Regulation of RNA Polymerase III Transcription by DNA Methylation and Chromatin

By

Dhaval Varshney

Thesis submitted for the degree of Doctor of Philosophy

UNIVERSITY *of* GLASGOW

October 2011

Institute of Cancer Sciences
University of Glasgow

Summary

A million copies of short interspersed nuclear elements (SINEs) comprise 11% of the human genome, but despite this abundance, they are expressed at a very low level. Alu elements are estimated to harbour 33% of CpG sites in the human genome and DNA methylation of these sites is believed to silence their transcription. However, this study shows the presence of RNA polymerase III (Pol III) transcription machinery on methylated SINEs alongside methylated DNA-binding proteins. Methylation of A- and B-block elements on the Pol III promoter is unable to inhibit Pol III loading at SINEs. Loss of DNA methylation in DNA methyl transferase1-null fibroblasts or following 5-azacytidine treatment neither changes expression levels nor the occupancy of the Pol III machinery on SINEs. H3K9-trimethylation, along with SUV39H1 and associated heterochromatin protein 1 (HP1), is found to be enriched at SINEs. Treatment with chaetocin, a SUV39H1 H3K9-methyltransferase inhibitor, elevates Pol III loading and expression of SINEs. Thus, H3K9 methylation, and not DNA methylation, is responsible for SINE transcriptional inhibition.

This suggests an alternative role for DNA methylation on these repetitive sequences. Homologous recombination events between Alu elements are implicated in several human diseases, including some cancers. Reduction in DNA methylation causes a two-fold increase in the rate of Alu-driven inter-chromosomal translocation events. Hence, DNA methylation serves not to inhibit transcription of SINEs but instead it plays a role in inhibiting homologous recombination between these elements.

Chromatin can pose a physical barrier to transcription by limiting polymerase accessibility to DNA. Studies showing effects of chromatin modifications caused by c-MYC and of remodeling by various complexes, clearly highlight the significance of chromatin in the regulation of Pol III transcription. Recently published ChIP-sequencing data also indicate that tRNA genes may be subject to the same chromatin-mediated transcriptional control as is seen for Pol II-transcribed genes. SWI/SNF is an evolutionarily conserved ATP-dependent chromatin remodeling complex. SNF5, a core SWI/SNF subunit, is a bona fide tumour suppressor and is commonly lost or mutated in malignant rhabdoid tumours.

SNF5, along with SWI/SNF ATP-hydrolysis subunits, BRG1 and BRM, is found to occupy Pol III-transcribed loci. SNF5 knockdown or a BRG1/BRM dual-knockdown leads to a significant increase in tRNA expression. However, a corresponding increase in Pol III transcription apparatus occupancy at tRNA genes is not observed. c-Myc null rat fibroblasts have elevated levels of SWI/SNF enrichment at tRNA genes compared to c-Myc wild-type cells, suggesting that Pol III induction by c-MYC evicts the SWI/SNF complex. Thus, SWI/SNF chromatin remodeling complex represses Pol III transcription and this repression may have to be overcome in order to increase Pol III transcriptional output. SWI/SNF subunits are lost and Pol III transcription is elevated in many human cancers, suggesting that these findings may be of clinical significance.

The mechanistic details of this repression are still unclear. SWI/SNF subunits co-immunoprecipitate with TFIIIC subunits, suggesting that Pol III transcriptional machinery is responsible for its recruitment. However, SWI/SNF occupancy at tRNA genes is highly correlated with Pol III enrichment, indicating that the mode of repression is not through inhibition of polymerase loading. SNF5 knockdown leads to higher levels of Pol III enrichment downstream of a tRNA locus. Thus, SWI/SNF may spatially limit the area available to Pol III transcriptional apparatus and, therefore, inhibit transcriptional elongation, termination or facilitated recycling.

Table of Contents

List of Tables

Table of Figures

Abbreviations

5mC	5-methylcytosine
5mCpG	5-methyl CpG dinucleotide
Adomet	S-adenosylmethionine
Ago	Argonaute
Apo-E	Apolipoprotein-E
ARPs	Actin-related proteins
BAF	Brahma associated factor
Bdp1	B double prime 1
BER	Base excision repair
BRF1	TFIIB-related factor1
BRG1	Brahma-related gene 1
BRM	Brahma
CBP	CREB-binding protein
Cdks	Cyclin-dependent kinases
CGIs	CpG islands
ChIP	Chromatin Immunoprecipitation
CIN	Chromosomal instability
CoIP	Co-immunoprecipitation
CTCF	CCCTC-binding factor
DMSO	Dimethylsulphoxide
Dnmt	DNA methyltransferase
DSB	Double strand break
dsDNA	Double-stranded DNA
DSE	Distal sequence element
dsRNA	Double-stranded RNA
ECL	Enhanced chemiluminescence
ES cells	Embryonic Stem cells
FACT	Facilitates chromatin transcription
GNAT	Gcn5-related N-acetyltransferase
H3K27me3	Histone H3 Lysine 27 trimethylation
H3K9me3	Histone H3 Lysine 9 trimethylation
HATs	Histone acetyl-transferases
HDACs	Histone deacetylases
HIV1	Human immunodeficiency virus-type 1
HKMT	Histone lysine methyltransferase
hmC	5-hydroxymethylcytosine
HP1	Heterochromatin Protein 1
HR	Homologous repair
IAP	Intercisternal A particle
ICF	immunodeficiency, centromeric instability and facial anomalies syndrome
ICR	Internal control region
IE	Intermediate element
Ini1	Integrase interactor 1
IP	Immunoprecipitation
K9MT	Histone H3 lysine 9 (K9) methyl transferase
LINE	Long interspersed nuclear element
LSD1	Lysine-specific demethylase 1
MBD	Methyl CpG-binding domain
MBP	Methylated DNA-binding protein
MBT	Mid-blastula transition

MEB	Microextraction buffer
meCpG	Methyl-CpG
MEFs	Mouse embryonic fibroblasts
MLL	Mixed lineage leukaemia
MNase	Micrococcal nuclease
mRNA	Messenger RNA
MRP RNA	Mitochondrial RNA processing RNA
MRT	Malignant rhabdoid tumour
NCoR-1	Nuclear receptor corepressor-1
NHEJ	Non-homologous end joining
NLS	Nuclear localisation signal
NUMAC	Nucleosomal methylation activated complex
PBS	Phosphate-buffered saline
PCG	Polycomb group
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PIC	Preinitiation complex
PKR	dsRNA regulated protein kinase
Plk1	Polo-like kinase 1
Pol I	RNA polymerase I
Pol II	RNA polymerase II
Pol III	RNA polymerase III
Pol IV	RNA polymerase IV
Pol V	RNA polymerase V
PSE	Proximal sequence element
qPCR	Quantitative polymerase chain reaction
RDRC	RNA-directed RNA polymerase complex
RE	Recognition element
RITS	RNA-induced transcriptional gene silencing
RNP	Ribonucleoproteins
RPA	Replication protein A
rRNA	Ribosomal RNA
RSC	remodels the structure of chromatin
sc	Scrambled
scRNA	Small cytoplasmic RNA
SINE	Short interspersed nuclear element
$SNAP_C$	snRNA activator protein complex
SNF	Sucrose non-fermenting
snoRNA	Small nucleolar RNA
snRNA	Small nuclear RNA
snRNP	Small nuclear riboprotein
SRP	Signal recognition particle
SSA	Single strand annealing
ssDNA	Single-stranded DNA
ssRNA	Single-stranded RNA
TBP	TATA-binding protein
TF	Transcription factor
TGD	Thymine DNA glycosylase
TPRT	Target-primed reverse transcription
TRD	Transcriptional repression domain
tRNA	Transfer RNA
TSA	Trichostatin A
TSS	Transcription start site

Declaration

I hereby declare that the thesis which follows is my own composition, that it is a record of he work done myself, and that is has not been presented in any previous application for a higher degree.

Dhaval Varshney

Acknowledgments

First and foremost I would like to thank my supervisor Prof. Robert J. White for providing me with this great opportunity and also for his generous support and guidance throughout my PhD. Without his help this project would not have been possible.

I would like to dedicate this work to my family and thank them for their neverending support in every endeavour of my life. My special thanks to Desiree Piscitello for her support during the final challenging months of my PhD.

I would also like to thank every past and present member of White Lab for their help in the lab and also for making the lab a wonderfull place to work in. I would like to especially thank Dr. Louise Mitchell for her help throughout my PhD and for her delightful friendship.

I would also like to thank CRUK for the financial support and the Beatson for providing the fantastic facilities. I would like to thank Dr. Andrew Oler and Prof. Bradle Cairns for the ChIP-bis-seq data; Dr. Patricia Sung and Dr. Maria Jasin for help with the Alu translocation assay. My special thanks to Prof. David Gillespie for his invaluable input on the project and to everone else who contributed time and resources to the project.

Chapter 1 Introduction

1.1 Transcription by RNA polymerase III

The central dogma of molecular biology was first articulated by Francis Crick, where he described how the hereditary DNA information within a cell is recognised by RNA polymerases and transcribed into complementary RNA molecules. These RNA molecules are further recognised by the cellular apparatus that translates them to produce proteins. Thus the transcription of DNA to RNA is an essential part of the central dogma (Crick, 1970). In eukaryotes, the task of transcribing nuclear genes is divided between three highly-related enzymes, RNA polymerase I, II and III (hereon referred to as Pol I, Pol II and Pol III) (Roeder and Rutter, 1969). Derivatives of Pol II that transcribe short interfering RNAs (Pol IV) or of noncoding RNAs (Pol V) have also been identified in Arabidopsis (Ream et al., 2009). An RNA polymerase of the mitochondrial origin, named snRNAP-IV, has also been found to transcribe mRNA of certain mammalian protein coding genes in the nucleus (Kravchenko et al., 2005).

Each of the three classical RNA polymerases is devoted to the transcription of specific genes. Pol I distinctively transcribes only one type of gene, the large tandemly repeated, ribosomal RNA (rRNA) genes. The untranslated RNA products of Pol I transcription are essential components of the cellular translation apparatus [reviewed in (McStay and Grummt, 2008; Russell and Zomerdijk, 2006)]. Pol II transcribes genes encoding messenger RNAs (mRNAs) that are used as templates for protein synthesis. Pol II is also responsible for transcription of most small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs) and micro RNAs [reviewed in (Baumann et al., 2010)]. The genes transcribed by Pol III are classified as class III genes. These encode a diverse group of small untranslated RNAs that are involved in various cellular processes like transcription, RNA splicing and translation. Pol III-transcribed genes include the 5S rRNA and tRNA genes (White, 2001).

Gene transcription is subject to a high degree of regulation that allows the cell to constantly adjust its RNA and protein content in response to environmental changes and metabolic requirements (White, 2001). Pol III-transcribed RNAs are important contributors to cellular protein synthesis and are thus considered crucial for the regulation of cell growth and proliferation. The deregulation of

Pol III transcription is observed in human diseases like cancer and is, therefore, an area of active research (White, 2004).

1.1.1 Class III genes

With 17 subunits, Pol III is the largest eukaryotic nuclear RNA polymerase, and is responsible for 10% of all nuclear transcription. Pol III transcription gives rise to small untranslated RNAs, which are usually shorter than 400nt in length (White, 2002). Some of the Pol III-transcribed RNAs are described below.

1.1.1.1 5S ribosomal RNA (rRNA)

Eukaryotic ribosomes are composed of two unequal subunits, which are made up of four RNA molecules (28S, 5.8S, 5S and 18S) and approximately 80 protein subunits. The four rRNAs are required in equal stoichiometry, one molecule of each per ribosome (Phillips and McConkey, 1976). 5S rRNA is produced by Pol III, whereas the other rRNAs are transcribed by Pol I. 5S rRNA is approximately 120nt long and is found associated with the large ribosomal subunit. Eukaryotes contain multiple copies of the 5S rRNA genes, for example the human genome contains 200 to 300 5S genes, many of which are found arranged in tandem arrays (Lander et al., 2001).

1.1.1.2 Transfer RNAs (tRNAs)

Transfer RNAs are 70-90nt long adaptor molecules, which facilitate the translation of mRNA molecules. Each tRNA recognises a specific three nucleotide codon on the mRNA and translates that codon to a specific amino acid (Crick, 1968). Pol III transcribes pre-tRNA molecules which, when processed, adopt an L-shaped secondary structure. The 3' end of the tRNA comprises the amino acid attachment region, whereas the opposite end contains a three nucleotide anticodon that recognises and binds to the appropriate triplet codon on the mRNA (Hopper and Phizicky, 2003). There are 506 tRNA genes and 110 predicted pseudogenes annotated in the hg19 human genome database (Chan and Lowe, 2009). Different tRNAs recognising the same amino acid but varying triplet codons are termed ‘isoacceptors’ and each isoacceptor group is named after the amino acid it recognises. The distribution of tRNA genes is mainly random

throughout the genome; however some clustering is observed (Lander et al., 2001).

1.1.1.3 U6 small nuclear RNA (snRNA)

Small nuclear riboproteins (snRNPs) are structurally-related RNA-protein complexes found within eukaryotic nuclei. U6 snRNA, together with the Pol II-transcribed U1, U2, U4 and U5 snRNAs, is an essential part of the most abundant snRNP called the spliceosome. This multisubunit ribonucleoprotein is essential for pre-mRNA splicing. There are multiple copies of U6 snRNA genes in the human genome, which show a high degree of evolutionary conservation (Valadkhan, 2005).

1.1.1.4 7SL RNA and short interspersed nuclear elements

The 7SL RNA forms the structural backbone of the signal recognition particle (SRP) and provides the scaffold for its six protein components. SRP is responsible for the transport of nascent polypeptide chains to the endoplasmic reticulum. In eukaryotes, 7SL RNA contains a small Alu domain and a large S-domain, which are separated by a long linker region (Batey et al., 2000). The Alu domain, which binds SRP9/14, retards the ribosomal elongation of the peptide before its association with the endoplasmic reticulum. The S-domain binds the signal peptide and targets it to the ER membrane-bound SRP receptor (Mason et al., 2000). The 7SL RNA is ~300nt long and is highly conserved through evolution. Some SINEs are 7SL derived repeat elements found within mammalian genomes and will be discussed in section 1.5.

1.1.1.5 Other class III genes encoding ribonucleoprotein components

Various ribonucleoproteins (RNPs) contain Pol III-transcribed RNAs as their RNA components, for example 7SK, H1 and mitochondrial RNA processing (MRP) RNA. 7SK is an abundant 330nt long Pol III-transcribed RNA that has only been identified in vertebrates to date. 7SK functions as a negative regulator of Pol II transcription by binding to the elongation factor P-TEFb and repressing the phosphorylation of Pol II CTD (Yang et al., 2001). H1 RNA is a component of RNase P endoribonuclease that processes the 5' termini of pre-tRNAs

(Bartkiewicz et al., 1989), whereas the structurally related MRP RNase is mainly involved in pre-RNA processing (Gold et al., 1989).

Thus, RNA Pol III transcribes multiple RNA species, which together contribute to the regulation of global protein synthesis. In addition to the above mentioned RNAs, Pol III is also found to transcribe less well known RNAs such as vault RNAs (van Zon et al., 2003), Y RNAs (Deutscher et al., 1988) and the neural specific BC1/BC200 RNAs (Cao et al., 2006). Pol III transcription is also involved in production of various viral RNAs such as the adenoviral VA and Epstein-Barr viral EBER RNAs (Rosa et al., 1981). The size of the Pol III transcriptome has expanded considerably in recent years and it now also includes snoRNAs, microRNAs, stem-bulge RNAs and other unclassified non-coding RNAs (Dieci et al., 2007).

1.1.2 Class III gene promoters and assembly of the transcription complex

The selective transcription of genes by Pol I, Pol II and Pol III is dictated by the identification of distinct promoters by their specific transcription factors.

1.1.2.1 Class III gene promoters

The majority of class III gene promoters comprise of regulatory elements that are within the transcribed region, downstream of the transcription start site (TSS). These promoter elements, known as internal control regions (ICRs), were identified by mutation analysis in yeast and are found to be present as discontinuous conserved sequences separated by non-essential regions (Koski et al., 1980; Kurjan and Hall, 1982). Three different types of promoters are identified by Pol III transcription apparatus, type I and type II promoters that are found within the gene, and type III promoters that are located entirely upstream of the TSS [reviewed in (Schramm and Hernandez, 2002)](See Figure 1.1).

Type I promoters are unique to 5S rRNA genes and consist of three internal elements; an A-block (+50 to +64), an intermediate element (+67 to +72) and a C-block (+80 to +92). These promoter sequences are highly conserved between different species and mutations within the A- and C-block sequences are detrimental to transcription (Pieler et al., 1985). A change in the spacing

between these elements also leads to diminished transcription (Pieler et al., 1987).

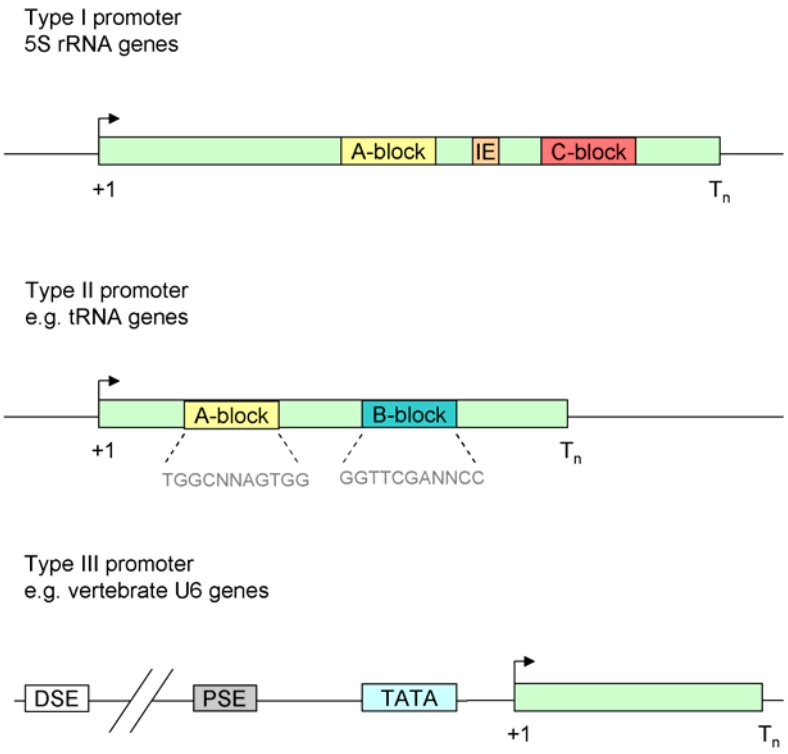


Figure 1.1 Schematic diagram of general types of Pol III-transcribed promoters
The TSS is indicated by +1 and the site of transcription termination is indicated by T_n. The approximate positions of various promoter elements are depicted including the intermediate element (IE), proximal sequence element (PSE) and distal sequence element (DSE). The known consensus sequences for A-block and B-block elements are also shown (Galli et al., 1981) [adapted from (Schramm and Hernandez, 2002)]

The most common Pol III promoters are found within tRNA genes and are termed type II promoters. This promoter class consists of two highly conserved A-block and B-block sequences that are separated by a variable distance. Interblock separation of ~30 to 60bp is thought to be optimal for transcription, however a distance as great as 365bp can be tolerated (Baker et al., 1987; Fabrizio et al., 1987). The A-block sequences of the type I and II promoters are homologous, but in the latter case they are much closer to the TSS (Ciliberto et al., 1983).

In contrast to type I and II promoters, type III promoters are external and lie upstream of the transcribed region. The best characterised type III promoter belongs to the human U6 gene where the regulatory sequences include a TATA box (-30 to -25), a proximal sequence element (PSE, -66 to -47) and a distal

sequence element (DSE, -244 to -214) (Das et al., 1988). The U6 PSE and DSE are homologous and interchangeable with elements found upstream of the Pol II-transcribed U2 gene. Surprisingly, it is the presence of the TATA box (generally a feature of Pol II-transcribed genes) that defines U6 as a Pol III-transcribed locus (Lobo and Hernandez, 1989). Other genes containing type III promoters include the 7SK RNA, H1 RNA and MRP RNA genes (White, 2002).

Some promoters recognised by the Pol III machinery cannot be classified as any of the conventional promoter types. For instance, the EBER2 gene requires all the elements of a type II promoter as well as those of a type III promoter for its transcription by Pol III (Howe and Shu, 1989). The optimal transcription of the 7SL gene requires the internal A- and B-box elements alongside a 37bp sequence found upstream of the TSS (Ullu and Weiner, 1985).

1.1.2.2 Preinitiation complex formation

Transcription by an RNA polymerase is preceded by the formation of a preinitiation complex (PIC). Different Pol III promoter types require a different cocktail of transcription factors for PIC formation [reviewed in (Schramm and Hernandez, 2002) and (White, 2002)]. The core transcription apparatus required for Pol III transcription is conserved from yeast to mammals, however certain transcription factor subunits in mammals have diverged considerably from their yeast counterparts (Huang and Maraia, 2001). The PIC formation of various promoter types is discussed below and depicted in Figure 1.2.

Type II promoters found at tRNA genes require a large multi-subunit transcription factor complex called TFIIIC. Human TFIIIC, with an aggregate mass of more than 500kDa (Geiduschek and Kassavetis, 2001), comprises of TFIIIC220, TFIIIC110, TFIIIC102, TFIIIC90, TFIIIC63 subunits (Kovelman and Roeder, 1992; Yoshinaga et al., 1989) and the recently identified TFIIIC35 subunit (Dumay-Odelot et al., 2007). The primary binding affinity of TFIIIC to the Pol III promoter is determined by the binding of TFIIIC220 to the B-block element. However, this alone is not sufficient for the recruitment of TFIIIC, which also requires IIIC220 to interact with TFIIIC110 (Shen et al., 1996; Yoshinaga et al., 1987).

TFIIIC63 recognises the A-block sequence and contributes to the recruitment of TFIIIC (Hsieh et al., 1999b), however the B-block binding is the major determinant of promoter efficiency (Baker et al., 1986). TFIIIC90 forms a flexible linker between the two TFIIIC sub-complexes (TFIIIC63/102 and TFIIIC110/220) and forms what appears as a dumbbell shaped transcription factor under the electron microscope (Schultz et al., 1989). This structure and the flexible central linker allow the recruitment of TFIIIC to promoters with variable distances between the A- and B-blocks. On promoters with large inter-block separation, the intervening DNA can be looped out to allow TFIIIC binding (Baker et al., 1987). Six TFIIIC subunits have been identified in yeast and five of them were found to be conserved in humans. The human homologue to the sixth yeast TFIIIC subunit, TFIIIC35, has been recently identified. Even though TFIIIC35 shows considerable sequence divergence from its yeast conterpart (*Tfc7),* the interactions with other TFIIIC subunits are conserved from yeast to man. TFIIIC35 is found to interact with TFIIIC63 and, to a lesser extent, with TFIIIC90 *in vitro.* TFIIIC35 also localises within active transcription complexes (Dumay-Odelot et al., 2007).

Once recruited, TFIIIC further recruits another transcription factor TFIIIB, which comprises of the TATA-binding protein (TBP), TFIIB-related factor 1 (BRF1) and B double prime 1 (Bdp1) (Schramm and Hernandez, 2002). BRF1 forms a stable complex with TBP (Kassavetis et al., 1991; Khoo et al., 1994), whereas Bdp1 interacts very weakly with this complex in the absence of a DNA template (Huet et al., 1994). An N-terminal TFIIB-like domain of BRF1 interacts with TBP, although the primary binding site was recognised to lie within the C-terminal part of BRF1 (Khoo et al., 1994). In addition, BRF1 is also found to interact with TFIIIC102 (Kassavetis et al., 1992b), which is thought to be the primary point of contact of TFIIIB and TFIIIC (Schramm and Hernandez, 2002). Further interactions of BRF1 with TFIIIC63 and TFIIIC90, alongside the binding of TBP to TFIIIC63 and TFIIIC102 contribute to the recruitment of TFIIIB to the TFIIIC-bound promoters (Hsieh et al., 1999a; Hsieh et al., 1999b).

The recruitment of TFIIIB is followed by the recruitment of Pol III via protein-protein interactions. TBP and BRF1 interact with multiple Pol III subunits and these interactions are crucial for the recruitment of the polymerase. The role of Bdp1 in the recruitment of Pol III is less defined, however it contains a SANT

domain which is essential for TFIIIC-dependent, but not for TFIIIC-independent Pol III transcription on type III promoters (Schramm et al., 2000). Moreover, TFIIIC subunits are also capable of interacting with the polymerase, which contributes to the polymerase recruitment and stabilises the PIC (Schramm and Hernandez, 2002).

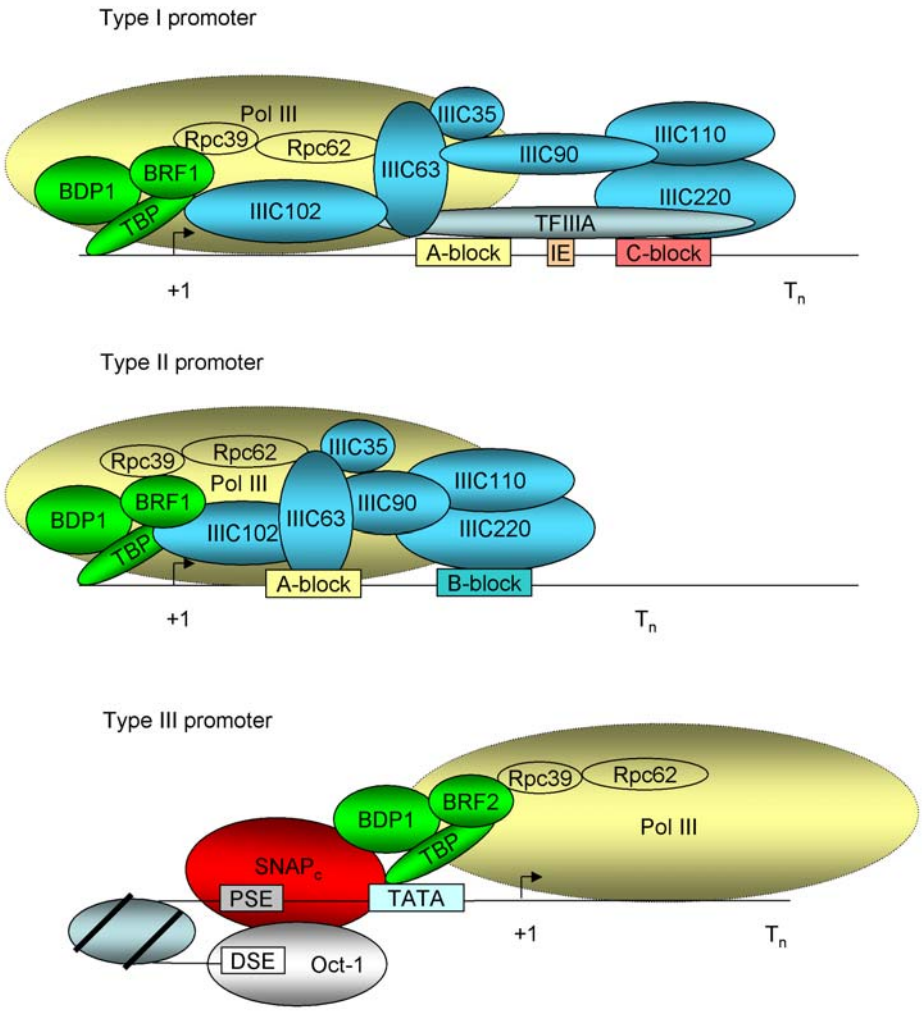


Figure 1.2: Transcription complex assembly on class III gene promoters
The promoter elements are recognised by the transcription factors shown associated with them. All TFIIIC subunits are depicted in blue, TFIIIB subunits in green and Pol III subunits in yellow. [Figue adapted from (Dumay-Odelot et al., 2010))

5S rRNA genes lack the B-block sequence and, therefore, cannot be recognised by TFIIIC. TFIIIC recruitment on type I promoters is dependent on a 40kDa adapter transcription factor called TFIIIA, which was the first transcription factor to be purified to homogeneity (Engelke et al., 1980). This single polypeptide contains nine zinc finger domains which bind to the A-block, IE and

the C-block promoter elements (Miller et al., 1985). Similar to B-block binding by TFIIIC on type II promoters, the C-block binding by TFIIIA on type I promoters determines the overall binding affinity (Hanas et al., 1983). It is not completely clear how TFIIIC binds TFIIIA, but together they form a complex that recruits TFIIIB and consequently Pol III to the 5S rRNA gene promoters (Bieker et al., 1985).

The PIC formation at type III promoters requires a different set of transcription factors. The PSE is identified by snRNA activator protein complex ($SNAP_c$) and the TATA box is recognised by TBP (Schramm et al., 2000). Both $SNAP_c$ and TBP can bind weakly to these promoter elements; however co-operative binding greatly enhances their promoter affinity (Mittal and Hernandez, 1997). TBP bound to the TATA box is found within an alternative form of TFIIIB that contains BRF2 instead of BRF1 (Schramm et al., 2000). The DSE is recognised by Oct-1, which also binds and promotes $SNAP_c$ promoter occupancy, thereby stimulating transcription (Mittal et al., 1996). However, Oct-1 is not essential for basal *in vitro* transcription (Hu et al., 2003). Given the distance between the two elements, a positioned nucleosome may be required to juxtapose PSE and DSE in order to facilitate Oct-1 and $SNAP_c$ interaction (Stunkel et al., 1997). The PIC formation is completed by the recruitment of Pol III to the promoter by $SNAP_c$ and TFIIIB (Schramm and Hernandez, 2002).

1.1.3 Pol III: transcription initiation, elongation and termination

With an aggregate mass of 600-700 kDa, Pol III is the largest eukaryotic RNA polymerase. Pol III consists of 17 subunits in humans and yeast, all of which are found to be essential for its function and for cell viability (Geiduschek and Kassavetis, 2001). Five of the subunits are shared by all three classical RNA polymerases, whereas an additional two are shared between Pol I and Pol III. Ten of the seventeen subunits are unique to Pol III (Breant et al., 1983; Buhler et al., 1980). The Pol III-specific subunits are thought to be responsible for its nuclear localisation and binding specificity for Pol III-specific transcription factors, as well as for its elongation and termination properties.

Following the successful formation of PIC, the DNA at the transcription initiation site is melted by Pol III, forming a transcription bubble. TFIIIB is responsible for

correctly positioning the polymerase at the TSS (Kassavetis et al., 1990). Moreover, mutations of specific domains in BRF1 and Bdp1 can impair the transcription initiation, even though the polymerase recruitment is unaffected. This suggests that TFIIIB may also be involved in transcription bubble formation. The presence of A/T rich sequences found around the TSS is thought to promote the formation of the transcription bubble, due to the lower thermodynamic stability of A:T bonds (Kassavetis et al., 1992a). The recruitment of Pol III and the formation of the transcription bubble are the rate-limiting steps for Pol III transcription, whereas the synthesis of an initial 17bp transcript takes mere seconds (White, 2002).

Once transcription is initiated, the 14bp long transcription bubble moves along the gene as transcription proceeds. Pol III dissociates from TFIIIB without a considerable delay in elongation (Kassavetis et al., 1992a). Moreover, the presence of TFIIIC does not seem to significantly delay the progress of Pol III either. It is presently unclear how Pol III passes through DNA that is bound by TFIIIC; however some experiments propose the transient dissociation of TFIIIC to allow the polymerase to pass (Bardeleben et al., 1994). Unlike the other polymerases, Pol III does not require specific elongation factors, possibly due to the small size of the Pol III-transcribed genes (White, 2002). Pol III transcription proceeds until it reaches a stretch of four thymidine residues, where the transcription terminates (Galli et al., 1981). It has been proposed that the La autoantigen and NF1 polypeptides are required for efficient termination of Pol III transcription (Maraia, 2001).

Once transcription is terminated, Pol III does not dissociate from the template, but is recycled to the TSS. The repeat rounds of transcription by the recycled polymerase occur much more rapidly than the initial round (Dieci and Sentenac, 1996). The La protein is also thought to promote polymerase recycling (Maraia et al., 1994). Thus once assembled, the Pol III transcription complexes can produce multiple transcripts with remarkable efficiency due to polymerase recycling.

1.1.4 Regulation of Pol III transcription

The synthesis of rRNA (including 5S rRNA) is an important step in ribosome production and consequently protein translation. Moreover, the cellular tRNA

levels also affect the rate of protein translation. Thus, Pol III transcriptional products are vital for cellular protein synthesis and, thus, cell growth. Therefore, Pol III transcription within a cell is subject to tight regulation (White, 2002).

The first evidence for the regulation of Pol III transcription in eukaryotic cells came from the 5S rRNA genes in *Xenopus* oocytes. 5S rRNA is highly expressed in the oocyte, whereas its expression is greatly diminished following oogenesis. The *Xenopus laevis* genome contains 21300 copies of oocyte-specific 5S rRNA genes and 400 copies of somatic 5S rRNA genes. The oocyte genes are only expressed during oogenesis, whereas the somatic genes are expressed both before and after oogenesis. This differential regulation of these two gene types is in part due to their different affinities to TFIIIA. Following oogenesis, a considerable drop in the TFIIIA protein levels is observed, and the limited amount of TFIIIA binds the high affinity somatic 5S genes (Ginsberg et al., 1984; White, 2002). Initial studies also found Pol III transcription to be regulated by various viral proteins including the adenoviral E1A protein (Berger and Folk, 1985) and the SV40 small τ antigen (Loeken et al., 1988).

The regulation of Pol III transcription is now known to occur in three primary ways: i) changes in transcription factor availability (Figure 1.3a), ii) post-translational modification (primarily phosphorylation) of transcription factors (Figure 1.3b), and iii) regulation by chromatin. The first two modes of regulation are discussed briefly below, whereas the third mode of regulation is discussed in chapter 6.

1.1.4.1 Changes in transcription factor levels

The first example for this mode of transcriptional regulation was obtained from studies investigating the E1A viral protein. The infection of cells by E1A was found to upregulate the cellular TFIIIC110 protein levels and cause increased Pol III transcription (Sinn et al., 1995). The hepatitis B virus X protein was also shown to upregulate Pol III transcription by causing an increase in TBP protein levels (Wang et al., 1995). Cell lines transformed by either SV40 or Py were shown to overexpress Bdp1, whereas the human papillomavirus 16 was found to

cause elevated BRF1 levels in cervical carcinomas (Daly et al., 2005; Felton-Edkins and White, 2002; Larminie et al., 1999).

JNK1 and JNK2 can modulate the levels of ELK1 occupancy on BRF1 and TBP promoters, thereby inducing or suppressing their expression respectively. The levels of Bdp1 respond similarly to JNK1 and JNK2 protein levels due to alterations in TBP occupancy at the Bdp1 gene promoter. Thus JNKs regulate Pol III transcription by altering the levels of all three TFIIIB subunits (Zhong and Johnson, 2009). c-MYC is also seen to bind E-boxes within the BRF1 gene promoter and upregulate its expression levels (Unpublished Data, Lynne Marshall). Thus, multiple proteins are capable of elevating levels of Pol III TFs and thereby induce Pol III transcription.

The levels of Pol III transcription can also be limited by reducing the availability of its TFs. Rb and p53 proteins were found to target BRF1 and TBP respectively, and sequester them away from the transcription complexes, thereby reducing Pol III transcriptional output (Crighton et al., 2003; Larminie et al., 1997). Rb-family proteins p107 and p130 have been shown to have similar properties (Sutcliffe et al., 1999). Recently, BRCA1 has been shown to cause reduced BRF1 and BRF2 proteins levels and repress Pol III transcription, however it still remains to be resolved whether this is a direct or indirect effect (Veras et al., 2009).

1.1.4.2 Post translational modification of transcription factors

Over the past decade Pol III transcriptional regulation has been shown to be a battle ground for kinases. Both activating and inhibitory phosphorylation events on Pol III transcription factors contribute to determining the overall transcriptional output of Pol III. CK2 (Johnston et al., 2002) and ERK (Felton-Edkins et al., 2003a) can directly interact with and phosphorylate BRF1, thereby facilitating its recruitment to promoters. Contrary to this, the phosphorylation of BRF1 by cyclin-dependent kinases (Cdks) can lead to transcriptional inhibition. Cdc2-cyclin B kinase was found to be sufficient for repressing transcription from *Xenopus* 5S rRNA genes (Gottesfeld et al., 1994). BRF1 was also shown to be hyperphosphorylated by kinases other than cdc2-cyclin B during mitosis, which leads to dissociation of Bdp1 from the gene promoters and subsequent repression of transcription (Fairley et al., 2003). Recent findings attribute this mitotic

repression of Pol III transcription to Polo-like kinase 1 (Plk1) [(Fairley et al., 2011), submitted].

Maf1 is a target for multiple signalling cascades including the PI3K/AKT, TORC1 and PP2A pathways. Maf1 was found to associate with Pol III-transcribed genes and interacts with TFIIIC, TFIIIB and Pol III. Repression by MAF1 leads to reduced PIC formation at Pol III transcribed promoters (Desai et al., 2005; Goodfellow et al., 2008). However, phosphorylated Maf1 is unable to repress Pol III transcription, which makes Pol III transcription sensitive to Maf1 phosphorylation by TORC1 or PKA and dephosphorylation by PP2A (Kantidakis et al., 2010; Moir et al., 2006; Oficjalska-Pham et al., 2006). The regulation of Pol III transcription by all these phosphorylation events makes it highly sensitive to the metabolic state of the cell.

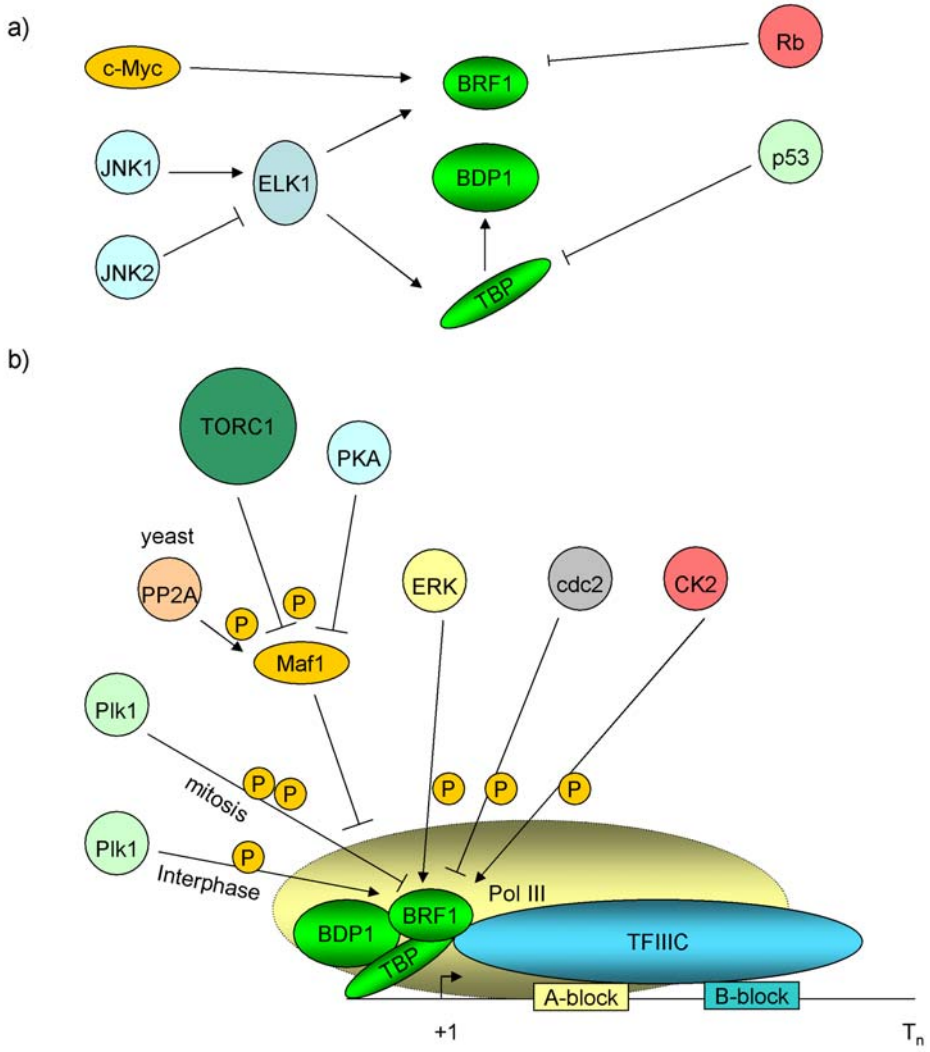


Figure 1.3: Schematic for the regulation of Pol III transcription
a) Mechanisms that influence cellular transcription factor availability. b) Kinases and phosphorylation events that influence Pol III transcriptional output. The arrows do not indicate direct protein-protein interactions. Ⓟ indicates a phosphorylation event.

1.1.5 Pol III and cancer

It is clear that the control of Pol III transcription is of vital importance for the correct regulation of cell growth and metabolism. Therefore, in order to upregulate its growth and proliferation, a cell must elevate the Pol III transcriptional output. Indeed, many studies have reported the overexpression of Pol III-transcribed RNAs in various tumour types [reviewed in (White, 2004; White, 2008)]. The levels of various Pol III transcription factors are also found to be upregulated in multiple cancers (White, 2004). TFIIIC is overexpressed in ovarian carcinomas (Winter et al., 2000), whereas BRF2 is found to be overexpressed in human breast, bladder and lung cancers (Lockwood et al., 2010; Melchor et al., 2007). Recent analysis of tissue microarrays has found elevated levels of BRF1 in prostate tumours (Noor Nam, unpublished data).

Thus, elevated Pol III transcription seems to be a feature of tumours; however this may simply be the outcome of deregulation of multiple oncoproteins and tumour suppressors. The question whether elevated Pol III transcription could cause cancer was answered recently by Marshall *et al*, where the authors demonstrate that a modest increase in BRF1 protein levels and a consequent increase in Pol III transcriptional output is sufficient for oncogenic transformation. Surprisingly, a small increase in initiator tRNA ($tRNA_i^{Met}$) was sufficient to reproduce these effects. An increase in c-MYC and cyclin D1 protein translation was observed as a result of elevated Pol III transcription. It is still unclear whether the observed oncogenic effects of BRF1 and $tRNA_i^{Met}$ are dependent on upregulation of c-MYC and cyclin D1 (Marshall et al., 2008). But it is clear that an elevation in Pol III transcriptional output can cause cellular transformation.

1.1.6 Recent insights

Recently, multiple studies have analysed Pol III and its transcriptional targets by ChIP-sequencing (Barski et al., 2010; Canella et al., 2010; Moqtaderi et al., 2010; Oler et al., 2010; Raha et al., 2010). These have revealed many aspects of Pol III transcription that were previously unknown and have shed light on the complexity of its regulation in mammalian cells (White, 2011). The first unexpected finding was the presence of Pol II 200bp upstream of Pol III-bound

loci, which in most cases does not correspond to the presence of a Pol II transcription unit (Barski et al., 2010; Oler et al., 2010; Raha et al., 2010). Basal Pol II transcription factors, such as TFIIB, are also present (Barski et al., 2010). It is not known what brings Pol II to these loci; however one can hypothesise that it may be the common transcription factor TBP or common transcriptional regulators like Myc, which are found to be associated with both Pol II and Pol III transcription complexes. Further biochemical and genetic analyses are required in order to understand if Pol II-enrichment at Pol III-transcribed loci is of functional relevance. It was also observed that active tRNA genes, i.e. those highly enriched for Pol III, often reside within or close to annotated Pol II promoters (Moqtaderi et al., 2010; Oler et al., 2010). This may reflect a preference for the open chromatin environment associated with Pol II genes.

Another unexpected finding was the presence of TFIIIC on its own at multiple intergenic loci (Moqtaderi et al., 2010; Oler et al., 2010). These loci were first identified in yeast and were named ETC (extra TFIIIC) loci (Moqtaderi and Struhl, 2004). ETC loci are often found near binding sites for the CCCTC-binding factor (CTCF) and, thus, may have a role in genome organisation. ETC sites are also often found in intergenic sequences between evenly spaced, but divergently transcribed Pol II genes (Moqtaderi et al., 2010). Since TFIIIC-binding sites have been shown to have barrier or insulator activities, these ETC sites may serve a similar purpose and regulate the differential transcription of neighbouring genes (Simms et al., 2008; Valenzuela et al., 2009).

The ChIP-seq data also show remarkable differences in Pol III-loading at tRNA genes between different cell lines. Despite having similar promoters and using the same core transcription apparatus, 26% of tRNA loci showed cell-type specific differences in Pol III occupancy (Barski et al., 2010). However, recent ChIP-seq. analysis in 6 different species has revealed that despite differences at individual tRNA genes, the additive Pol III-loading at all tRNA genes coding for a particular isotype is highly conserved (Kutter et al., 2011). These data suggest that the regulation of Pol III transcription is not as simple as it has been thought to be.

1.2 Chromatin

The average diameter of the human nucleus is only about 5 microns and 2 metres of DNA has to be contained within this small space. This is achieved by wrapping 146bp of the negatively charged DNA around an octamer of positively charged histone proteins. This forms the basic monomer of the chromatin, a nucleosome. These monomers are held together by linker DNA to form the classic 'beads on a string' conformation as viewed by X-ray crystallography (Thoma et al., 1979). Linker histone (H1/H5) binding organises a further 20bp of DNA into the nucleosome and also defines the geometry of the DNA entering and exiting the nucleosome. The angles of entry and exit further organise the nucleosomes into a 30-nm fibre which is then tethered onto an axial scaffold protein core to form higher order structures (Moser and Swedlow, 2011). Depending on the extent of chromatin compaction, the eukaryotic genome can be categorised into two distinct environments. The regions with relatively relaxed environment which permit gene transcription are classed as 'euchromatin'. In contrast, the more tightly packaged non-permissive regions with inactive genes are classed as 'heterochromatin' (Kouzarides, 2007).

The nucleosome core comprises of dimers of core histone proteins H2A, H2B, H3 and H4. These histone proteins have highly basic N-terminal tails which protrude from the core nucleosome and can make contacts with the nearby nucleosomes and other proteins. Histone tails are known to be targets of post-translational modifications such as acetylation, phosphorylation, ubiquitilation, methylation, deamination, ADP ribosylation, ß-N-acetylglucosamination, sumoylation and proline isomerisation. These histone tail modifications can directly influence the structure of the chromatin by altering the binding affinity of the nucleosome to the DNA. Moreover, they also act as signals for the recruitment of effector protein complexes (Bannister and Kouzarides, 2011). Only the post-translational histone modifications considered during this study will be discussed in detail.

1.2.1 Histone Acetylation

Histone acetylation is the characteristic euchromatic mark and, since its discovery, is almost invariably associated with transcriptional activation (Allfrey et al., 1964). Acetylation reduces the overall positive charge of histones, thus

reducing the affinity of the nucleosome to the negatively charged DNA. This in turn enables easier access by transcription factors to their target sequences. Acetylation occurs at numerous lysine residues on histone tails, including H3K9, H3K14, H3K18, H4K5, H4K8 and H4K12. Multiple lysine residues can get acetylated at the same time giving rise to hyper-acetylated regions of the genome. The genomic histone acetylation levels are highly dynamic and are regulated by opposing activity of the histone acetyl-transferases (HATs) and histone deacetylases (HDACs) (Kouzarides, 2007).

HATs catalyse the transfer of the acetyl group from acetyl-CoA to an amino group in the lysine side chain. HATs can be divided into three main families based on their acetyltransferase domains, GNAT superfamily (Gcn5-related N-acetyltransferase), MYST family (named for its founding members MOZ, Ybf2/Sas3, Sas2, and Tip60) and p300/CBP (CREB-binding protein) family (Sterner and Berger, 2000). In addition, multiple other proteins including (Hsieh et al., 1999a) and circadian rhythm protein CLOCK (Doi et al., 2006) have been shown to have HAT activity. To complicate the picture, HATs have also been shown to acetylate numerous non-histone proteins (Lee and Workman, 2007).

HATs are found to exist as multisubunit complexes in yeast, like the SAGA (spt-Ada-Gcn5) complex, the ADA complex (contains Ada but none of the other SAGA subunits), the NuA4 and the NuA3 complexes. The primary HAT complexes in humans are the GCN5/PCAF complex and Tip60 complex (Sterner and Berger, 2000). Other human HAT complexes like STAGA and TFTC have also been identified. The unique individual subunit compositions of these complexes define their targeting and specific acetylation activity. For example, the SAGA complex preferentially acetylates H3K9 and to a lesser extent H3K14, whereas the NuA3 complex prefers to acetylate H3K14. The Elongator complex (another GNAT family complex in yeast) prefers to acetylate H3K9, just like the SAGA complex. But unlike SAGA, which is targeted to gene promoters, the Elongator is targeted to gene bodies. The targeting of HAT complexes to active genes is also mediated by transcriptional regulators like c-MYC and TRRAP (Lee and Workman, 2007). Lysine acetylation is recognised by proteins containing a recognition domain called a bromodomain. For example, Pol II transcription factor TFIID subunit1 contains a bromodomain, thus directly linking histone acetylation to transcriptional activation. Certain HATs themselves contain bromodomains and

this maintains their occupancy at acetylated regions of the genome (Kouzarides, 2007).

Counteracting the activity of the HAT complexes, are the HDAC complexes which deacetylate histone tail lysines and repress transcription. Four different classes of HDACs have been identified: Class I (containing scRpd3 histone deacetylase-like proteins including HDAC1 and HDAC2), class II (containing scHda1 histone deacetylase-like proteins), class III (are homologous to scSir2 and referred to as Sirtuins) and class IV (with only one member HDAC11). Sertuins are different from the other classes because they require NAD^+ as a cofactor (Cress and Seto, 2000). HDACs recognise an acetylated aminoalkyl group and catalyse the removal of the acetyl group by cleaving an amide bond (Leipe and Landsman, 1997). HDACs on their own show little substrate specificity, but their presence within co-repressor complexes like NuRD, Sin3a or Co-REST provides specificity to their activity. HDAC-containing complexes can be recruited by DNA-binding proteins such as MBDs and p53 (Cress and Seto, 2000).

1.2.2 Histone methylation

Histones can be methylated at lysine and arginine residues. Unlike lysine acetylation, its methylation does not alter the charge of the residue, but changes its hydrophobic and steric properties. Lysines can be mono-, di- or tri-methylated depending on the functional properties of the histone lysine methyltransferase (HKMT) responsible. Each histone can be methylated at multiple sites and, depending on the location of the methylation, this has a positive or negative influence on transcription. In general, H3K4, H3K36 and H3K79 methylation are associated with transcriptionally active chromatin, whereas H3K9, H3K27 and H3K20 methylation marks inactive chromatin (Upadhyay and Cheng, 2011).

The majority of HKMTs, contain a highly conserved, 130 amino acid SET domain that uses S-adenosylmethionine (Adomet) as a methyl donor. HKMT are relatively specific with respect to the lysine side chain they can methylate and how many methyl moieties they can place upon that particular side chain. Well-studied human HKMTs include SUV39H1, G9a, SETDB1 (H3K9 methylases), EZH2 (H3K27 methylase), MLL1-5, SET1A, SET1B (H3K4 methylases), SET2, NSD1 (H3K36

methylases), SpSet9, SUV420H1, SUV420H2 (H3K20 methylases) and DOT1 (H3K79 methylase) (Kouzarides, 2007).

Depending on its position, histone lysine methylation is recognised by specific proteins. For example, Heterochromatin Protein 1 (HP1) chromodomain recognises H3K9 methylation and promotes the formation of repressive chromatin (Bannister et al., 2001). On the other hand, H3K4 methylation is recognised by bromodomain-PHD-transcription-factors (Li et al., 2006), CHD1 (an ATP-dependent chromatin remodelling enzyme) (Sims et al., 2005), as well as HAT complexes (Saksouk et al., 2009), thus associating this modification with active chromatin. H3K27 methylation is recognised by the polycomb group proteins (PCG), which are known to be global epigenetic transcriptional repressors. PRC1 (Polycomb Repressive Complex 1), once recruited in a H3K27-dependent or -independent manner, can in turn interact with EZH2 and propagate this repressive mark further (Margueron et al., 2009). Thus, the overall effect of a histone lysine methylation mark is defined by the effector proteins it recruits.

Histone methylation was considered to be an irreversible modification until the discovery of lysine-specific demethylase 1 (LSD1) (Shi et al., 2004). Multiple members of the jumonji C (jmjC)-domain-family, like JHDM1, JHDM2, JHDM3 and JMJD2 have also been shown to have histone demethylase activity. LSD1 and jmjC enzymes also demonstrate a level of substrate-specificity, where LSD1 primarily demethylates H3K4, JHDM1 demethylates H3K36, whereas JHDM2 and JHDM3 demethylate H3K9 and H3K36 residues (Kouzarides, 2007).

1.2.3 Chromatin remodelers

In addition to the effects of post-translational histone modification, the chromatin fibre is also actively reorganised by ATP-dependent remodelling complexes. The hydrolysis of ATP to ADP provides the energy required to move a nucleosome along the DNA *in cis* or completely reposition the nucleosome *in trans*. Thus chromatin remodelers make DNA/chromatin available to proteins that need to access DNA or histones directly for their function. Moreover, chromatin remodelling can also lead to nucleosome positioning and make specific regions of the DNA inaccessible (Hargreaves and Crabtree, 2011).

1.3 SWI/SNF chromatin remodelling complex

A screen for sucrose non-fermenting (SNF) yeast mutants identified the first components of the SWI/SNF complex, which were later characterised as chromatin remodelling proteins (Carlson et al., 1984). The SWI genes were later identified from mutants defective in mating type switching (Peterson and Herskowitz, 1992). Strains containing mutations in histone H2A and H2B failed to show the effects of SNF gene deletions. Moreover, swi2/snf2 and snf5 mutants showed altered sensitivity of the *suc2* gene to nuclease digestion, suggesting that these proteins altered the nucleosome occupancy at their target loci (Hirschhorn et al., 1992).

Brahma (BRM) was identified as an activator of homeotic genes in *Drosophila* and was found to be a homolog of the yeast swi2/snf2 gene (Tamkun et al., 1992). BRG1 was identified by a cDNA screen as the human BRM homolog and was later found to be involved in regulation of murine Hox genes (Khavari et al., 1993; Randazzo et al., 1994). *INI1* (integrase interactor 1), the human homolog of yeast snf5, was first identified in a yeast two hybrid screen as a binding partner of human immunodeficiency virus-type 1 (HIV1) integrase (Kalpana et al., 1994).

Swi2/snf2, a highly conserved member of the SWI/SNF family, was found to exhibit DNA-stimulated ATPase activity. Mutations in the nucleoside-binding site impaired snf2 activity, indicating that ATP hydrolysis by snf2 was necessary for transcriptional activation by SWI/SNF (Laurent et al., 1993). Purified yeast SWI/SNF complex was shown to exhibit ATPase activity, which was required to drive transcription from a nucleosomal GAL4 template (Cote et al., 1994). Partially-purified human homolog of the yeast SWI/SNF complex was also shown to mediate ATP-dependent nucleosome disruption and the binding of transcription factors to nucleosomal templates (Kwon et al., 1994). Thus, it was proposed that the primary role of the SWI/SNF complex was to promote transcriptional activation via nucleosomal eviction driven by ATP hydrolysis.

Microarray data analysis on snf2Δ or swi1Δ yeast strains revealed that 5% of yeast genes are regulated by the SWI/SNF complex. Moreover, the data also demonstrated that SWI/SNF represses more genes than it activates (Sudarsanam et al., 2000). Microarray data from SNF5 null mouse embryonic fibroblasts (MEFs)

also suggested a repressive role of SWI/SNF complex. The number of genes repressed by SWI/SNF is three-fold higher than the number of activated genes (Isakoff et al., 2005). Deletion of BRG1 in mouse ES cells followed by transcription analysis also revealed that BRG1 represses most of its direct targets (Ho et al., 2009a). Thus, despite its original identification as a transcriptional activator, a genome-wide repressive role is now attributed to the SWI/SNF chromatin remodelling complex.

A novel 15-subunit complex with the capacity to remodel the structure of chromatin (RSC) was later identified in yeast. The Sth1, Rsc6 and Rsc8 components of RSC were found to be significantly similar to the swi2/snf2, Swp73 and Swi3p components of the SWI/SNF complex respectively. It was observed that unlike the SWI/SNF complex, RSC is abundant in a yeast cell and is essential for cell viability (Cairns et al., 1996). Sth1 was later recognised as the ATPase essential for RSC activity (Du et al., 1998). The human homolog of the yeast RSC, the SWI/SNF-B complex, was termed the PBAF complex, since it contains a unique factor called BAF180 that shows homology to the chicken polybromo gene (Xue et al., 2000). RSC has been shown to have roles in DNA repair, sister chromatid cohesion, chromosome segregation and ploidy maintenance (Hargreaves and Crabtree, 2011).

1.3.1 Composition of the SWI/SNF complex

Purified yeast SWI/SNF complex, with an estimated molecular mass of ~1MDa, was shown to be composed of 10 subunits including swi1, swi2/swi2, swi3, snf5 and snf6 (Cote et al., 1994; Smith et al., 2003). However, its mammalian counterpart is twice as large in size and shows considerable heterogeneity in its subunit composition between different cell lines and tissues (Wang et al., 1996a; Wang et al., 1996b). Mammalian SWI/SNF complex normally consists of 9-12 subunits. Only four of these, called the core subunits, are required for remodelling activity *in vitro*. Reconstitution experiments showed that BRG1 or BRM can promote nucleosome remodelling independently *in vitro*. However, the addition of BRM-associated factors, BAF155 and BAF170, along with SNF5 (also known as BAF47) increases this basal activity to levels similar to that of immunoprecipitated SWI/SNF complexes (Phelan et al., 1999).

An individual SWI/SNF complex can only contain one of the two ATPases, such that BRM/BAF complexes are structurally distinct from BRG1/BAF complexes (Wang et al., 1996a; Wang et al., 1996b). The functional differences between these two complex types are still unclear and are an area of active investigation. A subset of SWI/SNF subunits, including BRG1 and BRM, can also be found as part of alternative chromatin remodelling complexes, such as the WINAC nucleosome assembly complex, the nucleosomal methylation activated complex (NUMAC) and the nuclear receptor corepressor-1 (NCoR-1) complex [for review see (Trotter and Archer, 2008)]

SNF5 is present in all the SWI/SNF complex variants and is thought to be essential for SWI/SNF function. SNF5 contains a highly-conserved central functional domain, which comprises of acidic repeat sequences. Mutations within this functional domain were found to influence the complex assembly step or the post-recruitment catalytic remodelling step (Geng et al., 2001). The mechanistic role of SNF5 in the remodelling step is still unclear, but it may be required for binding of SWI/SNF to a nucleosome (Dechassa et al., 2008). Homozygous loss of SNF5 has been shown to cause preimplantation embryonic lethality in mice (Guidi et al., 2001). Core subunits, BAF155 and BAF170, are thought to provide the protein scaffold that is important for complex assembly (Tang et al., 2010). In yeast, inactivation of any single subunit leads to dissolution and subsequent inactivation of the complex (Hargreaves and Crabtree, 2011).

Other known components of the mammalian SWI/SNF complex include BAF250a/b, BAF200, BAF180, BAF60a/b/c, BAF57, and BAF53A/B (Reisman et al., 2009). Unlike yeast SWI/SNF, the mammalian complex also contains actin or actin-related proteins (ARPs). Actin forms a tight complex with BRG1 via the conserved N-terminal HAS domain, however its precise function within the complex is still unclear (Szerlong et al., 2008). Actin can bind to myosin and act as an ADP-exchange factor, thus increase the rate of ATP-hydrolysis by myosin (Lymn and Taylor, 1971). Actin and ARPs may serve a similar role within the SWI/SNF complex.

The subunit composition of the human SWI/SNF complex varies dramatically in a tissue-specific manner (Wang et al., 1996a). Moreover, using low stringency immunoprecipitations and sequential chromatin immunoprecipitations, a recent

study has shown that SWI/SNF subunit composition can vary to a great degree even within the same cell type. Five different promoters regulated by the SWI/SNF complex showed varying subunit occupancy, leading to the proposition that each complex may be tailored to specific target requirements (Ryme et al., 2009).

This variability is thought to contribute significantly to the regulation of transcriptional programs during development. For example, pluripotent ES cells contain the esBAF version of the SWI/SNF complex, which comprises of BRG1, BAF155 and BAF53a, but not BRM, BAF170 or BAF53b. This complex is a part of the core pluripotency transcriptional network and is essential for ES cell self-renewal and pluripotency (Ho et al., 2009a; Ho et al., 2009b). Since BRG1 plays an essential role in the ES cell compartment, its deletion causes early embryonic lethality (Bultman et al., 2000). In contrast, BRM null mice are viable, however they show significant elevation in the levels of BRG1, indicating that BRG1 may compensate for the loss of BRM (Reyes et al., 1998).

Differentiation of ES cells into neuronal precursors is associated with an activation of BRM and BAF170 and the repression of BAF60b. In post mitotic neurons, repression of BAF53a and the activation of BAF53b, BAF45b and BAF45c is observed. Preventing this subunit switch impairs neuronal differentiation. Thus, varying the subunit composition of SWI/SNF complexes can lead to specific alterations in the transcriptional program which is essential for cellular differentiation (Lessard et al., 2007).

1.3.2 ATP-hydrolysis, structure of SWI/SNF and mechanisms of remodelling

Human BRG1 is approximately 74% identical to human BRM, 52% identical to *Drosophila* BRM and 33% identical to yeast swi2/snf2 (Khavari et al., 1993; Reisman et al., 2009). All snf2-family proteins contain a DNA-dependent ATPase domain. Early experiments showed that the addition of dsDNA causes more than a five-fold increase in the ATPase activity of recombinant swi2/snf2 (Laurent et al., 1993). The snf2 gene family is homologous to many helicases of the DEAD/H family and all its members contain seven conserved helicase-related motifs within the ATPase domain. These motifs enable snf2 proteins to specifically

recognise regions of double-stranded to single-stranded DNA transition, thus allowing SWI/SNF to localise to regions of active replication, repair and transcription (Muthuswami et al., 2000). A recent study has shown that the ATPase domain is capable of binding independently to DNA or ATP. However, only the ATPase domains which bind DNA first are capable of ATP hydrolysis (Nongkhlaw et al., 2009).

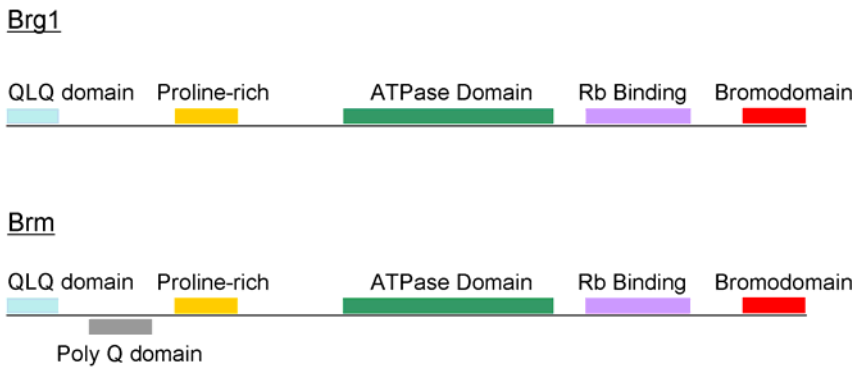


Figure 1.4: Schematic depicting the domains in the BRG1 **and** BRM **ATPases.**
Human BRG1 and BRM show 74% sequence identity. Multiple minigenes occur within both ORFs, however the functions of the resulting proteins are unclear. Both BRG1 and BRM also have an alternatively spliced exon [Figure adapted from (Reisman et al., 2009)]

In addition to the ATPase domain, snf2 family members also contain a C-terminal bromodomain that binds acetylated lysines on histone tails. However, the presence of this bromodomain is not essential for function, since its deletion does not affect BRM-mediated transcriptional activation (Muchardt and Yaniv, 1993; Trouche et al., 1997). Furthermore, a BRM mutant lacking the bromodomain was able to fully rescue the BRM knockout phenotype in flies (Elfring et al., 1998). BRG1 and BRM also contain a QLQ domain and an Rb-binding (LxCxE) domain, both of which are involved in protein-protein interactions. BRM also contains a polyQ expansion repeat domain, which is not found in BRG1. The functional properties of this domain are still unknown (Reisman et al., 2009).

Multiple attempts have been made to determine the structure of the SWI/SNF and RSC complexes. The yeast and human RSC complexes are shown to have a large central cavity surrounded by four regions of high density. The central

cavity is large enough to accommodate an entire nucleosome (Tang et al., 2010). However, the yeast SWI/SNF seems to be structurally different, as it contains an asymmetric shallow surface trough with dimensions matching the contour of a nucleosome. DNA footprinting and protein crosslinking experiments revealed that SWI/SNF subunits within and around the surface trough make extensive contacts with the DNA and the histone octamer. The catalytic Swi2/snf2 subunit contacts the nucleosomal DNA two helical turns from the dyad axis and also interacts with the histone octamer. SNF5 binds the DNA only weakly, but makes multiple contacts with the histone core (Dechassa et al., 2008; Tang et al., 2010).

Based on the structural and interaction data, a model for the remodelling activity of the SWI/SNF complex was proposed (Figure 1.5). The remodelling begins with a ~20bp translocation of SWI/SNF along the nucleosomal DNA towards the dyad axis. The DNA between the translocation site and an anchor point on the other side of the trough is displaced from the nucleosome, forming a loop. The translocation and the loop formation steps are thought to require ATP hydrolysis. This loop then propagates through the nucleosomal DNA causing the nucleosome to slide along the DNA (Dechassa et al., 2008). In addition to this, the extensive contacts made by the SWI/SNF complex with the nucleosome may also cause large disruptions in the nucleosome-DNA interactions and may facilitate the complete dissociation of the nucleosome by histone chaperones (Tang et al., 2010).

1.3.3 Transcriptional regulation by SWI/SNF

SWI/SNF has been linked to the activities of a large number of transcriptional regulators including Rb, c-MYC, p53, AP-1, EKLF, TGFβ, Myo-D, Sox2, Utf1 and Oct4. SWI/SNF has also been functionally linked to transcriptional regulation by various steroid receptors (Reisman et al., 2009). The best characterised interaction of SWI/SNF is with the retinoblastoma protein (Rb). As previously mentioned, the ATPases BRG1 and BRM both contain an Rb-binding (LxCxE) domain. BRG1 was shown to bind directly and specifically to hypo-phosphorylated Rb in a yeast two-hybrid screen and by co-immunoprecipitation experiments (Dunaief et al., 1994).

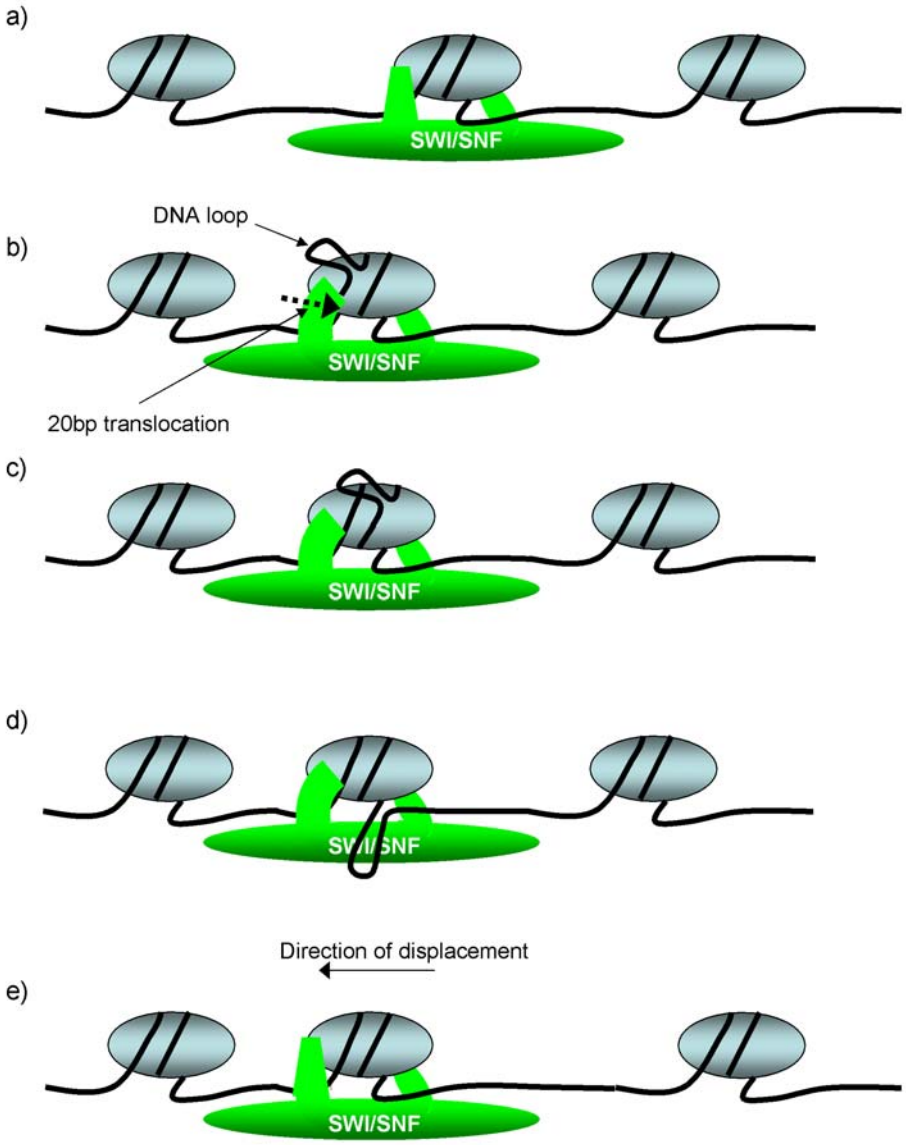


Figure 1.5: Schematic model for the chromatin remodelling activity of the SWI/SNF complex. The SWI/SNF complex contains a central trough that makes multiple contacts with the nucleosome. The left arm of the complex depicts the higher wall of the trough (a). A 20bp translocation of the left arm leads to disruptions in the DNA-histone contacts causing the formation of a loop (b). This step is thought to require ATP hydrolysis and is necessary for the remodelling activity. The loop propagates around through the nucleosome and exits through the lower wall of the trough depicted by the right arm (c-d). This causes a unidirectional displacement of the nucleosome along the DNA (e)(Dunaief et al., 1994).

The interaction with BRG1/BRM is required for Rb induced cell cycle arrest. Constitutively active Rb is unable to cause cell cycle arrest in BRG1- and BRM-deficient cells. This defect can be rescued by co-expressing BRG1, but not an LxCxE mutant form of BRG1 (Zhang et al., 2000). Moreover, SNF5-induced G1 arrest in SNF5-deficient cells was also found to depend upon the presence of functional Rb (Versteege et al., 2002). Rb co-operates with SWI/SNF to repress the activity of E2F1 transcription factors (Trouche et al., 1997). Conditional inactivation of SNF5 was shown to alter expression of various Rb-E2F-regulated genes (Guidi et al., 2006). The Rb repressor complex is found to contain HDACs and SWI/SNF, which together repress the transcription of cyclins-E and -A. Hyper-phosphorylation of Rb by cyclin-D-CDK4 causes the disruption of the Rb repressor complex, allowing E2F1 to induce cyclin-E, which enables S-phase entry (Zhang et al., 2000).

Despite this, direct interaction with Rb may not be necessary for SWI/SNF to regulate the cell cycle. Re-expression of BRG1 and SNF5 in deficient cells was found to induce p21 and p16 respectively, both of which inhibit Rb phosphorylation by CDKs (Kang et al., 2004; Oruetxebarria et al., 2004). There is also functional interaction between SWI/SNF and p53. BAF53 binds p53, and SWI/SNF activity is required for p53-mediated cell cycle control (Wang et al., 2007). Knockdown of SNF5 was found to cause cell cycle arrest and apoptosis in a p53-dependent manner (Kato et al., 2007).

SWI/SNF is also required for transcriptional transactivation by c-MYC. A repeat domain of SNF5 was found to directly interact with the C-terminal region of c-MYC. Mutations in either SNF5 or BRG1 abrogate the ability of c-MYC to transactivate its target loci (Cheng et al., 1999). Moreover, BAF53 also binds c-MYC and is critical for c-MYC oncogenic activity (Park et al., 2002). Thus, SWI/SNF complex is involved in genome-wide transcriptional regulation and has a critical role in cell growth and proliferation.

ChIP-sequencing analysis has shown that SWI/SNF localises near transcription regulatory elements, including enhancer sequences and regions critical for chromosome organisation, for example CTCF binding sites and replication origins. Gene-specific analysis and genome-wide assays suggest the involvement of SWI/SNF complex in cell cycle control and chromosomal organisation

(Euskirchen et al., 2011). Taken together, the available data suggests that the SWI/SNF complex is a master regulator of gene expression.

1.3.4 SWI/SNF and cancer

Many components of the SWI/SNF complex are lost, mutated or silenced in multiple cancers. SNF5 is a *bona fide* tumour suppressor, the homozygous loss of which leads to highly aggressive paediatric rhabdoid tumours (Versteege et al., 1998). SNF5 is also found deleted in both chronic and acute phase CML (Grand et al., 1999). Heterozygous knockout of SNF5 causes tumours in mice that are similar to human malignant rhabdoid tumours (MRT). Conditional deletion of SNF5 in mice leads to lymphomas or rhabdoid tumours with 100% penetrance within 10 weeks (Roberts et al., 2002).

Relative to SNF5, the ATPase subunits of SWI/SNF are lost in a greater number of human tumours. BRG1 and BRM expression is lost in lung, bladder, colon, breast, melanoma, esophageal, head/neck, pancreas and ovarian cancers [reviewed in (Reisman et al., 2009)]. Consistent with these data, $BRG1^{+/-}$ mice acquire mammary tumours within one year (Bultman et al., 2008). However, the tumour phenotype observed is significantly different from that in the $SNF5^{+/-}$ mice, indicating that these two proteins are not functionally equivalent. Loss of BRM, although non-transforming in mice, was found to cause distinct abnormalities in cell cycle control. Moreover, BRM heterozygote and null mice showed a significant increase in the number of lung tumours when treated with the carcinogen urethane (Glaros et al., 2007). The interplay of SWI/SNF with various tumour suppressors and oncoproteins may contribute significantly to its tumour suppressor activity. However, the mechanistic details of this activity are still unclear and more work is required in order to understand these.

1.4 DNA Methylation

DNA methylation is an epigenetic modification found in both prokaryotic and eukaryotic genomes. The occurrence of a methyl group at the 5th position of the cytosine nucleotide ring was first proposed by Wheeler and Johnson in 1904 (Figure 1.6). Naturally occurring 5-methylcytosine (5mC) was first discovered by Johnson and Coghill in 1925, following hydrolysis of nucleic acids from *Tubercule Bacillus* (Johnson and Coghill, 1925). In prokaryotes, both cytosine and adenine residues are found to be methylated, whereas in multicellular eukaryotes only cytosine bases are methylated. In human somatic cells, 5mC constitutes 1% of all DNA bases and usually occurs in the context of CpG dinucleotides (meCpG) (Ehrlich et al., 1982; Ehrlich and Wang, 1981; Gruenbaum et al., 1981).

Non-CpG methylation is widely prevalent in plants and forms a part of the host restriction system which protects the plant genome from foreign DNA (Finnegan et al., 1998). Non-CpG methylation in mammals is limited to the ES cell compartment and was initially recognised in mouse ES cells (Ramsahoye et al., 2000). Recently, the single-base-resolution map of the human DNA methylome revealed that nearly a quarter of DNA methylation in human ES cells exists at non-CpGs (Lister et al., 2009). This methylation is lost in differentiated cells, and is thought to have a role in maintaining stem cell pluripotency. The recent discovery of 5-hydroxymethylcytosine (hmC)(Figure 1.6), which occurs by hydroxylation of 5mC by Tet enzymes, adds to the repertoire of identified cytosine modifications (Tahiliani et al., 2009). Preliminary evidence suggests that hmC is involved in pluripotency, differentiation and carcinogenesis (Munzel et al., 2011). However, only meCpG and the associated literature will be discussed here.

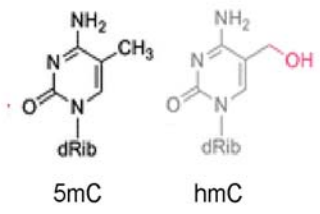


Figure 1.6: Molecular structures for 5-methylcytosine and 5-hydroxymethylcytosine

CpG methylation is essential for genomic imprinting, X-chromosome inactivation and spatiotemporal regulation of transcription during embryogenesis. It is also thought to be necessary for immobilization of mammalian retrotransposons, minimising genomic transcriptional noise and regulation of tissue specific transcriptional programmes (Bird and Wolffe, 1999). Despite these useful properties, the major hazard of the presence of 5mC is its mutability. Spontaneous deamination of cytosine results in uracil, which being an alien DNA base, is recognised by uracil glycosylases and corrected back to cytosine. However, deamination of 5mC leads to the formation of thymidine, which being a genuine DNA component is inefficiently repaired (Duncan and Miller, 1980). Glycosylases that can act on T:G mismatches have been identified, but due to the high frequency of deamination (approximately four C to T transitions per day in germ cells) mismatches are not always repaired (Millar et al., 2002).

Due to this high mutability, most regions of vertebrate genomes are distinctly devoid of CG dinucleotides. However, there are punctuated CpG-rich regions called CpG islands (CGIs). These ~1000 bp sequences have 10 times higher CpG content than the bulk genomic DNA, but these CpGs are distinctly devoid of methylation (Bird et al., 1985). This lack of DNA methylation has recently been attributed to the action of Tet1 enzymes and hmC, which were shown to preferentially associate with CGIs in mouse ES cells (Wu et al., 2011). CGIs contain unstable nucleosomes, and are found to be associated with origins of replication and transcription start sites. Proteins containing a "CXXC" motif specifically recognise unmethylated CpGs. Cfp1 (CXXC finger protein 1), a component of the Set1/COMPASS complex, localises at the vast majority of CGIs and causes high levels of H3K4me3 at these elements. This H3K4 methylation allows transcription initiation even at 'orphan' promoter-less CGIs (Deaton and Bird, 2011). Adrian Bird described CGIs as 'beacons for transcriptional initiation' (Personal communication).

1.4.1 DNA methyltransferases

Methylated CpGs are not randomly distributed throughout genomes and this observation suggests that DNA methylation is an enzymatic process. *In vitro* assays revealed that DNA polymerases are unable to distinguish between unmethylated and methylated cytosines, indicating that cytosines already

incorporated into DNA are methylated by an enzyme. The first DNA methylase was identified in *E.coli* and was shown to use AdoMet as the methyl donor (Gold et al., 1966).

It is now well-established that two different types of DNA methyltransferases (DNMTs) exist within mammalian cells: *de novo* methyltransferases and maintenance methyltransferases. Timothy Bestor and co-workers purified the first mammalian DNA methyltransferase from Friend murine erythroleukemia cells and named it DNMT1 (Bestor and Ingram, 1983). Initially its *de novo* methylation activity was identified, however further biochemical studies revealed that it has an increased preference for hemimethylated DNA. This was unlike the previously identified prokaryotic DNA methylases, which do not distinguish between hemimethylated and unmethylated DNA. The cloning and sequencing of mouse DNMT1 revealed a C-terminal catalytic domain which is similar to its prokaryotic counterparts. However, the mammalian enzyme also has a large N-terminal domain which is absent from the bacterial enzyme (Bestor et al., 1988; Lei et al., 1996). This N-terminal domain is thought to be responsible for the preference for hemimethylated DNA that has earned DNMT1 its classification as the maintenance methyltransferase.

Methylation maintenance refers to the process that maintains methylation patterns following DNA replication. The parental DNA strand retaining its methylation is used as a template for the methylation of the newly synthesised daughter strand. The palindromic nature of the CpG dinucleotide allows for the faithful reproduction of the methylation patterns onto the nascent daughter strands. The N-terminal domain of DNMT1 contains a replication fork targeting domain, which is responsible for its association with the replication fork (Liu et al., 1998). DNMT1 binds proliferating cell nuclear antigen (PCNA), an auxiliary factor for DNA replication, thus tightly coupling DNA methylation with replication (Chuang et al., 1997).

DNMT1 knockout (*Dnmt1*$^{-/-}$) mice were reported to be embryonically lethal, however ES cells carrying this mutation showed no detectable effect on viability or proliferation in culture (Li et al., 1992). In contrast to ES cells, differentiated cells do not tolerate the loss of DNMT1 well. Deletion of DNMT1 in proliferating neural cells leads to functional impairment and poor survival (Fan et al., 2001).

Fibroblast cells undergo apoptosis in the absence of DNMT1, through a p53-dependent mechanism (Jackson-Grusby et al., 2001).

DNMT2 is the most conserved of all known cytosine methyltransferases. It lacks the N-terminal recognition domains and more closely resembles bacterial DNA methylases. The strong conservation of all catalytic motifs indicates that DNMT2 should be a DNA cytosine methyltransferase; however, it does not show any *in vitro* or *in vivo* DNA methylation activity (Okano et al., 1998). There is recent evidence to suggest that DNMT2 may be involved in tRNA methylation (Goll et al., 2006) and RNA processing during cell stress (Thiagarajan et al., 2011). Moreover, DNMT2 has also been shown to have a role in retrotransposon silencing and telomere integrity in Drosophila somatic cells (Phalke et al., 2009).

Pre-implantation embryos show a dramatic drop in 5mC levels between the 8-cell and blastocyst stages, which is a consequence of the exclusion of DNMT1 from the nucleus. Following this, there is a rapid gain in 5mC levels and adult levels are reached by gastrulation (Carlson et al., 1992). The *de novo* DNMTs (mammalian DNMT3 family), as their name suggests, are responsible for establishing novel genomic DNA methylation patterns. DNMT3a and DNMT3b are highly expressed during early embryogenesis when most of the *de novo* methylation occurs and their deletion leads to impaired embryonic development (Lei et al., 1996). DNMT3a and DNMT3b do not show any preference for hemimethylated over fully unmethylated DNA substrate, which reflects their *de novo* methylation function (Okano et al., 1998). It has been proposed that they may also play a role in the maintenance of methylation (Hsieh, 2005; Liang et al., 2002). DNMT3a and DNMT3b double-knockout ES cells show disruption of *de novo* methylation; however, this does not cause demethylation at imprinted loci (Okano et al., 1999).

The third member of the DNMT3 family, DNMT3L, is expressed specifically in germ cells and lacks the methyltransferase activity. However, DNMT3L was shown to be crucial for establishment of genomic methylation imprints. DNMT3L-null females show a loss of methylation from imprinted regions of the genome, whereas null males are sterile with a complete absence of germ cells (Bourc'his et al., 2001). DNMT3L has been shown to stimulate the activities of DNMT3a and DNMT3b by direct interaction (Suetake et al., 2004). DNMT3L was shown to

interact with histone H3 tails with unmethylated K4 residues and target *de novo* methylation (Ooi et al., 2007). DNMT1 and DNMT3 family members are thought to interact with each other and together they establish genomic methylation patterns (Kim et al., 2002).

1.4.2 DNA methylation and transcriptional repression

Evidence accumulated over the past three decades suggests that DNA methylation plays an important role in transcriptional silencing. There are two primary ways in which the presence of meCpG is thought to suppress transcription. The first way is by reducing the affinity of transcription factors for their cognate DNA recognition sequences (Bird and Wolffe, 1999; Watt and Molloy, 1988). Many transcription factors have GC-rich binding sites and the presence of a methyl moiety may sterically hinder the specific contacts required for their recruitment to this target DNA (Figure 1.7a). Even though it doesn't alter the structure of the double helix significantly, the presence of the methyl group reduces the flexibility of the DNA molecule. Many transcription factors require bending of the DNA for successful binding and the increased inflexibility may impede bending of methylated DNA (Derreumaux et al., 2001). However, certain transcription factors, for example SP1 and CTF, are unperturbed by methylation of their recognition sequences (Ben-Hattar et al., 1989).

The second way in which DNA methylation can inhibit transcription is through the recruitment of methylated DNA-binding proteins (MBPs) (Figure 1.7 b). MeCP2 was the founding member of this family of proteins and was demonstrated to possess a methyl CpG-binding domain (MBD) sequence motif that is sufficient and necessary for 5mC binding (Lewis et al., 1992; Nan et al., 1993). MeCP2 is found to be concentrated on methylation-rich pericentromeric heterochromatin in mouse cells. On the other hand, MeCP2 shows a much more diffused chromosomal localisation in human or rat cells which do not contain highly methylated satellite DNA (Nan et al., 1997). This chromosomal localisation of MeCP2 is disrupted in Dnmt1 mutant mouse ES cells, indicating that it specifically binds methylated DNA (Nan et al., 1996).

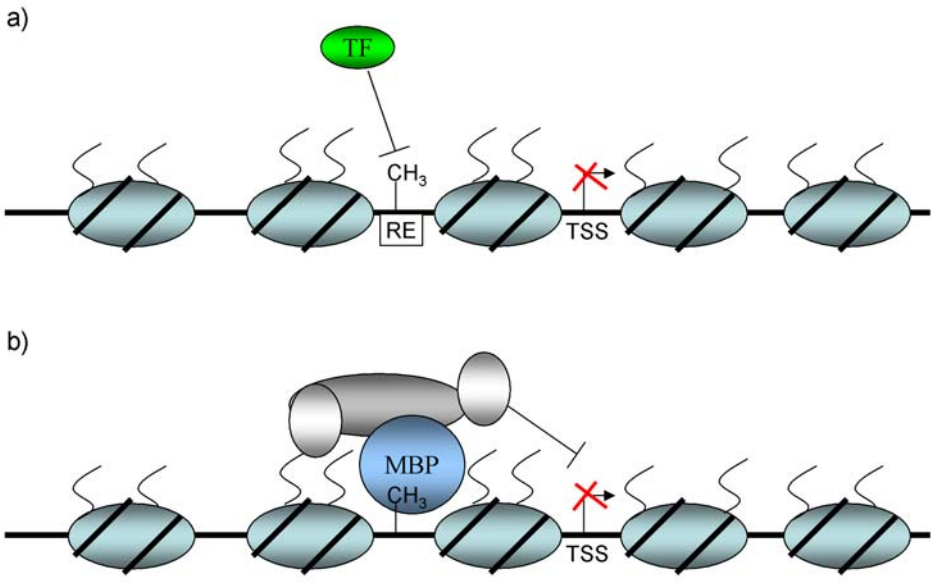


Figure 1.7: Modes of transcriptional repression by DNA methylation
DNA methylation can inhibit transcription by a) inhibiting transcription factor (TF) binding to its cognate recognition element (RE) and/or b) recruiting methylated DNA-binding proteins and associated inhibitory complexes [Adapted from (Klose and Bird, 2006)]

The role of MeCP2 in transcriptional repression was first demonstrated *in vitro* where it significantly repressed transcription from methylated promoters. This repression seemed to depend on the presence of its central transcriptional repression domain (TRD)(Nan et al., 1997). MeCP2 was also found to associate with DNMT1, suggesting a close co-ordination between DNA methylation and transcriptional repression following DNA replication (Kimura and Shiota, 2003). Despite this, imprinted genes that are known to be misexpressed in the absence of DNMTs are not misexpressed in the absence of MeCP2. This suggests that MeCP2 is not the only MBP recruited by DNA methylation patterns.

Indeed a sequence homology search with the MBD domain of MeCP2 revealed the presence of MBD1 in 1997 (Cross et al., 1997), followed by MBD2, MBD3 and MBD4 in 1998. All these new members, except MBD3, were shown to have specific methylated DNA-binding activity *in vitro* and *in vivo*. None or very low levels of MBP transcripts are detected in ES cells, except MBD3 transcripts can clearly be detected in all cell types (Hendrich and Bird, 1998). *MBD3*-null mice show early embryonic lethality, whereas MBD1 and MBD2 is dispensable for

mouse development and viability (Hendrich et al., 2001). MBD3 seems different to other MBPs and how MBD3 is targeted to its chromosomal loci was a mystery for a long time. However, some years ago, MBD3 was shown to interact with c-Jun in a yeast two hybrid screen (Nateri et al., 2004). Recently, the same group went on to show that unphosphorylated c-Jun recruits MBD3 to AP-1 target genes to inhibit expression, and this repression is relieved by c-Jun N-terminal phosphorylation by Jun-kinases (Aguilera et al., 2011).

MeCP2 has a 75 amino acid MBD domain, the disruption of which eliminates the chromosomal localisation of the protein. The MBD domain is the only common sequence feature amongst the MBP family members, the exception being MBD2 and MBD3 which are highly homologous (Dhasarathy and Wade, 2008). The MBD motifs from various species form a very similar, wedge-shaped structure. One side of the wedge is formed of beta sheets and the other side contains a short alpha helix (Dhasarathy and Wade, 2008). The recognition of meCpGs is due to four highly conserved residues: two arginines, an aspartic acid and a tyrosine. These form a hydrophilic patch that makes contacts within the major grove of the DNA. These contacts were found to be mediated by several highly structured water molecules. Thus, the MBD domain recognises hydration of the major groove of methylated DNA rather than cytosine methylation per se (Ho et al., 2008). MBD2 was recently shown to bind its target sequence in a single orientation. This binding orientation was reversed by partially reversing the sequence surrounding the meCpG. Moreover, a guanine residue directly following the meCpG significantly reduced the binding affinity. These data suggest the binding preferences of MBPs may also depend on sequences surrounding the meCpG (Scarsdale et al., 2011).

MBD1 was found to be recruited to both methylated and non-methylated CpGs. In addition to the MBD domain, MBD1 was found to contain three CXXC motifs that enable its binding to non-methylated CpGs. However, only one of these CXXC copies is conserved and able to bind DNA, while the other two may have roles in protein-protein interactions (Jorgensen et al., 2004). Point mutations in the CXXC domains that completely abolish unmethylated CpG binding, do not alter MBD1 binding to its target genes via the MBD domain. Thus, despite having binding preferences distinct from MBD2 and MeCP2, these findings suggest that MBD1 is primarily a methyl CpG binding protein (Clouaire et al., 2010).

There are two reports in the literature where MBD4 has been associated with transcriptional repression of methylated $p16^{INK4a}$ and hMLH1 genes (Fukushige et al., 2006; Kondo et al., 2005). However, the role of MBD4 in genome-wide transcriptional repression is still questionable (Hendrich and Bird, 1998). Aside from the MBD domain, it has a glycosylase domain which was shown to be involved in repair of the G:T and G:U mismatches resulting from deamination of 5mC and C residues. Due to this function, MBD4 is considered a crucial component of the base excision repair machinery (Aziz et al., 2009).

The Kaiso family of transcription factors are the newest members of the MBP family. Kaiso was identified as a p120 catenin binding partner using a yeast two-hybrid screen. Sequence analysis revealed that the Kaiso protein contained an N-terminal BTB/POZ protein-protein interaction domain and three C-terminal DNA-binding zinc fingers (Daniel and Reynolds, 1999). Further studies have shown that Kaiso behaves as a methylation-dependent transcriptional repressor and preferentially recognises two concurrent meCpG residues (Prokhortchouk et al., 2001). *In vivo,* Kaiso is required to maintain DNA methylation-dependent transcriptional silencing during early *Xenopus* development (Ruzov et al., 2004). p120-catenin is usually bound to E-cadherin and regulates its function and stability. However, following Wnt-stimulated phosphorylation, p120 is released from E-cadherin and associates with Kaiso (Del Valle-Perez et al., 2011). This association leads to nuclear export of Kaiso and enabling its binding to methylated CpGs at repressed promoters (Zhang et al., 2011). Thus, Kaiso provides a bridge between membrane signalling and DNA methylation-dependent transcriptional repression. Moreover, Kaiso has been shown to localise at the mitotic spindle and the centrosome, suggesting that Kaiso's transcriptional regulation may be linked to the control of the cell cycle (Soubry et al., 2010).

1.4.3 DNA methylation and chromatin

There is a clear bidirectional relationship between the DNA methylation and histone modification systems. It has been known for a long time that DNA methylation can influence histone modification patterns. DNA methylation levels of human HGMT gene variants demonstrated a direct correlation with their promoter accessibility and inverse correlation with their expression. This early

observation suggested that DNA methylation could effect the chromatin environment and thus influence transcription (Davey et al., 1997).

DNA methylation was shown to directly influence the positioning of nucleosomes at specific sequences *in vitro* (Davey et al., 1997). The transcription inhibition activity of MBPs is closely linked with their ability to recruit inhibitory chromatin complexes through their TRDs. The transcriptional repression by MeCP2 was found to involve Sin3A/HDAC co-repressor complex (Nan et al., 1998). MeCP2 TRD directly interacts with and recruits Sin3A, HDAC1 and HDAC2 to target loci. The transcriptional repression by MeCP2 can be alleviated by treatment with HDAC inhibitor trichostatin A (TSA) (Jones et al., 1998). MeCP2 has also been shown to co-immunoprecipitate with the BRM subunit of the SWI/SNF chromatin remodelling complex, which subsequently inhibits the transcription of the methylated genes (Harikrishnan et al., 2005).

MBD3 was identified as one of the seven subunits of the Mi-2/NuRD HDAC complex along with Swi2/Snf2 helicase/ATPase domain containing Mi2, HDAC1 and HDAC2 (Zhang et al., 1998). MBD2 has also been shown to interact with Mi-2/NuRD (Zhang et al., 1999) and Sin3A complexes (Boeke et al., 2000). Kaiso co-purifies with NCoR complex containing HDAC3, whereas Kaiso-like protein ZBTB38 has been shown to interact with several HDACs and co-repressor CtBP (Clouaire and Stancheva, 2008).

MeCP2 TRD was also shown to interact with proteins showing histone methyltransferase activity (Fuks et al., 2003b). However, this interaction could be mediated by DNMT1, since MeCP2 interacts with DNMT1 (Kimura and Shiota, 2003) and DNMT1 has been shown to interact with SUV39H1 (Fuks et al., 2003a). Treatment with 5-azacytidine, a cytosine analogue that inhibits DNMTs, was shown to cause a loss of H3K9 methylation at the $p14^{ARF}/p16^{INK4\alpha}$ locus (Nguyen et al., 2002). MBD1 has also been shown to associate with SETDB1, SUV39H and HP1 (Fujita et al., 2003). MBD1 was found to recruit SETDB1 to newly replicated DNA and facilitate methylation of H3-K9, thus coupling DNA and histone methylation (Sarraf and Stancheva, 2004). There seems to be a direct link between DNA methylation and H3K9 methylation.

Genomic analysis has revealed strong inverse correlation of DNA methylation with H3K4 methylation, while finding no correlation with methylation of H3K27 (Laurent et al., 2010). In contrast to meCpG, unmethylated CGIs are found to be enriched for H3K4 methylation and lack H3K9 methylation (Deaton and Bird, 2011). The maintenance of this reciprocal methylation of H3K4 and H3K9 in response to the DNA methylation state is intriguing. It could be due to different preferences in histone methylases recruited by MBD and CXXC protein families. However, Jumonji proteins are thought to contribute considerably to this inverse correlation. For example, PHF8 contains an N-terminal PHD domain that binds H3K4me3 and a Jumonji domain that demethylates H3K9me2 (Loenarz et al., 2010). JARID Jumoji family proteins bind H3K9me3 and demethylate H4K4Me3 (Iwase et al., 2007).

Thus DNA methylation recruits chromatin-associated repressors such as HDACs, HKMTS and HP1 through MBPs. However, the inverse influence of chromatin modification on DNA methylation has also become clear (Cheng and Blumenthal, 2010). H3K9 methylation has been shown to be required for recruitment of DNMTs via their interactions with G9a, GLP, SUV39H1 and HP1 at various loci (Cheng and Blumenthal, 2010; Fuks et al., 2003a; Lehnertz et al., 2003). A loss of DNA methylation in DNMT knockout cells does not lead to reduction in H3K9me3 levels (Dong et al., 2008; Lehnertz et al., 2003; Matsui et al., 2010), implying that H3K9 methylation is not dependent on the presence of meCpG as previously observed following 5-azacytidine treatment (Nguyen et al., 2002). In addition to this, a Snf2 family ATPase, LSH, was also found to be required for efficient DNA methylation. LSH was found to cooperate with G9a and facilitate histone and DNA methylation at target loci (Myant et al., 2011).

Methylation of H3K4 is thought to inhibit *de novo* methylation by suppressing the chromatin binding of DNMT3L (Ooi et al., 2007), DNMT3a and DNMT3b (Zhang et al., 2010). The *de novo* DNMTs were also shown to be catalytically less active on H3K4 methylated chromatin compared to unmodified or H3K9 methylated chromatin (Zhang et al., 2010). Despite this, orphan CGIs that show high levels of H3K4me3 are only partially resistant to DNA methylation, indicating that H3K4 methylation may not be sufficient for protecting gene promoters from DNA methylation (Thomson et al., 2010). The presence of transcriptional apparatus may also be required to provide complete immunity from DNA methylation

(Deaton and Bird, 2011). Thus, the relationship between DNA methylation and histone modifications is a complex bidirectional affair.

1.4.4 DNA Demethylases

DNA methylation levels in somatic cells do not fluctuate much. However, during embryonic development and reprogramming during gametogenesis, there is a requirement for DNA methylation to be in a spatiotemporal flux. DNA demethylation is also thought to be important during immune response, neurogenesis and tumorigenesis. This defined demethylation must require an active removal of the 5mC mark (Zhu, 2009). In contrast to the well known processes of DNA methylation, the enzymatic chemistry responsible for the removal of the methyl moiety from DNA was unknown for a long time. The chemistry necessary for active demethylation is energetically challenging and requires the disruption of a carbon-carbon bond.

Passive loss of meCpG through mutagenesis and spontaneous deamination were the only candidate mechanisms for DNA demethylation until the discovery of the first mammalian demethylase, MBD2b, in 1999 (Bhattacharya et al., 1999). MBD2b was shown to have DNA demethylase activity *in vitro*, however, this turned out to be "a red herring" and the findings could not be replicated by other groups (Boeke et al., 2000). Moreover, mice with mutant MBD2b show normal embryonic demethylation phases (Santos et al., 2002).

Another proposed mechanism for DNA demethylation is through base excision repair (BER), which is initiated by DNA glycosylases. MBD4 has been found to have DNA glycosylase activity that cleaves the glycosidic bond between the 5mC base and the deoxyribose. An endonuclease removes the deoxyribose and the gap is then filled in by the BER machinery. A thymine DNA glycosylase (TGD) has also been shown to have DNA glycosylase activity. However, both MBD4 and TGD show low *in vitro* glycosylase activity and much stronger G:T mismatch repair activity. Furthermore, no defect in DNA methylation was observed in MBD4 knockout mice (Zhu, 2009). However unlike other DNA glycoslases, knockout or catalytic inactivation of TDG leads to enryonic lethality in mice (Cortellino et al., 2011). *TGD*$^{-/-}$ MEFs show imbalanced histone modification and CpG

methylation at promoters of multiple genes with altered transcription (Cortazar et al., 2011).

Active enzymatic deamination of 5mC may also contribute to DNA demethylation. As previously stated, Tet1 enzymes deaminate 5mC and the hmC produced by this action can be reverted back to cytosine by DNA glycosylases and BER (Guo et al., 2011; Wu et al., 2011). DNMT1 cannot recognise hydroxymethylation a property which will also facilitate passive DNA demethylation (Zhu, 2009). Both DNMT3a and DNMT3b have also been found to deaminate meCpGs *in vitro*, thus contributing to DNA demethylation (Metivier et al., 2008). Moreover, DNMT3b was recently shown to interact with TGD and MBD4 and enhance their mismatch repair efficiency (Boland and Christman, 2008). Thus, *de novo* methyltransferases seem to have yet another function. Deamination coupled with mismatch repair seems to be the two step mechanism required for active demethylation in cells.

1.5 Short Interspersed nucleotide elements

The earliest reports of mammalian repetitive DNA date back to 1959, when Paul Doty and colleagues fractionated mammalian genomic DNA by centrifugation through density gradients and reported the presence of discrete satellite fractions (Sueoka et al., 1959). Following these initial studies, multiple groups embarked on the quest to characterise this fraction (Kit, 1961; Waring and Britten, 1966). Repetitive sequences were consequently shown to occupy large proportions of mammalian genomes (Britten and Kohne, 1970). Electron microscopy (Wu et al., 1972) and CsCl equilibrium sedimentation (Kram et al., 1972) indicated an interspersion of repetitive sequences in *Drosophila* DNA. However, it was not until 1973 that the existence of interspersed genomic repeats was clearly demonstrated by Roy Britten and co-workers. Using various physiochemical techniques that exploited the different thermodynamic properties of repeat sequences, they provided evidence for the presence of interspersed repeat sequences in *Xenopus* (Davidson et al., 1973) and sea urchin genomes (Graham et al., 1974).

In 1975, Carl Schmid entered the stage and published back-to-back reports with Paul Britten studying the interspersion of unique single copy and repetitive DNA

in sea urchin and *Drosophila* genomes (Davidson et al., 1975; Manning et al., 1975). By the end of that year, Schmid *et al*, using C_0t renaturation values, thermal elution and thermal hyperchromism, had demonstrated that at least 52% of the human genome consists of long interspersed single copy sequences of an average length of 2kb and their shorter counterparts, with the average length of 0.4kb (Schmid and Deininger, 1975). Approximately a third of the short interspersed repeats where shown to exist as inverted repeats. HeLa nuclear RNA was also found to contain 300nt inverted repeats that are resistant to single-strand specific nucleases (Fedoroff et al., 1977; Jelinek et al., 1974a); these were later attributed to transcription from the inverted nucleotide repeats. At least half of the human 300bp inverted repeats were found to have a site for *AluI* restriction endonuclease; thus, they were named the Alu family of SINEs. Early estimates predicted that the Alu SINEs populate at least 3% of the human genome and with 500,000 copies were dispersed over 30-60% of the genome (Houck et al., 1979). Following the publication of the partial nucleotide sequence of Alu SINEs in 1980 (Rubin et al., 1980), various primate genomes were also shown to contain related repeat families (Houck and Schmid, 1981).

At the same time, Kramerov *et al* screened for DNA sequences complementary to pre-mRNA with double-stranded 'fold back' structures and discovered the murine B1 and B2 repeats (Kramerov et al., 1979). They predicted that these repeats were present in excess of 10^5 copies in the mouse genome. Nucleotide sequence of B1 SINEs was found to be similar to that of the human Alu sequences (Krayev et al., 1980), whereas B2 repeats were found to have a different sequence. The authors suggested that B1 and B2 repeats may have possible roles in genome organisation, DNA replication and pre-mRNA processing (Krayev et al., 1982). William Ramsey and co-workers used a computational approach to show that murine B2 elements were derived from a tRNA sequence. The B2 elements have a 5' tRNA lysine-like sequence and terminate in a relatively divergent 3' A-rich sequence. The presence of the internal tRNA promoter was suggested to drive transcription by Pol III (Lawrence et al., 1985).

Sequence comparisons showed that T1 RNase resistant double stranded 'fold back' RNA are comprised in part of Alu transcripts (Jelinek et al., 1974b; Rubin et al., 1980). Alu sequences were estimated to comprise 18-25% of HeLa nuclear RNA (Fedoroff et al., 1977). The presence of Alu sequences in the human v-sis

gene (Dalla-Favera et al., 1981) and two B1 sequences in rat growth hormone genes (Gutierrez-Hartmann et al., 1984; Page et al., 1981) provided specific examples where SINEs are transcribed by Pol II. Thus, it was clear that SINEs could be 'read through' by Pol II as a part of an mRNA molecule.

However, previous *in vitro* transcription assays had shown that cloned human DNA containing ß-globin genes could be transcribed by Pol III from a transcription unit located approximately 1500bp upstream of the TSS (Duncan et al., 1979). Sequence comparison analysis revealed that this transcription unit belonged to the Alu SINE family, thus indicating that Alus were transcribed by Pol III (Jelinek et al., 1980). Later that year, Alan Weiner reported that 7SL RNA could form strong hybrids with Alu DNA; however, these hybrids were imperfect and could be digested by T1 RNase (Weiner, 1980). These sequence similarities were confirmed when the 7SL cDNA was cloned and sequenced two years later. The authors also reported differences between Alu and 7SL sequences; however at this point 7SL was thought to be a conserved subset of Alu repeats and the observed differences were due to divergent Alu family members (Ullu et al., 1982). Following two years of sequence homology and conservation analysis it was concluded that 7SL was an older conserved parent element from which Alu repeats had arisen (Ullu and Tschudi, 1984).

Our knowledge of the SINE family of repeats has come on in leaps and bounds since these early seminal works. It is now well-established that most mammalian SINE elements are derived from tRNA genes, except the human Alu and murine B1 families, which are derived from the 7SL locus (Daniels and Deininger, 1985; Okada, 1991b). SINEs are found throughout mammalian genomes; however, they are enriched in the gene-coding euchromatic regions and reside in the negative bands in Giemsa/Quinacrine metaphase staining. This is contrary to long interspersed nucler repeats (LINEs) that reside in the Giemsa/Quinacrine positive bands (Korenberg and Rykowski, 1988). Despite this relative abundance in gene-rich regions, very few SINEs are found within the 5' noncoding or coding regions of exons, where insertions may prove very deleterious. However, there are multiple examples of *Drosophila*, sea urchin, human and mouse SINEs that have provided an invaluable source of regulatory variation in evolution by insertions into vital regions of the genome (Britten, 1997; Deininger and Batzer, 1999).

1.5.1 SINEs and transposition

Individual Alu repeats are surrounded by short direct repeats and terminate in a polyA-rich 3' end. These observations suggest that Alu elements are dispersed through an RNA intermediate via retrotransposition (Jagadeeswaran et al., 1981). With over 1 million insertions, Alu elements are the most abundant retrotransposons in the human genome. The completion of whole genome sequencing has revealed that Alus comprise around 10% of the human genome (Lander et al., 2001; Smit, 1999). On average, there is one Alu insertion approximately every 3kb. Surprisingly, Alu elements are estimated to have originated only 65 million years ago, and are considered a young retrotransposon family. MIR elements, an older SINE family, have been estimated to have arisen 130 million years ago. However, despite their age, they only show 120,000 insertions within the human genome (Jurka et al., 1995). Comparatively, Alus have shown a phenomenal rate of expansion. Alu family shows 18% sequence divergence from the parental 7SL sequence and there is a further 14% sequence divergence of various Alu subfamilies from the Alu consensus sequence (Deininger and Daniels, 1986).

B1 and B2 are the most abundant murine SINE families. The mouse SINEs are relatively less abundant than the human counterparts, but show higher divergence from the parental genes. Mouse B1 elements, with 564,000 repeats, show 25% divergence from the murine 7SL sequence and various B1 subfamilies show a further 8% divergence from the consensus. Mouse B2 sequences, with 348,000 repeats, show 34-39% divergence from the parental tRNA, and B2 subfamilies show a further 10% sequence variability from the consensus (Deininger and Daniels, 1986; Waterston et al., 2002).

It is well-established that the propagation of Alu sequences occurs through the process of reverse transcription of a Pol III transcribed Alu RNA via retrotransposition. Alu sequences lack protein-coding ORFs. It is believed that they piggyback on the LINE retrotransposition machinery. LINE elements code for a functional reverse transcriptase, which also shows endonuclease activity (Feng et al., 1996; Mathias et al., 1991). This provides the infrastructure required for SINE retrotransposition by target-primed reverse transcription (TPRT). The endonuclease domain nicks one strand and the exposed 3'-OH is used to prime

the reverse transcription. Once the transposon cDNA is fully synthesised, the second strand is nicked and the transposition event is completed by ligating the 5' end of the transposon to the genomic DNA. The short tandem repeats found at 3' ends of many SINE insertion sites are thought to arise due to addition of untemplated nucleotides followed by template switching (Batzer and Deininger, 2002; Cost et al., 2002; Luan et al., 1993).

Even though LINE L1 reverse transcriptase exhibits a strong preference for L1 RNA, Alu elements have been remarkably successful in hijacking it for their own means, thus earning the name "a parasite's parasite" (Schmid, 2003; Wei et al., 2001). But with 1 million insertions in 65 million years, Alus are more successful at transposition compared to L1 elements, which have 500,000 insertions in 160 million years. So why is Alu 'in-trans' transposition better than LINE 'in-cis' transposition? Firstly, Alu flanking sequences have a conserved endonuclease cleavage motif that is highly homologous to the L1 endonuclease motif (Jurka, 1997). Moreover, the size of the A-rich tail was found to be highly correlated with increased transposition of Alu subfamilies (Roy-Engel et al., 2002). Alu RNA is known to interact with SRP9 and SRP14 with high affinity (Chang et al., 1996), which would enable its interaction with the large ribosomal subunit. This could increase the chances of Alu RNA interacting with the nascent L1 reverse transcriptase before its L1 RNA counterpart. A longer A-rich tail hanging out of the ribosome would improve these chances further (Boeke, 1997; Schmid, 2003). LINE L1 codes for two proteins, ORF1p and ORF2p, which are both deemed necessary for LINE retrotransposition. However, Alus are only thought to require ORF2p, even though the presence of ORF1p is beneficial to their transposition (Wallace et al., 2008). This may also provide SINEs with an advantage and explain their augmented prevalence compared to LINEs.

Apart from the Alu family, the MIR family is the other major SINE family resident in the human genome. This older family is thought to be transcriptionally and transpositionally silent (Smit and Riggs, 1995). MIRs are 260bp tRNA-derived sequences and comprise 0.2-0.3% of the human genome (Lander et al., 2001). These old elements provide a fossilised record of the period when, prior to the expansion of the LINE L1 and its Alu parasite, the genome had to withstand the expansion of the LINE L2 and its MIR parasite. The impact of these transposable

elements on the evolution of the human genome should not be underestimated (Cordaux and Batzer, 2009).

The rate of Alu expansion has varied throughout primate evolution (Shen et al., 1991). 65 million years of mutagenesis, methyl CpG deamination (Cooper and Krawczak, 1989) and ADAR editing of Alu RNA (Chen and Carmichael, 2008) have caused many Alu sequences to diverge significantly from the parental consensus. Despite this there is 70% homology between all Alu family members. This indicates that not all Alu sequences are capable of transposition, and only a few 'master Alu genes' were responsible for the spread of this SINE family. These master genes can be classified into subfamilies depending on 18 individual diagnostic mutations. The evolution of each subfamily can be explained by the sequential accumulation of alterations and these in turn indicate the age of the subfamily (Shen et al., 1991).

The nomenclature for the Alu subfamilies follows a alphabetical progression from the oldest (J) to the intermediate (S) and the young (Y) (Batzer et al., 1996). Two of the oldest and largest subfamilies, Alu-J and Alu-S, represent approximately 83% of all Alu sequences. Alu-S is further sub-divided into Alu-Sx, -Sp, -Sq, -Sc and –Sb sub-classes (Jurka and Milosavljevic, 1991; Jurka and Smith, 1988). Alu-Y is the youngest subfamily and shows the least amount of sequence divergence from the master gene. Studying the prevalence of the various subfamilies in various primate species has revealed that most Alus originated 60 to 40 million years ago and the rate of transposition has decreased dramatically since then (Shen et al., 1991).

1.5.2 SINE DNA and RNA structures

A typical Alu element has a dimeric sequence, with the left and the right monomeric arms linked by a variable length central A-rich linker. Both arms are believed to have evolved separately and fused together later to form the contemporary dimeric element. Since retrotransposons frequently insert into the A-rich tail of other retrotransposons, such a fusion can be easily conceived (Okada, 1991a). The sequence of the right arm is very similar to the 7SL sequence, except for the loss of 155bp that form the S domain of the 7SL molecule. The left arm has the same 155 nucleotides missing, but has undergone a further 31bp deletion (Weiner et al., 1986). The 3' end comprises of an A-rich tail of variable length (Figure 1.8a). The left arm of the consensus Alu sequence has functional A and B-boxes required for the recruitment of Pol III transcription apparatus. Moreover, the left arm lacks a terminator sequence and thus allows the transcription of the whole element, since the A and B-boxes in the right arm are degenerate and non-functional (Batzer and Deininger, 2002).

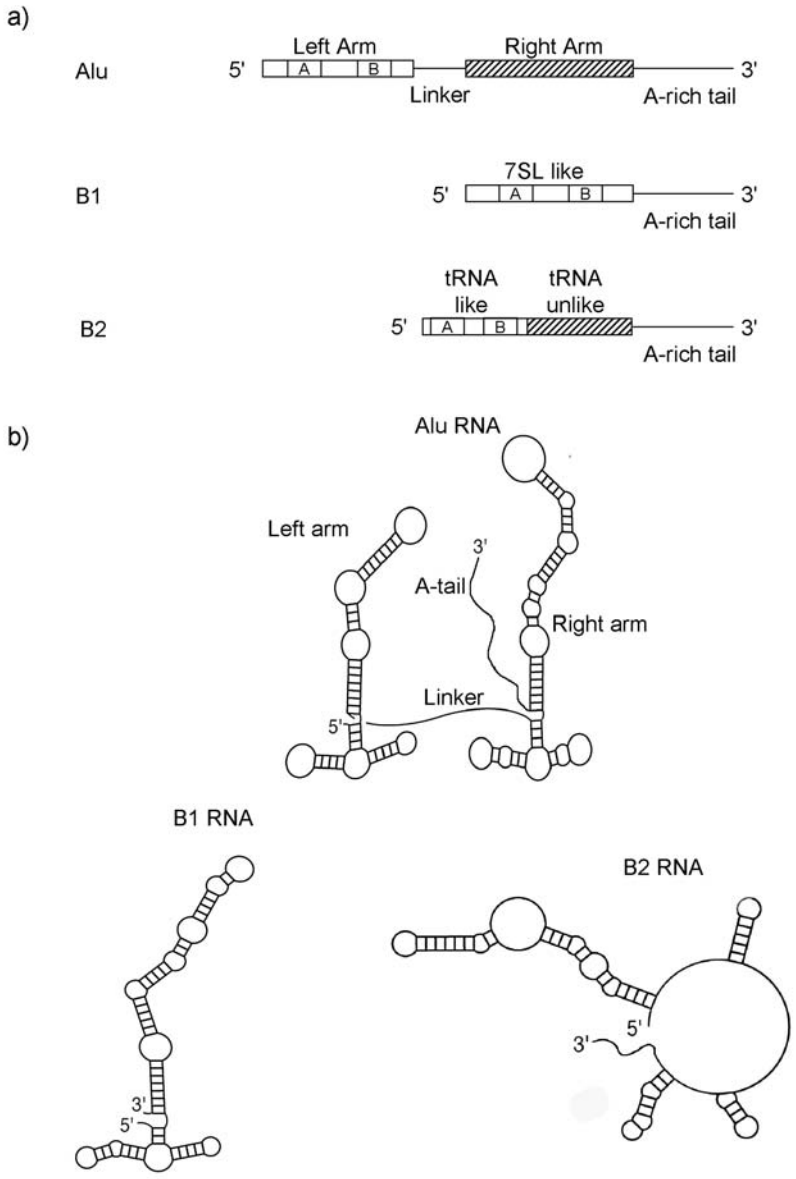


Figure 1.8: DNA sequence schematics and RNA secondary structures of SINEs.
a) Schematic depicting the DNA sequence and domains of human Alu and murine B1 and B2 SINEs. Functional A and B-boxes of a Pol III promoter are shown (adapted from (Batzer and Deininger, 2002; Krayev et al., 1980; Maraia, 1991)). b) Models for the secondary structures of SINE RNA molecules (adapted from (Mariner et al., 2008)).

The monomeric mouse B1 sequence is similar to the human Alu right monomer, however the internal deletion of the 7SL sequence extends 14nt further. The B2 SINEs are generally ~190bp long, with a 5' tRNA-like region, followed by a tRNA unrelated sequence and an AT-rich tail (Figure 1.8a). Both B1 and B2 elements contain functional Pol III promoters; however, they lack conventional Pol III transcription terminator sequences (Krayev et al., 1980; Krayev et al., 1982; Maraia, 1991; Singh et al., 1985).

As previously mentioned, SINEs can be transcribed by Pol II as a part of mRNAs. Alu sequences were identified in more than 80 fully spliced cDNA sequences, 82% of which were located in the 3'-UTRs and 14% in the 5'-UTRs (Yulug et al., 1995). It is widely believed that SINE transcription by Pol III is a rare event *in vivo* (discussed further in chapter 3). Human Alu RNA is a ~300nt molecule and its secondary structure consists of two 7SL-like arms linked by an A-rich linker (Sinnett et al., 1991). Mouse B1 RNA is the monomeric equivalent of one of the Alu arms and shows the highly conserved 7SL folding pattern (Labuda et al., 1991; Maraia, 1991). The B2 RNA has a more complex secondary structure consisting of a central bulge, one long arm and three short arms. The long arm is made up of three stems, two bulges and a loop (Espinoza et al., 2007) (Figure 1.8b).

1.5.3 Functional significance of SINEs

For a long time, SINEs were considered to be junk filler DNA that was of no functional significance; however, this point of view has dramatically altered. Several groups have observed that Pol III-derived SINE transcripts are elevated following cell stress, indicating that the transcripts may serve a purpose (Li et al., 1999; Liu et al., 1995; Rudin and Thompson, 2001; Schutz et al., 2005). The prevalence of SINEs in mammalian genomes also indicates that they may provide a selective advantage and may have one or more cellular functions.

Heat-shock response is an elegant example of a rapid and robust mechanism of cellular defence against environmental stress. A well-known feature of heat-shock response is the dramatic change in cellular transcription, whereby the transcription of heat-shock response genes is upregulated and the transcription of a multitude of housekeeping genes is downregulated. The heat-shock

transcription factor 1, elongation factor eEF1A and non-coding heat-shock RNA1 play important roles in inducing expression of heat-shock proteins (Sarge et al., 1991; Shamovsky et al., 2006). However, the mechanism for the transcriptional inhibition of housekeeping genes remained elusive until recently.

Both human and mouse SINEs show dramatic increases in Pol III-mediated transcription during heat-shock response. However, this increase is unique to SINEs and other Pol III-transcribed RNAs remain unaffected (Li et al., 1999; Liu et al., 1995). James Goodrich and co-workers have shown that Alu RNA and B2 RNA specifically bind to Pol II and inhibit mRNA transcription by interfering with preinitiation complex formation (Allen et al., 2004; Espinoza et al., 2007; Mariner et al., 2008). Only one arm of Alu RNA seems to show the transcriptional inhibition, even though both the arms can bind the polymerase. The inhibitory effect is attributed to two loosely-structured repression domains, which can transmit their inhibitory function in *trans* to other Pol II-binding RNA molecules. Both B1 and B2 RNAs can bind Pol II; however, only B2 RNA contains the inhibitory domains required to inhibit transcription (Shamovsky and Nudler, 2008).

The increase in levels of Pol III-transcribed SINE transcripts is very transient and upon recovery from stress, the abundance of SINE RNA rapidly decreases to its basal level (Liu et al., 1995). This is thought to be due to the very short half-life of Alu transcripts. In K562 cells, the half life of Alu RNA following actinoMycin block was calculated to be 0.5hrs (Li et al., 2000). This is extremely short compared to the parental 7SL RNA, which has a half life of ~10hrs. Using transiently-expressed Alu deletion constructs, it has been shown that the right Alu monomer is inherently unstable and causes the full-length transcript to be rapidly degraded (Li and Schmid, 2004). Moreover, Alu RNA was recently shown to be a direct target of DICER1-dependent degradation in human and mouse retinal pigmented epithelium. The elevated levels of Alu RNA in the absence of DICER1 were shown to directly contribute to disease progression in geographic atrophy, an untreatable advanced form of age-related macular degeneration (Kaneko et al., 2011).

SINE transcripts exist as full length nuclear and 3' processed cytoplasmic small cytoplasmic RNAs (scRNAs) (Maraia, 1991; Matera et al., 1990). However, it was

observed that very high expression of Alu RNA by transient transfections was unable to increase the abundance of scAlu RNA. This suggests that the level of these scRNAs is very tightly regulated by factors other than the abundance of the primary transcript (Chu et al., 1995). These cytoplasmic transcripts have been shown to associate with SRP9 and SRP14 with high affinity (Chang et al., 1996). The Alu domain of the 7SL RNA is thought to be responsible for blocking protein chain elongation until the protein reaches the ER membrane (Siegel and Walter, 1988). As assumed previously, Alu-SRP9/14 complexes were shown to inhibit protein translation *in vitro*. However, Alu-SRP9/14 particles seemed to act at the chain initiation step, rather than chain elongation, and resulted in lower polysome levels (Hasler and Strub, 2006).

Full length Alu transcripts have also been shown to interact with dsRNA-regulated protein kinase (PKR), which upon binding dsRNA, autophosphorylates and inhibits translation initiation via phosphorylation of eIF2α (Williams, 1999). Low level Alu transcripts activate PKR, whereas high levels of Alu transcripts have a strong inhibitory effect on PKR activity thus elevating global translation (Chu et al., 1998). Moreover, it was shown that the right arm of Alu RNA increase translation of specific mRNAs in a PKR-independent manner. In this case, levels of global translation are not affected (Rubin et al., 2002). The interaction of Alu RNA with poly(A)-binding protein, a regulator of translation, could also contribute to its translational regulation (Kondrashov et al., 2005). Therefore, SINE RNA seems to be capable of modulating cellular protein translation in multiple ways.

Adenosine-to-inosine (A-to-I) RNA editing is a post-transcriptional alteration of dsRNA by adenosine deaminases ADAR1, ADAR2 and ADAR3 (Bass, 2002). It has been observed that the primary targets of this process are 3' UTRs containing reversely oriented Alu elements (Athanasiadis et al., 2004). It was recently reported that there are 333 human genes which contain such Alu sequences. Edited mRNAs have been shown to be retained in the nucleus, thus inhibiting the protein expression (Chen et al., 2008). Varying degrees of A-to-I editing has also been thought to contribute to transcript diversity (Barak et al., 2009).

When Krayev *et al* sequenced mouse B1 SINE, they also observed that two regions within the B1 sequence were homologous to intron-exon junctions. This

was one of the earliest suggestions that SINE sequences may be involved in pre-mRNA splicing (Krayev et al., 1982). We know now that internal exons containing SINEs, predominantly in antisense orientation, are generally alternatively spliced (Sorek et al., 2002). SINEs are responsible for generating new exons by a process termed as exonisation. There are 7810 Alus within the human genome that are prone to exonisation; these may contribute considerably to protein diversity and also potentially lead to genetic disorders (Sorek et al., 2004).

SINEs have also been implicated in driving human miRNA transcription (Borchert et al., 2006; Monteys et al., 2010) and influence targeting of mRNAs by RNAi (Smalheiser and Torvik, 2006). Alus have also been shown to be sites of cohesin loading onto DNA, thus contributing potentially to genome organisation (Hakimi et al., 2002). Transcription at SINEs is also thought to influence the local chromatin environment. By serving as boundary elements, SINEs are thought to block the spread of heterochromatin and thus help regulate transcription (Lunyak et al., 2007). However, the boundary element function may be a general feature of Pol III transcribed loci, where the binding of TFIIIC is seen to be sufficient for stopping heterochromatin spreading (Raab et al., 2011). Thus, SINEs seem to provide modulatory and evolutionary contributions to mammalian genomes and are a lot more than a mere nuisance.

However, SINEs and their transcripts also have a darker side to them. SINE transcript overexpression has recently been shown to be cytotoxic in retinal cells and cause geographic atrophy, an untreatable form of human blindness (Kaneko et al., 2011). Moreover, SINE retrotransposition has the potential to disrupt vital genetic pathways and be of great detriment to the cell. SINE sequences are also responsible for chromosomal instability through transpositions and deletions caused by erroneous repair [reviewed in (Konkel and Batzer, 2011) and Discussed further in chapter 5].

Increasing amounts of evidence are emerging for the role of SINE transcripts in modulating various cellular processes. It seems that SINE elements and their transcripts also play a role in human diseases, and therefore it is vital to understand transcriptional regulation of SINEs (chapter3 and chapter4).

1.6 General Aims

The general aim of this study was to study the effects of DNA methylation on transcription by Pol III at SINEs. The effect of chromatin remodelling by SWI/SNF at tRNA genes was also investigated.

Chapter 2 Methods and materials

2.1 Cell Culture and treatments

All cell culture was performed in a class II hood, using aseptic techniques, sterile equipment and reagents. All cell lines were cultured in humid incubators set at 37°C and 5% CO_2.

HeLa, A31, MEF and IMR90 cells were cultured in DMEM (Dulbecco's Modified Eagle Medium) supplemented with 10% foetal bovine serum (FBS), 2mM L-glutamine, 100U/mL penicillin-streptoMycin. Dmnt1 wild-type and knockout MEFs were cultured in DMEM supplemented with 10% FBS, 2mM L-glutamine, 100U/ml penicillin-streptoMycin, 2mM sodium pyruvate, 1% non-essential amino acids and 0.01% ß-mercaptoethanol. Hom6.9 cells were cultured in DMEM supplemented with 10% ES batch-tested FBS, 2mM L-glutamine, 100U/ml penicillin-streptoMycin, 1% non-essential amino acids, $5x10^5$ U leukemia inhibitory factor (ESGRO® LIF, Millipore) and 0.01% ß-mercaptoethanol. ES cells were grown on plates coated with 0.1% gelatine.

For cell passaging, the media was aspirated from a 75-85% confluent flask/plate and the adhered cells were washed with 2ml of buffered trypsin-EDTA (0.05% trypsin, 0.02% EDTA). A further 2ml of buffered trypsin-EDTA was added after the wash and the flask was left at 37°C for approximately 2 mins. Following trypsinisation, fresh medium was immediately added to the dissociated cells in order to neutralise the trypsin and cell suspensions were transferred to new flasks/plates.

2.1.1 Cryo-freezing and recovery

Cryo-freezing was used for storage of all cell lines. Cells were trypsinised as described, pelleted by centrifugation at 500g for 5mins, and resuspended in 70% FBS, 20% DMEM and 10% dimethyl sulphoxide (DMSO, Sigma). Cells were frozen gradually by submerging the cryo-vials in propanol and placing the propanol container in -80°C overnight. The cells were transferred to liquid nitrogen storage the following day. For cell recovery after liquid nitrogen storage, the cryo-tubes were rapidly thawed at 37°C and the cells were washed in media at 37°C to remove the DMSO. The cells were then cultured in small $25cm^2$ flasks.

2.1.2 Drug treatments

Cells were treated with 4μM 5-azacytidine (Sigma) in DMEM for 16 – 72hrs depending on the cell type and 5ug/ml bleoMycin (Calbiochem) in water for 16 hours. Indicated concentrations of Chaetocin (Sigma) in dH_2O and 50μg/ml α-amanitin (Sigma) in dH_2O were used to treat cells for 24 hours.

2.2 Translocation assay

Maria Jasin generously provided Hom6.9 cells on which the translocation assays were conducted as previously [Elliott, 2005 #2797]. 7 x 10^6 cells were plated per 10cm gelatinised dish 16hrs prior and the growth media was changed 2hrs prior to the experiment. The cells were washed quickly with PBS and trypsinised with 2ml of buffered trypsin-EDTA (0.05% trypsin, 0.02% EDTA). These cells detach relatively quickly and should not be trypsinised for longer than 1min. The trypsin was neutralised using 8ml of media and the cells were passed gently through a 10ml pipette to ensure a single cell suspension. The cells were pelleted by centrifugation at 500g for 5mins at room-temperature and washed once with 10ml PBS. The cell pellet was then resuspended in 2ml PBS. 900μl of this cell suspension was transferred to two sterile microfuge tubes and 25μg of pTK-hyg and pCBAS (provided by Maria Jasin) plasmids were added to the tubes. The cell suspension was then transferred into 4mm electroporation cuvettes without introducing bubbles. The cells were elecroporated using a GenePulser Xcell™ electroporator (BIO-RAD) at 250V, 950μF and infinite resistance settings. The cells were then allowed to recover in the cuvettes for 20 minutes. Cells were treated at room-temperature throughout the electroporation procedure.

The cells were then resuspended in appropriate volume of warm growth medium and plated on labelled gelatinised plates. Cells were allowed to recover for 10hrs after electroporation and treated for 16 hrs with 4μM 5-azacytidine or 5μg/ml bleoMycin for 16hrs. Following drug treatment, the cells were washed with PBS and provided with fresh growth medium. 24hrs following media replacement, selection in 200μg/ml neoMycin was performed for 8-10days. One plate for each treatment was left without selection in order to ascertain loss of cell viability. The colonies obtained following selection were fixed for 10

minutes with methanol and stained with 1:40 GIEMSA stain. The clones were then counted manually.

2.3 RNAi and protein overexpression

Transient transfections were performed using Lipofectamine2000® (Invitrogen) and Opti-MEM® reduced serum media (Invitrogen). 2 x 10^5 cells were seeded in 10cm plates 16-24 hrs prior to transfection. The growth medium aspirated off and the cells were washed in PBS twice. After removing the PBS, 5ml of warm Opti-MEM® was added to the medium. The transfection solution per plate was prepared by adding siRNA (Table 2.3) or 5-10µg plasmid to 1.5ml Opti-MEM® and 20µl of Lipofectamine2000® to 1.5ml Opti-MEM® in separate tubes. Following 5min incubation at room-temperature, these two solutions were mixed and incubated at room-temperature for 20mins to allow complex formation. This transfection solution was then added to the 10cm plate with cells in 5ml Opti-MEM® and the cells were incubated at 37°C 5% CO_2 for 6hours. The Opti-MEM® was then replaced with growth medium and the cells were left to grow for 48hrs before harvesting. For knockdown ChIP assays, the transfection process was repeated twice on consecutive days to ensure a reproducible knockdown.

2.4 Preparation of protein extracts

Cells were cultured in either 10cm plates or 6-well plates to about 80% confluency before harvesting. Preparation was performed on ice, as rapidly as possible, in order to avoid protein denaturation. The maintenance medium was aspirated and cells were washed twice with ice-cold PBS. They were then scraped into microextraction buffer (MEB) (20mM HEPES pH7.8, 150mM NaCl, 25% glycerol, 50mM NaF, 0.2mM EDTA, 0.5% Triton X-100, 0.5% NP40, 1mM PMSF, 1mM DTT, 1µg/ml protease and phosphatase inhibitor cocktail) and transferred to sterile microfuge tubes. 100-500µl of buffer was used per plate depending on cell type and plate size. The cell lysates were then passed through a 26G needle five times and centrifuged at 13000g for 10mins at 4°C. The supernatants were aliquoted and snap-frozen on dry ice, before being stored at -80°C.

2.5 Determination of protein concentrations

The protein concentrations of protein extracts were determined using Bradford's reagent diluted 1 in 5 with dH_2O. For each experiment, a standard curve was constructed by measuring the absorbance at 595nm of 1, 2, 4, 6, 8 and 10µg of BSA in 1ml of diluted Bradford's reagent. The absorbance readings of protein extracts were performed in duplicates at 595nm, and the protein concentration of each sample was determined by plotting the average reading on the standard curve.

2.6 Sodium-dodecyl sulphate polyacrylamide gel electrophoresis

Protein extracts containing 10-20µg of protein per lane were resolved by denaturing SDS-PAGE typically on a 7.8% or 10% polyacrylamide (375 mM Tris pH8.8, 0.1% SDS) minigel with a 4% polyacrylamide (120mM Tris pH6.8, 0.1% SDS) stacking gel. Prior to loading, samples were boiled at 100^{o}C for 2 mins in 1x protein sample buffer (62.5mM Tris pH6.8, 0.5% SDS, 5% ß-mercaptoethanol, 10% glycerol, 0.125% bromophenol blue). Electrophoresis was performed in 1x SDS running buffer (0.1% SDS, 76.8 mM glycine, 10mM Tris pH8.3) at 150V.

2.7 Western blot analysis

After separation by SDS-PAGE, the resolved proteins were transferred to nitrocellulose membranes using a Bio-Rad Mini Trans-Blot Electrophoretic Transfer Cell system. The transfer was carried out in 1x transfer buffer (76.8mM glycine, 10mM Tris pH8.3, 20% methanol) at 50V for 2hrs at 4^{o}C for proteins smaller than 100 kDa and overnight for bigger proteins. The proteins transferred to the nitrocellulose were visualised by staining with 1x Ponceau S solution (Sigma) to ensure their efficient transfer. Subsequently, the membranes were washed with PBS and then blocked in milk buffer (32.5mM Tris, 150mM NaCl, 0.2% Tween-20, 5% skimmed milk powder) for 30-60mins at room-temperature. The membranes were then incubated in the presence of the appropriate primary antibody (Table 2.2) diluted in milk buffer for 2 hours at room-temperature or overnight at 4^{o}C. Following incubation with the primary antibody, the membranes were washed three times for 5mins in PBS 0.5% Tween-20, in order

to remove any unbound primary antibody. The membranes were then incubated for 45mins at room-temperature in the appropriate secondary antibody at a dilution of 1:1000 in milk buffer. The membranes were then washed for approximately 45mins with PBS 0.5% Tween-20 with regular changes of wash buffer. The bound antibodies were then detected with enhanced chemiluminescence (ECL) reagents according to manufacturer's instructions (Amersham).

2.8 Co-immunoprecipitation assay

50µg of protein extracts were used per reaction. The protein extract was incubated overnight at 4^{o}C with 5µg of IgG or appropriate antibody (Table 2.2) and 1µg/ml protease inhibitor cocktail made up to the total volume of 250µl with PBS/NP-40 (PBS, 0.015% NP40). 10Units of DNaseI or heat-denatured DNaseI were added prior to the overnight incubation. 30µl of Dynal® protein G magnetic beads were used for each reaction. The beads were washed three times in PBS/NP-40 using a magnetic rack and then blocked with 50µl of PBS/NP-40 1mg/ml BSA for 15mins at room-temperature using a rotating wheel. Once the blocking buffer was removed from the beads, the appropriate IP sample was added and incubated with the beads for 30mins at room-temperature or 2hrs at 4^{o}C on a rotating wheel. The beads were then washed 4-6 times in PBS/NP-40. To elute the proteins, 50µl of protein sample buffer was added and the beads were incubated for 10mins at 70^{o}C on a shaker. The co-IP samples were collected into fresh microfuge tubes and were resolved next to 20% input (i.e. 10µg protein extract). SDS-PAGE and Western blotted were performed as described above.

2.9 Gene expression analysis

2.9.1 RNA extraction

Total RNA was extracted using TRI reagent (Sigma-Aldrich) according to manufacturer's specifications. The maintenance medium was aspirated and the cells were washed two times in ice-cold PBS. 1ml or 500µl of TRI was used per 10cm plate or 6-well plate respectively, and the cells were harvested with a cell scraper. The samples were incubated for 5mins at room-temperature to allow

complete dissolution of nucleoprotein complexes, before addition of 0.2ml of chloroform per 1ml of TRI. Each sample was thoroughly mixed using a vortex for 1min. The samples were then centrifuged at 13000g for 15mins at 4°C, which resulted in the separation of the samples into three phases. The top aqueous phase containing the RNA was carefully transferred into new microfuge tubes and 0.5ml isopropanol was added to precipitate the RNA. The samples were mixed by vortexing briefly and centrifuged at 13000g for 10mins at 4°C. Once the supernatant had been removed, the RNA pellet was washed in 70% ethanol and briefly dried, before resuspension in 10-15µl RNase-free water. The RNA concentration was determined by spectrophotometry at 260nm, considering that OD of 1 at 260nm corresponds to 40µg/ml of RNA.

2.9.2 cDNA synthesis

cDNA was synthesised using the SuperScriptIII reverse transcriptase (Invitrogen). 0.2µg and 0.4µg of RNA was mixed with 2µl of 1x hexanucleotide mix (Roche) and made up to a final volume of 25µl with RNase-free water. This mix was then incubated at 80°C for 10mins and cooled down rapidly on ice to allow primer annealing. 14µl of reaction mix (8µl of First Strand Buffer (Invitrogen), 4µl of 0.1M DTT and 2µl of 10mM dNTP mix) with 1µl of SuperScriptIII was added to the random primed RNA mix. This was then incubated at 42°C for 1hour to allow cDNA synthesis and the reaction was stopped by boiling the samples at 70°C for 15mins. cDNA was stored at -20°C. In order to rule out DNA contamination of the RNA samples, a duplicate tube for each sample was incubated with the reaction mix without the SuperScriptIII. DNA contamination could be detected in the PCR reaction for the SuperScriptIII negative samples. cDNA prepared with 0.2 and 0.4µg RNA determined the quantitative nature of the cDNA synthesis reaction. RT- PCR primers are described in (Table 2.1).

2.10 Chromatin immunoprecipitation (ChIP) assays.

Cells were grown in 10cm plates to about 80% confluency and approximately one 10cm plate of cells was used per IP. Formaldehyde, to the final concentration of 1%, was added to the cell maintenance medium for 10mins at room-temperature in order to cross-link the protein-DNA complexes. A 15min cross-linking step was performed in ChIP assays for SWI/SNF subunits, SUV39H1 and HP1 proteins. The

cross-linking was stopped by the addition of 0.125M final concentration of glycine and the plates were left at room-temperature for 5mins. The plates were then transferred to ice and the cells were scraped in the culture media/formaldehyde/glycine solution into a 50ml Falcon® tube. The cells were pelleted by centrifugation at 500g for 5mins and washed twice in ice-cold PBS and twice in ice-cold PBS 0.5% NP-40. The cells were then resuspended in high salt buffer (PBS, 0.5% NP-40, 1M NaCl) and incubated in ice for 30mins. This was followed with one wash in PBS 0.5% NP-40 and incubation in low salt buffer (10mM Tris-CL pH8.0, 1mM EDTA, 0.1M NaCl, 0.5% NP-40) in ice for 30mins.

Following this incubation, the cells were pelleted as previously, resuspended in 1ml of low salt buffer, and passed five times through a 26G needle. The final volume was made up to 2.7ml with low salt buffer and 300µl 20% $^{w}/_{v}$ N-Laurylsarcosine sodium salt was added to the cell suspension. Once mixed, this solution was then layered carefully on top of a 40ml LSB/100mM sucrose cushion and centrifuged at 4000g for 10mins at 4°C. The resultant pellet was resuspended in 2ml TE (10mM Tris pH8.0, 1mM EDTA) and passed through another sucrose cushion. The pelleted chromatin was then resuspended in 2ml TE and was sheared into fragments smaller than 0.5kb by sonicating in a water bath sonicator (Bioruptor™, Diagenode) at full power for 30 (30sec on/off) cycles at 4°C. 200µl 11xNET buffer (1.56M NaCl, 5.5mM EDTA, 5.5% NP-40, 550mM Tris-Cl pH7.4) was added to the sonicated material, which was then centrifuged at 13000g for 5mins at 4°C. The supernatant was aliquoted evenly into microfuge tubes and 10% of this was kept as the input sample. 5-25µg of the appropriate antibody (Table 2.2) was added to each labelled aliquot tube and this mix was then incubated overnight at 4°C to allow the antibody protein interaction.

The next day, 50µl of protein A/G sepharose beads (washed three times in 1xNET buffer) were added to the IP samples and were allowed to incubate for a further 2hrs at 4°C. The beads were then recovered on 5ml polypropylene columns, washed twice in ice-cold RIPA buffer (50mM Tris-Cl pH8.0, 150mM NaCl, 0.1% SDS, 0.5% deoxycholate, 1% NP-40), twice with ice-cold LiCl buffer (10mM Tris-Cl pH8.0, 250mM LiCl, 0.5% deoxycholate, 0.5% NP-40, 1mM EDTA) and twice with ice-cold TE. Following removal of excess TE from the last wash, the bound chromatin was eluted with 400µl TE/1%SDS at room-temperature. The eluted

material was digested overnight with 0.125mg/ml proteinase K at 42^0C. DNA purification was performed with Qiagen PCR purification kit using manufacturer's instructions. The ChIP DNA was then analysed by PCR.

2.10.1 Sequential ChIP assay

For sequential ChIP analysis, the primary ChIP antibodies were freshly crosslinked to the sepharose beads using DMP (Pierce) the day before each sequential ChIP reaction. This was done in order to reduce background caused by primary antibody elution following the first ChIP. 100µl of protein G sepharose beads per IP were washed three times in 1ml HEPES lysis buffer (20mM HEPES-NaOH pH8.5, 150mM NaCl, 1% NP-40, 2mM EDTA). 5µg of antibody was added to the beads and the solution was made up to 1ml with HEPES lysis buffer. The beads and the antibody were incubated at 4^oC overnight on a rotating wheel to allow antibody-binding.

The next day, the beads were washed three times in 1ml HEPES lysis buffer and two times in 1ml 100mM HEPES-NaOH pH8.5. The beads were then incubated in 1ml of freshly prepared cross-linking buffer (100mM HEPES-NaOH pH8.5, 10mg/ml DMP) for 1hr at room-temperature on a rotating wheel. Following cross-linking, the beads were washed twice in 1ml HEPES cross-linking wash buffer and the reaction was then quenched with 100µl 1M glycine pH7.5 for 30mins at room-temperature. The beads were then washed twice in HEPES lysis buffer and used for the ChIP reaction in 100mM HEPES-NaOH pH 8.5. Following the first ChIP reaction the eluate in TE/1% SDS was diluted 1 in 10 before the re-ChIP in order to reduce the amount of SDS in the solution.

2.10.2 ChIP-ChOP assay

ChIP DNA was spiked with 100ng of unmethylated PCR product containing a HpaII/MspI site, which was obtained by PCR using pCDNA3 primers on a empty pCDNA3 vector. This cocktail was digested for 1 hr with 1µg of HpaII (Invitrogen) or MspI (NEB). Mock digestions with buffer only were also performed in order to ascertain the digestion efficiency. The digested samples were normalized to the unmethylated DNA in order to normalise for the differences in digestion efficiency between the two enzymes.

2.10.3 ChIP-sequencing

ChIP-bisulphite sequencing for Pol III-binding at Alu loci was performed by Andrew Oler (NIH) using the Illumina Bisulfite sequencing protocol. The Qiagen EpiTech kit was used for bisulfite conversion. ChIP-sequencing data for the Pol III, TFIIIC and SWI/SNF subunits was mined from the UCSC hg18 genome database.

2.11 Methylcollector assay

Separation of genomic DNA according to CpG methylation status was achieved by affinity chromatography with immobilised recombinant MBD2b and MBD3L1 using a MethylCollector™ Ultra kit (Active Motif), according to the manufacturer's specifications.

2.12 Polymerase chain reaction (PCR)

Each PCR reaction had a total volume of 20µl and contained 1µl of template DNA, 20pmol of the appropriate primers, 0.5U of Taq DNA polymerase (Promega), 1 x Taq DNA polymerase buffer (Promega), 1.5mM $MgCl_2$ and 0.2mM dNTPs. The cycling parameters and primers used are listed in Table 2.1. The PCRs were performed using a Dyad® Peltier thermal Cycler (BIO-RAD). The samples were resolved on 5% polyacrylamide native gels next to a 100bp DNA ladder.

Quantitative PCR (qPCR) was performed using the C1000™ Thermal Cycler (BIO-RAD). The qPCR reaction had a total volume of 10µl and contained 0.5µl of template DNA, 5µl of 2 x PerfeCTa™ SYBR® Green FastMix™ and 0.25µM of appropriate primers. An appropriate standard curve encompassing the samples within the linear range was constructed for each qPCR. Each sample was loaded in duplicate to avoid pipetting errors. The expression levels in qRT-PCR assays were obtained using ΔΔC(t) of the average of duplicate samples and the average of loading control, e.g. ARPP P0. The ChIP signal was quantified with the formula (Avg. IP/Avg. Input)-(Avg. Neg. IP/Avg. Input). Data from multiple experiments were then used to calculate overall average and standard deviation values.

Table 2.1: PCR primers for various analyses in mouse and human cells
PCR amplification was performed with the following cycling parameters: 95°C for 30sec, T(a) for 30sec, 72°C for 30sec for n cycles, where T(a) and n are specified for each primer pair. The same qPCR programme was used for all primer pairs, i.e. 95°C for 30sec, 60°C for 30sec, 72°C for 30sec for 45 cycles. The species specificity of a primer pair is mentioned next to its name, i.e. [m] for mouse, [r] for rat and [h] for human primers.

Locus name [Species]		Primers	T(a)	N
28S rRNA [m]	F	CCCGACGTACGCAGTTTTAT	58 °C	23-25
	R	CCTTTTCTGGGGTCTGATGA		
5S rRNA [h, m]	F	GGCCATACCACCCTGAACGC	58 °C	16-18
	R	CAGCACCCGGTATTCCCAGG		
7SL [h, m]	F	GTGTCCGCACTAAGTTCGGCATCAATATGG	70 °C	25-30
	R	TATTCACAGGCGCGATCCCACTACTGATC		
7SL set1 [h]	F	CCGTGGCCTCCTCTACTTG	60 °C	22-27
	R	TTTACCTCGTTGCACTGCTG		
7SL set2 [h]	F	CGTCACCATACCACAGCTTC	60 °C	22-27
	R	CGGGAGGTCACCATATTGA		
7SL set3 [h]	F	GTTGCCTAAGGAGGGGTGA	60 °C	22-27
	R	TCTCTTGAGAGTCCAAAATTAA		
7SL set4 [h]	F	TTTTTGACACACTCCTCCAAGA	60 °C	22-27
	R	ATCTGGTCAAAGCAACATACACTG		
7SL set5 [h]	F	TGCCTCCAGATAAAACTGCTC	60 °C	22-27
	R	ACCCCACTAGAACCCTGACA		
Alu (ch6) [h]	F	CCAGAAAAATTACCAATTAGTTC	53 °C	25-30
	R	GGGCCTATTGACTATGCTTAC		
Alu (ch10) [h]	F	GATTCTCAACAGCAGAATTCCA	53 °C	25-30
	R	CATGTTTGAGAATGTCTACTTC		
Alu (ch19) [h]	F	CCACGTGTTTATCTGTAAGGTG	53 °C	25-30
	R	GTTAGGAGCTAGAAGGAGCCT		
Alu (C19J) [h]	F	CTACTCAAAATATTAAACATAGGC	53 °C	25-30
	R	GCTGCAACGCTGCTATGAAC		
Alu (ch22) [h]	F	GTTCTGACACACTTGGAGAAA	53 °C	25-30
	R	GTTGTTGTTATTGCACAACTCA		
Alu RNA [h]	F	CTTACACGTGTCATCCCAGC	58 °C	30-32
	R	GACAGTGTCTCACTCTGCTACC		
Alu (MLL) [m]	F	CGATTACCCTGTTATCCCTAGGCTGGGCACAGTGGT	60 °C	25-30
	R	AAGCTAGCGGCTGAAATTCTCCTCTTC		
APC [h]	F	GAGGAAGGTGAAGCACTCAGTT	60 °C	23-25
	R	AGGGTGAGACATGGAGAGAAGA		
Apo-E [m]	F	TTCGGAAGGAGCTGGTAAGAC	57 °C	23-25
	R	CGACAGTCCCGTACTCCTTC		
Apo-E [h]	F	CAGCGGAGGTGAAGGACGTC	57 °C	23-25
	R	CTCCTCCTCTCCCCAAG		
Apo-E mRNA [m]	F	GTTTCGGAAGGAGCTGACTG	57 °C	30-32
	R	AGCGCAGGTAATCCCAGAAG		
Apo-E mRNA [h]	F	GGTCGCTTTTGGATTACCT	57 °C	30-32
	R	TTCCTCCAGTTCCGATTTGT		
ARPP P0 [h, m]	F	GCACTGGAAGTCCAACTACTTC	58 °C	16-20
	R	TGAGGTCCTCCTTGGTGAACAC		
B1 [m]	F	TGGTGGTGCATGCCTTTAAT	58 °C	10-12
	R	CCTGGTGTCCTGGAACTCACT		
B2 [m]	F	GGGGCTGGAGAGATGGCT	58 °C	10-12
	R	CCATGTGGTTGCTGGGAT		
B1 (c9) [m]	F	GCATGCATACCACTCCACAC	58 °C	25-30
	R	CAGAGAATCTGCAGTCGTATTTCC		

B2 (c9) [m]	F R	CTGCCTTCAGACACACCAGAAG GATGGAAGAGGTTTTGCCAAG	58 °C	25-30
BRG1 cDNA [m]	F R	TACAGGCTTCAGGCTCGAAT TCTCCAGGGCTGTGTCTCTT	60 °C	
BRM cDNA [m]	F R	TATGTCGCCAATCTGACCAA AGGTCACTCATCTGGCTGCT	60 °C	
GAPDH mRNA [m]	F R	TCCACCACCCTGTTGCTGTA ACCACAGTCCATGCCATCAC	60 °C	23-25
p53BP2 mRNA [m]	F R	GTTGGTTTCGGCGAGAAGG GAAGCCAAGCGAGAACGAG	60 °C	23-25
p16 cDNA [m]	F	TCTGGAGCAGCATGGAGTCC TCGCAGTTCGAATCTGCACC	58 °C	22-25
p21 promoter [m]	F R	CTCTGGGAAGCCAGAAGTTGTT GGTCCAGTCCTGCATCTAAGT	58 °C	25-30
p21 promoter [h]	F R	TATTGTGGGGCTTTTCTG CTGTTAGAATGAGCCCCCTTT	58 °C	25-30
pre-tRNA^{Leu} [h, r, m]	F R	GTCAGGATGGCCGAGTGGTCTAAGGCGCC CCACGCCTCCATACGGAGACCAGAAGACCC	68 °C	20-25
pre-tRNA^{Tyr} [h, r, m]	F R	CCTTCGATAGCTCAGCTGGTAGAGCGGAGG CGGAATTGAACCAGCGACCTAAGGATCTCC	65 °C	20-25
pre-tRNA$_i^{Met}$ [m]	F R	CTGGGCCCATAACCCAGAG TGGTAGCAGAGGATGGTTTC	55°C	20-25
pCDNA3	F R	ATTATGCAGAGGCCGAGG CCATCTTGTTCAATCATGCG	58 °C	20
SNF5 cDNA [m]	F R	AGCGTGTCATCATCAAGCTG CACTGTGGGAAGTGGGTTCT	58 °C	22-25
tRNA^{Leu} downstream	F R	CTTGGGAAGAAAACACTGGCAGTGGT CAGACAGCTTGGTAGTGTGGCCG	60 °C	22-25

Table 2.2 Antibodies used for various analyses in human and mouse cells
The supplier and the catalogue numbers are listed for the antibodies used, alongside the application for which that particular antibody was used. WB stands for Western Blotting. Most antibodies were used at a 1:1000 dilution, except the actin and Tubulin antibodies that were used at a 1:400 dilution for Western blotting.

Protein recognised	Supplier	Cat./I.D. No.	Application
TFIIA	SantaCruz Biotechnology	sc-25365	ChIP
TAF_I48	SantaCruz Biotechnology	sc-6571	ChIP
SUV39H1	SantaCruz Biotechnology	sc-25366	ChIP
SUV39H1	Upstate Biotechnology	05-615	ChIP
HP1	SantaCruz Biotechnology	sc-28735	ChIP
Histone H3	Abcam	ab1791	ChIP
Histone H4	Upstate Biotechnology	07-108	ChIP
BRF1	In house	128	ChIP, co-IP
BRF1	In house	482	WB, co-IP
TFIIIC110	In house	3208	WB, co-IP
TFIIIC220	In house	Ab7	WB, ChIP, co-IP
TFIIIC220	In house	Ab2	WB, ChIP, co-IP
Rpc155	In house	1900	WB, ChIP
MBD1	Imgenex	IMG-306A	ChIP
MBD2	Sigma-Aldrich	M7318	ChIP
MeCP2	Sigma-Aldrich	M9317	ChIP
H3K9me3	Cell Signalling Technology	9754S	ChIP
H3K27me3	Cell Signalling Technology	9756S	ChIP
Ini1	SantaCruz Biotechnology	sc-13055	ChIP, co-IP
Ini1	Abcam	ab12167	WB
BRG1	SantaCruz Biotechnology	sc-10768	ChIP, co-IP
BRG1	Abcam	ab4081	WB
BRM	SantaCruz Biotechnology	sc-28710	ChIP, co-IP
BRM	Abcam	ab15597	WB, co-IP
HDAC1	SantaCruz Biotechnology	sc-7872	ChIP
HDAC2	SantaCruz Biotechnology	sc-7899	ChIP
ac-H3	Upstate Biotechnology	06-599	ChIP
Actin	SantaCruz Biotechnology	sc-1615	WB
TFIIB	SantaCruz Biotechnology	sc-225	WB
α-Tubulin	Sigma-Aldrich	T9026	WB
HA	SantaCruz Biotechnology	sc-805	WB
TBP	In house	4C8/26	WB
c-MYC	SantaCruz Biotechnology	sc-764	ChIP, WB
DNMT1	Abcam	ab87656	WB

Table 2.3: siRNA oligos used for protein knockdowns in mouse cells
The oligo pairs were ordered from Invitrogen and resuspended in RNase-free water. The scrambled control oligos were also ordered from Invitrogen.

Oligo pair name	5' to 3' sequence	
SNF5 siRNA1	Sense oligo	GCUCCGAGGUGGGAAACUAtt
	Antisense oligo	UAGUUUCCCACCUCGGAGCcg
SNF5 siRNA2	Sense oligo	GCUUUUACCUGGAACAUGAtt
	Antisense oligo	UCAUGUUCCAGGUAAAAGCgt
SNF5 siRNA3	Sense oligo	CGAGGUUCUCUGUACAAGAtt
	Antisense oligo	UCUUGUACAGAGAACCUCGga
BRM siRNA1	Sense oligo	GAGCGAAUCCGUAAUCAUAtt
	Antisense oligo	UAUGAUUACGGAUUCGCUCct
BRM siRNA2	Sense oligo	GCUUCUCUGUCACAACGCAtt
	Antisense oligo	UGCGUUGUGACAGAGAAGCat
BRG1 siRNA1	Sense oligo	GGUCAACGGUGUCCUCAAAtt
	Antisense oligo	UUUGAGGACACCGUUGACCat
BRG1 siRNA2	Sense oligo	CACCUAACCUCACCAAGAAtt
	Antisense oligo	UUCUUGGUGAGGUUAGGUGgg

Chapter 3 DNA methylation and Pol III transcription of SINEs

3.1 Introduction

The large numbers of SINEs present within mammalian genomes could potentially provide a huge reservoir for Pol III transcription factors, thus contributing significantly to the global regulation of essential Pol III-transcribed genes. The first clear evidence of Alu transcription by Pol III was provided by Peter Geiduschek and co-workers in 1981, when they *in vitro* transcribed a series of clones containing Alu sequences using a soluble Pol III system (Fuhrman et al., 1981). Despite their relative genomic abundance, very low levels of Pol III-transcribed SINE RNA can be detected in cells (Maraia et al., 1993; Matera et al., 1990; Paulson and Schmid, 1986; Sinnett et al., 1992). Many Alus have highly degenerate promoters and thus were thought to be incapable of recruiting Pol III transcription apparatus. It has been reported that transcription from Alu Pol III promoters may require the presence of conserved 7SL upstream sequence (the first 37 nucleotides upstream of 7SL TSS) (Ullu and Weiner, 1985). A 7SL-like terminator sequence was also shown to increase Alu transcription in cells, thus suggesting that the lack of an efficient terminator may also reduce SINE transcription (Chu et al., 1995). In contradiction with these reports, it has been shown that Alu elements can be strongly transcribed both *in vitro* and *in vivo* (Fuhrman et al., 1981; Liu et al., 1995).

However, DNA methylation has been thought to be primarily responsible for the inhibition of Pol III-driven SINE transcription (Liu et al., 1994; Liu and Schmid, 1993; Schmid, 1991). The Alu consensus sequence contains 24 CpG sites and Alus account for up to one-third of all CpG sites within the human genome. In most normal tissues, these Alu CpGs are found to be methylated (Rubin et al., 1994; Schmid, 1991). *In vitro* transcription assays using nuclear extracts showed that DNA methylation inhibits Alu transcription at low template concentrations, however no such inhibition was observed at higher template concentrations. This lack of inhibition at higher Alu concentrations was attributed to limiting concentrations of MBPs in the extracts (Liu and Schmid, 1993). Methylation of a single CpG within the A-box was seen to inhibit transcription.

Further evidence came from treatment of HeLa cells with 5-azacytidine, a DNA demethylating agent. 5-azacytidine is a cytosine analogue, which incorporates into DNA during replication and inhibits the methylation of cytosine residues by

DNA methyltransferases (Jones and Taylor, 1980). Treatment with 5-azacytidine was shown to cause a 5-to 8-fold increase in full length Alu transcripts (Liu et al., 1994). K562 cells, which show considerably lower level of DNA methylation on Alus than other cell types, have highly elevated levels of Alu transcripts. ActinoMycin D treatment was used to show that this abundance is due to increased levels of SINE transcription and not increased transcript stability. Two independent studies have also shown that methylated Alu sequences transiently transfected into cells are transcribed at a reduced rate compared to their unmethylated counterparts (Li et al., 2000; Yu et al., 2001). SINEs have also been shown to be target sequences for MBPs, which may contribute to the transcriptional repression. The fraction of human genomic DNA fragments that was retained by an MBD column was found to be highly enriched for Alu sequences (Brock et al., 1999). MeCP2 was also shown to selectively bind Alu sequences by ChIP assays in MCF7 cells (Koch and Stratling, 2004).

Most lines of evidence for DNA methylation-mediated inhibition of SINE transcription arise from *in vitro* and artificial systems. Not only is there a distinct lack of endogenous data, but also there are many unexplained curiosities within the existing literature. For example, an unmethylated PV Alu repeat stably transfected into mouse cells was found to be transcriptionally silent (Leeflang et al., 1992). Demethylated Alu templates transiently transfected into HeLa cells were inactive, whereas these templates were highly expressed in 293 cells (Liu et al., 1995). Co-transfection of methylated Alu sequences with the MeCP2 TRD relieved the methylation mediated transcriptional repression, rather than compounding it (Yu et al., 2001). Cell stress cannot alleviate the transcriptional repression of transiently transfected methylated Alus (Li et al., 2000), although cell stress is able to elevate endogenous SINE transcription without altering DNA methylation (Chu et al., 1998; Liu et al., 1995).

Therefore, it seems we still lack a clear understanding of how DNA methylation regulates SINE transcription in cells. Increasing amounts of evidence are emerging for the role of SINE transcripts in modulating various cellular processes. Therefore it is increasingly important we try to understand the transcriptional regulation of these sequences. This chapter will try and address this question in detail.

3.2 Results

3.2.1 Alu sequences are occupied by Pol III transcription apparatus and MBPs in multiple human cell lines

Since low levels of Pol III-synthesised SINE transcripts are clearly detected in cells (Paulson and Schmid, 1986), ChIP assays for Pol III transcription machinery were performed in order to investigate the polymerase loading on these elements. HeLa cells were subjected to ChIP analysis with antibodies against the Rpc155 subunit of Pol III, the BRF1 subunit of TFIIIB and the 220kDa subunit of TFIIIC (Figure 3.1). ChIP with an antibody against histone H3 was performed to provide a positive control IP. ChIPs with beads only (mock) and an antibody against RNA polymerase I specific transcription factor TAF_I48 (Rudloff et al., 1994) were performed to provide the negative controls.

Analysis was performed at five randomly-chosen Alu loci at five independent genomic locations. The primers were designed within the flanking unique DNA sequences to allow specific amplification of these loci. Alu(ch19) is a primer set aimed at an Alu located centrally within a long stretch of tandem Alu repeats. The selected Alu elements all belong to the AluS subfamily and have 7-13 CpGs each. Primers that amplify two copies of the human 7SL gene, an actively-transcribed Pol III gene, provide the positive control for Pol III transcription apparatus occupancy. The apolipoprotein E precursor (Apo-E) gene, which is a Pol II-transcribed locus and is known to be silenced by DNA methylation (Ballestar et al., 2003), provides a negative control for Pol III transcription apparatus occupancy.

Surprisingly, the enrichment of Pol III transcription machinery at Alu loci is clearly above background. The Apo-E gene does not show any enrichment for Pol III transcription apparatus, showing the specificity of the ChIP reaction. The 7SL locus shows a considerably higher enrichment for Pol III and TFIIIB compared to the Alu loci. The Alu consensus sequence shows 67% reduction ($p<0.05$) in Pol III occupancy compared to the 7SL loci (data not shown). Individual Alu sequences show much lower polymerase occupancy compared to the 7SL loci, whereas no such differences are observed for TFIIIC occupancy.

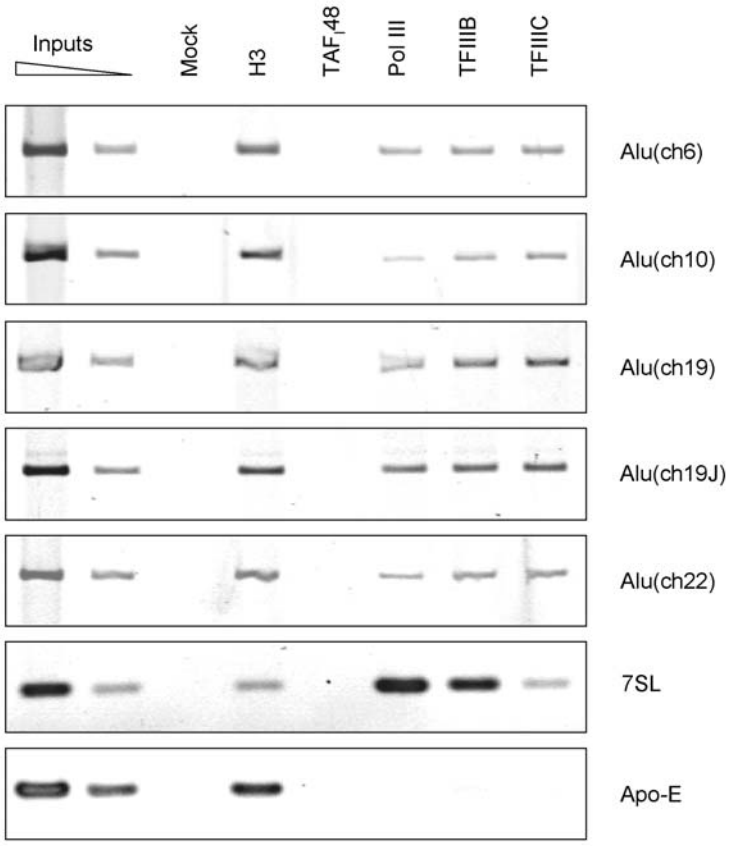


Figure 3.1: Pol III transcription apparatus occupies Alu SINEs in HeLa cells
ChIP assay in HeLa cells with antibodies against Rpc155, BRF1 and TFIIIC220 at Alu loci on chromosome 6, 10, 19 and 22. Alu(ch19J) is also on chromosome 19. 7SL and Apo-E loci respectively provide the positive and negative controls for Pol III transcription apparatus occupancy. ChIPs with beads only (Mock) and TAF$_I$48 antibody provide the negative controls, whereas ChIP with histone H3 antibody provides the positive control. Serial dilutions of the inputs show the quantitative nature of the PCR.

The presence of Pol III transcription apparatus at various SINEs was startling, seeing that the transcription of these sequences is believed to be inhibited by DNA methylation. Previous studies have suggested that this inhibition may be in-part due to the presence of MBPs at SINEs (Brock et al., 1999; Koch and Stratling, 2004). In order to investigate this possibility, HeLa cells were subjected to ChIP assays with antibodies against MBD1, MBD2 and MeCP2 (Figure 3.2). Beads only and TAF_I48 ChIPs provide the negative control IPs, whereas histone H3 ChIP provides the positive control. The Apo-E locus provides a positive control for the presence of MBPs (Ballestar et al., 2003), whereas 7SL locus provides the negative control.

Alu loci, previously showing the occupancy of Pol III transcription apparatus, also seem to be enriched for all three MBPs. The 7SL locus does not show any MBP enrichment, indicating the specificity of the assay. MeCP2 occupancy on Alus is higher than MBD1 and MBD2. However, this was not due to higher IP efficiency of the MeCP2 antibody since the opposite is observed on the Apo-E locus, which shows considerably higher MBD1 occupancy. The presence of MBPs also suggests that these specific Alu loci are methylated.

The enrichment of Pol III transcription apparatus and MBPs at these Alus is intriguing, however this could be a cell type-specific effect. In order to rule out this possibility, ChIP assays with antibodies against Pol III transcription apparatus and MBPs were performed in IMR90 human diploid fibroblasts (Figure 3.3), HCT116 colon carcinoma (data not shown) and HEK293 embryonic kidney cells (data not shown). All human cells lines analysed show the occupancy of Pol III transcription apparatus and MBPs on Alu SINEs. The Pol III enrichment at Alus is consistently lower than at the 7SL locus, whereas no such difference is observed for transcription factor occupancy. MBPs show variable enrichment at SINEs in different cell lines, however MBP occupancy is always observed at SINEs.

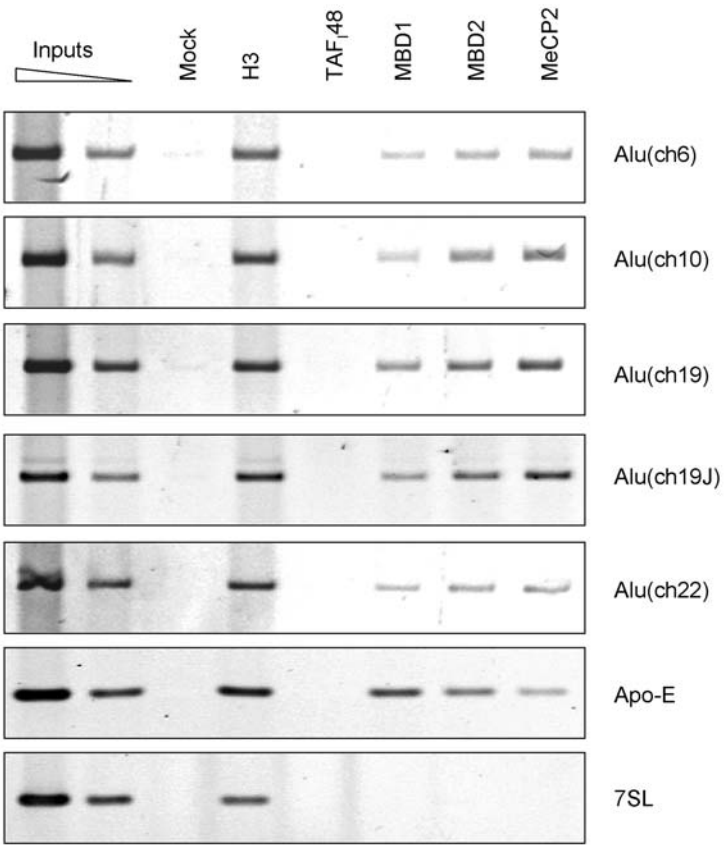


Figure 3.2: MBPs occupy Alu SINEs in HeLa cells
ChIP assay in HeLa cells with antibodies against MBD1, MBD2 and MeCP2 at Alu loci on chromosome 6, 10, 19 and 22. Alu(ch19J) is also on chromosome 19. Apo-E and 7SL loci respectively provide the positive and negative controls for MBP occupancy. ChIPs with beads only (Mock) and TAF$_I$48 antibody provide the negative controls, whereas ChIP with histone H3 antibody provides the positive control. Serial dilutions of the inputs show the quantitative nature of the PCR.

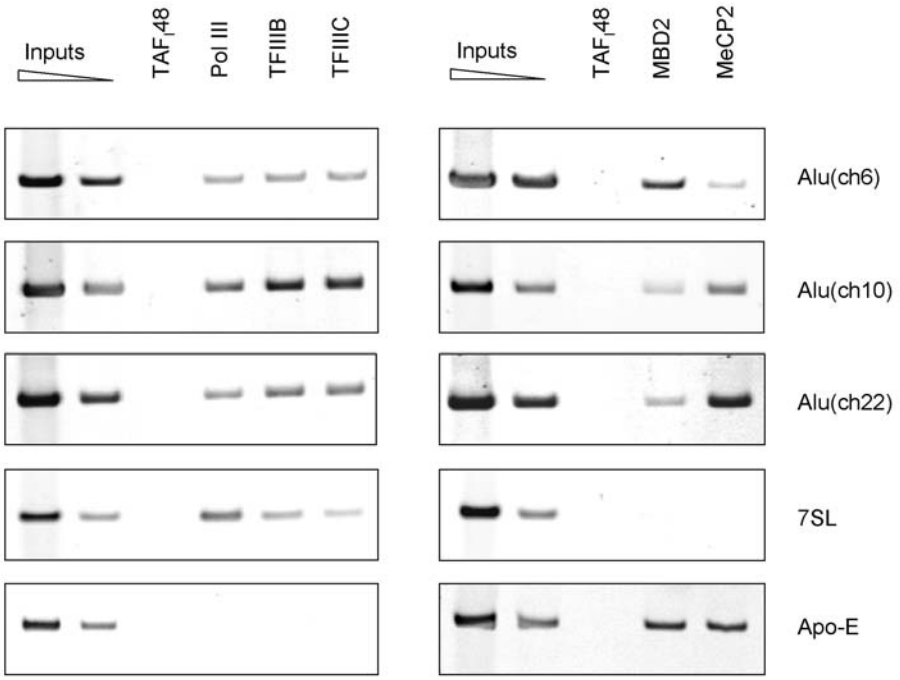


Figure 3.3: Pol III transcription apparatus and MBPs occupy Alu SINEs in IMR90 cells
ChIP assay in IMR90 cells with antibodies against Rpc155, BRF1, TFIIIC220, MBD2 and MeCP2 at Alu loci on chromosome 6, 10, 19 and 22. 7SL and Apo-E loci respectively provide the positive and negative controls for Pol III transcription apparatus occupancy and visa versa for MBP occupancy. ChIP for TAF$_I$48 provides the negative control. Serial dilutions of the inputs show the quantitative nature of the PCR.

3.2.2 B1 and B2 SINEs are occupied by Pol III transcription apparatus and MBPs in multiple mouse cell lines

Steady state expression of mouse B1 RNA is low in cultured cells (Maraia, 1991), whereas B2 transcripts are a little more abundant (Ryskov et al., 1983). There are some reports suggesting that B1 sequences in the mouse genome are methylated (Jeong and Lee, 2005; Yates et al., 1999), however there is not much evidence in the literature for B2 sequence methylation. Both B1 and B2 SINE transcript levels are upregulated in response to cell stress (Li et al., 1999). This indicates that murine SINEs may be regulated in the same manner as their human counterparts. However, nothing is known about the regulation of B1 and B2 SINE transcription by DNA methylation.

Since Pol III transcribed B1 and B2 RNA can be clearly detected in cells, the occupancy of Pol III transcription apparatus was investigated on these mouse SINEs. Mouse fibroblasts were subjected to ChIP assays with antibodies against Pol III transcription apparatus (Figure 3.4). Primers against the consensus B1 and B2 sequences were used to detect enrichment on most mouse SINEs. Moreover, enrichment at two specific SINEs, B1(ch9) and B2(ch9), was also analysed using primers designed against flanking DNA. Primers amplifying three different copies of the murine 7SL gene provide the positive control for Pol III transcription machinery.

As observed on human Alu SINEs, B1 and B2 SINES also show clear enrichment of Pol III and the transcription factors TFIIIB and TFIIIC. The negative control Apo-E shows no enrichment for the transcription apparatus, indicating the specificity of the assay. Pol III enrichment is considerably lower on SINEs compared to the actively transcribed 7SL loci, however no such difference is seen in TFIIIB and TFIIIC occupancy. B2 consensus sequence shows higher Pol III loading compared to its B1 counterpart ($p<0.05$). This indicates that the comparatively higher levels of B2 transcripts seen in mouse cells (Ryskov et al., 1983) could be due to increased Pol III loading onto these elements. The presence of Pol III on B1 and B2 sequences was also confirmed in mouse ES cells (Figure 3.14).

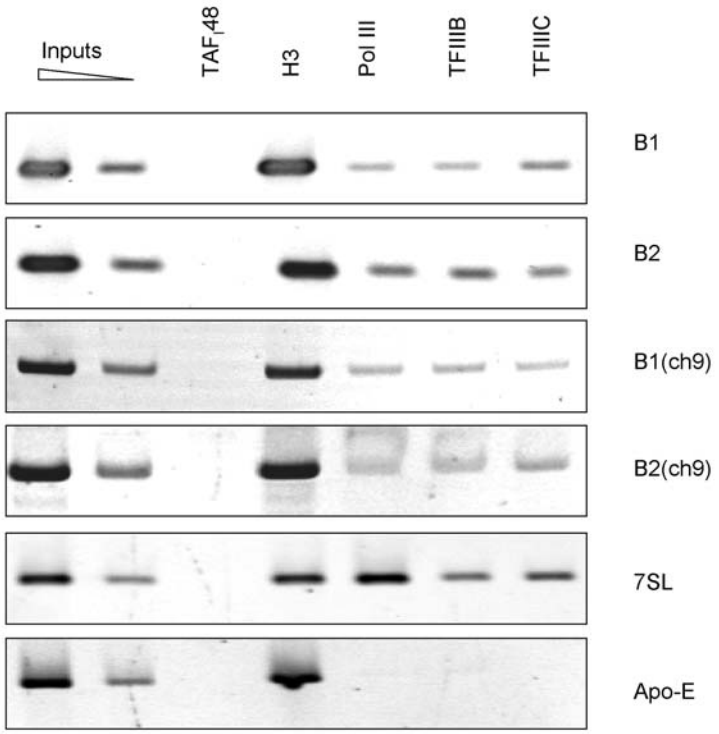


Figure 3.4: Pol III transcription apparatus occupies B1 and B2 SINEs in MEFs
ChIP assay in MEFs with antibodies against Rpc155, BRF1 and TFIIIC220 at B1 and B2 genomic consensus sequences. B1(ch9) and B2(ch9) are specific SINEs on chromosome 9. 7SL and Apo-E loci respectively provide the positive and negative controls for Pol III transcription apparatus occupancy. ChIPs with TAF$_I$48 antibody and histone H3 antibody provide the negative and positive controls, respectively. Serial dilutions of the inputs show the quantitative nature of the PCR.

The binding of MBPs to B1 and B2 sequences in mouse fibroblasts was also investigated using ChIP assays (Figure 3.5). Clear enrichment of all three MBPs is detected on mouse SINEs, suggesting that both murine SINE families are methylated. The 7SL locus does not show any enrichment, indicating the specificity of the reaction. Compared to MBD1 and MBD2, MeCP2 showed higher enrichment on SINEs. However the positive control, Apo-E, shows higher MBD1 binding and lower MBD2 and MeCP2 occupancy. Elevated MeCP2 occupancy seems to be a feature of SINEs. No significant differences are observed in MBP loading between B1 and B2 sequences ($p>0.1$). So in spite of having similar levels of MBPs, B2 SINEs are still capable of higher polymerase loading.

Thus it seems that both human and murine SINEs are occupied by Pol III transcription apparatus and MBPs.

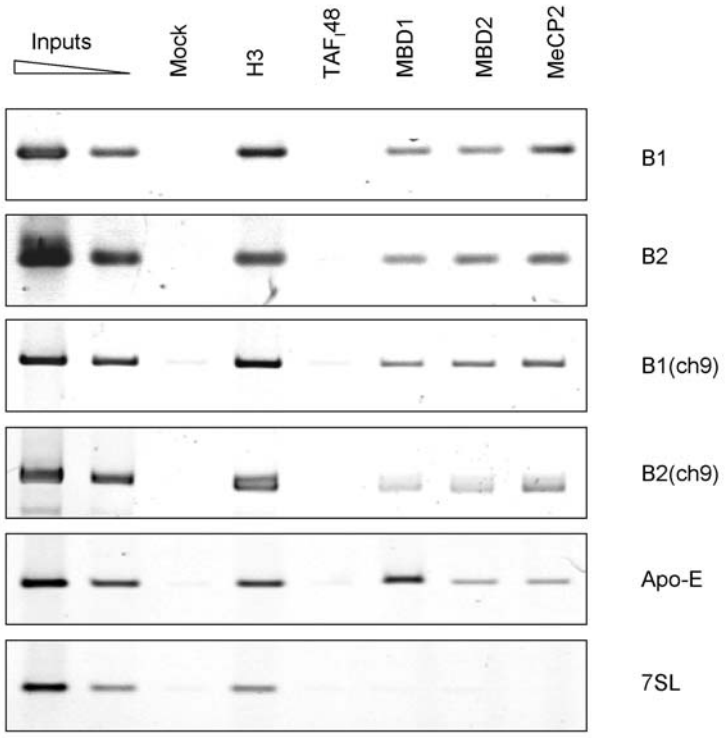


Figure 3.5: MBPs occupy B1 and B2 SINEs in MEFs
ChIP assay in MEFs with antibodies against MBD1, MBD2 and MeCP2 at B1 and B2 genomic consensus sequences. B1(ch9) and B2(ch9) are specific SINEs on chromosome 9. Apo-E and 7SL loci respectively provide the positive and negative controls for MBP occupancy. ChIPs with beads only and TAF$_I$48 antibody provide the negative controls, whereas histone H3 antibody provides the positive control. Serial dilutions of the inputs show the quantitative nature of the PCR.

3.2.3 Human and mouse cells contain sufficient Pol III transcription factors to occupy SINEs

SINEs are very abundant in human and mouse genomes. If all these loci are enriched for Pol III transcription apparatus, then the cell would require a huge reservoir of these transcription factors. Previous estimates in yeast have shown TBP, BRF1 and Bdp1 to be equimolar with ~1000 molecules per cell each (Sethy-Coraci et al., 1998). In HeLa cells, the number of TBP molecules has been estimated to be 1.9 $x10^5$ per cell (Kimura et al., 1999). So there seems to be sufficient TBP in a cell. But are there enough of Pol III specific transcription factor molecules per cell?

Table 3.1: TBP, BRF1 **and TFIIIC110 protein quantification.**
HeLa and MEF total cell extracts were subjected to quantitative western blot analysis. The number of molecules per cell was estimated using total cell extracts from a known cell number and a recombinant protein titration. Errors represent standard deviations. n=3 [TFIIIC110 (n=1) data provided by Damian Grazcyk]

Molecules/Cell	HeLa	MEF
TBP	$5.1x10^5 \pm 1.6x10^5$	$2.9x10^6 \pm 0.2x10^5$
BRF1	$6.4x10^5 \pm 0.9x10^5$	$1.4x10^6 \pm 2.8x10^5$
TFIIIC110	$5x10^5$	

In order to quantify the number of molecules of Pol III transcription factors per cell, HeLa and MEF total cell extracts obtained from a known number of cells were used alongside titrations of purified recombinant proteins in Western blot analysis. Table 3.1 shows the estimated molecules per cell counts for TBP, BRF1 and TFIIIC110. All three proteins have roughly half a million molecules per HeLa cell. MEFs have an estimated 3million TBP molecules and 1.4million BRF1 molecules. Thus it seems that these transcription factors have enough protein molecules per cell to occupy all SINE sequences.

3.2.4 The presence of MBPs and DNA methylation does not inhibit polymerase loading onto SINEs

The data so far suggest that SINEs in human and mouse cells occupied by Pol III transcription apparatus are also enriched for MBPs. However, these two protein subsets could be binding to the same locus at different times or in different cells, i.e. not simultaneously. In order to investigate this possibility, sequential ChIP assay was performed on HeLa cells (Figure 3.6). The primary ChIP was performed with an antibody against the 155kDa subunit of Pol III. TAF_I48 antibody provided the negative control for this IP and did not show any enrichment at Pol III transcribed sequences, indicating the specificity of the primary reaction (data not shown).

The chromatin obtained from the Pol III ChIP was then subjected to sequential ChIP with antibodies against MBPs. 100% input was taken from the primary ChIP material in order to make the input dilutions. Secondary ChIPs with Pol III and BRF1 antibodies provide the positive controls. The Pol III re-ChIP signal is nearly as strong as the undiluted input, indicating the specificity of the primary ChIP and the efficiency of the secondary reaction.

Sequential ChIPs show clear enrichment of MBPs on Pol III-bound Alu loci, indicating their co-occupancy. The 7SL locus, which is clearly enriched for BRF1 and Pol III in the sequential ChIP, remains negative for MBP occupancy. Alu(ch22) shows a distinct lack of MBD1 co-occupancy. This was due to experimental variation, and other repeats showed MBD1 co-occupancy at this locus. Different Alu loci show variable MBP co-occupancy in different experiments; however MBP co-occupancy is almost invariably seen. Thus it seems that the presence of MBPs is not sufficient to exclude Pol III from Alu sequences.

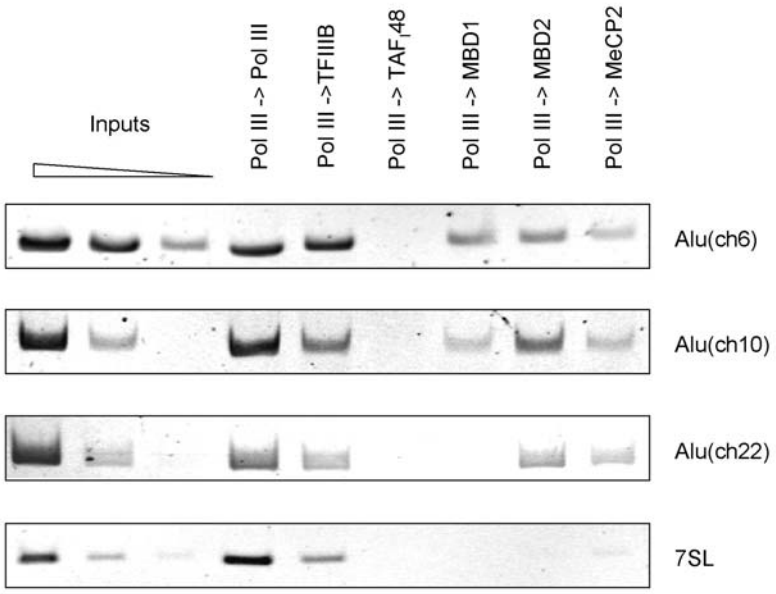


Figure 3.6: Pol III transcription apparatus and MBPs co-occupy Alu SINEs
Sequential ChIP assay in HeLa cells where the first ChIP performed with antibody against Rpc155 was followed with a second ChIP with antibodies against MBD1, MBD2 and MeCP2. TAF_I48 provides the negative control for the first (data not shown) and second ChIPs. Re-ChIPs with Rpc155 and BRF1 antibodies provide the positive controls. 7SL locus provides the negative control for MBP occupancy. Serial dilutions of the inputs show the quantitative nature of the PCR.

DNA methylation can directly inhibit the binding of transcription factors to their cognate sequences (Watt and Molloy, 1988). Since methylation of a single CpG within the A-box was seen to inhibit transcription (Liu and Schmid, 1993), this could be the mode of transcriptional inhibition at SINEs. In order to find out if Pol III transcription apparatus can occupy methylated DNA, material obtained from Pol III and TFIIIC ChIP was subjected to methylation-sensitive digests (Figure 3.7). Restriction endonucleases HpaII and MspI are isoschizomers that cleave at CCGG sequences. Cleavage by HpaII is prevented by the presence of a 5-methyl group at the internal cytosine of its recognition sequence, whereas MspI can cleave the DNA irrespective of its methylation status (Waalwijk and Flavell, 1978).

HeLa DNA obtained from Pol III, TFIIIC, MBD2 and MeCP2 ChIPs was spiked with 100ng of an unmethylated PCR product containing a central CCGG sequence. This cocktail was then digested with HpaII and MspI for 1 hour. PCRs against the unmethylated 'spike' were used to normalise for the differences in the digestion efficiency between the two enzymes. These normalised samples were then subjected to PCRs targeting an Alu sequence that contains two CCGG sites, one near the 5' end and another towards the 3' end of its sequence.

It can be seen that the Alu(ch6) DNA in the input sample is resistant to HpaII digestion compared to MspI digestion, indicating that this sequence is methylated in HeLa cells. As expected, the DNA from the positive control MBD2 and MeCP2 ChIPs is also methylated. Moreover, this resistance to HpaII digestion remains following Pol III and TFIIIC ChIPs. Surprisingly, this suggests that Pol III and TFIIIC are capable of occupying methylated Alu sequences.

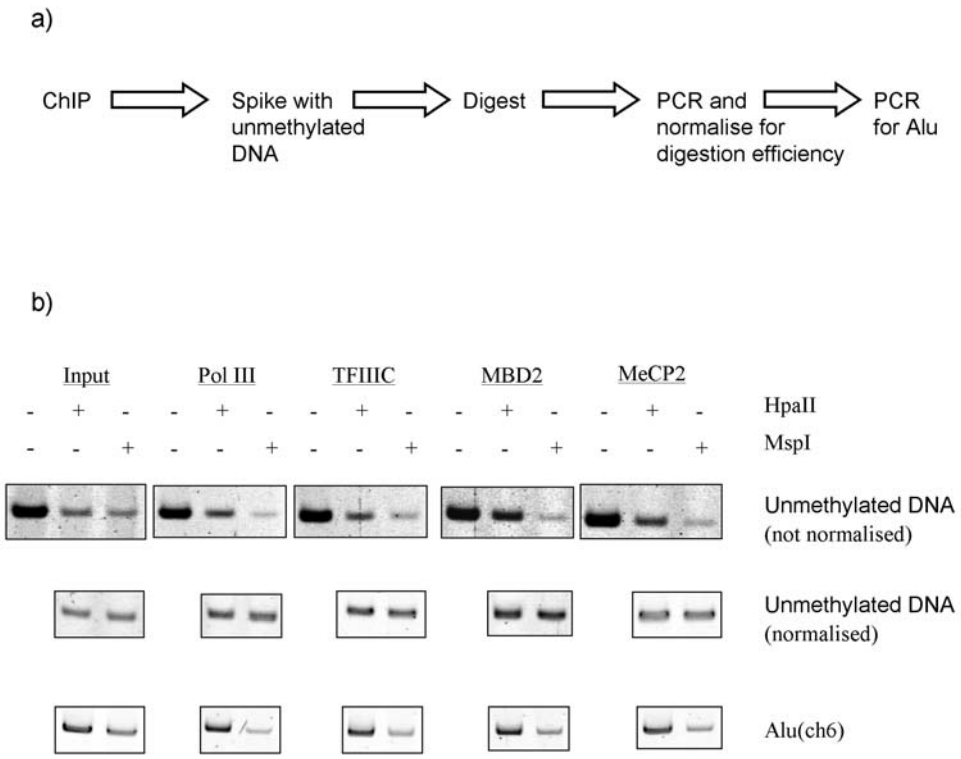


Figure 3.7: Pol III transcription apparatus occupies methylated Alu DNA
a) Schemaic for the ChIP-ChOP assay. b) ChIP-ChOP assay in HeLa cells where ChIP with antibodies against Rpc155, TFIIIC220, MBD2 and MeCP2 was followed by restrictions digests with DNA methylation-sensitive endonuclease HpaII and its methylation-insensitive isoschizomer MspI. The ChIP DNA was spiked with 100ng of unmethylated PCR product prior to digestion, which was then used to normalise for digestion efficiency. PCR for Alu(ch6) was performed using the normalised sample dilutions.

It is unlikely that only the CpGs within the restriction endonuclease recognition sites are methylated and the surrounding CpGs are unmethylated. However, the enzymatic approach is a crude one and in order to thoroughly investigate the methylation status of the polymerase-bound SINEs, ChIP bisulphite sequencing was performed on input DNA and Pol III ChIP DNA (Figure 3.8) by Andrew Oler from Bradley Cairns Lab.

Sodium bisulphite treatment of fully denatured DNA causes the conversion of cytosine residues to uracil. However, 5-methylcytosine residues are non-reactive to sodium bisulfite. Upon sequencing the DNA, one could determine the methylation status of the CpGs by looking for unconverted cytosine residues (Grunau et al., 2001). With the advent of high-throughput and massively parallel sequencing techniques, bisulphite sequencing has become a powerful technique and has been successfully used for methylation analysis of various genomes (Bernstein et al., 2007; Smith et al., 2009). ChIP-bisulphite sequencing is a variation on this technique, where DNA obtained from ChIP is converted with sodium bisulphite and sequenced. This is used to shed light on the methylation status of DNA bound to particular proteins.

Prior to bisulphite treatment, both input and Pol III ChIP DNA were spiked with unmethylated *lambda* DNA. The efficiency of bisulphite conversion was determined to be 99.87% based on the conversion of cytosine residues in this 'spike' DNA. Bisulfite conversion has been previously shown to introduce single strand breaks, thus making the DNA prone to breaking (Munson et al., 2007). Bisulfite treatment is also thought to cause DNA degradation through 5mC-selective hydrolysis and depyrimidination (Tanaka and Okamoto, 2007). Due to these technical shortcomings, small insert size distribution (median 80bp, peak 60bp) was observed in the ChIP bisulfite sequencing data. Unique alignments depend on fragments covering unique flanking regions alongside the consensus sequence. The small insert size resulted in low numbers of unique alignments for Alu sequences, which meant the overall coverage of Alus was low. Due to these technical limitations the sequencing data was aligned to 14 consensus sequences for different Alu families (Figure 3.8).

All Alu families show higher level of DNA methylation (5% to 14%) compared tRNA genes (0% to 5%). Alus are also generally less enriched for Pol III than tRNA genes, which is in agreement with the data previously obtained by ChIP-PCR. As expected, Pol III ChIP DNA does not show a decrease in Alu methylation compared to the Input sample. Input DNA also shows the presence of CpG methylation on A- and B-box sequences. However, Pol III ChIP sample also shows the presence of DNA methylation at these crucial regulatory sequences.

The newer AluY elements have higher methylation than their older AluJ or AluS counterparts. The same pattern is observed for A- and B-box methylation, where B-box methylation is 13% in AluJ, 30-50% in AluS and 55-65% in AluY. Thus it seems Alu methylation diminishes with age. However, the younger Alu elements also show higher Pol III loading compared to the older families. AluY families show more than two-fold higher Pol III occupancy than AluJ families (data not shown). Thus the newer Alu families are able to load more Pol III despite having a higher level of methylation.

These data strengthen the opinion that DNA methylation, including methylation within the A- and B-box sequences, is unable to inhibit Pol III loading onto SINEs.

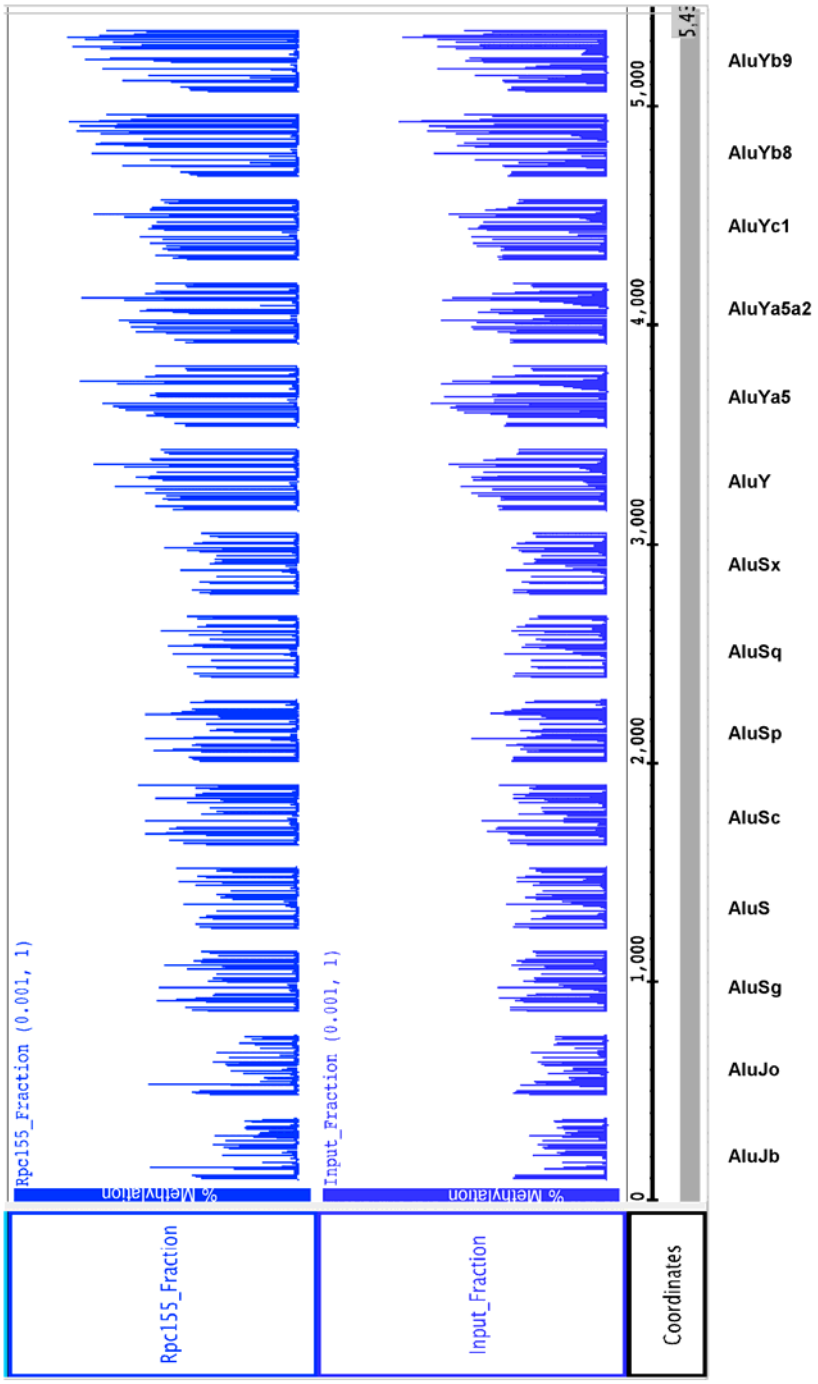


Figure 3.8: ChIP-bisulphite sequencing analysis of Alu consensus sequences
Bisulphite sequencing of input and Rpc155 ChIP DNA from HeLa cells aligned to 14 consensus sequences for various SINE families. 99.87% efficiency was calculated for the bisulphite conversion reaction. The Y-axis represents % of unconverted cytosine residues obtained in the Input fraction and the Pol III ChIP fraction (i.e. Rpc155_Fraction) respectively (Data provided by Andrew Oler).

3.2.5 Loss of DNA methylation neither increases Pol III loading onto SINEs, nor does it elevate SINE expression

From the evidence so far, it is clear that human and mouse SINEs are occupied by Pol III transcription apparatus and this polymerase loading is not prevented by the presence of MBPs or DNA methylation. However, this is contrary to what is suggested by the existing literature, where SINE transcription is reported to be inhibited by DNA methylation (see section 3.1). So what does the absence of DNA methylation on SINEs do to the polymerase loading onto these elements? And how does it affect their cellular transcript levels?

Dnmt1$^{-/-}$ mice were reported to be embryonically lethal, however ES cells carrying this mutation showed no detectable effect on viability or proliferation in culture (Li et al., 1992). Deletion of the first exon of *Dnmt1* led to 80% genomic demethylation, where the residual DNA methylation was attributed to activity of *de novo* methyltransferases DNMT3a and DNMT3b (Lei et al., 1996). It has been observed that only Dnmt1 mutants show this marked loss of genomic methylation. *Dnmt3a* and *Dnmt3b* double knockout ES cells show disruption of *de novo* methylation however this does not cause demethylation at imprinted loci (Okano et al., 1999), whereas loss of DNMT2 shows little effect on genomic methylation (Okano et al., 1998).

In cortrast to ES cells, differentiated cells do not tolerate the loss of DNMT1 well. Deletion of *Dnmt1* in proliferating neural cells lead to functional impairment and poor survival (Fan et al., 2001). Fibroblast cells undergo premature senescence in the absence of DNMT1 through a p53-dependent mechanism (Jackson-Grusby et al., 2001). Lande-Diner *et al* obtained a viable *Dnmt1* null MEF cell line by crossing *Dnmt1*$^{-/-}$ animals to a *p53*$^{-/-}$ null background (Lande-Diner et al., 2007). These cells grow at normal exponential rates and do not show senescence even after 100 passages. Since these cells are differentiated, they lack *de novo* methyltransferase activity and have no residual maintenance methyltransferase. Due to progressive loss of DNA methylation, their genome was shown to be severely hypomethylated compared to control *p53*$^{-/-}$ fibroblasts. Thus, these cells provide an ideal 'methylation free' environment to study the effect of DNA methylation on SINE transcription.

The levels of genomic methylation in *p53*$^{-/-}$ and *Dnmt1*$^{-/-}$ *p53*$^{-/-}$ MEFs were analysed by methylation-sensitive restriction digests (Figure 3.9). 1µg of genomic DNA obtained from both cell types was subjected to 30 minute digestion with HpaII and MspI. The digested material was then resolved on a 1% agarose gel alongside mock-digested DNA. The undigested material stays at the top of the lane, whereas MspI-digested DNA forms a smear in the entire lane lacking a clear 'table' at the top. It can be clearly seen that the genomic DNA from *Dnmt1* null MEFs is readily digested by HpaII, unlike its counterpart from the control cells. Moreover, the smear formed by HpaII digest is very similar to the one formed by MspI digestion, suggesting that the *Dnmt1*$^{-/-}$ genome is almost completely demethylated.

In order to study the effect of the loss of DNA methylation on SINEs, both the *p53*$^{-/-}$ and *Dnmt1*$^{-/-}$ *p53*$^{-/-}$ cells were subjected to ChIP assay with antibodies against Pol III transcription apparatus (Figure 3.10). ChIP with MBD2 antibody provided the control for the DNA methylation status. The ChIPs from the two different cell types were normalised to serial input dilutions. Apo-E locus was unoccupied by Pol III in both the control and null cells. Moreover, MBD2 occupancy on the Apo-E locus shows a dramatic reduction in the *Dnmt1* null cells, indicating the loss of DNA methylation. The occupancy of transcription apparatus on the positive control, 7SL, does not alter upon DNMT1 loss. This suggests that there is no overall change in Pol III loading at generic Pol III transcribed loci.

Surprisingly, neither B1 nor B2 sequences show any change in polymerase loading. This is also confirmed on the specific B1 and B2 loci on chromosome 9. There are no alterations in TFIIIB or TFIIIC loading on SINEs either, whereas MBD2 occupancy on SINEs is clearly diminished, indicating that these sequences are demethylated. This result, contrary to expectations, suggests that loss of DNA methylation does not cause any elevation in Pol III loading onto SINEs.

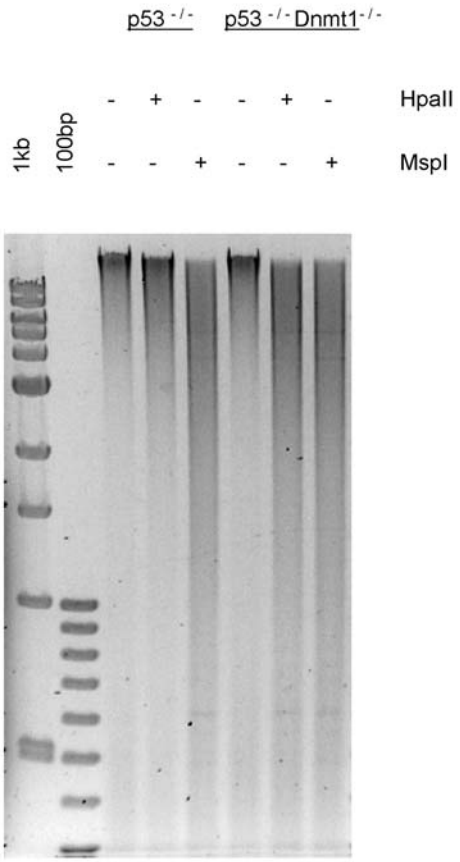


Figure 3.9: *Dnmt1*$^{-/-}$ cells have markedly reduced DNA methylation
Restriction digests on 1µg genomic DNA from *p53*$^{-/-}$ and *p53*$^{-/-}$ *Dnmt1*$^{-/-}$ cells with DNA methylation-sensitive endonuclease HpaII and its methylation-insensitive isoschizomer MspI.

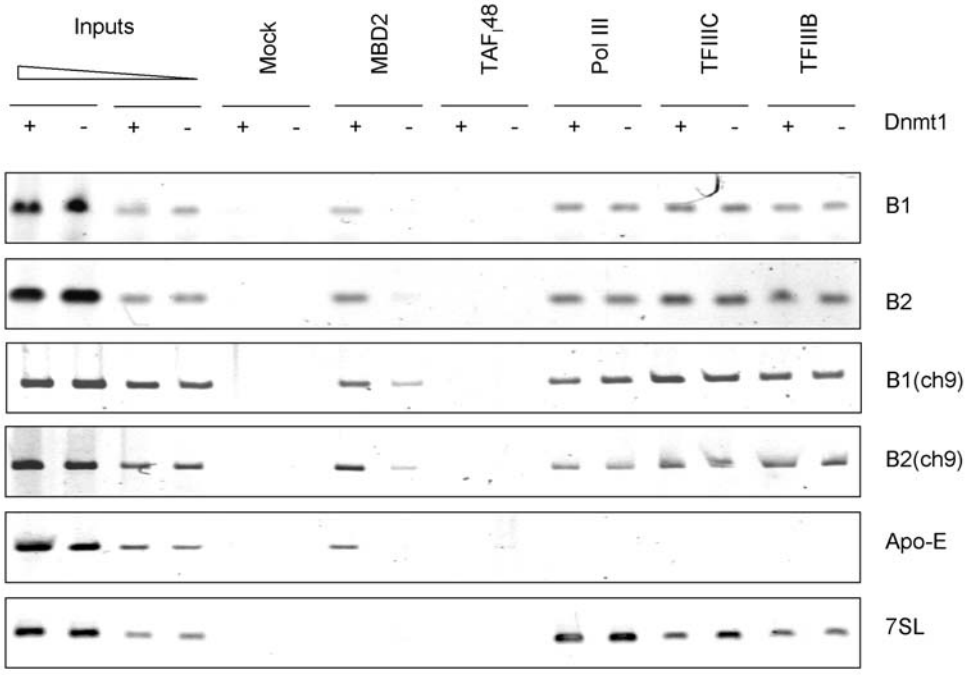


Figure 3.10: Loss of DNA methylation and MBP occupancy does not alter Pol III transcription apparatus occupancy on SINEs
ChIP assay in *p53*$^{-/-}$ and *p53*$^{-/-}$ *Dnmt1*$^{-/-}$ MEFs with antibodies against Rpc155, TFIIIC220 and BRF1 at B1 and B2 genomic consensus sequences. B1(ch9) and B2(ch9) are specific SINEs on chromosome 9. ChIP with MBD2 antibody indicates the DNA methylation status of the locus. Apo-E and 7SL loci respectively provide the positive and negative controls for MBP occupancy and visa versa for Pol III transcription apparatus. ChIPs with beads only and TAF$_I$48 antibody provide the negative controls. The ChIPs were normalised to serial dilutions of the inputs.

It is possible that a few of the B1 and B2 SINEs in the *Dnmt1*$^{-/-}$ cells have elevated polymerase loading, but this increase can not be detected by the consensus primers. This is a limitation of the ChIP PCR approach and ChIP sequencing is required to detect these genome wide locu-specific changes. However, increased transcription at these few SINEs would still lead to a detectable increase in cellular transcript levels.

SINEs are widely transcribed by Pol II as a part of mRNA transcripts (Yulug et al., 1995). In order to study the effects of DNA methylation loss on Pol III-driven SINE transcription, one must discriminate between Pol II-transcribed and Pol III-transcribed RNA. Traditional approaches have used primer extension technique, where the larger Pol II-transcribed RNA extension products can be resolved from the smaller Pol III-transcribed RNA products (Liu et al., 1994).

At low concentrations, α-amanitin specifically inhibits Pol II transcription whilst allowing normal Pol I and Pol III transcription (Lindell et al., 1970). Therefore in order to discriminate between Pol II and Pol III driven SINE transcription, *p53*$^{-/-}$ and *p53*$^{-/-}$ *Dnmt1*$^{-/-}$ cells were treated with 50μg/ml α-amanitin for 24hrs. Any changes observed in the SINE transcript levels following this treatment cannot be a response of genes transcribed by Pol II. Total RNA from mock-treated and α-amanitin-treated cells was subjected to RT-PCR analysis using primers against B1 and B2 consensus transcripts (Figure 3.11). The reaction was normalised to Pol I-transcribed 28S rRNA levels. c-MYC mRNA provides the control for the successful ablation of Pol II transcription. Apo-E mRNA shows elevated transcription in *Dnmt1*$^{-/-}$ cells, indicating that the DNA methylation-dependent transcriptional repression of this locus is lost in these cells. However, the Apo-E transcript levels are not reduced following α-amanitin treatment, which could be due to higher transcript stability.

Contrary to expectations, SINE transcript levels are not elevated in the absence of DNA methylation. This result remains unchanged following α-amanitin treatment. Therefore, we conclude that a genetic ablation of DNA methylation neither elevates Pol III loading onto SINEs, nor does it elevate cellular SINE transcript levels.

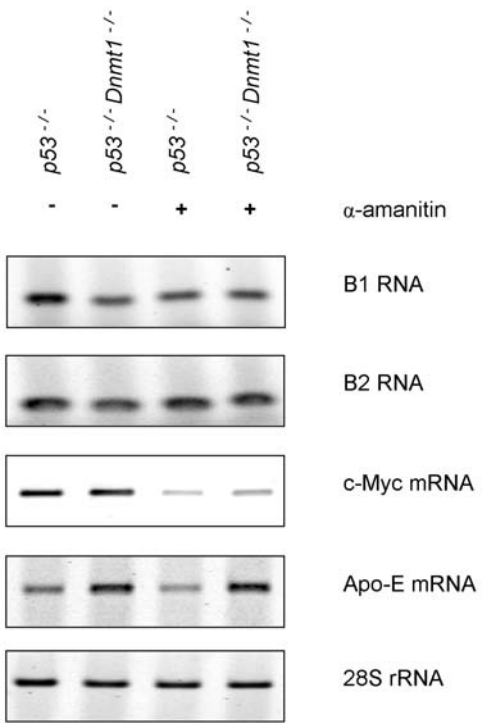


Figure 3.11: Loss of DNA methylation does not elevate SINE expression in *Dnmt1*$^{-/-}$ cells
Analysis by RT-PCR of expression levels of B1 and B2 transcripts in *p53*$^{-/-}$ and *p53*$^{-/-}$ *Dnmt1*$^{-/-}$ MEFs. Apo-E mRNA provides a positive control for loss of DNA methylation. Cells were treated with 50μg/ml α-amanitin for 24hrs. c-MYC mRNA provides a control for α-amanitin treatment. 28S rRNA provides the loading control.

The major drawback of using a genetic approach for reducing DNA methylation is that the cells may compensate for the alterations caused by the genetic deletion and use alternative pathways for silencing SINEs. It was therefore deemed necessary to use an acute pharmacological approach to reduce DNA methylation and study alteration in SINE transcription. Levels of Alu transcripts were shown to be highly elevated following 5-azacytidine treatment of HeLa cells for eight days (Liu et al., 1994). However, this long treatment may be unnecessary since clear demethylation of HeLa DNA could be seen following 72hr treatment of HeLa cells with 4μM 5-azacytidine (Figure 3.12).

Methylcollector assay (Active Motif) was performed on sonicated genomic DNA from HeLa cells that were mock-treated or treated with 4μM 5-azacytidine for 72hrs. The assay uses his-tagged recombinant MBD2b/MBD3L1 protein, which shows remarkable affinity to meCpG (Jiang et al., 2004) and specifically binds methylated DNA. PCRs were then performed on the collected (methylated) and wash-through (unmethylated) fractions. 7SL and APC loci, which provide controls for unmethylated DNA, are eluted in the wash-through fraction in both untreated and 5-azacytidine treated cells. Apo-E locus, which provides the positive control for DNA methylation, is present in the collected fraction in the untreated cell, whereas, upon 5-azacytidine treatment it is eluted in the wash-through, suggesting that it was successfully demethylated.

It can be clearly seen that Alus, which elute in the methylated fraction in untreated cells, are demethylated by 5-azacytidine treatment and elute in the unmethylated fraction. Thus, it is sufficient to treat HeLa cells for 72 hours with 4μM of 5-azacytidine in order to demethylate the majority of Alu repeats. This treatment was used in the following experiments to study the SINE transcript levels and promoter Pol III loading.

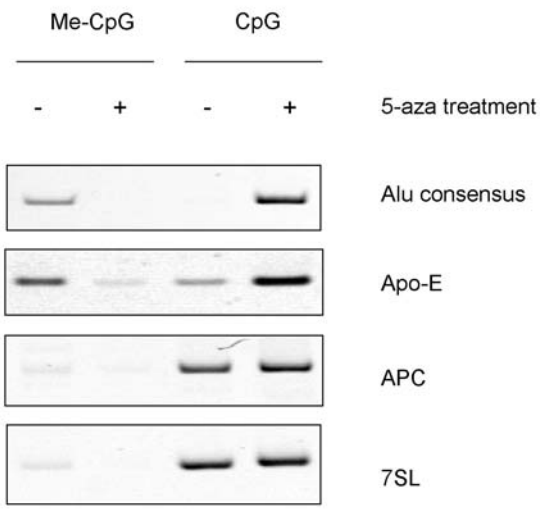


Figure 3.12: 5-azacytidine treatment reduces DNA methylation on Alus
Methylcollector assay on genomic DNA from HeLa cells treated with 4μM 5-azacytidine for 72hrs. me-CpG lanes contain methylated DNA retained by the methylcollector and the CpG lanes contain unmethylated DNA collected from the wash-through. Apo-E PCR provides the positive control for methylated DNA, whereas APC and 7SL PCRs provide the negative controls.

Total RNA from mock-treated or 5-azacytidine-treated MEF and HeLa cells was subjected to RT-PCR analysis in order to detect the levels of cellular SINE transcripts (Figure 3.13). ARPP P0 mRNA was used to normalise the PCR reactions (Laborda, 1991). As expected, Apo-E mRNA levels are increased following 5-azacytidine treatment and DNA demethylation, whereas no change is observed in the levels of pre-tRNA^{Leu} in HeLa cells. However, contrary to previous reports, neither mouse nor human SINE transcript levels show any increase after demethylation.

ChIP assay with antibodies against MBPs and Pol III was then performed on chromatin from mock-treated and 5-azacytidine-treated mouse ES cells (Figure 3.14). It was observed that 16hrs of 5-azacytidine treatment was enough to demethylate SINEs in this cell type. The ChIPs from the untreated and treated cells were normalised to serial input dilutions. 7SL sequence provides the negative control for MBP occupancy and positive control for Pol III occupancy. 5-azacytidine treatment leads to a clear eviction of MBD2 and MeCP2 from B1 and B2 consensus sequences, indicating that this DNA has been successfully demethylated. However, as observed in the *Dnmt1*$^{-/-}$ cells, 5-azacytidine treatment does not lead to increased polymerase loading onto murine SINEs. Alu(MLL) is a transgenic Alu sequence, which is introduced into these mouse ES cells (Elliott et al., 2005). This ectopic sequence also shows Pol III loading and loss of methylation is unable to elevate this polymerase occupancy either.

Thus, it can be concluded that genetically or pharmacologically stimulated loss of DNA methylation does not alleviate SINE transcriptional repression.

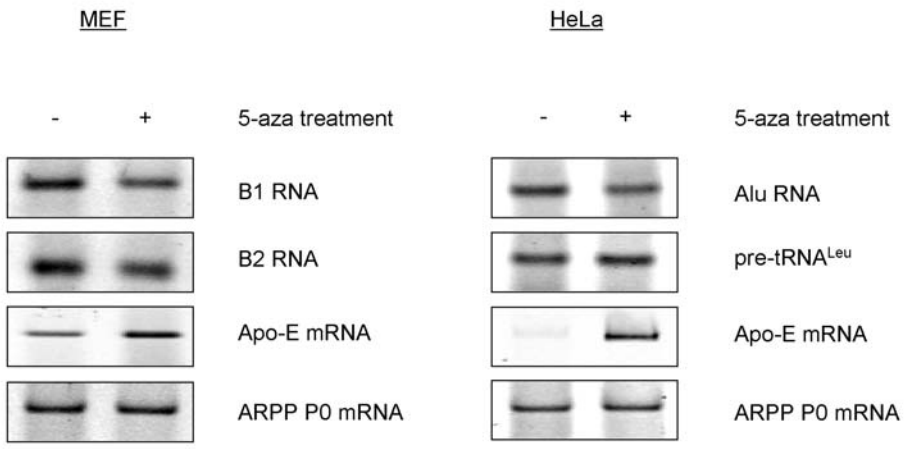


Figure 3.13: 5-azacytidine treatment does not elevate SINE expression
Analysis by RT-PCR of expression levels of B1 and B2 transcripts in MEFs and Alu transcripts in HeLa cells following treatment with 4μM 5-azacytidine for 16hrs and 72hrs respectively. Apo-E mRNA provides the positive control and ARPP P0 mRNA provides the loading control.

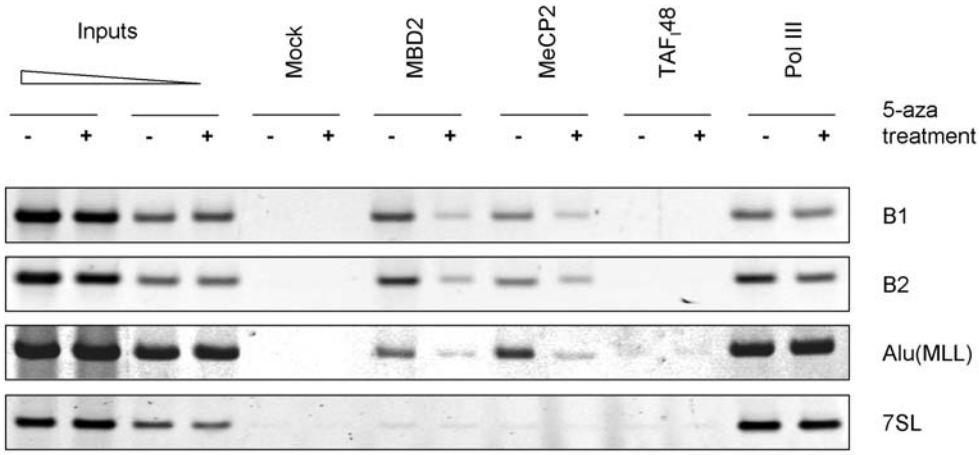


Figure 3.14: 5-azacytidine treatment does not alter Pol III occupancy on SINEs
ChIP assay in mouse ES cells with antibody against Rpc155 at B1 and B2 genomic consensus sequences following treatment with 4μM or without 5-azacytidine for 16hrs. ChIP with beads only and TAF$_I$48 antibody provide the negative controls. ChIPs with antibodies against MBD2 and MeCP2 indicate the DNA methylation status of the loci. 7SL locus provides positive control for Pol III binding and negative control for MBP occupancy. Alu(MLL) is an ectopic Alu sequence introduced into the mouse genome (Refer to chapter 4). The ChIPs were normalised to serial dilutions of the inputs.

3.3 Discussion

Recently published genome wide studies have suggested the presence of Pol III on SINES. Using ChIP-sequencing for Pol III in K562 cells, Kevin Struhl and co-workers reported ~1000 Pol III enriched loci that had not been previously described. 90% of these sites are found to be located near SINEs and comprise ~61% of total Pol III binding sites in the genome. These sites showed considerably reduced polymerase loading compared to tRNA genes, however enrichment levels were clearly above background (Moqtaderi et al., 2010). This could be unique to K562 cells, which have reduced DNA methylation on SINEs (Li et al., 2000). However, ChIP-sequencing in IMR90 cells (Canella et al., 2010) and HeLa cells (Oler and Cairns) has also revealed the clear presence of Pol III on Alus sequences.

In agreement with these genome-wide approaches and contrary to expectations, this chapter provides clear evidence for the presence of Pol III transcription apparatus on human and mouse SINE sequences. SINEs show a distinct reduction in Pol III loading compared to the actively transcribed 7SL locus. No such difference is observed in TFIIIB or TFIIIC occupancy. As yet the reason for this defect in Pol III recruitment on SINEs is unclear.

In order for transcription apparatus to occupy the sheer number of SINES, a cell would require a large amount of Pol III transcription factors. Estimates from quantitative western blots suggest that TBP, BRF1 and TFIIIC110 are equimolar in MEF ($2x10^6$ mol./cell) and HeLa cells ($5x10^5$ mol./cell). Since TBP is required for all nuclear transcription (Cormack and Struhl, 1992) it could be reasoned that a cell would require a higher amount of this transcription factor. The relatively high amounts Pol III specific transcription factors means that the cell possesses sufficient quantities of these molecules to quench the thirst of SINEs for these transcription factors. MEFs show higher amount of all three transcription factors quantified compared to HeLa cells, which may be to maintain equal concentrations over the larger nuclear volume observed in MEFs (personal observation).

As very low levels of SINE transcripts are present in a cell (Maraia et al., 1993; Paulson and Schmid, 1986), it seems that the preassembled transcription complexes on SINEs are transcriptionally inactive or have very low transcriptional output. The innate unstable nature of SINE transcripts (Li and Schmid, 2004) and degradation by DICER1 (Kaneko et al., 2011) may also contribute to the low transcript levels. The quick transcriptional response of SINEs to cell stresses (Li et al., 1999; Liu et al., 1995) may require the presence of these poised transcription complexes.

For more than a decade, DNA methylation has been thought to be primarily responsible for inhibiting SINE transcription. One mode of transcriptional repression by DNA methylation is through the recruitment of MBPs, which can repress transcription through their transcription repression domains (Hendrich and Bird, 1998). This study demonstrates the enrichment of MBPs on both human and mouse SINEs. B2 elements also seem to be occupied by MBPs, for which there is no previous precedent. MBD1, MBD2 and MeCP2 have been previously shown to have differing binding preferences. For example, high affinity binding of MeCP2 to DNA requires a stretch of four or more A/Ts adjacent to the meCpG, whereas MBD1 prefers to bind meCpGs within TCGCA and TGCGCA sequences (Clouaire et al., 2010). MeCP2 shows higher occupancy on SINEs compared to MBD1 and MBD2, however the reason for this is unclear since the SINEs studied are not distinctly enriched for CpGs flanked by A/Ts. However, the presence of various MBPs is not enough to inhibit polymerase recruitment to these sequences. Sequential ChIP assays revealed that Pol III is able to co-occupy SINE sequences alongside MBPs.

DNA methylation can also repress transcription by directly inhibiting transcription factor binding to their recognition sequences (Watt and Molloy, 1988). A considerable number of DNA-binding proteins require contact with cytosine in the major groove of the double helix. Most mammalian transcription factors have recognition sites which are within GC-rich sequences (Bird and Wolffe, 1999). Indeed, the consensus B-block binding sequence in Pol III promoters has a central CG dinucleotide (Galli et al., 1981). The methylation of an A-block CpG was shown to be deleterious for Alu transcription *in vitro,* although the methylation of surrounding CpGs was found to be equally disruptive (Liu and Schmid, 1993).

This study clearly demonstrates that TFIIIC and Pol III are capable of recruitment to methylated DNA in cells. It is possible that TFIIIC averts the binding inhibition by meCpG due to the numerous points of contacts it makes with the DNA (Bartholomew et al., 1990; Braun et al., 1992; Gabrielsen et al., 1989; Kovelman and Roeder, 1992; Yoshinaga et al., 1989). However, as B-block binding determines the promoter affinity of this multi-subunit transcription factor, it is possible that B-block CpG methylation would be more deleterious to TFIIIC and therefore Pol III recruitment.

Bisulfite sequencing revealed the presence of DNA methylation on Pol III occupied Alu sequences. Analysis of 14 Alu family consensus sequences shows that there is no difference in the DNA methylation levels between the input and Pol III ChIP samples. Pol III ChIP sample also contains Alu sequences methylation within the A- and B-boxes. These data suggest that Pol III loading is not affected by DNA methylation, even when A- and B-box regulatory elements are methylated.

Moreover, the young AluY elements show higher DNA methylation compared to their older AluJ and AluS counterparts, which may have lower CpG content due to a loss of meCpGs through deamination (Duncan and Miller, 1980). Despite having higher DNA methylation, AluY families show higher Pol III loading compared to AluJ or AluS elements. This may be due to the more intact A- and B-boxes found within these younger transposons which show the least amount of sequence divergence from the master Alu gene (Shen et al., 1991). Therefore, DNA methylation is not found to negatively correlate with RNA Pol III loading at SINEs.

This study finds no evidence to support the idea that DNA methylation inhibits SINE transcription. In *Dnmt1*$^{-/-}$ cells, loss of DNA methylation does not elevate Pol III loading onto SINEs, neither does it increase SINE expression. This result is also confirmed with 5-azacytidine treatment of human and mouse cells, which is surprising and contradictory to the existing literature. 5-azacytidine treatment of HeLa cells was previously shown to cause 5- to 8-fold elevation in Alu transcripts (Liu et al., 1994). However, the cells were treated for 8 days with 5-azacytidine, and the transcriptional effect observed may have been due to downstream effects of this long treatment. The authors confirmed that Alus

were demethylated, but this does not prove that demethylation was directly responsible for upregulation of Alu expression. 5-azacytidine has been shown to cause DNA damage (Juttermann et al., 1994; Murakami et al., 1995) and SINE transcription is upregulated in response to DNA damaging agents (Rudin and Thompson, 2001). The present study found that 16 – 72hr treatment of mouse and human cells was sufficient to cause SINE demethylation, but does not lead to elevation of SINE expression. Moreover, this expression data is strongly supported by the promoter occupancy analysis.

There have been multiple studies showing DNA methylation dependent inhibition of SINE transcription *in vitro* (Kochanek et al., 1993; Liu and Schmid, 1993) and following transient transfections (Koch and Stratling, 2004; Yu et al., 2001). DNA methylation inhibited Alu transcription at low template concentration, whereas no such inhibition was observed at higher template concentrations which would more accurately represent the high SINE copy number found *in vivo* (Liu and Schmid, 1993). Co-transfection of methylated Alu sequences with the MeCP2 TRD relieved the methylation-mediated repression, rather than compounding it (Yu et al., 2001). In agreement with this, there is evidence in our lab that a genetic loss of MeCP2 does not lead to elevation in cellular SINE transcript levels (Unpublished data, Jana Vávrová). Cell stress response is able to elevate SINE transcription without affecting DNA methylation status (Chu et al., 1998; Liu et al., 1995), suggesting that transcription through methylated SINE DNA may be possible. Thus, it is possible that SINE transcriptional inhibition is achieved through chromatin modifications. This possibility has been investigated in Chapter 4.

This chapter demonstrates that human and mouse SINEs are occupied by Pol III transcription apparatus and this occupancy is unaffected by the presence of DNA methylation and MBPs. Moreover, genetically or pharmacologically stimulated loss of DNA methylation does not elevate SINE transcriptional repression.

Chapter 4 Chromatin and Pol III mediated transcription of SINEs

4.1 Introduction

As previously described, human and mouse SINE elements are occupied by Pol III transcription apparatus. However, Pol III loading on SINEs is considerably diminished compared to actively transcribed 7SL genes. Loss of DNA methylation neither leads to increased cellular SINE transcript levels, nor does it cause increased Pol III loading at SINEs. The presented data, along with the observation that cell stresses can alleviate SINE transcription from methylated templates (Chu et al., 1998; Liu et al., 1995), suggest that SINE transcription is being inhibited by mechanisms other than DNA methylation.

Reconstitution experiments have shown that an Alu element can direct the rotational positioning of a nucleosome. Moreover, it was also observed that the methylation of this Alu template enhanced its capacity to position nucleosomes (Englander et al., 1993). This positioned nucleosome completely abrogated the Pol III-dependent *in vitro* transcription of the Alu template. The positional information was found to be an intrinsic property of the Alu family and could influence the conformation of the chromatin surrounding the element (Englander and Howard, 1995). These studies suggest that repressive chromatin could be an intrinsic feature of SINEs and may be responsible for their transcriptional repression.

There are clear structural and functional differences between Alu chromatin obtained from unstressed cells and that from cells stressed with viral infections. Stressed Alu chromatin is more readily transcribed *in vitro* and is more accessible to enzymatic digestion (Li et al., 2000; Russanova et al., 1995). There is genome-wide increase in chromatin accessibility of Alu and other repetitive sequences following heat shock (Kim et al., 2001), whereas the levels of DNA methylation are unaltered (Chu et al., 1998; Liu et al., 1995). These observations suggest that transcriptional repression at SINEs is primarily maintained by repressive chromatin.

The relationship between DNA methylation and chromatin modifications is well established. Chromatin modifications are well-known downstream effectors of DNA methylation (see section 1.4.3). However, there is increasing evidence to suggest that chromatin modifications are required for the maintenance of DNA methylation at many locations. For example, H3K9 methylation by G9a, GLP and SUV39H have been shown to be necessary for maintaining DNA methylation at retrotransposons and satellite repeats (Dong et al., 2008; Lehnertz et al., 2003). Loss of DNA methylation does not lead to a decrease in H3K9me3 (Gilbert et al., 2007; Lehnertz et al., 2003), whereas loss of K9MTs leads to a loss of DNA methylation (Dong et al., 2008; Lehnertz et al., 2003).

Alu sequences have been shown to be enriched for H3K9 methylation (Kondo and Issa, 2003), which is a mark for repressive chromatin and transcriptional repression. H3K9me3 modification is responsible for the recruitment of the heterochromatin protein 1 (HP1) family of highly conserved non-histone proteins, which were identified as a component of heterochromatin in *Drosophila*. However, more recently HP1 has also been shown to cause silencing of euchromatic genes (Fanti and Pimpinelli, 2008).

HP1 contains an N-terminal chromodomain which recognises H3K9 methylation. However, this recognition is not sufficient for the recruitment of HP1 to chromatin. SUV39H1 was the first K9MT identified to interact with HP1 and this protein-protein interaction was found to be crucial for HP1 targeting. Once recruited to a target HP1 can in turn recruit more SUV39H1 and propagate the H3K9me3 modification (Stewart et al., 2005). HP1 is now known to recognise the SET domain found in all K9MTs (Fritsch et al., 2010). Thus, HP1 and SUV39H1 can together mediate the formation of non-permissive chromatin. HP1α and HP1β isoforms are associated with gene repression, whereas HP1γ is associated with transcriptional activation (Kwon and Workman, 2010).

Recently, Barski *et al* showed that levels of H3K9me3 displayed a clear negative correlation with Pol III loading at tRNA loci (Barski et al., 2010). This chapter investigates the possibility that H3K9me3 may be responsible for the repression of Pol III-mediated SINE transcription.

4.2 Results

4.2.1 SINEs are enriched for H3K9me3, SUV39H1 and HP1.

H3K9me3 at SINEs was shown to be sensitive to 5-azadeoxycytidine treatment, indicating that the presence of this modification depends on DNA methylation (Kondo and Issa, 2003). However, there is contradictory evidence that the genetic loss of DNMTs does not lead to a reduction of H3K9me3 from repeat sequences (Lehnertz et al., 2003). H3K9me3 along with H3K27me3, which causes transcriptional repression through recruitment of the polycomb group of proteins, are enriched at silent tRNA loci and correlate with diminished Pol III loading (Barski et al., 2010). To investigate the role of these repressive histone modifications in transcriptional silencing of SINEs, comparative ChIP assays were performed at SINEs in *Dnmt1* wild-type and null MEFs (Figure 4.1).

ChIPs with beads only (mock) and TAF_I48 antibody provide the negative controls. ChIP with histone H3 provides the positive control and also controls for the nucleosome occupancy. The ChIP DNA was normalised to serial input dilutions. 7SL gene, which was previously shown to be highly enriched for Pol III, shows a complete absence of both H3K9me3 and H3K27me3. Apo-E locus only shows the presence of H3K27me3. Trimethylation at H3K9 and H3K27 residues is observed on both mouse SINE families in control cells. Alu SINEs are also enriched for H3K9me3 in HeLa cells (Figure 4.2).

Histone H3 occupancy is unaltered following the loss of DNA methylation in *Dnmt1*$^{-/-}$ cells. Thus, any change observed in the levels of histone modifications is not a function of altered nucleosome occupancy. There is a clear reduction in H3K27me3 at SINEs and Apo-E gene in *Dnmt1*$^{-/-}$ cells, whereas, the H3K9me3 at SINEs is unperturbed. This is contradictory to what was previously observed with 5-azadeoxycytidine treatment (Kondo and Issa, 2003) and suggests that H3K9me3 at SINEs is not dependent on DNA methylation. SINEs were found to have very low enrichment for active marks such as H3 acetylation, H4 acetylation and H3K4me3, indicating a repressive chromatin environment (data not shown).

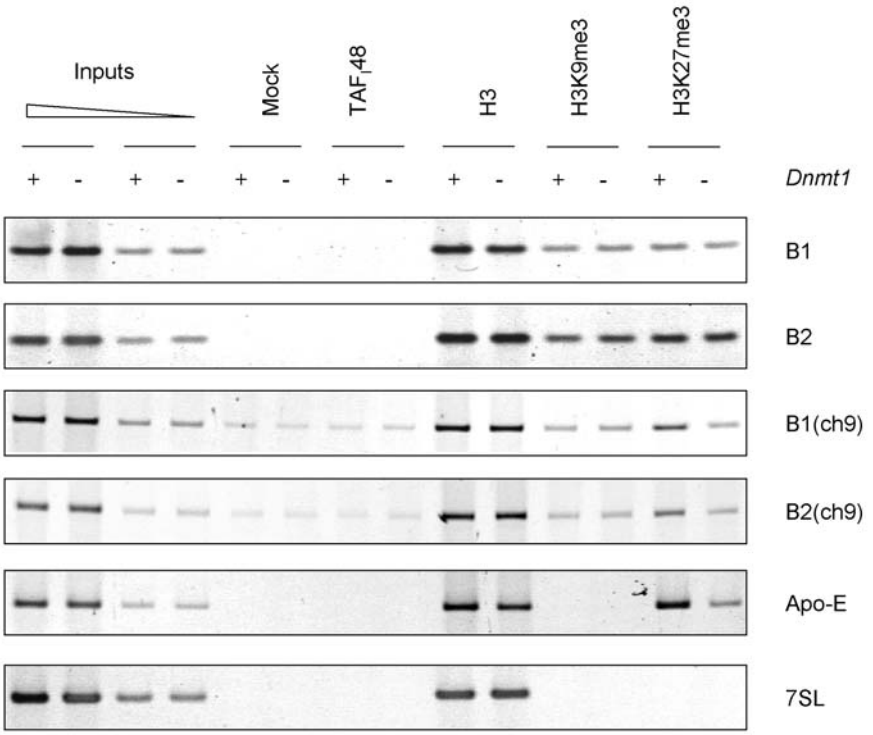


Figure 4.1: Histone H3 Lysine 9 trimethylation on SINEs is unaltered in *Dnmt1* null cells
ChIP assay in *p53*$^{-/-}$ and *p53*$^{-/-}$ *Dnmt1*$^{-/-}$ MEFs with antibodies against histone H3, histone H3 Lys9 trimethylation (H3K9me3) and histone H3 Lys27 trimethylation (H3K27me3) at B1 and B2 genomic consensus sequences. B1(ch9) and B2(ch9) are specific SINEs on chromosome 9. ChIPs with beads only and TAF$_I$48 antibodies provide the negative controls. ChIPs were normalised to serial dilutions of the inputs.

Therefore, SINEs are enriched for H3K9me3 and the presence of this histone modification is independent of DNA methylation. Histone lysine methyltransferases responsible for the methylation of H3K9 (K9MTs) include SUV39H1, G9a and SETDB1 (Kouzarides, 2007). SUV39H is thought to be responsible for H3K9me3 at pericentric repeat sequences (Lehnertz et al., 2003). Moreover, SUV39H knockout mouse ES cells show elevated levels of B1 and B2 SINE transcripts (Martens et al., 2005). These observations suggest that SUV39H may have a role in SINE transcriptional repression.

In order to investigate the presence of SUV39H1 and associated HP1 at SINEs, chromatin from HeLa cells was subjected to ChIP assays with an antibody against H3K9me3, two antibodies against SUV39H1 and an antibody against all HP1 isoforms (Figure 4.2). Histone H3 antibody provides a positive control, whereas TAF_I48 antibody provides a negative control. Promoter of the human p21 gene, which has previously been shown to be occupied by SUV39H1, provides a positive control PCR (Cherrier et al., 2009). 7SL locus, that was previously seen to be not enriched by H3K9me3, provides a negative control PCR. Serial input dilutions show the quantitative nature of the PCRs.

Three of the four Alu loci tested show the clear presence of both SUV39H1 and HP1. Alu(ch6) is enriched for H3K9me3, however it is not enriched for SUV39H1. The lack of HP1 at this site indicates that SUV39H1 occupancy may be necessary for HP1 recruitment at SINEs. It has been previously shown that the presence of H3K9me3 is insufficient for the recruitment of HP1 to its target loci. The presence of SUV39H1, but not G9a, alongside H3K9me3 was found to be required for successful HP1 recruitment (Stewart et al., 2005). Thus, the H3K9me3 at Alu(ch6) may be due to the activity of a alternative K9MT.

Thus, it can be concluded that SINEs are enriched for H3K9me3, SUV39H1 and HP1.

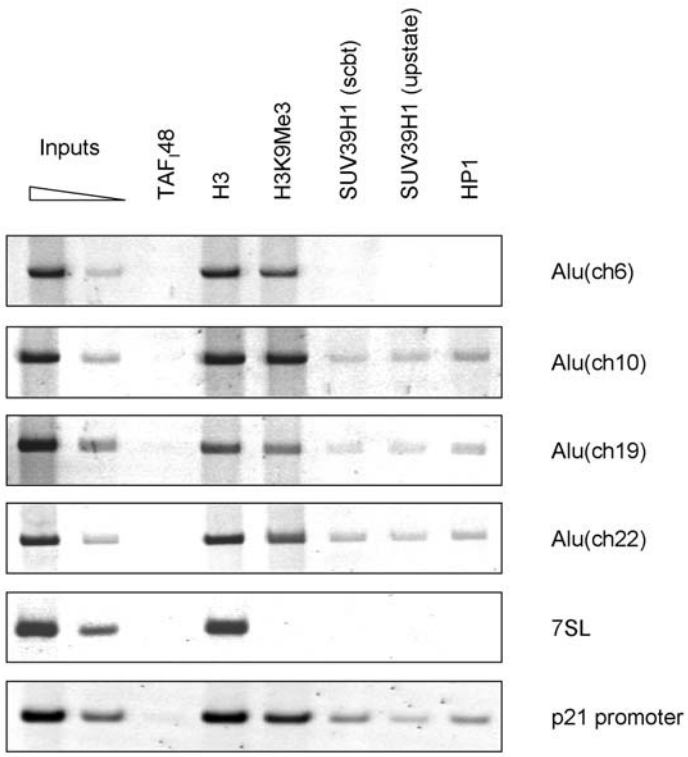


Figure 4.2: Histone H3 Lysine 9 methyltransferase SUV39H1 and HP1 occupy Alu sequences ChIP assay in HeLa cells with antibodies against H3K9me3, two different antibodies against SUV39H1 and an antibody against HP1 at Alu loci on chromosome 6, 10, 19 and 22. ChIP with TAF$_I$48 antibody provides the negative control. 7SL and p21 promoter respectively provide the negative and positive controls for the ChIP. Serial dilutions of the inputs show the quantitative nature of the PCR.

4.2.2 SUV39H1 inhibits SINE transcription irrespective of DNA methylation status

Chaetocin, a fungal metabolite from *chaetomium minutum*, was initially described in the 1970s. It was identified as a specific inhibitor of SUV39H methyltransferase both *in vitro* and *in vivo* (Greiner et al., 2005). More recently, chaetocin has been used to demonstrate SUV39H1 mediated repression of the $p21^{WAF1}$ gene in microglial cells (Lakshmikuttyamma et al., 2010) and repression of $p15^{INK4B}$ and E-cadherin genes in myeloid leukaemia cell lines (Cherrier et al., 2009). The anti-myeloma activity of chaetocin was initially reported to be mediated via imposition of oxidative stress, however the apoptosis observed has now been shown to be also due to derepression of the p21 gene promoter (Cherrier et al., 2009).

In order to investigate the role of H3K9me3 and SUV39H1 in SINE transcriptional repression, control and *Dnmt1*$^{-/-}$ cells were treated with increasing concentrations of chaetocin for 24hrs. Total RNA obtained from mock and chaetocin treated cells was subjected to RT-PCR analysis using primers against B1 and B2 consensus transcripts (Figure 4.3). Expression levels were normalised to ARPP P0 mRNA levels. Since the Apo-E locus was not found to be enriched for H3K9me3 (Figure 4.1), Apo-E mRNA levels do not respond to chaetocin treatment as expected. This demonstrates the specificity of the response.

A dose-dependent increase in both B1 and B2 expression is observed following chaetocin treatment of *Dnmt1* wild-type and null cells. An increase in levels of Alu transcripts was also observed in response to chaetocin treatment of HeLa cells (data not shown). These results suggest that SUV39H1 is responsible for SINE transcriptional inhibition. Moreover, chaetocin is able to derepress SINE transcription regardless of SINE DNA methylation status.

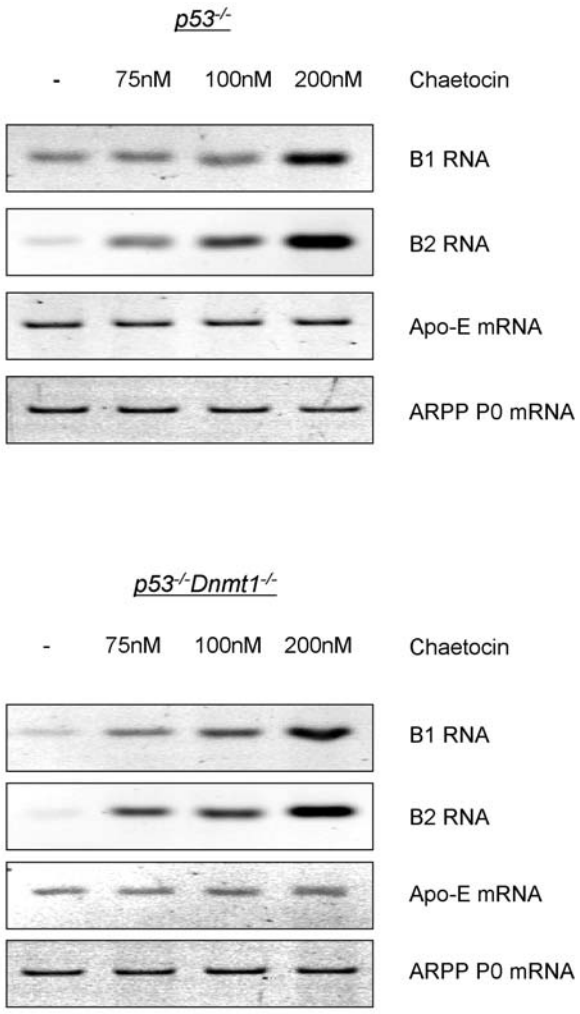


Figure 4.3: SUV39H **inhibitor, Chaetocin, de-represses SINE transcription irrespective of the DNA methylation state**
Analysis by RT-PCR of expression levels of B1 and B2 transcripts in *p53*$^{-/-}$ and *p53*$^{-/-}$ *Dnmt1*$^{-/-}$ MEFs following treatment with the shown concentrations of chaetocin for 24hrs. ARPP P0 mRNA provides the loading control.

The increase observed in SINE transcript levels in response to chaetocin treatment could be a function of increased transcript stability or reduced degradation and not elevated transcription. Since, SINEs show reduced Pol III occupancy compared to actively transcribed Pol III genes (Chapter 3), the release from transcriptional repression may lead to elevation in Pol III loading. Thus, the Pol III loading at SINEs could be considered a surrogate for the transcriptional activity.

Therefore, in order to establish if the observed effect was transcriptional, the levels of Pol III occupancy at SINEs were analysed following chaetocin treatment. HeLa cells treated with 100nM chaetocin were subjected to ChIP assays with antibodies against H3K9me3, Pol III and TFIIIC (Figure 4.4). TAF_I48 antibody provides the negative control, whereas, histone H3 antibody provides the positive control. 7SL gene provides the positive control PCR for Pol III transcription factor occupancy, whereas Apo-E gene provides the negative control PCR. ChIPs were normalised to serial input dilutions.

As expected, chaetocin treatment leads to a decrease in H3K9me3 enrichment at most Alu sequences analysed. No change in H3 enrichment is observed, suggesting that the decrease in H3K9mes is not due to an overall loss of nucleosomes but a specific reduction in the post-translational histone modification. Alu(ch6) does not show a decrease in H3K9me3 levels, but this locus is also not enriched for SUV39H1 (Figure 4.2). Since, at concentrations used, chaetocin is a specific inhibitor of SUV39H (Greiner et al., 2005), this suggests that H3K9 at Alu(ch6) is methylated by an alternative K9MT. As expected, the Pol III and TFIIIC occupancy on Alu(ch6) is not affected by chaetocin treatment.

All Alu loci showing a decrease in H3K9me3 in response to chaetocin treatment show a concordant elevation in Pol III loading. This suggests that the elevated SINE expression observed by RT-PCR is due to an increase in Pol III-mediated transcription at SINEs. No increase in TFIIIC occupancy is observed following chaetocin treatment, suggesting that H3K9me3 specifically represses polymerase recruitment at SINEs. Chaetocin causes no increase in Pol III loading at the 7SL locus, which indicates that the elevated Pol III-loading on SINEs is not part of a global increase in cellular Pol III transcription.

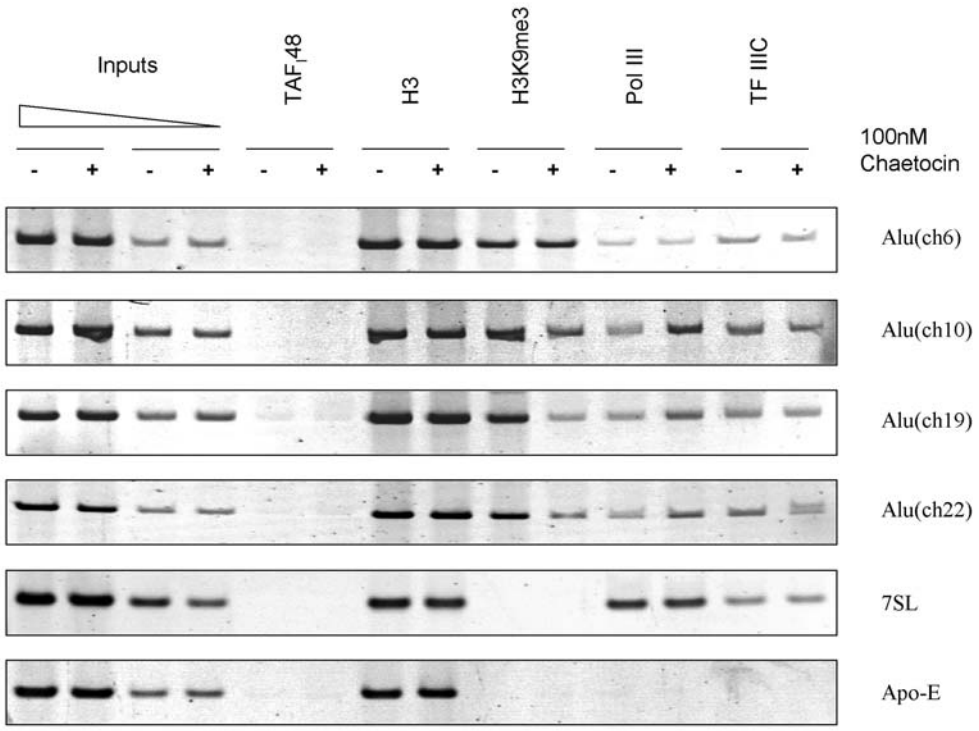


Figure 4.4: Chaetocin treatment leads to increased polymerase loading onto selected SINEs without altering TFIIIC occupancy

ChIP assay in HeLa cells with antibodies against H3K9me3, RPC155 and TFIIIC220 at Alu loci on chromosomes 6, 10, 19 and 22 following treatment with 100nM chaetocin for 24hrs. ChIPs with TAF$_I$48 and H3 antibodies provide the negative and positive controls respectively. 7SL and Apo-E loci respectively provide the positive and negative controls for polymerase occupancy. The samples were normalised to serial dilutions of the inputs.

ChIP assays in *Dnmt1* wild-type and null MEFs with antibodies against H3K9me3, MBD2 and Pol III were performed in order to investigate the interplay between DNA methylation and H3K9 methylation at SINEs (Figure 4.5). Just like the human SINEs, both B1 and B2 SINE sequences show reduced H3K9me3 and elevated Pol III occupancy in response to chaetocin treatment. However, the Pol III loading on a $tRNA^{Leu}$ gene, negative for H3K9me3, is unaltered.

Chaetocin is able to elevate Pol III loading onto SINEs both in the presence and absence of DNA methylation. H3K9me3 can recruit DNMTs and induce methylation of the target sequences (Cheng and Blumenthal, 2010). The DNA methylation at SINEs could be a result of such targeting. However, this is not found to be the case since MBD2 occupancy in *Dnmt1* wild-type cells is unaltered following the loss of H3K9me3. Moreover, this MBD2 occupancy and the implied DNA methylation at B1 and B2 SINEs is unable to perturb the chaetocin-mediated increase in Pol III transcription, indicating that Pol III may be able to transcribe through methylated DNA at SINEs.

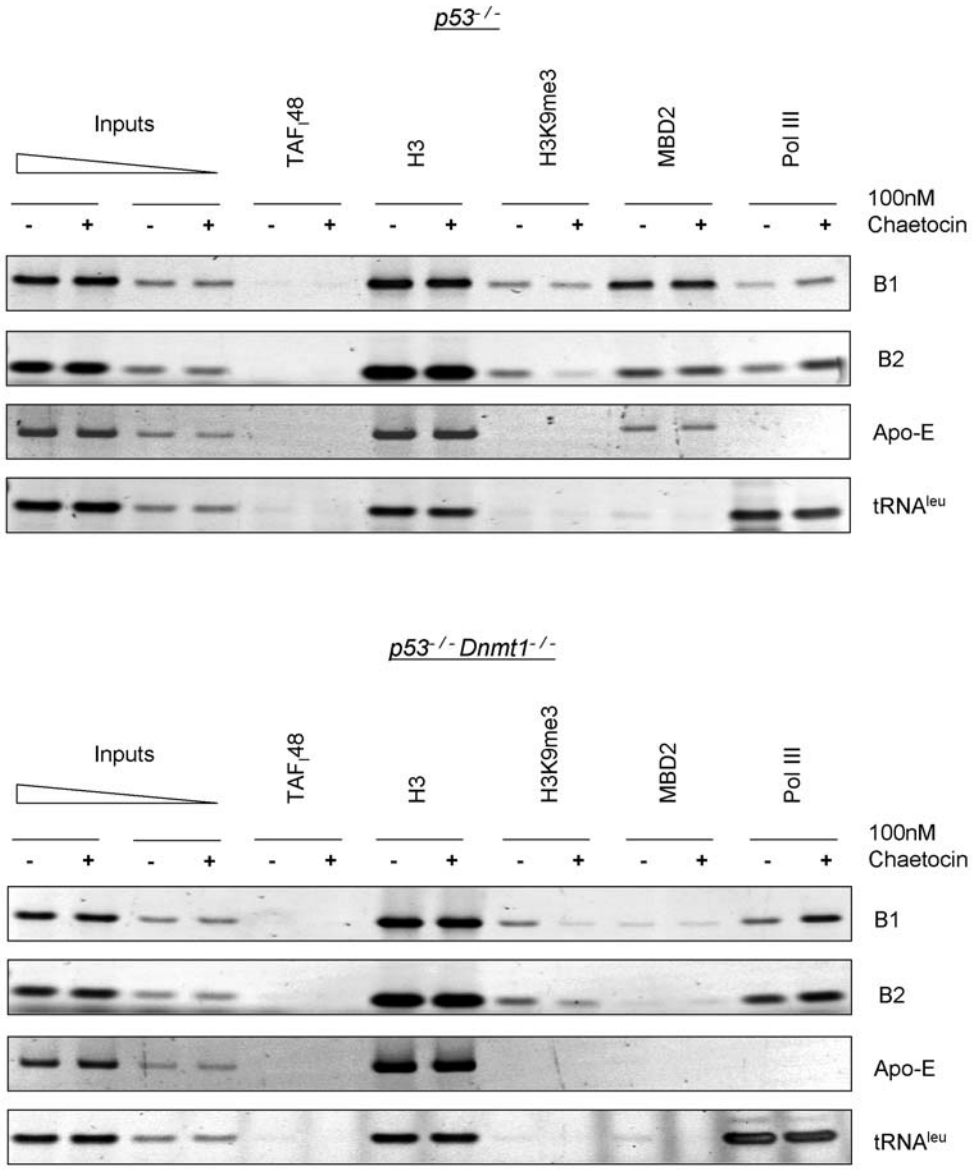


Figure 4.5: Chaetocin treatment leads to increased polymerase loading onto SINEs without altering and irrespective of the DNA methylation state

ChIP assay in *p53*$^{-/-}$ and *p53*$^{-/-}$ *Dnmt1*$^{-/-}$ MEFs with antibodies against H3K9me3, MBD2 and RPC155 at B1 and B2 genomic consensus sequences following treatment with 100nM chaetocin for 24hrs. ChIPs with TAF$_I$48 and H3 antibodies provide the negative and positive controls respectively. Apo-E and tRNA^{Leu} genes respectively provide the negative and positive controls for polymerase occupancy and visa versa for MBD2 occupancy. ChIPs were normalised to serial dilutions of the inputs.

4.2.3 SINE transcription and histone deacetylation

The data so far indicate that in human and mouse cells the transcriptional repression of SINEs is due to the trimethylation of H3K9 by SUV39H1 and this repression is independent of DNA methylation. H3K9 methylation is thought to impose transcriptional silencing through the recruitment of HP1. However, the presence of H3K9 methylation has also been shown to cause deacetylation of H3 and H4 in an HP1-independent manner (Stewart et al., 2005). There is a functional interaction of SUV39H1 with HDAC1, HDAC2 and HDAC3, which causes this HP1-independent silencing (Vaute et al., 2002). Moreover, DNA methylation also recruits HDACs via MBPs (Zhang et al., 1999) and therefore the involvement of HDACs in SINE transcriptional silencing cannot be excluded.

To investigate this, total RNA from *Dnmt1* wild-type and null cells was subjected to RT-PCR analysis following 6hr treatment with 100nM TSA (Figure 4.6). The expression levels were normalised to ARPP P0 mRNA levels. The levels of B1 and B2 consensus transcripts are elevated two-fold ($p<0.05$) following TSA treatment of wild-type cells, which suggests that HDACs may contribute to SINE transcriptional repression. In stark contrast, SINE expression is not increased following TSA treatment of *Dnmt1* null cells ($p>0.1$). This surprisingly suggests that there is a switch from HDAC-dependent SINE repression in wild-type cells to a HDAC-independent repression in *Dnmt1* null cells. The DNA methylation coupled HDAC-mediated repression indicates that DNA methylation may have a minor contribution in SINE transcriptional repression.

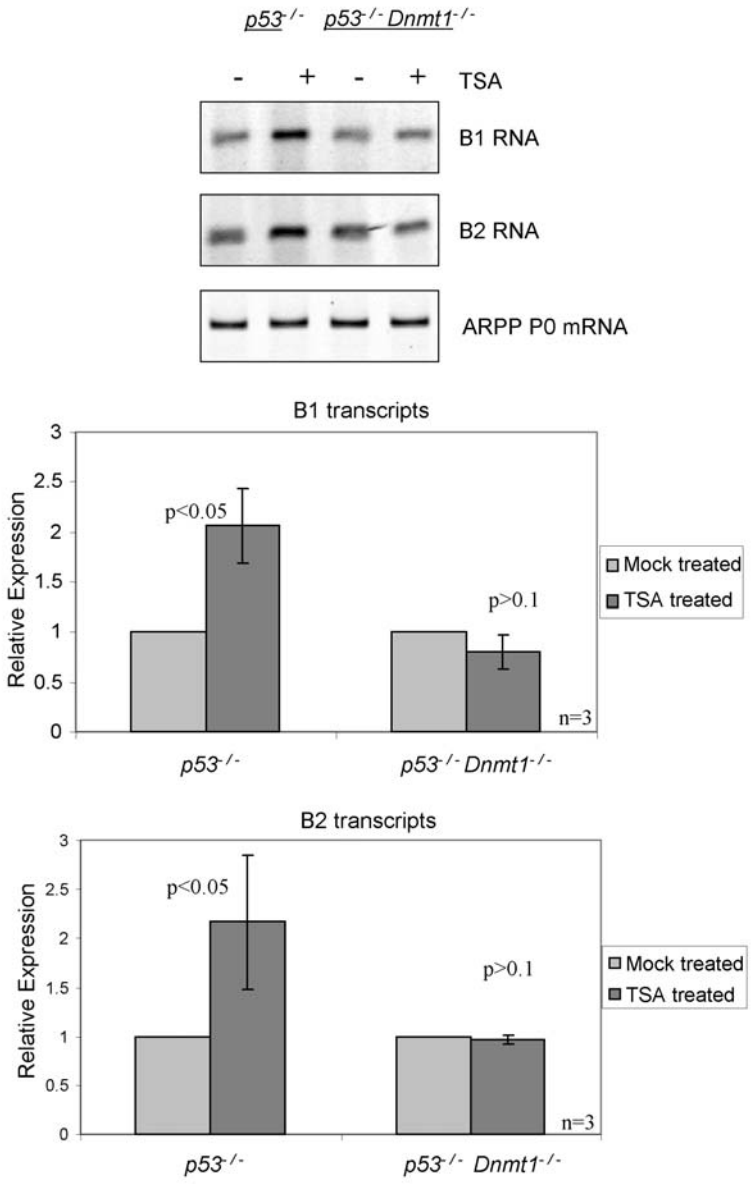


Figure 4.6: SINE expression is sensitive to TSA treatment in a DNA methylation-dependent manner.
Analysis by RT-PCR of expression levels of B1 and B2 transcripts in *p53*$^{-/-}$ and *p53*$^{-/-}$ *Dnmt1*$^{-/-}$ MEFs following treatment with 100nM TSA for 6hrs. ARPP P0 mRNA provides the loading control. The quantification was performed using ImageJ.

To further study the involvement of DNA methylation and HDACs in SINE silencing, chromatin from *Dnmt1* wild-type and null cells was subjected to ChIP assays with antibodies against HDAC1, HDAC2 and Pol III (Figure 4.7). TAF_I48 antibody provides the negative control. Sequential input dilutions demonstrate the quantitative nature of the PCRs and $tRNA^{Leu}$ gene provides the negative control for HDAC occupancy. B1, B2 and Apo-E genes are clearly enriched for HDAC1 and HDAC2 in wild-type cells. Since the Apo-E gene does not show H3K9me3 enrichment, HDACs at this locus are recruited by MBPs only. A reduction in HDAC1 and HDAC2 enrichment is observed following the loss of DNA methylation, whereas no resultant increase in Pol III loading is seen at SINEs.

The release of HDACs in *Dnmt1*$^{-/-}$ cells does not cause elevation in histone acetylation at SINEs (data not shown). Residual HDAC occupancy can still be seen in *Dnmt1*$^{-/-}$ cells, which may be recruited by SUV39H1. However, the SINEs which are still occupied by HDACs seem to be insensitive to TSA treatment and the reason for this is unclear. SINE transcription might be more sensitive to an acute loss of HDAC activity following TSA treatment of wild type cells. However, long term removal of HDAC-dependent repression in *Dnmt1* null cells may be compensated for by a switch to 'HP1 only' repression, which causes the SINE transcription to become HDAC-independent in *Dnmt1*$^{-/-}$ cells.

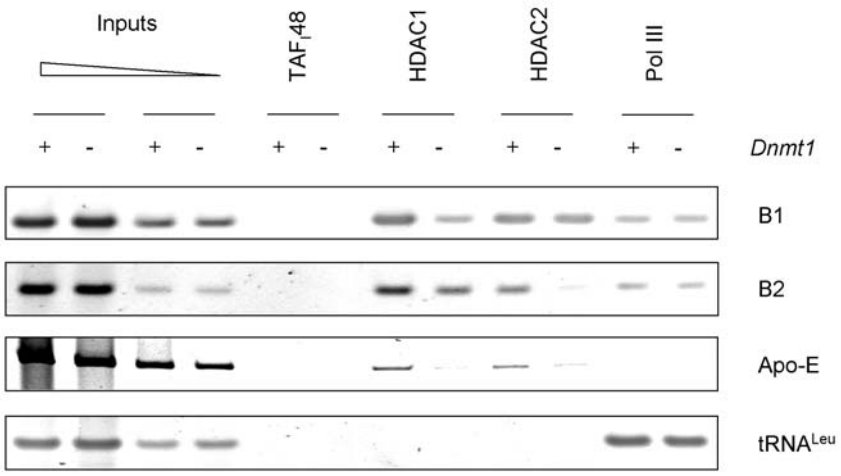


Figure 4.7: Loss of DNA methylation leads to reduced HDAC occupancy at SINEs
ChIP assay in *p53*$^{-/-}$ and *p53*$^{-/-}$ *Dnmt1*$^{-/-}$ MEFs with antibodies against HDAC1 and HDAC2 at B1 and B2 genomic consensus sequences. ChIPs with TAF$_I$48 and Pol III antibodies provide the negative and positive controls respectively. Apo-E and tRNA^{leu} genes respectively provide the positive and negative controls for HDAC occupancy and visa versa for Pol III occupancy. The ChIPs were normalised to serial dilutions of the inputs.

SINE expression shows clear differences in the sensitivity to TSA treatment in the presence and absence of DNA methylation. In order to look for any effect of DNA methylation on SINE transcriptional sensitivity to chaetocin treatment, RT-PCR data in Figure 4.3 was further analysed by ImageJ densitometry (Figure 4.8). The expression data were normalised to ARPP P0 mRNA levels. Both B1 and B2 expression shows an increased sensitivity to chaetocin treatment in *Dnmt1* null cells compared to the wild-type controls. These results collectively suggest that the DNA methylation may be contributing to transcriptional inhibition of SINEs, possibly through HDACs. However, this inhibition is superseded by HDAC-independent repression by H3K9me3-mediated repressive chromatin.

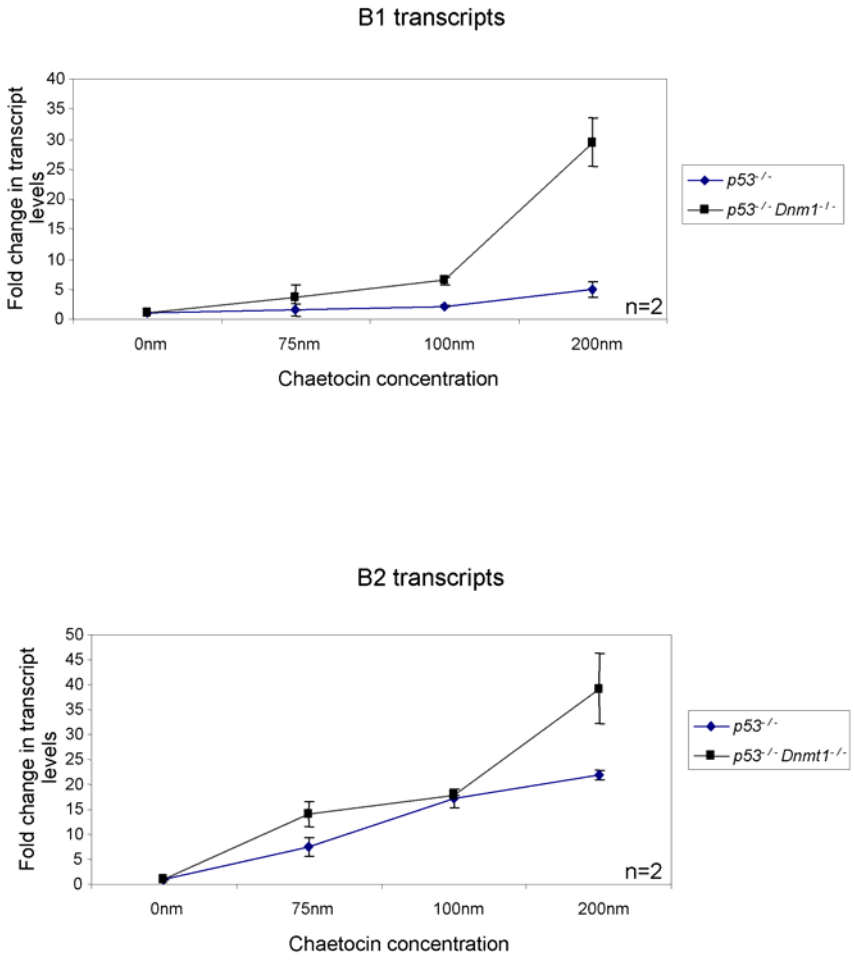


Figure 4.8: SINE expression shows increased sensitivity to chaetocin treatment in absence of DNA methylation
Quantification by ImageJ of expression levels of B1 and B2 transcripts in *p53*$^{-/-}$ and *p53*$^{-/-}$ *Dnmt1*$^{-/-}$ MEFs following treatment with shown concentrations of chaetocin for 24hrs. The expression was normalised to ARPP P0 mRNA levels.

4.3 Discussion

There is a considerable amount of evidence in the literature correlating DNA methylation with H3K9 methylation. DNA methylation can recruit K9MTs via MeCP2, whereas K9MTs can in turn lead to DNA methylation by recruiting DNMTs. These two represent conserved pathways that cooperate to achieve transcriptional repression (Cheng and Blumenthal, 2010).

K9MTs G9a and GLP have been shown to be required for DNA methylation at various loci, including retrotransposons and satellite repeats. $G9a^{-/-}$ cells show reduced DNA methylation, whereas no reduction in H3K9me3 or HP1 occupancy is observed. Despite the loss of DNA methylation, $G9a^{-/-}$ cells do not show elevated expression of these repeat sequences. Therefore, transcriptional repression at repeats in $G9a^{-/-}$ cells is dependent on H3K9 methylation (Dong et al., 2008). Similarly, H3K9 methylation is sufficient to maintain endogenous LTR retroviral silencing in *Dnmt1/3a/3b* triple knockout mouse ES cells which lack DNA methylation (Matsui et al., 2010).

SUV39H was also shown to be required for directing H3K9me3 and DNMT3b-dependent DNA methylation at pericentric repeats. Dnmt1 single or DNMT3a/DNMT3b double-deficient mouse ES cells neither show reduced H3K9me3 nor show increased transcription at these repeats. However, unlike the $G9a^{-/-}$ cells, $SUV39H^{-/-}$ cells show loss of H3K9me3 and elevated repeat transcription (Lehnertz et al., 2003). From these studies, the presence of H3K9me3 seems to be responsible for transcriptional repression, whereas the presence of DNA methylation is optional.

CTIP2, which gets recruited to the $p21^{WAF1}$ promoter through interactions with DNMTs and HDACs, cooperates with SUV39H1 to cause transcriptional repression in a DNA methylation-dependent manner (Cherrier et al., 2009). At the $p16^{INK4a}$ and the E-cadherin loci in myeloid leukaemia cells, H3K9 methylation and transcriptional silencing also seem dependent on DNA methylation. 5-azacytidine treatment of these cells is sufficient to dissociate SUV39H1 from these loci and cause transcriptional de-repression (Lakshmikuttyamma et al., 2010). However, elegant time-course experiments with colorectal cancer cells show that, even though the loss of DNA methylation leads to reduction in H3K9 methylation at

the p16 locus, it's re-silencing following multiple passages is first established by H3K9 trimethylation and then followed by DNA methylation (Bachman et al., 2003). In these examples, transcriptional repression is lost by the removal of DNA methylation, however the initial establishment of repression may be due to H3K9 methylation. It is becoming clearer that histone modification are not mere effectors of DNA methylation mediated silencing, but, on the contrary, may be responsible for the establishment of silencing. DNA methylation simply seems to be responsible for 'locking in' the repressed state per se.

SINEs have been previously shown to be enriched for H3K9 methylation (Kondo and Issa, 2003). H3K9me3 and H3K27me3 marks also correlate with low Pol III loading on inactive tRNA genes (Barski et al., 2010). The data presented in this chapter show SINEs to be enriched for H3K9me3 and H3K27me3. Although both H3K9me and H3K27me pathways are often found overlapping in transcriptionally silenced regions, the crosstalk between these pathways is not well understood. *Dnmt1* null cells show a loss of H3K27me3 modification, suggesting that polycomb group proteins that recognise this mark may be recruited to SINEs in a DNA methylation-dependent manner. However, loss of DNA methylation does not lead to alleviation of SINE transcriptional repression (Chapter 3). Since the reason for repression in the absence of DNA methylation was being investigated, H3K27me3 was not studied further.

Kondo *et al* had observed that H3K9me3 at SINEs is sensitive to 5-azadeoxycytidine treatment, suggesting that the presence of this modification at SINEs is dependent on DNA methylation (Kondo and Issa, 2003). However, contrary to these findings, *Dnmt1* null cells do not show a reduction in H3K9me3 at SINEs. It is possible that an acute removal of DNA methylation by 5-azadeoxycytidine causes a temporary decrease in H3K9me3 by disturbing the repression equilibrium at these loci, whereas this cannot be seen in stable genetic deletions. 5-azacytidine may also be causing a reduction of H3K9me3 due to associated cytotoxicity and cell stress responses (Juttermann et al., 1994).

The H3K9me3 observed at SINEs is accompanied by K9MT SUV39H1 and associated HP1. Since, H3K9me3 occupancy is not sensitive to loss of DNA methylation, the same can be assumed for SUV39H1 and HP1. Apo-E locus is not

enriched for H3K9me3 and can therefore be transcriptionally de-repressed following loss of DNA methylation (Chapter3).

Chaetocin, a specific SUV39H inhibitor, has recently been used to show the SUV39H1-mediated transcriptional repression of p16, p21 and E-cadherin loci (Cherrier et al., 2009; Lakshmikuttyamma et al., 2010). In the data presented here, chaetocin treatment causes a clear de-repression of SINE transcription, which is accompanied by a decrease in H3K9me3 and increased Pol III occupancy at SINEs. In 2005, Martens *et al* observed that B1 and B2 SINE transcript levels were elevated following SUV39H1h2 knockout in mouse ES cells. However, *Dnmt1*$^{-/-}$ and Dnmt3ab$^{-/-}$ cells did not show a similar elevation in SINE transcript levels (Martens et al., 2005). The data obtained in Chapters 3 and 4 agree with these findings.

No increase in TFIIIC enrichment is observed following chaetocin treatment, suggesting that H3K9me3 inhibits transcription by specifically limiting the polymerase recruitment at SINEs. B1 and B2 expression levels increase in response to chaetocin treatment even in *Dnmt1*$^{-/-}$ cells, indicating that SUV39H1 and H3K9me3 can maintain SINE transcriptional inhibition even in the absence of DNA methylation. Moreover, in wild-type cells, a reduction in H3K9me3 does not cause reduced MBD2 occupancy, indicating that these epigenetic modifications are not dependent on each other at SINEs. However, these data are obtained from a 24hr drug treatment that does not allow for a passive loss of DNA methylation and a genetic deletion of SUV39H1 may show a loss of DNA methylation at SINEs.

HDACs are associated with both DNA methylation and H3K9me3-mediated transcriptional silencing (Vaute et al., 2002; Zhang et al., 1999). B1 and B2 expression is found to be sensitive to TSA treatment in control cells. However, this is found not to be the case in absence of DNA methylation in *Dnmt1*$^{-/-}$ cells. *Dnmt1*$^{-/-}$ cells also show a reduction in HDAC1 and HDAC2 occupancy at SINEs, whereas, no consequent elevation in Pol III loading is observed. These data together suggest that HDAC-dependent silencing of SINEs is reliant on DNA methylation; however this repression is superseded by an alternative mechanism in the absence of DNA methylation, involving H3K9me3 and HP1.

This chapter demonstrates that DNA methylation and H3K9me3 provide two different mechanisms for SINE transcriptional repression, the former being HDAC-dependent and the later being HDAC-independent. However, H3K9 methylation is the primary cause of SINE transcriptional repression that supersedes DNA methylation.

Chapter 5 DNA methylation and recombination at SINEs

5.1 Introduction

Alu sequences show a high level of CpG methylation in most somatic tissues (Rubin et al., 1994; Schmid, 1991). DNA methylation also seems to be feature of murine SINEs (Jeong and Lee, 2005; Yates et al., 1999). The data obtained during this study clearly demonstrate that a loss of DNA methylation does not cause elevated SINE expression. Even though DNA methylation seems to have a contribution to SINE transcriptional inhibition, this contribution is relatively minor and is superseded by repression due to H3K9 trimethylation. So why is DNA methylation so prevalent on SINEs and does it serve any other purpose?

Hypomethylation is common in solid tumours such as metastatic hepatocellular carcinomas (Lin et al., 2001), cervical cancers (Kim et al., 1994) and prostate tumours (Bedford and van Helden, 1987), and is also found in non-solid tumours such as B-cell chronic lymphocytic leukaemias (Wahlfors et al., 1992). Genomic hypomethylation correlates with increased disease progression and poor prognosis (Ehrlich, 2002). Hypomethylation is also often linked to chromosomal instability (CIN) in multiple human diseases (Ehrlich, 2005).

Loss of methylation on centromeric satellite repeats is observed in multiple tumour types, including Wilms tumours, breast adenocarcinomas, ovarian epithelial carcinomas and hepatocellular carcinomas (Ehrlich, 2005). Pericentric repeat hypomethylation has been shown to be highly correlative with CIN in hepatocellular carcinomas (Wong et al., 2001). The best example for the link between DNA hypomethylation and CIN is provided by the rare recessive autosomal disorder known as immunodeficiency, centromeric instability and facial anomalies syndrome ICF). Biallelic mutations in DNMT3B cause hypomethylation at centromeric repeats leading to CIN in ICF (Xu et al., 1999).

The loss of DNA methylation has been thought to lead to elevated transcription of transposable elements and thus contribute to genomic instability via an increase in transposition. This may be true for LINE and IAP (intercisternal A particle) elements (Bourc'his and Bestor, 2004; Howard et al., 2008; Jurgens et al., 1996), but no such evidence for Alu elements has been published. Moreover, the data presented in the previous chapters would argue against this possibility,

since the loss of DNA methylation does not lead to an elevation in SINE expression.

Double-strand breaks are a threat to genomic stability and the two primary cellular mechanisms for dealing with this threat are homologous repair (HR) and non-homologous end joining (NHEJ). HR uses an identical sequence from the sister chromatid as a template for repair, whereas NHEJ simply ligates the broken ends back together. The choice between HR and NHEJ can depend on multiple factors such as the stage of the cell cycle, the DNA structure and the difficulty of repair. NHEJ is thought to be the preferred repair pathway for DSBs, but when the repair is more complicated HR machinery is engaged (Helleday et al., 2007).

A third repair pathway called single strand annealing (SSA) is used particularly if two adjacent repeat sequences are involved. During SSA, the double-stranded DNA on both the repeat sequences is resected to give single-stranded DNA (ssDNA), which are then aligned and ligated. One of the repeat alleles and any intermediate DNA are always lost during SSA. Therefore, this repair pathway is highly mutagenic and can lead to permanent large deletions or chromosomal translocations (Haber, 2006).

Alu sequences are homologous repeats and are, therefore, prime substrates for SSA. Erroneous repair between Alu repeats can be quite deleterious, since Alus are spread throughout the protein-coding regions of the genome (Korenberg and Rykowski, 1988). Areas of higher than average Alu density appear to be particularly associated with genomic instability. Alus are often found in the vicinity or even within the breakage points of translocations (Kolomietz et al., 2002). Recombination between Alu sequences has been implicated in the etiology of several human genetic diseases, including α-thalassemia (Nicholls et al., 1987), hypercholesterolemia (Lehrman et al., 1987) and many cancers (Kolomietz et al., 2002).

Erroneous repair at Alu sequences and DNA hypomethylation are both directly linked to genomic instability. So, does DNA methylation at Alus repress erroneous repair?

A model system designed by Elliot *et al* was employed to answer this question (Figure 5.1). Two homologous Alu elements have been stably inserted at the *Rb* locus on chromosome 14 and the *Pim1* locus on chromosome 17 in a mouse ES cell line called Hom6.9 (Elliott et al., 2005). The Alu was derived from the mixed lineage leukaemia gene (*MLL*), where the erroneous recombination between Alu sequences in introns 6 and 1 gives rise to a MLL self-fusion protein (So et al., 1997). These Alu sequences are inserted 3' and 5' of the neoMycin phosphotransferase gene splice donor and splice acceptor portions. An I-SceI site is inserted 5' and 3' of the Alu sequences respectively (Elliott et al., 2005).

A targeted double-strand break can be introduced by electroporating the cells with a plasmid coding for the I-SceI restriction endonuclease. It was observed that an erroneous recombination event occurred with the frequency of 5.0 ± 4.0 x 10^{-5}, resulting in chromosomal translocation. The chromosome derived as a result of this translocation contained the functional neoMycin phosphotransferase gene. The translocation frequency was scored by counting the number of resistant colonies following neoMycin selection and dividing the colony count by the total number of cells seeded (Elliott et al., 2005). So what effect would demethylation of these Alu elements have on translocation frequency?

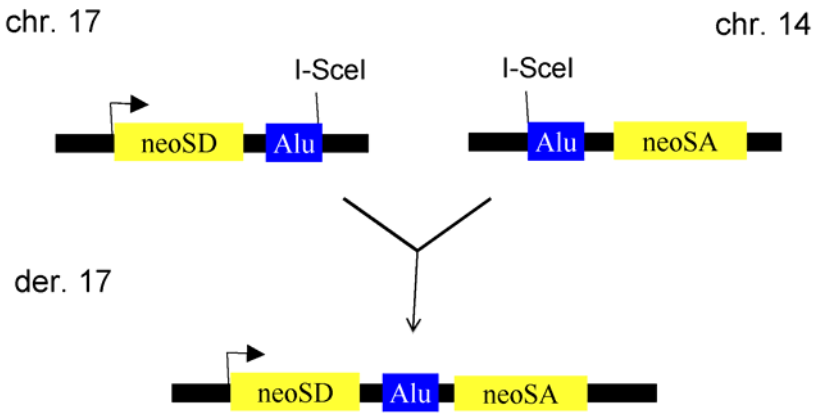


Figure 5.1: Schematic – Translocation assay
Homologous Alu elements (blue boxes) from the human MLL gene are inserted on chromosomes 17 and 14 in mouse ES cells (Hom6.9 cells). These homologous Alu elements are 3' and 5' of the splice donor and splice acceptor portions of neoMycin resistance gene (neoSD and neoSA, yellow boxes). Double strand breaks (DSB) are generated by the I-SceI endonuclease. Erroneous repair can give rise to chromosomal translocations, where the derivative chromosome causes the cells to become neoMycin resistant. The frequency of this event was quantified to be 5.0 ± 4.0 translocations/10^5 cells. (Elliott et al., 2005).

5.2 Results

5.2.1 Ectopic Alus in mouse ES cells are methylated and sensitive to 5-azacytidine

In order to investigate the effect of DNA methylation on translocation frequency at the Alu(*MLL*) element in the Hom6.9 cell line, it was first necessary to determine the DNA methylation status of these Alu sequences. The data in previous chapters show that the MBP enrichment at Alu elements is a clear indication of their DNA methylation status. Thus, in order to establish the methylation status of Alu(*MLL*), Hom6.9 cells were subject to ChIP assay with antibodies against MBD2 and MeCP2 (Figure 5.2). The polymerase occupancy at these ectopic Alu loci was also investigated using an antibody against the RPC155 subunit of Pol III. Beads-only and TAF_I48 antibody provide the negative controls. The pol II-transcribed p21 gene promoter provides the positive control for MBP occupancy (Cherrier et al., 2009) and negative control for Pol III occupancy, whereas, the $tRNA^{Leu}$ gene provides the reverse control.

PCR with primers recognising both the ectopic Alu(*MLL*) sequences shows the clear presence of MBD2 and MeCP2, indicating that these sequences are methylated. Higher enrichment of MeCP2 is observed at Alu(*MLL*) loci compared to MBD2, which is in agreement with the previous findings at endogenous Alu sites in human cells. Moreover, Pol III occupancy is also observed; however Alu(*MLL*) shows considerably less Pol III occupancy compared to $tRNA^{Leu}$ gene (Note: The Pol III band in the $tRNA^{Leu}$ PCR is overexposed).

Treatment of Hom6.9 cells for 16hrs with 5-azacytidine is able to demethylate their genomes and cause considerable reduction in MBP enrichment at B1 and B2 SINEs (Figure 3.13). Clear eviction of MBPs from Alu(*MLL*) loci is also observed, however, 5-azacytidine does not cause an increase in Pol III occupancy at SINEs. These results confirm that Alu(*MLL*) behaves in a manner similar to Alu sequences within human cells, and DNA methylation alongside Pol III occupancy are inherent features of SINEs. Moreover, these data suggest that any findings using this artificial system can be extrapolated to Alu sequences in human cells.

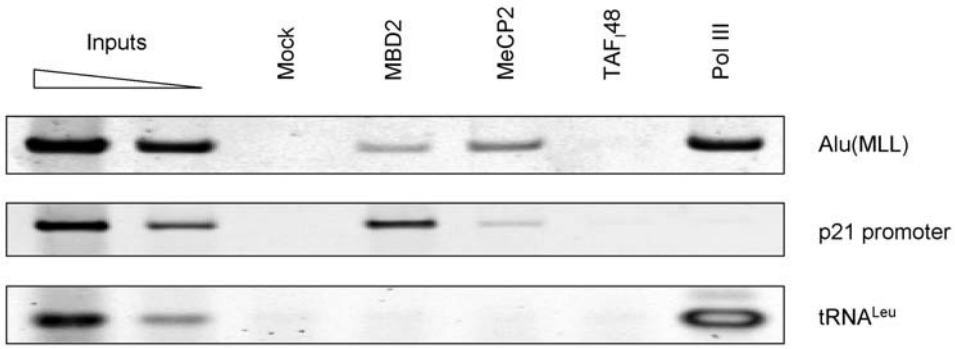


Figure 5.2: Ectopic Alus in the mouse genome behave similarly to Alus in the human genome
ChIP assay in Hom6.9 cells with antibodies against MBD2, MeCP2 and RPC155 at Alu(MLL) sequences on chromosomes 17 and 14. ChIPs with beads only and TAF$_I$48 antibody provide the negative controls. tRNA^{leu} and p21 promoter, respectively, provide the positive and negative controls for polymerase occupancy and visa versa for MBP occupancy. Serial dilutions of the inputs show the quantitative nature of the PCR.

5.2.2 DNA methylation suppresses rearrangement at SINEs

The translocation assay was performed on Hom6.9 cells as described in Elliot *et al,* 2005 (Figure 5.4). Cells were allowed to recover for 10hrs following electroporation with I-SceI vector (pCBAS) and then were treated with 5-azacytidine for 16 hours . This duration of treatment is sufficient to demethylate Alu(MLL) sequences as it leads to clear eviction of MBPs (Figure3.14). However, any alterations observed in the translocation frequency could result from the cytotoxic effects of 5-azacytidine (Juttermann et al., 1994; Murakami et al., 1995). To rule out this possibility, cells were alternatively treated with bleoMycin, an agent that causes double strand breaks. Cells treated with 5-azacytidine and bleoMycin showed the presence of γ-H2Ax foci as viewed by immunofluorescence (data not shown).

Negative control cells electroporated with pTK hyg plasmid do not survive selection with neoMycin, since there is no targeted DSB at the Alu(*MLL*) loci. Moreover, 5-azacytidine or bleoMycin treatments do not alter this result. These data indicate that the drug treatment on its own is not sufficient to cause the translocation event and that the number of colonies obtained in the translocation experiment reflects the frequency of erroneous repair at the Alu(*MLL*) loci.

The untreated cells electroporated with pCBAS (I-SceI) show the clear presence of colonies following 8-10 day selection with neoMycin. This indicates that the translocation assay was successful and, unlike pTK hyg, pCBAS can lead to targeted DSB at Alu(*MLL*) loci which can then be subjected to erroneous repair.

The translocation frequency is calculated to be 3.5 ± 2 translocations/10^5 cells, which is similar to what was previously reported by Elliot *et al*. 5-azacytidine treatment leads to a small but insignificant increase in the total number of colonies observed following neoMycin selection. On the other hand, a dramatic reduction in the total colony count is observed following bleoMycin treatment.

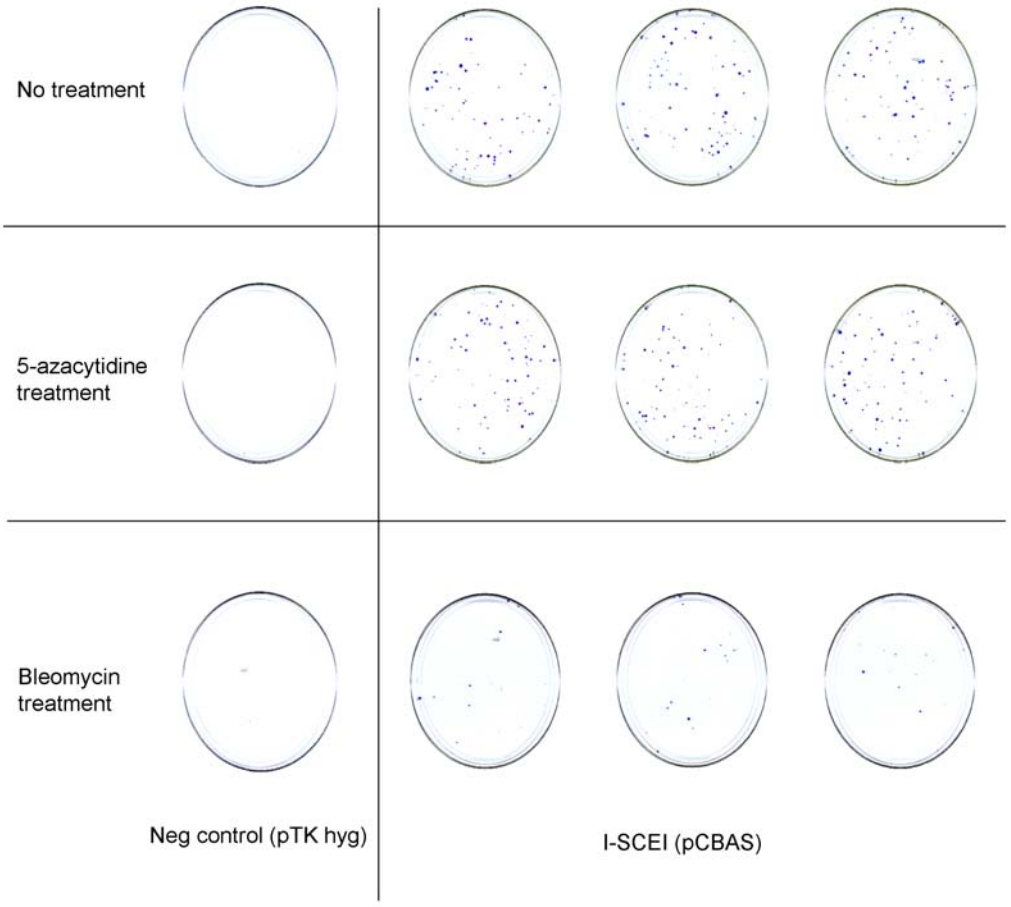


Figure 5.3: Translocation assay following 5-azacytidine and bleoMycin treatments
Translocation assay was performed in Hom6.9 cells with 25µg of I-SceI plasmid (pCBAS) and negative control plasmid (pTK hyg). The cells were then treated with 4µM 5-azacytidine and 5µg/ml bleoMycin for 16hours. To obtain resistant colonies, selection for 8 to 10 days with 200µg/ml neoMycin was performed. Colony counts were performed following GIEMSA staining.

However, the total colony count is not a true reflection of the translocation frequency following drug treatment, since both the drugs are cytotoxic and lead to a substantial reduction in cell numbers on control plates not selected with neoMycin. 5-azacytidine and bleoMycin cause ~65% and ~85% loss in cell viability respectively. Therefore, the colony counts must be normalised to this loss in cell viability.

Once normalised, the translocation frequency was plotted relative to the translocation frequency observed in control cells (Figure 5.4). 5-azacytidine shows a significant elevation in chromosomal translocation frequency compared to untreated control cells ($p<0.01$), whereas bleoMycin is unable to cause any alteration. These data indicate that the increased translocation is due to the demethylating property of 5-azacytidine and not its cytotoxicity. Moreover, an overall elevation in cellular DNA repair following bleoMycin treatment does not lead to an increase in erroneous repair at Alu(*MLL*).

Therefore, it can be concluded that DNA methylation inhibits erroneous repair at SINEs, thus contributing to genomic stability.

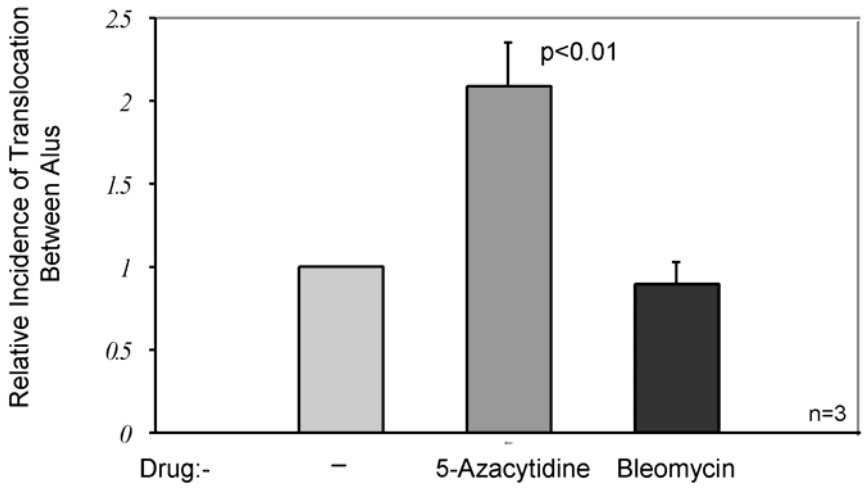


Figure 5.4: Loss of DNA methylation elevates the rate of recombination at SINEs
Relative incidence of translocation between homologous Alu elements in Hom6.9 cells was calculated by counting the number of neoMycin resistant colonies. The count was obtained from three independent biological repeats each performed in triplicates. This count was first normalised to the average loss of cell viability following drug treatment and then to the average number of colonies on the untreated plates.

As previously stated, Alu sequences provide prime targets for SSA-mediated mutagenesis. Elliot *et al* had previously reported that Hom6.9 cells undergo chromosomal translocation via the SSA repair pathway (Elliott et al., 2005). However, the dramatic increase in translocation frequency following the loss of DNA methylation could be due to an alteration in the repair pathway used.

In order to investigate this, seven clones were picked from the control and 5-azacytidine-treated Hom6.9 plates. The genomic DNA from these clones was subjected to PCR with primers flanking the Alu(*MLL*) sequences on different chromosomes (Figure 5.5). Since SSA would lead to the loss of one of the Alu(*MLL*) allele and NHEJ will not, the size of the PCR product obtained would be different depending on the repair pathway used for the translocation. Alus recombined as a result of SSA would be expected to give a 1kb PCR product, whereas Alus recombined as a result of NHEJ would result in a 1.3kb PCR product. DNA from Hom6.9 cells not subjected to the translocation protocol provides the negative control and demonstrates the specificity of the PCR reaction.

A 1kb product is obtained from the PCR reaction on DNA from all the untreated clones, indicating that SSA is the repair pathway that causes the translocation. Moreover, 5-azacytine treated clones show no deviation from this pattern, indicating that a loss of DNA methylation does not lead to alteration in the repair pathway responsible for the erroneous repair. It is possible that the observed increase in translocation frequency is due to an increase in the kinetics of SSA repair. However, the details of how DNA methylation represses SSA repair are still unclear.

Thus, DNA methylation serves to suppress erroneous repair at SINEs, possibly by suppressing error-prone repair pathways like SSA.

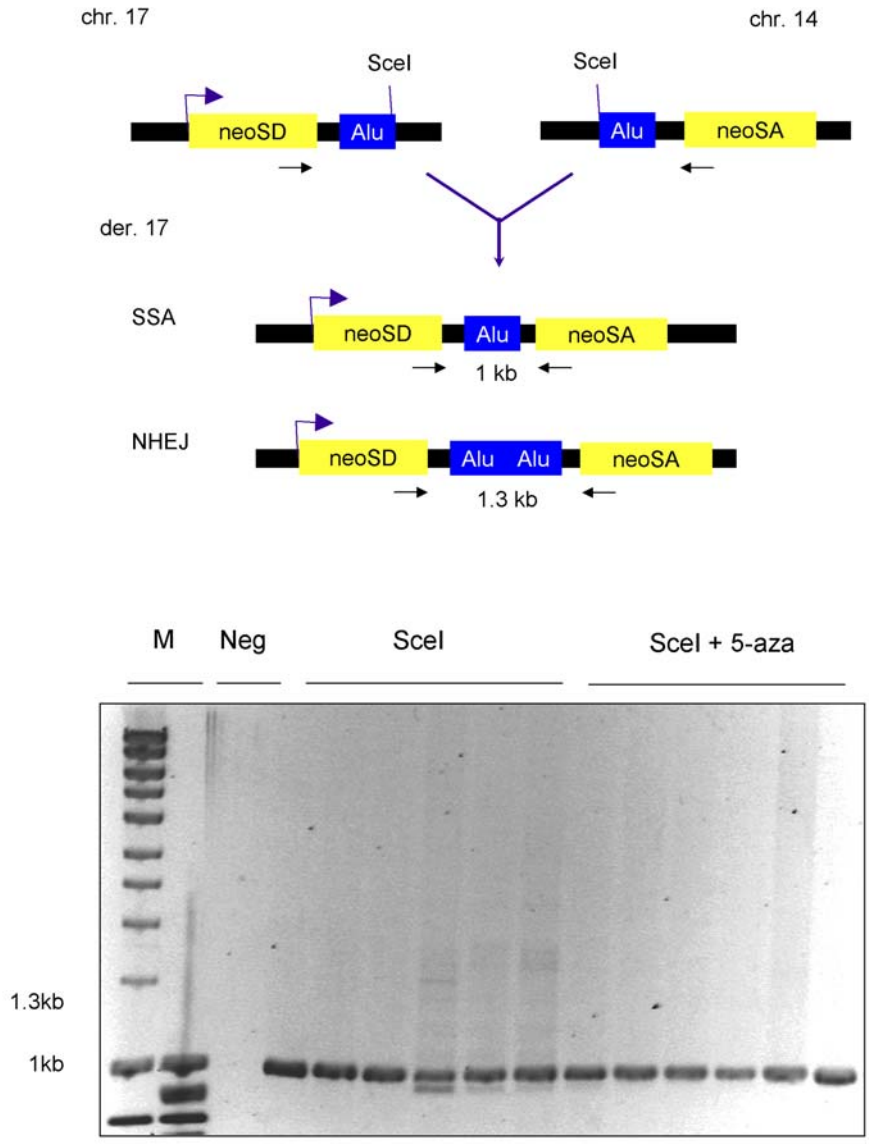


Figure 5.5: Increased recombination in the absence of DNA methylation is not due to a change in the repair pathway used
PCR on DNA from clones picked following translocation assay in Hom6.9 cells with or without 5-azacytidine treatment. The schematic on top depicts the location of the primer pair used (black arrows) and the different possible PCR product sizes depending on the repair pathway. Neg lane shows the PCR reaction on DNA from Hom6.9 cells that did not undergo the translocation protocol. M lanes are 1kb and 100bp DNA ladders.

5.3 Discussion

Repetitive elements comprise at least 45% of the human genome and present ample opportunity for genomic recombination. With 1 million copies, Alu elements are the most abundant repeats in the human genome. Since Alus are spread throughout the protein-coding regions, rearrangements due to Alus can have devastating effects on genomic stability and have been implicated in many human diseases (Deininger et al., 2003; Konkel and Batzer, 2011; Lander et al., 2001).

Using a model for rearrangements driven by identical Alu elements, this chapter has shown that DNA methylation significantly suppresses interchromosomal translocation between identical Alu SINEs ($p<0.01$). It has been previously reported that translocations between identical Alu elements arise due to erroneous SSA repair (Elliott et al., 2005). A loss of DNA methylation does not lead to an alteration in the repair pathway used, but causes an increase in the translocation frequency. However, the reason for elevated SSA at the Alu(*MLL*) loci in the absence of DNA methylation is still unclear.

RPA32 is a component of the replication protein A (RPA) complex, which binds single-stranded DNA obtained following strand resection. RPA binding is thought to facilitate single-strand annealing and thus promote SSA (Helleday et al., 2007). Preliminary ChIP experiments on HeLa cells with an antibody against RPA32 show that 5-azacytidine treated Alus are more abundantly bound by this ssDNA-binding protein compared to untreated or bleoMycin-treated Alus (data not shown). These data, although preliminary, suggest that a loss of DNA methylation may lead to increased amounts of RPA-bound ssDNA within the cell. Thus DNA methylation may either inhibit the strand-resection kinetics or RPA-binding to ssDNA.

The interplay between DNA methylation and repair is not a new finding. DNMT1 is recruited to sites of DSB repair via its interaction with PCNA and DMAP1 (Lee et al., 2010; Mortusewicz et al., 2005). This is thought to cause *de novo* methylation at recombined loci and silence them (Cuozzo et al., 2007). However, depletion of DMAP1 leads to elevated levels of HR, suggesting that the

recruitment of DNMT1 to the sites of DSB repair may somehow throttle the repair process (Lee et al., 2010).

As previously mentioned, DNA hypomethylation has been correlated with CIN in multiple human diseases. *Dnmt1*-null cells show a high amount of chromosomal translocations as seen by chromosome painting (Karpf and Matsui, 2005). Demethylation by 5-azacytidine has also been shown to induce sister chromatid exchange *in vitro* and *in vivo*. This effect of 5-azacytidine has been shown to be due to its demethylating activity rather than the associated cytotoxicity (Albanesi et al., 1999; Morales-Ramirez et al., 2007). The data presented in this chapter demonstrate that Alu sequences may contribute to the genomic instability observed in the absence of DNA methylation.

5-azacytidine and its analogues are FDA-approved treatments for myelodysplastic syndrome and other leukemias, where demethylation is thought to reactivate repressed tumour suppressor genes and cause tumour regression (Christman, 2002). Clinical trials with 5-azacytidine in patients with melanoma and renal cell carcinoma have also been reported (Gollob et al., 2006). However, 5-azacytidine treatment has also been associated with genome-wide DNA damage, CIN and reversal of differentiation in ES cells (Christman, 2002; Tsuji-Takayama et al., 2004). CIN caused by DNA demethylation could lead to multi-drug resistance and tumour relapse by causing tumour heterogeneity (Lee et al., 2011; Sotillo et al., 2010). In addition to these existing studies, the data obtained in this chapter would advise a careful review of the use of DNA demethylating agents in therapy.

It must be noted that a 20% sequence divergence between Alu sequences causes a shift from SSA to NHEJ as the preferred mode of repair. This, however, does not lead to any alterations in the overall translocation frequency (Elliott et al., 2005). The effects of DNA demethylation on NHEJ may not be similar to its effects on SSA. Crucially, NHEJ does not require the strand resection step (Helleday et al., 2007), which, according to preliminary data, seems to be the process influenced by DNA methylation.

Since different Alu sequences show ~14% sequence divergence from the Alu consensus sequence (Deininger and Daniels, 1986), it is likely that the majority

of recombination events between Alus *in vivo* result from NHEJ as opposed to SSA. Elliot *et al* have also designed a system to study chromosomal translocation events between heterologous Alu elements. Initial attempts to use this Het-Alu system were unsuccessful due to extremely high cell death following electroporation. However, the effect of DNA demethylation on the translocation frequency in the Het-Alu cell line must be analysed in order to fully understand the mechanisms of erroneous repair.

The Alu(MLL) elements used in the translocation assay in the Hom6.9 cells are occupied by Pol III. This indicates that, even in this artificial system, the recruitment of Pol III is an inherent feature of SINEs. Moreover, these ectopic Alus show the same preference for MeCP2 binding over MBD2 binding that is seen at endogenous Alus in the human genome. Thus, this ectopic Alu sequence seems to behave in a manner very similar to endogenous SINEs. So is its transcription inhibited by H3K9me3? What effects would treatments with chaetocin or TSA have on the translocation frequency at these ectopic Alus? In addition to this, does the Pol III machinery interfere with the translocation event? Would B-block mutations or knockdown of TFIIIC elevate the translocation frequency by making more of the DNA available for erroneous repair? All these questions still remain to be answered.

From the data obtained so far, it is clear that DNA methylation negatively regulates erroneous repair at homologous Alu elements and protects the genome from instability. But how this is achieved is still an unanswered question.

Chapter 6 Chromatin remodelling and Pol III transcription

6.1 Introduction

Pol III transcriptional output is of vital importance for cellular growth and proliferation, and therefore, is subjected to tight regulation (see section 1.1.4). The previous chapters reveal that transcription of mammalian SINE sequences by Pol III is inhibited by H3K9 trimethylation. The first identified example of regulation of Pol III transcription by chromatin comes from the *Xenopus* 5S rRNA genes. The oocyte 5S genes show lower affinity for TFIIIA compared to the somatic 5S genes, and therefore less stable transcription complexes are formed. Following mid-blastula transition (MBT), the cellular levels of TFIIIA drop and the oocyte genes lose their expression due to their lower TFIIIA affinity. Introduction of excess TFIIIA is able to raise repression of oocyte 5S genes, but this effect is transient. The organisation of the oocyte 5S genes into compact chromatin by histone H1 is thought to contribute significantly to their silencing following MBT (Wolffe and Brown, 1988).

Indeed, a positioned nucleosome near the 5S ICR was found to inhibit transcription initiation and downstream nucleosomes were shown to block elongation in an oocyte nuclear extract (Morse, 1989). Multiple other studies have shown that repressive chromatin structure, once formed over the 5S gene, can inhibit the binding of TFIIIA and inhibit transcription (White, 2002). However, TFIIIA can easily bind 5S ICR upon removal or acetylation of histone tails, suggesting that post-translational histone modifications play a decisive role in Pol III transcriptional regulation (Lee et al., 1993). A positioned nucleosome is also thought to be required to juxtapose PSE and DSE within the U6 promoter in order to facilitate the Oct-1 and $SNAP_c$ interaction (Stunkel et al., 1997). Alu sequences have been shown to be able to position nucleosomes, where a positioned nucleosome is able to repress Pol III-dependent Alu transcription (Englander and Howard, 1995).

In contrast, the literature suggests that tRNA transcription is not influenced by chromatin (White, 2002). Manipulating the levels of histone H1 in *Xenopus* oocytes causes alterations in 5S rRNA expression, but tRNA expression is unaffected (Bouvet et al., 1994). Moreover, *Drosophila* tRNA genes were found to have increased sensitivity to micrococcal nuclease (MNase) and DNaseI digestion, indicating that tRNA genes are distinctly devoid of nucleosomes

(DeLotto and Schedl, 1984). tRNA genes are found to be transcriptionally active in hepatocytes and inactive in erythrocytes. Here, the correspondence between chromatin accessibility and transcription seems to break down, since tRNA genes show similar DNaseI sensitivity in both these cell types (Coveney and Woodland, 1982).

This insensitivity of tRNA loci to chromatin regulation has been attributed to the intrinsic HAT activity of TFIIIC. Immunopurified TFIIIC was found to be able to bind A- and B-block sequences and initiate transcription of a tRNA gene within an *in vitro* reconstituted chromatin template. Two TFIIIC subunits, TFIIIC90 and TFIIIC110 were found to have intrinsic HAT activity, where the former preferentially acetylates H3K14 residue and the latter acetylates H3, H4 and H2B (Hsieh et al., 1999a; Kundu et al., 1999). In addition to this the small size of tRNA genes could also account for their insensitivity to chromatin-mediated transcriptional inhibition. PIC assembly is thought to be the major rate-limiting step in Pol III transcription, and, once the PIC is in place, the small size of the tRNA gene means that transcriptional elongation is not a substantial hurdle (White, 2002). Pol II transcription can be repressed by chromatin during the elongation step, whereas Pol III may be able to escape this repression due to the relatively short elongation step. Thus, the ability of TFIIIC to form a PIC in a repressive chromatin environment and the relatively short size of tRNA genes may explain their insensitivity to chromatin-mediated transcriptional repression.

Contrary to what a substantial amount of literature suggested at the time, tRNA genes were found to be incorporated into inactive chromatin in chick embryos and tRNA gene transcription was found to be influenced by nucleosome positioning. It was also shown that when a tRNA gene is located at the edge of a nucleosome, its transcription is drastically repressed. Intriguingly, when the tRNA gene is positioned in the middle or close to the dyad axis, transcription is elevated 30-fold (Wittig and Wittig, 1982). These data suggest that nucleosome positioning may play an important role in tRNA gene transcription. In addition to this, a yeast tRNA gene was also found to be repressed when inserted into the silent HMR mating-type locus (Schnell and Rine, 1986).

More recently, various other studies have also indicated that chromatin may in fact influence the transcription of tRNA genes. Different tRNA genes in the

mulberry silkworm *Bombyx mori* were found to be occupied by a nucleosome within the 5'-upstream region and this nucleosome occupancy correlated with poor expression. The positioned nucleosome was found to impair TFIIIB recruitment and the addition of excess TFIIIC with intrinsic HAT activity was able to relieve transcription. Moreover, deletion of the upstream sequence also led to the loss of transcriptional repression. This suggests that, despite the similarity of their ICR sequences, differential expression of tRNA isoacceptor genes in cells may be achieved by their unique flanking sequences and associated chromatin (Parthasarthy and Gopinathan, 2005).

c-MYC was found to induce tRNA expression by inducing Gcn5-mediated histone acetylation. Once it is recruited to the promoters, potentially through its binding with TFIIIC (unpublished data, Kirsteen Campbell), c-MYC was found to recruit Gcn5 via its interaction with TRRAP. Gcn5 acetylates histone H3 and promotes transcription at tRNA genes by promoting TFIIIB and Pol III recruitment. Histone H4 acetylation and TFIIIC enrichment does not change in response to c-MYC, indicating that TFIIIC may be able to occupy tRNA genes within a repressive chromatin environment, potentially by acetylating H4. However, TFIIIC occupancy on its own may not be sufficient for optimal PIC formation in a repressive chromatin environment (Kenneth et al., 2007). The initial evidence for TFIIIC itself having HAT activity was weak and it has now been shown that TFIIIC directly interacts with p300 and recruits it to tRNA genes. The recruitment of p300 not only acetylates the histones, but also stabilizes the binding of TFIIIC to the core promoter elements (Mertens and Roeder, 2008).

The mammalian histone chaperone FACT (facilitates chromatin transcription) subunit SSRP1 was found to associate with tRNA, 5S rRNA and 7SL loci. Knockdown of SSRP1 leads to a reduction in tRNA and 7SL rRNA expression, indicating that FACT facilitates Pol III transcription (Birch et al., 2009). Loss of Sth1, the ATPase subunit of the yeast RSC complex, was also found to cause a reduction in Pol III transcriptional output. This is accompanied by a loss of Pol III loading and an elevation in nucleosome density (Parnell et al., 2008). Recently, FACT and RSC complexes were shown to regulate the transcription of a yeast $tRNA^{Tyr}$ gene by specific deposition of H2A.Z-containing nucleosomes that flank the transcribed region (Mahapatra et al., 2011).

In addition to these studies, recent ChIP-seq and RNA-seq analyses have revealed that active tRNA genes are marked by histone modifications associated with active transcription of Pol II-transcribed genes. Active tRNA loci are enriched for histone acetylation, H3K4me1/2/3 and H2A.Z, whereas inactive tRNA genes show inactive chromatin marks such as H3K9me3 and H3K27me2/3 (Barski et al., 2010).

Thus, it seems that tRNA genes are not exempt from chromatin-mediated regulation. However, the effects of chromatin on tRNA transcription are still poorly understood and there is a distinct lack of studies that look at these. Regulation by chromatin may contribute significantly to defining Pol III transcriptional output and understanding these regulatory mechanisms may be of vital importance.

6.2 Results

6.2.1 A defined chromatin structure and SWI-SNF chromatin remodelling complex at Pol III-transcribed genes

Before chromatin-mediated transcriptional regulation could be studied, it was necessary to define the chromatin structure of a Pol III-transcribed gene. This has been done in yeast, however a mammalian locus has not been investigated so far. The small size of tRNA genes makes it difficult to perform a detailed analysis of the chromatin structure and associated histone marks. However, the 7SL locus is 300bp long and can be utilised for such analysis. ChIP assays in conjunction with mapping PCRs were used to map the location of the Pol III transcription apparatus, nucleosomes and histone modifications. Such an approach has been previously used to map Maf1 localisation on the 7SL gene locus (Goodfellow et al., 2008).

HeLa chromatin was subjected to ChIP assays with antibodies against RPC155, BRF1 and TFIIIC. Antibodies against core histones H2A, H2B and H3 were used to map the nucleosome occupancy. Active chromatin marks acetyl-H3 and H3K4 trimethylation were also investigated in order to gain an idea about the state of the chromatin. Beads-only and TAF_I48 antibody provided the negative controls for each ChIP. The mapping primers amplify five different regions within and around the 7SL locus (Figure 6.1).

Pol III and TFIIIB occupancy peaks at the beginning of the gene (primer set 2), whereas the TFIIIC is more evenly distributed across the locus. This is in agreement with the known model for the PIC formation on type II promoters (Figure 1.2) and with what was observed previously (Goodfellow et al., 2008). The 7SL promoter region shows a distinct decrease in core histone occupancy and an increase in the acetyl-H3 and H3K4me3.

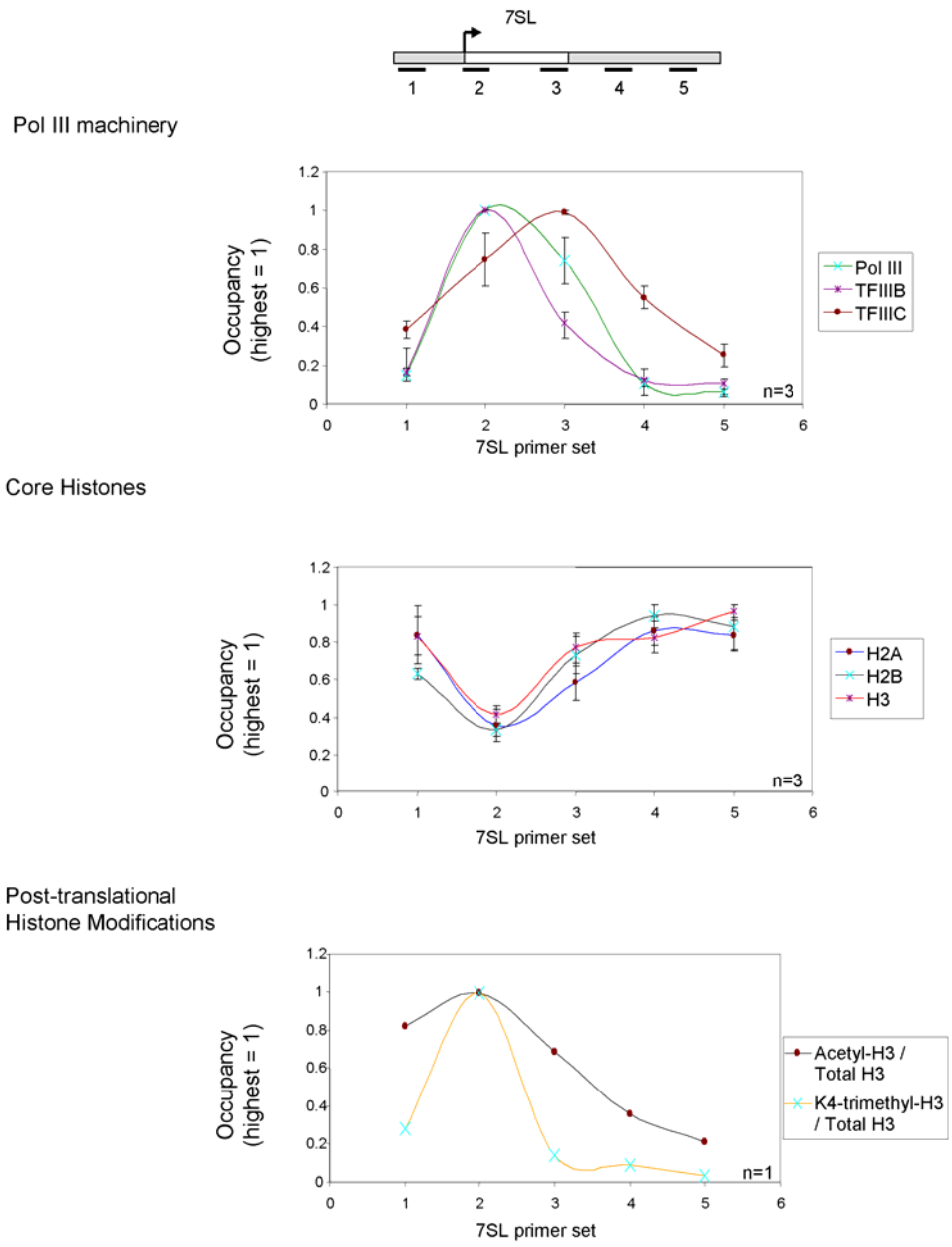


Figure 6.1: Chromatin structure of a Pol III-transcribed 7SL locus
ChIP in HeLa cells with antibodies against RPC155, BRF1, TFIIIC220; core histones H2A, H2B, H3; activating histone modifications acetyl-histone H3 and lysine4 trimethyl-histone H3. PCRs performed with mapping primers (top schematic) at 7SL locus. ChIPs with beads only and TAF_I48 antibodies provide the negative controls. The plotted values were calculated using the following equation y=[(IP/Input)-(Avg. Neg. IP/Input)] and highest y=1. ChIP values for histone modifications were further normalised to total histone H3 levels.

These data show that the 7SL locus has a well-defined chromatin structure with a nucleosome-free region over the transcribed sequence. MNase digestions coupled with scanning PCRs also confirmed these findings, where the DNA amplified by the 7SL set2 primers was found to be hypersensitive to digestion (data not shown). Thus, the 7SL gene seems to have a chromatin structure similar to the tRNA genes in this respect and can be used as a model (Barski et al., 2010).

The presence of such defined chromatin structure could be a result of active transcription and highly organised transcription apparatus. However, the presence of acetyl-H3 and H3K4me3 marks suggest that some degree of chromatin remodelling may be involved. In addition to this, there was previous evidence in the lab to suggest the presence of SWI/SNF chromatin remodelling complex at tRNA genes (Niall Kenneth, unpublished data). The presence of such a remodelling complex could explain the highly ordered chromatin structure observed on the 7SL gene.

Chromatin from A31 mouse fibroblasts was subjected to ChIP assays with antibodies against SNF5, BRG1 and BRM subunits of the SWI/SNF complex (Figure 6.2). Beads only (Mock) and TAF_I48 antibodies provide the negative controls, whereas, a Pol III antibody provides the positive control. The Pol II-transcribed ARPP P0 locus provides the negative control PCR. All three SWI/SNF subunits were found to occupy multiple tRNA and 5S rRNA genes. BRG1 and BRM ATPase subunits do not show considerable differences in enrichment levels.

A literature search revealed that SWI/SNF occupancy on tRNA genes has also been observed in ChIP-seq experiments on HeLa chromatin (Euskirchen et al., 2011). Further data mining revealed that 33.5% of all annotated tRNA genes were occupied by at least one SWI/SNF subunit (Figure 6.3). Core subunits BAF155 and BAF170 were found to be enriched on 146 and 193 tRNA loci, respectively. SNF5 was found to occupy 150 tRNA loci, whereas BRG1 was seen to occupy 55 tRNA genes. This relatively low gene count for BRG1 occupancy could be due to lower antibody efficiency or the presence of BRM as the ATPase subunit at many loci. BRM ChIP-seq. was not performed by the authors. Thus, the core subunits of the SWI/SNF chromatin remodelling complex are present at multiple tRNA genes and may have a role in their transcriptional regulation.

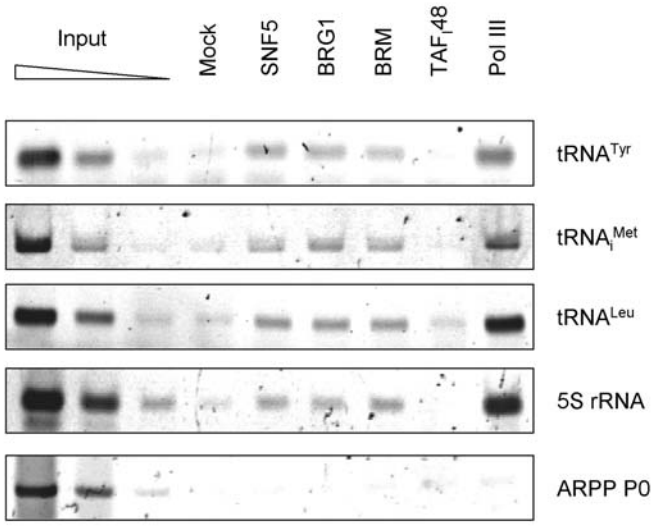


Figure 6.2: SWI/SNF subunits occupy Pol III-transcribed genes
ChIP in A31 mouse fibroblasts with antibodies against SNF5, BRG1 and BRM at tRNA and 5S rRNA loci. ChIPs with beads only and TAF$_I$48 antibodies provide the negative controls, whereas ChIP with RPC155 antibody provides the positive control. ARPP P0 gene provides the negative control for the ChIP. Serial dilutions of inputs show the quantitative nature of the PCR.

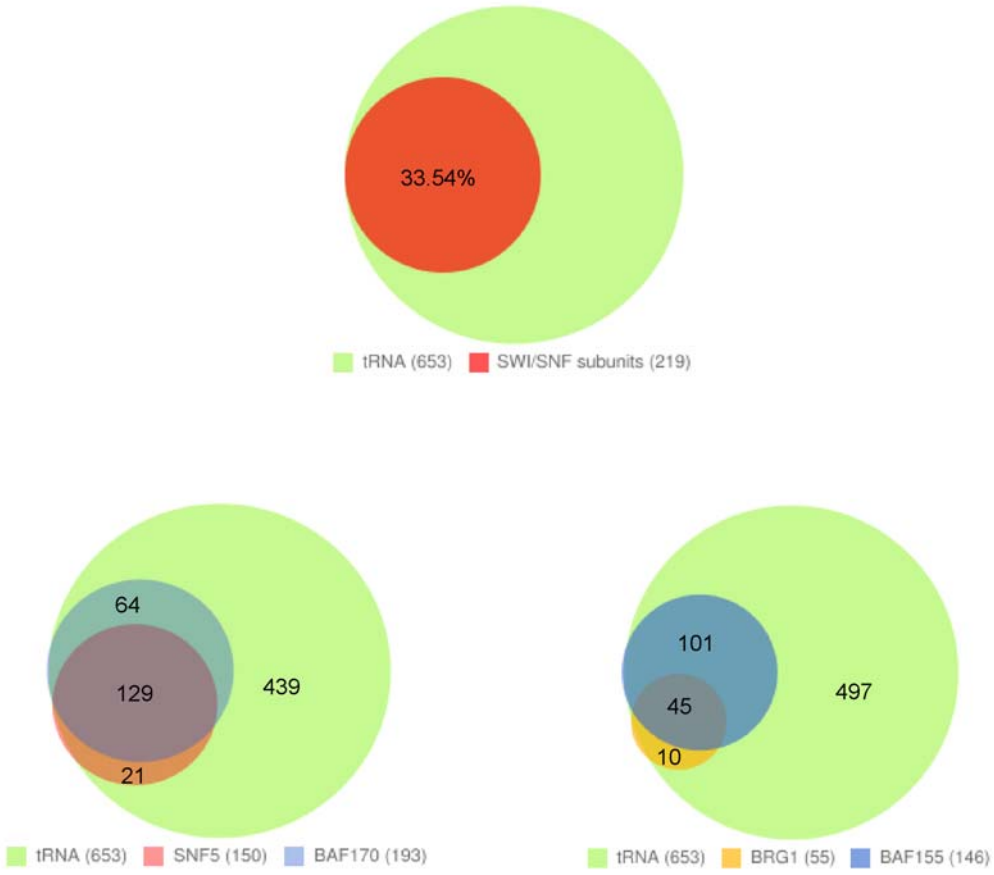


Figure 6.3: Considerable proportion of human tRNA loci are occupied by SWI/SNF subunits
Venn diagrams with ChIP sequencing data for SNF5, BRG1, BAF155 and BAF170 subunits at tRNA loci annotated in hg18 human genomic database (data mined from (Euskirchen et al., 2011) by Ann Hedley].

6.2.2 SWI/SNF directly represses transcription at tRNA and 5S rRNA genes

Since a substantial number of tRNA genes are enriched for the SWI/SNF complex, one can predict that this complex may play a role in their transcriptional regulation. SWI/SNF has been implicated in the etiology of many human cancers (see section 1.3.4). The core SWI/SNF subunit SNF5 is a *bona fide* tumour suppressor (Versteege et al., 1998), whereas other SWI/SNF subunits are also lost or silenced in multiple tumours (Reisman et al., 2009). Pol III transcription is also found to be elevated through various mechanisms in many human tumours. The regulation of Pol III transcriptional output is of high importance for cell growth and proliferation (see section 1.1.5). Therefore, studying the regulation of Pol III transcription by the SWI/SNF chromatin remodelling complex may be of therapeutic importance.

SNF5 is one of the core SWI/SNF subunits that is essential for the chromatin remodelling activity (Phelan et al., 1999). SNF5 is thought to contribute to SWI/SNF recruitment to its target loci, and once recruited it also makes contacts with the histone core of the nucleosome (Dechassa et al., 2008; Geng et al., 2001). Homozygous loss of SNF5 has been shown to cause preimplantation embryonic lethality in mice (Guidi et al., 2001). Thus, in order to investigate the role of SWI/SNF in Pol III transcriptional regulation, this essential subunit was knocked-down using RNAi in A31 mouse fibroblasts.

Western blots were performed on protein extracts from A31 cells 72 hours following transient transfection with scrambled siRNA (sc siRNA) and SNF5 siRNA (Figure 6.4a). A substantial reduction in SNF5 protein levels is observed. TFIIB and actin provide the loading controls. The SNF5 antibody recognises a doublet in the Western blot since SNF5 has two splice variants, which result from alternative splicing at the end of exon 2 (Bruder et al., 1999). The functional differences between the splice variants are not known, but both the splice variants are knocked-down with the siRNA. The RNA from these cells was subjected to RT-PCR analysis (Figure 6.4b). ARPP P0 mRNA provides the negative control, since microarray data from SNF5 null MEFs did not show alteration in ARPP P0 transcript levels (Isakoff et al., 2005). SNF5 mRNA levels show a considerable decrease following the knockdown.

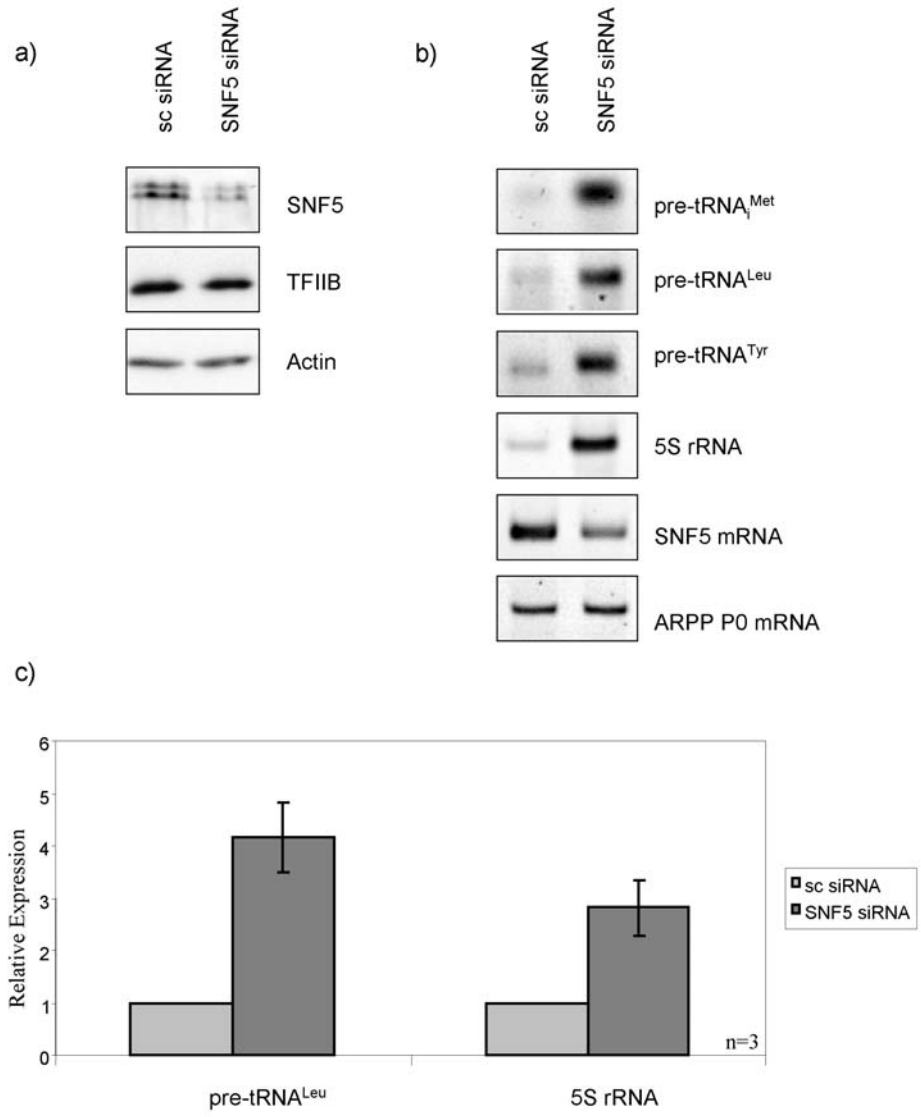


Figure 6.4: SNF5 **knockdown elevates tRNA and 5S rRNA expression**
a) Analysis of SNF5 protein levels by Western blot in A31 mouse fibroblasts following transient transfections with 10nM scrambled (sc) siRNA or SNF5 siRNA for 72hrs. TFIIB and actin serve as loading controls. b) Analysis by RT-PCR of pre-tRNA and 5S rRNA levels following knockdown of SNF5 protein. SNF5 mRNA demonstrates the knockdown and ARPP P0 mRNA serves as loading control. c) Analysis by qRT- PCR of pre-tRNA^{leu} and 5S rRNA expression levels. The expression was normalised to ARPP P0 mRNA levels.

Primers designed within the tRNA introns specifically amplify unprocessed pre-tRNA molecules. Since the half-life of rat liver tRNAs was estimated to be 5 days (Hanoune and Agarwal, 1970), specific amplification of pre-tRNA molecules provides a better estimate of the current rate of transcription. Interestingly, various pre-tRNA and 5S rRNA transcripts show a three to four fold elevation in expression following SNf5 knockdown. It must be noted that the primers against 5S rRNA detect the steady state levels. This elevation in expression was also confirmed using real-time RT-PCR (Figure 6.4c).

An Rsc4 mutant yeast strain shows a two-fold increase in tRNA expression compared to wild-type (Mahapatra et al., 2011). However, the authors were looking at steady-state tRNA levels and may have underestimated the increase. Nonetheless, just like the yeast RSC complex, the mammalian SWI/SNF complex seems to be repressing Pol III transcription. A knockdown of SNF5 in HeLa cells is associated with transcriptional upregulation of p21 causing cell cycle arrest and apoptosis (Kato et al., 2007). In contrast, I observed no evidence of cell cycle arrest or apoptosis upon SNF5 knockdown in A31 cells. A31 cells have been shown to illicit a normal p53 and p21 response to γ-irradiation (Nozaki et al., 1997). The data here would argue that the previous observations are cell type-specific.

However, the data obtained from RNAi-mediated knockdown of SNF5 must be treated with caution, since the biggest drawbacks of this technology are its off-target effects [reviewed in (Echeverri and Perrimon, 2006)]. In order to ascertain the specificity of the observed results, two alternative siRNA sequences targeting SNF5 were used alongside the siRNA previously used. Variable levels of SNF5 knockdown were achieved with the three different siRNAs. A dose-dependent increase in pre-tRNA^{Leu} expression is observed, such that the siRNA showing the highest level of knockdown shows the greatest elevation in tRNA expression (Figure 6.5). These data indicate that the increase in Pol III transcriptional output due to SNF5 knockdown is not due to off-target effects, but is target-specific.

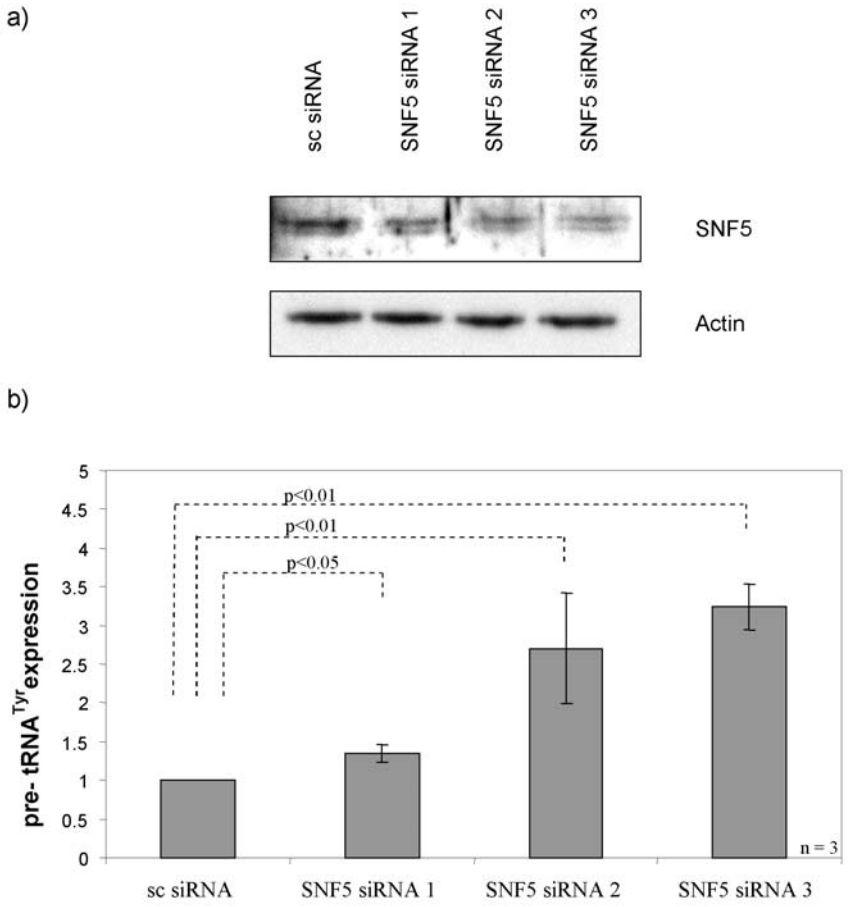


Figure 6.5: SNF5 **knockdowns with multiple siRNAs elevate tRNA expression**
a) Analysis by Western blot of SNF5 protein levels in A31 mouse fibroblasts following transient transfections with 10nM sc siRNA or three different SNF5 siRNAs for 72hrs. Actin provides the loading control b) Analysis by qRT-PCR of pre-tRNA^{Tyr} expression levels. Data were normalised to ARPP P0 mRNA levels.

From the ChIP data, it is clear that the SWI/SNF complex occupies tRNA loci. Therefore, one can assume that the effect of SNF5 knockdown on tRNA expression is direct. However, the interaction between SWI/SNF and Rb is well established. Loss of SNF5 directly or indirectly causes hyperphosphorylation of Rb, thus inactivating it (Oruetxebarria et al., 2004). Since Rb represses Pol III transcription, its inactivation could in turn cause elevation in tRNA expression (Larminie et al., 1997). Therefore, increased tRNA expression following SNF5 knockdown could be Rb-mediated and not direct.

The simian virus 40 oncoprotein large T antigen can bind and neutralise Rb protein. Cell lines transformed with the large T antigen show elevated tRNA expression compared to control cells (Felton-Edkins and White, 2002). Cl38 is one such SV40-transformed A31 mouse fibroblast cell line, which can be used to investigate the involvement of the Rb pathway in SWI/SNF-mediated repression of Pol III transcription.

SNF5 protein was knocked-down in Cl38 cells using two different siRNAs (Figure 6.6). Whole cell extracts from cells transfected with sc siRNA and SNF5 siRNAs were subjected to Western blot analysis. A subsantial decrease in SNF5 protein levels is observed 72hrs post transfection. α-Tubulin and actin provide the loading controls. Total RNA from the control and knockdown cells was then subjected to RT-PCR analysis. ARPP P0 provides the loading control. Similar to A31 fibroblasts, Cl38 cells show elevated expression of pre-tRNAs and 5S rRNA following SNF5 knockdown.

Despite the inactivation of Rb in Cl38 cells, loss of SNF5 still causes a substantial increase in the expression of Pol III products. These data suggest that the inhibition of Pol III transcription by SWI/SNF is independent of the Rb pathway.

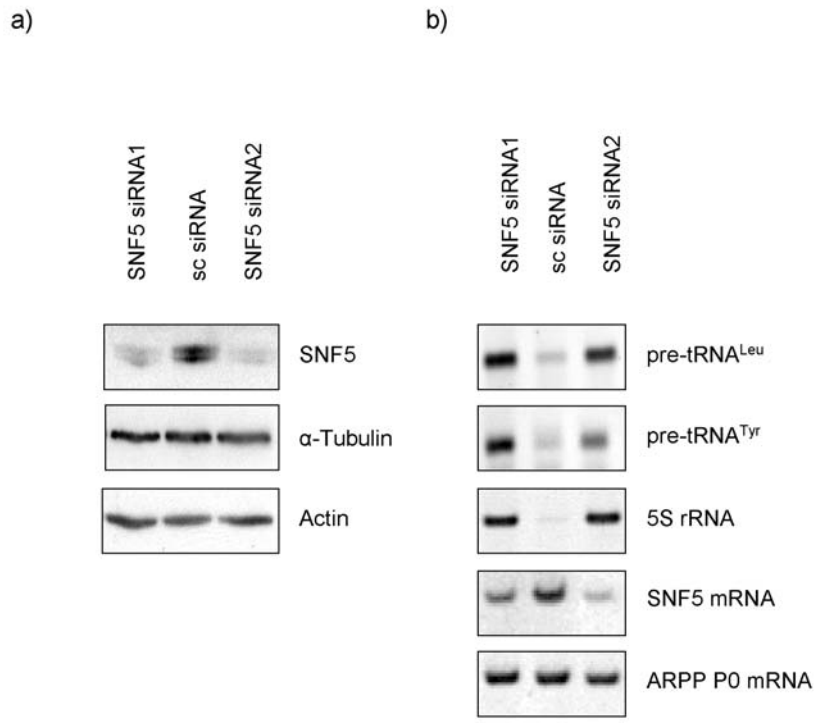


Figure 6.6: SNF5 **knockdown elevates tRNA and 5S rRNA expression in SV40-transformed cells**
a) Analysis by Western blot of SNF5 protein levels in SV40-transformed Cl38 mouse fibroblasts following transient transfections with 10nM sc siRNA or two different SNF5 siRNAs for 72hrs. α-Tubulin and actin provide the loading controls. b) Analysis by RT-PCR of pre-tRNA and 5S rRNA levels following SNF5 knockdown in Cl38 mouse fibroblasts. SNF5 mRNA demonstrates the knockdown and ARPP P0 mRNA provides the loading control.

SWI/SNF is considered a master regulator of gene transcription and shows functional interactions with many other transcriptional regulators, including c-MYC (see section 1.3.3). c-MYC may influence the expression of BRF1 by binding the E-boxes upstream of the BRF1 gene (unpublished data, Lynne Marshall). Therefore, the transcriptional regulation by SWI/SNF may be the result of altered expression of Pol III or its transcription factors.

In order to rule out this possibility, levels of various subunits of the Pol III transcriptional apparatus were analysed by Western blot and RT-PCR analysis following SNF5 knockdown in A31 cells (Figure 6.7). Actin and ARPP P0 provide the loading controls for the Western blot and RT-PCR analysis respectively. No reproducible differences are observed in the protein or mRNA levels of RPC155, BRF1 and TFIIIC110. TBP protein levels do not change either. Even though an exhaustive analysis of all the subunits was not performed, the data obtained so far from the ChIP and knockdown assays suggest a direct involvement of SWI/SNF in regulating transcription at tRNA promoters.

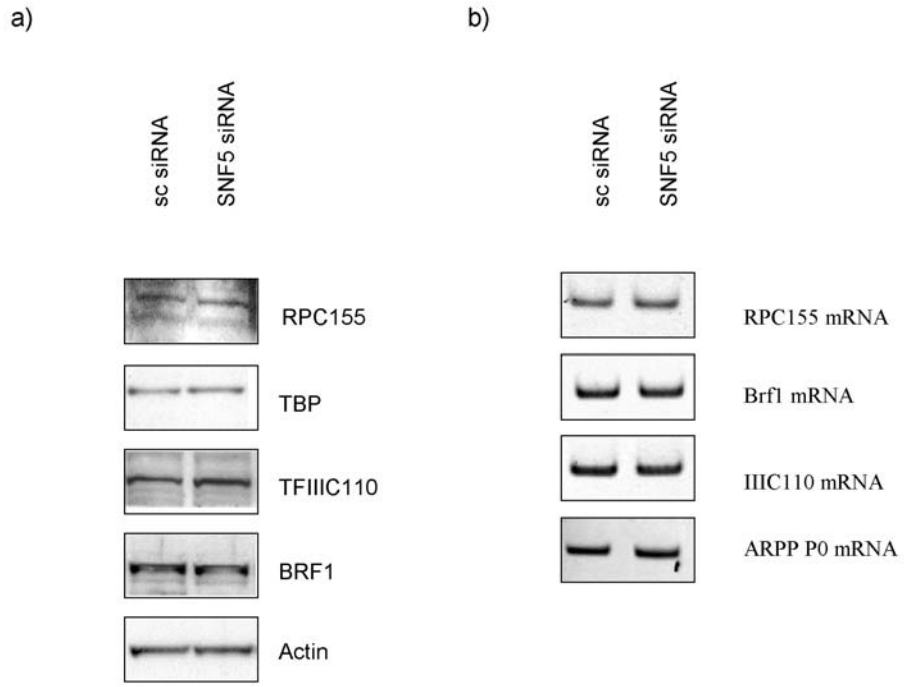


Figure 6.7: SNF5 **knockdown does not alter levels of Pol III transcription apparatus**
a) Analysis by Western blot of RPC155, TBP, TFIIIC110 and BRF1 protein levels in A31 mouse fibroblasts following transient transfections with 10nM sc siRNA or SNF5 siRNA for 72hrs. Actin provides the loading control. b) Analysis by RT-PCR of RPC155, BRF1 and TFIIIC110 mRNA levels following SNF5 knockdown. ARPP P0 mRNA provides the loading control.

6.2.3 BRG1 and BRM are both involved in repressing tRNA expression

An individual SWI/SNF complex can only contain one of the two ATPases, such that BRM/BAF complexes are structurally distinct from BRG1/BAF complexes (Wang et al., 1996a; Wang et al., 1996b). The ChIP data indicate that both BRM/BAF and BRG1/BAF complexes occupy tRNA genes (Figure 6.2). However, these two different complex types may have different effects on tRNA transcription. Moreover, our conclusion that SWI/SNF represses Pol III transcription has been drawn from the knockdown of a single subunit. In yeast, the inactivation of any single subunit leads to dissolution and subsequent inactivation of the whole complex (Hargreaves and Crabtree, 2011). Therefore, would the reduction in the ATPase subunits of the SWI/SNF complex lead to a similar increase in tRNA expression, as observed following SNF5 knockdown?

In order to test this, A31 cells were subjected to RNAi with two different siRNAs targeting each ATPase. A substanial decrease in BRG1 protein levels and a 50% decrease in BRG1 mRNA levels could be achieved with both the BRG1 siRNAs (Figure 6.8). However, BRG1 knockdown results in an elevation of BRM transcript levels. A substantial knockdown in BRM protein levels results in a similar increase in BRG1 protein and mRNA levels (Figure 6.9).

A homozygous loss of BRG1 leads to embryonic lethality in mice, whereas BRM null mice are viable (Bultman et al., 2000; Reyes et al., 1998). This may reflect the important role of BRG1 in the maintenance of the ES cell compartment (Ho et al., 2009a; Ho et al., 2009b). However, BRM null mice show an upregulation in BRG1 protein levels in a manner similar to what is observed in Figure 6.10 (Reyes et al., 1998). Put together, these data indicate that a differentiated cell may compensate for the loss of one ATPase by upregulation of the other, although the presence of BRG1 is essential for ES cells.

No elevation in tRNA expression is observed following BRG1 or BRM knockdowns. This suggests that the loss of any one ATPase is not sufficient to cause transcriptional upregulation at tRNA loci, and that the two ATPases can substitute for each other within BAF complexes associated with tRNA genes.

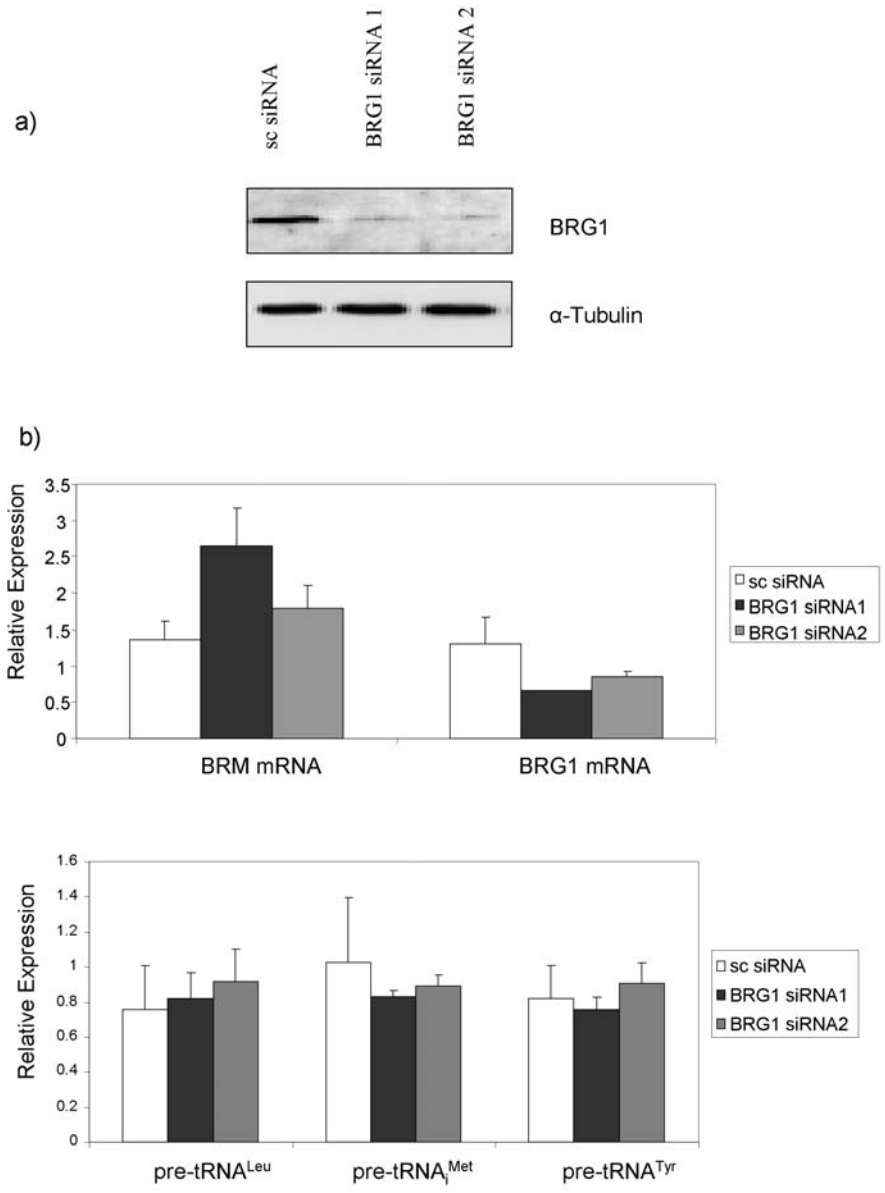


Figure 6.8: BRG1 **knockdown does not alter tRNA expression**
a) Analysis by Western blot of BRG1 protein levels in A31 mouse fibroblasts following transient transfections with 10nM sc siRNA or two different BRG1 siRNAs for 72hrs. α-Tubulin provides the loading control. b) Analysis by qRT-PCR of BRM mRNA and pre-tRNA levels following BRG1 knockdown. BRG1 mRNA demonstrates the knockdown. The data were normalised to ARPP P0 mRNA levels.

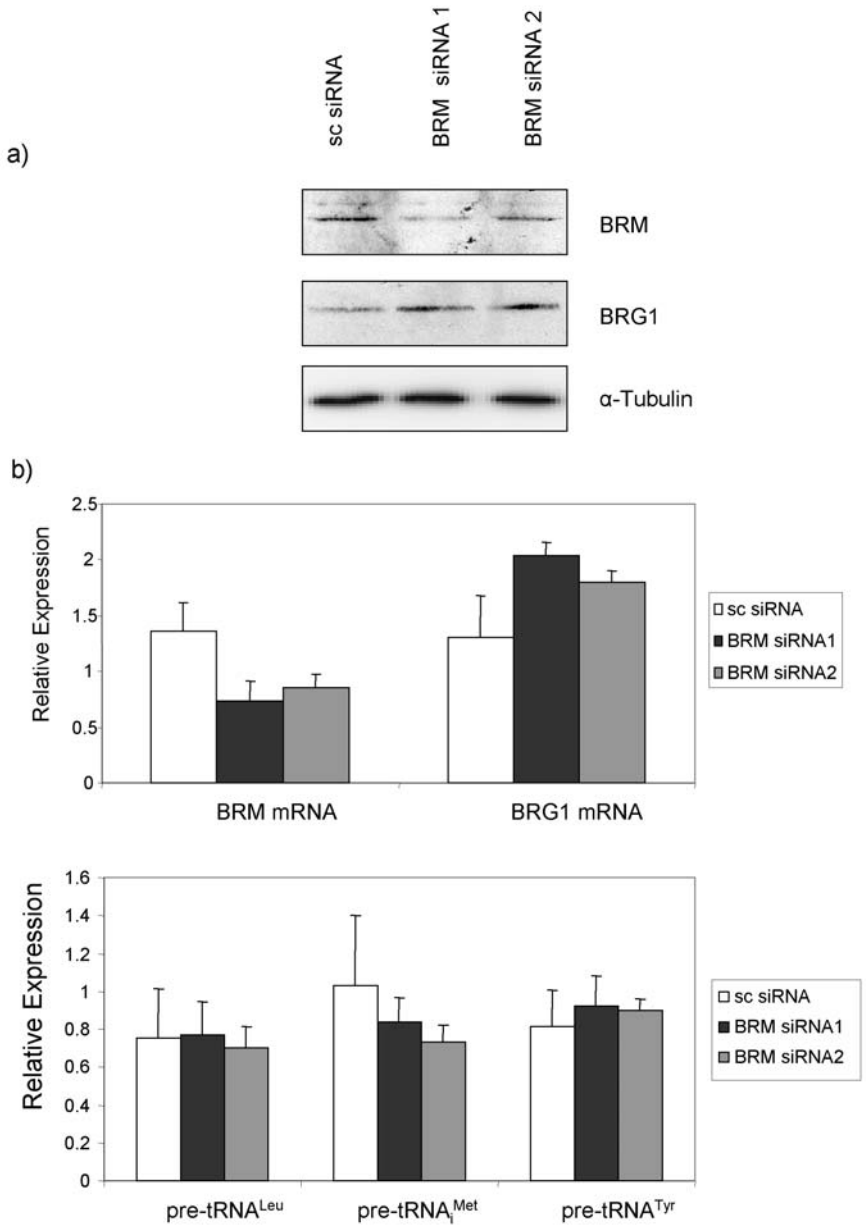


Figure 6.9: BRM **knockdown does not alter tRNA expression**
a) Analysis by Western blot of BRM or BRG1 protein levels in A31 mouse fibroblasts following transient transfections with 10nM sc siRNA and two different BRM siRNAs for 72hrs. α-Tubulin provides the loading control. b) Analysis by qRT-PCR of BRG1 mRNA and pre-tRNA levels following BRM knockdown. BRM mRNA demonstrates the knockdown. The data were normalised to ARPP P0 mRNA levels.

Since BRG1 and BRM compensate for each others activity at tRNA genes, a dual-knockdown is required to investigate the effects of the loss of SWI/SNF ATPase subunits. siRNAs targeting both ATPase subunits were used in tandem in order to achieve a dual knockdown. However, since single ATPase knockdown upregulates the other ATPase subunit, a dual knockdown could not be achieved easily.

In order to overcome this problem, a pool of two siRNAs against each ATPase was used. A31 cells were transfected with sc siRNA, BRG1 siRNA, BRM siRNA or a mixture of four different siRNAs, two against each ATPase (Figure 6.10). Even though BRG1 siRNA works quite efficiently on its own, the BRG1 siRNA pool is only able to reduce BRG1 mRNA and protein levels by 40% when BRM is knocked-down in tandem. The BRM knockdown on the other hand is more successful and 80% reduction in BRM protein levels can be observed in the dual-knockdown sample.

Despite the poor knockdown of BRG1 protein, a two-fold increase in pre-tRNA and 5S rRNA expression can be seen following the dual-knockdown. These results indicate that BRM/BAF and BRG1/BAF complexes at tRNA genes are functionally equivalent and consolidate the hypothesis that SWI/SNF represses Pol III transcription. Moreover, the ATPase subunits are required for SWI/SNF-mediated repression of Pol III transcription and, therefore, this repression may be achieved through ATPase-dependent chromatin remodelling.

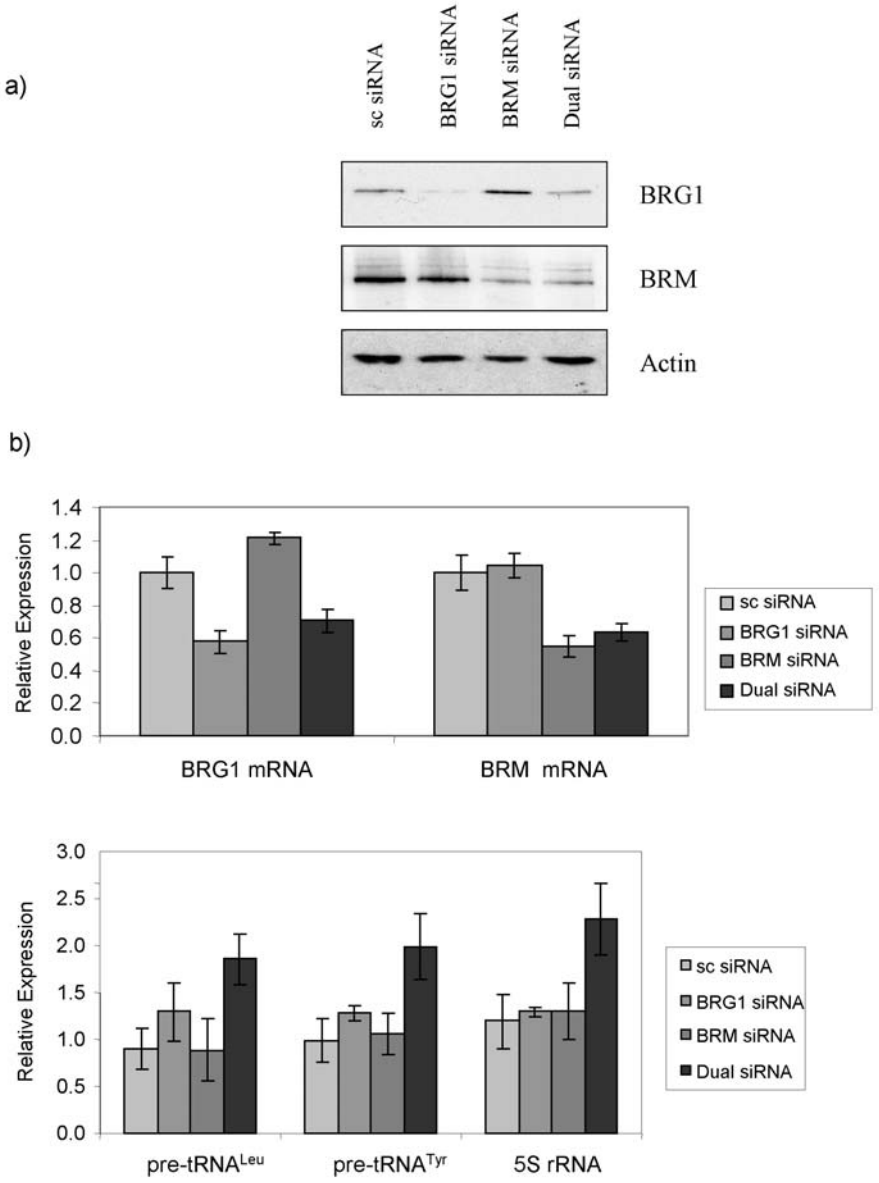


Figure 6.10: BRG1 **and** BRM **dual-knockdown elevates tRNA expression**
a) Analysis by Western blot of BRG1 and BRM protein levels in A31 mouse fibroblasts 72hrs following transient transfections with 10nM sc siRNA, BRG1 siRNA, BRM siRNA or 5nm BRG1 siRNA pool mixed with 5nM BRM siRNA pool. Actin provides the loading control. b) Analysis by qRT-PCR of BRG1 mRNA, BRM mRNA, pre-tRNA and 5S rRNA levels following BRG1, BRM and dual knockdowns. The data were normalised to ARPP P0 mRNA levels.

6.2.4 SWI/SNF subunits interact and co-localise with TFIIIC

A reduction in the cellular SNF5 protein level causes an increase in Pol III transcription. The loss of SNF5 may either lead to the dissociation of the whole complex or reduce the specific targeting of the SWI/SNF complex to tRNA and 5S rRNA loci. However, SNF5 may just be a bystander in regulating Pol III transcription, and other subunits of the SWI/SNF complex may be responsible for its targeting.

In order to investigate the specific role of SNF5 in the repression of SWI/SNF complex, a vector constitutively expressing HA-tagged SNF5 from a CMV promoter was transiently transfected into A31 cells (Figure 6.11). Western blots on protein extracts indicate a mere 25% increase in SNF5 protein levels in the pCDNA3HA.SNF5 transfected cells compared to the empty vector controls. HA-SNF5 migrates as a third (higher) band on the SNF5 blot due to the presence of the three N-terminal HA moieties. The HA blot confirms the identity of this third band as HA-SNF5. TFIIB serves as a loading control.

Despite this modest elevation in cellular SNF5 levels, RT-PCR expression analysis reveals a dramatic reduction in pre-tRNA and 5S rRNA expression. ARPP P0 serves as a loading control. A substantial increase in SNF5 mRNA is also observed, but the cells do not express such high amounts of SNF5 protein. This may be because the cells that express higher amounts of SNF5 do not survive the drastic reduction in Pol III products.

Reintroduction of SNF5 into malignant rhabdoid tumour (MRT) cells causes an upregulation of p16 causing cellular senescence (Oruetxebarria et al., 2004), whereas overexpression of SNF5 in A31 cells did not show changes in the p16 transcript levels. However, this experiment was performed over a relatively short time frame of 72hrs and maybe a longer time is required for SNF5 to trigger the senescence program. Moreover, the effects observed in the MRT cells may be specific to cells which lack SNF5 and may not apply to cells with endogenous SNF5.

The data obtained by this experiment show that a modest increase in SNF5 protein levels is sufficient to suppress Pol III transcriptional output.

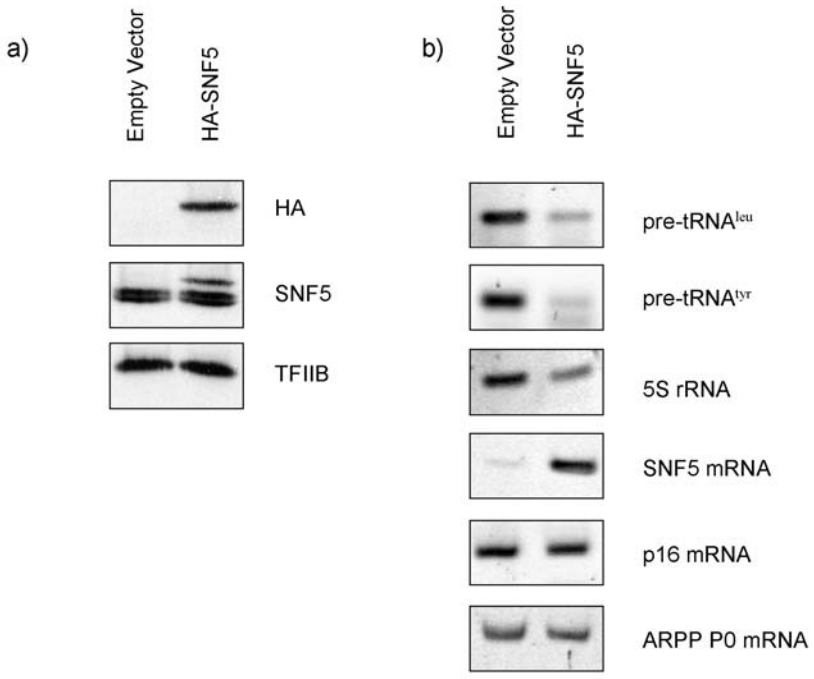


Figure 6.11: SNF5 **overexpression represses tRNA and 5S rRNA expression**
a) Analysis by Western blot of HA and SNF5 protein levels in A31 mouse fibroblasts following transient transfections with 5µg empty vector or pCDNA3HA.SNF5 for 72hrs. TFIIB provides the loading control. b) Analysis by RT-PCR of pre-tRNA and 5S rRNA levels following SNF5 overexpression. SNF5 mRNA demonstrates the successful overexpression. ARPP P0 mRNA provides the loading control.

The data obtained so far indicate that SWI/SNF complex specifically represses tRNA expression. But how this is achieved is still unclear. c-MYC has been shown to upregulate tRNA expression by the recruitment of Gcn5 and histone acetylation. Myc-mediated induction leads to a change in the chromatin environment of tRNA genes (Kenneth et al., 2007). Moreover, SWI/SNF has been linked to Myc-mediated transcriptional transactivation (Cheng et al., 1999). However, in the case of tRNA transcription, Myc and SWI/SNF seem to have contradictory effects.

In order to resolve the interplay between c-MYC and SWI/SNF at tRNA loci, c-MYC wild-type and c-MYC null rat fibroblasts were subjected to ChIP assays with antibodies against SWI/SNF subunits. $Myc^{-/-}$ cells show a dramatic drop in the protein levels of Myc, as seen by Western blot (Figure 6.12a). However, a faint band can be observed in the blot, but the reason for this is unknown. But, since the wild-type and null cells have drastic differences in the cellular c-MYC levels, the system can still be used to investigate the influence of c-MYC on SWI/SNF occupancy at tRNA genes.

Pol III ChIP was also performed in order to confirm the induction of transcription at the tRNA genes by c-MYC. TAF_I48 antibody provided the negative control for the ChIP and the data were normalised to serial input dilutions. c-MYC wild-type cells show 2-3 fold increase in Pol III enrichment at $tRNA^{Leu}$ and $tRNA^{Tyr}$ genes compared to null cells (Figure 6.12b). It must be noted that each tRNA primer set detects multiple genes coding for that tRNA isotype. This agrees with the published observation that c-MYC induces Pol III recruitment at tRNA genes (Kenneth et al., 2007).

The transcriptional induction by c-MYC is accompanied by an eviction of SWI/SNF subunits from tRNA genes. Considerable reduction in SNF5, BRG1 and BRM enrichment at $tRNA^{Leu}$ and $tRNA^{Tyr}$ genes is observed in *c-myc*$^{+/+}$ cells compared to *c-myc*$^{-/-}$ cells. These data indicate that the eviction of SWI/SNF complex by c-MYC may be necessary in order to overcome the transcriptional repression. A recent study has shown that Gcn5 can also acetylate Swi2/Snf2 subunit of the SWI/SNF complex (Kim et al., 2010). This acetylation was found to cause the dissociation of SWI/SNF complex from the chromatin and could be the mechanism by which c-MYC evicts the SWI/SNF complex from tRNA loci.

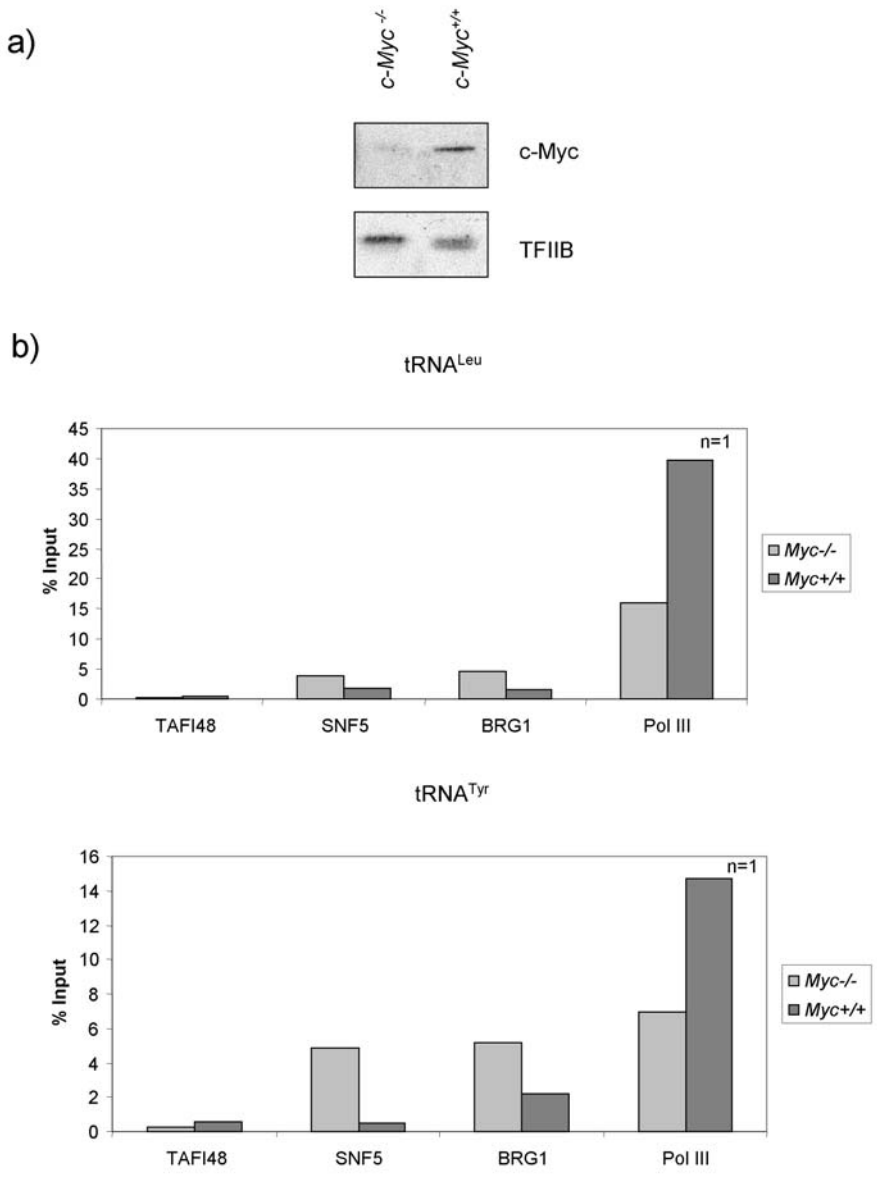
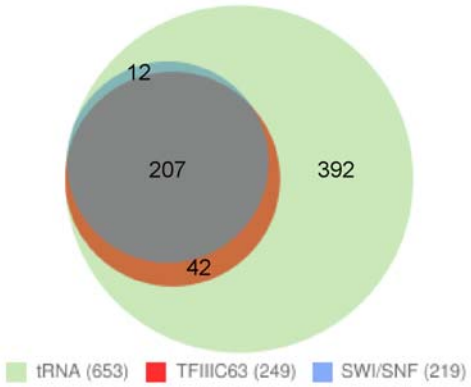


Figure 6.12: Myc evicts SWI/SNF subunits from Pol III transcribed loci
a) Analysis by Western blot of c-MYC protein levels in *c-myc*$^{+/+}$ and *c-myc*$^{-/-}$ rat fibroblasts. TFIIB provides the loading control. b) ChIP assay in *c-myc*$^{+/+}$ and *c-myc*$^{-/-}$ rat fibroblasts with antibodies against SNF5, BRG1 and BRM at tRNA loci. ChIP with RPC155 antibody provides the positive control. ChIP with TAF_I48 provided the negative control.

c-MYC was reported to cause elevated TFIIIB and Pol III occupancy at tRNA genes. However, no differences in the TFIIIC enrichment could be detected in response to Myc (Kenneth et al., 2007). TFIIIC, due to its associated HAT activity, is thought to be capable of penetrating repressive chromatin structures and binding its target sequences (Hsieh et al., 1999a; Kundu et al., 1999; Mertens and Roeder, 2008). Despite this, the binding of TFIIIC on its own may not be sufficient for optimal TFIIIB and Pol III recruitment, since Myc can readily open up the chromatin structure further to elevate TFIIIB and Pol III occupancy (Kenneth et al., 2007). This suggests that TFIIIC may occupy Pol III transcribed genes within repressed chromatin and prime them for transcriptional activation by proteins like Myc. This is reflected in the ChIP-seq analysis, where less than 50% of TFIIIC occupied sites are occupied by Pol III and BRF1 (Oler et al., 2010).

Since SWI/SNF represses transcription in the absence of Myc, and TFIIIC is present prior to induction by Myc, is there a correlation between their enrichment at tRNA genes? Analysis performed on the existing ChIP-seq data sets for TFIIIC (Oler et al., 2010) and SWI/SNF subunits (Euskirchen et al., 2011) reveal a striking correlation between their occupancy at tRNA genes (Figure 6.13). All but 12 tRNA loci enriched for at least one SWI/SNF subunit are also enriched for TFIIIC. Moreover, SNF5 occupies 145 of the 249 tRNA genes occupied by TFIIIC. The 12 tRNA genes which show SWI/SNF occupancy and no TFIIIC may be artefacts due to epitope masking during the immunoprecipitation. Thus, the majority of tRNA genes repressed by the SWI/SNF chromatin remodelling complex seem to be occupied by TFIIIC.



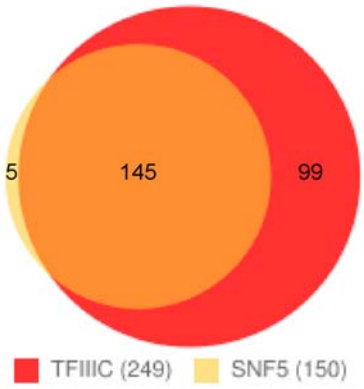


Figure 6.13: SWI/SNF occupancy on tRNA genes is highly correlative with TFIIIC occupancy
Venn diagrams with ChIP sequencing data for SWI/SNF subunits (SNF5, BRG1, BAF155 and BAF170) mined from (Euskirchen et al., 2011) and TFIIIC63 data mined from (Oler et al., 2010) at tRNA loci annotated in hg18 human genomic library. SWI/SNF (blue circle) indicates tRNAs occupied by one or more of the SWI/SNF subunits (data mined by Ann Hedley).

Since there is a significant correlation between TFIIIC and SWI/SNF occupancy at tRNA genes, it is fair to ask the question whether TFIIIC may be involved in SWI/SNF recruitment. In order to address this issue, co-immunoprecipitation (co-IP) assays were performed on protein extracts from A31 fibroblasts. In the forward reaction, IP was performed with antibodies against the SWI/SNF subunits and the Western blot was performed with antibody against TFIIIC220 subunit (Figure 6.14). The IP samples are resolved next to 20% input and rabbit IgG, where the latter provides the negative control IP. IPs against all three SWI/SNF subunits successfully co-IP TFIIIC220 and the signals obtained were consistently above background. The BRM antibody shows the best co-IP with TFIIIC220; however this may not mean that BRM is responsible for complex recruitment. The better quality of the BRM co-IP may just be due to better antibody affinity or its better suitability for the assay.

The reverse IP reaction was performed with an antibody against an alternative TFIIIC subunit, TFIIIC110. Since BRM provided the best co-IP for the forward reaction, the Western blot for the reverse reaction was performed against BRM. A clear co-IP of BRM with TFIIIC110 is observed in the reverse reaction. Thus, SWI/SNF and TFIIIC can be co-immunoprecipitated with antibodies against different subunits in both directions, indicating that these two multisubunit complexes interact with each other.

However, this interaction may be mediated by DNA, where both the complexes are recruited to the same DNA sequence, but do not physically interact. In order to rule out this possibility, co-IP was performed in the absence and presence of DNaseI (Figure 6.14). DNaseI would digest any common DNA that may be holding the two complexes together. TFIIIC220 antibody was used for the co-IP, where rabbit IgGs and an alternative BRM antibody provide the negative and positive control IPs respectively. The addition of DNaseI does not abolish the co-IP of BRM with TFIIIC220, however it does cause a small reduction in the co-IP signal. This suggests that TFIIIC and SWI/SNF complexes primarily interact with each other independently of DNA; however the presence of DNA may promote the interaction slightly. The data together indicate that TFIIIC and SWI/SNF may co-occupy tRNA genes, where the former may recruit the latter through protein-protein interactions. Thus, TFIIIC may not only be priming tRNA genes for induction, but also keeping them repressed by recruiting the SWI/SNF complex.

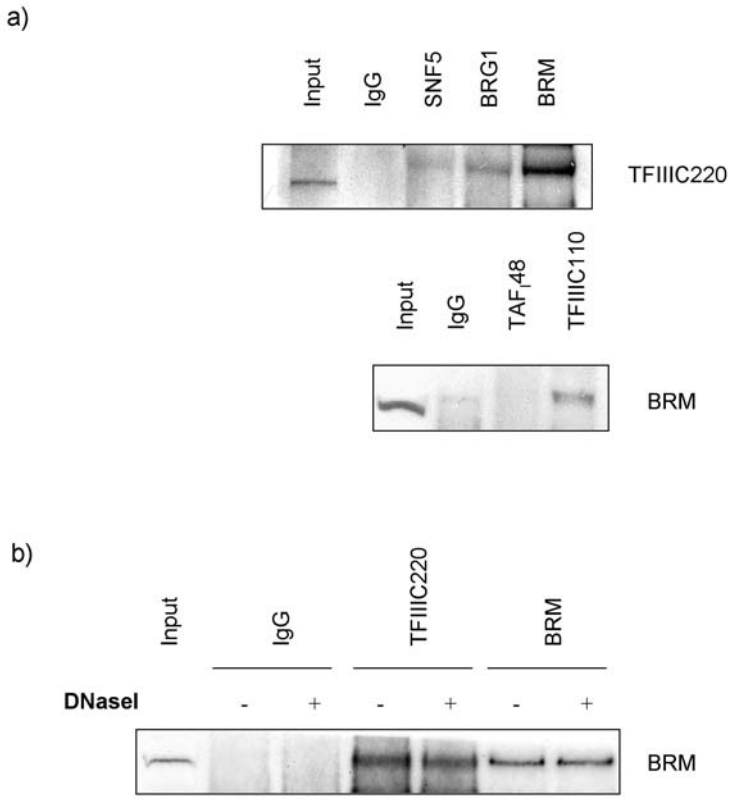


Figure 6.14: SWI/SNF subunits co-immunoprecipitate with TFIIIC
a) Analysis by Western blot for co-immunoprecipitation of SNF5, BRG1 and BRM with TFIIIC220 in A31 total cell extracts. The reverse co-immunoprecipitation of TFIIIC110 with BRM is shown in the bottom panel. IgG and TAF_I48 serve as negative controls. b) Analysis of Western blot for co-immunoprecipitation of TFIIIC220 with BRM in A31 total cell extract following overnight treatment with DNase. IgG and a second BRM antibody provide the negative and positive controls respectively.

6.2.5 SWI/SNF co-localises with Pol III and limits tRNA transcription

The data obtained so far indicate that the SWI/SNF complex is recruited by TFIIIC and inhibits tRNA gene transcription. The polymerase loading on tRNA genes is generally considered to be a reliable surrogate measure for their transcriptional output. Since SNF5 knockdown leads to a drastic upregulation of tRNA expression, it may also be expected to cause increased Pol III loading at tRNA genes. Moreover, since SWI/SNF is a chromatin remodelling complex, an alteration in core histone occupancy at tRNA genes may also be expected upon reduction in SNF5.

In order to study the effects of SNF5 knockdown upon Pol III transcription factor and histone occupancies at tRNA genes, A31 cells were subjected to ChIP assays following transfection with scrambled and SNF5 siRNAs. The knockdown of SNF5 was verified by Western blot for each ChIP repeat, an example of which is shown in Figure 6.15a. ChIP for SWI/SNF subunits reveals a drop in SNF5, BRG1 and BRM occupancy at $tRNA^{Leu}$ and $tRNA^{Tyr}$ genes following SNF5 knockdown. This indicates that a reduction in SNF5 leads to the loss of the entire SWI/SNF complex from tRNA genes. This may either be due to dissolution of the entire complex or just its dissociation from TFIIIC. Variability in the levels of knockdown obtained by transient transfections led to relatively large error bars. This may have resulted in the higher than expected BRG1 and BRM enrichment in the absence of SNF5 on the $tRNA^{Leu}$ genes.

In order to circumvent this problem, stable cell lines expressing multiple short-hairpins against SNF5 were constructed. However, despite showing 50% reduction in SNF5 mRNA level, the cells did not show any reduction in SNF5 protein levels (data not shown). Therefore, these cell lines could not be used for further experiments. G401 cells containing an inducible SNF5 expression system in a SNF5 null background were also obtained (Oruetxebarria et al., 2004). However, these cells failed to show any Pol III transcriptional response to SNF5 induction (data not shown). Since these cells are obtained from an MRT cancer cell line, it is possible that the lack of Pol III response to SNF5 induction is due to some unknown genetic alterations. Therefore, this cell line could not be used for further experiments either.

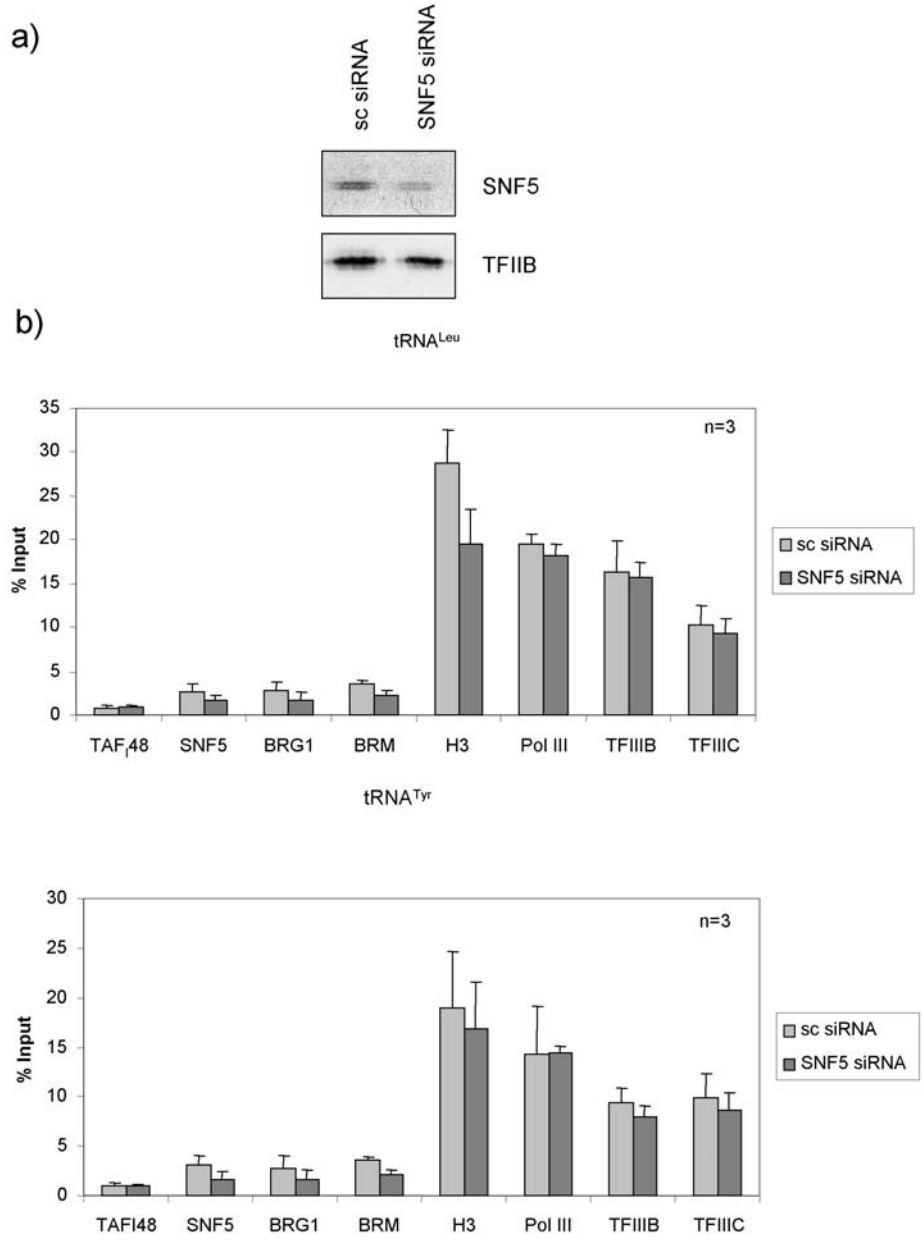


Figure 6.15: SNF5 knockdown does not significantly alter histone levels or transcription apparatus occupancy at tRNA genes

a) Analysis by Western blot of SNF5 protein levels in A31 mouse fibroblasts following transient transfections with 10nM sc siRNA and SNF5 siRNA for 72hrs. TFIIB provides the loading control. b) ChIP assay in A31 mouse fibroblasts following SNF5 knockdown with antibodies against histone H3, RPC155, BRF1 and TFIIIC220 at tRNA loci. ChIPs with SNF5, BRG1 and BRM antibodies provide the controls for the knockdown. ChIP with TAF_I48 antibody provided the negative control.

Thus, the best system available was the transient transfection system. $tRNA^{Leu}$ genes show considerable reduction in histone H3 occupancy following SNF5 knockdown, whereas this reduction is less apparent at $tRNA^{Tyr}$ loci. However, the reduction in histone occupancy is consistently seen over multiple experiments Further experimentation is required to consolidate these data.

Surprisingly, despite the dramatic elevation in $tRNA^{Leu}$ and $tRNA^{Tyr}$ expression following SNF5 knockdown, no increase in the loading of Pol III transcription factors at the respective genes is observed. This was an unexpected result since Pol III occupancy is generally concordant with tRNA expression, for example in response to c-MYC (Figure 6.12) and (Kenneth et al., 2007). The same experiment was performed with lower antibody concentration, reduced crosslinking and increased sonication in order to improve sensitivity, but the results remained unchanged. This could be a short-coming of the transient transfection system, as previously mentioned; however, this result prompted further analysis of the available ChIP-seq datasets (Figure 6.16).

Surprisingly, there is a high degree of correlation between the ChIP-seq datasets for SWI/SNF subunits and Pol III. A vast majority of the tRNA genes occupied by SWI/SNF subunits are also utilised by Pol III. Thus, SWI/SNF seems to localise to genes that are likely to be transcribed. These data, contradict the repressive role of SWI/SNF complex revealed by SNF5 and BRG1+BRM knockdown, as well as the SNF5 overexpression experiments.

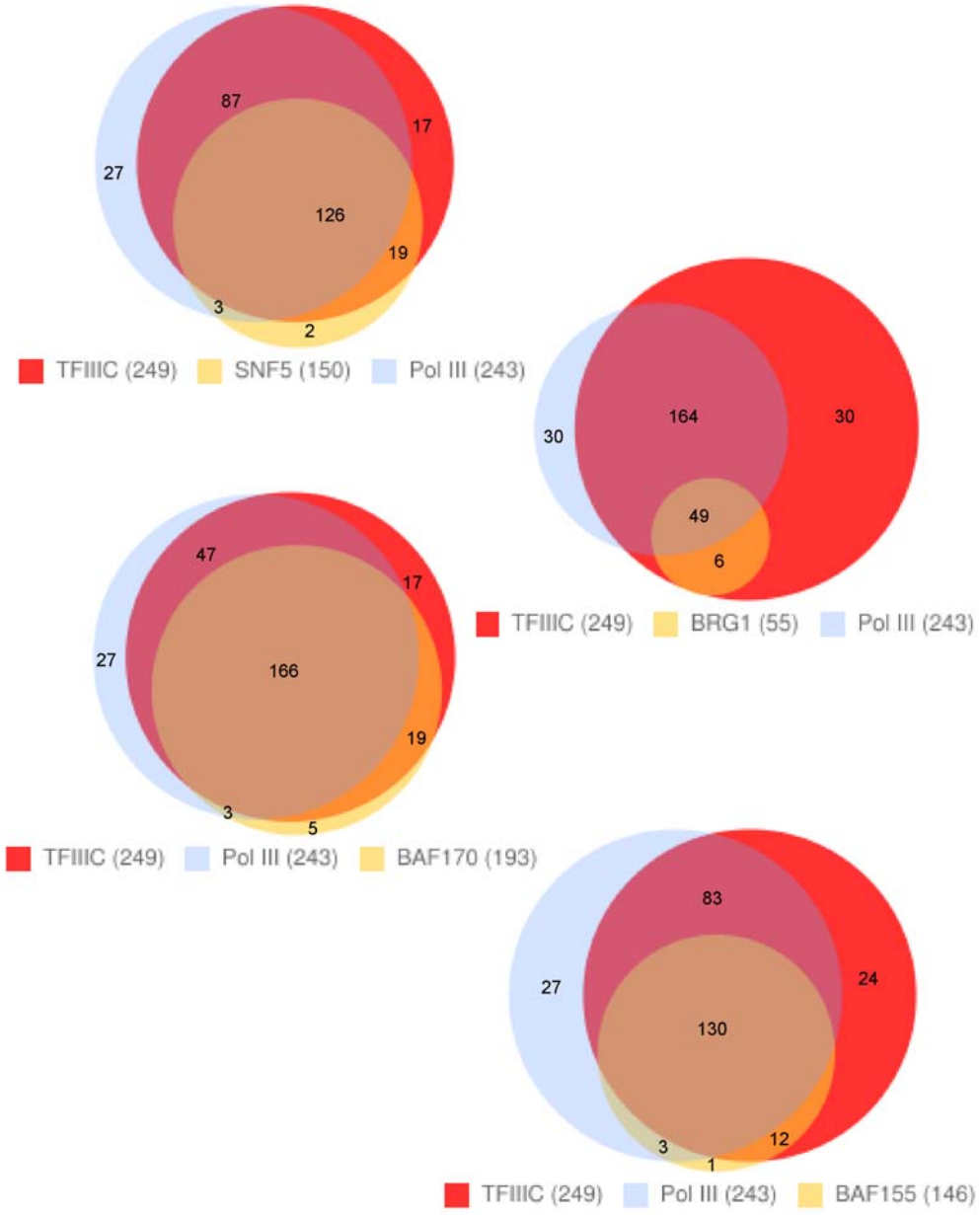


Figure 6.16: SWI/SNF occupancy at tRNA genes correlates well with polymerase loading
Venn diagrams with ChIP sequencing data for SWI/SNF subunits (SNF5, BRG1, BAF155 and BAF170) mined from (Euskirchen et al., 2011), TFIIIC63 and RPC32 data mined from (Oler et al., 2010) at tRNA loci annotated in hg18 human genomic library (data mined by Ann Hedley).

In yeast, Rsc reduces the production of spurious transcripts from the *SUP4* tRNA gene by defining the chromatin structure around the tRNA locus and providing a defined area for transcription. A nucleosome positioned downstream of the tRNA gene is also thought to throttle transcription, possibly by repressing efficient termination and facilitated recycling of the polymerase (Mahapatra et al., 2011). A similar situation may also exist in mammalian cells.

In order to test this, locus specific primers were designed against a sequence 200bp downstream of an actively transcribed $tRNA^{Leu}$ gene which is occupied by SWI/SNF. Since the resolution of a ChIP assay is higher than 200bp, these primers will detect differences in the polymerase enrichment toward the end of the tRNA gene. Thus the ChIP signal obtained by PCR with these primers is likely to represent the polymerase molecules that are successful in transcriptional elongation. A significant increase in Pol III enrichment is detected by these primers following SNF5 knockdown, suggesting an elevation in transcriptional elongation (Figure 6.17).

The data put together suggest that, even though SWI/SNF does not inhibit the overall Pol III loading at tRNA genes, it spatially limits the area bound by Pol III. The elevation in Pol III downstream of the tRNA locus also suggests that, SWI/SNF may repress tRNA transcription by inhibiting elongation or efficient termination by Pol III. Thus, SWI/SNF may be toning down transcription of active tRNA genes to maintain it at homeostatic levels.

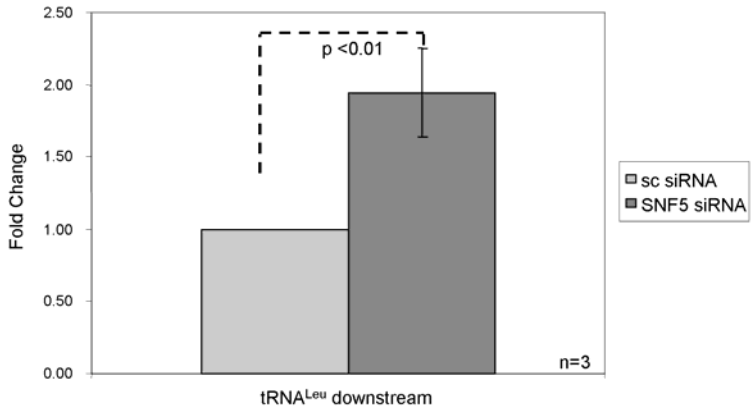


Figure 6.17: SNF5 knockdown causes increased Pol III occupancy towards the end of a tRNA locus

ChIP assay in A31 mouse fibroblasts following SNF5 knockdown with antibody against RPC155 at $tRNA^{Leu}$ 200bp downstream sequence. TAF_I48 provided the negative control (not shown). The plotted values were calculated using the % Input values for Pol III ChIP.

6.3 Discussion

The SWI/SNF chromatin remodelling complex is found to be deregulated in many human cancers. SNF5 and many other SWI/SNF subunits are considered to be tumour suppressors and are lost in multiple tumour types. Despite this, the molecular mechanisms which cause cellular transformation in the absence of SWI/SNF are still not clear (Reisman et al., 2009). SWI/SNF regulates the cell cycle through its physical and functional interactions with proteins like Rb, p53 and c-MYC, and these interactions may be crucial for its tumour suppressor activity (Cheng et al., 1999; Isakoff et al., 2005; Kato et al., 2007).

Elevated Pol III transcription is also a characteristic of cancers. In addition to this, a modest increase in Pol III transcriptional output, by itself, can be tumorigenic (Marshall et al., 2008). Thus, the discovery that SWI/SNF inhibits Pol III transcription may be of considerable therapeutic significance. The loss of SWI/SNF in human cancers could lead to elevated Pol III transcriptional output, which in turn can promote cellular transformation. Thus, regulation of Pol III transcription may provide another dimension to SWI/SNF's tumour suppressor potential. However, there are no published data studying the levels of Pol III transcription in tumours lacking SWI/SNF subunits. This is required in order to fathom the full implications of the findings in this chapter. Moreover, further experiments are required to ascertain whether elevated Pol III transcription is an integral part of the tumorigenic response in the absence of SWI/SNF.

SWI/SNF is known to be required for Rb and p53-mediated tumour suppression and both these proteins also inhibit Pol III transcription (Felton-Edkins et al., 2003b; Kato et al., 2007; Versteege et al., 2002). However, the effect of SWI/SNF on Pol III transcription seems independent of these two pathways, since cells transformed with the SV40 large T antigen still show elevated tRNA expression upon SNF5 knockdown. Thus, cells seem to employ multiple redundant pathways in order to keep the Pol III transcriptional output in check, SWI/SNF being one of them.

SWI/SNF was found to be required for activation by c-MYC, since dominant-negative mutants of SNF5 and BRG1 inhibit c-MYC-mediated transactivation (Cheng et al., 1999). However, c-MYC and SWI/SNF seem to have opposing

activities at tRNA genes, where transcriptional induction by c-MYC leads to eviction of SWI/SNF. c-MYC has been shown to recruit Gcn5 via its interaction with TRRAP at Pol III-transcribed genes. It has been recently shown that Gcn5 acetylates Snf2 at two lysine residues, which upon acetylation interact with the Snf2 bromodomain and cause its eviction from the nucleosome (Kim et al., 2010). This may be the mechanism by which Myc evicts the SWI/SNF complex from tRNA genes. Thus, SWI/SNF may play a role in counteracting the oncogenic effects of c-MYC, at least as far as Pol III transcription is concerned.

Mice with single copy loss of different SWI/SNF subunits become tumour-prone. $SNF5^{+/-}$ mice acquire lymphomas and rhabdoid tumours with extremely high penetrance within 10 weeks, whereas $BRG1^{+/-}$ mice develop mammary tumours within a year (Roberts et al., 2002). But $BRG1^{+/-}$ tumours are not similar in morphology to the $SNF5^{+/-}$ tumours. Firstly, they are carcinomas and not sarcomas. Secondly, they arise due to haploinsufficiency of BRG1, rather than the loss of heterozygosity as seen in $SNF5^{+/-}$ tumours. In addition to this, $BRG1^{+/-}$ $BRM^{-/-}$ double mutant mice do not develop tumours similar to $Snf^{+/-}$ mice either. This suggests that the observed phenotypic differences are not simply due to the fact that BRG1 has a closely related paralog, whereas, SNF5 does not (Bultman et al., 2008). BRM null mice do not develop tumours spontaneously, but BRM loss does sensitise them to carcinogens (Glaros et al., 2007). Therefore mice lacking different SWI/SNF subunits show differing phenotypes, and accordingly, it has been suggested that there are functional differences between the tumour suppression properties of these proteins.

However, as far as Pol III transcription is concerned, BRG1 and BRM can compensate for the loss of each other and the BRG1/BRM dual-knockdown produces the same effect as SNF5 knockdown. Therefore, both the BRG1/BAF and BRM/BAF complexes seem to behave in a similar manner. Moreover, the reduction in BRM leads to the upregulation of BRG1, and vice versa. This phenomenon was previously observed in MEFs obtained from $BRM^{-/-}$ mice; however, the upregulation of BRM mRNA following BRG1 reduction has not been previously reported (Glaros et al., 2007). This compensatory feedback loop depicts the importance of these ATPases for cellular homeostasis; however the mechanisms of this feedback loop still remain unclear.

The recruitment of the SWI/SNF complex to tRNA genes seems to occur via its interaction with TFIIIC. Not only is there a high degree of correlation between their enrichment at tRNA genes, but SWI/SNF subunits also co-IP with more than one TFIIIC subunit. Attempts to co-IP BRF1 and SWI/SNF were unsuccessful. Moreover, c-MYC does not alter the TFIIIC occupancy at tRNA genes, but it evicts SWI/SNF and causes the recruitment of TFIIIB and Pol III. This inverse correlation consolidates the hypothesis that SWI/SNF is recruited by TFIIIC, and not by TFIIIB or Pol III. It is still unclear which SWI/SNF subunit is responsible for its interaction with TFIIIC or if the presence of the whole complex is required. There is a considerable degree of overlap between the ChIP-seq enrichment for all the four subunits tested at tRNA genes, suggesting that the presence of the whole complex may be required for recruitment. Pull-down assays with SNF5 and BRG1 dominant negative mutants defective for complex formation may provide further insights into the role of particular subunits in SWI/SNF at tRNA genes (de La Serna et al., 2000; Geng et al., 2001).

In yeast, the loss of RSC was initially found to cause an increase in nucleosome density at tRNA genes (Parnell et al., 2008). This suggests that RSC may be an activator of Pol III transcription. However, recent experiments have shown that a knockdown of RSC causes elevated tRNA expression. RSC deposits nucleosomes containing H2A.Z flanking the yeast *SUP4* tRNA gene. The nucleosome positioned downstream of the transcribed region represses Pol III transcription, possibly by inhibiting facilitated polymerase recycling. This was proposed to help tone down transcription levels to maintain a balanced output (Mahapatra et al., 2011). The ChIP-seq data from HeLa cells indicate that SWI/SNF may occupy actively transcribed tRNA loci. In addition to this, no elevation in Pol III occupancy at tRNA genes is observed following SNF5 knockdown. Therefore, SWI/SNF may not behave like conventional inhibitors that repress the recruitment of the polymerase, but instead may somehow modulate transcription of active tRNA genes. Loss of SNF5 leads to elevated Pol III occupancy at the end of a tRNA gene, indicating that there may be increased transcriptional elongation. As observed previously, when a tRNA gene is located at the edge of a nucleosome, it transcription is dramatically reduced (Wittig and Wittig, 1982). One can envisage that SWI/SNF may achieve transcriptional repression via a nucleosome

positioned towards the end of tRNA genes, as is observed for the yeast *SUP4* gene.

Therefore, SWI/SNF inhibits Pol III transcription, however, detailed analysis of nucleosome positioning, along with promoter and terminator accessibility assays, are required in order to ascertain its precise mechanisms. The role of Pol III transcription in formation of tumours following loss of SWI/SNF subunits also needs to be examined further.

Chapter 7 Final Discussion

SINE transcription can put a tremendous load on cellular metabolism by engaging a large fraction of the available Pol III transcriptional apparatus. Moreover, an abundance of SINE transcripts may impede essential cellular process like protein translation and may also be toxic to the cell (Hasler and Strub, 2006; Kaneko et al., 2011). Thus, it is important for the cell to silence SINE transcription and this has been thought to be achieved via DNA methylation (Liu et al., 1994; Liu and Schmid, 1993; Schmid, 1991). Contrary to expectations and common belief, this study finds the presence of Pol III transcriptional machinery at SINE loci enriched for DNA methylation and MBPs. Methylation of A- and B- block promoter sequences is found not to impede Pol III recruitment to SINEs.

From the data presented, it is clear that silencing at SINEs is achieved by H3K9 trimethylation and SUV39H1. However, it is still unclear how this is targeted to SINEs, especially in the absence of DNA methylation. HP1 is thought to be recruited to its target sequences via the RNAi machinery (Pal-Bhadra et al., 2004) and an RNA component is required for HP1 localisation (Kwon and Workman, 2010). Alu transcripts have recently been shown to be targets of RNA processing by DICER1 (Kaneko et al., 2011). I have preliminary data to suggest that knockdown of DICER1 leads to elevated Pol III occupancy at Alu loci. Thus, it can be hypothesised that SINE RNA, once processed by the RNAi machinery, is responsible for targeting HP1 to SINE sequences. HP1, once recruited at SINEs, establishes H3K9 methylation and repressive chromatin via recruitment of SUV39H1. DNMTs are then recruited to fill in the last piece of the puzzle and lock down SINE transcription (Figure 7.1). The basal level of Pol III found at SINEs may be required to produce sufficient transcripts in order to maintain the hypothesised RNA-directed transcriptional repression.

The RITS (RNA-induced transcriptional gene silencing) complex is thought to be vital for the establishment of RNA-directed transcriptional repression. RITS comprises of Argonaute (Ago1) protein, a chromodomain protein Chp1 and a novel protein Tas3. The slicer activity of Ago1 is thought to be crucial for siRNA processing, whereas Chp1 contributes to the H3K9me-binding ability of RITS (Grewal and Jia, 2007). RITS also recruits the RNA-directed RNA polymerase complex (RDRC), which is responsible for the production of dsRNA molecules that can be processed by DICER (Sugiyama et al., 2005). The involvement of RITS and RDRC in SINE transcriptional silencing is still unexplored. However, the

secondary structure of SINE RNA contains dsRNA regions that can be processed by DICER. Thus, RDRC may not be required for RNA-directed SINE transcriptional repression.

Younger AluY elements, which are closer to the Alu consensus sequence, show higher levels of DNA methylation compared to their older AluJ and AluS counterparts. This could be due to such RNA-directed active targeting of DNA methylation via HP1, since the younger elements have higher sequence similarity to the consensus Alu sequence. However, the mechanism for the recruitment of DNMTs is still unclear from the data available and the methylation status of SINEs in SUV39H1h2 knockout cells needs to be analysed.

The data obtained in chapters 3 and 4 suggest the following model for SINE transcriptional repression. In the presence of DNA methylation, transcription is repressed by HP1, H3K9 methylation, MBPs and HDACs. Transcription is sensitive to TSA since HDACs contribute to both MBP and SUV39H1-dependent repression. In the absence of DNA methylation, a loss of MBP and HDACs is observed, however HP1 and H3K9 methylation can still keep transcription repressed. Here the repression is insensitive to TSA treatment due to the reduced HDAC occupancy and a switch away from HDAC-dependent repression. Moreover, when control cells are treated with chaetocin, HDAC-dependent silencing only allows for a modest increase in transcription. But when DNA methylation and HDACs are lost in *Dnmt1*$^{-/-}$ cells, SINEs become more sensitive to chaetocin treatment.

Further work is required to fully establish this as a model for SINE transcriptional repression. It may be necessary to detect micro- and pi-RNA derivatives of SINE RNA in cellular extracts and their association with RITS complex. The effects of DICER1-knockdown on SINEs also need to be studied thoroughly, including changes in HP1, SUV39H1 and H3K9me3 enrichments. Antisense oligonucleotides targeting SINE RNA may also be able to alleviate RNA-directed transcriptional repression and cause increased Pol III occupancy at SINEs.

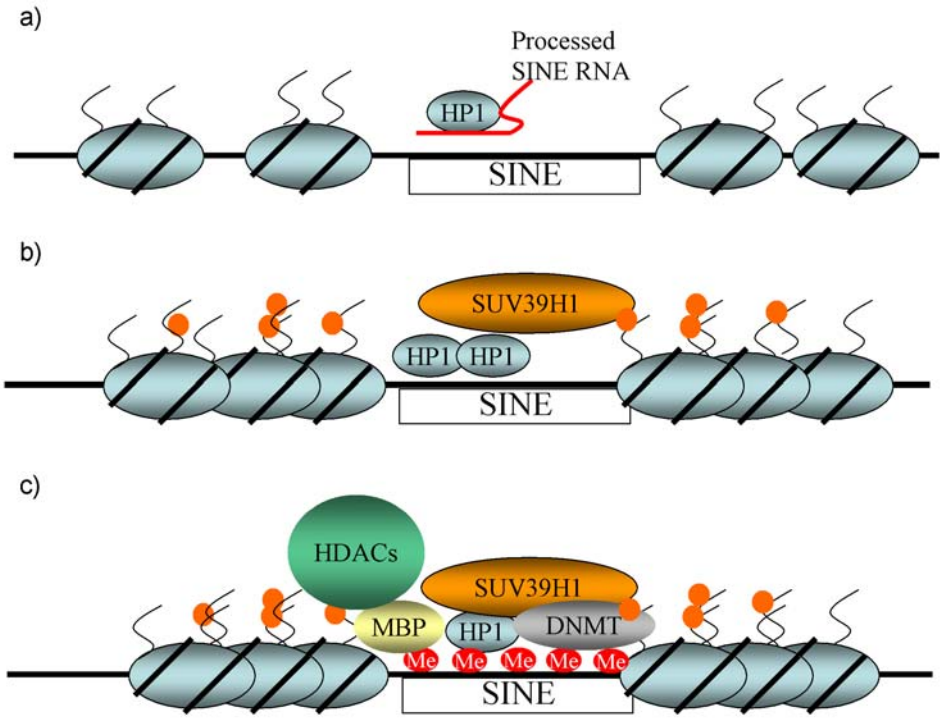


Figure 7.1: Model for SINE transcriptional repression.
a) Hypothetical HP1 targeting by processed SINE RNA b) HP1 recruits SUV39H1 and H3K9me3. c) DNMTs, MBPs and HDACs are then recruited.

HP1α isoform associates with heterochromatic regions and promotes gene silencing, whereas HP1γ associates with euchromatin and is thought to be involved in gene activation. In *Drosophila,* HP1c associates with heat-shock induced genes, where it colocalises with elongating Pol II and histone chaperone FACT (Kwon and Workman, 2010). SINEs are heat-shock induced genes, and therefore HP1 isoform switching may be involved in SINE transcriptional regulation. The presence of limiting levels of Pol III at suppressed SINE loci indicates that these sites are poised for transcriptional activation. However, the involvement of HP1 isoforms in the rapid transcriptional activation of SINEs following heat-shock still needs to be investigated.

Repetitive elements provide ample opportunity for erroneous recombination events and these, if left unchecked, can threaten genomic stability. Recombination between SINEs can be particularly deleterious, since these elements populate the gene-rich regions of the genome (Konkel and Batzer, 2011). The Alu consensus sequence contains 24 CpG sites and Alus account for up to one-third of all CpG sites within the human genome. In most normal tissues, these Alu CpGs are found to be methylated (Rubin et al., 1994; Schmid, 1991). From the data obtained during this study, it seems that the purpose of DNA methylation at SINEs is to ensure genomic stability by suppressing erroneous recombination between these sequences. Moreover, the data also imply that SINEs may contribute to the hypomethylation-associated genomic instability observed in many human diseases (Ehrlich, 2005).

LINE transcription and retrotransposition has previously been shown to be suppressed either directly by DNA methylation or by MeCP2 overexpression (Yu et al., 2001). Since SINE retrotransposition depends on LINE proteins, by suppressing LINE transcription, DNA methylation may be able to suppress SINE transposition and further ensure genomic stability (Schmid, 2003).

The mechanism by which DNA methylation suppresses recombination is not yet understood. Preliminary data suggest that DNA methylation may somehow reduce the amount of ssDNA generated in the cell following DNA damage. However, these data were obtained from 5-azacytidine treated HeLa cells, where damage may be being specifically targeted to methylated sequences, for example Alus. This experiment must be repeated in the *Dnmt1*$^{-/-}$ and wild-type

MEFs in order to consolidate the findings. Further work is also required to fully understand how DNA methylation can affect the availability of ssDNA and the kinetics of repair. Moreover, would NHEJ-mediated repair be affected by DNA methylation as is observed for SSA-mediated repair at SINEs?

Recent publication of ChIP-seq analysis at Pol III-transcribed genes has revealed that tRNA genes may be subject to the same chromatin-mediated transcriptional control as is observed for Pol II-transcribed genes (White, 2011). However, it is still unclear whether tRNA loci have the permissive chromatin due to their proximity to Pol II-transcribed genes and enhancer elements, or whether the Pol III transcription apparatus establishes the permissive chromatin environment (Oler et al., 2010).

However, the presence of SWI/SNF at tRNA loci was found not to correlate with the presence of Pol II (Euskirchen et al., 2011). In contrast, SWI/SNF and Pol III transcription apparatus occupancy at tRNA genes is highly correlative, suggesting that the former may be recruited by the latter. In addition to this, multiple SWI/SNF subunits co-IP with more than one TFIIIC subunit, suggesting there is direct interaction with the transcription complex.

SWI/SNF is found to suppress Pol III transcription, where both BRG1/BAF and BRM/BAF complexes have similar activities. Despite this, no elevation in the enrichment for Pol III transcriptional apparatus is observed on tRNA loci following SNF5 knockdown. However, SNF5 knockdown does lead to an elevation in Pol III occupancy downstream of a tRNA locus. In addition to this, transcriptional activation by c-MYC leads to eviction of the SWI/SNF complex from the tRNA loci. These two observations consolidate the hypothesis that SWI/SNF directly represses transcription at tRNA loci, but it is still unclear how this is achieved. It may be achieved by limiting the space available to the transcriptional apparatus, possibly through positioned nucleosomes. Since there is more polymerase present downstream in the absence of SNF5, the presence of SWI/SNF may impede transcript elongation, efficient termination or facilitated recycling. Changes in DNA accessibility following SNF5 knockdown must be analysed by MNase/DNaseI digestion. Fine mapping the changes in the positioning of nucleosomes and Pol III transcriptional machinery following the

knockdown of SWI/SNF subunits will be required to resolve the mechanistic details of this repression.

43% of SWI/SNF-enriched loci are found to correspond to enhancer regions and 17% correspond to CTCF sites (Euskirchen et al., 2011). TFIIIC binding sites, especially ETC loci, are also found to localise near CTCF binding sites (Carriere et al., 2011; Moqtaderi et al., 2010). In addition to this, Pol III-occupied loci, when not near Pol II promoters, also have enhancer-like chromatin (Oler et al., 2010). The correlation between TFIIIC, CTCF and SWI/SNF suggests a functional link between these complexes.

Recent ChIP-seq analysis in 6 different species has revealed that, despite differences at individual tRNA genes, Pol III loading is highly conserved among amino acid isotypes, i.e. the total Pol III loading at the members of a particular tRNA isotype is highly conserved between species (Kutter et al., 2011). This observation suggests that higher order 3D organisation between tRNA isotype genes may be involved in the regulation of Pol III loading at individual tRNA loci. Moreover, this also indicates that there is a requirement to modulate individual tRNA gene expression in order to regulate overall cellular availability of each tRNA isotype. SWI/SNF may be involved in such 3D organisation and fine-tuning.

The Pol III loading at tRNA isotypes was also found to correlate with the codon usage by the cellular transcriptome, i.e. the Pol III loading at a less-used isotype was lower and vice-versa (Kutter et al., 2011). H3K9me3 is found to negatively correlate with Pol III loading at tRNA genes (Barski et al., 2010). Thus, RNA-directed transcriptional silencing may also be involved in the overall regulation of Pol III loading at tRNAs. One can envisage the scenario where, tRNA isotypes not used by the translational apparatus accumulate in the cell and feed back into the RNAi machinery and repress expression from their parent genes through the HP1-H3K9me3 pathway.

Thus, the regulation of Pol III transcription is more complex than previously thought and a considerable amount of work is still required to fully understand it.

References

Aguilera, C., Nakagawa, K., Sancho, R., Chakraborty, A., Hendrich, B., and Behrens, A. (2011). c-Jun N-terminal phosphorylation antagonises recruitment of the Mbd3/NuRD repressor complex. Nature *469*, 231-235.

Albanesi, T., Polani, S., Cozzi, R., and Perticone, P. (1999). DNA strand methylation and sister chromatid exchanges in mammalian cells in vitro. Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis *429*, 239-248.

Allen, T. A., Von Kaenel, S., Goodrich, J. A., and Kugel, J. F. (2004). The SINE-encoded mouse B2 RNA represses mRNA transcription in response to heat shock. Nat Struct Mol Biol *11*, 816-821.

Allfrey, V. G., Faulkner, R., and Mirsky, A. E. (1964). Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis. Proc Natl Acad Sci U S A *51*, 786-794.

Athanasiadis, A., Rich, A., and Maas, S. (2004). Widespread A-to-I RNA editing of Alu-containing mRNAs in the human transcriptome. PLoS Biol *2*, e391.

Aziz, M. A., Schupp, J. E., and Kinsella, T. J. (2009). Modulation of the activity of methyl binding domain protein 4 (MBD4/MED1) while processing iododeoxyuridine generated DNA mispairs. Cancer Biol Ther *8*, 1156-1163.

Bachman, K. E., Park, B. H., Rhee, I., Rajagopalan, H., Herman, J. G., Baylin, S. B., Kinzler, K. W., and Vogelstein, B. (2003). Histone modifications and silencing prior to DNA methylation of a tumor suppressor gene. Cancer Cell *3*, 89-95.

Baker, R. E., Camier, S., Sentenac, A., and Hall, B. D. (1987). Gene size differentially affects the binding of yeast transcription factor tau to two intragenic regions. Proc Natl Acad Sci U S A *84*, 8768-8772.

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.

Ballestar, E., Paz, M. F., Valle, L., Wei, S., Fraga, M. F., Espada, J., Cigudosa, J. C., Huang, T. H., and Esteller, M. (2003). Methyl-CpG binding proteins identify novel sites of epigenetic inactivation in human cancer. Embo J *22*, 6335-6345.

Bannister, A. J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell Res *21*, 381-395.

Bannister, A. J., Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature *410*, 120-124.

Barak, M., Levanon, E. Y., Eisenberg, E., Paz, N., Rechavi, G., Church, G. M., and Mehr, R. (2009). Evidence for large diversity in the human transcriptome created by Alu RNA editing. Nucleic Acids Res *37*, 6905-6915.

Bardeleben, C., Kassavetis, G. A., and Geiduschek, E. P. (1994). Encounters of Saccharomyces cerevisiae RNA polymerase III with its transcription factors during RNA chain elongation. J Mol Biol *235*, 1193-1205.

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.

Bartholomew, B., Meares, C. F., and Dahmus, M. E. (1990). Photoaffinity labeling of RNA polymerase III transcription complexes by nascent RNA. J Biol Chem *265*, 3731-3737.

Bartkiewicz, M., Gold, H., and Altman, S. (1989). Identification and characterization of an RNA molecule that copurifies with RNase P activity from HeLa cells. Genes Dev *3*, 488-499.

Bass, B. L. (2002). RNA editing by adenosine deaminases that act on RNA. Annu Rev Biochem *71*, 817-846.

Batey, R. T., Rambo, R. P., Lucast, L., Rha, B., and Doudna, J. A. (2000). Crystal structure of the ribonucleoprotein core of the signal recognition particle. Science *287*, 1232-1239.

Batzer, M. A., and Deininger, P. L. (2002). Alu repeats and human genomic diversity. Nat Rev Genet *3*, 370-379.

Batzer, M. A., Deininger, P. L., Hellmann-Blumberg, U., Jurka, J., Labuda, D., Rubin, C. M., Schmid, C. W., Zietkiewicz, E., and Zuckerkandl, E. (1996). Standardized nomenclature for Alu repeats. J Mol Evol *42*, 3-6.

Baumann, M., Pontiller, J., and Ernst, W. (2010). Structure and basal transcription complex of RNA polymerase II core promoters in the mammalian genome: an overview. Mol Biotechnol *45*, 241-247.

Bedford, M. T., and van Helden, P. D. (1987). Hypomethylation of DNA in pathological conditions of the human prostate. Cancer Res *47*, 5274-5276.

Ben-Hattar, J., Beard, P., and Jiricny, J. (1989). Cytosine methylation in CTF and Sp1 recognition sites of an HSV tk promoter: effects on transcription in vivo and on factor binding in vitro. Nucleic Acids Res *17*, 10179-10190.

Berger, S. L., and Folk, W. R. (1985). Differential activation of RNA polymerase III-transcribed genes by the polyomavirus enhancer and the adenovirus E1A gene products. Nucleic Acids Res *13*, 1413-1428.

Bernstein, B. E., Meissner, A., and Lander, E. S. (2007). The mammalian epigenome. Cell *128*, 669-681.

Bestor, T., Laudano, A., Mattaliano, R., and Ingram, V. (1988). Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. J Mol Biol *203*, 971-983.

Bestor, T. H., and Ingram, V. M. (1983). Two DNA methyltransferases from murine erythroleukemia cells: purification, sequence specificity, and mode of interaction with DNA. Proc Natl Acad Sci U S A *80*, 5559-5563.

Bhattacharya, S. K., Ramchandani, S., Cervoni, N., and Szyf, M. (1999). A mammalian protein with specific demethylase activity for mCpG DNA. Nature *397*, 579-583.

Bieker, J. J., Martin, P. L., and Roeder, R. G. (1985). Formation of a rate-limiting intermediate in 5S RNA gene transcription. Cell *40*, 119-127.

Birch, J. L., Tan, B. C., Panov, K. I., Panova, T. B., Andersen, J. S., Owen-Hughes, T. A., Russell, J., Lee, S. C., and Zomerdijk, J. C. (2009). FACT

facilitates chromatin transcription by RNA polymerases I and III. Embo J *28*, 854-865.

Bird, A., Taggart, M., Frommer, M., Miller, O. J., and Macleod, D. (1985). A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. Cell *40*, 91-99.

Bird, A. P., and Wolffe, A. P. (1999). Methylation-induced repression--belts, braces, and chromatin. Cell *99*, 451-454.

Boeke, J., Ammerpohl, O., Kegel, S., Moehren, U., and Renkawitz, R. (2000). The minimal repression domain of MBD2b overlaps with the methyl-CpG-binding domain and binds directly to Sin3A. J Biol Chem *275*, 34963-34967.

Boeke, J. D. (1997). LINEs and Alus--the polyA connection. Nat Genet *16*, 6-7.

Boland, M. J., and Christman, J. K. (2008). Characterization of Dnmt3b:thymine-DNA glycosylase interaction and stimulation of thymine glycosylase-mediated repair by DNA methyltransferase(s) and RNA. J Mol Biol *379*, 492-504.

Borchert, G. M., Lanier, W., and Davidson, B. L. (2006). RNA polymerase III transcribes human microRNAs. Nat Struct Mol Biol *13*, 1097-1101.

Bourc'his, D., and Bestor, T. H. (2004). Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. Nature *431*, 96-99.

Bourc'his, D., Xu, G. L., Lin, C. S., Bollman, B., and Bestor, T. H. (2001). Dnmt3L and the establishment of maternal genomic imprints. Science *294*, 2536-2539.

Bouvet, P., Dimitrov, S., and Wolffe, A. P. (1994). Specific regulation of Xenopus chromosomal 5S rRNA gene transcription in vivo by histone H1. Genes Dev *8*, 1147-1159.

Braun, B. R., Kassavetis, G. A., and Geiduschek, E. P. (1992). Bending of the Saccharomyces cerevisiae 5S rRNA gene in transcription factor complexes. J Biol Chem *267*, 22562-22569.

Breant, B., Huet, J., Sentenac, A., and Fromageot, P. (1983). Analysis of yeast RNA polymerases with subunit-specific antibodies. J Biol Chem *258*, 11968-11973.

Britten, R. J. (1997). Mobile elements inserted in the distant past have taken on important functions. Gene *205*, 177-182.

Britten, R. J., and Kohne, D. E. (1970). Repeated segments of DNA. Sci Am *222*, 24-31.

Brock, G. J., Charlton, J., and Bird, A. (1999). Densely methylated sequences that are preferentially localized at telomere-proximal regions of human chromosomes. Gene *240*, 269-277.

Bruder, C. E., Dumanski, J. P., and Kedra, D. (1999). The mouse ortholog of the human SMARCB1 gene encodes two splice forms. Biochem Biophys Res Commun *257*, 886-890.

Buhler, J. M., Huet, J., Davies, K. E., Sentenac, A., and Fromageot, P. (1980). Immunological studies of yeast nuclear RNA polymerases at the subunit level. J Biol Chem *255*, 9949-9954.

Bultman, S., Gebuhr, T., Yee, D., La Mantia, C., Nicholson, J., Gilliam, A., Randazzo, F., Metzger, D., Chambon, P., Crabtree, G., and Magnuson, T. (2000).

A Brg1 null mutation in the mouse reveals functional differences among mammalian SWI/SNF complexes. Mol Cell *6*, 1287-1295.

Bultman, S. J., Herschkowitz, J. I., Godfrey, V., Gebuhr, T. C., Yaniv, M., Perou, C. M., and Magnuson, T. (2008). Characterization of mammary tumors from Brg1 heterozygous mice. Oncogene *27*, 460-468.

Cairns, B. R., Lorch, Y., Li, Y., Zhang, M., Lacomis, L., Erdjument-Bromage, H., Tempst, P., Du, J., Laurent, B., and Kornberg, R. D. (1996). RSC, an essential, abundant chromatin-remodeling complex. Cell *87*, 1249-1260.

Canella, D., Praz, V., Reina, J. H., Cousin, P., and Hernandez, N. (2010). Defining the RNA polymerase III transcriptome: Genome-wide localization of the RNA polymerase III transcription machinery in human cells. Genome Res *20*, 710-721.

Cao, X., Yeo, G., Muotri, A. R., Kuwabara, T., and Gage, F. H. (2006). Noncoding RNAs in the mammalian central nervous system. Annu Rev Neurosci *29*, 77-103.

Carlson, L. L., Page, A. W., and Bestor, T. H. (1992). Properties and localization of DNA methyltransferase in preimplantation mouse embryos: implications for genomic imprinting. Genes Dev *6*, 2536-2541.

Carlson, M., Osmond, B. C., Neigeborn, L., and Botstein, D. (1984). A suppressor of SNF1 mutations causes constitutive high-level invertase synthesis in yeast. Genetics *107*, 19-32.

Carriere, L., Graziani, S., Alibert, O., Ghavi-Helm, Y., Boussouar, F., Humbertclaude, H., Jounier, S., Aude, J. C., Keime, C., Murvai, J., Foglio, M., Gut, M., Gut, I., Lathrop, M., Soutourina, J., Gerard, M., and Werner, M. (2011). Genomic binding of Pol III transcription machinery and relationship with TFIIS transcription factor distribution in mouse embryonic stem cells. Nucleic Acids Res.

Chan, P. P., and Lowe, T. M. (2009). GtRNAdb: a database of transfer RNA genes detected in genomic sequence. Nucleic Acids Res *37*, D93-97.

Chang, D. Y., Hsu, K., and Maraia, R. J. (1996). Monomeric scAlu and nascent dimeric Alu RNAs induced by adenovirus are assembled into SRP9/14-containing RNPs in HeLa cells. Nucleic Acids Res *24*, 4165-4170.

Chen, L. L., and Carmichael, G. G. (2008). Gene regulation by SINES and inosines: biological consequences of A-to-I editing of Alu element inverted repeats. Cell Cycle *7*, 3294-3301.

Chen, L. L., DeCerbo, J. N., and Carmichael, G. G. (2008). Alu element-mediated gene silencing. Embo J *27*, 1694-1705.

Cheng, S. W., Davies, K. P., Yung, E., Beltran, R. J., Yu, J., and Kalpana, G. V. (1999). c-MYC interacts with INI1/hSNF5 and requires the SWI/SNF complex for transactivation function. Nat Genet *22*, 102-105.

Cheng, X., and Blumenthal, R. M. (2010). Coordinated chromatin control: structural and functional linkage of DNA and histone methylation. Biochemistry *49*, 2999-3008.

Cherrier, T., Suzanne, S., Redel, L., Calao, M., Marban, C., Samah, B., Mukerjee, R., Schwartz, C., Gras, G., Sawaya, B. E., Zeichner, S. L., Aunis, D., Van Lint, C., and Rohr, O. (2009). p21(WAF1) gene promoter is epigenetically silenced by CTIP2 and SUV39H1. Oncogene *28*, 3380-3389.

Christman, J. K. (2002). 5-Azacytidine and 5-aza-2'-deoxycytidine as inhibitors of DNA methylation: mechanistic studies and their implications for cancer therapy. Oncogene *21*, 5483-5495.

Chu, W. M., Ballard, R., Carpick, B. W., Williams, B. R., and Schmid, C. W. (1998). Potential Alu function: regulation of the activity of double-stranded RNA-activated kinase PKR. Mol Cell Biol *18*, 58-68.

Chu, W. M., Liu, W. M., and Schmid, C. W. (1995). RNA polymerase III promoter and terminator elements affect Alu RNA expression. Nucleic Acids Res *23*, 1750-1757.

Chuang, L. S., Ian, H. I., Koh, T. W., Ng, H. H., Xu, G., and Li, B. F. (1997). Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. Science *277*, 1996-2000.

Ciliberto, G., Raugei, G., Costanzo, F., Dente, L., and Cortese, R. (1983). Common and interchangeable elements in the promoters of genes transcribed by RNA polymerase iii. Cell *32*, 725-733.

Clouaire, T., de Las Heras, J. I., Merusi, C., and Stancheva, I. (2010). Recruitment of MBD1 to target genes requires sequence-specific interaction of the MBD domain with methylated DNA. Nucleic Acids Res *38*, 4620-4634.

Clouaire, T., and Stancheva, I. (2008). Methyl-CpG binding proteins: specialized transcriptional repressors or structural components of chromatin? Cell Mol Life Sci *65*, 1509-1522.

Cooper, D. N., and Krawczak, M. (1989). Cytosine methylation and the fate of CpG dinucleotides in vertebrate genomes. Hum Genet *83*, 181-188.

Cordaux, R., and Batzer, M. A. (2009). The impact of retrotransposons on human genome evolution. Nat Rev Genet *10*, 691-703.

Cormack, B. P., and Struhl, K. (1992). The TATA-binding protein is required for transcription by all three nuclear RNA polymerases in yeast cells. Cell *69*, 685-696.

Cortazar, D., Kunz, C., Selfridge, J., Lettieri, T., Saito, Y., MacDougall, E., Wirz, A., Schuermann, D., Jacobs, A. L., Siegrist, F., Steinacher, R., Jiricny, J., Bird, A., and Schar, P. (2011). Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability. Nature *470*, 419-423.

Cortellino, S., Xu, J., Sannai, M., Moore, R., Caretti, E., Cigliano, A., Le Coz, M., Devarajan, K., Wessels, A., Soprano, D., Abramowitz, L. K., Bartolomei, M. S., Rambow, F., Bassi, M. R., Bruno, T., Fanciulli, M., Renner, C., Klein-Szanto, A. J., Matsumoto, Y., Kobi, D., Davidson, I., Alberti, C., Larue, L., and Bellacosa, A. (2011). Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. Cell *146*, 67-79.

Cost, G. J., Feng, Q., Jacquier, A., and Boeke, J. D. (2002). Human L1 element target-primed reverse transcription in vitro. Embo J *21*, 5899-5910.

Cote, J., Quinn, J., Workman, J. L., and Peterson, C. L. (1994). Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex. Science *265*, 53-60.

Coveney, J., and Woodland, H. R. (1982). The DNase I sensitivity of Xenopus laevis genes transcribed by RNA polymerase III. Nature *298*, 578-580.

Cress, W. D., and Seto, E. (2000). Histone deacetylases, transcriptional control, and cancer. J Cell Physiol *184*, 1-16.

Crick, F. (1970). Central dogma of molecular biology. Nature *227*, 561-563.

Crick, F. H. (1968). The origin of the genetic code. J Mol Biol *38*, 367-379.

Crighton, D., Woiwode, A., Zhang, C., Mandavia, N., Morton, J. P., Warnock, L. J., Milner, J., White, R. J., and Johnson, D. L. (2003). p53 represses RNA polymerase III transcription by targeting TBP and inhibiting promoter occupancy by TFIIIB. Embo J *22*, 2810-2820.

Cross, S. H., Meehan, R. R., Nan, X., and Bird, A. (1997). A component of the transcriptional repressor MeCP1 shares a motif with DNA methyltransferase and HRX proteins. Nat Genet *16*, 256-259.

Cuozzo, C., Porcellini, A., Angrisano, T., Morano, A., Lee, B., Di Pardo, A., Messina, S., Iuliano, R., Fusco, A., Santillo, M. R., Muller, M. T., Chiariotti, L., Gottesman, M. E., and Avvedimento, E. V. (2007). DNA damage, homology-directed repair, and DNA methylation. PLoS Genet *3*, e110.

Dalla-Favera, R., Gelmann, E. P., Gallo, R. C., and Wong-Staal, F. (1981). A human onc gene homologous to the transforming gene (v-sis) of simian sarcoma virus. Nature *292*, 31-35.

Daly, N. L., Arvanitis, D. A., Fairley, J. A., Gomez-Roman, N., Morton, J. P., Graham, S. V., Spandidos, D. A., and White, R. J. (2005). Deregulation of RNA polymerase III transcription in cervical epithelium in response to high-risk human papillomavirus. Oncogene *24*, 880-888.

Daniel, J. M., and Reynolds, A. B. (1999). The catenin p120(ctn) interacts with Kaiso, a novel BTB/POZ domain zinc finger transcription factor. Mol Cell Biol *19*, 3614-3623.

Daniels, G. R., and Deininger, P. L. (1985). Repeat sequence families derived from mammalian tRNA genes. Nature *317*, 819-822.

Das, G., Henning, D., Wright, D., and Reddy, R. (1988). Upstream regulatory elements are necessary and sufficient for transcription of a U6 RNA gene by RNA polymerase III. Embo J *7*, 503-512.

Davey, C., Pennings, S., and Allan, J. (1997). CpG methylation remodels chromatin structure in vitro. J Mol Biol *267*, 276-288.

Davidson, E. H., Hough, B. R., Amenson, C. S., and Britten, R. J. (1973). General interspersion of repetitive with non-repetitive sequence elements in the DNA of Xenopus. J Mol Biol *77*, 1-23.

Davidson, E. H., Hough, B. R., Klein, W. H., and Britten, R. J. (1975). Structural genes adjacent to interspersed repetitive DNA sequences. Cell *4*, 217-238.

de La Serna, I. L., Carlson, K. A., Hill, D. A., Guidi, C. J., Stephenson, R. O., Sif, S., Kingston, R. E., and Imbalzano, A. N. (2000). Mammalian SWI-SNF complexes contribute to activation of the hsp70 gene. Mol Cell Biol *20*, 2839-2851.

Deaton, A. M., and Bird, A. (2011). CpG islands and the regulation of transcription. Genes Dev *25*, 1010-1022.

Dechassa, M. L., Zhang, B., Horowitz-Scherer, R., Persinger, J., Woodcock, C. L., Peterson, C. L., and Bartholomew, B. (2008). Architecture of the SWI/SNF-nucleosome complex. Mol Cell Biol *28*, 6010-6021.

Deininger, P. L., and Batzer, M. A. (1999). Alu repeats and human disease. Mol Genet Metab *67*, 183-193.

Deininger, P. L., and Daniels, G. R. (1986). The recent evolution of mammalian repetitive DNA elements. Trends in Genetics *2*, 76-80.

Deininger, P. L., Moran, J. V., Batzer, M. A., and Kazazian, H. H., Jr. (2003). Mobile elements and mammalian genome evolution. Curr Opin Genet Dev *13*, 651-658.

Del Valle-Perez, B., Casagolda, D., Lugilde, E., Valls, G., Codina, M., Dave, N., de Herreros, A. G., and Dunach, M. (2011). Wnt controls the transcriptional activity of Kaiso through CK1{varepsilon}-dependent phosphorylation of p120-catenin. J Cell Sci *124*, 2298-2309.

DeLotto, R., and Schedl, P. (1984). Internal promoter elements of transfer RNA genes are preferentially exposed in chromatin. J Mol Biol *179*, 607-628.

Derreumaux, S., Chaoui, M., Tevanian, G., and Fermandjian, S. (2001). Impact of CpG methylation on structure, dynamics and solvation of cAMP DNA responsive element. Nucleic Acids Res *29*, 2314-2326.

Desai, N., Lee, J., Upadhya, R., Chu, Y., Moir, R. D., and Willis, I. M. (2005). Two steps in Maf1-dependent repression of transcription by RNA polymerase III. J Biol Chem *280*, 6455-6462.

Deutscher, S. L., Harley, J. B., and Keene, J. D. (1988). Molecular analysis of the 60-kDa human Ro ribonucleoprotein. Proc Natl Acad Sci U S A *85*, 9479-9483.

Dhasarathy, A., and Wade, P. A. (2008). The MBD protein family-reading an epigenetic mark? Mutat Res *647*, 39-43.

Dieci, G., Fiorino, G., Castelnuovo, M., Teichmann, M., and Pagano, A. (2007). The expanding RNA polymerase III transcriptome. Trends Genet *23*, 614-622.

Dieci, G., and Sentenac, A. (1996). Facilitated recycling pathway for RNA polymerase III. Cell *84*, 245-252.

Doi, M., Hirayama, J., and Sassone-Corsi, P. (2006). Circadian regulator CLOCK is a histone acetyltransferase. Cell *125*, 497-508.

Dong, K. B., Maksakova, I. A., Mohn, F., Leung, D., Appanah, R., Lee, S., Yang, H. W., Lam, L. L., Mager, D. L., Schubeler, D., Tachibana, M., Shinkai, Y., and Lorincz, M. C. (2008). DNA methylation in ES cells requires the lysine methyltransferase G9a but not its catalytic activity. Embo J *27*, 2691-2701.

Du, J., Nasir, I., Benton, B. K., Kladde, M. P., and Laurent, B. C. (1998). Sth1p, a Saccharomyces cerevisiae Snf2p/Swi2p homolog, is an essential ATPase in RSC and differs from Snf/Swi in its interactions with histones and chromatin-associated proteins. Genetics *150*, 987-1005.

Dumay-Odelot, H., Durrieu-Gaillard, S., Da Silva, D., Roeder, R. G., and Teichmann, M. (2010). Cell growth- and differentiation-dependent regulation of RNA polymerase III transcription. Cell Cycle *9*, 3687-3699.

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. J Biol Chem *282*, 17179-17189.

Dunaief, J. L., Strober, B. E., Guha, S., Khavari, P. A., Alin, K., Luban, J., Begemann, M., Crabtree, G. R., and Goff, S. P. (1994). The retinoblastoma

protein and BRG1 form a complex and cooperate to induce cell cycle arrest. Cell *79*, 119-130.

Duncan, B. K., and Miller, J. H. (1980). Mutagenic deamination of cytosine residues in DNA. Nature *287*, 560-561.

Duncan, C., Biro, P. A., Choudary, P. V., Elder, J. T., Wang, R. R., Forget, B. G., de Riel, J. K., and Weissman, S. M. (1979). RNA polymerase III transcriptional units are interspersed among human non-alpha-globin genes. Proc Natl Acad Sci U S A *76*, 5095-5099.

Echeverri, C. J., and Perrimon, N. (2006). High-throughput RNAi screening in cultured cells: a user's guide. Nat Rev Genet *7*, 373-384.

Ehrlich, M. (2002). DNA methylation in cancer: too much, but also too little. Oncogene *21*, 5400-5413.

Ehrlich, M. (2005). DNA methylation and cancer-associated genetic instability. Adv Exp Med Biol *570*, 363-392.

Ehrlich, M., Gama-Sosa, M. A., Huang, L. H., Midgett, R. M., Kuo, K. C., McCune, R. A., and Gehrke, C. (1982). Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. Nucleic Acids Res *10*, 2709-2721.

Ehrlich, M., and Wang, R. Y. (1981). 5-Methylcytosine in eukaryotic DNA. Science *212*, 1350-1357.

Elfring, L. K., Daniel, C., Papoulas, O., Deuring, R., Sarte, M., Moseley, S., Beek, S. J., Waldrip, W. R., Daubresse, G., DePace, A., Kennison, J. A., and Tamkun, J. W. (1998). Genetic analysis of brahma: the Drosophila homolog of the yeast chromatin remodeling factor SWI2/SNF2. Genetics *148*, 251-265.

Elliott, B., Richardson, C., and Jasin, M. (2005). Chromosomal translocation mechanisms at intronic alu elements in mammalian cells. Mol Cell *17*, 885-894.

Engelke, D. R., Ng, S. Y., Shastry, B. S., and Roeder, R. G. (1980). Specific interaction of a purified transcription factor with an internal control region of 5S RNA genes. Cell *19*, 717-728.

Englander, E. W., and Howard, B. H. (1995). Nucleosome positioning by human Alu elements in chromatin. J Biol Chem *270*, 10091-10096.

Englander, E. W., Wolffe, A. P., and Howard, B. H. (1993). Nucleosome interactions with a human Alu element. Transcriptional repression and effects of template methylation. J Biol Chem *268*, 19565-19573.

Espinoza, C. A., Goodrich, J. A., and Kugel, J. F. (2007). Characterization of the structure, function, and mechanism of B2 RNA, an ncRNA repressor of RNA polymerase II transcription. Rna *13*, 583-596.

Euskirchen, G. M., Auerbach, R. K., Davidov, E., Gianoulis, T. A., Zhong, G., Rozowsky, J., Bhardwaj, N., Gerstein, M. B., and Snyder, M. (2011). Diverse roles and interactions of the SWI/SNF chromatin remodeling complex revealed using global approaches. PLoS Genet *7*, e1002008.

Fabrizio, P., Coppo, A., Fruscoloni, P., Benedetti, P., Di Segni, G., and Tocchini-Valentini, G. P. (1987). Comparative mutational analysis of wild-type and stretched tRNA3(Leu) gene promoters. Proc Natl Acad Sci U S A *84*, 8763-8767.

Fairley, J. A., Mitchell, L. E., Berg, T., Kenneth, N. S., Schubert, C. V., Sillje, H. H. W., Medema, R. H., Nigg, E. A., and White, R. J. (2011). Direct regulation of

tRNA and 5S rRNA gene transcription by polo-like kinase 1. Submitted to Mol Cell.

Fairley, J. A., Scott, P. H., and White, R. J. (2003). TFIIIB is phosphorylated, disrupted and selectively released from tRNA promoters during mitosis in vivo. Embo J *22*, 5841-5850.

Fan, G., Siniaia, M., Poon, C. S., and Jaenisch, R. (2001). Effect of DNA hypomethylation on neural control of respiration: a genetic model. Adv Exp Med Biol *499*, 195-200.

Fanti, L., and Pimpinelli, S. (2008). HP1: a functionally multifaceted protein. Curr Opin Genet Dev *18*, 169-174.

Fedoroff, N., Wellauer, P. K., and Wall, R. (1977). Intermolecular duplexes in heterogeneous nuclear RNA from HeLa cells. Cell *10*, 597-610.

Felton-Edkins, Z. A., Fairley, J. A., Graham, E. L., Johnston, I. M., White, R. J., and Scott, P. H. (2003a). The mitogen-activated protein (MAP) kinase ERK induces tRNA synthesis by phosphorylating TFIIIB. Embo J *22*, 2422-2432.

Felton-Edkins, Z. A., Kenneth, N. S., Brown, T. R., Daly, N. L., Gomez-Roman, N., Grandori, C., Eisenman, R. N., and White, R. J. (2003b). Direct regulation of RNA polymerase III transcription by RB, p53 and c-Myc. Cell Cycle *2*, 181-184.

Felton-Edkins, Z. A., and White, R. J. (2002). Multiple mechanisms contribute to the activation of RNA polymerase III transcription in cells transformed by papovaviruses. J Biol Chem *277*, 48182-48191.

Feng, Q., Moran, J. V., Kazazian, H. H., Jr., and Boeke, J. D. (1996). Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. Cell *87*, 905-916.

Finnegan, E. J., Genger, R. K., Peacock, W. J., and Dennis, E. S. (1998). DNA Methylation in Plants. Annu Rev Plant Physiol Plant Mol Biol *49*, 223-247.

Fritsch, L., Robin, P., Mathieu, J. R., Souidi, M., Hinaux, H., Rougeulle, C., Harel-Bellan, A., Ameyar-Zazoua, M., and Ait-Si-Ali, S. (2010). A subset of the histone H3 lysine 9 methyltransferases Suv39h1, G9a, GLP, and SETDB1 participate in a multimeric complex. Mol Cell *37*, 46-56.

Fuhrman, S. A., Deininger, P. L., LaPorte, P., Friedmann, T., and Geiduschek, E. P. (1981). Analysis of transcription of the human Alu family ubiquitous repeating element by eukaryotic RNA polymerase III. Nucleic Acids Res *9*, 6439-6456.

Fujita, N., Watanabe, S., Ichimura, T., Tsuruzoe, S., Shinkai, Y., Tachibana, M., Chiba, T., and Nakao, M. (2003). Methyl-CpG binding domain 1 (MBD1) interacts with the Suv39h1-HP1 heterochromatic complex for DNA methylation-based transcriptional repression. J Biol Chem *278*, 24132-24138.

Fuks, F., Hurd, P. J., Deplus, R., and Kouzarides, T. (2003a). The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. Nucleic Acids Res *31*, 2305-2312.

Fuks, F., Hurd, P. J., Wolf, D., Nan, X., Bird, A. P., and Kouzarides, T. (2003b). The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. J Biol Chem *278*, 4035-4040.

Fukushige, S., Kondo, E., Gu, Z., Suzuki, H., and Horii, A. (2006). RET finger protein enhances MBD2- and MBD4-dependent transcriptional repression. Biochem Biophys Res Commun *351*, 85-92.

Gabrielsen, O. S., Marzouki, N., Ruet, A., Sentenac, A., and Fromageot, P. (1989). Two polypeptide chains in yeast transcription factor tau interact with DNA. J Biol Chem *264*, 7505-7511.

Galli, G., Hofstetter, H., and Birnstiel, M. L. (1981). Two conserved sequence blocks within eukaryotic tRNA genes are major promoter elements. Nature *294*, 626-631.

Geiduschek, E. P., and Kassavetis, G. A. (2001). The RNA polymerase III transcription apparatus. J Mol Biol *310*, 1-26.

Geng, F., Cao, Y., and Laurent, B. C. (2001). Essential roles of Snf5p in Snf-Swi chromatin remodeling in vivo. Mol Cell Biol *21*, 4311-4320.

Gilbert, N., Thomson, I., Boyle, S., Allan, J., Ramsahoye, B., and Bickmore, W. A. (2007). DNA methylation affects nuclear organization, histone modifications, and linker histone binding but not chromatin compaction. J Cell Biol *177*, 401-411.

Ginsberg, A. M., King, B. O., and Roeder, R. G. (1984). Xenopus 5S gene transcription factor, TFIIIA: characterization of a cDNA clone and measurement of RNA levels throughout development. Cell *39*, 479-489.

Glaros, S., Cirrincione, G. M., Muchardt, C., Kleer, C. G., Michael, C. W., and Reisman, D. (2007). The reversible epigenetic silencing of BRM: implications for clinical targeted therapy. Oncogene *26*, 7058-7066.

Gold, H. A., Topper, J. N., Clayton, D. A., and Craft, J. (1989). The RNA processing enzyme RNase MRP is identical to the Th RNP and related to RNase P. Science *245*, 1377-1380.

Gold, M., Gefter, M., Hausmann, R., and Hurwitz, J. (1966). Methylation of DNA. J Gen Physiol *49*, 5-28.

Goll, M. G., Kirpekar, F., Maggert, K. A., Yoder, J. A., Hsieh, C. L., Zhang, X., Golic, K. G., Jacobsen, S. E., and Bestor, T. H. (2006). Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2. Science *311*, 395-398.

Gollob, J. A., Sciambi, C. J., Peterson, B. L., Richmond, T., Thoreson, M., Moran, K., Dressman, H. K., Jelinek, J., and Issa, J. P. (2006). Phase I trial of sequential low-dose 5-aza-2'-deoxycytidine plus high-dose intravenous bolus interleukin-2 in patients with melanoma or renal cell carcinoma. Clin Cancer Res *12*, 4619-4627.

Goodfellow, S. J., Graham, E. L., Kantidakis, T., Marshall, L., Coppins, B. A., Oficjalska-Pham, D., Gerard, M., Lefebvre, O., and White, R. J. (2008). Regulation of RNA polymerase III transcription by Maf1 in mammalian cells. J Mol Biol *378*, 481-491.

Gottesfeld, J. M., Wolf, V. J., Dang, T., Forbes, D. J., and Hartl, P. (1994). Mitotic repression of RNA polymerase III transcription in vitro mediated by phosphorylation of a TFIIIB component. Science *263*, 81-84.

Graham, D. E., Neufeld, B. R., Davidson, E. H., and Britten, R. J. (1974). Interspersion of repetitive and non-repetitive DNA sequences in the Sea Urchin genome. Cell *1*, 127-137.

Grand, F., Kulkarni, S., Chase, A., Goldman, J. M., Gordon, M., and Cross, N. C. (1999). Frequent deletion of hSNF5/INI1, a component of the SWI/SNF complex, in chronic myeloid leukemia. Cancer Res *59*, 3870-3874.

Greiner, D., Bonaldi, T., Eskeland, R., Roemer, E., and Imhof, A. (2005). Identification of a specific inhibitor of the histone methyltransferase SU(VAR)3-9. Nat Chem Biol *1*, 143-145.

Grewal, S. I., and Jia, S. (2007). Heterochromatin revisited. Nat Rev Genet *8*, 35-46.

Gruenbaum, Y., Stein, R., Cedar, H., and Razin, A. (1981). Methylation of CpG sequences in eukaryotic DNA. FEBS Lett *124*, 67-71.

Grunau, C., Clark, S. J., and Rosenthal, A. (2001). Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. Nucleic Acids Res *29*, E65-65.

Guidi, C. J., Mudhasani, R., Hoover, K., Koff, A., Leav, I., Imbalzano, A. N., and Jones, S. N. (2006). Functional interaction of the retinoblastoma and Ini1/Snf5 tumor suppressors in cell growth and pituitary tumorigenesis. Cancer Res *66*, 8076-8082.

Guidi, C. J., Sands, A. T., Zambrowicz, B. P., Turner, T. K., Demers, D. A., Webster, W., Smith, T. W., Imbalzano, A. N., and Jones, S. N. (2001). Disruption of Ini1 leads to peri-implantation lethality and tumorigenesis in mice. Mol Cell Biol *21*, 3598-3603.

Guo, J. U., Su, Y., Zhong, C., Ming, G. L., and Song, H. (2011). Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. Cell *145*, 423-434.

Gutierrez-Hartmann, A., Lieberburg, I., Gardner, D., Baxter, J. D., and Cathala, G. G. (1984). Transcription of two classes of rat growth hormone gene-associated repetitive DNA: differences in activity and effects of tandem repeat structure. Nucleic Acids Res *12*, 7153-7173.

Haber, J. E. (2006). Transpositions and translocations induced by site-specific double-strand breaks in budding yeast. DNA Repair (Amst) *5*, 998-1009.

Hakimi, M. A., Bochar, D. A., Schmiesing, J. A., Dong, Y., Barak, O. G., Speicher, D. W., Yokomori, K., and Shiekhattar, R. (2002). A chromatin remodelling complex that loads cohesin onto human chromosomes. Nature *418*, 994-998.

Hanas, J. S., Bogenhagen, D. F., and Wu, C. W. (1983). Cooperative model for the binding of Xenopus transcription factor A to the 5S RNA gene. Proc Natl Acad Sci U S A *80*, 2142-2145.

Hanoune, J., and Agarwal, M. K. (1970). Studies on the half life time of rat liver transfer RNA species. FEBS Lett *11*, 78-80.

Hargreaves, D. C., and Crabtree, G. R. (2011). ATP-dependent chromatin remodeling: genetics, genomics and mechanisms. Cell Res *21*, 396-420.

Harikrishnan, K. N., Chow, M. Z., Baker, E. K., Pal, S., Bassal, S., Brasacchio, D., Wang, L., Craig, J. M., Jones, P. L., Sif, S., and El-Osta, A. (2005). Brahma links the SWI/SNF chromatin-remodeling complex with MeCP2-dependent transcriptional silencing. Nat Genet *37*, 254-264.

Hasler, J., and Strub, K. (2006). Alu RNP and Alu RNA regulate translation initiation in vitro. Nucleic Acids Res *34*, 2374-2385.

Helleday, T., Lo, J., van Gent, D. C., and Engelward, B. P. (2007). DNA double-strand break repair: from mechanistic understanding to cancer treatment. DNA Repair (Amst) *6*, 923-935.

Hendrich, B., and Bird, A. (1998). Identification and characterization of a family of mammalian methyl-CpG binding proteins. Mol Cell Biol *18*, 6538-6547.

Hendrich, B., Guy, J., Ramsahoye, B., Wilson, V. A., and Bird, A. (2001). Closely related proteins MBD2 and MBD3 play distinctive but interacting roles in mouse development. Genes Dev *15*, 710-723.

Hirschhorn, J. N., Brown, S. A., Clark, C. D., and Winston, F. (1992). Evidence that SNF2/SWI2 and SNF5 activate transcription in yeast by altering chromatin structure. Genes Dev *6*, 2288-2298.

Ho, K. L., McNae, I. W., Schmiedeberg, L., Klose, R. J., Bird, A. P., and Walkinshaw, M. D. (2008). MeCP2 binding to DNA depends upon hydration at methyl-CpG. Mol Cell *29*, 525-531.

Ho, L., Jothi, R., Ronan, J. L., Cui, K., Zhao, K., and Crabtree, G. R. (2009a). An embryonic stem cell chromatin remodeling complex, esBAF, is an essential component of the core pluripotency transcriptional network. Proc Natl Acad Sci U S A *106*, 5187-5191.

Ho, L., Ronan, J. L., Wu, J., Staahl, B. T., Chen, L., Kuo, A., Lessard, J., Nesvizhskii, A. I., Ranish, J., and Crabtree, G. R. (2009b). An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency. Proc Natl Acad Sci U S A *106*, 5181-5186.

Hopper, A. K., and Phizicky, E. M. (2003). tRNA transfers to the limelight. Genes Dev *17*, 162-180.

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.

Houck, C. M., and Schmid, C. W. (1981). The evolution of a family of short interspersed repeats in primate DNA. J Mol Evol *17*, 148-155.

Howard, G., Eiges, R., Gaudet, F., Jaenisch, R., and Eden, A. (2008). Activation and transposition of endogenous retroviral elements in hypomethylation induced tumors in mice. Oncogene *27*, 404-408.

Howe, J. G., and Shu, M. D. (1989). Epstein-Barr virus small RNA (EBER) genes: unique transcription units that combine RNA polymerase II and III promoter elements. Cell *57*, 825-834.

Hsieh, C. L. (2005). The de novo methylation activity of Dnmt3a is distinctly different than that of Dnmt1. BMC Biochem *6*, 6.

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.

Hu, P., Wu, S., and Hernandez, N. (2003). A minimal RNA polymerase III transcription system from human cells reveals positive and negative regulatory roles for CK2. Mol Cell *12*, 699-709.

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.

Huet, J., Conesa, C., Manaud, N., Chaussivert, N., and Sentenac, A. (1994). Interactions between yeast TFIIIB components. Nucleic Acids Res *22*, 3433-3439.

Isakoff, M. S., Sansam, C. G., Tamayo, P., Subramanian, A., Evans, J. A., Fillmore, C. M., Wang, X., Biegel, J. A., Pomeroy, S. L., Mesirov, J. P., and Roberts, C. W. (2005). Inactivation of the Snf5 tumor suppressor stimulates cell cycle progression and cooperates with p53 loss in oncogenic transformation. Proc Natl Acad Sci U S A *102*, 17745-17750.

Iwase, S., Lan, F., Bayliss, P., de la Torre-Ubieta, L., Huarte, M., Qi, H. H., Whetstine, J. R., Bonni, A., Roberts, T. M., and Shi, Y. (2007). The X-linked mental retardation gene SMCX/JARID1C defines a family of histone H3 lysine 4 demethylases. Cell *128*, 1077-1088.

Jackson-Grusby, L., Beard, C., Possemato, R., Tudor, M., Fambrough, D., Csankovszki, G., Dausman, J., Lee, P., Wilson, C., Lander, E., and Jaenisch, R. (2001). Loss of genomic methylation causes p53-dependent apoptosis and epigenetic deregulation. Nat Genet *27*, 31-39.

Jagadeeswaran, P., Forget, B. G., and Weissman, S. M. (1981). Short interspersed repetitive DNA elements in eucaryotes: transposable DNA elements generated by reverse transcription of RNA pol III transcripts? Cell *26*, 141-142.

Jelinek, W., Mollov, G., Salditt, M., Wall, R., Sheiness, D., and Darnell, J. E., Jr. (1974a). Origin of mRNA in HeLa cells and the implications for chromosome structure. Cold Spring Harb Symp Quant Biol *38*, 891-898.

Jelinek, W., Molloy, G., Fernandez-Munoz, R., Salditt, M., and Darnell, J. E. (1974b). Secondary structure in heterogeneous nuclear RNA: involvement of regions from repeated DNA sites. J Mol Biol *82*, 361-370.

Jelinek, W. R., Toomey, T. P., Leinwand, L., Duncan, C. H., Biro, P. A., Choudary, P. V., Weissman, S. M., Rubin, C. M., Houck, C. M., Deininger, P. L., and Schmid, C. W. (1980). Ubiquitous, interspersed repeated sequences in mammalian genomes. Proc Natl Acad Sci U S A *77*, 1398-1402.

Jeong, K. S., and Lee, S. (2005). Estimating the total mouse DNA methylation according to the B1 repetitive elements. Biochem Biophys Res Commun *335*, 1211-1216.

Jiang, C. L., Jin, S. G., and Pfeifer, G. P. (2004). MBD3L1 is a transcriptional repressor that interacts with methyl-CpG-binding protein 2 (MBD2) and components of the NuRD complex. J Biol Chem *279*, 52456-52464.

Johnson, T. B., and Coghill, R. D. (1925). Researches on pyrimidines. C111. The discovery of 5-methylcytosine in tuberculinic acid, the nucleic acid of the tubercle bacillus. J Am Chem Soc *47*, 2838-2844.

Johnston, I. M., Allison, S. J., Morton, J. P., Schramm, L., Scott, P. H., and White, R. J. (2002). CK2 forms a stable complex with TFIIIB and activates RNA polymerase III transcription in human cells. Mol Cell Biol *22*, 3757-3768.

Jones, P. A., and Taylor, S. M. (1980). Cellular differentiation, cytidine analogs and DNA methylation. Cell *20*, 85-93.

Jones, P. L., Veenstra, G. J., Wade, P. A., Vermaak, D., Kass, S. U., Landsberger, N., Strouboulis, J., and Wolffe, A. P. (1998). Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nat Genet *19*, 187-191.

Jorgensen, H. F., Ben-Porath, I., and Bird, A. P. (2004). Mbd1 is recruited to both methylated and nonmethylated CpGs via distinct DNA binding domains. Mol Cell Biol *24*, 3387-3395.

Jurgens, B., Schmitz-Drager, B. J., and Schulz, W. A. (1996). Hypomethylation of L1 LINE sequences prevailing in human urothelial carcinoma. Cancer Res *56*, 5698-5703.

Jurka, J. (1997). Sequence patterns indicate an enzymatic involvement in integration of mammalian retroposons. Proc Natl Acad Sci U S A *94*, 1872-1877.

Jurka, J., and Milosavljevic, A. (1991). Reconstruction and analysis of human Alu genes. J Mol Evol *32*, 105-121.

Jurka, J., and Smith, T. (1988). A fundamental division in the Alu family of repeated sequences. Proc Natl Acad Sci U S A *85*, 4775-4778.

Jurka, J., Zietkiewicz, E., and Labuda, D. (1995). Ubiquitous mammalian-wide interspersed repeats (MIRs) are molecular fossils from the mesozoic era. Nucleic Acids Res *23*, 170-175.

Juttermann, R., Li, E., and Jaenisch, R. (1994). Toxicity of 5-aza-2'-deoxycytidine to mammalian cells is mediated primarily by covalent trapping of DNA methyltransferase rather than DNA demethylation. Proc Natl Acad Sci U S A *91*, 11797-11801.

Kalpana, G. V., Marmon, S., Wang, W., Crabtree, G. R., and Goff, S. P. (1994). Binding and stimulation of HIV-1 integrase by a human homolog of yeast transcription factor SNF5. Science *266*, 2002-2006.

Kaneko, H., Dridi, S., Tarallo, V., Gelfand, B. D., Fowler, B. J., Cho, W. G., Kleinman, M. E., Ponicsan, S. L., Hauswirth, W. W., Chiodo, V. A., Kariko, K., Yoo, J. W., Lee, D. K., Hadziahmetovic, M., Song, Y., Misra, S., Chaudhuri, G., Buaas, F. W., Braun, R. E., Hinton, D. R., Zhang, Q., Grossniklaus, H. E., Provis, J. M., Madigan, M. C., Milam, A. H., Justice, N. L., Albuquerque, R. J., Blandford, A. D., Bogdanovich, S., Hirano, Y., Witta, J., Fuchs, E., Littman, D. R., Ambati, B. K., Rudin, C. M., Chong, M. M., Provost, P., Kugel, J. F., Goodrich, J. A., Dunaief, J. L., Baffi, J. Z., and Ambati, J. (2011). DICER1 deficit induces Alu RNA toxicity in age-related macular degeneration. Nature *471*, 325-330.

Kang, H., Cui, K., and Zhao, K. (2004). BRG1 controls the activity of the retinoblastoma protein via regulation of p21CIP1/WAF1/SDI. Mol Cell Biol *24*, 1188-1199.

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.

Karpf, A. R., and Matsui, S. (2005). Genetic disruption of cytosine DNA methyltransferase enzymes induces chromosomal instability in human cancer cells. Cancer Res *65*, 8635-8639.

Kassavetis, G. A., Bartholomew, B., Blanco, J. A., Johnson, T. E., and Geiduschek, E. P. (1991). Two essential components of the Saccharomyces cerevisiae transcription factor TFIIIB: transcription and DNA-binding properties. Proc Natl Acad Sci U S A *88*, 7308-7312.

Kassavetis, G. A., Blanco, J. A., Johnson, T. E., and Geiduschek, E. P. (1992a). Formation of open and elongating transcription complexes by RNA polymerase III. J Mol Biol *226*, 47-58.

Kassavetis, G. A., Braun, B. R., Nguyen, L. H., and Geiduschek, E. P. (1990). S. cerevisiae TFIIIB is the transcription initiation factor proper of RNA polymerase III, while TFIIIA and TFIIIC are assembly factors. Cell *60*, 235-245.

Kassavetis, G. A., Joazeiro, C. A., Pisano, M., Geiduschek, E. P., Colbert, T., Hahn, S., and Blanco, J. A. (1992b). The role of the TATA-binding protein in the assembly and function of the multisubunit yeast RNA polymerase III transcription factor, TFIIIB. Cell *71*, 1055-1064.

Kato, H., Honma, R., Sanda, T., Fujiwara, T., Ito, E., Yanagisawa, Y., Imai, J., Okamoto, T., and Watanabe, S. (2007). Knock down of hSNF5/Ini1 causes cell cycle arrest and apoptosis in a p53-dependent manner. Biochem Biophys Res Commun *361*, 580-585.

Kenneth, N. S., Ramsbottom, B. A., Gomez-Roman, N., Marshall, L., Cole, P. A., and White, R. J. (2007). TRRAP and GCN5 are used by c-Myc to activate RNA polymerase III transcription. Proc Natl Acad Sci U S A *104*, 14917-14922.

Khavari, P. A., Peterson, C. L., Tamkun, J. W., Mendel, D. B., and Crabtree, G. R. (1993). BRG1 contains a conserved domain of the SWI2/SNF2 family necessary for normal mitotic growth and transcription. Nature *366*, 170-174.

Khoo, B., Brophy, B., and Jackson, S. P. (1994). Conserved functional domains of the RNA polymerase III general transcription factor BRF. Genes Dev *8*, 2879-2890.

Kim, C., Rubin, C. M., and Schmid, C. W. (2001). Genome-wide chromatin remodeling modulates the Alu heat shock response. Gene *276*, 127-133.

Kim, G. D., Ni, J., Kelesoglu, N., Roberts, R. J., and Pradhan, S. (2002). Co-operation and communication between the human maintenance and de novo DNA (cytosine-5) methyltransferases. Embo J *21*, 4183-4195.

Kim, J. H., Saraf, A., Florens, L., Washburn, M., and Workman, J. L. (2010). Gcn5 regulates the dissociation of SWI/SNF from chromatin by acetylation of Swi2/Snf2. Genes Dev *24*, 2766-2771.

Kim, Y. I., Giuliano, A., Hatch, K. D., Schneider, A., Nour, M. A., Dallal, G. E., Selhub, J., and Mason, J. B. (1994). Global DNA hypomethylation increases progressively in cervical dysplasia and carcinoma. Cancer *74*, 893-899.

Kimura, H., and Shiota, K. (2003). Methyl-CpG-binding protein, MeCP2, is a target molecule for maintenance DNA methyltransferase, Dnmt1. J Biol Chem *278*, 4806-4812.

Kimura, H., Tao, Y., Roeder, R. G., and Cook, P. R. (1999). Quantitation of RNA polymerase II and its transcription factors in an HeLa cell: little soluble holoenzyme but significant amounts of polymerases attached to the nuclear substructure. Mol Cell Biol *19*, 5383-5392.

Kit, S. (1961). Equilibrium sedimentation in density gradients of DNA preparations from animal tissues. J Mol Biol *3*, 711-716.

Klose, R. J., and Bird, A. P. (2006). Genomic DNA methylation: the mark and its mediators. Trends Biochem Sci *31*, 89-97.

Koch, C., and Stratling, W. H. (2004). DNA binding of methyl-CpG-binding protein MeCP2 in human MCF7 cells. Biochemistry *43*, 5011-5021.

Kochanek, S., Renz, D., and Doerfler, W. (1993). DNA methylation in the Alu sequences of diploid and haploid primary human cells. Embo J *12*, 1141-1151.

Kolomietz, E., Meyn, M. S., Pandita, A., and Squire, J. A. (2002). The role of Alu repeat clusters as mediators of recurrent chromosomal aberrations in tumors. Genes Chromosomes Cancer *35*, 97-112.

Kondo, E., Gu, Z., Horii, A., and Fukushige, S. (2005). The thymine DNA glycosylase MBD4 represses transcription and is associated with methylated p16(INK4a) and hMLH1 genes. Mol Cell Biol *25*, 4388-4396.

Kondo, Y., and Issa, J. P. (2003). Enrichment for histone H3 lysine 9 methylation at Alu repeats in human cells. J Biol Chem *278*, 27658-27662.

Kondrashov, A. V., Kiefmann, M., Ebnet, K., Khanam, T., Muddashetty, R. S., and Brosius, J. (2005). Inhibitory effect of naked neural BC1 RNA or BC200 RNA on eukaryotic in vitro translation systems is reversed by poly(A)-binding protein (PABP). J Mol Biol *353*, 88-103.

Konkel, M. K., and Batzer, M. A. (2011). A mobile threat to genome stability: The impact of non-LTR retrotransposons upon the human genome. Semin Cancer Biol *20*, 211-221.

Korenberg, J. R., and Rykowski, M. C. (1988). Human genome organization: Alu, lines, and the molecular structure of metaphase chromosome bands. Cell *53*, 391-400.

Koski, R. A., Clarkson, S. G., Kurjan, J., Hall, B. D., and Smith, M. (1980). Mutations of the yeast SUP4 tRNATyr locus: transcription of the mutant genes in vitro. Cell *22*, 415-425.

Kouzarides, T. (2007). Chromatin modifications and their function. Cell *128*, 693-705.

Kovelman, R., and Roeder, R. G. (1992). Purification and characterization of two forms of human transcription factor IIIC. J Biol Chem *267*, 24446-24456.

Kram, R., Botchan, M., and Hearst, J. E. (1972). Arrangement of the highly reiterated DNA sequences in the centric heterochromatin of Drosophila melanogaster. Evidence for interspersed spacer DNA. J Mol Biol *64*, 103-117.

Kramerov, D. A., Grigoryan, A. A., Ryskov, A. P., and Georgiev, G. P. (1979). Long double-stranded sequences (dsRNA-B) of nuclear pre-mRNA consist of a few highly abundant classes of sequences: evidence from DNA cloning experiments. Nucleic Acids Res *6*, 697-713.

Kravchenko, J. E., Rogozin, I. B., Koonin, E. V., and Chumakov, P. M. (2005). Transcription of mammalian messenger RNAs by a nuclear RNA polymerase of mitochondrial origin. Nature *436*, 735-739.

Krayev, A. S., Kramerov, D. A., Skryabin, K. G., Ryskov, A. P., Bayev, A. A., and Georgiev, G. P. (1980). The nucleotide sequence of the ubiquitous repetitive DNA sequence B1 complementary to the most abundant class of mouse fold-back RNA. Nucleic Acids Res *8*, 1201-1215.

Krayev, A. S., Markusheva, T. V., Kramerov, D. A., Ryskov, A. P., Skryabin, K. G., Bayev, A. A., and Georgiev, G. P. (1982). Ubiquitous transposon-like repeats B1 and B2 of the mouse genome: B2 sequencing. Nucleic Acids Res *10*, 7461-7475.

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.

Kurjan, J., and Hall, B. D. (1982). Mutations at the Saccharomyces cerevisiae SUP4 tRNA(Tyr) locus: isolation, genetic fine-structure mapping, and correlation with physical structure. Mol Cell Biol *2*, 1501-1513.

Kutter, C., Brown, G. D., Goncalves, A., Wilson, M. D., Watt, S., Brazma, A., White, R. J., and Odom, D. T. (2011). Pol III binding in six mammals shows conservation among amino acid isotypes despite divergence among tRNA genes. Nat Genet *43*, 948-955.

Kwon, H., Imbalzano, A. N., Khavari, P. A., Kingston, R. E., and Green, M. R. (1994). Nucleosome disruption and enhancement of activator binding by a human SW1/SNF complex. Nature *370*, 477-481.

Kwon, S. H., and Workman, J. L. (2010). The changing faces of HP1: From heterochromatin formation and gene silencing to euchromatic gene expression: HP1 acts as a positive regulator of transcription. Bioessays *33*, 280-289.

Laborda, J. (1991). 36B4 cDNA used as an estradiol-independent mRNA control is the cDNA for human acidic ribosomal phosphoprotein PO. Nucleic Acids Res *19*, 3998.

Labuda, D., Sinnett, D., Richer, C., Deragon, J. M., and Striker, G. (1991). Evolution of mouse B1 repeats: 7SL RNA folding pattern conserved. J Mol Evol *32*, 405-414.

Lakshmikuttyamma, A., Scott, S. A., DeCoteau, J. F., and Geyer, C. R. (2010). Reexpression of epigenetically silenced AML tumor suppressor genes by SUV39H1 inhibition. Oncogene *29*, 576-588.

Lande-Diner, L., Zhang, J., Ben-Porath, I., Amariglio, N., Keshet, I., Hecht, M., Azuara, V., Fisher, A. G., Rechavi, G., and Cedar, H. (2007). Role of DNA methylation in stable gene repression. J Biol Chem *282*, 12194-12200.

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J. P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J. C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R. H., Wilson, R. K., Hillier, L. W., McPherson, J. D., Marra, M. A., Mardis, E. R., Fulton, L. A., Chinwalla, A. T., Pepin, K. H., Gish, W. R., Chissoe, S. L., Wendl, M. C., Delehaunty, K. D., Miner, T. L., Delehaunty, A., Kramer, J. B., Cook, L. L., Fulton, R. S., Johnson, D. L., Minx, P. J., Clifton, S. W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N.,

Cheng, J. F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., Gibbs, R. A., Muzny, D. M., Scherer, S. E., Bouck, J. B., Sodergren, E. J., Worley, K. C., Rives, C. M., Gorrell, J. H., Metzker, M. L., Naylor, S. L., Kucherlapati, R. S., Nelson, D. L., Weinstock, G. M., Sakaki, Y., Fujiyama, A., Hattori, M., Yada, T., Toyoda, A., Itoh, T., Kawagoe, C., Watanabe, H., Totoki, Y., Taylor, T., Weissenbach, J., Heilig, R., Saurin, W., Artiguenave, F., Brottier, P., Bruls, T., Pelletier, E., Robert, C., Wincker, P., Smith, D. R., Doucette-Stamm, L., Rubenfield, M., Weinstock, K., Lee, H. M., Dubois, J., Rosenthal, A., Platzer, M., Nyakatura, G., Taudien, S., Rump, A., Yang, H., Yu, J., Wang, J., Huang, G., Gu, J., Hood, L., Rowen, L., Madan, A., Qin, S., Davis, R. W., Federspiel, N. A., Abola, A. P., Proctor, M. J., Myers, R. M., Schmutz, J., Dickson, M., Grimwood, J., Cox, D. R., Olson, M. V., Kaul, R., Raymond, C., Shimizu, N., Kawasaki, K., Minoshima, S., Evans, G. A., Athanasiou, M., Schultz, R., Roe, B. A., Chen, F., Pan, H., Ramser, J., Lehrach, H., Reinhardt, R., McCombie, W. R., de la Bastide, M., Dedhia, N., Blocker, H., Hornischer, K., Nordsiek, G., Agarwala, R., Aravind, L., Bailey, J. A., Bateman, A., Batzoglou, S., Birney, E., Bork, P., Brown, D. G., Burge, C. B., Cerutti, L., Chen, H. C., Church, D., Clamp, M., Copley, R. R., Doerks, T., Eddy, S. R., Eichler, E. E., Furey, T. S., Galagan, J., Gilbert, J. G., Harmon, C., Hayashizaki, Y., Haussler, D., Hermjakob, H., Hokamp, K., Jang, W., Johnson, L. S., Jones, T. A., Kasif, S., Kaspryzk, A., Kennedy, S., Kent, W. J., Kitts, P., Koonin, E. V., Korf, I., Kulp, D., Lancet, D., Lowe, T. M., McLysaght, A., Mikkelsen, T., Moran, J. V., Mulder, N., Pollara, V. J., Ponting, C. P., Schuler, G., Schultz, J., Slater, G., Smit, A. F., Stupka, E., Szustakowski, J., Thierry-Mieg, D., Thierry-Mieg, J., Wagner, L., Wallis, J., Wheeler, R., Williams, A., Wolf, Y. I., Wolfe, K. H., Yang, S. P., Yeh, R. F., Collins, F., Guyer, M. S., Peterson, J., Felsenfeld, A., Wetterstrand, K. A., Patrinos, A., Morgan, M. J., de Jong, P., Catanese, J. J., Osoegawa, K., Shizuya, H., Choi, S., and Chen, Y. J. (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860-921.

Larminie, C. G., Cairns, C. A., Mital, R., Martin, K., Kouzarides, T., Jackson, S. P., and White, R. J. (1997). Mechanistic analysis of RNA polymerase III regulation by the retinoblastoma protein. Embo J *16*, 2061-2071.

Larminie, C. G., Sutcliffe, J. E., Tosh, K., Winter, A. G., Felton-Edkins, Z. A., and White, R. J. (1999). Activation of RNA polymerase III transcription in cells transformed by simian virus 40. Mol Cell Biol *19*, 4927-4934.

Laurent, B. C., Treich, I., and Carlson, M. (1993). The yeast SNF2/SWI2 protein has DNA-stimulated ATPase activity required for transcriptional activation. Genes Dev *7*, 583-591.

Laurent, L., Wong, E., Li, G., Huynh, T., Tsirigos, A., Ong, C. T., Low, H. M., Kin Sung, K. W., Rigoutsos, I., Loring, J., and Wei, C. L. (2010). Dynamic changes in the human methylome during differentiation. Genome Res *20*, 320-331.

Lawrence, C. B., McDonnell, D. P., and Ramsey, W. J. (1985). Analysis of repetitive sequence elements containing tRNA-like sequences. Nucleic Acids Res *13*, 4239-4252.

Lee, A. J., Endesfelder, D., Rowan, A. J., Walther, A., Birkbak, N. J., Futreal, P. A., Downward, J., Szallasi, Z., Tomlinson, I. P., Howell, M., Kschischo, M., and Swanton, C. (2011). Chromosomal instability confers intrinsic multidrug resistance. Cancer Res *71*, 1858-1870.

Lee, D. Y., Hayes, J. J., Pruss, D., and Wolffe, A. P. (1993). A positive role for histone acetylation in transcription factor access to nucleosomal DNA. Cell *72*, 73-84.

Lee, G. E., Kim, J. H., Taylor, M., and Muller, M. T. (2010). DNA methyltransferase 1-associated protein (DMAP1) is a co-repressor that stimulates DNA methylation globally and locally at sites of double strand break repair. J Biol Chem *285*, 37630-37640.

Lee, K. K., and Workman, J. L. (2007). Histone acetyltransferase complexes: one size doesn't fit all. Nat Rev Mol Cell Biol *8*, 284-295.

Leeflang, E. P., Liu, W. M., Hashimoto, C., Choudary, P. V., and Schmid, C. W. (1992). Phylogenetic evidence for multiple Alu source genes. J Mol Evol *35*, 7-16.

Lehnertz, B., Ueda, Y., Derijck, A. A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T., and Peters, A. H. (2003). Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. Curr Biol *13*, 1192-1200.

Lehrman, M. A., Russell, D. W., Goldstein, J. L., and Brown, M. S. (1987). Alu-Alu recombination deletes splice acceptor sites and produces secreted low density lipoprotein receptor in a subject with familial hypercholesterolemia. J Biol Chem *262*, 3354-3361.

Lei, H., Oh, S. P., Okano, M., Juttermann, R., Goss, K. A., Jaenisch, R., and Li, E. (1996). De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. Development *122*, 3195-3205.

Leipe, D. D., and Landsman, D. (1997). Histone deacetylases, acetoin utilization proteins and acetylpolyamine amidohydrolases are members of an ancient protein superfamily. Nucleic Acids Res *25*, 3693-3697.

Lessard, J., Wu, J. I., Ranish, J. A., Wan, M., Winslow, M. M., Staahl, B. T., Wu, H., Aebersold, R., Graef, I. A., and Crabtree, G. R. (2007). An essential switch in subunit composition of a chromatin remodeling complex during neural development. Neuron *55*, 201-215.

Lewis, J. D., Meehan, R. R., Henzel, W. J., Maurer-Fogy, I., Jeppesen, P., Klein, F., and Bird, A. (1992). Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. Cell *69*, 905-914.

Li, E., Bestor, T. H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell *69*, 915-926.

Li, H., Ilin, S., Wang, W., Duncan, E. M., Wysocka, J., Allis, C. D., and Patel, D. J. (2006). Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. Nature *442*, 91-95.

Li, T., Spearow, J., Rubin, C. M., and Schmid, C. W. (1999). Physiological stresses increase mouse short interspersed element (SINE) RNA expression in vivo. Gene *239*, 367-372.

Li, T. H., Kim, C., Rubin, C. M., and Schmid, C. W. (2000). K562 cells implicate increased chromatin accessibility in Alu transcriptional activation. Nucleic Acids Res *28*, 3031-3039.

Li, T. H., and Schmid, C. W. (2004). Alu's dimeric consensus sequence destabilizes its transcripts. Gene *324*, 191-200.

Liang, G., Chan, M. F., Tomigahara, Y., Tsai, Y. C., Gonzales, F. A., Li, E., Laird, P. W., and Jones, P. A. (2002). Cooperativity between DNA methyltransferases in the maintenance methylation of repetitive elements. Mol Cell Biol *22*, 480-491.

Lin, C. H., Hsieh, S. Y., Sheen, I. S., Lee, W. C., Chen, T. C., Shyu, W. C., and Liaw, Y. F. (2001). Genome-wide hypomethylation in hepatocellular carcinogenesis. Cancer Res *61*, 4238-4243.

Lindell, T. J., Weinberg, F., Morris, P. W., Roeder, R. G., and Rutter, W. J. (1970). Specific inhibition of nuclear RNA polymerase II by alpha-amanitin. Science *170*, 447-449.

Lister, R., Pelizzola, M., Dowen, R. H., Hawkins, R. D., Hon, G., Tonti-Filippini, J., Nery, J. R., Lee, L., Ye, Z., Ngo, Q. M., Edsall, L., Antosiewicz-Bourget, J., Stewart, R., Ruotti, V., Millar, A. H., Thomson, J. A., Ren, B., and Ecker, J. R. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. Nature *462*, 315-322.

Liu, W. M., Chu, W. M., Choudary, P. V., and Schmid, C. W. (1995). Cell stress and translational inhibitors transiently increase the abundance of mammalian SINE transcripts. Nucleic Acids Res *23*, 1758-1765.

Liu, W. M., Maraia, R. J., Rubin, C. M., and Schmid, C. W. (1994). Alu transcripts: cytoplasmic localisation and regulation by DNA methylation. Nucleic Acids Res *22*, 1087-1095.

Liu, W. M., and Schmid, C. W. (1993). Proposed roles for DNA methylation in Alu transcriptional repression and mutational inactivation. Nucleic Acids Res *21*, 1351-1359.

Liu, Y., Oakeley, E. J., Sun, L., and Jost, J. P. (1998). Multiple domains are involved in the targeting of the mouse DNA methyltransferase to the DNA replication foci. Nucleic Acids Res *26*, 1038-1045.

Lobo, S. M., and Hernandez, N. (1989). A 7 bp mutation converts a human RNA polymerase II snRNA promoter into an RNA polymerase III promoter. Cell *58*, 55-67.

Lockwood, W. W., Chari, R., Coe, B. P., Thu, K. L., Garnis, C., Malloff, C. A., Campbell, J., Williams, A. C., Hwang, D., Zhu, C. Q., Buys, T. P., Yee, J., English, J. C., Macaulay, C., Tsao, M. S., Gazdar, A. F., Minna, J. D., Lam, S., and Lam, W. L. (2010). Integrative genomic analyses identify BRF2 as a novel lineage-specific oncogene in lung squamous cell carcinoma. PLoS Med *7*, e1000315.

Loeken, M., Bikel, I., Livingston, D. M., and Brady, J. (1988). trans-activation of RNA polymerase II and III promoters by SV40 small t antigen. Cell *55*, 1171-1177.

Loenarz, C., Ge, W., Coleman, M. L., Rose, N. R., Cooper, C. D., Klose, R. J., Ratcliffe, P. J., and Schofield, C. J. (2010). PHF8, a gene associated with cleft lip/palate and mental retardation, encodes for an Nepsilon-dimethyl lysine demethylase. Hum Mol Genet *19*, 217-222.

Luan, D. D., Korman, M. H., Jakubczak, J. L., and Eickbush, T. H. (1993). Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. Cell *72*, 595-605.

Lunyak, V. V., Prefontaine, G. G., Nunez, E., Cramer, T., Ju, B. G., Ohgi, K. A., Hutt, K., Roy, R., Garcia-Diaz, A., Zhu, X., Yung, Y., Montoliu, L., Glass, C. K.,

and Rosenfeld, M. G. (2007). Developmentally regulated activation of a SINE B2 repeat as a domain boundary in organogenesis. Science *317*, 248-251.

Lymn, R. W., and Taylor, E. W. (1971). Mechanism of adenosine triphosphate hydrolysis by actomyosin. Biochemistry *10*, 4617-4624.

Mahapatra, S., Dewari, P. S., Bhardwaj, A., and Bhargava, P. (2011). Yeast H2A.Z, FACT complex and RSC regulate transcription of tRNA gene through differential dynamics of flanking nucleosomes. Nucleic Acids Res *39*, 4023-4034.

Manning, J. E., Schmid, C. W., and Davidson, N. (1975). Interspersion of repetitive and nonrepetitive DNA sequences in the Drosophila melanogaster genome. Cell *4*, 141-155.

Maraia, R. J. (1991). The subset of mouse B1 (Alu-equivalent) sequences expressed as small processed cytoplasmic transcripts. Nucleic Acids Res *19*, 5695-5702.

Maraia, R. J. (2001). La protein and the trafficking of nascent RNA polymerase iii transcripts. J Cell Biol *153*, F13-18.

Maraia, R. J., Driscoll, C. T., Bilyeu, T., Hsu, K., and Darlington, G. J. (1993). Multiple dispersed loci produce small cytoplasmic Alu RNA. Mol Cell Biol *13*, 4233-4241.

Maraia, R. J., Kenan, D. J., and Keene, J. D. (1994). Eukaryotic transcription termination factor La mediates transcript release and facilitates reinitiation by RNA polymerase III. Mol Cell Biol *14*, 2147-2158.

Margueron, R., Justin, N., Ohno, K., Sharpe, M. L., Son, J., Drury, W. J., 3rd, Voigt, P., Martin, S. R., Taylor, W. R., De Marco, V., Pirrotta, V., Reinberg, D., and Gamblin, S. J. (2009). Role of the polycomb protein EED in the propagation of repressive histone marks. Nature *461*, 762-767.

Mariner, P. D., Walters, R. D., Espinoza, C. A., Drullinger, L. F., Wagner, S. D., Kugel, J. F., and Goodrich, J. A. (2008). Human Alu RNA is a modular transacting repressor of mRNA transcription during heat shock. Mol Cell *29*, 499-509.

Marshall, L., Kenneth, N. S., and White, R. J. (2008). Elevated tRNA(iMet) synthesis can drive cell proliferation and oncogenic transformation. Cell *133*, 78-89.

Martens, J. H., O'Sullivan, R. J., Braunschweig, U., Opravil, S., Radolf, M., Steinlein, P., and Jenuwein, T. (2005). The profile of repeat-associated histone lysine methylation states in the mouse epigenome. Embo J *24*, 800-812.

Mason, N., Ciufo, L. F., and Brown, J. D. (2000). Elongation arrest is a physiologically important function of signal recognition particle. EMBO J *19*, 4164-4174.

Matera, A. G., Hellmann, U., and Schmid, C. W. (1990). A transpositionally and transcriptionally competent Alu subfamily. Mol Cell Biol *10*, 5424-5432.

Mathias, S. L., Scott, A. F., Kazazian, H. H., Jr., Boeke, J. D., and Gabriel, A. (1991). Reverse transcriptase encoded by a human transposable element. Science *254*, 1808-1810.

Matsui, T., Leung, D., Miyashita, H., Maksakova, I. A., Miyachi, H., Kimura, H., Tachibana, M., Lorincz, M. C., and Shinkai, Y. (2010). Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. Nature *464*, 927-931.

McStay, B., and Grummt, I. (2008). The epigenetics of rRNA genes: from molecular to chromosome biology. Annu Rev Cell Dev Biol *24*, 131-157.

Melchor, L., Garcia, M. J., Honrado, E., Pole, J. C., Alvarez, S., Edwards, P. A., Caldas, C., Brenton, J. D., and Benitez, J. (2007). Genomic analysis of the 8p11-12 amplicon in familial breast cancer. Int J Cancer *120*, 714-717.

Mertens, C., and Roeder, R. G. (2008). Different functional modes of p300 in activation of RNA polymerase III transcription from chromatin templates. Mol Cell Biol *28*, 5764-5776.

Metivier, R., Gallais, R., Tiffoche, C., Le Peron, C., Jurkowska, R. Z., Carmouche, R. P., Ibberson, D., Barath, P., Demay, F., Reid, G., Benes, V., Jeltsch, A., Gannon, F., and Salbert, G. (2008). Cyclical DNA methylation of a transcriptionally active promoter. Nature *452*, 45-50.

Millar, C. B., Guy, J., Sansom, O. J., Selfridge, J., MacDougall, E., Hendrich, B., Keightley, P. D., Bishop, S. M., Clarke, A. R., and Bird, A. (2002). Enhanced CpG mutability and tumorigenesis in MBD4-deficient mice. Science *297*, 403-405.

Miller, J., McLachlan, A. D., and Klug, A. (1985). Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes. Embo J *4*, 1609-1614.

Mittal, V., Cleary, M. A., Herr, W., and Hernandez, N. (1996). The Oct-1 POU-specific domain can stimulate small nuclear RNA gene transcription by stabilizing the basal transcription complex SNAPc. Mol Cell Biol *16*, 1955-1965.

Mittal, V., and Hernandez, N. (1997). Role for the amino-terminal region of human TBP in U6 snRNA transcription. Science *275*, 1136-1140.

Moir, R. D., Lee, J., Haeusler, R. A., Desai, N., Engelke, D. R., and Willis, I. M. (2006). Protein kinase A regulates RNA polymerase III transcription through the nuclear localization of Maf1. Proc Natl Acad Sci U S A *103*, 15044-15049.

Monteys, A. M., Spengler, R. M., Wan, J., Tecedor, L., Lennox, K. A., Xing, Y., and Davidson, B. L. (2010). Structure and activity of putative intronic miRNA promoters. Rna *16*, 495-505.

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.

Morales-Ramirez, P., Rodriguez-Reyes, R., Toribio-Escobedo, E., Olvera-Nestor, C., and Garcia-Firo, B. (2007). Mechanism of in vivo sister-chromatid exchange induction by 5-azacytidine. Mutagenesis *22*, 177-181.

Morse, R. H. (1989). Nucleosomes inhibit both transcriptional initiation and elongation by RNA polymerase III in vitro. Embo J *8*, 2343-2351.

Mortusewicz, O., Schermelleh, L., Walter, J., Cardoso, M. C., and Leonhardt, H. (2005). Recruitment of DNA methyltransferase I to DNA repair sites. Proc Natl Acad Sci U S A *102*, 8905-8909.

Moser, S. C., and Swedlow, J. R. (2011). How to be a mitotic chromosome. Chromosome Res *19*, 307-319.

Muchardt, C., and Yaniv, M. (1993). A human homologue of Saccharomyces cerevisiae SNF2/SWI2 and Drosophila brm genes potentiates transcriptional activation by the glucocorticoid receptor. Embo J *12*, 4279-4290.

Munson, K., Clark, J., Lamparska-Kupsik, K., and Smith, S. S. (2007). Recovery of bisulfite-converted genomic sequences in the methylation-sensitive QPCR. Nucleic Acids Res *35*, 2893-2903.

Munzel, M., Globisch, D., and Carell, T. (2011). 5-hydroxymethylcytosine, the sixth base of the genome. Angew Chem Int Ed Engl *50*, 6460-6468.

Murakami, T., Li, X., Gong, J., Bhatia, U., Traganos, F., and Darzynkiewicz, Z. (1995). Induction of apoptosis by 5-azacytidine: drug concentration-dependent differences in cell cycle specificity. Cancer Res *55*, 3093-3098.

Muthuswami, R., Truman, P. A., Mesner, L. D., and Hockensmith, J. W. (2000). A eukaryotic SWI2/SNF2 domain, an exquisite detector of double-stranded to single-stranded DNA transition elements. J Biol Chem *275*, 7648-7655.

Myant, K., Termanis, A., Sundaram, A. Y., Boe, T., Li, C., Merusi, C., Burrage, J., de Las Heras, J. I., and Stancheva, I. (2011). LSH and G9a/GLP complex are required for developmentally programmed DNA methylation. Genome Res *21*, 83-94.

Nan, X., Campoy, F. J., and Bird, A. (1997). MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. Cell *88*, 471-481.

Nan, X., Meehan, R. R., and Bird, A. (1993). Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2. Nucleic Acids Res *21*, 4886-4892.

Nan, X., Ng, H. H., Johnson, C. A., Laherty, C. D., Turner, B. M., Eisenman, R. N., and Bird, A. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature *393*, 386-389.

Nan, X., Tate, P., Li, E., and Bird, A. (1996). DNA methylation specifies chromosomal localization of MeCP2. Mol Cell Biol *16*, 414-421.

Nateri, A. S., Riera-Sans, L., Da Costa, C., and Behrens, A. (2004). The ubiquitin ligase SCFFbw7 antagonizes apoptotic JNK signaling. Science *303*, 1374-1378.

Nguyen, C. T., Weisenberger, D. J., Velicescu, M., Gonzales, F. A., Lin, J. C., Liang, G., and Jones, P. A. (2002). Histone H3-lysine 9 methylation is associated with aberrant gene silencing in cancer cells and is rapidly reversed by 5-aza-2'-deoxycytidine. Cancer Res *62*, 6456-6461.

Nicholls, R. D., Fischel-Ghodsian, N., and Higgs, D. R. (1987). Recombination at the human alpha-globin gene cluster: sequence features and topological constraints. Cell *49*, 369-378.

Nongkhlaw, M., Dutta, P., Hockensmith, J. W., Komath, S. S., and Muthuswami, R. (2009). Elucidating the mechanism of DNA-dependent ATP hydrolysis mediated by DNA-dependent ATPase A, a member of the SWI2/SNF2 protein family. Nucleic Acids Res *37*, 3332-3341.

Nozaki, T., Masutani, M., Sugimura, T., Takato, T., and Wakabayashi, K. (1997). Abrogation of G1 arrest after DNA damage is associated with constitutive overexpression of Mdm2, Cdk4, and Irf1 mRNAs in the BALB/c 3T3 A31 variant 1-1 clone. Biochem Biophys Res Commun *233*, 216-220.

Oficjalska-Pham, D., Harismendy, O., Smagowicz, W. J., Gonzalez de Peredo, A., Boguta, M., Sentenac, A., and Lefebvre, O. (2006). General repression of RNA polymerase III transcription is triggered by protein phosphatase type 2A-mediated dephosphorylation of Maf1. Mol Cell *22*, 623-632.

Okada, N. (1991a). SINEs. Curr Opin Genet Dev *1*, 498-504.

Okada, N. (1991b). SINEs: Short interspersed repeated elements of the eukaryotic genome. Trends Ecol Evol *6*, 358-361.

Okano, M., Bell, D. W., Haber, D. A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell *99*, 247-257.

Okano, M., Xie, S., and Li, E. (1998). Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells. Nucleic Acids Res *26*, 2536-2540.

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., and Cairns, B. R. (2010). Human RNA polymerase III transcriptomes and relationships to Pol II promoter chromatin and enhancer-binding factors. Nat Struct Mol Biol *17*, 620-628.

Oler, A. J., and Cairns, B. R. Personal communication. Unpublised Data.

Ooi, S. K., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument-Bromage, H., Tempst, P., Lin, S. P., Allis, C. D., Cheng, X., and Bestor, T. H. (2007). DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature *448*, 714-717.

Oruetxebarria, I., Venturini, F., Kekarainen, T., Houweling, A., Zuijderduijn, L. M., Mohd-Sarip, A., Vries, R. G., Hoeben, R. C., and Verrijzer, C. P. (2004). P16INK4a is required for hSNF5 chromatin remodeler-induced cellular senescence in malignant rhabdoid tumor cells. J Biol Chem *279*, 3807-3816.

Page, G. S., Smith, S., and Goodman, H. M. (1981). DNA sequence of the rat growth hormone gene: location of the 5' terminus of the growth hormone mRNA and identification of an internal transposon-like element. Nucleic Acids Res *9*, 2087-2104.

Pal-Bhadra, M., Leibovitch, B. A., Gandhi, S. G., Rao, M., Bhadra, U., Birchler, J. A., and Elgin, S. C. (2004). Heterochromatic silencing and HP1 localization in Drosophila are dependent on the RNAi machinery. Science *303*, 669-672.

Park, J., Wood, M. A., and Cole, M. D. (2002). BAF53 forms distinct nuclear complexes and functions as a critical c-Myc-interacting nuclear cofactor for oncogenic transformation. Mol Cell Biol *22*, 1307-1316.

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

Parthasarthy, A., and Gopinathan, K. P. (2005). Modulation of differential transcription of tRNA genes through chromatin organization. Biochem J *391*, 371-381.

Paulson, K. E., and Schmid, C. W. (1986). Transcriptional inactivity of Alu repeats in HeLa cells. Nucleic Acids Res *14*, 6145-6158.

Peterson, C. L., and Herskowitz, I. (1992). Characterization of the yeast SWI1, SWI2, and SWI3 genes, which encode a global activator of transcription. Cell *68*, 573-583.

Phalke, S., Nickel, O., Walluscheck, D., Hortig, F., Onorati, M. C., and Reuter, G. (2009). Retrotransposon silencing and telomere integrity in somatic cells of Drosophila depends on the cytosine-5 methyltransferase DNMT2. Nat Genet *41*, 696-702.

Phelan, M. L., Sif, S., Narlikar, G. J., and Kingston, R. E. (1999). Reconstitution of a core chromatin remodeling complex from SWI/SNF subunits. Mol Cell *3*, 247-253.

Phillips, W. F., and McConkey, E. H. (1976). Relative stoichiometry in ribosomal proteins in HeLa cell nucleoli. J Biol Chem *251*, 2876-2881.

Pieler, T., Appel, B., Oei, S. L., Mentzel, H., and Erdmann, V. A. (1985). Point mutational analysis of the Xenopus laevis 5S gene promoter. Embo J *4*, 1847-1853.

Pieler, T., Hamm, J., and Roeder, R. G. (1987). The 5S gene internal control region is composed of three distinct sequence elements, organized as two functional domains with variable spacing. Cell *48*, 91-100.

Prokhortchouk, A., Hendrich, B., Jorgensen, H., Ruzov, A., Wilm, M., Georgiev, G., Bird, A., and Prokhortchouk, E. (2001). The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor. Genes Dev *15*, 1613-1618.

Raab, J. R., Chiu, J., Zhu, J., Katzman, S., Kurukuti, S., Wade, P. A., Haussler, D., and Kamakaka, R. T. (2011). Human tRNA genes function as chromatin insulators. Embo J.

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.

Ramsahoye, B. H., Biniszkiewicz, D., Lyko, F., Clark, V., Bird, A. P., and Jaenisch, R. (2000). Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. Proc Natl Acad Sci U S A *97*, 5237-5242.

Randazzo, F. M., Khavari, P., Crabtree, G., Tamkun, J., and Rossant, J. (1994). brg1: a putative murine homologue of the Drosophila brahma gene, a homeotic gene regulator. Dev Biol *161*, 229-242.

Ream, T. S., Haag, J. R., Wierzbicki, A. T., Nicora, C. D., Norbeck, A. D., Zhu, J. K., Hagen, G., Guilfoyle, T. J., Pasa-Tolic, L., and Pikaard, C. S. (2009). Subunit compositions of the RNA-silencing enzymes Pol IV and Pol V reveal their origins as specialized forms of RNA polymerase II. Mol Cell *33*, 192-203.

Reisman, D., Glaros, S., and Thompson, E. A. (2009). The SWI/SNF complex and cancer. Oncogene *28*, 1653-1668.

Reyes, J. C., Barra, J., Muchardt, C., Camus, A., Babinet, C., and Yaniv, M. (1998). Altered control of cellular proliferation in the absence of mammalian brahma (SNF2alpha). Embo J *17*, 6979-6991.

Roberts, C. W., Leroux, M. M., Fleming, M. D., and Orkin, S. H. (2002). Highly penetrant, rapid tumorigenesis through conditional inversion of the tumor suppressor gene Snf5. Cancer Cell *2*, 415-425.

Roeder, R. G., and Rutter, W. J. (1969). Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. Nature *224*, 234-237.

Rosa, M. D., Gottlieb, E., Lerner, M. R., and Steitz, J. A. (1981). Striking similarities are exhibited by two small Epstein-Barr virus-encoded ribonucleic acids and the adenovirus-associated ribonucleic acids VAI and VAII. Mol Cell Biol *1*, 785-796.

Roy-Engel, A. M., Salem, A. H., Oyeniran, O. O., Deininger, L., Hedges, D. J., Kilroy, G. E., Batzer, M. A., and Deininger, P. L. (2002). Active Alu element "A-tails": size does matter. Genome Res *12*, 1333-1344.

Rubin, C. M., Houck, C. M., Deininger, P. L., Friedmann, T., and Schmid, C. W. (1980). Partial nucleotide sequence of the 300-nucleotide interspersed repeated human DNA sequences. Nature *284*, 372-374.

Rubin, C. M., Kimura, R. H., and Schmid, C. W. (2002). Selective stimulation of translational expression by Alu RNA. Nucleic Acids Res *30*, 3253-3261.

Rubin, C. M., VandeVoort, C. A., Teplitz, R. L., and Schmid, C. W. (1994). Alu repeated DNAs are differentially methylated in primate germ cells. Nucleic Acids Res *22*, 5121-5127.

Rudin, C. M., and Thompson, C. B. (2001). Transcriptional activation of short interspersed elements by DNA-damaging agents. Genes Chromosomes Cancer *30*, 64-71.

Rudloff, U., Eberhard, D., Tora, L., Stunnenberg, H., and Grummt, I. (1994). TBP-associated factors interact with DNA and govern species specificity of RNA polymerase I transcription. Embo J *13*, 2611-2616.

Russanova, V. R., Driscoll, C. T., and Howard, B. H. (1995). Adenovirus type 2 preferentially stimulates polymerase III transcription of Alu elements by relieving repression: a potential role for chromatin. Mol Cell Biol *15*, 4282-4290.

Russell, J., and Zomerdijk, J. C. (2006). The RNA polymerase I transcription machinery. Biochem Soc Symp, 203-216.

Ruzov, A., Dunican, D. S., Prokhortchouk, A., Pennings, S., Stancheva, I., Prokhortchouk, E., and Meehan, R. R. (2004). Kaiso is a genome-wide repressor of transcription that is essential for amphibian development. Development *131*, 6185-6194.

Ryme, J., Asp, P., Bohm, S., Cavellan, E., and Farrants, A. K. (2009). Variations in the composition of mammalian SWI/SNF chromatin remodelling complexes. J Cell Biochem *108*, 565-576.

Ryskov, A. P., Ivanov, P. L., Kramerov, D. A., and Georgiev, G. P. (1983). Mouse ubiquitous B2 repeat in polysomal and cytoplasmic poly(A)+RNAs: uniderectional orientation and 3'-end localization. Nucleic Acids Res *11*, 6541-6558.

Saksouk, N., Avvakumov, N., Champagne, K. S., Hung, T., Doyon, Y., Cayrou, C., Paquet, E., Ullah, M., Landry, A. J., Cote, V., Yang, X. J., Gozani, O., Kutateladze, T. G., and Cote, J. (2009). HBO1 HAT complexes target chromatin throughout gene coding regions via multiple PHD finger interactions with histone H3 tail. Mol Cell *33*, 257-265.

Santos, F., Hendrich, B., Reik, W., and Dean, W. (2002). Dynamic reprogramming of DNA methylation in the early mouse embryo. Dev Biol *241*, 172-182.

Sarge, K. D., Zimarino, V., Holm, K., Wu, C., and Morimoto, R. I. (1991). Cloning and characterization of two mouse heat shock factors with distinct inducible and constitutive DNA-binding ability. Genes Dev *5*, 1902-1911.

Sarraf, S. A., and Stancheva, I. (2004). Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly. Mol Cell *15*, 595-605.

Scarsdale, J. N., Webb, H. D., Ginder, G. D., and Williams, D. C., Jr. (2011). Solution structure and dynamic analysis of chicken MBD2 methyl binding domain bound to a target-methylated DNA sequence. Nucleic Acids Res.

Schmid, C. W. (1991). Human Alu subfamilies and their methylation revealed by blot hybridization. Nucleic Acids Res *19*, 5613-5617.

Schmid, C. W. (2003). Alu: a parasite's parasite? Nat Genet *35*, 15-16.

Schmid, C. W., and Deininger, P. L. (1975). Sequence organization of the human genome. Cell *6*, 345-358.

Schnell, R., and Rine, J. (1986). A position effect on the expression of a tRNA gene mediated by the SIR genes in Saccharomyces cerevisiae. Mol Cell Biol *6*, 494-501.

Schramm, L., and Hernandez, N. (2002). Recruitment of RNA polymerase III to its target promoters. Genes Dev *16*, 2593-2620.

Schramm, L., Pendergrast, P. S., Sun, Y., and Hernandez, N. (2000). Different human TFIIIB activities direct RNA polymerase III transcription from TATA-containing and TATA-less promoters. Genes Dev *14*, 2650-2663.

Schultz, P., Marzouki, N., Marck, C., Ruet, A., Oudet, P., and Sentenac, A. (1989). The two DNA-binding domains of yeast transcription factor tau as observed by scanning transmission electron microscopy. Embo J *8*, 3815-3824.

Schutz, E., Urnovitz, H. B., Iakoubov, L., Schulz-Schaeffer, W., Wemheuer, W., and Brenig, B. (2005). Bov-tA short interspersed nucleotide element sequences in circulating nucleic acids from sera of cattle with bovine spongiform encephalopathy (BSE) and sera of cattle exposed to BSE. Clin Diagn Lab Immunol *12*, 814-820.

Sethy-Coraci, I., Moir, R. D., Lopez-de-Leon, A., and Willis, I. M. (1998). A differential response of wild type and mutant promoters to TFIIIB70 overexpression in vivo and in vitro. Nucleic Acids Res *26*, 2344-2352.

Shamovsky, I., Ivannikov, M., Kandel, E. S., Gershon, D., and Nudler, E. (2006). RNA-mediated response to heat shock in mammalian cells. Nature *440*, 556-560.

Shamovsky, I., and Nudler, E. (2008). Modular RNA heats up. Mol Cell *29*, 415-417.

Shen, M. R., Batzer, M. A., and Deininger, P. L. (1991). Evolution of the master Alu gene(s). J Mol Evol *33*, 311-320.

Shen, Y., Igo, M., Yalamanchili, P., Berk, A. J., and Dasgupta, A. (1996). DNA binding domain and subunit interactions of transcription factor IIIC revealed by dissection with poliovirus 3C protease. Mol Cell Biol *16*, 4163-4171.

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J. R., Cole, P. A., Casero, R. A., and Shi, Y. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell *119*, 941-953.

Siegel, V., and Walter, P. (1988). Each of the activities of signal recognition particle (SRP) is contained within a distinct domain: analysis of biochemical mutants of SRP. Cell *52*, 39-49.

Simms, T. A., Dugas, S. L., Gremillion, J. C., Ibos, M. E., Dandurand, M. N., Toliver, T. T., Edwards, D. J., and Donze, D. (2008). TFIIIC binding sites function as both heterochromatin barriers and chromatin insulators in Saccharomyces cerevisiae. Eukaryot Cell *7*, 2078-2086.

Sims, R. J., 3rd, Chen, C. F., Santos-Rosa, H., Kouzarides, T., Patel, S. S., and Reinberg, D. (2005). Human but not yeast CHD1 binds directly and selectively to histone H3 methylated at lysine 4 via its tandem chromodomains. J Biol Chem *280*, 41789-41792.

Singh, K., Carey, M., Saragosti, S., and Botchan, M. (1985). Expression of enhanced levels of small RNA polymerase III transcripts encoded by the B2 repeats in simian virus 40-transformed mouse cells. Nature *314*, 553-556.

Sinn, E., Wang, Z., Kovelman, R., and Roeder, R. G. (1995). Cloning and characterization of a TFIIIC2 subunit (TFIIIC beta) whose presence correlates with activation of RNA polymerase III-mediated transcription by adenovirus E1A expression and serum factors. Genes Dev *9*, 675-685.

Sinnett, D., Richer, C., Deragon, J. M., and Labuda, D. (1991). Alu RNA secondary structure consists of two independent 7 SL RNA-like folding units. J Biol Chem *266*, 8675-8678.

Sinnett, D., Richer, C., Deragon, J. M., and Labuda, D. (1992). Alu RNA transcripts in human embryonal carcinoma cells. Model of post-transcriptional selection of master sequences. J Mol Biol *226*, 689-706.

Smalheiser, N. R., and Torvik, V. I. (2006). Alu elements within human mRNAs are probable microRNA targets. Trends Genet *22*, 532-536.

Smit, A. F. (1999). Interspersed repeats and other mementos of transposable elements in mammalian genomes. Curr Opin Genet Dev *9*, 657-663.

Smit, A. F., and Riggs, A. D. (1995). MIRs are classic, tRNA-derived SINEs that amplified before the mammalian radiation. Nucleic Acids Res *23*, 98-102.

Smith, C. L., Horowitz-Scherer, R., Flanagan, J. F., Woodcock, C. L., and Peterson, C. L. (2003). Structural analysis of the yeast SWI/SNF chromatin remodeling complex. Nat Struct Biol *10*, 141-145.

Smith, Z. D., Gu, H., Bock, C., Gnirke, A., and Meissner, A. (2009). High-throughput bisulfite sequencing in mammalian genomes. Methods *48*, 226-232.

So, C. W., Ma, Z. G., Price, C. M., Dong, S., Chen, S. J., Gu, L. J., So, C. K., Wiedemann, L. M., and Chan, L. C. (1997). MLL self fusion mediated by Alu repeat homologous recombination and prognosis of AML-M4/M5 subtypes. Cancer Res *57*, 117-122.

Sorek, R., Ast, G., and Graur, D. (2002). Alu-containing exons are alternatively spliced. Genome Res *12*, 1060-1067.

Sorek, R., Lev-Maor, G., Reznik, M., Dagan, T., Belinky, F., Graur, D., and Ast, G. (2004). Minimal conditions for exonization of intronic sequences: 5' splice site formation in alu exons. Mol Cell *14*, 221-231.

Sotillo, R., Schvartzman, J. M., Socci, N. D., and Benezra, R. (2010). Mad2-induced chromosome instability leads to lung tumour relapse after oncogene withdrawal. Nature *464*, 436-440.

Soubry, A., Staes, K., Parthoens, E., Noppen, S., Stove, C., Bogaert, P., van Hengel, J., and van Roy, F. (2010). The transcriptional repressor Kaiso localizes at the mitotic spindle and is a constituent of the pericentriolar material. PLoS One *5*, e9203.

Sterner, D. E., and Berger, S. L. (2000). Acetylation of histones and transcription-related factors. Microbiol Mol Biol Rev *64*, 435-459.

Stewart, M. D., Li, J., and Wong, J. (2005). Relationship between histone H3 lysine 9 methylation, transcription repression, and heterochromatin protein 1 recruitment. Mol Cell Biol *25*, 2525-2538.

Stunkel, W., Kober, I., and Seifart, K. H. (1997). A nucleosome positioned in the distal promoter region activates transcription of the human U6 gene. Mol Cell Biol *17*, 4397-4405.

Sudarsanam, P., Iyer, V. R., Brown, P. O., and Winston, F. (2000). Whole-genome expression analysis of snf/swi mutants of Saccharomyces cerevisiae. Proc Natl Acad Sci U S A *97*, 3364-3369.

Sueoka, N., Marmur, J., and Doty, P., 2nd (1959). Dependence of the density of deoxyribonucleic acids on guanine-cytosine content. Nature *183*, 1429-1431.

Suetake, I., Shinozaki, F., Miyagawa, J., Takeshima, H., and Tajima, S. (2004). DNMT3L stimulates the DNA methylation activity of Dnmt3a and Dnmt3b through a direct interaction. J Biol Chem *279*, 27816-27823.

Sugiyama, T., Cam, H., Verdel, A., Moazed, D., and Grewal, S. I. (2005). RNA-dependent RNA polymerase is an essential component of a self-enforcing loop coupling heterochromatin assembly to siRNA production. Proc Natl Acad Sci U S A *102*, 152-157.

Sutcliffe, J. E., Cairns, C. A., McLees, A., Allison, S. J., Tosh, K., and White, R. J. (1999). RNA polymerase III transcription factor IIIB is a target for repression by pocket proteins p107 and p130. Mol Cell Biol *19*, 4255-4261.

Szerlong, H., Hinata, K., Viswanathan, R., Erdjument-Bromage, H., Tempst, P., and Cairns, B. R. (2008). The HSA domain binds nuclear actin-related proteins to regulate chromatin-remodeling ATPases. Nat Struct Mol Biol *15*, 469-476.

Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L. M., Liu, D. R., Aravind, L., and Rao, A. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science *324*, 930-935.

Tamkun, J. W., Deuring, R., Scott, M. P., Kissinger, M., Pattatucci, A. M., Kaufman, T. C., and Kennison, J. A. (1992). brahma: a regulator of Drosophila homeotic genes structurally related to the yeast transcriptional activator SNF2/SWI2. Cell *68*, 561-572.

Tanaka, K., and Okamoto, A. (2007). Degradation of DNA by bisulfite treatment. Bioorg Med Chem Lett *17*, 1912-1915.

Tang, L., Nogales, E., and Ciferri, C. (2010). Structure and function of SWI/SNF chromatin remodeling complexes and mechanistic implications for transcription. Prog Biophys Mol Biol *102*, 122-128.

Thiagarajan, D., Dev, R. R., and Khosla, S. (2011). The DNA methyltranferase Dnmt2 participates in RNA processing during cellular stress. Epigenetics *6*, 103-113.

Thoma, F., Koller, T., and Klug, A. (1979). Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. J Cell Biol *83*, 403-427.

Thomson, J. P., Skene, P. J., Selfridge, J., Clouaire, T., Guy, J., Webb, S., Kerr, A. R., Deaton, A., Andrews, R., James, K. D., Turner, D. J., Illingworth, R., and Bird, A. (2010). CpG islands influence chromatin structure via the CpG-binding protein Cfp1. Nature *464*, 1082-1086.

Trotter, K. W., and Archer, T. K. (2008). The BRG1 transcriptional coregulator. Nucl Recept Signal *6*, e004.

Trouche, D., Le Chalony, C., Muchardt, C., Yaniv, M., and Kouzarides, T. (1997). RB and hbrm cooperate to repress the activation functions of E2F1. Proc Natl Acad Sci U S A *94*, 11268-11273.

Tsuji-Takayama, K., Inoue, T., Ijiri, Y., Otani, T., Motoda, R., Nakamura, S., and Orita, K. (2004). Demethylating agent, 5-azacytidine, reverses differentiation of embryonic stem cells. Biochem Biophys Res Commun *323*, 86-90.

Ullu, E., Murphy, S., and Melli, M. (1982). Human 7SL RNA consists of a 140 nucleotide middle-repetitive sequence inserted in an alu sequence. Cell *29*, 195-202.

Ullu, E., and Tschudi, C. (1984). Alu sequences are processed 7SL RNA genes. Nature *312*, 171-172.

Ullu, E., and Weiner, A. M. (1985). Upstream sequences modulate the internal promoter of the human 7SL RNA gene. Nature *318*, 371-374.

Upadhyay, A. K., and Cheng, X. (2011). Dynamics of histone lysine methylation: structures of methyl writers and erasers. Prog Drug Res *67*, 107-124.

Valadkhan, S. (2005). snRNAs as the catalysts of pre-mRNA splicing. Curr Opin Chem Biol *9*, 603-608.

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

van Zon, A., Mossink, M. H., Scheper, R. J., Sonneveld, P., and Wiemer, E. A. (2003). The vault complex. Cell Mol Life Sci *60*, 1828-1837.

Vaute, O., Nicolas, E., Vandel, L., and Trouche, D. (2002). Functional and physical interaction between the histone methyl transferase Suv39H1 and histone deacetylases. Nucleic Acids Res *30*, 475-481.

Veras, I., Rosen, E. M., and Schramm, L. (2009). Inhibition of RNA polymerase III transcription by BRCA1. J Mol Biol *387*, 523-531.

Versteege, I., Medjkane, S., Rouillard, D., and Delattre, O. (2002). A key role of the hSNF5/INI1 tumour suppressor in the control of the G1-S transition of the cell cycle. Oncogene *21*, 6403-6412.

Versteege, I., Sevenet, N., Lange, J., Rousseau-Merck, M. F., Ambros, P., Handgretinger, R., Aurias, A., and Delattre, O. (1998). Truncating mutations of hSNF5/INI1 in aggressive paediatric cancer. Nature *394*, 203-206.

Waalwijk, C., and Flavell, R. A. (1978). MspI, an isoschizomer of hpaII which cleaves both unmethylated and methylated hpaII sites. Nucleic Acids Res *5*, 3231-3236.

Wahlfors, J., Hiltunen, H., Heinonen, K., Hamalainen, E., Alhonen, L., and Janne, J. (1992). Genomic hypomethylation in human chronic lymphocytic leukemia. Blood *80*, 2074-2080.

Wallace, N., Wagstaff, B. J., Deininger, P. L., and Roy-Engel, A. M. (2008). LINE-1 ORF1 protein enhances Alu SINE retrotransposition. Gene *419*, 1-6.

Wang, H. D., Yuh, C. H., Dang, C. V., and Johnson, D. L. (1995). The hepatitis B virus X protein increases the cellular level of TATA-binding protein, which mediates transactivation of RNA polymerase III genes. Mol Cell Biol *15*, 6720-6728.

Wang, M., Gu, C., Qi, T., Tang, W., Wang, L., Wang, S., and Zeng, X. (2007). BAF53 interacts with p53 and functions in p53-mediated p21-gene transcription. J Biochem *142*, 613-620.

Wang, W., Cote, J., Xue, Y., Zhou, S., Khavari, P. A., Biggar, S. R., Muchardt, C., Kalpana, G. V., Goff, S. P., Yaniv, M., Workman, J. L., and Crabtree, G. R. (1996a). Purification and biochemical heterogeneity of the mammalian SWI-SNF complex. Embo J *15*, 5370-5382.

Wang, W., Xue, Y., Zhou, S., Kuo, A., Cairns, B. R., and Crabtree, G. R. (1996b). Diversity and specialization of mammalian SWI/SNF complexes. Genes Dev *10*, 2117-2130.

Waring, M., and Britten, R. J. (1966). Nucleotide sequence repetition: a rapidly reassociating fraction of mouse DNA. Science *154*, 791-794.

Waterston, R. H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J. F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., An, P., Antonarakis, S. E., Attwood, J., Baertsch, R., Bailey, J., Barlow, K., Beck, S., Berry, E., Birren, B., Bloom, T., Bork, P., Botcherby, M., Bray, N., Brent, M. R., Brown, D. G., Brown, S. D., Bult, C., Burton, J., Butler, J., Campbell, R. D., Carninci, P., Cawley, S., Chiaromonte, F., Chinwalla, A. T., Church, D. M., Clamp, M., Clee, C., Collins, F. S., Cook, L. L., Copley, R. R., Coulson, A., Couronne, O., Cuff, J., Curwen, V., Cutts, T., Daly, M., David, R., Davies, J., Delehaunty, K. D., Deri, J., Dermitzakis, E. T., Dewey, C., Dickens, N. J., Diekhans, M., Dodge, S., Dubchak, I., Dunn, D. M., Eddy, S. R., Elnitski, L., Emes, R. D., Eswara, P., Eyras, E., Felsenfeld, A., Fewell, G. A., Flicek, P., Foley, K., Frankel, W. N., Fulton, L. A., Fulton, R. S., Furey, T. S., Gage, D., Gibbs, R. A., Glusman, G., Gnerre, S., Goldman, N., Goodstadt, L., Grafham, D., Graves, T. A., Green, E. D., Gregory, S., Guigo, R., Guyer, M., Hardison, R. C., Haussler, D., Hayashizaki, Y., Hillier, L. W., Hinrichs, A., Hlavina, W., Holzer, T., Hsu, F., Hua, A., Hubbard, T., Hunt, A., Jackson, I., Jaffe, D. B., Johnson, L. S., Jones, M., Jones, T. A., Joy, A., Kamal, M., Karlsson, E. K., Karolchik, D., Kasprzyk, A., Kawai, J., Keibler, E., Kells, C., Kent, W. J., Kirby, A., Kolbe, D. L., Korf, I., Kucherlapati, R. S., Kulbokas, E. J., Kulp, D., Landers, T., Leger, J. P., Leonard, S., Letunic, I., Levine, R., Li, J., Li, M., Lloyd, C., Lucas, S., Ma, B., Maglott, D. R., Mardis, E. R., Matthews, L., Mauceli, E., Mayer, J. H., McCarthy, M., McCombie, W. R., McLaren, S., McLay, K., McPherson, J. D., Meldrim, J., Meredith, B., Mesirov, J. P., Miller, W., Miner, T. L., Mongin, E., Montgomery, K. T., Morgan, M., Mott, R., Mullikin, J. C., Muzny, D. M., Nash, W. E., Nelson, J. O., Nhan, M. N., Nicol, R., Ning, Z., Nusbaum, C., O'Connor, M. J., Okazaki, Y., Oliver, K., Overton-

Larty, E., Pachter, L., Parra, G., Pepin, K. H., Peterson, J., Pevzner, P., Plumb, R., Pohl, C. S., Poliakov, A., Ponce, T. C., Ponting, C. P., Potter, S., Quail, M., Reymond, A., Roe, B. A., Roskin, K. M., Rubin, E. M., Rust, A. G., Santos, R., Sapojnikov, V., Schultz, B., Schultz, J., Schwartz, M. S., Schwartz, S., Scott, C., Seaman, S., Searle, S., Sharpe, T., Sheridan, A., Shownkeen, R., Sims, S., Singer, J. B., Slater, G., Smit, A., Smith, D. R., Spencer, B., Stabenau, A., Stange-Thomann, N., Sugnet, C., Suyama, M., Tesler, G., Thompson, J., Torrents, D., Trevaskis, E., Tromp, J., Ucla, C., Ureta-Vidal, A., Vinson, J. P., Von Niederhausern, A. C., Wade, C. M., Wall, M., Weber, R. J., Weiss, R. B., Wendl, M. C., West, A. P., Wetterstrand, K., Wheeler, R., Whelan, S., Wierzbowski, J., Willey, D., Williams, S., Wilson, R. K., Winter, E., Worley, K. C., Wyman, D., Yang, S., Yang, S. P., Zdobnov, E. M., Zody, M. C., and Lander, E. S. (2002). Initial sequencing and comparative analysis of the mouse genome. Nature *420*, 520-562.

Watt, F., and Molloy, P. L. (1988). Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. Genes Dev *2*, 1136-1143.

Wei, W., Gilbert, N., Ooi, S. L., Lawler, J. F., Ostertag, E. M., Kazazian, H. H., Boeke, J. D., and Moran, J. V. (2001). Human L1 retrotransposition: cis preference versus trans complementation. Mol Cell Biol *21*, 1429-1439.

Weiner, A. M. (1980). An abundant cytoplasmic 7S RNA is complementary to the dominant interspersed middle repetitive DNA sequence family in the human genome. Cell *22*, 209-218.

Weiner, A. M., Deininger, P. L., and Efstratiadis, A. (1986). Nonviral retroposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. Annu Rev Biochem *55*, 631-661.

White, R. J. (2001). Gene transcription, mechanisms and control. Blackwell Science Ltd.

White, R. J. (2002). RNA polymerase III transcription, 3rd Edition. Landes Bioscience.

White, R. J. (2004). RNA polymerase III transcription and cancer. Oncogene *23*, 3208-3216.

White, R. J. (2008). RNA polymerases I and III, non-coding RNAs and cancer. Trends Genet *24*, 622-629.

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

Williams, B. R. (1999). PKR; a sentinel kinase for cellular stress. Oncogene *18*, 6112-6120.

Winter, A. G., Sourvinos, G., Allison, S. J., Tosh, K., Scott, P. H., Spandidos, D. A., and White, R. J. (2000). RNA polymerase III transcription factor TFIIIC2 is overexpressed in ovarian tumors. Proc Natl Acad Sci U S A *97*, 12619-12624.

Wittig, S., and Wittig, B. (1982). Function of a tRNA gene promoter depends on nucleosome position. Nature *297*, 31-38.

Wolffe, A. P., and Brown, D. D. (1988). Developmental regulation of two 5S ribosomal RNA genes. Science *241*, 1626-1632.

Wong, N., Lam, W. C., Lai, P. B., Pang, E., Lau, W. Y., and Johnson, P. J. (2001). Hypomethylation of chromosome 1 heterochromatin DNA correlates with q-arm copy gain in human hepatocellular carcinoma. Am J Pathol *159*, 465-471.

Wu, H., D'Alessio, A. C., Ito, S., Xia, K., Wang, Z., Cui, K., Zhao, K., Sun, Y. E., and Zhang, Y. (2011). Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. Nature *473*, 389-393.

Wu, J. R., Hurn, J., and Bonner, J. (1972). Size and distribution of the repetitive segments of the Drosophila genome. J Mol Biol *64*, 211-219.

Xu, G. L., Bestor, T. H., Bourc'his, D., Hsieh, C. L., Tommerup, N., Bugge, M., Hulten, M., Qu, X., Russo, J. J., and Viegas-Pequignot, E. (1999). Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. Nature *402*, 187-191.

Xue, Y., Canman, J. C., Lee, C. S., Nie, Z., Yang, D., Moreno, G. T., Young, M. K., Salmon, E. D., and Wang, W. (2000). The human SWI/SNF-B chromatin-remodeling complex is related to yeast rsc and localizes at kinetochores of mitotic chromosomes. Proc Natl Acad Sci U S A *97*, 13015-13020.

Yang, Z., Zhu, Q., Luo, K., and Zhou, Q. (2001). The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. Nature *414*, 317-322.

Yates, P. A., Burman, R. W., Mummaneni, P., Krussel, S., and Turker, M. S. (1999). Tandem B1 elements located in a mouse methylation center provide a target for de novo DNA methylation. J Biol Chem *274*, 36357-36361.

Yoshinaga, S. K., Boulanger, P. A., and Berk, A. J. (1987). Resolution of human transcription factor TFIIIC into two functional components. Proc Natl Acad Sci U S A *84*, 3585-3589.

Yoshinaga, S. K., L'Etoile, N. D., and Berk, A. J. (1989). Purification and characterization of transcription factor IIIC2. J Biol Chem *264*, 10726-10731.

Yu, F., Zingler, N., Schumann, G., and Stratling, W. H. (2001). Methyl-CpG-binding protein 2 represses LINE-1 expression and retrotransposition but not Alu transcription. Nucleic Acids Res *29*, 4493-4501.

Yulug, I. G., Yulug, A., and Fisher, E. M. (1995). The frequency and position of Alu repeats in cDNAs, as determined by database searching. Genomics *27*, 544-548.

Zhang, H. S., Gavin, M., Dahiya, A., Postigo, A. A., Ma, D., Luo, R. X., Harbour, J. W., and Dean, D. C. (2000). Exit from G1 and S phase of the cell cycle is regulated by repressor complexes containing HDAC-Rb-hSWI/SNF and Rb-hSWI/SNF. Cell *101*, 79-89.

Zhang, P. X., Wang, Y., Liu, Y., Jiang, G. Y., Li, Q. C., and Wang, E. H. (2011). p120-catenin isoform 3 regulates subcellular localization of Kaiso and promotes invasion in lung cancer cells via a phosphorylation-dependent mechanism. Int J Oncol *38*, 1625-1635.

Zhang, Y., Jurkowska, R., Soeroes, S., Rajavelu, A., Dhayalan, A., Bock, I., Rathert, P., Brandt, O., Reinhardt, R., Fischle, W., and Jeltsch, A. (2010). Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. Nucleic Acids Res *38*, 4246-4253.

Zhang, Y., LeRoy, G., Seelig, H. P., Lane, W. S., and Reinberg, D. (1998). The dermatomyositis-specific autoantigen Mi2 is a component of a complex

containing histone deacetylase and nucleosome remodeling activities. Cell *95*, 279-289.

Zhang, Y., Ng, H. H., Erdjument-Bromage, H., Tempst, P., Bird, A., and Reinberg, D. (1999). Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. Genes Dev *13*, 1924-1935.

Zhong, S., and Johnson, D. L. (2009). The JNKs differentially regulate RNA polymerase III transcription by coordinately modulating the expression of all TFIIIB subunits. Proc Natl Acad Sci U S A *106*, 12682-12687.

Zhu, J. K. (2009). Active DNA demethylation mediated by DNA glycosylases. Annu Rev Genet *43*, 143-166.