

**Epigenetic regulation of  
Mash1 expression**

by

Chiara Beretta

A thesis submitted to Imperial College London

for the degree of Master of Philosophy

MRC Clinical Sciences Centre

Imperial College School of Medicine

## Abstract

Mash1 is a proneural gene important for specifying the neural fate. The Mash1 locus undergoes specific epigenetic changes in ES cells following neural induction. These include the loss of repressive H3K27 trimethylation and acquisition of H3K9 acetylation at the promoter, switch to an early replication timing and repositioning of the locus away from the nuclear periphery. Here I examine the relationship between nuclear localization and gene expression during neural differentiation and the role of the neuronal repressor REST in silencing Mash1 expression in ES cells. Following neural induction of ES cells, I observed that relocation of the Mash1 locus occurs from day 4-6 whereas overt expression begins at day 6. Mash1 expression was unaffected by REST removal in ES cells as well as the locus localization at the nuclear periphery. In contrast bona fide REST target genes were upregulated in REST  $-/-$  cells. Interestingly, among REST targets, loci that were more derepressed upon REST removal showed an interior location (Sthatmin, Synaptophysin), while those more resistant to REST withdrawal, showed a peripheral location (BDNF, Calbindin, Complexin). To ask whether the insulator protein CTCF together with the cohesin complex might be involved in regulating Mash1 in ES cells, I performed ChIP analysis of CTCF and cohesin binding across the Mash1 locus in ES cells and used RNAi to deplete CTCF and cohesin expression. A slight increase in the transcription of Mash1 was seen in cells upon Rad21 knock down, although it was not possible to exclude this was a consequence of delayed cell cycle progression. Finally ES cell lines that carried a Mash1 transgene were created as a tool to look at whether activation of Mash1 can affect the epigenetic properties of neighbouring genes.

## **Acknowledgements**

I would like to thank Mandy for giving me the possibility to work on this project and for supervising me with criticism and encouragement during the past few years. Thank you to Mandy and Matthias for making the Lymphocyte Development Group an exciting group where to have stimulating conversations and fun at the same time.

I would like to thank you all the members, past and present, of the group, in particular Ruth, Sara, Helle, Mikhail, Maria, Ben and Hegias for the help received all over this years. Thanks you to my friends and family for their support and love.

## Table of contents

<b>Abstract</b>	<b>2</b>
<b>Acknowledgements</b>	<b>3</b>
<b>Table of contents</b>	<b>4</b>
<b>List of figures</b>	<b>8</b>
<b>Abbreviation</b>	<b>9</b>
<b>Chapter 1. Introduction</b>	<b>12</b>
<u>1.1. Epigenetic aspects of gene regulation</u>	12
1.1.1 Chromatin and histone modifications	12
<i>Histone acetylation</i>	13
<i>Histone methylation</i>	14
<i>Histone variants</i>	15
<i>Replication timing</i>	15
1.1.2 Higher order chromatin organisation	16
1.1.3 The nuclear periphery	18
<i>The nuclear periphery as a repressive compartment</i>	18
<i>Molecular mechanism underlying repression at the nuclear edge</i>	20
<i>Gene activity at the Nuclear Pore Complex</i>	20
<u>1.2. Mash1 and the specification of a neural fate</u>	21
1.2.1 Proneural genes and models of mammalian neurogenesis	21
1.2.2 In vitro neural differentiation of ES cells	22
1.2.3 Mash1	24
1.2.4 Role of REST in silencing Mash1 and neuronal genes in ES cells	26

<i>REST silencing of neuronal genes in non neural cells</i>	26
<i>Role of REST in ES cells and neural differentiation</i>	26
<b><u>1.3 Role of CTCF and Cohesin in gene regulation</u></b>	<b>28</b>
1.3.1 The cohesin complex and associated proteins	28
<i>Cohesin recruitment to chromatin</i>	28
<i>Establishment of cohesion during S phase</i>	29
<i>Cohesin dissociation from the chromatin</i>	29
1.3.2 Evidence of cohesin role in gene regulation	30
1.3.3 Cohesin binding to chromosomal arms	30
1.3.4 Role of CTCF and cohesin in regulating gene expression	31
<b><u>1.4 Aims of the study</u></b>	<b>32</b>
<b>Chapter 2. Mash1 regulation in ES cells</b>	<b>34</b>
<b><u>2.1. Introduction</u></b>	<b>34</b>
<b><u>2.2. Mash1 expression and nuclear location in ES cells following neural induction</u></b>	<b>34</b>
2.2.1. In vitro neural differentiation of embryonic stem cells	35
2.2.2. 2D FISH analysis of Mash1 nuclear position in ES cells before and after neural induction	37
<b><u>2.3. Potential role of the REST protein in regulating Mash1 in ES cells</u></b>	<b>39</b>
2.3.1. REST is not require to maintain the peripheral localisation of the Mash1 locus in ES cells	39
2.3.2. The epigenetic state of Mash1 is unchanged in REST deficient cells	42
<b><u>2.4. Role of the REST protein in regulating neuronal genes in ES cells</u></b>	<b>42</b>

<b>Chapter 3. Role of CTCF and cohesin in the regulation of Mash1</b>	<b>48</b>
<u>3.1. Introduction</u>	48
<u>3.2. CTCF and cohesin binding at the Mash1 locus in ES cells</u>	48
<u>3.3. Analysis of Mash1 regulation by siRNA of CTCF and cohesin in ES cells</u>	52
3.3.1 Gene expression profile at the Mash1 locus upon neural differentiation	52
3.3.2. Mash1 expression profile in cells depleted of CTCF and cohesins	54
<b>Chapter 4. Generating ES cells transgenic for Mash1 locus</b>	<b>59</b>
<u>4.1. Introduction</u>	59
<u>4.2. Creation of ES cell lines carrying low copy number Mash1</u>	59
4.2.1 Screening of positive clones	61
<b>Chapter 5. General discussion</b>	<b>63</b>
<b>Chapter 6. Material and Methods</b>	<b>69</b>
6.1 Cell culture	69
6.2 Neural differentiation	69
6.3 Preparation of primary embryonic fibroblast (PEFs) for ES cell culture	69
6.4 Immunofluorescence staining	70
6.5 RT-PCR and Real Time PCR	70
6.6 Preparation of probes for fluorescence in situ hybridisation (FISH) and metaphase spread preparation	73
6.7 Fluorescence in situ hybridisation (FISH)	74
6.8 Microscopy and Measurements	75
6.9 CTCF and Rad21 knock down in ES cells	75

6.10 Western blot analysis	77
6.11 Mash1 BAC electroporation in ES cells	77
<b>References</b>	<b>79</b>

## List of figures

1.1. In vitro systems used to promote neural differentiation of ES cells.	23
1.2. Summary of the changes in the replication timing and nuclear positioning of the Mash1 locus observed upon neural induction.	25
2.1. Kinetics of MR-7 neural differentiation in co-culture with PA6 stromal cell line.	36
2.2. Kinetics of Mash1 induction and locus repositioning in the MR-7 neural differentiation system.	38
2.3. Mash1 nuclear localisation is unchanged in REST deficient cells.	41
2.4. Analysis of nuclear localisation and gene expression of REST targets in REST deficient cells.	44
S.1. Analysis of Mash1 epigenetic features in REST knock out cells	46
3.1. CTCF and Rad21 binding within a 1.2 Mb region surrounding the Mash1 locus.	50
3.2. Binding of CTCF and Rad21 at specific sites within the Mash1 locus in ES cells.	51
3.3. Gene expression profile of genes surrounding the Mash1 locus in ES cells and ES-derived NSC.	53
3.4. Western Blot analysis of CTCF and Rad21 knock down in ES cells.	54
3.5. Gene expression analysis of Mash1 after CTCF and Rad21 knock down in ES cells.	56-57
4.1. Studying the effect that an extra copy of the Mash1 has on the neighbouring genes.	60
4.2. Preparation of a Mash1 transgene to engineer ES cells.	61
4.3. Screening of Mash1-Neo positive clones by PCR.	62



## Abbreviations

3C	chromosome conformation capture assay
AT	adenine-thymine
ATP	adenosine tri-phosphate
bp	base pair
BSA	bovine serum albumin
cDNA	complementary DNA
ChIP	chromatin immunoprecipitation
ChIP-ChIP	chromatin immunoprecipitation on chip
ChIP-Seq	chromatin immunoprecipitation sequencing
Chr	chromosome
CpG	cytosine-guanine dinucleotide
C(T)	threshold cycle
CTCF	CCCTC binding factor
DAPI	4',6-diamino-2-phenylindole
DBS	double strand break
DMEM	Dulbecco's modified Eagle's medium
DNA	deoxyribonucleic acid
DHS	DNase 1 hypersensitive site
dsRNA	double strand RNA
<i>E. coli</i>	<i>Escherichia coli</i>
EGF	epidermal growth factor
EGFP	enhanced green fluorescence protein
EGTA	ethylene glycol tetraacetic acid
ES	embryonic stem (cells)
FACS	fluorescence-activated cell sorter
FCS	foetal calf serum
FISH	fluorescence <i>in situ</i> hybridisation

FITC	fluorescein isothiocyanate
FGF	fibroblast growth factor
(i)FRAP	fluorescence recovery after photobleaching
g	gram
GC	guanine-cytosine
GFP	green fluorescence protein
h	hour
H2AK119	ubiquitination of lysine 119 of histone H2A
H3K27me3	trimethylation of lysine 27 histone H3
H3K4me2	dimethylation of lysine 4 histone H3
H3K4me3	trimethylation of lysine 4 histone H3
H3K9ac	acetylation of lysine 9 of histone H3
H3K9me3	trimethylation of lysine 9 of histone H3
HAT	histone acetyl-transferase
HDAC	histone deacetylase
HMT	histone methyl-transferase
ICM	inner cell mass
IF	immunofluorescence
kb	kilobase pairs
KCl	potassium chloride
KH <sub>2</sub> PO <sub>4</sub>	potassium dihydrogen phosphate
KO	knock out
KSR	knockout serum replacement
L	litre
LADs	lamina associated domains
LIF	leukaemia inhibitory factor
m	milli
M	molar
Mb	megabase pairs
MgCl <sub>2</sub>	magnesium chloride
μ	micro

min	minutes
mRNA	messenger RNA
NaCl	sodium chloride
NP	neural progenitors
NS	neural stem cells
O/N	over night
PBS	phosphate buffered saline
PcG	Polycomb group
PCR	polymerase chain reaction
PEF	primary embryonic fibroblast
qPCR	real time PCR
RNA	ribonucleic acid
RNAi	RNA interference
rpm	rotation per minute
RT	room temperature
<i>S. Cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
<i>S. Pombe</i>	<i>Schizosaccharomyces pombe</i>
siRNA	small interfering RNA
Tris	tris(hydroxymethyl)aminomethane
TSS	transcription start site
U	units

# Chapter 1

## Introduction

### 1.1. Epigenetic aspects of gene regulation

#### 1.1.1 Chromatin and histone modifications

Chromatin is composed of DNA and associated proteins such as histones and regulatory proteins. The basic unit of chromatin is the nucleosome that consists of 146bp DNA wrapped around an octamer of four histones. The histone core is formed by two copies of each histone H2A, H2B, H3 and H4 organised in dimers H2A-H2B and H3-H4. The nucleosomal array visible by electron microscopy has been described as 'beads on a string', where beads are represented by histones and the string being the linker DNA. Each nucleosome is then linked by DNA of varying length decorated by external histones called linker H1 and H5 which can further compact the nucleosome array in a higher order chromatin conformation (Kornberg and Lorch 1999).

Histones are basic proteins with a conserved histone fold necessary for histone-histone interactions and an amino-terminal tail that lies outside the nucleosome. These amino-terminal domains are lysine-rich and can be modified post-translationally.

These modifications, which include acetylation, methylation, phosphorylation, sumoylation, ubiquitination and ADP-ribosylation, play a role in the structural organisation of the chromatin and influence gene expression (Santos-Rosa and Caldas 2005; Kouzarides 2007). One consequence of modifications is to recruit additional proteins that play a role in the regulation of chromatin configuration (for instance, HP1 to methylated H3K9, PRC1 to methylated H3K27 and bromodomain proteins to acetylated residues). Another, more structural role has been implicated for some modifications, such as acetylation, making the nucleosomal array more loosely packed and hence the DNA more accessible to transcription factors. An interplay between the different modifications have been described and the histone code hypothesis postulates that a combination of specific post-translation modifications create a particular chromatin state that results in either gene activity or silencing (Jenuwein and Allis 2001).

Different forms of chromatin are present in the nucleus, characterised by a different level of compaction of the chromatin fibre. Euchromatin describes regions where the chromatin fibre is decondensed with loosely or irregularly packed nucleosome. Euchromatin is generally rich in genes

that are highly transcribed or ready to be expressed (Arney and Fisher 2004). Epigenetic features of actively transcribed genes are H3K9 acetylation, H3K4 mono, di and trimethylation in the promoter regions and H3K36 trimethylation in the transcribed regions. Also the monomethylation of H3K27, H3K9, H4K20, H3K79 and H2BK5 are linked to gene activation (Bernstein, Kamal et al. 2005; Barski, Cuddapah et al. 2007). The histone variant H2A.Z is found associated to euchromatic region and may be important in 'poising' genes for transcription (Creighton, Markoulaki et al. 2008).

Heterochromatin is characterised by regularly packed nucleosomes and manifests as bright regions in the nuclei after staining with DNA binding dyes such as 4',6-diamino-2-phenylindole (DAPI). Two different types of heterochromatin are present in the nucleus. Constitutive heterochromatin is present in all cell types and is composed of gene-poor regions and repetitive DNA. This includes centromeric repeats that are characterised by trimethylation of H3K9/H4K20 and the presence of the heterochromatin-associated protein HP1. Facultative heterochromatin, on the other end, is made of DNA that is heterochromatic in some cell types but actively transcribed in others. One example is the X chromosome which is randomly inactivated in females and can be seen as the barr body, whereas other regions of facultative heterochromatin are not discernible by DAPI staining (Arney and Fisher 2004). Epigenetic modifications of repressed genes include trimethylation of H3K27, H3K9 and H3K79 (Barski, Cuddapah et al. 2007). Recently, developmentally regulated genes that are important for the early specification of embryonic development were demonstrated to carry the opposite modifications H3K4 trimethylation and H3K27 trimethylation (Azua, Perry et al. 2006; Bernstein, Mikkelsen et al. 2006).

### *Histone acetylation*

The promoter regions of active genes is characterised by low nucleosome occupancy and histone acetylation ((Kurdistani, Tavazoie et al. 2004; Pokholok, Harbison et al. 2005; Sekinger, Moqtaderi et al. 2005). Histone acetylation is found at several residues of the four core histones and it is proposed to affect gene expression by creating a more open chromatin structure. The addition of an acetyl group prevents protonation of lysine residues thereby decreasing the affinity of histones for DNA, making nucleosome more loosely packed and the DNA more accessible to transcription factors (Hong, Schroth et al. 1993). In addition the acetyl group can act as a bridge for the binding of SWI/SNF chromatin remodelling complexes that are involve in gene activation (Loyola and Almouzni 2004). Histone acetylation is mediated by histone acetyltransferase (HAT) that transfer acetyl groups from acetyl-coenzymeA to lysine residues. Transcriptional activators such as Gcn5/PCAF, CBP/p300, SRC-1 has HAT activity. Histone acetylation can be reverse by the activity of histone deacetylases (HDAC's). Many transcriptional corepressors such as mSin3a, NcoR/SMRT, NURD/Mi-2 contains subunits with HDAC activity (Bhaumik, Smith et al. 2007). The constant interplay between histone

acetyltransferase complexes (HAT) and histone deacetylase complexes (HDAC) regulate the cellular histone acetylation level (Wade and Wolffe 1997).

### *Histone methylation*

Histones 3 and histone 4 can be methylated at arginine or lysine residues. Arginine can be mono- or dimethylated while lysines can be mono-, di- or trimethylated. Histone methylation is associated both with silencing and activation of chromatin. The transcriptional outcome is dependent upon the site and type of residue methylated. Trimethylation of lysine H3K4, H3K36, H3K79 is associated with gene activity while di-/trimethylation of lysine H3K9, H4K20, H3K27 with gene silencing (Peters, Kubicek et al. 2003). Methylation of arginine also correlates with gene expression and activation or repression of transcription varies according to the type of residue modified and the modification introduced. Arginine methylation is catalysed by PMRT enzymes (Pal and Sif 2007)

The effect of histone methylation in regulating the chromatin fibre is thought to involve the recruitment of additional proteins and spreading of the repressive marks (Daniel, Pray-Grant et al. 2005).

Almost all the lysine methyl transferases characterised so far contain a SET domain and are named after the *D. Melanogaster* proteins Su(var)3-9, Enhancer of Zeste (E(z)), Trithorax (trx). Su(var)3-9 and its mammalian homologues is responsible for methylation of histone 3 lysine-9 (H3K9). HP1 heterochromatic protein is recruited to sites of H3K9 methylation and act as a bridge for the recruitment of other molecules of Suv39 and compaction of the chromatin fibre (Bannister, Zegerman et al. 2001).

The polycomb group proteins are responsible for gene silencing mediated by histone 3 lysine-27 (H3K27) methylation (Ringrose, Ehret et al. 2004). PcG proteins were first discovered in *Drosophila* as repressors of homeotic genes that are responsible for embryo segmentation (Lewis 1978). Two different polycomb complexes modify the chromatin structure to mediate silencing. The PRC2 complex, made of EZH2, SUZ12 and EED, catalyses the methylation of histone H3 lysine-27 (H3K27) with the EZH2 component being the catalytic subunit. The other core components SUZ12/EED are necessary for this function (Czermin, Melfi et al. 2002; Muller, Hart et al. 2002). Methylated H3K27 is recognised by the PRC1 complex that facilitates condensation of the chromatin fibre and inhibit chromatin remodelling complexes to favour silencing (Cao, Wang et al. 2002; Wang, Wang et al. 2004). Additionally, the core subunit Ring1A/B of the PRC1 complex catalyses the ubiquitination of lysine 119 of H2A, and it is thought to prevent active transcription of some developmentally regulated genes in ES cells (Guenther, Levine et al. 2007; Stock, Giadrossi et al. 2007).

Trimethylation of H3K4 is found at the 5' UTR of most of the transcribed genes in different species and it is catalysed by the Tritorax/MLL complex. H3K4 methylation can recruit transcriptional activators and remodelling complexes, such as histone acetyltransferases and the NURF complex (Ruthenburg, Allis et al. 2007).

Methylation was considered a stable modification until the discovery of histone lysine demethylases. Two families of enzymes have been shown to have demethylase activity, amine oxidase such as LSD1 and hydroxylases of the JmjC type. LSD1 was shown to demethylate mono- and dimethylation while JumonjiC domain-containing proteins were shown to catalyse the removal also of trimethylation, making methylation a fully reversible modification (Bhaumik, Smith et al. 2007).

### *Histone variants*

Another epigenetic modification of chromatin that has an effect on gene expression is the incorporation of histone variants to replace canonical histones. The histone variant H3.3 is associated with active genes and replaces the H3 histone. Rather than being incorporated only at the time DNA replicate, H3.3 can be replaced throughout the cell cycle (Mito, Henikoff et al. 2005). The histone variant H2A.Z is associated with transcribed genes in yeast and functions in part to prevent spreading of heterochromatin in euchromatic regions (Meneghini, Wu et al. 2003). H2A.Z is enriched at genes repositioned at the nuclear periphery upon transcription together with SWI/SNF chromatin remodelling complexes. Localisation at the nuclear periphery is maintained when genes become repressed and helps the genes to reactivate more easily (Brickner, Cajigas et al. 2007). In mammalian system H2A.Z association to chromatin is more complex and recent findings demonstrate its association with both euchromatin and facultative heterochromatin. Monoubiquitylation marks the association of H2A.Z with the different chromatin states and in particular its role in compaction of the chromatin fibre (Sarcinella, Zuzarte et al. 2007). Genome-wide studies show that H2A.Z occupies the promoters of developmentally important genes targeted also by the Polycomb group. Depletion of H2A.Z affects the differentiation of ES cells and suggests a role in establishment of an ES chromatin state necessary for the establishment of differentiation pathway (Creyghton, Markoulaki et al. 2008).

### *Replication Timing*

The timing at which a locus replicates has also been considered important for maintaining expression of genes through cell division (Gilbert 2002). Replication timing has been demonstrated to broadly correlate with the chromatin status of genes: constitutive heterochromatin and some facultative heterochromatin are late replicating, whereas transcriptionally active euchromatic regions are generally early replicating (Schubeler, Scalzo et al. 2002). Replication timing analysis has been used to study the chromatin profile of genes during differentiation and a genome wide analysis has

highlighted the relationship between switches in replication timing and genomic context (Hiratani, Leskovar et al. 2004; Perry, Sauer et al. 2004).

### **1.1.2 Higher order chromatin organisation**

The interphase nucleus is highly compartmentalized, with chromosomes occupying discrete positions and regulatory proteins organised in nuclear organelles or interspersed between chromosome territories (Kumaran, Thakar et al. 2008). The position occupied by specific chromosomes in the nucleus is not random, rather, radial positioning of chromosome relative to the nuclear periphery has been observed in different cell types. Gene-poor chromosomes localize closer to the nuclear edge while gene-rich chromosome have a more internal location (Croft, Bridger et al. 1999; Boyle, Gilchrist et al. 2001; Cremer and Cremer 2001). Also within the chromosome territories (CT), there is a radial distribution of chromatin domains relative to their CT with gene-poor chromatin located towards the nuclear periphery and gene-rich domains more interior (Bolzer, Kreth et al. 2005; Kupper, Kolbl et al. 2007).

Chromatin is not constrained within the chromosome territory but looping out of genes from the chromosome territory has been described in relation to gene activation (Volpi, Chevret et al. 2000; Williams, Broad et al. 2002). This looping out has been interpreted as either a way to protect genes within the loop from the spreading of silenced chromatin that resides outside the loop, or to enhance local concentrations of the transcriptional machinery necessary for gene expression (Chambeyron and Bickmore 2004; Osborne, Chakalova et al. 2004). However, looping out is not always observed and not all the active genes loop out their chromosome territories (Morey, Da Silva et al. 2007). Also the intermingling of different chromosomes has been related to change in the gene expression (Branco and Pombo 2006).

Chromatin movement is observed in relation to specific nuclear organelles or the nuclear periphery and it has been related to the transcriptional status of genes. Examples of nuclear bodies are nuclear speckles that are enriched of pre-mRNA splicing factors, Cajal bodies, promyelocytic leukemia (PML) bodies, PcG bodies, and the perinucleolar compartment (Spector 2003). Gene recruitment to centromeric heterochromatin has been associated with permanent silencing and allele recombination (Brown, Guest et al. 1997; Skok, Brown et al. 2001; Kosak, Skok et al. 2002).

Upon activation different genes can loop out of their chromatin territories to share the same transcription factory (Fraser and Bickmore 2007). In erythroid progenitor cells active genes present on different chromosomes have been shown to localise at sites of RNA polymerase and associated regulatory factors (Osborne, Chakalova et al. 2004). However gene association is not driven by



transcription factories but the frequency of association between genes is influenced by the regional genomic context and transcriptional status. Association is a stochastic event and it occurs more frequently for those genes located in decondensed chromatin stretches that cluster around common nuclear speckles (Brown, Green et al. 2008).

Different examples of chromatin association between loci on the same chromosome (in cis) and loci belonging to different chromosomes (in trans) have been described. These long-range intra and inter-chromosomal interactions are thought to put in close proximity genes and regulatory elements located far away (Williams, Spilianakis et al. 2010). A well documented example of promoter-enhancer communication in cis is the  $\beta$ -globin locus where 3C technology has been used to study the association between hypersensitive sites of the LCR regulatory region and active genes. DNA looping creates an active chromatin domain that facilitate the clustering of transcription factors and high level of transcription (Tolhuis, Palstra et al. 2002).

One of the first reported example of interchromosomal interactions is the association between the regulatory regions of the Th2 cytokine locus on chromosome 11 and the promoter of the *Ifng* gene on chromosome 10 (Spilianakis, Lalioti et al. 2005). This interaction is seen only in naive T cells, where none of the genes are expressed and it is lost during the differentiation in TH1 or TH2 cells with the expression of interleukin genes in TH2 cells and interferon gamma in TH1 cells. Association is thought to increase the efficiency of transcription and to poise the two classes of genes for immediate expression (Spilianakis, Lalioti et al. 2005).

Interchromosomal interaction has also been implicated in the regulation of imprinted genes. Association between the imprinting control region (ICR) of the *Igf2/H19* locus on the maternal chromosome 7 and the *Wsb1/Nsf1* locus on paternal chromosome 11 has been described. This interaction is dependent on the presence of the maternal ICR on chromosome 7 and it seems to be mediated by the CTCF protein. Abrogation of this interaction by knock down of CTCF or deletion of the ICR reduces the expression of *Wsb1* and *Nsf1* on chromosome 11, suggesting a role for this interaction in the interchromosomal gene regulation (Ling, Li et al. 2006). However it is important to remember that CTCF is able to interact and to recruit Pol II at CTCF binding site (Chernukhin, Shamsuddin et al. 2007). It is possible that knock down of CTCF may affect gene transcription and therefore the reduction in the interactions might be just a consequence of a decrease in transcription (Williams, Spilianakis et al.). More than 100 chromosomal fragments have been demonstrated to associate with the maternal allele of the H19 ICR, including many imprinted regions. In general, association of these regions is dependent on the presence of intact CTCF sites at the ICR, although its role is not completely clear (Zhao, Tavoosidana et al. 2006).

In the case of the 1300 olfactory receptor genes, the choice of expression of a specific olfactory gene is functionally dependent on an interaction with the enhancer element H on chromosome 14. The H element stochastically interact with one of the receptor genes located on the same chromosome or with the promoter of olfactory receptor genes located on different chromosomes resulting in the exclusive expression of that receptor (Lomvardas, Barnea et al. 2006). However following studies reveal that abrogation of the H element influence gene expression only of those genes located in cis. Moreover, in heterozygous mice the H- phenotype is not rescued by the presence of an intact H allele in trans (Fuss, Omura et al. 2007; Nishizumi, Kumasaka et al. 2007). It is therefore possible that additional enhancer element exist interspersed in the genome (Williams, Spilianakis et al. 2010).

Gene association is not necessarily linked to interchromosomal gene regulation. Particular interesting is the association between human  $\alpha$ -globin and  $\beta$ -globin. Although it is tempting to speculate that this association is functional to gene regulation, different evidences suggest it is more likely to be a consequence of sharing common resources (Williams, Spilianakis et al. 2010). In the case of co-transcribed erythroid genes, association seems in fact to be stochastic and determined by the sharing of common SC35-enriched splicing speckles (Brown, Green et al. 2008).

Finally association at the X inactivation centre (XIC) has been described for the two X chromosomes before silencing of one of the two chromosomes. Association has been linked to the counting of the two chromosomes before stochastically inactivation of one of the two X (Bacher, Guggiari et al. 2006). Association is however not essential for the X inactivation per se and seems to be a consequence of transcriptional activation of Xist and related genes at the XIC that may result in the pairing (Barakat, Jonkers et al.).

### **1.1.3 The nuclear periphery**

#### *The nuclear periphery as a repressive compartment*

From yeast to mammals the nuclear periphery has been generally referred to as a repressive compartment. In yeast silent regions such as the silenced mating type loci and telomeres are associated with the nuclear periphery (Andrulis, Neiman et al. 1998; Feuerbach, Galy et al. 2002). In the mammalian system inactive heterochromatin is visible as electron-dense material close to the nuclear periphery and around the nucleolus in electron microscopy studies.

Experiments of RNA FISH show that the nuclear periphery is generally depleted of transcribed genes and transcribed genes are found in the interior of the nucleus (Kosak, Scalzo et al. 2007; Levsky, Shenoy et al. 2007). Analysis of the regions interacting with the nuclear periphery by an approach called DamID revealed that these are associated with epigenetic features characteristic of repressed

chromatin. Lamina interacting domains (LADs) are in fact generally large regions that are gene poor, late replicating and depleted of active histone marks and RNA Pol II. The borders of the LADs contain CpG islands and CTCF binding site that marks, and are thought to be important for, the transition between active and inactive chromatin domains. These chromatin features are conserved between *Drosophila* and Humans (Pickersgill, Kalverda et al. 2006; Guelen, Pagie et al. 2008).

Repositioning of loci from the nuclear periphery has been correlated in some cases with gene activation and chromatin movements has been seen following cell differentiation (Pickersgill, Kalverda et al. 2006). IgH loci are usually associated with the nuclear periphery in hematopoietic progenitors and pro-T cells where they are not expressed, but relocate to the interior of the nucleus in pro-B cells. Moving away from the nuclear periphery may allow rearrangement of the IgH loci during B cell commitment (Kosak, Skok et al. 2002). Also the CFTR and associated genes are inactive when associated at the nuclear periphery and relocate into the nuclear interior when active (Zink, Amaral et al. 2004). However changes in the gene expression and nuclear localisation are not always associated. During T cell differentiation *Ifng* does not change its localisation at the nuclear periphery when expressed although recruitment of the relevant genes at the nuclear periphery is associated with silencing of cytokine regulators (Hewitt, High et al. 2004). The *Mash1* locus, relocates away from the nuclear periphery upon neural differentiation thereby driving the relocation of all the genes present in a 2Mb genomic context, but not all the genes that relocate to the nuclear interior are transcribed. (Williams, Azuara et al. 2006). Finally, although the beta-globin locus relocates away the nuclear periphery during erythroid maturation, transcription is visible while the locus is still located at the periphery, and notably increases when it relocate to the interior (Ragoczy, Bender et al. 2006).

To understand the relationship between gene expression and the nuclear periphery, different groups have artificially tethered genes to the nuclear periphery. Fusion proteins between the Lac repressor and integral nuclear membrane proteins or lamin B were used to target LacO sites integrated into human or mouse chromosome to the nuclear membrane. This targeting at the nuclear membrane was dependent on breakdown and subsequent formation of the nuclear envelope and in the cases where gene repression was observed, silencing did not involve the recruitment to heterochromatic regions. The conclusion of these experiments was that generally tethering causes silencing of the targeted genes, with different degrees of repression according to the targeted site but that transcription is not incompatible with the repositioning at the nuclear periphery. (Finlan, Sproul et al. 2008; Kumaran and Spector 2008; Reddy, Zullo et al. 2008).

As mentioned above, LADs are often late replicating (Guelen, Pagie et al. 2008). Genome wide analysis of the replication timing during differentiation of embryonic stem cells into neural progenitors show that changes in the replication timing of large megabase-size domains are accompanied by their nuclear repositioning relative to the nuclear periphery. Generally late replicating

domain are present at the nuclear periphery, while early replicating regions relocate to a more internal location (Hiratani, Ryba et al. 2008).

#### *Molecular mechanism underlying repression at the nuclear edge*

Gene silencing at the nuclear periphery can be the result of different mechanisms, for example a reduced availability of transcription factors or a limited mobility of loci restricted at the nuclear edge (Chubb, Boyle et al. 2002). Several lines of evidence pointed out a role for HDACs and low level of histone acetylation in keeping genes repressed while at the periphery. The nuclear periphery is in fact usually depleted of acetylated histones and TSA treatment of cells cause a global increase of histone acetylation at the nuclear periphery (Gilchrist, Gilbert et al. 2004; Bartova, Pachernik et al. 2005). Prolonged TSA treatment caused a rearrangement of the chromatin close to the nuclear edge (Brown et al., 2008), and reduce the repression of tethered loci (Finlan, Sproul et al. 2008; Reddy, Zullo et al. 2008). HDAC have also been shown to interact with the nuclear membrane protein LAP2b and mediate histone H4 deacetylation (Somech, Shaklai et al. 2005).

#### *Gene activity at the NPC*

Recent reports have uncovered a role for transcriptional activity at the nuclear pore complexes (NPC) in yeast. Different studies show that active genes can be associated with the nuclear periphery (Brickner and Walter 2004; Taddei, Van Houwe et al. 2006). In particular, gene activation requires the association of loci with the nuclear transport machinery and nuclear pore proteins that mediate their recruitment to the nuclear periphery (Casolari, Brown et al. 2004). In the case of activation of the Gal locus activation movement becomes confined to a bidirectional sliding at the nuclear edge. This limited diffusion is mediated by members of the SAGA acetyltransferase complexes and by a member of the RNA export machinery, that links the Gal gene to the nuclear pore (Cabal, Genovesio et al. 2006). Active genes can remain at the periphery even when they are repressed. Positioning at the nuclear edge have been demonstrate to facilitate their reactivation and this particular transcriptional status is epigenetically characterised by the presence of histone H2AZ variant (Brickner, Cajigas et al. 2007).

Also in *Drosophila*, reports uncovered an interaction between SAGA acetyltransferase factors with the nuclear pore complex that allow association of active genes with the nuclear periphery (Kurshakova, Krasnov et al. 2007). The presence of different domains of active and repressive chromatin at the nuclear periphery is suggested also in mammalian systems where different domains of lamin A or lamin B exist and are associated with gene-rich/active or gene-poor/silent chromatin, respectively. In particular, lamin A microdomains are also characterized by the presence of RNA pol and active marks but transcription is reduced, 'as a result of stalling of the RNA PolIII (Shimi, Pflgebraar et al. 2008). 2008)

## **1.2. Mash1 and the specification of a neural fate**

### **1.2.1 Proneural genes and models of mammalian neurogenesis**

A multitude of different neurons is produced in the embryo during the neurogenesis process. Multipotent neural progenitors are formed along the neural tube during a patterning process that creates progenitors with different potential according to the location the cells have along the tube (anteroposterior or dorsoventral axes). Cells present in the ventral spinal cord produce a population of neurons different from the one produced by progenitors in a more dorsal position. Later, multipotent progenitors become progressively committed to produce specific neuronal subtypes and the same progenitors give rise to different types of neurons and glial cells at different times of development. Generally, neurons are produced first, followed by oligodendrocytes and then astrocytes (Guillemot 2007).

The identity of multipotent neural progenitors along the neural tube and their subsequent commitment to a specific neuronal subtype is regulated by the interplay of different families of transcription factors. Patterning proteins, such as Pax6, Olig2 and Nk2.2, are expressed first to generate the multipotent progenitors with different identities along the neural tube. Their expression is maintained during further neuronal commitment where it is integrated with the expression of other factors such as proneural proteins (Sugimori, Nagao et al. 2007).

Proneural genes, such as Mash1, Ngn1-3 and Math1, are necessary and sufficient for the commitment of neural progenitors into different neural subtypes (Bertrand, Castro et al. 2002). Proneural genes were first identified in *Drosophila* as a group of four genes, achaete (ac), scute (sc), lethal of scute (sc) and asense (ase) that share sequence similarity and a bHLH domain, necessary for DNA-binding and dimerization (Villares and Cabrera 1987). Later, another gene atonal (ato) was identified and related to the achaete-scute complex (asc) (Jarman, Brand et al. 1993). The vertebrate homologues to the asc family include Ash1 (Mash1 in mouse) while Math1 and the family of Ngn, NeuroD and Olig are related to the ato family.

Proneural genes have different roles in the induction of neural specification. Proneural genes induce several helix-loop-helix genes that are involved in the neural differentiation, inhibit glial differentiation by different mechanisms and promote cell cycle exit by induction of cyclin-dependent kinase inhibitors (Bertrand, Castro et al. 2002). Differentiation in fact requires first inhibition of self-renewal by blocking the activity of genes of the SoxB1 family, that are expressed in the primitive neuroectoderm (Bylund, Andersson et al. 2003). The neuronal commitment involves then the exit

from the cell cycle and the induction of a cascade of downstream transcription factors such as neuronal proteins and the inhibition of astrocyte differentiation (Sun, Nadal-Vicens et al. 2001). Combinations of different transcription factors at different times allow the production of different type of neurons.

### **1.2.2 In vitro neural differentiation of embryonic stem cells**

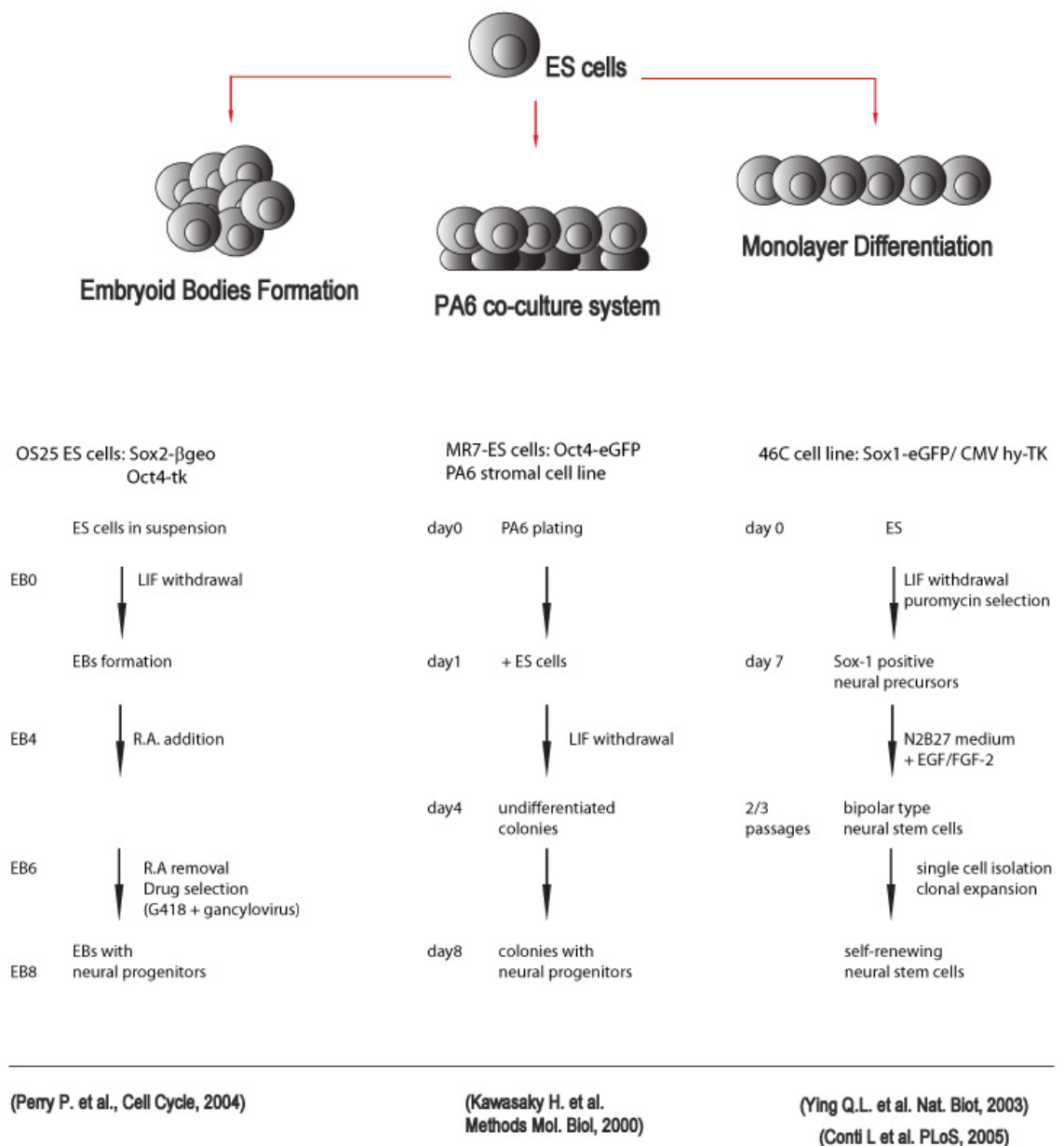
Mouse embryonic stem cells are pluripotent cells that can differentiate into a multitude of different cell type in vitro. ES cells are derived from the inner cell mass of the blastocyst and can be maintained in an undifferentiated state in the presence of cytokines such as the leukaemia inhibitor factor (LIF) (Ying, Nichols et al. 2003). Upon withdrawal of the self-renewing stimulus, ES cells can differentiate into neural, muscular, haematopoietic, epidermal and other cell types according to the culture conditions used (Giadrossi, Dvorkina et al. 2007).

Different protocols have been established for the neural differentiation of ES cells in vitro (Ying, Stavridis et al. 2003). One of the first protocols established involved the use of retinoic acid. According to this protocol, ES cells are grown in suspension culture to allow the creation of multicellular aggregates called embryoid bodies with the induction of derivatives from the three embryonic germ layers. Neural differentiation is then triggered by the use of retinoic acid that results in the production of neural progenitors (Bain, Kitchens et al. 1995). An alternative protocol involves the production of embryoid bodies without the addition of retinoic acid and their further plating in a defined serum-free media to eliminate non-neural cells (Lendahl, Zimmerman et al. 1990). Cells can be further propagated and differentiated into a neuronal phenotype after plating the embryoid bodies on laminin- or gelatine-coated plates and using additional morphogens and growth factors such as Shh, FGF-2 and FGF-8 (Lee, Lumelsky et al. 2000).

A different system that does not require the formation of aggregates is the co-culture of ES cells with bone marrow-derived stromal cells (PA6). Although the molecular mechanism is unknown, this system allows the specific induction of the neural lineage (Kawasaki, Mizuseki et al. 2000). Alternatively, ES cells can be cultured in adherent monoculture and allow differentiation in a defined media. The process required withdrawal of LIF, inhibition of alternative cell fates in serum-free conditions and the use of FGF (Ying, Stavridis et al. 2003). In all these cases the heterogeneity in the population of cells produced is high and neural progenitors can be purified using selectable markers such as Sox1-GFP with a dual selection/reporter cassette (Fig. 1.1). Isolation of Sox1 neural progenitors can be achieved by fluorescence activated cell sorting or drug selection. ES-derived neural cells expressed specific markers such as nestin for neural progenitors and  $\beta$ III tubulin,  $\gamma$ -aminobutyric acid (GABA), tyrosine hydroxylase for neurons (Ying, Stavridis et al. 2003).

Neural stem cells have the ability to self-renew and generate all the different types of neurons. In vivo their ability to self-renew is maintained in a complex niche environment of the mammalian brain. In vitro neural stem cells were first isolated in neurosphere, floating cell clusters containing a mixed of committed progenitors and a few neural stem cells (Reynolds and Weiss 1992). A recent protocol has now established the derivation of neural stem cells from embryonic stem cells using a monolayer differentiation, with the addition of fibroblast growth factor 2 (FGF-2) and epidermal growth factor (EGF). These cells are functional as seen by electrophysiology studies and are able to integrate in the host brain after transplantation.(Conti, Pollard et al. 2005).

Figure 1.1



1.1. In vitro systems used to promote neural differentiation of ES cells. A schematic representation of retinoic acid differentiation, PA6-mediated neural induction and monolayer differentiation in serum-free condition is shown in the top panel. Below, details of the three protocols, including the use of specific ES cell lines, drug selection and references.

### 1.2.3 Mash1

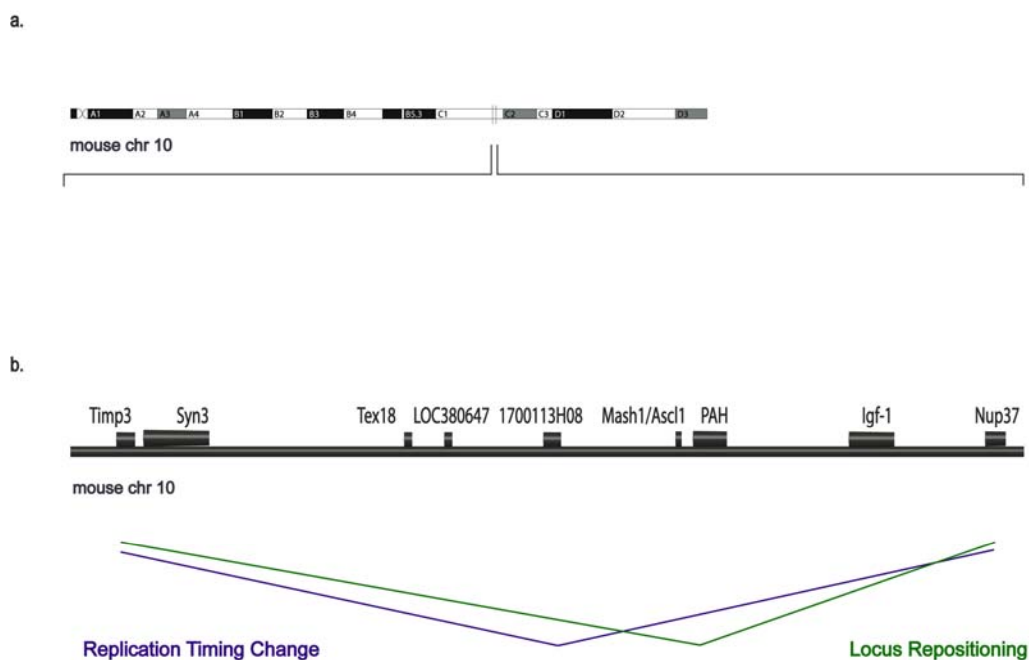
*Mash1* (*Ascl1*) is a mammalian proneural gene that was identified through its homology to the *Drosophila* Achaete-Scute family (Johnson, Birren et al. 1990). During the early stages of mouse embryogenesis *Mash1* is expressed in a tissue and temporal-specific pattern. *Mash1* mRNA and protein were detected in the ventricular zone of the central nervous system, within the innermost layer of the neural tube as well as in the olfactory epithelium and the autonomic nervous system (Cau, Casarosa et al. 2002). Mice homozygous for a null mutation in the *Mash1* gene die at birth. Although the general structure of brain and spinal cord appear normal, the neurogenesis process is deeply affected. In particular, in the olfactory epithelium, neuronal progenitors die at an early stage, while in the sympathetic ganglia the neuronal precursors are present but fail to differentiate. In both these tissues the non-neural supporting cells are present and unaffected (Guillemot, Lo et al. 1993). These observations, together with the neural specific expression pattern of *Mash1* suggest that it is required in the early stages of embryogenesis for the formation of neural precursors and for the commitment of these multipotent progenitors into different neural cell types (Parras, Galli et al. 2004; Pattyn, Simplicio et al. 2004). Elucidating the mechanisms that control the spatial and temporal expression patterns of *Mash1* and other proneural factors may lead to a better understanding of pattern formation in the developing nervous system as well as gaining further insights into general mechanisms of gene regulation during cell commitment.

Using the retinoic acid based protocol to promote neural differentiation of ES cells, our laboratory has documented changes in the replication timing and chromatin status of *Mash1* and other developmental-specific genes upon neural differentiation. The *Mash1* locus replicates late in ES cell and becomes early replicating upon neural induction (Perry, Sauer et al. 2004). Changes in the replication timing of *Mash1* occur concomitant with its upregulation and histone modifications. At the promoter, a high level of repressive Me<sub>3</sub>H3K27 in undifferentiated ES cells is replaced by a high level of AcH3K9 upon their differentiation towards the neural lineage. Moreover, a striking change in the nuclear positioning of *Mash1* is observed: from the periphery in ES cells the locus is found to undergo a large-scale chromatin reorganisation resulting in its relocation away the nuclear periphery upon



neural differentiation (Williams, Azuara et al. 2006). These chromatin events have been demonstrated to be specific for the neural fate, since they are observed both in *ex vivo* neural cells and in two different ES-derived neural progenitor cells, but not in *ex vivo* T cells, keratinocytes or ES-derived mesoderm cells. The genes surrounding the *Mash1* locus also relocate away from the nuclear periphery upon neural differentiation. However none of them is neural and the different epigenetic changes seem to be centred at the *Mash1* locus (Fig.1.2) (Williams, Azuara et al. 2006).

Figure 1.2



1.2. Summary of the changes in the replication timing and nuclear positioning of the *Mash1* locus observed upon neural induction. a. Mouse chromosome 10 with G banding pattern and *Mash1* localisation at 10C1. b. *Mash1* and neighbouring genes shown in further details (2Mb genomic region). Highest switch in the replication timing between ES and neural progenitors focuses in the *Mash1*-*LOC380647* region and extends to *Timp3* (1.2 Mb centromeric) and *Nup37* (0.65 Mb telomeric). The whole region relocates away from the nuclear periphery after neural induction with *Mash1* showing the highest degree of repositioning.

#### **1.2.4 Role of REST in silencing Mash1 and neuronal genes in ES cells**

##### *REST silencing of neuronal genes in non neural cells*

REST/NRSF (RE1 silencing transcription factor) is a transcriptional repressor of neuronal genes first identified in non-neuronal cell types. The REST protein contains a zinc-finger domain that interact with the DNA and at least two repressor domains (N-terminal and C-terminal), which are necessary for the recruitment of a number of corepressors (Chong, Tapia-Ramirez et al. 1995; Schoenherr and Anderson 1995). REST binding to DNA is specific to the 21-23 bp RE-1 responsive element which is found near a large number of neuronal genes (Sun, Greenway et al. 2005; Otto, McCorkle et al. 2007; Johnson, Teh et al. 2008). Two different mechanisms have been proposed for REST mediated silencing: an active repression via recruitment of mSin3 and histone deacetylase 1,2 (HDAC1,2) by the NH<sub>2</sub>-terminal repressor domain (Ballas, Battaglioli et al. 2001) and silencing with the recruitment of the corepressor coREST by the C-terminal domain. coREST act as a platform for the binding of the silencing machinery that results in the CpG methylation and recruitment of methyl DNA binding protein MeCP2, histone H3K9 methyltransferase SUV39H1, heterochromatin protein HP1 and spreading of the silencing marks (Lunyak, Burgess et al. 2002).

REST is required for the correct development of the nervous system in mice and chicken embryos. REST is expressed ubiquitously in the mouse embryo at the embryonic day 8.5 and 9.5 (E 8.5, E 9.5). Deletion of REST affects the embryonic development of mice, with growth retardation and forebrain malformation visible at embryonic day 9.5 (E.9.5). Embryonic lethality is observed at embryonic stage 9.5/10 (E.9.5/10) and it is preceded by general apoptotic death. Inhibition of REST function in chicken embryos causes the derepression of tubulin and other neuronal genes in non neural tissues (Chen, Paquette et al. 1998). Also the overexpression of REST in spinal cord of developing chicken embryo caused impairment of the nervous system (Paquette, Perez et al. 2000). The protein REST is abundantly expressed during embryogenesis in non-neural tissue and neural progenitors suggesting a role for REST in restricting neuronal expression in mature neurons (Chong, Tapia-Ramirez et al. 1995; Schoenherr and Anderson 1995).

##### *Role of REST in ES cells and neural differentiation*

Genes containing a REST binding site have been identified with computational studies (Bruce, Donaldson et al. 2004). Genome-wide chromatin immunoprecipitation studies have demonstrated REST binding close to many neuronal genes in ES cell, neural stem cells, mature neurons and non-neuronal cells. Most target genes are expressed at low level in ES cells or neural stem cells and become highly transcribed in mature neurons (Sun, Greenway et al. 2005; Otto, McCorkle et al. 2007; Johnson, Teh et al. 2008). REST target genes include developmentally regulated genes common to all the cell type and cell type specific including a class of ESC specific genes where binding of REST is

detected only in ES cells (Johnson, Teh et al. 2008). Binding of REST at a RE1 responsive element 50Kb upstream the Mash1 promoter has been identified in ES cells. The amount of REST protein bound at the Mash1 locus decreases in cortical progenitors concomitant with the upregulation of the gene. No changes in the REST binding are observed at other neuron specific genes in cortical progenitors until fully differentiation in neurons (Ballas, Grunseich et al. 2005).

The role of REST during neural differentiation has been linked to its ability to recruit HDAC and repressor proteins at RE1 repressive elements (Ballas, Grunseich et al. 2005). The mechanism proposed is independent of DNA methylation and creates a repressed chromatin state where genes are poised for transcription. During neural differentiation REST protein is degraded posttranslationally first at a low level in neural progenitors and then it is completely abolished in neurons with the derepression of target genes. Two different classes of genes have been classified according to their response at low or high concentration of the REST protein (Ballas, Grunseich et al. 2005). REST protein is degraded with the progression of the neural differentiation. The degradation is regulated by a ubiquitin-mediated proteolysis and the E3 ubiquitin ligase responsible for REST degradation is part of a complex containing the F-box protein  $\beta$ -TRCP (SCF <sup>$\beta$ -TRCP</sup>) (Westbrook, Hu et al. 2008).

REST role in recruiting co-repressor factors at RE1-containing genes has also been shown in neural stem cells. Epigenetic features of a repressed chromatin state such as low level of histone H4 and low level of H3K9 acetylation are present together with the active mark H3K4 methylation at REST repressed neuronal genes in embryonic hippocampal neural stem cells (Greenway, Street et al. 2007).

REST protein is abundant in ES cells, and pluripotency-associated factors Oct4 and Nanog bind to the REST promoters in human and mouse ES cells (Boyer, Lee et al. 2005; Loh, Wu et al. 2006; Kim, Chu et al. 2008). The role of REST in ES cells is controversial. The protein has been implicated in the repression of neural specific gene (Mash1, Ngn2, Bry, Gata4, Sox18) and in the expression of pluripotency associated genes. Loss of REST determines the expression of lineage specific markers and differentiation of ES cells (Singh, Kagalwala et al. 2008). However, these findings were not confirmed by several other studies in which the partial or complete abrogation of the REST protein did not impair the characteristics features of ES cells. No precocious expression of early differentiation markers was observed as well as expression of pluripotency associated genes was unaffected suggesting REST specificity in regulating neuronal genes (Buckley, Johnson et al. 2009; Jorgensen, Chen et al. 2009).

## 1.3. Role of CTCF and Cohesin in gene regulation

### 1.3.1 The cohesin complex and associated proteins

After DNA replication, cohesion is established between sister chromatids to avoid aberrant separation of chromosome during mitosis and to prevent double strand breaks that would determine loss of chromosome. The cohesin complex and associated proteins are responsible for the cohesive forces that allow the correct localisation of the chromosomes on the metaphase plate. Subsequent loss of cohesion determine the separation of sister chromatids that, pulled away by the shortening of microtubules, migrate to the opposite pole of the cell (Skibbens 2008). The cohesion proteins were identified in yeast screening looking for mutants that were able to separate sister chromatids precociously before anaphase. Four proteins were shown to be part of a complex called cohesin to indicate its essential role in cohesion (Guacci, Koshland et al. 1997; Michaelis, Ciosk et al. 1997). The structure of the complex is highly conserved. Two subunits, Smc1 and Smc3, are part of the 'structural maintenance of chromosome' family of proteins and are characterised by a long polypeptide chain that folds back on itself at a hinge domain and terminate through an anti-parallel coiled-coil structure with a globular domain. This end domain is an ATPase head formed by the two N and C terminals (Michaelis, Ciosk et al. 1997). The two subunits Smc1 and Smc3 are attached at their hinge domain and are connected at their ATPase head by the subunit Scc1 that is part of the kleisin family. The three subunits form a ring-like structure with a diameter of 50 nm that is thought to embrace the DNA (Haering and Nasmyth 2003). A fourth protein Scc3, is linked to the cohesin complex by the Scc1 protein and in somatic vertebrate cells is represented by the related homologues SA1 and SA2. Binding of ATP at the Smc1 subunit is necessary for Scc1 recruitment (Arumugam, Gruber et al. 2003).

Other proteins associated to the complex seem to be important in protein-protein interaction and are thought to be required for loading or removal of cohesin from the chromosomes. Pds5 has numerous HEAT repeats known to be important in protein-protein interactions and associate in vertebrates with SA (Sumara, Vorlaufer et al. 2000). Sororin, identified only in vertebrate, is the substrate of ubiquitin ligase anaphase-promoting complex/cyclosome (APC/C) (Rankin, Ayad et al. 2005).

#### *Cohesin recruitment to chromatin*

An indirect role in sister chromatids cohesion is mediated by proteins that recruit the cohesin ring to chromosomes. Scc2p and Scc4p proteins, first identified in yeast, form soluble complexes called the loading complex that bind chromatin and are necessary for the loading of the cohesin complex on the DNA (Ciosk, Shirayama et al. 2000). In Scc2 and Scc4 mutants, the cohesin ring can form correctly but fails to bind the chromatin fibre. The same phenotype is observed when ATP binding or hydrolysis at the cohesin heads is inhibited, suggesting that hydrolysis of ATP might allow the

opening of the ring, with the temporal detachment of Scc1 subunit and the entering of DNA in the ring (Ciosk, Shirayama et al. 2000). In vertebrates, experiments in *Xenopus* eggs, showed that recruitment of cohesins to DNA is still dependent on Scc2 but it also requires proper assembly of pre-replicative complexes (pre-RC) at origins of DNA replication (Takahashi, Yiu et al. 2004; Ryu, Kim et al. 2006).

#### *Establishment of cohesion during S phase*

Cohesin complex alone is not sufficient for sister chromatid cohesion in S phase. Eco1/ctf7 cohesion factor is not part of the cohesin complex but is necessary for the establishment of cohesion during S phase and becomes dispensable for its maintenance later on in mitosis (Skibbens, Corson et al. 1999; Toth, Ciosk et al. 1999). Eco1/Ctf is an acetyltransferase whose function is to acetylate two lysine residues located on the ATPase head of the Smc3 subunit that are thought to mediate the formation of stable links between the newly replicated sister chromatids (Unal, Heidinger-Pauli et al. 2008). Two different types of cohesin are in fact present during the cell cycle in term of mobility and affinity for chromatin as visualised by fluorescence recovery after photobleaching (FRAP). Cohesin unstably bound to DNA is present in G1 and is replaced by a stably associated form after DNA replication in G2 (Gerlich, Koch et al. 2006). In vertebrate, the sororin protein seem also to be important in the establishment of cohesion (Schmitz, Watrin et al. 2007).

#### *Cohesin dissociation from the chromatin*

In yeast, removal of cohesin from DNA coincides with the onset of anaphase. A protease called Esp1 promote the proteolytic cleavage of the Scc1 subunit with consequent separation of the two sister chromatids (Uhlmann, Lottspeich et al. 1999). The APC complex (anaphase promoting complex) has a ubiquitin-protein ligase that targets Pds1p for degradation, and releases the anaphase activator Esp1p (Ciosk, Zachariae et al. 1998). Scc1 levels are cell cycle regulated: visible in G1, the protein peaks in S phase and declines in anaphase (Michaelis, Ciosk et al. 1997).

In vertebrates, removal of cohesins from the chromosome arms happens at two distinct moments and is regulated by two different mechanisms. The prophase pathway removes the majority of cohesins that are unstably bound to DNA. This process is independent from the APC complex and no cleavage of Scc1 is observed (Sumara, Vorlaufer et al. 2000). It relies on the activity of the enzyme Polo-like kinase 1 (Plk1) and in *Xenopus* egg, Plk1 activity is required for phosphorylation of Scc1 and SA1/2 subunits, decreasing the affinity of cohesin to chromatin (Sumara, Vorlaufer et al. 2002). The remaining pool of cohesins, only about 10% of all cohesin complexes, is mainly bound to centromeric region and it is removed later on at the metaphase-anaphase transition by separate cleavage of the Scc1 subunit (Waizenegger, Hauf et al. 2000).

### **1.3.2 Evidence of cohesin role in gene regulation**

The first evidence suggesting cohesin has an additional role in regulating gene expression came from experiments in yeast and *Drosophila*. In yeast, mutations in the Smc1/3 subunits affected the function of boundary elements at the HMR mating type locus (Donze, Adams et al. 1999). In *Drosophila*, heterozygous mutations in the Nipped-B gene, the orthologue of *S. Cerevisiae* Scc2, increase the enhancer-blocking activity of a gypsy insulator (Rollins, Morcillo et al. 1999). On the other hand, mutations in Scc3/SA diminish the enhancer-blocking activity of the same gypsy insulator, suggesting a role in blocking enhancer promoter interaction (Rollins, Korom et al. 2004). In zebrafish, Rad21 and Smc3 are important for the regulation of the Runx family of transcription factors and Rad21 depletion cause impairment in the development of the hemotopoietic and nervous system (Horsfield, Anagnostou et al. 2007).

A role of cohesin in gene regulation is suggested also in vertebrates by the presence of cohesin on chromosomal arms in interphase and by the presence of cohesin in post-mitotic cells like neurons (Wendt, Yoshida et al. 2008). Moreover, mutations in the NIPBL gene, the human orthologue of *S. Cerevisiae* Scc2, are found in patients affected by the Cornelia de Lange Syndrome, characterized by developmental defects, such as craniofacial abnormalities, growth and mental retardation, but mild or no alterations in chromatid cohesion is observed (Tonkin, Wang et al. 2004). Mutations associated to the disease were also found in the Smc1 and Smc3 subunits (Revenkova, Focarelli et al. 2009)), while in the related Robert's syndrome mutations affect the ESCO2 gene, an homologue of the yeast cohesion factor Eco1 (Vega, Waisfisz et al. 2005). The absence of Pds5b in mice cause developmental defects similar to the ones observed in the Cornelia de Lange syndrome. Pds5b *-/-* mice die shortly after birth without showing cohesion defects (Zhang, Jain et al. 2007).

### **1.3.3 Cohesin binding to chromosomal arms**

Insights into the role of cohesin in regulating gene expression came from genome-wide studies of cohesin binding to DNA.

In yeast the presence of cohesins along the chromosomal arms correlates with transcription. Cohesin is enriched at the 3' UTR of genes and at intergenic sites between convergent transcription units (Glynn, Megee et al. 2004). Cohesin bound sites do not correlate with regions where Scc2/Scc4 complexes are present. The molecular mechanism proposed suggests that cohesin is relocated from sites of loading by the transcription machinery and therefore accumulates at the end of genes (Lengronne, Katou et al. 2004). Cohesin is also enriched at the silent chromatin of centromeric regions, where it is recruited by Swi6/HP1 complex (Nonaka, Kitajima et al. 2002) and at boundaries

elements between active and repressed domains (Blat and Kleckner 1999; Laloraya, Guacci et al. 2000).

In *Drosophila*, the loading factor Nipped-B and cohesin colocalise genome-wide and they preferentially bind transcribed genes. The bound regions are wider than in yeast and binding occurs usually at 5'UTR or intronic regions, where they colocalise with RNA PolIII. Differences in cohesin binding across different cell lines usually correlate with differences in transcription (Misulovin, Schwartz et al. 2008).

Differently in mammalian system cohesin binding does not correlate with transcription but it is influenced by chromatin structure and DNA sequence content. Cohesin binding has been detected preferentially at a subset of constitutive hypersensitive sites where it is recruited by the CTCF factor (Parelho, Hadjur et al. 2008; Stedman, Kang et al. 2008; Wendt, Yoshida et al. 2008). No preference for intergenic region between convergent transcript is observed but highest binding is seen close to the TSS of genes. Cohesin binding is compatible with locus remodelling and a high level of transcription (Parelho, Hadjur et al. 2008).

#### **1.3.4 Role of CTCF and cohesin in regulating gene expression**

The colocalisation of cohesin and CTCF at many sites across the genome of mammalian cells suggests these proteins cooperate to regulate gene expression.

CTCF 11 zinc finger protein binds DNA and can act as a repressor or activator of gene expression. Binding to DNA is methylation sensitive and is inhibited when CpG dinucleotides are methylated (Bell and Felsenfeld 2000). CTCF binding has been detected at insulators where it has enhancer-blocking functions and at boundary elements which are involved in demarcating different chromatin domains. The role of CTCF in regulating gene expression has been demonstrated at the H19/IGF2 imprinted locus,  $\beta$ -globin locus, X chromosome and boundary elements between active and inactive domains (Zlatanova and Caiafa 2009).

In the case of the H19/Igf2 locus, CTCF mediates the formation of long range chromatin loops that result in the differential expression of H19 and IGF2 as shown by 3C experiments (Kurukuti, Tiwari et al. 2006). Interactions between the imprinting control region (ICR) and other differentially methylated regions in the 5' and 3' UTR of the IGF2 gene allow the creation of a transcriptional active domain for IGF2 on the paternal allele and a repressive loop on the maternal allele (Murrell, Heeson et al. 2004). At the chicken beta-globin locus, CTCF binding has been detected at the 5'HS4 insulator region and implicated in both barrier and enhancer blocking activity (Chung, Bell et al. 1997). The molecular mechanism proposed involves the formation of long range chromatin

interactions within the locus for further recruitment of transcription factors that activate the locus (Splinter, Heath et al. 2006). Deletion of CTCF does not alter the expression of the globin genes, but CTCF seems to have a more structural role (Splinter, Heath et al. 2006; de Laat, Klous et al. 2008).

Cohesin cooperation in the enhancer blocking activity of CTCF has been demonstrated at the H19/IGF2 and beta globin locus (Wendt 2008; Parhelo 2008). In luciferase assay, where the ICR is placed between H19 enhancer and a reporter gene, siRNA mediated reduction of cohesin level has the same effect in increasing the activity of the reporter gene as knock down of CTCF (Wendt, Yoshida et al. 2008). Likewise, at the beta-globin 5' HS4 insulator, depletion of cohesin has the same effect as CTCF, in increasing the reporter expression when two copies of insulator region are placed between enhancer and reporter gene (Parelho, Hadjur et al. 2008).

In *Drosophila* the role of cohesin in gene regulation has been observed in post mitotic neurons where its inactivation leads to defects in neuron pruning (Pauli, Althoff et al. 2008; Schuldiner, Berdnik et al. 2008).

The molecular mechanism underlying the role of cohesin in regulating gene expression is not known but experiments at the *Ifng* locus suggest its requirement in mediating long range chromosomal interactions. During differentiation of CD4 T cells in Th1 cells the *Ifng* undergoes locus remodelling that results in its expression. This remodelling is accompanied by an enhanced association of CTCF and cohesin at a set of conserved sequences and it has been suggested that the spatial constraining imposed by cohesin binding may affect the probability with which regulatory regions interact with each other (Hadjur, Williams et al. 2009).

## **1.4 Aims of the study**

In ES cells *Mash1*, an important proneural gene, is not expressed. Upon neural differentiation the locus undergoes specific epigenetic changes in histone modifications, replication timing and nuclear positioning that result in the gene expression. Here I want to look at the dynamics of such changes and at the molecular mechanism underlying the large scale chromatin reorganisation of the locus.

Using an in vitro system of ES differentiation into neural progenitors, I have looked at the timing of nuclear repositioning and gene expression during neural differentiation to understand whether relocalisation is a consequence of gene expression or precedes *Mash1* transcription.

I have also looked at the role of REST repressor in silencing *Mash1* in ES cells. REST has been proposed to be a master regulator of ES cell differentiation into neural cells and a REST binding site



has been identified 50Kb downstream the Mash1 promoter (Ballas et al., 2005). Here I have looked at the effect that REST depletion has on the nuclear localisation of Mash1 and on its expression, to address whether REST is important in maintain Mash1 in a repressed state at the nuclear periphery in ES cells. I have also looked at the nuclear positioning of other REST targets in REST depleted cells to address whether the localisation at the nuclear periphery is a general mechanism of keeping genes in a repressed state in ES cells.

I have then looked at the role of CTCF and cohesin in the regulation of Mash1 in ES cells. CTCF and cohesin has been implicated in the formation of long range chromatin interactions that result in the creation of domains permissive for transcription or repression. Binding of CTCF is present at the borders of LADs and at the nuclear periphery at sites of differential expression when HDAC is inhibited (Guelen, Brown, Kennedy et al. 2008). Here I have looked at the binding of CTCF and cohesin across the locus in ES cells, to see if binding is present in the region that is shown to switch its replication timing and have a differential expression site upon neural differentiation. I have then used RNAi to deplete cells of CTCF and cohesin to ask whether these proteins are important for the repression of Mash1 in ES cells and its localisation at the nuclear periphery.

Finally we have generated ES cells that carry a few copy number of Mash1 transgene at different integration sites. We aim to dissect whether a 140Kb region encompassing the Mash1 locus is able to localise at the nuclear periphery and to relocate upon neural induction. We also want to dissect whether the transgene has the ability to influence the 'epigenetic status' of neighbouring genes at different site of integration. If so, the responsive elements will be mapped by deletion approaches.

## Chapter 2

### Mash1 regulation in ES cells

#### 2.1. Introduction

The nuclear periphery has been generally referred to as a repressive compartment for higher eukaryotes. This domain of repressive chromatin is characterized by gene poor regions and the absence of active histone marks as well as RNA pol as demonstrated in *Drosophila* and Humans (Kupper, Kolbl et al. 2007; Guelen, Pagie et al. 2008). The periphery is usually depleted of transcribed genes as shown by multiplex RNA FISH analysis (Levsky, Shenoy et al. 2007) and repositioning to or away from the nuclear periphery coincides in some cases with, respectively, repression or activation during differentiation (Hewitt, High et al. 2004; Williams, Azuara et al. 2006). However it is still not clear what role the nuclear periphery plays in the regulation of gene expression: whether it is necessary for transcriptional repression or it has just a passive role in which gene transcription tends to drive the locus away from it (Deniaud and Bickmore 2009).

#### 2.2. Mash1 expression and nuclear location in ES cells following neural induction

In ES cells Mash1 is not abundantly expressed and the locus has features associated with repressive chromatin: low levels of histone acetylation, low transcript expression, replication timing late in S-phase and the preferential localisation of the locus at the nuclear periphery. In neural progenitors, where the gene is expressed, markers of an active chromatin state are present: high level of histone acetylation, early replication timing and the gene preferentially localise in the nuclear interior (Williams, Azuara et al. 2006). These changes in replication timing and locus repositioning, in response to neural induction of ES cells, extends also to neighbouring genes (Fig. 1.2). To further investigate the relationship between Mash1 transcription and location at the nuclear periphery, I examined the gene nuclear location and expression profile at different days of ES cells neural differentiation.

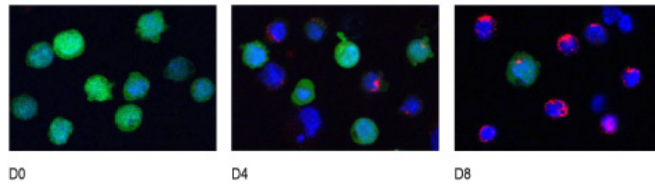
### **2.2.1. In vitro neural differentiation of embryonic stem cells**

Different protocols have been established to differentiate ES cells towards the neural lineage (Fig 1.1). Culturing of ES cells on a layer of PA6 stromal cells under serum free condition was shown to produce neural progenitors. This differentiation is specific for the neural lineage, with minimal induction of other germ layers, and does not require retinoic acid or the use of selective markers (Kawasaki, Mizuseki et al. 2000).

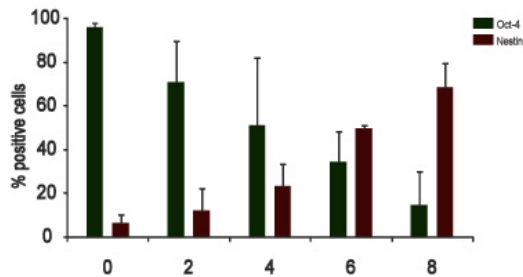
ES cells with an Oct4-GFP transgene (MR-7) were plated on a layer of irradiated PA6 stromal cell line and allowed to differentiate into neural cells under serum free conditions. ES cells form characteristic colonies that change morphology gradually with the progression of the differentiation: from rounded and raised shaped colonies at day 4 of differentiation to the more flat and irregular colonies at day 8 upon induction of neural progenitors. Neural progression was assessed at different days by immunofluorescence for the proneural marker Nestin and by Oct4-GFP expression analysis. The number of cells that were positive for Oct4 (GFP) or Nestin (Alexa labelled) was counted and represented in graph as a percentage of the total. At day 8 of neural induction around 70% of cells were positive for Nestin while the remaining ~30% were still Oct-4 positive (15%) or non-neural (Fig. 2.1 a and b). Differentiation was also assessed at the level of transcription for the pluripotency gene Oct4 and neural specific marker Sox1. Gene expression profile by Real Time RT-PCR showed a progressive increase of Sox1 mRNA level and decrease of Oct4 transcripts confirming that most cells had undergone neural differentiation (Fig. 2.1 c).

Figure 2.1

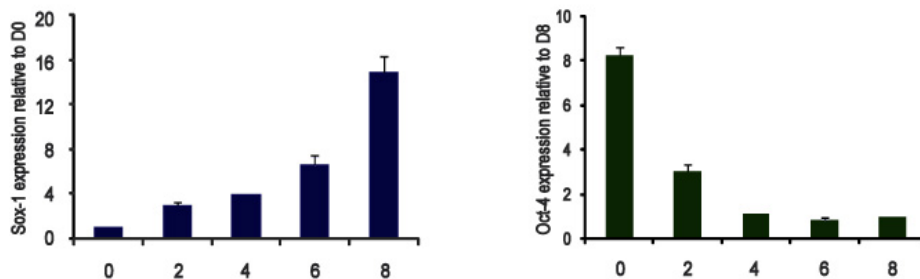
a.



b.



c.



2.1. Kinetics of MR-7 neural differentiation in co-culture with PA6 stromal cell line. a. Representative images of undifferentiated ES cells (day0) and differentiated cells at day 4 and day 8 of neural differentiation. Oct-4 (marker for pluripotency) in green, Nestin (neural specific marker) in red. DAPI-staining of DNA in blue. b. Time course of neural induction by immunofluorescence analysis for the Oct4-GFP positive cells in green and Nestin positive cells in red. c. Time course of neural induction by gene expression analysis. Sox1 and Oct-4 mRNA levels, normalised for HMB5 and

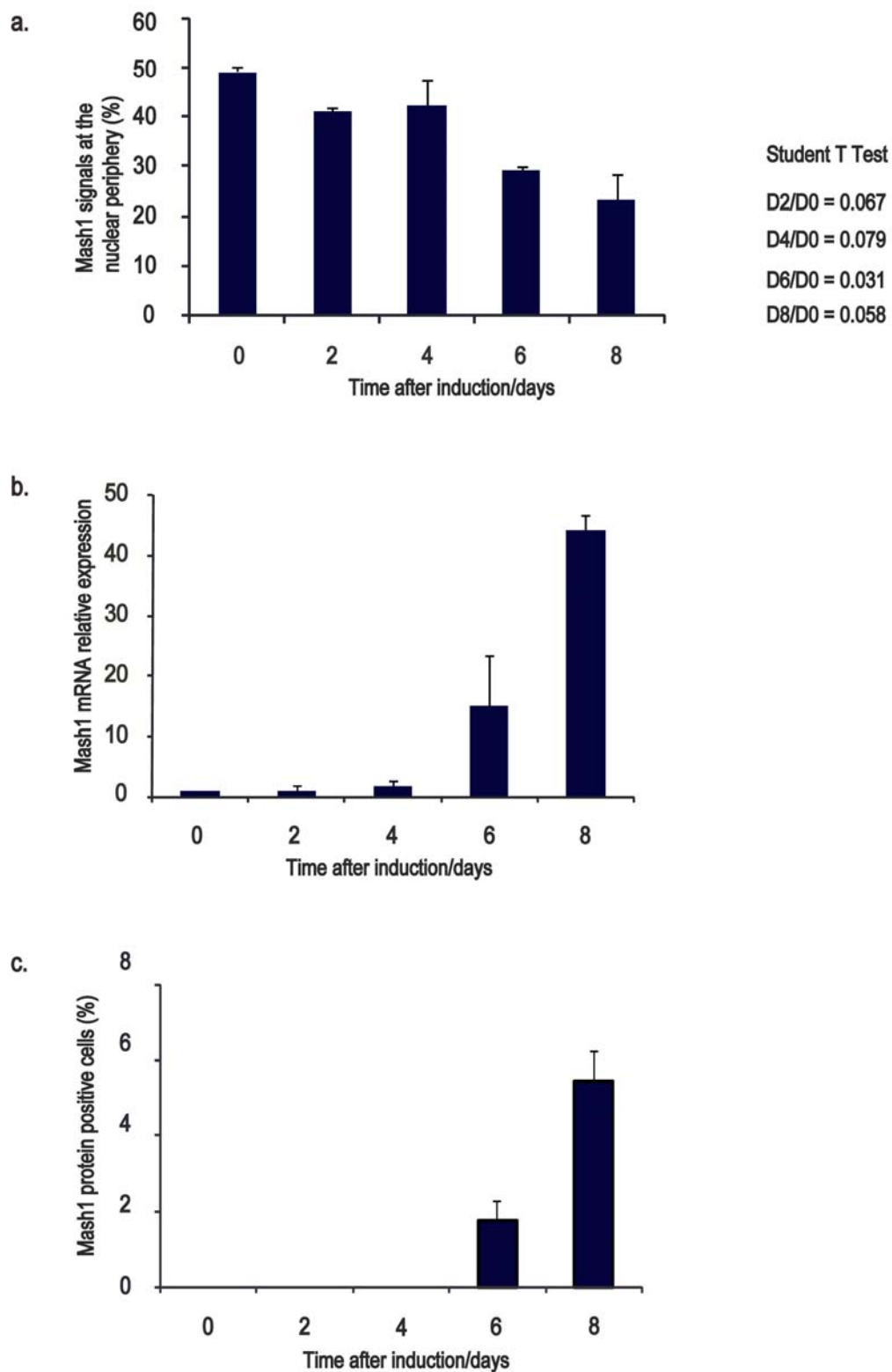
UBC housekeeping genes, are shown at different days of neural differentiation. Error bars indicate s.d. between two different experiments.

### **2.2.2. 2D FISH analysis of Mash1 nuclear position in ES cells before and after neural induction**

We previously showed that Mash1 is transcribed in neural progenitors where the locus is located away from the nuclear periphery (Williams, Azuara et al. 2006). However the dynamics of its repositioning is not clear, whether transcription is possible only once the gene is located away from the repressive nuclear compartment or if it is the process of transcription itself that drive the locus away. FISH analysis in 2D was used to measure the frequency with which the Mash1 gene resides in a small percentage of the nuclear volume close to the nuclear periphery. In this assay, cells are fixed with methanol plus acetic acid, and the FISH signals scored as a ratio between the distance (d) nuclear centre to FISH signal and the distance (r) nuclear centre to periphery. Only alleles where this ratio is above 0.8 are considered peripheral. Since 50% of alleles reside in this small portion of the nucleus at day 0 of differentiation the gene is considered peripheral. If the gene was located randomly in the nucleus we would expect to find a theoretical 22% of peripheral alleles in this analysis (Kosak, Skok et al. 2002). FISH using a BAC probe spanning 140 Kb of the Mash1 locus was performed at different days of neural differentiation to determine the timing of Mash1 gene repositioning in the nucleus. A change in the nuclear localisation was observed at day 6 that further increased at day 8, where the locus is preferentially located away from the nuclear periphery. No striking change was observed in the initial days of neural differentiation. Statistical analysis using the Student T test was performed to assess the statistical significance of the changes and confirm the change in nuclear positioning visible at day 6 (Fig. 2.2 a). In parallel with the FISH experiments, gene expression analysis as well as Mash1 protein detection by immunofluorescence showed rise in Mash1 transcript level initiated at day 6 and increased to day 8 in concomitant with the visible production of neural progenitors in the ES-derived colonies (Fig. 2.2 b and c).

Although no clear movement is observed during the first days of neural differentiation before active expression, by day 6, as Mash1 transcripts and protein are detected, the majority of alleles (70%) are no longer peripheral, confirming that gene expression occurs away from the nuclear periphery.

Figure 2.2



2.2. Kinetics of Mash1 induction and locus repositioning in the MR-7 neural differentiation system. a. Mash1 locus association with the nuclear periphery was assessed by 2D-FISH. The bar chart shows the percentage of peripheral alleles at different days. b. Time course analysis of Mash1 mRNA level at day 0, 2, 4, 6 and 8 of neural differentiation. c. Mash1 protein analysis was assessed by immunofluorescence. The percentage of positive cells is shown. Error bars indicate s.d. between two different experiments.

### **2.3. Potential role of the REST in regulating Mash1 in ES cells**

The molecular mechanism underlying Mash1 repositioning away from the nuclear periphery during neural differentiation is not known. One possibility is that in ES cells repressor proteins actively tethers the locus close to the nuclear membrane.

REST protein was identified as a transcriptional repressor in non-neural cells with the role of preventing the inappropriate expression of neuronal genes in non-neural cell type (Schoenherr and Anderson 1995). Mash1 was shown to be a REST target in ES cells, together with other neuronal genes. Loss of REST binding to a RE1-containing REST element 50Kb downstream of the transcriptional start site was observed following neural induction and this was hypothesised to be linked to the activation of the gene (Ballas, Grunseich et al. 2005). We therefore asked whether REST has a direct role in the repression of Mash1, keeping the gene in a repressive compartment close to the nuclear periphery.

#### **2.3.1. REST is not required to maintain the peripheral location of the Mash1 locus in ES cells**

To test the importance of REST on Mash1 subnuclear localisation and gene expression, we used ES cells derived from embryos that were wild type or homozygous for a targeted deletion of REST. 3D FISH was used to measure Mash1 location in ES cells, neural stem cells REST knock out and wild type cells. Two different experiments were performed for each cell type but only cells from one experiment were analysed by confocal microscopy. The distance between each allele and the nuclear periphery was measured in single optical focal sections (Fig. 2.3a). FISH signals were scored as for the 2D FISH analysis where only ratios above 0.8 are considered peripheral. The percentage of alleles at the periphery per cell type is shown (Fig. 2.3b). As previously demonstrated Mash1 alleles preferentially (~50%) localise at the nuclear periphery in ES cells. The localisation is changed in ES derived neural stem cells (which were used here as a positive control) where the gene is uniformly

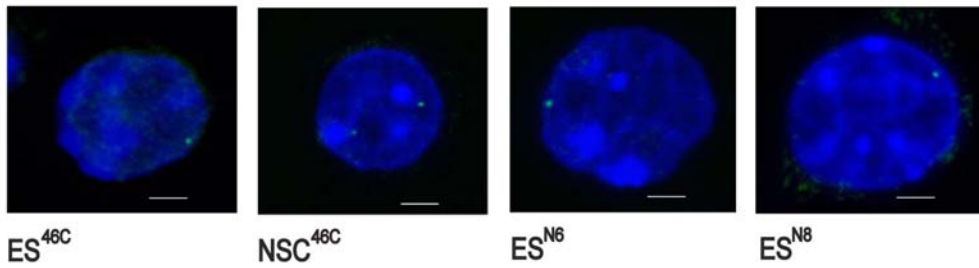
expressed. No difference in the nuclear localisation was observed between REST knock out ES cells and correspondent wt cells. Mash1 was peripherally located in both cell lines (Fig 2.3b).

We then re-analysed the data considering the position of both Mash1 alleles in each single cell to see if there were any changes of interest. No differences were detected between ES wt and REST knock out cells as in both cell type the majority of cells have at least one allele at the periphery. Again, the positive control, neural stem cells, is significantly different from ES cells with a higher number of cells with both internal alleles (~50%), consistent with the gene expression (Fig. 2.3c).

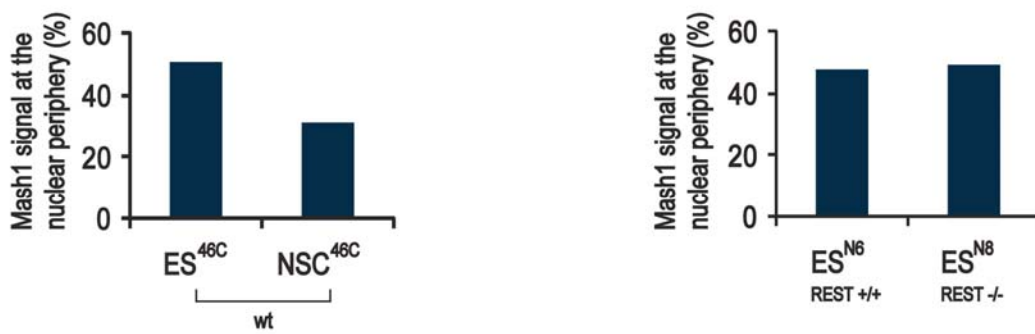


Figure 2.3

a.



b.



c.



2.3. Mash1 nuclear localisation is unchanged in REST deficient cells. The nuclear location of Mash1 was determined by 3D-FISH in embryonic stem cells, ES-derived neural stem cells and in REST

knock-out cells (Rest  $-/-$ , N8 and Rest  $+/+$ , N6 as control cells). a. Representative images of single focal sections for each cell type. Scale bars: 2 $\mu$ m. b. The bar chart shows the percentage of peripheral alleles in each cell type. c. The bar chart represents the percentage of cells with both alleles scored as peripheral (P.P.), peripheral/internal (P.I.) or both internal (I.I.). The data are representative of a single experiment.

### **2.3.2. The epigenetic state of Mash1 is unchanged in REST deficient cells**

Late replication timing and low histone acetylation levels are also characteristics of the repressive chromatin state found at the Mash1 locus in ES cells. We asked whether REST was important in the regulation of such local chromatin marks or the large scale feature of the timing at which the locus replicate. As shown in the paper by Jorgensen and colleagues, no differences in the replication timing or histone modifications pattern were detected at the Mash1 gene between REST knock out and their wild type counterpart ES cells (Fig. S.1). Moreover, no binding of the REST protein was detected at the responsive element 50Kb downstream the promoter region in ES cells (Jorgensen, Terry et al. 2009). In conclusion our analysis does not support the claim that Mash1 is regulated by REST in ES cells and further studies show that REST protein is not necessary in maintaining also other neural genes repressed in ES cells (Jorgensen, Terry et al. 2009).

## **2.4. Role of the REST protein in regulating neuronal genes in ES cells**

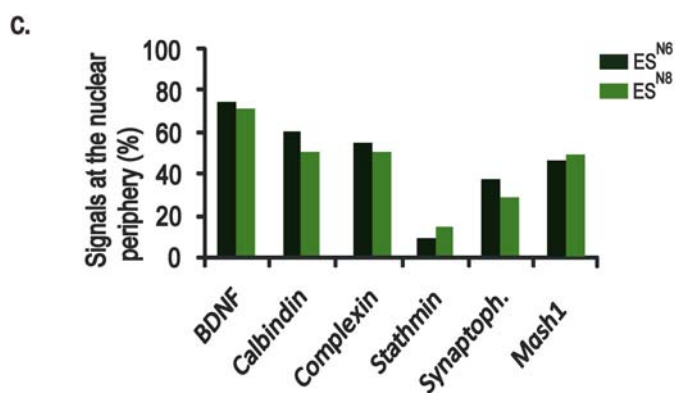
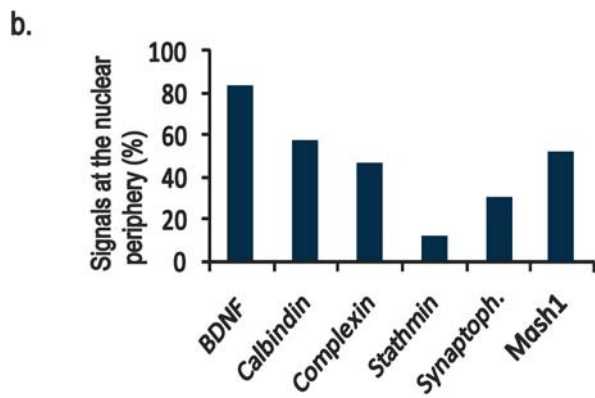
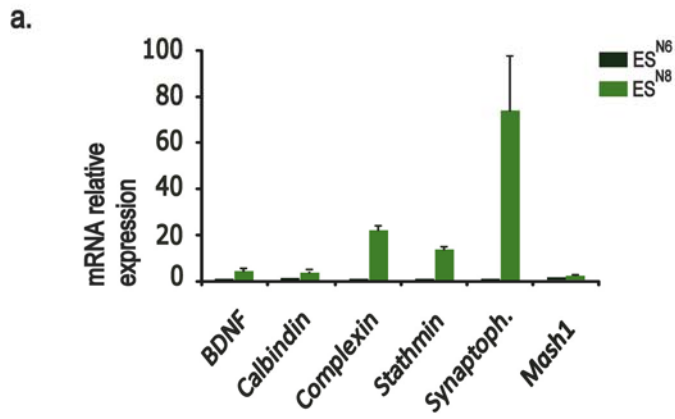
REST has also been implicated in the regulation of neuronal specific and pluripotency genes in ES cells (Wu and Xie 2006; Johnson, Teh et al. 2008) and the effect of REST deficiency in ES cells is controversial. Reduction of REST levels in ES cells has been claimed to cause loss of pluripotency and expression of early markers associated with early differentiation towards the three germ layers (Singh, Kagalwala et al. 2008). We thus asked what was the role of REST in regulating neuronal genes in ES cells, and whether it has a role in their nuclear organisation relative to the periphery.

First, canonical neuronal specific REST targets were considered. Gene expression profile in REST knock out and REST knock down cells in combination with analysis of published data on REST binding in ES cells showed that REST does not bind early markers of neural differentiation and it is not required for their repression (Jorgensen, Terry et al. 2009). The genome-wide expression profiles were then used to identify REST targets genes. Genes with an altered expression profile both in REST knock out as well as knock down cells were mainly brain related genes and genes with a REIelement

close to the TSS (transcription start site) (Jorgensen, Terry et al. 2009). Some of these genes were tested in gene expression analysis by qPCR to verify that these are upregulated in REST knock out cells as compared to wild type (Fig. 2.4a). Although REST depletion in ES cells determine misregulation of genes important in neuronal differentiation, REST depletion does not seem to impair stem cell function as these ES cells still express pluripotency marker and maintain functional properties such as multilineage potential and reprogramming capacity (Jorgensen, Terry et al. 2009).

To ask whether REST depletion affects the nuclear location of these REST target genes, I performed 3D FISH analysis in wild type and REST  $-/-$  ES cells. Using gene specific FISH probes, we found that these genes have a different location relative to the nuclear periphery, with BDNF being more closely located to the nuclear periphery and Stathmin positioned mostly away from that compartment (Fig. 2.4b). To ask whether REST was important for the peripheral localisation of BDNF, Calbindin and Complexin, 3D FISH analysis was performed in REST knock out cells and their counterpart wild type ES cells. No differences were detected in the nuclear localisation between the two cell types (Fig. 2.4c). We also look at the nuclear location of Stathmin and Synaptophysin, more centrally located in ES cells to ask whether there was a change in their nuclear location after REST depletion. Similarly to the previously genes, no changes were observed. We conclude that REST is not important for the nuclear positioning of neuronal genes in ES cells.

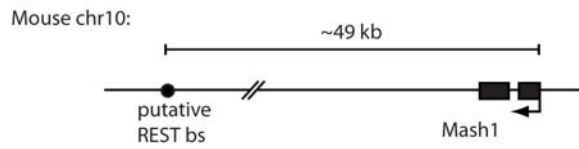
Figure 2.4



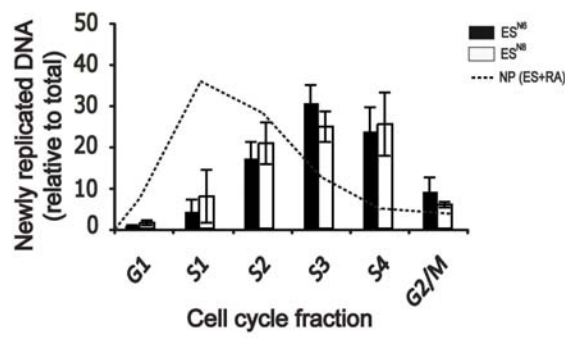
2.4. Analysis of nuclear localisation and gene expression of REST targets in REST deficient cells. a. RT-PCR analysis of increase in mRNA levels of RE1-containing genes in REST knock out cells ( $ES^{N8}$ ) relative to REST wild type cells ( $ES^{N6}$ ). The expression level was normalised to housekeeping genes and is shown relative to wt ES cells. Error bars indicate s.d. between two different experiments. b. Nuclear localisation of Mash1 and RE1-containing neuronal genes in ES cells. The bar chart shows the percentage of peripheral alleles for each gene in 46C ES cell line. The data are representative of a single experiment. c. Nuclear localisation of Mash1 and a subset of RE1-containing neuronal genes is not affected in REST deficient cells. The bar chart shows the percentage of peripheral alleles in REST wild type cells ( $ES^{N6}$ ) and knock out cells ( $ES^{N8}$ ) for each gene. The data are representative of a single experiment.

Figure S.1 (Jorgensen, Terry et al., 2009)

a.



b.



c.

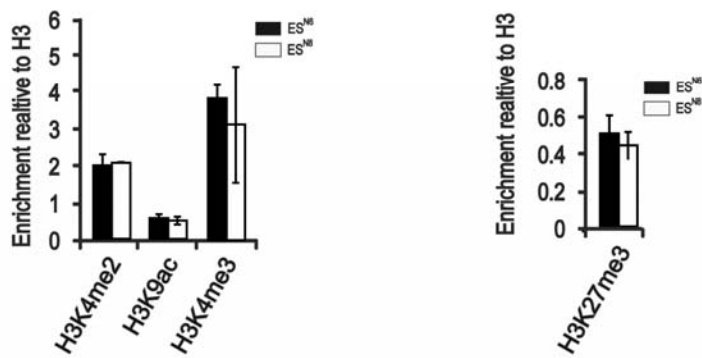


Fig. S.1: Analysis of Mash1 epigenetic features in REST knock out cells. a. Schematic representation of the Mash1 gene and 50 Kb genomic region downstream its TSS. The arrow indicates the TSS and the black boxes the exons. The putative binding site for REST is indicated. b. Replication timing analysis for Mash1 in REST knock out cells (ES<sup>N8</sup>) and wild type (ES<sup>N6</sup>). The fractions of newly synthesized DNA from G1, several phases of S, G2/M is represented. c. ChIP analysis assessing the level of modified histones for the Mash1 locus are represented in wild type (ES<sup>N6</sup>) and knock out cells (ES<sup>N8</sup>). Error bars represent the deviation standard between two different experiments.

## Chapter 3

### Role of CTCF and cohesin in the regulation of Mash1

#### 3.1. Introduction

Higher order conformations of chromatin modulate the binding of transcription factors and create environments that are permissive or repressive for transcription (Chambeyron and Bickmore 2004). The formation of such structures is supported by loops between distant regulatory elements that can reside on the same chromosome (in cis) or on different chromosomes (in trans) and are mediated by specific proteins such as transcription factors or insulators (Dorman, Bushey et al. 2007).

CTCF is an insulator protein that has been shown to mediate the formation of loops between distant regulatory elements. Such looping can serve to prevent the access of an enhancer region to its regulatory promoter (enhancer-blocking activity) or the spreading of silencing heterochromatin (barrier insulator activity) (Zlatanova and Caiafa 2009). The ability of CTCF to create loops has been linked to the presence of cohesin that can constrain chromatin in higher order chromatin structures (Hadjur, Williams et al. 2009). Cohesin generally colocalises at sites of CTCF binding in the mammalian system and is required for the insulator binding activity of CTCF (Parelho, Hadjur et al. 2008; Wendt, Yoshida et al. 2008). The nuclear periphery is generally depleted of CTCF proteins but enrichment is seen at the borders of LADs and at sites of differential expression when HDACs activity is inhibited (Brown, Kennedy et al. 2008; Guelen, Pagie et al. 2008).

Here I look at the role of CTCF and cohesin in the regulation of the Mash1/Ascl1 locus in ES cells. When ES cells differentiate into neural progenitors, differential expression of genes is seen within the locus with genes downstream of Mash1 being transcribed while those upstream, remaining silent (Fig. 3.1a). This behaviour suggest the presence of an insulator between the Mash1 gene and the PAH (Williams, Azuara et al. 2006). We ask whether CTCF and cohesin binding is present across the locus and it is responsible for keeping Mash1 in a repressed state in ES cells close to the nuclear periphery.

#### 3.2. CTCF and cohesin binding at the Mash1 locus in ES cells

To identify potential binding sites for CTCF and cohesin at the Mash1 locus in ES cells, previous data from ChIP/CHIP and ChIP/Seq experiments were aligned (Fig. 3.1b and c). Data for CTCF binding



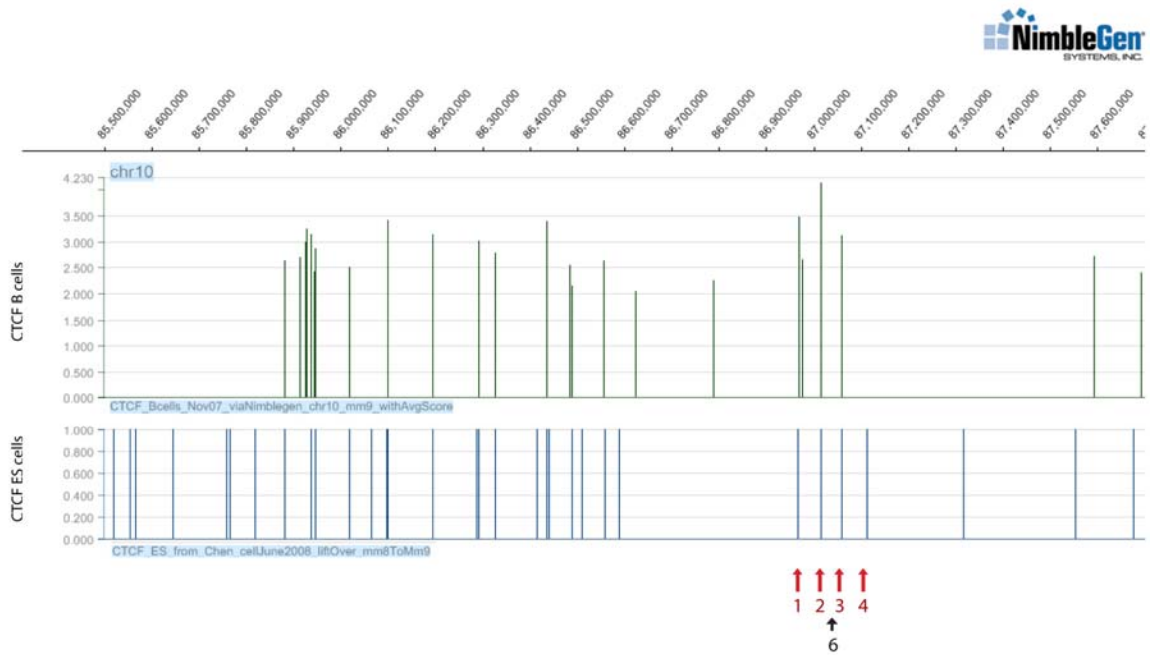
was available in B cells and ES cells (Chen, Xu et al. 2008; Parelho, Hadjur et al. 2008)) while data for Rad21 binding was available for B cells and T cells (Parelho, Hadjur et al. 2008) (Fig. 3.1 b and c). Primers were designed for the regions 1-4 (red), 6-7 (black) and used to assess CTCF and cohesion binding in ES cells by ChIP analysis. Anti-IgG was used as a negative control. 46C ES cells were then harvested and fixed for CTCF and cohesin cross-linking. DNA precipitation and sonication was followed by immunoprecipitation of those fragments where CTCF and cohesin protein was bound. PCR amplification was performed using the primers designed in the region of interest and negative controls in regions where we do not expect binding. To confirm the efficiency of the immunoprecipitation, control primers, previously verified (Parelho, Hadjur et al. 2008) were used as positive and negative control for CTCF and cohesin binding. In conclusion, I was able to detect both CTCF and Rad21 at the peaks of interest 1-4 while peak 6 was a negative control in 46C ES cells (Fig. 3.2).

Figure 3.1

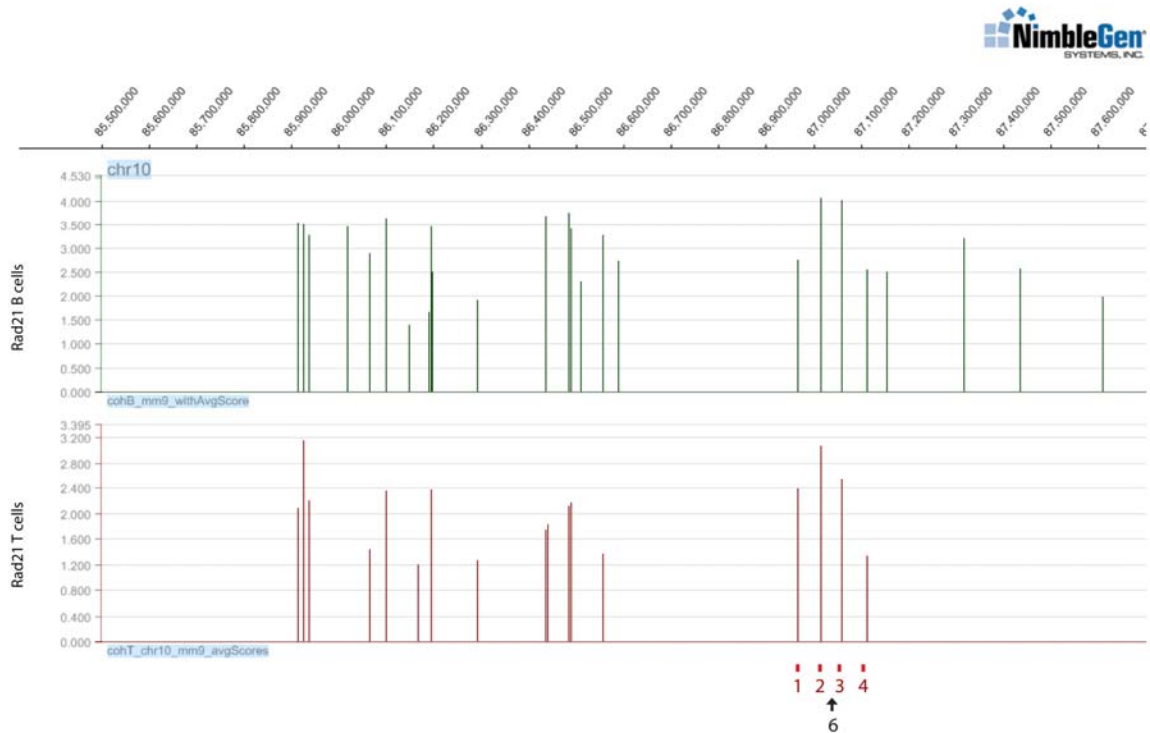
a.



b.

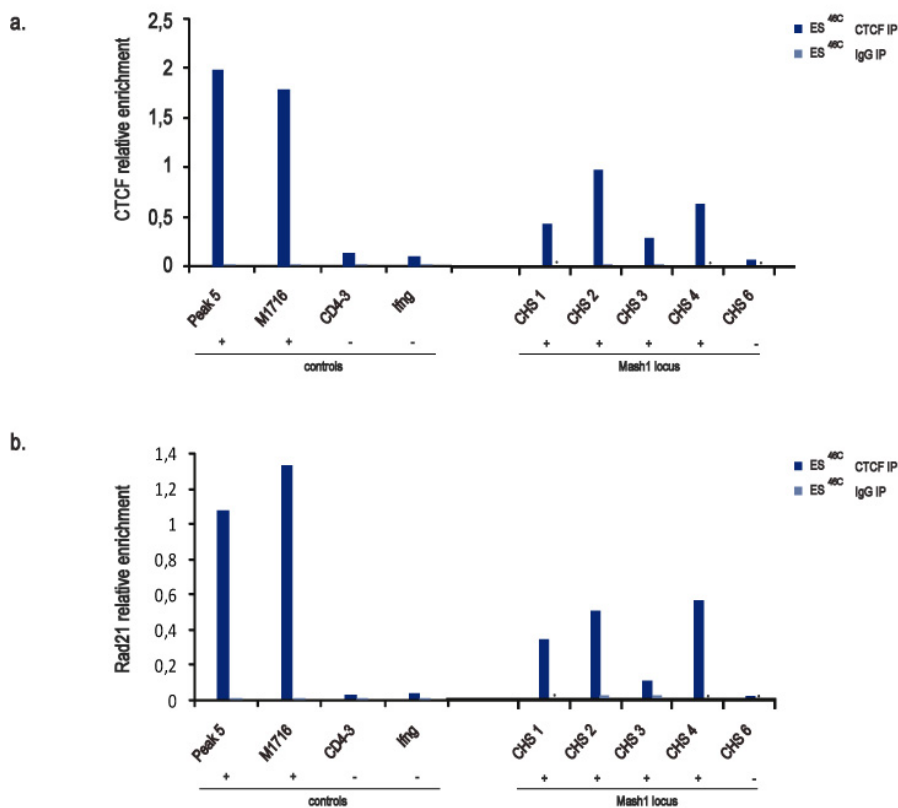


c.



3.1. CTCF and Rad21 binding within a 1.2 Mb region surrounding the Mash1 locus. a. Peaks represent CTCF binding sites as determined by ChIP-CHIP data in B cells (Parelho et al., 2008, Cell) and ChIP -Seq data in ES cells (Chen et al., 2008, Cell). b. Rad21 binding sites as determined by ChIP-CHIP data in B cells and T cells ( Parelho et al., 2008, Cell). The position of PCR primers amplifying regions of interest for CTCF or Rad21 binding (red arrows) and region of non-binding as negative controls are shown (black arrows). In the upper panel dark blocks are a schematic representation of genes in the Mash1 locus.

Figure 3.2



3.2. Binding of CTCF and Rad21 at specific sites within the Mash1 locus in ES cells. a-b. 46C ES cells were fixed with formaldehyde, the chromatin sonicated and immunoprecipitated with anti-CTCF, anti-Rad21, anti-IgG (control). The final DNA was amplified by qPCR with the primers shown in fig. 3.1. Bar chart representing the enrichment of CTCF (a) and Rad21 (b) relative to Input in ES cells (blue). Control IgG for CTCF ChIP (red) and Rad21 ChIP (violet) are shown with a mark when not detected.)

### **3.3. Analysis of Mash1 regulation by siRNA of CTCF and cohesins in ES cells**

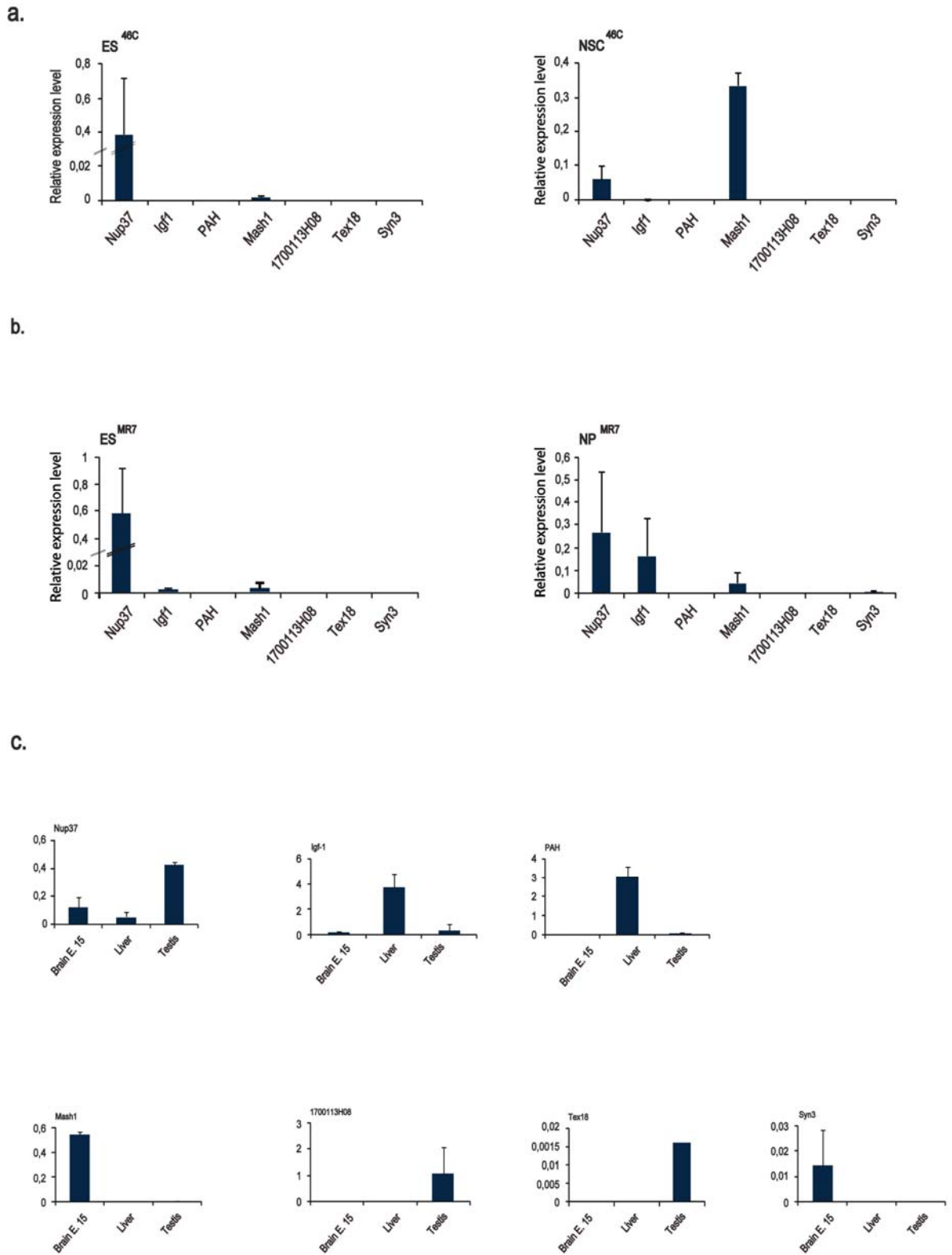
To address the roles of CTCF and cohesin in the regulation of Mash1 I decided to knock down protein expression of these two genes using siRNA (previously shown in (Parelho, Hadjur et al. 2008)). We aim to see whether CTCF and cohesin are responsible for maintaining Mash1 in a repressive chromatin state in ES cells and to see whether the depletion of these important proteins might lead to a misregulation of the gene.

#### **3.3.1 Gene expression profile at the Mash1 locus upon neural differentiation**

During retinoic acid differentiation, ES cells formed multicellular aggregates with the induction of progenitors from all the three germ layers. Retinoic acid and drug selection using specific marker for neural induction allow the isolation of neural progenitors (Billon, Jolicoeur et al. 2002). Mash1 is transcriptionally upregulated during retinoic acid differentiation and a differential expression is observed for the genes present in a 1.2 Mb region, with upregulation of genes downstream the Mash1 TSS and no change in the transcription of Igf1 and PAH upstream (Williams, Azuara et al. 2006). This behaviour suggests the presence of an insulator region between the Mash1 gene and the upstream gene PAH, where the CTCF and cohesin proteins might exploit their insulator activity. We aim to see whether the same expression profile was observed in neural stem cells, a self-renewing population that uniformly express Mash1 and in neural progenitors derived with an alternative method where only neural specific genes were induced. Gene expression analysis shows upregulation of Mash1 both in neural stem cells and neural progenitors derived with the PA6-coculture system, but no upregulation of the testis specific genes 1700113H08 and Tex18 downstream Mash1. A slight upregulation was observed for Synaptotagmin 3 (Syn3) in neural progenitors. Also the Igf1 gene, in contrast with the result observed with the retinoic acid differentiation, was expressed in neural progenitors. However, PAH was found silent in both neural stem cells and neural progenitors as in

retinoic acid differentiation confirming a differential expression pattern between Mash1 and PAH and suggesting the presence of an insulator in this region (Fig. 3.3).

Figure 3.3



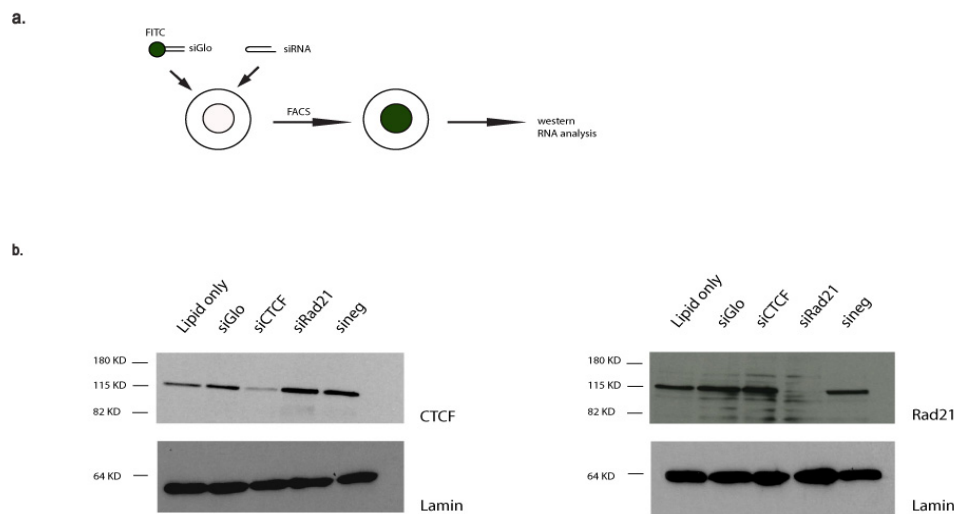
3.3. Gene expression profile of genes surrounding the Mash1 locus in ES cells, ES-derived NSC and ES-derived NP. Bar chart represent the mRNA level of genes relative to housekeeping HMBS-5 and YWHAZ in 46C ES cell line and ES-derived neural stem cells (a), MR7-ES cell line and ES-derived Neural Progenitors (b), control tissue ex-vivo neural progenitors, liver and testis (c). Error bars represent the standard deviation of duplicate experiments.

### **3.3.2. Mash1 expression profile in cells depleted of CTCF and cohesins**

Transfection of ds-RNA oligos can be used to efficiently knock down mRNA production by a process called RNA interference. We aim to knock down CTCF and cohesin in 46C cell line where Mash1 is expressed at very low levels.

Indicators to evaluate the efficiency of siRNA are ds-oligos carrying a fluorescence tag FITC or Alexa (siGlo) that can be used to sort only the cells where transfection had occurred. 46C cells were plated 4 hours before transfection on gelatine coated dishes and then transfected with siGlo indicator plus si-oligos directed against CTCF, the Rad21 subunit of the cohesin complex and a control gene. After 48hours, almost 80% of the cells were positive for FITC expression, via analysis by fluorescence-activated cell sorting (FACS) and collected for western blot and RNA analysis (Fig. 3.4a). Immunoblotting using a mouse antibody against CTCF and Rad21 show efficient protein knock down 48hours after transfection. Control cells were unaffected and only a low amount of CTCF protein was detectable after CTCF knock down(Fig 3.4b).

Figure 3.4

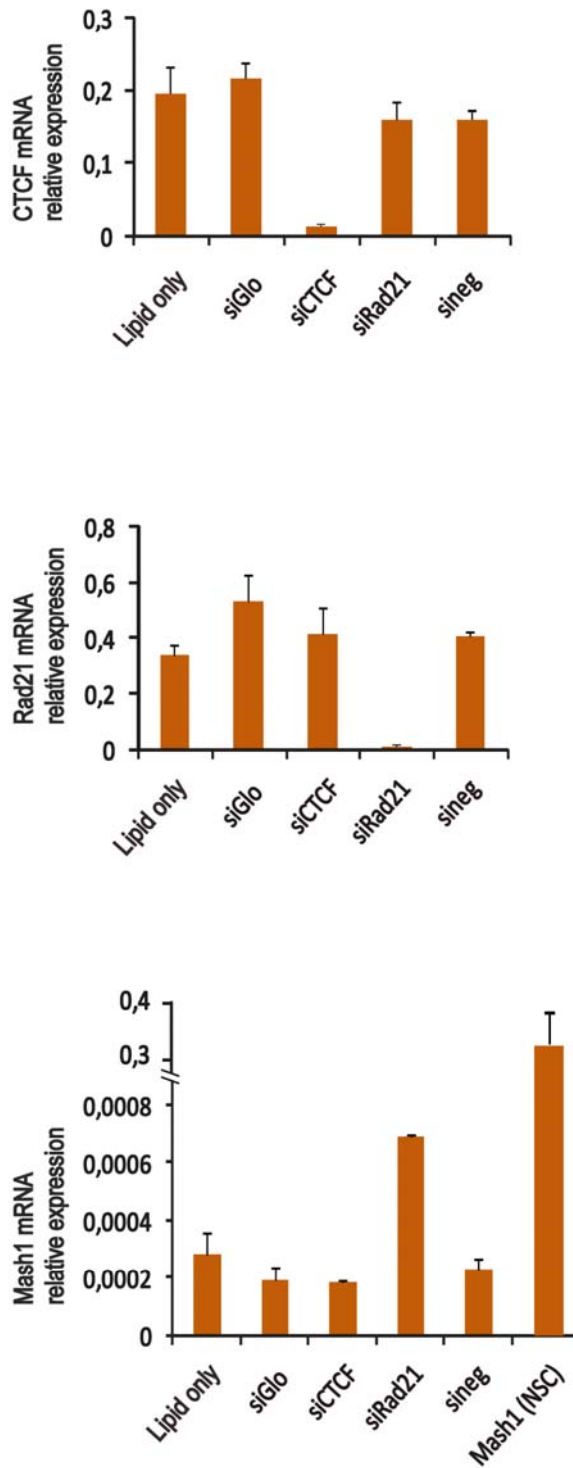


3.4. CTCF and Rad21 knock down in 46C ES cell line. A. Schematic representation of the procedure used to knock down CTCF and cohesin. SiGlo indicators shown in green (FITC labelled). b. Western Blot analysis of CTCF and Rad21 knock down in ES cells. Whole extracts from 46C ES cells transfected with Dharmafect lipid only, FITC-siGlo indicator, FITC-siGlo indicator plus siCTCF, siRad21 and sicontrol oligonucleotides were subjected to SDS-PAGE electrophoresis. Membranes were hybridized with anti-CTCF, anti-Rad21 and anti-lamin as loading control.

Analysis of the mRNA level show a decrease in mRNA around 90% for both CTCF and Rad21, while Mash1 transcripts are 2-3 fold upregulated after Rad21 knock down (Fig. 3.5a). To verify this effect was specific for the Rad21 knock down and to reduce the amount of protein left, sorted cells were replated for a second round of transfection. Looking at the rate of cell division by microscopy, a decrease in the efficiency of cell replicating was observed after Rad21 knock down while CTCF knock down cells divide normally. A second transfection with FITC-siGlo indicator and siCTCF, Rad21 and sicontrol was performed and cells collected for FACS analysis. A high number of dead cells were present when Rad21 was knocked down but also a decrease in the number of viable cells was observed with CTCF knock down, compared to control. In conclusion, CTCF and Rad21 mRNA level was reduced by around 90% and a higher upregulation of the Mash1 transcript was observed after Rad21 knock down (Fig. 3.5b).

Figure 3.5

a.



Student TTest

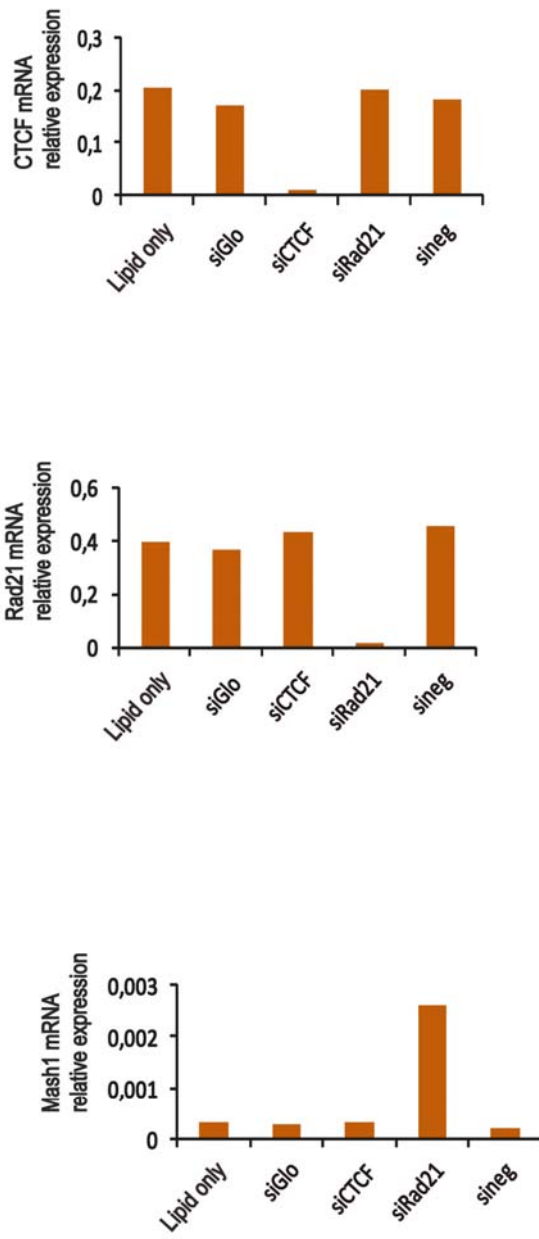
siCTCF/siNeg = 0.118

siRad21/siNeg = 0.015



Figure 3.5

b.



3.5. Gene expression analysis of Mash1 after CTCF and Rad21 knock down in ES cells. a. Bar chart representing the mRNA level of CTCF (top), Rad21 (middle) and Mash1 (bottom) relative to housekeeping genes HMBS-5 and YWHAZ. 46C ES cell line transiently transfected with Dharmafect lipid only, FITC-siGlo indicator alone and in cotransfection with siCTCF, siRad21 and sicontrol oligonucleotides. Cells were sorted for FITC expression and samples analysed for gene expression. Student TTest for the Mash1 upregulation in CTCF and Rad21 knock down is shown. b. Sorted cells for FITC-siGlo indicator were replated and subjected to a second CTCF, Rad21 knock down. Gene expression analysis for CTCF, Rad21 and Mash1 relative to housekeeping HMBS-5 and YWHAZ is shown for this second knock down.

## Chapter 4

### Generating and characterising ES cells transgenic for the Mash1 locus

#### 4.1. Introduction

A 3 Mb genomic region surrounding the endogenous Mash1 gene repositioned in the nucleus of ES cells upon neural differentiation. None of the genes in this domain encode neural specific proteins. However the epigenetic changes that occur are orchestrated at the *Mash1* locus (Fig.1.2) (Williams, Azuara et al. 2006). Studies on mice carrying a *Mash1*-GFP transgene showed that 140 Kb of genomic region surrounding the *Mash1* locus is sufficient for correct expression of the transgene in the embryo (Parras, Hunt et al. 2007). Here we aim to verify if this genomic region contains also the information required for the peripheral localisation of the transgene when not expressed and its movement away the nuclear envelope when activated.

Cells derived from the Mash1-GFP transgenic mice showed a single multicopy integration site of the transgene in the genome. Preliminary data suggested a similar behaviour between transgene and endogenous, with the transgene located at the nuclear periphery in B cells where Mash1 is not expressed and relocated to the nuclear interior in ex vivo neural progenitors (Chiara Beretta, data not shown). Here we aim to create ES cells with a single copy integration of the Mash1-GFP transgene, to assess whether the transgene confers the ability to influence the ‘epigenetic status’ of neighbouring genes at different sites of integration. We will then aim to delete the responsible regulatory regions by a deletion approach (Fig. 4.1).

#### 4.2. Creation of ES cell lines carrying a transgenic copy of Mash1

A BAC vector, RP24-130P7, encompassing 140Kb the genomic region at the Mash1 gene was used to engineer a 46C ES cell line. The first step involved the insertion of a Neo cassette necessary for selection of positive clones in mammalian cells. For this purpose Mash1-BAC was transformed into a recombineering competent strain SW102 that was used to exchange the CM(R) with the NEO resistance (Simone Alves’ personal work) (Fig. 4.2).

I then linearised the Mash1-BAC carrying a NEO resistance using a rare-cutting enzyme and I electroporated the BAC into 46C ES cells line. Single cells carrying Mash1-BAC integrated in their

genome were able to survive culture with neomycin selection and formed colonies. At day 10, colonies were picked and a cell suspension from each colony was plated in multi-well plate. Clone expansion was carried out for several days without neomycin selection to avoid losing putative clones where the transgene integrated in a silencing region.

## Figure 4.1

### I. Insertion of a mammalian selectable marker into Mash1-BAC (Simone Alves' work)

- Transformation of Mash1-BAC into SW102 strain (recombineering-competent)
- Preparation of NEO cassette
- Recombineering process
- Selection of recombinants and isolation of DNA

### II. Electroporation of Mash1-BAC into ES cells

- Mash1-BAC linearisation with P1-Sce1 rare cutting enzyme
- Mash1-BAC electroporation into 46C ES cell line
- Selection of positive clones

### III. Copy-number screening of positive clones

- PCR analysis
- Southern blot analysis

### IV. Identification of the transgene insertion sites

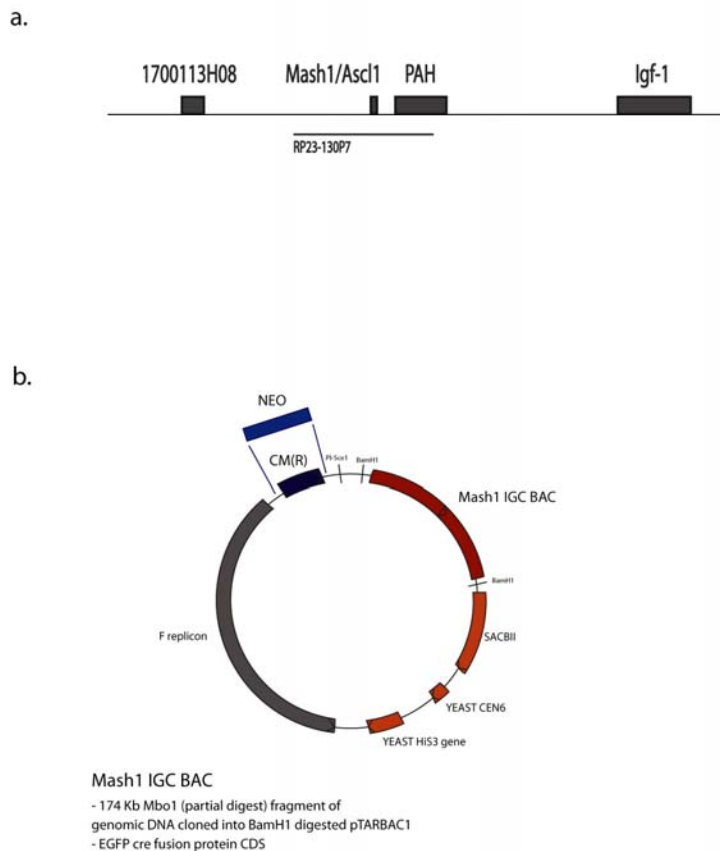
- Nested PCR

### V. 3D-FISH analysis

- Double probe labeling to discriminate between transgene and endogenous

4.1. Studying the effect that an extra copy of the Mash1 has on the neighbouring genes. Key steps for the generation of Mash1 transgenic lines and analysis.

Figure 4.2



4.2. Preparation of a Mash1 transgene to engineer ES cells. a. Schematic representation of Mash1 genomic context with the chosen BAC for electroporation. b. Mash1 BAC map with the NEO marker, suitable for selection in mammalian cells and the replaced CM(R) bacterial marker.

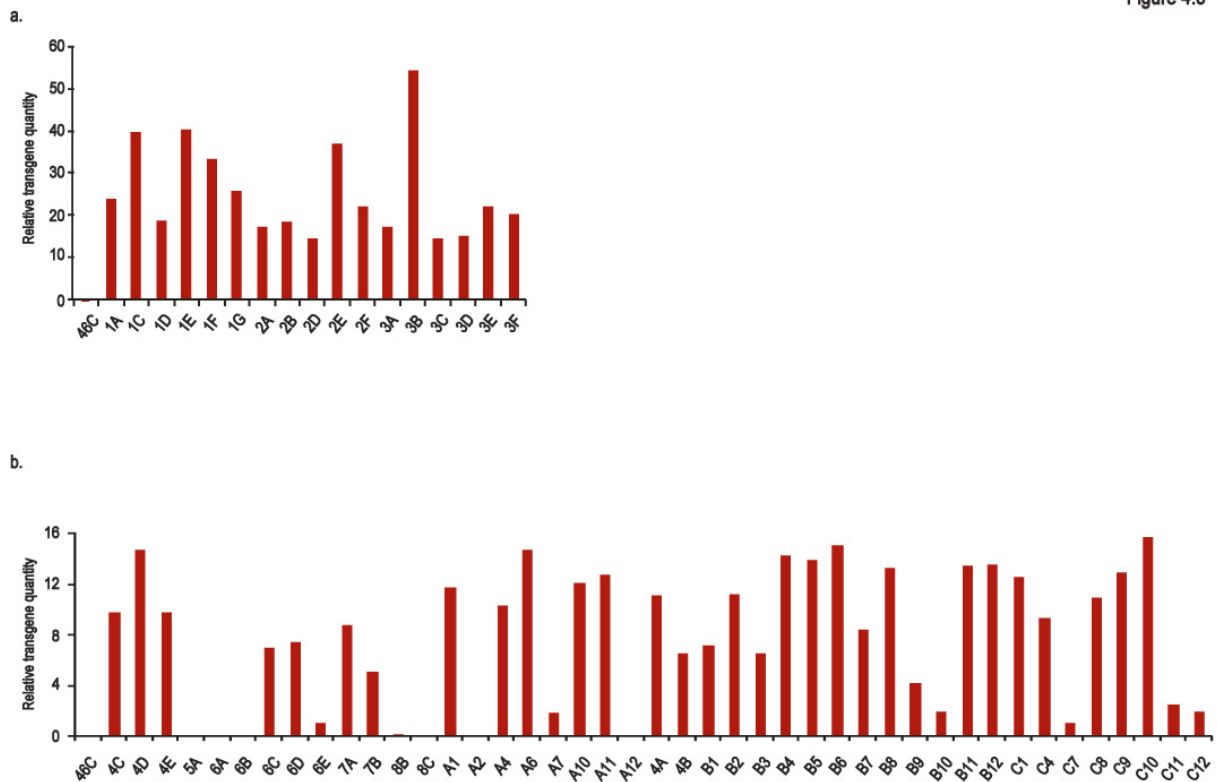
#### 4.2.1 Screening of positive clones

I performed two rounds of electroporation with two different DNA preparations. The first time 17 clones were obtained while the second electroporation gave 43 clones. DNA was then extracted from each clone, purified and analysed by PCR. Primers designed on the neomycin resistance sequence were used to selectively amplify the transgene. Clone products were normalised for the housekeeping gene Flk-1, present in single copy in the genome and 46C ES cell line was used as a negative control.

Some false positive were present as not all the clones give a product by PCR, but most of them were carrying the transgene. This analysis does not provide information on the exact number of transgenic

copies integrated in each clone because the set of primers designed on Neo resistance and the set of primers on Flk1 have different efficiencies. However it is likely that clones from the first round of electroporation have integrated more copies in their genome in a single or multiple site compared to the clones obtained in the second electroporation. In particular, clones 6E, A7, B10, C7, C11 and C12 could have a single or a few copies of transgene and therefore they are interesting for a further analysis (Fig. 4.3).

Figure 4.3



4.3. Screening of Mash1-Neo positive clones by PCR. Set of primers were designed to amplify unique sequences present on the transgene only. Values for each clone were normalised for the single-copy gene Flk-1. 46C line is used as negative control. Bar chart represents the relative transgene quantity in the clones obtained from a first (a) and second (b) round of electroporation.

## Chapter 5

### General discussion

Differentiation of embryonic stem cells can recapitulate in vitro the early steps of cell fate and lineage commitment. During ES differentiation developmentally regulated loci undergoes specific changes at the level of chromatin and nuclear organisation that result in their expression or repression (Arney and Fisher 2004). Here I looked at the regulation of Mash1, a proneural gene, in ES cells and upon ES differentiation into neural cells. During neural differentiation specific epigenetic changes occur at the Mash1 locus in concomitant with the gene activation: acquisition of active histone marks at the promoter region, switch to an early replication timing and repositioning of the locus away the nuclear periphery (Williams, Azuara et al. 2006). It is not known the dynamic of such changes, in particular whether it is the transcription of the genes that drive the locus away the periphery or the relocation precede the gene transcription. It is also not known whether specific repressors keep the locus tethered at the nuclear edge. Artificial tethering of a locus to the nuclear periphery can in fact induce repression with a different degree according to the locus tethered (Finlan, Sproul et al. 2008; Kumaran and Spector 2008; Reddy, Zullo et al. 2008). How the nuclear periphery is involved in gene repression is not clear but it could involve the recruitment of specific repressors or chromatin remodelling complexes. It has been shown that for example an integral membrane Lap $\beta$  interacts with HDAC3 and induces histone deacetylation at the nuclear periphery (Somech, Shaklai et al. 2005). In the case of the Mash1 gene, however, the removal of important chromatin silencing complexes such as Ezh2/Eed HMTase complex does not release the locus from its peripheral positioning (Williams, Azuara et al. 2006).

Here I tried to understand the timing of locus repositioning in relation to gene activation and I look at candidate proteins that might keep the Mash1 locus repressed and at the nuclear periphery in ES cells. First I looked at Mash1 location and gene expression during the progression of neural differentiation to address whether the locus relocates before gene transcription. I used an in vitro system to differentiate ES cells into neural progenitors. At the end of the differentiation 70% of the cells were positive for the neural marker Nestin while the remaining were still Oct-4 positive or non neural. Analysis of Mash1 positioning at different days of neural differentiation was performed in 2D FISH as previously reported for Mash1 location (Williams, Azuara et al. 2006). In ES cells (day 0) 50% of the Mash1 alleles reside in a small portion of the nucleus close to the nuclear edge and the locus is considered peripheral. If the locus was randomly distributed in the nucleus the proportion of alleles in

this peripheral region would be around 22% (Kosak, Skok et al. 2002). When the cells differentiate the number of peripheral alleles decreases. This is consistent with the reported relocation of Mash1 upon neural differentiation and the percentage of alleles still at the periphery could be due to the presence of non neural cells at the end of the differentiation (30%).

The change in the nuclear localisation is specific for the neural fate since it is not observed in T cells or keratinocytes where the gene is silent (Williams, Azuara et al. 2006). A change in morphology was observed when ES cells differentiate to neural progenitors, with the latter having a smaller diameter. It is unlikely that this decrease in nuclear volume in neural cells could determine a shift to a more internal location. Mash1 was in fact observed in a peripheral location in ES cells as well as T lymphocyte or keratinocytes that have a much smaller diameter than ES cells (Williams, Azuara et al. 2006). Therefore Mash1 relocates in the nucleus upon neural induction but from my data in 2D FISH, it is not clear whether relocation precedes gene activation. Proper relocation was in fact observed only at day 6 when cells start to express Mash1. A definitive answer to the question would have come from RNA FISH experiments to see where nascent RNA is present in the nucleus but several attempts to set up the technique failed. Moreover, an analysis in 3D FISH and at smaller intervals could have been more accurate. However 2D FISH was preferred because the fixation of neural progenitors during differentiation is a critical step and it was very difficult to preserve the structure of the nucleus as required in 3D FISH.

In conclusion these results are consistent with the nuclear periphery being a repressive compartment and confirm that Mash1 relocates to the nuclear interior upon neural differentiation. Although it is not certain whether relocation precedes gene expression, my results suggest that Mash1 upregulation does not start at the nuclear periphery since at day 6 when upregulation is observed almost 70% of the alleles are no longer peripheral. This is different from what was shown at the beta-globin locus during erythroid differentiation. Transcription at the beta-globin locus starts at the nuclear periphery and further increases when the locus relocates to a more internal position (Ragoczy, Bender et al. 2006).

REST has been proposed to repress neural genes in ES cells, including Mash1, via the recruitment of different corepressors such as the histone deacetylase complex HDAC. A REST binding site has been identified 50Kb downstream the Mash1 promoter (Ballas, Grunseich et al. 2005). Here I looked at the role of REST in regulating Mash1 and other neural genes in ES cells. I asked if REST is important in keeping Mash1 repressed and at the nuclear periphery in ES cells. Data in 3D FISH showed no changes in Mash1 peripheral location when REST was depleted and no changes were also observed in the gene transcription. Analysis of peripheral location was first performed considering the % of alleles in the population residing at the nuclear periphery to see whether lack of REST would determine a



shift in the number of peripheral alleles. The analysis was then repeated considering the position of the two alleles in each cell, in this case to ask whether REST confers a specific distribution of the two alleles inside the cell, with both alleles located at the periphery or one allele more internally located. Both analysis showed no changes in Mash1 distribution in REST depleted cells compared to wild type. Moreover, as shown in Jorgensen's paper, no change in replication timing and chromatin modifications were observed at the Mash1 locus in REST depleted cells. These results, together with a lack of REST binding at the Mash1 locus suggest that REST is not important in the regulation of Mash1 in ES cells (Jorgensen, Terry et al. 2009). This is consistent with genome-wide studies of REST binding in ES cells and neural cells that failed to identify Mash1 as a REST target gene (Sun, Greenway et al. 2005; Otto, McCorkle et al. 2007). It is important to mention here that the locus is late replicating in ES cells and REST knock out cells where 50% of the alleles reside at the nuclear periphery. It is possible that such % of alleles at the nuclear periphery is sufficient to confer general late replication of the locus, or that the late replication timing is in the case of Mash1 independent from its nuclear location.

Further studies carried out in our laboratory showed also that REST does not target other early neural specifying gene in ES cells, such as Sox1, Ngn1-2, and it is not required for maintaining ES state of pluripotency. In ES cells, REST targets are specific neuronal genes that are silenced in ES cells and express later on in mature neurons (Jorgensen, Terry et al. 2009). ES cells lacking REST misexpressed genes important in terminal neuronal differentiation but still maintain their stem cell identity as demonstrated by the expression of pluripotency markers, potential to differentiate in different lineages and ability in reprogramming (Jorgensen, Terry et al. 2009). This is in line with other analysis of REST depletion in ES cells and neural cells where no changes in the expression of pluripotency markers or early specifying genes were observed (Buckley, Johnson et al. 2009).

We asked how REST regulates neuronal genes in ES cells, whether it targets them at the nuclear periphery. First, I looked at the nuclear positioning of some of the identified REST target genes in ES cells to see if they were preferentially located at the nuclear periphery, but while BDNF, Calbindin and Complexin were more peripheral, Stathmin and Synaptophysin, were internally located. When REST was depleted I observed no changes in their nuclear localisation, but among those that were more derepressed upon REST withdrawal Stathmin and Synaptophysin had an internal location. Complexin, on the other hand, showed derepression albeit located at the nuclear periphery. This is consistent with the nuclear interior being favourable for transcription and with the presence of different domains at the nuclear periphery (Deniaud and Bickmore 2009). While BDBF and Calbindin are located in repressive domains, Complexin could be situated in a more permissive domain at the nuclear periphery.

CTCF and cohesin have been implicated in the formation of long range chromatin conformations and the generation of domains that favours or represses transcription (Wendt and Peters 2009; Zlatanova and Caiafa 2009). Binding of CTCF has been detected at the borders of LADs domains and at differentially expressed genes at the nuclear periphery when HDAC is inhibited (Brown, Kennedy et al. 2008; Guelen, Pagie et al. 2008). A differential expression region is present at the Mash1 locus between Mash1 and PAH: upon neural induction Mash1 and genes downstream are upregulated while PAH and Igf-1, upstream, remain silent, suggesting the presence of an insulator (Williams, Azuara et al. 2006). Here I asked whether CTCF and cohesin bind to the Mash1 locus, in particular between Mash1 and PAH and whether they are responsible for keeping Mash1 in a repressed chromatin environment at the nuclear periphery in ES cells.

First I aligned data for CTCF and cohesin binding from ChIP/CHIP and ChIP/Seq data previously reported (Chen, Xu et al. 2008; Parelho, Hadjur et al. 2008) and CTCF and cohesin binding at the Mash1 locus was confirmed in 46C ES cell line by ChIP analysis.

I asked whether depletion of CTCF or cohesin by RNAi might result in the Mash1 derepression in ES cells or Mash1 misregulation in neural progenitors/neural stem cells. I first performed gene expression analysis of the ES and neural cells I used in my previous experiments. MR7-PA6 differentiation was chosen because it specifically induces neural genes, while NSC derived from the monolayer differentiation of 46C give rise to self-renewing cells that uniformly express Mash1. It is important to highlight that neural progenitors derived with the first system are multipotent progenitors, that will give rise to dopaminergic neurons (Kawasaki, Mizuseki et al. 2000) while neural stem cells on the other end are multipotent progenitors that can self-renew indefinitely and can give rise to all the types of neurons and glia (Conti, Pollard et al. 2005). Here I aimed to see if Mash1 exhibit the differential expression profile observed with retinoic acid differentiation in both the two differentiation system.

Gene expression profiles reveal differences with the data obtained with the retinoic acid differentiation, with the upregulation only of those genes that are specific or important for the neural specification. *Tex18* and *1700113H08* resulted to be testis specific and not expressed in neural progenitors or neural stem cells derived with the two methods. It is possible that retinoic acid receptors are present close to these genes and the induction of the three germ layers during the first days of retinoic acid differentiation lead to the gene upregulation. *Igf-1* on the other end was expressed in neural progenitors and at low level in neural stem cells. This is in line with the role of *Igf-1* in neural cells (Ye and D'Ercole 2006) and the requirement of insulin in the growth medium of neural stem cells. This gene expression profile suggest the presence of more than one differential expressed site at the Mash1 locus and confirm the difference in expression between Mash1 and PAH in all the three different differentiation systems, suggesting the presence of an insulator in this region.

It is important to remember here that this region showed the highest switch in replication timing and nuclear positioning.

To see if CTCF and cohesin binding in this insulator region could determine the silencing of the locus and the positioning at the nuclear periphery in ES cells, I performed knock down of CTCF and cohesin in 46C ES cells. A change in gene expression was observed when Rad21, a subunit of the cohesin complex, was knocked down. Mash1 transcript was slightly upregulated compared to negative control, suggesting a role for cohesin in regulating Mash1 expression. The same small upregulation was not observed after CTCF knock down but western blot analysis showed that there was still CTCF protein left after knock down. It could be possible that the amount of CTCF protein left is responsible for the different behaviour observed between CTCF and cohesin knock down. The upregulation of Mash1 after Rad21 knock down was very small compared to the expression level it is observed in neural progenitors, but significant as confirmed by Student T Test (Fig. 3.5a). However, it is not possible here to rule out here an effect of a delayed cell cycle on Mash1 derepression in Rad21 knock down since cells stop cycling. Cohesin play in fact an important role during cell cycle and depletion of cohesin result in interruption of the cell cycle (Uhlmann, Lottspeich et al. 1999). Moreover, Mash1 expression is sensitive to the oxido-reductive state of the cells (Prozorovski, Schulze-Topphoff et al. 2008). A good control to address this possibility would be to block the cell cycle in ES cells and see whether the same activation is visible.

Finally, it would be also interesting to look for changes in the nuclear positioning of Mash1 when CTCF and cohesin are knocked down to see whether Mash1 is released from the nuclear periphery after depletion of such important regulators.

In collaboration with Simone Alves, we generated ES cells carrying low copy number of Mash1 transgene integrated at different sites of the genome. We aimed to look at the ability of the transgene to influence the 'epigenetic status' of neighbouring genes and identify the responsible regions by deletion approach. Previous work on transgenic mice suggested that the transgene contains all the information required for the correct expression in the embryo (F. Guillemot personal communication) and preliminary analysis in cells derived from transgenic mice showed that the transgene is able to locate at the periphery in B cells where Mash1 is not expressed (data not shown). This transgene contain 140Kb of genomic region surrounding the Mash1 gene while the same epigenetic features that characterise the Mash1 gene are present in a 3Mb region of the Mash1 locus. We asked whether the transgene was able to locate at the nuclear periphery in ES cells independently from its integration site and it was then able to relocate to the nuclear interior upon neural differentiation. We also wanted to test whether this 140Kb region could exert its influence on a random chosen genomic context.

In the case of X-chromosome inactivation, spreading of silencing from the XIC (X inactivation centre) is thought to be facilitated by the presence of silencing elements along the chromosome. These elements are rich in LINE-1 repeats and are enriched on the X chromosome. Silencing by the XIC can also be induced on autosomes by the integration of a Xist transgene. Gene-rich regions poor in LINES are more resistant to silencing compared to gene-poor region rich in LINES (Tang, Huntley et al.). We asked whether Mash1 could represent a repressive element that maintains its repressive potential in ES cells also when integrated in another part of the genome and that is able to influence its genomic surrounding. Future works will involve the characterisation of the ES transgenic cell line and the dissection of the effects that the transgene has on the integration regions by ChIP, FISH and replication timing analysis.

## Chapter 6

### Material and methods

#### 6.1 Cell culture

MR-7 ES cells (Oct4GOF18ΔPE), with a GFP inserted downstream the Oct4 gene were maintained on irradiated primary embryonic fibroblasts on gelatine coated dishes (0.1% gelatine in PBS) in Knock-out D-MEM supplemented with 10% FCS, 2mM glutamine, 0.1mM non essential amino acid, 1mM pyruvate, 0.1mM 2-mercaptoethanol, penicillin-streptomycin and 1000U/uL leukaemia inhibitory factor (LIF) (ES Medium plus supplements). Primary embryonic fibroblasts were cultured in D-MEM with 10% FCS, 2mM glutamine, 0.1mM non essential amino acid and penicillin-streptomycin. PA6 cells, a stromal cell line derived from newborn mouse calvaria, were cultured in D-MEM with 10% FCS.

46C ES cell line (Sox1-GFP-ires-Pac) with a double cassette for GFP expression and puromycin selection, under the control of Sox1 promoter, were cultured in gelatine coated dishes and maintained in ES Medium plus supplements and 1000U/uL LIF. Neural stem cells (NS<sup>MK1</sup>) derived from 46C ES cells using the monolayer differentiation system (Conti et al., 2005) were maintained in NS-expansion medium (Euromed-N medium supplemented with modified N-2 (Ying et al., 2003) and 10 ng/mL FGF2, 10 ng/mL EGF).

#### 6.2 Neural differentiation

Neural differentiation of MR-7 ES cells was performed as previously described (Kawasaki et al. 2000). Briefly, the day before differentiation, PA6 stromal cells were irradiated, plated on 10 cm<sup>2</sup> gelatine coated dishes and allowed to attach over-night. At day 0 of differentiation, ES cells were collected eliminating the irradiated PEFs, washed and 2x10<sup>3</sup> cells were plated on PA6 feeders in D-MEM supplemented with 10% KNOCKOUT serum replacement, 2mM glutamine, 0.1mM non essential amino acid, 1mM sodium pyruvate, 0.1mM 2-mercaptoethanol and penicillin-streptomycin (Differentiation medium). After day 4, the medium was replaced daily. Cells were harvested by trypsination, after eliminating irradiated PA6, and collected at day 2, 4, 6 and 8 for FISH, immunofluorescence and RNA extraction.

#### 6.3 Preparation of primary embryonic fibroblast (PEFs) for ES cell culture

Mouse embryos from embryonic day 15 (E15) were used to derive primary embryonic fibroblast for ES cell culture. Embryos were dissected, organs and head removed, cell suspension created with a

syringe and plated in DMEM (GIBCO) supplemented with 10% FCS, 2mM glutamine, 0.1mM non essential amino acid, 1mM pyruvate, 0.1mM 2-mercaptoethanol and penicillin-streptomycin. After a few passages cells were collected and frozen. PEFs were expanded in culture for a maximum of 4 passages and irradiated in IBL cell irradiator (3000 Rad). Feeder cells were plated on gelatine-coated dishes at least 4 hours before plating ES cells.

#### **6.4 Immunofluorescence staining**

Cells were trypsinised, washed in PBS and resuspended in PBS or the appropriate cell medium. About 60uL of cell suspension ( $0.2 \times 10^6$  cells) were let attached to Poly-L-Lysine coated coverslips for a few minutes, washed and fixed in 2% paraformaldehyde/PBS for 20' at room temperature. After the fixation cells were washed twice in Ca+/Mg PBS and immunofluorescence was carried out as previously described (Perry et al., 2005). Briefly, cells were permeabilised in 0.4% Triton X-100/PBS for 5', followed by washes in PBS and washing buffer (0.2% BSA, 0.05% Tween-20 in PBS). The coverslips were then incubated in a humid chamber in blocking buffer (2.5% BSA, 0.05% Tween-20, 10% NGS in PBS) for half an hour before incubation with specific antibodies diluted in blocking buffer. Cells were stained with the following primary antibodies for an hour and half at room temperature: mouse anti-Oct4 (1:100, BD Biosciences), mouse anti-Nestin (1:100, BD Pharmigen) and mouse anti-Mash1 (1:3, kindly provided by F. Guillemot, NIMR, London, UK). Coverslips were washed three times in washing buffer before being incubated with the secondary antibody, Alexa568-conjugated goat anti-mouse (1:1000, Molecular Probe), for 45' at room temperature. Finally cells, were washed twice in washing buffer and once in PBS and coverslips mounted in Vectashield containing 1ug/mL DAPI to be analysed under a Leica SP1 confocal microscope.

#### **6.5 RT-PCR and Real Time PCR**

RNA extraction was performed using RNA-Bee according to the manufacturer protocol (Tel-Test inc.).  $10^6$  cells were resuspended in 0.5mL of RNA-Bee, mixed thoroughly, subjected to phenol-chloroform extraction and then precipitated with isopropanol. RNA was resuspended in 20uL of RNase-free water and 1µgr of total RNA was reverse transcribed using Superscript First-Strand Synthesis System (Quiagen) with oligo (dT)<sub>12-18</sub> (Invitrogen). cDNA was then analysed by semi-quantitative PCR or real-time PCR. The following oligonucleotides were used for gene amplification by semi-quantitative PCR:

Mash1\_forward: TGGAGACGCTGCGCTCGGC

Mash1\_reverse: CGTTGCTTCAATGGAGGCAAATG

Sox1\_forward: CATCTGCCCCATCACCTTC

Sox1\_reverse: CAACCAACCCAAAAGAGCGG

Oct4\_forward: AAGAGGTTGGGTGTGACTGG

Oct4\_reverse: GGTATCCACTCGCACCTTGT

c-DNA was amplified in a 50L reaction, using 0.4uM each primer, 1.25 U HotStarTaq DNA Polimerase (Quiagen) with the following program: 95° for 15', then 34 cycles (Mash1), 32 cycles (Sox1), 28 cycles (Oct4) at 94° for 15'', 60° for 30'', 72° for 30'' and final extension at 72° for 10'.

PCR products were visualised by electrophoresis on 1% gel with ethidium bromide staining.

To analyse cDNA with real-time PCR, 0.2ug RNA was retrotranscribed and cDNA diluted 1:10. The following primers were used:

HMBS-5\_forward: ACTGGTGGAGTCTGGAGTCTAGATGGC

HMBS-5\_reverse: GCCAGGCTGATGCCCAGG

YWHAZ\_forward: CGTTGTAGGAGCCCGTAGGTCAT

YWHAZ\_reverse: TCTGGTTGCGAAGCATTGGG

Nup37\_forward: TGGAACCTGGAAGGAAAGCAGACAG

Nup37\_reverse: CGGATTGTTCCATTCTTCTCTGCAA

Igf1\_forward: TCAGAAGCGATGGGGAAAATCAG

Igf1\_reverse: GCCAGGTAGAAGAGGTGTGAAGACG

PAH\_forward: TGTGGAGTTTGGGCTTTGCAAGG

PAH\_reverse: GCAGGAGCTTTGGCTTGTCTGATAA

Mash1\_forward: GGTCTCCGGTCTCGTCCTACTCCT

Mash1\_reverse: CCATTTGACGTCGTTGGCGAG

1700113H01\_forward: ATTTGTAATAATGCCCGGAGTGTGG

1700113H01\_reverse: GACTGACACTTCTGCAGAGGCCAAT

Tex18\_forward: ACTGTTCCGTGATGTCAAACCCTGT

Tex18\_reverse: CAATAATTCCAGGCCCAAGCTGTCT

Syn3\_forward: CTCCACATGCCCCAAATGAATTG

Syn3\_reverse: CCATGAGAAGTGGACTTCCCAGGAC

Mash1-Vector\_forward: GTCATGAGCAACAGTTTCAATGGCC

Mash1-Vector\_reverse: GGCTCTGCACCGTATTGAAACTGAG

Flk1\_forward: GGAAACCGGGAAACCCAAAC

Flk1\_reverse: GGAAACACAGCTTACTCTCTTGGG

Real-Time PCR analysis was carried out on Chromo4 DNA engine using an Opticon DNA engine (MJ Research Inc.), running the following program: 95° for 15', then 40 cycles at 94° for 15'', 60° for 30'', followed by read-plate. PCR reactions were in 20uL with 2X Syber-Green PCR Mastermix (Quiagen), 500nM primers and 2uL of template. Each experiment was performed at least in duplicate. The same conditions were used to analysed DNA for CTCF and cohesins chromatin immunoprecipitation (ChIP). The following primers were used:

CHS-1\_forward: TGGACCACACAACCCTCTAGTGACA

CHS-1\_reverse: CATTTGCCCTGGAGGTTATTTTAG

CHS-2\_forward: ACTTCCTTCTGTCCTTCCTCACAGC

CHS-2\_reverse: GGCAGCTGATGTTTTCTGAAGCATT

CHS-3\_forward: AGAGTACTTGGAGAGTCACACGGGC

CHS-3\_reverse: GAATAAATGGCTGGCCAGTGGTCT

CHS-4\_forward: TCATAAATGCTCCTTGAGGTCCAGG

CHS-4\_reverse: TTTCCCCCTCTGTCCCAAGT

CHS-5\_forward: AAAAATGGCCAAGGACTCATAGGAA

CHS-5\_reverse: TGAGGGAGCTCTGAAGATGTGGTAA

CHS-6\_forward: CTCATTACACAGCAGAACAGGGC

CHS-6\_reverse: TGCAAAGCCCAATATGGTTCTGG

PEAK 5\_forward: GAGAAGGAGGCACCCTACCCT

PEAK 5\_reverse: GGTCCCTACATACAGAAAGGTGTGCAG

M17:16298\_forward: AGGACATAGTCGCTTGAGTGATGG

M17:16298\_reverse: TTGGGCCAGGCTGGTACTTT

CD4-3\_forward: GGATCCTGTCAGCTTTGCTCTCTC

CD4-3\_reverse: GAGGTAGCATGTCTCAGGGTTCAG

Ifng\_forward: AGAAATCTGTCTGAACGGGTGGAG

Ifng\_reverse: GGTCGGTCCTCTGTAAGACAAAACCTAC



## 6.6 Preparation of probes for fluorescence in situ hybridisation (FISH) and metaphase spread preparation

RP-24 130P7 BAC (kindly provided by F. Guillemot, NIMR, London, UK) was used as a probe for the *Mash1* locus in 2D and 3D FISH experiments. For the probe preparation 1µg of DNA was labelled with Digoxigenin-11-dUTP (Roche) using the Nick Translation System (Invitrogen). Fragments of 400bp were obtained and the DNA probe was purified using Microspin S-300 HR Columns (Amersham Biosciences).

The following BAC were used as a probe for BDNF, Calbindin, Complexin, Stathmin, Synaptophysin. BAC DNA was checked by PCR with the following primers:

BDNF\_forward: GCC GGC TGG TGC AGA A

BDNF\_reverse: GCC TTG TCA AGC TAG GGC G

Calbindin\_forward: GCT CCG CGC ACT CTC AAA

Calbindin\_reverse: GAG ATG ACT GCA GGT GGGATT C

Complexin\_forward: CTGAAAGGCGCTGTCCATCG

Complexin\_reverse: AGGCTTTGTTCCCACACCCA

Stathmin\_forward: CTCCCCTCCACAGCATTGCA

Stathmin\_reverse: ACCCACTGCGGTCCCATGAT

Synaptophysin\_forward: GACTGGGCTGTTCCGACGAT

Synaptophysin\_reverse: TCAGGGTTCGAGGGCCAAAG

Probe was prepared as previously described and tested on metaphase spread analysis by 2D-FISH analysis. Briefly, 46C ES cells were incubated for 2 hours at 37°C with ethidium bromide at 1.5µg/mL to stretch the chromosomes and then with colchicine at 0.1µg/mL for 40' at 37°C to induce chromosome stretching and mitotic arrest. Cells were trypsinased, collected by centrifugation at 1200rpm for 5' and resuspended in hypotonic solution (0.56% KCl) for 5' at room temperature. Cells were then washed in PBS and fixed in ice-cold methanol:acetic acid (3:1) solution through three subsequent round of washes. Fixed cells were dropped onto slides, dried and used for 2D-FISH.

## 6.7 Fluorescence in situ hybridisation (FISH)

For 2D-FISH  $10^6$  cells were trypsinised, washed with PBS and fixed as previously described for metaphase spreads preparation. Slides were denatured, hybridised and washed as previously described (Williams et al. 2002). 0.1 µg of probe was used for each slide with 3 µg Cot1 DNA and 30 µg Salmon Sperm to avoid non-specific binding. The DNA probe was precipitated with Sodium Acetate (1:10) and 2 volumes of pure ethanol in a vacuum centrifuge for half an hour at high speed and temperature. The DNA pellet was suspended in 10 µL of hybridisation mix (2xSSC/50% formamide/10% dextran sulphate/1% tween-20). Slides were washed in PBS and increasing series of ethanol dilutions at 70%, 90%, 100% for 5' at room temperature to dehydrate the cells. Slides were let to dry at room temperature before denaturation in denaturing solution for 3' at 65°C (70% formamide/2x SSC/pH7). Slide denaturation was followed by ice-cold series of increasing ethanol dilutions at 70%, 90%, 100% for 5' each wash to finally let the slides dry at room temperature. 10 µL of probe was applied to each slide, covered by a coverslip and sealed with rubber. Incubation was carried out in waterbath in a humid chamber at 37°C. The following day three rounds of washes of 5' each were performed in First washing solution at 45°C (50% formamide/2x SSC/pH7) and Second washing solution at 55°C (1x SSC/pH7). A first wash in 4x SSCT was made at room temperature before incubating the slides in milk in a humid chamber at room temperature. The following antibodies were used diluted in milk: primary antibody mouse anti-DIG (1:30) and secondary antibody FITC conjugated anti-DIG (1:100). First slides were incubated with the primary antibody anti-DIG for an hour in dark humid chamber followed by three washes of 5' in 4X SSCT. Then incubation with the second antibody FITC conjugated anti-DIG was carried out in humid chamber for an hour at room temperature. Slides were let to dry at room temperature and finally mounted in Vectashield containing 1 µg/mL DAPI.

For 3D-FISH cells were trypsinised, washed with PBS and  $0.4 \times 10^6$  cells were let to attach for 5' on a poly-L-lysine-coated coverslip. Cells were specifically treated to preserve the integrity of the nuclei as previously described and DNA FISH was performed as previously described (Brown et al. 1997). Briefly, cells were fixed for 30' at room temperature in fixing solution (20mM  $\text{KH}_2\text{PO}_4$ , 130mM NaCl, 20mM KCl, 10mM EGTA, 2mM  $\text{MgCl}_2$ , 0.1% Triton X-100, 0.5% Gluteraldehyde at pH 7.3), followed by three washes in PBS. Coverslips were then incubated in sodium borohydride solution twice for 15' to reduce cellular autofluorescence. Three washes in PBS for 3' were followed by an incubation in blocking buffer for half an hour in humid chamber. Cells were incubated in EGS diluted in DMSO and added to PBS to precipitate crystals for half an hour at 37. Three washes in PBS were made before incubation in RNase solution at 100 µg/mL. The probe was prepared as described in 2D DNA FISH and suspended in 10 µL of hybridisation mix. Three washes in PBS were followed by DNA denaturation in NaOH solution at pH 13.5. Before applying the probe, cells were washed three times in ice-cold PBS. Hybridisation was carried out in waterbath and in a humid chamber at 37°C

over-night. The following day cells were washed for half an hour in 2x SCC at 37°C, in 2xSCC and 1x SCC at room temperature.

Coverslips were incubated in blocking buffer (4XSSC, 0.05% Tween-20, 3% BSA) for half an hour in humid chamber before incubation with antibodies diluted in blocking buffer. The following antibodies were used: primary antibody mouse anti-DIG (1:30,) and secondary antibody FITC conjugated anti-DIG (1:100). Coverslips were first incubated with primary antibody anti-DIG for an hour in a dark humid chamber followed by washes in washing buffer (4XSSC, 0.05% Tween-20). Coverslips were then incubated with secondary antibody FITC-conjugated anti-mouse for another hour. Coverslips were finally mounted in Vectashield containing 1 µg/mL DAPI.

## **6.8 Microscopy and measurements**

FISH on 2D fixed nuclei was analysed under a Leica microscope using a 100x oil-immersion object. Images were captured with a CDD camera and analysed using IP lab software. The position of loci relative to the nuclear periphery was determined as a ratio between the distance centre-periphery and centre-signal. FISH values >0.8 were considered peripheral (Kosak et al. 2002). Two independent experiments were performed for each sample and 50 nuclei were scored for each sample.

FISH on 3D preserve nuclei was analysed using a Leica laser scanning confocal microscope with a 100x oil-immersion object. Optical slices along the Z-axis of the nuclei were taken every 0.34µm to create Z-stacks for analysis. Image analysis was performed using ImageJ software and the position of loci relative to the nuclear periphery was determined as a ratio between the distance centre-periphery and centre-signal. At least 25 nuclei were scored per each sample.

## **6.9 CTCF and Rad21 chromatin immunoprecipitation (ChIP)**

46C ES cells were trypsinased, counted and collected in ES media.  $10^8$  cells were fixed in formaldehyde solution 11% for 10' at 37°C. Fixation was stopped with 0.125M glycine final concentration for 5' at RT. To extract chromatin, cells were first washed twice in cold PBS. Cells lysis was carried out in wash buffer 1 (10mM Hepes pH 7.5, 10mM EDTA, 0.5mM EGTA, 0.75% Triton X-100) by rotating the cells for 10' at 4°. Nuclei pellet were collected by centrifugation and resuspended in Wash Buffer 2 (10mM Hepes pH 7.5, 200mM NaCl, 1mM EDTA, 0.5mM EGTA) for nuclei lysis. Chromatin pellet was collected by centrifugation and resuspended in Lysis Buffer (150mM NaCl, 25mM Tris pH 7.5, 5mM EDTA, 1% Triton, 0.1% SDS, 0.5% Deoxycholate). Chromatin was sonicated with Bioruptor with max output and 30'' on-off cycles to generate

chromatin size of 400bp. Chromatin was then precipitate by centrifugation to clear away any debris and concentration was measured at nanodrop. Chromatin immunoprecipitation was performed with magnetic beads. For every IP 25uL of magnetic beads were used with 150µg chromatin. Beads were wash buffer A (50mM Tris ph 8, 150mM NaCl, 0,1% SDS, 0,5% deoxycholate, 1% NP40, 1mM EDTA) twice to equilibrate. 150µgr of chromatin was pre-cleared in 500µL final volume with 25uL of beads for 3h on rotor at 4°. At the same time antibody were bind to beads: 5µL (5µg) Ab (anti-CTCF, anti-Rad21, anti-IgG) was incubated with 25µL beads in Buffer A 500µL final volume for 3h at 4°. Beads and antibody were washed twice with buffer A. 1/100 of chromatin was taken from the preclearing as the input while the rest was incubated with the Ab conjugated with beads. IP was carried out over night on rotor at 4°. The day after beads were washed in Buffer A twice, Buffer B (50mM Tris ph 8, 500mM NaCl, 0,1% SDS, 0,5% deoxycholate, 1% NP40, 1mM EDTA) once, Buffer C (50mM Tris pH 8, 250mM LiCl, 0,5% deoxycholate, 1% NP40, 1mM EDTA) once and rinse in TE. Each wash was carried out for 10' on rotor at 4°. DNA was eluted from beads in 450µL Elution Buffer (1% SDS, 0,1M NaHCO<sub>3</sub> in water) with 22µL protease K (10mg/mL) and 5µL RNase A (10mg/mL) shaking for 30'' at 100rpm and 1' at 0rpm, 2h at 37° and over night at 65°. DNA was purified with phenol-chloroform extraction and precipitated in ethanol, NaAcetate, Glycoblu. DNA pellet was resuspend in 40µL TE and analysed by PCR.

## 6.10 CTCF and Rad21 knock down in ES cells

46C ES cell line were trypsinized, counted and plated at the concentration of  $0.3 \times 10^6$  in gelatine-coated 6 well plates. Cells were let to attach for at least 4 hours at 37°C before transfection. The following RNAi oligonucleotides were used for the knock down: set of 4 siRNA for CTCF, GAUGAUAUGUCACACCUUA, GACGAUACCCAGAUCAUAA, GGUCGAAGAUCAGAAUACA, AAAUUUGGAUCGUCACAUG (Dharmacon), set of 4 siRNA for Rad21, GAGCCCAGCUUAGCGAUUA, UGGAAUAACCGGCUCACUGA, CGGAAUGGAUGACCGUGAA, GAUGACGACAUGUUAGUGA (Dharmacon), non-targeting siRNA2 for the firefly luciferase sequence as a negative control (Dharmacon). siGlo-FITC conjugated was used to assess the efficiency of the transfection (Dharmacon). Tranfection of RNAi oligonucleotides using Dharmafect reagent was performed according to the manufacturer's instruction (Dharmafect). Briefly, in a first set of tubes, 100µL siGlo-FITC conjugated (final concentration 25nM ) were mixed with serum free medium or RNAi oligos for CTCF, Rad21 (final concentration 25nM for each RNAi oligos) or negative control (final concentration 100nM) to a final volume of 200 µL. At the same time, in a second set of tubes, 6uL Dharmafect 1 transfection reagent (Dharmacon) for each reaction were incubated for 5' at room temperature in 200uL of serum free

medium. The content of the second tube was added to the first tubes and incubated for 20' at room temperature. 400uL final volume of reaction mix was added in 2mL ES medium without penicillin-streptomycin to the cells and incubated over night at 37°C. The following day medium was replaced with normal ES medium and cells were harvested after 48hours for cell sorting. Cells positive for FITC signal were sorted using a Becton Dickinson Vantage/Diva using the FACS DIVA software.

## 6.11 Western Blot Analysis

Samples from 46C ES cell line parental and transfected were prepared by resuspending  $1 \times 10^6$  cells in 25 $\mu$ L PBS and 25 $\mu$ L of 2X loading buffer (0.1M TRIS ph 6.8, 2% (w/v) SDS, 20% (v/v) glycerol, 0.005% bromophenol blue, 10% b-mercaptoethanol in distilled water). Samples were frozen and denaturated at 95<sup>o</sup> for 5' before loading on a acrylamide gel. Loading in a 4% stacking gel (4% w/v acrylamide, 0.125M TRIS ph 6.8, 0.1% (w/v) SDS, 0.07% (w/v) ammonium persulphate, 0.12% N,N,N',N'-tetramethylethylenediamine in distilled water) was followed by protein separation in a 8% running gel (8% w/v acrylamide, 0.4M TRIS ph 8.8, 0.1% (w/v) SDS, 0.05% (w/v) ammonium persulphate, 0.05% N,N,N',N'-tetramethylethylenediamine in distilled water) with Tris-glycine electrophoresis buffer (0.025M TRIS, 0.192M glycine, 0.1% SDS). The benchmark non-stained protein ladder (Invitrogen) was also applied to the gel. Acrylamide gels were blotted into a Protan nitrocellulose membrane using the Trans-Blot Semi-dry Electrophoretic Transfer apparatus (Biorad) following the manufacturer's instructions in Transfer Buffer (25mM Trizma base, 0.192M glycine, 0.075% (w/v) SDS, 20% (v/v) methanol in distilled water). The membranes were incubated in 10% milk in agitation for half an hour and then with primary antibodies for an hour at room temperature, washed three times for 5' in washing buffer TBST and then incubated with secondary antibodies for another hour. The following antibodies were used diluted in 5% milk: anti-CTCF (1:1000, UPSTATE), anti-Rad21(1:1000, Abcam), anti-LaminB (1:5000, Santa Cruz Biotechnology), anti-rabbit IgG/HPR linked (1:5000, GE Healthcare UK), donkey anti-goat IgG/HPR linked (1:5000, Santa Cruz Biotechnology). After three washes in washing buffer, detection of proteins was done with the ECL western blot detection kit (Amersham) following the manufacturer's instructions.

## 6.12 Mash1 BAC electroporation in ES cells

BAC DNA was prepared according to the Chori BACPAC resources' protocol. A NEO cassette flanked by the 3' and 5' regions of the CMr was excised from a plasmid (gift from S. Sauer) and purified. Recombineering of the BAC DNA to replace the Cmr with the NEO cassette was carried out in competent SW102 E. Coli cells (Simone Alves's work)

BAC DNA was purified and linearised using the rare cutting enzyme PI-Sce1 (20uL DNA with 5u enzyme in 1xPI-SCE1 Buffer (NEB) at 37 over-night. Enzyme inactivation was carried out at 65°C for 20'. DNA was precipitated with 1/10 NaAcetate and 2 volume Ethanol 100%. DNA was resuspended in 10uL distilled water and quantified at the nanodrop. On the day of electroporation, 46C ES cells were collected, counted and centrifuged at 1200rpm for 5'. The cell pellet was washed twice with PBS and 10x10<sup>6</sup> cells were resuspended in 800uL PBS. 60 and 120ugr of DNA were added to the cells and transferred to a pre-chilled electroporation cuvette. After 5' in ice cells were electroporated using Biorad GenePulser at 200V and 960F, and replaced in ice for another 5'. Cells were plated in 5mL medium and subjected to G418 antibiotic selection at 400ugr/mL\_after 48h. Clones were picked after 11 days using a microscope. Single cell suspension of single clone was obtained in 96-well plates using trypsin/EDTA/2%Chicken Serum for 45' at 37. Clones were expanded for a few passages and cells collected for RNA analysis or froze down.

## References

- Andrulis, E. D., A. M. Neiman, et al. (1998). "Perinuclear localization of chromatin facilitates transcriptional silencing." *Nature* **394**(6693): 592-5.
- Arney, K. L. and A. G. Fisher (2004). "Epigenetic aspects of differentiation." *J Cell Sci* **117**(Pt 19): 4355-63.
- Arumugam, P., S. Gruber, et al. (2003). "ATP hydrolysis is required for cohesin's association with chromosomes." *Curr Biol* **13**(22): 1941-53.
- Azuara, V., P. Perry, et al. (2006). "Chromatin signatures of pluripotent cell lines." *Nat Cell Biol* **8**(5): 532-8.
- Bacher, C. P., M. Guggiari, et al. (2006). "Transient colocalization of X-inactivation centres accompanies the initiation of X inactivation." *Nat Cell Biol* **8**(3): 293-9.
- Bain, G., D. Kitchens, et al. (1995). "Embryonic stem cells express neuronal properties in vitro." *Dev Biol* **168**(2): 342-57.
- Ballas, N., E. Battaglioli, et al. (2001). "Regulation of neuronal traits by a novel transcriptional complex." *Neuron* **31**(3): 353-65.
- Ballas, N., C. Grunseich, et al. (2005). "REST and its corepressors mediate plasticity of neuronal gene chromatin throughout neurogenesis." *Cell* **121**(4): 645-57.
- Bannister, A. J., P. Zegerman, et al. (2001). "Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain." *Nature* **410**(6824): 120-4.
- Barakat, T. S., I. Jonkers, et al. "X-changing information on X inactivation." *Exp Cell Res* **316**(5): 679-87.
- Barski, A., S. Cuddapah, et al. (2007). "High-resolution profiling of histone methylations in the human genome." *Cell* **129**(4): 823-37.
- Bartova, E., J. Pachernik, et al. (2005). "Nuclear levels and patterns of histone H3 modification and HP1 proteins after inhibition of histone deacetylases." *J Cell Sci* **118**(Pt 21): 5035-46.
- Bell, A. C. and G. Felsenfeld (2000). "Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene." *Nature* **405**(6785): 482-5.
- Bernstein, B. E., M. Kamal, et al. (2005). "Genomic maps and comparative analysis of histone modifications in human and mouse." *Cell* **120**(2): 169-81.
- Bernstein, B. E., T. S. Mikkelsen, et al. (2006). "A bivalent chromatin structure marks key developmental genes in embryonic stem cells." *Cell* **125**(2): 315-26.
- Bertrand, N., D. S. Castro, et al. (2002). "Proneural genes and the specification of neural cell types." *Nat Rev Neurosci* **3**(7): 517-30.
- Bhaumik, S. R., E. Smith, et al. (2007). "Covalent modifications of histones during development and disease pathogenesis." *Nat Struct Mol Biol* **14**(11): 1008-16.
- Billon, N., C. Jolicoeur, et al. (2002). "Normal timing of oligodendrocyte development from genetically engineered, lineage-selectable mouse ES cells." *J Cell Sci* **115**(Pt 18): 3657-65.
- Blat, Y. and N. Kleckner (1999). "Cohesins bind to preferential sites along yeast chromosome III, with differential regulation along arms versus the centric region." *Cell* **98**(2): 249-59.
- Bolzer, A., G. Kreth, et al. (2005). "Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes." *PLoS Biol* **3**(5): e157.
- Boyer, L. A., T. I. Lee, et al. (2005). "Core transcriptional regulatory circuitry in human embryonic stem cells." *Cell* **122**(6): 947-56.
- Boyle, S., S. Gilchrist, et al. (2001). "The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells." *Hum Mol Genet* **10**(3): 211-9.
- Branco, M. R. and A. Pombo (2006). "Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations." *PLoS Biol* **4**(5): e138.

- Brickner, D. G., I. Cajigas, et al. (2007). "H2A.Z-mediated localization of genes at the nuclear periphery confers epigenetic memory of previous transcriptional state." *PLoS Biol* **5**(4): e81.
- Brickner, J. H. and P. Walter (2004). "Gene recruitment of the activated INO1 locus to the nuclear membrane." *PLoS Biol* **2**(11): e342.
- Brown, C. R., C. J. Kennedy, et al. (2008). "Global histone acetylation induces functional genomic reorganization at mammalian nuclear pore complexes." *Genes Dev* **22**(5): 627-39.
- Brown, J. M., J. Green, et al. (2008). "Association between active genes occurs at nuclear speckles and is modulated by chromatin environment." *J Cell Biol* **182**(6): 1083-97.
- Brown, K. E., S. S. Guest, et al. (1997). "Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin." *Cell* **91**(6): 845-54.
- Bruce, A. W., I. J. Donaldson, et al. (2004). "Genome-wide analysis of repressor element 1 silencing transcription factor/neuron-restrictive silencing factor (REST/NRSF) target genes." *Proc Natl Acad Sci U S A* **101**(28): 10458-63.
- Buckley, N. J., R. Johnson, et al. (2009). "Is REST a regulator of pluripotency?" *Nature* **457**(7233): E5-6; discussion E7.
- Bylund, M., E. Andersson, et al. (2003). "Vertebrate neurogenesis is counteracted by Sox1-3 activity." *Nat Neurosci* **6**(11): 1162-8.
- Cabal, G. G., A. Genovesio, et al. (2006). "SAGA interacting factors confine sub-diffusion of transcribed genes to the nuclear envelope." *Nature* **441**(7094): 770-3.
- Cao, R., L. Wang, et al. (2002). "Role of histone H3 lysine 27 methylation in Polycomb-group silencing." *Science* **298**(5595): 1039-43.
- Casolari, J. M., C. R. Brown, et al. (2004). "Genome-wide localization of the nuclear transport machinery couples transcriptional status and nuclear organization." *Cell* **117**(4): 427-39.
- Cau, E., S. Casarosa, et al. (2002). "Mash1 and Ngn1 control distinct steps of determination and differentiation in the olfactory sensory neuron lineage." *Development* **129**(8): 1871-80.
- Chambeyron, S. and W. A. Bickmore (2004). "Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription." *Genes Dev* **18**(10): 1119-30.
- Chambeyron, S. and W. A. Bickmore (2004). "Does looping and clustering in the nucleus regulate gene expression?" *Curr Opin Cell Biol* **16**(3): 256-62.
- Chen, X., H. Xu, et al. (2008). "Integration of external signaling pathways with the core transcriptional network in embryonic stem cells." *Cell* **133**(6): 1106-17.
- Chen, Z. F., A. J. Paquette, et al. (1998). "NRSF/REST is required in vivo for repression of multiple neuronal target genes during embryogenesis." *Nat Genet* **20**(2): 136-42.
- Chernukhin, I., S. Shamsuddin, et al. (2007). "CTCF interacts with and recruits the largest subunit of RNA polymerase II to CTCF target sites genome-wide." *Mol Cell Biol* **27**(5): 1631-48.
- Chong, J. A., J. Tapia-Ramirez, et al. (1995). "REST: a mammalian silencer protein that restricts sodium channel gene expression to neurons." *Cell* **80**(6): 949-57.
- Chubb, J. R., S. Boyle, et al. (2002). "Chromatin motion is constrained by association with nuclear compartments in human cells." *Curr Biol* **12**(6): 439-45.
- Chung, J. H., A. C. Bell, et al. (1997). "Characterization of the chicken beta-globin insulator." *Proc Natl Acad Sci U S A* **94**(2): 575-80.
- Ciosk, R., M. Shirayama, et al. (2000). "Cohesin's binding to chromosomes depends on a separate complex consisting of Scc2 and Scc4 proteins." *Mol Cell* **5**(2): 243-54.
- Ciosk, R., W. Zachariae, et al. (1998). "An ESP1/PDS1 complex regulates loss of sister chromatid cohesion at the metaphase to anaphase transition in yeast." *Cell* **93**(6): 1067-76.
- Conti, L., S. M. Pollard, et al. (2005). "Niche-independent symmetrical self-renewal of a mammalian tissue stem cell." *PLoS Biol* **3**(9): e283.
- Cremer, T. and C. Cremer (2001). "Chromosome territories, nuclear architecture and gene regulation in mammalian cells." *Nat Rev Genet* **2**(4): 292-301.
- Creyghton, M. P., S. Markoulaki, et al. (2008). "H2AZ is enriched at polycomb complex target genes in ES cells and is necessary for lineage commitment." *Cell* **135**(4): 649-61.
- Croft, J. A., J. M. Bridger, et al. (1999). "Differences in the localization and morphology of chromosomes in the human nucleus." *J Cell Biol* **145**(6): 1119-31.
- Czermin, B., R. Melfi, et al. (2002). "Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites." *Cell* **111**(2): 185-96.



- Daniel, J. A., M. G. Pray-Grant, et al. (2005). "Effector proteins for methylated histones: an expanding family." *Cell Cycle* **4**(7): 919-26.
- de Laat, W., P. Klous, et al. (2008). "Three-dimensional organization of gene expression in erythroid cells." *Curr Top Dev Biol* **82**: 117-39.
- Deniaud, E. and W. A. Bickmore (2009). "Transcription and the nuclear periphery: edge of darkness?" *Curr Opin Genet Dev* **19**(2): 187-91.
- Donze, D., C. R. Adams, et al. (1999). "The boundaries of the silenced HMR domain in *Saccharomyces cerevisiae*." *Genes Dev* **13**(6): 698-708.
- Dorman, E. R., A. M. Bushey, et al. (2007). "The role of insulator elements in large-scale chromatin structure in interphase." *Semin Cell Dev Biol* **18**(5): 682-90.
- Feuerbach, F., V. Galy, et al. (2002). "Nuclear architecture and spatial positioning help establish transcriptional states of telomeres in yeast." *Nat Cell Biol* **4**(3): 214-21.
- Finlan, L. E., D. Sproul, et al. (2008). "Recruitment to the nuclear periphery can alter expression of genes in human cells." *PLoS Genet* **4**(3): e1000039.
- Fraser, P. and W. Bickmore (2007). "Nuclear organization of the genome and the potential for gene regulation." *Nature* **447**(7143): 413-7.
- Fuss, S. H., M. Omura, et al. (2007). "Local and cis effects of the H element on expression of odorant receptor genes in mouse." *Cell* **130**(2): 373-84.
- Gerlich, D., B. Koch, et al. (2006). "Live-cell imaging reveals a stable cohesin-chromatin interaction after but not before DNA replication." *Curr Biol* **16**(15): 1571-8.
- Giadrossi, S., M. Dvorkina, et al. (2007). "Chromatin organization and differentiation in embryonic stem cell models." *Curr Opin Genet Dev* **17**(2): 132-8.
- Gilbert, D. M. (2002). "Replication timing and transcriptional control: beyond cause and effect." *Curr Opin Cell Biol* **14**(3): 377-83.
- Gilchrist, S., N. Gilbert, et al. (2004). "Nuclear organization of centromeric domains is not perturbed by inhibition of histone deacetylases." *Chromosome Res* **12**(5): 505-16.
- Glynn, E. F., P. C. Megee, et al. (2004). "Genome-wide mapping of the cohesin complex in the yeast *Saccharomyces cerevisiae*." *PLoS Biol* **2**(9): E259.
- Greenway, D. J., M. Street, et al. (2007). "RE1 Silencing transcription factor maintains a repressive chromatin environment in embryonic hippocampal neural stem cells." *Stem Cells* **25**(2): 354-63.
- Guacci, V., D. Koshland, et al. (1997). "A direct link between sister chromatid cohesion and chromosome condensation revealed through the analysis of MCD1 in *S. cerevisiae*." *Cell* **91**(1): 47-57.
- Guelen, L., L. Pagie, et al. (2008). "Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions." *Nature* **453**(7197): 948-51.
- Guenther, M. G., S. S. Levine, et al. (2007). "A chromatin landmark and transcription initiation at most promoters in human cells." *Cell* **130**(1): 77-88.
- Guillemot, F. (2007). "Spatial and temporal specification of neural fates by transcription factor codes." *Development* **134**(21): 3771-80.
- Guillemot, F., L. C. Lo, et al. (1993). "Mammalian achaete-scute homolog 1 is required for the early development of olfactory and autonomic neurons." *Cell* **75**(3): 463-76.
- Hadjur, S., L. M. Williams, et al. (2009). "Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus." *Nature* **460**(7253): 410-3.
- Haering, C. H. and K. Nasmyth (2003). "Building and breaking bridges between sister chromatids." *Bioessays* **25**(12): 1178-91.
- Hewitt, S. L., F. A. High, et al. (2004). "Nuclear repositioning marks the selective exclusion of lineage-inappropriate transcription factor loci during T helper cell differentiation." *Eur J Immunol* **34**(12): 3604-13.
- Hiratani, I., A. Leskovar, et al. (2004). "Differentiation-induced replication-timing changes are restricted to AT-rich/long interspersed nuclear element (LINE)-rich isochores." *Proc Natl Acad Sci U S A* **101**(48): 16861-6.
- Hiratani, I., T. Ryba, et al. (2008). "Global reorganization of replication domains during embryonic stem cell differentiation." *PLoS Biol* **6**(10): e245.

- Hong, L., G. P. Schroth, et al. (1993). "Studies of the DNA binding properties of histone H4 amino terminus. Thermal denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 "tail" to DNA." *J Biol Chem* **268**(1): 305-14.
- Horsfield, J. A., S. H. Anagnostou, et al. (2007). "Cohesin-dependent regulation of Runx genes." *Development* **134**(14): 2639-49.
- Jarman, A. P., M. Brand, et al. (1993). "The regulation and function of the helix-loop-helix gene, *asense*, in *Drosophila* neural precursors." *Development* **119**(1): 19-29.
- Jenuwein, T. and C. D. Allis (2001). "Translating the histone code." *Science* **293**(5532): 1074-80.
- Johnson, J. E., S. J. Birren, et al. (1990). "Two rat homologues of *Drosophila* *achaete-scute* specifically expressed in neuronal precursors." *Nature* **346**(6287): 858-61.
- Johnson, R., C. H. Teh, et al. (2008). "REST regulates distinct transcriptional networks in embryonic and neural stem cells." *PLoS Biol* **6**(10): e256.
- Jorgensen, H. F., Z. F. Chen, et al. (2009). "Is REST required for ESC pluripotency?" *Nature* **457**(7233): E4-5; discussion E7.
- Jorgensen, H. F., A. Terry, et al. (2009). "REST selectively represses a subset of RE1-containing neuronal genes in mouse embryonic stem cells." *Development* **136**(5): 715-21.
- Kawasaki, H., K. Mizuseki, et al. (2000). "Induction of midbrain dopaminergic neurons from ES cells by stromal cell-derived inducing activity." *Neuron* **28**(1): 31-40.
- Kim, J., J. Chu, et al. (2008). "An extended transcriptional network for pluripotency of embryonic stem cells." *Cell* **132**(6): 1049-61.
- Kornberg, R. D. and Y. Lorch (1999). "Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome." *Cell* **98**(3): 285-94.
- Kosak, S. T., D. Scalzo, et al. (2007). "Coordinate gene regulation during hematopoiesis is related to genomic organization." *PLoS Biol* **5**(11): e309.
- Kosak, S. T., J. A. Skok, et al. (2002). "Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development." *Science* **296**(5565): 158-62.
- Kouzarides, T. (2007). "Chromatin modifications and their function." *Cell* **128**(4): 693-705.
- Kumaran, R. I. and D. L. Spector (2008). "A genetic locus targeted to the nuclear periphery in living cells maintains its transcriptional competence." *J Cell Biol* **180**(1): 51-65.
- Kumaran, R. I., R. Thakar, et al. (2008). "Chromatin dynamics and gene positioning." *Cell* **132**(6): 929-34.
- Kupper, K., A. Kolbl, et al. (2007). "Radial chromatin positioning is shaped by local gene density, not by gene expression." *Chromosoma* **116**(3): 285-306.
- Kurdistani, S. K., S. Tavazoie, et al. (2004). "Mapping global histone acetylation patterns to gene expression." *Cell* **117**(6): 721-33.
- Kurshakova, M. M., A. N. Krasnov, et al. (2007). "SAGA and a novel *Drosophila* export complex anchor efficient transcription and mRNA export to NPC." *Embo J* **26**(24): 4956-65.
- Kurukuti, S., V. K. Tiwari, et al. (2006). "CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to *Igf2*." *Proc Natl Acad Sci U S A* **103**(28): 10684-9.
- Laloraya, S., V. Guacci, et al. (2000). "Chromosomal addresses of the cohesin component *Mcd1p*." *J Cell Biol* **151**(5): 1047-56.
- Lee, S. H., N. Lumelsky, et al. (2000). "Efficient generation of midbrain and hindbrain neurons from mouse embryonic stem cells." *Nat Biotechnol* **18**(6): 675-9.
- Lendahl, U., L. B. Zimmerman, et al. (1990). "CNS stem cells express a new class of intermediate filament protein." *Cell* **60**(4): 585-95.
- Lengronne, A., Y. Katou, et al. (2004). "Cohesin relocation from sites of chromosomal loading to places of convergent transcription." *Nature* **430**(6999): 573-8.
- Levsky, J. M., S. M. Shenoy, et al. (2007). "The spatial order of transcription in mammalian cells." *J Cell Biochem* **102**(3): 609-17.
- Lewis, E. B. (1978). "A gene complex controlling segmentation in *Drosophila*." *Nature* **276**(5688): 565-70.
- Ling, J. Q., T. Li, et al. (2006). "CTCF mediates interchromosomal colocalization between *Igf2/H19* and *Wsb1/Nf1*." *Science* **312**(5771): 269-72.

- Loh, Y. H., Q. Wu, et al. (2006). "The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells." Nat Genet **38**(4): 431-40.
- Lomvardas, S., G. Barnea, et al. (2006). "Interchromosomal interactions and olfactory receptor choice." Cell **126**(2): 403-13.
- Loyola, A. and G. Almouzni (2004). "Bromodomains in living cells participate in deciphering the histone code." Trends Cell Biol **14**(6): 279-81.
- Lunyak, V. V., R. Burgess, et al. (2002). "Corepressor-dependent silencing of chromosomal regions encoding neuronal genes." Science **298**(5599): 1747-52.
- Meneghini, M. D., M. Wu, et al. (2003). "Conserved histone variant H2A.Z protects euchromatin from the ectopic spread of silent heterochromatin." Cell **112**(5): 725-36.
- Michaelis, C., R. Ciosk, et al. (1997). "Cohesins: chromosomal proteins that prevent premature separation of sister chromatids." Cell **91**(1): 35-45.
- Misulovin, Z., Y. B. Schwartz, et al. (2008). "Association of cohesin and Nipped-B with transcriptionally active regions of the *Drosophila melanogaster* genome." Chromosoma **117**(1): 89-102.
- Mito, Y., J. G. Henikoff, et al. (2005). "Genome-scale profiling of histone H3.3 replacement patterns." Nat Genet **37**(10): 1090-7.
- Morey, C., N. R. Da Silva, et al. (2007). "Nuclear reorganisation and chromatin decondensation are conserved, but distinct, mechanisms linked to Hox gene activation." Development **134**(5): 909-19.
- Muller, J., C. M. Hart, et al. (2002). "Histone methyltransferase activity of a *Drosophila* Polycomb group repressor complex." Cell **111**(2): 197-208.
- Murrell, A., S. Heeson, et al. (2004). "Interaction between differentially methylated regions partitions the imprinted genes *Igf2* and *H19* into parent-specific chromatin loops." Nat Genet **36**(8): 889-93.
- Nishizumi, H., K. Kumasaka, et al. (2007). "Deletion of the core-H region in mice abolishes the expression of three proximal odorant receptor genes in cis." Proc Natl Acad Sci U S A **104**(50): 20067-72.
- Nonaka, N., T. Kitajima, et al. (2002). "Recruitment of cohesin to heterochromatic regions by Swi6/HP1 in fission yeast." Nat Cell Biol **4**(1): 89-93.
- Osborne, C. S., L. Chakalova, et al. (2004). "Active genes dynamically colocalize to shared sites of ongoing transcription." Nat Genet **36**(10): 1065-71.
- Otto, S. J., S. R. McCorkle, et al. (2007). "A new binding motif for the transcriptional repressor REST uncovers large gene networks devoted to neuronal functions." J Neurosci **27**(25): 6729-39.
- Pal, S. and S. Sif (2007). "Interplay between chromatin remodelers and protein arginine methyltransferases." J Cell Physiol **213**(2): 306-15.
- Paquette, A. J., S. E. Perez, et al. (2000). "Constitutive expression of the neuron-restrictive silencer factor (NRSF)/REST in differentiating neurons disrupts neuronal gene expression and causes axon pathfinding errors in vivo." Proc Natl Acad Sci U S A **97**(22): 12318-23.
- Parelho, V., S. Hadjur, et al. (2008). "Cohesins functionally associate with CTCF on mammalian chromosome arms." Cell **132**(3): 422-33.
- Parras, C. M., R. Galli, et al. (2004). "Mash1 specifies neurons and oligodendrocytes in the postnatal brain." Embo J **23**(22): 4495-505.
- Parras, C. M., C. Hunt, et al. (2007). "The proneural gene Mash1 specifies an early population of telencephalic oligodendrocytes." J Neurosci **27**(16): 4233-42.
- Pattyn, A., N. Simplicio, et al. (2004). "Ascl1/Mash1 is required for the development of central serotonergic neurons." Nat Neurosci **7**(6): 589-95.
- Pauli, A., F. Althoff, et al. (2008). "Cell-type-specific TEV protease cleavage reveals cohesin functions in *Drosophila* neurons." Dev Cell **14**(2): 239-51.
- Perry, P., S. Sauer, et al. (2004). "A dynamic switch in the replication timing of key regulator genes in embryonic stem cells upon neural induction." Cell Cycle **3**(12): 1645-50.
- Peters, A. H., S. Kubicek, et al. (2003). "Partitioning and plasticity of repressive histone methylation states in mammalian chromatin." Mol Cell **12**(6): 1577-89.
- Pickersgill, H., B. Kalverda, et al. (2006). "Characterization of the *Drosophila melanogaster* genome at the nuclear lamina." Nat Genet **38**(9): 1005-14.

- Pokholok, D. K., C. T. Harbison, et al. (2005). "Genome-wide map of nucleosome acetylation and methylation in yeast." *Cell* **122**(4): 517-27.
- Prozorovski, T., U. Schulze-Toppoff, et al. (2008). "Sirt1 contributes critically to the redox-dependent fate of neural progenitors." *Nat Cell Biol* **10**(4): 385-94.
- Ragoczy, T., M. A. Bender, et al. (2006). "The locus control region is required for association of the murine beta-globin locus with engaged transcription factories during erythroid maturation." *Genes Dev* **20**(11): 1447-57.
- Rankin, S., N. G. Ayad, et al. (2005). "Sororin, a substrate of the anaphase-promoting complex, is required for sister chromatid cohesion in vertebrates." *Mol Cell* **18**(2): 185-200.
- Reddy, K. L., J. M. Zullo, et al. (2008). "Transcriptional repression mediated by repositioning of genes to the nuclear lamina." *Nature* **452**(7184): 243-7.
- Revenkova, E., M. L. Focarelli, et al. (2009). "Cornelia de Lange syndrome mutations in SMC1A or SMC3 affect binding to DNA." *Hum Mol Genet* **18**(3): 418-27.
- Reynolds, B. A. and S. Weiss (1992). "Generation of neurons and astrocytes from isolated cells of the adult mammalian central nervous system." *Science* **255**(5052): 1707-10.
- Ringrose, L., H. Ehret, et al. (2004). "Distinct contributions of histone H3 lysine 9 and 27 methylation to locus-specific stability of polycomb complexes." *Mol Cell* **16**(4): 641-53.
- Rollins, R. A., M. Korom, et al. (2004). "Drosophila nipped-B protein supports sister chromatid cohesion and opposes the stromalin/Scc3 cohesion factor to facilitate long-range activation of the cut gene." *Mol Cell Biol* **24**(8): 3100-11.
- Rollins, R. A., P. Morcillo, et al. (1999). "Nipped-B, a Drosophila homologue of chromosomal adherins, participates in activation by remote enhancers in the cut and Ultrabithorax genes." *Genetics* **152**(2): 577-93.
- Ruthenburg, A. J., C. D. Allis, et al. (2007). "Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark." *Mol Cell* **25**(1): 15-30.
- Ryu, M. J., B. J. Kim, et al. (2006). "Direct interaction between cohesin complex and DNA replication machinery." *Biochem Biophys Res Commun* **341**(3): 770-5.
- Santos-Rosa, H. and C. Caldas (2005). "Chromatin modifier enzymes, the histone code and cancer." *Eur J Cancer* **41**(16): 2381-402.
- Sarcinella, E., P. C. Zuzarte, et al. (2007). "Monoubiquitylation of H2A.Z distinguishes its association with euchromatin or facultative heterochromatin." *Mol Cell Biol* **27**(18): 6457-68.
- Schmitz, J., E. Watrin, et al. (2007). "Sororin is required for stable binding of cohesin to chromatin and for sister chromatid cohesion in interphase." *Curr Biol* **17**(7): 630-6.
- Schoenherr, C. J. and D. J. Anderson (1995). "Silencing is golden: negative regulation in the control of neuronal gene transcription." *Curr Opin Neurobiol* **5**(5): 566-71.
- Schubeler, D., D. Scalzo, et al. (2002). "Genome-wide DNA replication profile for Drosophila melanogaster: a link between transcription and replication timing." *Nat Genet* **32**(3): 438-42.
- Schuldiner, O., D. Berdnik, et al. (2008). "piggyBac-based mosaic screen identifies a postmitotic function for cohesin in regulating developmental axon pruning." *Dev Cell* **14**(2): 227-38.
- Sekinger, E. A., Z. Moqtaderi, et al. (2005). "Intrinsic histone-DNA interactions and low nucleosome density are important for preferential accessibility of promoter regions in yeast." *Mol Cell* **18**(6): 735-48.
- Shimi, T., K. Pflieger, et al. (2008). "The A- and B-type nuclear lamin networks: microdomains involved in chromatin organization and transcription." *Genes Dev* **22**(24): 3409-21.
- Singh, S. K., M. N. Kagalwala, et al. (2008). "REST maintains self-renewal and pluripotency of embryonic stem cells." *Nature* **453**(7192): 223-7.
- Skibbens, R. V. (2008). "Mechanisms of sister chromatid pairing." *Int Rev Cell Mol Biol* **269**: 283-339.
- Skibbens, R. V., L. B. Corson, et al. (1999). "Ctf7p is essential for sister chromatid cohesion and links mitotic chromosome structure to the DNA replication machinery." *Genes Dev* **13**(3): 307-19.
- Skok, J. A., K. E. Brown, et al. (2001). "Nonequivalent nuclear location of immunoglobulin alleles in B lymphocytes." *Nat Immunol* **2**(9): 848-54.
- Somech, R., S. Shaklai, et al. (2005). "The nuclear-envelope protein and transcriptional repressor LAP2beta interacts with HDAC3 at the nuclear periphery, and induces histone H4 deacetylation." *J Cell Sci* **118**(Pt 17): 4017-25.

- Spector, D. L. (2003). "The dynamics of chromosome organization and gene regulation." *Annu Rev Biochem* **72**: 573-608.
- Spilianakis, C. G., M. D. Lalioti, et al. (2005). "Interchromosomal associations between alternatively expressed loci." *Nature* **435**(7042): 637-45.
- Splinter, E., H. Heath, et al. (2006). "CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus." *Genes Dev* **20**(17): 2349-54.
- Stedman, W., H. Kang, et al. (2008). "Cohesins localize with CTCF at the KSHV latency control region and at cellular c-myc and H19/Igf2 insulators." *Embo J* **27**(4): 654-66.
- Stock, J. K., S. Giadrossi, et al. (2007). "Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells." *Nat Cell Biol* **9**(12): 1428-35.
- Sugimori, M., M. Nagao, et al. (2007). "Combinatorial actions of patterning and HLH transcription factors in the spatiotemporal control of neurogenesis and gliogenesis in the developing spinal cord." *Development* **134**(8): 1617-29.
- Sumara, I., E. Vorlaufer, et al. (2000). "Characterization of vertebrate cohesin complexes and their regulation in prophase." *J Cell Biol* **151**(4): 749-62.
- Sumara, I., E. Vorlaufer, et al. (2002). "The dissociation of cohesin from chromosomes in prophase is regulated by Polo-like kinase." *Mol Cell* **9**(3): 515-25.
- Sun, Y., M. Nadal-Vicens, et al. (2001). "Neurogenin promotes neurogenesis and inhibits glial differentiation by independent mechanisms." *Cell* **104**(3): 365-76.
- Sun, Y. M., D. J. Greenway, et al. (2005). "Distinct profiles of REST interactions with its target genes at different stages of neuronal development." *Mol Biol Cell* **16**(12): 5630-8.
- Taddei, A., G. Van Houwe, et al. (2006). "Nuclear pore association confers optimal expression levels for an inducible yeast gene." *Nature* **441**(7094): 774-8.
- Takahashi, T. S., P. Yiu, et al. (2004). "Recruitment of Xenopus Scc2 and cohesin to chromatin requires the pre-replication complex." *Nat Cell Biol* **6**(10): 991-6.
- Tang, Y. A., D. Huntley, et al. "Efficiency of Xist-mediated silencing on autosomes is linked to chromosomal domain organisation." *Epigenetics Chromatin* **3**(1): 10.
- Tolhuis, B., R. J. Palstra, et al. (2002). "Looping and interaction between hypersensitive sites in the active beta-globin locus." *Mol Cell* **10**(6): 1453-65.
- Tonkin, E. T., T. J. Wang, et al. (2004). "NIPBL, encoding a homolog of fungal Scc2-type sister chromatid cohesion proteins and fly Nipped-B, is mutated in Cornelia de Lange syndrome." *Nat Genet* **36**(6): 636-41.
- Toth, A., R. Ciosk, et al. (1999). "Yeast cohesin complex requires a conserved protein, Eco1p(Ctf7), to establish cohesion between sister chromatids during DNA replication." *Genes Dev* **13**(3): 320-33.
- Uhlmann, F., F. Lottspeich, et al. (1999). "Sister-chromatid separation at anaphase onset is promoted by cleavage of the cohesin subunit Scc1." *Nature* **400**(6739): 37-42.
- Unal, E., J. M. Heidinger-Pauli, et al. (2008). "A molecular determinant for the establishment of sister chromatid cohesion." *Science* **321**(5888): 566-9.
- Vega, H., Q. Waisfisz, et al. (2005). "Roberts syndrome is caused by mutations in ESCO2, a human homolog of yeast ECO1 that is essential for the establishment of sister chromatid cohesion." *Nat Genet* **37**(5): 468-70.
- Villares, R. and C. V. Cabrera (1987). "The achaete-scute gene complex of D. melanogaster: conserved domains in a subset of genes required for neurogenesis and their homology to myc." *Cell* **50**(3): 415-24.
- Volpi, E. V., E. Chevret, et al. (2000). "Large-scale chromatin organization of the major histocompatibility complex and other regions of human chromosome 6 and its response to interferon in interphase nuclei." *J Cell Sci* **113** ( Pt 9): 1565-76.
- Wade, P. A. and A. P. Wolffe (1997). "Histone acetyltransferases in control." *Curr Biol* **7**(2): R82-4.
- Waizenegger, I. C., S. Hauf, et al. (2000). "Two distinct pathways remove mammalian cohesin from chromosome arms in prophase and from centromeres in anaphase." *Cell* **103**(3): 399-410.
- Wang, H., L. Wang, et al. (2004). "Role of histone H2A ubiquitination in Polycomb silencing." *Nature* **431**(7010): 873-8.
- Wendt, K. S. and J. M. Peters (2009). "How cohesin and CTCF cooperate in regulating gene expression." *Chromosome Res* **17**(2): 201-14.

- Wendt, K. S., K. Yoshida, et al. (2008). "Cohesin mediates transcriptional insulation by CCCTC-binding factor." Nature **451**(7180): 796-801.
- Westbrook, T. F., G. Hu, et al. (2008). "SCFbeta-TRCP controls oncogenic transformation and neural differentiation through REST degradation." Nature **452**(7185): 370-4.
- Williams, A., C. G. Spilianakis, et al. (2010) "Interchromosomal association and gene regulation in trans." Trends Genet **26**(4): 188-97.
- Williams, R. R., V. Azuara, et al. (2006). "Neural induction promotes large-scale chromatin reorganisation of the Mash1 locus." J Cell Sci **119**(Pt 1): 132-40.
- Williams, R. R., S. Broad, et al. (2002). "Subchromosomal positioning of the epidermal differentiation complex (EDC) in keratinocyte and lymphoblast interphase nuclei." Exp Cell Res **272**(2): 163-75.
- Wu, J. and X. Xie (2006). "Comparative sequence analysis reveals an intricate network among REST, CREB and miRNA in mediating neuronal gene expression." Genome Biol **7**(9): R85.
- Ye, P. and A. J. D'Ercole (2006). "Insulin-like growth factor actions during development of neural stem cells and progenitors in the central nervous system." J Neurosci Res **83**(1): 1-6.
- Ying, Q. L., J. Nichols, et al. (2003). "BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3." Cell **115**(3): 281-92.
- Ying, Q. L., M. Stavridis, et al. (2003). "Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture." Nat Biotechnol **21**(2): 183-6.
- Zhang, B., S. Jain, et al. (2007). "Mice lacking sister chromatid cohesion protein PDS5B exhibit developmental abnormalities reminiscent of Cornelia de Lange syndrome." Development **134**(17): 3191-201.
- Zhao, Z., G. Tavoosidana, et al. (2006). "Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions." Nat Genet **38**(11): 1341-7.
- Zink, D., M. D. Amaral, et al. (2004). "Transcription-dependent spatial arrangements of CFTR and adjacent genes in human cell nuclei." J Cell Biol **166**(6): 815-25.
- Zlatanova, J. and P. Caiafa (2009). "CTCF and its protein partners: divide and rule?" J Cell Sci **122**(Pt 9): 1275-84.