TFIIIC subunits employ different modes of action for regulating N-MYC

TFIIIC Untereinheiten verwenden unterschiedliche Wirkungsweisen zur Regulierung von N-MYC

Doctoral Thesis

for a doctoral degree
at the Graduate School of Life Sciences,
Julius-Maximilians-Universität Würzburg
Section Biomedicine

submitted by

Ka Yan MAK

from Hong Kong

Würzburg 2019

Members of the Thesis Committee

Chairperson: Prof. Alexander Buchberger

Primary supervisor: Prof. Martin Eilers
Second supervisor: Prof. Manfred Gessler
Third supervisor: Dr. Sonja Lorenz

Submitted on:

Date of Public Defence:

Date of Receipt of Certificates:

Table of contents

Summary

Amplification of N-MYC is a poor prognostic and survival marker of neuroblastoma. To broaden the scope of knowledge in N-MYC cancer biology, interactors of N-MYC should be investigated. TFIIIC complex was identified as a new protein interacting partner of N-MYC. TFIIIC is a core component of RNAPIII transcription machinery which is important for the synthesis of tRNA genes. TFIIIC recognizes and binds to B-box located internal of tRNA genes which subsequently initiate the RNAPIII transcription process. Apart from the role in RNAPIII transcription machinery, TFIIIC is an architectural protein. TFIIIC binds to thousands of sites across the genome without RNAPIII and TFIIIB. These binding loci are known as Extra TFIIIC (ETC) sites at which TFIIIC performs its role in genome organization. However, knowledge of TFIIIC is mostly restricted to studies conducted in yeasts, the exact function of TFIIIC and how it regulates N-MYC remains to be elucidated.

To obtain a better overview about TFIIIC functions, two TFIIIC subunits (TFIIIC5 and TFIIIC2) which represent sub-complexes A and B were chosen for investigation. ChIP-seq experiment of RNAPIII transcription machinery was performed. It showed that both TFIIIC subunits functioned together as a complex. Next, joint binding sites of two TFIIIC subunits and N-MYC were identified. The data revealed that co-occupancies between N-MYC and TFIIIC subunits had different preference on genomic distribution. Furthermore, TFIIIC5 exhibited strong binding association with architectural proteins RAD21 and CTCF whereas TFIIIC2 was only modestly enriched with these two proteins. Both TFIIIC subunits showed equal but weak enrichment with accessory protein CAPH2. Despite the weak association with other architectural proteins, TFIIIC2 binds preferentially to repetitive elements SINE.

In order to understand how TFIIIC5 affects other architectural proteins in chromatin binding, cells were depleted of TFIIIC protein upon doxycycline induction of shRNA. N-MYC binding was not affected. Yet, 50% reduction of RAD21 binding to joint N-MYC/TFIIIC sites was noticed. CAPH2 binding was increased at some joint sites while some did not respond. Lastly, CTCF did not show changes in binding under the effect of TFIIIC5 knockdown.

In summary, the data indicated TFIIIC subunits from different sub-complexes diverge in functions other than tRNA synthesis. The association of TFIIIC5 with architectural proteins and that of TFIIIC2 with SINE elements were suggested to be distinct mechanisms to regulate N-Myc directly or indirectly.

Zusammenfassung

Die Amplifikation von N-MYC ist ein schlechter Prognose- und Überlebensmarker für Neuroblastome. Um den Kenntnisstand über die Krebsbiologie von N-MYC zu erweitern, Interaktoren von N-MYC sollten untersucht werden. Der TFIIIC-Komplex wurde als neuer interaktiver Partner von N-MYC identifiziert. TFIIIC ist eine Kernkomponente der RNAPIII-Transkriptionsmaschinerie, die für die Synthese von tRNA-Genen wichtig ist. TFIIIC erkennt und bindet an B-Box innerhalb von tRNA-Genen, die anschließend den RNAPIII-Transkriptionsprozess initiieren. Abgesehen von der Rolle in der RNAPIII-Transkriptionsmaschinerie ist TFIIIC ein Architekturprotein. TFIIIC bindet an Tausende von Stellen im gesamten Genom ohne RNAPIII und TFIIIB. Diese Bindungsorte sind als Extra TFIIIC (ETC) -Stellen bekannt, an denen TFIIIC seine Rolle bei der Genomorganisation spielen kann. Das Wissen über TFIIIC beschränkt sich jedoch meist auf Studien, die mit Hefen durchgeführt werden. Die genaue Funktion von TFIIIC und die Art seiner Regulierung von N-MYC sind noch zu klären.

Um einen besseren Überblick über die TFIIIC-Funktionen zu erhalten, wurden zwei TFIIIC-Untereinheiten (TFIIIC5 und TFIIIC2) ausgewählt, die die Unterkomplexe A und B repräsentieren. Es wurde ein ChIP-seq-Experiment der RNAPIII-Transkriptionsmaschinerie durchgeführt. Es zeigte sich, dass beide TFIIIC-Untereinheiten zusammen als Komplex fungierten. Als nächstes wurden gemeinsame Bindungsstellen von zwei TFIIIC-Untereinheiten und N-MYC identifiziert. Die Daten zeigten, dass Co-Besetzungen zwischen N-MYC- und TFIIIC-Untereinheiten unterschiedliche Präferenzen bei der Verteilung von Genom hatten. Darüber hinaus zeigte TFIIIC5 eine starke Bindungsassoziation mit den Architekturproteinen RAD21 und CTCF, während TFIIIC2 mit diesen beiden Proteinen nur wenig angereichert war. Beide TFIIIC-Untereinheiten zeigten eine gleiche, aber schwache Anreicherung mit dem Zusatzprotein CAPH2. Trotz der schwachen Assoziation mit anderen Architekturproteinen bindet TFIIIC2 bevorzugt an repetitive Elemente SINE.

Um zu verstehen, wie TFIIIC5 andere Architekturproteine bei der Chromatinbindung beeinflusst, wurden die Zellen bei der Doxycyclin-Induktion von shRNA an TFIIIC-Protein aufgebraucht. Die N-MYC-Bindung war nicht betroffen. Es wurde jedoch eine Verringerung der Bindung von RAD21 an gemeinsame N-MYC / TFIIIC-Stellen um 50% festgestellt. Die CAPH2-Bindung war an einigen gemeinsamen Stellen erhöht, während einige nicht reagierten. Schließlich zeigte CTCF keine Bindungsänderungen unter dem Einfluss von TFIIIC5-Knockdown.

Zusammenfassend zeigen die Daten, dass TFIIIC-Untereinheiten aus verschiedenen Unterkomplexen in anderen Funktionen als der tRNA-Synthese voneinander abweichen. Es wurde vermutet, dass die Assoziation von TFIIIC5 mit Architekturproteinen und TFIIIC2 mit SINE-Elementen unterschiedliche Mechanismen sind, um N-Myc direkt oder indirekt zu regulieren.

1 Introduction

1.1 Transcription factor IIIC (TFIIIC) complex

TFIIIC complex is the basal unit of RNA polymerase III (RNAPIII) transcription machinery. It is composed of six subunits which can be partitioned into sub-complexes A and B. The subunit composition is conserved from yeast to human (Table 1.1). Nomenclature of human TFIIIC subunits is based on their protein molecular size in kDa or numbered in reverse order of their relative protein size. Although other components of RNAPIII transcription machinery tend to share high homology between different species, all human TFIIIC subunits have markedly different protein sequences, with less than 30% sequence identity when compared to that of yeast strains (Fig. 1.1) (Huang and Maraia, 2001).

Sub-complex	*Homo sapiens*	*Saccharomyces cerevisiae*	*Schizosaccharomyces prombe*
	Subunit		
A	TFIIIC63 (GTF3C5)	Tfc1 (τ95)	Sfc1
	TFIIIC102 (GTF3C3)	Tfc4 (τ131)	Sfc4
	TFIIIC35 (GTF3C6)	Tfc7 (τ55)	Sfc7
B	TFIIIC220 (GTF3C1)	Tfc3 (τ138)	Sfc3
	TFIIIC110 (GTF3C2)	Tfc6 (τ91)	Sfc6
	TFIIIC90 (GTF3C4)	Tfc8 (τ60)	Sfc9

Table 1.1: Subunit composition of TFIIIC in *Homo sapiens*, *Saccharomyces cerevisiae* and *Schizosaccharomyces prombe*
TFIIIC complex from different species follow the same subunit composition.

RNAPIII accounts for transcription of small RNA molecules like small 5S rRNA, U6 splicesosomal RNA, transfer RNAs (tRNAs). They are classified into three major types based on the promoter in use (Fig. 1.2). Three transcription factors TFIIIA, TFIIIB, and TFIIIC are involved in different combinations depending on the type of promoter for transcription. TFIIIC is required for both type 1 and 2 transcriptions and essentially accounts for tRNA synthesis which is the major RNAPIII gene product.

RNAPIII basal transcription machinery for tRNA genes constitutes RNAPIII and two initiation transcription factor complexes TFIIIB and TFIIIC. Binding of TFIIIC to its conserved DNA sequences initiates the transcription process. Sub-complex B of TFIIIC first recognizes internal promoter B-box sequence and binds strongly to it, which then enables sub-complex A to bind to the less conserved A-box with a low affinity. Afterwards, TFIIIC recruits TFIIIB upstream of transcription start site (TSS) where TFIIIB then recruits RNAPIII by direct protein-protein interaction to drive tRNA synthesis (Taylor et al., 2013).

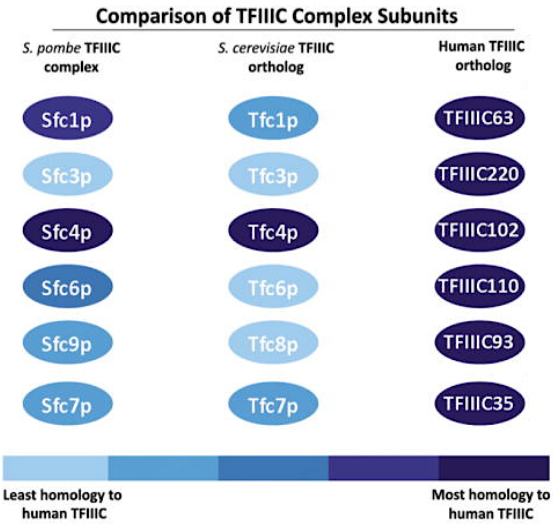


Figure 1.1: Schematic representation of relative sequence homology comparison between human TFIIIC subunits and *S. pombe* and *S. cerevisiae*
Different TFIIIC subunits possess varying degree of sequence conservation from yeasts to human. TFIIIC102 (GTF3C3) and TFIIIC220 (GTF3C1) exhibit the highest and lowest sequence homology respectively. Adopted from (Lunyak and Atallah, 2011).

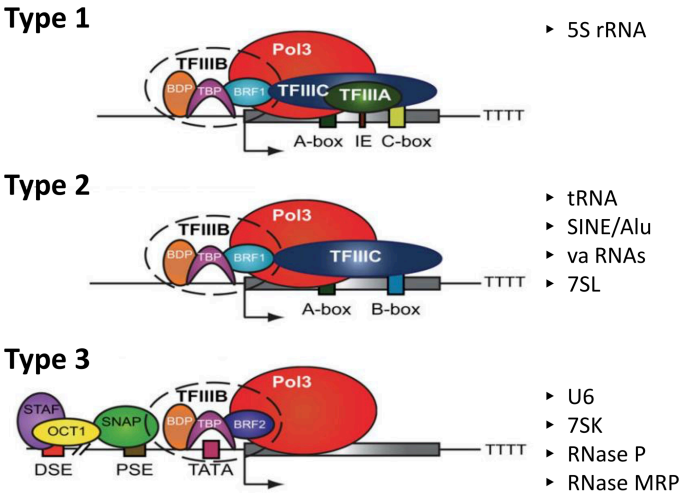


Figure 1.2: Three types of promoters for RNAPIII transcription
Basal transcription components of RNAPIII using different promoters and the corresponding gene products are shown. TFIIIC is required for synthesis of 5S rRNA, tRNA, SINE/*Alu*, etc. In type 2 promoter, both A-box and B-box are internally located. Selenocysteine tRNA gene has a type 3 promoter and is noted as an exception. Modified from (Oler et al., 2010).

TFIIIC is a prerequisite to initiate RNAPIII transcription but is not necessarily for RNAPIII re-initiation (Dieci et al., 2000; Ferrari et al., 2004). Following the first transcription cycle, RNAPIII is recycled after passing TES and recruited again to the same transcription unit. To transcribe short genes (approx. 100 bp), recycled RNAPIII only needs to interact with TFIIIB. However, for genes longer than 300 bp, RNAPIII located at transcription end site (TES) is too

remoted from TSS to interact with TFIIIB. In this case, polymerase recycling no longer permits re-initiation of RNAPIII. Therefore, TFIIIC is essential to RNAPIII re-initiation in a gene-length dependent manner.

Subunits of TFIIIC complex

TFIIIC has mostly been described as a complex. Less is known about the functions of individual human TFIIIC subunit. The knowledge is mainly limited to the subunit purification and characterization one to two decades ago. Lack of protein structure of human TFIIIC subunits and low sequence conservation between different species also impede further understanding of each subunit (Fig. 1.1). Comparative analysis between homologs in yeasts and human TFIIIC102 and TFIIIC63 showed they are both conserved in structure and can interact with TFIIIB and RNAPIII *in vitro* (Hsieh et al., 1999). However, in spite of null sequence homology, GTF3C1 have analogous functions to its orthologs (Huang et al., 2000).

The most prominent functional feature of individual TFIIIC subunit is the intrinsic histone acetyltransferase (HAT) activities of GTF3C2 and GTF3C4 (Hsieh et al., 1999; Kundu et al., 1999). In particular, GTF3C4 has a substrate specificity for histone H3. GTF3C1, which also belongs to sub-complex B was suggested to have HAT activities in in-gel assays (Kundu et al., 1999).

Extra TFIIIC (ETC) sites

Surprisingly, TFIIIC binding sites across the genome significantly outweigh the number of tRNA genes. These “additional” TFIIIC binding sites are exclusive of other RNAPIII transcription components and are defined as ETC sites, which were first reported in *S. cerevisiae* (Moqtaderi and Struhl, 2004).

TFIIIC recognizes B-box sequence for binding. B-box sequence at ETC loci is highly similar to that of at tRNA genes (GGTTCGANYCY, underlined C is essential for TFIIIC binding and tRNA synthesis) (Baker, Gabrielsen, and Hall 1986; Marzouki et al. 1986; Moqtaderi et al. 2010; Newman, Ogden, and Abelson 1983). In addition, *de novo* motif search identified an ETC-specific motif which is almost mutually exclusive to B-box (Fig. 1.3).

The occurrence of ETC sites is proportional to genome size, ranging from less than ten in *S. cerevisiae*, to about 2000 in human K562 leukemia cells (Moqtaderi et al., 2010). TFIIIC binding to ETC sites is assumed to be determined by cell-type specific factors due to variations of TFIIIC binding in different cell lines (Moqtaderi et al. 2010).

ETC sites in yeasts were characterized to be barriers. It is a marker of repressive chromatin domain and B-box deletion resulted in invasion of heterochromatin spreading into euchromatic domain (Noma et al., 2006). Multimerized ETC sites in *S. cerevisiae* demonstrated insulator activity of TFIIIC, which can act as both enhancer-blocker and heterochromatin barrier (Valenzuela et al., 2009). Moreover, sfc3- and sfc6-mapped ETC sites clustered at nuclear periphery and were suggested to be chromosome-organizing clamps (Noma et al., 2006) (see section 1.2.2). However, this mechanism remains open in human since immunofluorescence of human TFIIIC paralogs did not show similar peripheric enrichment.

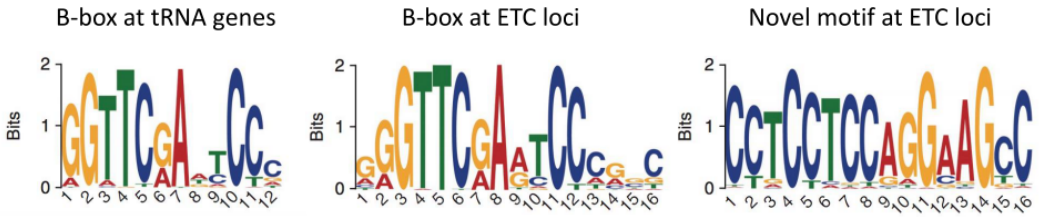


Figure 1.3: ETC sites are represented by B-box and a novel motif sequence
Motif search based on ChIP-seq of subunit GTF3C2 in K562 human leukemia cell line. tRNA genes contain both A- and B-boxes whilst ETC sites are characterized by the standard B-box and a novel motif. Adopted from (Moqtaderi et al., 2010).

The exact functions of ETC sites in human are not completely understood but reports have shown strong associations of ETC sites with the following factors/aspects:

i. Association with RNAPII genes/at TSS

Genome-wide distribution of ETC sites showed positional bias towards TSS of RNAPII genes within the window of 1 kb from TSS. ETC sites at *S. cerevisiae* are a few hundred bases upstream of the RNAPII promoters (Noma et al., 2006). Similar observations were reported in human ETC sites (Moqtaderi et al., 2010; Oler et al., 2010). In particular, human neuroblastoma cells showed an overall displacement of 85 bp of TFIIIC 5' to N-MYC binding sites (Büchel et al., 2017). Gene ontology analysis revealed an enrichment of nuclear-localized proteins at RNAPII genes located adjacent to ETC sites (Moqtaderi et al., 2010).

ii. Association with CTCF

CTCF (CCCTC-binding factor) is a major insulator of the mammalian genome, which has functions similar to ETC sites in yeasts as aforementioned (Felsenfeld et al., 2004). Correlation between CTCF and ETC sites was established in human cancer cells and mouse embryonic stem cells (Carrière et al., 2012; Moqtaderi et al., 2010; Oler et al., 2010). Their close proximity highly suggests TFIIIC at ETC sites may adopt similar functions as CTCF. More details in section 1.2.

iii. Association with histone modifications

The binding patterns of histone marks associated with functional RNAPII promoters (e.g. H3K27ac, H3K4me2, etc) in ETC sites are distinct from that of expressed RNAPIII genes (Moqtaderi et al., 2010). Histone modifications of ETC sites peak at the TFIIIC binding loci. However, nucleosome is lost for active transcription of RNAPIII genes, resulting in lower binding of the histone modifications of RNAPIII expressed genes at the TFIIIC peaks (Moqtaderi et al., 2010; Oler et al., 2010).

iv. Association with SINEs/*Alu*

Same as tRNAs, SINEs/*Alu* elements are transcribed by RNAPIII with type II promoter (Figure 1.2). *Alu* repetitive elements are a subgroup of SINEs (short interspersed elements) and occupy up to 10% of the human genome (Lander et al., 2001). SINEs and *Alu* are both enriched in ETC sites (Oler et al., 2010; Sun et al., 2016). Hela cells have at least 181 SINEs with ETC sites (Oler et al., 2010). Recent study in neurons reveals SINEs located downstream to neuronal activity-dependent genes are ETC sites (Crepaldi et al., 2013a). Binding of TFIIIC at those SINEs increases upon depolarization which serves as mechanism to control gene expression for neurons.

TFIIIC and cancer co-relation

Although TFIIIC has not yet been shown to be oncogenic or transforming to cells, available data have demonstrated its co-relation with cancers. Over-expression of TFIIIC subunits 1-5 was reported in ovarian tumor biopsies. Elevated TFIIIC transcripts leads to increase in DNA-binding activity of TFIIIC complex and facilitate RNAPIII transcription (Winter et al., 2000). Higher rate of tRNA synthesis favors protein production which then support cancer growth. In addition, TFIIIC subunits are over-expressed in different primary patient samples (Table 1.2). Some even confer unfavorable prognosis.

Subunits	Over-expression by immunostaining	Unfavorable prognostic marker
GTF3C1	Colorectal cancer, breast cancer, prostate cancer, lung cancer, liver cancer	Liver cancer
GTF3C2	N/A	Liver cancer, endometrial cancer
GTF3C3	N/A	N/A
GTF3C4	N/A	N/A
GTF3C5	Colorectal cancer, breast cancer, prostate cancer, lung cancer, liver cancer	N/A
GTF3C6	Colorectal cancer, breast cancer, prostate cancer, lung cancer, skin cancer	Liver cancer

Table 1.2: Pathology status of human TFIIIC subunits in different cancer types
Information is obtained in Human Protein Atlas (www.proteinatlas.org) (Pontén et al., 2008).

1.2 Role of TFIIIC in transcriptional regulation

1.2.1 Genome organization by architectural proteins

Vital nuclear processes such as transcription and replication often require regulatory elements for tight and precise regulation. Promoters, enhancers and/or insulators can be brought together by non-random organization of the genome. The myth of packing two-meter long DNA into the tiny space of nucleus can be explained by genome organization in a hierarchical manner: chromosome territories A/B compartments, topologically associated domains (TADs) and DNA-loops (Fig. 1.4) (Dekker et al., 2017; Ea et al., 2015; Ruiz-Velasco and Zaugg, 2017). Chromosome territories (CT) refer to the compartmentalization of nucleus during interphase (Cremer and Cremer, 2010). This form of three-dimensional organization of chromatin fibers separates active genomic regions from the inactive one, which are also known as A and B compartments. Another scale of nuclear architecture is called TADs, mega-base-sized domains within which a high degree of chromatin activities take place (Dixon et al., 2012). Long-range interactions under non-stochastic genome architecture are based on intra-chromosomal DNA loops. They bring distant regulatory elements together for gene regulation irrespective of activation or repression (Cavalli and Misteli, 2013).

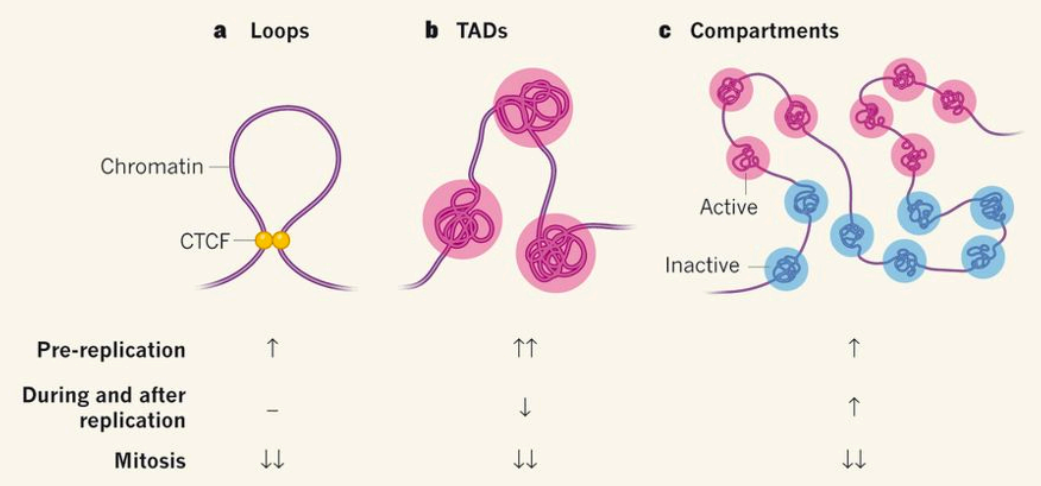


Figure 1.4: Schematic diagram of different domains of genome organization and their changes during cell cycle
Genome architecture is organized into three scales of domains. In ascending order of their sizes are chromatin loops, TADs and compartments. CTCF loop is the most well-studied example and served as an example of chromatin loop. These domains are dynamic throughout the cell cycles. In particular, TADs and compartments are organized independently of each other. Adopted from (Beagrie and Pombo, 2017).

Topological organization of genome within the nucleus is controlled and governed by architectural proteins (Ong and Corces, 2014). Other than insulator functions, they can also facilitate enhancer-promoter interactions. Examples of architectural proteins have been reported from yeasts to human. Among them, the most extensive studies of architectural

proteins were conducted in *Drosophila* (Table 1.3). Architectural proteins can be divided into two types – DNA binding protein and accessory protein. The latter does not recognize any specific DNA motif and is recruited to chromatin by DNA binding protein. Although studies with depletion of accessory protein has shown dynamic changes in genome architecture, accessory protein alone is not sufficient to regulate nuclear architecture but require the DNA-binding protein to function together.

Architectural proteins	
DNA-binding proteins	Accessory proteins
CTCF	Rad21 (cohesin)
Su(Hw)	Cap-H2 (condensin II)
BEAF-32	Cap-H/Barren (condensin I)
DREF	Mod/mdg4
TFIIIC	CP190
	Chromator
	L(3)mbt
	Fs(1)h-L
	Zw5
	GAF

Table 1.3: List of architectural proteins identified in *Drosophila*
Architectural proteins are divided into DNA-binding and accessory proteins. TFIIIC is one of the DNA-binding architectural proteins. It is noted that yeasts only have one condensin in comparison to two condensins in higher organisms.

Accessory proteins from SMC family

Among the list of accessory proteins, Rad21 (cohesin), Cap-H2 (condensin II) and Cap-H (condensin I) are members of structural maintenance of chromosomes (SMC) family (Losada and Hirano, 2005). In particular, they are the kleisin subunits of their corresponding SMC protein complex (in bracket) (Fig. 1.5). Structurally, SMC complexes share two symmetrical intramolecular coiled coil arms which form a hinge at one end and an ATP-binding head domain at the other end. Kleisin proteins accounts for the closing of the ring-like structure at the head domain. This specific structure of SMC family enables bundling of chromosomal strands which is conserved from bacteria to eukaryotes.

The unique structural feature of SMC proteins allows DNA strand to pass through its ring-like structure and thus DNA loops can be formed. Based on studies with CTCF and Rad21, loop extrusion is suggested to be the working model (Sanborn et al., 2015). Loop extruding factors such as Rad21 can translocate along the chromosome and they hold different parts of the genomic loci together. Chromatin loops halts when loop extruding factors encounter boundary elements such as CTCF. When only one side of the chromosome is halted, loop extruding factors still can continue to extrude the chromosome from the other side. The process of loop extrusion stops when loop extruding factors come into contact with boundary elements of both

sides of chromatin fiber. The extruded loop is eventually released when the loop extruding factors are dissociate from the chromosome.

Loop extrusion model describes only intra-chromosomal loops and it remains unknown how many CTCF and cohesin proteins participate. A recent report visualizes DNA loop extrusion by condensin in *S. cerevisiae* using real-time imaging (Ganji et al., 2018). However, this model is based solely on one condensin complex for both DNA anchor and sliding. How condensin recognizes the site to start or stop loop extrusion has not been addressed.

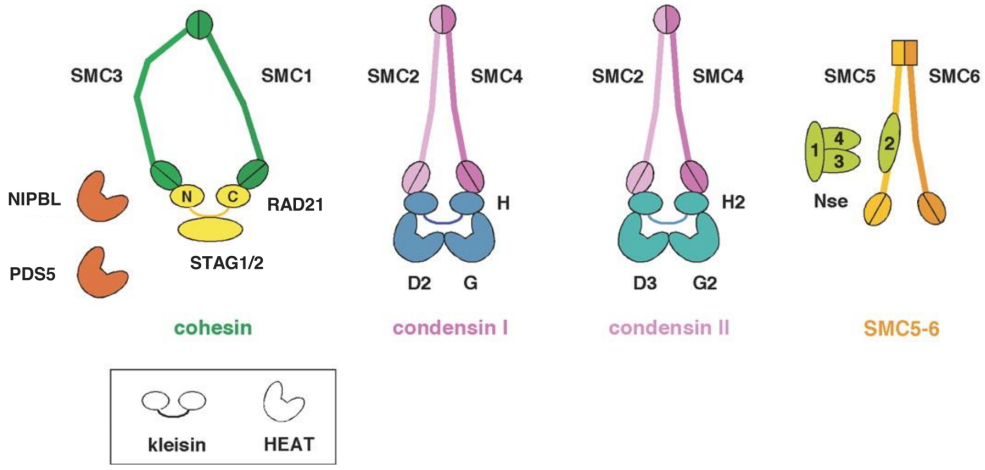


Fig. 1.5 Members of SMC family and their structure
Cartoons depicting the two-armed structure of SMC proteins. RAD21 (cohesin), CAPH (condensin I), and CAPH2 (condensin II) are the kleisin subunits. Modified from (Losada and Hirano, 2005).

Genome wide binding of architectural proteins

Architectural proteins are found throughout the genome. For instance, CTCF has about 14000 binding sites in human genome (Kim et al., 2007). Architectural proteins can co-occupy genomic loci in multiple combinations. These joint-binding sites are known as "architectural protein binding sites" (APBS) (Van Bortle et al., 2014). APBS can be classified as high-, medium- or low-occupancy depending on the number of DNA binding proteins and accessory proteins present. When the APBS has at least one protein from each type, it can be TAD border. Higher APBS occupancy confers stronger border strength, which may reflect the intra-TAD interaction frequencies and/or interaction stability. Architectural proteins at TAD border restrain interactions between genes and regulatory elements at adjacent TADs. On the other hand, APBS within a TAD can still mediate interactions. Taking CTCF as an example, 15% of its binding sites localizes at TADs and the remaining 85% is present within a TAD (Dixon et al., 2012). Depletion of CTCF resulted in decrease in intra-TAD interactions and an increase in

interactions in the adjacent TAD (Zuin et al., 2014). However, loss of CTCF is not sufficient to affect compartmentalization of the genome (Nora et al., 2017).

Genome architecture dynamics

Although it has been reported that up to 70% of TADs was conserved in embryonic stem cells and differentiated cells, and in mouse and human, chromosome conformation is highly dynamic and influenced by various factors (Dixon et al., 2012). One example is cell cycle (Fig. 1.4). It is widely known that DNA compacts when going through cell cycle until mitosis. Afterwards, DNA unfolds to allow replication which then condenses immediately again for the next mitosis. In accordance to the temporal change in chromatin compaction, genome architecture demonstrates dynamics and distinct patterns of organization at different cell cycle phases (Kakui et al., 2017; Lazar-Stefanita et al., 2017; Nagano et al., 2017). Compartmentalization increases as cell cycle progresses until it reaches the highest level at the end of S-phase. On the contrary, TADs are observed in S-phase but the borders are weakened when replication begins. But TADs expand during transition from mitosis to G1 and diminish afterwards. DNA loops remain relatively stable throughout interphase. However, all three hierarchical domains of genome organization cannot be observed at mitosis during which the chromatin is most compact (Dileep et al., 2015; Naumova et al., 2013).

A number of studies conducted in yeasts sought to investigate the role of SMC complexes in the dynamics of genome architecture at different cell cycle phases (Kakui et al., 2017; Lazar-Stefanita et al., 2017; Schalbetter et al., 2017). However, cohesin and condensin employ different mechanisms. Cohesin is responsible for long-range interaction related to DNA replication whereas condensin is more important for the structure of ribosomal DNA locus. Nonetheless, it requires further investigation to confirm if this knowledge can be applied to higher eukaryotes.

Genome architecture is also modified when cells are under different environmental factors. Temperature stress can trigger redistribution of architectural proteins from TAD borders to inside TADs (Li et al., 2015). Low- and medium occupancy APBS at TAD borders increases, thus lowering TAD border strength. Long-range interactions between enhancers and promoters within a TAD are induced upon heat shock. Another reported factor for change in TAD structure is B-cell activation (Kieffer-Kwon et al., 2017). Intra-chromosomal loops are also doubled upon activation.

While these examples have shed light on some factors that can modulate plasticity of nuclear architecture, the whole picture is still far from complete.

1.2.2 TFIIIC is an architectural protein

A study in *S. prombe* provided the first evidence of TFIIIC function as a topological device (Noma et al., 2006). TFIIIC binds at inverted repeats (IR) which are boundary elements of heterochromatic domains and prevents the spread of heterochromatin spreading. On the other hand, TFIIIC forms clusters at nuclear periphery which are known as chromosome-organizing clamps (COC). When TFIIIC is recruited to IR and COC, both of which are ETC sites, these chromosomal loci are tethered to nuclear periphery. Majority of COC sites are present close to gene promoters in divergent directions. They may separate genes under different regulatory mechanisms into different domains. Thus, TFIIIC at COC acts similarly to APBS at TAD border.

Research in mouse neurons offered new information about architectural role of TFIIIC. Neurons employs a mechanism in which TFIIIC tethers gene loci to transcription factories for gene expression regulation (Crepaldi et al., 2013). Upon depolarization, TFIIIC binds to SINEs downstream to promoters to activity-dependent neuronal genes. Inducible, but not housekeeping genes, are relocated to transcription factories for transcription. This process is controlled by TFIIIC and gene expression in response to synapsis is therefore coordinated.

CTCF and Rad21 are the most well-known architectural protein partners (Cubeñas-Potts and Corces, 2015). Literature has shown the co-relations and interactions between TFIIIC and different architectural proteins. Yet, it is less explored if TFIIIC couples with an accessory protein for architectural function. And if yes, which accessory protein is the partner of TFIIIC. ETC sites are in close proximity to CTCF binding sites and this association is not found when RNAPIII-bound TFIIIC sites are also included (Moqtaderi et al. 2010). Also, CTCF motif is strikingly enriched at ETC sites in a neuroblastoma cell line (Büchel et al., 2017). Despite TFIIIC and CTCF are in the same protein complex with PRDM5 as shown by co-immunoprecipitation, there is no evidence of direct protein-protein interactions between TFIIIC and CTCF (Galli et al. 2013).

On the other hand, TFIIIC shows direct protein-protein interactions with both DNA-binding and accessory architectural proteins. Examples include BEAF-32, CP190 and Mod/mdg4 in *Drosophila* and CAPH2, RAD21 and PRDM5 in mammalian cell lines (Van Bortle et al. 2014; Galli et al. 2013; Büchel et al. 2017; Yuen, Slaughter, and Gerton 2017). In addition, TFIIIC can be co-immunoprecipitated with other SMC proteins in HEK293 cells (Yuen et al., 2017).

Pertaining to chromatin binding, TFIIIC is a DNA-binding architectural protein and recognizes its canonical B-box. It was first reported that TFIIIC colocalizes with cohesin and condensin in yeasts (D'Ambrosio et al., 2008). Genome-wide ChIP-seq analysis in *Drosophila* has

demonstrated co-occupancy of TFIIIC with other DNA-binding (CTCF, BEAF-32, CP190, etc.) and accessory proteins (Rad21, Cap-H1, Cap-H2) (Van Bortle et al., 2014). A recent paper showed TFIIIC recruits CAPH2 to their joint binding sites which are frequently located at TAD borders in mESC (Yuen et al., 2017). Also, these sites are demarcated by H3K4me3 marks which accounts for open promoters.

1.3 The transcription factor N-MYC

1.3.1 The MYC family

MYC is a textbook example of oncogene due to its wide-spread involvement in various cancer types (Dang, 2012). *MYC* was first identified 40 years ago in avian tumors as the cellular homolog of retroviral transforming gene, *v-myc* (Sheiness and Bishop, 1979). Chromosomal translocation that juxtaposes *MYC* to immunoglobin enhancers in Burkitt's lymphoma first established *MYC* as a human oncogene (Dalla-Favera et al., 1982; Taub et al., 1982). Other members of the MYC family, N-MYC and L-MYC, were afterwards identified in neuroblastoma and small cell lung cancer respectively (Kohl et al., 1983; Nau et al., 1985; Schwab et al., 1983a).

All members of MYC family share as high as 95% sequence homology at five conserved domains termed MYCBoxes, namely MBI, MBII, MBIIIa, MBIIIb, and MBIV (except the missing MBIIIa in L-MYC) (Fig. 1.6). These MBs are sites for direct protein-protein interactions between MYC and other interacting partners to regulate MYC degradation and gene transcription. MYC protein is unstable as reflected by the short half-lives of MYC and N-MYC which are about 20 and 65 minutes respectively (Choi et al., 2010; Hann and Eisenman, 1984; Ikegaki et al., 1986; Ingvarsson et al., 1988). Its turnover is tightly regulated by ubiquitin-proteasomal degradation system (UPS) which recognizes phosphorylated threonine-58 and serine-62 within MBI (Liu and Eisenman, 2012). Another pathway for MYC degradation utilizes the mechanism of ubiquitin-independent proteolysis via the D-element domain at MBIIIa (Herbst et al., 2004). PEST sequence (rich in proline, glutamic acid, serine and threonine) covering MBIIIb is another domain for MYC turnover by UPS (Gregory and Hann, 2000). Also, MBIIIb allows interaction with WDR5 which facilitates MYC binding to chromatin (Thomas et al., 2015).

MYC is a transcription factor under the basic helix-loop-helix-leucine zipper (HLH-LZ) family due to its carboxy-terminal basic HLH-LZ domain. It binds specifically to canonical E-box (CACGTG) upon dimerization with MAX which is another HLH-LZ protein (Blackwell et al., 1993). In the absence of MAX, MYC alone or as homodimer is deprived of any DNA-binding

property. This C-terminal region also allows interaction with MIZ1 (MYC-interacting zinc finger protein 1) and p300 and CBP for gene regulation (Peukert, 1997; Vervoorts et al., 2003). A lysine rich nuclear localization signal (NLS) can be found downstream of MBIV. It directs predominant localization of MYC in cell nucleus. Phosphorylated serine-62 MYC is lately shown to be preferentially located at nuclear pore basket (Su et al., 2018).

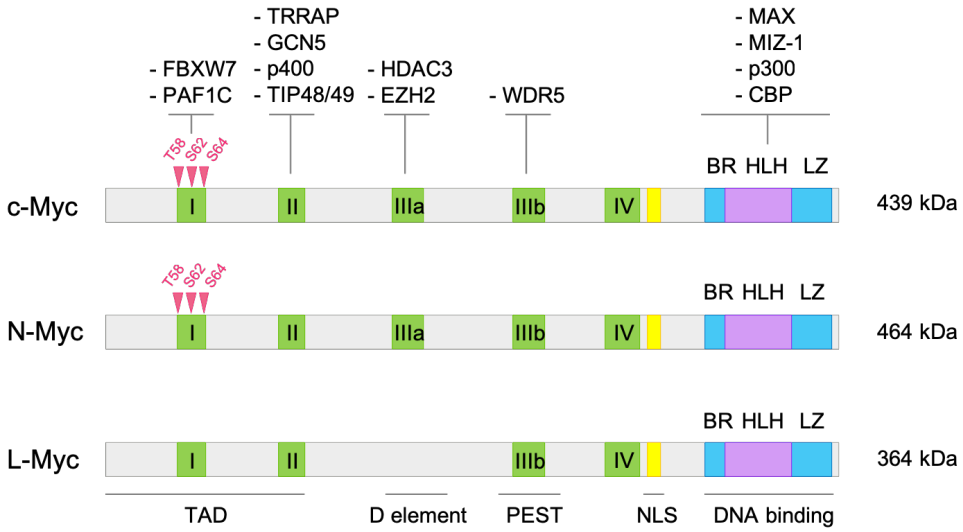


Figure 1.6: Structural and functional domains of human MYC family and the interacting proteins of corresponding domains
All MYC proteins share several highly conserved MYC box (MB) (colored in green), basic region helix-loop-helix-leucine zipper (BR-HLH-LZ) domain (colored in blue and purple) and nuclear localization signal (NLS) (colored in yellow). Reported protein interacting partners with certain MBs are as indicated on top. Functional domains are highlighted at the bottom. Conserved phosphorylation sites threonine-58 (T58), serine-62 (S62) and serine-64 (S64) in c-MYC and N-MYC are labelled in red. S64 is Aurora-A kinase site. TAD = Transactivation domain.

1.3.2 Transcription regulation by MYC family

The essentiality of MYC lies with its role as a master gene regulator. As a transcription factor, MYC modulates the transcriptomes extensively. This does not limit to mRNAs of RNAPII transcription, but also includes ribosomal RNAs of RNAPI and tRNAs of RNAPIII transcriptions (Arabi et al., 2005; Gomez-Diaz and Corces, 2014; Grandori et al., 2005; Steiger et al., 2008). The category of MYC target genes spans from cell growth, apoptosis, cell cycle, to angiogenesis, inflammation, and metabolism, which governs most, if not all, vital cellular processes (Chen et al., 2018; Eilers and Eisenman, 2008). MYC was also reported to be a general amplifier of transcription, which globally enhance transcription of the cell's gene expression program (Lin et al., 2012).

Transcriptional activation by MYC

Transcription activation requires binding to canonical E-box element CACGTG or non-canonical CANNTG (N = any nucleotide) upon MYC/MAX dimerization (Blackwell et al.,

1993; Blackwood and Eisenman, 1991). MYC binds to open promoters with histone marks H3K4me3 and binding sites of WDR5 across the genome (Guccione et al., 2006; Thomas et al., 2015).

First mechanism for MYC-mediated transcription activation is recruitment of co-activators. One example is TRRAP, which can form complex with MYC-recruited HATs such as TIP60 and GCN5. This macromolecule complex collaborate to increase acetylation of H3 and H4, thus forming euchromatin for active transcription (Bouchard et al., 2001; Frank et al., 2003; McMahon et al., 2000). MYC also recruits proteins with chromatin remodeling activities as co-activators, for instance, ATPase TIP48 and TIP49 (Wood et al., 2000).

Second mechanism is the interaction of MYC with RNAPII basal transcription machinery at different stages of a transcription cycle. Studies have shown that MYC activates transcription at loading of RNAPII to promoters (de Pretis et al., 2017), pause-release of RNAPII from promoters (de Pretis et al., 2017; Rahl et al., 2010; Walz et al., 2014) and elongation (Rahl and Young, 2014).

Third mechanism is about occupancies of genes weakly bound by MYC. In comparison to physiological MYC level, over-expression of MYC in tumors increases MYC binding at genes that are weakly bound in normal cells (Lorenzin et al., 2016). These genes are also activated upon elevated MYC level. Therefore, the high MYC level under tumor situation results in more MYC-bound genes which are subjected to transcriptional activation by MYC.

Transcriptional repression by MYC

MYC adopts mechanism to counteract transcriptional activation aforementioned for repression. MYC displaces co-activators p300 and CBP bound at its HLH-LZ domain upon formation of repressive complex with MIZ1. It then recruits co-repressors, histone deacetylases (HDACs) and DNA methyltransferases (e.g. DNMT3a), for a heterochromatic structure (Brenner et al., 2005; Peukert, 1997; Staller et al., 2001). MYC/MIZ1 complex attenuates MYC-mediated transcriptional activation and represses MIZ1-induced target genes. The direction of transcriptional output depends on cellular MYC level. Elevated MYC level favors complex formation with MIZ1 and cooperative binding to MYC-regulated promoters, therefore hindering gene activation by MYC (Walz et al., 2014; Wiese et al., 2015).

Formation of inhibitory complex with PAF1C (polymerase-associated factor 1 complex) also represses MYC-mediated transcription. This repressive complex antagonizes effective elongation upon interaction between MYC and p-TEFb during transcriptional activation (Jaenicke et al. 2016).

Another repressive mechanism is direct recruitment of HDAC3 via MBIIIa to deacetylate H3 and H4 at promoters which have E-box (Kurland and Tansey 2008).
For indirect mechanism, MYC has self-regulatory mechanism for repression via microRNAs (miRNAs). MYC induces miR17-92 which in turn inhibits translation of proteins pivotal for inducing senescence or apoptosis (Li et al., 2014).

1.3.3 Biological function of N-MYC and its dysregulation in neuroblastoma

N-MYC was identified to be an amplified homolog of *v-myc* that is different from *c-MYC* in neuroblastoma (Kohl et al. 1983; Schwab et al. 1983). Some canonical properties of MYC family such as MBs, dimerization with MAX for E-box binding, roles in transformation, proliferation, apoptosis and stem-like state, are well-shared between N-MYC and MYC. Indeed, *MYCN* knock-in mice can rescue the lethality of *MYC* knockout mice and even restore immunity, yet only partially compensate the growth phenotype (Malynn et al., 2000). Furthermore, embryos with homozygous deletion of either *MYC* or *MYCN* cannot survive through gestation (Charron et al., 1992; Davis et al., 1993).

Despite the redundancy in many functions, N-MYC diverges from MYC. One key fact is their distinct spatiotemporal expression patterns (Zimmerman et al., 1986) (Fig. 1.6). N-MYC expression pattern is specific to tissues during early developmental stages. High expression of N-MYC in forebrain and hindbrain during infancy highlights its importance to neurogenesis. N-MYC is responsible for proliferation of progenitor cells and granule neural precursors in the central nervous system (Hatton et al., 2006; Knoepfler et al., 2002). A six-fold size reduction in cerebellum upon conditional deletion of *MYCN* in neural progenitor cells was observed. Importantly, this effect is restrained to N-MYC as deletion of MYC cannot reproduce the same phenotype. An explanation to the discrepancy of functions is that expansion of cerebellar granule neural precursors by sonic hedgehog (SHH) signaling is only associated with *MYCN* transcription (Kenney et al., 2003).

N-MYC is important to neurogenesis. This is reflected by its role in neuronal cell expansion and inhibition to neuronal differentiation. *In vivo* experiment showed highest N-MYC expression in immature cells and the lowest in differentiated adult tissues. Deletion of N-MYC is accompanied with a gain in neuronal differentiation (Knoepfler et al., 2002).

Figure 1.6: MYC and MYCN expression levels in different tissues of newborn and adults
MYC is expressed in a ubiquitous pattern whereas MYCN is preferentially expressed in neural tissue. Both proteins in general have reduced expression in adulthood. Adopted from (Huang and Weiss, 2013).

Preferential expression of N-MYC facilitate neurological development of newborns, however over-expression of N-MYC is a driver event for tumors in nervous and hematological systems as well as neuroendocrine tumors in other parts of the body (Rickman et al., 2018). The most well-studied cancer type for N-MYC dysregulation is neuroblastoma. It is an paediatric extracranial cancer that is most prevalent in children under the age of five (American Cancer Society, 2018). It has been known for more than 30 years that *MYCN* amplification confers poor prognosis and survival of neuroblastoma patients (Brodeur et al. 1984; Seeger et al. 1985; Schwab et al. 1983). It remains until today to be an effective biomarker for neuroblastoma risk stratification.

Differentiation of neuroblastoma cells is corelated to a decrease in N-MYC level (Cinatl et al., 1993; Han et al., 2001; Matsumoto et al., 1989). Neuroblastoma regression is observed upon induction of differentiation, for example, via NGF/trk-A signaling (Matsushima and Bogenmann, 1993; Nakagawara and Brodeur, 1997; Nakagawara et al., 1993). The use of all-trans retinoic acid is therefore one of the options for differentiation therapy (Matthay et al., 2009).

N-MYC is stabilized by Aurora-A in *MYCN*-amplified neuroblastoma (Brockmann et al., 2013; Otto et al., 2009). Aurora-A binds to N-MYC and protects it from proteasomal degradation by FBXW7 (Otto et al., 2009). In addition, gene-encoded by Aurora-A, *AURKA*, is a gene target of N-MYC and its expression in elevated in *MYCN*-amplified neuroblastoma (Berwanger et al., 2002). Interaction between N-MYC and Aurora-A is specific to S-phase at which Aurora-A inhibits N-MYC dependent-pause-release of RNAPII (Büchel et al., 2017). This N-MYC/Aurora-A interaction can in turn prevent the conflicts between transcription and replication. Different Aurora-A inhibitors are developed for therapeutic purposes. Catalytic

inhibitor MLN8237 or “Alisertib” can induce N-MYC degradation (Brockmann et al., 2013). It has entered clinical trial phase II to treat neuroblastoma patients in combination therapy (DuBois et al., 2018; Felgenhauer et al., 2018).

1.4 Objectives

Dysregulation of N-MYC is the driver event of various tumor entities. Therefore, broadening the scope of knowledge in N-MYC cancer biology helps better understanding of tumor pathogenesis and may eventually coming up with new therapeutic windows. Identification of TFIIIC as a new N-MYC interacting partner has opened up an exciting direction for N-MYC cancer biology.

TFIIIC complex is known as a core component of RNAPIII transcription machinery. It is also reported to be an architectural protein in yeasts and *Drosophila*. Intriguingly, TFIIIC binds to thousands of ETC sites across the human genome without well-defined functions. Previous findings have demonstrated that subunit TFIIIC5 co-occupies with N-MYC at sites where CTCF motif is strongly enriched.

As a RNAPIII transcription factor, MYC can activate RNAPIII transcription. However, whether a RNAPIII transcription factor would in return influence RNAPII transcription is still open. Studies in neurons provided evidence that TFIIIC binding to SINEs in close proximity to RNAPII gene promoters can regulate RNAPII transcription of activity-dependent genes (Crepaldi et al., 2013a; Policarpi et al., 2017). This suggests that at least some ETC sites are distributed to sites close to RNAPII TSS for regulatory functions.

This thesis aims to deepen the knowledge of TFIIIC complex and understand how TFIIIC is involved in N-MYC transcriptional regulation. Two TFIIIC subunits, TFIIIC5 and TFIIIC2 which belongs to sub-complex A and B respectively were selected for investigation. The latter contains intrinsic HAT activities. The objectives were to characterize both subunits for their binding to ETC sites. It should be analyzed whether N-MYC interacts with both TFIIIC subunits in the same way. Finally, relationships between TFIIIC, N-MYC and other architectural proteins should be deciphered.

2 Materials

2.1 Nucleic acids

2.1.1 Primers

Primers and oligos used in this thesis were synthesized by Sigma-Aldrich and listed as follows:

Table 2.1 Primers used for cloning

fw = forward, rev = reverse

Name	Application	Sequence (5' to 3')
sgRNA-T1-fw	knockout*	CACCGCCCTGCCAGACGCACAGGGA
sgRNA-T1-rev	knockout*	AAACTCCCTGTGCGTCTGGCAGGG
sgRNA-T2-fw	knockout*	CACCGGCTCATGCTCCGGCCCGAGA
sgRNA-T2-rev	knockout*	AAACTCTCGGGCCGGAGCATGAGC
sgRNA-T3-fw	knockout*	CACCGGAATCCATAGGCTGCGCCAG
sgRNA-T3-rev	knockout*	AAACCTGGCGCAGCCTATGGATTC
sgRNA-T4-fw	knockout*	CACCGCGTGGACCGGAAGGTGGAGG
sgRNA-T4-rev	knockout*	AAACCCTCCACCTTCCGGTCCACG
sgRNA-T5-fw	knockout*	CACCGCATTTCGGGACCAGATGGGA
sgRNA-T5-rev	knockout*	AAACTCCCATCTGGTCCCGAAATG
sgRNA-T6-fw	knockout*	CACCGGAGAACGAGGCGGCAGAAAG
sgRNA-T6-rev	knockout*	AAACCTTTCTGCCGCCTCGTTCTC
sgRNA-T7-fw	knockout*	CACCGATGGTGTGCGTGGAGTACCC
sgRNA-T7-rev	knockout*	AAACGGGTACTCCACGCACACCAT
sgRNA-T8-fw	knockout*	CACCGACCGACCAGAGACCCAGCAC
sgRNA-T8-rev	knockout*	AAACGTGCTGGGTCTCTGGTCGGT
sgRNA-T9-fw	knockout*	CACCGACCATCCGCTCCAAGAGGCC
sgRNA-T9-rev	knockout*	AAACGGCCTCTTGGAGCGGATGGT
sgRNA-T10-fw	knockout*	CACCGCATCTGCCGATAGGGTGGCA
sgRNA-T10-R	knockout*	AAACTGCCACCCTATCGGCAGATG
sgRNA-T11-fw	knockout*	CACCGTGTCAAATGTGACCTCGGAG
sgRNA-T11-rev	knockout*	AAACCTCCGAGGTCACATTTGACA
sgRNA-T12-fw	knockout*	CACCGCGTCAGTGCGGTTGCCCAGC
sgRNA-T12-rev	knockout*	AAACGCTGGGCAACCGCACTGACG
sgRNA-#1-fw	knockout^	CACCGAGCGAGTTAGTCCACCGGGC
sgRNA-#1-rev	knockout^	CACCGAGCGAGTTAGTCCACCGGGC
sgRNA-#4-fw	knockout^	CACCGCGGGCCCTGCCAGACGCACA
sgRNA-#4-rev	knockout^	AAACTGTGCGTCTGGCAGGGCCCG
TFIIIC5-shRNA-#5	TFIIIC5 knockdown$^{\Psi}$	TGCTGTTGACAGTGAGCGCAGCGCAGCA CCTACAACTACATAGTGAAGCCACAGAT GTATGTAGTTGTAGGTGCTGCGCTTTGCC TACTGCCTCGGA
TFIIIC2-shRNA-#2	TFIIIC2 knockdown$^{\Psi}$	TGCTGTTGACAGTGAGCGCCCGTAGAGA TGTCATTACCTATAGTGAAGCCACAGATG TATAGGTAATGACATCTCTACGGATGCCT ACTGCCTCGGA

miRE-Xho-fw	mirE-shRNA generation[#]	TGAACTCGAGAAGGTATATTGCTGTTGACAGTGAGCG
miRE-EcoOligo-rev	mirE-shRNA generation[#]	TCTCGAATTCTAGCCCCTTGAAGTCCGAGGCAGTAGGC
sgTFIIIC5-#1-edit-F	Validation of genomic editing	TCGAGGGAGGTTTGTGTCTT
sgTFIIIC5-#1-edit-R	Validation of genomic editing	GCTCGCGTCATCCTAACTCC
sgTFIIIC5-#4-edit-pp2-F	Validation of genomic editing	GGGAGTTAGGATGACGCGAG
sgTFIIIC5-#4-edit-pp2-R	Validation of genomic editing	CCAGAGTCGGCAGCATCTTA
pIn20_HA-TFIIIC5_fwd	TFIIIC5 over-expression	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCGAAGGAGATAGAACCATGTACCCTTACGACGTGCCCGAC
pIn20_TFIIIC5_no-stop_rev	TFIIIC5 over-expression	GGGGACCACTTTGTACAAGAAAGCTGGGTCCACGTAGTCCAGAATCTCTG

Remarks

* Oligo sequences targeting exons for TFIIIC5 knockout* were adopted from (Hart et al., 2015).

^ Oligo sequences targeting 5' UTR sequence for TFIIIC5 knockout^ were manually designed.

Ψ Oligo sequences targeting exons for TFIIIC knockdownΨ were adopted from (Fellmann et al., 2013).

mirE sequences were based on (Fellmann et al., 2013).

Table 2.2 Primers used for qPCR

Name	**Application**	**Sequence (5' to 3') forward primer**	**Sequence (5' to 3') reverse primer**
b2M *(LJ290 & LJ291)*	qPCR	GTGCTCGCGCTACTCTCTC	GTGCTCGCGCTACTCTCTC
BIRC5	ChIP	CTTTGAAAGCAGTCGAGGGG	CTTTGAAAGCAGTCGAGGGG
CAPH2 intergenic region 83-84 (chr.11: 3602387-3602407; 3602754-3602774)	ChIP	CACGTATCCGGAACACATTG	CACGTATCCGGAACACATTG
CAPH2 intergenic region 87-88 (chr.13: 31439581-31439601; 31439756-31439776)	ChIP	CAGGCCTGTTTGTGCTGTAA	CAGGCCTGTTTGTGCTGTAA
CAPH2 intergenic region 89-90 (chr.14: 102780573-102780593; 102780837-102780857)	ChIP	TAAGACGGGGTGAGAAGTGG	TAAGACGGGGTGAGAAGTGG

CAPH2 pp4	ChIP	TGTGGTGGGGGATTGATATG	CCCTCAGCAGGTTGTCACAT
CAPH2 pp5	ChIP	CCCTCAGCAGGTTGTCACAT	TAGTGCTGCCTCACAGTTGG
CAPH2 pp6	ChIP	GCCATTTTGTGCCTCTGG	GGGAGGGGAGCGATATCTTA
CAPH2 pp7	ChIP	ACTGGTCCTGCCCCACAC	GTTCCTCCTCCGCCAGGT
CAPH2 pp8	ChIP	AGCGAGTGCACCGTTCCT	TAAATAGATTTTTGGAGCAGGGAGA
DPYSL2	ChIP	CACACCAACGGAAACAAGTG	CACACCAACGGAAACAAGTG
EIF2B5	ChIP	TTTTCGTTCCGCACCCTAAC	TTTTCGTTCCGCACCCTAAC
EIF3A	ChIP	GAGAGGAGACGAAGGGGAAC	GAGAGGAGACGAAGGGGAAC
EIF4H	ChIP	CAGCTCTCCAGGTCACCTC	CAGCTCTCCAGGTCACCTC
GALNT14 pp1	ChIP	AATGTGCTCGTCCTACCACA	AATGTGCTCGTCCTACCACA
GALNT14 pp2	ChIP	CTAGACCCAGGATCCGGTTG	CTAGACCCAGGATCCGGTTG
ID2 *(LJ269 & LJ270)*	qPCR	ATATCAGCATCCTGTCCTTGC	ATATCAGCATCCTGTCCTTGC
Intergenic region (chr.16)	ChIP	CACACGAGGGTCCATAACGT	CACACGAGGGTCCATAACGT
negative region *(GB_98_ChIP_negChr1_f_SW & GB_99_ChIP_negChr1_r_SW)*	ChIP	GCAGTTCAACCTACAAGCCAATAGAC	GCAGTTCAACCTACAAGCCAATAGAC
NME1	ChIP	GGGGTGGAGAGAAGAAAGCA	GGGGTGGAGAGAAGAAAGCA
PCDHA	ChIP	GCCCGCTGCTACTGGAGACA	GCCCGCTGCTACTGGAGACA
PLK1	ChIP	GTTTGAATTCGGGGAGGAGC	GTTTGAATTCGGGGAGGAGC
PPRC1 pp1	ChIP	GTGAGGATTAGCGCTTGGAG	GTGAGGATTAGCGCTTGGAG
PPRC1 pp2	ChIP	GAAGGCTGAGACCTCCATGT	GAAGGCTGAGACCTCCATGT
PRPS2 *(FL_215_PRPS2_exp_2_f & FL_216_PRPS2_exp_2_r)*	qPCR	AAACACAATTCCGCAAGAGG	AAACACAATTCCGCAAGAGG
RCC1 pp1	ChIP	AGTGGTCGCTTCTTCTCCTT	AGTGGTCGCTTCTTCTCCTT
RCC1 pp2	ChIP	GTAGCTGGGACTGGAGGTG	GTAGCTGGGACTGGAGGTG
SLC1A5 *(FD_RT_SLC1A5_for & FD_RT_SLC1A5_rev)*	qPCR	ACCATGGTTCTGGTCTCCTG	ACCATGGTTCTGGTCTCCTG

TFAP4 *(LJ391 & LJ392)*	qPCR	ACGGAGAGAAGC TCAGCAAG	ACGGAGAGAAGCTC AGCAAG

Remarks: some primers were obtained from oligo collection. Their original oligo names were marked italic.

2.1.2 Plasmids

Table 2.3: Empty vectors used in this study

Vector	Description
LentiCRISPRv2	lentiviral vector constitutively expressing *S. pyogenes* CRISPR-Cas9
pDONR221	vector for gateway cloning
pINDUCER11	dox-inducible lentiviral expression vector for shRNA with IRES RFP
pINDUCER20	dox-inducible lentiviral expression vector for cDNA with neomycin resistance
pJET1.2	positive selection vector for blunt-ended PCR product inserts

Table 2.4: Packaging vectors used for lentivirus production

Vector	Description
psAX.2	plasmid for lentivirus production, encoding for virion packaging system
pMD2.G	plasmid for lentivirus production, encoding for virion envelope

Table 2.5: Plasmids used in this study

Vector	Description
LentiCRISPRv2-sgTFIIIC5	lentiviral vector constitutively expressing *S. pyogenes* CRISPR-Cas9 and sgRNA targeting CDS or 5' UTR of human *TFIIIC5*
pINDUCER11-shTFIIIC5	dox-inducible lentiviral expression vector with shRNA against *TFIIIC5* mRNA
pINDUCER11-shTFIIIC2	dox-inducible lentiviral expression vector with shRNA against *TFIIIC2* mRNA
pINDUCER20-TFIIIC5-HA	dox-inducible lentiviral expression vector with CDS of human *TFIIIC5* tagged with HA at C-terminal

2.2 Chemicals

All chemicals were purchased from Roth, Sigma-Aldrich, Invitrogen, Merck, Calbiochem and Applichem. Solutions and buffers were prepared in ddH2O unless otherwise specified.

2.3 Buffers and solutions

Blocking solution for PVDF membrane	5% (w/v) skim milk powder in TBS-T
Bradford solution	0.01% (w/v) Coomassie Brilliant Blue G250 8.5% phosphoric acid

	4.75% ethanol solution was filtered and stored in the dark
BSA PBS	0.5% (w/v) BSA in PBS
ChIP elution buffer	1% (v/v) SDS 0.1 M $NaHCO_3$ prepared fresh in aqua dest
ChIP lysis buffer I (Nuclear extract lysis buffer I)	5 mM PIPES pH 8.0 85 mM KCl 0.5% (v/v) NP-40 freshly added protease inhibitor mix (Sigma) at 1:1000 dilution
ChIP lysis buffer II	50 mM HEPES pH 7.9 140 mM NaCl 1 mM EDTA 1% (v/v) TritonX-100 0.1% (w/v) deoxycholic acid sodium salt 0.1% (v/v) SDS freshly added protease inhibitor mix (Sigma) at 1:1000 dilution
ChIP wash buffer I	20 mM Tris HCl pH 8.1 150 mM NaCl 2 mM EDTA 0.1% (v/v) SDS 1% (v/v) TritonX-100
ChIP wash buffer II	20 mM Tris HCl pH 8.1 500 mM NaCl 2 mM EDTA 0.1% (v/v) SDS 1% (v/v) TritonX-100
ChIP wash buffer III	10 mM Tris HCl pH 8.1 250 mM LiCl 1 mM EDTA 1% (v/v) NP-40 1% (v/v) deoxycholic acid sodium salt
Crystal violet solution	0.1% (w/v) crystal violet 20% (v/v) ethanol
DNA loading buffer 6x	40% (w/v) sucrose 0.2% (w/v) bromophenol Blue 0.2% (w/v) xylene cyanol 10 mM EDTA pH 8.0

EDTA 0.5M	0.5 M EDTA adjusted to pH 8.0 using 10 M NaOH autoclaved
HEGN lysis buffer 2x	40 mM HEPES pH 7.8 20% (v/v) glycerol 0.4 mM EDTA 0.2% (v/v) NP-40 2 mM Na-glycerophosphate 20 mM NaF 20 mM $Na_4P_2O_7$ filtered sterile (0.2 µm) and stored at 4 °C freshly added 0 - 450 mM KCl and aqua dest ad 1x and freshly added protease inhibitor mix (Sigma) at 1:100 dilution and phosphatase inhibitors (Sigma) each at 1:1000 dilution
NP-40 lysis buffer	150 mM NaCl 1% (v/v) NP-40 50 mM Tris-HCl pH 8.0 10 mM $Na_4P_2O_7$ 100 mM NaF 2 mM Na_3VO_4 filtered sterile (0.2 µm) and stored at 4 °C freshly added protease inhibitor mix (Sigma) at 1:1000 dilution
Nuclear extract lysis buffer II	50 mM Hepes (pH 7.9) 5 mM $MgCl_2$ 0.2% Triton X-100 20% glycerol 300 mM NaCl freshly added protease inhibitor mix (Sigma) at 1:1000 dilution
Nuclear extract elution buffer	50 mM tris (pH 8.3) 1 mM EDTA 1% SDS
PBS 1x	137 mM NaCl 2.7 mM KCl 10.1 mM Na_2HPO_4 1.76 mM KH_2PO_4 autoclaved
Plasmid prep buffer 1	TE with RNase A at 1:1000 dilution
Plasmid prep buffer 2	200 mM NaOH 1% (w/v) SDS

Materials

Plasmid prep buffer 3	3.1 M potassium acetate (trihydrate) adjusted to pH 4.8 using acetic acid
Sample buffer 6x	1.2 g SDS 6 mg bromophenol blue 4.7 ml 100% glycerol 1.2 ml 0.5 M Tris, pH 6.8 2.1 ml ddH_2O the solution was heated up, 0.93 g DTT was dissolved, solution was aliquotted and keep frozen at -20 °C
SDS running buffer (1X)	25 mM Tris base 250 mM glycine 0.1% (v/v) SDS
SDS separating gel	7.5 - 12.5% (v/v) acrylamide/bisacrylamide 375 mM Tris HCl pH 8.8 0.1% (w/v) SDS 0.1% (w/v) APS 0.1% (v/v) TEMED
SDS stacking gel	4% (v/v) acrylamide/bisacrylamide 125 mM Tris HCl pH 6.8 0.1% (w/v) SDS 0.1% (w/v) APS 0.1% (v/v) TEMED
TNN lysis buffer 1x	50 mM Tris-HCl, pH 7.5 120 mM NaCl 5 mM EDTA 0.5% (v/v) NP-40 10 mM $Na_4P_2O_7$ 2 mM Na_3VO_4 100 mM NaF freshly added protease inhibitor mix (Sigma) at 1:1000 dilution and phosphatase inhibitors each at 1:500 dilution
Transfer buffer 10x	250 mM Tris base 1.5 M glycine
Transfer buffer 1x	prepared by diluting Transfer buffer 10x with 15% (v/v) methanol
TBS 20x	500 mM Tris base 2.8 M NaCl adjusted to pH 7.4 using concentrated HCl
TBS-T	prepared by diluting TBS 20x 0.2% (v/v) Tween-20

	25 mM Tris pH 7.4 140 mM NaCl
TE	10 mM Tris pH 8.0 1 mM EDTA pH 8.0 autoclaved
Trypsin solution	0.25% trypsin 5 mM EDTA 22.3 mM Tris pH 7.4 125 mM NaCl

2.4 Standards, enzymes and kits

2.4.1 Standards

DNA marker	Gene Ruler 1 kb Plus DNA ladder (Thermo Scientific)
Protein marker	PageRuler^{TM} Prestained Protein Ladder (Thermo Scientific)

2.4.2 Enzymes

DNase	Applichem
M-MLV Reverse Transcriptase	Promega
Phusion HF DNA polymerase	Thermo Scientific
Phusion Hot Start HF DNA polymerase	Thermo Scientific
Restriction endonuclease	Thermo Scientific, NEB
RNase-free DNase	Qiagen
RNaseA	Roth
SYBR Green qPCR Master Mix	Thermo Scientific
T4 DNA ligase	NEB

2.4.3 Beads for purification

Dynabeads® Protein A/G	Life Technologies

2.4.4 Kits

ABsolute qPCR SYBR Green Mix	Thermo Scientific
CloneJET PCR Cloning Kit	Thermo Scientific
(DNF-474) High Sensitivity NGS Fragment Analysis Kit (1 bp – 6,000 bp)	Advanced Analytical
Duolink® In Situ PLA Technology [PLA® Probes, Detection Reagents, Wash Buffers (Fluorescence)]	Sigma Aldrich
GeneJET Gel Extraction Kit	Thermo Scientific
MinElute® PCR Purification Kit	Qiagen
NEBNext® ChIP-Seq Library	NEB

NEBNext® Multiplex Oligos Index Set 1 and 2	NEB
PureLink® HiPure Plasmid DNA Purification Kit for MAXiprep	Invitrogen
QIAquick® PCR Purification Kit	Qiagen
Qiagen Gel extraction Kit	Qiagen
Quant-iT™ PicoGreen® dsDNA Kit	Invitrogen

2.5 Antibodies

Table 2.6 List of primary antibodies

WB = western blot, ChIP = chromatin immuno-precipitation, IF = immunofluorescence, IP = immuno-precipitation

Antibody	Host/Isotype	Application	Description
acetylated SMC	rabbit, polyclonal IgG	WB	MBL-PD040 (MBL International)
Aurora-A	rabbit, polyclonal IgG	WB	3092 (Cell Signaling)
BRF1	rabbit, polyclonal IgG	ChIP	A301-227A (Bethyl)
CAPH2	rabbit, polyclonal IgG	ChIP, IP	A302-276A (Bethyl)
CAPH2	rabbit, polyclonal IgG	WB	A302-275A (Bethyl)
CTCF	rabbit, polyclonal IgG	WB, ChIP	ab70303 (Abcam)
H3K27ac	rabbit, polyclonal IgG	ChIP	ab4729 (Abcam)
H3K4me1	rabbit, polyclonal IgG	ChIP	ab8895 (Abcam)
H3K4me3	rabbit, polyclonal IgG	ChIP	ab8580 (Abcam)
H4K20me1	rabbit, polyclonal IgG	ChIP	ab9051 (Abcam)
HA	mouse, monoclonal IgG	WB	901502 (BioLegend)
N-MYC	mouse, monoclonal IgG	WB, ChIP	B8.4.B (Santa cruz)
RAD21	rabbit, polyclonal IgG	WB, ChIP	A300-080A (Bethyl)
RPC32	rabbit, polyclonal IgG	ChIP	sc-21754 (Santa cruz)
SMC	rabbit, polyclonal IgG	WB	A300-060A (Bethyl)
TFIIIC1	rabbit, polyclonal IgG	WB	A301-291A (Bethyl)
TFIIIC2	mouse, polyclonal IgG	WB, ChIP, IP	ab89113 (Abcam)
TFIIIC2	mouse, monoclonal IgG	IF	sc-81406 (santa cruz)
TFIIIC5	rabbit, polyclonal IgG	WB, ChIP, IP	A301-242A (Bethyl)
VINCULIN	mouse, monoclonal IgG	WB	V9131 (Sigma-Aldrich)

Table 2.7: List of secondary antibodies

WB = western blot, IF = immunofluorescence, IP = immuno-precipitation

Antibody	Host/Isotype	Application	Description
Alexa Fluor 647 anti mouse	goat	IF	A-21235, invitrogen
Alexa Fluor 647 anti rabbit	goat	IF	A-21244, invitrogen
anti-mouse IgG-HRP	donkey	WB	sc-2314, Santa Cruz

anti-rabbit IgG-HRP	donkey	WB	sc-2313, Santa Cruz
FITC mouse anti-BrdU	mouse	FACS	B44, 347583, BD Biosciences
Mouse TrueBlot ULTRA	rat	WB (IP)	18-8817-33, Rockland
Rabbit TrueBlot	mouse	WB (IP)	18-8816-33, Rockland

2.6 Strains and cell lines

2.6.1 Bacterial strains

DH5α	*Escherichia coli*, genotype F– Φ80lacZΔM15 Δ(lacZYA-argF) U169 recA1 endA1 hsdR17 (rK–, mK+) phoA supE44 λ– thi-1 gyrA96 relA1; used for plasmid amplification
XL1 Blue	*Escherichia coli*, genotype recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F´ proAB lacIqZΔM15 Tn10 (Tetr)], used for amplification of lentiviral plasmids

2.6.2 Human cell lines

HEK293TN	human embryonic kidney cell line (ATCC)
IMR-5	human *MYCN*-amplified neuroblastoma cell line, stable transfected with the murine ecotropic receptor and Hygromycin resistance gene (kindly provided by Angelika Eggert)

2.7 Cultivation media and supplements

2.7.1 Media and antibiotics for bacterial cell culture

LB-medium	10% (w/v) Bacto tryptone 0.5% (w/v) yeast extract 1% (w/v) NaCl
LB-agar	LB-medium with 1.2% (w/v) Bacto agar was autoclaved and cooled down to 50 °C, antibiotics were then added and 20 ml was poured into 10 cm dishes
Antibiotics	Antibiotics were added to LB-medium or LB-agar: Ampicillin 100 mg/ml Kanamycin 100 mg/ml

2.7.2 Media for mammalian cell culture

Culture medium were prepared from basal medium DMEM and RPMI-1640 with 0.584 g/l L-glutamine (Sigma), FBS (Biochrom, heat inactivated at 56 °C for 30 min before use) and penicillin/ streptomycin (Sigma, 100.000 U/ml) as follows:

HEK293TN	DMEM with 10% (v/v) FBS and 1% (v/v) penicillin/ streptomycin
IMR-5	RPMI-1640 with 10% (v/v) FBS and 1% (v/v) penicillin/ streptomycin

2.7.3 Antibiotics and compounds for mammalian cell culture

CD532	1 μM in DMSO
Doxycycline (Dox)	1 μg/ml in absolute ethanol
G418	400 μg/ml in sterile water
Polybrene	4 μg/ml in sterile water
Puromycin	0.25 μg/ml in absolute ethanol

2.8 Consumables

Consumables such as cell culture dishes, reaction tubes and other disposable plastic items were purchased from the companies Applied Biosystems, Eppendorf, ibidi, Greiner, Kimberley- Clark, Nunc, Sarstedt and VWR.

2.9 Equipment and membranes

Automated capillary electrophoresis	Fragment Analyzer™ (Advanced Analytical)
Chemiluminescence imaging	LAS-4000 mini (Fujifim)
Cell culture incubator	BBD 6220 (Heraeus)
Cell Counter	Casy® cell counter (Innovatis)
Centrifuges	Avanti J-26 XP (Backman Coulter) Eppendorf 5417 R, 5425 and 5430 (Eppendorf) Galaxy MiniStar (VWR) Multifuge 1S-R (Heraeus) Optima L-90K (Beckman Coulter)
Flow cytometer	BD FACS CantoTM II (BD Biosciences)
Heating block	Dry Bath System (Starlab) Thermomixer® comfort (Eppendorf)
Incubator shaker	Model G25 (New Brunswick Scienti□c)
Microscopes	Axiovert 40CFL (Zeiss) Eclipse Ti-E (Nikon) TCS SP5 (Leica)
PCR thermal cycler	Mastercycler pro S (Ependorf)
Photometer	Spectrofluorometer NanoDrop 1000 (Thermo Scientific) UltrospecTM 3100 pro UV/Visible (Amersham Biosciences)
Power supply	Power Pac (Bio-Rad)
PVDF transfer membrane	Immobilon-P transfer membrane (Millipore)
Quantitative RT-PCR machine	StepOne™ Realtime Cycler (Applied Biosystems)
SDS-PAGE system	Mini-PROTEAN Tetra Cell (Bio-Rad)
Sterile bench	HeraSafe (Heraeus)
Ultrasonifier	Digital Sonifier® W-250 D (Branson)
UV fluorescent table	Maxi UV fluorescent table (Peqlab)
Vortex mixer	Vortex-Genie 2 (Scientific Industries)
Water bath	Julabo ED-5M water bath (Julabo)

Immunoblot transfer chamber	PerfectBlue Tank Electro Blotter Web S (Peqlab)
Whatman filter paper	Gel Blotting Paper (Schleicher and Schuell)

2.10 Software and online programs

Acrobat™ 9 Pro v9.5.5	Adobe Inc.
BD FACSDiva 6.1.2	BD Biosciences
Bedtools v2.19.1	(Quinlan and Hall, 2010)
Bowtie v1.1.1	(Langmead et al., 2009)
DeepTools	(Ramírez et al., 2014)
FastQC	www.bioinformatics.babraham.ac.uk/projects/fastqc/
Illustrator™ v14.0.0	Adobe Inc.
ImageJ	NIH
JASPAR database	(Khan et al., 2018)
Integrated Genome Browser	(Freese et al., 2016)
Mac OS High Sierra	Apple Inc.
MACS v1.4.2	(Zhang et al., 2008)
MEME Suite	(Bailey et al., 2009)
Microsoft Office 2016 Mac	Microsoft Inc.
Multi Gauge v3.2	Fujifilm
Photoshop™ v11.0.2	Apple Inc.
Primer3 v0.4.0	http://frodo.wi.mit.edu/
PROSize®	Advanced analytical
Sicher	(Xu et al., 2014)
StepOne™ Software v2.3	Applied Biosystems
UCSC Genome Bioinformatics	http://genome.ucsc.edu

3 Methods

3.1 Molecular biology methods

3.1.1 Transformation of competent cells with plasmid DNA and plasmid amplification

Chemically competent bacteria were thawed on ice and mixed with 1 µg plasmid DNA or ligation reaction. Bacteria were incubated on ice for 30 min and heat shocked for 90 s at 45 °C. Bacteria were then recovered on ice for 5 min and supplied with 1 ml LB medium. The bacteria suspension was shaked at 37 °C for 1 h before plated on LB agar plates with appropriate antibiotics for selection. LB agar plate were incubated overnight at 37 °C.

3.1.2 Isolation of plasmid DNA from bacteria

Large-scale purification of plasmid DNA required 200 ml bacterial culture and was done according to manufacturer's instructions (PureLink HiPure Plasmid Maxiprep Kit, Life Technologies). Purified plasmid DNA was dissolved in Ampuwa water and adjusted to 1 mg/ml.

Mini-preparation of plasmid DNA required 3-5 ml bacterial culture and was done according to manufacturer's instructions (PureYield™ Plasmid Miniprep System, Promega). Purified plasmid DNA was dissolved in Ampuwa water.

3.1.3 Gel Electrophoresis of DNA fragments

Solution of 1-2% agarose in 1X TAE was prepared depending on the size of DNA fragment. The solution was boiled and supplemented with 0.3 µl/ml ethidium bromide, then poured into a gel chamber with combs. Samples were mixed with DNA loading buffer and loaded into the wells of the gel. DNA ladder (Thermo Scientific) was loaded next to the samples for size determination of DNA. Gel electrophoresis was run at 120 V for one hour and DNA was visualized using a UV transilluminator.

3.1.4 DNA extraction and purification

Purification of DNA fragments excised from agarose gel was performed with a gel extraction kit according to manufacturer's instructions (GeneJET Gel Extraction Kit, Thermo Scientific). PCR products was purified likewise according to manufacturer's instructions (GeneJET PCR Purification Kit, Thermo Scientific).

3.1.5 Cloning of shRNA mediated gene knockdown

shRNA mediated gene knockdown was employed to study the biological role of gene-of-interest. mir-RNA embedded shRNAs were improved to have higher knockdown efficiency and lower toxicity to cells (Fellmann et al., 2013). The 97-mer sequences of shRNAs reported from (Fellmann et al., 2013) were converted to mir-E shRNAs by PCR with mir-E primers and PCR product of about 125 bp was cloned into blunt-end pJET1.2 vector. Afterwards, mir-E shRNAs were restriction digested by EcoRI and XhoI and ligated into sticky-end pINDUCER11 vector.

1. Cloning of shRNA into pJET1.2

(i) Generation of mir-E shRNA

PCR recipe

Primer mir E	1 µl
100 µM shRNA oligo	1 µl
10mM dNTPs	1 µl
DMSO	1 µl
GC buffer	10 µl
Water	36 µl
Phusion Hi-Fi DNA polymerase	0.5 µl
Total volume	50 µl

PCR

Cycle step	Temperature	Time	Cycle
Initial denaturation	98 ºC	3 min	1 x
Denaturation	98 ºC	25 s	25 x
Annealing	54 ºC	30 s	
Extension	72 ºC	60 s	

(ii) Ligation of shRNA into pJET1.2

Purified PCR product	15.5 µl
pJET1.2 blunt (Thermo Scientific)	0.5 µl
T4 ligase (Fermentas)	2 µl
T4 ligase buffer	2 µl

Ligation reaction was performed with T4 ligase with 15 min incubation at RT.

2. Cloning of shRNA into pINDUCER11

(i) Restriction digest of insert

Maxi-prep (shRNA in pJET1.2)	26 µl
EcoRI	2 µl
XhoI	2 µl
Tango buffer (2X)	10 µl
Water	10 µl

(ii) Restriction digest of vector

Maxi-prep (pINDUCER11)	15 µl
EcoRI	2 µl
XhoI	2 µl
Tango buffer (2X)	10 µl
Water	10 µl

Restriction cut with EcoRI and XhoI was performed at 37 ºC for 1 h. Digested fragment was extracted after gel electrophoresis and eluded in 20 µl Ampuwa water.

(iii) Ligation of shRNA into pINDUCER11

pINDUCER vector	1 µl
insert	7 µl
T4 ligase (Fermentas)	1 µl
T4 ligase buffer	1 µl

Ligation reaction was performed with T4 ligase with 15 min incubation at RT.

3.1.6 Cloning of CRISPR-Cas9 mediated gene knockout

1. **Design of single guide RNA (sgRNA)**

sgRNAs targeting gene-of-interest were either designed manually or by online tool CHOP-CHOP (http://chopchop.cbu.uib.no/index.php). For manual design, the sequence targeted should contain a PAM sequence at 3' end. An overhang which is compatible to the vector backbone upon BsmBI digestion should be included in the sgRNA sequence (Cong and Zhang, 2015; Sanjana et al., 2014). For computationally predicted sgRNA sequences, Doench 2016 cleavage efficiency should be considered for sgRNA selection. All sgRNAs were designed to target 5' UTR of TFIIIC5 which allows the differentiation between endogenous and exogenous TFIIIC5.

2. **CRISPR cloning**

Cloning of sgRNA into lentiCRISPRv2 followed the protocol published by Zhang lab (Cong and Zhang, 2015; Sanjana et al., 2014). Briefly, each pair of sgRNA oligos was phosphorylated and annealed by PNK. LentiCRISPRv2 was digested by BsmBI and dephosphorylated. The digested vector (about 13 kb) was gel purified which then ligated with 1:200 diluted oligos. The ligation product was transformed into XL-1 Blue.

3.1.6 Cloning of exogenous TFIIIC5

Construct of exogenous TFIIIC5 in pINDUCER21 was prepared for exogenous expression of TFIIIC5. Cloning of cDNA employed Gateway cloning technology and was done according to manufacturer's protocol (Life Technologies). Briefly, specific primers which contain *att*B1 and

*att*B2 sites were designed for amplification of cDNA to be cloned. First recombination reaction (BP) was done to insert PCR product with *att*B sites into pDONRTM vector. This allowed the second recombination reaction (LR) which transfer the entry clone from BP reaction to pINDUCER vector.

3.1.7 Nucleic acid quantification

3.1.7.1 Nanodrop

DNA and RNA concentration and quality were determined by NanoDrop 1000 (PeqLab). Absorbance at 260 nm and ratio of absorbance at 260 and 280 nm (A260/280) was measured for the quantification and purity of nucleic acid respectively. A260/280 for pure DNA is about 1.8 and for pure RNA is about 2.

3.1.7.2 PicoGreen

Concentration of double-stranded DNA (dsDNA) was determined by Quant-iT PicoGreen dsDNA reagent (Invitrogen) according to manufacturer's instructions. PicoGreen is a fluorescent dye that intercalates into dsDNA and the fluorescence intensity can be determined at wavelength 485/535 nm. Concentration of chromatin samples from ChIP were measured by PicoGreen before the samples were proceeded to library preparation for ChIP-sequencing.

3.1.7.3 Fragment Analyzer

Size and precise quantification of DNA used for ChIP-sequencing or 4C-sequencing was determined by Fragment Analyzer (Advanced Analytical) according to manufacturer's instructions.

3.1.8 RNA isolation

Total RNA was isolated from cultured cells using TriFast reagent (Peqlab). Cells were harvested by cold PBS and pelleted at 1500 rpm at 4 °C for 5 min. Cell pellet was resuspended in 1 ml TriFast. Samples were processed immediately or stored at –80 °C. Upon 5 min incubation at RT, 200 µl chloroform was added and the mixture was vortexed thoroughly for 15 s. The mixture was incubated at RT for 5 min followed by centrifugation at 14,000 rpm at 4 °C for 10 min. Aqueous and organic phases were then separated and the upper aqueous phase was transferred to a new reaction tube. 500 µl isopropanol and 1 µl GlycoBlue (15 µg/µl stock solution, Thermo Scientific) were added to precipitate RNA in the isolated aqueous phase. Samples were incubated on ice for 20 min and centrifugated at 14,000 rpm at 4 °C for 10 min.

The RNA pellet was washed twice with 75% ethanol, then air-dried and solubilized in 20 µl Ampuwa water. The RNA was used for cDNA synthesis, otherwise stored at – 80 °C.

3.1.9 cDNA synthesis

To analyze the expression of a specific gene, RNA was reversely transcribed into complementary DNA (cDNA) with random hexanucleotide primers. 2 µg total RNA diluted in 10 µl Ampuwa was first incubated at 65 °C for 1 min to dissolve secondary structures and then cooled down quickly on ice. 40 µl cDNA synthesis mix was added per reaction which was then incubated stepwise at RT for 10 min, 37 °C for 50 min and 70 °C for 15 min.

cDNA synthesis mix	
5x First strand reaction buffer (Promega)	10 µl
dNTPs (2.5 mM, Roth)	5 µl
Random primer (2 mg/ml, Roche)	2 µl
Ribolock (Fermentas)	0.2 µl
M-MLV reverse transcriptase (200 U/ml, Promega)	1 µl
Ampuwa water	ad 40 µl

50 µl cDNA product was further diluted in 450 µl Ampuwa water and 10 µl of diluted cDNA was used per reaction of qPCR (see section 3.1.10).

3.1.10 Quantitative reverse real-time PCR (qPCR)

The abundance of mRNAs or DNA fragments after chromatin immunoprecipitation was measured by quantitative PCR. Amplification of a specific part of the template DNA was coupled with intercalation of a fluorescent dye into the amplified DNA. Fluorescence intensity was quantified during the PCR process in real time. Technical triplicates for each sample were performed and the reactions were set up as follows:

Component	**Amount**
Diluted cDNA/chromatin	10 µl
SYBRGreen Mix (Thermo Scientific)	20 µl
Forward primer	5 pmol
Reverse primer	5 pmol
Ampuwa water	Ad 20 µl

The qPCR reaction was performed using StepOne™ Realtime Cycler (Applied Biosystems) with the following thermal cycling profile:

Temperature	Time	Cycle
95 °C	15 min	1 x
95 °C	15 s	40 x
60 °C	20 s	
72 °C	15 s	
95 °C	15 s	1x
60 °C	1 min	1x
60-95 °C; stepwise	15 s	0.3 °C step
95 °C	15 s	1x

The relative abundance of mRNAs or DNA enrichment was calculated using the ΔΔ-CT method (Applied Biosystems User Bulletin 2). House-keeping gene β2M and input sample were used for the normalization of RNA and ChIP samples respectively.

3.2 Cell biology methods

3.2.1 Cultivation of eukaryotic cell lines

Eukaryotic cell lines used in this study were grown and maintained in a cell incubator at 37 °C, 5% CO_2, and relative humidity of 95%.

3.2.1.1 Passaging cells

Cell line was passaged when the dish is 90% confluent. Cultivation medium was removed, cells were washed with PBS and trypsinized to detach from cell culture dish upon incubation in the cell incubator for about 2 min. The enzymatic activity of trypsin was quenched by adding serum-containing medium. Cells were replated with pre-warmed medium at a desired density on a new cell culture dish. Neubauer counting chamber or CASY cell counter was used if a specific cell number was to be plated.

3.2.1.2 Freezing and thawing cells

Cells were cryopreserved in liquid nitrogen for long-term storage. Cells were trypsinized from cell culture dish as described above, resuspended in fresh medium and pelleted (1200 rpm, 5 min, RT). Cell pellet was then resuspended in 1 ml freezing medium, transferred to cryo vials and slowly frozen (1 °C per min) to -80 °C. After that, vials were kept at liquid nitrogen storage tanks.

Frozen cells were rapidly thawed in a 37 °C water bath and cells were resuspended with pre-warmed medium. Cell resuspension was transferred to a new cell culture dish with pre-warmed medium.

3.2.2 Transfection of plasmid DNA

Cells were transfected with plasmid DNA with polyethylenimin (PEI). For that, cells were seeded at a specific density which gave around 50% confluency on the next day. Shortly prior to transfection, cells were washed with PBS and replenished with antibiotics-free medium of low serum (2%). Transfection reaction was prepared by incubating 30 µl polyethylenimine (PEI) in 700 µl Opti-MEM medium for 5 min at RT. In a separate tube, 10 µg of plasmid DNA was diluted in 700 µl Opti-MEM medium. DNA-containing Opti-MEM medium was added dropwise to PEI-Opti-MEM solution. The mixture was mixed thoroughly by inverting the tubes and incubated at RT for 20 min. Afterwards the transfection complex was added to the cells. Cells were washed with PBS and replenished with cultivation medium with antibiotics 4 to 5 h following transfection.

3.2.3 Lentivirus production

4.5 – 5 million HEK293T cells were plated per 10 cm dish one day prior to transfection. Cells were replenished with antibiotics-free medium of low serum (2%) shortly before transfection. Transfection was performed as described in 3.2.2 with 2.8 µg packaging vector psPAX2, 1.4 µg envelop vector pMD2.G. and 11.1 µg lentiviral expression plasmid. After 14 – 16 h of transfection, medium was removed and fresh medium was added. Lentivirus-containing supernatant was harvested 48 h and 60 h following transfection. Pooled viral supernatant was filtered sterile using a syringe with a 0.45 µm filter and stored in aliquots at –80 °C.

3.2.4 Ultra-centrifugation of lentivirus (concentration of lentivirus)

High titer of virus was obtained by concentration of virus by ultra-centrifugation. Filtered lentivirus-containing supernatant was centrifuged at 25,000 rpm at 4 °C for 90 min (Optima L-90K, Beckman Coulter). Supernatant was aspirated carefully and pellet was air-dried for 30 min. Concentrated lentivirus was dissolved in 50 – 100 µl PBS and stored at –80 °C.

3.2.5 Lentiviral infection

Cells were seeded at a specific density which gave about 90% 48 h after infection. For each 10 cm dish, a total of 6 ml infection mixture was composed of 5 ml medium, 1 ml virus, and 4 µg/ml polybrene. The amount of virus applied to each infection depended on the lentiviral backbone and virus titer. Infection was repeated on the next day and medium was changed 24 h after the second infection. Selection with appropriate antibiotics or by FACS was performed when cells were recovered 48h post-infection.

3.2.6 Colony formation assay (crystal violet staining)

Cell growth and survival of adherent cell lines were determined by crystal violet staining. Same number of cells were plated on 10 cm petri dishes and cultured. Cells were fixed by adding 10 ml formaldehyde (37%) with 10-min incubation at room temperature. The dishes were dried overnight. Crystal violet solution was then added to stain the cells for 30 min. Excessive dye was washed away with desalted water and dishes were allowed to dry overnight.

3.2.7 Cumulative growth curve

Proliferation capability of a cell line was analyzed with cumulative growth curve. Equal cell number (1 x 10^6 for IMR-5) was plated in triplicate on 10 cm dish. Cells were trypsinized every three days and total number of living cells was determined by CASY cell counter. The original cell number was plated again. Increase in cell number (R) (R=X/1,000,000, with X being the total viable cell counts) was determined. Cumulative cell number (Y) of each passage (p) was calculated ($Y_{(p)}=Y_{(p-1)}$ x R) and graphically illustrated.

3.2.8 Propidium iodide staining for flow cytometry (PI-FACS)

Cell cycle distribution and DNA content were determined by propidium iodide (PI) staining of DNA followed by FACS (fluorescence-activated cell sorting) analysis. Adherent cells were trypsinized and centrifuged at 400 g at 4 °C for 5 min. Cell pellet was washed with cold PBS and again centrifuged which was then resuspended in 1 ml cold PBS in a 15 ml polystyrene tube. Cells were fixed by adding 4 ml ice-cold absolute ethanol dropwise with vortexing. Fixed cells were kept at –20 °C overnight. On the day of FACS analysis, cells were centrifuged at 400 g at 4 °C for 10 min and washed twice with 5 ml ice-cold PBS. Washed cells were centrifuged at 400 g at 4 °C for 5 min and resuspended in 400 µl 38 mM sodium citrate with 1 µl RNase A (stock 10 mg/ml) and 15 µl PI (stock 1 mg/ml) in a 1,5 ml reaction tube. Samples were incubated at 37 °C for 30 min in dark and afterwards transferred to FACS tubes for measurement (excitation at 488 nm; longpass- and bandpass-filter at 556 nm and 585/42 nm for propidium iodide respectively.). The cell cycle distribution was analyzed with BD FACSDiva 6.1.2 software.

3.2.9 Indirect immunofluorescence for super-resolution fluorescence microscopy using (dSTORM)

To analyze subcellular localization of proteins, cells were plated in chambered slides (Sarstedt) coated with Lab-Tex or collagen. On the following day, cells were fixed with 4% paraformaldehyde (pH 7.4) for 10 min at 37 °C. Cells were then further fixed with absolute ice-

cold methanol for 20 min at -20 °C. After that, cells were permeabilized by incubating them with 0.1% Triton X-100 in PBS for 15 min at RT. Permeabilized cells were then blocked by 4% BSA in PBS solution for 30 min at RT, followed by incubation with primary antibody diluted in 4% BSA in PBS solution overnight at 4 °C. On the following day, the chambered slides were washed three times with PBS, each wash for 5 min and incubated with fluorescence-labelled secondary antibody for 60 min at RT, followed by Hoechst solution diluted in PBS for 10 min at RT with light protection. Chambered slides were washed three times with PBS and air-dried which then mounted by ibidl mounting solution. Samples were stored at 4 °C light-protected until analysed with dSTORM (magnification 100x).

3.2.10 In situ proximity ligation assay (PLA)

To demonstrate that two proteins-of-interest are in close proximity, Duolink® PLA Technology (Sigma-Aldrich) was applied. Proximity ligation assay gives fluorescent signal when two proteins-of-interest are closer than 40 nm. PLA was performed according to manufacturer's instructions. Briefly, an appropriate cell number was seeded in µ-slide 18 well flat chamber (ibidl) and allowed for overnight cell attachment. On the next day, cells were fixed by 4% paraformaldehyde, permeabilized by methanol and blocked with 5% BSA in PBS. Primary antibodies were diluted in 5% BSA and applied to cells for overnight incubation at 4 °C. On the next day, PLA probes which can recognize the host of the primary antibodies were applied. Ligation and amplification reactions were then performed using the kit reagents. Hoechst solution diluted in PBS (1:250, stock conc. ???) was applied to stain the nucleus for 10 min at RT. Chambered slides were air-dried and mounted by ibidl mounting solution. Samples were stored at 4 °C light-protected for up to three days until analyzed with confocal microscope (Nikon Ti-Eclipse, 60x magnification).

3.3 Protein biochemistry methods

3.3.1 Preparation of whole cell protein extracts

Whole cell protein extracts were prepared on ice. Cells were washed twice with ice-cold PBS and scraped. Cell pellet was obtained from centrifugation at 1200 rpm at 4 °C for 5 min. Total protein was isolated by resuspending the cell pellet with 50 – 100 µl NP-40 lysis buffer with 1: 1000 freshly added protease inhibitor mix. Cells were shock frozen in liquid nitrogen for 30 s followed by 2 min-incubation at 37 °C water bath. This freezing/thawing step was done for three times. Samples were then incubated on ice for 10 min and centrifuged at 13000 rpm at 4

°C for 5 min. The supernatant was transferred to a new reaction and protein concentration was determined by Bradford assay (see 3.3.3).

3.3.2 Preparation of nuclear protein extracts

Whole procedure of nuclear protein extraction was kept cold. Cell suspension collected from after two times washing with PBS was pelleted at 1200 rpm at 4 °C for 5 min. Cell pellet was then resuspended in ice-cold nuclear extract lysis buffer I and incubated on ice for 10 min. Afterwards, suspension was centrifuged at 500 x *g* at 4 °C for 10 min. Isolated nuclear pellet was resuspended in lysis buffer II and incubated on ice for 30 min. Sample was then centrifuged at 12000 x *g* at 4 °C for 20 min. Supernatant was collected as nuclear protein extract and concentration was determined by Bradford assay (see 3.3.3).

3.3.3 Determination of protein concentration by Bradford assay

Protein lysates were quantified according to the Bradford method (Bradford, 1976). 1 µl of protein lysate was mixed with 900 µl Bradford reagent and 100 µl 150 mM NaCl. Absorbance was measured at 595 nm with a reference using lysis buffer instead of protein lysate. Protein concentrations were calculated using a standard calibration curve.

3.3.4 SDS polyacrylamide gel eletrophoresis (SDS-PAGE)

Proteins were separated based on molecular weight by discontinuous SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis). Protein lysates were boiled in 3x sample buffer at 95 °C for 5 min. Equal amounts of protein were loaded on a SDS polyacrylamide gel consisting of a 10% stacking gel and a 4% resolving gel. PageRuler Pre-stained Protein Ladder (Fermentas) was used as a size marker. Gel electrophoresis was performed in Bio-Rad SDS-PAGE chamber with 1x SDS running buffer at 80-120 V.

3.3.5 Western Blot

Following separation of proteins by SDS-PAGE, proteins were transferred onto a polyvinylidene difluoride (PVDF) membrane. The membrane was incubated in methanol for 30 s and equilibrated in transfer buffer. The gel and membrane were layered on one another and fixed between Whatman filter papers in a tank blot transfer chamber filled with 1x transfer buffer. Transfer was carried out in a cold room at 250 mA for 2-3 h depending on protein size. The membrane with immobilized proteins was blocked with blocking solution for at least 10 min followed by overnight incubation with primary antibody at 4 °C. On the next day, the membrane was washed three times with TBS-T and then incubated with secondary antibody

for 1 h at RT. Following further three-time washing with TBS-T, the membrane was subjected to Western Chemiluminescent HRP Substrate (Millipore) for signal detection with LAS-4000 mini (Fujifilm).

3.3.6 Co-immunoprecipitation (Co-IP)

3.3.6.1 Co-IP with whole cell lysate

Whole cell lysate was harvested and resuspended in 1.5 ml TNN lysis buffer. Cell resuspension was incubated on ice for 30 min with occasional vortex followed by sonification (4 x 5 sec with 10 sec pause, 20% amplitude). Sonified suspension was centrifuged for 5 min at 18000 g at 4 °C. Supernatant was kept for Co-IP. 1 mg lysate was used per IP and 4 µg antibody or immunoglobin control. The mixture was incubated overnight at 4 °C with rotation. On the next day, 50 µl of sepharose beads or magnetic A or G Dynabeads was prepared for each IP. Beads were washed twice with 1 ml lysis buffer three times by spinning at 300 g for 5 min at 4 °C before being transferred to cell lysate with antibody. The mixture was then incubated for further two hours at 4 °C with rotation. Samples were then centrifuged down and supernatant was removed. Beads were washed as follows: 1 ml TNN lysis buffer twice, high salt TNN lysis buffer (200 mM) once, 1 ml TNN lysis buffer once and 1x PBS once. Precipitated proteins were eluted in 40 µl 3x SDS sample buffer by boiling at 95 °C for 5 min. 15 µl of eluate per IP and 1% input were loaded to SDS-PAGE gel for further analysis.

3.3.6.2 Co-IP with nuclear extract

1 mg nuclear protein obtained from 3.3.2 was incubated with antibody overnight at 4 °C with rotation. On the next day, 50 µl of magnetic A or G Dynabeads were prepared by washing with PBS followed by lysis buffer II. Each washing step lasted 5 min, lysis and samples were rotated. Washed Dynabeads were transferred to the nuclear extract coupled with antibody and incubated with rotation for two hours at 4 °C. Afterwards, samples were washed three times with lysis buffer II with rotation at 4 °C, each wash 5 min. To elute the protein, 40 µl elution buffer was added and samples were subjected to 1400 rpm shaking at a thermo-shaker at 65 °C for 10 min. 10 µl of eluate per IP and 1% input were loaded to SDS-PAGE gel for further analysis.

3.3.7 Chromatin immunoprecipitation (ChIP)

Chromatin immunoprecipitation was carried out to study interactions between DNA and protein. DNA and its associated proteins were crosslinked. Sheared DNA fragments associated

with protein-of-interest were enriched by immunoprecipitation. Precipitated DNA was purified and quantified by qPCR.

Formaldehyde fixation and chromatin isolation

Cells were cultured on 15 cm dishes and should achieve 80% confluency on the day of harvest per IP. In order to crosslink proteins to DNA, cells were fixed with 1% formaldehyde and incubated at RT for 10 min with gentle shaking. Fixation was quenched by addition of glycine with a final concentration of 20 mM and plates were incubated at RT for 5 min with gentle shaking. Cells were then washed twice with ice-cold PBS and scraped off with PBS containing 1:1000 protease inhibitor mix. Harvested cells were pelleted at 1200 rpm at 4 °C for 5 min. First lysis step was done by resuspending cell pellet in 3 ml ChIP lysis buffer I with 1:1000 protease inhibitor mix followed by 20 min incubation on ice. After centrifugation at 1200 rpm at 4 °C for 5 min, second lysis step was performed by resuspending cell pellet in 2 ml ChIP lysis buffer II with I with 1:1000 protease inhibitor mix followed by 10 min incubation on ice for complete disruption of cellular membranes. Samples were then sonified under cell line-specific conditions. Chromatin of IMR-5 and SHEP cells was fragmented for 20 min at 20% amplitude. Each sonification lasts for 10 s and is followed by 30 s pause. Size of the fragmented chromatin was then determined.

Chromatin size check

To check the size of sonified fragment, 25 µl chromatin was diluted with 475 µl TE buffer. 160 mM NaCl and 20 µg/ml RNase were added to diluted chromatin before reverse-crosslinking step which consisted of 1h incubation at 37 °C and overnight shaking at 65 °C. Subsequently, 5 mM EDTA and 200 µg/ml protease K were added to the samples which were then incubated at 45 °C for 2 h. Chromatin was purified by phenol-chloroform extraction (as described below). Purified DNA was solubilized in 25 µl TE buffer and was loaded on 2% agarose gel for size determination. Sonified chromatin was centrifuged at 13,000 rpm for 15 min and the supernatant was transferred to a new reaction tube. This centrifugation step was performed twice.

Coupling of antibodies to protein A or G Dynabeads

30 µl of protein A or G Dynabeads was used per immunoprecipitation. Beads were washed three times with 1 ml of 5 mg/ml BSA in PBS solution and resuspended with 1 ml of the same bead-washing solution. 3 µg antibodies were incubated with the beads overnight at 4 °C on a

rotating wheel. On the following day, the beads were washed three times with 1 ml of 5 mg/ml BSA in PBS solution and resuspended with 30 µl of the same bead-washing solution.

Immunoprecipitation

Concentration of chromatin obtained from different treatments was measured by NanoDrop measurement using the sonified chromatin after purification. Same amount of chromatin was incubated with prepared Dynabeads overnight at 4 °C on a rotating wheel. 1% of chromatin used for an IP was kept as input. Beads were subsequently washed three times with cold ChIP wash buffer I, II, and III (washing step with wash buffer III requires 5 min incubation at 4 °C on a rotating wheel). An additional washing step with TE buffer was done before transferring the mixture to a new reaction tube.

Elution and decrosslinking

The precipitated DNA was eluded from Dynabeads by incubating the magnetic beads with 150 µl freshly prepared ChIP elution buffer on a rotating wheel at RT for 15 min. The elution step was performed twice and the eluates were pooled. Input was supplemented with 300 µl ChIP elution buffer. 160 mM NaCl and 20 µg/ml RNase were added to the eluted and input samples which were then reverse-crosslinked by 1 h incubation at 37 °C followed by overnight shaking at 65 °C. On the next day, 5 mM EDTA and 200 µg/ml protease K were added to the samples which were then incubated at 45 °C for 2 h.

DNA purification and qPCR

DNA was purified by phenol-chloroform extraction with subsequent ethanol precipitation. 300 µl of the lower phase from phenol-chloroform-isoamyl alcohol mixture (25:24:1) was added to samples and vortexed for 15 s. Phase separation was done by centrifugation at 14,000 rpm for 5 min at RT. The upper aqueous layer was transferred to a new reaction tube and mixed with 1 ml ice-cold absolute ethanol, 50 µl 3M sodium acetate at pH 5.2 and 1 µl Glycoblue. Upon a minimum of 30 min incubation at -20 °C, DNA was pelleted by centrifugation at 13,000 rpm at 4 °C for 30 min. DNA pellet was washed with 500 µl 70% ethanol and air-dried before resuspension in 500 µl Ampuwa water. 10 µl DNA was used for each qPCR reaction.

3.3.8. ChIP-sequencing library preparation

3.3.8.1 ChIP for deep sequencing

For ChIP-sequencing, a standard ChIP was performed as described in section 3.3.6. with some modifications. The cell number for one ChIP was increased by 5-fold. The amounts of antibodies and Dynabeads were increased to 10 µg and 100 µl respectively. Washing step with Wash Buffer III was lengthened from 5 min to 15 min. Elution was performed with 250 µl elution buffer twice. Chromatin was reversely crosslinked at 37 °C for 2 h followed by 65 °C overnight. Phenol-chloroform extraction was carried out with one extra washing step. Purified DNA was solubilized in 30 µl Ampuwa water and quantified using Quant-iT PicoGreen dsDNA reagent (Invitrogen).

3.3.8.2 Library preparation

ChIP-seq library was prepared according to the manufacturer's protocol (NEBNext® ChIP-Seq Library Prep Master Mix Set for Illumina). Libraries were size-selected (~ 200 bp) by separation on agarous gel and purified by gel extraction (Qiagen). Libraries were amplified with 16 PCR cycles. Quantification and size determination were carried out using Fragment Analyzer (Advanced Analytical).

3.4 Sequencing analysis

Bioinformatic analyses of ChIP-seq data were performed by Dr. Susanne Walz (Core Unit Bioinformatics. Comprehensive Cancer Center Mainfranken).

Quality control and mapping

Base calling was done using Illuminas software pipeline CASAVA and overall sequencing quality was checked by FastQC. Mapping to human genome (hg 19) was done with default parameters with Bowtie v1.1.1 and samples were normalized to the sample with the smallest number of mapped reads (sequencing depth-normalization). Density files (.bedgraph) were generated using the "genomeCoverage" function from BEDTools v2.19.1 and visualized with the Integrated Genome Browser.

Peak calling and annotation

Peak calling for factors generating distinct, sharp peaks (TF3C5 with CD532 treatment, TF3C2, RAD21, CTCF, CapH2, BRF1, RPC32, MYCN) was done with MACS v1.4.2 and peak calling for broad enrichments (histone modifications) was done with SICER. Corresponding input samples for each experiment were used as controls for peak calling. For MACS peak calling, the "–keep-dup" parameter was set to 1 for TF3C5, TF3C2, CapH2, BRF1, RPC32 and MYCN and 15 for RAD21 and CTCF. The p-value cut-off was adjusted for each sample: MYCN/TF3C5/BRF1: 1e-6, CapH2/RPC32: 1e-7, TF3C2: 1e-10, RAD21: 1e-11, CTCF: 1e-

12. Peak calling for histone modifications was done with the following parameters: H3K4me1 (W=400, G=2000, FDR=1e-9), H3K4me3 (W=200, G=600, FDR=1e-8), H3K27ac (W=200, G=600, FDR=1e-3). Enhancers were defined as in (Walz et al., 2014b). Briefly, enhancers are regions with H3K4me1 and overlapping H3K27ac binding sites without H3K4me3 and at least 1kb away from RNAPII transcriptional start sites (TSS). Open and active promoters are defined as regions with H3K4me3 and overlapping H3K27ac binding sites without H3K4me1 and within +/-1kb around RNAPII TSS.

Peaks were assigned to the next RNAPIII (UCSC: hg19_tRNA) and RNAPII (UCSC: hg19_RefSeq) TSS with the "closestBed" function from BEDTools and peaks located in a region +/-1kb around TSSs are defined as being promoter-proximal. Joint binding sites are determined by "intersectBed" from BEDTools with at least 1bp overlap.

Heat maps and read density profiles

Heat maps and density profiles were generated using DeepTools. Briefly, bam files were converted to bigwig files with "bamCoverage" at a resolution of 1bp and read density matrices were calculated with "computeMatrix reference-point" at a resolution of 10bp. Matrices were plotted with "plotMatrix" and mean read density was calculated with "plotProfile" using the standard error of mean as measurement for the variance within each bin.

Motif analysis

Distribution of DNA motifs in peaks was analyzed using the CENTRIMO tool from the MEME Suite with published letter frequency matrices (CTCF: JASPAR database MA0139.1, A-/B-box: (Büchel et al., 2017) or a consensus E-box (CACGTG) frequency matrix. Motif frequencies were calculated in a window of +/-400bp around the peak summit and curves were smoothed by a running mean over 50bp.

Published datasets

The following published ChIP-seq datasets were obtained from the Gene Expression Omnibus (GEO): [N-MYC (GSM2082059), TFIIIC5 with CD532 treatment (GSM2082060), TFIIIC5 with DMSO treatment (GSM2082061) and input (GSM2082057)] and analyzed as described above.

4 Results

4.1 TFIIIC subunits diverge in B-box distribution

Proteomics and interactome of c-MYC and N-MYC revealed multiple subunits of TFIIIC complex as interacting protein partners (Büchel et al., 2017; Kalkat et al., 2018; Koch et al., 2007). TFIIIC is known as a multi-subunit protein complex essential for initiating RNAPIII transcription for tRNA synthesis. Majority of studies of TFIIIC did not clarify if there could be functional discrepancy in two sub-complexes. Genome-wide studies of ETC sites are often based on only one TFIIIC subunit or subunits from one sub-complex.

In this project, one TFIIIC subunit from each sub-complex was chosen for investigation. TFIIIC5 was the first subunit confirmed as N-MYC interacting partner by co-immunoprecipitation experiment, therefore this subunit represents sub-complex A (Büchel et al., 2017). For sub-complex B, TFIIIC2 was selected instead of the largest subunit TFIIIC1 due to the availability of an efficient antibody for ChIP application which the latter lacks.

First, chromatin binding of RNAPIII transcription machinery was investigated by genome-wide chromatin immunoprecipitation coupled with high throughput sequencing (ChIP-seq) experiment.

4.1.1 TFIIIC5 and TFIIIC2 have similar enrichment with RNAPIII transcription machinery

Both TFIIIC5 and TFIIIC2 dispersed throughout the nucleus (Fig. 4.1). ChIP-seq analysis revealed components of RNAPIII transcription machinery bound to 54% to 72% of total 625 human tRNA genes (Fig. 4.2 A). N-MYC bound to tRNA genes to a lesser extent, with 26% of tRNA genes intersected with N-MYC peaks.

All ChIP-seq samples were prepared by untreated *MYCN*-amplified neuroblastoma cells IMR-5, except N-MYC (with DMSO treatment) and TFIIIC5 (with Aurora-A inhibitor CD532 treatment). This inhibitor decreased N-MYC protein level and chromatin binding. Yet, it increased binding strength and number of binding sites of TFIIIC5 across genome. ChIP experiment validated this increase in TFIIIC5 binding intensity was lowered upon doxycycline induction of shRNA against TFIIIC5 at N-MYC/TFIIIC joint binding sites (Fig. 4.2 B). This result confirmed the effect of CD532 was specific to TFIIIC5 binding at N-MYC/ TFIIIC joint binding sites. ChIP-seq sample of TFIIIC5 with CD532 treatment had stronger binding which allowed better peak calling. It was therefore used for all bioinformatic analyses in this thesis.

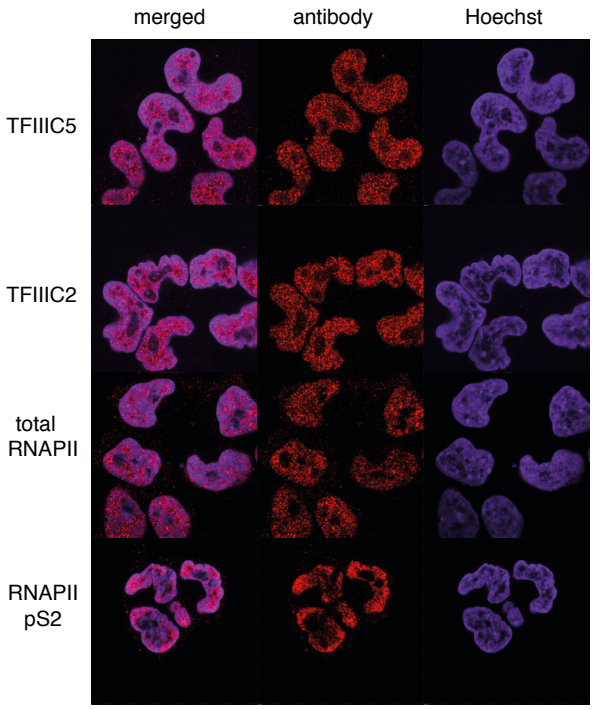
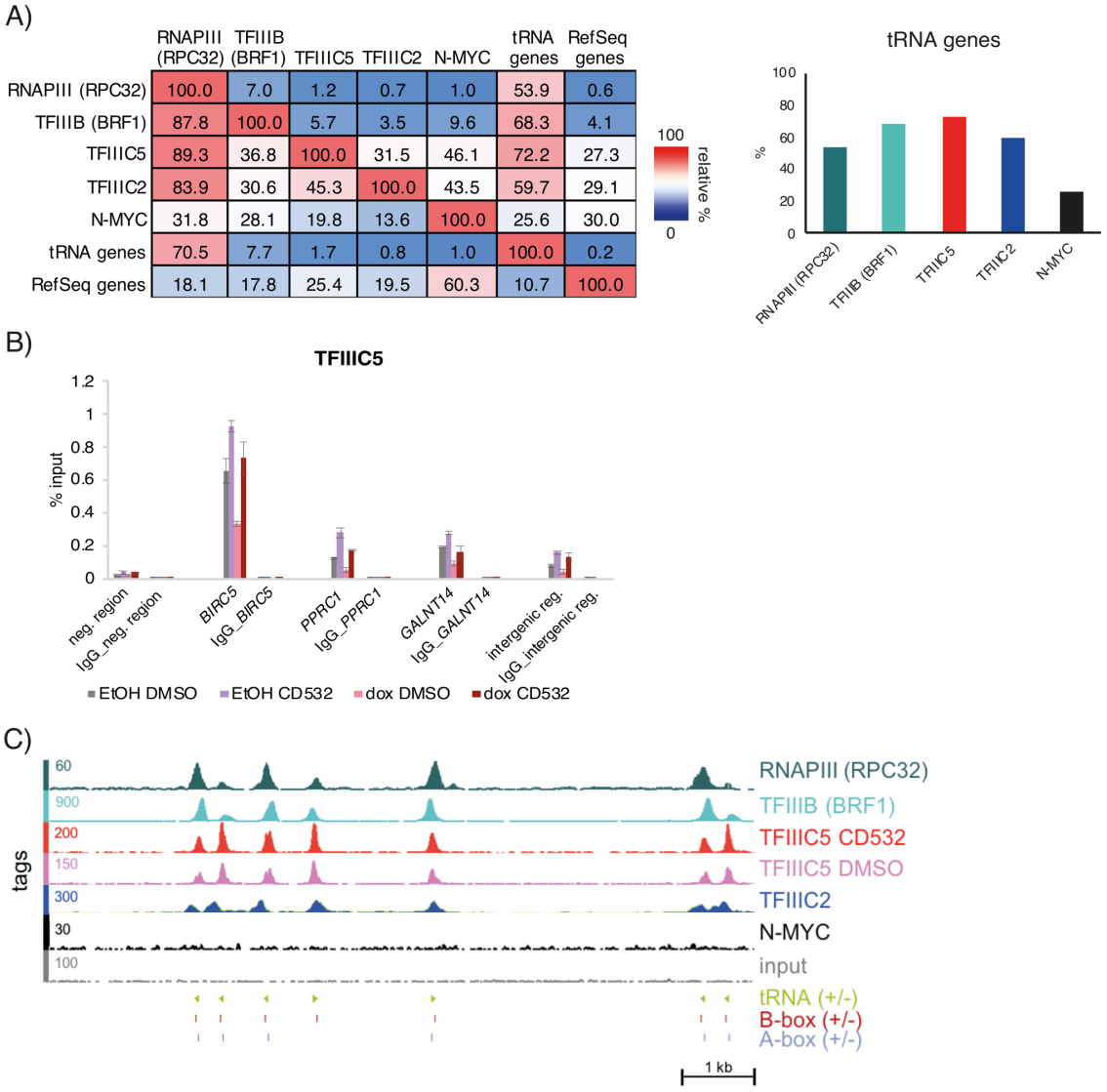


Figure 4.1: Subcellular localization of endogenous TFIIIC5 and TFIIIC2
Immunofluorescence staining of endogenous TFIIIC5 and TFIIIC2 by super resolution dSTORM microscopy. *MYCN*-amplified neuroblastoma cells IMR-5 were seeded on cover slips and the subcellular localizations of TFIIIC5 and TFIIIC2 were detected by specific antibodies. Total RNAPII and serine 2 phosphorylation of RNAPII served as positive control for nucleus localization. Nuclei were stained with Hoechst (Magnification 60x).

Joint occupancies of RPC32 (RNAPIII subunit G, also known as POLR3G), BRF1 (subunit of TFIIIB complex), and TFIIIC subunits were observed at tRNA genes (Fig. 4.2 B). As shown in the representative genome browser region, tRNA genes often occur in clusters (Fig. 4.2 B). Binding intensities of RNAPIII transcription machinery at different tRNA genes varied, with RNAPIII and TFIIIB exhibited higher degree of variation. A- and B-box motifs were aligned close to each other and were localized at tRNA genes. One tRNA shown (the fourth tRNA gene from left) only had B-box.

Global binding of the entire RNAPIII transcription machinery showed slight displacement between each basal component in tRNA binding (Fig. 4.2 C-E). Figure 4.2 C showed the binding of RNAPII transcription machinery to tRNA genes was not exactly aligned to one another. BRF1 peaked slightly upstream of tRNA TSS whereas the others were all downstream (Fig. 4.2 D). TFIIIC2 peak was asymmetric and located further downstream of RNAPIII and spread across a region of about 250 bp. Both TFIIIC5 and TFIIIC2 peaks exactly aligned with RPC32 peaks at tRNA gene promoters (Fig. 4.2 E). But TFIIIC5 presented a sharper peak than TFIIIC2. BRF1 binding was symmetrical except the peak with strongest intensity showed an inclination towards 3’ of RPC peaks (Fig. 4.2 E).

BRF1 displayed a unique enrichment pattern in comparison to other components of RNAPIII transcription machinery (Fig. 4.2 D & E). In addition to the highest peak at TSS of tRNA genes, two BRF1 peaks with lower binding intensity were observed at regions approximately 600 bp upstream, as well as 400 bp and 600 bp downstream to TSS of tRNA genes (Fig. 4.2 D). Since tRNA genes are 80 – 90 bp in length, the BRF1 peaks at further up- and downstream may account for other RNAPIII transcripts at type II promoter, such as SINE elements which are 100 – 700 bp in size. A similar trend can be observed at RPC32 peaks in tRNA gene promoters (Fig. 4.2 E). Also, BRF1 had the strongest binding to tRNA genes among the RNAPIII transcription machinery (Fig. 4.2 D & E).



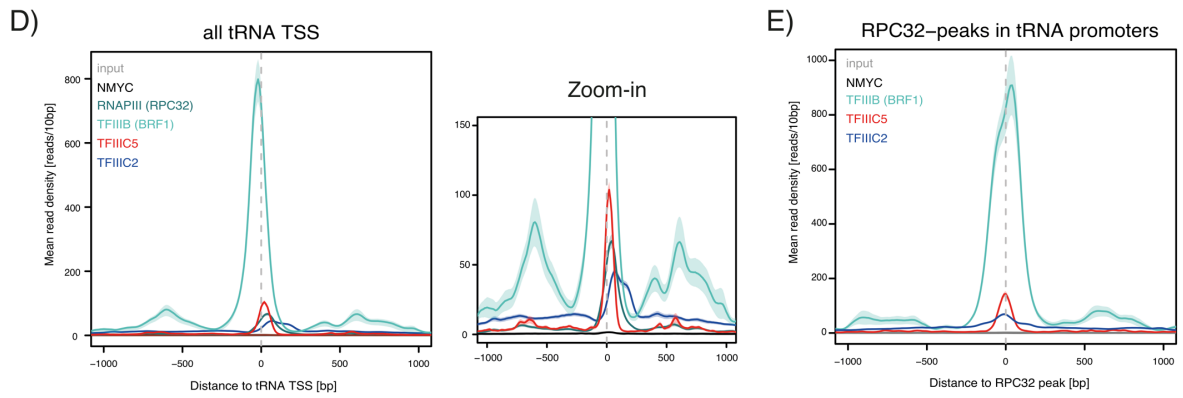


Figure 4.2: RNAPIII transcription machinery and their genome wide binding profiles at tRNA genes

A) Left: Overlap enrichment between RNAPIII transcription machinery, N-MYC, tRNA and RefSeq genes. Numbers reflect percentage of overlapping peaks. Red to blue color scheme represents high (red) or low (blue) percentage.
Right: Graphical illustration depicting percentage binding of RNAPIII transcription machinery and N-MYC to total tRNA genes. Data are from the sixth column of the left table. tRNA and RefSeq genes are defined as ±1 kb within TSS.

B) Chromatin-immunoprecipitation of TFIIIC5 depleted IMR-5 cells. Chromatin of IMR-5 upon 48 hr doxycycline induction for shRNA against TFIIIC5 and 4 hr CD532 treatment before harvest was immunoprecipitated with TFIIIC5 antibody or an unspecific IgG as control. Precipitated and purified DNA was analyzed by qPCR with primers amplifying the TSS of joint NMYC/TFIIIC binding sites or an intergenic control region or a negative control region. Data are represented as mean ± SD.

C) Genomic browser picture displaying ChIP-seq tracks for RNAPIII (subunit RPC32), TFIIIB (subunit BRF1), TFIIIC5 with CD532 or DMSO treatment, TFIIIC2, N-MYC and input sample. Shown is a genomic region in chromosome 1 with several tRNA genes co-bound by the whole RNAPIII transcription machinery.

D) Tag density enrichment profiles for RNAPIII transcription machinery over TSS of all tRNA genes. Right is a zoom-in of the same diagram. Plotted is the mean for each bin and the shadow indicates SEM.

E) Enrichment of TFIIIB and TFIIIC subunits at RNAPIII peak.

Remarks for D) & E): No. of peaks analyzed are as follows: 15403 peaks for N-MYC, 35736 peaks for TFIIIC5, 39923 peaks for TFIIIC2, 4690 peaks for BRF1 and 403 peaks for RPC32. Tag density profiles are in 50 bp resolution.

4.1.2 TFIIIC subunits are distributed around B-box distinctly

Distribution of motifs around RNAPIII machinery showed divergent patterns and frequencies (Fig. 4.3). Both A- and B-boxes were most frequently found at RPC32 peaks. A-box was found at the center of BRF1 peaks but B-box was found at about ± 60 bp away from BRF1 peaks. For TFIIIC5 and TFIIIC2, A-box was distributed around ± 60 – 80 bp away from their peaks. A-box was further found ± 200 bp from TFIIIC5 peaks. However, distributions of B-box around TFIIIC5 and that of TFIIIC2 were totally different. TFIIIC2 showed the same pattern as BRF whereas TFIIIC5 had most of the B-box about 40 bp upstream of its peak.

Both TFIIIC subunits demonstrated unified enrichment at RPC32 peaks in tRNA promoters, suggesting they function together for tRNA synthesis. However, the overlap of total binding sites between them is relatively low (less than 50%) (Fig. 4.2 A). Also, the distinct distribution of B-box around them was an interesting observation. B-box is not only located within tRNA genes but at ETC sites genome-wide. It therefore raised the question of whether TFIIIC5 and TFIIIC2 have unique ETC binding sites which account for any discrepancies in non-canonical functions of TFIIIC.

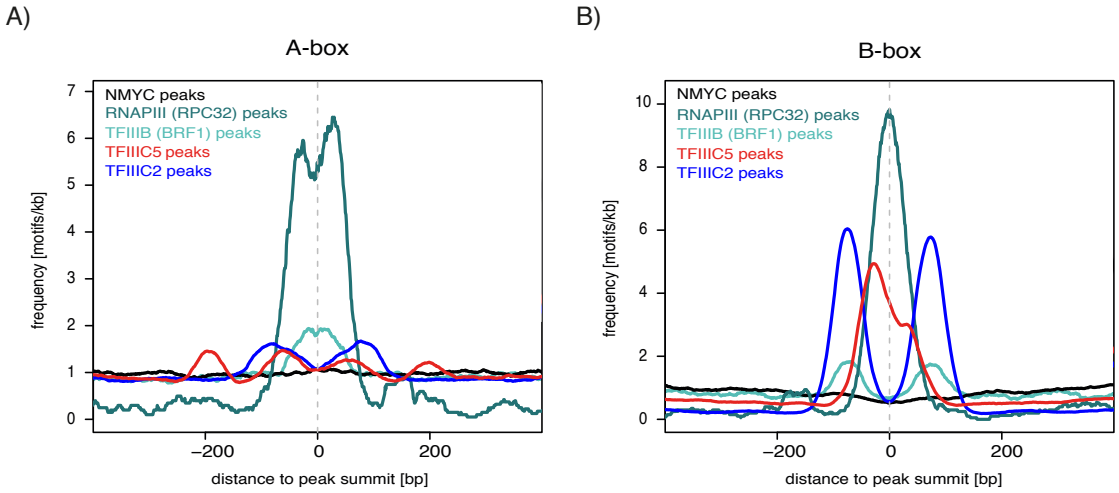


Figure 4.3: Distribution of motifs around RNAPIII transcription machinery
A-box (A) and B-box (B) motifs are distributed 200 bp within peaks of RNAPIII transcription machinery. Curves are smoothened by 50 bp. No. of peaks analyzed are as follows: 15403 peaks for N-MYC, 35736 peaks for TFIIIC5, 39923 peaks for TFIIIC2, 4690 peaks for BRF1 and 403 peaks for RPC32.

4.2 Co-occupancies of N-MYC with TFIIIC subunits shows different genomic distributions

Joint binding sites of TFIIIC5 and N-MYC has previously been reported (Büchel et al., 2017). But whether TFIIIC2 followed the same co-occupancy of TFIIIC5 with N-MYC was not known (Fig. 4.4 A). Thus, TFIIIC5, TFIIIC2 and N-MYC binding sites were compared. About 60% of total N-MYC binding sites overlapped with either or both TFIIIC subunits. Number of joint binding sites between TFIIIC5 and N-MYC was 7418 and that of between TFIIIC2 and N-MYC was 6778. Among which 4791 binding sites were co-occupied by N-MYC and both TFIIIC subunits.

Majority of TFIIIC binding sites were found at intergenic sites whereas about 60% N-MYC binding sites were located at promoters (Fig. 4.4 B). Considering the joint binding sites, different combinations of co-occupancy exhibited different preference on genomic distribution.

N-MYC/TFIIIC5/TFIIIC2 joint binding sites were preferably found at RNAPII promoters (61%) and this preferred distribution was further increased to 77% when TFIIIC5 was excluded (Fig. 4.4 C). Yet, N-MYC/TFIIIC5 joint binding sites exclusive of TFIIIC2 was slightly more frequently found at intergenic sites (57%) than promoters (43%). The sound correlation between N-MYC and TFIIIC2 in genomic distribution was reflected by the shift of predominant intergenic localization of TFIIIC2 to promoters in N-MYC/TFIIIC2 joint binding sites (Fig. 4.4 B & C).

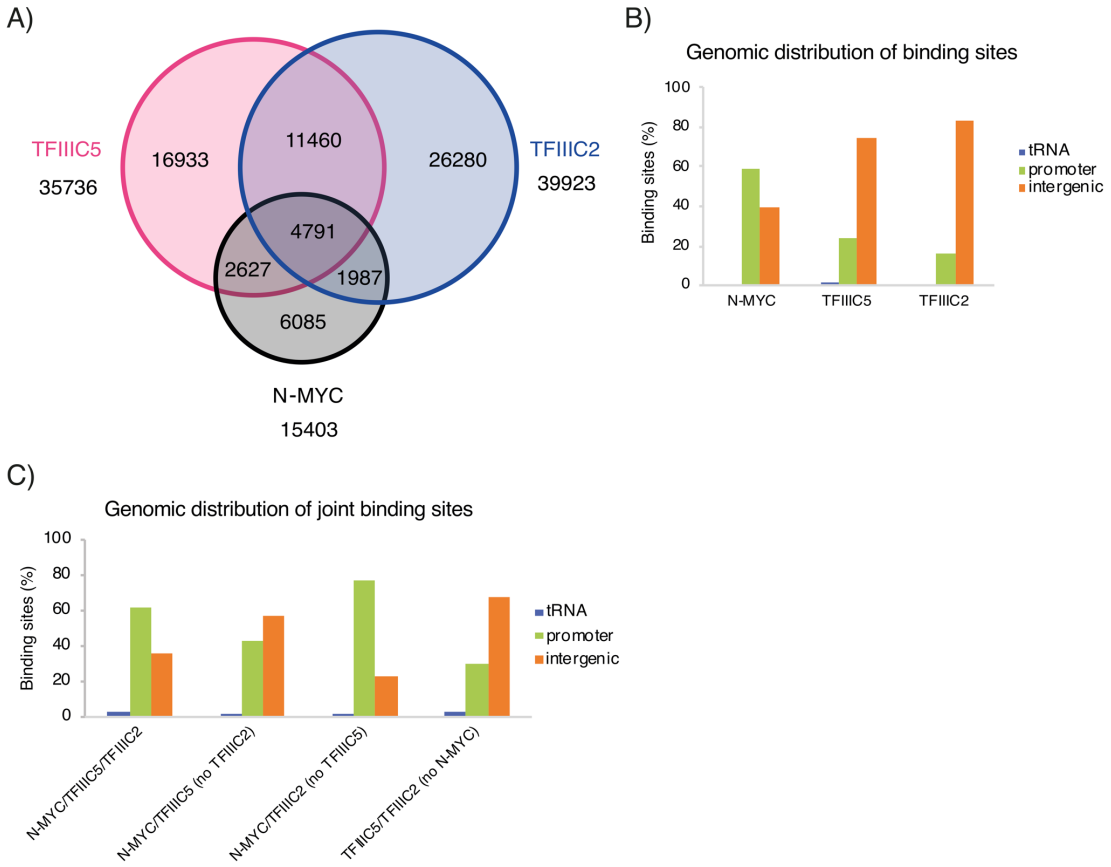


Figure 4.4: TFIIIC5/N-MYC and TFIIIC2/N-MYC joint binding sites

A) Area proportional Venn diagram of overlapping N-MYC, TFIIIC5 and TFIIIC2 total binding sites.
B) Genomic distribution of total binding sites of N-MYC, TFIIIC5 and TFIIIC2.
C) Genomic distribution of joint binding sites of N-MYC, TFIIIC5 and TFIIIC2.

4.3 TFIIIC5 and N-MYC are corelated with other architectural proteins

Architectural proteins CTCF, RAD21 and CAPH2 were selected for comparison with TFIIIC subunits. CTCF and RAD21 are very well-known examples of architectural proteins in mammalian systems. Majority of the knowledge about genomic organization in mammals are based on studies of either one or both of them (Nora et al., 2017; Rao et al., 2017; Tang et al., 2015; Zuin et al., 2014). Accessory protein CAPH2 shares the structural feature with RAD21

for being the kleisin subunit of SMC family protein. It is recently reported to interact with TFIIIC and share similar binding patterns in mESC (Yuen et al., 2017). ChIP-seq of these three proteins was performed and analyzed.

First, overlap of binding sites between TFIIIC subunits and architectural proteins was analyzed (Fig. 4.5 A). As seen in the area proportional Venn diagram, each architectural protein overlapped with one another in varying degree. About 40% of TFIIIC5 binding sites was shared with other architectural proteins, in comparison to 25% of TFIIIC2 binding sites that overlapped with the other three architectural proteins studied. All in all, there were approximately 1300 genomic loci co-occupied by TFIIIC5 and other architectural proteins, which is slightly higher than that of TFIIIC2.
This overlap analysis also revealed a huge proportion (97%) of CTCF overlapped with RAD21. RAD21 was the architectural protein with the highest binding sites (more than the sum of CTCF and CAPH2) and 60% of which was joint binding sites with others. Therefore, it appeared to be the major player for collaboration with other architectural proteins.

Next, global enrichment profiles of architectural proteins at TFIIIC subunits located at RNAPII promoters (Fig. 4.5 B) and that of outside promoters (Fig. 4.5 C) were evaluated. The plots revealed genome-wide binding association between two TFIIIC subunits relative to each other. TFIIIC5 was not enriched at TFIIIC2 peak center but ± 100 bp distant from peak center (Fig. 4.5 B & C left). Similarly, TFIIIC2 was enriched at ± 125 bp distant from TFIIIC2 peak center (Fig. 4.5 B & C right).
As for other architectural proteins, TFIIIC5 exhibited strong enrichment of CTCF and RAD21 (Fig. 4.5 B & C left). Magnitude of CTCF enrichment was similar at both TFIIIC5 peaks at promoters and outside promoters. However, RAD21 binding was stronger at TFIIIC5 peaks outside promoters. CAPH2 was relatively weak but still observable. It had stronger binding in TFIIIC5 peak at promoters.
In sharp contrast, TFIIIC2 had much weaker enrichment with CTCF and RAD21 than TFIIIC5 (Fig. 4.5 B & C right). Also, TFIIIC2 had a pattern of enrichment with CTCF and RAD21 differed from that of TFIIIC5. Instead of one single sharp peak, CTCF and RAD21 enrichment appeared dually ± 125 bp away from TFIIIC2 peak center. Minimal enrichment of CAPH2 was noticed at TFIIIC2 peak at promoters.
This analysis was repeated in ETC sites. ETC sites were identified by TFIIIC binding sites without RPC32 nor BRF1 (Fig. 4.6). TFIIIC2-only ETC showed similar enrichment pattern and

intensity (Fig. 4.7 A & B). However, TFIIIC5/TFIIIC2 common ETC sites followed the association of TFIIIC5 with other architectural proteins (Fig. 4.7 C & D). This strongly argued that only a subset of TFIIIC2 sites which were co-localized with TFIIIC5 could be associated with CTCF and RAD21, whereas other TFIIIC2 binding sites had no relationship with architectural proteins.

Taken together, the analysis of TFIIIC subunits and architectural proteins demonstrated that TFIIIC5 and TFIIIC2 were not enriched with architectural proteins to the same extent. TFIIIC5 had a higher percentage of binding sites co-occupied with other architectural proteins and the enrichment strength was multi-fold higher than that of TFIIIC2.

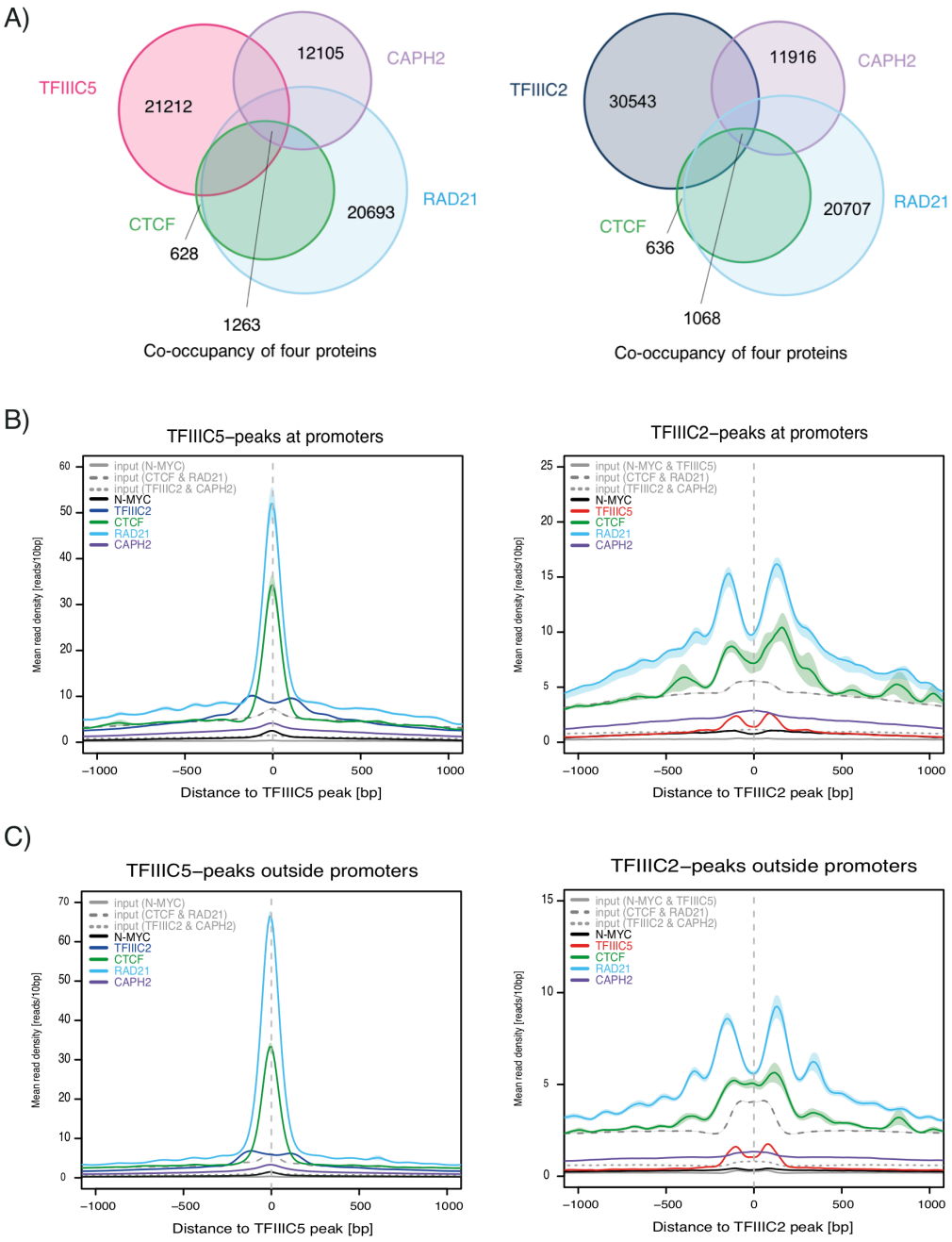


Figure 4.5: Global analysis of TFIIIC subunits with architectural proteins

A) Area proportional Venn diagram of overlapping TFIIIC subunits, CTCF, RAD21 and CAPH2 total binding sites (TFIIIC5: left; TFIIIC2; right). Numbers shown within the circles are the non-overlapped peaks of the corresponding protein.

B - C) Tag density enrichment profiles for architectural proteins over peak of TFIIIC subunits at RNAPII promoters (B) and outside RNAPII promoters (C). (TFIIIC5: left; TFIIIC2; right).

Curves are smoothened by 50 bp. Plotted is the mean for each bin and the shadow indicates SEM. No. of peaks analyzed are as follows: 35736 peaks for TFIIIC5, 39923 peaks for TFIIIC2, 15403 peaks for N-MYC, 23479 peaks for CTCF, 52109 peaks for RAD21 and 23174 peaks for CAPH2.

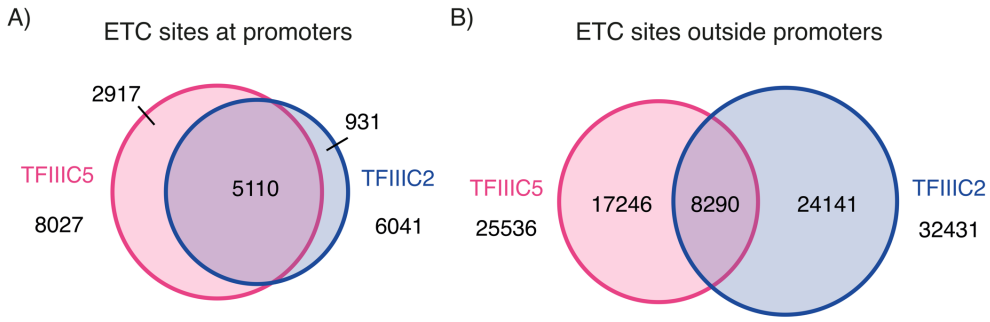


Figure 4.6: ETC sites of TFIIIC5 and TFIIIC2

Area proportional Venn diagram of overlapping TFIIIC5 and TFIIIC2 ETC sites at RNAPII promoters (A) and outside RNAPII promoters (B).

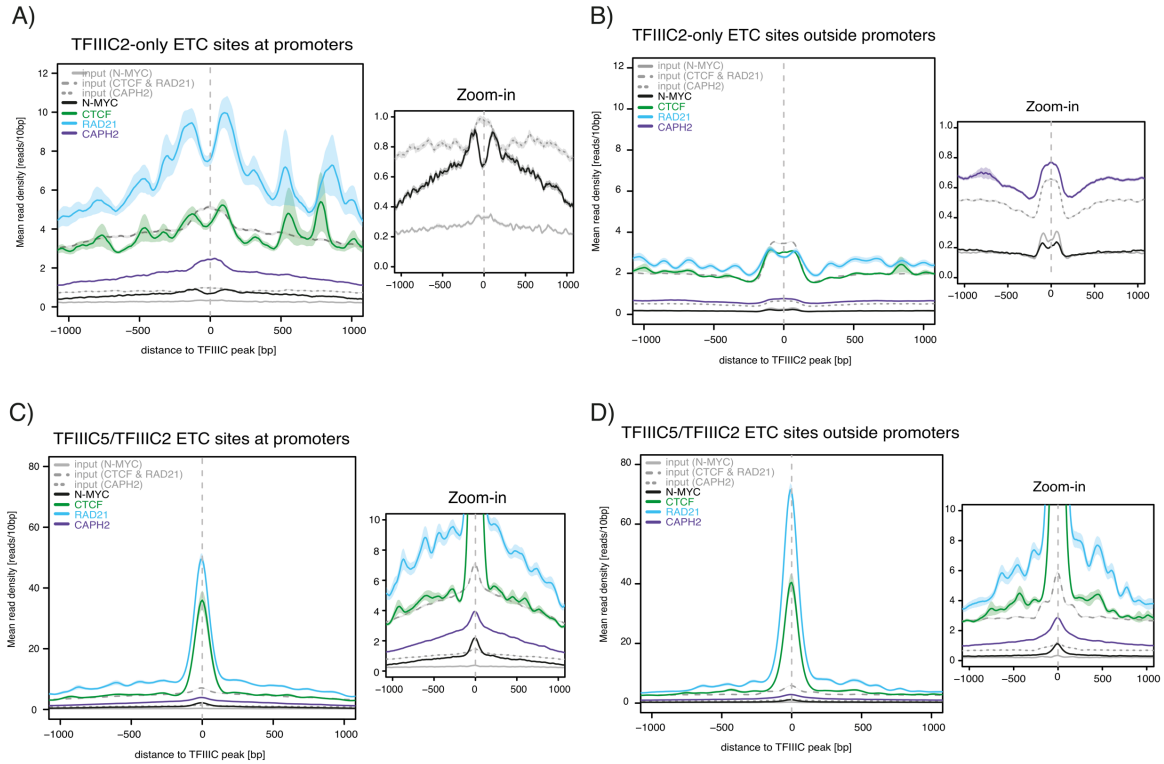


Figure 4.7: ETC sites at promoters

A & B) Tag density enrichment profiles for N-MYC and architectural proteins over peak of TFIIIC2-only ETC sites at RNAPII promoters (A) and outside RNAPII promoters (B).

C & D) Tag density enrichment profiles for N-MYC and architectural proteins over peak of TFIIIC5 at TFIIIC5/TFIIIC2 common ETC sites at RNAPII promoters (C) and outside RNAPII promoters (D).

Remarks: Right is a zoom-in of the corresponding diagram. Plotted is the mean for each bin and the shadow indicates SEM. Curves are smoothened by 50 bp. No. of peaks analyzed are as follows: 6041 peaks for TFIIIC2-only ETC sites at promoters, 32431 peaks for TFIIIC2-only ETC sites outside promoters, 5110 peaks for TFIIIC5/TFIIIC2 ETC sites at promoters, 8290 peaks for TFIIIC5/TFIIIC2 ETC sites outside promoters, 15403 peaks for N-MYC, 23479 peaks for CTCF, 52109 peaks for RAD21 and 23174 peaks for CAPH2.

Furthermore, enrichment of architectural proteins at N-MYC was analyzed. Heat maps demonstrated the tag distributions across sites with different strengths of N-MYC occupancies (Fig. 4.8 A). In other words, this analysis described binding of architectural proteins at relatively highly occupied N-MYC sites (upper part of the heat map) and that of relatively weakly occupied N-MYC sites (lower part of the heat map). The binding strength of architectural proteins at N-MYC was reflected by the density plots (Fig. 4.8 B).

Consistent with the previous analysis on genomic distribution of joint binding sites (Fig. 4.4 C), more TFIIIC5 binding at N-MYC was observed outside promoters than that of at promoters (Fig. 4.8 A). Also, TFIIIC5 binding strength was stronger at N-MYC sites outside promoters (Fig. 4.8 B). In contrast, TFIIIC2 showed stronger enrichment to N-MYC peaks at promoters irrespective of N-MYC occupancy. Yet, the binding strength was weak as shown in the density plot.

For CTCF and RAD21 ChIP-seq, the input was slightly contaminated as shown by the weak enrichment on the heat map (Fig. 4.8 A). In comparison to enrichment outside promoters, CTCF enriched to N-MYC peaks at promoters was not very prominent (Fig. 4.8 A). Despite the small number of binding sites, CTCF bound at N-MYC peaks at promoters with high intensity (Fig. 4.8 B).

Enrichment of RAD21 to N-MYC was in a similar vein to that of CTCF, but with a much stronger degree of enrichment and binding strength. Both RAD21 and CTCF were more enriched outside promoters than at promoters and they were preferentially found at sites with lower N-MYC occupancy (Fig. 4.8 A). Tags were more concentrated to the middle at the lower part of the heat map whereas tags were more dispersed throughout the 2 kb window at the upper part of the heat map. Binding strength of RAD21 was the highest among all the architectural proteins studied here and it was stronger at N-MYC peaks outside promoters than that of at promoters (Fig. 4.8 B).

CAPH2 was more prominently enriched at N-MYC peaks outside promoters than at promoters (Fig. 4.8 A). At N-MYC peaks outside promoters, CAPH2 demonstrated a tendency of enrichment at sites with higher N-MYC occupancy (Fig. 4.8 A). Similar to RAD21 and CTCF, CAPH2 bound more strongly at N-MYC peaks outside promoters than that of at promoters (Fig. 4.8 B).

In short, enrichment of architectural proteins except TFIIIC2 was stronger at N-MYC peaks outside promoters. This was supported by both binding sites (Fig. 4.8 A) and binding strength (Fig. 4.8 B).

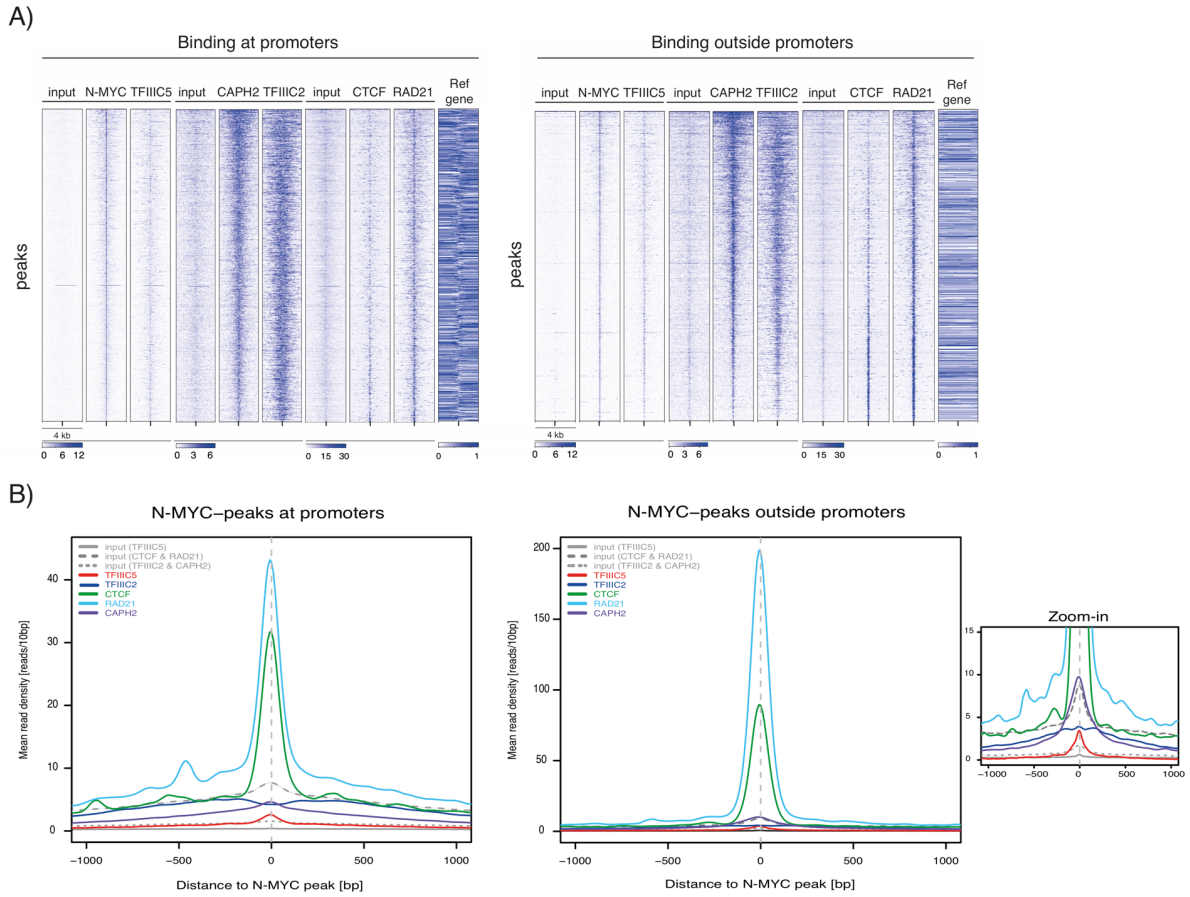


Figure 4.8: Global analysis of N-MYC with architectural proteins

A) Heat map exhibiting the distribution of N-MYC at architectural protein binding sites outside (right) and at RNAPII promoters of all annotated human RefSeq TSS (left) in a window of ± 2 kb. Tag densities are sorted by H3K27ac (50 bp resolution).

B) Tag density enrichment profiles for architectural proteins over peak of N-MYC at RNAPII promoters and outside RNAPII promoters. Right is a zoom-in of the corresponding diagram. Plotted is the mean for each bin and the shadow indicates SEM. No. of peaks analyzed are same as in Fig. 4.5.

TAD border was enriched with TSS of coding genes (Dixon et al., 2012). Therefore, global binding of architectural proteins to all RefSeq genes was investigated. In this density plot, all samples were normalized to same sequencing depth and the mean of the all binding sites was taken for analysis. TFIIIC5 showed two peaks – one downstream of TSS and was associated with N-MYC and RAD21 albeit to a lesser extent (marked by orange single arrow); the other one upstream of TSS and was associated with RAD21 and CTCF (marked by orange double arrow) (Fig. 4.9). On the other hand, TFIIIC2 and CAPH2 did not showcase particular associations with the others but could be associated with each other downstream of TSS (marked by yellow solid arrow). Yet, it was a relatively broad peak, especially in the case of CAPH2. Nonetheless, each of them presented a non-overlapped 200 bp-sized broad peak upstream of TSS.

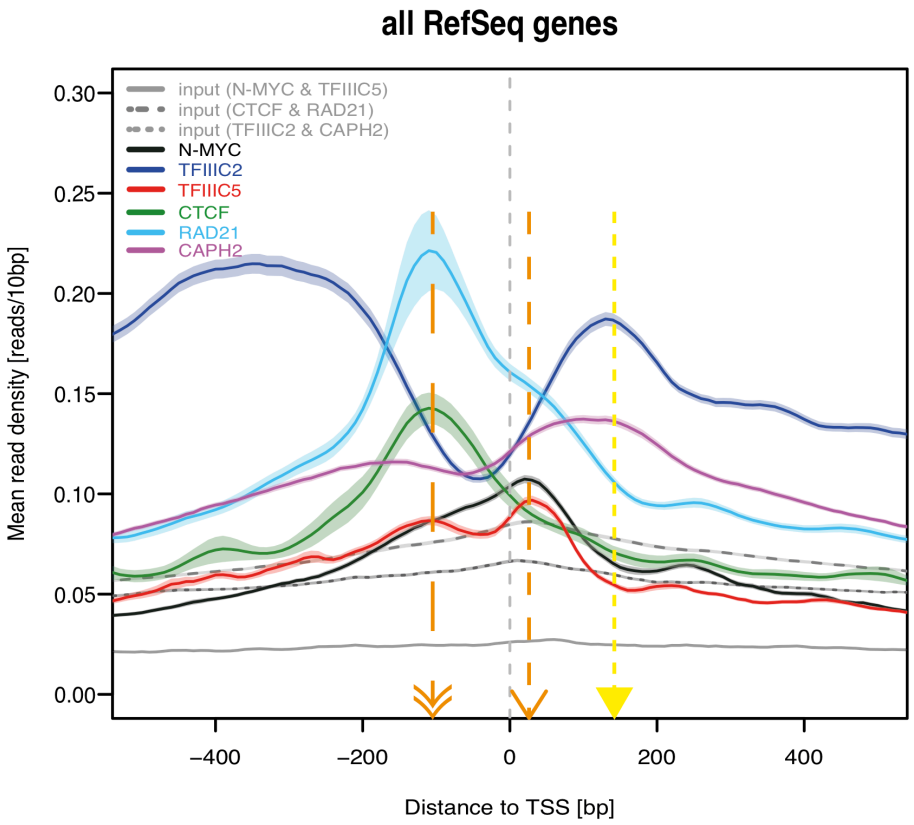


Figure 4.9: Global analysis of architectural proteins and N-MYC of all RefSeq genes
Tag density enrichment profiles for architectural proteins and N-MYC over TSS of all RefSeq genes. Samples are all normalized to the same read and plotted at a resolution of 10 bp. Two orange lines indicate two peaks of TFIIIC5 overlapping with other samples. Two TFIIIC5 peaks are marked as single orange arrow (downstream of TSS) and double orange arrow (upstream of TSS). A sharp TFIIIC2 peak was downstream of TSS (marked by yellow solid arrow). Plotted is the mean for each bin and the shadow indicates SEM. No. of peaks analyzed is same as in Fig. 4.5 but all normalized to the same sequencing depth.

Proteins investigated here recognized their own specific binding sequence, except accessory protein RAD21 and CAPH2. With the aforementioned results that revealed enrichment of the proteins-of-interest, distribution of motif was studied to investigate if the consensus binding sequence governed binding patterns. Three motifs, B-box for TFIIIC subunits, E-box for N-MYC and CTCF motif for CTCF were investigated (Fig. 4.10). Distribution of B-box was only found at or around TFIIIC subunits but not other architectural proteins nor N-MYC (Fig. 4.10 A & Fig. 4.3 B). E-box was distributed at the center of N-MYC, both of TFIIIC subunits, CTCF and RAD21 peaks (Fig. 4.10 B). Moreover, E-box could be found at some distance away from peaks of TFIIIC subunits summit (about ± 150 bp for TFIIIC5 and ± 100 bp for TFIIIC2). For CTCF motif, it is located at peak center of CTCF and RAD21 with high frequency as expected (Fig. 4.10 C). Also, CTCF motif was found at N-MYC, TFIIIC5 and CAPH2 peaks. TFIIIC2, however, was peaked ± 75 bp away from CTCF motif. This unique distribution pattern of motifs in which the motif was found symmetrically upstream and downstream from TFIIIC2 peak

center was observed for all motifs studied, including A-box which was shown in earlier section (Fig. 4.10 A – C & Fig. 4.3 A).

Genome-wide binding association with other architectural proteins further elucidated the divergence of both TFIIIC subunits. TFIIIC5 binding sites were enriched with other architectural proteins, some of which were also associated with N-MYC whereas TFIIIC2 displayed low enrichment. In addition, TFIIIC2 presented unique motif distribution which strongly distinguished itself from TFIIIC5.

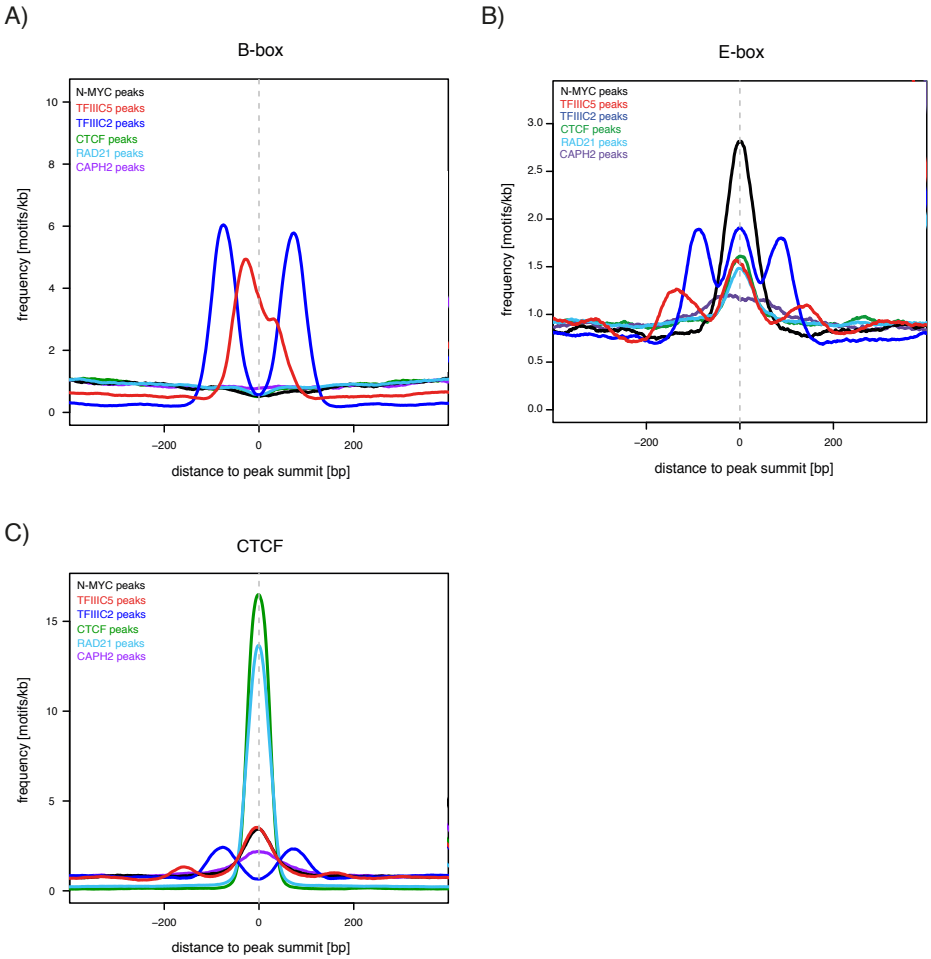


Figure 4.10: Distribution of motifs around TFIIIC subunits, N-MYC and other architectural proteins

B-box, E-box and CTCF motifs are distributed 200 bp within peaks of TFIIIC subunits, N-MYC and other architectural proteins. Curves are smoothened by 50 bp. No. of peaks analyzed are same as in Fig. 4.5.

4.4 TFIIIC2 is associated with SINE elements

The weak enrichment of TFIIIC2 with other architectural proteins posed the question of the binding preference of TFIIIC2 across the genome. The association between TFIIIC and SINE elements was under the spotlight owing to the A- and B-boxes possessed by SINE elements. Enrichment profile exhibited completely different patterns of TFIIIC5 and TFIIIC2 (Fig. 4.11

A). TFIIIC5 showed one peak 50 bp downstream of SINE TSS. Another TFIIIC5 peak at 200 bp downstream of SINE TSS did not show enrichment over input. TFIIIC2 enrichment was stronger than TFIIIC5 and presented as one single broad summit peaked at about 150 bp downstream of SINE TSS. TFIIIC2 peak spanned across an area covering both TFIIIC5 peaks. RPC32 and TFIIIB did not show enrichment, indicating TFIIIC-bound SINEs are ETC sites.

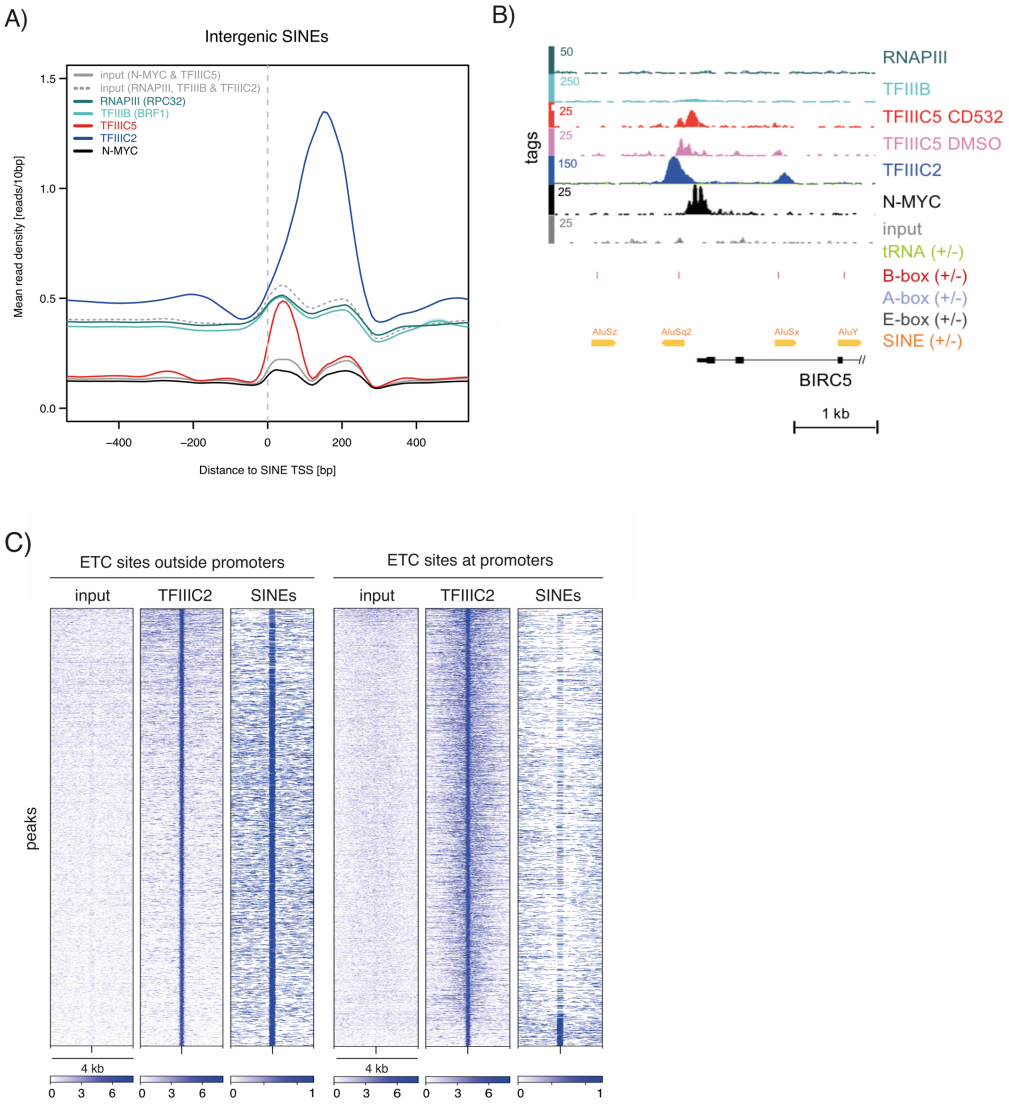


Figure 4.11: Enrichment of SINEs

A) Tag density enrichment profiles for RNAPIII transcription machinery and N-MYC over TSS of all intergenic SINEs (50 bp resolution). Plotted is the mean for each bin and the shadow indicates SEM. No. of peaks analyzed are same as Fig. 4.2.

B) Genomic browser picture displaying ChIP-seq tracks for RNAPIII (subunit RPC32), TFIIIB (subunit BRF1), TFIIIC5 with CD532 or DMSO treatment, TFIIIC2, N-MYC and input sample. Shown is a genomic region in chromosome 17 with several SINEs/*Alu* elements, with one close to TSS of RNAPII gene and one within the gene body.

C) Heat map exhibiting the distribution of TFIIIC at TFIIIC2-only ETC sites outside (left) and at RNAPII promoters of all annotated human RefSeq TSS (right) in a window of ± 2 kb. Tag densities are in 50 bp resolution. 24141 and 931 peaks of TFIIIC2-only ETC outside and at RNAPII promoters were analyzed.

Fig. 4.11 B was a representative example depicting the differential binding pattern of two TFIIIC subunits to SINEs. TFIIIC2 completely covered *AluSq2* while TFIIIC5 was slightly displaced towards the downstream of TFIIIC2, which was then closer to the *AluSq2* TSS. *AluSx* inside the gene body of *BIRC5* exhibited similar binding patterns of TFIIIC5 and TFIIIC2. Strong enrichment of TFIIIC2 ETC sites with SINEs was reflected in heatmap (Fig. 4.11 C). All TFIIIC2-only ETC sites outside RNAPII promoters had binding association with SINEs. However, this association was limited to a small fraction of TFIIIC2-only ETC sites at RNAPII promoters.

4.5 TFIIIC5 recruits RAD21 to chromatin at joint N-MYC/TFIIIC loci

4.5.1 Establishment of a protein depletion cell system

With the global analysis, next step was to study the mechanisms of how TFIIIC subunits interact with N-MYC and how other architectural proteins were involved. A stable cell line system with TFIIIC protein depletion was necessary. However, CRISPR/Cas9 knockout of TFIIIC5 subunit using 12 single guide RNAs (sgRNAs) from (Hart et al., 2015) in *MYCN*-amplified cells IMR-5 resulted in cell death (Fig. 4.12 A). The knockout approach was modified to use sgRNA specifically targeting endogenous TFIIIC while rescuing cells with exogenous TFIIIC. Exogenous TFIIIC5 tagged with HA at the C-terminal could be distinguished from endogenous TFIIIC5 on immunoblot based on size difference (Fig. 4.12 B). Total TFIIIC5 protein level did not increase upon doxycycline induction of the exogenous expression system, in comparison to parental IMR-5 cells and the infected cells without doxycycline treatment (Fig. 4.12 B). This system was subjected to CRISPR/Cas9 knockout using two sgRNAs (#1 and #4) targeting 5'UTR of endogenous TFIIIC5. Single clones were isolated from infected cell pool and the two with strongest depletion on immunoblot were shown (Fig. 4.12 C left). Despite more than 85% depletion of TFIIIC5 was obtained, homozygous knock-out was not achieved (Fig. 4.12 C right). Wild-type and genetic editing events such as deletion, insertion or missense mutation were confirmed by Sanger sequencing. More than one type of mutations identified in one cell clone implied contamination with other cells. It was also noted that sgRNA #1 did not work because only wild-type sequence was found in sequence confirmation of cell clones (data not shown).

Due to the complexity of this "knockout with rescue" system, RNA interference using shRNA was employed. Stable cell lines infected with doxycycline-inducible shRNA against TFIIIC subunits based on (Fellmann et al., 2013) were established. Upon 48 hours doxycycline

induction, more than 85% depletion of both subunits was obtained (Fig. 4.13 A). No significant difference in cell proliferation and apoptosis was found when TFIIIC subunit was depleted for two days (Fig. 4.13 B).

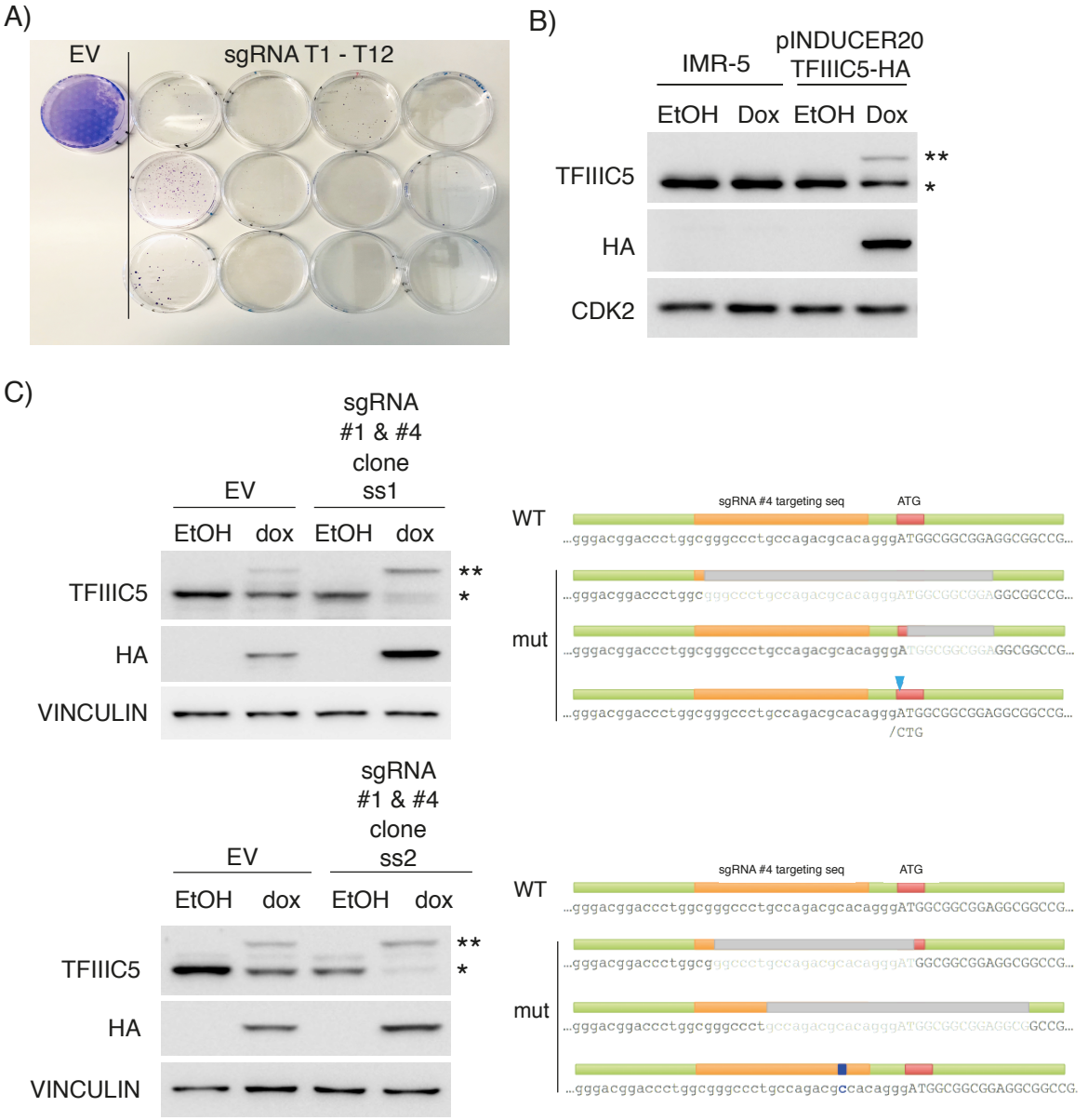


Figure 4.12: TFIIIC5 is essential to cells

A) Crystal violet staining of IMR-5 cells infected with sgRNAs targeting *TFIIIC5* coding sequence in lentiCRISPRv2 backbone and empty vector control.

B) Stable cell line of IMR-5 with doxycycline-inducible expression of HA-tagged exogenous TFIIIC5. Immunoblot of TFIIIC expression upon 48 hr doxycycline induction.

C) CRISPR/Cas9 knock-out of endogenous TFIIIC5 in stable cells expressing HA-tagged exogenous TFIIIC5. Best two cell clones ss1 (upper panel) and ss2 (lower panel) are shown. Left is the immunoblot of TFIIIC5 and HA expression of single clone.
Right is the graphical illustration of Sanger sequencing results. Targeting sequence of sgRNA #4 (orange box) is upstream of start codon of *TFIIIC5* (red box). Mutations such as deletion (grey box), single nucleotide mutation (light blue triangle) and insertion (blue box) were shown.
Remarks: ** refers to HA-tagged exogenous TFIIIC5 and * refers to endogenous TFIIIC5.

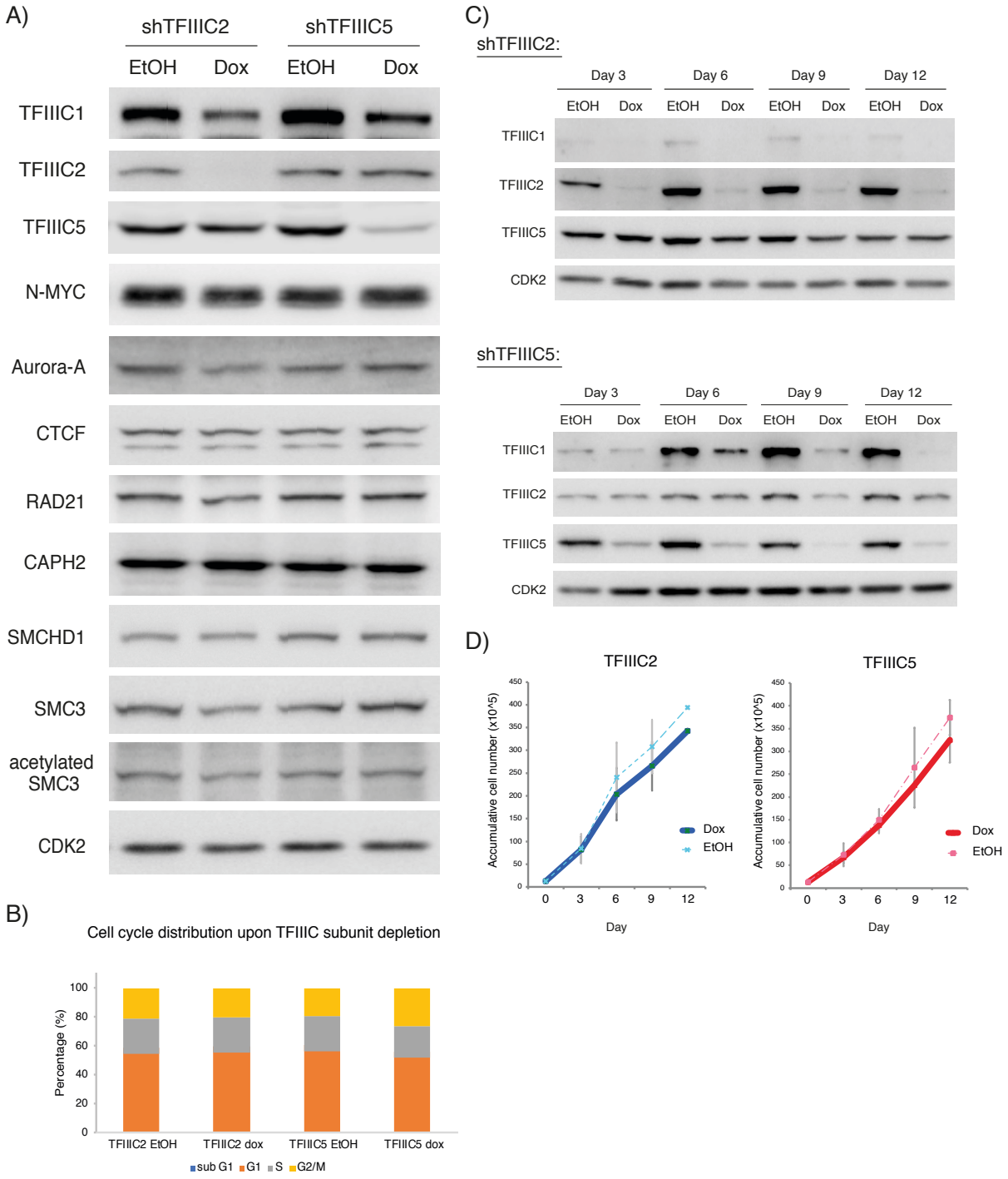


Figure 4.13: Stable cell system of doxycycline inducible shRNA against TFIIIC2 and TFIIIC5

A) Western blot analysis of proteins upon 48 hr doxycycline treatment of stable cell lines with shRNA against TFIIIC2 and TFIIIC5.

B) BrdU-PI FACS upon 48 hr doxycycline treatment of stable cell lines with shRNA against TFIIIC2 and TFIIIC5.

C) Western blot analysis of TFIIIC subunits for knock-down for consecutive 12 days of doxycycline induction.

D) Growth curve of stable cell lines with shRNA against TFIIIC5 and TFIIIC2 for consecutive 12 days of doxycycline induction.

4.5.2 Effects of one TFIIIC subunit to another

Concerning the effect of one TFIIIC subunit on the others and other proteins-of-interest, depletion of TFIIIC5 had no effect on TFIIIC2 protein steady level, but had mild effect on TFIIIC1 which is also part of sub-complex B as TFIIIC2 (Fig. 4.13 A). On the other hand, TFIIIC1 but not TFIIIC5 responded upon depletion of TFIIIC2. About 50% reduction of TFIIIC1 protein level was observed. However, N-MYC, Aurora-A, other architectural proteins and members of SMC family which are non-architectural proteins did not show any changes in response to doxycycline induction of shRNA against TFIIIC5 or TFIIIC2.

Effectiveness of TFIIIC subunit knockdown in stable cell lines was confirmed by induction of doxycycline for up to 12 days (Fig. 4.13 C). Western blot revealed the continued protein depletion without rebound. The effect of depletion to other TFIIIC subunits remained unchanged except TFIIIC2. Long-term depletion of TFIIIC5 led to reduction of TFIIIC2 starting from day 9. Meanwhile, cell growth was not impaired when TFIIIC subunits were depleted in a time course of 12 days (Fig. 4.13 D).

Direct protein-protein interaction between TFIIIC subunits and N-MYC was evaluated by Co-IP experiment (Fig. 4. 14). TFIIIC is a complex and its subunits could be immunoprecipitated with each other. However, only TFIIIC5, but not TFIIIC2 could pull down N-MYC, even though the amounts of both subunits immunoprecipitated by TFIIIC2 were comparable.

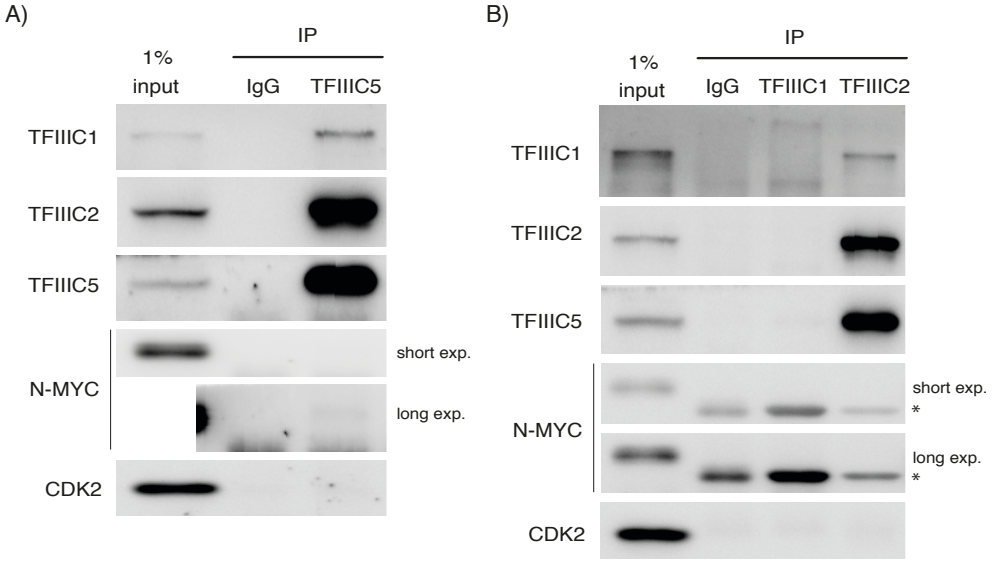


Figure 4.14: TFIIIC5 and TFIIIC2 can form a complex but only TFIIIC5 can immunoprecipitate N-MYC

Endogenous co-immunoprecipitation of TFIIIC5 (A) and TFIIIC2 (B). IMR-5 whole cell lysate was prepared and TFIIIC5 and TFIIIC2 were immunoprecipitated or an immunoprecipitation with unspecific rabbit or mouse anti-IgG was performed. Precipitated proteins were detected by Western blotting with the indicated antibodies. TFIIIC1 immunoprecipitation in (B) did not work.

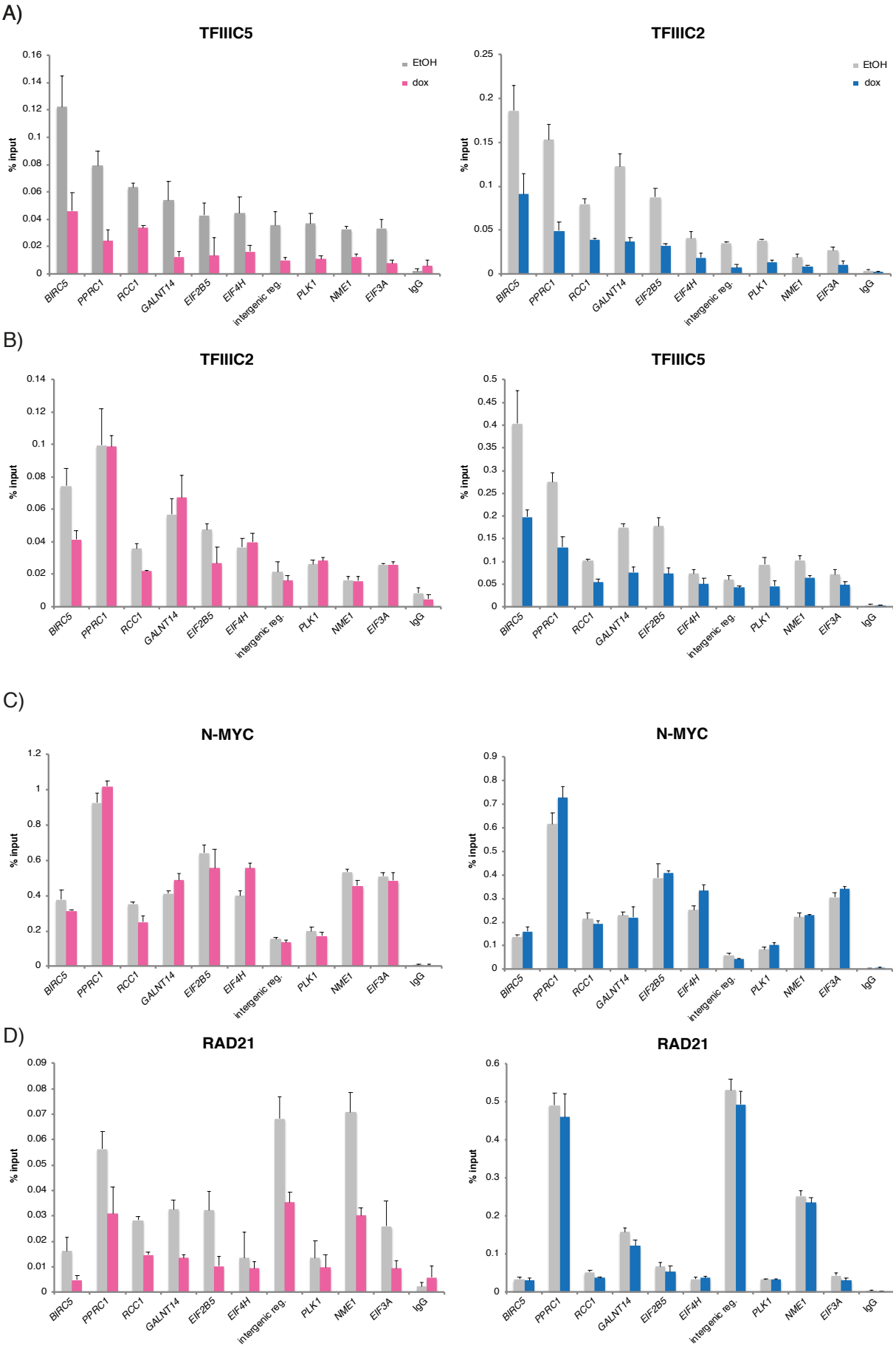
Remarks: * marks the band for heavy chain of the N-MYC antibody.

To follow, influence of TFIIIC subunit to chromatin binding at joint N-MYC/TFIIIC sites was explored by ChIP experiment (Fig. 4.15 A & B). Chromatin was prepared from cells treated with doxycycline for 48 hour to achieve satisfactory protein knock-down as shown in Figure 4.13 A. Both TFIIIC5- and TFIIIC2-depleted cells showed more than 50% reduction of their chromatin bindings at almost all target sites (Fig. 4.15 A, red and blue panel for shRNA against TFIIIC5 and TFIIIC2 respectively). In a parallel experiment, chromatin with depletion of one TFIIIC subunit was immunoprecipitated with another subunit (Fig. 4.15 B). Although knockdown of TFIIIC5 and TFIIIC2 did not influence protein levels of the other subunit, TFIIIC2 depletion could bring TFIIIC5 chromatin binding down to half at majority of sites. However, decrease in TFIIIC5 protein level did not affect TFIIIC2 bindings in general, except three out of ten target sites checked exhibited not more than 45% reduction. In addition, N-MYC binding was not influenced by depletion of either subunit (Fig. 4.15 C).

4.5.3 TFIIIC subunits and RAD21

RAD21 was previously shown to be N-MYC interacting partner (Büchel et al., 2017). Section 4.3 has shown the genome-wide enrichment of RAD21 at N-MYC and TFIIIC subunits binding sites though in varying degree. In order to understand the mechanism of how TFIIIC interacts with RAD21 on chromatin binding, ChIP experiment was performed with cells induced with doxycycline for the shRNA against TFIIIC subunits as aforementioned. RAD21 binding decreased more than 50% at most of genomic loci in response to TFIIIC5 but not TFIIIC2 depletion (Fig. 4.15 D).

Based on the ChIP results, it may be possible that TFIIIC first coupled with RAD21 and then loaded this complex to chromatin at N-MYC/TFIIIC sites. Therefore, the reduction in TFIIIC5 protein but not its chromatin binding resulted in decrease in RAD21 binding to chromatin.



(Legend in the next page)

Figure 4.15: TFIIIC5 recruits RAD21 to chromatin

Chromatin-immunoprecipitation of TFIIIC5- (red) and TFIIIC2- (blue) depleted IMR-5 cells. Chromatin of IMR-5 upon 48 hr doxycycline induction for TFIIIC5 or TFIIIC2 shRNA was immunoprecipitated respectively with antibodies or an unspecific IgG as control. Precipitated and purified DNA was analyzed by qPCR with primers amplifying the TSS of joint NMYC/TFIIIC binding sites or an intergenic control region or a negative control region. Data are represented as mean ± SD of technical triplicates. Shown is the representative result from three biological replicates.

4.5.4 TFIIIC subunits and CAPH2

Next, another accessory protein which also belongs to SMC family, CAPH2 was investigated. Although multiple members of SMC family were found at N-MYC mass spectrometry, CAPH2 was not identified as a hit. Recent publication from (Yuen et al., 2017a) and genome-wide binding association from section 4.3 have shown enrichment of CAPH2 at TFIIIC5 and TFIIIC2 sites, though the ChIP-seq results for this thesis did not demonstrate as strong enrichment as what the publication stated.

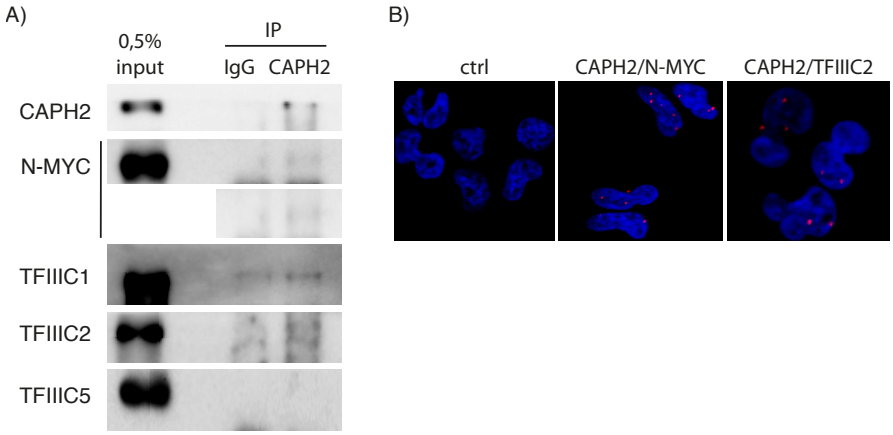


Figure 4.16: Protein-protein interaction between CAPH2, N-MYC and TFIIIC

A) Endogenous co-immunoprecipitation of CAPH2. IMR-5 nuclear lysate was prepared and CAPH2 was immunoprecipitated or an immunoprecipitation with unspecific rabbit anti-IgG was performed. Precipitated proteins were detected by Western blotting with the indicated antibodies.

B) Proximity ligation assay of endogenous CAPH2/N-MYC and CAPH2/TFIIIC2 interactions in IMR-5 cells. Red dots show PLA signals resulting from interactions. Nuclei were stained with Hoechst (Magnification 60x).

Protein-protein interaction between CAPH2, N-MYC and TFIIIC was first investigated with Co-IP using nuclear lysate of IMR-5 (Fig. 4.16 A). Immunoprecipitation of CAPH2 was weak even after optimization of experimental conditions, as shown by the amount of CAPH2 protein immunoprecipitated. Despite the technical hurdle, N-MYC could be pulled down. For TFIIIC subunits, TFIIIC1 showed a band at correct size for both IgG and CAPH2 IP samples. The contamination in IgG sample failed the control. TFIIIC5 did not have detectable

immunoprecipitation. Nonetheless, TFIIIC2 showed weak interaction. *In situ* proximity ligation assay (PLA) detects endogenous proteins that are located within 40 nm. This may not indicate direct-direct protein interaction but the close proximity between proteins-of-interest. Signals were detected in CAPH2/N-MYC and CAPH2/TFIIIC2 (Fig. 4.16 B). This assay requires two antibodies raised from different species. This limited the PLA detection between CAPH2 and TFIIIC5 and that of TFIIIC1 since all working antibodies were raised in rabbit.

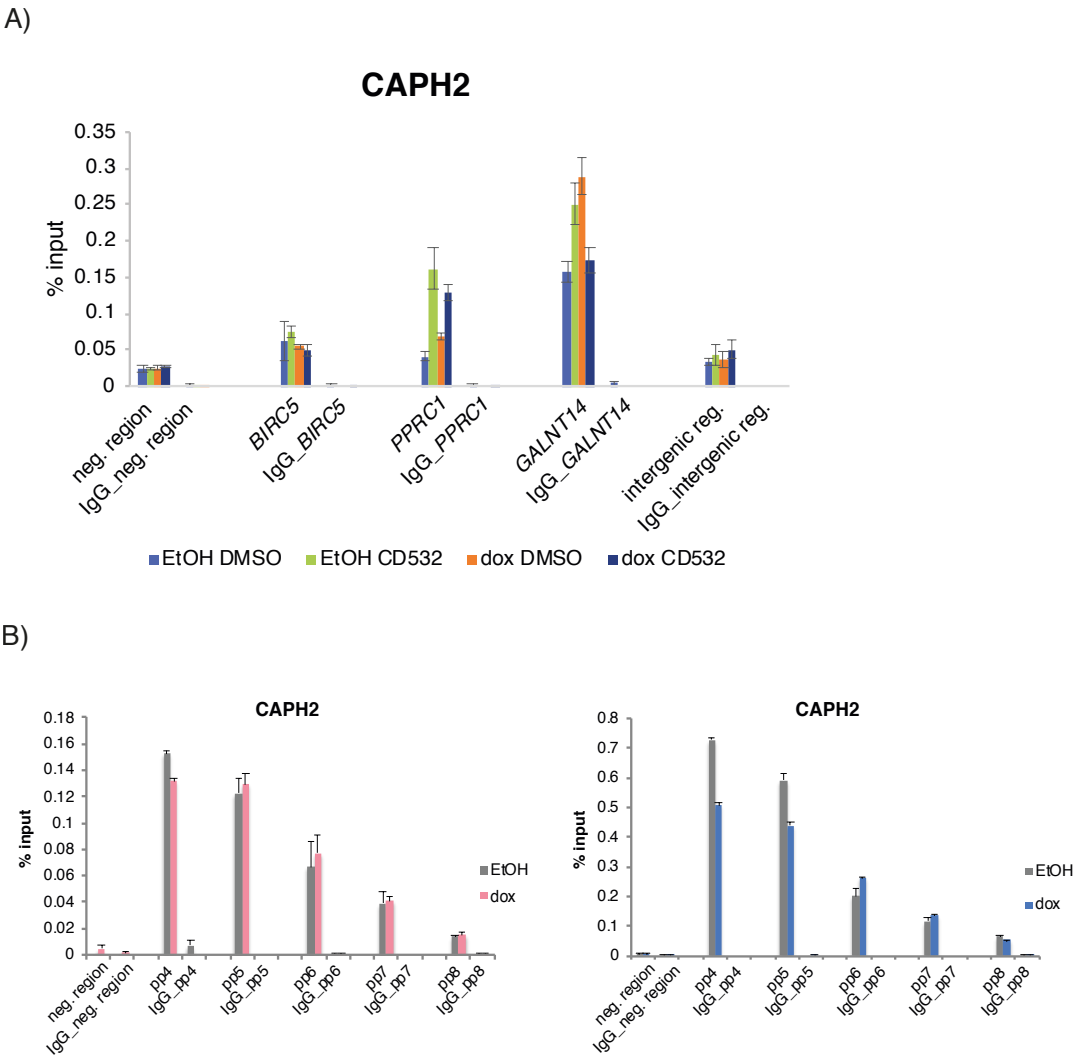


Figure 4.17: TFIIIC unloads CAPH2 to chromatin upon CD532 treatment

A) Chromatin-immunoprecipitation of TFIIIC5 depleted IMR-5 cells upon CD532 treatment. Chromatin of IMR-5 upon 48 hr doxycycline induction for TFIIIC5 and 4 hr CD532 treatment before harvest was immunoprecipitated with CAPH2 antibody or an unspecific IgG as control. Precipitated and purified DNA was analyzed by qPCR with primers amplifying the TSS of joint N-MYC/TFIIIC binding sites or a negative control region.

B) Chromatin-immunoprecipitation of TFIIIC5- (red) and TFIIIC2- (blue) depleted IMR-5 cells. Chromatin of IMR-5 upon 48 hr doxycycline induction for TFIIIC5 or TFIIIC2 shRNA was immunoprecipitated with CAPH2 antibody or an unspecific IgG as control. Precipitated and purified DNA was analyzed by qPCR with primers amplifying the non-TSS sites with CAPH2-only (pp4 & pp5) and that of with N-MYC/TFIIIC/CAPH2 joint binding sites (pp6, pp7 & pp8) or a negative control region.

Remarks: Data are represented as mean ± SD of technical triplicates.

Since Co-IP experiment exhibited interaction between CAPH2 and N-MYC, ChIP experiment was performed upon treatment with Aurora-A inhibitor CD532 which can reduce N-MYC protein level and chromatin binding. CAPH2 binding increased upon CD532 treatment at the target sites used in Fig. 4.15 for evaluation with varying magnitude (Fig. 4.17 A). Under the effect of TFIIIC5 knockdown, CAPH2 showed increase in binding at two sites whereas the other two sites were not affected. However, no conclusion could be made in combined treatment of CD532 and TFIIIC5 knockdown because the binding at all four sites varied.
In addition, other genomic loci were checked upon knockdown of TFIIIC5 (Fig. 4.17 B left) and TFIIIC2 (Fig. 4.17 B right). None of the subunits could affect or cause significant changes to CAPH2 binding at CAPH2-specific binding sites (pp4 & pp5) and joint N-MYC/TFIIIC/CAPH2 binding sites (pp6 – pp8).

4.5.5 TFIIIC5 and CTCF

TFIIIC and CTCF are both DNA binding architectural proteins and CTCF motif was strongly enriched in N-MYC/TFIIIC5 joint binding sites (Büchel et al., 2017). CTCF peaks were mostly co-occupied with RAD21. However, despite the effect of TFIIIC5 on RAD21 chromatin binding, TFIIIC5 depletion did not affect CTCF binding at intergenic or intragenic regions jointly bound by N-MYC/TFIIIC/RAD21/CTCF (Fig. 4.18).

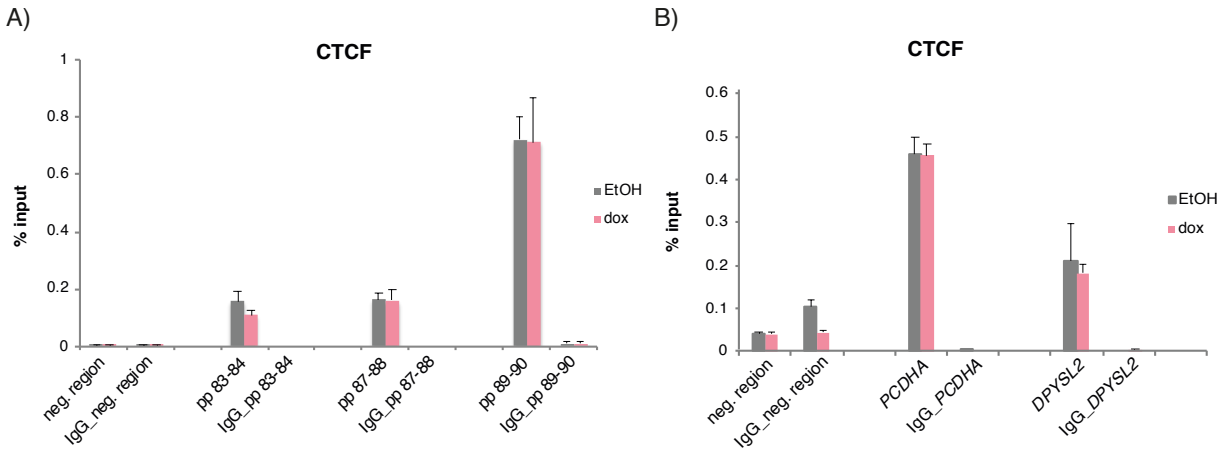


Figure 4.18: TFIIIC and CTCF at chromatin
Chromatin-immunoprecipitation of TFIIIC5 depleted IMR-5 cells. Chromatin of IMR-5 upon 48 hr doxycycline induction for TFIIIC5 was immunoprecipitated with CTCF antibody or an unspecific IgG as control. Precipitated and purified DNA was analyzed by qPCR with primers amplifying the intergenic regions (A) or intragenic regions (B) of joint N-MYC/TFIIIC/CTCF/RAD21 binding sites or a negative control region.
Remarks: Data are represented as mean ± SD of technical triplicates.

4.6 TFIIIC subunits predominantly bind at open promoters over enhancers

Regulatory elements such as enhancers were reported to bind architectural proteins to mediate chromatin interactions. Chromatin environment of TFIIIC subunits and other architectural proteins was therefore investigated. Enhancers and open promoters were identified by ChIP-seq of histone mark modifications. DNA-binding architectural proteins CTCF and TFIIIC subunits predominantly bound open promoters (Fig. 4.19 A). In contrast, accessory architectural proteins had more binding to enhancers (Fig. 4.19 A & B). Previous report has shown TFIIIC binding to a subgroup of SINE elements with same histone mark modifications as enhancers (eSINE). However, number of eSINE across the genome may be very limited which made it hard to notice in a heat map (Fig. 4.19 B).

Lastly, effect of intrinsic HAT activity of TFIIIC2 on N-MYC target genes was investigated. Upon 48 hours of TFIIIC2 knock-down, no change in mRNA expression of four N-MYC target genes was observed (Fig. 4.19 C).

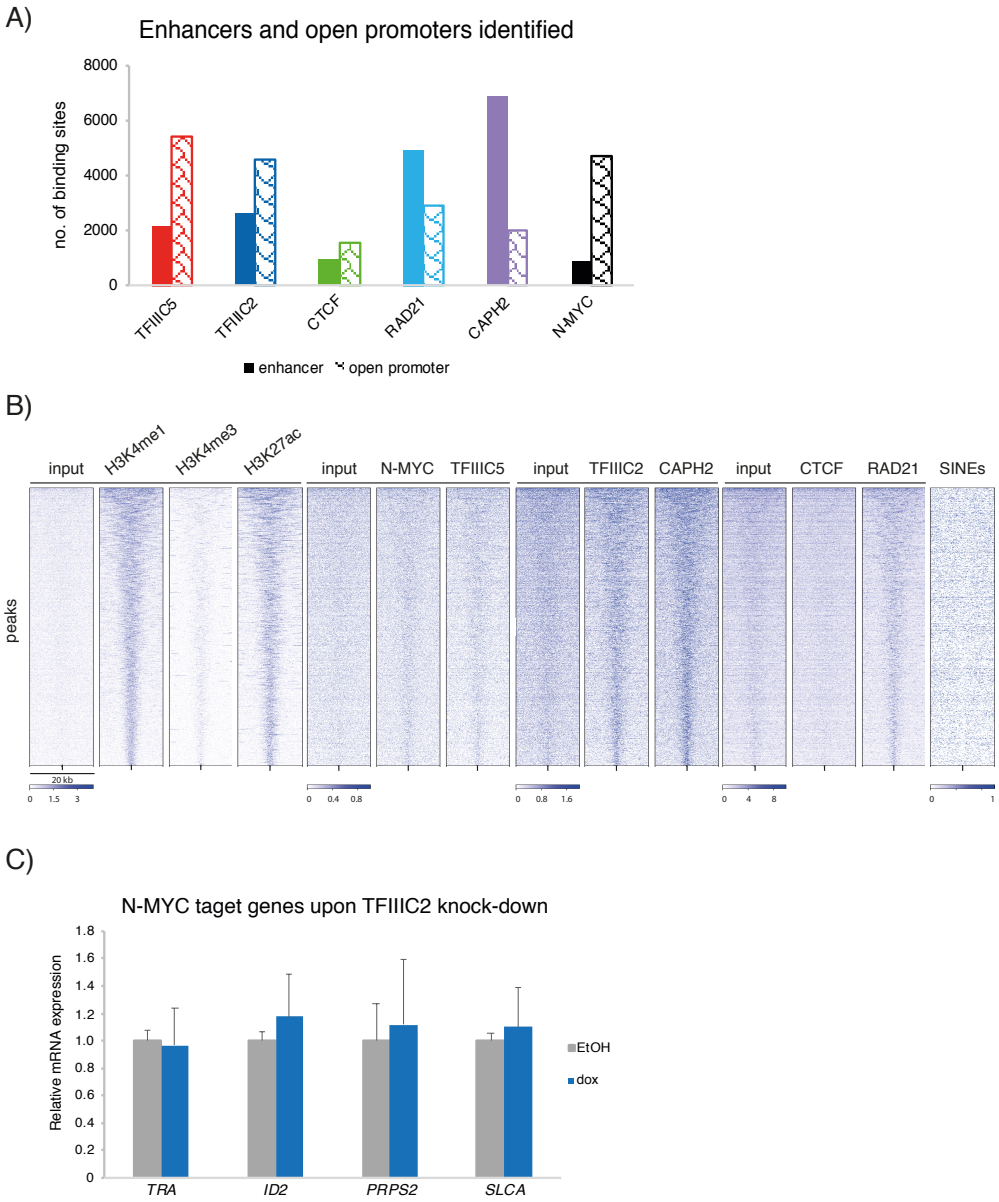


Figure 4.19: Chromatin environment and architectural proteins

A) Graphical illustration of number of architectural proteins and N-MYC bound in open promoters and enhancers.

B) Heat map displaying the distribution of 8532 enhancers at N-MYC and architectural proteins binding sites of all annotated human RefSeq TSS in a window of ± 10 kb. A track illustrating the presence of annotated SINE elements was included. Tag densities are sorted by H3K4me1 (50 bp resolution). No. of peaks analyzed are same as in Fig. 4.5, plus 29223 peaks for H3K27ac, 19233 for H3K3me1 and 17234 peaks for H3K4me3.

C) qPCR of N-MYC target genes upon TFIIIC2 depletion. Data are represented as mean ± SD of technical triplicates.

5 Discussion

Identification of TFIIIC complex as interacting partner of N-MYC guided the basic research of N-MYC into uncharted waters. Importantly, genome-wide N-MYC/TFIIIC5 joint binding sites founded the ground which strongly suggests N-MYC binds not only to E-box sequences, but also to non-E-box sequences for cellular or regulatory purposes. TFIIIC complex was first identified as core component of RNAPIII transcription machinery and is known for its essential role in initiating transcription for tRNA genes. Early studies of TFIIIC in yeasts discovered ETC sites which eventually unveiled the additional role of TFIIIC as an architectural protein. However, what is known about TFIIIC complex or its individual subunits is still very limited. Not to mention the biological relevance of N-MYC/TFIIIC5 co-occupancies across the genome. Therefore, the main objective of this thesis was to elucidate the functional importance of TFIIIC on N-MYC by studying two TFIIIC subunits from different sub-complexes.

5.1 TFIIIC subunits act as a single complex in tRNA synthesis but become two different entities for non-canonical functions

In this study, TFIIIC5 and TFIIIC2 which are subunits from sub-complex A and B respectively were selected. The aim was to obtain a better overview about the TFIIIC complex by comparison between two subunits. Due to the availability of working antibodies and/or the assumption that all TFIIIC subunits work together as one single complex, studies on ETC sites were mostly based on the binding of only one TFIIIC subunit (Moqtaderi et al., 2010; Oler et al., 2010). One comprehensive report in *Drosophila* studied all subunits in sub-complex B which demonstrated similar binding patterns (Carrière et al., 2012). Nonetheless, many other reports targeted different subunits in different experiments and regarded them to be interchangeably the same (Crepaldi et al., 2013; Policarpi et al., 2017; Yuen et al., 2017). No study by far compared subunits from different sub-complexes and it is unknown if discrepancy between them exists. Here, ChIP-seq of RNAPIII transcription machinery was performed and analyzed. Apart from displaying how the two TFIIIC subunits involve in tRNA synthesis (Fig. 4.2), it allowed the identification of ETC sites (Fig. 4.5).

The binding patterns of RNAPIII transcription machinery align with the previous findings. For instance, the percentage of tRNA bound by RNAPIII (Fig. 4.2 A) and the distribution of RNAPIII transcription machinery relative to tRNA genes (Fig. 4.2 D) are in line with what was reported (Carrière et al., 2012; Moqtaderi et al., 2010).

As observed in genome browser picture (Fig. 4.2 C), TFIIIC5 and TFIIIC2 binding was slightly displaced from each other. This is mainly due to the distance between A- and B-boxes which are located internally of tRNA genes. The relative distance between two boxes hinges on the length of tRNA genes. The minimal distance reported is 40 bp (Arimbasseri et al., 2014). Presence of intron within tRNA gene will further separate the boxes.
Binding data of RNAPIII transcription machinery to tRNA genes did not show any particular or unusual variations between TFIIIC5 and TFIIIC2. Therefore, it confirms both TFIIIC subunits function as a complex for tRNA synthesis.

However, further analysis focusing on A- and B-box (Fig. 4.3) and other motif distributions (Fig. 4.10), N-MYC/TFIIIC joint binding sites (Fig 4.4), and association with other architectural proteins (Fig. 4.5 & 4.7) revealed unprecedented discrepancies between two TFIIIC subunits.
A- and B-boxes refer to the motifs bound by sub-complex A and B respectively. Surprisingly, instead of at the peak center of TFIIIC5, A-box was distributed at ± 60 bp away from TFIIIC5 peak (Fig. 4.3 A). This differs from the notion that TFIIIC5 is the A-box binding subunit of sub-complex A. One of the possible explanations is about the structure of the complex. Since TFIIIC complex is quite large in size, the organization of different subunits may result in the distance between TFIIIC5 peaks and A-box. Protein structure of the entire TFIIIC complex is by far still lacking. Otherwise, it would provide more in-depth insights to answer this question.
Moreover, B-box was distributed ± 60 bp around TFIIIC5 asymmetrically, with an inclination to localize upstream of TFIIIC5 (Fig. 4.3 B). This B-box distribution pattern was even more astonishing in the case of TFIIIC2. It was further deviated from TFIIIC2 peaks than TFIIIC5 despite the fact that TFIIIC2 should be B-box binding. B-box is found both in tRNA genes and ETC sites. While A-box was distributed with modest similarity between two TFIIIC subunits, the distinct distribution of B-box around TFIIIC5 and TFIIIC2 strongly suggests they do not function together at binding sites other than tRNA genes.
Since it was already known that N-MYC and TFIIIC5 jointly bind to multiple loci across the genome, the co-occupancy between N-MYC and TFIIIC2 was investigated (Fig 4.4). Co-binding sites of N-MYC with TFIIIC2 were mostly at promoters (Fig. 4.4 C). This tendency is sharply different from that of N-MYC/TFIIIC5 joint binding sites. This was further confirmed with the enrichment association with N-MYC at TFIIIC2-only ETC sites (Fig. 4.7 A & B). Another important evidence that distinguishes one TFIIIC subunit from another was the association with architectural proteins (Fig. 4.7). TFIIIC5 was strongly enriched with CTCF

and RAD21 which is a prominent feature known as architectural protein binding sites (APBS) (more discussion in 5.2). However, TFIIIC2 did not demonstrate this property.

In short, these results expose the previously unappreciated discrepancies in non-canonical functions of TFIIIC between the two sub-complexes. Co-IP data could only show interaction between N-MYC and TFIIIC5 but not that of N-MYC and TFIIIC2 (Fig. 4.14). This was supported by unpublished data from N-MYC/TFIIIC co-purification experiment which showed that only sub-complex A was co-expressed with N-MYC (Eoin Leen, personal communication). All these data pinpoint TFIIIC5 and TFIIIC2 for being two different functional entities with N-MYC.

5.2 Profiling of architectural proteins in neuroblastoma cell line IMR-5

Genomic binding profile of an architectural protein can give a lot of insights about how the protein functions. In this study, binding profiles of both TFIIIC subunits and other architectural proteins (CTCF, RAD21, and CAPH2) in untreated *MYCN*-amplified neuroblastoma cells IMR-5 were generated by ChIP-seq. These binding profiles are therefore unique to this cell line. Comparison between binding sites of one architectural protein in different cell lines affords the association with tissue-specificity. A study compared CTCF binding in 19 cell lines and revealed as high as 64% of binding sites is not conserved in at least one cell type (Wang et al., 2012). However, cohesin has shown colocalization with estrogen receptor in breast cancer, liver-specific transcription factors in hepatocellular carcinomas and pluripotency factors in embryonic stem cells (Faure et al., 2012; Nitzsche et al., 2011; Schmidt et al., 2010). Hence, it was expected that N-MYC, the predominant tissue-specific factor in IMR-5, may demonstrate colocalization with architectural proteins in addition to TFIIIC5.

Indeed, ChIP-seq data showed N-MYC enrichment with all architectural proteins, albeit in varying intensities (Fig. 4.8). RAD21 exhibited the strongest binding strength among five architectural proteins studied and the binding intensity was stronger at sites with lower N-MYC occupancy. The architectural proteins with second strongest binding strength was CTCF. CTCF was found mainly at N-MYC sites outside promoters. Similar to RAD21, CTCF binding was stronger at sites with lower N-MYC occupancy. In a similar vein to RAD21 and CTCF, CAPH2 was more enriched at N-MYC peaks outside promoters. However, CAPH2 had stronger binding strength at sites with higher N-MYC occupancy, which was opposite to the case of CTCF and RAD21. TFIIIC2 had the weakest enrichment with N-MYC among all architectural proteins.

Total TFIIIC5 binding and TFIIIC5/TFIIIC2 common ETC sites were also enriched with other architectural proteins in similar pattern to what was observed in N-MYC (Fig. 4.7). However,

the enrichment at total TFIIIC2 binding sites and TFIIIC2-only ETC sites was very weak (Fig. 4.6 & 4.7). This result highlights that TFIIIC5 and N-MYC share similar enrichment pattern. This implicates functional linkage between them and this excludes TFIIIC2.
Moreover, the tight association with tissue-specific factor may also explain why the binding patterns of TFIIIC5 and CAPH2 obtained in this study is not as strong as recent publication in which the ChIP-seq was performed in mESC (Yuen et al., 2017). It also implies the interaction between TFIIIC5 and CAPH2 may not be a general mechanism for all cell types.

The second importance of binding profiles is to study APBS (architectural protein binding sites). Architectural proteins often co-localize at TAD border and the degree of APBS occupancy, i.e. number and intensity of all architectural proteins present, can define TAD border strength (Li et al., 2015; Van Bortle et al., 2014). It does not require the presence of certain protein.
As shown in the overlap of binding sites between TFIIIC subunits, RAD21, CTCF and CAPH2, both TFIIIC subunits reported more than 1000 binding sites shared with other three architectural proteins (Fig. 4.5 A). Although the enrichment strength of architectural proteins at TFIIIC2 was multi-fold weaker than that of TFIIIC5, the number of joint binding sites (1263 in the case of TFIIIC5 and 1068 in the case of TFIIIC2) did not differ hugely. If the threshold for peak calling can be further elevated, TFIIIC subunits would be expected to have bigger difference in number of co-occupancies with other architectural proteins.
TFIIIC5 is enriched with the architectural proteins to a higher degree than that of TFIIIC2 (Fig. 4.5 B & C). This entails the joint binding sites between TFIIIC5 and other architectural proteins would be APBS with higher occupancies and offer stronger border strength than that of TFIIIC2. If Hi-C data in IMR-5 cells were available, these APBS sites could be aligned to the TAD and give a more complete picture of how N-MYC is involved.
TAD organization is highly elastic. Architectural proteins can be redistributed from TAD borders to intra-TAD regions or change in number and level of binding sites under different conditions. Various factors have been shown to contribute to the genome dynamics (Kakui et al., 2017; Kieffer-Kwon et al., 2017; Lazar-Stefanita et al., 2017; Li et al., 2015; Nagano et al., 2017). The binding profiles were based on untreated IMR-5 cells, which had majority of cell population at G1 phase. Complex formation between N-MYC, TFIIIC5, and RAD21 occurs at G1 and G2 phases (Büchel et al., 2017). It is therefore meaningful to evaluate the binding profiles and TAD organization at different cell cycle phases. Hi-C data from cells depleted of cohesin and CTCF are already available (Nora et al., 2017; Rao et al., 2017). During the

preparation of this thesis, a manuscript describing how TFIIIC organizes genome architecture in breast cancer was available on preprint server (Ferrari et al., 2018). It shows TFIIIC first binds to *Alu* elements and promotes DNA loop between *Alu* and promoters. This in turn controls expression of cell-cycle dependent genes (further discussion about *Alu* in section 5.4).

All in all, the profiling of architectural proteins provides the first insight into how N-MYC is associated with different architectural proteins. Also, the data identified different combinations of APBS which lay the ground for future study of TAD border. Hi-C experiment would bring the understanding of architectural proteins in relation to N-MYC a step forward.

5.3 TFIIIC5 recruits RAD21 but not CAPH2 to chromatin

In order to understand functional relevance of joint N-MYC/TFIIIC sites, TFIIIC subunits were depleted and the chromatin was subjected ChIP experiment. A panel of genomic loci with co-occupancies of N-MYC/TFIIIC were chosen as target sites. These target sites also share RAD21, CTCF, and CAPH2 binding, except *PLK1* and *EIF4H* were not enriched with CTCF. Overlap of architectural proteins binding sites showed as high as 97% of CTCF binding sites was jointly bound by RAD21 (Fig. 4.5 E). This was even higher than other reports which ranged from 80 – 85% (Faure et al., 2012; Schmidt et al., 2010). However, RAD21 had more than 20000 sites independent of any architectural proteins. Majority of joint sites between RAD21 and TFIIIC5 involve CTCF. On the other hand, there is another group of co-occupancies made up of TFIIIC5/RAD21/CAPH2. APBS can be found at TAD borders as well as inside TADs for *cis*-interaction. In particular, cohesin binding is not only for long-range interaction but has been linked to transcription (Busslinger et al., 2017; Davidson et al., 2016). Cohesin is loaded to DNA by Nibpl at sites which are mostly TSS for active genes. Transcription is the driving force to translocate cohesin and provides directionality to the movement. This translocation also helps to position cohesin to CTCF sites by allowing cohesin to slide along the DNA until it encounters CTCF. Until that point, cohesin functions as an architectural protein and mediates long-range chromatin *cis*-interactions. Data from other report has shown the cohesin sites for transcription and that of long-range interaction do not overlap (Kagey et al., 2010). Nonetheless, presence of joint N-MYC/TFIIIC/RAD21/CTCF/CAPH2 sites may point to unexplored functions.

ChIP experiments upon depletion of TFIIIC5 subunit revealed N-MYC requires TFIIIC5 to determine chromatin binding of RAD21. RAD21 was recruited by TFIIIC5 to chromatin (Fig.

4.15 D). Further ChIP experiment upon N-MYC depletion revealed the hierarchical order that N-MYC recruits TFIIIC5 which eventually recruits RAD21 to joint N-MYC/TFIIIC5 binding sites (Büchel et al., 2017). However, CAPH2 binding upon TFIIIC5 depletion increased at some target sites (*PPRC1* and *GALNT14*) while the other two sites did not respond. The different effects of TFIIIC5 on RAD21 and CAPH2 binding may be reminiscent of the antagonizing effect between RAD21 and CAPH2 on enhancer-promoter contacts (Li et al., 2015). RAD21 facilitates interactions between regulatory elements such as enhancer-promoter, enhancer-enhancer, and promoter-promoter interactions (Seitan et al., 2013; Whalen et al., 2016). Yet, CAPH2 is suggested to be a negative regulator of enhancer-promoter *cis*-interactions (Bauer et al., 2012; Hartl et al., 2008; Joyce et al., 2012; Li et al., 2015). With that, cohesin and condensin II complexes operate together to guarantee accurate and precise enhancer-promoter communication. Although the data presented is not strong enough to claim TFIIIC5 applies opposing mechanisms on RAD21 and CAPH2, it should still be noted that these two accessory proteins may have antagonizing function.

RNA-seq experiment upon depletion of TFIIIC, RAD21, and N-MYC elucidates gene expression changes are corelated to MYCN amplification and advanced tumor stage of neuroblastoma (Büchel et al., 2017). RAD21 and TFIIIC5 are responsible for a subset of N-MYC target gene expression.

Previous reports have shown mild effects of RAD21 depletion on gene expression (Rao et al., 2017; Zuin et al., 2014). Upon acute depletion of RAD21, only 1.2% of total expressed genes showed expression changes for more than 1.75 fold (Rao et al., 2017). In another *in vitro* system, around 50 genes (with false discovery rate less than 5%) were reported differentially expressed after RAD21 depletion and similar result was observed upon CTCF depletion (Zuin et al., 2014). Genes from Hox gene family are among list with lower expression upon RAD21 depletion and they have been shown to be involved in topological organization and contact to distant enhancers (Zuin et al., 2014). Therefore, depletion of architectural proteins may not trigger dramatic changes in global expression profile but only a certain subgroup of genes. This indicates the control of genome architecture is a very precise and delicate mechanism on gene expression.

Based on the division of labor between cohesin and condensin II aforementioned, it is postulated that CAPH2 may counter the effect of RAD21 on N-MYC target gene expression. This can be evaluated with RNA-seq experiment to profile the role of CAPH2 in gene expression in IMR-5 cells.

CTCF binding to chromatin was also studied (Fig. 4.18). No change was observed with TFIIIC5 knockout. Similar to TFIIIC, CTCF is a DNA-binding architectural protein. It may therefore only couple with accessory proteins for regulatory purposes. Association with TFIIIC is more likely as APBS at TAD borders for maintenance of border strength (Fig. 4.7). Hi-C experiment provides information about TAD and is suggested to validate this statement.

5.4 TFIIIC is not an adaptor protein for CAPH2

In accordance to the classic mode of functions employed by architectural proteins, one DNA-binding protein pairs with an accessory protein. Data from studies in yeasts have established the interaction relationship between TFIIIC and CAPH2. TFIIIC recruits condensin in yeasts through its B-box motif (D'Ambrosio et al., 2008b; Haeusler et al., 2008). Inactivation of TFIIIC subunit reduces condensing binding at chromosome. Condensin can mediate clustering of tRNA in yeast nucleolus (Haeusler et al., 2008). Condensin is associated with both RNAPII- and RNAPIII-transcribed genes through interaction with TATA-box binding protein (TBP) which is common for all RNAPIII transcription machinery. The predominant localization of condensin at centromeres may explain why RNAPIII genes and active RNAPII genes are usually associated with centromere (Iwasaki et al., 2015; Nakazawa et al., 2008). Apart from that, condensin and TFIIIC share the chromatin bookmark feature. Unlike cohesin which only binds to chromatin during interphase, condensin is found on DNA throughout cell cycle (Gause et al., 2010; Wang et al., 2005). TFIIIC stands out from other transcription factors because it does not leave the chromatin at any time within a cell cycle (Donze, 2012). This unified presence at chromatin therefore suggests condensin remains at TFIIIC sites for structural maintenance of chromosomes.

Despite the strong evidence that suggests TFIIIC and CAPH2 to be a perfect match, the experimental data did not follow the expectation. As discussed in section 5.3, TFIIIC5 depletion resulted in unloading of CAPH2 from some of the N-MYC/TFIIIC joint sites (Fig. 4.17 A). In addition, CAPH2-specific binding sites were not influenced (Fig. 4.17 B). No enrichment of CAPH2 was observed at B-box (Fig. 4.10 A) nor tRNA genes (data not shown). The data clearly states that TFIIIC is not the recruiting factor of CAPH2 to chromatin. Also, CAPH2 is only responsive to TFIIIC at a subset of genomic loci and the effect of TFIIIC on CAPH2 is not global.

Difference between species would be a possible reason. One on hand, there is only one condensin complex in yeasts whereas it was evolved to two in higher organisms. On the other

hand, TFIIIC paralogs are only conserved in subunit composition but not protein sequences. Many properties observed in yeast TFIIIC may not still hold true for the paralogs in human. For instance, subcellular localization of TFIIIC subunits sfc3 and sfc6 is predominantly at nuclear periphery. However, their corresponding paralogs TFIIIC1 (Human protein atlas) and TFIIIC2 (Fig. 4.1) localizes throughout the nucleus. Nonetheless, TFIIIC3 shows peripherical distribution in non-*MYCN* amplified neuroblastoma cell line SH-SY5Y, Hela and U2OS (Human protein atlas). On top of that, CTCF, the most well-known architectural protein in mammals, does not exist in yeasts. In this case, responsibility of genome organization in yeasts may all fall on TFIIIC. In higher organisms which bear more sophisticated biological systems, several architectural proteins can cooperate together and develop more complicated approaches for genome organization. It should also be emphasized that function of an architectural protein is not always fully conserved. Although cohesin makes use of CTCF for sequence specificity for binding in mammals, cohesin mainly associates with Nipbl at active genes and do not share strong co-occupancies with CTCF in *Drosophila* (Schaaf et al., 2013; Schmidt et al., 2010b). The contradiction between the results from this study and another mammalian system from Yuen et al. has already been discussed in section 5.2.

The classic paradigm that accessory protein needs an adaptor protein is recently challenged by a real-time imaging study of condensin in yeasts (Ganji et al., 2018b). It demonstrates how condensin alone is capable of anchoring DNA for loop formation. Condensin undergoes loop extrusion without TFIIIC or any other adaptor proteins. If this is a general mechanism that also applies to mammals, it could explain the function of about one-third of CAPH2 sites that was not shared with other architectural proteins (Fig. 4.7 E).

To sum up, the data ascertain that TFIIIC is not the recruiting factor of CAPH2 in neuroblastoma cell line. On the contrary, TFIIIC5 could unload CAPH2 from chromatin at some N-MYC/TFIIIC joint sites while CAPH2 binding at its unique sites remained unchanged. Recent finding in yeasts that shows DNA loop formation solely by condensin without help from adaptor protein offers a clue that architectural proteins can have different modes of action for genome organization. Many of which still await discovery.

5.5 CAPH2 is novel protein interacting partner of N-MYC

Mass-spectrometry of N-MYC not only identified TFIIIC subunits but also many members from SMC family as new interacting partners of N-MYC (Fig. 5.1) (Büchel et al., 2017). Despite majority of the family is on the list, CAPH2 is not called as a hit from the proteomics. However, Co-IP experiment could still show the interaction between CAPH2 and N-MYC (Fig. 4.16 A). A further confirmation of whether it is a protein-protein or DNA-protein interaction requires benzonase nuclease treatment of IP samples.

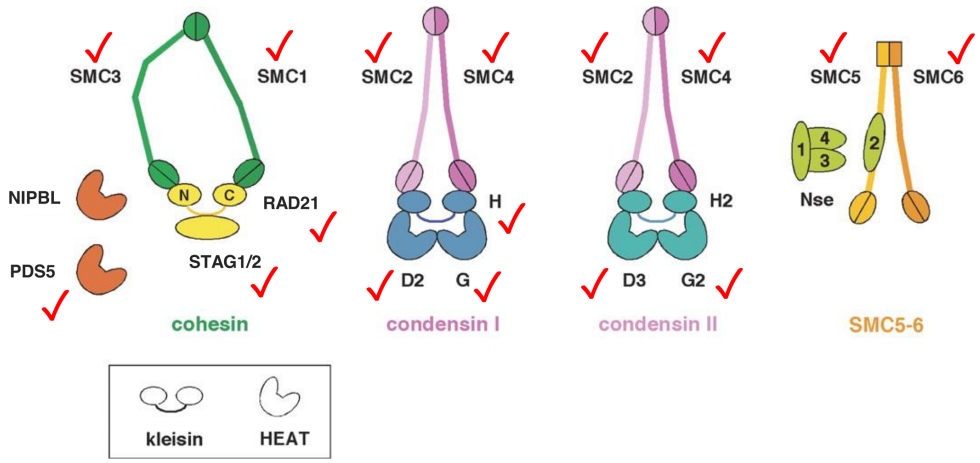


Fig. 5.1 Cartoons showing mammalian members of SMC protein family and their structures
Proteins identified in N-MYC mass spectrometry of (Büchel et al., 2017) were marked with a tick. Diagram is modified from (Losada and Hirano, 2005).

Interaction between CAPH2 and N-MYC was also studied by Aurora-A inhibitor CD532. CAPH2 binding to chromatin is responsive to CD532 treatment (Fig. 4.17 A). The elevation in binding is similar to the response of TFIIIC5 and RAD21. In house unpublished data showed that CD532 is a CDK9 inhibitor which prevents RNAPII to go through efficient elongation (Morales and Giordano, 2016). Since TFIIIC5 has been shown to be involved in N-MYC-dependent promoter escape of RNAPII and the subsequent pause release, it would be worth investigating if treatment of transcription inhibitor such as flavopiridol or DRB could result in similar effect as CD532 and if the combined treatment will generate stronger effect. The answer can help to explain the increase in TFIIIC5 binding site and intensity upon CD532 treatment and may eventually provide other clues for the effects on CAPH2 and RAD21.

RAD21 forms a complex with N-MYC and TFIIIC5 which is validated by Co-IP and *in vitro* binding assay (Büchel et al., 2017). CAPH2 co-immunoprecipitated with N-MYC and TFIIIC2. Yet, no direct interaction with TFIIIC5 was observed (Fig. 4.16). Since TFIIIC2 could not pulldown N-MYC (Fig. 4.14 B), CAPH2 may form separate complex with N-MYC and

TFIIIC2 or other protein also exists in the complex which hinders efficient detection by Co-IP. Further investigation is essential to understand how N-MYC/CAPH2 complex. With the information about the binding site of TFIIIC5, RAD21 and Aurora-A on N-MYC, further knowledge regarding the docking site for N-MYC/CAPH2 interaction can definitely offer better the understanding of N-MYC biology.

Another important aspect of N-MYC biology is the cell-cycle dependency. N-MYC/TFIIIC5/RAD21 complex exists at G1 and G2 phases and is replaced by N-MYC/Aurora-A during S-phase (Büchel et al., 2017). Interaction between N-MYC and CAPH2 may also be cell-cycle dependent. Because RAD21 and CAPH2 interacted with TFIIIC5 at the same genomic loci, it is important to know whether both events take place simultaneously or controlled by other factors such as cell cycle phases.

Recent finding show activated B-cells which have higher MYC level are corelated to decompaction of nucleosome (Kieffer-Kwon et al., 2017b). Intriguingly, depletion of condensin II subunits results in increase in nuclear size, presumably due to condensin-mediated chromatin compaction (George et al., 2014). Dramatic changes in nuclear shape are observed in both cases but in opposite directions. Condensin II regulates chromatin compaction meanwhile MYC regulates chromatin decompaction. Although the underlying molecular mechanisms for the observations in B-cells has not been unveiled, it is persuasive that the SMC family may be involved. A myriad of proteins from SMC family are noted as N-MYC interactors and that their interaction with N-MYC may serve to antagonize each other for DNA packaging. By far, very little is known about MYC or N-MYC on genome architecture organization. But the interaction between N-MYC and CAPH2 or other SMC proteins is certainly worthy of more in-depth exploration.

5.6 TFIIIC2 has weak association with other architectural proteins

One of the key findings of this study was the association of two TFIIIC subunits with other architectural proteins (Fig. 4.5). TFIIIC2 and TFIIIC5 were in sharp contrast to each other – TFIIIC5 showed strong association with CTCF and RAD21 whereas TFIIIC2 was only modestly enriched with CTCF and RAD21. At TFIIIC2-only ETC sites, only weak enrichment of RAD21 could be observed (Fig. 4.7 A & B). In comparison to CTCF and RAD21, association with CAPH2 was weak for both TFIIIC subunits (Fig. 4.5).

TFIIIC2 shared ETC sites with TFIIIC5 and those common ETC sites were enriched with other architectural proteins as in the case of total TFIIIC5 peaks (Fig. 4.7 C & D). This indicates that a group TFIIIC2 sites cooperate with TFIIIC5 for APBS. Another group of TFIIIC2 sites are distributed across the genome with unknown functions. Localization of ETC sites is believed to be non-stochastic. But what determines those TFIIIC2 sites for co-occupancy with TFIIIC5, whereas a bigger proportion of its sites are located elsewhere requires more in-depth understanding of this protein. One intriguing direction to look into is CAPH2, as both proteins peak at a defined distance downstream of RefSeq gene TSS (Fig. 4.9).

TFIIIC2 was able to alter TFIIIC5 chromatin binding but the reverse was not true. This was shown in the ChIP experiment using cells with TFIIIC subunit knockdown (Fig. 4.15 A & B). TFIIIC2 binding remained unchanged when TFIIIC5 was depleted. On the other hand, TFIIIC5 binding decreases by 50% at many target sites upon the influence of TFIIIC2 knockdown. These results appear that TFIIIC2 recruits TFIIIC5 to chromatin. However, TFIIIC2-depleted samples did not lead to reduction of RAD21 binding, as observed in the case of TFIIIC5 depletion (Fig. 4.15 D). N-MYC is the major contributor to the hierarchical chromatin binding of TFIIIC5 and RAD21. It would therefore be meaningful to perform a TFIIIC2 ChIP when cells are depleted of N-MYC and evaluate if the binding of TFIIIC2 changes.

In order to understand the complications between TFIIIC5 and TFIIIC2, investigation of one more TFIIIC subunit may shed some lights on this puzzle. TFIIIC4 would be an ideal candidate. It has been shown to interact with TFIIIC1, TFIIIC2 and TFIIIC5 (Hsieh et al., 1999). The interaction linkage with subunits from both sub-complexes bestows its alias as “the bridge between two sub-complexes”. In addition, TFIIIC4 has intrinsic HAT activity, which is similar to TFIIIC2 (Hsieh et al., 1999). It is therefore a good target to investigate the distinct roles of TFIIIC5 and TFIIIC2.

5.7 TFIIIC2 preferentially binds SINE elements

Modest association between TFIIIC2 and other architectural proteins is astonishing. The immediate questions would be where TFIIIC2 binds to and what its function is. Two publications regarding TFIIIC in neurons highlighted the link between TFIIIC and enhancer SINE elements (eSINEs) which are SINE elements containing same histone modifications as enhancers. The data showed that TFIIIC makes use of eSINEs to bridge RNAPII and RNAPIII transcription machinery together for a tight expression regulation of activity-dependent genes

(Crepaldi et al., 2013; Policarpi et al., 2017). These exciting data directed the investigation of TFIIIC2 to SINE elements binding.

SINEs are transposable elements which accounts for more than one-tenth of eukaryotic genome, in comparison to a mere 2% of protein-coding genes (Lander et al., 2001). They share A- and B-box motifs like tRNA genes (Fig. 5.2). They are mainly distributed at gene-rich regions, in contrary to LINEs (long interspersed elements) which are enriched in intergenic regions (Medstrand et al., 2002). Being repetitive elements, SINEs were once dubbed as the junk DNA. But new discoveries have established their importance, in particular a subgroup called *Alu* elements, in pre-mRNA splicing (Lev-Maor et al., 2003), nuclear localization (Lubelsky and Ulitsky, 2018) and transcriptional activation as enhancer (Tashiro and Lanctôt, 2015).

TFIIIC2 indeed showed strong enrichment to SINE elements and it spanned through a region up to 300 bp from SINE TSS (Fig. 4.11). This binding pattern is compatible with a previous report which demonstrates all subunits from sub-complex B are enriched with RNAPIII-bound SINEs in (Carrière et al., 2012). The data here additionally shows the enrichment does not depend on transcription status of SINEs since total SINEs were considered in the analysis (Fig. 4.11). In particular, the average size of *Alu* elements is around 300 bp, in comparison to SINEs which range from 100 – 700 bp (Fig. 5.2). This may imply TFIIIC2 bind specifically to *Alu* elements.

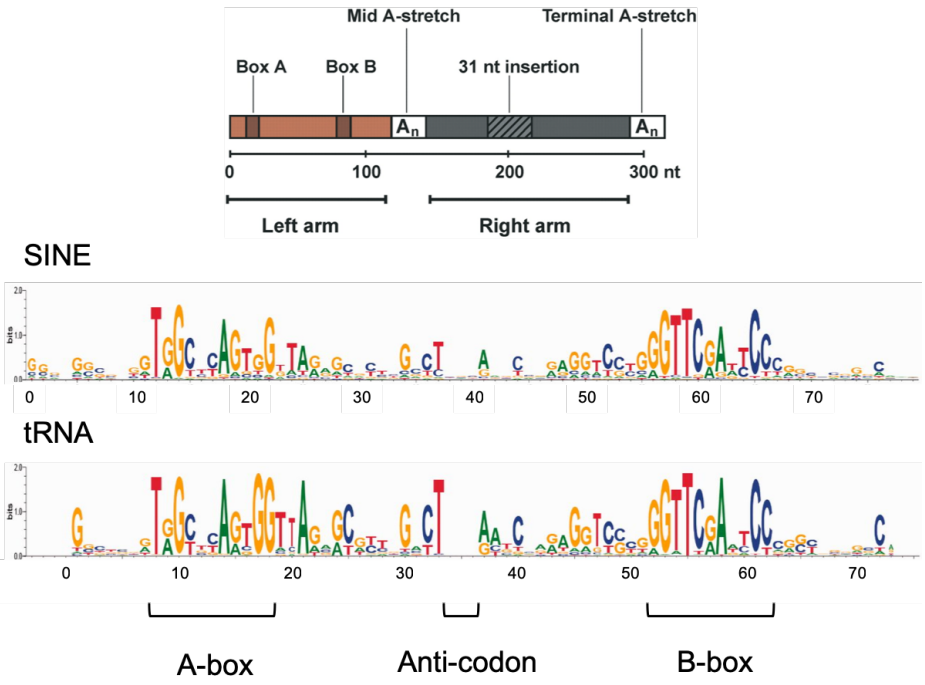


Fig. 5.2 Conserved SINEs sequence from tRNA
Upper panel is the architecture of *Alu* elements, which is a subgroup of SINE elements. Lower panel is sequences of SINE and human tRNA gene. It demonstrated the conserved domains and sequences between both transcripts. Modified from (Häsler and Strub, 2006; Vassetzky and Kramerov, 2013)

Binding to SINE or *Alu* elements is believed to be a unique regulatory mechanism of TFIIIC. In addition to the studies in neurons (Crepaldi et al., 2013b; Policarpi et al., 2017), a manuscript to be peer-reviewed about genome organization by TFIIIC upon serum starvation was available online (Ferrari et al., 2018). Their data demonstrate how *Alu*-bound TFIIIC accounts for long-range intra-TAD interactions and consequently modulates gene expression of cell-cycle dependent genes. In particular, TFIIIC harnesses binding to *Alu* elements that are close to promoters of target genes.

The new information about TFIIIC and SINE/*Alu* elements is very exciting and all point to the importance of regulation of a subset of genes that are cell-type specific. It may also apply to the case of neuroblastoma in which N-MYC is the master gene regulator. SINE/*Alu*-bound TFIIIC would be a feasible mechanism utilized by N-MYC for maintaining gene expression profile in neuroblastoma. eSINEs were not noticed in Fig. 4.19 B. However, since this regulation is specific to a subgroup of N-MYC target genes, it certainly requires sophisticated algorithms to identify eSINE.

5.8 HAT function of TFIIIC2 on N-MYC transcription regulation

Literature has long suggested that subunits from sub-complex B contain HAT activities which is absent in sub-complex A (Fukuda, 2006). HATs are epigenetic regulators of the genome by modifying histone chromatins. Acetylation of lysine residues on N-terminal tails of histones favors gene expression. Native HATs mostly form multi-peptide complexes of which other components can determine the preference of HAT substrates or target genes. TFIIIC also follows this general role and all subunits from sub-complex B have been reported to have HAT activities. TFIIIC2 can take advantage of its intrinsic HAT property to activate tRNA transcription. However, whether this also applies to RNAPII transcription has never been explored (Hsieh et al., 1999).

MYC interacts with proteins with HAT activities such as CBP and p300 to regulate MYC-mediated gene activation (Ogryzko et al., 1996). Therefore, question of whether TFIIIC2 regulate N-MYC gene expression was raised.

Expression of N-MYC target genes was evaluated upon depletion of TFIIIC2 for 48 hours (Fig. 4.19 C). No significant change was observed under the influence of TFIIIC2 knockdown. Although N-MYC gene expression was not affected, HAT has also been suggested to have many other functions, such as cell-cycle progression and DNA damage repair (Carrozza et al., 2003). The study of TFIIIC2 functions in this thesis is however not exhaustive.

As discussed in section 5.9 about TFIIIC2 binding to SINE elements, it is suggested that binding of TFIIIC to *Alu* elements increases H3K18 acetylation of the *Alu* elements (Ferrari et al., 2018). The HAT activity of TFIIIC2 may contribute to N-MYC gene expression in an very indirect manner.

5.9 Protein depletion system for study of essential genes

Gene silencing is a powerful tool to study the function of protein-of-interest. Different methods of protein depletion in mammalian cells have been developed and they each have pros and cons. In this thesis, gene knock-out and shRNA-mediated gene knockdown were employed. Gene knock-out is based on the principle that mutations introduced to gene sequence will results in pre-mature stop codons or a non-functional protein. This approach can specifically and completely remove the protein. However, this process is not reversible. Gene knockdown using shRNA relies on the degradation of mRNA which eventually results in protein loss. The use of shRNA sequences generated by computational algorithms is straight-forward but the protein is only partially depleted and problem with off-target silencing exists. Essential genes such as TFIIIC subunits are particularly difficult to target with knock-out approach. sgRNA targeting 5' UTR of TFIIIC gene sequence was adopted so as to knock-out specifically the endogenous TFIIIC while cell could be rescued with exogenous TFIIIC upon doxycycline treatment. This experimental design is theoretically feasible but practically tedious. Doxycycline induction has to be maintained from infection of cells with sgRNA, selection with selection marker, single clone isolation and cell expansion. The whole procedure until the establishment of cell population sufficient for functional experiments requires more than one month. It is technically more challenging for IMR-5 because single IMR-5 cell is difficult to survive.

The measure adopted for investigation of TFIIIC subunits was shRNA. Satisfactory protein depletion (more than 80%) was obtained and cells did not confer any phenotypic difference upon doxycycline induction. An inducible system also allows ease of use. But this pINDUCER11 system has certain degree of leakiness in which the knock-down efficiency may be lowered over time. This requires selection of cells again.

An improved way for acute depletion of protein level is by auxin-inducible degron (AID) degradation (Nishimura et al., 2009). Protein-of-interest needs an AID tag which can interact with OsTIR1, a F-box protein, in the presence of auxin. The protein is eventually degraded by ubiquitin proteasomal degradation system. This method offers a huge advantage of its reversibility and the speed of achieving complete protein elimination. It allows elimination of protein within short time and before any cellular phenotype occur.

6 Bibliography

American Cancer Society (2018). Cancer Facts & Figures 2018.

Anderson, D.E., Losada, A., Erickson, H.P., and Hirano, T. (2002). Condensin and cohesin display different arm conformations with characteristic hinge angles. J. Cell Biol. *156*, 419–424.

Arabi, A., Wu, S., Ridderstråle, K., Bierhoff, H., Shiue, C., Fatyol, K., Fahlén, S., Hydbring, P., Söderberg, O., Grummt, I., et al. (2005). c-Myc associates with ribosomal DNA and activates RNA polymerase I transcription. Nat. Cell Biol. *7*, 303–310.

Baker, R.E., Gabrielsen, O., and Hall, B.D. (1986). Effects of tRNATyr point mutations on the binding of yeast RNA polymerase III transcription factor C. J. Biol. Chem. *261*, 5275–5282.

Barrington, C., Pezic, D., and Hadjur, S. (2017). Chromosome structure dynamics during the cell cycle: a structure to fit every phase. EMBO J. *36*, 2661–2663.

Barski, A., Chepelev, I., Liko, D., Cuddapah, S., Fleming, A.B., Birch, J., Cui, K., White, R.J., and Zhao, K. (2010). Pol II and its associated epigenetic marks are present at Pol III-transcribed noncoding RNA genes. Nat. Struct. Mol. Biol. *17*, 629–634.

Beagrie, R.A., and Pombo, A. (2017). Cell cycle: Continuous chromatin changes. Nature *547*, 34–35.

Beroukhim, R., Mermel, C.H., Porter, D., Wei, G., Raychaudhuri, S., Donovan, J., Barretina, J., Boehm, J.S., Dobson, J., Urashima, M., et al. (2010). The landscape of somatic copy-number alteration across human cancers. Nature *463*, 899–905.

Berwanger, B., Hartmann, O., Bergmann, E., Bernard, S., Nielsen, D., Krause, M., Kartal, A., Flynn, D., Wiedemeyer, R., Schwab, M., et al. (2002). Loss of a FYN-regulated differentiation and growth arrest pathway in advanced stage neuroblastoma. Cancer Cell *2*, 377–386.

Blackwell, T.K., Huang, J., Ma, A., Kretzner, L., Alt, F.W., Eisenman, R.N., and Weintraub, H. (1993). Binding of myc proteins to canonical and noncanonical DNA sequences. Mol. Cell. Biol. *13*, 5216–5224.

Blackwood, E.M., and Eisenman, R.N. (1991). Max: a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with Myc. Science *251*, 1211–1217.

Bouchard, C., Dittrich, O., Kiermaier, A., Dohmann, K., Menkel, A., Eilers, M., and Lüscher, B. (2001). Regulation of cyclin D2 gene expression by the Myc/Max/Mad network: Myc-dependent TRRAP recruitment and histone acetylation at the cyclin D2 promoter. Genes Dev. *15*, 2042–2047.

Bouchard, C., Marquardt, J., Brás, A., Medema, R.H., and Eilers, M. (2004). Myc-induced proliferation and transformation require Akt-mediated phosphorylation of FoxO proteins. EMBO J. *23*, 2830–2840.

Brenner, C., Deplus, R., Didelot, C., Loriot, A., Viré, E., De Smet, C., Gutierrez, A., Danovi, D., Bernard, D., Boon, T., et al. (2005). Myc represses transcription through recruitment of DNA methyltransferase corepressor. EMBO J. *24*, 336–346.

Brockmann, M., Poon, E., Berry, T., Carstensen, A., Deubzer, H.E., Rycak, L., Jamin, Y., Thway, K., Robinson, S.P., Roels, F., et al. (2013). Small molecule inhibitors of aurora-a induce proteasomal degradation of N-myc in childhood neuroblastoma. Cancer Cell *24*, 75–89.

Brodeur, G.M., Seeger, R.C., Schwab, M., Varmus, H.E., and Bishop, J.M. (1984). Amplification of N-myc in untreated human neuroblastomas correlates with advanced disease stage. Science *224*, 1121–1124.

Büchel, G., Carstensen, A., Mak, K.-Y., Roeschert, I., Leen, E., Sumara, O., Hofstetter, J., Herold, S., Kalb, J., Baluapuri, A., et al. (2017). Association with Aurora-A Controls N-MYC-Dependent Promoter Escape and Pause Release of RNA Polymerase II during the Cell Cycle. Cell Reports *21*, 3483–3497.

Busslinger, G.A., Stocsits, R.R., van der Lelij, P., Axelsson, E., Tedeschi, A., Galjart, N., and Peters, J.-M. (2017). Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl. Nature *544*, 503–507.

Campbell, K.J., and White, R.J. (2014). MYC regulation of cell growth through control of transcription by RNA polymerases I and III. Cold Spring Harb Perspect Med *4*.

Carrière, L., Graziani, S., Alibert, O., Ghavi-Helm, Y., Boussouar, F., Humbertclaude, H., Jounier, S., Aude, J.-C., Keime, C., Murvai, J., et al. (2012). Genomic binding of Pol III transcription machinery and relationship with TFIIS transcription factor distribution in mouse embryonic stem cells. Nucleic Acids Res. *40*, 270–283.

Cavalli, G., and Misteli, T. (2013). Functional implications of genome topology. Nat. Struct. Mol. Biol. *20*, 290–299.

Charron, J., Malynn, B.A., Fisher, P., Stewart, V., Jeannotte, L., Goff, S.P., Robertson, E.J., and Alt, F.W. (1992). Embryonic lethality in mice homozygous for a targeted disruption of the N-myc gene. Genes Dev. *6*, 2248–2257.

Chen, H., Liu, H., and Qing, G. (2018). Targeting oncogenic Myc as a strategy for cancer treatment. Signal Transduct Target Ther *3*, 5.

Chen, K., Hu, J., Moore, D.L., Liu, R., Kessans, S.A., Breslin, K., Lucet, I.S., Keniry, A., Leong, H.S., Parish, C.L., et al. (2015). Genome-wide binding and mechanistic analyses of Smchd1-mediated epigenetic regulation. Proc Natl Acad Sci U S A *112*, E3535-44.

Choi, S.H., Wright, J.B., Gerber, S.A., and Cole, M.D. (2010). Myc protein is stabilized by suppression of a novel E3 ligase complex in cancer cells. Genes Dev. *24*, 1236–1241.

Ciesla, M., Skowronek, E., and Boguta, M. (2018). Function of TFIIIC, RNA polymerase III initiation factor, in activation and repression of tRNA gene transcription. Nucleic Acids Res.

Cinatl, J., Cinatl, J., Mainke, M., Weissflog, A., Rabenau, H., Kornhuber, B., and Doerr, H.W. (1993). In vitro differentiation of human neuroblastoma cells induced by sodium phenylacetate. Cancer Lett. *70*, 15–24.

Conacci-Sorrell, M., McFerrin, L., and Eisenman, R.N. (2014). An overview of MYC and its interactome. Cold Spring Harb Perspect Med *4*, a014357.

Cong, L., and Zhang, F. (2015). Genome engineering using CRISPR-Cas9 system. Methods Mol Biol *1239*, 197–217.

Conti, A., Carnevali, D., Bollati, V., Fustinoni, S., Pellegrini, M., and Dieci, G. (2015). Identification of RNA polymerase III-transcribed *Alu* loci by computational screening of RNA-Seq data. Nucleic Acids Res. *43*, 817–835.

Cremer, T., and Cremer, M. (2010). Chromosome Territories. Cold Spring Harbor Perspectives in Biology *2*, a003889–a003889.

Crepaldi, L., Policarpi, C., Coatti, A., Sherlock, W.T., Jongbloets, B.C., Down, T.A., and Riccio, A. (2013). Binding of TFIIIC to SINE Elements Controls the Relocation of Activity-Dependent Neuronal Genes to Transcription Factories. PLoS Genetics *9*, e1003699.

Cubeñas-Potts, C., and Corces, V.G. (2015). Architectural proteins, transcription, and the three-dimensional organization of the genome. FEBS Letters *589*, 2923–2930.

Cubeñas-Potts, C., Rowley, M.J., Lyu, X., Li, G., Lei, E.P., and Corces, V.G. (2017). Different enhancer classes in Drosophila bind distinct architectural proteins and mediate unique chromatin interactions and 3D architecture. Nucleic Acids Res. *45*, 1714–1730.

Dalla-Favera, R., Bregni, M., Erikson, J., Patterson, D., Gallo, R.C., and Croce, C.M. (1982). Human c-myc onc gene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells. Proc. Natl. Acad. Sci. U.S.A. *79*, 7824–7827.

D'Ambrosio, C., Schmidt, C.K., Katou, Y., Kelly, G., Itoh, T., Shirahige, K., and Uhlmann, F. (2008). Identification of cis-acting sites for condensin loading onto budding yeast chromosomes. Genes Dev. *22*, 2215–2227.

Davis, A.C., Wims, M., Spotts, G.D., Hann, S.R., and Bradley, A. (1993). A null c-myc mutation causes lethality before 10.5 days of gestation in homozygotes and reduced fertility in heterozygous female mice. Genes Dev. *7*, 671–682.

Deininger, P. (2011). *Alu* elements: know the SINEs. Genome Biol. *12*, 236.

Dekker, J., Belmont, A.S., Guttman, M., Leshyk, V.O., Lis, J.T., Lomvardas, S., Mirny, L.A., O'Shea, C.C., Park, P.J., Ren, B., et al. (2017). The 4D nucleome project. Nature *549*, 219–226.

Dieci, G., Percudani, R., Giuliodori, S., Bottarelli, L., and Ottonello, S. (2000). TFIIIC-independent in vitro transcription of yeast tRNA genes. J. Mol. Biol. *299*, 601–613.

Dieci, G., Conti, A., Pagano, A., and Carnevali, D. (2013). Identification of RNA polymerase III-transcribed genes in eukaryotic genomes. Biochim Biophys Acta *1829*, 296–305.

Dileep, V., Ay, F., Sima, J., Vera, D.L., Noble, W.S., and Gilbert, D.M. (2015). Topologically associating domains and their long-range contacts are established during early G1 coincident with the establishment of the replication-timing program. Genome Research *25*, 1104–1113.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376–380.

Donze, D. (2012). Extra-transcriptional functions of RNA Polymerase III complexes: TFIIIC as a potential global chromatin bookmark. Gene *493*, 169–175.

Doudna, J.A., and Charpentier, E. (2014). Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science *346*, 1258096.

Dubois, N.C., Adolphe, C., Ehninger, A., Wang, R.A., Robertson, E.J., and Trumpp, A. (2008). Placental rescue reveals a sole requirement for c-Myc in embryonic erythroblast survival and hematopoietic stem cell function. Development *135*, 2455–2465.

DuBois, S.G., Mosse, Y.P., Fox, E., Kudgus, R.A., Reid, J.M., McGovern, R., Groshen, S., Bagatell, R., Maris, J.M., Twist, C.J., et al. (2018). Phase II Trial of Alisertib in Combination with Irinotecan and Temozolomide for Patients with Relapsed or Refractory Neuroblastoma. Clin. Cancer Res. *24*, 6142–6149.

Dumay-Odelot, H., Marck, C., Durrieu-Gaillard, S., Lefebvre, O., Jourdain, S., Prochazkova, M., Pflieger, A., and Teichmann, M. (2007). Identification, Molecular Cloning, and Characterization of the Sixth Subunit of Human Transcription Factor TFIIIC. Journal of Biological Chemistry *282*, 17179–17189.

Duttke, S.H. (2014). RNA polymerase III accurately initiates transcription from RNA polymerase II promoters in vitro. J Biol Chem *289*, 20396–20404.

Ea, V., Baudement, M.-O., Lesne, A., and Forné, T. (2015). Contribution of Topological Domains and Loop Formation to 3D Chromatin Organization. Genes *6*, 734–750.

Eberhardy, S.R., and Farnham, P.J. (2002). Myc recruits P-TEFb to mediate the final step in the transcriptional activation of the cad promoter. J. Biol. Chem. *277*, 40156–40162.

Eilers, M., and Eisenman, R.N. (2008). Myc's broad reach. Genes Dev. *22*, 2755–2766.

Facchini, L.M., and Penn, L.Z. (1998). The molecular role of Myc in growth and transformation: recent discoveries lead to new insights. FASEB J. *12*, 633–651.

Faure, A.J., Schmidt, D., Watt, S., Schwalie, P.C., Wilson, M.D., Xu, H., Ramsay, R.G., Odom, D.T., and Flicek, P. (2012). Cohesin regulates tissue-specific expression by stabilizing highly occupied cis-regulatory modules. Genome Res. *22*, 2163–2175.

Fay, A., Misulovin, Z., Li, J., Schaaf, C.A., Gause, M., Gilmour, D.S., and Dorsett, D. (2011). Cohesin selectively binds and regulates genes with paused RNA polymerase. Curr. Biol. *21*, 1624–1634.

Felgenhauer, J., Tomino, L., Selich-Anderson, J., Bopp, E., and Shah, N. (2018). Dual BRD4 and AURKA inhibition is synergistic against MYCN-amplified and nonamplified neuroblastoma. Neoplasia *20*, 965-974.

Fellmann, C., Hoffmann, T., Sridhar, V., Hopfgartner, B., Muhar, M., Roth, M., Lai, D.Y., Barbosa, I.A.M., Kwon, J.S., Guan, Y., et al. (2013). An Optimized microRNA Backbone for Effective Single-Copy RNAi. Cell Reports *5*, 1704–1713.

Felsenfeld, G., Burgess-Beusse, B., Farrell, C., Gaszner, M., Ghirlando, R., Huang, S., Jin, C., Litt, M., Magdinier, F., Mutskov, V., et al. (2004). Chromatin Boundaries and Chromatin Domains. Cold Spring Harbor Symposia on Quantitative Biology *69*, 245–250.

Ferrari, R., Rivetti, C., Acker, J., and Dieci, G. (2004). Distinct roles of transcription factors TFIIIB and TFIIIC in RNA polymerase III transcription reinitiation. Proc. Natl. Acad. Sci. U.S.A. *101*, 13442–13447.

Frank, S.R., Parisi, T., Taubert, S., Fernandez, P., Fuchs, M., Chan, H.-M., Livingston, D.M., and Amati, B. (2003). MYC recruits the TIP60 histone acetyltransferase complex to chromatin. EMBO Rep. *4*, 575–580.

Galli, G.G., Carrara, M., Francavilla, C., de Lichtenberg, K.H., Olsen, J.V., Calogero, R.A., and Lund, A.H. (2013). Genomic and proteomic analyses of Prdm5 reveal interactions with insulator binding proteins in embryonic stem cells. Mol Cell Biol *33*, 4504–4516.

Ganji, M., Shaltiel, I.A., Bisht, S., Kim, E., Kalichava, A., Haering, C.H., and Dekker, C. (2018). Real-time imaging of DNA loop extrusion by condensin. Science *360*, 102–105.

Gartel, A.L., Ye, X., Goufman, E., Shianov, P., Hay, N., Najmabadi, F., and Tyner, A.L. (2001). Myc represses the p21(WAF1/CIP1) promoter and interacts with Sp1/Sp3. Proc. Natl. Acad. Sci. U.S.A. *98*, 4510–4515.

Gartenberg, M.R., and Merkenschlager, M. (2008). Condensin goes with the family but not with the flow. Genome Biol. *9*, 236.

George, C.M., Bozler, J., Nguyen, H.Q., and Bosco, G. (2014). Condensins are Required for Maintenance of Nuclear Architecture. Cells *3*, 865–882.

Ghavi-Helm, Y., Michaut, M., Acker, J., Aude, J.-C., Thuriaux, P., Werner, M., and Soutourina, J. (2008). Genome-wide location analysis reveals a role of TFIIS in RNA polymerase III transcription. Genes Dev. *22*, 1934–1947.

Gomez-Diaz, E., and Corces, V.G. (2014). Architectural proteins: regulators of 3D genome organization in cell fate. Trends Cell Biol *24*, 703–711.

Gomez-Roman, N., Grandori, C., Eisenman, R.N., and White, R.J. (2003). Direct activation of RNA polymerase III transcription by c-Myc. Nature *421*, 290–294.

Grandori, C., Cowley, S.M., James, L.P., and Eisenman, R.N. (2000). The Myc/Max/Mad network and the transcriptional control of cell behavior. Annu. Rev. Cell Dev. Biol. *16*, 653–699.

Grandori, C., Gomez-Roman, N., Felton-Edkins, Z.A., Ngouenet, C., Galloway, D.A., Eisenman, R.N., and White, R.J. (2005). c-Myc binds to human ribosomal DNA and stimulates transcription of rRNA genes by RNA polymerase I. Nat. Cell Biol. *7*, 311–318.

Gregory, M.A., and Hann, S.R. (2000). c-Myc proteolysis by the ubiquitin-proteasome pathway: stabilization of c-Myc in Burkitt's lymphoma cells. Mol. Cell. Biol. *20*, 2423–2435.

Guccione, E., Martinato, F., Finocchiaro, G., Luzi, L., Tizzoni, L., Dall' Olio, V., Zardo, G., Nervi, C., Bernard, L., and Amati, B. (2006). Myc-binding-site recognition in the human genome is determined by chromatin context. Nat. Cell Biol. *8*, 764–770.

Haeusler, R.A., Pratt-Hyatt, M., Good, P.D., Gipson, T.A., and Engelke, D.R. (2008). Clustering of yeast tRNA genes is mediated by specific association of condensin with tRNA gene transcription complexes. Genes Dev. *22*, 2204–2214.

Han, S., Wada, R.K., and Sidell, N. (2001). Differentiation of Human Neuroblastoma by Phenylacetate Is Mediated by Peroxisome Proliferator-activated Receptor γ. Cancer Res *61*, 3998.

Hann, S.R., and Eisenman, R.N. (1984). Proteins encoded by the human c-myc oncogene: differential expression in neoplastic cells. Mol. Cell. Biol. *4*, 2486–2497.

Hart, T., Chandrashekhar, M., Aregger, M., Steinhart, Z., Brown, K.R., MacLeod, G., Mis, M., Zimmermann, M., Fradet-Turcotte, A., Sun, S., et al. (2015). High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities. Cell *163*, 1515–1526.

Hatton, B.A., Knoepfler, P.S., Kenney, A.M., Rowitch, D.H., de Alborán, I.M., Olson, J.M., and Eisenman, R.N. (2006). N-myc is an essential downstream effector of Shh signaling during both normal and neoplastic cerebellar growth. Cancer Res. *66*, 8655–8661.

He, T.C., Sparks, A.B., Rago, C., Hermeking, H., Zawel, L., da Costa, L.T., Morin, P.J., Vogelstein, B., and Kinzler, K.W. (1998). Identification of c-MYC as a target of the APC pathway. Science *281*, 1509–1512.

Herbst, A., Salghetti, S.E., Kim, S.Y., and Tansey, W.P. (2004). Multiple cell-type-specific elements regulate Myc protein stability. Oncogene *23*, 3863–3871.

Hiraga, S., Botsios, S., Donze, D., and Donaldson, A.D. (2012). TFIIIC localizes budding yeast ETC sites to the nuclear periphery. Mol. Biol. Cell *23*, 2741–2754.

Houck, C.M., Rinehart, F.P., and Schmid, C.W. (1979). A ubiquitous family of repeated DNA sequences in the human genome. J. Mol. Biol. *132*, 289–306.

Hsieh, Y.J., Kundu, T.K., Wang, Z., Kovelman, R., and Roeder, R.G. (1999a). The TFIIIC90 subunit of TFIIIC interacts with multiple components of the RNA polymerase III machinery and contains a histone-specific acetyltransferase activity. Mol. Cell. Biol. *19*, 7697–7704.

Hsieh, Y.J., Wang, Z., Kovelman, R., and Roeder, R.G. (1999b). Cloning and characterization of two evolutionarily conserved subunits (TFIIIC102 and TFIIIC63) of human TFIIIC and

their involvement in functional interactions with TFIIIB and RNA polymerase III. Mol. Cell. Biol. *19*, 4944–4952.

Huang, M., and Weiss, W.A. (2013). Neuroblastoma and MYCN. Cold Spring Harbor Perspectives in Medicine *3*, a014415–a014415.

Huang, Y., and Maraia, R.J. (2001). Comparison of the RNA polymerase III transcription machinery in Schizosaccharomyces pombe, Saccharomyces cerevisiae and human. Nucleic Acids Res. *29*, 2675–2690.

Huang, Y., Hamada, M., and Maraia, R.J. (2000). Isolation and cloning of four subunits of a fission yeast TFIIIC complex that includes an ortholog of the human regulatory protein TFIIICbeta. J. Biol. Chem. *275*, 31480–31487.

Ikegaki, N., Bukovsky, J., and Kennett, R.H. (1986). Identification and characterization of the NMYC gene product in human neuroblastoma cells by monoclonal antibodies with defined specificities. Proc. Natl. Acad. Sci. U.S.A. *83*, 5929–5933.

Ingvarsson, S., Sundaresan, S., Jin, P., Francke, U., Asker, C., Sümegi, J., Klein, G., and Sejersen, T. (1988). Chromosome localization and expression pattern of Lmyc and Bmyc in murine embryonal carcinoma cells. Oncogene *3*, 679–685.

Iwasaki, O., Tanizawa, H., Kim, K.-D., Yokoyama, Y., Corcoran, C.J., Tanaka, A., Skordalakes, E., Showe, L.C., and Noma, K.-I. (2015). Interaction between TBP and Condensin Drives the Organization and Faithful Segregation of Mitotic Chromosomes. Mol. Cell *59*, 755–767.

Jaenicke, L.A., von Eyss, B., Carstensen, A., Wolf, E., Xu, W., Greifenberg, A.K., Geyer, M., Eilers, M., and Popov, N. (2016). Ubiquitin-Dependent Turnover of MYC Antagonizes MYC/PAF1C Complex Accumulation to Drive Transcriptional Elongation. Mol. Cell *61*, 54–67.

Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S., et al. (2010). Mediator and cohesin connect gene expression and chromatin architecture. Nature *467*, 430–435.

Kakui, Y., Rabinowitz, A., Barry, D.J., and Uhlmann, F. (2017). Condensin-mediated remodeling of the mitotic chromatin landscape in fission yeast. Nature Genetics *49*, 1553–1557.

Kalkat, M., Resetca, D., Lourenco, C., Chan, P.-K., Wei, Y., Shiah, Y.-J., Vitkin, N., Tong, Y., Sunnerhagen, M., Done, S.J., et al. (2018). MYC Protein Interactome Profiling Reveals Functionally Distinct Regions that Cooperate to Drive Tumorigenesis. Mol. Cell *72*, 836-848.e7.

Kantidakis, T., Ramsbottom, B.A., Birch, J.L., Dowding, S.N., and White, R.J. (2010). mTOR associates with TFIIIC, is found at tRNA and 5S rRNA genes, and targets their repressor Maf1. Proc. Natl. Acad. Sci. U.S.A. *107*, 11823–11828.

Kenney, A.M., Cole, M.D., and Rowitch, D.H. (2003). Nmyc upregulation by sonic hedgehog signaling promotes proliferation in developing cerebellar granule neuron precursors. Development *130*, 15–28.

Kieffer-Kwon, K.-R., Nimura, K., Rao, S.S.P., Xu, J., Jung, S., Pekowska, A., Dose, M., Stevens, E., Mathe, E., Dong, P., et al. (2017). Myc Regulates Chromatin Decompaction and Nuclear Architecture during B Cell Activation. Molecular Cell *67*, 566-578.e10.

Kim, T.H., Abdullaev, Z.K., Smith, A.D., Ching, K.A., Loukinov, D.I., Green, R.D., Zhang, M.Q., Lobanenkov, V.V., and Ren, B. (2007). Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. Cell *128*, 1231–1245.

Knoepfler, P.S., Cheng, P.F., and Eisenman, R.N. (2002). N-myc is essential during neurogenesis for the rapid expansion of progenitor cell populations and the inhibition of neuronal differentiation. Genes Dev. *16*, 2699–2712.

Koch, H.B., Zhang, R., Verdoodt, B., Bailey, A., Zhang, C.-D., Yates, J.R., Menssen, A., and Hermeking, H. (2007). Large-scale identification of c-MYC-associated proteins using a combined TAP/MudPIT approach. Cell Cycle *6*, 205–217.

Kohl, N.E., Kanda, N., Schreck, R.R., Bruns, G., Latt, S.A., Gilbert, F., and Alt, F.W. (1983). Transposition and amplification of oncogene-related sequences in human neuroblastomas. Cell *35*, 359–367.

Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., Ziller, M.J., et al. (2015). Integrative analysis of 111 reference human epigenomes. Nature *518*, 317–330.

Kundu, T.K., Wang, Z., and Roeder, R.G. (1999). Human TFIIIC relieves chromatin-mediated repression of RNA polymerase III transcription and contains an intrinsic histone acetyltransferase activity. Mol. Cell. Biol. *19*, 1605–1615.

Kurland, J.F., and Tansey, W.P. (2008). Myc-mediated transcriptional repression by recruitment of histone deacetylase. Cancer Res. *68*, 3624–3629.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860–921.

Lazar-Stefanita, L., Scolari, V.F., Mercy, G., Muller, H., Guérin, T.M., Thierry, A., Mozziconacci, J., and Koszul, R. (2017). Cohesins and condensins orchestrate the 4D dynamics of yeast chromosomes during the cell cycle. EMBO J. *36*, 2684–2697.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.

Li, L., Lyu, X., Hou, C., Takenaka, N., Nguyen, H.Q., Ong, C.-T., Cubeñas-Potts, C., Hu, M., Lei, E.P., Bosco, G., et al. (2015a). Widespread rearrangement of 3D chromatin organization underlies polycomb-mediated stress-induced silencing. Mol. Cell *58*, 216–231.

Li, W., Hu, Y., Oh, S., Ma, Q., Merkurjev, D., Song, X., Zhou, X., Liu, Z., Tanasa, B., He, X., et al. (2015b). Condensin I and II Complexes License Full Estrogen Receptor α-Dependent Enhancer Activation. Mol. Cell *59*, 188–202.

Li, Y., Choi, P.S., Casey, S.C., Dill, D.L., and Felsher, D.W. (2014). MYC through miR-17-92 suppresses specific target genes to maintain survival, autonomous proliferation, and a neoplastic state. Cancer Cell *26*, 262–272.

Lin, C.Y., Lovén, J., Rahl, P.B., Paranal, R.M., Burge, C.B., Bradner, J.E., Lee, T.I., and Young, R.A. (2012). Transcriptional Amplification in Tumor Cells with Elevated c-Myc. Cell *151*, 56–67.

Liu, C., Li, S., Dai, X., Ma, J., Wan, J., Jiang, H., Wang, P., Liu, Z., and Zhang, H. (2015). PRC2 regulates RNA polymerase III transcribed non-translated RNA gene transcription through EZH2 and SUZ12 interaction with TFIIIC complex. Nucleic Acids Res. *43*, 6270–6284.

Losada, A., and Hirano, T. (2005). Dynamic molecular linkers of the genome: the first decade of SMC proteins. Genes Dev. *19*, 1269–1287.

Lunyak, V.V., and Atallah, M. (2011). Genomic relationship between SINE retrotransposons, Pol III-Pol II transcription, and chromatin organization: the journey from junk to jewel. Biochem Cell Biol *89*, 495–504.

Male, G., von Appen, A., Glatt, S., Taylor, N.M.I., Cristovao, M., Groetsch, H., Beck, M., and Müller, C.W. (2015). Architecture of TFIIIC and its role in RNA polymerase III pre-initiation complex assembly. Nat Commun *6*, 7387.

Malynn, B.A., de Alboran, I.M., O'Hagan, R.C., Bronson, R., Davidson, L., DePinho, R.A., and Alt, F.W. (2000). N-myc can functionally replace c-myc in murine development, cellular growth, and differentiation. Genes Dev. *14*, 1390–1399.

Marzouki, N., Camier, S., Ruet, A., Moenne, A., and Sentenac, A. (1986). Selective proteolysis defines two DNA binding domains in yeast transcription factor tau. Nature *323*, 176–178.

Matsumoto, M., Akiyama, T., Miyatake, S., Oda, Y., Kikuchi, H., Hanaoka, M., and Namba, Y. (1989). Expression of proto-oncogene products during drug-induced differentiation of a neuroblastoma cell line SK-N-DZ. Acta Neuropathol. *79*, 217–221.

Matsushima, H., and Bogenmann, E. (1993). Expression of trkA cDNA in neuroblastomas mediates differentiation in vitro and in vivo. Mol. Cell. Biol. *13*, 7447–7456.

Matthay, K.K., Reynolds, C.P., Seeger, R.C., Shimada, H., Adkins, E.S., Haas-Kogan, D., Gerbing, R.B., London, W.B., and Villablanca, J.G. (2009). Long-term results for children with high-risk neuroblastoma treated on a randomized trial of myeloablative therapy followed by 13-cis-retinoic acid: a children's oncology group study. J. Clin. Oncol. *27*, 1007–1013.

McMahon, S.B., Wood, M.A., and Cole, M.D. (2000). The essential cofactor TRRAP recruits the histone acetyltransferase hGCN5 to c-Myc. Mol. Cell. Biol. *20*, 556–562.

Meerbrey, K.L., Hu, G., Kessler, J.D., Roarty, K., Li, M.Z., Fang, J.E., Herschkowitz, J.I., Burrows, A.E., Ciccia, A., Sun, T., et al. (2011). The pINDUCER lentiviral toolkit for inducible RNA interference in vitro and in vivo. Proc. Natl. Acad. Sci. U.S.A. *108*, 3665–3670.

Moqtaderi, Z., and Struhl, K. (2004). Genome-wide occupancy profile of the RNA polymerase III machinery in Saccharomyces cerevisiae reveals loci with incomplete transcription complexes. Mol. Cell. Biol. *24*, 4118–4127.

Moqtaderi, Z., Wang, J., Raha, D., White, R.J., Snyder, M., Weng, Z., and Struhl, K. (2010). Genomic binding profiles of functionally distinct RNA polymerase III transcription complexes in human cells. Nat Struct Mol Biol *17*, 635–640.

Nagano, T., Lubling, Y., Várnai, C., Dudley, C., Leung, W., Baran, Y., Mendelson Cohen, N., Wingett, S., Fraser, P., and Tanay, A. (2017). Cell-cycle dynamics of chromosomal organization at single-cell resolution. Nature *547*, 61–67.

Nakagawara, A., and Brodeur, G.M. (1997). Role of neurotrophins and their receptors in human neuroblastomas: a primary culture study. Eur. J. Cancer *33*, 2050–2053.

Nakagawara, A., Arima-Nakagawara, M., Scavarda, N.J., Azar, C.G., Cantor, A.B., and Brodeur, G.M. (1993). Association between high levels of expression of the TRK gene and favorable outcome in human neuroblastoma. N. Engl. J. Med. *328*, 847–854.

Nakazawa, N., Nakamura, T., Kokubu, A., Ebe, M., Nagao, K., and Yanagida, M. (2008). Dissection of the essential steps for condensin accumulation at kinetochores and rDNAs during fission yeast mitosis. The Journal of Cell Biology *180*, 1115–1131.

Nau, M.M., Brooks, B.J., Battey, J., Sausville, E., Gazdar, A.F., Kirsch, I.R., McBride, O.W., Bertness, V., Hollis, G.F., and Minna, J.D. (1985). L-myc, a new myc-related gene amplified and expressed in human small cell lung cancer. Nature *318*, 69–73.

Naumova, N., Imakaev, M., Fudenberg, G., Zhan, Y., Lajoie, B.R., Mirny, L.A., and Dekker, J. (2013). Organization of the Mitotic Chromosome. Science *342*, 948–953.

Nesbit, C.E., Tersak, J.M., and Prochownik, E.V. (1999). MYC oncogenes and human neoplastic disease. Oncogene *18*, 3004–3016.

Newman, A.J., Ogden, R.C., and Abelson, J. (1983). tRNA gene transcription in yeast: effects of specified base substitutions in the intragenic promoter. Cell *35*, 117–125.

Nicol, J.W., Helt, G.A., Blanchard, S.G., Raja, A., and Loraine, A.E. (2009). The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. Bioinformatics *25*, 2730–2731.

Nikiforov, M.A., Chandriani, S., Park, J., Kotenko, I., Matheos, D., Johnsson, A., McMahon, S.B., and Cole, M.D. (2002). TRRAP-dependent and TRRAP-independent transcriptional activation by Myc family oncoproteins. Mol. Cell. Biol. *22*, 5054–5063.

Nishimura, K., Fukagawa, T., Takisawa, H., Kakimoto, T., and Kanemaki, M. (2009). An auxin-based degron system for the rapid depletion of proteins in nonplant cells. Nat. Methods *6*, 917–922.

Nitzsche, A., Paszkowski-Rogacz, M., Matarese, F., Janssen-Megens, E.M., Hubner, N.C., Schulz, H., de Vries, I., Ding, L., Huebner, N., Mann, M., et al. (2011). RAD21 cooperates with pluripotency transcription factors in the maintenance of embryonic stem cell identity. PLoS ONE *6*, e19470.

Nora, E.P., Goloborodko, A., Valton, A.-L., Gibcus, J.H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L.A., and Bruneau, B.G. (2017). Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. Cell *169*, 930-944.e22.

Oler, A.J., Alla, R.K., Roberts, D.N., Wong, A., Hollenhorst, P.C., Chandler, K.J., Cassiday, P.A., Nelson, C.A., Hagedorn, C.H., Graves, B.J., et al. (2010). Human RNA polymerase III transcriptomes and relationships to Pol II promoter chromatin and enhancer-binding factors. Nat Struct Mol Biol *17*, 620–628.

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

Otto, T., Horn, S., Brockmann, M., Eilers, U., Schüttrumpf, L., Popov, N., Kenney, A.M., Schulte, J.H., Beijersbergen, R., Christiansen, H., et al. (2009). Stabilization of N-Myc Is a Critical Function of Aurora A in Human Neuroblastoma. Cancer Cell *15*, 67–78.

Palomero, T., Lim, W.K., Odom, D.T., Sulis, M.L., Real, P.J., Margolin, A., Barnes, K.C., O'Neil, J., Neuberg, D., Weng, A.P., et al. (2006). NOTCH1 directly regulates c-MYC and activates a feed-forward-loop transcriptional network promoting leukemic cell growth. Proc. Natl. Acad. Sci. U.S.A. *103*, 18261–18266.

Peukert, K. (1997). An alternative pathway for gene regulation by Myc. The EMBO Journal *16*, 5672–5686.

Policarpi, C., Crepaldi, L., Brookes, E., Nitarska, J., French, S.M., Coatti, A., and Riccio, A. (2017). Enhancer SINEs Link Pol III to Pol II Transcription in Neurons. Cell Reports *21*, 2879–2894.

Pontén, F., Jirström, K., and Uhlen, M. (2008). The Human Protein Atlas--a tool for pathology. J. Pathol. *216*, 387–393.

Popov, N., Schulein, C., Jaenicke, L.A., and Eilers, M. (2010). Ubiquitylation of the amino terminus of Myc by SCF(beta-TrCP) antagonizes SCF(Fbw7)-mediated turnover. Nat Cell Biol *12*, 973–981.

Pugh, T.J., Morozova, O., Attiyeh, E.F., Asgharzadeh, S., Wei, J.S., Auclair, D., Carter, S.L., Cibulskis, K., Hanna, M., Kiezun, A., et al. (2013). The genetic landscape of high-risk neuroblastoma. Nature Genetics *45*, 279–284.

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

Raha, D., Wang, Z., Moqtaderi, Z., Wu, L., Zhong, G., Gerstein, M., Struhl, K., and Snyder, M. (2010). Close association of RNA polymerase II and many transcription factors with Pol III genes. Proc Natl Acad Sci U S A *107*, 3639–3644.

Rahl, P.B., Lin, C.Y., Seila, A.C., Flynn, R.A., McCuine, S., Burge, C.B., Sharp, P.A., and Young, R.A. (2010). c-Myc regulates transcriptional pause release. Cell *141*, 432–445.
Ramani, V., Deng, X., Qiu, R., Gunderson, K.L., Steemers, F.J., Disteche, C.M., Noble, W.S., Duan, Z., and Shendure, J. (2017). Massively multiplex single-cell Hi-C. Nat. Methods *14*, 263–266.

Richards, M.W., Burgess, S.G., Poon, E., Carstensen, A., Eilers, M., Chesler, L., and Bayliss, R. (2016). Structural basis of N-Myc binding by Aurora-A and its destabilization by kinase inhibitors. Proc Natl Acad Sci U S A *113*, 13726–13731.

Rickman, D.S., Schulte, J.H., and Eilers, M. (2018). The Expanding World of N-MYC–Driven Tumors. Cancer Discovery *8*, 150–163.

Ruiz-Velasco, M., and Zaugg, J.B. (2017). Structure meets function: How chromatin organisation conveys functionality. Current Opinion in Systems Biology *1*, 129–136.

Sanborn, A.L., Rao, S.S.P., Huang, S.-C., Durand, N.C., Huntley, M.H., Jewett, A.I., Bochkov, I.D., Chinnappan, D., Cutkosky, A., Li, J., et al. (2015). Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. Proc. Natl. Acad. Sci. U.S.A. *112*, E6456-6465.

Sanjana, N.E., Shalem, O., and Zhang, F. (2014). Improved vectors and genome-wide libraries for CRISPR screening. Nat Methods *11*, 783–784.

Sawai, S., Shimono, A., Wakamatsu, Y., Palmes, C., Hanaoka, K., and Kondoh, H. (1993). Defects of embryonic organogenesis resulting from targeted disruption of the N-myc gene in the mouse. Development *117*, 1445–1455.

Schaaf, C.A., Kwak, H., Koenig, A., Misulovin, Z., Gohara, D.W., Watson, A., Zhou, Y., Lis, J.T., and Dorsett, D. (2013). Genome-wide control of RNA polymerase II activity by cohesin. PLoS Genet *9*, e1003382.

Schalbetter, S.A., Goloborodko, A., Fudenberg, G., Belton, J.-M., Miles, C., Yu, M., Dekker, J., Mirny, L., and Baxter, J. (2017). SMC complexes differentially compact mitotic chromosomes according to genomic context. Nat. Cell Biol. *19*, 1071–1080.

Schmidt, D., Schwalie, P.C., Ross-Innes, C.S., Hurtado, A., Brown, G.D., Carroll, J.S., Flicek, P., and Odom, D.T. (2010). A CTCF-independent role for cohesin in tissue-specific transcription. Genome Res. *20*, 578–588.

Schwab, M., Alitalo, K., Klempnauer, K.H., Varmus, H.E., Bishop, J.M., Gilbert, F., Brodeur, G., Goldstein, M., and Trent, J. (1983). Amplified DNA with limited homology to myc cellular oncogene is shared by human neuroblastoma cell lines and a neuroblastoma tumour. Nature *305*, 245–248.

Sears, R., Nuckolls, F., Haura, E., Taya, Y., Tamai, K., and Nevins, J.R. (2000). Multiple Ras-dependent phosphorylation pathways regulate Myc protein stability. Genes Dev. *14*, 2501–2514.

Seeger, R.C., Brodeur, G.M., Sather, H., Dalton, A., Siegel, S.E., Wong, K.Y., and Hammond, D. (1985). Association of multiple copies of the N-myc oncogene with rapid progression of neuroblastomas. N. Engl. J. Med. *313*, 1111–1116.

Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelsen, T.S., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G., et al. (2014). Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. Science *343*, 84–87.

Shandilya, J., and Roberts, S.G.E. (2012). The transcription cycle in eukaryotes: from productive initiation to RNA polymerase II recycling. Biochim. Biophys. Acta *1819*, 391–400.

Sharma, V.M., Calvo, J.A., Draheim, K.M., Cunningham, L.A., Hermance, N., Beverly, L., Krishnamoorthy, V., Bhasin, M., Capobianco, A.J., and Kelliher, M.A. (2006). Notch1 contributes to mouse T-cell leukemia by directly inducing the expression of c-myc. Mol. Cell. Biol. *26*, 8022–8031.

Sheiness, D., and Bishop, J.M. (1979). DNA and RNA from uninfected vertebrate cells contain nucleotide sequences related to the putative transforming gene of avian myelocytomatosis virus. J. Virol. *31*, 514–521.

Shukla, S., Kavak, E., Gregory, M., Imashimizu, M., Shutinoski, B., Kashlev, M., Oberdoerffer, P., Sandberg, R., and Oberdoerffer, S. (2011). CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. Nature *479*, 74–79.

Smith, S.T., Wickramasinghe, P., Olson, A., Loukinov, D., Lin, L., Deng, J., Xiong, Y., Rux, J., Sachidanandam, R., Sun, H., et al. (2009). Genome wide ChIP-chip analyses reveal important roles for CTCF in Drosophila genome organization. Dev. Biol. *328*, 518–528.

Souleimani, A., and Asselin, C. (1993). Regulation of c-myc expression by sodium butyrate in the colon carcinoma cell line Caco-2. FEBS Lett. *326*, 45–50.

Staller, P., Peukert, K., Kiermaier, A., Seoane, J., Lukas, J., Karsunky, H., Möröy, T., Bartek, J., Massagué, J., Hänel, F., et al. (2001). Repression of p15INK4b expression by Myc through association with Miz-1. Nat. Cell Biol. *3*, 392–399.

Steiger, D., Furrer, M., Schwinkendorf, D., and Gallant, P. (2008). Max-independent functions of Myc in Drosophila melanogaster. Nat. Genet. *40*, 1084–1091.

Stine, Z.E., Walton, Z.E., Altman, B.J., Hsieh, A.L., and Dang, C.V. (2015). MYC, Metabolism, and Cancer. Cancer Discovery *5*, 1024–1039.

Su, Y., Pelz, C., Huang, T., Torkenczy, K., Wang, X., Cherry, A., Daniel, C.J., Liang, J., Nan, X., Dai, M.-S., et al. (2018). 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 Dev. *32*, 1398–1419.

Sun, Y., Zhang, H., Kazemian, M., Troy, J.M., Seward, C., Lu, X., and Stubbs, L. (2016). ZSCAN5B and primate-specific paralogs bind RNA polymerase III genes and extra-TFIIIC (ETC) sites to modulate mitotic progression. Oncotarget *7*, 72571–72592.

Takahashi, K., and Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. Cell *126*, 663–676.

Tashiro, S., and Lanctôt, C. (2015). The International Nucleome Consortium. Nucleus *6*, 89–92.

Taub, R., Kirsch, I., Morton, C., Lenoir, G., Swan, D., Tronick, S., Aaronson, S., and Leder, P. (1982). Translocation of the c-myc gene into the immunoglobulin heavy chain locus in human Burkitt lymphoma and murine plasmacytoma cells. Proc. Natl. Acad. Sci. U.S.A. *79*, 7837–7841.

The ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.

Thomas, L.R., Wang, Q., Grieb, B.C., Phan, J., Foshage, A.M., Sun, Q., Olejniczak, E.T., Clark, T., Dey, S., Lorey, S., et al. (2015). Interaction with WDR5 promotes target gene recognition and tumorigenesis by MYC. Mol. Cell *58*, 440–452.

Valenzuela, L., Dhillon, N., and Kamakaka, R.T. (2009). Transcription independent insulation at TFIIIC-dependent insulators. Genetics *183*, 131–148.

Van Bortle, K., Nichols, M.H., Li, L., Ong, C.-T., Takenaka, N., Qin, Z.S., and Corces, V.G. (2014). Insulator function and topological domain border strength scale with architectural protein occupancy. Genome Biol. *15*, R82.

Vassetzky, N.S., and Kramerov, D.A. (2013). SINEBase: a database and tool for SINE analysis. Nucleic Acids Res. *41*, D83-89.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A., et al. (2001). The Sequence of the Human Genome. Science *291*, 1304–1351.

Vervoorts, J., Lüscher-Firzlaff, J.M., Rottmann, S., Lilischkis, R., Walsemann, G., Dohmann, K., Austen, M., and Lüscher, B. (2003). Stimulation of c-MYC transcriptional activity and acetylation by recruitment of the cofactor CBP. EMBO Rep. *4*, 484–490.

Walz, S., Lorenzin, F., Morton, J., Wiese, K.E., von Eyss, B., Herold, S., Rycak, L., Dumay-Odelot, H., Karim, S., Bartkuhn, M., et al. (2014). Activation and repression by oncogenic MYC shape tumour-specific gene expression profiles. Nature *511*, 483–487.

Wang, H., Maurano, M.T., Qu, H., Varley, K.E., Gertz, J., Pauli, F., Lee, K., Canfield, T., Weaver, M., Sandstrom, R., et al. (2012). Widespread plasticity in CTCF occupancy linked to DNA methylation. Genome Res. *22*, 1680–1688.

Weng, A.P., Millholland, J.M., Yashiro-Ohtani, Y., Arcangeli, M.L., Lau, A., Wai, C., Del Bianco, C., Rodriguez, C.G., Sai, H., Tobias, J., et al. (2006). c-Myc is an important direct

target of Notch1 in T-cell acute lymphoblastic leukemia/lymphoma. Genes Dev. *20*, 2096–2109.

White, R.J. (2011). Transcription by RNA polymerase III: more complex than we thought. Nat Rev Genet *12*, 459–463.

Wiese, K.E., Haikala, H.M., von Eyss, B., Wolf, E., Esnault, C., Rosenwald, A., Treisman, R., Klefstrom, J., and Eilers, M. (2015). Repression of SRF target genes is critical for Myc-dependent apoptosis of epithelial cells. The EMBO Journal *34*, 1554–1571.

Wood, M.A., McMahon, S.B., and Cole, M.D. (2000). An ATPase/helicase complex is an essential cofactor for oncogenic transformation by c-Myc. Mol. Cell *5*, 321–330.

Wright, J.B., Brown, S.J., and Cole, M.D. (2010). Upregulation of c-MYC in cis through a Large Chromatin Loop Linked to a Cancer Risk-Associated Single-Nucleotide Polymorphism in Colorectal Cancer Cells. Molecular and Cellular Biology *30*, 1411–1420.

Yuen, K.C., and Gerton, J.L. (2018). Taking cohesin and condensin in context. PLoS Genet. *14*, e1007118.

Yuen, K.C., Slaughter, B.D., and Gerton, J.L. (2017). Condensin II is anchored by TFIIIC and H3K4me3 in the mammalian genome and supports the expression of active dense gene clusters. Science Advances *3*, e1700191.

Zimmerman, K.A., Yancopoulos, G.D., Collum, R.G., Smith, R.K., Kohl, N.E., Denis, K.A., Nau, M.M., Witte, O.N., Toran-Allerand, D., and Gee, C.E. (1986). Differential expression of myc family genes during murine development. Nature *319*, 780–783.

Zuin, J., Dixon, J.R., van der Reijden, M.I.J.A., Ye, Z., Kolovos, P., Brouwer, R.W.W., van de Corput, M.P.C., van de Werken, H.J.G., Knoch, T.A., van IJcken, W.F.J., et al. (2014). Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. Proceedings of the National Academy of Sciences *111*, 996–1001.

7 Appendix

7.1 Abbreviations

Prefixes

p Pico- (10^{-12})
n Nano- (10^{-9})
µ Micro- (10^{-6})
m Milli- (10^{-3})
c Centi- (10^{-2})
k Kilo- (10^{3})

Units

A ampere
Da dalton
g gram
h hour
l liter
m meter
M mol/l
min minute
s second
U unit
v/v volume per volume
w/v weight per volume
°C degree celsius

Proteins, protein domains and other biomolecules

A alanine
ATP adenosine-triphosphate
aa amino acid
bp basepair(s)
cDNA complementary DNA
CDK cyclin-dependent kinase
CTD C-terminal domain
CRISPR/Cas9 clustered regularly interspaced short palindromic repeats/CRISPR associated protein 9
DNA deoxyribonucleic acid
dNTPs deoxyribonucleoside-5'-triphosphate (dATP, dCTG, dGTP, dTTP)
MB Myc-box
MIZ1 E3 SUMO-protein ligase PIAS2; old name: Msx-interacting zinc finger protein
mRNA messenger RNA
RNAP RNA polymerase
pT58/S62 phosphorylated aa T58/S62

RNA ribonucleic acid
S serine
shRNA short hairpin RNA
SINE short interspersed nuclear element
T threonine
TF transcription factor
tRNA transfer RNA

Chemicals and solutions

APS ammoniumpersulfate
aqua dest destilled water
H_2O water
DMEM Dulbeccos Modified Eagle-Medium
DMSO dimethylsulfoxide
DTT dithiothreitol
EDTA ethylendiamintetraacetate
FBS fetal bovine serum
PBS phosphate-buffered saline
PVDF polyvinylidenfluoride
SDS sodium dodecyl sulfate
TAE Tris-acetate-EDTA
TBS Tris-buffered saline
TBS-T Tris-buffered saline with Tween-20
TE Tris-EDTA-buffer
TEMED N,N,N',N'-tetramethylethylendiamine
TES transcriptional end site
Tris Tris-(hydroxymethyl)-aminomethan
TSS transcriptional start site

Other abbreviations

ChIP Chromatin immunoprecipitation
Co-IP Co-precipitation of a protein interacting with the precipitated one
ECL enhanced chemoluminescence
E. coli Escherichia coli
G1/G2 phase Gap 1/2 phase of the cell cycle
Fig. Figure
Inc. Incorporated
IP immunoprecipitation
M phase mitotic phase of the cell cycle
nt nucleotides
PAGE polyacrylamide-gelelectrophoresis
PCR polymerase chain reaction
qPCR quantitative PCR
qRT-PCR quantitative real time PCR
rpm rotations per minute
RT room temperature
S1 security level 1
S2 security level 1
S phase synthetic phase of the cell cycle
O/N over night; 16-20 h
wt wild type

7.2 Acknowledgements

First of all, I would like to express my gratitude to my primary supervisor, Prof. Martin Eilers for offering me the opportunity to join his team and supervising this PhD thesis.
I would also like to thank Prof. Manfred Gessler and Dr. Sonja Lorenz for their kind support and commitment as the thesis committee.
My heart-felt thank also goes to the Graduate School of Life Sciences for their help throughout my PhD study and for offering professional training.
This thesis is completed with the assistance and advice of many helpful members of the department. I am truly thankful to Dr. Anne Carstensen for initiating the TFIIIC project and her support especially during the beginning stage of my study here. I am particularly grateful to Dr. Susanne Walz who offered me generous help on bioinformatic analysis of all ChIP-seq experiments. I would also like to thank Andreas Kurz, the PhD student from Prof. Markus Sauer's lab for the dSTORM images.
I am honored to have met inspiring, smart and aspired people along the path of my PhD: especially Cristian for sharing both the good and bad times with me and giving me honest advice; Guan-Nan and Sana for not only impressing me with their talents but also enlightening me with their attitudes towards hurdles.
Special thanks to the words of Prof. Robert Peters, of which backed me up over these four and a half years.
My final note of acknowledgement goes to my parents and my dear friends Christopher, Alden, Cherry, Nga-Man and Chong-Ka. Thank you for fueling me with energy, letting me fly high and always being there whenever I need.

7.3 Publications

Büchel, G.*, Carstensen, A.*, ***Mak, K.****, Roeschert, I, Leen, E., Sumara, O., Hofstetter, J., Herold, S., Baluapuri, A., Poon, E., Kwok, C., Chesler, L., Maric, H.M., Rickman, D.S., Wolf, E., Bayliss, R., Walz, S. & Eilers, M. Association with Aurora-A controls N-MYC-dependent promoter escape and pause release of RNA polymerase II during the cell cycle. Cell Reports, vol.21, issue 12, (2017).

* co-first authorship

7.4 Curriculum Vitae

7.5 Affidavit

I hereby confirm that my thesis entitled “TFIIIC subunits employ different modes of action for regulating N-MYC” is the result of my own work. I did not receive any help or support from commercial consultants. All sources and/or materials applied are listed and specified in the thesis.

Furthermore, I confirm that this thesis has not yet been submitted as part of another examination process neither in identical nor in similar form.

_____ _____

Place, Date Signature