From the Department of Oncology-Pathology
Karolinska Institutet, Stockholm, Sweden

BEYOND TRANSCRIPTION:

A POST-TRANSCRIPTIONAL ROLE OF 3D CHROMATIN CROSSTALK IN ONCOGENE REGULATION

Johanna Vestlund

Karolinska Institutet

Stockholm 2022

Published by Karolinska Institutet.
Printed by Universitetsservice US-AB, 2022

ISBN 978-91-8016-760-4
Cover illustration: Chromatin interactions impinging on the OSE generated by Nodewalk experiments, plot made by Dr Deeksha Bhartiya.

Beyond transcription: A post-transcriptional role of 3D chromatin crosstalk in oncogene regulation
THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Johanna Vestlund

The thesis will be defended in public at Björvellrummet NKS D4:52, Karolinska Universitetssjukhuset Solna, Stockholm, on October 3rd 2022 at 09:00 AM.

Principal Supervisor:
Anita Göndör, MD, PhD, Associate Professor
Karolinska Institutet
Department of Oncology-Pathology

Co-supervisor(s):
Professor Camilla Björkegren
Karolinska Institutet
Department of Cell and Molecular Biology
Department of Biosciences and Nutrition

Professor Lars Holmgren
Karolinska Institutet
Department of Oncology-Pathology

Opponent:
Associate Professor Francisca Lottersberger
Linköping University
Department of Biomedical and Clinical Sciences

Examination Board:
Professor Ann-Kristin Östlund Farrants
Stockholm University
Department of Molecular Biosciences

Docent Andreas Lennartsson
Karolinska Institutet
Department of Biosciences and Nutrition

Docent Padraig D'Arcy
Linköping University
Department of Biomedical and Clinical Sciences

Till min familj, som alltid har stöttat mig.

Perheelleni, joka on aina minua tukenut.

POPULAR SCIENCE SUMMARY OF THE THESIS

English

DNA contains the genetic information that is passed on from parents to their offspring, and is located in the cell nucleus. DNA molecules are long, and one human cell contains 2 m of DNA, while the DNA molecules of all the cells in the human body would be able to reach twice across the solar system if laid out in a straight line. These large amounts of DNA are compactly packed and organized in a highly regulated manner inside the cell nucleus. The compacted DNA is called chromatin, and the 3D organization of chromatin inside the nucleus is able to regulate the activity (or "expression") of genes. This regulation of gene expression is important for cells to be able to activate the right genes at the right time. The gene regulation entails cell-specific gene expression patterns, which allow the first single cell ("zygote"), formed by a sperm cell and an egg cell in humans, to develop into hundreds of different kinds of cells, such as liver cells or brain cells. If the cells are not able to properly regulate the activity of certain genes, diseases such as cancer might develop. Genes that may cause cancer when mutated or over-expressed are called "oncogenes".

In this thesis we have invented a novel assay enabling us to uncover the cell-to-cell dynamics of chromatin interactions between the *MYC* oncogene and its enhancers (regulatory elements of DNA which may regulate gene expression) at high resolution. Moreover, we have discovered a novel mechanism underlying oncogene regulation within the 3D nucleus of the cancer genome. We thus identified the molecular determinants of the gene gating process of *MYC*; where the *MYC* gene is physically recruited by a colorectal cancer super-enhancer to the nuclear pores at the periphery of the nucleus. Proximity to nuclear pores thus facilitated the export of *MYC* mRNA molecules (the product of gene activity) to the cytoplasm to thereby increase *MYC* expression that provides the cancer cells with a growth advantage compared to normal colon cells. Our findings thus increase the understanding of how cancer cells may gain advantages over normal cells during disease development, and open up new avenues for future diagnostic or therapeutic approaches.

Svenska

DNA finns i cellkärnan och innehåller den genetiska information som förs vidare från förälder till barn. DNA-molekyler är långa, och i en mänsklig cell finns det 2 m DNA, medan DNA-molekylerna från alla kroppens celler skulle räcka fram och tillbaka över hela solsystemet om deras längd lades ihop i en rak linje. Dessa stora mängder av DNA är kompakt packade och organiserade på ett strikt reglerat vis i cellkärnan. Det kompakterade DNAt kallas för kromatin, och 3D organiseringen av kromatin inuti cellkärnan kan reglera olika geners aktivitet (”genuttryck”). Denna reglering av genuttryck är viktig för cellers förmåga att aktivera rätt gener vid rätt tillfälle. Regleringen av gener medför cell-specifika genuttrycksprogram, som tillåter den allra första cellen (”zygoten”), som bildas av en spermie

och en äggcell hos människor, att utvecklas till flera hundra olika typer av celler, t.ex. leverceller eller hjärnceller. Om cellerna inte kan reglera aktiviteten av särskilda gener så kan detta leda till att sjukdomar såsom cancer utvecklas. Gener som har förmågan att orsaka cancer då de är muterade eller överuttryckta i tumörvävnader kallas för ”onkogener”.

I denna avhandling har vi utvecklat en ny teknik som möjliggör upptäckandet av cell-till-cell-dynamik hos kromatininteraktioner mellan *MYC* onkogenen och dess *enhancers* (regulatoriska DNA-element som kan reglera genuttryck). Dessutom har vi upptäckt en ny mekanism som ligger bakom regleringen av onkogener i 3D utrymmet i cellkärnans cancergenom. Vi har sålunda identifierat de molekylära determinanterna för *gene gating* processen av *MYC*; där *MYC*-genen fysiskt rekryteras av en kolorektalcancer *super-enhancer* till kärnporerna i cellkärnans periferi. Närheten till kärnporer faciliterade således exporten av *MYC* mRNA-molekyler (produkten av genaktivitet) till cytoplasman, för att därigenom öka uttrycket av *MYC* som ger cancercellerna en tillväxtfördel jämfört med normala kolonceller. Våra fynd ökar därmed förståelsen för hur cancerceller skaffa sig fördelar jämfört med normala celler under sjukdomsutvecklingen av cancer, och öppnar nya möjligheter för framtida diagnostiska eller terapeutiska strategier.

ABSTRACT

This thesis explores how stochastic chromatin fibre interactions, chromatin organization in the 3D nuclear architecture, and environmental signals collaborate to regulate *MYC* oncogene expression in human colon cancer cells. In Paper I, we employ the ultra-sensitive Nodewalk technique to uncover the dynamic and stochastic nature of chromatin networks impinging on *MYC*. The analyses revealed that the *MYC* interactome mainly consists of stochastic pairwise interactions between *MYC* and its flanking enhancers in two neighbouring topologically associated domains (TADs), which are insulated self-interacting genomic domains. The limits of Nodewalk were also pushed to enable the detection of interactions in very small cell populations, corresponding to the genomic content of ~7 cells. Comparing the frequency of interactions detected in such small input samples with ensemble interactomes of large cell populations uncovered that the enhancer hubs of the ensemble interactomes that appear to simultaneously interact with *MYC* likely represent virtual events, which are not present in reality at the single cell level. These data support a model where *MYC* interacts with its enhancers in a mutually exclusive way, with *MYC* screening for enhancer contacts, rather than the other way around.

Paper II provides a detailed understanding of a novel post-transcriptional mechanism of enhancer action on *MYC* expression. We have thus uncovered that the cancer-specific recruitment of the *MYC* gene to nuclear pores and ensuing rapid nuclear export of *MYC* transcripts - a process that increases MYC expression by enabling the escape of *MYC* mRNAs from rapid decay in the nucleus - require a CTCF binding site positioned within the colorectal oncogenic super-enhancer. Genetic editing by CRISPR-Cas9 was thus commissioned to establish two clones of human colon cancer cells with a mutated sequence in the OSE-specific CTCFBS. Comparing the mutant cells to the parental cell line, we uncovered that the WNT-dependent increase in the nuclear export rate of *MYC* transcripts was abrogated in the CTCFBS mutant clones, providing the first genetic evidence of super-enhancer-mediated gene gating in human cells. In line with this finding, the OSE-specific CTCFBS thus conferred a significant growth advantage to the parental colon cancer cells, compared to the mutant clones. Moreover, we found that WNT-dependent *CCAT1* eRNA transcription is mediated by the OSE-specific CTCFBS that is required for recruitment of AHCTF1 to the OSE to mediate the positioning of the OSE to the nuclear periphery, enabling the subsequent facilitation of *MYC* mRNA export. A multistep molecular process including WNT signalling and the OSE-specific CTCFBS thus underlies the gene gating of *MYC* in human colon cancer cells, and could potentially be targeted for diagnostic or therapeutic uses.

In summary, this thesis explores the dynamics of the stochastic interactomes impinging on the *MYC* oncogene, and provides new insights on the role of 3D chromatin orchestration in the transcriptional regulation of *MYC*. Our analyses uncovered the molecular factors involved in the gene gating of *MYC*, and thus increase our understanding of tumour development. These findings could potentially be beneficial for future diagnostic approaches, or for targeted therapeutic strategies in the treatment of cancer.

LIST OF SCIENTIFIC PAPERS

I. Noriyuki Sumida, Emmanouil G Sifakis, Narsis A Kiani, Anna Lewandowska Ronnegren, Barbara A Scholz, **Johanna Vestlund**, David Gomez-Cabrero, Jesper Tegner, Anita Göndör and Rolf Ohlsson.

MYC as a driver of stochastic chromatin networks: Implications for the fitness of cancer cells. Nucleic Acids Research, 2020 Oct 13, gkaa817.

II. Ilyas Chachoua*, Ilias Tzelepis*, Hao Dai*, Jia Pei Lim*, Anna Lewandowska-Ronnegren*, Felipe Beccaria Casagrande, Shuangyang Wu, **Johanna Vestlund**, Carolina Diettrich Mallet de Lima, Deeksha Bhartiya, Barbara A. Scholz, Mirco Martino, Rashid Mehmood and Anita Göndör.

Canonical WNT signaling-dependent gating of MYC requires a noncanonical CTCF function at a distal binding site. Nature Communications, 2022 Jan 11, 10.1038/s41467-021-27868-3.

CONTENTS

LIST OF ABBREVIATIONS

3C	Chromatin Conformation Capture
3D	Three-dimensional
4C	Circular Chromatin Conformation Capture
5C	Chromosome Conformation Capture Carbon Copy
AHCTF1	AT-hook containing transcription factor 1
APC	Adenomatous polyposis coli
Asf1	Anti-silencing factor 1
ATP	Adenosine 5'-triphosphate
BAC	Bacterial artificial chromosome
BENC	Blood enhancer cluster
BET	Bromodomain and extraterminal domain
BMAL1	Brain and muscle Arnt-like protein-1
Bp/Kbp/Mbp	Base pairs/ Kilo-base pairs/ Million base pairs
BRD2/4	Bromodomain-containing protein 2/4
BRG1	Brahma-related gene 1
CARLo-5	Cancer–associated region lncRNA 5
Cas9	CRISPR-associated protein 9
CBP	CREB-binding protein
CCAT1	Colon cancer associated transcript 1
cDNA	Complementary DNA
ChIP	Chromatin immunoprecipitation
ChIP-seq	Chromatin immunoprecipitation-sequencing
ChrISP	Chromatin *in situ* proximity
CK1	Casein kinase 1
cLADs	Constitutive lamina-associated domains
CLOCK	Circadian locomotor output cycles kaput

Co-IP	Co-immunoprecipitation
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CT	Chromosomal territory
CTCF	CCCTC-binding factor
CTCFBS	CTCF-binding site
DamID	DNA adenine methylation identification
DAPI	4,6-diamidino-2-phenylindole
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
DNA/RNA FISH	DNA/RNA fluorescence *in situ* hybridization
DSB	Double stranded DNA-break
EnhD	Enhancer D
eRNA	Enhancer RNA
ESCs	Embryonic stem cells
FACS	Fluorescence activated cell sorting
FAM49B	Family with sequence similarity 49 member B
FBN2	Fibrillin 2
FBS	Fetal bovine serum
FGF5	Fibroblast growth factor 5
fLADs	Facultative lamina associated domains
FRZ	Frizzled protein
GAL	Galanin And GMAP Prepropeptide
GFP	Green fluorescent protein
GSK3β	Glycogen synthase kinase-3β
GWAS	Genome-wide association studies
H1, H2A/B, H3, H4	Histone 1/linker histone, Histone 2A/B, Histone 3, Histone 4

H3K122ac	Histone 3 Lysine 122 acetylation
H3K27ac	Histone 3 Lysine 27 acetylation
H3K27me3	Histone 3 Lysine 27 tri-methylation
H3K36me3	Histone 3 Lysine 36 tri-methylation
H3K4me1/2/3	Histone 3 Lysine 4 mono/di/tri-methylation
H3K9me2/3	Histone 3 Lysine 9 di/tri-methylation
HCEC	Human colon epithelial cells
HCT116	Human colon cancer cells
HMGB1	High mobility group box 1
HP1	Heterochromatin protein 1
ICR	Imprinting control region
IGF2	Insulin-like growth factor 2
INO1	Inositol-3-phosphate synthase 1
ISPLA	*In situ* proximity ligation assay
JMJD6	Jumonji Domain Containing 6
KCNA4	Potassium voltage-gated channel subfamily A member 4
LADs	Lamina associated domains
LLPS	Liquid-liquid phase separation
lncRNA	Long non-coding RNA
LRP5/6	LDL Receptor Related Protein 5/6
LOCK	Large organized chromatin K9 modification
LSD1	Lysine-specific histone demethylase 1
MAX	MYC-associated factor X
MC-4C	Multi-contact 4C
MED1	Mediator subunit 1
mRNA	Messenger RNA
miRNA	Micro RNA

mRNP	Messenger ribonucleoprotein
MRS	Memory recruitment sequence
MYC	Myelocytomatosis proto-oncogene
NADs	Nucleoulus associated domains
ncRNA	Non-coding RNA
NDR	Nucleosome depleted region
NF-kB	Nuclear factor-κB
NGS	Next generation sequencing
$NHEIII_1$	Nuclease hypersensitivity element III_1
NL	Nuclear lamina
NPC	Nuclear pore complex
NUP	Nucleoporin
nt	Nucleotides
NuRD	Nucleosome remodelling and deacetylase
NXF1	Nuclear RNA export factor 1
OCT4	Octamer-binding transcription factor 4
OSE	Oncogenic super-enhancer
PARD3	Partitioning defective 3 homolog
PARP1	Poly-ADP-ribose polymerase 1
PARylation	Poly(ADP)ribosylation
PBAP	Poly-bromo containing Brm-associated protein
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PHF19	PHD Finger Protein 19
PIN1	Peptidylprolyl *cis/trans* isomerase, NIMA-Interacting 1
POM	Pore membrane protein
PRDM14	PR/SET Domain 14

PTM	Post-translational modification
qPCR	Quantitative PCR
qRT-PCR (RT-qPCR)	Quantitative reverse transcription PCR
RARα	Retinoic acid receptor-α
RBP	RNA binding protein
RNA	Ribonucleic acid
RNA pol I/II	RNA polymerase I/II
RNA-seq	RNA-sequencing
rRNA	Ribosomal RNA
RSC	Remodels the structure of chromatin complex
SAGA	Spt-Ada-Gcn5 acetyltransferase
scHi-C	Single cell Hi-C
SE	Super-enhancer
SENP1/2	Sentrin-specific protease 1/2
sgRNA	Single guide RNA
siRNA	Small interfering RNA
SLCO5A1	Solute carrier organic anion transporter family member 5A1
SNF2	Sucrose non-fermentable 2
SNP	Single nucleotide polymorphism
SOX2	SRY-box 2
SRSF1/2	Serine/arginine-rich splicing factor 1/2
SUMO	Small ubiquitin-like modifier
SWI/SNF	SWItch/Sucrose Non-Fermentable
SWR1	SWI2/SNF2-related 1
TADs	Topologically associated domains
TARDBP	TAR DNA-binding protein 43
TBC1D16	TBC1 Domain Family Member 16

TCF/LEF	Transcription factor/lymphoid enhancer-binding factor
TF	Transcription factor
TGF-β	Transforming growth factor- β
ThPOK	T-helper-inducing POZ/Krueppel-like factor
TOP1	Topoisomerase 1
TPR	Translocated promoter region, nuclear basket protein
TREX-2	Transcription and Export-2
tRNA	Transfer RNA
TSS	Transcription start site
UTR	Untranslated region
WNT	Wingless/Integrated
WRE	WNT response element
WT	Wild type
XIST	X inactive specific transcript

1 INTRODUCTION

1.1 PRINCIPLES OF EPIGENETICS

The ecosystems surrounding and within us, both in micro and macro, are ever-changing. Consequentially, it has been a requisite for organisms to be able to adapt and respond to internal and external stimuli in order to overcome the challenges imposed by environmental fluctuations. Phenotypic plasticity, *i.e.* the ability of cells to acclimate and assume a certain phenotype in response to such internal and external cues[1], plays an important role during development as well as in mature organisms by enabling stem cells sharing one identical genome to establish a vast number of different cell types and specialized functions[2-4]. Such responses to environmental fluxes are mainly achieved through epigenetic regulation, which mediates mitotically, or rarely even meiotically, inheritable, and yet reversible, changes to the cell's phenotype, without the entailing differences being attributed to alterations of the cell's genome[1,5,6]. Thus, acquiescent epigenetics and the ensuing phenotypic plasticity comprise important mechanisms of protection against perturbations of the phenotype[7,8]. As this plasticity is essential for maintaining homeostasis beneficial for organisms, the notion that disruptions of such protective mechanisms could entail pathological events seems foreseeable. Indeed, impaired epigenetic regulation have been found to play a significant role in several complex diseases and ailments, such as cancer, where epigenetic alterations have not only been found to contribute to disease progression and pathology, but contrary to former belief, have also been implied as disease-initiating events[9,10]. Thus, the capacity of the cell to adapt to environmental fluxes is involved in both maintenance of robust homeostasis and in the pathophysiology of disease - in the case of a perturbed phenotypic adaptivity - making the increased understanding of epigenetic regulation elemental for gaining new insights into the evolution of diseases ranging from diabetes to cancer.

1.1.1 Phenotypic plasticity and canalization

The term "epigenetics" was originally introduced by Conrad Waddington in 1942 in his article *The Epigenotype*[11], in which interactions between environmental influences and the genome were hypothesized to establish distinct phenotypes. Waddington's renowned theory about the effects of epigenetics on phenotype is known as the epigenetic landscape model (or simply the "Waddington landscape"), which describes how embryonic stem cells are able to differentiate into a vast number of different cell types through canalization. The term canalization stems from the fact that Waddington imagined the persuasion of differentiation as a developmental landscape adorned by valleys with several divergences, in which a metaphoric marble will take different paths depending on the slope and tilt of the landscape[12]. Thus, the model depicts a scenario where a marble is driven towards a mature cell fate at the end of the slope by gravitational forces representing regulatory factors of epigenetics and differentiation. As the marble travels through an increasing number of divergences on its path, the number of potential diverse cell fates it might assume diminishes on the journey

towards differentiation. In order for this developmental plasticity to take place, Waddington recognized that there had to be an active exchange of influences between external cues and genotype, and thus coined the concept of epigenetics, although, the precise biomolecular mechanisms underlying epigenetic regulation were not known at the time. Since the first depictions of epigenetics emerged, the field has vastly developed. It is now known that epigenetic perturbations are implicated in the loss of maintenance of mature cell fates, causing the cell to stray from its path of normal development[9]. Under this scenario, epigenetic states also constitute the walls dictating the direction of the marble's movement. Consequently, epigenetic dysregulation either reduces the walls or establishes new walls that re-direct the marble towards pathological states.

During development, canalization and plasticity thus work in synchrony to regulate cell fate through heritable and reversible chromatin modifications that possess the ability to alter gene expression while remaining responsive to certain environmental signals. Consequentially, these two mechanisms, which at first glance might appear significantly different from one another, balance the two sides of epigenetic regulation to provide an adaptive response system which allows the cell to respond to environmental cues, so that it can best adjust and maintain a robust, well acclimatized phenotype.

Since Waddington initially portrayed the epigenetic landscape model, this system has been developed further to also account for stochasticity and buffering effects of epigenetic regulation[7,13]. The original Waddington model held a more deterministic approach that primarily focused on regulation by particular genes, whereas the revised model, proposed by Pujadas and Feinberg, also further highlights the role of epigenetic regulation of gene expression variability during development and in response to external stimuli[7]. Hence, the chromatin organization affects transcriptional regulation and cellular states during development, in health and in disease, as they are able to tweak and pull at the epigenetic landscape model, thereby altering the width and depth of its gorges.

1.1.2 Chromatin: a platform for epigenetic regulation

On account of such epigenetic adaptability, a human stem cell sharing one genome can give rise to hundreds of different cell types by establishing cell type-specific gene expression patterns by responding to environmental and developmental stimuli[4]. These internal and external cues are able to regulate gene activity and phenotype by coordinating the effects of transcription factors (TFs), signaling pathways, stochastic events and other factors on chromatin, which serves as an fundamental platform for the regulation of genetic information[4,14].

Chromatin is present in all eukaryotic cells, and consists of DNA and proteins. The diploid human genome comprises 6 billion bases of which only a meager 2% are genes coding for proteins[4,15]. The remaining 98% of the genome consists of non-coding DNA that abounds with regulatory elements that support the establishment of cell type-specific gene expression

patterns[4]. The DNA molecules of a diploid human cell reach a length of ca 2 m in total[16]. In order for the cell to accommodate the long DNA molecules, it is enfolded around 30 million nucleosomes, which are formed by DNA segments of approximately 145-147 base pairs (bp) wrapping around an octamer formed by two copies each of four positively charged histone proteins: H2A, H2B, H3, and H4[15,17-19]. An additional histone variant, linker histones (H1), occupy the nucleosomes' DNA entry and exit sites, entailing a stabilizing function for both the structure of nucleosomes and higher-order chromatin architecture[20]. The nucleosomes are separated by sequences of DNA approximately 20-40 bp in length, termed linker regions, giving the nucleosomes an average span of ~200 bp, while also composing the characteristic "beads on a string" appearance of the primary chromatin fibre[14,15,18,21].

Chromatin can be modified through both acquisition and removal of chromatin marks that leave the DNA sequence itself unaltered[14]. Such modifications are referred to as epigenetic marks when mitotically heritable, and may consist of DNA methylation, histone acetylation and other post-translational modifications (PTMs) of histones, binding of transcription factors, and non-coding RNAs[14,21]. The epigenetic marks thus add further shape to the structure of the primary chromatin fibre, and the different ensuing combinations of chromatin modifications create a vast number of conceivable variants, which benefit the establishment of different structural chromatin states[18,22].

In brief, chromatin can be categorized into two structural types: euchromatin and heterochromatin. The former consists of more loosely packed chromatin, which accordingly is more accessible for DNA-templated activities such as transcription, while heterochromatin is more compact and transcriptionally repressed[18,23]. Euchromatin is enriched in activating epigenetic marks, such as histone acetylation of H3 and H4 (*e.g.* H3K27ac and H3K122ac) and methylation of H3 (*e.g.* H3K4me1, H3K4me2, H3K4me3)[24-26]. The acetylation of the histones' lysine residues neutralizes their positive charge and consequently causes an increased accessibility of the negatively charged DNA through chromatin decompaction, hence allowing for transcriptional activation[22]. In opposition, heterochromatin is abounding with repressive histone modifications, such as H3K9me2/3 in the case of constitutive heterochromatin which remains compacted throughout the cell cycle, and H3K27me3 in the case of facultative heterochromatin, which is favorably, although not exclusively, heterochromatic[18,27,28].

Apart from chemical modifications, also nucleosome positioning and density are essential components of transcriptional regulation, although the precise underlying mechanisms remain only partially understood[8,29,30]. Organization of the genome by enfolding around nucleosomes thus not only permits compact packaging of the DNA, but also entails a reduction of accessibility of chromatin for DNA-templated activities such as transcription. The number of nucleosomes is reduced in aging yeast and mammalian cells, and has been associated with aberrant expression of the histone chaperone Asf1, with ensuing downregulation of histone H3 and H4 biosynthesis[31,32]. As a reduced number of histone proteins and nucleosomes renders the compaction of chromatin more open, it is conceivable that such changes to

chromatin structure may alter transcriptional activity. Indeed, Celona *et al.* used siRNA transfection targeting high mobility group box 1 (*Hmgb1*), encoding a protein which facilitates nucleosome assembly, and were able to document that the abundance of histones and nucleosomes in *Hmgb1-/-* mammalian HeLa cells was reduced by approximately 20%[30]. Moreover, lack of HMGB1 protein was associated with a global increase of transcription, with *Hmgb1-/-* cells containing approximately 1.3 times more RNA transcripts than the control HeLa cells as quantified by FACS[30]. Furthermore, reduction of the amount of histones and nucleosomes has also been implicated in the regulation of coordination between replication and transcription, as depletion of H1 caused an accumulation of stalled replication forks and DNA damage due to transcription-replication conflicts[33].

Additionally, histone exchange plays a key role in regulation of transcription[34]. Chromatin decompaction is essential for transcription initiation, and chromatin remodellers such as the ATP-dependent Remodels the Structure of Chromatin (RSC) complex of the Snf2-family, a paralogue of the SWI/SNF complex in *S. cerevisiae*[35], maintain nucleosome-depleted regions (NDRs, also known as nucleosome free regions [NFRs]) by sliding nucleosomes along the chromatin in order to aid recruitment of RNA polymerase II (RNA pol II) to gene promoter regions[36-38]. Moreover, depletion of RSC by using a heat-sensitive degron system targeting the catalytic subunit protein Sth1 of the RSC complex results in global transcriptional downregulation[37,39]. Although the formation of NDRs alone is not sufficient for RNA pol II recruitment and transcription initiation, it facilitates SWR1-dependent incorporation of the histone variant H2A.Z into the nucleosomes that flank NDRs (termed +1 and -1 according to their positioning either upstream or downstream, respectively, of the NDR) at promoters[40]. Subsequentially, the exchange of the canonical H2A-H2B dimer and incorporation of H2A.Z at the +1 nucleosome facilitates recruitment of transcription factors and chromatin remodellers, with ensuing transcriptional activation[41]. Accordingly, Draker *et al.* discovered that upon hormone activation, the bromodomain and extraterminal domain (BET)-family protein bromodomain containing 2 (Brd2) is recruited to androgen receptor (AR) regulated genes in a H2A.Z dependent manner, and that JQ1, a BET inhibitor which disrupts the interaction between BRD2 and H2A.Z nucleosomes, inhibits cell proliferation and thus also has profound effects on gene expression[41].

While histone modifications characterize chromatin states that dictate gene activity, they also distinguish regulatory elements of gene expression, such as promoters, enhancers and super-enhancers (SEs). However, the picture is complicated by the fact that there are dynamic and context-dependent factors influencing the tenure of epigenetic marks, and some regulatory elements can also be bereft of otherwise characteristic histone modifications[26,42]. Thus, histone modifications, repressive and activating marks alike, are not permanent and may be rendered flexible by environmental cues and fluxes. Therefore, the nuclear compartmentalization of active and inactive chromatin domains appears to constitute one of the forces that contribute to a buffering function, enabling transcriptional processes to withstand dysregulation due to stochastic events[8].

1.2 THE NUCLEAR ARCHITECTURE AND 3D CHROMATIN ORGANIZATION

1.2.1 Nuclear compartmentalization and separation of active and inactive chromatin

Spatiotemporal orchestration of chromatin has emerged as an important facet of transcriptional regulation, although the precise underlying mechanisms are not yet fully understood. In concordance with this notion, the nuclear organization of the genome is not random; and since further back than a century ago, perturbations of chromatin organization have been observed in several ailments and diseases, such as cancer[43,44].

One of the first examples of spatial chromatin organization is represented by the discovery of the nucleolus, which comprises genomic loci from several different chromosomes juxtaposed to the vicinity of one other, and that are involved in the transcription of genes coding for ribosomal DNA and synthesis of ribosomal RNA (rRNA)[45,46]. The prominence of spatial chromatin organization can also be observed during the interphase of the cell cycle, as chromosomes are folded to form chromosome territories (CTs), prior to their compaction into mitotic chromosomes to enable proper segregation during mitosis[47,48]. These spatial clusters of chromosomes may intermingle with one another, and also exhibit a cell-type specific preference regarding their organization in the nucleus that also contributes to the stability of gene expression patterns[49]. Within the CTs, the organization of chromatin is highly regulated and correlated with transcriptional activity, with transcriptionally permissive regions typically being located close to the CT borders[50,51]. Chromosome conformation capture and its derivate techniques, such as Hi-C, which detects physical long-range interactions of the genome in an all-to-all manner, have generated heat maps documenting that CTs can be further divided into active “A compartments” and inactive “B compartments”, with the former being distinguished by open chromatin and the latter by condensed chromatin[51-53]. The mechanisms behind the formation of such compartments remain not fully understood, however, it seems to potentially be mediated at least in part through liquid-liquid phase separation (LLPS) of chromatin-associated proteins[54].The existence of such nuclear compartmentalization has also been documented by multiplexed FISH-methods and super-resolution microscopy-based approaches[55,56]. Accordingly, interactions within compartmentalized territories sharing similar chromatin marks, proteins and transcriptional states that share affinity for each other, are thought to contribute to the formation of membrane-less compartments such as the nucleolus[57], through the LLPS phenomenon, with distinct liquid-like condensates with well mixed, highly concentrated and uniformly distributed content[58-60]. Several different kinds of molecules may contribute to the formation of these condensates, and include chromatin modifications, PTMs, transcription factors or proteins (*e.g.* HP1, Polycomb proteins, Mediator, RNA polymerase II, bromodomain-containing protein 4 (BRD4), OCT4, SRSF1/2 and PTBP1)[61].

The nuclear interior also houses modules important for transcription and posttranscriptional modulation; such as transcription factories and interchromatin granule clusters, also known as splicing factor compartments or splicing speckles[62-64]. Chromatin organization is not limited to intrachromosomal compartmentalization, and hubs of active chromatin have been found located around such nuclear speckles, while inactive hubs were organized around the nucleolus[65]. However, it is likely that interactions of such hubs are represented by transient stochastic events, as pulse-labeling experiments have shown that in interphase cells only 1% of chromatin from separate chromosomes co-localizes[66].

Another level of genome organization revealed by heat maps generated by using the "C family" techniques, such as Hi-C and carbon copy chromosome conformation capture (5C), are the sub-nuclear compartments termed topologically associated domains (TADs) that compose A and B compartments, and which consist of loops of self-interacting regions of the genome[51]. Similar to LAD boundaries, the TAD borders are also demarcated by the architectural protein CTCF, as well as cohesin[67], and are further characterized by the presence of histone marks typical for transcriptionally active chromatin, such as H3K4me3 and H3K36me3, and housekeeping genes[51,68] (**Figure 1**).

Layered on top of this organization, chromatin is also radially arranged, with A compartment and gene-rich regions of the genome frequently localized in the nuclear interior. Conversely, gene-poor and AT-enriched regions of B compartments are typically positioned at the nuclear periphery, preferentially associating with the nuclear lamina and forming structures termed lamina-associated domains (LADs), or they localize to the nucleolus[50], forming nucleolus-associated domains (NADs)[47,65,69]. LADs consist of 100 kb-10 Mb sized regions bordered by repressive H3K27me3 histone marks and overlap with repressive H3K9me2 Large Organized Chromatin K9 modifications (LOCKs) that have an average size of roughly 100 kb[50,70-72]. In mammalian model organisms LAD borders are clearly demarcated, with an enrichment of CCCTC-binding factor (CTCF) binding motifs, and with 9% containing CTCF within 10 kb from the LAD boundaries[73]. The LAD borders are also abound in CpG islands, and intriguingly, also active gene promoters[74]. Moreover, LADs constitute a highly distinctive feature of the epigenome, as they can comprise up to 40% of the genome in certain cell types, including both constitutive and facultative heterochromatin, with the former (cLADs) consisting of cell-type-independent silenced regions, and the latter (fLADs) of developmentally silenced genes[69,75,76]. The nuclear periphery also plays an important role during development, where LOCKs constitute <5% of the genome in undifferentiated embryonic stem (ES) cells, while composing > 30% of the genome of some differentiated cells [71]. Given the link between the nuclear periphery and gene regulation, it will be important to uncover the factors that mediate peripheral localisation and affect gene expression at the lamina, as well as their mechanism of action.

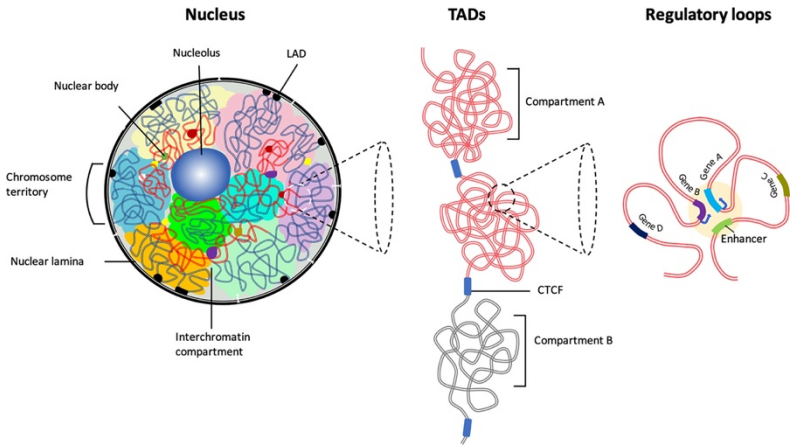

***Figure 1**: Hierarchical chromatin organization in the interphase nucleus. Reprinted under the Creative Commons Attribution 4.0 International License, https://creativecommons.org/licenses/by/4.0/, by Frontiers in Immunology, 2021;12: 633825. Jagan M.R. Pongubala, Cornelis Murre. Spatial Organization of Chromatin: Transcriptional Control of Adaptive Immune Cell Development [77].*

1.2.2 Transcriptional control by 3D chromatin organization and the nuclear periphery

The radial positioning of genes within the nucleus has thus been implicated in transcriptional regulation, with LADs located at the nuclear periphery being associated with transcriptional repression; a phenomenon that has been attributed to gene silencing mediated by the nuclear lamina (NL)[78,79].

The NL comprises a fibrous layer that lines the inner nuclear membrane, and which consists of type V intermediate filament proteins termed lamins[69]. The genomic regions that are anchored to the NL thus constitute LADs: repressive chromatin domains which possess several characteristics that distinguish heterochromatin[69]. The vast majority of genes positioned within LADs are transcriptionally silenced, or expressed only at low levels, and overlap with regions that replicate late during S-phase[72,74]. Many so called "gene deserts", *i.e.* >1Mb regions bereft of protein-coding genes, as well as human pericentromeric heterochromatin, and some telomeric regions, are also frequently located in LADs[71,74].

The underlying mechanisms contributing to the repressive effects of the LAD environment remain elusive, however, it seems that both structural characteristics of LAD chromatin and NL interactions may play roles in the process[69]. Several repressive histone modifications (*e.g.* H3K9me2/3 and H3K27me3) distinguish LADs, and deletion of the histone methyltransferase G9a, which mediates H3K9me2, has been shown to cause a preferential upregulation of genes in LADs in mouse embryonic stem cells (mESCs)[80]. Moreover, a study where an integrated reporter gene was tethered to the NL demonstrated that such repositioning was associated with a 2 – 3-fold decrease in gene expression, and during *Drosophila* neuroblast differentiation, depletion of lamin entails incomplete silencing of the *hunchback* (*Hb*) gene, whose recruitment to the nuclear periphery is also prevented[79,81].

A conceptual question that arises is whether transcriptional attenuation is a consequence of gene recruitment to LADs, or if LADs are formed by repressed gene expression. Our group has previously documented how the genome organizer CTCF, together with poly-ADP-ribose polymerase 1 (PARP1), is able to facilitate rhythmic recruitment of circadian loci to LADs, where they undergo a gradual transcriptional attenuation and acquire repressive histone marks[82]. These findings would argue for the first possibility, which is further supported by a study that randomly inserted thousands of identical reporter genes at different locations, and discovered that the reporter genes were >5-fold less active when inserted into LADs compared to inter-LAD regions[83]. However, the juxta-positioning of genes to the nuclear periphery on its own does not exclusively entail silencing of gene activity[84,85]. Already in 2008, Kumaran and Spector were able to stably target genetic loci to the lamina by using an inducible lac operator-repressor-system to target loci to the nuclear periphery[85]. The authors also performed transfection of cells with a dual promoter vector, expressing Tet-On and MS2-YFP, and subsequent visualization by 3D time-lapse microscopy, which revealed that a significant percentage of both targeted and non-targeted Lamin B1-associated regions were inducible at the nuclear lamina, adding another layer of complexity and indicating that other environmental and contextual factors influence transcriptional regulation by genome organization[85].

Hence, plasticity and adequate responses to external stimuli are of great importance during dynamic processes, such as during development. Interestingly, the formation of LADs displays highly flexible features, and genome-wide maps of NL associating regions have been uncovered by DNA adenine methylation identification (DamID) using Lamin B1 in embryonic stem cells (ESCs), multipotent neural progenitor cells and terminally differentiated astrocytes; exposing a basal chromosome orchestration present in ESCs which is altered at hundreds of regions in a cumulative manner during differentiation[72,74]. Moreover, many genes that moved away from the nuclear lamina were concomitantly activated, however, many also remained repressed and became unlocked for transcriptional activation in a subsequent differentiation step, hence suggesting an essential role for lamina-genome interactions in the regulation of gene expression programs during lineage commitment[72].

1.2.3 Chromatin mobility

DNA-templated activities, such as replication, transcription and DNA damage repair, constitute highly dynamic processes that require plastic changes of chromatin fibre structure and mobility to and from structural hallmarks of the nucleus such as the lamina[86,87]. Chromatin as an entity in organisms is under constant motion, and in yeast increased chromatin mobility both locally at sites of double stranded DNA-breaks (DSBs), as well as globally, has been documented by several authors[86,87]. Moreover, chromatin fibres have been found to travel vast distances in yeast interphase nuclei, and, in contrast to the polewards movements of chromosomes occurring during mitosis that are motorically driven by microtubuli, these movements seem to lack a clear directionality[88]. However, as the mobility

of the chromatin fibres was highly sensitive to metabolic states and ATP depletion, such movements are likely not entirely random or resulting only through diffusion[88,89]. Actions of ATP-dependent enzymes facilitating chromatin remodelling or transcriptional activities have been suggested to facilitate such movements, a notion supported by the fact that stationary phase cells with lower degrees of transcriptional activity display reduced chromatin mobility[90]. Intriguingly, a recent study by Gu *et al.* provided further evidence for this hypothesis, showing that the mobility of distinct loci correlated with transcriptional activation[91]. Thus, the study demonstrated that active transcription of the *Fgf5* enhancer resulted in enhancer exploring a larger nuclear space compared to in cells lacking ongoing *Fgf5* transcription, and that the peak of mobility coincided with the timing of active transcription at the enhancer alleles[91].

Another prominent example of chromatin mobility is represented by the inactivation of one of the X-chromosomes which takes place in mammalian organisms during lineage commitment and early development[92,93]. The random inactivation of the X chromosome in female mammals is initiated through upregulation of the non-coding RNA (ncRNA) X inactive specific transcript (*XIST*), with *XIST* subsequently inducing the recruitment of the inactive X chromosome to the nuclear lamina[92-94].

In addition, chromatin fibre movements have also been documented to be under circadian regulation. Zhao *et al.* thus documented that upon entrainment of circadian rhythm by serum shock, several regions including the *IGF2/H19* and *PARD3* loci were recruited to the nuclear lamina in a circadian manner facilitated by rhythmic protein-complex formation between the genome organizer CTCF and the DNA repair and transcriptional regulator PARP1. Interestingly, recruitment of circadian genes to the lamina resulted in the transient formation of an inter-chromosomal interactome between active circadian loci and LADs[82]. Combining 3D DNA FISH and RNA FISH with chromatin *in situ* proximity assay (ChrISP) experiments that translate proximity between labelled DNA FISH signals and H3K9me2 repressive modifications into light signals, revealed that the genetic loci were recruited to the nuclear periphery in a transcriptionally active state, and remained active at the lamina for several hours prior to undergoing transcriptional attenuation and concomitant gradual acquisition of repressive H3K9me2 modifications[82]. These experiments thus not only support a causal role of peripheral recruitment in circadian transcriptional repression, but they also evoke the notion of loci potentially arriving first at a locally transcriptionally permissive milieu, such as nuclear pore complexes (NPCs).

1.2.4 The nuclear pore complex: a platform for transcriptional regulation, 3D genome organization and regulation of lineage commitment

Although the nuclear periphery is generally characterized by a repressive environment, it has long been known from electron microscopy images that the condensed heterochromatin at the lamina is frequently interrupted by specks of decondensed chromatin[95]. These fragments,

scattered across the nuclear envelope, proved to comprise nuclear pore complexes (NPCs), macromolecular protein channels that perforate the bi-layered nuclear membrane[96,97]. The structure of the NPC has an octagonal symmetry and approximately 30 different nucleoporins constitute the larger subunits of the NPC; the nuclear basket, nuclear ring, cytoplasmic ring, central pore, and cytoplasmic filaments[95,96,98]. The general structure of the NPCs is highly conserved among eukaryotes, however, there is a considerable variation amongst different species when it comes to the size and composition of the NPC, ranging from 60 kDa in *S. cerevisiae* to approximately 90-120 MDa in human cells[99,100]. There is emerging evidence that the composition of the NPC varies greatly in human and yeast cells. The stoichiometry of NPCs in yeast cells has been explored using high-resolution quantitative live-cell imaging to define the composition of Nups in the NPC by using yeast-strains with GFP-tagged Nups, showing that although the eightfold symmetry of the NPC is preserved, there is a significant difference in the cytoplasmic and nucleoplasmic rings where in yeast the NPC contains only 16 copies of most Nups, whereas the human NPC comprises a significantly higher amount of Nup copies, generally the double number of copies or more[101,102].

1.2.4.1 NPCs and nuclear export of mRNA

The NPC permits compartmentalization between the nucleus and cytoplasm, and functions as a selective barrier where ions and other smaller molecules (mass below ~30 KDa or size < 3 nm in diameter in human cells) can freely diffuse through[103]. However, bigger molecules, such as some RNAs (*e.g.* tRNA, rRNA and miRNA), require transport receptors that translocate their cargo through the NPC, with many proteins and RNAs being assisted by karyopherins that also control transport directionality facilitated by the Ran GTPase system[104]. Contrastingly, export of mRNA is independent of karyopherin proteins and Ran, and instead the dedicated heterodimeric transport receptor Nxf1/Nxt1 or TAP/p15 in mammalian cells, or Mtr2/Mex67 in yeast, are used[105,106]. After initial processing of the pre-mRNA and mRNA in the nucleus, through splicing and formation of mRNPs by association with RNA binding proteins (RBPs) at localization elements (also known as RNA transport signals (RTSs)) mainly in the 3' untranslated region (3'UTR), incompletely processed or incorrectly assembled mRNPs are targeted by quality control mechanisms that direct the defective mRNPs to the nuclear exosome for decay[107-109]. Ultimately, export receptors in the correctly assembled and processed mRNPs associate with specific Nups in the NPC, thus facilitating the export of transcripts out of the nucleus[106,109].

1.2.4.2 The nuclear pore complex, nucleoporins and transcriptional control

While the function of the evolutionarily conserved NPCs as bilateral transport channels between the cytoplasm and nucleus has been known since a long time, there is an accumulating amount of evidence of nucleoporins (NUPs) and NPCs, despite their frequent localization at the generally repressive nuclear periphery, also being directly involved in the regulation of chromatin states and transcription in a wide range of organisms, including mammalian cells[110-112]. For instance, Kuhn *et al.* have documented that ectopic tethering of several nucleoporins, and in particular Sec13, to chromatin in *Drosophila*, is associated with

PBAP-dependent chromatin decondensation and reduced histone density, thus entailing gene activation[111].

While some Nups, termed pore membrane Nups ('Poms'), continuously reside at the nuclear envelope, the majority of nucleoporins are mobile, soluble peripheral proteins, which may also be found in the nuclear interior[109]. As an example, the nucleoporin AHCTF1 (also known as ELYS) shows considerable mobility within the nucleus, and is part of both the nuclear pore and kinetochore, with RNAi experiments documenting that AHCTF1 is required for both cell division and NPC assembly at the nuclear envelope[113]. Moreover, using immunofluorescence imaging, luciferase assays, and chromatin immunoprecipitation (ChIP) assays, Gao *et al.* recently showed that in rat cardiomyocytes the mobile Nup50 selectively induces transcriptional activity of the K^+ voltage gated-channel subfamily A member 4 (*Kcna4*) gene through direct binding of a phenylalanine and glycine (FG)-repeat domain within Nup50 and the *Kcna4* promoter[114]. Interestingly, in yeast the vast majority of interactions between Nups and genes take place at the nuclear envelope, whereas the nucleoporins possess a higher degree of mobility in metazoan organisms, where they frequently interact with chromatin also in the nucleoplasm[110]. To explore the dynamics of Nups, Rabut *et al.* used GFP-tagged nucleoporins to follow their localization *in situ*, and uncovered that the central components of the NPCs remain soundly positioned at the nuclear envelope, whilst more peripheral Nups, such as Nup98, Nup50, and Nup153 were more mobile and dynamic in their association with NPCs[115]. Intriguingly, the highly mobile nucleoporins Nup50, Nup98, and Nup153 have been further linked to transcriptional regulation through their association with RNA polymerase II (RNA pol II) activity, as shown by Griffis *et al.* in a study where the mobility of Nup98 was lost in tsBN2 cells upon treatment with actinomycin-D and other inhibitors of RNA pol I and II[116], indicating that the dynamic mobility of Nup98 back and forth between the nuclear interior and the NPC requires active transcription.

DNA topoisomerases are enzymes which control DNA topology through relieving torsional stress, and are therefore essential for DNA-templated activities[117]. Nups have been found to share a connection with the activity of topoisomerase 1 (TOP1), as some Nups, such as Nup153, have been shown to be able to interact with SUMO proteases SENP1 and SENP2[118], and TOP1 activity is inhibited at transcriptionally active chromatin by SUMOylation[119]. Moreover, binding sites for Nups have been observed at active gene bodies, implying a role in transcription elongation or relocation of polymerases to the gene promoters in order for subsequent transcriptional re-activation[120]. Additionally, certain Nups, such as Nup153 and Nup93, have been found to bind to enhancers and promoters of genes significant for development, thereby regulating their expression[112,121]. The NPC basket protein TPR and Nup153 were also shown to associate with the transcription factor MYC (c-Myc) after its recruitment to the nuclear periphery in mitogen-activated cells that exhibit high proliferation[122]. The authors further revealed that activation of MYC through serine-62 phosphorylation or PIN1-mediated polymerization facilitated the association of MYC with the nuclear basket of the NPC, and promoted the establishment of a transcriptionally

permissive milieu at the nuclear periphery, which included the SAGA complex, acetyltransferases and target genes of MYC[122]. These findings thus imply that gene positioning and NPCs play essential parts in transcriptional regulation.

1.2.4.3 Nuclear pore plasticity in transcriptional memory and cell development

NPCs have been described to play a role in the rapid re-activation of inducible genes in yeast upon environmental cues[123]. As an example, fluxes in temperature and metabolic changes can induce transcription of genes such as INO1, HXK1, and GAL genes in *S. cerevisiae* if tethered to NPCs[110]. NPCs have also been implicated in the regulation of 3D chromatin organization in *S. cerevisiae* and *Drosophila melanogaster*, by forming chromatin loops between promoters and 3' terminal ends of genes, consequently tethering either, or both, ends of the gene to components of the NPCs[124,125]. This loop-formation accordingly provides transcriptional regulation in a process known as transcriptional memory, which promotes rapid recruitment of RNA pol II and subsequent rapid transcriptional re-activation after briefer periods of repression[126]. Transcriptional memory allows for precision of control of gene expression programs during development through transmission of active transcriptional states from mother cell to daughter cell, and the importance of such epigenetic memory during development has been particularly well-documented in the *Drosophila* embryo[127]. Moreover, early proteomics studies from two decades ago uncovered a vast variation of Nup levels (such as Nup37, Nup50, TPR, Nup210, Nup214, and Pom121) among different human cell types and cancer cell lines[128-130]. Thus, the notion of cell-type specific expression levels of certain Nups being able to facilitate changes in gene expression and protein export during development emerged. Indeed, examples of such developmental regulation by Nups and NPCs have been documented, and an acclaimed example is the NPC transmembrane ring component Nup210[104]. Nup210 expression is cell-type specific during organogenesis in mice[131], and although Nup210 is absent in proliferating myoblasts and ESCs, its expression level increases during differentiation in the mouse C2C12 *in vitro* myogenic model system to incorporate Nup210 into NPCs, subsequently inducing gene expression programs vital for cell differentiation[132]. Additionally, Nup210 RNAi did not affect nucleo-cytoplasmic transport, but blocked both myogenic and neuronal development, hence suggesting that Nup210 is required for induction of gene expression programs mediating cell fate decisions[132]. While Nup210 levels have been shown to increase during differentiation, levels of the nuclear basket component Nup153 have been observed to decrease during neural differentiation, with levels anti-correlating with the degree of cellular plasticity[104,133,134]. In addition, Nup153 has been documented to bind to silenced developmental genes in the nucleoplasm and at NPCs in mESCs, suggesting that Nup153 is important for maintenance of undifferentiated cell states[133]. These data indicate that the levels of Nups and the composition of NPCs can vary in different developmental stages and cell types. Moreover, Nups and NPCs may direct gene expression patterns to either promote or hinder cell differentiation, deducing the idea that specialized NPCs with diverse features might ultimately lead to cell-type specific functions[104,132].

The inducible *INO1* gene in yeast illustrates another well-studied example of transcriptional regulation and transcriptional memory by NPCs, as its gene promoter associates with NPC components upon transcriptional activation[135,136]. The recruitment of *INO1* to the NPC is controlled by two upstream 8 bp and 20 bp *cis*-acting DNA elements referred to as 'DNA zip codes', and after repression *INO1* remains poised for transcriptional activation at the NPC for 3-4 mitotic cycles through transcriptional memory facilitated by a Memory Recruitment Sequence (MRS)[135-138]. The involvement of such DNA sequences in the regulation of association of loci to NPCs thus suggests that regulation of spatial chromatin organization and transcriptional memory in yeast is encoded by the genome itself - in addition to being facilitated by Nups and NPCs[135].

1.2.4.4 Role of nuclear pores in spatial genome organization

Evidence for the role of NPCs in 3D genome organization has been emerging in several model organisms. For example, tethering of chromatin bound Nups to the NPCs in *Drosophila* has been shown to regulate chromatin organization at the NPCs, and has been implied to hold the ability to affect global chromatin organization through a negative regulatory loop consisting of Nup155 and Nup62[139,140]. In yeast the NPC proteins Mlp1 and Mlp2 are essential for maintaining genome stability, as they facilitate physical contact with messenger ribonucleoprotein particles (mRNPs) and chromatin proximity to the NPC, thus preventing detrimental R-loop formation, a process that occurs during transcription as mRNA hybridizes back to DNA and can be both physiological and pathological[141]. Moreover, the mobile Nup98 has been observed to associate with CTCF, the master organizer of 3D chromatin conformation, upon ecdysone induction[142].

In summary, these findings suggest a principal regulative role of NPCs and Nups in regards of 3D genome orchestration, as a platform for transcriptional regulation, and in developmental linage commitment, however, the precise underlying mechanisms remain not fully elucidated.

1.2.5 The gene gating hypothesis and its mechanism in yeast and mammalian cells

In eukaryotes, the exit of mRNPs from the nucleus to the cytoplasm is essential for their localization to sites of translation, and for further downstream effects. In 1985, Günter Blobel proposed a more sophisticated model of gene regulation by NPCs in yeast[143]. This mechanism was termed "gene gating", and depicted the anchoring of inducible genes to the NPCs that he postulated to serve as organelles that coordinated transcription and RNA maturation with a rapid export of mRNAs to the cytoplasm through the NPCs. Following genome-wide studies in yeast supported the notion of Blobel's gene gating hypothesis, as they demonstrated that some Nups and NPC-associated transport factors, such as karyopherins, preferentially associate with active genes, and uncovered the presence of a vast number of interactions between gene promoters and components of the NPC nuclear

basket[120,144]. However, while such regulated export of transcripts has also been observed in *Drosophila* and *C. elegans*, it has not been explored and described in mammalian model organisms until recently[145-147].

A recent study from our group documented a novel form of gene gating mechanism in human colon cancer cells, by which *MYC* expression is post-transcriptionally increased through recruitment of *MYC* to the nuclear periphery, where facilitated export of *MYC* mRNAs through NPCs enables the transcripts' escape from rapid decay inside the nucleus. This gene gating process was found to be mediated by an oncogenic super-enhancer (OSE) present in the HCT116 cancer cells, but absent in normal colon epithelial cells[147]. There are both shared and distinct roles for Nups in gene gating across species. The mammalian genomes and nuclei possess some significant differences compared to the ones of *S. cerevisiae* and *Drosophila*. For instance, the human nucleus is ~10 μm in diameter and with a genome consisting of ~3 billion base pairs (or approximately 2865 Mb)[148,149], while the *Drosophila* and *S. cerevisiae* nuclei comprises significantly smaller nuclear volumes and genome sizes (~175 Mb and ~12 Mb respectively)[150,151]. Such differences make it foreseeable that there might be significant variances in the mechanism(s) of gene gating amongst different species. In mammalian cells, RNPs travel through channels in the interchromatin compartments on their way to the NPC at the nuclear envelope, a feature that has not been observed in a similar manner in yeast or *Drosophila*[109], potentially due to the significantly smaller nuclear volumes of the latter species. However, there are also a number of similarities, all disparities aside. In both mammalian model systems and in *Drosophila*, the highly mobile Nup98 impacts transcriptional events in the nucleoplasm, such as regulation of transcriptional memory of gamma interferon (IFN-γ)-induced genes in human cells and stress-induced genes in *Drosophila*[109]. Moreover, in *S. cerevisiae* Nups are required for the recruitment of several genes to the nuclear periphery in order for optimal gene expression, a feature that has recently been documented also in human cells. Scholz *et al.* has thus demonstrated that downregulation of *AHCTF1/ELYS* caused a significant reduction not only in the percentage of *MYC* alleles that are recruited to the lamina but also in the rate of nuclear export of *MYC* transcripts in HCT116 human colon cancer cells[147].

Many questions remain unanswered regarding the regulation and effects of gene gating; for instance, little is known about potential mechanisms and stimuli that control the gating process, and whether gene gating is limited to a few selective genes or acts genome-wide. However, data regarding the regulation of gene gating is emerging, and gene gating can be regulated through signalling pathways such as the WNT cascade, or by PTMs of NPC components. A recent study from our group uncovered that WNT signalling plays a central role in the gating of *MYC* in HCT116 cells, as β-catenin was found necessary for interactions between AHCTF1 and TCF4, which are required for anchoring of the OSE to the NPC[147]. Moreover, in yeast the lysine acetyltransferase Esa1 of the NuA4 complex was recently found by Gomar-Alba *et al.*[152] to mediate acetylation of Nup60 of the nuclear basket, entailing recruitment of the mRNA export factor Sac3; a scaffolding subunit of the Transcription and Export-2 (TREX-2) complex, to the nuclear basket, thus promoting export of mRNA

transcripts. Intriguingly, gene gating and acetylation of the NPC is coupled with cell cycle regulation, as the export of mRNAs has been shown to promote entry from G1- into S-phase[104,152,153], hence supporting the notion that gene gating can also be influenced by factors such as metabolic states and external stimuli.

Perturbed gene gating might contribute to cancer development *via* several mechanisms. Although the recruitment of the gene encoding the tumor suppressor p21 - which also functions as a cell-cycle inhibitor and anti-proliferative effector in normal cells[154] - to NPCs has not been demonstrated yet, its RNA and protein levels are regulated by Nup98, which associates with the 3' UTR of p21 mRNA, thus protecting the transcript from exosomal degradation[155]. Moreover, in the aforementioned study by Scholz *et al.*, the authors documented a mechanism of how colon cancer cells seem to have developed a specific gene gating mechanism, which involves an oncogenic super-enhancer (OSE)-mediated tethering of active *MYC* alleles to NPCs, consequently enhancing export of *MYC* mRNAs as compared to normal colon epithelial cells[147]. There is also existing data indicating that similar processes might be present in lung cancer tissue, as a TPR-dependent and NXF1-mediated nuclear export of tRNAs has been described in yeast and lung cancer cell lines, and knockdown of TPR in the lung cancer cells caused inhibition of nuclear export of tRNA, cell growth, and protein synthesis[156]. Nup98 has also been shown to play a role in oncogenesis, as a recent study uncovered that overexpression of Nup98 in aggressive breast cancer unblocks nucleocytoplasmic transport of several transcription factors, including β-catenin, and promotes metastasis in mice[157].

Thus, although gene gating first emerged as a hypothesis almost four decades ago, it is only recently that it has been described in human cells. NPCs and Nups show a vast variety in compositions and cellular functions, which have been implied in both development and cancer, with recent studies providing valuable clues to how malignant cells may benefit from recruitment of active genes to the nuclear envelope or an enhanced nucleo-cytoplasmic transport. Emerging data also supports the notion that gene gating might be influenced by factors such as metabolic states, PTMs and external stimuli, hence adding a probabilistic component to the regulation and intricacy of gene gating. However, many questions remain unanswered, such as how wide the extent of this phenomenon is in different cell types and genomes.

1.3 CHROMATIN CROSSTALK IN 3D

While the nuclear architecture, chromatin compartmentalization and NPC/Nup-mediated processes play important roles in transcriptional regulation and development, yet another layer of complexity and control is added through transient interactions of chromatin fibres in *cis* and *trans*, with *cis* interactions being more frequent and principally based on chromatin looping, whereas *trans* interactions rely on chromatin fibre bridging[158].

1.3.1 CTCF: the master organizer of 3D chromatin conformation

Spatiotemporal regulation of chromatin orchestration requires architectural proteins and 3D genome organizers, such as cohesin, condensin, CTCF and PARP1. CTCF is an evolutionarily conserved transcription factor with house-keeping functions, and was initially discovered as a regulator of *c-Myc* expression[159]. Due to its 11 zinc-fingers, CTCF is able to bind directly to DNA in a versatile manner[160]. However, this ability can be regulated through DNA methylation and post-transcriptional modifications, such as PARylation or SUMOylation[161-163]. Interestingly, CTCF is able to inhibit DNA methylation of its binding sites by interacting with PARP1, due to the ability of PARP1 to interfere with the activity of DNA methyl transferase-1 (DNMT1)[164]. Moreover, CTCF-binding motifs are commonly found in enhancers, and CTCF may promote transcriptional activation through its capacity to interact with the RNA pol II complex[165,166]. CTCF is, however, perhaps most well-known for its functions as an insulator protein and the ensuing blocking of enhancer-promoter interactions[161]. Yet, as CTCF, with or without the cohesin complex, also allows for establishment of chromatin loop formation, CTCF may likewise contribute to long-range chromatin interactions between enhancers and promoters within TADs, thus positively regulating transcriptional activity[73]. Moreover, CTCF is involved in the demarcation of TAD borders, which are enriched in CTCF-binding motifs[167]. In conclusion, CTCF holds important functions related to the regulation of chromatin organization and interactomes.

1.3.2 Chromatin interactomes

1.3.2.1 Capturing chromosome conformation and the 'C'-family of technologies

The 'Chromosome Conformation Capture (C)-family' of techniques are based on the original Chromosome Conformation Capture (3C)-technique developed by Dekker *et al.* in 2002 [168], and which paved the way for the development of a large number of technologies that complement the microscopy based methods, such as 3D DNA FISH and super-resolution microscopy, to study chromosomal structures[169]. In brief, the 3C method encompasses initial formaldehyde (FA) cross-linking of chromatin, followed by restriction enzyme digestion and subsequent re-ligation under conditions that promote intra-molecular ligation between the covalently joined DNA fragments. Subsequently, the chimeric DNA fragments can be used for semi-quantitative PCR to visualize chromatin interactions in a one *vs.* one manner[168]. There has been a considerable and rapid progression in the field of the "C-family" of technologies over the past two decades, giving rise to 3C-derived assays such as 4C (Circular Chromosome Conformation Capture), 5C (Chromosome Conformation Capture Carbon Copy), Hi-C, and scHi-C (single cell Hi-C), that have provided unprecedented possibilities for exploring mechanistic connections between nuclear biology and chromatin structure[169-171]. The Hi-C technique is based on the 3C assay, and incorporates a biotin-labelled nucleotide at the ligation junction, enabling enrichment of the chimeric biotin-labelled junctions during library preparation before implementing deep next generation sequencing (NGS),

consequently envisaging "all *vs.* all" chromatin interactions[170]. Hi-C has thus uncovered several features of the folded genome, and the immense amounts of data generated by the ever-growing C-family of technologies and other capture-based methods have opened doors to novel ways to explore the structural orchestration and functions of the genome.

1.3.2.2 Topologically associated domains and the loop extrusion model

Mapping of genome-wide chromatin interactomes by Hi-C has shown that active regions of the genome preferentially interact with other active regions, and transcriptionally inactive chromatin with other repressed regions, in a mutually exclusive manner correspondingly representing functional A and B compartments described earlier in the introduction[52,142]. In 2012, two studies exploiting 5C and Hi-C were published back-to-back by Nora *et al.* and Dixon *et al.* respectively, and uncovered that at a shorter scale (< 1Mb) chromosomes fold into domains of 200 kb-1Mb in size, which showed preferential intra-domain interactions[67,172]. The former study revealed that the *Xist* locus was distributed into such self-interacting regions and named these "topologically associated domains" (TADs), while the latter study, utilizing Hi-C and deep (for the time) sequencing with binning of ca 40kb fragments to generate contact matrices, documented self-interacting regions that were conserved amongst different tissues, and also between human and mouse[67,172]. Thus, while A and B compartments preferentially interact with other regions that share similar transcriptional states, TADs represent insulated regions in which chromatin fragments favourably interact with other segments localized within the same region[142,173]. However, stochastic clusters of contacts can occur between TAD boundaries, although they average into TADs[142,173]. Thus, the precise definition of TADs is ambiguous, partially due to technical resolution constraints and the complexity of sub-megabase interaction patterns, although a functional classification has been proposed to comprise domains whose boundaries are most conserved during cell differentiation[169,174,175]. Additionally, as the resolution of several methods has increased, the classification TADs have been able to be further divided into smaller sub-TADs, supporting the concept of a hierarchically arranged genome[176,177].

TADs and functional A/B compartments may thus overlap and share certain features, such as overall transcriptional states[55]. However, most evidence seem to point to distinct mechanisms underlying TAD and A/B compartment formation, where the stratification into epigenomic A/B compartments are formed from both intra- and interchromosomal interactions with other compartments sharing similar transcriptional and structural states, whereas TADs represent local, *cis*-interacting domains[178]. While the mechanisms of TAD boundary formation are not fully understood, it is currently best depicted by the chromatin loop extrusion hypothesis[178-180]. This widely accepted model depicts how chromatin loops are mediated by *cis*-acting loop-extruding factors (LEFs) such as cohesin (proposed to be able to use ATPase activity to extrude chromatin loops) and CTCF[181,182]. Accordingly, chromatin is extruded bidirectionally through a cohesin ring which, together with forces such as transcription induced supercoiling, facilitates a progressively growing chromatin loop, eventually stalling at TAD-borders upon encountering CTCF-occupied insulator DNA elements arranged in a forward-reverse

orientation[181-184]. Such convergent CTCF-binding sites (CTCFBSs) demarcate the vast majority (>80%) of TAD borders in mammalian cells, and function as contact points for cohesin, thus mediating long-range chromosomal *cis*-interactions[67,185-187]. However, in a recent study by Dequeker *et al.* it was intriguingly shown that loop extrusion might be impeded before reaching CTCF. Using scHi-C, the authors observed that loading of the minichromosome maintenance (MCM) complex reduces CTCF-anchored loops and decreases TAD-boundary insulation in mouse zygots. This finding was extended also to human HCT116 cells where MCMs affect the number of CTCF-anchored chromatin loops and gene expression, thus suggesting that MCM functions as a barrier that restricts loop extrusion[188]. Nevertheless, the loop extrusion model is supported by the observation that depletion of cohesin abrogates close to all TADs and chromatin loop domains, while loss of CTCF affects many, but not all, TAD boundaries and loop domains[179,187,189]. Moreover, a recent study demonstrated that mutation of zink finger protein 8 (ZF8) of CTCF impairs its chromatin residence time, consequentially causing widespread weakening of TADs, aberrant gene expression and increased global DNA methylation[190]. Additionally, deletion or inversion of CTCFBSs may merge TADs, or establish new TAD boundaries, which further supports the notion of the loop extrusion model[182,184,191,192]. Thus, CTCF and cohesin act in concert to regulate chromatin looping and TAD structures, and cohesin has been shown to be positioned in the mammalian genome by CTCF, transcriptional activity, and the cohesin-unloading factor Wapl[185]. Interestingly, in CTCF-depleted fibroblasts, cohesin is not properly recruited to CTCFBSs, instead accumulating at transcription start sites (TSSs), whereas in cells depleted in both CTCF and Wapl, cohesin accumulates at the 3' ends of active genes[185]. Moreover, stabilization of cohesin by knockout of Wapl entails reduced intra-TAD interactions as well as caused an accumulation of interactions at TAD boundaries, hence implying that the dynamic properties of cohesin-mediated loop extrusion are essential for TAD formation[193].

TADs have been found to be highly conserved amongst species, and a study by Wang *et al*. found several characteristics of TADs conserved in mammalian and *Drosophila* genomes[194]. Moreover, the authors used a high restriction site-limited map resolution (~200 bp) for visualization of Hi-C reads, and found a significantly higher number of TADs in *Drosophila* (cells arrested in G1/S) than previously documented, interestingly, uncovering that the *Drosophila* genome is fully partitioned into contiguous TADs[194]. Contrastingly, a recent study from Gabriele *et al.* was able to visualize chromatin looping at the *Fbn2* TAD in mECSs by employing super-resolution live-cell imaging and fluorescent labelling of the CTCFBSs at each side of the *Fbn2* TAD, revealing that the looping was highly dynamic and rare, with a looped fraction of only 3-6.5% and an average loop life time of 10-30 minutes[195]. Taken together these results imply that TADs visualized with non-live imaging microscopy-based methods or by spatial proximity heat maps only represent a snapshot of the reality of all possible *bona fide* chromatin interactions at single cell level[195]. The reality of chromatin crosstalk in individual cells is likely highly complex, and interaction mapping or traditional microscopy-based methods showing TADs likely represent highly dynamic, potential

interactions within a cell population, rather than actual events. Indeed, such stochastic aspects have also been observed when comparing results from Hi-C with single-nucleus Hi-C where TADs in single cells are seemingly non-present, however, upon averaging of several single cells, TAD structures began to emerge[173]. A recent study which used high resolution snHi-C in order to construct maps of individual *Drosophila* genomes uncovered that up to 40% of TAD borders are preserved between individual nuclei, and are occupied by a high number of active histone marks[196]. However, long-distance interactions showed a prominent variation between individual cells, suggesting an important role of stochastic processes in the folding of the fly genome[196].

Given their role in genome organization, TADs are also implicated in development and transcriptional regulation. For instance, during the maturation of mouse oocytes, a decrease in TAD, loop and compartment strength has been observed, possibly due to transcriptional silencing and visual detachment of chromatin from the nuclear envelope[173]. TADs also hold a role in transcription regulation, by functioning as regulatory frameworks that facilitate enhancer-promoter interactions and prevent ectopic enhancer activation[175]. In HCT116 cancer cells, cohesin-mediated TAD formations were abolished after acute depletion of the cohesin subunit RAD21 (also known as Scc1) by using an auxin-inducible degron system[179]. The authors also employed precision nuclear run-on sequencing (PRO-Seq) to examine the effects of cohesin depletion on nascent transcription, and intriguingly found a strong down-regulation of genes near super-enhancers (SEs), whereas no widespread ectopic activation could be observed[179]. Additionally, a study by Rhodes *et al.* found that in mESCs, cohesin disrupts polycomb-dependent chromosomal interactions to modulate gene expression, with observed repression at polycomb target genes with increased interactions[197]. However, these studies only showed a modest effect on global transcription, suggesting that the presence of TADs might play a partial role in transcription control, with other factors contributing to its maintenance. Interestingly, disruption of TAD boundaries has proven to entail more profound effects on gene expression. Removal of a CTCF-cohesin boundary upstream of a 80 kb sub-TAD containing the mouse α-globin gene cluster in erythroid cells extends the sub-TAD upstream to contiguous CTCF-cohesin binding sites, and caused strong (up to 10-fold) ectopic transcriptional activation of adjacent genes[198]. Moreover, dramatic effects on global transcription were observed after deletion of the cohesin-loading factor *Nipbl* (yeast homologue *Scc2*) in mouse hepatocytes, where approximately a 1000 genes were found to be either upregulated (487) or downregulated (637) [199]. A recent study also documented how the positioning of *Hoxd* genes in certain TADs are crucial for enhancer-mediated gene expression during limb development[200].

Consequentially, these interactomes, although constrained by TAD boundaries, provide a vast diversification of transcriptional activity and expression patterns, and play a significant role during development and the maintenance of particular phenotypes[201].

1.3.2.3 Cis-regulatory elements and super-enhancers

Chromatin interactions are mainly represented by *cis*-regulatory elements such as enhancers, promoters, silencers and insulators, which play important roles in the formation of chromatin loops. Promoters have traditionally been defined as DNA elements where transcription is initiated, with the transcription start site (TSS) sometimes referred to as the "core promoter", while enhancers consist of a couple of hundred bps long non-coding DNA fragments that function as positive transcriptional regulators by acting as platforms for binding of transcription factors (TFs) and coactivators, subsequently delivering these factors to target promoters to assist the formation of the pre-initiation complex, and loading of RNA polymerases[202,203]. Enhancers and promoters share particular similarities, though enhancers are characterized by enrichment of certain chromatin modifications such as H3K4me1 that is present at lower levels at promoters, H3K4me3 – a mark also found at active promoters but at significantly higher levels –, H3K4me2 and binding of RNA pol II, which occupy promoters and enhancers alike[204,205]. Moreover, enhancers exhibit DNAse I hypersensitivity, and the level of the active histone mark H3K27ac can be used to discriminate between poised and disengaged enhancer states[206,207]. Certain enhancers that have co-bound JMJD6 and BRD4 have also been shown to be involved in the release of RNA pol II in promoter-proximal pausing to allow a transition to transcriptional elongation[208]. Given that enhancers determine cell type- and differentiation stage-specific gene expression patterns[201], an important question concerns the identity of the factors that determine the mechanism of gene targeting by specific enhancers. In lower organisms, physical proximity of regulatory elements along the linear chromosome poses the mechanism of E-P specificity and regulation of global gene expression patterns. In *S. cerevisiae*, the vast majority of upstream activating elements are thus located only a few ~100 bps from their target promoters[209]. However, in the genomes of *Drosophila* and mammals, enhancers are located much further away (in general up to several kbs) from their target promoters. Unlike promoters, enhancers are thus able to exert their regulatory functions over target genes over large distances, ranging from hundreds up to thousands of kbs in length, and independently of the enhancer's positioning and distance relative to its target gene[201,202]. It has been estimated that in *Drosophila* and mammals only 79-88% and 27-60% of enhancers, respectively, contact the nearest gene promoter[210-213]. However, local chromosome folding into TADs may limit the contact potential of enhancer-promoter (E-P) interactions, and examples of how local gene expression is altered by TAD structures in a manner consistent with E-P specificity have been well documented[198,214,215]. A third mechanism that contributes to E-P interactions and specificity in eukaryotes is through specific motifs at promoters; as certain enhancers may preferentially interact with particular core promoter elements (CPEs), which consist of ~6-12 bp long sequences surrounding TSSs to facilitate binding of TFs, or with "tethering elements" located proximal to the promoter[29,216-218]. Taken together, these findings indicate that E-P specificity is determined by a combination of factors including chromosome topology, linear proximity and core promoter identity, which is facilitated through several local promoters competing for the activity of a particular enhancer.

The “active chromatin hub” hypothesis proposes that several active enhancers are able to simultaneously contact genes, while inactive regulatory elements and genes are looped out from the physical interactions of such enhancer hubs[219]. Hence, not only Hi-C experiments performed on large cell populations but also targeted multi-contact 4C (MC-4C) analyses designed to detect multi-way interactions between enhancers and promoters, support the notion of active enhancer hubs[220], for example at the beta-globin locus. Likewise, the Tri-C method, which also detects multi-way 3C concatemers, has uncovered preferential multi-way E-P interactions within the alpha-globin cluster of an erythroid-specific sub-TAD[221]. However, as the chromatin interactions detected within the TAD that contains the alpha-globin locus were highly variable, the authors contemplate that such heterogeneity could reflect a dynamic process underlying contacts formed by chromatin loop extrusion[221]. Moreover, as multi-way contacts have been found to comprise rare events in chromatin networks by DNA FISH experiments[82], it is thus possible that the enhancer hubs detected by the “C” techniques might represent rare events, an ensembled snapshot of transient and stochastic contacts occurring at a level of bigger cell populations, rather than actual events on single-cell level. However, the generality and inner workings of hub formation are not fully understood.

The amount of enhancers in the genome is vast, with a total number estimated to be >1 million, with several thousands of active enhancers in specific cell types[222]. Moreover, enhancers are able to form large clusters, termed “super-enhancers” (SEs) (for a schematic comparison between enhancers and SEs, please see **Figure 2**), that span tens to hundreds of kbs, and that are characterized by high levels of the Mediator subunit Med1, BRD4, p300, cohesin, LSD1-NuRD complexes and master TFs such as OCT4, NANOG and SOX2 [223,224]. SEs frequently regulate expression of pluripotency and cell fate-determining genes, with generally higher transcriptional activity as compared to genes which are regulated by normal enhancers[223]. In addition, SEs more frequently bind terminal TFs of the Wnt, TGF-β, and leukemia-inhibitory factor (LIF) pathways, with SE-driven genes also being more sensitive to perturbations of associated enhancer-binding transcriptional regulator genes than genes driven by typical enhancers[225,226].

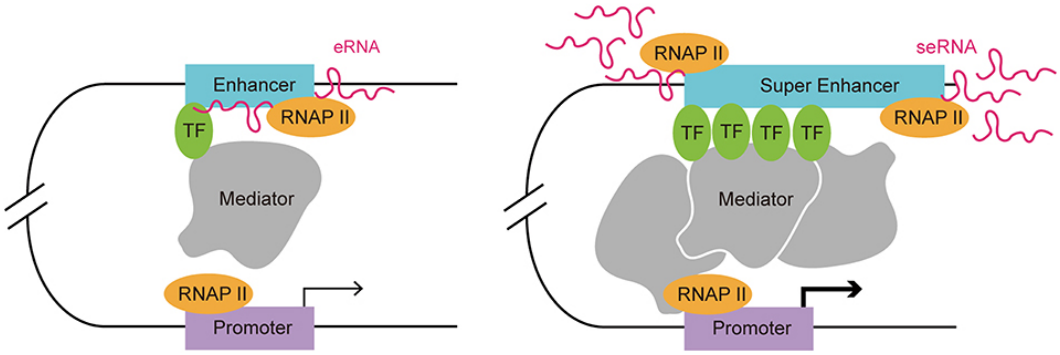

Figure 2: *Schematic comparison between a typical enhancer and super-enhancer. Reprinted under the Creative Commons Attribution 4.0 International License, https://creativecommons.org/licenses/by/4.0/, by Frontiers in Oncology, 2019;9:1307. Wu M., Shen J. From Super-Enhancer Non-coding RNA to Immune Checkpoint: Frameworks to Functions[227].*

Since SEs were first described approximately a decade ago, their involvement in cancer development and progression has been increasingly recognized. Thus, it has been suggested that acquisition of SEs might play a role in the development of complex diseases such as malignancies, by cells acquiring SEs at regions with functions related to the hallmarks of cancer: angiogenesis, proliferation, apoptosis, cell growth, and factors important for invasiveness and metastasis[224,228]. For instance, vast oncogenic SEs in the gene desert adjacent to the well-known *Myc* oncogene have been found in several cancers such as colon cancer, lung cancer, multiple myeloma, and acute lymphoblastic lymphoma (ALL), but not in non-cancerous cells of the same cell type as the proposed origin of the cancer[224]. Moreover, a recent study that employed genome-wide profiling using ChIP-seq, RNA-seq and whole genome sequencing in colorectal cancer paired patient tissues was able to identify two oncogenic SEs (OSEs) (*PHF19* and *TBC1D16*) in the cancer cells, but not in the paired cells from adjacent tissue from the same patients, thus providing important clues to the molecular mechanisms by which SEs are formed and contribute to oncogenesis[229]. Further clinical evidence of the role of SEs in malignancies exist, as analysis of the SE landscape of 66 patients with acute myeloid leukaemia (AML) intriguingly uncovered 6 novel epigenetic subtypes with associated regulatory loci, of which one expressed a RARα dependency that proved to be targetable by a selective RARα agonist[230]. These results further underline the complexity of SE characteristics in oncogenesis, in concordance with other studies that have characterized SE circuitries in triple-negative breast cancer, neuroblastoma, and ovarian cancer[231-234].

Somatic and germline cells seem to acquire oncogenic SEs through several different mechanisms such as genomic duplications, insertions, deletions, inversions, translocations and single nucleotide polymorphisms (SNPs), which may cause alterations of SE copy numbers, TF-binding sites or spatial alterations of SE location by structural changes in 3D chromatin organization, consequently causing changes of gene expression patterns that drive oncogenesis[225]. For instance, both genetic and epigenetic perturbations of TAD boundaries permit new enhancer/SE-promoter contacts that may drive cancer development and progression[235,236]. Hence, TAD borders that function to insulate E-P interactions, thereby preventing ectopic enhancer-mediated activation, insulate typical enhancers and SEs alike[225]. Accordingly, SE functions are also associated with spatial chromatin organization. For instance, cohesin depletion by an auxin-inducible degron system targeting RAD21 causes SEs to co-localize in HCT116 cells. In these cells, SEs thus form hundreds of intra- and interchromosomal contacts, strengthening the links between distant SEs, which is in contrast to the situation in the presence of cohesin, where the majority (68.7%) of SEs were positioned in the interior of cohesin-associated loops and long-distance cohesin-independent contacts were much weaker[179]. Interestingly, a recent study found that a SE encompassing the *Prdm14* pluripotency gene in mESCs mediated recruitment and deposition of cohesin onto its flanking CTCFBSs. Moreover, the insulation of the SE was not strictly controlled by looping, as inversion and deletion of the CTFBSs of the left border and right border respectively did not cause upregulation of the adjacent *Slco5a1* gene[237]. Additionally, the authors discovered that

long-range activation of *Slco5a1* by the *Prdm14* SE benefitted from CTCF recruitment to its promoter, although it did not solely depend on it. Finally, the *Prdm14* SE appeared to form a "hub" that is able to simultaneously contact both the *Slco5a1* and *Prdm14* genes, in concordance with previous evidence that SEs are able to contact multiple genes at once. For example, in the case of the mouse β-globin SE neighbouring chromatin domain loops were documented to form rosette-like structures through collision of their CTCF-bound anchors, thus creating enhancer hubs that can accommodate more than one gene[220,237].

1.3.2.4 Contribution of non-coding RNAs to chromatin crosstalk

Interestingly, non-coding enhancer DNA sequences are often transcribed by RNA pol II to produce short non-coding RNAs (eRNAs), which have been used to infer enhancer activity[238,239]. Although the eRNAs' precise functions and mechanisms of action remain not fully understood, there is evidence suggesting a role in the stabilization of chromatin looping and interactions in both *cis* and *trans*, thus also implying a role in transcriptional regulation[238,240,241]. Accordingly, there are several examples of how eRNAs and seRNAs interact with either proteins, DNA or RNA to regulate gene expression in both *cis* and *trans*. In myoblasts, the master TF MyoD induces expression of long non-coding RNAs (lncRNAs) from SEs (seRNAs) during myogenic differentiation, thereby regulating the expression of target genes in *cis* through their interaction with heterogeneous ribonucleoprotein L (hnRNPL)[242]. Additionally, the seRNA *PAM* has been found to control skeletal muscle satellite cell proliferation and aging through regulation in *trans* of *Timp2* expression[243]. Moreover, eRNAs are able to interact with CBP or BRD4 within TADs in a localized manner, and BRD4-eRNA interactions increased BRD4-binding to acetylated histones *in vitro*, while also enhancing enhancer recruitment and transcriptional cofactor activities to regulate gene expression[244]. Moreover, the *lincRNA-p21* has been found to regulate transcription in both *cis* and *trans*, as it acts in *trans* through mediating the recruitment of heterogeneous nuclear ribonucleoprotein K (hnRNPK) to its target promoter[245], and is able to activate *Cdkn1a* in *cis*[246]. Taken together, these findings suggest an interesting additional function of enhancers and SEs in their control of gene activity, consequently bridging intra- and interchromosomal transcriptional regulation.

1.3.2.5 Interactions in trans

Most chromatin interactions are represented by intrachromosomal contacts, as has been observed with several chromatin conformation techniques. However, these assays have also confirmed the existence of interactions in *trans*, previously also observed by microscopy-based methods, although such long-range contacts tend to consist of rare and dynamic events[247]. One of the first prominent examples of functional mammalian interchromosomal contacts was described in 2005, and depicts interactions between the promoter of the IFN-γ gene on chromosome 10 and regulatory regions of the Th2 (T-helper 2) cytokine locus on chromosome 11. This interaction was regulated by DNAse I hypersensitivity sites present during development at the Th2 LCR (Locus Control Region) and affected the dynamics of IFN-γ expression upon Th1 cell differentiation [248]. The authors observed that the interactions

seemed to exhibit a cell-type-specific dynamic, with interchromosomal interactions being lost in favor of intrachromosomal contacts upon transcriptional activation, thus suggesting a function of *trans*-interactions in the coordination of gene expression[248]. Later studies support the notion of cell-type-specific dynamics of chromatin interactomes, especially in the case of long-range intrachromosomal as well as interchromosomal interactions, which have been shown to be more frequent in mouse sperm as compared to zygots and somatic cells[249]. Moreover, stochastic expression of the interferon-β gene *IFBN1* upon viral infection requires interchromosomal contacts between particular Alu-like DNA elements termed NRCs (NF-kB reception centers), which capture and deliver NF-kB to the *IFBN1* promoter through stochastic *trans*-interactions that are mediated by cooperative DNA binding by NF-kB and the mammalian GAGA-binding protein ThPOK[250].

While the above examples depict non-homologous chromosomal contacts, another well studied example of transcriptional regulation by long-range interactions in *trans* is transvection, which has been extensively studied in *Drosophila*. Transvection is an epigenetic phenomenon that allows pairing of maternal and paternal homologs, which may permit enhancers on homologs to contact the promoter of the other, thus positively regulating gene expression[251,252]. A study by Lim *et al.* which visualized transvection in living *Drosophila* embryos intriguingly also discovered that a shared developmental enhancer was able to coactivate a *cis*-linked PP7-reporter gene while simultaneously activating a MS2-*lacZ* reporter gene on the other homolog[253]. Such coactivation was found to be promoted by insulating DNA sequences, which stabilized the homolog-pairing, but did not increase its frequency[253]. In addition to being involved in transcriptional regulation, transvection has also been implied in control of 3D genome orchestration, as different kinds of homolog-pairing, tight pairing (that span contiguous small domains) *vs.* loose pairing (consisting of larger single domains), are associated with different kinds of chromatin compartments, where tight pairing was correlated with transcriptionally active A compartments[254]. Although transvection is most well-studied in *Drosophila*, the phenomenon also exists in mammalian organisms. The first example of non-allelic transvection in mammals was discovered at a genomically imprinted region at the *H19-Igf2* locus[255], where the *H19* ICR (Imprinting Control Region) regulates monoallelic expression of *H19* and *Igf2*[247,255]. CTCFBSs within the *H19* ICR thus not only determine physical proximity among imprinted regions, but also transvect allele-specific epigenetic states to the interacting non-allelic ICRs during germline development[247,255].

In summary, long-range chromatin interactions in *trans* comprise dynamic and rare occurrences, which add another layer of complexity to regulation of transcription through their stochastic nature. These interchromosomal interactions provide means for fine-tuning gene expression patterns, and seem to preferentially occur in mammalian organisms during developmental windows of opportunity[247]. Thus, it is not surprising that chromatin crosstalk - both in *cis* and *trans* - and particular regions involved in such contacts have been implied in several diseases, such as cancer.

1.4 THE ROLE OF STOCHASTICITY AND 3D GENOME ORGANISATION IN CANCER DEVELOPMENT

Cancers constitute a vast group of heterogeneous diseases, where malignant cells may develop distinct mechanisms to provide selective advantages driving clonality and disease progression. Whereas cancer has initially been considered a genetic disease, recent findings have highlighted that epigenetic aberrations play prominent roles in its pathophysiology[256]. Epigenetic mechanisms thus protect and stabilize gene expression patterns and cellular states while also facilitating pertinent responses to environmental and developmental cues[8]. However, chromatin homeostasis can be perturbed by genetic, metabolic or environmental insults that may cause alterations entailing either more compacted and repressed chromatin - or overly permissive chromatin states. In the latter case, the excessively permissive chromatin results in epigenetic plasticity, which in turn allows for stochastic activation of alternate gene expression patterns, of which a few might confer fitness under changing selection pressure and become drivers of disease, such as cancer[4].

1.4.1 Regulation of *MYC* expression in 3D

Myelocytomatosis proto-oncogene (*MYC*, also known as "*c-MYC"*) is a gene encoding the MYC oncoprotein, which functions as a master regulator of transcription and as a potent driver of transformation[257]. *MYC* plays a crucial role in many cancers, and is elevated or dysregulated in up to 70% of all human malignancies, and has therefore received considerable attention in scientific studies[258-260]. However, *MYC* is considered a pharmacologically "undruggable target"[259,261], consequentially shifting the attention of the research to epigenetic mechanisms regulating *MYC* expression in cancer cells, as well as downstream signalling pathways of MYC, in order to explore potential therapeutic strategies.

1.4.1.1 The MYC "enhancer-ome"

Although specific oncogenic SEs are the main regulators of *MYC* expression in several tumour types[224,258], the complexity of *MYC* expression is further increased by its intricate chromatin fibre interactome containing numerous regulatory elements. The *MYC* locus itself is thus positioned within a TAD corresponding to a ~3 Mb region at the human chromosomal band 8q24, with several sub-TADS demarcated by CTCFBSs that promote intra-TAD interactions[259,262]. This region contains enhancer clusters on both sides of *MYC,* mediating tumour-specific chromatin loops between *cis*-regulatory elements[260]. Transcription of *MYC* is mediated by four alternative promoters termed P0, P1, P2, and P3, as well as by several of the bilaterally flanking enhancers located in the same TAD that *MYC* is embedded in and in neighbouring TADs[260]. A region located 142-115 bps upstream of P1 is also of particular interest, as this nuclease hypersensitivity element III_1 ($NHEIII_1$) constitutes a negative regulator of *MYC* expression, which has been found to account for up to a staggering 90% of

the transcriptional control of *MYC* through a mechanism encoded by the genome itself *via* formation of a G-quadruplex (G4) structure[263-265].

As discussed above, CTCF is a well-studied TF and 3D genome organizer, often called "the master weaver of the genome", that is involved in mediating chromosomal contacts that regulate gene expression. There are several CTCFBSs upstream of *MYC,* of which one specific conserved CTCFBS positioned 2 kb upstream of the locus functions as an important enhancer-docking site that is necessary for enhancer-promoter loop formation[260]. Genome-wide association studies (GWAS) have shown that many SNPs positioned in the gene desert at Chr:8q24 are associated with aberrant *MYC* expression and susceptibility to malignancies in the breast, prostate, oesophagus, ovaries, colon and pancreas, due to perturbations of enhancer-promoter interactions[262,266-268]. Moreover, there is a known SE located 1.7 Mb downstream of *MYC*, which plays a pivotal role in haematopoiesis and in leukaemia stem cell (SC) hierarchies, with disruption of this particular SE in mice entailing a loss of *Myc* expression in haematopoietic SCs coupled with an accumulation of differentiation-arrested progenitor cells[269]. Interestingly, the SE, which has been named "blood enhancer cluster" (BENC), is conserved in mice and humans, and is able to cause leukaemia through altered *MYC* expression[259,269]. Thus, SEs have been commonly associated with *MYC*-driven malignancies, and transcriptional dysregulation of *MYC* in cancers is often accomplished through acquisition of a SE located in the 2.8 Mb *MYC* TAD. A study from Schuijers *et al.* intriguingly found that cancer cell-type-specific SEs in colorectal cancer, breast cancer, AML and prostate cancer loop to a common enhancer-docking site **(Figure 3)**, containing a conserved CTCFBS, at the *MYC* locus[270]. Moreover, genetic and epigenetic perturbation of the enhancer-docking site using CRISPR/Cas9 and dCas9/DNMT3A-3L systems decreased CTCF-binding to the enhancer-docking site, and reduced SE interactions, *MYC* transcription and cell proliferation, thus describing a mechanism by which oncogenes are able to hijack SEs[270]. Interestingly, the finding of a common enhancer-docking site extended to other genes as well, of which several were also found to be involved in cancer[270].

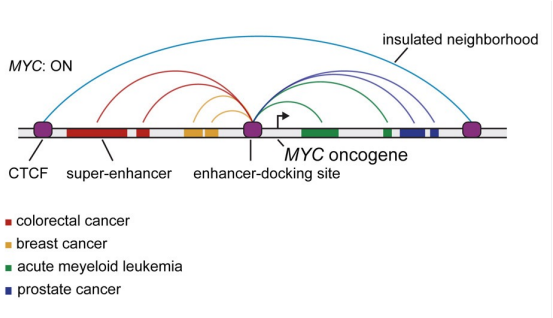

***Figure 3**: Overview of tumour-specific SEs that regulate MYC gene expression. Adapted from Schuijers J, Manteiga JC, Weintraub AS, et al. Transcriptional Dysregulation of MYC Reveals Common Enhancer-Docking Mechanism. Cell Rep. 2018;23(2):349-360, doi:10.1016/j.celrep.2018.03.056. [270] Reprinted under the Creative Commons Attribution 4.0 International License, https://creativecommons.org/licenses/by/4.0/*

1.4.1.1 Control of MYC transcription by lncRNA

Concurrently, lncRNAs are also able to control the expression of *MYC*. For instance, the Colon Cancer Associated Transcript 1 (CCAT1) is a 2628 nt long transcript, located in a CRC-specific SE 500 kb upstream of *MYC*, that was first identified as a CRC biomarker, and which is consistently strongly expressed in colorectal adenocarcinoma while being largely undetectable in normal tissue[224,271]. Intriguingly, in HCT116 cells a nuclear retained isoform of this transcript, named CCAT1-L, interacts with CTCF to facilitate *MYC* transcription by long-range chromatin interactions between the colorectal SE and the *MYC* promoter[272]. The same study also found that overexpression of CCAT1-L promoted tumorigenesis through enhanced expression of *MYC,* while depletion of CCAT1-L reduced MYC expression[272]. CCAT1-L also has shorter isoforms, of which the promoter of *Cancer-Associated Region lncRNA-5 (CARLo-5)* interacts with a *MYC* enhancer region and is involved in tumorigenesis by regulating cell cycle progression through inhibition of *CDKN1A* mRNA levels, which is a critical regulator of G1 arrest[273]. Moreover, similar to CCAT1-L, other lncRNAs have been found to interact with CTCF. In pancreatic ductal adenocarcinoma (PDAC) the lncRNA LINC00346 interacts with CTCF, which prevents CTCF binding to the *MYC* promoter to subsequently reduce CTCF-mediated repression of *MYC*[274]. The same study also found that knockout (KO) of LINC00346 caused impaired proliferation, migration, tumorigenesis and invasion ability, and that these phenotypes could be restored upon re-expression of LINC00346 in KO PDAC cells through rescue-experiments[274].

Taken together, *MYC* exerts pleiotropic effects in oncogenic transformation to control gene expression patterns, and although *MYC* has been extensively studied, many aspects of its regulation as well as mechanism of action in oncogenesis remain poorly understood [275]. An important future direction in cancer treatment involves targeting the cancer-specific mechanisms of *MYC* over-expression in tumours. For example, the finding that the OSE-mediated gating of *MYC* is specific for HCT116 cells and is absent from HCECs suggests that regulators of this process, such as AHCTF1-binding to the OSE, might serve as potential novel therapeutic targets[147]. The extensive *MYC* – enhancer interactome, however, also highlights the complexity of *MYC* regulation in 3D, which likely provides opportunities for cancer evolution and resistance development under changing selection pressure, such as, for example, under the inhibition of gene gating. To efficiently target *MYC* over-expression in tumours, it will thus be necessary to decipher the mechanism underlying the division of labour among the potential regulatory elements *MYC* might contact and the mechanism underlying the formation of its interactome. It has already been shown that while functional OSE-*MYC* proximity is highest at the nuclear periphery, another enhancer called enhancer D that displays no NUP binding likely regulates *MYC* transcription in the interior of the nucleus[147]. This is in keeping with the observed two waves of *MYC* mRNA maturation with one peaking at the nuclear periphery while the other in more internal positions[147]. It will thus be necessary to understand also the dynamics of chromatin fibre mobility that likely underlies the formation of the ensemble *MYC* interactome in a cell population, as well as identifying the signalling pathways that might regulate its function. In this respect, the WNT pathway

seems to be particularly important due to the prominent presence of WNT responsive elements among the enhancers of *MYC*[276,277] as well as its impact on the OSE-mediated gating process[147].

1.4.2 WNT-signalling

MYC is not only one of the most frequently overexpressed genes in cancer cells, but it is also intricately connected in a network comprising several signalling processes such as signalling pathways related to cellular growth and apoptosis[278,279]. Growth factors are thus able to upregulate *MYC* expression, of which factors of the WNT/β-catenin pathway are arguably amongst the most famous[278,280].

The WNT signalling pathway comprises an integral cell-to-cell signalling hub that is evolutionarily conserved and which is essential for maintenance of tissue homeostasis through its coordination of differentiation, cell proliferation, cell polarity, cell motility and stem cell renewal[281]. Unsurprisingly, abnormal WNT signalling is thus involved in a variety of disorders including autoimmune diseases, diabetes, embryonic anomalies, and cancers[282]. The WNT pathway is divided into two classes: the first is the β-catenin-dependent pathway, which is also referred to as the canonical pathway (**Figure 4**), while the second branch consists of β-catenin-independent signaling also known as the non-canonical pathway[283]. The non-canonical pathway is further stratified into two groups: WNT/planar cell polarity (PCP) and WNT/calcium (WNT/Ca^{2+})[284]. In brief, the canonical WNT signalling pathway is primarily involved in the control of cell proliferation, whereas the non-canonical pathway regulates cell polarity and motility[283]. As the work included in the current thesis mainly concentrates on the canonical pathway, the focus will lie on this particular branch of WNT signalling.

To activate the WNT signalling cascade through the canonical pathway, secreted WNT glycoproteins bind to a transmembrane receptor protein named Frizzled (FRZ) and low-density lipoprotein receptor-related protein 5 and 6 (LRP5 and LRP6) at the plasma membrane[281,285]. In the absence of WNT proteins, a destruction complex including tumour suppressors such as adenomatous polyposis coli (APC), Axin, casein kinase 1 (CK1), and glycogen synthase kinase-3β (GSK3β), degrades β-catenin through phosphorylation of its key serine (Ser) and threonine (Thr) residues (Ser33, Ser37, Ser 45, and Thr41), resulting in ubiquitination and subsequent proteasomal degradation[285,286]. Contrariwise, the presence of WNTs activate FRZ and LRP5/6 plasma-membrane receptors to permit binding of DVL proteins, thus causing phosphorylation of cytoplasmic motifs of LRP5 and LRP6, which in turn mediates interaction with Axin and subsequent destabilization of the β-catenin destruction complex [281,285]. Unphosphorylated β-catenin is then translocated into the nucleus where it interacts with transcription factor/lymphoid enhancer-binding factor (TCF/LEF) family proteins to initiate transcription of WNT-target genes, such as *MYC*[281,286,287]. Moreover, a subset of the TCF/LEF target genes are constitutively inhibited by nuclear TCF

proteins which recruit transcriptional corepressors to WNT response elements (WREs)[288]. Upon WNT ligand binding and subsequent destabilisation of the β-catenin degradation complex, the repressive cofactors are replaced by β-catenin which is tethered to WREs by TCF that thus functions as a scaffold for recruitment of an auxiliary machinery of cofactors that regulate chromatin remodelling factors and RNA pol II to initiate transcription[289]. Interestingly, perturbations rendering the β-catenin degradation complex dysfunctional, or lack of phosphorylation of key residues of the β-catenin molecule causes accumulation of dephosphorylated β-catenin, which has been correlated with malignancies such as colorectal cancer (CRC)[289-291].

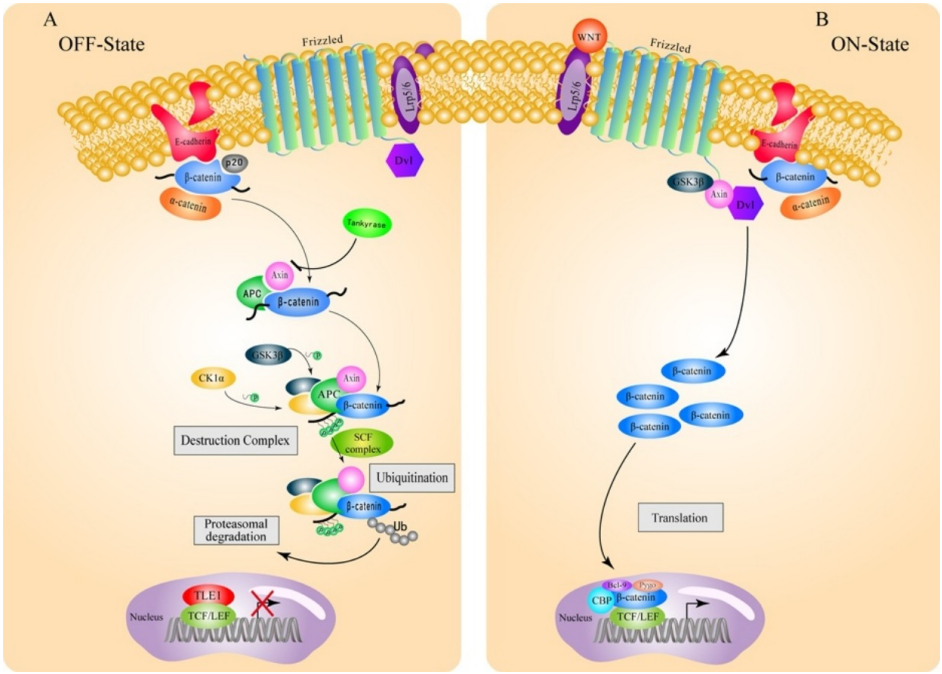

Figure 4: *Overview of the canonical WNT/β-catenin signalling pathway. Reprinted under the Creative Commons Attribution 4.0 International License, https://creativecommons.org/licenses/by/4.0/, by Biomedicine and Pharmacotherapy, 2020;132:110851. He S., Tang S. WNT/β-catenin signaling in the development of liver cancers*[292].

WNT signalling has been widely implied in cancers, and plays a central role in CRC where its activation constitutes one of the earliest events. It is estimated that up to 80% of all CRC patients feature *APC* LOFs, while half of the remaining cases harbour activating β-catenin mutations, which cause an upregulation of the WNT signalling pathway[293]. Moreover, WNT signalling plays an important role in the self-renewal of CRC stem cells (CSCs) in the intestinal crypt[294], and may favour either cell proliferation or differentiation, depending on whether the recruited β-catenin cofactor is p300 (promoting CSC differentiation) or cAMP Responsive Element Binding Protein (CREB)-binding protein (CBP) (favouring maintenance of CSC potency)[293], highlighting the complexity of interactions in the WNT signalling pathway. Additionally, WNT signalling has been found to drive tumorigenesis in CRC cells through transmembrane 4 L6 family member 1 (TM4SF1), which modulates expression of the pluripotency gene *SOX2* in a WNT/β-catenin-MYC-dependent manner to maintain cancer

cell stemness and promote EMT[295]. The same study documented that in CRC cell lines, knockdown of TM4SF1 resulted in reduced expression of *MYC*, in turn decreasing MYC-binding to the *SOX2* promoter, whilst TM4SF1 depletion inhibited metastasis and tumour growth in a mouse xenograft model[295]. Another example of how coordinated WNT signalling and *MYC*-expression drives tumorigenesis has been demonstrated in lung cancer. A study by Xiong *et al*. showed that serine-threonine kinase 31 (STK31) regulates proliferation and cell cycle progression of lung cancer cells through activation of the WNT/β-catenin pathway, and downregulation of STK31 significantly inhibited cell proliferation through G1-arrest of the cell cycle, concurrent with decreased protein levels of β-catenin, MYC and cyclin D1[296]. Moreover, MYC was found to regulate *STK31* expression through direct binding to its promoter region, thus, the authors also inferred that MYC might exert effects upstream of STK31, accordingly forming a positive feed-back loop[296].

Interestingly, 3D genome organization also seems to play a role in the WNT-mediated regulation of *MYC*-expression in CRC. WREs are located both distal and proximal of the *MYC* locus, and the *MYC* 3' WRE, *MYC* -335 WRE and distal SE may be juxtaposed to the 5' *MYC* WRE, positioned within the proximal promoter region, through long-range chromatin loops[277]. Thus, it has been proposed that distal "hijacked" WREs interact with the proximal *MYC* promoter to locally increase the concentration of β-catenin/TCF complexes to enhance oncogenic *MYC* expression in CRC[277]. Moreover, the drug BC21, which inhibits WNT signalling through disruption of interactions between β-catenin and TCF4, has been observed to affect the gating of *MYC* in HCT116 cells, where BC21 treatment reduced the potential for interaction between NUP133 and the OSE, thus insulting the anchoring of the OSE to the nuclear pores[147]. Consequentially, BC21 administration significantly decreased both the nuclear export rate of nascent *MYC* transcripts, as well as total *MYC* mRNA cytoplasmic levels in HCT116 cells[147].

In summary, the WNT signalling pathway, transcriptional and post-transcriptional regulation of *MYC* expression, and 3D chromatin organization within the nuclear architecture are intricately intertwined in signalling networks, which suffer from frequent insults in diseases such as cancer. However, the scope of these signalling cascades is vast, and the full extent of the crosstalk between these features is not yet fully understood.

2 RESEARCH AIMS

The overall aim of this thesis was to explore the connections between 3D genome organization and chromatin crosstalk within the nuclear architecture underlying the regulation of oncogene expression in 3D, in response to cancer-relevant signalling pathways. Accordingly, two independent studies were implemented in order to answer the following research questions:

1. What is the connection between stochastic chromatin crosstalk, ensuing dynamic 3D chromatin states and trasncriptional activation? More specifically, how does *MYC* communicate with its flanking enhancers, and how do these interactions relate to the transcriptional activity of *MYC*?

2. What are the underlying mechanisms and molecular factors that organise the spatial localisation of functional long-range enhancer-promoter interactions in response to external signals? More specifically, does CTCF, the master weaver of the genome, orchestrate WNT-mediated gene gating of the oncogene *MYC*?

3 MATERIALS AND METHODS

3.1 CELL CULTURE AND TREATMENTS

Human colon cancer cells (HCT116) were kindly provided as a gift by Dr B. Vogelstein, and cultured in complete growth medium (McCoy's 5A modified medium, Thermo Fisher Scientific, 26600023, with 10% fetal bovine serum (FBS), Thermo Fisher Scientific, 16141079, and 1% penicillin-streptomycin, Gibco, Invitrogen, Thermo Fisher Scientific, 11548876). Primary cultures of normal human colon epithelial cells were commercially obtained from ScienCell (HCEC, HCoEpiC, ScienCell, 2950) and were maintained in CoEpiC medium (CoEpiCM, ScienCell, 2951). Both cell lines were cultured at 37°C under 5% CO_2. Drosophila Schneider 2 (S2) cells were commercially obtained from Thermo Fisher Scientific (R69007) and grown in Schneider's Drosophila medium (Thermo Fisher Scientific, 21720024) at the ambient temperature. Cell cultures were regularly screened for Mycoplasma contamination with the EZ-PCR Mycoplasma Test Kit (Biological Industries, Cromwell, 20-700-20).

HCT116 cells were treated with 10 μM of the β-catenin and TCF V Inhibitor (BC21, Merck Millipore, 219334) or the equivalent amount of the vector DMSO as a control, for 16 hours.

Downregulation of CTCF and AHCTF1 was accomplished through transfection with short interfering RNA (siRNA). HCT116 cells were transfected with 20 nM CTCF siRNA (Santa Cruz, sc-35124), AHCTF1 siRNA (Santa Cruz), or GFP siRNA (Santa Cruz, sc-45924). After 6 hours of lipofection using Lipofectamine RNAiMAX transfection reagent (Life Technologies, 13778075) the medium was replaced with complete McCoy's 5A modified medium with 10% FBS and 1% penicillin-streptomycin, followed by subsequent harvesting of cells after 48 hours. The efficiency of siRNA-mediated downregulation was validated through RT-qPCR analysis.

3.2 CRISPR-CAS9-MEDIATED EDITING OF THE OSE-SPECIFIC CTCFBS

For paper II, CRISPR-Cas9 was used to create HCT116 cell lines with a mutated OSE-specific CTCFBS. Key sequences within the main CTCFBS embedded in the OSE (chr8:128,219,114-128,219,767) were modified using CRISPR-Cas9 technology custom service provided by Synthego (CA, USA). Specific guide RNA (sgRNA) targeting the CTCFBS were complexed with the *Streptococcus pyogenes* Cas9 (spCas9) nuclease to form a ribonucleoprotein complex (RNP), which was delivered together with donor DNA to the cells through electroporation. The sequence within the OSE-specific CTCFBS was modified from CTCACCATTGGAGGGCATTG to TTCATTATTTTATTTCATTG. The donor DNA sequence was:
TTCTCACTGACTCTAAAACCTATCCATGCTCCTAAACCTCTTCATTATTTTATTT
CATTGCTGTTTACCCTTTCAGTTTCAGCTGTACTATCAAAAGCAG. After 48 hours of recovery, the created modifications were evaluated by PCR amplification of the edited site,

followed by Sanger sequencing. Single cells for clonal expansion were seeded from the CRISPR-Cas9 edited cell pool, and each seeded well was rigorously tracked and imaged every 2-3 days to confirm that the population was truly clonal. Two clones were selected and expanded (D3 and E4), without the use of any selective agents to enrich for edited populations. Resulting clones were verified using Sanger sequencing. As a control for the D3 and E4 edited clones, a HCT116 cell population, which was not exposed to any reagents used for the CRISPR-Cas9-mediated editing process, was used and referred to as "WT".

Co-culture experiments were performed in order to distinguish edited mutants (D3 and E4) from the control (WT) cells. Specific primers for the mutated sequence of the CTCFBS, as well as primers designed to anneal specifically to the non-edited CTCFBS sequence were used.

3.3 *IN SITU* METHODS

3.3.1 Wide-field microscopy

For both paper I and II, a Leica DMi8 microscope, equipped with a HC PL APO 63X NA 1.4 oil objective and DFC9000 camera, and using the Instant Computational Clearing Method of the Thunder Imaging System (Leica Microsystems), was used for cell imaging and generation of optical sections in 3D. Stacks were taken using the software system optimized intervals in the z-axis. Acquired images were analyzed with the use of the Imaris or Leica Application Suite X (LasX) softwares. Due to the limitations in the resolution of the fluorophores (with CY3 at 239.6 nm), the distance data were stratified using 240 nm as the first cut-off.

3.3.2 3D DNA FISH

In paper I, DNA FISH probes were prepared from a pool of PCR products spanning 8–10 kb regions of Hind III sites encompassing the *MYC* promoter and gene body (chr8:128,746,000–128,756,177), the OSE (positioned at chr8:128,216,526–128,225,855), or an in-between enhancer (EnhD, positioned at chr8:128413009–128414109), and the PCR products were sonicated to 500-2000 bps range and labeled with green 496-dUTP (Enzo, 42831), Cy3-dCTP (GE Healthcare, PA53021)and Cy5-dCTP (GE Healthcare, PA55021) respectively. The genomic regions surrounding the *MYC* loci and OSE were visualized by the bacterial artificial chromosome (BAC) clone CTD-3066D1. 3D DNA FISH analyses were performed as previously described [147].

For paper II, DNA FISH probes for *MYC* promoter and the OSE were similarly prepared from a pool of PCR products consisting of 8–10 kb regions of Hind III sites containing the *MYC* promoter and gene body (chr8:128,746,000–128,756,177), and the OSE (positioned at chr8:128,216,526–128,225,855). The probes were labeled with either biotin-16-dUTP (Roche, 11093070910) or digoxigenin-11-dUTP (Roche, 11573179910), using the Bioprime

Array CGH kit (Life Technologies, 18095011), and hybridized to formaldehyde crosslinked cells as described before[147]. To ensure that the PCR probes identified the correct genomic region surrounding the OSE and *MYC*, the BAC clone CTD-3066D1 was routinely included. After hybridization, the cells were incubated with primary anti-biotin antibodies (Cell Signaling, 5597S; diluted 1:200) and anti-digoxigenin antibodies (Roche, 11333062910; diluted 1:200) at 4 °C, overnight. Following primary antibody incubation, the cells were washed with 0.05% Tween 20 in PBS, and incubated with secondary antibodies for 1 hour at room temperature, followed by washing with 0.05% Tween 20 in PBS. The cells were then counterstained with DAPI (ThermoFisher Scientific, 62248) and mounted with Vectashield mounting medium (Vector Labs, H-1900).

3.3.3 RNA FISH

RNA FISH experiments for paper II were performed on cells crosslinked at room temperature for 15 min, using 3% formaldehyde in PBS. Vanadyl-ribonucleoside complex was added to buffers in all steps to inhibit ribonuclease activity. The *CCAT1* RNA FISH probes were generated from two PCR products encompassing a region within the OSE (positioned at chr8:128,216,526- 128,225,855) and labeled with biotin-16-dUTP (Roche, 11093070910) as described above. RNA FISH analyses were performed as previously described[147].

3.3.4 *In situ* Proximity Ligation Assay (ISPLA)

For paper II, ISPLA analyses were performed using cells fixed with 1% formaldehyde in PBS, followed by blocking in 10% goat serum in PBS and subsequent overnight incubation at 4 °C with primary antibodies, as previously described[82]. Purified mouse anti-CTCF (BD Biosciences, 612148; diluted 1:40) and rabbit polyclonal anti-AHCTF1 (Novus Biologicals, NBP1-87952; diluted 1:100) were used as primary antibodies for the CTCF:AHCTF1 ISPLA experiments. For CTCF:NUP133 ISPLA experiments, purified mouse anti-CTCF (BD Biosciences, 612148; diluted 1:40) and rabbit monoclonal anti-NUP133 (Abcam, ab155990; diluted 1:100) were used. To generate rolling circle amplification, oligonucleotide-conjugated anti-rabbit and anti-mouse antibodies (termed “R+” and “M-”) were used as described previously[82]. Samples without incubation with the primary antibodies served as background control. Quantitation of the Cy3 fluorophore-marked ISPLA signals inside the nuclei (counterstained with DAPI) was performed on the Imaris v.8.1.2 software (Bitplane, Switzerland).

3.4 IMMUNOPRECIPITATION EXPERIMENTS

3.4.1 Co-immunoprecipitation (co-IP) assay and Western blot

For paper II, co-immunoprecipitation experiments were performed using the Nuclear Complex Co-IP kit, following the manufacturer's protocol (Active Motif, 54001). Pre-clearing was performed using 250 µg of protein from nuclear lysates, which were incubated with Dynabeads protein G (Thermo Fisher Scientific, 10004D) for 1 hour at 4 °C. The samples were then incubated overnight at 4 °C with 2.5 µg of anti-CTCF (Abcam, ab37477) or normal mouse IgG (Santa Cruz, sc-2025). The immunoprecipitated material was analyzed by Simple Western assay using an automated WES/JESS system (ProteinSimple, Bio-Techne). For Western blot analysis anti-CTCF (Cell Signaling, 2899S, rabbit; diluted 1:50), anti-NUP133 (Abcam, ab155990, rabbit; diluted 1:25), anti-AHCTF1 (Novus Bio, NB600-238, rabbit; diluted 1:50), anti-β-catenin (Cell Signaling, 8480S, rabbit; diluted 1:100) and anti-TATA binding protein (Abcam, ab51841, mouse; diluted 1:50) were used. Relative amounts of each protein were quantified using the detected peaks of the chemiluminescence electropherogram generated by the Compass for SW software (ProteinSimple, San Jose, CA).

3.4.2 ChIP-qPCR and ChIP-seq

For ChIP-qPCR and ChIP-seq experiments in paper II, cells were fixed with freshly prepared 1% formaldehyde in PBS. Fragmentation of chromatin was performed by either sonication or enzymatic digestion by micrococcal nuclease (Cell Signaling Technology). DNA-protein complexes were immuno-purified with antibodies against CTCF (Cell Signaling, CS 2899), AHCTF1/ELYS (Novusbio, NBP1-87952), TCF4 (Santa Cruz, sc-8631), or ß-catenin (Cell Signaling, 8480S), and Dynabeads protein G (Thermo Sciences, 10004D). After purification of immunoprecipitated DNA using the ChIP DNA Clean & Concentrator kit (Zymo Research, D5205), qPCR analysis was used to quantify target loci-protein associations, using primer sequences and qPCR conditions as previously described[147]. For the three independent ChIP-seq experiments of paper II exploring the binding patterns of CTCF, chromatin was prepared following the same protocol as for ChIP-qPCR experiments, however, the ChIP-seq experiments used 3 x the input material. This threefold scaling-up of the protocol was conducted in order to attain enough immunoprecipitated DNA for sequencing.

3.5 NODEWALK

Generation of chromatin networks impinging on *MYC* were performed using the Nodewalk technique as described previously[147,297,(Vestlund *et al.*, *Nat Prot*, in press)]. In paper I, Nodewalk experiments were employed to generate chromatin networks impinging on *MYC* and 9 other loci in HCT116 cells and HCEC, to compare the stochastic dynamics of interactomes in both smaller and larger cell populations. In order to do so, both the standard Nodewalk protocol and an adapted protocol for small input material were used. In paper II, chromatin networks

impinging on *MYC* were generated using the standard Nodewalk protocol, using WT HCT116 cells and the mutant clones D3 and E4 to explore the how dynamics of the interactomes are affected by abrogation of the OSE-specific CTCFBS.

In summary, Nodewalk is based on the original 3C protocol by Dekker *et al*.[168], but comprises several key modifications. One of the advantages of Nodewalk is its sensitivity, and it has been adapted for low input material, making it possible to attain reproducible results from input corresponding to genomic content of 10 cells. For Nodewalk experiments, cells are crosslinked using 1% freshly prepared formaldehyde in PBS, before enzymatic digestion with *Hind* III restriction enzyme (NEB, R3104M). The crosslinked chromatin is digested to near completion, to a digestion efficiency that should be >90%. Digestion efficiency is estimated through qPCR using primers annealing to sequences on each side of *Hind* III restriction sites flanking *MYC*, and one primer pair spanning a region of *MYC* exon 2. The chromatin is then ligated using T4 DNA ligase (NEB, M0202L), with the final concentration of DNA during ligation being <0.5 ng/µl in order to avoid random ligation events. Another key feature of Nodewalk is the inclusion of crosslinked and digested *Drosophila* S2 chromatin in the ligation step, which enables estimation of background artificial random ligation events, which usually are <1%. Ligation efficiency is assessed using the DNA High sensitivity kit (Agilent Technologies) for Bioanalyzer 2100 (Agilent Technologies). Following ligation, de-crosslinking, and purification, the Illumina DNA Prep, (M) Tagmentation kit (Illumina, 20018704) (or Nextera XT DNA Library Preparation kit [Illumina, FC131-1024] for the adapted low-input protocol) is used to simultaneously fragment chimeric DNA and introduce adapter sequences. Nodewalk also incorporates an *in vitro* transcription step, which enables linear amplification of 3C DNA and contributes to the technique's quantitative qualities. Specific primers targeting a region of interest (such as *MYC*, as in the case of both paper I and II) are then annealed to the RNAs, corresponding to the chimeric DNA sequences, followed by conversion into cDNA through RT-PCR using the Platinum Quantitative RT-PCR ThermoScript One-Step system (Invitrogen, 11731-015). The cDNA is then used for further library preparation, and Illumina sequence adapters are incorporated using the same primers (as used in the RNA annealing step), but equipped with a P7 sequence against the cDNA that already contains P5 sequences, thus enabling direct generation of double-stranded cDNA suitable for sequencing. Following purification with Ampure XP beads (Beckman Coulter, A63880), quantification of cDNA concentration, and assessment of average fragment size by the DNA High sensitivity kit for BioAnalyzer 2100 (Agilent Technologies), libraries are sequenced using the MiSeq Reagent Kit v2 (300-cycles) (Illumina, MS-102-2002) that generates 140–150 bp paired-end reads.

3.6 RNA ASSAYS

3.6.1 Pulse labelling of RNA and nuclear RNA export assay

For paper II, nascent mRNA was labeled by incubating cells for 30 min with 0.5 mM (final concentration) 5'-ethynyl uridine (Thermo Scientific, E10345). Cells were harvested immediately for experiments assessing transcriptional activity, while for pulse-chase experiments, EU was removed and cells were washed with PBS followed by incubation with normal growth medium for 1 hour. EU-labeled RNAs were captured using the Click-iT Nascent RNA Capture kit (Thermo Fisher, C10365), following the manufacturer's protocol. RNA was the converted into cDNA by using the SuperScript VILO cDNA synthesis kit (Life Technologies, 11754050). To determine cytoplasmic/nuclear ratios of exported mRNA, nascent RNAs were EU-labeled (as described above), and 1 hour after chase the nuclear and cytoplasmic fractions were separated using the Ambion PARIS kit (Thermo Fisher, AM1921) following the instructions provided by the manufacturer. The EU-labeled RNAs (nuclear and cytoplasmic) were then converted to cDNA as described above, followed by qRT-PCR analyses.

3.6.2 qRT-PCR analysis of transcription and RNA-seq

Transcription levels of *MYC*, *FAM49B* and *CCAT1* were evaluated by qRT-PCR analysis. RNA samples were lysed and purified with RNeasy Mini kit (Qiagen) according to the manufacturer's protocol, and RNA quality was assessed by the RNA Pico 6000 kit for Bioanalyzer 2100 (Agilent Technologies). RNA was then used for cDNA synthesis using the SuperScript VILO cDNA synthesis kit, as described above. qRT-PCRs were performed using 2 µl of diluted (1:50) cDNA per reaction, with 10 µl iTaq Universal SYBR Green Supermix (Bio-Rad, 1725125), 1 µl 10 mM forward primer, 1 µl of 10 mM reverse primer, and 6 µl nuclease free H_2O (total volume 20 µl) on the RotorGene 6000 PCR cycler (Corbett Research). Serial dilution of sonicated gDNA from HCT116 cells was used to confirm the linear range of amplification.

For precise quantification of *MYC* and *FAM49B* mRNA levels in paper II, 10^6 cells (from WT, D3 and E4 respectively) were counted and washed with pre-warmed PBS. Before lysis, 1 µl of diluted (1:10) ERCC ExFold RNA Spike-In Mix (Thermo Fisher Scientific) was added to each sample. Lysis, purification and cDNA synthesis were performed as described above.

The ERCC RNA spike-in mix was also used for RNA-seq samples in paper II. RNA-seq libraries were prepared with standard Illumina TruSeq Stranded mRNA kit with Poly-A selection, and samples were sequenced on NovaSeq6000 (NovaSeq Control Software 1.7.0/RTA v3.4.4) with a 51nt(Read1)-10nt(Index1)-10nt(Index2)-51nt(Read2) setup using "NovaSeqStandard" workflow using the mode for the "SP" flowcell.

4 RESULTS

4.1 PAPER I: *MYC* AS A DRIVER OF STOCHASTIC CHROMATIN NETWORKS: IMPLICATIONS FOR THE FITNESS OF CANCER CELLS

The "active chromatin hub" hypothesis[219] depicts how active enhancer hubs may contact multiple genes at the same time, however, due to technical limitations it has been challenging to explore the dynamics of stochastic events occurring at single-cell level. It has been argued that multi-way contacts between distinct loci might represent transiently stabilized chromatin interactions[221], however, the regulation and generality of such hub formation are not completely understood.

4.1.1 *MYC* drives the formation of chromatin networks

The active *MYC* gene is able to form chromatin interactomes that are abundant in connections with its flanking enhancers[147,297]. The Nodewalk technique was employed in order to explore the most important nodes of the *MYC* network in both normal human colon epithelial cells (HCEC), and in a human colon cancer cell line (HCT116). Two thirds of the interactions in HCT116 cells were sampled randomly in order to compensate for any bias due to the fact that *MYC* is triploid in HCT116 cells and diploid in HCECs[298]. Only minor differences were observed between the un-stratified interactomes of HCT116 cells and HCECs, however, the most connected nodes were substantially more prominent in the cancer cells (**Figure 5**). Moreover, increasing *k*-core values (indicating a higher node-connectivity) were correlated with enrichment of active or primed enhancers in HCT116 cells. We then aimed to identify the nodes with the highest dynamic indexes, in order to elucidate which interactors were the most important in regulating the topology of the network. Surprisingly, rather than the most well-connected enhancer nodes, *MYC* turned out to be the interactor with the highest dynamic index in both cancer cells and HCECs, suggesting that *MYC* is seeking out contacts with its flanking enhancers, as opposed to the other way around.

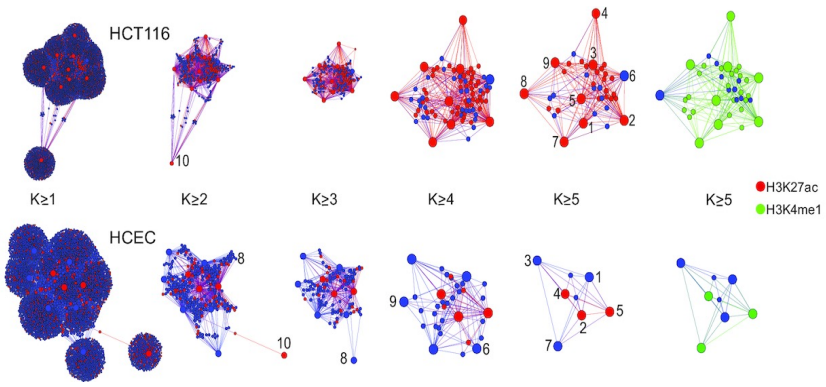

***Figure 5**: The chromatin network structures generated from 20 000 HCT116 cells and 20 000 HCEC cells respectively, stratified by k-core values. Adapted from Sumida N, Sifakis EG, Kiani NA, et al. MYC as a driver of stochastic chromatin networks: implications for the fitness of cancer cells. Nucleic Acids Res. 2020;48(19):10867-10876. doi:10.1093/nar/gkaa817 [299] . Reprinted under the Creative Commons Attribution 4.0 International License, https://creativecommons.org/licenses/by/4.0/*

4.1.2 Stochastic contacts establish the *MYC* interactome

To explore the significance of stochastic events in the generation of chromatin networks, and investigate whether the *MYC* network is detectable only in larger cell populations, masking the high biological variability in very small cell populations, the Nodewalk protocol was adapted for very low input samples. Then, three distinct sets of samples were prepared: a) one set of technical replicates with large input samples, b) one set of 10 technical replicate samples prepared from small input material, where each sample contained 0.88 ng (corresponding to 176 cells) from the same pool of 3C DNA from 10^6 cells, and c) one set of 9 biological replicate samples prepared from small cell populations (with input material corresponding to 177 cells). The majority of the *MYC* network consists of *cis*-interactions, and we therefore focused on such intrachromosomal interactors in our analysis. We found that the overlap of interactors in three technical replicates originating from the same RNA library, which had been prepared from 0.88 ng of ligated 3C DNA, generated a technical reproducibility >90%. Next, the reproducibility of chromatin interactions between the ten 0.88 ng technical replicates was compared to the reproducibility of the nine 177-cell biological replicate samples. Among the 177-cell samples, >70% of *MYC*-interactors were only detected in one of the libraries, whereas among the 0.88 ng-aliquots, >85% of the interactors were reproduced in two or more libraries. Moreover, when the 177-cell libraries were pooled together, there was a >91% observed overlap with the interactors detected in the network of the large input samples (corresponding to 10 000 cells), suggesting that the variability in the interactions is not due to technical variation. These results thus demonstrate that Nodewalk using small input is reliably able to identify interactors of an ensemble network present in larger cell populations. Thus, the analysis of the generated interactomes from small *vs*. large input samples indicate that the *MYC* chromatin network is established by stochastic interactions occurring at the level of individual cells.

4.1.3 Chromatin interactions between *MYC* and its enhancers are mutually exclusive

To further explore the importance of stochastic interactions in the *MYC* network, input for Nodewalk was reduced to even smaller amounts of 34.8 pg 3C DNA per aliquot, each taken from one 177-cell 3C sample and corresponding to 21 alleles in seven cells. The nine 177-cell samples were then compared to 23 aliquots of 34.8 pg 3C DNA samples as input material covering an entire 177-cell sample, and revealed that 6 out of 8 distinct interactors within the *MYC* flanking TADS (referred to as TAD 1 and 2), that were detected in the pooled libraries of the 23 x 34.8 pg samples, overlapped with interactors of the pooled 9 x 177-cell samples. Additionally, the number of enhancers that interacted with *MYC* in each of the twenty-three 34.8 pg aliquots ranged between 0 and 1. There was no significant change of the observed number of enhancers impinging on *MYC,* compared to an expected stochastic interactome, when including all *cis* and *trans* interactors of the ensemble network, and this observation, taken together with an average bait recovery of 36.2% in the 34.8 pg aliquots, made us infer

that on average 0.7 enhancers interact with 7.6 different *MYC* alleles. Moreover, the binned interactome data from samples using input of 10 000 cells, 0.88 ng, and 177 cells respectively, uncovered that the overall interaction pattern of the ensemble network could be reproduced in the smaller input samples, thus illustrating a network formed by stochastic and mutually exclusive contacts between *MYC* and its flanking enhancers. This conclusion was further strengthened by *in situ* experiments using 3D DNA FISH, which revealed that *MYC* and two major interactors (the OSE and Enhancer D [EnhD]), rarely, if at all, co-localize in the 3D nuclear space. Taken together, the findings demonstrate that although individual chromatin interactions are specific and highly reproducible in larger cell populations, the generated networks are merely virtual, and emerge as a result of accumulated stochastic events taking place in smaller cell populations.

4.2 PAPER II: CANONICAL WNT SIGNALING-DEPENDENT GATING OF *MYC* REQUIRES A NONCANONICAL CTCF FUNCTION AT A DISTAL BINDING SITE

Our group has previously uncovered a WNT-regulated mechanism by which human colon cancer cells are able to post-transcriptionally increase *MYC* expression through OSE-mediated recruitment of *MYC* to nuclear pores[147]. Several lines of evidence supported the notion that it was the OSE within the *MYC* interactome that facilitated the gating of *MYC* to NPCs. 3D DNA FISH analyses thus revealed that the OSE was generally localized closer to the lamina than *MYC,* specifically in HCT116 cells, indicating that *MYC* trails the OSE when within 1 μm from the nuclear periphery. Moreover, AHCTF1-mediated anchoring of the OSE to NPCs was found to be under the control of WNT signalling, as BC21 treatment disrupted TCF4-AHCTF1 interactions and decreased the nuclear export rates of *MYC* mRNAs in HCT116 cells[147]. Importantly, AHCTF1 was found to bind primarily to the OSE but not to *MYC*. While these findings represent circumstantial evidence that strongly supports a central role of the OSE in the gating of *MYC*, they did not provide any genetic evidence for the OSE-mediated gating. Therefore, we set out to explore the underlying mechanisms of this gene gating phenomenon in human cancer cells.

4.2.1 CTCF binding to the *MYC* OSE confers a proliferative advantage to colon cancer cells

To provide genetic evidence for the OSE-specific gating of *MYC*, we chose to focus on the role of CTCF in the OSE-mediated gene gating process, as CTCF has been shown to be involved in the regulation of enhancer-promoter interactions, transcription of *MYC*, and chromatin mobility within the 3D nuclear architecture[82,300]. Moreover, there is a distinguished CTCFBS positioned within the OSE-specific eRNA gene *CCAT1*.

First, the binding of CTCF to the motif sequence within the eRNA gene *CCAT1* was examined using ChIP-qPCR in HCT116 cells and normal colon epithelial cells (HCECs), and

revealed that CTCF binding to the OSE-specific site is prominent in HCT116 cells, whilst being completely absent from the corresponding region in HCECs. This finding strengthened the hypothesis that the *CCAT1* CTCFBS might be involved in the gene gating of *MYC* in HCT116 cells, as the gating process of *MYC* is not found in HCECs. Therefore, a sequence of 8 bps within the OSE-specific CTCFBS was genomically edited using CRISPR-Cas9, and two mutant clones (D3 and E4) were generated, while parental HCT116 cells were referred to as WT (please see Materials and methods). ChIP-qPCR experiments were then employed to investigate whether the edited CTCFBSs maintained the ability to bind CTCF. As expected, the mutant clones D3 and E4 showed a significant decrease of CTCF binding recovery (>90%) compared to the WT cells, while CTCF binding to the positive internal controls at the *MYC* promoter and the *H19* ICR were largely unaffected amongst WT, D3, E4, and HCECs alike. To exclude direct or indirect off-target effects of the CRISPR-Cas9-mediated editing process, ChIP-seq experiments were performed in order to get an overview of the CTCF binding patterns genome wide, with a particular focus on the *MYC* TAD and its neighboring TADs harboring key *MYC* regulatory elements. Apart from the reduced binding at the *CCAT1*-specific CTCFBS already confirmed by ChIP-qPCR, all other CTCFBSs within the *MYC*-OSE region retained their CTCF-binding ability equally well in WT, D3 and E4 cells. ChIP-seq analysis also revealed that the global CTCF-binding patterns were not significantly altered in WT compared to the mutant clones, which together with implemented whole-genome sequencing of WT, D3 and E4 indicated no of off-target effects.

To investigate whether the altered CTCFBS also conferred a phenotypical distinction, co-culture experiments were performed, and demonstrated that the WT *CCAT1*-specific CTCFBS provided an excessive growth advantage to the WT cells, which had outgrown the mutant D3/E4 clones within a week. Given the WNT-dependency of the gating process, we sought to address the potential role of WNT signalling in the observed growth advantage of WT cells. To this end, the agent BC21 was used, as it has been shown to disrupt interaction between TCF4 and β-catenin, as well as evicting β-catenin from chromatin proximal to the CTCFBS within the OSE[147]. Strikingly, the administration of BC21 in co-cultures of WT/D3 and WT/E4 was found to reduce the WT growth advantage by ~6-fold during 2 weeks of co-culturing. Thus, our findings suggest that CTCF binding to the *CCAT1*-specific motif within the OSE confers a WNT signalling-dependent growth advantage to human colon cancer cells.

4.2.2 *MYC* and *FAM49B* expression is increased through facilitated mRNA export regulated by the OSE-specific CTCFBS

Our research group has previously documented facilitated nuclear export of *MYC* mRNAs under the control of the OSE in HCT116 cells[147]. To investigate the potential involvement of the *CCAT1*-specific CTCFBS in this process, we employed 5-EU pulse-chase and RNA export assay experiments to explore the rates of nuclear mRNA export in WT cells and the mutant clones D3/E4. We thus determined the export rates of both *MYC* and *FAM49B*

derived transcripts, the latter constituting a gene with a proximal enhancer that has been found to interact with the OSE in Nodewalk analyses[299] and whose gene products functionally interact with those of *MYC*[301,302]. Results of these experiments showed that there was a more than threefold reduction of the nuclear export rate of *MYC* and *FAM49B* mRNAs in both mutant clones (D3/E4) compared to the WT cells, while no significant differences in overall transcriptional rates were observed. Moreover, treatment with the BC21 drug did not affect nuclear export rates in the D3 and E4 clones, while the export rates of *MYC* and *FAM49B* mRNAs in WT HCT116 cells were significantly decreased, suggesting that WNT signaling impinges on the *CCAT1*-specific CTCFBS. Additionally, the reduced nuclear export rates of *MYC* and *FAM49B* transcripts in the mutant clones correlated with a significantly reduced total mRNA expression levels of *MYC* and *FAM49B* in the D3/E4 mutant clones compared to WT HCT116 cells. These results thus demonstrate that although the *CCAT1*-specific CTCFBS does not control the transcription of *MYC* and *FAM49B*, the CTCFBS exerts its control of *MYC* and *FAM49B* expression through regulating the nuclear export rates of their respectively derived mRNAs. This conclusion was further strengthened by computer simulations, which confirmed that the OSE-specific CTCFBS increases *MYC* expression in HCT116 WT cells solely through facilitating nuclear export of *MYC* mRNA, consequently conferring a strong proliferative advantage to the colon cancer cells.

4.2.3 CTCF and β-catenin collaborate to recruit AHCTF1 to the OSE

As AHCTF1 holds an important role in the gating of *MYC* in colon cancer cells by mediating the tethering of the OSE to nuclear pores[147], next we investigated the potential relationship between AHCTF1 and CTCF through implementing co-immunoprecipitation analyses. These experiments uncovered that CTCF physically interacts not only with AHCTF1, but also with β-catenin and NUP133, which are also involved in the gene gating of *MYC*. In line with the strong CTCF-AHCTF1 interaction, the mutated CTCFBS in both the D3 and E4 mutant clones was found to convey a reduced AHCTF1 binding to the CTCFBS. Moreover, the interaction between CTCF and AHCTF1 was found to likely be under the control of WNT signalling, as administration of BC21 disrupted the AHCTF1:CTCF complex, reducing the complex formation by approximately 50%. Similarly, BC21 treatment strongly reduced the presence of AHCTF1 at the OSE while leaving CTCF binding to the *CCAT1*-specific CTFBS unaffected, suggesting that the recruitment of AHCTF1 to the OSE is jointly facilitated by its interaction with CTCF and WNT signalling. To further support the role of CTCF in the recruitment of AHCTF1 to the WT CTCFBS, CTCF was down-regulated using siRNA treatment. The results showed reduced binding of both AHCTF1 and CTCF to the OSE-specific CTCFBS in WT HCT116 cells, suggesting that the effect of CTCFBS on AHCTF1 recruitment to the OSE is mediated *via* CTCF. Collectively, these results thus indicate a collaborative function between the CTCF-CTCFBS complex and WNT signalling in the recruitment of AHCTF1 to the OSE.

To explore the potential role of AHCTF1 in the regulation of spatial positioning of the OSE within the nucleus, 3D DNA FISH analyses were carried out on WT HCT116 cells treated with either siAHCTF1 or siGFP. These analyses showed an altered distribution of the OSE's localisation in the nucleus upon reduced expression of AHCTF1, which was most prominent within 0.7 μm from the periphery. Downregulation of AHCTF1 thus impaired the mobility of the OSE in the nuclear space closest to the periphery, hence inhibiting the recruitment of the OSE to NPCs. Moreover, ISPLA experiments revealed that the potential for interactions between AHCTF1 and CTCF was the highest in the region around 1 μm distal to the nuclear periphery, similarly to ISPLA-signals between CTCF and NUP133. Interestingly, the proximities between AHCTF1 and CTCF in WT cells were significantly reduced upon treatment with BC21, strengthening the notion that collaboration between CTCF and the TCF4:β-catenin complex is required for the establishment of its interaction also with AHCTF1.

4.2.4 The *CCAT1*-specific CTCFBS coordinates the proximity between the OSE and *MYC* at the nuclear periphery, but does not directly mediate their interactions

To understand further the role of the CTCFBS in the regulation of *MYC* gating, we explored the effects of the CTCFBS mutation on the sub-nuclear localisation of the OSE and *MYC* regions as well as on their interaction. First, 3D DNA FISH experiments were thus employed on WT, D3 and E4 cells in order to examine the nuclear localisation of these regions. In order to visualize the relationship between the generated FISH-signals, a "c"-value was calculated based on the formula $c = b - a$, where "a" is the distance between the *MYC* locus and nuclear periphery (in μm), and "b" represents the distance between the OSE and the periphery. The data from the 3D DNA FISH experiments was further stratified based on cell cycle phase, as determined by the presence of single-single (G1), single-double (early S), or double-double (late S/G2) DNA FISH-signals. The results revealed that in WT HCT116 cells, *MYC* and the OSE approach the nuclear periphery in a coordinated manner, corresponding to c-values around 0 close to the periphery, in concordance with previous findings showing that the proximity between *MYC* and the OSE is most frequent at the nuclear periphery[147]. Interestingly, both the relative positions of *MYC* and the OSE to the nuclear periphery, as well as their coordinated recruitment to the periphery, were significantly reduced in the D3/E4 mutant clones compared to WT cells. However, only cells with un-replicated (single-single) alleles were found to show a significant difference in the proximity of the OSE to the nuclear envelope within a distance of 1 μm from the periphery, thus indicating that the gating process may occur specifically during the G1-phase of the cell cycle. Additionally, the percentage of *MYC* alleles present at the periphery was significantly reduced in both mutant clones, reinforcing the earlier observation that it is the OSE that leads the recruitment of *MYC* to the nuclear pore, and not the other way around[147].

Finally, Nodewalk experiments were implemented in order to further explore the role of the OSE-specific CTCFBS and *MYC*-OSE interactions in the gating process. However, the analyses demonstrated that abrogation of the *CCAT1*-specific CTCFBS did not affect the overall frequency or pattern of chromatin interactions between the OSE and *MYC* in the D3 and E4 mutant cells compared to WT HCT116 cells. Taken together, these findings thus suggest that CTCF binding to the CCAT1-specific CTCFBS promotes proximity between *MYC* and the OSE at the periphery, although without directly influencing their overall interaction frequency, consequentially conveying a non-canonical function of CTCF in the gating process of *MYC*.

4.2.5 WNT signalling activates *CCAT1* eRNA transcription via the OSE-specific CTCFBS to promote juxtapositioning of the OSE to the nuclear periphery

There is emerging evidence of *CCAT1* eRNA mediating interactions between *MYC* and the OSE, which has been suggested to be facilitated through RNA-RNA interactions[272,303]. Hence, we chose to further explore the role of the *CCAT1* eRNA in the gating of *MYC*. Employing an RNA-seq analysis, we first identified isoforms of the *CCAT1* eRNA present in WT HCT116 cells, and found that while the *CCAT1*-5L version is not expressed, WT HCT116 cells prominently express the long *CCAT1*-L isoform. Next, RNA FISH of *CCAT1* and subsequent DNA FISH experiments, targeting *MYC* and the OSE, were performed in order to investigate the potential of the transcriptionally active *CCAT1* and its derived eRNA to mediate *MYC*-OSE interactions. RNA FISH analysis revealed that *CCAT1* alleles with low transcriptional activity were positioned more distal to the nuclear periphery than alleles with higher transcriptional activity, with FISH signal distribution peaking at a position proximal to but not at the periphery (<1 μm from the lamina). Importantly, at very small distances (<0.5 μm) from the nuclear periphery/pores, the *CCAT1* RNA FISH signal declined with a concomitant increase in OSE-*MYC* proximity[147]. Moreover, we did not find a correlation between the localization or degree of transcriptional activity of *CCAT1* with proximities between *MYC* and the OSE DNA FISH signals. Taken together, these findings suggest that *CCAT1* is likely not mediating the interaction between *MYC* and the OSE in HCT116 cells. Importantly, *CCAT1* expression was correlated with the pattern of OSE distribution in the WT cells, raising the question whether the CTCFBS within the OSE might regulate *CCAT1* transcription to fine-tune the juxtapositioning of the OSE to the nuclear periphery. To this end, we examined the total and nascent levels of *CCAT1* eRNA by qRT-PCR analysis, which revealed a significant ~4-fold decrease of total and newly synthesized eRNA levels in both mutant clones compared to the WT HCT116 cells. Moreover, treatment with BC21 substantially reduced the levels of newly synthesized *CCAT1* eRNA in WT cells, whereas no significant effect was observed in the D3/E4 mutant cells. In summary, these findings indicate that WNT-dependent activation of *CCAT1* eRNA transcription requires a functional OSE-specific CTCFBS in order to potentially facilitate the step-wise recruitment of the OSE-*MYC* complex to peripheral positions, as schematically illustrated in **Figure 6**.

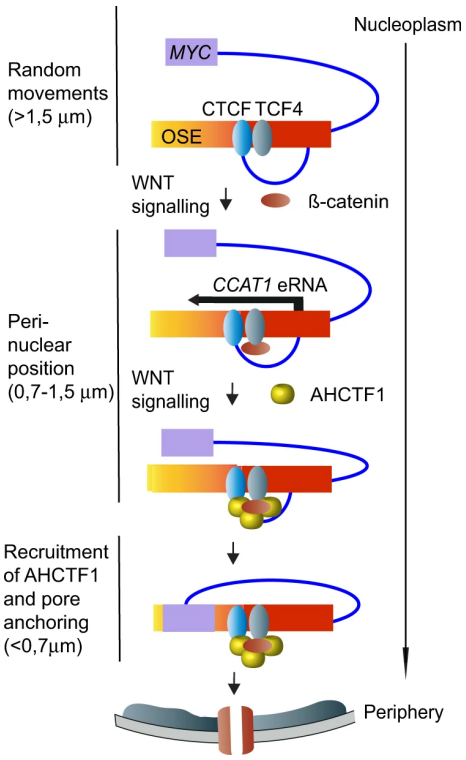

***Figure 6**: A schematic model depicting the role of CTCF in the WNT-dependent, OSE-mediated recruitment of MYC to NPCs. Reprinted under the Creative Commons Attribution 4.0 International License, https://creativecommons.org/licenses/by/4.0/ . By Chachoua I, Tzelepis I, Dai H, et al. Canonical WNT signaling-dependent gating of MYC requires a noncanonical CTCF function at a distal binding site. Nat Commun. 2022;13(1):204. Published 2022 Jan 11. doi:10.1038/s41467-021-27868-3* [304].

5 DISCUSSION

5.1 DYNAMICS OF ENHANCER-PROMOTER INTERACTIONS IMPINGING ON THE *MYC* ONCOGENE

MYC is able to form extensive chromatin networks with prevalent enhancer connections. However, the inner workings underlying the coordination of enhancer-promoter interactions, transcriptional states, and subnuclear localization of distinct genomic regions, remain largely unknown.

In Paper I, our results demonstrate that *MYC* is the driving force in the establishment of its impinging chromatin network, as *MYC* emerged as the node with the highest dynamic index in both HCT116 colon cancer cells and normal colon epithelial cells (HCECs). Thus, rather than being contacted by its enhancers to promote transcriptional activity, it seems that *MYC* itself seeks out interactions with its flanking enhancers in a stochastic and highly dynamic manner, which entails mainly pairwise interactions between *MYC* and one enhancer at a time. The dynamic features of the chromatin crosstalk between *MYC* and its enhancers were highlighted by Nodewalk analyses, which revealed that at any given time <10% of *MYC* alleles interacted with an enhancer, while >95% of *MYC* alleles have been shown to be transcriptionally active in HCT116 cells[147]. This might reflect a way for the cancer cells to diversify functional enhancer-gene interactions, and suggests that once *MYC* transcription has been activated, the enhancers no longer interact with the *MYC* gene. Indeed, previous observations from our lab indicate that different enhancers may have distinct functions depending on their subnuclear localization when interacting with *MYC*. For instance, Enhancer D (EnhD), located ~300 kb upstream of *MYC*, preferentially interacts with *MYC* in the nuclear interior where it likely promotes transcriptional activation (as indicated by largely overlapping EnhD-*MYC* ChrISP signals and RNA FISH signals of *MYC* intron 1), whereas the OSE mainly interacts with *MYC* closer to the nuclear periphery/pores, where it mediates the facilitated nuclear export of *MYC* mRNAs[147].

Transcriptional activation of *MYC* might also influence its subsequent recruitment to the nuclear periphery. A recent study thus revealed a correlation between the transcriptional states of regulatory elements and nuclear mobility, with increased mobility occurring concomitantly with the activation of transcription at enhancer alleles[91]. Interestingly, however, Flavopiridol treatment, which inhibits positive transcriptional elongation factor B (P-TEFb), has been observed to speed up the recruitment of circadian loci to the nuclear periphery[82]. Considering that *MYC* transcription is activated in two waves, one in the nuclear interior and one at the periphery[147], transcriptional elongation of active *MYC* alleles in the nuclear interior might be inhibited in order to speed up their migration to nuclear pores. At the same time, the OSE-mediated relocation of the *MYC* locus to the nuclear periphery might be necessary for the processing of *MYC* transcripts, as RNA FISH signals indicating partly or completely processed *MYC* transcripts were enriched at the nuclear periphery, in the vicinity of the *MYC* locus itself (as visualized by *MYC*-OSE ChrISP signals)[147]. Thus, these observations demonstrate enhancers with diversified functions in the control of *MYC*

transcription, and which imply that enhancers may partially divide labour amongst themselves to regulate *MYC* expression within the 3D space of the nucleus.

One possibility as to why *MYC* comprised the node with the highest dynamic index in both HCT116 cells and HCECs, might be that *MYC* was found to reside at, or close to, an inter-TAD-boundary region, which may enable *MYC* to screen for and interact with several distinct enhancers in both of its flanking TADs. Taken together with the above example of divided and diversified labour amongst *MYC*'s enhancers, this feature might have evolved to provide an enhanced plasticity of the transcriptional regulation of *MYC*. Moreover, the plasticity underlying such varied regulatory mechanisms likely allows for selective advantages that could drive the development and progression of disease, as several signalling pathways congregate on different sets of *MYC* enhancers in different cancers[305]. However, the interactions of *MYC* with its two flanking enhancers do not explain the observed high dynamic index of the OSE, which might instead be related to the gating of *MYC* mRNA transcripts that post-transcriptionally increases its expression[147].

TADs are domains that are distinguished by preferential self-interactions within its boundaries, although stochastic interactions between adjacent TADs do occur on single cell level, and more commonly so in pluripotent embryonic stem cells as compared to differentiated cells[167,306]. Given the high reproducibility of TAD structures between different studies and even certain cell types[67,167,307], TAD boundaries have been considered highly regulated and permanent to their nature[178]. However, considering our observation that MYC seems to reach out to contact interactors in both of its flanking TADs, as well as the strong variation of chromatin interactions at single cell level, the question arises whether TADs and chromatin loops are stable constructions or not? Biochemical techniques such as Hi-C, as well as microscopy and FISH based methods, have been integral in the uncovering of genome folding and organization. However, both approaches face technical limitations, with Hi-C contact maps representing static snapshots of population-averaged networks, and while FISH-based methods are able to explore interactions at single-cell level, they are still only able to provide terminal snapshots with little information about dynamic and stochastic events[178]. There is evidence that TAD formation is indeed dynamic; in mammalian cells it has been shown that TADs and loops are absent during mitosis, but most prominent in G1-phase, hence implying that TADs are formed and dissolved at least once during each cell cycle (approximately every 15 to 30 hours)[308,309]. However, a study using live cell imaging tracking of cohesin and CTCF suggested that chromatin loop and TAD boundary formation might form and disrupt even faster, as often as every ~20-25 minutes, in mammalian cells[310]. This notion is further supported by experiments where cohesin subunits were acutely depleted using an auxin-inducible degron system, showing near complete loss of loop domains, and >50% loss of TADs 20-40 min after auxin-treatment, with TADs completely disappearing at later time points[179,189]. Additionally, upon removal of auxin, Hi-C experiments revealed that loop domains were close to fully re-established 1 hour later[179]. Interestingly, the same study was able to document that cohesin loss entailed a formation of SE clusters, which established hundreds of interactions in *cis* and *trans*[179]. One might therefore speculate that enhancer hubs

might form on a single cell level due to stochastic and transient events, such as signalling cues or cell states, that could affect the presence of cohesin at CTCFBSs, or metabolic states (such as ATP availability), which by extrapolation could be hypothesized to interfere with the chromatin loop extrusion ability of cohesin.

In summary, these finding are in line with our reasoning that the *MYC* network consists of highly dynamic and stochastic interactions, with transient contacts between *MYC* and its enhancers. With *MYC* being the most important node in its impinging interactome, it seems to be able to act as a platform to converge several signalling pathways, which potentially might promote selective advantages to drive carcinogenesis.

5.2 STRENGTHS AND LIMITATIONS OF THE NODEWALK TECHNIQUE

Under the presumption that stochastic and dynamic events make up chromatin interactomes, such as the *MYC* network, increased variability is expected in smaller samples, which are necessary for providing insights on random events occurring on single cell level. However, many 3D chromatin conformation assays rely on input material from large cell populations, or face limitations related to resolution and quantification. Therefore, our lab sought to develop an assay that could address these matters; leading to the implementation of the Nodewalk assay, which has been optimised to approach the challenges related to small input material and quantifying precise frequencies of chromatin fibre interactions.

Unique features of the Nodewalk include an incorporated *in vitro* transcription step allowing for linear amplification (thus overcoming the need for extensive logarithmic PCRs characteristic of other methods such as 4C), which allows the Nodewalk assay to quantitatively detect pairwise interactions in very small cell populations (using input material corresponding to genomic content of ~7 cells). Furthermore, using Nodewalk makes it possible to discover 3-way contacts by inference, and we therefore argue that Nodewalk is highly suitable for the exploration of the dynamics of chromatin networks. Other targeted 3C-derived methods, such as the targeted circular chromosome conformation capture (4C)[311], multi-contact 4C (MC-4C)[220] and Tri-C[221] assays are also able to detect multiway contacts between chromatin fibres that represent single cell events, and have been used to explore interaction patterns denoting regulatory enhancer hubs and chromatin "rosette" formation between TADs[220,221]. While an advantage of these techniques lies in their ability to detect multiway contacts, from which it is possible to infer the dynamics of single alleles, the techniques require a large amount of input material (10^6, $10x10^6$, and $20x10^6$ cells respectively for 4C, Tri-C and MC-4C)[220,221,311]. Moreover, the size selection for larger concatemers in assays like Tri-C and MC-4C makes the approximation of pairwise vs. multiway contacts challenging. Indeed, previous results from our group have shown that simultaneous contacts between multiple chromatin fibres comprise rare events, which constitute only a few percent of the total reads generated by 4C high-throughput sequencing[82]. Considering our observation that the *MYC* interactome seems to be formed by

stochastic, mainly pairwise, interactions rather than multiway contacts, by extrapolation, other chromatin fibre networks might be distinguished by similar dynamics.

However, the Nodewalk technique also faces certain technical limitations. For instance, the T7 RNA polymerase might introduce errors due to base mismatch incorporation, although such events are rare and would most likely be corrected for in most cases (except for rare interactions) by multiple ligation events between more common contacts. There might also be restrictions during the reverse transcription reaction regarding the synthesis of longer cDNAs (longer than approximately 12kb according to the manufacturer's instructions). Generally, this should not impede Nodewalk analyses to a significant extent, as the thorough fragmentation by Nextera tagmentation generates fragments which are in the range of ca 200 bps up to 1-1.5 kbs. Nevertheless, it is possible that either the limitations during the cDNA synthesis step or the extensive tagmentation might affect the detection of rare, long multi-way concatemers. Finally, a potential limitation of the Nodewalk assay, using large input samples, is represented by potential underestimation of the recovery of bait alleles in the case that the same tagmented end is generated multiple times, due to a finite number of possible tagmentation ends of each chimeric fragment. A solution to this potential problem in larger input samples is to use unique molecular identifiers (UMIs) with the target-specific Nodewalk probes for quantitative purposes instead of using the approach based on ligation events.

Taken together, Nodewalk comprises a highly sensitive method for detection of pairwise interactions in very small cell populations, making it ideal for the exploration of dynamic chromatin fibre interactomes in small samples.

5.3 DISTINCT ROLES OF CTCF IN THE REGULATION OF CHROMATIN FIBRE INTERACTIONS AND GENE GATING OF *MYC*

CTCF is an evolutionarily conserved transcription factor and is mainly known as a master orchestrator of the genome. In Paper II, we have uncovered a novel, noncanonical, role of CTCF in gene regulation, as a *CCAT1*-specific CTCFBS is involved in the regulation of OSE-mediated gene gating of *MYC* mRNA transcripts[147], thus post-transcriptionally increasing its expression. Several lines of evidence highlight that the role of CTCF in regulating gene gating is distinct from its previously well-documented function in mediating chromatin loop formation. A single OSE-specific CTCFBS was thus shown to promote proximities between *MYC* and the OSE to the nuclear periphery, however, mutation of the CTCFBS did not affect overall pattern of *MYC*-OSE interactions themselves, thus indicating that the CTCFBS does not directly mediate such contacts. Indeed, the *MYC* bait established tight interactions with several positions within the OSE even outside of the CTCFBS, which might be responsible to mediate OSE-*MYC* contacts in the absence of CTCF binding (ref). Our data also contrast previous observations showing that overexpression of the *CCAT1* eRNA facilitates chromatin interactions between *MYC* and the OSE through promoted

CTCF-binding[272]. We could thus not observe any effect on the binding of CTCF to the WT OSE region upon RNAse treatment, which might be due to differences in the level and the sequence of the *CCAT1* transcript variant in our model system. However, the OSE-specific CTCFBS might still indirectly play a role in promoting interactions between the OSE and *MYC*, by facilitating the recruitment of these genomic regions to a shared nuclear compartment, *i.e.* the nuclear periphery. Further supporting a role for CTCF in the recruitment of the OSE-*MYC* complex to the lamina, our findings established that the efficient binding of the pore-anchor AHCTF1 to the OSE[147] requires not only β-catenin but also CTCF-binding at the *CCAT1*-specific CTCFBS, and, thus strengthening the concept of CTCF indirectly promoting OSE-*MYC* interactions. Additionally, the repositioning of the OSE close to, but not exactly at, the nuclear periphery was found to be associated with WNT-dependent transcription of CCAT1 eRNA mediated by CTCF, highlighting the intricate and potentially sub-nuclear position-dependent role of CTCF in the gating process of *MYC*, as CTCF seems to regulate transcriptional states, potential for chromatin fibre interactions and nuclear localization of specific loci in the multi-step process of *MYC* gating. This is in line with previous evidence of CTCF being able to regulate the mobility of certain loci within the nucleus, as demonstrated by the CTCF/PARP1 mediated oscillating recruitment of circadian genes to the nuclear periphery[82]. Given the distinct roles of CTCF in the regulation of both transcriptional activity and chromatin mobility, as well as the oscillating complex formation between CTCF and its binding partner PARP1 in synchronized HCT116 cells[82], it is plausible to speculate that the gene gating of *MYC* – and potentially other genes as well – might be under circadian control. Preliminary results from our group support this notion, as circadian Nodewalk experiments indicate a variation in the number of interactors as well as the "flexibility" regarding the range of chromatin fibre interactions impinging on the OSE during the circadian cycle. Additionally, ChIP analyses have shown that both CTCF and PARP1 appear to bind to the OSE and the *MYC* promoter in an oscillating manner (unpublished, JV *et al.*).

Interestingly, CTCF-binding to the *CCAT1*-specific CTCFBS emerged as a focal point in the WNT signalling pathway that facilitated the juxtapositioning of the OSE to the nuclear periphery. Moreover, as WNT signalling is induced in breast cancer cells by *MYC* expression[312], it is plausible that there are intricate feed-back loops between the WNT pathway and transcriptional regulation of *MYC*. The convergence of different signalling pathways – such as the WNT cascade - on *MYC* might confer a way for cancer cells to increase their plasticity of transcriptional regulation, thus enabling diverse and rapid responses to external stimuli, consequentially conferring a selective advantage to drive disease in responses to changes in the tumour microenvironment.

5.4 DYSREGULATION OF ONCOGENES IN THE DEVELOPMENT OF CANCER

Oncogenes such as *MYC* have been extensively studied in malignancies. However, many aspects of their functional roles in the development and progression of cancer remain not

fully understood. The studies of the current thesis provide insights on the dynamics underlying the chromatin networks impinging on *MYC*, as well as on a novel function of CTCF in the regulation of *MYC* gene gating in colon cancer cells. These observations have thus uncovered several steps of the gene trafficking process to the nuclear pores, which underlies the gating of *MYC*, and which might potentially allow for future diagnostic and therapeutic strategies.

However, our findings also raise several questions: for instance, does gating of *MYC* - or other oncogenes - occur in additional cancer types, other than colon cancer? Considering that mutations or altered expression of *MYC* is involved in the pathophysiology of up to 70% of all human cancers[260], it is not unreasonable to consider that gene gating mechanisms of oncogenes may exist also in other types of malignancies. The nucleoporin NUP133 has been found to interact at NPCs with centromere protein F (CENP-F)[313], which has been proposed to comprise an oncogene implied to have functional roles in cancers such as prostate cancer[314], head and neck squamous carcinomas[315], and breast adenocarcinoma, where its expression is positively correlated with poor prognosis[316]. These observations thus strengthen the concept of potential gene gating processes of oncogenes in several types of cancers, rather than gene gating comprising a unique feature of colon cancer cells.

The findings in Paper I indicate that *MYC* acquires an ability to get dynamically activated by diverse environmental stimuli and signalling cascades during cancer development. This highlights the intricate role of chromatin fibre interactions between regulatory elements such as enhancers and super-enhancers (SEs) in tumorigenesis. SEs are commonly acquired by cancer cells during disease development[223], and in an interesting study by Schuijers *et al*. a single conserved CTCFBS located ~2 kb upstream of the *MYC* promoter was found to play a key role in regulating enhancer-promoter chromatin looping with cancer-specific SEs in several tumour types[270]. Moreover, abrogation of the CTCF binding site within the enhancer docking site close to *MYC* caused decreased CTCF binding, SE-interactions, *MYC* expression, and cell proliferation rates, thus emphasizing the central functions that SEs may have in tumorigenesis. As the cancer-specific gene gating of *MYC* in HCT116 cells is also mediated by an oncogenic SE, which is not present in normal colon epithelial cells[147], one might thus speculate that cancer-specific SEs in other tumours might mediate similar gating functions. It remains to be explored how factors that have been shown to mediate the activity of SEs on transcription might affect the gating process. SEs are thus characterized by strong enrichment of transcription factor and Mediator binding, and were found to be more sensitive to loss of co-activators, such as OCT4, than other genes during development[223]. This feature suggests that the transcriptional activity of SE-associated genes can be rapidly and preferentially reduced during differentiation or in response to specific stimuli. The observed sensitivity of ESC SEs to environmental fluxes contributes to the notion of SEs being able to diversify transcriptional regulation, and converge responses to several signalling pathways, in a highly dynamic manner, which may contribute to selective advantages amongst cancer cells to drive disease. It remains to be seen, however, if the Mediator complex is involved in the

OSE-mediated gating process and how the OSE-mediated effects on transcription are coordinated with its role in gene gating.

SEs occupied by master TFs may also be present at genes that define cell identity in more differentiated cell types[223], suggesting that SEs may play important roles also in normal cells. An additional question which remains unanswered is thus whether gene gating might exist in normal human cells? Given that *MYC* expression is under circadian control, it is plausible that the gating mechanism might also be regulated by the circadian clock. By extrapolation, it could therefore be conjectured that the circadian clock might regulate gene gating in some normal cell types, outside the scope of disease. For instance, our research group previously found that CTCF and PARP1 mediate the oscillating recruitment of circadian loci, such as *PARD3* and *TARDBP*, to the nuclear periphery, where these loci remain transcriptionally active for several hours prior to transcriptional attenuation[82]. This elicited the idea that the circadian loci might arrive at specks of transcriptionally active environments at the lamina, such as NPCs, which was succeeded by the work that would eventually uncover the gating of *MYC*[147]. However, given that circadian loci are trafficked to the lamina, where they seemingly first arrive at non-repressive environments, it is thus possible that also circadian genes might be recruited to nuclear pores for gene gating.

6 CONCLUSIONS

Paper I describes how the chromatin fibre network impinging on the oncogene *MYC* is characterized by highly dynamic and stochastic interactions with its flanking enhancers, located in both of its neighbouring TADs. *MYC* was thus found to be positioned at, or close to, an inter-TAD boundary region, and intriguingly emerged as the most important node (as indicated by the highest dynamic index) in both colon cancer cells and normal colon epithelial cells. This suggests that *MYC* screens for its enhancer contacts rather than the other way around. These findings are in contrast with the common notion that hubs of enhancers seek out promoters to concurrently activate transcription. Finally, through optimising the Nodewalk technique for very small input samples, we were able to uncover that the ensemble interactomes generated from large cell populations likely only represent virtual networks, rather than actual events at single ell level. This is demonstrated by the *MYC* network consisting of mainly pairwise, stochastic interactions, which may provide a diversification and labour division of enhancers in the regulation of transcription. Such plasticity in the regulation of oncogene transcription might thus represent a feature beneficial for cancer cells.

In Paper II, we provide the first genetic evidence of a gene gating mechanism in human colon cancer cells. The multistep process of *MYC* gating depicts how a single CTCF binding site within the colorectal oncogenic super-enhancer functions as an integral point of WNT signalling, and increases *MYC* expression post-transcriptionally. The binding of CTCF to the CTCFBS promotes transcription of *CCAT1* eRNA, which likely mediates the recruitment of the OSE to positions close to but not at the nuclear periphery, where CTCF and β-catenin collaborate to recruit AHCTF1 to the OSE, subsequently facilitating its tethering to the nuclear pore. Moreover, employing CRISPR-Cas9-based genetic editing of the CTFBS, we demonstrate how the CTCFBS confers a WNT-dependent growth advantage in WT colon cancer cells, compared to CTCFBS mutant clones.

In summary, our findings contribute to an increased understanding of mechanisms by which cancer cells may diversify their responses to external stimuli in order to regulate oncogene expression. Hence, we present evidence for a previously unknown noncanonical function of CTCF in gene gating. These findings could thus potentially provide new possibilities to target *MYC* diagnostically or therapeutically, without altering the normal functions of *MYC*. However, it remains to be seen whether gene gating of oncogenes occurs also in other tumours other than colon cancer.

7 POINTS OF PERSPECTIVE

It will be interesting to see whether or not the mechanism of gene gating holds true also for other cancers, cell types, and genes. Hence, future research addressing these questions is required. Additional studies will also be necessary in order to investigate potential diagnostic and therapeutic targets of the multistep gene gating process, which might constitute the most important conceivable future application of the current studies. The relevance of targeting MYC as a strategy for cancer therapies has long been acknowledged[260]. However, MYC has been considered an "undruggable" target[260,261], and so far only one drug, Omomyc, has entered clinical trials[305]. However, it was recently shown that Omomyc exerts its actions both through Omomyc dimers competing with the MYC:MAX complex for binding at E-boxes, and by preferentially binding to MAX rather than MYC, thus competing with MYC to form Omomyc:MAX complexes that inhibit MYC-mediated transcription[317]. As MAX has been shown to regulate the circadian clock[318], and perturbations of diurnal rhythms itself has been found to contribute to tumour development[319], it will be important to evaluate both short- and long-term results of clinical Omomyc trials on circadian homeostasis. Moreover, given the roles of the genome organizers CTCF and PARP1 in the regulation of circadian transcription and chromatin mobility, which includes recruitment of circadian genes to the nuclear periphery, there is a possibility of gene gating also being under circadian control. Although preliminary results from our group supports such a notion, further studies are required in order to explore this hypothesis.

8 ACKNOWLEDGEMENTS

This PhD has been a long journey, and I feel fortunate to have met so many wonderful and bright people who have helped me grow and made this time so treasurable. I would like to warmly thank all of you, especially:

First and foremost, my main supervisor, Associate Professor **Anita Göndör**, for all the support and guidance you have provided since the first day I joined your group. I greatly appreciate all the time and effort you have put into helping me grow, and I've cherished your advice and understanding in both scientific and more private matters during these years. You have always been there for discussions, whether about troubleshooting methods or providing guidance about non-work-related issues. I'm impressed with your comprehensive knowledge - there never seems to be a question which you can't answer or form an interesting hypothesis about -, and how much patience and kindness you show to those who want to learn from you. It is a privilege to have worked with a supervisor who cares so strongly for the science, but also for her team. I wish you the best for the future and many thanks for teaching me about so many wonderful topics, which I know will be useful both in clinics and science!

Professor **Rolf Ohlsson**, thank you for sharing your vast scientific knowledge and valuable ideas during group meetings. You have a brilliant mind, and it is always interesting to listen to your broad scientific understanding as well as non-scientific stories.

Professor **Camilla Björkegren**, thank you for agreeing to be my co-supervisor and for the scientific input regarding our projects. Thank you also for being so helpful and showing interest in the progress of my PhD during these years. Your broad knowledge and academic experience have been really inspiring to me.

Professor **Lars Holmgren**, thank you for your support as a co-supervisor during my PhD, and for all the help during the move to BioClinicum, especially for so kindly letting us work in CCK during this time. You are a brilliant scientist, and you always raise the spirit during OnkPat breakfast meetings!

I'd also like to thank Dr **Martin Jädersten** for agreeing to be my mentor and for all the valuable advice and opportunities during all these years. You are an amazing physician and scientist, and were the one who first made me fall in love with science by seeing how you could apply your research for your patients. You have been a great inspiration to me since I met you in medical school, many thanks again for giving me a chance to work with you and for your support.

Thank you also Professor **Maria Larsson**, for your mentorship during the PhD, and for the interesting discussions, support and advice that I received from you. Your academic experience is impressive, but also your kindness to those around you.

I'd also like to thank all of my lab friends from the bottom of my heart, even during hard times, it's because of you that these years have been the best of my life:

Jia Pei, thank you so much for all your help and valuable scientific discussions. You're an incredibly sweet and kind person, who is always ready help everyone around you. Your great spirits are contagious, and you always create a happy and friendly atmosphere in the group. Thank you for your support and our off-work mini adventures, they have really meant a lot to me.

Mirco, you're a truly amazing person, with so much kindness and remarkable knowledge. I feel like you know how to do anything in the lab, but even though you work so hard, you still always find time to help others (both in our group and in all of KI!) and make their days enjoyable with your company and humour. You're a great person, and I think that you will succeed at anything!

Barbara, it's been a pleasure to work with you. You are a great scientist with incredible skills, and who is a role model for everyone in the lab. Thank you for all the times that you have helped me and for teaching me so many new things when sharing your broad scientific knowledge. Your guidance and support during good and trying times has really moved me, and I wish you and your lovely family all the best for the future.

Felipe, you are an amazingly kind-hearted and positive person. I have really enjoyed working with you as neighbours on the lab bench and listening to your interesting stories. Your meticulousness when working in the lab is inspiring, but even more so how helpful and kind you are. Thank you for all the support, I will really miss working with you in the lab and exploring the city with you!

Ilyas, thank you so much for all the scientific advice and discussions. You always have interesting and helpful ideas to share about results or troubleshooting methods. You have a really bright mind, and with your passion for science I'm sure that you will go really far!

Andrej, it's been a pleasure to work with you! You have a really warm personality and a great sense of humour, always making the best jokes. With your broad knowledge on many topics, I enjoyed discussing many things with you.

Rashid, I was always impressed by your knowledge and programming skills. You are such a positive, helpful and friendly person, and I'm thankful for all your efforts to teach us bioinformatics!

Narsis, we have never met in person, but I really enjoyed your talks and discussions during our online group meetings. You are a really kind and humble person, and your scientific knowledge is striking!

I also would like to thank former members of our group: **Marta**, thank you for your support since day one and for introducing me to the lab. Your warmth and huge efforts to help all of us in the lab everyday – always with a big smile – really made the group feel like a big family. **Anna**, your skills and varied knowledge are really impressive, you are a really bright and friendly person, and I miss working with you in the lab in the early mornings. **Ilias**, thank

you for sharing your interesting scientific ideas and for always being kind and helpful. You have so much ambition, and your love for science is obvious. I wish you all the best for your new chapter in the U.S.! **Carolina**, you have so much kindness and talent. Thank you for all your help and what you taught me, especially in the beginning when I joined the group and felt I knew nothing. **Noriyuki**, you are such an incredible person! Your truly brilliant mind and nice sense of humour are amazing. I can't thank you enough for your enormous patience and kindness when teaching me Nodewalk. **Honglei**, it was always great to work with you, you have such a sweet and positive personality. Thank you for your support all these years and for helping me in the very beginning with the ISPLAs and the microscope. **Deeksha**, I really enjoyed or discussions on and off from work. You're such an intelligent, warm and supportive person, I don't think anyone else could be so positive and nice while trying to explain me bioinformatics for the umpteenth time. **Manos**, it was a great pleasure to work with you. You are one of the most positive and friendly people I've ever met! Thank you for all your efforts to teach us the amazing secrets of bioinformatics. **Hao**, it was a pleasure to work with you, you have such a encouraging and energetic attitude! **Shuangyang**, you are a really smart person. I am amazed with your ideas and your hard-working nature. I wish you all the best in the future!

I would also like to thank all **co-authors**, **collaborators**, and **scientific colleagues** for their great efforts and work to improve the knowledge about human disease, and I apologize to those whose work I failed to cite in this thesis.

A big thank you also to **Vicky**, **Pedro, Aarren** and all of the **Holmgren group**, you are always so kind and helpful, it's been a real pleasure to work as your "neighbour" and sharing the cell culture, you always greeted with a smile even in the early mornings.

Thank you also to my amazing clinical colleagues and friends **Soma**, **Erika**, **Katarina** and **Sam**, who I had the privilege to work with in the hospital during the Covid-19 pandemic. You are truly amazing doctors, scientist and mothers, and I really appreciated your support and working together. I will miss our talks during lunches or when running into each other in the hospital. You are really inspiring people and I wish for you all to get the ST-positions of your dreams in the (close) future!

I would also like to thank my close friends **Agnes**, **Jenni** and **Ulrika**. You are all so kind and warm-hearted people, and I feel truly lucky to have you as friends for more than 20 years now (and I look forward to many more adventures together!). Thank you for all your support in good and trying times, in such moments it means everything.

Last but not least, I would like to thank my family. **Äidilleni**, joka on aina minua tukenut. Kiitos että olet myös aina uskonut minuun ja olet niin iso osa elämässäni. Tiedän että voin aina sinun kanssa puhua tai kysyä tarvittaessa apua. **Pappa**, tack för allt stöd och hjälp du gett mig genom åren, oavsett vad jag har valt för väg har du alltid funnits där för mig. **Carlos**, gracias por apoyar mis sueños durante todos estos años. Hemos vivido momentos felices y difíciles, pero los hemos superado juntos, y me siento afortunada de conocerte. Gracias por

cuidarme cuando me enfermé, y por darme lo más preciado de mi vida: **Arvid**. Du är mitt hjärtas fröjd och glädje, och mitt livs största motivation. Jag är så tacksam och glad över att ha fått lära känna en sådan härlig och fin liten person, och det finns ingenting jag hellre vill än att du ska växa upp och alltid känna dig trygg, lycklig och älskad. Olet elämäni suurin ilo, ja tuen sinua aina.

9 REFERENCES

1. Feinberg, A.P. Phenotypic plasticity and the epigenetics of human disease. *Nature.* **447**, 433-440 (2007).
2. Perino, M. & Veenstra, G.J.C. Chromatin Control of Developmental Dynamics and Plasticity. *Developmental cell.* **38**, 610-620 (2016).
3. O'Brien-Ball, C. & Biddle, A. Reprogramming to developmental plasticity in cancer stem cells. *Developmental biology.* **430**, 266-274 (2017).
4. Flavahan, W.A., Gaskell, E. & Bernstein, B.E. Epigenetic plasticity and the hallmarks of cancer. *Science.* **357**(2017).
5. Waddington, Dynamic Systems, and Epigenetics. *Frontiers in behavioral neuroscience.* **10**(2016).
6. Epigenetics in comparative physiology. *Journal of experimental biology.* **218**(2015).
7. Pujadas, E. & Feinberg, A.P. Regulated noise in the epigenetic landscape of development and disease. *Cell.* **148**, 1123-1131 (2012).
8. Feinberg, A.P., Koldobskiy, M.A. & Göndör, A. Epigenetic modulators, modifiers and mediators in cancer aetiology and progression. *Nature reviews.* **17**, 284-299 (2016).
9. Feinberg, A.P., Ohlsson, R. & Henikoff, S. The epigenetic progenitor origin of human cancer. *Nature reviews.* **7**, 21-33 (2006).
10. Mack, S.C. *et al.* Epigenomic alterations define lethal CIMP-positive ependymomas of infancy. *Nature.* **506**, 445-450 (2014).
11. Waddington, C.H. The epigenotype. 1942. *International journal of epidemiology.* **41**, 10-13 (2012).
12. Noble, D. Conrad Waddington and the origin of epigenetics. *Journal of experimental biology.* **218**, 816-818 (2015).
13. Yadav, T., Quivy, J.-P. & Almouzni, G. Chromatin plasticity: A versatile landscape that underlies cell fate and identity. *Science.* **361**, 1332-1336 (2018).
14. Allis, C.D. & Jenuwein, T. The molecular hallmarks of epigenetic control. *Nature reviews.* **17**, 487-500 (2016).
15. Zhou, K., Gaullier, G. & Luger, K. Nucleosome structure and dynamics are coming of age. *Nature structural & molecular biology.* **26**, 3-13 (2019).
16. Lawrence, M., Daujat, S. & Schneider, R. Lateral Thinking: How Histone Modifications Regulate Gene Expression. *Trends in genetics.* **32**, 42-56 (2016).
17. Suganuma, T. & Workman, J.L. Signals and combinatorial functions of histone modifications. *Annual review of biochemistry.* **80**, 473-499 (2011).
18. Klein, D.C. & Hainer, S.J. Chromatin regulation and dynamics in stem cells. *Current topics in developmental biology.* **138**, 1-71 (2020).

19. Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F. & Richmond, T.J. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature.* **389**, 251-260 (1997).

20. Hergeth, S.P. & Schneider, R. The H1 linker histones: multifunctional proteins beyond the nucleosomal core particle. *EMBO reports.* **16**, 1439-1453 (2015).

21. Luger, K., Dechassa, M.L. & Tremethick, D.J. New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? *Nature reviews.* **13**, 436-447 (2012).

22. Bowman, G.D. & Poirier, M.G. Post-translational modifications of histones that influence nucleosome dynamics. *Chemical reviews.* **115**, 2274-2295 (2015).

23. Allis, C.D. & Grewal, S.I. Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. *Science.* **293**, 1150-1155 (2001).

24. Strahl, B.D. & Allis, C.D. The language of covalent histone modifications. *Nature.* **403**, 41-45 (2000).

25. Rothbart, S.B. & Strahl, B.D. Interpreting the language of histone and DNA modifications. *Biochimica et biophysica acta* **1839**, 627-643 (2014).

26. Atlasi, Y. & Stunnenberg, H.G. The interplay of epigenetic marks during stem cell differentiation and development. *Nature reviews.* **18**, 643-658 (2017).

27. Nishibuchi, G. & Déjardin, J. The molecular basis of the organization of repetitive DNA-containing constitutive heterochromatin in mammals. *Chromosome research.* **25**, 77-87 (2017).

28. Jamieson, K. *et al.* Loss of HP1 causes depletion of H3K27me3 from facultative heterochromatin and gain of H3K27me2 at constitutive heterochromatin. *Genome research.* **26**, 97-107 (2016).

29. Batut, P.J. *et al.* Genome organization controls transcriptional dynamics during development. *Science.* **375**, 566-570 (2022).

30. Celona, B. *et al.* Substantial histone reduction modulates genomewide nucleosomal occupancy and global transcriptional output. *PLoS biology.* **9**, e1001086 (2011).

31. Feser, J. *et al.* Elevated histone expression promotes life span extension. *Molecular cell.* **39**, 724-735 (2010).

32. O'Sullivan, R.J., Kubicek, S., Schreiber, S.L. & Karlseder, J. Reduced histone biosynthesis and chromatin changes arising from a damage signal at telomeres. *Nature structural & molecular biology.* **17**, 1218-1225 (2010).

33. Almeida, R. *et al.* Chromatin conformation regulates the coordination between DNA replication and transcription. *Nature communications.* **9**, 1590 (2018).

34. Venkatesh, S. & Workman, J.L. Histone exchange, chromatin structure and the regulation of transcription. *Nature reviews.* **16**, 178-189 (2015).

35. Cairns, B.R. *et al.* RSC, an essential, abundant chromatin-remodeling complex. *Cell* **87**, 1249-1260 (1996).

36. Ye, Y. *et al.* Structure of the RSC complex bound to the nucleosome. *Science.* **366**, 838-843 (2019).

37. Hartley, P.D. & Madhani, H.D. Mechanisms that specify promoter nucleosome location and identity. *Cell* **137**, 445-458 (2009).

38. Ganguli, D., Chereji, R.V., Iben, J.R., Cole, H.A. & Clark, D.J. RSC-dependent constructive and destructive interference between opposing arrays of phased nucleosomes in yeast. *Genome research.* **24**, 1637-1649 (2014).

39. Parnell, T.J., Huff, J.T. & Cairns, B.R. RSC regulates nucleosome positioning at Pol II genes and density at Pol III genes. *The EMBO journal.* **27**, 100-110 (2008).

40. Ranjan, A. *et al.* Nucleosome-free region dominates histone acetylation in targeting SWR1 to promoters for H2A.Z replacement. *Cell* **154**, 1232-1245 (2013).

41. Draker, R. *et al.* A combination of H2A.Z and H4 acetylation recruits Brd2 to chromatin during transcriptional activation. *PLOS genetics.* **8**, e1003047 (2012).

42. Göndör, A. & Ohlsson, R. Enhancer functions in three dimensions: beyond the flat world perspective. *F1000Research.* **7**(2018).

43. Reddy, K.L. & Feinberg, A.P. Higher order chromatin organization in cancer. *Seminars in cancer biology.* **23**, 109-115 (2013).

44. Zink, D., Fischer, A.H. & Nickerson, J.A. Nuclear structure in cancer cells. *Nature reviews.* **4**, 677-687 (2004).

45. Cremer, T. & Zakhartchenko, V. Nuclear architecture in developmental biology and cell specialisation. *Reproduction, fertility, and development.* **23**, 94-106 (2011).

46. Pederson, T. The nucleolus. *Cold Spring Harbor perspectives in biology.* **3**(2011).

47. Cremer, T. & Cremer, M. Chromosome territories. *Cold Spring Harbor perspectives in biology.* **2**, a003889 (2010).

48. Cremer, T. & Cremer, C. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nature reviews.* **2**, 292-301 (2001).

49. Parada, L.A., McQueen, P.G. & Misteli, T. Tissue-specific spatial organization of genomes. *Genome biology.* **5**, R44 (2004).

50. Bickmore, W.A. & van Steensel, B. Genome architecture: domain organization of interphase chromosomes. *Cell.* **152**, 1270-1284 (2013).

51. Zheng, H. & Xie, W. The role of 3D genome organization in development and cell differentiation. *Nature reviews.* **20**, 535-550 (2019).

52. Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science.* **326**, 289-293 (2009).

53. Fortin, J.-P. & Hansen, K.D. Reconstructing A/B compartments as revealed by Hi-C using long-range correlations in epigenetic data. *Genome biology.* **16**, 180 (2015).

54. Wang, L. *et al.* Histone Modifications Regulate Chromatin Compartmentalization by Contributing to a Phase Separation Mechanism. *Molecular cell.* **76**, 646-659.e6 (2019).

55. Wang, S. *et al.* Spatial organization of chromatin domains and compartments in single chromosomes. *Science.* **353**, 598-602 (2016).

56. Bintu, B. *et al.* Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. *Science.* **362**(2018).

57. Lafontaine, D.L.J., Riback, J.A., Bascetin, R. & Brangwynne, C.P. The nucleolus as a multiphase liquid condensate. *Nature reviews.* **22**, 165-182 (2021).

58. Singh, P.B. & Newman, A.G. On the relations of phase separation and Hi-C maps to epigenetics. *Royal Society open science.* **7**, 191976 (2020).

59. Erdel, F. & Rippe, K. Formation of Chromatin Subcompartments by Phase Separation. *Biophysical journal* **114**, 2262-2270 (2018).

60. Ghosh, R.P. & Meyer, B.J. Spatial Organization of Chromatin: Emergence of Chromatin Structure During Development. *Annual review of cell and developmental biology.* **37**, 199-232 (2021).

61. Bhat, P., Honson, D. & Guttman, M. Nuclear compartmentalization as a mechanism of quantitative control of gene expression. *Nature reviews.* **22**, 653-670 (2021).

62. Papantonis, A. & Cook, P.R. Transcription factories: genome organization and gene regulation. *Chemical reviews.* **113**, 8683-8705 (2013).

63. Sutherland, H. & Bickmore, W.A. Transcription factories: gene expression in unions? *Nature reviews.* **10**, 457-466 (2009).

64. Galganski, L., Urbanek, M.O. & Krzyzosiak, W.J. Nuclear speckles: molecular organization, biological function and role in disease. *Nucleic acids research.* **45**, 10350-10368 (2017).

65. Pontvianne, F. *et al.* Identification of Nucleolus-Associated Chromatin Domains Reveals a Role for the Nucleolus in 3D Organization of the A. thaliana Genome. *Cell reports.* **16**, 1574-1587 (2016).

66. Olivares-Chauvet, P., Fennessy, D., Jackson, D.A. & Maya-Mendoza, A. Innate structure of DNA foci restricts the mixing of DNA from different chromosome territories. *PloS one.* **6**, e27527 (2011).

67. Dixon, J.R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature.* **485**, 376-380 (2012).

68. Rao, S.S.P. *et al.* A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665-1680 (2014).

69. van Steensel, B. & Belmont, A.S. Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. *Cell.* **169**, 780-791 (2017).

70. Luperchio, T.R., Wong, X. & Reddy, K.L. Genome regulation at the peripheral zone: lamina associated domains in development and disease. *Current opinion in genetics & development.* **25**, 50-61 (2014).

71. Wen, B., Wu, H., Shinkai, Y., Irizarry, R.A. & Feinberg, A.P. Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nature genetics.* **41**, 246-250 (2009).

72. Peric-Hupkes, D. *et al.* Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Molecular cell.* **38**, 603-613 (2010).

73. Ong, C.-T. & Corces, V.G. CTCF: an architectural protein bridging genome topology and function. *Nature reviews.* **15**, 234-246 (2014).

74. Guelen, L. *et al.* Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature.* **453**, 948-951 (2008).

75. Trojer, P. & Reinberg, D. Facultative heterochromatin: is there a distinctive molecular signature? *Molecular cell.* **28**, 1-13 (2007).

76. Lochs, S.J.A., Kefalopoulou, S. & Kind, J. Lamina Associated Domains and Gene Regulation in Development and Cancer. *Cells.* **8**(2019).

77. Pongubala, J.M.R. & Murre, C. Spatial Organization of Chromatin: Transcriptional Control of Adaptive Immune Cell Development. *Frontiers in immunology.* **12**, 633825 (2021).

78. Hübner, M.R., Eckersley-Maslin, M.A. & Spector, D.L. Chromatin organization and transcriptional regulation. *Current opinion in genetics & development.* **23**, 89-95 (2013).

79. Reddy, K.L., Zullo, J.M., Bertolino, E. & Singh, H. Transcriptional repression mediated by repositioning of genes to the nuclear lamina. *Nature.* **452**, 243-247 (2008).

80. Yokochi, T. *et al.* G9a selectively represses a class of late-replicating genes at the nuclear periphery. *Proceedings of the National Academy of Sciences of the United States of America.* **106**, 19363-19368 (2009).

81. Kohwi, M., Lupton, J.R., Lai, S.-L., Miller, M.R. & Doe, C.Q. Developmentally regulated subnuclear genome reorganization restricts neural progenitor competence in Drosophila. *Cell.* **152**, 97-108 (2013).

82. Zhao, H. *et al.* PARP1- and CTCF-Mediated Interactions between Active and Repressed Chromatin at the Lamina Promote Oscillating Transcription. *Molecular cell.* **59**, 984-997 (2015).

83. Akhtar, W. *et al.* Chromatin position effects assayed by thousands of reporters integrated in parallel. *Cell.* **154**, 914-927 (2013).

84. Pollex, T. & Heard, E. Nuclear positioning and pairing of X-chromosome inactivation centers are not primary determinants during initiation of random X-inactivation. *Nature genetics.* **51**, 285-295 (2019).

85. Kumaran, R.I. & Spector, D.L. A genetic locus targeted to the nuclear periphery in living cells maintains its transcriptional competence. *The journal of cell biology.* **180**, 51-65 (2008).

86. Zimmer, C. & Fabre, E. Chromatin mobility upon DNA damage: state of the art and remaining questions. *Current genetics.* **65**, 1-9 (2019).

87. Lamm, N., Rogers, S. & Cesare, A.J. Chromatin mobility and relocation in DNA repair. *Trends in cell biology.* **31**, 843-855 (2021).

88. Heun, P., Laroche, T., Shimada, K., Furrer, P. & Gasser, S.M. Chromosome dynamics in the yeast interphase nucleus. *Science.* **294**, 2181-2186 (2001).

89. Gasser, S.M. Visualizing chromatin dynamics in interphase nuclei. *Science.* **296**, 1412-1416 (2002).

90. DeRisi, J.L., Iyer, V.R. & Brown, P.O. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science.* **278**, 680-686 (1997).

91. Gu, B. *et al.* Transcription-coupled changes in nuclear mobility of mammalian cis-regulatory elements. *Science.* **359**, 1050-1055 (2018).

92. Robert Finestra, T. & Gribnau, J. X chromosome inactivation: silencing, topology and reactivation. *Current opinion in cell biology.* **46**, 54-61 (2017).

93. Chen, C.-K. *et al.* Xist recruits the X chromosome to the nuclear lamina to enable chromosome-wide silencing. *Science.* **354**, 468-472 (2016).

94. Giorgetti, L. *et al.* Structural organization of the inactive X chromosome in the mouse. *Nature.* **535**, 575-579 (2016).

95. Capelson, M. & Hetzer, M.W. The role of nuclear pores in gene regulation, development and disease. *EMBO reports.* **10**, 697-705 (2009).

96. Yang, J. *et al.* Gating pluripotency via nuclear pores. *Trends in molecular medicine* **20**, 1-7 (2014).

97. Sun, J., Shi, Y. & Yildirim, E. The Nuclear Pore Complex in Cell Type-Specific Chromatin Structure and Gene Regulation. *Trends in genetics.* **35**, 579-588 (2019).

98. Alber, F. *et al.* The molecular architecture of the nuclear pore complex. *Nature.* **450**, 695-701 (2007).

99. Lin, D.H. & Hoelz, A. The Structure of the Nuclear Pore Complex (An Update). *Annual review of biochemistry.* **88**, 725-783 (2019).

100. Maimon, T., Elad, N., Dahan, I. & Medalia, O. The human nuclear pore complex as revealed by cryo-electron tomography. *Structure.* **20**, 998-1006 (2012).

101. Rajoo, S., Vallotton, P., Onischenko, E. & Weis, K. Stoichiometry and compositional plasticity of the yeast nuclear pore complex revealed by quantitative fluorescence microscopy. *Proceedings of the National Academy of Sciences of the United States of America.* **115**, E3969-E3977 (2018).

102. Kim, S.J. *et al.* Integrative structure and functional anatomy of a nuclear pore complex. *Nature.* **555**, 475-482 (2018).

103. Mohr, D., Frey, S., Fischer, T., Güttler, T. & Görlich, D. Characterisation of the passive permeability barrier of nuclear pore complexes. *The EMBO journal.* **28**, 2541-2553 (2009).

104. Gomar-Alba, M. & Mendoza, M. Modulation of Cell Identity by Modification of Nuclear Pore Complexes. *Frontiers in genetics.* **10**, 1301 (2019).

105. Natalizio, B.J. & Wente, S.R. Postage for the messenger: designating routes for nuclear mRNA export. *Trends in cell biology.* **23**, 365-373 (2013).

106. De Magistris, P. The Great Escape: mRNA Export through the Nuclear Pore Complex. *International journal of molecular sciences.* **22**(2021).

107. Tutucci, E. & Stutz, F. Keeping mRNPs in check during assembly and nuclear export. *Nature reviews.* **12**, 377-384 (2011).

108. Schmid, M. & Jensen, T.H. Transcription-associated quality control of mRNP. *Biochimica et biophysica acta* **1829**, 158-168 (2013).

109. Burns, L.T. & Wente, S.R. From hypothesis to mechanism: uncovering nuclear pore complex links to gene expression. *Molecular and cellular biology : MCB.* **34**, 2114-2120 (2014).

110. Ibarra, A. & Hetzer, M.W. Nuclear pore proteins and the control of genome functions. *Genes & development.* **29**, 337-349 (2015).

111. Kuhn, T.M., Pascual-Garcia, P., Gozalo, A., Little, S.C. & Capelson, M. Chromatin targeting of nuclear pore proteins induces chromatin decondensation. *The journal of cell biology.* **218**, 2945-2961 (2019).

112. Capelson, M. *et al.* Chromatin-bound nuclear pore components regulate gene expression in higher eukaryotes. *Cell.* **140**, 372-383 (2010).

113. Rasala, B.A., Orjalo, A.V., Shen, Z., Briggs, S. & Forbes, D.J. ELYS is a dual nucleoporin/kinetochore protein required for nuclear pore assembly and proper cell division. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 17801-17806 (2006).

114. Gao, X., Yu, S., Guan, Y., Shen, Y. & Xu, L. Nucleoporin 50 mediates Kcna4 transcription to regulate cardiac electrical activity. *Journal of cell science.* **134**(2021).

115. Rabut, G., Doye, V. & Ellenberg, J. Mapping the dynamic organization of the nuclear pore complex inside single living cells. *Nature cell biology.* **6**, 1114-1121 (2004).

116. Griffis, E.R., Altan, N., Lippincott-Schwartz, J. & Powers, M.A. Nup98 is a mobile nucleoporin with transcription-dependent dynamics. *Molecular biology of the cell.* **13**, 1282-1297 (2002).

117. Bush, N.G., Evans-Roberts, K. & Maxwell, A. DNA Topoisomerases. *Ecosal plus.* **6**(2015).

118. Chow, K.-H., Elgort, S., Dasso, M. & Ullman, K.S. Two distinct sites in Nup153 mediate interaction with the SUMO proteases SENP1 and SENP2. *Nucleus.* **3**, 349-358 (2012).

119. Li, M., Pokharel, S., Wang, J.-T., Xu, X. & Liu, Y. RECQ5-dependent SUMOylation of DNA topoisomerase I prevents transcription-associated genome instability. *Nature communications.* **6**, 6720 (2015).

120. Casolari, J.M. *et al.* Genome-wide localization of the nuclear transport machinery couples transcriptional status and nuclear organization. *Cell.* **117**, 427-439 (2004).

121. Ibarra, A., Benner, C., Tyagi, S., Cool, J. & Hetzer, M.W. Nucleoporin-mediated regulation of cell identity genes. *Genes & development.* **30**, 2253-2258 (2016).

122. Su, Y. *et al.* Post-translational modification localizes MYC to the nuclear pore basket to regulate a subset of target genes involved in cellular responses to environmental signals. *Genes & development.* **32**, 1398-1419 (2018).

123. Kalverda, B., Pickersgill, H., Shloma, V.V. & Fornerod, M. Nucleoporins directly stimulate expression of developmental and cell-cycle genes inside the nucleoplasm. *Cell.* **140**, 360-371 (2010).

124. Arib, G. & Akhtar, A. Multiple facets of nuclear periphery in gene expression control. *Current opinion in cell biology.* **23**, 346-353 (2011).

125. Lainé, J.-P., Singh, B.N., Krishnamurthy, S. & Hampsey, M. A physiological role for gene loops in yeast. *Genes & development.* **23**, 2604-2609 (2009).

126. Tan-Wong, S.M., Wijayatilake, H.D. & Proudfoot, N.J. Gene loops function to maintain transcriptional memory through interaction with the nuclear pore complex. *Genes & development.* **23**, 2610-2624 (2009).

127. Ferraro, T. Transcriptional Memory in the Drosophila Embryo. *Current biology.* **26**, 212-218 (2015).

128. Guan, T. *et al.* Nup50, a nucleoplasmically oriented nucleoporin with a role in nuclear protein export. *Molecular and cellular biology : MCB.* **20**, 5619-5630 (2000).

129. Ori, A. *et al.* Cell type-specific nuclear pores: a case in point for context-dependent stoichiometry of molecular machines. *Molecular systems biology.* **9**, 648 (2013).

130. Cho, A.R. *et al.* Tissue-specific expression and subcellular localization of ALADIN, the absence of which causes human triple A syndrome. *Experimental and molecular medicine.* **41**, 381-386 (2009).

131. Olsson, M., Schéele, S. & Ekblom, P. Limited expression of nuclear pore membrane glycoprotein 210 in cell lines and tissues suggests cell-type specific nuclear pores in metazoans. *Experimental cell research.* **292**, 359-370 (2004).

132. D'Angelo, M.A., Gomez-Cavazos, J.S., Mei, A., Lackner, D.H. & Hetzer, M.W. A change in nuclear pore complex composition regulates cell differentiation. *Developmental cell.* **22**, 446-458 (2012).

133. Jacinto, F.V., Benner, C. & Hetzer, M.W. The nucleoporin Nup153 regulates embryonic stem cell pluripotency through gene silencing. *Genes & development.* **29**, 1224-1238 (2015).

134. Toda, T. *et al.* Nup153 Interacts with Sox2 to Enable Bimodal Gene Regulation and Maintenance of Neural Progenitor Cells. *Cell stem cell.* **21**, 618-634.e7 (2017).

135. Sood, V. & Brickner, J.H. Nuclear pore interactions with the genome. *Current opinion in genetics & development.* **25**, 43-49 (2014).

136. Brickner, D.G. *et al.* Transcription factor binding to a DNA zip code controls interchromosomal clustering at the nuclear periphery. *Developmental cell.* **22**, 1234-1246 (2012).

137. Brickner, J.H. & Walter, P. Gene recruitment of the activated INO1 locus to the nuclear membrane. *PLoS biology.* **2**, e342 (2004).

138. Brickner, D.G. & Brickner, J.H. Interchromosomal clustering of active genes at the nuclear pore complex. *Nucleus.* **3**, 487-492.

139. Van de Vosse, D.W., Wan, Y., Wozniak, R.W. & Aitchison, J.D. Role of the nuclear envelope in genome organization and gene expression. *Wiley interdisciplinary reviews.* **3**, 147-166.

140. Breuer, M. & Ohkura, H. A negative loop within the nuclear pore complex controls global chromatin organization. *Genes & development.* **29**, 1789-1794 (2015).

141. García-Benítez, F., Gaillard, H. & Aguilera, A. Physical proximity of chromatin to nuclear pores prevents harmful R loop accumulation contributing to maintain genome stability. *Proceedings of the National Academy of Sciences of the United States of America.* **114**, 10942-10947 (2017).

142. Rowley, M.J. & Corces, V.G. Organizational principles of 3D genome architecture. *Nature reviews.* **19**, 789-800 (2018).

143. Blobel, G. Gene gating: a hypothesis. *Proceedings of the National Academy of Sciences of the United States of America* **82**, 8527-8529 (1985).

144. Schmid, M. *et al.* Nup-PI: the nucleopore-promoter interaction of genes in yeast. *Molecular cell.* **21**, 379-391 (2006).

145. Ben-Yishay, R., Ashkenazy, A.J. & Shav-Tal, Y. Dynamic Encounters of Genes and Transcripts with the Nuclear Pore. *Trends in genetics.* **32**, 419-431 (2016).

146. Pascual-Garcia, P. *et al.* Metazoan Nuclear Pores Provide a Scaffold for Poised Genes and Mediate Induced Enhancer-Promoter Contacts. *Molecular cell.* **66**, 63-76.e6 (2017).

147. Scholz, B.A. *et al.* WNT signaling and AHCTF1 promote oncogenic MYC expression through super-enhancer-mediated gene gating. *Nature genetics.* **51**, 1723-1731 (2019).

148. Sun, H.B., Shen, J. & Yokota, H. Size-dependent positioning of human chromosomes in interphase nuclei. *Biophysical journal* **79**, 184-190 (2000).

149. Nurk, S. *et al.* The complete sequence of a human genome. *Science.* **376**, 44-53 (2022).

150. Cuomo, C.A. & Birren, B.W. The fungal genome initiative and lessons learned from genome sequencing. *Methods in enzymology.* **470**, 833-855 (2010).

151. Bennett, M.D., Leitch, I.J., Price, H.J. & Johnston, J.S. Comparisons with Caenorhabditis (approximately 100 Mb) and Drosophila (approximately 175 Mb) using flow cytometry show genome size in Arabidopsis to be approximately 157 Mb and thus approximately 25% larger than the Arabidopsis genome initiative estimate of approximately 125 Mb. *Annals of botany.* **91**, 547-557 (2003).

152. Gomar-Alba, M. *et al.* Nuclear pore complex acetylation regulates mRNA export and cell cycle commitment in budding yeast. *The EMBO journal.*, e2021110271 (2022).

153. Kumar, A. *et al.* Daughter-cell-specific modulation of nuclear pore complexes controls cell cycle entry during asymmetric division. *Nature cell biology.* **20**, 432-442 (2018).

154. Shamloo, B. & Usluer, S. p21 in Cancer Research. *Cancers.* **11**(2019).

155. Singer, S. *et al.* Nuclear pore component Nup98 is a potential tumor suppressor and regulates posttranscriptional expression of select p53 target genes. *Molecular cell.* **48**, 799-810 (2012).

156. Chen, M. *et al.* Nucleoporin TPR promotes tRNA nuclear export and protein synthesis in lung cancer cells. *PLOS genetics.* **17**, e1009899 (2021).

157. Nataraj, N.B. *et al.* Nucleoporin-93 reveals a common feature of aggressive breast cancers: robust nucleocytoplasmic transport of transcription factors. *Cell reports.* **38**, 110418 (2022).

158. Göndör, A. & Ohlsson, R. Chromosome crosstalk in three dimensions. *Nature.* **461**, 212-217 (2009).

159. Klenova, E.M. *et al.* CTCF, a conserved nuclear factor required for optimal transcriptional activity of the chicken c-myc gene, is an 11-Zn-finger protein differentially expressed in multiple forms. *Molecular and cellular biology : MCB.* **13**, 7612-7624 (1993).

160. Hashimoto, H. *et al.* Structural Basis for the Versatile and Methylation-Dependent Binding of CTCF to DNA. *Molecular cell.* **66**, 711-720.e3 (2017).

161. Yu, W. *et al.* Poly(ADP-ribosyl)ation regulates CTCF-dependent chromatin insulation. *Nature genetics.* **36**, 1105-1110 (2004).

162. MacPherson, M.J., Beatty, L.G., Zhou, W., Du, M. & Sadowski, P.D. The CTCF insulator protein is posttranslationally modified by SUMO. *Molecular and cellular biology : MCB.* **29**, 714-725 (2009).

163. Merkenschlager, M. & Nora, E.P. CTCF and Cohesin in Genome Folding and Transcriptional Gene Regulation. *Annual review of genomics and human genetics* **17**, 17-43 (2016).

164. Zampieri, M. *et al.* ADP-ribose polymers localized on Ctcf-Parp1-Dnmt1 complex prevent methylation of Ctcf target sites. *The biochemical journal.* **441**, 645-652 (2012).

165. Chernukhin, I. *et al.* CTCF interacts with and recruits the largest subunit of RNA polymerase II to CTCF target sites genome-wide. *Molecular and cellular biology : MCB.* **27**, 1631-1648 (2007).

166. Zuin, J. *et al.* Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *PNAS : Proceedings of the National Academy of Sciences of the United States of America.* **111**, 996-1001 (2014).

167. Barrington, C. *et al.* Enhancer accessibility and CTCF occupancy underlie asymmetric TAD architecture and cell type specific genome topology. *Nature communications.* **10**, 2908 (2019).

168. Dekker, J., Rippe, K., Dekker, M. & Kleckner, N. Capturing chromosome conformation. *Science.* **295**, 1306-1311 (2002).

169. McCord, R.P., Kaplan, N. & Giorgetti, L. Chromosome Conformation Capture and Beyond: Toward an Integrative View of Chromosome Structure and Function. *Molecular cell.* **77**, 688-708 (2020).

170. Belton, J.-M. *et al.* Hi-C: a comprehensive technique to capture the conformation of genomes. *Methods.* **58**, 268-276 (2012).

171. Lafontaine, D.L., Yang, L., Dekker, J. & Gibcus, J.H. Hi-C 3.0: Improved Protocol for Genome-Wide Chromosome Conformation Capture. *Curr Protoc* **1**, e198 (2021).

172. Nora, E.P. *et al.* Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature.* **485**, 381-385 (2012).

173. Flyamer, I.M. *et al.* Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. *Nature.* **544**, 110-114 (2017).

174. Zhan, Y. *et al.* Reciprocal insulation analysis of Hi-C data shows that TADs represent a functionally but not structurally privileged scale in the hierarchical folding of chromosomes. *Genome research.* **27**, 479-490 (2017).

175. de Wit, E. TADs as the Caller Calls Them. *Journal of molecular biology.* **432**, 638-642 (2020).

176. Phillips-Cremins, J.E. *et al.* Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* **153**, 1281-1295 (2013).

177. Gibcus, J.H. & Dekker, J. The hierarchy of the 3D genome. *Molecular cell.* **49**, 773-782 (2013).

178. Hansen, A.S., Cattoglio, C., Darzacq, X. & Tjian, R. Recent evidence that TADs and chromatin loops are dynamic structures. *Nucleus.* **9**, 20-32 (2018).

179. Rao, S.S.P. *et al.* Cohesin Loss Eliminates All Loop Domains. *Cell.* **171**, 305-320.e24 (2017).

180. Hansen, A.S. CTCF as a boundary factor for cohesin-mediated loop extrusion: evidence for a multi-step mechanism. *Nucleus.* **11**, 132-148 (2020).

181. Fudenberg, G. *et al.* Formation of Chromosomal Domains by Loop Extrusion. *Cell reports.* **15**, 2038-2049 (2016).

182. Sanborn, A.L. *et al.* Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proceedings of the National Academy of Sciences of the United States of America.* **112**, E6456-E6465 (2015).

183. Racko, D., Benedetti, F., Dorier, J. & Stasiak, A. Transcription-induced supercoiling as the driving force of chromatin loop extrusion during formation of TADs in interphase chromosomes. *Nucleic acids research.* **46**, 1648-1660 (2018).

184. Guo, Y. *et al.* CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function. *Cell.* **162**, 900-910 (2015).

185. Busslinger, G.A. *et al.* Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl. *Nature.* **544**, 503-507 (2017).

186. Xi, W. & Beer, M.A. Loop competition and extrusion model predicts CTCF interaction specificity. *Nature communications.* **12**, 1046 (2021).

187. Nora, E.P. *et al.* Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell.* **169**, 930-944.e22 (2017).

188. Dequeker, B.J.H. *et al.* MCM complexes are barriers that restrict cohesin-mediated loop extrusion. *Nature.* **606**, 197-203 (2022).

189. Wutz, G. *et al.* Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. *The EMBO journal.* **36**, 3573-3599 (2017).

190. Soochit, W. *et al.* CTCF chromatin residence time controls three-dimensional genome organization, gene expression and DNA methylation in pluripotent cells. *Nature cell biology.* **23**, 881-893 (2021).

191. Narendra, V. *et al.* CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. *Science.* **347**, 1017-1021 (2015).

192. de Wit, E. *et al.* CTCF Binding Polarity Determines Chromatin Looping. *Molecular cell.* **60**, 676-684 (2015).

193. Haarhuis, J.H.I. *et al.* The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. *Cell.* **169**, 693-707.e14 (2017).

194. Wang, Q., Sun, Q., Czajkowsky, D.M. & Shao, Z. Sub-kb Hi-C in D. melanogaster reveals conserved characteristics of TADs between insect and mammalian cells. *Nature communications.* **9**, 188 (2018).

195. Gabriele, M. *et al.* Dynamics of CTCF- and cohesin-mediated chromatin looping revealed by live-cell imaging. *Science.* **376**, 496-501 (2022).

196. Ulianov, S.V. *et al.* Order and stochasticity in the folding of individual Drosophila genomes. *Nature communications.* **12**, 41 (2021).

197. Rhodes, J.D.P. *et al.* Cohesin Disrupts Polycomb-Dependent Chromosome Interactions in Embryonic Stem Cells. *Cell reports.* **30**, 820-835.e10 (2020).

198. Hanssen, L.L.P. *et al.* Tissue-specific CTCF-cohesin-mediated chromatin architecture delimits enhancer interactions and function in vivo. *Nature cell biology.* **19**, 952-961 (2017).

199. Schwarzer, W. *et al.* Two independent modes of chromatin organization revealed by cohesin removal. *Nature.* **551**, 51-56 (2017).

200. Bolt, C.C. *et al.* Context-dependent enhancer function revealed by targeted inter-TAD relocation. *Nature communications.* **13**, 3488 (2022).

201. Calo, E. & Wysocka, J. Modification of enhancer chromatin: what, how, and why? *Molecular cell.* **49**, 825-837 (2013).

202. Shlyueva, D., Stampfel, G. & Stark, A. Transcriptional enhancers: from properties to genome-wide predictions. *Nature reviews.* **15**, 272-286 (2014).

203. Andersson, R. & Sandelin, A. Determinants of enhancer and promoter activities of regulatory elements. *Nature reviews.* **21**, 71-87 (2020).

204. Koch, F. & Andrau, J.-C. Initiating RNA polymerase II and TIPs as hallmarks of enhancer activity and tissue-specificity. *Transcription.* **2**, 263-268.

205. Djebali, S. *et al.* Landscape of transcription in human cells. *Nature.* **489**, 101-108 (2012).

206. Heintzman, N.D. *et al.* Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature.* **459**, 108-112 (2009).

207. Bonn, S. *et al.* Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. *Nature genetics.* **44**, 148-156 (2012).

208. Liu, W. *et al.* Brd4 and JMJD6-associated anti-pause enhancers in regulation of transcriptional pause release. *Cell.* **155**, 1581-1595 (2013).

209. Dobi, K.C. & Winston, F. Analysis of transcriptional activation at a distance in Saccharomyces cerevisiae. *Molecular and cellular biology : MCB.* **27**, 5575-5586 (2007).

210. Kvon, E.Z. *et al.* Genome-scale functional characterization of Drosophila developmental enhancers in vivo. *Nature.* **512**, 91-95 (2014).

211. Sanyal, A., Lajoie, B.R., Jain, G. & Dekker, J. The long-range interaction landscape of gene promoters. *Nature.* **489**, 109-113 (2012).

212. Li, G. *et al.* Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell.* **148**, 84-98 (2012).

213. van Arensbergen, J., van Steensel, B. & Bussemaker, H.J. In search of the determinants of enhancer-promoter interaction specificity. *Trends in cell biology.* **24**, 695-702 (2014).

214. Hnisz, D. *et al.* Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science.* **351**, 1454-1458 (2016).

215. Kragesteen, B.K. *et al.* Dynamic 3D chromatin architecture contributes to enhancer specificity and limb morphogenesis. *Nature genetics.* **50**, 1463-1473 (2018).

216. Juven-Gershon, T. & Kadonaga, J.T. Regulation of gene expression via the core promoter and the basal transcriptional machinery. *Developmental biology.* **339**, 225-229 (2010).

217. Arnold, C.D. *et al.* Genome-wide assessment of sequence-intrinsic enhancer responsiveness at single-base-pair resolution. *Nature biotechnology.* **35**, 136-144 (2017).

218. Kwon, D. *et al.* Enhancer-promoter communication at the Drosophila engrailed locus. *Development.* **136**, 3067-3075 (2009).

219. de Laat, W. & Grosveld, F. Spatial organization of gene expression: the active chromatin hub. *Chromosome research.* **11**, 447-459 (2003).

220. Allahyar, A. *et al.* Enhancer hubs and loop collisions identified from single-allele topologies. *Nature genetics.* **50**, 1151-1160 (2018).

221. Oudelaar, A.M. *et al.* Single-allele chromatin interactions identify regulatory hubs in dynamic compartmentalized domains. *Nature genetics.* **50**, 1744-1751 (2018).

222. An integrated encyclopedia of DNA elements in the human genome. *Nature.* **489**, 57-74 (2012).

223. Whyte, W.A. *et al.* Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell.* **153**, 307-319 (2013).

224. Hnisz, D. *et al.* Super-enhancers in the control of cell identity and disease. *Cell.* **155**, 934-947 (2013).

225. Jia, Q., Chen, S., Tan, Y., Li, Y. & Tang, F. Oncogenic super-enhancer formation in tumorigenesis and its molecular mechanisms. *Experimental and molecular medicine.* **52**, 713-723 (2020).

226. Hnisz, D. *et al.* Convergence of developmental and oncogenic signaling pathways at transcriptional super-enhancers. *Molecular cell.* **58**, 362-370 (2015).

227. Wu, M. & Shen, J. From Super-Enhancer Non-coding RNA to Immune Checkpoint: Frameworks to Functions. *Frontiers in oncology.* **9**, 1307 (2019).

228. Hanahan, D. & Weinberg, Robert A. Hallmarks of Cancer: The Next Generation. *Cell.* **144**, 646-674 (2011).

229. Li, Q.-L. *et al.* Genome-wide profiling in colorectal cancer identifies PHF19 and TBC1D16 as oncogenic super enhancers. *Nature communications.* **12**, 6407 (2021).

230. McKeown, M.R. *et al.* Superenhancer Analysis Defines Novel Epigenomic Subtypes of Non-APL AML, Including an RARα Dependency Targetable by SY-1425, a Potent and Selective RARα Agonist. *Cancer discovery* **7**, 1136-1153 (2017).

231. Wu, T., Huang, H. & Wang, X. Dissecting super-enhancer heterogeneity: time to re-examine cancer subtypes? *Trends in genetics.* (2022).

232. Huang, H. *et al.* Defining super-enhancer landscape in triple-negative breast cancer by multiomic profiling. *Nature communications.* **12**, 2242 (2021).

233. van Groningen, T. *et al.* Neuroblastoma is composed of two super-enhancer-associated differentiation states. *Nature genetics.* **49**, 1261-1266 (2017).

234. Kelly, M.R. *et al.* A multi-omic dissection of super-enhancer driven oncogenic gene expression programs in ovarian cancer. *Nature communications.* **13**, 4247 (2022).

235. Spielmann, M., Lupiáñez, D.G. & Mundlos, S. Structural variation in the 3D genome. *Nature reviews.* **19**, 453-467 (2018).

236. Furlong, E.E.M. & Levine, M. Developmental enhancers and chromosome topology. *Science.* **361**, 1341-1345 (2018).

237. Vos, E.S.M. *et al.* Interplay between CTCF boundaries and a super enhancer controls cohesin extrusion trajectories and gene expression. *Molecular cell.* **81**, 3082-3095.e6 (2021).

238. Natoli, G. & Andrau, J.-C. Noncoding transcription at enhancers: general principles and functional models. *Annual review of genetics.* **46**, 1-19 (2012).

239. Chen, H. & Liang, H. A High-Resolution Map of Human Enhancer RNA Loci Characterizes Super-enhancer Activities in Cancer. *Cancer cell.* **38**, 701-715.e5 (2020).

240. Plank, J.L. & Dean, A. Enhancer function: mechanistic and genome-wide insights come together. *Molecular cell.* **55**, 5-14 (2014).

241. Pefanis, E. *et al.* RNA exosome-regulated long non-coding RNA transcription controls super-enhancer activity. *Cell.* **161**, 774-789 (2015).

242. Zhao, Y. *et al.* MyoD induced enhancer RNA interacts with hnRNPL to activate target gene transcription during myogenic differentiation. *Nature communications.* **10**, 5787 (2019).

243. So, K.K.H. *et al.* seRNA PAM controls skeletal muscle satellite cell proliferation and aging through trans regulation of Timp2 expression synergistically with Ddx5. *Aging cell.*, e13673 (2022).

244. Rahnamoun, H. *et al.* RNAs interact with BRD4 to promote enhanced chromatin engagement and transcription activation. *Nature structural & molecular biology.* **25**, 687-697 (2018).

245. Huarte, M. *et al.* A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell.* **142**, 409-419 (2010).

246. Dimitrova, N. *et al.* LincRNA-p21 activates p21 in cis to promote Polycomb target gene expression and to enforce the G1/S checkpoint. *Molecular cell.* **54**, 777-790 (2014).

247. Göndör, A. Dynamic chromatin loops bridge health and disease in the nuclear landscape. *Seminars in cancer biology.* **23**, 90-98 (2013).

248. Spilianakis, C.G., Lalioti, M.D., Town, T., Lee, G.R. & Flavell, R.A. Interchromosomal associations between alternatively expressed loci. *Nature.* **435**, 637-645 (2005).

249. Ke, Y. *et al.* 3D Chromatin Structures of Mature Gametes and Structural Reprogramming during Mammalian Embryogenesis. *Cell.* **170**, 367-381.e20 (2017).

250. Nikopoulou, C. *et al.* The Transcription Factor ThPOK Orchestrates Stochastic Interchromosomal Interactions Required for IFNB1 Virus-Inducible Gene Expression. *Molecular cell.* **71**, 352-361.e5 (2018).

251. King, T.D., Johnson, J.E. & Bateman, J.R. Position Effects Influence Transvection in Drosophila melanogaster. *Genetics.* **213**, 1289-1299 (2019).

252. Bateman, J.R. & Johnson, J.E. Altering enhancer-promoter linear distance impacts promoter competition in cis and in trans. *Genetics.* (2022).

253. Lim, B., Heist, T., Levine, M. & Fukaya, T. Visualization of Transvection in Living Drosophila Embryos. *Molecular cell.* **70**, 287-296.e6 (2018).

254. AlHaj Abed, J. *et al.* Highly structured homolog pairing reflects functional organization of the Drosophila genome. *Nature communications.* **10**, 4485 (2019).

255. Sandhu, K.S. *et al.* Nonallelic transvection of multiple imprinted loci is organized by the H19 imprinting control region during germline development. *Genes & development.* **23**, 2598-2603 (2009).

256. Sandoval, J. Cancer epigenomics: beyond genomics. *Current opinion in genetics & development.* **22**, 50-55 (2012).

257. Lourenco, C. *et al.* MYC protein interactors in gene transcription and cancer. *Nature reviews.* **21**, 579-591 (2021).

258. Thandapani, P. Super-enhancers in cancer. *Pharmacology & therapeutics.* **199**, 129-138 (2019).

259. Lancho, O. & Herranz, D. The MYC Enhancer-ome: Long-Range Transcriptional Regulation of MYC in Cancer. *Trends in Cancer* **4**, 810-822 (2018).

260. Amjadi-Moheb, F., Paniri, A. & Akhavan-Niaki, H. Insights into the Links between MYC and 3D Chromatin Structure and Epigenetics Regulation: Implications for Cancer Therapy. *Cancer research.* **81**, 1925-1936 (2021).

261. Dang, C.V., Reddy, E.P., Shokat, K.M. & Soucek, L. Drugging the 'undruggable' cancer targets. *Nature reviews.* **17**, 502-508 (2017).

262. Huppi, K., Pitt, J.J., Wahlberg, B.M. & Caplen, N.J. The 8q24 gene desert: an oasis of non-coding transcriptional activity. *Frontiers in genetics.* **3**, 69 (2012).

263. Davis, T.L., Firulli, A.B. & Kinniburgh, A.J. Ribonucleoprotein and protein factors bind to an H-DNA-forming c-myc DNA element: possible regulators of the c-myc gene. *Proceedings of the National Academy of Sciences of the United States of America* **86**, 9682-9686 (1989).

264. Berberich, S.J. & Postel, E.H. PuF/NM23-H2/NDPK-B transactivates a human c-myc promoter-CAT gene via a functional nuclease hypersensitive element. *Oncogene.* **10**, 2343-2347 (1995).

265. DesJardins, E. & Hay, N. Repeated CT elements bound by zinc finger proteins control the absolute and relative activities of the two principal human c-myc promoters. *Molecular and cellular biology : MCB.* **13**, 5710-5724 (1993).

266. Sotelo, J. *et al.* Long-range enhancers on 8q24 regulate c-Myc. *Proceedings of the National Academy of Sciences of the United States of America.* **107**, 3001-3005 (2010).

267. Wasserman, N.F., Aneas, I. & Nobrega, M.A. An 8q24 gene desert variant associated with prostate cancer risk confers differential in vivo activity to a MYC enhancer. *Genome research.* **20**, 1191-1197 (2010).

268. Grisanzio, C. & Freedman, M.L. Chromosome 8q24-Associated Cancers and MYC. *Genes & cancer.* **1**, 555-559 (2010).

269. Bahr, C. *et al.* A Myc enhancer cluster regulates normal and leukaemic haematopoietic stem cell hierarchies. *Nature.* **553**, 515-520 (2018).

270. Schuijers, J. *et al.* Transcriptional Dysregulation of MYC Reveals Common Enhancer-Docking Mechanism. *Cell reports.* **23**, 349-360 (2018).

271. Nissan, A. *et al.* Colon cancer associated transcript-1: a novel RNA expressed in malignant and pre-malignant human tissues. *International journal of cancer.* **130**, 1598-1606 (2012).

272. Xiang, J.-F. *et al.* Human colorectal cancer-specific CCAT1-L lncRNA regulates long-range chromatin interactions at the MYC locus. *Cell research.* **24**, 513-531 (2014).

273. Kim, T. *et al.* Long-range interaction and correlation between MYC enhancer and oncogenic long noncoding RNA CARLo-5. *Proceedings of the National Academy of Sciences of the United States of America.* **111**, 4173-4178 (2014).

274. Peng, W.-X., He, R.-Z., Zhang, Z., Yang, L. & Mo, Y.-Y. LINC00346 promotes pancreatic cancer progression through the CTCF-mediated Myc transcription. *Oncogene.* **38**, 6770-6780 (2019).

275. McKeown, M.R. & Bradner, J.E. Therapeutic strategies to inhibit MYC. *Cold Spring Harbor perspectives in medicine.* **4**(2014).

276. Kalkat, M. *et al.* MYC Deregulation in Primary Human Cancers. *Genes.* **8**(2017).

277. Rennoll, S. & Yochum, G. Regulation of MYC gene expression by aberrant Wnt/β-catenin signaling in colorectal cancer. *World journal of biological chemistry.* **6**, 290-300 (2015).

278. Park, J.H., Pyun, W.Y. & Park, H.W. Cancer Metabolism: Phenotype, Signaling and Therapeutic Targets. *Cells.* **9**(2020).

279. Hoffman, B. & Liebermann, D.A. Apoptotic signaling by c-MYC. *Oncogene.* **27**, 6462-6472 (2008).

280. He, T.C. *et al.* Identification of c-MYC as a target of the APC pathway. *Science.* **281**, 1509-1512 (1998).

281. Sharma, M. & Pruitt, K. Wnt Pathway: An Integral Hub for Developmental and Oncogenic Signaling Networks. *International journal of molecular sciences.* **21**(2020).

282. Clevers, H. & Nusse, R. Wnt/β-catenin signaling and disease. *Cell.* **149**, 1192-1205 (2012).

283. Hayat, R., Manzoor, M. & Hussain, A. Wnt signaling pathway: A comprehensive review. *Cell biology international.* **46**, 863-877 (2022).

284. Grumolato, L. *et al.* Canonical and noncanonical Wnts use a common mechanism to activate completely unrelated coreceptors. *Genes & development.* **24**, 2517-2530 (2010).

285. Logan, C.Y. & Nusse, R. The Wnt signaling pathway in development and disease. *Annual review of cell and developmental biology.* **20**, 781-810 (2004).

286. Klaus, A. & Birchmeier, W. Wnt signalling and its impact on development and cancer. *Nature reviews.* **8**, 387-398 (2008).

287. Huang, H. & He, X. Wnt/beta-catenin signaling: new (and old) players and new insights. *Current opinion in cell biology.* **20**, 119-125 (2008).

288. Cadigan, K.M. TCFs and Wnt/β-catenin signaling: more than one way to throw the switch. *Current topics in developmental biology.* **98**, 1-34 (2012).

289. Mosimann, C., Hausmann, G. & Basler, K. Beta-catenin hits chromatin: regulation of Wnt target gene activation. *Nature reviews.* **10**, 276-286 (2009).

290. Takayama, T. *et al.* Aberrant expression and phosphorylation of beta-catenin in human colorectal cancer. *British journal of cancer.* **77**, 605-613 (1998).

291. Wu, H. *et al.* TRAF6 inhibits colorectal cancer metastasis through regulating selective autophagic CTNNB1/β-catenin degradation and is targeted for GSK3B/GSK3β-mediated phosphorylation and degradation. *Autophagy.* **15**, 1506-1522 (2019).

292. He, S. & Tang, S. WNT/β-catenin signaling in the development of liver cancers. *Biomedicine & pharmacotherapy.* **132**, 110851 (2020).

293. Cheng, X., Xu, X., Chen, D., Zhao, F. & Wang, W. Therapeutic potential of targeting the Wnt/β-catenin signaling pathway in colorectal cancer. *Biomedicine & pharmacotherapy.* **110**, 473-481 (2019).

294. ten Berge, D. *et al.* Embryonic stem cells require Wnt proteins to prevent differentiation to epiblast stem cells. *Nature cell biology.* **13**, 1070-1075 (2011).

295. Tang, Q. *et al.* TM4SF1 promotes EMT and cancer stemness via the Wnt/β-catenin/SOX2 pathway in colorectal cancer. *Journal of experimental & clinical cancer research.* **39**, 232 (2020).

296. Xiong, J. *et al.* STK31 regulates the proliferation and cell cycle of lung cancer cells via the Wnt/β-catenin pathway and feedback regulation by c-myc. *Oncology reports.* **43**, 395-404 (2020).

297. Sumida, N. *et al.* The ultra-sensitive Nodewalk technique identifies stochastic from virtual, population-based enhancer hubs regulating MYC in 3D: Implications for the fitness of cancer cells. *bioRxiv*, 286583 (2018).

298. Langer, S., Geigl, J.B., Ehnle, S., Gangnus, R. & Speicher, M.R. Live cell catapulting and recultivation does not change the karyotype of HCT116 tumor cells. *Cancer genetics and cytogenetics.* **161**, 174-177 (2005).

299. Sumida, N. *et al.* MYC as a driver of stochastic chromatin networks: implications for the fitness of cancer cells. *Nucleic acids research.* **48**, 10867-10876 (2020).

300. Hyle, J. *et al.* Acute depletion of CTCF directly affects MYC regulation through loss of enhancer-promoter looping. *Nucleic acids research.* **47**, 6699-6713 (2019).

301. Fort, L. *et al.* Fam49/CYRI interacts with Rac1 and locally suppresses protrusions. *Nature cell biology.* **20**, 1159-1171 (2018).

302. Benitah, S.A., Frye, M., Glogauer, M. & Watt, F.M. Stem cell depletion through epidermal deletion of Rac1. *Science.* **309**, 933-935 (2005).

303. Cai, Z. *et al.* RIC-seq for global in situ profiling of RNA-RNA spatial interactions. *Nature.* **582**, 432-437 (2020).

304. Chachoua, I. *et al.* Canonical WNT signaling-dependent gating of MYC requires a noncanonical CTCF function at a distal binding site. *Nature communications.* **13**, 204 (2022).

305. Dhanasekaran, R. *et al.* The MYC oncogene - the grand orchestrator of cancer growth and immune evasion. *Nature reviews.* **19**, 23-36 (2022).

306. Finn, E.H. *et al.* Extensive Heterogeneity and Intrinsic Variation in Spatial Genome Organization. *Cell.* **176**, 1502-1515.e10 (2019).

307. Dixon, J.R. *et al.* Chromatin architecture reorganization during stem cell differentiation. *Nature.* **518**, 331-336 (2015).

308. Naumova, N. *et al.* Organization of the mitotic chromosome. *Science.* **342**, 948-953 (2013).

309. Nagano, T. *et al.* Cell-cycle dynamics of chromosomal organization at single-cell resolution. *Nature.* **547**, 61-67 (2017).

310. Hansen, A.S., Pustova, I., Cattoglio, C., Tjian, R. & Darzacq, X. CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *ELife.* **6**(2017).

311. Zhao, Z. *et al.* Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nature genetics.* **38**, 1341-1347 (2006).

312. Cowling, V.H. & Cole, M.D. Turning the tables: Myc activates Wnt in breast cancer. *Cell cycle.* **6**, 2625-2627 (2007).

313. Berto, A. *et al.* Disentangling the molecular determinants for Cenp-F localization to nuclear pores and kinetochores. *EMBO reports.* **19**(2018).

314. Zhuo, Y.-J. *et al.* Enhanced expression of centromere protein F predicts clinical progression and prognosis in patients with prostate cancer. *International journal of molecular medicine.* **35**, 966-972 (2015).

315. de la Guardia, C., Casiano, C.A., Trinidad-Pinedo, J. & Báez, A. CENP-F gene amplification and overexpression in head and neck squamous cell carcinomas. *Head & neck.* **23**, 104-112 (2001).

316. O'Brien, S.L. *et al.* CENP-F expression is associated with poor prognosis and chromosomal instability in patients with primary breast cancer. *International journal of cancer.* **120**, 1434-1443 (2007).

317. Demma, M.J. *et al.* Omomyc Reveals New Mechanisms To Inhibit the MYC Oncogene. *Molecular and cellular biology : MCB.* **39**(2019).

318. Blaževitš, O. *et al.* MYC-Associated Factor MAX is a Regulator of the Circadian Clock. *International journal of molecular sciences.* **21**(2020).

319. Sulli, G., Lam, M.T.Y. & Panda, S. Interplay between Circadian Clock and Cancer: New Frontiers for Cancer Treatment. *Trends in Cancer* **5**, 475-494 (2019).