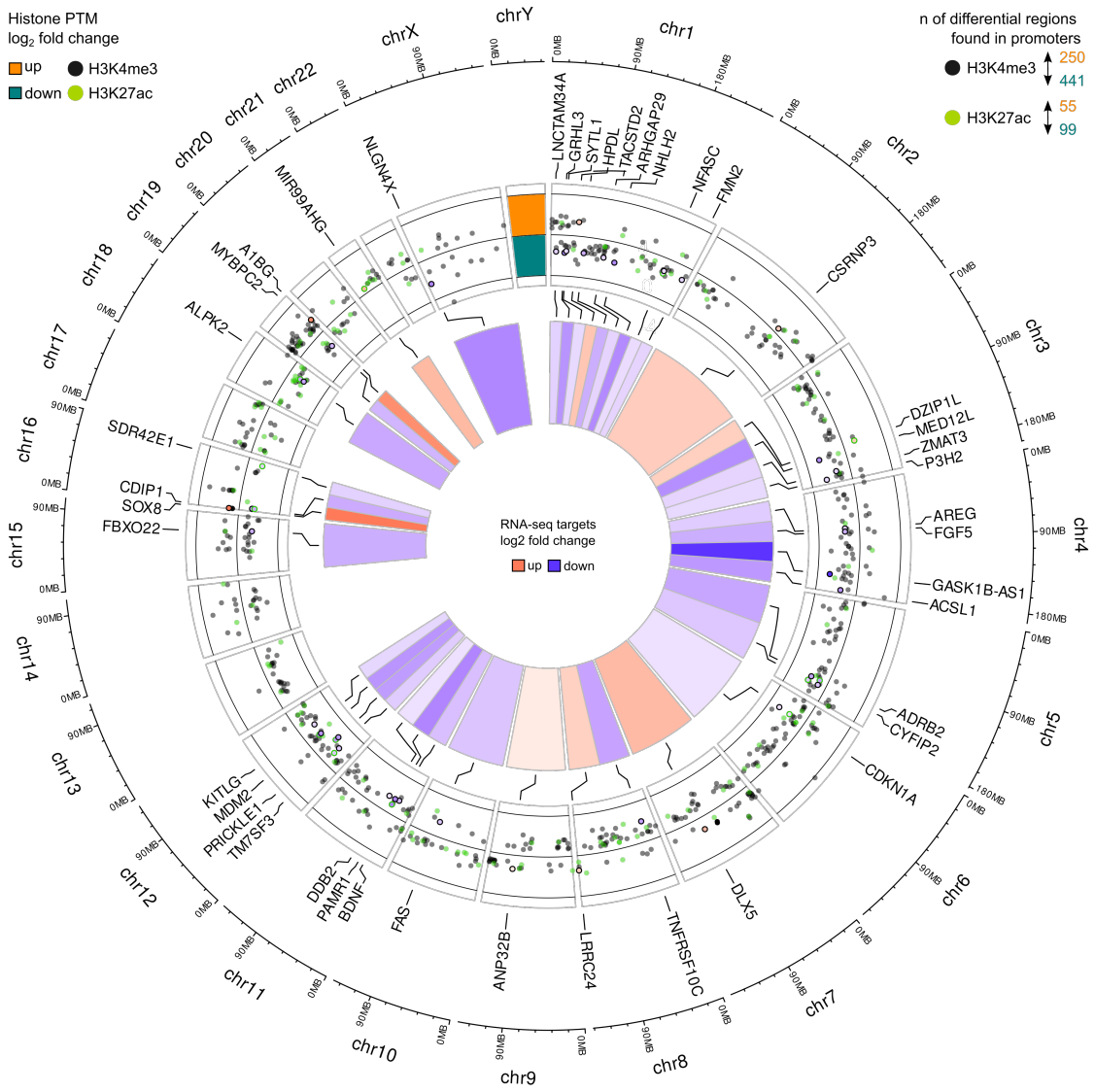


Figure 42: Genome-wide distribution of significant H3K4me3 and H3K27ac histone PTM changes. The circos plot shows the genome-wide significantly changed H3K4me3 and H3K27ac PTMs and their relative chromosomal position. Respective peak regions were identified using macs2 with narrow settings and potential differential values between 55K+ and 55K- were determined with the diffReps software. Annotated genes in the inner circle were found to have significantly changed mRNA expression levels (p-value ≤ 0.01). The absolute number of changes in promoters is annotated in the top right corner of the plot.

Here again, the regions were connected with the respective transcriptomic changes between the two conditions in order to contextualize the effect of differential histone PTM signals regarding its effect on the expression of the associated genes (Figure 45**a** and **b**), only allowing genes that have been changed significantly (p-value ≤ 0.01). Each dot represents a gained signal (orange) or lost signal (green) with its respective H3K4me3 or H3K27ac $\log_2$ fold change and the mRNA $\log_2$ fold change of its associated gene (black triangle). This representation shows the clear overabundance of repressed regions compared to regions with gained signal and visualizes that in some cases (e.g. ch13 and

chr21), whole regions containing multiple genes were either up- or downregulated. Next, to investigate the functional consequences of E1B-55K-induced activating histone PTM deregulation, the gained or loss histone PTM regions were overlain with the respective mRNA $\log_2$ fold change of the nearest genes (Figure 43). Heatmaps visualize all significantly changed H3K4me3 (Figure 43**a**) and H3K27ac (Figure 43**b**) PTMs in promoter regions in a 20 kbp window, sorted according to their respective mRNA $\log_2$ fold change from positive to negative values. I separated the differential regions into upregulated or downregulated groups, representing all regions that have either a positive or negative mRNA $\log_2$ fold change. A ratio-metric heatmap plot that compares the relationship between signal intensity and mRNA $\log_2$ fold changes of associated nearest genes on the right shows that the signal ratio ($\log_2$ ratio) in both H3K4me3 and H3K27ac correlates with the transcriptional outcome. A read count intensity blot with signal from all regions superimposed on the y-axis from either one of these groups is presented directly next to the heatmaps, with an average signal plot directly below, which is presented by a black line (μ), with up to two standard deviations (σ) visualized in shades of grey, capturing around 95% of all regions. The signal plots were overlapped on the far right, showing that E1B-55K-expressing hMSC (55K+, black track) display a stronger activating histone PTM signal on average in the upregulated genes and vice versa for downregulated genes, compared to E1B-55K-negative hMSC (55K-, orange track). As H3K27ac histone PTM is present on both promoters and enhancer regions, I created a set of differential regions that overlap with a human enhancer data set predicted by the FANTOM5 international research consortium (ANDERSSON *et al.*, 2014; LIZIO *et al.*, 2015) to investigate the observed difference between 55K+ and 55K- (Figure 43 and 44). A certain degree of unreliability comes from allocating regions that are potentially many kbp up- or downstream to a gene based on distance alone, which likely introduced a small bias. Yet, H3K27ac signal loss was related to a drop in $\log_2$ fold change of the nearest genes (Figure 44), just like observed in the promoter-associated regions.

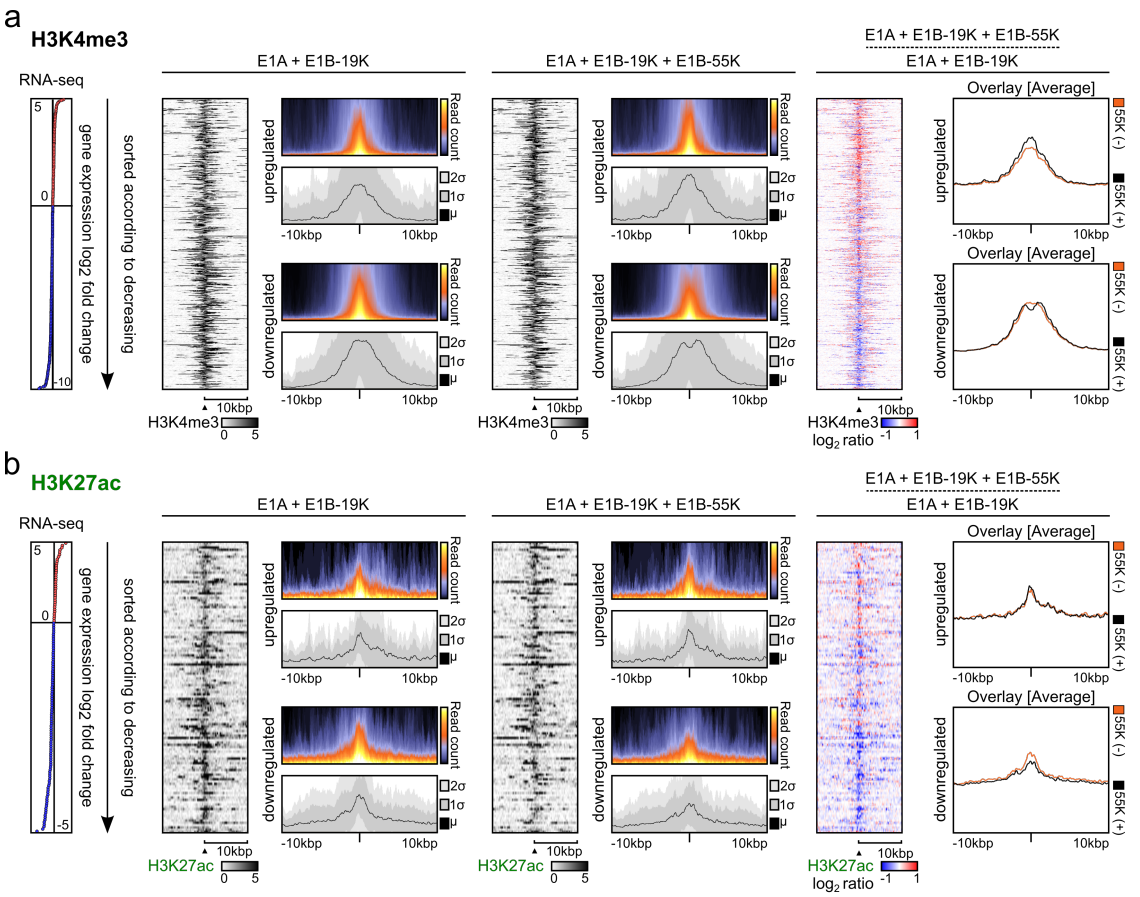


Figure 43: Significant H3K4me3 and H3K27ac differences in promoters predict associated gene transcriptional changes. MNase ChIP heatmaps of **a** H3K4me3 and **b** H3K27ac displaying all promoter regions identified by macs2-, diffReps- and MSPC in the comparison of 55K+ versus 55K-. In general, significantly different regions in the presence of E1B-55K are represented as the y-axes in the heatmaps, around a 20 kbp region. These regions were overlain with the RNA-seq $\log_2$ fold changes of the nearest genes, sorted from positive (red) to negative (blue) values, visualized in the dot plot on the right side of the figure with boundary values depicted at the edges. Regions that had a positive (negative) gene mRNA $\log_2$ fold change were gated into the upregulated (downregulated) subgroups and investigated regarding their respective read count signal intensity, which is superimposed on the y-axis. The average track (μ) of this projection is annotated directly below with up to two standard deviations (σ), capturing up to 95% of the signal from all regions, visualized as a color gradient. Illustrated on the right side is a ratiometric plot of the difference between the presence and absence of E1B-55K, with calculated signal intensity values depicted as the $\log_2$ ratio from blue (negative) over white (unchanged) to red (positive). A projection of the two average tracks from either upregulated- or downregulated groups is depicted on the far right. The black track presents the signal in hMSC that express E1B-55K while the orange track presents the signal in hMSC that lack E1B-55K.

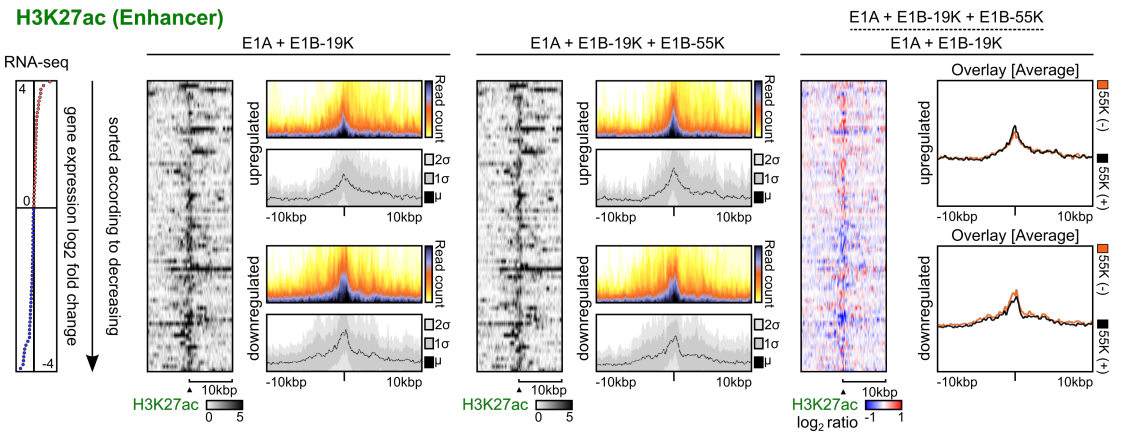


Figure 44: Significant H3K27ac differences in enhancer regions predict transcriptional changes of associated genes. MNase ChIP heatmaps displaying all regions identified by macs2-, diffReps- and MSPC that overlap with FANTOM5-predicted enhancer regions. Significantly different regions in the presence of E1B-55K are represented as the y-axes in the heatmaps, around a 20 kbp region. These regions were overlain with the mRNA-seq $\log_2$ fold changes of the nearest genes, sorted from positive (red) to negative (blue) values, visualized in the dot plot on the right side of the figure with boundary values depicted at the edges. Regions that had a positive (negative) gene mRNA $\log_2$ fold change were gated into the upregulated (downregulated) subgroups and investigated regarding their respective read count signal intensity, which is superimposed on the y-axis. The average track (μ) of this projection is annotated directly below with up to two standard deviations (σ), capturing up to 95% of the signal from all regions, visualized in shades of grey. Illustrated on the right side is a ratiometric plot of the ration between the presence and absence of E1B-55K, with calculated signal intensity values depicted as the $\log_2$ ratio from blue (negative) over white (unchanged) to red (positive). A projection of the two average tracks from either upregulated- or downregulated groups is depicted on the far right. The black track presents the signal in hMSC that express E1B-55K while the orange track presents the signal in hMSC that lack E1B-55K.

The Pearson coefficient (ρ) was calculated between these two numerical values, revealing that both activating histone marks and their associated gene expression were, as expected, moderately positively correlated with a ρ of 0.4201 (p-value 3.32 x 10^{-28}, Figure 45**a**) and 0.4122 (p-value 1.47 x 10^{-6}, Figure 45**b**), respectively. In the next step, I combined the genes that have gained or lost either or both H3K4me3 and H3K27ac with differentially expressed genes (adjusted p-value ≤ 0.01) and performed a metascape-based pathway analysis (Figure 45**c**). Interestingly, in presence of E1B-55K, many of the upregulated H3K4me3 signal-associated genes are categorized in diverse immune- and infection related pathways (Figure 45**c**, left side), like "Interferon type I signaling pathways", "IL-18 signaling pathway" and "miRNA involvement in the immune response in sepsis" - although no clear connection with upregulated mRNA could be identified. Pathways related to cellular metabolism and ribosomal biogenesis were also found to be enriched with H3K4me3 modification, a potential side effect of a non-functional p53 pathway, allowing cells to retain a mitogenic state explained by the increased need for enhanced stimulation of protein synthesis (DONATI *et al.*, 2012).

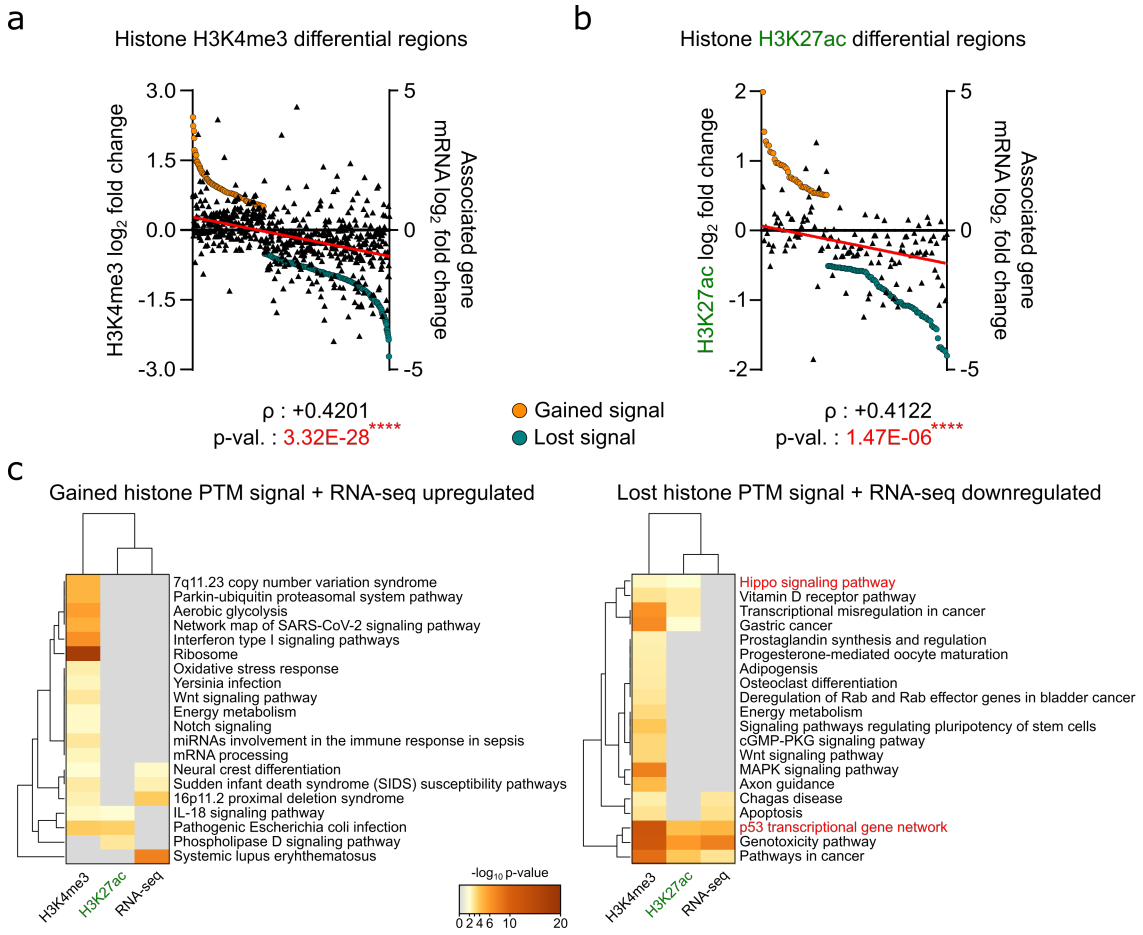


Figure 45: Quantitative relationship between differential H3K4me3 and H3K27ac PTM signal and gene mRNA expression changes. Ranked differential **a** H3K4me3 and **b** H3K27ac PTM regions combined with mRNA $\log_2$ fold changes of associated genes (black triangles) in the comparison between E1B-55K and its negative control. Linear regression analysis (red line) of mRNA $\log_2$ fold changes is included in the figure to visualize a correlation. The Pearson correlation coefficient (ρ) was calculated to quantify the relationship between histone PTM- and mRNA $\log_2$ fold changes. A positive ρ-value corresponds to a correlative relationship between both values, while the p-value presents the calculated significance of ρ using a two-tailed t-test (****$p \leq 0.0001$). **c** Metascape was used with genes that are either upregulated (left side) or downregulated (right side) in hMSC in the comparison between E1B-55K and its negative control. As a cut-off, only genes with an adjusted p-value of $\leq$ to 0.01 were included in this analysis. Pathway and process enrichment analyses were carried out with the KEGG Pathway, Reactome Gene Sets and WikiPathways ontology sources. Pathways involving p53 or Hippo signaling have been marked in red for the purpose of clarity. The significant pathways are colored based on their -$\log_{10}$ p-value from grey (0, not significant) to dark orange (20, very significant).

Besides, the most significant and prominent cluster of downregulated transcripts that correlates with loss of H3K4me3 and H3K27ac PTMs (Figure 45**c**, right side), consists of the closely-connected ”p53 transcriptional gene network”, ”Pathways in cancer” and ”Genotoxicity pathway”. The ”Hippo signaling pathway”, regulated by the TEAD transcription factor family, was also found to have less signal of both H3K4me3 and H3K27ac in promoter regions, but with no significant consequences to the transcriptome.

4.4.3 Early changes of polycomb-associated marker H3K27me3 and heterochromatin marker H3K9me3 are not associated with significant changes to mRNA expression

In the previous part, I primarily focused on histone PTMs that are traditionally associated with euchromatic, usually transcriptionally activated, regions of the genome. In the next section, I investigated whether short time expression of E1B-55K would induce changes in the host histone PTMs associated with polycomb-repression or constitutive heterochromatin.

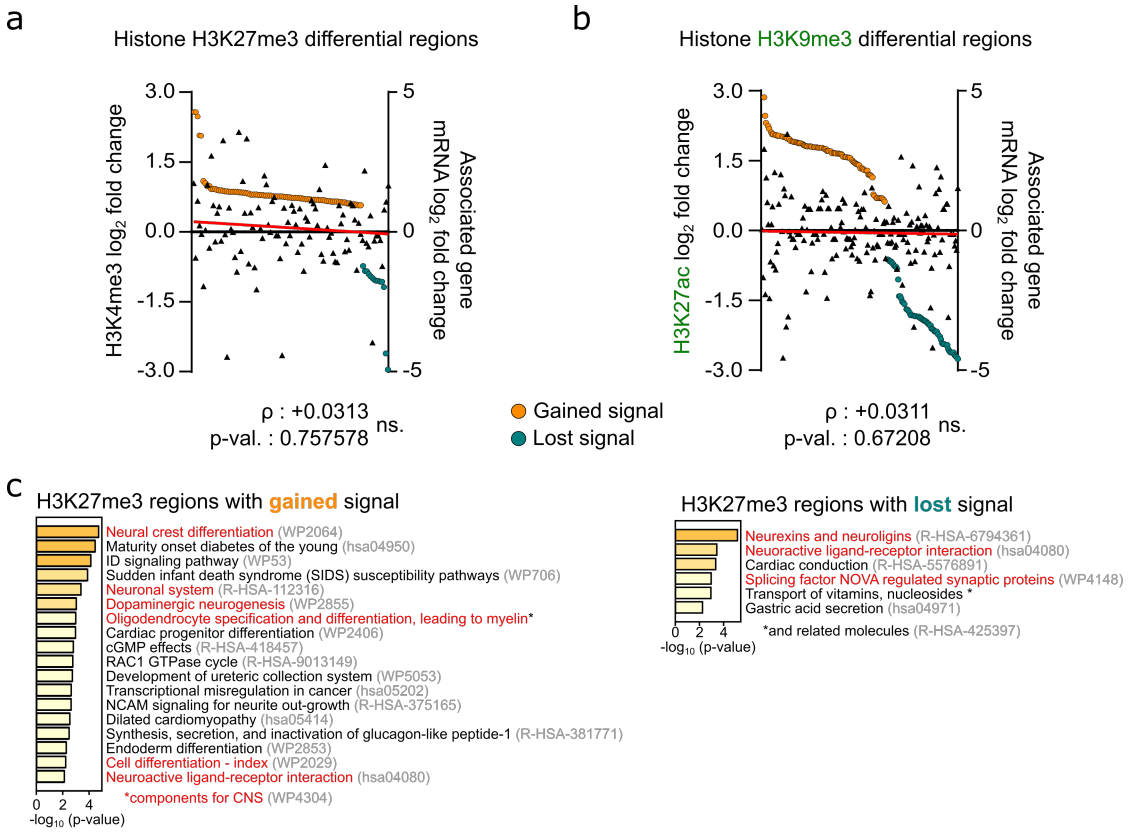


Figure 46: Changes to histone H3K27me3 and H3K9me3 PTM are not associated with differential mRNA expression but hint towards a neurogenic transformation process. Ranked differential **a** H3K27me3 and **b** H3K9me3 PTM regions combined with mRNA $\log_2$ fold changes of associated genes (black triangles) in the comparison between presence and absence of E1B-55K. Linear regression analysis (red line) of mRNA $\log_2$ fold changes is included in the figure to visualize a correlation. The Pearson correlation coefficient (ρ) was calculated to quantify the relationship between histone PTM and mRNA $\log_2$ fold changes. A positive ρ-value corresponds to a correlative relationship between both values, while the p-value presents the calculated significance of ρ using a two-tailed t-test (ns: not significant). **c** Metascape was used with genes that are either upregulated (left side) or downregulated (right side) in hMSC in the comparison between the presence and absence of E1B-55K. As a cut-off, only genes with an adjusted p-value of ≤ 0.01 were allowed in this analysis. Pathway and process enrichment analyses were carried out with the KEGG Pathway, Reactome Gene Sets and WikiPathways ontology sources. Pathways related to neurological mechanisms have been generally marked in red to highlight involvement. The significant pathways are colored based on their -$\log_{10}$ p-value from grey (0, not significant) to dark orange (20, very significant).

Histone H3 lysine 27 trimethylation (H3K27me3) marks facultative heterochromatin and is mediated through the PcG proteins which are crucial for developmental gene expression and maintenance of cell-type identity (reviewed in IRAGAVARAPU *et al.* (2021)). This form of heterochromatin preserves the potential to transition into euchromatin when specific signals are met (reviewed in ZYLICZ *et al.* (2020)). Histone H3 lysine 9 trimethylation (H3K9me3) is regulated through the SUV39H1/2 enzyme, which plays a vital role in the organization of constitutive heterochromatin (reviewed in PENAGOS-PUIG and FURLAN-MAGARIL (2020)). I hypothesised that E1B-55K could potentially interfere with the establishment or maintenance of both facultative and constitutive heterochromatin, with a focus on the former as it represents a more dynamic chromatin state. The majority of H3K27me3 and H3K9me3 PTM are distributed on broad and diffuse regions across the genome, justifying the usage of sicer/epic2 as a peak calling software in this case (STOVNER and SÆTROM, 2019) (Figure 41**a**). These peak regions were overlain with diffReps-identified differential regions, similarly to the other PTM, and combined with the mRNA-seq data of the nearest genes. The Pearson coefficient was calculated from these regions in the subsequent step (Figure 46**a** and **b**). This shows that there was no clear relationship between gain or loss of these histone PTMs and transcriptomic consequences for associated genes in this experiment. Nonetheless, I performed metascape biological pathway enrichment on the associated genes. Interestingly, regarding H3K9me3, no significant enrichment of any pathway could be achieved. When it comes H3K27me3, many pathways were significantly enriched in either direction that suggest involvement of neurological differentiation programs (Figure 46**c**, marked in red). However, whether these changes were induced directly by E1B-55K, akin to previous findings, remains unclear.

Chapter 5

Discussion

5.1 Ad5 E1B-55K represses DNA-bound p53 *in vivo*, independent from its SUMOylation state

Viruses are obligate intracellular parasites that have evolved various diverse responses to cellular antiviral systems that are inevitably induced by the infectious cycle (STERN-GINOSSAR *et al.*, 2019; TENTHOREY *et al.*, 2022). Ultimately, all viruses are completely reliant on the translation machinery of the host to produce the proteins required for viral particle production. In order to establish efficient replication, they have to exert some level of control over the signaling networks and pathways of the host infected cell. These range from the innate immune system and stress-activated host defenses to metabolic pathways and apoptotic responses, among many others. In LIU *et al.* (2020), the authors review multiple previously established viral transcriptional regulator (vTR)s, including the adenoviral E1A-, the EBV EBNA1- and the KSHV LANA proteins. Some of these proteins are able to directly interact with DNA through their DNA-binding domains, like the herpes simplex virus (HSV) infected cell protein (ICP)4 (TUNNICLIFFE *et al.*, 2017), while others interact indirectly with DNA through host chromatin remodeling complexes, like the Epstein-Barr virus nuclear antigen (EBNA)1 protein (WU *et al.*, 1996). One important host pathway induced by the aforementioned viruses is the transcriptional activator p53 that plays a key role in a multitude of host defense pathways and therefore has to be deregulated, inactivated or degraded by numerous viruses in some way or another (SATO and TSURUMI, 2013; ALONI-GRINSTEIN *et al.*, 2018). Prolonged interference can lead to unregulated proliferation of infected cells and thereby induces cellular transformation and subsequent tumor growth. As an example for this intrusion, the HTLV-1 Tax protein represses the transcriptional functionality of p53 through phosphorylation at the N-terminal Ser15 residue (TABAKIN-FIX *et al.*, 2006), while also competitively binding to

or actively sequestering the coactivator CREB-binding protein (CBP) (ARIUMI *et al.*, 2000). Adenoviral infection will inadvertently activate the p53 tumor suppressor through expression of the oncogenic E1A protein, since it induces cellular mitogenic networks to establish a pro-viral environment. In isolation, this would result in the induction of cell cycle arrest, the DNA damage response (DDR), cellular senescence and apoptosis (DEBBAS and WHITE, 1993). This interference is mediated by tremendous control over multiple cellular and viral pathways through its complex protein interaction networks (FERRARI *et al.*, 2008, 2014; KING *et al.*, 2018). Some of these detrimental effects are countered through in different ways by the E1B region proteins (SARNOW *et al.*, 1982), which was first shown two decades ago in a set of hallmark publications (MARTIN and BERK, 1998; YEW *et al.*, 1994). While no experimental data could show that E1B-55K interacts directly with host or viral DNA, YEW *et al.* (1994) have shown that a GAL4-E1B-55K fusion protein binding to a CAT promoter via its GAL4-domain and thereby represses transcription. The first evidence of actual *in vivo* association of large E1B proteins with host cell DNA was found by ZHAO *et al.* (2007b), where the presence of the Ad12 large E1B protein in the promoter region of the significant p53-downstream target gene CDKN1A was shown via ChIP-qPCR, however the authors did not establish a clear connection between this association and a transcriptional outcome. These findings suggested that E1B-55K is able to repress p53-mediated transcriptional activation not only through its interaction with the N-terminal region of p53, but that it can also act as a general transcriptional repressor, if present adjacent to or within gene regulatory regions. Additionally, potential consequences of p53-independent secondary interactions on the DNA are unknown so far. Consequently, one of the main aims of this thesis was to improve the present knowledge of E1B-55K as a secondary vTR by globally investigating its interaction with the host cellular DNA to reveal so far unknown consequences of E1B-55K mediated transcriptional regulation. Early ChIP-seq experiments, using formaldehyde as the sole fixation agent, were unable to provide any evidence on direct E1B-55K presence on the host genome (Figure 10), although other ChIP-seq experiments targeting various host cell transcription factors with similar conditions can reliably reveal DNA-binding (TSANKOV *et al.*, 2015). This finding consolidated the idea that E1B-55K would not, by itself, bind to host DNA and a secondary intermediate factor was required. I decided to use a secondary protein-protein fixation agent, the homo-bifunctional N-hydroxysuccinimide ester DSG, which creates approximately 7 Å long covalent bridges between protein complexes (TIAN *et al.*, 2012; SINGH *et al.*, 2019). Performing this slightly altered ChIP-seq protocol allowed to identify bind-

ing sites that were indirectly bound by E1B-55K on the BRK host genome (Figure 13**b**), most likely mediated through p53 (Figure 13**c**). These experiments validated and expanded the earlier findings of ZHAO *et al.* (2007b), although my data clearly showed a preference for an upstream putative regulatory region. As this CDKN1A-adjacent region was not yet described in the *rattus norvegicus* system, I compared the region to the nucleotide collection of the online BLAST *blastn*-suite (ALTSCHUL *et al.*, 1990; DUMONTIER and HOGUE, 2002) and identified an extremely high degree of homology (>90%) to the human Linc-p21 region (data not shown), which contains multiple enhancer elements (GROFF *et al.*, 2016). This data suggested that adenoviral E1B-55K would interact and interfere with this regulatory site of CDKN1A by binding to enhancer-bound p53 to deregulate expression of target downstream genes. Unfortunately, the other identified peaks were not significantly correlated with any specific transcription factor. In a bid to further increase the strength and sensitivity of the ChIP-assay in a more standardized setting, I adapted the experimental setup further by using a lentiviral transduction system to establish multiple BRK cell lines that were expressing both E1A- and E1B region variants. It is known that PTMs of TFs can influence their ability to regulate transcriptional activation. Early studies investigating the direct effect of SUMOylation on transcription factors E26 transformation-specific - like 1 (ELK-1) (YANG *et al.*, 2003) and specificity protein (Sp)3 (ROSS *et al.*, 2002) found a predominantly repressive effect regarding their transcriptional activation functionality. On the other hand, as reviewed in CHYMKOWITCH *et al.* (2015) and ROSONINA *et al.* (2017), SUMO-1 modification can have very diverse effects on transcriptional activity of transcription factors, including the activation of p53 (RODRIGUEZ *et al.*, 1999). It is thought that SUMOylation of transcription factors can inhibit transcription by recruiting specific corepressors like HDACs (YANG and SHARROCKS, 2004), or directly compete with transcription-promoting modifications (VAN NGUYEN *et al.*, 2012). Studies that investigated Ad5 E1B-55K SUMOylation by ENDTER *et al.* (2001) and PENNELLA *et al.* (2010) suggested that this PTM is necessary for maximal deregulation of p53 by inducing E1B-55K-mediated efficient relocalization of the tetrameric p53-protein complexes to PML-nuclear bodies and its abrogation has been shown to result in reduced transcriptional repression in luciferase assays (KOLBE *et al.*, 2022). I was therefore specifically interested in investigating the effect of Ad5 E1B-55K SUMOylation in the context of transcription factor-targeted ChIP- and mRNA-seq. Global transcriptome analysis revealed that all E1B-55K variants with deregulated SUMOylation were reliably repressing p53-responsive genes with the exception of the SUMOylation-deficient K104R (Figures 17**b**

and 18). Intriguingly, in the work of ENDTER *et al.* (2001), the authors could show that abrogation of SUMOylation was not alleviating E1B-55K's general transcriptional repressor functionality, as a GAL4-E1B-55K K104R fusion protein was still able to repress transcriptional activation. These findings suggest that the K104R mutant, while still generally being able to interact with p53 (ENDTER *et al.*, 2001), was not able to repress DNA-bound p53 in a similar manner to the wildtype. Analysis of highly validated p53-target genes (FISCHER, 2017) via unsupervised k-means clustering to extract and visualize the strongest difference of p53-responsive gene transcription between all E1B-55K mutants showed that, besides K104R, occupancy of the TSS of associated genes was elevated in the cluster of p53 genes that was found to be downregulated (Figure 19, left side), compared to the transcriptionally ambiguous cluster (Figure 19, right side). Both NES and NES/K104R mutants are equally more effective in repressing p53-target gene expression compared to the wildtype, which is visualized in the lower part of cluster 1. Despite this, ChIP-seq assays of the NES mutant, which localizes predominantly in the nucleus (Figure 15**i-l**) and is hyper-SUMOylated, repeatedly manifested reduced occupancy regarding its binding strength to DNA, similar to the hyper-SUMOylated K101R mutant (Figure 20, 21, 22). This observation regarding the K101R and NES proteins does coincide with a reduction in the absolute number of significantly downregulated genes (Figure 18), which also lowers the average significance of the associated mRNA-seq-derived biological pathway analysis (Figure 17**b**). This might be due to association of the majority of the E1B-55K proteins to PML-nuclear bodies, as shown by (WIMMER *et al.*, 2013) and (KOLBE *et al.*, 2022), which drastically reduces the sensitivity of the ChIP-assay, as the overall relative signal strength was clearly diminished but not completely abrogated (Figure 21). A similar effect was identified in the case of Sp100, which is massively SUMOylated during infection with KSHV and displaced from the chromatin after expression of the latency associated nuclear antigen (LANA) protein (GÜNTHER *et al.*, 2014). The general reduction in DNA occupancy of the NES protein attributed to the hyper-SUMOylation is rescued by the introduction of the K104R-mutant. It has been shown that the NES/K104R mutant is relocalizing p53 into nuclear foci independent from PML-nuclear bodies (PENNELLA *et al.* (2010), as seen in Figure 15**m**), whereas wildtype E1B-55K would predominantly relocalize the protein to PML-NB (ENDTER *et al.*, 2005) and subsequently export the protein to the cytoplasm. These findings, and previously published data, suggest that transcriptional repression via binding to DNA-bound p53 is a function that manifests in tandem with the relocalization of p53. E1B-55K has to be present in the nucleus to associate with the host transcription

factor directly at its responsive regulatory gene sites. This work is the first *in vivo* experimental evidence that could reliably establish a direct link between presence of E1B-55K on the promoters of multiple p53-regulated target genes and their transcriptional repression, thereby solidifying the role of E1B-55K as a secondary viral transcriptional regulator. Concurrently, SUMOylation is evidently not required for this mechanism, but rather dictates the duration of nuclear localization and therefore access of Ad5 E1B-55K to the host genome. The authors of KOLBE *et al.* (2022) come to a similar conclusion, attributing the effect of the modification primarily to protein localization.

5.2 Genome-wide occupancy of large E1B proteins is conserved and correlated with transcriptional repression

Many interactions between large E1B proteins of different Adenoviruses with cellular proteins are highly conserved, but subtle differences exist between the species (CHENG *et al.*, 2011; BLANCHETTE *et al.*, 2013). The establishment of a large E1B/E4orf6- E3 ubiquitin ligase complex represents a conserved mechanism between the species, while the exact composition of this complex can differ. For example, Ad5 forms a primarily Cullin 5-based complex, whereas Ad12 and F40 form a Cullin 2-based one (CHENG *et al.*, 2013). Likewise, not all interaction partners of large E1B proteins are found to be degraded by this complex, which is another major distinction trait. Variations also exist when it comes to localization, with Ad5 being predominantly cytoplasmic, while Ad12, Ad40 and Ad52 (data not shown) display a dominant nuclear localization (KOLBE *et al.*, 2022). These distinctions between the adenoviral species and the general importance of the E1 region in their transformative capabilities (as reviewed in IP and DOBNER (2020)) begs the question how this variance of the large E1B proteins is reflected in the interference with host gene transcription. Combining mRNA-seq- with ChIP-seq data from Ad5 E1B-55K with the findings derived from large E1B proteins of species A12, F40 and G52 vastly expanded our knowledge of the viral protein. The total number of peaks differed greatly between the species, ranging from approx. 800 for Ad5 to over 18.000 for Ad12, highlighting a further possible distinction trait for the species. For context, a systemic analysis of 33 different viral transcriptional regulators performed by LIU *et al.* (2020) found a median of 675 binding sites across the host genome. In this collective analysis, the authors identified 305.562 distinct binding sites in total, highlighting the large potential to influence host gene expression. There is not a singular unified consequence of these interactions,

the ramifications rather depend on the viral protein itself. In the review of CHIANG and LIU (2019), the authors describe multiple mechanisms through which HSV1, HIV1 and KSHV interfere with cellular IRF- and STAT-dependent immune responses, one of which is the antagonisation of the transcriptional complex formation by interacting and obstructing recruitment of p300 to the associated genes. Viral transcriptional regulators have found various ways to alter cellular gene expression and provide plenty examples and possibilities for mechanisms in which also conceivably interferes. By way of combining peak-associated genes with their respective mRNA-seq $\log_2$-fold changes, I could establish that binding events of either Ad5 wildtype- or NES/K104R are significantly associated with transcriptional repression on a global scale (Figure 25**c**), which provides evidence for potential intrusion into pathways beyond previously established p53-signaling. Analysis of NES/K104R procured a two times higher negative pearson correlation coefficient (-ρ) between peak score and transcriptomic consequence compared to the wildtype. It is likely that this finding correlates with the previously established stronger repression of singular p53 target genes by NES/K104R. The exact cause of this difference remains unclear as both proteins have an equal absolute number of peaks (Figure 21**b**) and share general genome occupancy (Figure 22). It can be argued that NES/K104R generally shows a much stronger nuclear phenotype compared to wildtype (Figure 15). Similarly, SUMOylation influences and affects a plethora of different mechanisms, including changes in the protein-protein interaction portfolio of the modulated proteins (ZHAO, 2018). Likewise, large E1B proteins from Ad12 and Ad52 were found to be negatively associated with gene expression as a consequence of their widespread global indirect DNA interaction sites (Figure 35**a-c**), while F40 was only very weakly associated with global repression with a non-significant p-value of 0.17. Group F Adenoviruses, to which F40 belongs, are causative agents of frequent outbursts of acute gastroenteritis (SONESON *et al.*, 2016; HASSOU *et al.*, 2019; DO NASCIMENTO *et al.*, 2022), and are known to have an extremely low transformation efficiency (MAUTNER *et al.*, 1995). Whether this finding is correlated with the here shown weak and non-significant repressive potential of the F40 large E1B protein is to be seen. As already noted, ChIP-seq of both Ad40 and Ad52 were only performed once, so additional experiments could provide more evidence to clarify the validity of these results. The apparent overlap between the different species, when it comes to genome occupancy, is a clear indicator of a generally conserved mechanism with subtle differences between the species (Figure 33). The most outstanding large E1B protein belongs to Ad12, which shows comparatively wide-spread coverage through numerous host genome bind-

ing sites (around 22 times the number of Ad5) over three different biological replicates, ruling out possible confounding factors like fixation time prior to the ChIP-seq experiment. Most investigated Adenovirus groups are largely non-oncogenic, but exceptions do exist in group A and D. Ad12 belongs to the former, a group that is known to be highly oncogenic when introduced into rodents (TRENTIN *et al.*, 1962; KUHLMANN *et al.*, 1982). This potent tumorigenicity is largely attributed to its E1A protein, which is uniquely able to repress transcription of MHC class I genes, enabling potent immune escape of the virus (BRICKELL *et al.*, 1983; TANAKA *et al.*, 1985). The latter group of oncogenic Adenoviruses is represented by Ad9, which was shown to introduce mammary fibroadenomas in rodents (JONSSON and ANKERST, 1977). This coverage and potent attribution to transcriptional repression of the Ad12 large E1B protein thereby presents a novel potential mechanism by which group A Adenoviruses present a highly oncogenic phenotype. Complex biological pathway- and network analysis of large E1Bs reveal that several oncogenesis-, stress response-, tumor growth- and p53/apoptosis-associated pathways are deregulated (Figure 24 and 34). These visualizations show that, in essence, all large E1B proteins are functionally similar to each other and overlap in their core terms. It has to be stressed that the maximum number of entries for the utilized pathway analysis in this thesis is 3.000 and both Ad12 and Ad52 exceed that cut-off, in which case only the most significant genes have been chosen, which raises the possibility that some less significant pathways may be excluded. Another indication for a general and global repressive mechanism by large E1B proteins expression is apparent in the overrepresentation of highly significant (and mostly shared) pathways in the downregulated pathways, compared to the upregulated ones. An example of a difference is the ”Signaling by Rho GTPase signaling” pathways, which was not found to be enriched by Ad52 in ChIP-seq and consequently not repressed in mRNA-seq. Ad12, Ad40 and Ad52 displayed ”TGFβ-signaling” pathways to be both bound and repressed, an observation that will become relevant later in the thesis. Unfortunately, exact connection between the terms is difficult to create in some cases, as mRNA-seq naturally captures primary and secondary effects, whereas ChIP-seq only shows primary target regions where binding occupancy could be established. Intriguingly, this interpretation of large E1B functionality is contrasting previously published findings by MILLER *et al.* (2009), in which the authors argue for a dominant immune-response targeted-, rather than a p53-responsive role during viral infection of normal human fibroblasts (HFFs). Multiple reasons are possible for these divergent findings, like the inclusion of other viral proteins that influence the setup or the general divergence introduced by adenoviral in-

fection in contrast to lentiviral-based transduction and transformation. In summation, the here presented data provides ample evidence that transcriptional repression by large E1B proteins of different species is a conserved mechanism by which the virus can further extend and modulate its influence on the host. The genome-wide coverage of the viral protein and the significance of the associated repression correlates slightly with the differing observed oncogenicity of the adenoviral species, which proposes a novel synergistic functionality between the different E1- and E4 region proteins. The data suggests that this functional capacity of large E1B proteins, at least in part, contributes to the transformative potential of the early adenoviral regions, as the core and consensus of affected genes are implicated in oncogenicity-, stress-response- and apoptosis- related pathways. Interference with these pathways is likewise an important part of viral infection to circumvent the various host defense mechanisms and to establish a pro-viral environment.

5.3 Adenovirus large E1B proteins target AP-1- and TEAD transcription factor family members

Recognition of specific nucleotide sequences in proximal or distal regulatory regions, so called TFBS, is a key feature of transcription factors through which they are able to induce transcription of target genes (reviewed in LAMBERT *et al.* (2018)). Regulatory regions can contain multiple different motifs directly adjacent to each other, allowing complex fine-tuning interplay of gene expression by multiple factors (WEIRAUCH and HUGHES, 2014). Viruses can express proteins, so called secondary viral transcriptional regulators (LIU *et al.*, 2020), that interact with these TFs to interfere with or modulate gene expression to their advantage. An example of an indirect viral regulation of a host transcription factor is the interplay between HPV E6 and cellular interferon regulatory factor (IRF)-3 (RONCO *et al.*, 1998; SHAH *et al.*, 2015). This interaction reduces the transactivation of IRF-3 by up to 85%. Naturally, direct interaction of a viral transcriptional regulator with host or viral DNA is also possible, like in the earlier-described case of EBV BamHI Z fragment leftward open reading frame 1 (BZLF1) (TILLO *et al.*, 2017). In order to improve interpretation of E1B-55K ChIP-seq data, I performed HOMER-mediated *de novo* motif enrichment (HEINZ *et al.*, 2010). Besides p53, as discussed earlier in detail, analysis of large E1B species revealed highly conserved interaction with factors of the AP-1 transcription factor complex. Other viruses have also shown deregulation of AP-1, notably the HTLV-1 Tax oncoprotein, a crucial component during the establishment of malignant transformation leading

to adult T-cell leukemia-lymphoma (ATL) (GRASSMANN *et al.*, 2005). Expression of Tax lead to increased expression of AP-1 members c-Jun, Fos and Fra-1, while the protein itself is able to activate transcription through the AP-1 binding site. Intriguingly, the HTLV-1 bZIP factor (HBZ) protein, which is similarly expressed by HTLV-1, can repress transcriptional activation of AP-1. This divergent interplay between Tax and HBZ could represent an attenuating control mechanism of the virus (BASBOUS *et al.*, 2003). In Adenovirus, expression of E1A also leads to activation of AP-1 (MULLER *et al.*, 1989), an activity that can potentially be repressed by E1B - a feed-back loop similar to activation and repression of p53. The AP-1 complex is formed by dimerization through a leucine-zipper motif and its components are therefore known as bZIP proteins. Their binding can have both oncogenic and anti-oncogenic effects through regulation of cellular proliferation, apoptosis, angiogenesis, differentiation and tumor invasion (reviewed in EFERL and WAGNER (2003)). Intriguingly, AP-1 preferentially binds to enhancer regions in human tumor cell lines (SEO *et al.*, 2021), whereas I could not observe a similar preference in the primary BRK cell data sets. In the global transcriptomic repression analysis, "PI3K-Akt signaling pathway", "Small cell lung cancer" and "Signaling by TGFB family members" were found to be repressed (Figure 17**b** and 31**b**), pathways that are also frequently associated with AP-1 inhibition in the literature (KIKUCHI *et al.*, 2008; FENG *et al.*, 2020; JOSUE RUIZ *et al.*, 2021). In general, tumor- and cancer-related pathways, found to be repressed in the transcriptome, can be regulated by AP-1 (EFERL and WAGNER, 2003; EL-FATTAH IBRAHIM *et al.*, 2018; CASALINO *et al.*, 2022; SHU *et al.*, 2022). Evidence exist that viruses have specific AP-1 binding sites within their viral genome, as is the case with the visna virus (SHIH *et al.*, 1992) or the gallid alpha-herpesvirus 1 (WANG *et al.*, 2021), which highlights the importance for potential deregulation of this transcription factor complex. The motif that is predominantly bound by the AP-1 complex depends on the composition of the dimer: Jun-Fos heterodimers prefer the TRE-motif, while Jun-Jun homodimers show drastically less enrichment for this motif (ISAKOVA *et al.*, 2017). My data therefore provides evidence that large E1B proteins might preferentially interact with the Fos transcription factor family. Differences in the statistical significance exist between the species, as interaction with AP-1 was predominant in Ad5, A12 and G52, but much less significant in F40. Whether this finding regarding F40 coincides with its non-significant repressive functionality is to be seen.

The other dominant motif which was found to be enriched by all large E1B proteins was the TEA-motif that is recognized and targeted by the TEAD transcription factor family

- a vital component of the Hippo signaling pathway. I found extensive overlap and relatively weak evidence of relocalization between the TEAD factor family and E1B-55K on the host genome (Figures 28**a** and 38**a**), as well as direct interaction between the tested proteins in direct co-immunoprecipitation assays (Figures 29 and 38**b**). Ad5 E1B-55K- and TEAD genome coverage co-occupancy data suggest that E1B-55K expression leads to some relocalization of TEAD on the host chromatin, as cells expressing K104R, which does not show any presence on the genome, display a unique distribution of TEAD not seen by the other E1B-55K variants (Figure 28**a**). The data presented in this work indicates interaction between E1B-55K and TEAD, as the significance of the TEAD motif correlated with the presence of E1B-55K on the genome (Figure 28**b**). In addition, while the AP-1 TRE-motif was one of the most significantly enriched motifs by viral transcriptional regulators, the TEA motif was not significantly enriched by any of the investigated viruses to a notable degree (LIU *et al.* (2020), Supplementary table S5), suggesting a distinctive trait of Adenoviruses. Recently, ZEMKE *et al.* (2019) have presented data that suggests dedifferentiation of HEK293 cells by adenoviral E1A-mediated sequestration of Hippo signaling effectors YAP and TAZ, hindering their association with DNA-bound TEAD at poised enhancers. Combining these findings with the here presented data suggest a synergistic effect of E1A and E1B-55K, as both functionalities would lead to repression of TEAD-mediated transcriptional activation. Regarding repression of TEAD-motifs, a notable difference between Ad5 and the other species lies in the proportion of repressed associated genes, as Ad12, Ad40 and Ad52 were repressing all of their primary motifs (Figure 36), while Ad5 only showed remarkable repression of AP-1 and p53 (Figure 26). AP-1 and TEAD are known to regularly co-occupy and associate with each other specifically at enhancer sites to regulate oncogenic growth (ZANCONATO *et al.*, 2015), cellular transformation (HE *et al.*, 2021) or cancer cell migration and invasion (LIU *et al.*, 2016), in addition to their isolated functions described earlier in this work. These data suggest that in principle, large E1B proteins are able to bind and repress AP-1- and TEAD-transcription factors in primary BRK cells, but differences in efficiency do exist between the species, as also noted in the previous part. ChIP-seq of Ad5 E1B-55K in p53-negative H1299 cells showed enrichment of AP-1, but none of either p53 or TEAD, while binding sites were remarkably restricted to promoter sites. Recent literature suggests that p53 (MELO *et al.*, 2013; YOUNGER and RINN, 2017), AP-1 (SEO *et al.*, 2021) and TEAD (ZANCONATO *et al.*, 2015) are preferentially bound to enhancer sites of the human genome. As neither p53- nor TEAD-motifs were found in these data, a shift of the previously observed balance between promoter and enhancer to

a primarily promoter-associated distribution is thinkable. Stringent data analysis focusing on AP-1 revealed TGFβ-signaling pathways as the most significant pathway that is affected. The authors of YAMAMURA *et al.* (2000), VERRECCHIA *et al.* (2001) and SUNDQVIST *et al.* (2020) have shown a specific enrichment of AP-1 around promoter sites of TGFβ-signaling associated genes, which would explain the aforementioned shift. This pathway was also found enriched by ChIP-seq and repressed in mRNA-seq of species A12, F40 and G52, solidifying the evidence for selective repression of the TGFβ-signaling pathways by large E1B proteins. The H1299 cell data provide strong evidence that host chromatin presence of the viral oncoprotein does not rely on p53- and p53-interacting proteins alone and rules out a potential false-positive artifact induced by DSG-fixation.

5.4 Ad5 E1B-55K expression in human primary cells deregulates specific epigenetic histone modifications

Cellular transformation of cultured human primary cells with adenoviral oncogenes has been achieved on multiple occasions, albeit at less efficiencies than compared to the rodent cell system. Several widely used cell lines have been created by different working groups, like the HEK293 cell line from human embryonic kidney cells (GRAHAM *et al.*, 1977) or the 911 cell line from human embryonic retinoblasts (FALLAUX *et al.*, 1996). In SPEISEDER *et al.* (2017), the authors were able to transform primary hMSC by transduction and subsequent continued expression of the early region E1 genes. I complemented this earlier work by creating three biological replicates of hMSC that have been transformed through expression of E1A and E1B-19K or E1A, E1B-19K and E1B-55K. The aim of this part of my thesis was to provide insight into the potential mechanism of E1B-55K-mediated gene repression, as influence on the epigenetic landscape has been suggested over 30 years ago by YEW *et al.* (1994) and MARTIN and BERK (1998). Several viral proteins are known to obstruct the p300 histone acetyltransferase (CHIANG and LIU, 2019) to interfere with antiviral-gene expression while others are able to actively modulate vast domains of the host chromatin (FERRARI *et al.*, 2008, 2012), like the adenoviral E1A protein. The interplay between the epigenetic landscape and transcription factors is a necessity of gene expression (STEIN *et al.*, 2009; AHSENDORF *et al.*, 2017; ONODERA *et al.*, 2018). An example of this reciprocity is found in the diverse pioneer transcription factors that are required to unmask heterochromatin domains to allow transcription in the correct and proper environment (MAYRAN and DROUIN, 2018). The p53 transcription factor has been

established as having pioneer-like traits (SAMMONS *et al.*, 2015). If adenoviral large E1B proteins directly bind to p53 and repress p53-mediated cell cycle arrest- and apoptotic responses, as described by YOUNGER and RINN (2017) and in detail in this thesis, then it would be possible to observe this footprint on epigenetic modifications of specific target gene regions. By analysis of canonically activating H3K4me3 and H3K27ac histone epigenetic modifications, I found a general overabundance of regions that had statistically lost signal when E1B-55K was expressed (64% for either modification), which further lends to the previously discussed hypothesis of E1B-55K being a genome-wide transcriptional repressor. Interestingly, some of these regions cluster together according to their gain-or-loss phenotype (Figure 42). This is exemplified by chromosome 21 with all regions having gained signal- or by chromosome 13 with all but one region having lost signal strength. This suggests extended distal influence of E1B-55K on genes that are organized in topologically associating domains (TAD)s and is a potential future direction for research (LONG *et al.*, 2022). Boundaries of TADs are enriched for CTCF binding sites (ROCHA *et al.*, 2015), a protein that has been specifically enriched by Ad12 in earlier experiments (Figure 36) and is known to play an important role in adenoviral DNA replication (KOMATSU *et al.*, 2013). The most prominent regions that have lost activating histone PTMs signal belong to p53-regulated pathways, which is visualized in Figure 41**b**, with the most significant regions being directly associated with histone PTM loss in the promoter regions of CDKN1A and MDM2. Transcriptional activation of p53-associated genes can be regulated by recruitment of the histone acetyltransferase p300 (LILL *et al.*, 1997) or a complex containing the histone methyltransferase MLL3 (LEE *et al.*, 2009). It is conceivable that E1B-55K binding to DNA-bound p53 inhibits recruitment of these factors or complexes and thereby leads to this marked drop of either one of these signals. A different repressive mechanism could be induced equivalent to the BTB/POZ and zinc finger domains factor on chromosome 1 (BOZF-1) protein, which binds to Sp1-rich GC boxes directly upstream of the CDKN1A TSS and thereby blocks recruitment of the p53-binding enhancer element (KIM *et al.*, 2013), as the extended Ad12 large E1B protein coverage revealed Sp2 as one of the strongest motifs. Significant statistical relationships between activating histone PTM gain or loss in promoter and enhancer regions and transcriptional up- and down-regulation could be established, with both PTMs displaying an equally strong positive correlation (Figures 43**a** and **b** and 44). Biological pathway analysis in hMSC suggests a significant degree of conservation between large E1B protein functions in rodents and humans, as it visualizes many similar pathways identified by earlier ChIP-seq and mRNA-

seq. In line with this, the core of the most significant pathways was focused around p53, cancer and genotoxicity. Regarding the upregulated pathways, the most significant pathway was represented by "Ribosome", which manifests through overexpression of several ribosomal proteins. Failure to repress an activated p53 response, which is the case in hMSC cells that are lacking E1B-55K expression, can lead to induction of senescence and blockage of cellular growth (reviewed in RUFINI *et al.* (2013)), therefore exhibiting a reduced expression of ribosomal genes when E1B-55K is not present (MAYER and GRUMMT, 2006; BERTOLI *et al.*, 2013). Overexpression of ribosomal genes is a regular observation I found as a consequence of E1B-55K-mediated repression of p53 (data not shown). Additionally, multiple immune-related pathways were found to harbor elevated H3K4me3 signal at their involved genes, albeit with comparatively low significance. I suspect that this phenotype is similarly a consequence of a deregulated p53-response, as p53 interacts and crosstalks with multiple downstream modulators implicated in the immune response (reviewed in UEHARA and TANAKA (2018)). It has been shown that p53 actively suppresses inflammation (ZHENG *et al.*, 2005) and that p53-negative (a phenotype induced by E1B-55K expression) mice are susceptible to premature death as a consequence of inflammation-induced abscesses (DONEHOWER *et al.*, 1992). Genes involved in the Hippo signaling pathway appeared to have lost both activating H3K4me3 and H3K27ac histone PTMs without a discernible change in mRNA expression, possibly as a consequence of interference with the TEAD and AP-1 transcription factors. This observed reduction further lends credibility to a possible synergistic effect of E1A and large E1B proteins. Regarding the heterochromatin-associated histone PTMs, while I could identify several regions that either gained or lost signal strength, a relationship between these differences and consequences for expression of associated genes could not be reliably established (Figures 46**a** and **b**). Depending on the regional context and developmental state of the cell, constitutive heterochromatin H3K9me3 width can vary tremendously (WEI *et al.*, 2021) compared to both H3K4me3 and H3K27ac, which predominantly mark sharp promoter- or enhancer regions. This distinction complicates causative connections between H3K9me3 and specific genes as no significant pathway could be enriched with genes that exclusively gained or lost H3K9me3. Pathway analysis of genes affected by differential gain and loss of H3K27me3 PTM in hMSC expressing E1B-55K exhibited traits of neuronal differentiation processes (Figure 46**c**). H3K27me3 has been repeatedly implicated in neuronal differentiation (CHEN *et al.*, 2019; BUONTEMPO *et al.*, 2022). Additionally, E1A-induced transformation of primary amniotic fluid-derived cells leads to morphological features re-

miniscent of epithelial cells and transformed cells expressed neuronal-, glial- and epithelial cell markers (ARNHOLD *et al.*, 2008). This cellular transformation process has been shown to be possible with either E1B-19K or E1B-55K, in addition to E1A, but the pathways they influence are thought to act in an additive manner (MCLORIE *et al.*, 1991), which implies that expression of both E1B proteins would allow for a more efficient and advanced transformation process that would sufficiently explain this phenotype. This interpretation therefore argues against an involvement of E1B-55K with heterochromatin-associated PTMs, and the above-described observation could be explained by an advanced stage of the transformation process. Regarding the investigated histone PTMs, these data suggest that E1B-55K influences activating epigenetic H3K4me3 and H3K27ac by interacting with transcription factor complexes bound at promoter- and enhancer regions, in addition to its other functions like the sequestration of p53, and suggest no directly induced deregulation of heterochromatin-associated regions, as described earlier. Further studies investigating these exact connections between ChIP- and mRNA-seq and their consequences are needed, preferentially performed within a confined *Homo sapiens*-based setup that includes large E1B host genome profiling, epigenetic- and transcriptomic consequences, as well as general chromatin accessibility studies. In addition, as there are numerous histone PTMs that have not been investigated in this thesis, is is thinkable that this experimental setup does not capture the exact mechanism of interference by large E1B protein expression.

To summarize, the data presented here illustrate and explore a previously only hypothesized function of adenoviral large E1B proteins interaction with DNA-bound transcription factors. The viral proteins establish genome-wide coverage on the host chromatin through a highly conserved interaction repertoire with specific cellular transcription factors. The most prominent factors include the p53 tumor suppressor, components of the AP-1 complex and members of the TEAD family. As a consequence of this interaction, I identified a prominent abrogation of activating histone modifications on the host genome that was accompanied with a reduction of gene expression. Large E1B proteins thereby deregulate the antiviral host response to establish an environment conducive to efficient transformation. The exact mechanism of this transcriptional repression remains an open question for future research.

Bibliography

Ablack J. N. G., Cohen M., Thillainadesan G., Fonseca G. J., Pelka P., Torchia J. and Mymryk J. S. (2012) Cellular GCN5 Is a Novel Regulator of Human Adenovirus E1A-Conserved Region 3 Transactivation, *Journal of Virology* **86**(15), 8198–8209.

Agger K., Cloos P. A., Christensen J., Pasini D., Rose S., Rappsilber J., Issaeva I., Canaani E., Salcini A. E. and Helin K. (2007) UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development, *Nature* **449**(7163), 731–734.

Ahsendorf T., Müller F.-J., Topkar V., Gunawardena J. and Eils R. (2017) Transcription factors, coregulators, and epigenetic marks are linearly correlated and highly redundant, *PLOS ONE* **12**(12), e0186324.

Akerblom I. E., Slater E. P., Beato M., Baxter J. D. and Mellon P. L. (1988) Negative regulation by glucocorticoids through interference with a cAMP responsive enhancer, *Science* **241**(4863), 350–353.

Allen M. A., Andrysik Z., Dengler V. L., Mellert H. S., Guarnieri A., Freeman J. A., Sullivan K. D., Galbraith M. D., Luo X., Lee Kraus W., Dowell R. D. and Espinosa J. M. (2014) Global analysis of p53-regulated transcription identifies its direct targets and unexpected regulatory mechanisms, *eLife* **2014**(3).

Allshire R. C. and Madhani H. D. (2018) Ten principles of heterochromatin formation and function.

Aloni-Grinstein R., Charni-Natan M., Solomon H. and Rotter V. (2018) p53 and the viral connection: Back into the future.

Altschul S. F., Gish W., Miller W., Myers E. W. and Lipman D. J. (1990) Basic local alignment search tool, *Journal of Molecular Biology* **215**(3), 403–410.

Amemiya H. M., Kundaje A. and Boyle A. P. (2019) The ENCODE Blacklist: Identification of Problematic Regions of the Genome, *Scientific Reports* **9**(1), 1–5.

Andersson R., Gebhard C., Miguel-Escalada I., Hoof I., Bornholdt J., Boyd M., Chen Y., Zhao X., Schmidl C., Suzuki T., Ntini E., Arner E., Valen E., Li K., Schwarzfischer L., Glatz D., Raithel J., Lilje B., Rapin N., Bagger F. O., Jørgensen M., Andersen P. R., Bertin N., Rackham O., Burroughs A. M., Baillie J. K., Ishizu Y., Shimizu Y., Furuhata E., Maeda S., Negishi Y., Mungall C. J., Meehan T. F., Lassmann T., Itoh M., Kawaji H., Kondo N., Kawai J., Lennartsson A., Daub C. O., Heutink P., Hume D. A., Jensen T. H., Suzuki H., Hayashizaki Y., Müller F., Forrest A. R., Carninci P., Rehli M. and Sandelin A. (2014) An atlas of active enhancers across human cell types and tissues, *Nature* **507**(7493), 455–461.

ANDREWS S. (2019) Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data.

ARAGONA M., PANCIERA T., MANFRIN A., GIULITTI S., MICHIELIN F., ELVASSORE N., DUPONT S. and PICCOLO S. (2013) A mechanical checkpoint controls multicellular growth through YAP/TAZ regulation by actin-processing factors, *Cell* **154**(5), 1047–1059.

ARIUMI Y., KAIDA A., LIN J. Y., HIROTA M., MASUI O., YAMAOKA S., TAYA Y. and SHIMOTOHNO K. (2000) HTLV-1 Tax oncoprotein represses the p53-mediated transactivation function through coactivator CBP sequestration, *Oncogene* **19**(12), 1491–1499.

ARNHOLD S., POST C., GLÜER S., HOOPMANN M., WENISCH S., VOLPERS C. and ADDICKS K. (2008) Neuronal characteristics of amniotic fluid derived cells after adenoviral transformation, *Cell Biology International* **32**(12), 1559–1566.

AVVAKUMOV N., KAJON A. E., HOEBEN R. C. and MYMRYK J. S. (2004) Comprehensive sequence analysis of the E1A proteins of human and simian adenoviruses, *Virology* **329**(2), 477–492.

AVVAKUMOV N., SAHBEGOVIC M., ZHANG Z., SHUEN M. and MYMRYK J. S. (2002) Analysis of DNA binding by the adenovirus type 5 E1A oncoprotein., *The Journal of general virology* **83**(Pt 3), 517–24.

AYLON Y. and OREN M. (2016) The Paradox of p53: What, how, and why?, *Cold Spring Harbor Perspectives in Medicine* **6**(10), a026328.

BALSALOBRE A. and DROUIN J. (2022) Pioneer factors as master regulators of the epigenome and cell fate.

BANNISTER A. J. and KOUZARIDES T. (2011) Regulation of chromatin by histone modifications.

BARLEV N. A., LIU L., CHEHAB N. H., MANSFIELD K., HARRIS K. G., HALAZONETIS T. D. and BERGER S. L. (2001) Acetylation of p53 activates transcription through recruitment of coactivators/histone acetyltransferases, *Molecular Cell* **8**(6), 1243–1254.

BARSKI A., CUDDAPAH S., CUI K., ROH T. Y., SCHONES D. E., WANG Z., WEI G., CHEPELEV I. and ZHAO K. (2007) High-Resolution Profiling of Histone Methylations in the Human Genome, *Cell* **129**(4), 823–837.

BARTER R. L. and YU B. (2018) Superheat: An R Package for Creating Beautiful and Extendable Heatmaps for Visualizing Complex Data, *Journal of Computational and Graphical Statistics* **27**(4), 910–922.

BASBOUS J., ARPIN C., GAUDRAY G., PIECHACZYK M., DEVAUX C. and MESNARD J. M. (2003) The HBZ Factor of Human T-cell Leukemia Virus Type I Dimerizes with Transcription Factors JunB and c-Jun and Modulates Their Transcriptional Activity, *Journal of Biological Chemistry* **278**(44), 43620–43627.

BEACON T. H., DELCUVE G. P., LÓPEZ C., NARDOCCI G., KOVALCHUK I., VAN WIJNEN A. J. and DAVIE J. R. (2021) The dynamic broad epigenetic (H3K4me3, H3K27ac) domain as a mark of essential genes.

BERGELSON J. M., CUNNINGHAM J. A., DROGUETT G., KURT-JONES E. A., KRITHIVAS A., HONG J. S., HORWITZ M. S., CROWELL R. L. and FINBERG R. W. (1997) Isolation of a common receptor for Coxsackie B viruses and adenoviruses 2 and 5, *Science* **275**(5304), 1320–1323.

BERGER S. L., KOUZARIDES T., SHIEKHATTAR R. and SHILATIFARD A. (2009) An operational definition of epigenetics, *Genes and Development* **23**(7), 781–783.

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 of the United States of America* **74**(8), 3171–3175.

BERTOLI C., SKOTHEIM J. M. and DE BRUIN R. A. (2013) Control of cell cycle transcription during G1 and S phases.

BESSON S., VRAGNIAU C., VASSAL-STERMANN E., DAGHER M. C. and FENDER P. (2020) The Adenovirus Dodecahedron: Beyond the Platonic Story, *Viruses* **12**(7), 718.

B.GREENBERG M. K. H. (2013) Adenoviridae (Fields virology), in *Fields Virology*, pp. 1347–1401, Wolters Kluwer Health/Lippincott Williams & Wilkins, Philadelphia, 6th ed. edn.

BIGGS H. M., LU X., DETTINGER L., SAKTHIVEL S., WATSON J. T. and BOKTOR S. W. (2018) Adenovirus-associated influenza-like illness among college students, Pennsylvania, USA.

BIOCONDUCTOR CORE TEAM AND BIOCONDUCTOR PACKAGE MAINTAINER (2022) TxDb.Rnorvegicus.UCSC.rn7.refGene: Annotation package for TxDb object(s).

BLACKFORD A. N. and GRAND R. J. A. (2009) Adenovirus E1B 55-Kilodalton Protein: Multiple Roles in Viral Infection and Cell Transformation, *Journal of Virology* **83**(9), 4000–4012.

BLANCHETTE P., WIMMER P., DALLAIRE F., CHENG C. Y. and BRANTON P. E. (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), 4872–4881.

BLANCO E., GONZÁLEZ-RAMÍREZ M., ALCAINE-COLET A., ARANDA S. and DI CROCE L. (2020) The Bivalent Genome: Characterization, Structure, and Regulation.

BLIGHE K., RANA S. and LEWIS M. (2022) EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling.

BOHMANN D., BOS T. J., ADMON A., NISHIMURA T., VOGT P. K. and TJIAN R. (1987) Human proto-oncogene c-jun encodes a DNA binding protein with structural and functional properties of transcription factor AP-1, *Science* **238**(4832), 1386–1392.

BORKENHAGEN L. K., FIELDHOUSE J. K., SETO D. and GRAY G. C. (2019) Are adenoviruses zoonotic? A systematic review of the evidence, *Emerging Microbes & Infections* **8**(1), 1679–1687.

BOYD J. M., MALSTROM S., SUBRAMANIAN T., VENKATESH L. K., SCHAEPER U., ELANGOVAN B., D'SA-EIPPER C. and CHINNADURAI G. (1994) Adenovirus E1B 19 kDa and Bcl-2 proteins interact with a common set of cellular proteins, *Cell* **79**(2), 341–351.

BOYER L. A., PLATH K., ZEITLINGER J., BRAMBRINK T., MEDEIROS L. A., LEE T. I., LEVINE S. S., WERNIG M., TAJONAR A., RAY M. K., BELL G. W., OTTE A. P., VIDAL M., GIFFORD D. K., YOUNG R. A. and JAENISCH R. (2006) Polycomb complexes repress developmental regulators in murine embryonic stem cells, *Nature* **441**(7091), 349–353.

BRATTAIN M. G., FINE W. D., KHALED F. M., THOMPSON J. and BRATTAIN D. E. (1981) Heterogeneity of Malignant Cells from a Human Colonic Carcinoma, *Cancer Research* **41**(5), 1751–1756.

BREMNER K. H., SCHERER J., YI J., VERSHININ M., GROSS S. P. and VALLEE R. B. (2009) Adenovirus transport via direct interaction of cytoplasmic dynein with the viral capsid hexon subunit, *Cell host & microbe* **6**(6), 523–535.

BRICKELL P. M., LATCHMAN D. S., MURPHY D., WILLISON K. and RIGBY P. W. (1983) Activation of a Qa/Tla class I major histocompatibility antigen gene is a general feature of oncogenesis in the mouse, *Nature* **306**(5945), 756–760.

BRIND'AMOUR J., LIU S., HUDSON M., CHEN C., KARIMI M. M. and LORINCZ M. C. (2015) An ultra-low-input native ChIP-seq protocol for genome-wide profiling of rare cell populations, *Nature Communications* **6**.

BUONTEMPO S., LAISE P., HUGHES J. M., TRATTARO S., DAS V., RENCUREL C. and TESTA G. (2022) EZH2-Mediated H3K27me3 Targets Transcriptional Circuits of Neuronal Differentiation, *Frontiers in Neuroscience* **16**, 267.

BÜRGLIN T. R. (1991) The TEA domain: A novel, highly conserved DNA-binding motif.

BURNETT R. M. (1985) The structure of the adenovirus capsid. II. The packing symmetry of hexon and its implications for viral architecture, *Journal of Molecular Biology* **185**(1), 125–143.

BURNETT R. M., GRÜTTER M. G. and WHITE J. L. (1985) The structure of the adenovirus capsid. I. An envelope model of hexon at 6 Å resolution, *Journal of Molecular Biology* **185**(1), 105–123.

CAGLIERO J., FORGET A., DALDELLO E., SILBER J. and ZIDER A. (2013) The hippo kinase promotes scalloped cytoplasmic localization independently of warts in a CRM1/exportin1-dependent manner in drosophila, *FASEB Journal* **27**(4), 1330–1341.

CAI L., SUTTER B. M., LI B. and TU B. P. (2011) Acetyl-CoA Induces Cell Growth and Proliferation by Promoting the Acetylation of Histones at Growth Genes, *Molecular Cell* **42**(4), 426–437.

CANZIO D., LIAO M., NABER N., PATE E., LARSON A., WU S., MARINA D. B., GARCIA J. F., MADHANI H. D., COOKE R., SCHUCK P., CHENG Y. and NARLIKAR G. J. (2013) A conformational switch in HP1 releases auto-inhibition to drive heterochromatin assembly, *Nature* **496**(7445), 377–381.

CAO R., WANG L., WANG H., XIA L., ERDJUMENT-BROMAGE H., TEMPST P., JONES R. S. and ZHANG Y. (2002) Role of histone H3 lysine 27 methylation in polycomb-group silencing, *Science* **298**(5595), 1039–1043.

CARLSON M. (2022a) org.Hs.eg.db: Genome wide annotation for Human.

CARLSON M. (2022b) org.Rn.eg.db: Genome wide annotation for Rat.

CASALINO L., TALOTTA F., CIMMINO A. and VERDE P. (2022) The Fra-1/AP-1 Oncoprotein: From the "Undruggable" Transcription Factor to Therapeutic Targeting.

CAVANAGH H. M., MAHONY T. J. and VANNIASINKAM T. (2012) Genetic characterization of equine adenovirus type 1, *Veterinary Microbiology* **155**(1), 33–37.

CELEN A. B. and SAHIN U. (2020) Sumoylation on its 25th anniversary: mechanisms, pathology, and emerging concepts, *The FEBS Journal* **287**(15), 3110–3140.

CENIK B. K. and SHILATIFARD A. (2021) COMPASS and SWI/SNF complexes in development and disease.

CENTERS FOR DISEASE CONTROL AND PREVENTION (2022) Adenovirus Outbreaks — CDC.

CHALLBERG M. D. and KELLY T. J. (1979) Adenovirus DNA replication in vitro, *Proceedings of the National Academy of Sciences of the United States of America* **76**(2), 655–659.

CHANG H., LIU Y., XUE M., LIU H., DU S., ZHANG L. and WANG P. (2016) Synergistic action of master transcription factors controls epithelial-to-mesenchymal transition, *Nucleic Acids Research* **44**(6), 2514–2527.

CHEN X., YE Y., GU L., SUN J., DU Y., LIU W. J., LI W., ZHANG X. and JIANG C. (2019) H3K27me3 Signal in the Cis Regulatory Elements Reveals the Differentiation Potential of Progenitors During Drosophila Neuroglial Development.

CHENG C. Y., GILSON T., DALLAIRE F., KETNER G., BRANTON P. E. and BLANCHETTE P. (2011) The E4orf6/E1B55K E3 Ubiquitin Ligase Complexes of Human Adenoviruses Exhibit Heterogeneity in Composition and Substrate Specificity, *Journal of Virology* **85**(2), 765–775.

CHENG C. Y., GILSON T., WIMMER P., SCHREINER S., KETNER G., DOBNER T., BRANTON P. E. and BLANCHETTE P. (2013) Role of E1B55K in E4orf6/E1B55K E3 Ligase Complexes Formed by Different Human Adenovirus Serotypes, *Journal of Virology* **87**(11), 6232–6245.

CHIANG H. S. and LIU H. M. (2019) The molecular basis of viral inhibition of IRF- and STAT-dependent immune responses.

CHINENOV Y. and KERPPOLA T. K. (2001) Close encounters of many kinds: Fos-Jun interactions that mediate transcription regulatory specificity.

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), 2400–2415.

CHING W., KOYUNCU E., SINGH S., ARBELO-ROMAN C., HARTL B., KREMMER E., SPEISEDER T., MEIER C. and DOBNER T. (2013) A ubiquitin-specific protease possesses a decisive role for adenovirus replication and oncogene-mediated transformation, *PLoS pathogens* **9**(3), e1003273.

CHINNADURAI G. (1983) Adenovirus 2 lp+ locus codes for a 19 kd tumor antigen that plays an essential role in cell transformation, *Cell* **33**(3), 759–766.

CHIOU S. K., TSENG C. C., RAO L. and WHITE E. (1994) Functional complementation of the adenovirus E1B 19-kilodalton protein with Bcl-2 in the inhibition of apoptosis in infected cells, *Journal of Virology* **68**(10), 6553–6566.

CHOW L. T., GELINAS R. E., BROKER T. R. and ROBERTS R. J. (1977) An amazing sequence arrangement at the 5- ends of adenovirus 2 messenger RNA, *Cell* **12**(1), 1–8.

CHOWDHURY S. and SARKAR R. R. (2015) Comparison of human cell signaling pathway databases—evolution, drawbacks and challenges, *Database* **2015**.

CHRISTENSEN J., AGGER K., CLOOS P. A., PASINI D., ROSE S., SENNELS L., RAPPSILBER J., HANSEN K. H., SALCINI A. E. and HELIN K. (2007) RBP2 Belongs to a Family of Demethylases, Specific for Tri-and Dimethylated Lysine 4 on Histone 3, *Cell* **128**(6), 1063–1076.

CHYMKOWITCH P., NGUÉA P A. and ENSERINK J. M. (2015) SUMO-regulated transcription: Challenging the dogma, *BioEssays* **37**(10), 1095–1105.

COHEN J. (1960) A Coefficient of Agreement for Nominal Scales, *Educational and Psychological Measurement* **20**(1), 37–46.

CONSTANTIN AHLMANN-ELTZE (2020) ggupset: Combination Matrix Axis for 'ggplot2' to Create 'UpSet' Plots.

CORE L. J., MARTINS A. L., DANKO C. G., WATERS C. T., SIEPEL A. and LIS J. T. (2014) Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers, *Nature Genetics* **46**(12), 1311–1320.

COSULICH S. C., WORRALL V., HEDGE P. J., GREEN S. and CLARKE P. R. (1997) Regulation of apoptosis by BH3 domains in a cell-free system, *Current Biology* **7**(12), 913–920.

COWIESON N. P., PARTRIDGE J. F., ALLSHIRE R. C. and MCLAUGHLIN P. J. (2000) Dimerisation of a chromo shadow domain and distinctions from the chromodomain as revealed by structural analysis, *Current Biology* **10**(9), 517–525.

COYNE C. B. and BERGELSON J. M. (2005) CAR: A virus receptor within the tight junction.

CRAWFORD-MIKSZA L. and SCHNURR D. P. (1996) Analysis of 15 adenovirus hexon proteins reveals the location and structure of seven hypervariable regions containing serotype-specific residues, *Journal of Virology* **70**(3), 1836–1844.

CREYGHTON M. P., CHENG A. W., WELSTEAD G. G., KOOISTRA T., CAREY B. W., STEINE E. J., HANNA J., LODATO M. A., FRAMPTON G. M., SHARP P. A., BOYER L. A., YOUNG R. A. and JAENISCH R. (2010) Histone H3K27ac separates active from poised enhancers and predicts developmental state, *Proceedings of the National Academy of Sciences of the United States of America* **107**(50), 21931–21936.

CRICK F. (1970) Central dogma of molecular biology, *Nature* **227**(5258), 561–563.

CUCONATI A. and WHITE E. (2002) Viral homologs of BCL-2: Role of apoptosis in the regulation of virus infection.

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), 7707–7716.

CURRAN T., PETERS G., VAN BEVEREN C., TEICH N. M. and VERMA I. M. (1982) FBJ murine osteosarcoma virus: identification and molecular cloning of biologically active proviral DNA, *Journal of Virology* **44**(2), 674–682.

CURREY L., THOR S. and PIPER M. (2021) TEAD family transcription factors in development and disease, *Development (Cambridge)* **148**(12).

CUSANOVICH D. A., PAVLOVIC B., PRITCHARD J. K. and GILAD Y. (2014) The Functional Consequences of Variation in Transcription Factor Binding, *PLoS Genetics* **10**(3), e1004226.

DALE R., GRÜNING B., SJÖDIN A., ROWE J., CHAPMAN B. A., ROWE J., KÖSTER J., TOMKINS-TINCH C. H., VALIERIS R. and THE BIOCONDUCTOR TEAM (2018) Bioconda: Sustainable and comprehensive software distribution for the life sciences, *Nature Methods* **15**(7), 475–476.

DANECEK P., BONFIELD J. K., LIDDLE J., MARSHALL J., OHAN V., POLLARD M. O., WHITWHAM A., KEANE T., MCCARTHY S. A., DAVIES R. M. and LI H. (2021) Twelve years of SAMtools and BCFtools, *GigaScience* **10**(2).

DAVISON A. J., WRIGHT K. M. and HARRACH B. (2000) DNA sequence of frog adenovirus, *Journal of General Virology* **81**(10), 2431–2439.

DE STANCHINA E., MCCURRACH M. E., ZINDY F., SHIEH S. Y., FERBEYRE G., SAMUELSON A. V., PRIVES C., ROUSSEL M. F., SHERR C. J. and LOWE S. W. (1998) E1A signaling to p53 involves the p19(ARF) tumor suppressor, *Genes and Development* **12**(15), 2434–2442.

DEBBAS M. and WHITE E. (1993) Wild-type p53 mediates apoptosis by E1A, which is inhibited by E1B, *Genes and Development* **7**(4), 546–554.

DELEU L., SHELLARD S., ALEVIZOPOULOS K., AMATI B. and LAND H. (2001) Recruitment of trrap required for oncogenic transformation by E1A, *Oncogene* **20**(57), 8270–8275.

DO NASCIMENTO L. G., FIALHO A. M., DE ANDRADE J. D. S. R., DE ASSIS R. M. S. and FUMIAN T. M. (2022) Human enteric adenovirus F40/41 as a major cause of acute gastroenteritis in children in Brazil, 2018 to 2020, *Scientific Reports* **12**(1), 1–12.

DONATI G., MONTANARO L. and DERENZINI M. (2012) Ribosome biogenesis and control of cell proliferation: p53 is not alone.

DONEHOWER L. A., HARVEY M., SLAGLE B. L., MCARTHUR M. J., MONTGOMERY C. A., BUTEL J. S. and BRADLEY A. (1992) Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours, *Nature* **356**(6366), 215–221.

DONOVAN-BANFIELD I., TURNELL A. S., HISCOX J. A., LEPPARD K. N. and MATTHEWS D. A. (2020) Deep splicing plasticity of the human adenovirus type 5 transcriptome drives virus evolution, *Communications Biology* **3**(1), 1–14.

DOSZPOLY A., HARRACH B., LAPATRA S. and BENKŐ M. (2019) Unconventional gene arrangement and content revealed by full genome analysis of the white sturgeon adenovirus, the single member of the genus Ichtadenovirus, *Infection, Genetics and Evolution* **75**.

DOUCAS V., ISHOV A. M., ROMO A., JUGUILON H., WEITZMAN M. D., EVANS R. M. and MAUL G. G. (1996) Adenovirus replication is coupled with the dynamic properties of the PML nuclear structure, *Genes and Development* **10**(2), 196–207.

DULL T., ZUFFEREY R., KELLY M., MANDEL R. J., NGUYEN M., TRONO D. and NALDINI L. (1998) A Third-Generation Lentivirus Vector with a Conditional Packaging System, *Journal of Virology* **72**(11), 8463–8471.

DUMONTIER M. and HOGUE C. W. (2002) NBLAST: A cluster variant of BLAST for NxN comparisons, *BMC Bioinformatics* **3**(1), 13.

DUPONT S., MORSUT L., ARAGONA M., ENZO E., GIULITTI S., CORDENONSI M., ZANCONATO F., LE DIGABEL J., FORCATO M., BICCIATO S., ELVASSORE N. and PICCOLO S. (2011) Role of YAP/TAZ in mechanotransduction, *Nature* **474**(7350), 179–184.

DURINCK S., SPELLMAN P. T., BIRNEY E. and HUBER W. (2009) Mapping identifiers for the integration of genomic datasets with the R/ Bioconductor package biomaRt, *Nature Protocols* **4**(8), 1184–1191.

DYBAS J. M., LUM K. K., KULEJ K., REYES E. D., LAUMAN R., CHARMAN M., PURMAN C. E., STEINBOCK R. T., GRAMS N., PRICE A. M., MENDOZA L., GARCIA B. A. and WEITZMAN M. D. (2021) Adenovirus Remodeling of the Host Proteome and Host Factors Associated with Viral Genomes, *mSystems* **6**(4).

EBNER K., PINSKER W. and LION T. (2005) Comparative Sequence Analysis of the Hexon Gene in the Entire Spectrum of Human Adenovirus Serotypes: Phylogenetic, Taxonomic, and Clinical Implications, *Journal of Virology* **79**(20), 12635–12642.

ECKNER R., EWEN M. E., NEWSOME D., GERDES M., DECAPRIO J. A., LAWRENCE J. B. and LIVINGSTON D. M. (1994) Molecular cloning and functional analysis of the adenovirus E1A- associated 300-kD protein (p300) reveals a protein with properties of a transcriptional adaptor, *Genes and Development* **8**(8), 869–884.

EFERL R. and WAGNER E. F. (2003) AP-1: A double-edged sword in tumorigenesis.

EISENBERG E. and LEVANON E. Y. (2013) Human housekeeping genes, revisited.

EL-DEIRY W. S., KERN S. E., PIETENPOL J. A., KINZLER K. W. and VOGELSTEIN B. (1992) Definition of a consensus binding site for p53, *Nature Genetics* **1**(1), 45–49.

EL-FATTAH IBRAHIM S. A., ABUDU A., JONHSON E., AFTAB N., CONRAD S. and FLUCK M. (2018) The role of AP-1 in self-sufficient proliferation and migration of cancer cells and its potential impact on an autocrine/paracrine loop, *Oncotarget* **9**(76), 34259–34278.

ELSEN P. V. D., HOUWELING A. and EB A. V. D. (1983) Expression of region E1b of human adenoviruses in the absence of region E1a is not sufficient for complete transformation, *Virology* **128**(2), 377–390.

ENDERS J. F., BELL J. A., DINGLE J. H., FRANCIS T., HILLEMAN M. R., HUEBNER R. J. and PAYNE A. M. (1956) "Adenoviruses": Group Name Proposed for New Respiratory-Tract Viruses, *Science* **124**(3212), 119–120.

ENDTER C., HÄRTL B., SPRUSS T., HAUBER J. and DOBNER T. (2005) Blockage of CRM1-dependent nuclear export of the adenovirus type 5 early region 1B 55-kDa protein augments oncogenic transformation of primary rat cells, *Oncogene* **24**(1), 55–64.

ENDTER C., KZHYSHKOWSKA J., STAUBER R. and DOBNER T. (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), 11312–11317.

EWELS P., MAGNUSSON M., LUNDIN S. and KÄLLER M. (2016) MultiQC: summarize analysis results for multiple tools and samples in a single report, *Bioinformatics* **32**(19), 3047–3048.

EWER K. J., LAMBE T., ROLLIER C. S., SPENCER A. J., HILL A. V. and DORRELL L. (2016) Viral vectors as vaccine platforms: From immunogenicity to impact.

FABREGAT A., JUPE S., MATTHEWS L., SIDIROPOULOS K., GILLESPIE M., GARAPATI P., HAW R., JASSAL B., KORNINGER F., MAY B., MILACIC M., ROCA C. D., ROTHFELS K., SEVILLA C., SHAMOVSKY V., SHORSER S., VARUSAI T., VITERI G., WEISER J., WU G., STEIN L., HERMJAKOB H. and D'EUSTACHIO P. (2018) The Reactome Pathway Knowledgebase, *Nucleic Acids Research* **46**(D1), D649–D655.

FALLAUX F. J., KRANENBURG O., CRAMER S. J., HOUWELING A., VAN ORMONDT H., HOEBEN R. C. and VAN DER EB A. J. (1996) Characterization of 911: a new helper cell line for the titration and propagation of early region 1-deleted adenoviral vectors., *Human gene therapy* **7**(2), 215–222.

FARRANCE I. K., MAR J. H. and ORDAHL C. P. (1992) M-CAT binding factor is related to the SV40 enhancer binding factor, TEF-1, *Journal of Biological Chemistry* **267**(24), 17234–17240.

FATTAEY A. R., HARLOW E. and HELIN K. (1993) Independent regions of adenovirus E1A are required for binding to and dissociation of E2F-protein complexes, *Molecular and Cellular Biology* **13**(12), 7267–7277.

FEGHOUL L., CHEVRET S., CUINET A., DALLE J. H., OUACHÉE M., YACOUBEN K., FAHD M., GUÉRIN-EL KHOUROUJ V., ROUPRET-SERZEC J., STERKERS G., BARUCHEL A., SIMON F. and LEGOFF J. (2015) Adenovirus infection and disease in paediatric haematopoietic stem cell transplant patients: Clues for antiviral pre-emptive treatment, *Clinical Microbiology and Infection* **21**(7), 701–709.

FENG Y., PAN L., ZHANG B., HUANG H. and MA H. (2020) BATF acts as an oncogene in non-small cell lung cancer, *Oncology Letters* **19**(1), 205–210.

FERRARI R., GOU D., JAWDEKAR G., JOHNSON S. A., NAVA M., SU T., YOUSEF A. F., ZEMKE N. R., PELLEGRINI M., KURDISTANI S. K. and BERK A. J. (2014) Adenovirus small E1A employs the lysine acetylases p300/CBP and tumor suppressor RB to repress select host genes and promote productive virus infection, *Cell Host and Microbe* .

FERRARI R., PELLEGRINI M., HORWITZ G. A., XIE W., BERK A. J. and KURDISTANI S. K. (2008) Epigenetic reprogramming by adenovirus e1a, *Science* **321**(5892), 1086–1088.

FERRARI R., SU T., LI B., BONORA G., OBERAI A., CHAN Y., SASIDHARAN R., BERK A. J., PELLEGRINI M. and KURDISTANI S. K. (2012) Reorganization of the host epigenome by a viral oncogene, *Genome Research* .

FIGGE J., WEBSTER T., SMITH T. F. and PAUCHA E. (1988) Prediction of similar transforming regions in simian virus 40 large T, adenovirus E1A, and myc oncoproteins, *Journal of Virology* **62**(5), 1814–1818.

FISCHER M. (2017) Census and evaluation of p53 target genes.

FLAUS A., LUGER K., TAN S. and RICHMOND T. J. (1996) Mapping nucleosome position at single base-pair resolution by using site-directed hydroxyl radicals, *Proceedings of the National Academy of Sciences of the United States of America* **93**(4), 1370–1375.

FLINT S. J. and GONZALEZ R. A. (2003) Regulation of mRNA production by the adenoviral E1B 55-kDa and E4 Orf6 proteins.

FRANCIS N. J., KINGSTON R. E. and WOODCOCK C. L. (2004) Chromatin compaction by a polycomb group protein complex, *Science* **306**(5701), 1574–1577.

FRIETZE S. and FARNHAM P. (2011) Transcription Factor Effector Domains, in *A Handbook of Transcription Factors*, vol. 52, pp. 261–277, Springer New York.

FRISCH S. M. and MYMRYK J. S. (2002) Adenovirus-5 E1A: Paradox and paradigm.

FRITSCH L., ROBIN P., MATHIEU J. R., SOUIDI M., HINAUX H., ROUGEULLE C., HAREL-BELLAN A., AMEYAR-ZAZOUA M. and AIT-SI-ALI S. (2010) A Subset of the Histone H3 Lysine 9 Methyltransferases Suv39h1, G9a, GLP, and SETDB1 Participate in a Multimeric Complex, *Molecular Cell* **37**(1), 46–56.

GAGGAR A., SHAYAKHMETOV D. M. and LIEBER A. (2003) CD46 is a cellular receptor for group B adenoviruses, *Nature Medicine* **9**(11), 1408–1412.

GALLARDO J., PÉREZ-ILLANA M., MARTÍN-GONZÁLEZ N. and MARTÍN C. S. (2021) Adenovirus structure: What is new?

GALLIMORE P. H. and TURNELL A. S. (2001) Adenovirus E1A: remodelling the host cell, a life or death experience, *Oncogene* **20**(54), 7824–7835.

GAO T. and QIAN J. (2020) EnhancerAtlas 2.0: An updated resource with enhancer annotation in 586 tissue/cell types across nine species, *Nucleic Acids Research* **48**(D1), D58–D64.

GAO Z., ZHANG J., BONASIO R., STRINO F., SAWAI A., PARISI F., KLUGER Y. and REINBERG D. (2012) PCGF Homologs, CBX Proteins, and RYBP Define Functionally Distinct PRC1 Family Complexes, *Molecular Cell* **45**(3), 344–356.

GARNETT C. T., ERDMAN D., XU W. and GOODING L. R. (2002) Prevalence and Quantitation of Species C Adenovirus DNA in Human Mucosal Lymphocytes, *Journal of Virology* **76**(21), 10608–10616.

GARNETT C. T., TALEKAR G., MAHR J. A., HUANG W., ZHANG Y., ORNELLES D. A. and GOODING L. R. (2009) Latent Species C Adenoviruses in Human Tonsil Tissues, *Journal of Virology* **83**(6), 2417–2428.

GARNIER S. (2021) {viridis} - Colorblind-Friendly Color Maps for R.

GAZON H., BARBEAU B., MESNARD J. M. and PELOPONESE J. M. (2018) Hijacking of the AP-1 signaling pathway during development of ATL.

GENTLEMAN R., CAREY V., HUBER W. and HAHNE F. (2022) genefilter: genefilter: methods for filtering genes from high-throughput experiments.

GLOVER J. N. and HARRISON S. C. (1995) Crystal structure of the heterodimeric bZIP transcription factor c-Fos–c-Jun bound to DNA.

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), 4507–4519.

GRAHAM F. L., SMILEY J., RUSSELL W. C. and NAIRN R. (1977) Characteristics of a human cell line transformed by DNA from human adenovirus type 5, *The Journal of general virology* **36**(1), 59–74.

GRASSMANN R., ABOUD M. and JEANG K. T. (2005) Molecular mechanisms of cellular transformation by HTLV-1 Tax, *Oncogene* **24**(39), 5976–5985.

GRAU D. J., CHAPMAN B. A., GARLICK J. D., BOROWSKY M., FRANCIS N. J. and KINGSTON R. E. (2011) Compaction of chromatin by diverse polycomb group proteins requires localized regions of high charge, *Genes and Development* **25**(20), 2210–2221.

GROFF A. F., SANCHEZ-GOMEZ D. B., SORUCO M. M., GERHARDINGER C., BARUTCU A. R., LI E., ELCAVAGE L., PLANA O., SANCHEZ L. V., LEE J. C., SAUVAGEAU M. and RINN J. L. (2016) In Vivo Characterization of Linc-p21 Reveals Functional cis-Regulatory DNA Elements, *Cell Reports* **16**(8), 2178–2186.

GUANGCHUANG YU (2022) ggimage: Use Image in 'ggplot2'.

GUIBINGA G. H., MIYANOHARA A., ESKO J. D. and FRIEDMANN T. (2002) Cell surface heparan sulfate is a receptor for attachment of envelope protein-free retrovirus-like particles and VSV-G pseudotyped MLV-derived retrovirus vectors to target cells, *Molecular Therapy* **5**(5), 538–546.

GÜNTHER T., SCHREINER S., DOBNER T., TESSMER U. and GRUNDHOFF A. (2014) Influence of ND10 Components on Epigenetic Determinants of Early KSHV Latency Establishment, *PLoS Pathogens* **10**(7).

HAFNER A., BULYK M. L., JAMBHEKAR A. and LAHAV G. (2019) The multiple mechanisms that regulate p53 activity and cell fate.

HALAZONETIS T. D., DAVIS L. J. and KANDIL A. N. (1993) Wild-type p53 adopts a 'mutant'-like conformation when bound to DNA, *EMBO Journal* **12**(3), 1021–1028.

HAN J., SABBATINI P., PEREZ D., RAO L., MODHA D. and WHITE E. (1996) The E1B 19K protein blocks apoptosis by interacting with and inhibiting the p53-inducible and death-promoting Bax protein, *Genes and Development* **10**(4), 461–477.

HANAHAN D. (1983) Studies on transformation of Escherichia coli with plasmids, *Journal of molecular biology* **166**(4), 557–580.

HARLOW E., FRANZA B. R. and SCHLEY C. (1985) Monoclonal antibodies specific for adenovirus early region 1A proteins: extensive heterogeneity in early region 1A products, *Journal of Virology* **55**(3), 533–546.

HARRACH B., TARJÁN Z. L. and BENKŐ M. (2019) Adenoviruses across the animal kingdom: a walk in the zoo.

HARVEY B. G., MARONI J., O'DONOGHUE K. A., CHU K. W., MUSCAT J. C., PIPPO A. L., WRIGHT C. E., HOLLMANN C., WISNIVESKY J. P., KESSLER P. D., RASMUSSEN H. S., ROSENGART T. K. and CRYSTAL R. G. (2002) Safety of local delivery of low- and intermediate-dose adenovirus gene transfer vectors to individuals with a spectrum of morbid conditions, *Human Gene Therapy* **13**(1), 15–63.

HASSOU N., BOUSEETTINE R., ABOUCHOAIB N. and ENNAJI M. M. (2019) Enteric adenoviruses: Emerging of a public health threat, in *Emerging and Reemerging Viral Pathogens: Volume 1: Fundamental and Basic Virology Aspects of Human, Animal and Plant Pathogens*, pp. 879–905, Elsevier.

HAUPT Y., MAYA R., KAZAZ A. and OREN M. (1997) Mdm2 promotes the rapid degradation of p53, *Nature* **387**(6630), 296–299.

HE L., PRATT H., GAO M., WEI F., WENG Z. and STRUHL K. (2021) Yap and taz are transcriptional co-activators of ap-1 proteins and stat3 during breast cellular transformation, *eLife* **10**.

HEINZ S., BENNER C., SPANN N., BERTOLINO E., LIN Y. C., LASLO P., CHENG J. X., MURRE C., SINGH H. and GLASS C. K. (2010) Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities, *Molecular Cell* **38**(4), 576–589.

HERNANDO-PÉREZ M., MARTÍN-GONZÁLEZ N., PÉREZ-ILLANA M., SUOMALAINEN M., CONDEZO G. N., OSTAPCHUK P., GALLARDO J., MENÉNDEZ M., GREBER U. F., HEARING P., DE PABLO P. J. and MARTÍN C. S. (2020) Dynamic competition for hexon binding between core protein VII and lytic protein VI promotes adenovirus maturation and entry, *Proceedings of the National Academy of Sciences of the United States of America* **117**(24), 13699–13707.

HIDALGO P., IP W. H., DOBNER T. and GONZALEZ R. A. (2019) The biology of the adenovirus E1B 55K protein, *FEBS Letters* **593**(24), 3504–3517.

HINRICHS A. S., KAROLCHIK D., BAERTSCH R., BARBER G. P., BEJERANO G., CLAWSON H., DIEKHANS M., FUREY T. S., HARTE R. A., HSU F., HILLMAN-JACKSON J., KUHN R. M., PEDERSEN J. S., POHL A., RANEY B. J., ROSENBLOOM K. R., SIEPEL A., SMITH K. E., SUGNET C. W., SULTAN-QURRAIE A., THOMAS D. J., TRUMBOWER H., WEBER R. J., WEIRAUCH M., ZWEIG A. S., HAUSSLER D. and KENT W. J. (2006) The UCSC Genome Browser Database: update 2006., *Nucleic acids research* **34**(Database issue), D590–D598.

HIWARKAR P., KOSULIN K., CESARO S., MIKULSKA M., STYCZYNSKI J., WYNN R. and LION T. (2018) Management of adenovirus infection in patients after haematopoietic stem cell transplantation: State-of-the-art and real-life current approach, *Reviews in Medical Virology* **28**(3), e1980.

HOFFMANN A., SINN E., YAMAMOTO T., WANG J., ROY A., HORIKOSHI M. and ROEDER R. G. (1990) Highly conserved core domain and unique N terminus with presumptive regulatory motifs in a human TATA factor (TFIID), *Nature* **346**(6282), 387–390.

HOLDEN J. K. and CUNNINGHAM C. N. (2018) Targeting the hippo pathway and cancer through the TEAD family of transcription factors.

HOPPE E., PAULY M., GILLESPIE T. R., AKOUA-KOFFI C., HOHMANN G., FRUTH B., KARHEMERE S., MADINDA N. F., MUGISHA L., MUYEMBE J. J., TODD A., PETRZELKOVA K. J., GRAY M., ROBBINS M., BERGL R. A., WITTIG R. M., ZUBERBÜHLER K., BOESCH C., SCHUBERT G., LEENDERTZ F. H., EHLERS B. and CALVIGNAC-SPENCER S. (2015) Multiple cross-species transmission events of human adenoviruses (HAdV) during hominine evolution, *Molecular Biology and Evolution* **32**(8), 2072–2084.

HORRIDGE J. J. and LEPPARD K. N. (1998) RNA-Binding Activity of the E1B 55-Kilodalton Protein from Human Adenovirus Type 5, *Journal of Virology* **72**(11), 9374–9379.

HORWITZ G. A., ZHANG K., MCBRIAN M. A., GRUNSTEIN M., KURDISTANI S. K. and BERK A. J. (2008) Adenovirus Small e1a Alters Global Patterns of Histone Modification, *Science* **321**(5892), 1084–1085.

HOUWELING A., VAN DEN ELSEN P. J. and VAN DER EB A. J. (1980) Partial transformation of primary rat cells by the leftmost 4.5% fragment of adenovirus 5 DNA, *Virology* **105**(2), 537–550.

HOWE J. A., MYMRYK J. S., EGAN C., BRANTON P. E. and BAYLEY S. T. (1990) Retinoblastoma growth suppressor and a 300-kDa protein appear to regulate cellular DNA synthesis, *Proceedings of the National Academy of Sciences of the United States of America* **87**(15), 5883–5887.

HOWLEY P. M. and LIVINGSTON D. M. (2009) Small DNA tumor viruses: Large contributors to biomedical sciences.

HUA-VAN A., LE ROUZIC A., BOUTIN T. S., FILÉE J. and CAPY P. (2011) The struggle for life of the genome's selfish architects, *Biology Direct* **6**(1), 19.

HUANG S., REDDY V., DASGUPTA N. and NEMEROW G. R. (1999) A Single Amino Acid in the Adenovirus Type 37 Fiber Confers Binding to Human Conjunctival Cells, *Journal of Virology* **73**(4), 2798–2802.

HUBER W., VON HEYDEBRECK A., SULTMANN H., POUSTKA A. and VINGRON M. (2002) Variance stabilization applied to microarray data calibration and to the quantification of differential expression, *Bioinformatics* **18**(Suppl 1), S96–S104.

HUH H. D., KIM D. H., JEONG H. S. and PARK H. W. (2019) Regulation of TEAD transcription factors in cancer biology.

HUNT R. W. and WASSERMAN W. W. (2014) Non-targeted transcription factors motifs are a systemic component of ChIP-seq datasets, *Genome Biology* **15**(7), 412.

IP W. H. and DOBNER T. (2020) Cell transformation by the adenovirus oncogenes E1 and E4, *FEBS Letters* **594**(12), 1848–1860.

IRAGAVARAPU A. G., YAO L. and KASINATH V. (2021) Structural insights into the interactions of Polycomb Repressive Complex 2 with chromatin.

ISAKOVA A., GROUX R., IMBEAULT M., RAINER P., ALPERN D., DAINESE R., AMBROSINI G., TRONO D., BUCHER P. and DEPLANCKE B. (2017) SMiLE-seq identifies binding motifs of single and dimeric transcription factors, *Nature Methods* **14**(3), 316–322.

ISHOV A. M. and MAUL G. G. (1996) The periphery of nuclear domain 10 (ND10) as site of DNA virus deposition, *Journal of Cell Biology* **134**(4), 815–826.

ISMAIL A. M., LEE J. S., DYER D. W., SETO D., RAJAIYA J. and CHODOSH J. (2016) Selection Pressure in the Human Adenovirus Fiber Knob Drives Cell Specificity in Epidemic Keratoconjunctivitis, *Journal of Virology* **90**(21), 9598–9607.

JAIN D., BALDI S., ZABEL A., STRAUB T. and BECKER P. B. (2015) Active promoters give rise to false positive 'Phantom Peaks' in ChIP-seq experiments, *Nucleic Acids Research* **43**(14), 6959–6968.

JALILI V., MATTEUCCI M., MASSEROLI M. and MORELLI M. J. (2015) Using combined evidence from replicates to evaluate ChIP-seq peaks, *Bioinformatics* **31**(17), 2761–2769.

JALILI V., MATTEUCCI M., MORELLI M. J. and MASSEROLI M. (2016) MuSERA: Multiple Sample Enriched Region Assessment, *Briefings in Bioinformatics* **18**(3), bbw029.

JAMES L., VERNON M. O., JONES R. C., STEWART A., LU X., ZOLLAR L. M., CHUDOBA M., WESTERCAMP M., ALCASID G., DUFFEE-KERR L., WOOD L., BOONLAYANGOOR S., BETHEL C., RITGER K., CONOVER C., ERDMAN D. D. and GERBER S. I. (2007) Outbreak of Human Adenovirus Type 3 Infection in a Pediatric Long-Term Care Facility–Illinois, 2005, *Clinical Infectious Diseases* **45**(4), 416–420.

JIN C., ZANG C., WEI G., CUI K., PENG W., ZHAO K. and FELSENFELD G. (2009) H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions, *Nature Genetics* **41**(8), 941–945.

JOERGER A. C. and FERSHT A. R. (2016) The p53 Pathway: Origins, Inactivation in Cancer, and Emerging Therapeutic Approaches, *Annual Review of Biochemistry* **85**, 375–404.

JOHNSON W. E. (2019) Origins and evolutionary consequences of ancient endogenous retroviruses.

JONES M. S., HARRACH B., GANAC R. D., GOZUM M. M. A., DELA CRUZ W. P., RIEDEL B., PAN C., DELWART E. L. and SCHNURR D. P. (2007) New Adenovirus Species Found in a Patient Presenting with Gastroenteritis, *Journal of Virology* **81**(11), 5978–5984.

JONES N. and SHENK T. (1979) An adenovirus type 5 early gene function regulates expression of other early viral genes, *Proceedings of the National Academy of Sciences of the United States of America* **76**(8), 3665–3669.

JONSSON N. and ANKERST J. (1977) Studies on adenovirus type 9-induced mammary fibroadenomas in rats and their malignant transformation, *Cancer* **39**(6), 2513–2519.

JOSUE RUIZ E., LAN L., DIEFENBACHER M. E., RIISING E. M., DA COSTA C., CHAKRABORTY A., HOECK J. D., SPENCER-DENE B., KELLY G., DAVID J. P., NYE E., DOWNWARD J. and BEHRENS A. (2021) JunD, not c-Jun, is the AP-1 transcription factor required for Ras-induced lung cancer, *JCI Insight* **6**(13).

JUMPER J., EVANS R., PRITZEL A., GREEN T., FIGURNOV M., RONNEBERGER O., TUNYASUVUNAKOOL K., BATES R., Z'IDEK A., POTAPENKO A., BRIDGLAND A., MEYER C., KOHL S. A. A., BALLARD A. J., COWIE A., ROMERA-PAREDES B., NIKOLOV S., JAIN R., ADLER J., BACK T., PETERSEN S., REIMAN D., CLANCY E., ZIELINSKI M., STEINEGGER M., PACHOLSKA M., BERGHAMMER T., BODENSTEIN S., SILVER D., VINYALS O., SENIOR A. W., KAVUKCUOGLU K., KOHLI P. and HASSABIS D. (2021) Highly accurate protein structure prediction with AlphaFold., *Nature* .

KANEHISA M. and GOTO S. (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes.

KAPLAN J. (2005) Adenovirus-Based Cancer Gene Therapy, *Current Gene Therapy* **5**(6), 595–605.

KHERADPOUR P. and KELLIS M. (2014) Systematic discovery and characterization of regulatory motifs in ENCODE TF binding experiments, *Nucleic Acids Research* **42**(5), 2976–2987.

KIDD A. H., CHROBOCZEK J., CUSACK S. and RUIGROK R. W. (1993) Adenovirus type 40 virions contain two distinct fibers, *Virology* **192**(1), 73–84.

KIDDER B. L., HU G. and ZHAO K. (2011) ChIP-Seq: Technical considerations for obtaining high-quality data.

KIKUCHI J., KINOSHITA I., SHIMIZU Y., OIZUMI S., NISHIMURA M., BIRRER M. J. and DOSAKA-AKITA H. (2008) Simultaneous blockade of AP-1 and phosphatidylinositol 3-kinase pathway in non-small cell lung cancer cells, *British Journal of Cancer* **99**(12), 2013–2019.

KILLERBY M. E., ROZWADOWSKI F., LU X., CAULCRICK-GRIMES M., MCHUGH L., HALDEMAN A. M., FULTON T., SCHNEIDER E., SAKTHIVEL S. K., BHATNAGAR J., RABENECK D. B., ZAKI S., GERBER S. I. and WATSON J. T. (2019) Respiratory Illness Associated With Emergent Human Adenovirus Genome Type 7d, New Jersey, 2016–2017, *Open Forum Infectious Diseases* **6**(2).

KIM M. K., JEON B. N., KOH D. I., KIM K. S., PARK S. Y., YUN C. O. and HUR M. W. (2013) Regulation of the cyclin-dependent kinase inhibitor 1A gene (CDKN1A) by the repressor BOZF1 through inhibition of p53 acetylation and transcription factor Sp1 binding, *Journal of Biological Chemistry* **288**(10), 7053–7064.

KIM S., YU N. K. and KAANG B. K. (2015) CTCF as a multifunctional protein in genome regulation and gene expression.

KINDSMÜLLER K., GROITL P., HÄRTL B., BLANCHETTE P., HAUBER J. and DOBNER T. (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), 6684–6689.

KING C. R., ZHANG A., TESSIER T. M., GAMEIRO S. F. and MYMRYK J. S. (2018) Hacking the cell: Network intrusion and exploitation by adenovirus E1A.

KNUDSEN E. S., NAMBIAR R., ROSARIO S. R., SMIRAGLIA D. J., GOODRICH D. W. and WITKIEWICZ A. K. (2020) Pan-cancer molecular analysis of the RB tumor suppressor pathway, *Communications Biology* **3**(1), 1–12.

KOLBE V., IP W. H., KIEWEG-THOMPSON L., LANG J., GRUHNE J., MEYER T., WILKENS B., SCHIE M., THÜNAUER R., SCHREINER S., BERTZBACH L. D., RODRÍGUEZ E. and DOBNER T. (2022) Conserved E1B-55K SUMOylation in Different Human Adenovirus Species Is a Potent Regulator of Intracellular Localization, *Journal of Virology* **96**(3).

KOLDE R. (2019) pheatmap: Pretty Heatmaps.

KOMATSU T., SEKIYA T. and NAGATA K. (2013) DNA replication-dependent binding of CTCF plays a critical role in adenovirus genome functions, *Scientific Reports* **3**(1), 2187.

KOSULIN K., GEIGER E., VÉCSEI A., HUBER W. D., RAUCH M., BRENNER E., WRBA F., HAMMER K., INNERHOFER A., PÖTSCHGER U., LAWITSCHKA A., MATTHES-LEODOLTER

S., FRITSCH G. and LION T. (2016) Persistence and reactivation of human adenoviruses in the gastrointestinal tract, *Clinical Microbiology and Infection* **22**(4), 1–381.

KOSULIN K., HABERLER C., HAINFELLNER J. A., AMANN G., LANG S. and LION T. (2007) Investigation of Adenovirus Occurrence in Pediatric Tumor Entities, *Journal of Virology* **81**(14), 7629–7635.

KOVESDI I., REICHEL R. and NEVINS J. R. (1986) Identification of a cellular transcription factor involved in E1A trans-activation, *Cell* **45**(2), 219–228.

KRÄTZER F., ROSORIUS O., HEGER P., HIRSCHMANN N., DOBNER T., HAUBER J. and STAUBER R. H. (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), 850–857.

KREMER E. J. (2021) What is the risk of a deadly adenovirus pandemic?, *PLOS Pathogens* **17**(9), e1009814.

KRIPPL B., FERGUSON B., ROSENBERG M. and WESTPHAL H. (1984) Functions of purified E1A protein microinjected into mammalian cells, *Proceedings of the National Academy of Sciences of the United States of America* **81**(22 I), 6988–6992.

KRUMP N. A. and YOU J. (2018) Molecular mechanisms of viral oncogenesis in humans.

KUHLMANN I., ACHTEN S., RUDOLPH R. and DOERFLER W. (1982) Tumor induction by human adenovirus type 12 in hamsters: loss of the viral genome from adenovirus type 12-induced tumor cells is compatible with tumor formation., *The EMBO journal* **1**(1), 79–86.

KUNDHAVAI NATCHIAR S., VENKATARAMAN S., MULLEN T. M., NEMEROW G. R. and REDDY V. S. (2018) Revised Crystal Structure of Human Adenovirus Reveals the Limits on Protein IX Quasi-Equivalence and on Analyzing Large Macromolecular Complexes, *Journal of Molecular Biology* **430**(21), 4132–4141.

LACHNER M., O'CARROLL D., REA S., MECHTLER K. and JENUWEIN T. (2001) Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins, *Nature* **410**(6824), 116–120.

LAMBERT S. A., JOLMA A., CAMPITELLI L. F., DAS P. K., YIN Y., ALBU M., CHEN X., TAIPALE J., HUGHES T. R. and WEIRAUCH M. T. (2018) The Human Transcription Factors.

LANE D. P. (1992) p53, guardian of the genome.

LANGE C., CAKIROGLU F., SPIESS A. N., CAPPALLO-OBERMANN H., DIERLAMM J. and ZANDER A. R. (2007) Accelerated and safe expansion of human mesenchymal stromal cells in animal serum-free medium for transplantation and regenerative medicine, *Journal of Cellular Physiology* **213**(1), 18–26.

LANGMEAD B. and SALZBERG S. L. (2012) Fast gapped-read alignment with Bowtie 2, *Nature Methods* **9**(4), 357–359.

LAWRENCE M., HUBER W., PAGÈS H., ABOYOUN P., CARLSON M., GENTLEMAN R., MORGAN M. T. and CAREY V. J. (2013) Software for Computing and Annotating Genomic Ranges, *PLoS Computational Biology* **9**(8), e1003118.

LEE B., DAMON C. F. and PLATTS-MILLS J. A. (2020) Pediatric acute gastroenteritis associated with adenovirus 40/41 in low-income and middle-income countries.

LEE J., KIM D. H., LEE S., YANG Q. H., DONG K. L., LEE S. K., ROEDER R. G. and LEE J. W. (2009) A tumor suppressive coactivator complex of p53 containing ASC-2 and histone H3-lysine-4 methyltransferase MLL3 or its paralogue MLL4, *Proceedings of the National Academy of Sciences of the United States of America* **106**(21), 8513–8518.

LERDRUP M., JOHANSEN J. V., AGRAWAL-SINGH S. and HANSEN K. (2016) An interactive environment for agile analysis and visualization of ChIP-sequencing data, *Nature Structural and Molecular Biology* **23**(4), 349–357.

LEVINE A. J. (2009) The common mechanisms of transformation by the small DNA tumor viruses: The inactivation of tumor suppressor gene products: p53.

LI E., STUPACK D., KLEMKE R., CHERESH D. A. and NEMEROW G. R. (1998) Adenovirus Endocytosis via α v Integrins Requires Phosphoinositide-3-OH Kinase , *Journal of Virology* **72**(3), 2055–2061.

LIFTON R. P., GOLDBERG M. L., KARP R. W. and HOGNESS D. S. (1977) The organization of the histone genes in Drosophila melanogaster: Functional and evolutionary implications, *Cold Spring Harbor Symposia on Quantitative Biology* **42**(2), 1047–1051.

LILL N. L., GROSSMAN S. R., GINSBERG D., DECAPRIO J. and LIVINGSTON D. M. (1997) Binding and modulation of p53 by p300/CBP coactivators, *Nature* **387**(6635), 823–827.

LIN D. L. and CHANG C. (1996) p53 is a mediator for radiation-repressed human TR2 orphan receptor expression in MCF-7 cells, a new pathway from tumor suppressor to member of the steroid receptor superfamily, *The Journal of biological chemistry* **271**(25), 14649–14652.

LION T. (2014) Adenovirus infections in immunocompetent and immunocompromised patients, *Clinical Microbiology Reviews* **27**(3), 441–462.

LION T., KOSULIN K., LANDLINGER C., RAUCH M., PREUNER S., JUGOVIC D., PÖTSCHGER U., LAWITSCHKA A., PETERS C., FRITSCH G. and MATTHES-MARTIN S. (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), 706–714.

LIU H., JIN L., KOH S. B. S., ATANASOV I., SCHEIN S., WU L. and ZHOU Z. H. (2010) Atomic structure of human adenovirus by Cryo-EM reveals interactions among protein networks, *Science* **329**(5995), 1038–1043.

LIU X., HONG T., PARAMESWARAN S., ERNST K., MARAZZI I., WEIRAUCH M. T. and FUXMAN BASS J. I. (2020) Human Virus Transcriptional Regulators.

LIU X., LI H., RAJURKAR M., LI Q., COTTON J. L., OU J., ZHU L. J., GOEL H. L., MERCURIO A. M., PARK J. S., DAVIS R. J. and MAO J. (2016) Tead and AP1 Coordinate Transcription and Motility, *Cell Reports* **14**(5), 1169–1180.

LIU X. and MARMORSTEIN R. (2007) Structure of the retinoblastoma protein bound to adenovirus E1A reveals the molecular basis for viral oncoprotein inactivation of a tumor suppressor, *Genes and Development* **21**(21), 2711–2716.

LIU Y., SHEVCHENKO A., SHEVCHENKO A. and BERK A. J. (2005) Adenovirus Exploits the Cellular Aggresome Response To Accelerate Inactivation of the MRN Complex, *Journal of Virology* **79**(22), 14004–14016.

LIZIO M., HARSHBARGER J., SHIMOJI H., SEVERIN J., KASUKAWA T., SAHIN S., ABUGESSAISA I., FUKUDA S., HORI F., ISHIKAWA-KATO S., MUNGALL C. J., ARNER E., BAILLIE J. K., BERTIN N., BONO H., DE HOON M., DIEHL A. D., DIMONT E., FREEMAN T. C., FUJIEDA K., HIDE W., KALIYAPERUMAL R., KATAYAMA T., LASSMANN T., MEEHAN T. F., NISHIKATA K., ONO H., REHLI M., SANDELIN A., SCHULTES E. A., 'T HOEN P. A., TATUM Z., THOMPSON M., TOYODA T., WRIGHT D. W., DAUB C. O., ITOH M., CARNINCI P., HAYASHIZAKI Y., FORREST A. R. and KAWAJI H. (2015) Gateways to the FANTOM5 promoter level mammalian expression atlas, *Genome Biology* **16**(1), 22.

LOMONOSOVA E., SUBRAMANIAN T. and CHINNADURAI G. (2005) Mitochondrial localization of p53 during adenovirus infection and regulation of its activity by E1B-19K, *Oncogene* **24**(45), 6796–6808.

LONG H. S., GREENAWAY S., POWELL G., MALLON A. M., LINDGREN C. M. and SIMON M. M. (2022) Making sense of the linear genome, gene function and TADs, *Epigenetics and Chromatin* **15**(1), 4.

LONGWORTH M. S. and DYSON N. J. (2010) PRb, a local chromatin organizer with global possibilities.

LOVE M. (2022) tximportData: tximportData.

LOVE M. I., HUBER W. and ANDERS S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, *Genome Biology* **15**(12), 550.

LOVE M. I., SONESON C., HICKEY P. F., JOHNSON L. K., PIERCE N. T., SHEPHERD L., MORGAN M. and PATRO R. (2020) Tximeta: Reference sequence checksums for provenance identification in RNA-seq, *PLOS Computational Biology* **16**(2), e1007664.

LOW B. C., PAN C. Q., SHIVASHANKAR G., BERSHADSKY A., SUDOL M. and SHEETZ M. (2014) YAP/TAZ as mechanosensors and mechanotransducers in regulating organ size and tumor growth, *FEBS Letters* **588**(16), 2663–2670.

LOWE S. W., JACKS T., HOUSMAN D. E. and RULEY H. E. (1994) Abrogation of oncogene-associated apoptosis allows transformation of p53- deficient cells, *Proceedings of the National Academy of Sciences of the United States of America* **91**(6), 2026–2030.

LU Y., CHAN Y. T., TAN H. Y., LI S., WANG N. and FENG Y. (2020) Epigenetic regulation in human cancer: The potential role of epi-drug in cancer therapy.

LUGER K., MÄDER A. W., RICHMOND R. K., SARGENT D. F. and RICHMOND T. J. (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution, *Nature* **389**(6648), 251–260.

LYONS R. H., FERGUSON B. Q. and ROSENBERG M. (1987) Pentapeptide nuclear localization signal in adenovirus E1a, *Molecular and Cellular Biology* **7**(7), 2451–2456.

MA S., MENG Z., CHEN R. and GUAN K.-L. (2019) The Hippo Pathway: Biology and Pathophysiology, *Annual Review of Biochemistry* **88**(1), 577–604.

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), 4657–4661.

MAKI Y., BOS T. J., DAVIS C., STARBUCK M. and VOGT P. K. (1987) Avian sarcoma virus 17 carries the jun oncogene., *Proceedings of the National Academy of Sciences* **84**(9), 2848–2852.

MANGEL W. and SAN MARTÍN C. (2014) Structure, Function and Dynamics in Adenovirus Maturation, *Viruses* **6**(11), 4536–4570.

MARABINI R., CONDEZO G. N., KRUPOVIC M., MENÉNDEZ-CONEJERO R., GÓMEZ-BLANCO J. and MARTÍN C. S. (2021) Near-atomic structure of an atadenovirus reveals a conserved capsid-binding motif and intergenera variations in cementing proteins, *Science Advances* **7**(14).

MARGUERON R., LI G., SARMA K., BLAIS A., ZAVADIL J., WOODCOCK C. L., DYNLACHT B. D. and REINBERG D. (2008) Ezh1 and Ezh2 Maintain Repressive Chromatin through Different Mechanisms, *Molecular Cell* **32**(4), 503–518.

MARION S., SAN MARTÍN C. and ŠIBER A. (2017) Role of Condensing Particles in Polymer Confinement: A Model for Virus-Packed "Minichromosomes", *Biophysical Journal* **113**(8), 1643–1653.

MARTENS M., AMMAR A., RIUTTA A., WAAGMEESTER A., SLENTER D. N., HANSPERS K., MILLER R. A., DIGLES D., LOPES E. N., EHRHART F., DUPUIS L. J., WINCKERS L. A., COORT S. L., WILLIGHAGEN E. L., EVELO C. T., PICO A. R. and KUTMON M. (2021) WikiPathways: Connecting communities, *Nucleic Acids Research* **49**(D1), D613–D621.

MARTIN M. E. D. and BERK A. J. (1998) Adenovirus E1B 55K Represses p53 Activation In Vitro, *Journal of Virology* **72**(4), 3146–3154.

MATTHES-MARTIN S., FEUCHTINGER T., SHAW P., ENGELHARD D., HIRSCH H., CORDONNIER C. and LJUNGMAN P. (2012) European guidelines for diagnosis and treatment of adenovirus infection in leukemia and stem cell transplantation: summary of ECIL-4 (2011), *Transplant Infectious Disease* **14**(6), 555–563.

MAUTNER V., STEINTHORSDOTTIR V. and BAILEY A. (1995) Enteric adenoviruses.

MAYER C. and GRUMMT I. (2006) Ribosome biogenesis and cell growth: mTOR coordinates transcription by all three classes of nuclear RNA polymerases.

MAYRAN A. and DROUIN J. (2018) Pioneer transcription factors shape the epigenetic landscape.

MCFERRAN J. B., MCCRACKEN R. M., MCKILLOP E. R., MCNULTY M. S. and COLLINS D. S. (1978) Studies on a depressed egg production syndrome in northern ireland, *Avian Pathology* **7**(1), 35–47.

MCLORIE W., MCGLADE C. J., TAKAYESU D. and BRANTON P. E. (1991) Individual adenovirus E1B proteins induce transformation independently but by additive pathways, *Journal of General Virology* **72**(6), 1467–1471.

MEIER O., BOUCKE K., HAMMER S. V., KELLER S., STIDWILL R. P., HEMMI S. and GREBER U. F. (2002) Adenovirus triggers macropinocytosis and endosomal leakage together with its clathrin-mediated uptake, *The Journal of cell biology* **158**(6), 1119–1131.

MELO C. A., DROST J., WIJCHERS P. J., VAN DE WERKEN H., DE WIT E., VRIELINK J. A., ELKON R., MELO S. A., LÉVEILLÉ N., KALLURI R., DE LAAT W. and AGAMI R. (2013) ERNAs Are Required for p53-Dependent Enhancer Activity and Gene Transcription, *Molecular Cell* **49**(3), 524–535.

MERRICK R. M., GRAND R. J., BROWN J. C. and GALLIMORE P. H. (1991) The use of β-galactosidase fusion proteins encoding the early region 1 transforming proteins of adenovirus type 12 to examine the humoral response in tumour-bearing animals, *Journal of General Virology* **72**(4), 955–960.

MILLER D. L., RICKARDS B., MASHIBA M., HUANG W. and FLINT S. J. (2009) The Adenoviral E1B 55-Kilodalton Protein Controls Expression of Immune Response Genes but Not p53-Dependent Transcription, *Journal of Virology* **83**(8), 3591–3603.

MILLER T., KROGAN N. J., DOVER J., ERDJUMENT-BROMAGE H., TEMPST P., JOHNSTON M., GREENBLATT J. F. and SHILATIFARD A. (2001) COMPASS: A complex of proteins associated with a trithorax-related SET domain protein, *Proceedings of the National Academy of Sciences of the United States of America* **98**(23), 12902–12907.

MIRDITA M., SCHÜTZE K., MORIWAKI Y., HEO L., OVCHINNIKOV S. and STEINEGGER M. (2022) ColabFold: making protein folding accessible to all, *Nature Methods* **19**(6), 679–682.

MIRDITA M., STEINEGGER M. and SÖDING J. (2019) MMseqs2 desktop and local web server app for fast, interactive sequence searches, *Bioinformatics* **35**(16), 2856–2858.

MITSUDOMI T., STEINBERG S. M., NAU M. M., CARBONE D., D'AMICO D., BODNER S., OIE H. K., LINNOILA R. I., MULSHINE J. L., MINNA J. D. and GAZDAR A. F. (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**(1), 171–180.

MORAN E., GRODZICKER T., ROBERTS R. J., MATHEWS M. B. and ZERLER B. (1986) Lytic and transforming functions of individual products of the adenovirus E1A gene, *Journal of virology* **57**(3), 765–775.

MORGEN M. (2022) BiocManager: Access the Bioconductor Project Package Repository.

MORRISON O. and THAKUR J. (2021) Molecular complexes at euchromatin, heterochromatin and centromeric chromatin, *International Journal of Molecular Sciences* **22**(13), 6922.

MULLER S. and DOBNER T. (2008) The adenovirus E1B-55K oncoprotein induces SUMO modification of p53, *Cell Cycle* **7**(6), 754–758.

MULLER U., ROBERTS M. P., ENGEL D. A., DOERFLER W. and SHENK T. (1989) Induction of transcription factor AP-1 by adenovirus E1A protein and cAMP, *Genes and Development* **3**(12 A), 1991–2002.

NAKABEPPU Y., RYDER K. and NATHANS D. (1988) DNA binding activities of three murine Jun proteins: Stimulation by Fos, *Cell* **55**(5), 907–915.

NESTIĆ D., BOŽINOVIĆ K., PEHAR I., WALLACE R., PARKER A. L. and MAJHEN D. (2021) The revolving door of adenovirus cell entry: Not all pathways are equal.

NESTIĆ D., UIL T. G., MA J., ROY S., VELLINGA J., BAKER A. H., CUSTERS J. and MAJHEN D. (2019) αvβ3 Integrin Is Required for Efficient Infection of Epithelial Cells with Human Adenovirus Type 26, *Journal of Virology* **93**(1).

NEUWIRTH E. (2022) RColorBrewer: ColorBrewer Palettes.

NEVELS M., TÄUBER B., KREMMER E., SPRUSS T., WOLF H. and DOBNER T. (1999) Transforming Potential of the Adenovirus Type 5 E4orf3 Protein, *Journal of Virology* **73**(2), 1591–1600.

NEVINS J. R., GINSBERG H. S., BLANCHARD J.-M., WILSON M. C. and DARNELL J. E. (1979) Regulation of the Primary Expression of the Early Adenovirus Transcription Units, *Journal of Virology* **32**(3), 727–733.

NGUYEN T. A., JONES R. D., SNAVELY A. R., PFENNING A. R., KIRCHNER R., HEMBERG M. and GRAY J. M. (2016) High-throughput functional comparison of promoter and enhancer activities, *Genome Research* **26**(8), 1023–1033.

NISHIZAWA M., KATAOKA K., GOTO N., FUJIWARA K. T. and KAWAI S. (1989) v-maf, a viral oncogene that encodes a 'leucine zipper' motif, *Proceedings of the National Academy of Sciences of the United States of America* **86**(20), 7711–7715.

NORRBY E. and WADELL G. (1967) Soluble components of adenovirus type 4, *Virology* **31**(4), 592–600.

OHLENSCHLÄGER O., SEIBOTH T., ZENGERLING H., BRIESE L., MARCHANKA A., RAMACHANDRAN R., BAUM M., KORBAS M., MEYER-KLAUCKE W., DÜRST M. and GÖRLACH M. (2006) Solution structure of the partially folded high-risk human papilloma virus 45 oncoprotein E7, *Oncogene* **25**(44), 5953–5959.

OLINER J. D., PIETENPOL J. A., THIAGALINGAM S., GYURIS J., KINZLER K. W. and VOGELSTEIN B. (1993) Oncoprotein MDM2 conceals the activation domain of tumour suppressor p53, *Nature* **362**(6423), 857–860.

ONODERA A., KOKUBO K. and NAKAYAMA T. (2018) The Interplay between Transcription Factors and Epigenetic Modifications in Th2 Cells, in *Gene Expression and Regulation in Mammalian Cells - Transcription From General Aspects*, InTech.

OSTAPCHUK P., SUOMALAINEN M., ZHENG Y., BOUCKE K., GREBER U. F. and HEARING P. (2017) The adenovirus major core protein VII is dispensable for virion assembly but is essential for lytic infection, *PLOS Pathogens* **13**(6), e1006455.

PAGES H., CARLSON M., FALCON S. and LI N. (2022) AnnotationDbi: Manipulation of SQLite-based annotations in Bioconductor.

PAN M., CHEW T. W., PEI WONG D. C., XIAO J., ONG H. T., LI CHIN J. F. and LOW B. C. (2020) BNIP-2 retards breast cancer cell migration by coupling microtubule-mediated GEF-H1 and RhoA activation, *Science Advances* **6**(31).

PANIGRAHI A. and O'MALLEY B. W. (2021) Mechanisms of enhancer action: the known and the unknown.

PANTELIC R. S., LOCKETT L. J., ROTHNAGEL R., HANKAMER B. and BOTH G. W. (2008) Cryoelectron Microscopy Map of Atadenovirus Reveals Cross-Genus Structural Differences from Human Adenovirus , *Journal of Virology* **82**(15), 7346–7356.

PARK S. Y. and KIM J. S. (2020) A short guide to histone deacetylases including recent progress on class II enzymes.

PARKER J. B., PALCHAUDHURI S., YIN H., WEI J. and CHAKRAVARTI D. (2012) A Transcriptional Regulatory Role of the THAP11–HCF-1 Complex in Colon Cancer Cell Function, *Molecular and Cellular Biology* **32**(9), 1654–1670.

PATRO R., DUGGAL G., LOVE M. I., IRIZARRY R. A. and KINGSFORD C. (2017) Salmon provides fast and bias-aware quantification of transcript expression, *Nature Methods* **14**(4), 417–419.

PEAR W. S., NOLAN G. P., SCOTT M. L. and BALTIMORE D. (1993) Production of high-titer helper-free retroviruses by transient transfection, *Proceedings of the National Academy of Sciences of the United States of America* **90**(18), 8392–8396.

PELKA P., ABLACK J. N. G., FONSECA G. J., YOUSEF A. F. and MYMRYK J. S. (2008) Intrinsic Structural Disorder in Adenovirus E1A: a Viral Molecular Hub Linking Multiple Diverse Processes, *Journal of Virology* **82**(15), 7252–7263.

PENAGOS-PUIG A. and FURLAN-MAGARIL M. (2020) Heterochromatin as an Important Driver of Genome Organization.

PENNACCHIO L. A., BICKMORE W., DEAN A., NOBREGA M. A. and BEJERANO G. (2013) Enhancers: Five essential questions.

PENNELLA M. A., LIU Y., WOO J. L., KIM C. A. and BERK A. J. (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), 12210–12225.

PÉREZ-BERNÁ A. J., MARION S., CHICHÓN F. J., FERNÁNDEZ J. J., WINKLER D. C., CARRASCOSA J. L., STEVEN A. C., ŠIBER A. and SAN MARTÍN C. (2015) Distribution of DNA-condensing protein complexes in the adenovirus core, *Nucleic Acids Research* **43**(8), 4274–4283.

PÉREZ-ILLANA M., MARTÍNEZ M., CONDEZO G. N., HERNANDO-PÉREZ M., MANGROO C., BROWN M., MARABINI R. and MARTÍN C. S. (2021) Cryo-EM structure of enteric adenovirus HAdV-F41 highlights structural variations among human adenoviruses, *Science Advances* **7**(9).

PERRICAUDET M., AKUSJÄRVI G., VIRTANEN A. and PETTERSSON U. (1979) Structure of two spliced mRNAs from the transforming region of human subgroup C adenoviruses [23].

PERSSON H., KATZE M. G. and PHILIPSON L. (1982) Purification of a native membrane-associated adenovirus tumor antigen, *Journal of Virology* **42**(3), 905–917.

PESONEN S., KANGASNIEMI L. and HEMMINKI A. (2011) Oncolytic Adenoviruses for the Treatment of Human Cancer: Focus on Translational and Clinical Data, *Molecular Pharmaceutics* **8**(1), 12–28.

PFITZNER S., BOSSE J. B., HOFMANN-SIEBER H., FLOMM F., REIMER R., DOBNER T., GRÜNEWALD K. and FRANKEN L. E. (2021) Human adenovirus type 5 infection leads to nuclear envelope destabilization and membrane permeability independently of adenovirus death protein, *International Journal of Molecular Sciences* **22**(23).

PFITZNER S., HOFMANN-SIEBER H., BOSSE J. B., FRANKEN L. E., GRÜNEWALD K. and DOBNER T. (2020) Fluorescent protein tagging of adenoviral proteins pV and pIX reveals 'late virion accumulation compartment', *PLOS Pathogens* **16**(6), e1008588.

PHELPS W. C., YEE C. L., MÜNGER K. and HOWLEY P. M. (1988) The human papillomavirus type 16 E7 gene encodes transactivation and transformation functions similar to those of adenovirus E1A, *Cell* **53**(4), 539–547.

POBBATI A. V., CHAN S. W., LEE I., SONG H. and HONG W. (2012) Structural and functional similarity between the Vgll1-TEAD and the YAP-TEAD complexes, *Structure* **20**(7), 1135–1140.

POBBATI A. V. and HONG W. (2013) Emerging roles of TEAD transcription factors and its coactivators in cancers.

PRIER J. E. (1962) Canine hepatitis virus and human adenovirus., *Public health reports (Washington, D.C. : 1896)* **77**(4), 290–2.

QIAGEN (2017) CLC Main Workbench.

QUERIDO E., BLANCHETTE P., YAN Q., KAMURA T., MORRISON M., BOIVIN D., KAELIN W. G., CONAWAY R. C., CONAWAY J. W. and BRANTON P. E. (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), 3104–3117.

QUINLAN A. R. and HALL I. M. (2010) BEDTools: A flexible suite of utilities for comparing genomic features, *Bioinformatics* **26**(6), 841–842.

R CORE TEAM (2022) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing.

RAMÍREZ F., RYAN D. P., GRÜNING B., BHARDWAJ V., KILPERT F., RICHTER A. S., HEYNE S., DÜNDAR F. and MANKE T. (2016) deepTools2: a next generation web server for deep-sequencing data analysis, *Nucleic Acids Research* **44**(W1), W160–W165.

RASTI M., GRAND R. J. A., MYMRYK J. S., GALLIMORE P. H. and TURNELL A. S. (2005) Recruitment of CBP/p300, TATA-Binding Protein, and S8 to Distinct Regions at the N Terminus of Adenovirus E1A, *Journal of Virology* **79**(9), 5594–5605.

RAUSCHER F. J., VOULALAS P. J., FRANZA B. R. and CURRAN T. (1988) Fos and Jun bind cooperatively to the AP-1 site: reconstitution in vitro., *Genes & development* **2**(12 B), 1687–1699.

REA S., EISENHABER F., O'CARROLL D., STRAHL B. D., SUN Z. W., SCHMID M., OPRAVIL S., MECHTIER K., PONTING C. P., ALLIS C. D. and JENUWEIN T. (2000) Regulation of chromatin structure by site-specific histone H3 methyltransferases, *Nature* **406**(6796), 593–599.

REID J. L., BANNISTER A. J., ZEGERMAN P., MARTÍNEZ-BALBÁS M. A. and KOUZARIDES T. (1998) E1A directly binds and regulates the P/CAF acetyltransferase, *EMBO Journal* **17**(15), 4469–4477.

REIMAND J., ISSERLIN R., VOISIN V., KUCERA M., TANNUS-LOPES C., ROSTAMIANFAR A., WADI L., MEYER M., WONG J., XU C., MERICO D. and BADER G. D. (2019) Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap, *Nature Protocols* **14**(2), 482–517.

REITER F., WIENERROITHER S. and STARK A. (2017) Combinatorial function of transcription factors and cofactors.

RICHMOND T. J., FINCH J. T., RUSHTON B., RHODES D. and KLUG A. (1984) Structure of the nucleosome core particle at 7 resolution, *Nature* **311**(5986), 532–537.

RILEY T., SONTAG E., CHEN P. and LEVINE A. (2008) Transcriptional control of human p53-regulated genes, *Nature Reviews Molecular Cell Biology* **9**(5), 402–412.

ROBINSON J. T., THORVALDSDÓTTIR H., WINCKLER W., GUTTMAN M., LANDER E. S., GETZ G. and MESIROV J. P. (2011) Integrative genomics viewer.

ROCHA P. P., RAVIRAM R., BONNEAU R. and SKOK J. A. (2015) Breaking TADs: Insights into hierarchical genome organization.

RODRIGUEZ M. S., DESTERRO J. M., LAIN S., MIDGLEY C. A., LANE D. P. and HAY R. T. (1999) SUMO-1 modification activates the transcriptional response of p53, *EMBO Journal* **18**(22), 6455–6461.

ROELVINK P. W., LIZONOVA A., LEE J. G. M., LI Y., BERGELSON J. M., FINBERG R. W., BROUGH D. E., KOVESDI I. and WICKHAM T. J. (1998) The Coxsackievirus-Adenovirus Receptor Protein Can Function as a Cellular Attachment Protein for Adenovirus Serotypes from Subgroups A, C, D, E, and F, *Journal of Virology* **72**(10), 7909–7915.

RONCO L. V., KARPOVA A. Y., VIDAL M. and HOWLEY P. M. (1998) Human papillomavirus 16 E6 oncoprotein binds to interferon regulatory factor-3 and inhibits its transcriptional activity, *Genes and Development* **12**(13), 2061–2072.

ROPA J., SAHA N. and MUNTEAN A. G. (2018) SETDB1 Represses Hox Gene Expression and Suppresses Acute Myeloid Leukemia, *Blood* **132**(Supplement 1), 1320–1320.

ROSONINA E., AKHTER A., DOU Y., BABU J. and SRI THEIVAKADADCHAM V. S. (2017) Regulation of transcription factors by sumoylation.

ROSS S., BEST J. L., ZON L. I. and GILL G. (2002) SUMO-1 modification represses Sp3 transcriptional activation and modulates its subnuclear localization, *Molecular Cell* **10**(4), 831–842.

ROWE D. T., GRAHAM F. L. and BRANTON P. E. (1983) Intracellular localization of adenovirus type 5 tumor antigens in productively infected cells, *Virology* **129**(2), 456–468.

ROWE W. P., HUEBNER R. J., GILMORE L. K., PARROTT R. H. and WARD T. G. (1953) Isolation of a cytopathogenic agent from human adenoids undergoing spontaneous degeneration in tissue culture, *Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine* **84**(3), 570–573.

ROY S., CALCEDO R., MEDINA-JASZEK A., KEOUGH M., PENG H. and WILSON J. M. (2011) Adenoviruses in Lymphocytes of the Human Gastro-Intestinal Tract, *PLoS ONE* **6**(9), e24859.

ROY S., VANDENBERGHE L. H., KRYAZHIMSKIY S., GRANT R., CALCEDO R., YUAN X., KEOUGH M., SANDHU A., WANG Q., MEDINA-JASZEK C. A., PLOTKIN J. B. and WILSON J. M. (2009) Isolation and characterization of adenoviruses persistently shed from the gastrointestinal tract of non-human primates, *PLoS Pathogens* **5**(7).

RSTUDIO TEAM (2020) RStudio: Integrated Development Environment for R.

RUFINI A., TUCCI P., CELARDO I. and MELINO G. (2013) Senescence and aging: The critical roles of p53.

S. LEUNG B., C. MANAUGH L. and C. WOOD D. (1973) Estradiol receptors in benign and malignant disease of the breast, *Clinica Chimica Acta* **46**(1), 69–76.

SAFRIN S., CHERRINGTON J. and JAFFE H. S. (1997) Clinical uses of cidofovir, *Reviews in Medical Virology* **7**(3), 145–156.

SAMMONS M. A., ZHU J., DRAKE A. M. and BERGER S. L. (2015) TP53 engagement with the genome occurs in distinct local chromatin environments via pioneer factor activity, *Genome Research* **25**(2), 179–188.

SANTOS-ROSA H., SCHNEIDER R., BANNISTER A. J., SHERRIFF J., BERNSTEIN B. E., EMRE N. C., SCHREIBER S. L., MELLOR J. and KOUZARIDES T. (2002) Active genes are tri-methylated at K4 of histone H3, *Nature* **419**(6905), 407–411.

SARNOW P., HO Y. S., WILLIAMS J. and LEVINE A. J. (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), 387–394.

SATO Y. and TSURUMI T. (2013) Genome guardian p53 and viral infections.

SCHAACK J., BENNETT M. L., COLBERT J. D., TORRES A. V., CLAYTON G. H., ORNELLES D. and MOORHEAD J. (2004) E1A and E1B proteins inflammation by adenovirus, *Proceedings of the National Academy of Sciences of the United States of America* **101**(9), 3124–3129.

SCHECKEL C. and AGUZZI A. (2018) Prions, prionoids and protein misfolding disorders.

SCHIEDNER G., HERTEL S. and KOCHANEK S. (2000) Efficient Transformation of Primary Human Amniocytes by E1 Functions of Ad5: Generation of New Cell Lines for Adenoviral Vector Production, *Human Gene Therapy* **11**(15), 2105–2116.

SCHINDELIN J., ARGANDA-CARRERAS I., FRISE E., KAYNIG V., LONGAIR M., PIETZSCH T., PREIBISCH S., RUEDEN C., SAALFELD S., SCHMID B., TINEVEZ J. Y., WHITE D. J., HARTENSTEIN V., ELICEIRI K., TOMANCAK P. and CARDONA A. (2012) Fiji: An open-source platform for biological-image analysis.

SCHREINER S., BÜRCK C., GLASS M., GROITL P., WIMMER P., KINKLEY S., MUND A., EVERETT R. D. and DOBNER T. (2013) Control of human adenovirus type 5 gene expression by cellular Daxx/ATRX chromatin-associated complexes, *Nucleic Acids Research* **41**(6), 3532–3550.

SCHREINER S., WIMMER P., SIRMA H., EVERETT R. D., BLANCHETTE P., GROITL P. and DOBNER T. (2010) Proteasome-Dependent Degradation of Daxx by the Viral E1B-55K Protein in Human Adenovirus-Infected Cells, *Journal of Virology* **84**(14), 7029–7038.

SCHRÖDINGER L. and DELANO W. (2021) PyMOL.

SCHULTZ D. C., AYYANATHAN K., NEGOREV D., MAUL G. G. and RAUSCHER F. J. (2002) SETDB1: A novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins, *Genes and Development* **16**(8), 919–932.

SCHWARTZ R. A., LAKDAWALA S. S., ESHLEMAN H. D., RUSSELL M. R., CARSON C. T. and WEITZMAN M. D. (2008) Distinct Requirements of Adenovirus E1b55K Protein for Degradation of Cellular Substrates, *Journal of Virology* **82**(18), 9043–9055.

SEO J., KOÇAK D. D., BARTELT L. C., WILLIAMS C. A., BARRERA A., GERSBACH C. A. and REDDY T. E. (2021) AP-1 subunits converge promiscuously at enhancers to potentiate transcription, *Genome Research* **31**(4), 538–550.

SHAH M., ANWAR M. A., PARK S., JAFRI S. S. and CHOI S. (2015) In silico mechanistic analysis of IRF3 inactivation and high-risk HPV E6 species-dependent drug response, *Scientific Reports* **5**.

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), 905–916.

SHECHTER D., DORMANN H. L., ALLIS C. D. and HAKE S. B. (2007) Extraction, purification and analysis of histones, *Nature Protocols* **2**(6), 1445–1457.

SHEN L., SHAO N.-Y., LIU X., MAZE I., FENG J. and NESTLER E. J. (2013) diffReps: Detecting Differential Chromatin Modification Sites from ChIP-seq Data with Biological Replicates, *PLoS ONE* **8**(6), e65598.

SHIEH S. Y., IKEDA M., TAYA Y. and PRIVES C. (1997) DNA damage-induced phosphorylation of p53 alleviates inhibition by MDM2, *Cell* **91**(3), 325–334.

SHIEH W. J. (2022) Human adenovirus infections in pediatric population - An update on clinico–pathologic correlation.

SHIH D. S., CARRUTH L. M., ANDERSON C. and CLEMENTS J. E. (1992) Involvement of FOS and JUN in the activation of visna virus gene expression in macrophages through an AP-1 site in the viral LTR, *Virology* **190**(1), 84–91.

SHILATIFARD A. (2012) The COMPASS family of histone H3K4 methylases: Mechanisms of regulation in development and disease pathogenesis, *Annual Review of Biochemistry* **81**, 65–95.

SHU L., CHEN A., LI L., YAO L., HE Y., XU J., GU W., LI Q., WANG K., ZHANG T. and LIU G. (2022) NRG1 regulates Fra-1 transcription and metastasis of triple-negative breast cancer cells via the c-Myc ubiquitination as manipulated by ERK1/2-mediated Fbxw7 phosphorylation, *Oncogene* **41**(6), 907–919.

SIEBER T. and DOBNER T. (2007) Adenovirus Type 5 Early Region 1B 156R Protein Promotes Cell Transformation Independently of Repression of p53-Stimulated Transcription, *Journal of Virology* **81**(1), 95–105.

SINGH A. A., SCHUURMAN K., NEVEDOMSKAYA E., STELLOO S., LINDER S., DROOG M., KIM Y., SANDERS J., VAN DER POEL H., BERGMAN A. M., WESSELS L. F. and ZWART W. (2019) Optimized ChIP-seq method facilitates transcription factor profiling in human tumors, *Life Science Alliance* **2**(1).

SINGH A. K., BERBIS M. A., BALLMANN M. Z., KILCOYNE M., MENENDEZ M., NGUYEN T. H., JOSHI L., CANADA F. J., JIMENEZ-BARBERO J., BENKO M., HARRACH B. and VAN RAAIJ M. J. (2015) Structure and Sialyllactose Binding of the Carboxy-Terminal Head Domain of the Fibre from a Siadenovirus, Turkey Adenovirus 3, *PLOS ONE* **10**(9), e0139339.

SINGH A. K., MENENDEZ-CONEJERO R., SAN MARTIN C. and VAN RAAIJ M. J. (2014) Crystal Structure of the Fibre Head Domain of the Atadenovirus Snake Adenovirus 1, *PLoS ONE* **9**(12), e114373.

SLOWIKOWSKI K. (2021) ggrepel: Automatically Position Non-Overlapping Text Labels with 'ggplot2.

SMALE S. T. and KADONAGA J. T. (2003) The RNA polymerase II core promoter.

SONESON C., LOVE M. I. and ROBINSON M. D. (2016) Differential analyses for RNA-seq: Transcript-level estimates improve gene-level inferences, *F1000Research* **4**.

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), 395–414.

SPEISEDER T., HOFMANN-SIEBER H., RODRÍGUEZ E., SCHELLENBERG A., AKYÜZ N., DIERLAMM J., SPRUSS T., LANGE C. and DOBNER T. (2017) Efficient Transformation of Primary Human Mesenchymal Stromal Cells by Adenovirus Early Region 1 Oncogenes, *Journal of Virology* **91**(1), 01782–16.

STEIN G. S., ZAIDI S. K., STEIN J. L., LIAN J. B., VAN WIJNEN A. J., MONTECINO M., YOUNG D. W., JAVED A., PRATAP J., CHOI J. Y., ALI S. A., PANDE S. and HASSAN M. Q. (2009) Transcription-factor-mediated epigenetic control of cell fate and lineage commitment.

STEINEGGER M. and SÖDING J. (2017) MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets.

STEPHENS M., CARBONETTO P., GERARD D., LU M., SUN L., WILLWERSCHEID J. and XIAO N. (2022) ashr: Methods for Adaptive Shrinkage, using Empirical Bayes.

STERN M., JENSEN R. and HERSKOWITZ I. (1984) Five SWI genes are required for expression of the HO gene in yeast, *Journal of Molecular Biology* **178**(4), 853–868.

STERN-GINOSSAR N., THOMPSON S. R., MATHEWS M. B. and MOHR I. (2019) Translational control in virus-infected cells, *Cold Spring Harbor Perspectives in Biology* **11**(3), a033001.

STEWART P. L. and NEMEROW G. R. (2007) Cell integrins: commonly used receptors for diverse viral pathogens.

STOCK J. K., GIADROSSI S., CASANOVA M., BROOKES E., VIDAL M., KOSEKI H., BROCKDORFF N., FISHER A. G. and POMBO A. (2007) Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells, *Nature Cell Biology* **9**(12), 1428–1435.

STOVNER E. B. and SÆTROM P. (2019) Epic2 efficiently finds diffuse domains in ChIP-seq data, *Bioinformatics* **35**(21), 4392–4393.

STRACKER T. H., CARSON C. T. and WEILZMAN M. D. (2002) Adenovirus oncoproteins inactivate the Mre11-Rad50-NBs1 DNA repair complex, *Nature* **418**(6895), 348–352.

STRAHL B. D. and ALLIS C. D. (2000) The language of covalent histone modifications.

STRUHL K. and SEGAL E. (2013) Determinants of nucleosome positioning.

STRUNZE S., TROTMAN L. C., BOUCKE K. and GREBER U. F. (2005) Nuclear targeting of adenovirus type 2 requires CRM1-mediated nuclear export, *Molecular Biology of the Cell* **16**(6), 2999–3009.

SU C. (2011) Adenovirus-Based Gene Therapy for Cancer, in *Viral Gene Therapy*, InTech.

SUNDQVIST A., VASILAKI E., VOYTYUK O., BAI Y., MORIKAWA M., MOUSTAKAS A., MIYAZONO K., HELDIN C. H., TEN DIJKE P. and VAN DAM H. (2020) TGFβ and EGF signaling orchestrates the AP-1- and p63 transcriptional regulation of breast cancer invasiveness, *Oncogene* **39**(22), 4436–4449.

SUZEK B. E., WANG Y., HUANG H., MCGARVEY P. B. and WU C. H. (2015) UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches, *Bioinformatics* **31**(6), 926–932.

TABAKIN-FIX Y., AZRAN I., SCHAVINKY-KHRAPUNSKY Y., LEVY O. and ABOUD M. (2006) Functional inactivation of p53 by human T-cell leukemia virus type 1 Tax protein: mechanisms and clinical implications, *Carcinogenesis* **27**(4), 673–681.

TAMKUN J. W., DEURING R., SCOTT M. P., KISSINGER M., PATTATUCCI A. M., KAUFMAN T. C. and KENNISON J. A. (1992) brahma: A regulator of Drosophila homeotic genes structurally related to the yeast transcriptional activator SNF2 SWI2, *Cell* **68**(3), 561–572.

TANAKA K., ISSELBACHER K. J., KHOURY G. and JAY G. (1985) Reversal of oncogenesis by the expression of a major histocompatibility complex class I gene, *Science* **228**(4695), 26–30.

TANG Z., CHEN W. Y., SHIMADA M., NGUYEN U. T., KIM J., SUN X. J., SENGOKU T., MCGINTY R. K., FERNANDEZ J. P., MUIR T. W. and ROEDER R. G. (2013) XSET1 and p300 act synergistically, through coupled histone modifications, in transcriptional activation by p53, *Cell* **154**(2), 297.

TATSIS N. and ERTL H. C. (2004) Adenoviruses as vaccine vectors.

TAYLOR B. C. and YOUNG N. L. (2021) Combinations of histone post-Translational modifications.

TCHASOVNIKAROVA I. A., TIMMS R. T., MATHESON N. J., WALS K., ANTROBUS R., GÖTTGENS B., DOUGAN G., DAWSON M. A. and LEHNER P. J. (2015) Epigenetic silencing by the HUSH complex mediates position-effect variegation in human cells, *Science* **348**(6242), 1481–1485.

TEJERA B., LOPEZ R. E., HIDALGO P., CARDENAS R., BALLESTEROS G., RIVILLAS L., FRENCH L., AMERO C., PASTOR N., SANTIAGO A., GROITL P., DOBNER T. and GONZALEZ R. A. (2019) The human adenovirus type 5 E1B 55kDa protein interacts with RNA promoting timely DNA replication and viral late mRNA metabolism, *PLoS ONE* **14**(4).

TENTHOREY J. L., EMERMAN M. and MALIK H. S. (2022) Evolutionary Landscapes of Host-Virus Arms Races.

TEODORO J. G., HALLIDAY T., WHALEN S. G., TAKAYESU D., GRAHAM F. L. and BRANTON P. E. (1994) Phosphorylation at the carboxy terminus of the 55-kilodalton adenovirus type 5 E1B protein regulates transforming activity, *Journal of Virology* **68**(2), 776–786.

TESSIER T. M., DODGE M. J., MACNEIL K. M., EVANS A. M., PRUSINKIEWICZ M. A. and MYMRYK J. S. (2021) Almost famous: Human adenoviruses (and what they have taught us about cancer).

TIAN B., YANG J. and BRASIER A. R. (2012) Two-step cross-linking for analysis of protein-chromatin interactions, *Methods in Molecular Biology* **809**, 105–120.

TIE F., BANERJEE R., STRATTON C. A., PRASAD-SINHA J., STEPANIK V., ZLOBIN A., DIAZ M. O., SCACHERI P. C. and HARTE P. J. (2009) CBP-mediated acetylation of histone H3 lysine 27 antagonizes Drosophila Polycomb silencing, *Development* **136**(18), 3131–3141.

TILLO D., RAY S., SYED K. S., GAYLOR M. R., HE X., WANG J., ASSAD N., DURELL S. R., POROLLO A., WEIRAUCH M. T. and VINSON C. (2017) The Epstein-Barr Virus B-ZIP Protein Zta Recognizes Specific DNA Sequences Containing 5-Methylcytosine and 5-Hydroxymethylcytosine, *Biochemistry* **56**(47), 6200–6210.

TOLLEFSON A. E., RYERSE J. S., SCARIA A., HERMISTON T. W. and WOLD W. S. (1996a) The E3-11.6-kDa adenovirus death protein (ADP) is required for efficient cell death: Characterization of cells infected with adp mutants, *Virology* **220**(1), 152–162.

TOLLEFSON A. E., SCARIA A., HERMISTON T. W., RYERSE J. S., WOLD L. J. and WOLD W. S. (1996b) 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), 2296–2306.

TOMSO D. J., INGA A., MENENDEZ D., PITTMAN G. S., CAMPBELL M. R., STORICI F., BELL D. A. and RESNICK M. A. (2005) Functionally distinct polymorphic sequences in the human genome that are targets for p53 transactivation, *Proceedings of the National Academy of Sciences of the United States of America* **102**(18), 6431–6436.

TREMBLAY M. L., MCGLADE C. J., GERBER G. E. and BRANTON P. E. (1988) Identification of the phosphorylation sites in early region 1A proteins of adenovirus type 5 by amino acid sequencing of peptide fragments, *Journal of Biological Chemistry* **263**(13), 6375–6383.

TRENTIN J. J., YABE Y. and TAYLOR G. (1962) The quest for human cancer viruses, *Science* **137**(3533), 835–841.

TSANKOV A. M., GU H., AKOPIAN V., ZILLER M. J., DONAGHEY J., AMIT I., GNIRKE A. and MEISSNER A. (2015) Transcription factor binding dynamics during human ES cell differentiation, *Nature* **518**(7539), 344–349.

TSUJIMOTO Y. (1989) Overexpression of the human BCL-2 gene product results in growth enhancement of Epstein-Barr virus-immortalized B cells, *Proceedings of the National Academy of Sciences of the United States of America* **86**(6), 1958–1962.

TUNNICLIFFE R. B., LOCKHART-CAIRNS M. P., LEVY C., MOULD A. P., JOWITT T. A., SITO H., BALDOCK C., SANDRI-GOLDIN R. M. and GOLOVANOV A. P. (2017) The herpes viral transcription factor ICP4 forms a novel DNA recognition complex, *Nucleic Acids Research* **45**(13), 8064–8078.

TURNELL A. S. and MYMRYK J. S. (2006) Roles for the coactivators CBP and p300 and the APC/C E3 ubiquitin ligase in E1A-dependent cell transformation.

UEHARA I. and TANAKA N. (2018) Role of p53 in the Regulation of the Inflammatory Tumor Microenvironment and Tumor Suppression, *Cancers* **10**(7), 219.

VAN DEN HEUVEL S. J., THE S. I., KLEIN B., JOCHEMSEN A. G., ZANTEMA A. and VAN DER EB A. J. (1992) p53 shares an antigenic determinant with proteins of 92 and 150 kilodaltons that may be involved in senescence of human cells, *Journal of Virology* **66**(1), 591–595.

VAN NGUYEN T., ANGKASEKWINAI P., DOU H., LIN F. M., LU L. S., CHENG J., CHIN Y. E., DONG C. and YEH E. T. (2012) SUMO-Specific Protease 1 Is Critical for Early Lymphoid Development through Regulation of STAT5 Activation, *Molecular Cell* **45**(2), 210–221.

VAN OOSTRUM J., SMITH P. R., MOHRAZ M. and BURNETT R. M. (1987) The structure of the adenovirus capsid. III. Hexon packing determined from electron micrographs of capsid fragments, *Journal of Molecular Biology* **198**(1), 73–89.

VAUDIN P., DELANOUE R., DAVIDSON I., SILBER J. and ZIDER A. (1999) TONDU (TDU), a novel human protein related to the product of vestigial (vg) gene of Drosophila melanogaster interacts with vertebrate TEF factors and substitutes for Vg function in wing formation, *Development* **126**(21), 4807–4816.

VAUX D. L., CORY S. and ADAMS J. M. (1988) Bcl-2 gene promotes haemopoietic cell survival and cooperates with c-myc to immortalize pre-B cells.

VAYDA M. E., ROGERS A. E. and FLINT S. J. (1983) The structure of nucleoprotein cores released from adenovirions, *Nucleic Acids Research* **11**(2), 441–460.

VEESLER D., CUPELLI K., BURGER M., GRABER P., STEHLE T. and JOHNSON J. E. (2014) Single-particle EM reveals plasticity of interactions between the adenovirus penton base and integrin αVβ3, *Proceedings of the National Academy of Sciences of the United States of America* **111**(24), 8815–8819.

VELLINGA J., VAN DEN WOLLENBERG D. J. M., VAN DER HEIJDT S., RABELINK M. J. W. E. and HOEBEN R. C. (2005) The Coiled-Coil Domain of the Adenovirus Type 5 Protein IX Is Dispensable for Capsid Incorporation and Thermostability, *Journal of Virology* **79**(5), 3206–3210.

VERFAILLIE A., SVETLICHNYY D., IMRICHOVA H., DAVIE K., FIERS M., ATAK Z. K., HULSELMANS G., CHRISTIAENS V. and AERTS S. (2016) Multiplex enhancer-reporter assays uncover unsophisticated TP53 enhancer logic, *Genome Research* **26**(7), 882–895.

VERRECCHIA F., VINDEVOGHEL L., LECHLEIDER R. J., UITTO J., ROBERTS A. B. and MAUVIEL A. (2001) Smad3/AP-1 interactions control transcriptional responses to TGF-β in a promoter-specific manner, *Oncogene* **20**(26), 3332–3340.

VIEUX-ROCHAS M., FABRE P. J., LELEU M., DUBOULE D. and NOORDERMEER D. (2015) Clustering of mammalian Hox genes with other H3K27me3 targets within an active nuclear domain, *Proceedings of the National Academy of Sciences of the United States of America* **112**(15), 4672–4677.

VINCKEVICIUS A., PARKER J. B. and CHAKRAVARTI D. (2015) Genomic Determinants of THAP11/ZNF143/HCFC1 Complex Recruitment to Chromatin, *Molecular and Cellular Biology* **35**(24), 4135–4146.

WADELL G. (1984) Molecular epidemiology of human adenoviruses., *Current topics in microbiology and immunology* **110**, 191–220.

WALKER P. J., SIDDELL S. G., LEFKOWITZ E. J., MUSHEGIAN A. R., ADRIAENSSENS E. M., DEMPSEY D. M., DUTILH B. E., HARRACH B., HARRISON R. L., HENDRICKSON R. C., JUNGLEN S., KNOWLES N. J., KROPINSKI A. M., KRUPOVIC M., KUHN J. H., NIBERT M., ORTON R. J., RUBINO L., SABANADZOVIC S., SIMMONDS P., SMITH D. B., VARSANI A., ZERBINI F. M. and DAVISON A. J. (2020) Changes to virus taxonomy and the Statutes ratified by the International Committee on Taxonomy of Viruses (2020), *Archives of Virology* **165**(11), 2737–2748.

WANG Y., SCHWEDES J. F., PARKS D., MANN K. and TEGTMEYER P. (1995) Interaction of p53 with its consensus DNA-binding site, *Molecular and Cellular Biology* **15**(4), 2157–2165.

WANG Z., QIAO Y., CHEN Z., LIANG Y., CUI L., ZHANG Y., LI X., XU L., WEI P., LIU S. and LI H. (2021) Fos facilitates gallid alpha-herpesvirus 1 infection by transcriptional control of host metabolic genes and viral immediate early gene, *Viruses* **13**(6).

WASSERMAN W. W. and SANDELIN A. (2004) Applied bioinformatics for the identification of regulatory elements.

WEBER K., BARTSCH U., STOCKING C. and FEHSE B. (2008) A multicolor panel of novel lentiviral "gene ontology" (LeGO) vectors for functional gene analysis, *Molecular Therapy* **16**(4), 698–706.

WEI K. H., CHAN C. and BACHTROG D. (2021) Establishment of h3k9me3-dependent heterochromatin during embryogenesis in drosophila miranda, *eLife* **10**.

WEIRAUCH M. T. and HUGHES T. R. (2014) A catalogue of eukaryotic transcription factor types, their evolutionary origin, and species distribution, *Sub-Cellular Biochemistry* **52**, 25–73.

WELLEN K. E., HATZIVASSILIOU G., SACHDEVA U. M., BUI T. V., CROSS J. R. and THOMPSON C. B. (2009) ATP-citrate lyase links cellular metabolism to histone acetylation, *Science* **324**(5930), 1076–1080.

WHITE E., SABBATINI P., DEBBAS M., WOLD W. S., KUSHER D. I. and GOODING L. R. (1992) The 19-kilodalton adenovirus E1B transforming protein inhibits programmed cell death and prevents cytolysis by tumor necrosis factor alpha, *Molecular and Cellular Biology* **12**(6), 2570–2580.

WHYTE P., BUCHKOVICH K. J., HOROWITZ J. M., FRIEND S. H., RAYBUCK M., WEINBERG R. A. and HARLOW E. (1988) Association between an oncogene and an anti-oncogene: the adenovirus E1A proteins bind to the retinoblastoma gene product, *Nature* **334**(6178), 124–129.

WICKER T., SABOT F., HUA-VAN A., BENNETZEN J. L., CAPY P., CHALHOUB B., FLAVELL A., LEROY P., MORGANTE M., PANAUD O., PAUX E., SANMIGUEL P. and SCHULMAN A. H. (2007) A unified classification system for eukaryotic transposable elements.

WICKHAM H. (2022) forcats: Tools for Working with Categorical Variables (Factors).

WICKHAM H., FRANCOIS ROMAIN, HENRY L. and MÜLLER K. (2022) dplyr: A Grammar of Data Manipulation.

WICKHAM H. and GIRLICH MAXIMILIAN (2022) tidyr: Tidy Messy Data.

WICKHAM H. and SEIDEL D. (2022) scales: Scale Functions for Visualization.

WICKHAM T. J., MATHIAS P., CHERESH D. A. and NEMEROW G. R. (1993) Integrins αvβ3 and αvβ5 promote adenovirus internalization but not virus attachment, *Cell* **73**(2), 309–319.

WILSON M. C. and DARNELL J. E. (1981) Control of messenger RNA concentration by differential cytoplasmic half-life. Adenovirus messenger RNAs from transcription units 1A and 1B, *Journal of Molecular Biology* **148**(3), 231–251.

WIMMER P., BLANCHETTE P., SCHREINER S., CHING W., GROITL P., BERSCHEMINSKI J., BRANTON P. E., WILL H. and DOBNER T. (2013) Cross-talk between phosphorylation and SUMOylation regulates transforming activities of an adenoviral oncoprotein, *Oncogene* **32**(13), 1626–1637.

WRIGHT N. G., CORNWELL H. J., THOMPSON H., ARMITAGE A. and MORRISON I. (1972) Canine adenovirus respiratory disease: isolation of infectious canine hepatitis virus from natural cases and the experimental production of the disease., *The Veterinary record* **90**(15), 411–416.

WU D. Y., KALPANA G. V., GOFF S. P. and SCHUBACH W. H. (1996) Epstein-Barr virus nuclear protein 2 (EBNA2) binds to a component of the human SNF-SWI complex, hSNF5/Ini1, *Journal of Virology* **70**(9), 6020–6028.

WU Z., NICOLL M. and INGHAM R. J. (2021) AP-1 family transcription factors: a diverse family of proteins that regulate varied cellular activities in classical hodgkin lymphoma and ALK+ ALCL.

YAMAGUCHI N. (2020) Multiple Roles of Vestigial-Like Family Members in Tumor Development., *Frontiers in oncology* **10**, 1266.

YAMAMOTO K. and SONODA M. (2003) Self-interaction of heterochromatin protein 1 is required for direct binding to histone methyltransferase, SUV39H1, *Biochemical and Biophysical Research Communications* **301**(2), 287–292.

YAMAMURA Y., HUA X., BERGELSON S. and LODISH H. F. (2000) Critical role of Smads and AP-1 complex in transforming growth factor-β-dependent apoptosis, *Journal of Biological Chemistry* **275**(46), 36295–36302.

YANG S. H., JAFFRAY E., HAY R. T. and SHARROCKS A. D. (2003) Dynamic interplay of the SUMO and ERK pathways in regulating Elk-1 transcriptional activity, *Molecular Cell* **12**(1), 63–74.

YANG S. H. and SHARROCKS A. D. (2004) SUMO promotes HDAC-mediated transcriptional repression, *Molecular Cell* **13**(4), 611–617.

YANG X. J., OGRYZKO V. V., NISHIKAWA J. I., HOWARD B. H. and NAKATANI Y. (1996) A p300/CPB-associated factor that competes with the adenoviral oncoprotein E1A, *Nature* **382**(6589), 319–324.

YEW P. R., LIU X. and BERK A. J. (1994) Adenovirus E1B oncoprotein tethers a transcriptional repression domain to p53, *Genes and Development* **8**(2), 190–202.

YOU J. S., KELLY T. K., DE CARVALHO D. D., TABERLAY P. C., LIANG G. and JONES P. A. (2011) OCT4 establishes and maintains nucleosome-depleted regions that provide additional layers of epigenetic regulation of its target genes, *Proceedings of the National Academy of Sciences of the United States of America* **108**(35), 14497–14502.

YOUNGER S. T. and RINN J. L. (2017) P53 regulates enhancer accessibility and activity in response to DNA damage, *Nucleic Acids Research* **45**(17), 9889–9900.

YU G. and HE Q. Y. (2016) ReactomePA: An R/Bioconductor package for reactome pathway analysis and visualization, *Molecular BioSystems* **12**(2), 477–479.

YU G., WANG L.-G. and HE Q.-Y. (2015) ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization, *Bioinformatics* **31**(14), 2382–2383.

YU Y., OUYANG Y. and YAO W. (2018) shinyCircos: an R/Shiny application for interactive creation of Circos plot, *Bioinformatics* **34**(7), 1229–1231.

ZANCONATO F., FORCATO M., BATTILANA G., AZZOLIN L., QUARANTA E., BODEGA B., ROSATO A., BICCIATO S., CORDENONSI M. and PICCOLO S. (2015) Genome-wide association between YAP/TAZ/TEAD and AP-1 at enhancers drives oncogenic growth, *Nature Cell Biology* **17**(9), 1218–1227.

ZANTEMA A., FRANSEN J. A., DAVIS-OLIVIER A., RAMAEKERS F. C., VOOIJS G. P., DELEYS B. and VAN DER EB A. J. (1985) Localization of the E1 B proteins of adenovirus 5 in transformed cells, as revealed by interaction with monoclonal antibodies, *Virology* **142**(1), 44–58.

ZEMKE N. R., GOU D. and BERK A. J. (2019) Dedifferentiation by adenovirus E1A due to inactivation of hippo pathway effectors YAP and TAZ, *Genes and Development* **33**(13-14), 828–843.

ZHANG Y., LIU T., MEYER C. A., EECKHOUTE J., JOHNSON D. S., BERNSTEIN B. E., NUSSBAUM C., MYERS R. M., BROWN M., LI W. and SHIRLEY X. S. (2008) Model-based analysis of ChIP-Seq (MACS), *Genome Biology* **9**(9), R137.

ZHANG Y., XIONG Y. and YARBROUGH W. G. (1998) ARF promotes MDM2 degradation and stabilizes p53: ARF-INK4a locus deletion impairs both the Rb and p53 tumor suppression pathways, *Cell* **92**(6), 725–734.

ZHAO B., LI L., TUMANENG K., WANG C. Y. and GUAN K. L. (2010) A coordinated phosphorylation by Lats and CK1 regulates YAP stability through SCFβ-TRCP, *Genes and Development* **24**(1), 72–85.

ZHAO B., WEI X., LI W., UDAN R. S., YANG Q., KIM J., XIE J., IKENOUE T., YU J., LI L., ZHENG P., YE K., CHINNAIYAN A., HALDER G., LAI Z. C. and GUAN K. L. (2007a) Inactivation of YAP oncoprotein by the Hippo pathway is involved in cell contact inhibition and tissue growth control, *Genes and Development* **21**(21), 2747–2761.

ZHAO L. J., LOEWENSTEIN P. M. and GREEN M. (2017) Adenovirus E1A TRRAP-targeting domain-mediated enhancement of MYC association with the NuA4 complex activates a panel of MYC target genes enriched for gene expression and ribosome biogenesis, *Virology* **512**, 172–179.

ZHAO L. Y., SANTIAGO A., LIU J. and LIAO D. (2007b) Repression of p53-mediated transcription by adenovirus E1B 55-kDa does not require corepressor mSin3A and histone deacetylases, *Journal of Biological Chemistry* **282**(10), 7001–7010.

ZHAO X. (2018) SUMO-Mediated Regulation of Nuclear Functions and Signaling Processes.

ZHENG S. J., LAMHAMEDI-CHERRADI S. E., WANG P., XU L. and CHEN Y. H. (2005) Tumor suppressor p53 inhibits autoimmune inflammation and macrophage function, *Diabetes* **54**(5), 1423–1428.

ZHOU W., ZHU P., WANG J., PASCUAL G., OHGI K. A., LOZACH J., GLASS C. K. and ROSENFELD M. G. (2008) Histone H2A Monoubiquitination Represses Transcription by Inhibiting RNA Polymerase II Transcriptional Elongation, *Molecular Cell* **29**(1), 69–80.

ZHOU Y., ZHOU B., PACHE L., CHANG M., KHODABAKHSHI A. H., TANASEICHUK O., BENNER C. and CHANDA S. K. (2019) Metascape provides a biologist-oriented resource for the analysis of systems-level datasets, *Nature Communications* **10**(1).

ZHU A., IBRAHIM J. G. and LOVE M. I. (2019) Heavy-Tailed prior distributions for sequence count data: Removing the noise and preserving large differences, *Bioinformatics* **35**(12), 2084–2092.

ZUBIETA C., SCHOEHN G., CHROBOCZEK J. and CUSACK S. (2005) The structure of the human adenovirus 2 penton, *Molecular Cell* **17**(1), 121–135.

ZYLICZ J. J., ZYLICZ J. J. and HEARD E. (2020) Molecular Mechanisms of Facultative Heterochromatin Formation: An X-Chromosome Perspective.

List of Figures

Acknowledgements

I would like to express my gratitude to Prof. Thomas Dobner and Prof. Nicole Fischer for giving me the opportunity to write my doctoral thesis under their supervision. My deepest thanks and appreciation goes to the various Postdocs in our laboratory, especially Wing Hang Ip and Luca Bertzbach, who guided my work and tremendously assisted in the compilation of my thesis. Thank you Wingi, for all your time! I also would like to thank the examination committee of my disputation: Prof. Julia Kehr and Dr. Gabrielle Vieyres for the evaluation of my thesis defense. In addition, I want to thank Dr. Thomas Günther for an uncountable number of hours of discussion, interpretation and evaluation of my data, for various stupid questions from my side and generally for being the person-in-need whenever I required aid during my work. I want to extend this heart-felt gratitude to Christina Herrde, who helped me in my various ChIP-seq-related quests, to Arne Düsedau and Jana Hennesen for their help in FACS and to Simon Weißmann and Alexis Robitaille for the occasional in-between question. I want to thank all the members of the viral transformation department for the lively discussions, their advices and assistance, for all the extended coffee breaks and for the rare occasional beer right before the weekend. Thank you, Britta, Tina, Gabi, Paloma and Edda - and all former lab members - for your help and comradery! A special mention goes to Laura Seddar, who has to suffer through my continued presence every day! I have seen many people come and go in my time here, and I am happy to get to know each and everyone of you guys! I want to thank all the Bachelor- and Master students that were involved in my work in many ways: Eileen Dudda, Max Hüpper, Leah Kobza, Jennifer Rothe, Sophie Weinert and Alexandra Reinbrecht.

Mein größter Dank geht an meine Familie, die mich in jedem Schritt meines Lebens unterstützt und begleitet hat. Lieber Vater, liebe Mutter, ohne eure tagtägliche Hingabe und Mühe wäre ich nicht der Mensch geworden, der ich heute bin. Lieber Bruder, ich weiß dass ich immer auf dich zählen kann, du bist eine große Stütze in meinem Leben. Ich verdanke euch alles in meinem Leben, und widme euch aus diesem Grund diese Arbeit. Mein Dank geht auch an meine Tanten und Onkel, Cousinen und Cousins. Ihr alle habt meinem Leben einen Rahmen gegeben und mich zu einem großen Teil geprägt, dafür

danke ich euch von tiefstem Herzen. Danke an alle meine alten und neuen Freunde, die mir immer zur Seite gestanden und mich auf meinem Weg begleitet haben: Simon, Marco, Jesper, Matteo, Jule, Philipp, Michael, Steewen, Timo, Johannes, Søren und all die, die ich hier nicht aufzählen konnte, die Liste wäre viel zu groß.

LIV • Martinistraße 52 (N63) • 20251 Hamburg

Quantitative Virology

Dr. Timothy Soh, Ph.D.
Post-doctoral Fellow

Phone: +49 (0) 40 8998 87646
timothy.soh@leibniz-liv.de

Hamburg, 14.11.2022

The Ph.D. thesis from Konstantin von Stromberg entitled "**Adenovirus large E1B proteins regulate transcription through interaction with mammalian transcription factors**" is written in fluent English. I confirm that the language is clearly written and properly articulated.

Kind regards,

Dr. Timothy Soh, Ph.D.
Post-doctoral Fellow

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 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.

Date and Place	Signature