The role of high mobility nucleosomal binding protein (Hmgn2) in undifferentiated mouse epiblast carcinoma stem cells

MSc. Ohoud Mohammedsabri M. Rehbini

Submitted in fulfillment of the requirement for
The degree of philosophy

School of medicine
College of medical, veterinary and life sciences
Institution of cancer science

March/2016

Abstract

High mobility group nucleosome binding (HMGN) proteins belong to the superfamily of high mobility group (HMG) proteins. HMGN1 and HMGN2 are ubiquitously expressed in all vertebrates, and are most highly expressed in embryonic tissue. Moreover, HMGN1 and HMGN2 were found to be highly expressed in neural stem/progenitor cells in the mouse brain. Here, mouse embryonal carcinoma cells (P19 EC) were used as a model system to study the role of HMGN proteins in pluripotent stem cells. Previously, experiments using short interfering RNA (siRNA) technology to knockdown HMGN1 and HMGN2 have suggested that HMGN proteins are important for the expression of key pluripotent genes, *Oct4*, *Sox2* and *Nanog*, in P19 EC cells (Mohan, 2012).

The aim of this thesis was to develop a lentiviral system for the long term knockdown of Hmgn2, in order to investigate more fully the role of this protein in stem cell pluripotency and differentiation. Constitutive and inducible lentiviral shRNAmir systems were tested and optimized, and a constitutive system was chosen for further work. HMGN2 knockdown in undifferentiated P19 EC cells resulted in the down-regulation of Oct4 protein levels. ChIP assays showed that HMGN2 binding over the *Oct4* gene was absent in HMGN2 knockdown cells. Furthermore, binding of HMGN1 at this locus was increased in the absence of HMGN2. Consistent with the reduction in *Oct4* expression, levels of the active histone modification, H3K4me3, were also decreased at the *Oct4* gene. These results support a role for HMGN2 in the regulation of *Oct4* expression in P19 cells, and imply that HMGN2 may be important for maintaining stem cell pluripotency.

Table of Contents

List of Tables

List of Figures

Acknowledgement

I am highly grateful to my supervisor Dr. Katherine west for the guidance, advice and support provided throughout my PhD journey, especially during difficult time. A special thanks to Dr Adam west who support me during my PhD and give me invaluable contribution and experimental advices to the work presented in this thesis. Dr Adam West had helped me to overcome the lentivirus issues during data meetings and he also provided me with PSLIK-CD4 and let me share and use his lab facilities. Even though he left main campus to Wolfson whol cancer research building, he provided me with an access to his lab to attend the data meeting and keep using the facilities in the building such as Qrt-pcr, western blot scanner and DNA sonicator. He also provided me with protocols for nuclear extracts and helped me to solve the issues in the cell mediums Ph.

I also thank my advisor Dr. Tomoko Iwata, Dr. Sarah Meek, Dr. Claire Caney for providing a critical review of my work and advice on further experiment and support during the mini-viva and writing support during thesis writing. Also I would like to thank Prof. Helen Wheadon and Dr. Anuradha Tarafdar for providing me an opportunity to culture embryonic stem cells (E14) in the lab and develop the system in our lab. I thank Dr Andrew Hamilton and Dr Torstein Stain for valuable conversation about experimental plan and their advice.

I am grateful to everyone in the group of Drs Adam and Katherine west for their help, advice and encouragement. Special thanks to Dr. Grainne Barkass, Dr. Elaine Gourlay and Dr Gockola Mohan for invaluable discussion and motivating me during my first year in PhD and Dr. Sabarinadh Chilaka for teaching me techniques and motivating me during the writing up stages. Dr. Narrendra Kumar has helped me to understand ChIP results and manage to use UCSC browser and provided advice to present some ChIP data. And I special thanks to Alejandro Revira, Ileana Guerrini, Aliya Hussein and Abdulmajeed Sindi for their helps, supports, good laboratory practices, discussions and providing me an advice for strategies that had helped me in my experiments. Ileana Guerrini taught me how to insert shRNA into PGIPZ vector from PSLIK and Alejandro Revira helped me in how to design primer for QRT-PCR and how to do CO-IP experiment.

My studies were funded by King Abdullah scholarship programme who I gratefully acknowledge and proud to be part of this scholarship during King's Abdullah life. Finally, I give my thanks Allah (God) for his help and keeping my strong in faith and spirit during my PhD, then my family in Saudi Arabia who encourage me all the way. My mother and my dad supported me all my education life and never give up on me. My siblings were always a good motivated and encourage me to continue fulfilling my dream.

Author's Declaration

I declare that, except where explicit reference is made to the contribution of others, that this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

Ohoud Rehbini

Definitions/Abbreviations

A/B compartment = genome can be divided into two compartment based on cell-type specific and associate with close and open chromatin.
ACF = Asymmetric Crying Facies (Cayler Cardiofacial Syndrome)
ALT = alternative lengthening of telomeres
AP = (apurinic/apyrimidinic site), also known as an abasic site
APP = Amyloid Beta (A4) Precursor Protein
ART = ADP-ribosyltransferase
Asf1= Anti-silencing function protein 1
ATAC-seq = assay for transposase-accessible chromatin using sequencing
ATM = Ataxia telangiectasia mutated
ATP-dependent chromatin = ATPase subunit that belongs to the SNF2 superfamily of proteins.
AZaC = 5-aza-2'-deoxy-cytidine
BAC = bacterial artificial chromosome
BAP1 = BRCA1 Associated Protein-1 (Ubiquitin Carboxy-Terminal Hydrolase
BER = base excision repair
BLM = Bloom Syndrome, RecQ Helicase-Like
BMP = Bone morphogenetic proteins
BRCA1 = Breast Cancer 1, Early Onset
BRD4 = Bromodomain-containing protein 4
Brg1 = Brahma-related gene-1
Bromodomain = 110 amino acid protein domain that recognizes acetylated lysine residues
BTG4 = B-Cell Translocation Gene 4
CAP-G protein = Capping Protein (Actin Filament), Gelsolin-Like
CBP = CREB-binding protein
CCNB1= G2/mitotic-specific cyclin-B1 protein
CHD = Chromodomain Helicase DNA Binding Protein
Chd7 = chromodomain helicase DNA binding protein 7
ChIP-seq = combines chromatin immunoprecipitation (ChIP) with massively parallel DNA sequencing to identify the binding sites of DNA-associated proteins.
Chromodomain = chromatin organization modifier
CITRULLINE = α-amino acid
Cmv = Cytomegalovirus
C-Myc = V-Myc Avian Myelocytomatosis Viral Oncogene Homolog
CNTF = Ciliary Neurotrophic Factor
COMMD1 = Copper Metabolism (Murr1) Domain Containing 1
CpG = CG sites are regions of DNA where a cytosine nucleotide is followed by a guanine nucleotide in the linear sequence of bases along its 5' → 3' direction.
DICER = endoribonuclease Dicer or helicase with RNase motif
DNasel hypersensitive site (DHS)= are regions of chromatin that are sensitive to cleavage by the DNase I enzyme.
DNase-seq = (DNase I hypersensitive sites) sequencing
DORSHA = Class 2 ribonuclease III enzyme
DPN = Depleted Proximal Nucleosomes
DSB = double-strand break
DSRNA = Double-strand RNA
EB = embryo bodies
EC = embryonic carcinoma
ELISA = enzyme-linked immunosorbent assay
EpiSCs = epiblast stem cells
ERK = extracellular-signal-regulated kinases (ERKs)/ classical MAP kinases
ESCs = embryonic stem cells
ESRRB = Estrogen-Related Receptor Beta
EZh2 = Enhancer Of Zeste 2 Polycomb Repressive Complex 2 Subunit
FGF2 = Fibroblast Growth Factor 2 (Basic)
FGF4 = Fibroblast growth factor 4
FGF5 = Fibroblast growth factor 5
FIGLA = Folliculogenesis Specific BHLH Transcription Factor
footprinting techniques = A DNase footprinting assay
FOXA2 = Forkhead Box A2
Fpr4 = histone chaperone FK506-binding protein 4
Gag = Group-specific antigen
GAL4-VP16 = S. cerevisiae GAL4 [1-147] fused with HSV2 VP16 [411-490]
GATA6/4 = GATA Binding Protein
GBX2 = Gastrulation Brain Homeobox 2
Gcn5 = histone acetyltransferase GCN5
GEAP = Glial E2F1-Associated Proteins
GFAP = Glial fibrillary acidic protein
GSK3 = Glycogen synthase kinase-3
H2A = Histone H2A

H2A.Z = H2A Histone Family, Member Z
H2Aub = histone H2A monoubiquitylation
H2B = Histone H2B
H2Bub1 = H2B monoubiquitination
H3 = Histone H3
H3.3 = histone H3.3
H3K122ac = Histone H3 (acetyl K122) acetylated
H3K14aC = Histone H3 acetyl Lys14 acetylated
HDACs = Histone deacetylases
HDMs = Protein hold'em
Histone chaperones a group of proteins that bind histones and regulate nucleosome assembly
HIV = human immunodeficiency virus
HMG = high mobility group
HMOF = lysine acetyltransferase MOF (males absent on the first)
HMT = Histone methyltransferases
HOX = homebox gene
HP = heterochromatin protein
HP1= heterochromatin protein 1 alpha
HR = homologous recombination
HRE = hormone responsive element
HSP68 = Mouse Hsp68 minimal promoter.
ICM = inner cell mass
IFN-ALPHA/beta = Alpha/Beta interferon
iPSCs = induce pluripotent stem cells
IR = ion radiation
IRES = internal ribosome entry site
IRF1/2 = interferon regulatory factor 1/2
ISWI = imitation SWI of drosophila melanogaster
JAK SIGNALING = Janus kinase
Jmjd1a = Lysine (K)-Specific Demethylase 3A
Jmjd2c = Lysine (K)-Specific Demethylase 4C
JNK-mediated = The c-Jun N-terminal Kinase (JNK)
KCR = lysine crotonylation
KLF2/4/5 = Krüppel-like family of transcription factors
LANA = latency-associated nuclear antigen
LTR = Long terminal repeats
MAPK = Mitogen-activated protein kinases
MeCP2 = methyl CpG binding protein 2 (Rett syndrome)
MEFs = mouse embryonic fibroblasts
MEK = Mitogen-activated protein kinase
MIRNA = micro RNA
MNase = Micrococcal Nuclease
Moi = multiplicity of infuction
MSC = bone marrow stromal stem cells
Nanog = Homeobox Transcription Factor Nanog
NBD = nucleosome binding domain
NER = Nucleotide excision repair
Nestin = acronym for neuroectodermal stem cell marker
NCE = nuclear cell extract
NEUN = Feminizing Locus on X-3, Fox-3, Rbfox3, or Hexaribonucleotide Binding Protein-3
NFkB = nuclear factor kappa-light-chain-enhancer of activated B cells
NFRs = nucleosome free regions
NLRP9 = NOD-like receptor family pyrin domain containing 9
NLS = nuclear localization signal
NODAL = signal transduction pathway
NON-SET = non-SET domain-containing histone methyltransferase utilizes the enzyme Dot1
NOTCH = a family of transmembrane proteins
NPCs = neuro progenitor cells
NSC = neuro stem cell
nucleosome sliding is a result of energy-dependent nucleosome remodelling in vitro
Oct4/6 = Octamer-Binding Transcription Factor
O-GlcNAc = UDP-N-acetylglucosamine—peptide N-acetylglucosaminyltransferase
Olig2 = Oligodendrocyte transcription factor
OPN = occupied proximal-nucleosome
OTX2 = Homeobox protein OTX2
P21 = cyclin-dependent kinase inhibitor 1
P300 = transcriptional coactivator
P300/CBP = coactivator family
P38 = mitogen-activated protein kinases
P53 = tumor suppressor p53
PADI4 = Peptidyl arginine deiminase, type IV
PAF = Platelet-activating factor
Palmitoylated residues = covalent attachment of fatty acids
PARP = Poly (ADP-ribose) polymerase
Pax6 = Paired box protein Pax-6
PCNA = Proliferating cell nuclear antigen
PDGFRA = Platelet-derived growth factor receptor
PGC = Primordial Germ Cell
PHD fingers = plant homeodomain

PIAS = Protein inhibitor of activated STAT
Pioneer factors are transcription factors that can directly bind condensed chromatin
PITX2 = Paired-like homeodomain transcription factor 2
PKDMs = protein lysine demethylases
PKMTs = lysine methyltransferases
PKR = Protein kinase RNA-activated
Pol = polymerase
PR = progesterone receptor
PRC1/2 = Polycomb Repressive Complex
PrEn = primitive endoderm
PRMT5 = Protein arginine N-methyltransferase 5
PTM = post-translational modification
PURO = puromycine
RA = retinoic acid
Rad6 = Ubiquitin-conjugating enzyme E2
RCC1= Regulator of chromosome condensation
RD = regulatory domain
RING1B = Putative E3 ubiquitin-protein ligase RING1b
RISC = RNA-induced silencing complex
RNAP2 = RNA polymerase II
Rre = Rev response element
rRNA = ribosomal RNA
RT-qPCR = Quantitative reverse transcription PCR
Rtt109 = histone acetyltransferase Rtt109
SAXS = Small-angle X-ray scattering
SERINE/THREONINE Kinase Aurora-B = Serine/threonine-protein kinase component of the chromosomal passenger complex (CPC), a complex that acts as a key regulator of mitosis.
SET = SET Nuclear Proto-Oncogene
SFM = scanning-force microscopy
Smarca4 = ATP-dependent helicase SMARCA4
SMC protein = Structural Maintenance of Chromosomes.
Snf21/Snf2h = ATP-dependent DNA helicase
SONIC HEDGEHOG = shh
Sox1/2/7/17 = sex determining region Y
SSEA1= stage-specific embryonic antigen-1
STAT = signal transducer and activator of transcription
SIN = self-inactivating vector
STORM = stochastic optical reconstruction microscopy
SUMO-3 = Small ubiquitin-related modifier 3
Suz12 = Polycomb protein SUZ12
SV40 = Simian vacuolating virus 40
SWI/SNF = SWItch/Sucrose Non-Fermentable
TADs = Topologically associating domains
Tat = Tyrosine Aminotransferase
TBP = Transcription Initiation Factor TFIID TBP Subunit
TBX3 = T-Box 3
TCR = T cell receptor
TEAD4 = Transcriptional enhancer factor TEF-3
TFIIB/TFIIIA= Transcription Factor II/III B/A
T-helper1 = T helper cells
Tip6 = tension-induced/inhibited protein
TSA = trichostatin A inhibitor
TSS = transcription starting sites
Vsv-g = Vesicular stomatitis Indiana virus
WAS = Wiskott-Aldrich syndrome
WCE = whole cell extract
WNT = complex protein network
WRN = Werner syndrome RecQ like helicase
X-ray crystallography is a tool used for identifying the atomic and molecular structure of a crystal
Zfp42 = Zinc finger protein 42 homolog
ZP2/3/4 = Zona pellucida sperm-binding protein
λ5-VpreB1 is mouse genes express in B-cells

Chapter 1 Introduction

1.1 Chromatin structure in eukaryotic cells

Mammalian genomes are packed into nucleoprotein complex known as chromatin. Chromatin regulates all processes involving DNA, which include DNA synthesis, transcription, translation, DNA repair, and recombination (Tagami et. al., 2004). Chromatin can be described as euchromatin or heterochromatin. Euchromatin corresponds to the regions with high numbers of transcribed genes, whereas heterochromatin indicates the highly condensed regions of the genome and less transcribed genes.

1.1.1 Nucleosome structure

Chromatin contains a repeated subunit called the nucleosome, which consists of ~147 bp of DNA wrapped around an octamer of core histone proteins, and linker DNA that connects the nucleosomes to each other. The histone octamer contains two molecules of each of the four core histones, H3, H4, H2A, and H2B (Zlatanova et. al., 1998) (Figure 1.1). Core histones are highly conserved essential proteins that are found in all eukaryotes, and they contain a high amount of the amino acids lysine and arginine. Histones have a positive charge that permits them to connect with DNA that has a negative charge. All four cores histone proteins have a long N-terminal domain and shorter C-terminal domain. Histone octamer consists of a modular assembly of two stable heterodimers of H2A/H2B and one H3/H4 tetramer.

The nucleosome structure is very highly conserved in all eukaryotic cells. Klug and colleagues using X-ray crystallographic analysis first observed nucleosome core particle structure in 1977. The structure of nucleosome was explained by several studies using electron microscopy and nuclease digest for offering a histone-DNA repeat structure of chromatin and providing the basic unit of the chromatin (Bryan Turner, 2005) (Figure 1.1).

1.1.2 Linker Histones

In addition to the four core histones, a linker histone can bind to the outside of the nucleosome particle, where it also interacts with the linker DNA (Zhou et. al., 2015). To begin with, linker histone H1 was found as an additional protein within a array of nucleosomes compacted into the 30 nm of fiber (Figure 1.2). Linker histones H1 have many posttranslational modifications, and it can interact with non-histone proteins. Moreover, linker histones can regulate the essential cellular function such as gene transcription, mitotic chromosomes formation, and separation, embryonic stem cells, muscle differentiation, the genetic activity of heterochromatin and cell pluripotency (Zhou et. al., 2015). The role of linker histones was determined using the three-dimensional organization of the extended chromatin fiber. Histones were found to interact with linker DNA at or near the entry/exit point of the core particle, which close two turns of superhelical DNA around the histone octamer. Moreover, SFM image showed that eliminate linker histones caused an increase in linker length due to a release of some DNA from histone core (Zlatanova et. al., 1998).

1.1.3 Histone variants

In most eukaryotic cells, histone protein genes can be found in gene clusters, and some of them become active during S phase of cell cycle process. These four histones help to assemble nucleosomes and to pack new DNA synthesis. To start with histone H2A contain 129 amino acids residues. It includes H2A.Z and H2A.X. In yeast, H2A.Z was found to prevent involved in gene expression activities whereas H2A.X was found to contain in double-strand DNA breaks as results of phosphorylation at serine residue at the C-terminal region of H2A-X. Moreover, this phosphorylation was found to help proteins to enroll in DNA repair (Marino-Ramirez et. al., 2005). Histone H2B contains 125 amino acids residues, and it has a role in gametogenesis, male gametic cells, bovine, and human spermatozoa. Histone H3 consists of H3.3, CenH3, and H3.4. Histone H3.3 was not found in S phase, but it was found in active transcription chromatin whereas H3.4 was found in testis in

spermatocytes and CenH3 was found in centromeric chromatin. H4 is highly conserved histones, and it has contact with other histones. The protein contains 102 amino acids residues, and it is important for nucleosome assembly due to the formation of half of the H3-H4 tetramer of the core particle (Marino-Ramirez et. al., 2005) (Figure 1.2).

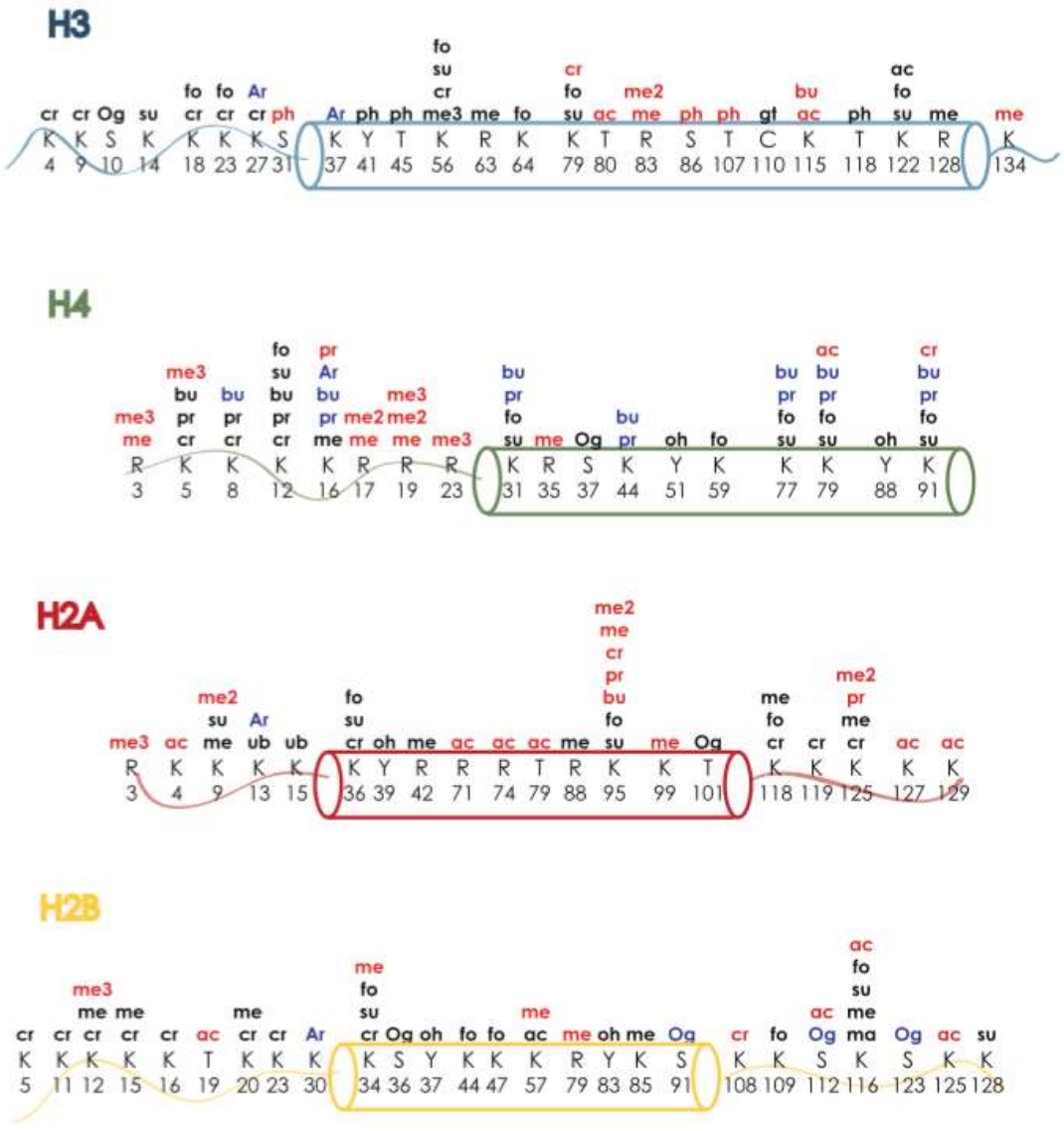


Figure 1.1 Histone modification patterns in vivo and in vitro
H3, H4, H2A and H2B histones are wrapped around 147 bp of DNA to form a nucleosome. The largest histone is H3 with 135 amino acid residues following H2A, H2B and H4 with 129,125 and 102 amino acids residues respectively. All four histones have long functional N-terminal region. This Histone region can undergo diverse post-translational modifications using the right combination and translation factors. These modifications cause a local and global change in chromatin condensation and gene expression. Modifications of histone are described in the text. This figure represents

histone variants and each histone H has amino acids residues in specific location that can be acetylated or methylated etc. on lysine (K) or serine (S) or tyrosine (T) or arginine (R). Black, modifications found in vivo in human; red, modifications found in mouse brain; blue, modifications found in vitro. Ac, acetylation; Ar, ADP-ribosylation; bu, butyrylation; cr, crotonylation; fo, formylation; gt, glutathionylation; ma, malonylation; me, methylation; Og, O-glcNAcylation; oh, hydroxylation; pr, propionylation; su, succinylation; ph, phosphorylation; ub, ubiquitination. Adapted from (Arnaudo and Garcia, 2013).

1.1.4 Structure of the chromatin fibre

Previously, chromatin fibre structure was studied based on in vitro experiment to describe nucleosome folding model called 30-nm fibre using physiological salt concentration. At low ionic strength, the use of the mathematical model for understanding extends chromatin fibre revealed that chromatin fibre depends on the range of DNA entering and existence the core particle and the linker length (Zlatanova et. al., 1998). Moreover, to determine the fiber structure of chromatin, it was shown that N and C-terminal of the core histone have associated with nucleosome and nucleosome interaction and interact with linker DNA and linker histones using imaging capabilities of SFM with conventional biochemical approaches. However, same experiments, which were in low ionic strength, were applied in high ionic strength to investigate the chromatin condensation (Zlatanova et. al., 1998). However, detecting 30-nm fiber was failed using small-angle X-ray scattering (SAXS) and electron spectroscopic imaging experiments (Nishino et. al., 2012) Using SFM technology allowed a quantitative measurement of single fiber level and applying three dimensions for each nucleosome showed centre-to-centre distance. However, these technologies have limitations in lacking the molecular specificity and histone visual activity in nuclei. Many methods were used to study the chromatin fiber such as super-resolution nanoscopy called stochastic optical reconstruction microscopy (STORM) (Rust et. al., 2006), cryo-electron microscopy, synchrotron X-ray scattering (Nishino et. al., 2012) and a combination of electron spectroscopic imaging with tomography, three-dimensional images (Fussner et. al., 2012) were applied to visualize the structure of chromatin

fiber at interphase nuclei. Nonetheless, these studies did not address the organization based on single nucleosome and the way of chromatin modification during cell pluripotent and differentiation (Ricci et. al., 2015).

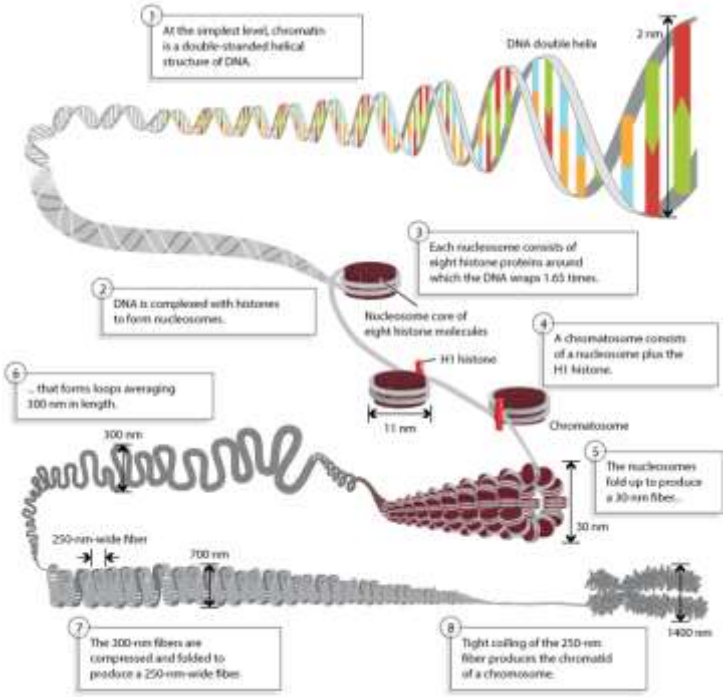


Figure 1.2 : the model of chromatin structure
A long DNA molecule with diameter 2 nm wrapped around a nucleosome, which contains core eight-histone octamer, with diameter 11 nm. The nucleosome and histone H1 form chromatosome. This nucleosome found to fold into 30 nm chromatin fibers. Then chromatin fibers formed loop 300 nm lengths. 300 nm length are more compressed fibers produced 700 nm fiber length, and subsequently, chromatid or chromosome was formed after tight coiling of 250 nm fiber which showed a high order organization 1400 nm duration of interphase nuclei or mitotic chromosome the picture adapted from (Anninziato, A., 2008)

Chromatin accessibility can indicate by active regulatory regions techniques. These techniques are DNAase-seq (deoxyribonuclease I) and ATAC-seq (transposase-accessible chromatin using sequencing). They have been used to measure the region of chromatin where it can be accessible to regulatory factors regardless of cell heterogeneity (Cusanovich et. al., 2015). A Recent study looked at cell-to-cell differences in chromatin using ATAC-seq technology for single cells from a mixture of mouse and human cell lines. From single cell interpretable results, it showed that NFkB

binding region was highly associated with chromatin accessibility, and it can be linked to extracellular signaling (Cusanovich et. al., 2015).

1.1.5 Chromatin condensation in mitosis

Changing the chromosome organization occurs during the cell cycle in the interphase chromosome. However, a study found that metaphase chromosome has same organization in different cell types (Naughton et. al., 2013 and Tatsuya Hirano, 2015). These results were found spatial segregation into cell type called A/B compartment and TADs (Topologically associating domains) are lost and no evidence for new compartments or interaction within the chromosomal band was found using homogeneous mitotic interaction map. Therefore, although it was indicated that epigenetic memory involved in transcription factors and chromatin formation, which will be explained later, at certain loci, the high order of chromatin has to form a new chromatin in early G1 with no epigenetic memory (Bell and Straight, 2015).

Chromosome condensation in bacterial and eukaryotic genome consists of two factors, DNA topoisomerase II and large protein complexes called condensins II and I. DNA topoisomerase II help to unwind the DNA strand during condensation process (Naughton et. al., 2013 and Tatsuya Hirano, 2014). Maintaining this compaction needs proteins called structural maintenance of chromosome (SMC) proteins, which are the core subunit of condensins II and I (Bell and Straight, 2015). The function of condensins in eukaryotic cell cycle process occurs during a conformational change of the chromosomes in the initiation of separation and during the suppression of gene activity (Naughton et. al., 2013 and Tatsuya Hirano, 2014). Moreover, the function of condensins was demonstrated in two models; first one is that the similar condensins structure in vivo acts as a bridge offering interaction system between distance chromatins; second is that condensins regulate the high-order of packed chromatin to derive the condensation. For example, human (Hela cells) and Chicken (DT40) cells were used antibody CAP-G protein against condensins I to perform chromatin immunoprecipitation

sequencing (ChIP-seq). This analysis revealed that condensins I was around centromere and 7,825 binding peaks overlapping with transcription starting sites (TSS) for several genes and nearly 70% of those peaks were on (RNA polymerase II) RNAP2 gene. In addition, Chip-seq and RT-qPCR in human cells showed condensin I peaks was located in prometaphase cells that linked to RNAP2 peaks and gene expression inhibitors did not find a relation with condensin I in the mitotic chromosome. Moreover, ChIP-seq results showed the mitotic effect of TATA-binding protein (TBP) depletion, which is important protein for the pre-initiation complex in transcription initiation process, did not change the condensin I binding to TSS. Another isoform of condensin protein complex involved in a large scale of transcription repression such as mouse condensin II was found to bind to active enhancers and promoters and to be important for ordinary gene expression in interphase embryonic stem cells (Sutani et. al., 2015).

1.2 Regulation of chromatin structure and gene expression

1.2.1 Nucleosome positioning and gene expression

Studies showed that nucleosome site around transcription starting site can modify the gene expression via repression or activation process (Buck and Leib, 2006). some factors, which can drive for easier access to their binding sites, are affected by nucleosome position such as free sites from nucleosome such as nucleosome close to linker DNA is more accessible than middle nucleosome. Transcription ability in many chromatin sites is influenced by many nucleosome properties including positioning, turnover and histone variation and modification, which play roles in gene regulation (Bai and Morozov, 2010). Nucleosome positions at specific region have been identified by the micrococcal (MNase) enzyme, which starts with chromatin crosslinked digested with micrococcal nuclease enzyme to generate mononucleosomes. This enzyme can digest linker DNA and remain the nucleosomal DNA undamaged. After MNase digest, the ligation-mediated PCR analysis was applied to map genomic, double-stranded MNase cleavages at the nucleotide level. Another application is that after isolating DNA, a

nucleosomal fragment is performed using microarray or ultra high-throughput sequencing. These applications offer an insight of chromatin structure and nucleosome position in different organisms. Genome-wide studies can provide a view of the nucleosome position in many genes. Promoter sites have less nucleosome due to its relation to transcription active sites; therefore, it is called nucleosome free regions (NFRs), which are located at the upstream of the transcription starting sites (TSS) such as high localizing position of +1 nucleosome at the downstream of TSS (Radman-Livaja and Rando, 2010 and Schones et. al., 2008).

However, the influence of the nucleosomal position in a different organism can be varied reflecting the differences in transcription process or the core transcription mechanism such as TFIIB (Radman-Livaja and Rando, 2010 and Schones et. al., 2008). For example, in human T cells, the 5' end of +1 nucleosome is located at 40 bp from TSS in genes with elongated RNA polymerase II and +10 bp in inactive genes with a delay in RNA polymerase II (Schones et. al., 2008).

Moreover, the regulation of gene expression through nucleosome can be divided into two classes. The first is the relationship between the nucleosome free region and transcriptional initiation, and the second one is the interaction between transcription starting site and intragenic nucleosome. Due to nucleosome conservation throughout the organisms, Tirosh and Barkai (2008) found two promoters linked to transcriptional plasticity, which identifies as changing the transcription level by owned gene under a different condition such as BRD4 in leukemia cells (Rathert et. al., 2015). These two promoters: one is large nucleosome free region (NFR) close to transcription starting site (TSS) with well-positioned nucleosome and the second one is the distribution of nucleosome dynamics close to starting place. These two promoters called depleted proximal nucleosome (DPN) and occupied proximal-nucleosome (OPN). DPN showed a less used region close to TSS such as histone H2A.Z, but more open region in far region whereas OPN showed high used region close to TSS, but less used area in the far region such as TATA box. For this reason; these strategies

were suggested for gene regulation in a different gene by chromatin (Tirosh and Barkai, 2008) (Figure 1.3)

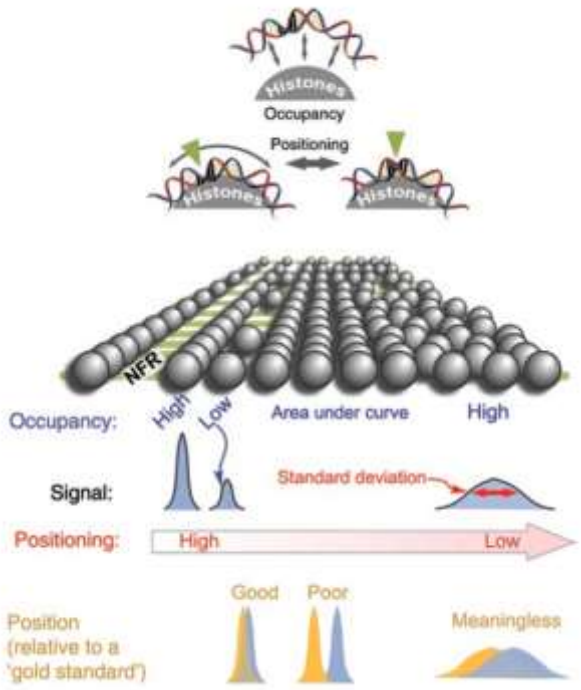


Figure 1.3 the differences between nucleosomal position and occupancy Upper panel presents the differences between occupancy and position in a cross-section of the nucleosome. Nucleosome occupancy determines the histone or nucleosome density away from nucleosome region whereas nucleosome position identifies nucleosome area with DNA. The lower panel presents the way that nucleosome occupancy and position measured. Nucleosome occupancy determines local density of nucleosome, which is the area under the curve whereas nucleosome position determines the standard deviation of the curve and how the spheres organized. Moreover, the nucleosome position determines how the two peaks separated. Low standard deviation means weak position whereas high position determines with two or more peak distance to each other (B Franklin Pugh, 2010).

Besides, to nucleosome roles in DNA packaging, nucleosome organization, and nucleosomal positioning affect gene regulation. For example, breast cancer cells study was found the small part of estrogen and progesterone receptors were found linking to the cytoplasmic side of the cell membrane through palmitoylated residues. Progesterone receptor (PR) and hormonal gene regulation study showed that PR bound to small fraction of hormone

responsive element (HRE), which was identified to have high sensitivity to DNasel hypersensitivity region. Moreover, HRE was enriched with nucleosome before hormone influence and HRE within these enriched nucleosomes were remodeled with no expelling from hormone induction. In addition to HRE, PR binding sites bind to open chromatin region known as DNasel hypersensitive sites (DHS) using MNase digest and cell sequencing to apply genome-wide nucleosome mapping before hormone induction. The results showed high nucleosome enrichment on the sites that have PR binding after hormone stimulation. Thus, a sustained and efficient PR binding is organized by nucleosomes enrichment in region lacking of PRE (Ballare et. al., 2012).

1.2.2 Nucleosome organization and gene expression in stem cells

Stem cell biological differences and cell fate decision has been linked to chromatin influencing gene expression. Teif and colleagues (2012) identify the position of nucleosome over genomic of three different mouse stem cells, ESCs, MEFs and NPCs deriving from ESCs by mapping the nucleosome occupancy al. This experiment was performed using genome-wide paired-end sequencing to map nucleosome position in mouse ESCs, NPCs and MEFs after MNase digest following ChIP-seq. in ESCs, some identified genes were nucleosome depleted region and nucleosome occupancy reduced to 40-80% in the same area in all participant cells. These transcription factors binding sites such as C-myc, Klf4, Stat3, Essrb, Sox2, Nanog and Oct4, and other factors such as p300 histone acetyltransferase, chromatin remodelers Chd7 and Brg1 and DNase I-hypersensitivity sites, determined by ChIP-seq in ESCs cells. they found that nucleosome position in all three cell lines is one of a factor that affects DNA-protein binding affinity known as “pioneer factors” which can start the nuclear program. Furthermore, they found nucleosomal depletion in some promoter regions at transcription starting sites such as a Smarca4 promoter (Teif et. al., 2012). Another researcher used BAC solution-hybridization technique to increase the genomic pool of mononucleosome targeting promoter Nestin, Oct4, Olig2, Pax6, Sox1 and Sox2 in vivo and in vitro nucleosome occupancy in embryonic and

differentiated stem cells. They found that Oct4 and Sox2 controlled the their target and other differentiated genes. Moreover, they identify nucleosomal occupancy at active and inactive Sox2 and Oct4 binding regions. They found no significant change in vivo and in-vitro although there are some different distributions of nucleosome occupancies (Sebeson et. al., 2015).

1.2.3 The dynamic nature of nucleosome

The active nature of nucleosomes was observed using X-ray crystallography and footprinting techniques (Cutter and Hayes, 2015). This technology revealed that nucleosome is highly dynamic and play essential roles in DNA accessibility by unwrapping and rebinding (breathing) transiently the surface of the nucleosome and regulating the position and density of nucleosome on the DNA sequence. Therefore, nucleosomes can provide thermodynamic as a barrier to binding factors (Cutter and Hayes, 2015). Besides thermodynamic for more access to DNA, remodelling factors such as nucleosome sliding and ATP-dependent chromatin, histone chaperones, chromatin-binding proteins such as non-histone proteins and linker histone H1 have roles in changing the nucleosomal structure and allow more accessibility to DNA sequence (Luger, 2006 and Luger and Hansen, 2005).

1.2.4 Nucleosome dynamic and stem cells

Nucleosome dissociation, which correlates with transcription factor binding, found to be linked with ATP-dependent remodeling systems such as nucleosome sliding and displacement, partial histone disassembly and substitution by histone variants (Voss and Hager, 2014). Moreover, ATP-dependent remodelling systems consist of enzymes that use energy from ATP hydrolysis to interrupt the association between histone protein and DNA and alter the nucleosome structure. Stem cell researchers have identified chromatin-remodelling enzymes to be regulating the expression or to collaborate with pluripotent factors for controlling chromatin structure and transcriptional machinery. For example, SWI/SNF, ISWI, and CHD are ATP-

dependent chromatin remodelling enzymes found to control many systems in embryonic stem cells (Saladi and Serna, 2010). Some of SWI/SNF and ISWI complexes contains Brig 1 subunit and Snf2h and Snf21 respectively. SW1/SNF has a role in self-renew and pluripotent maintenance and facilitate the process of differentiation whereas ISWI, Snf2h subunit can interact with nucleosome and unique biological function in heterochromatin replication and gene expression and Snf21 subunit has a role in fly development (Saladi and Serna, 2010). CHD is another ATPase subunit has roles in facilitating neuronal differentiation and affecting stem cell pluripotency state (Saladi and Serna, 2010).

1.3 Histone post-translational modification

An Important step in gene regulation and chromatin modification is posttranslational modification of histone. This modification in transcription and chromatin are affected by PTMs roles in recruitment specific factors or change the current location to erase previous interaction. From histone variants, the N-terminal domain is the most accessible site due to its localization in nucleosome core. Histone's tails are exposed to various modifications catalysed by writer enzymes such as acetylation, methylation, phosphorylation, ubiquitination, SUMOylation and other modifications (Bernstein et al., 2006 and Quina et al 2006).

Some modification can change the charge of histone and DNA by acetylation or phosphorylation in favour of nucleosome stabilization. The molecular weights of these modifications consider measuring from light such as acetylation, methylation and phosphorylation or heavy such as ubiquitination and poly ADP ribosylation (Bernstein et al., 2006 and Quina et al 2006). Furthermore, besides gene expression and chromatin modification, PTMs are found to be involved in many systems such as cell cycles, signal transduction, protein-protein and cell-cell interaction and intracellular and extracellular communication (Wang et. al., 2014). The main regulatory factors of histone modification include posttranslational modifications such as acetylation, phosphorylation, methylation and ubiquitination of histones.

These histones modification are involved in chromatin function and nuclear processes regulation such as RNA and DNA transcription, DNA replication and DNA repair (Xu et. al., 2013).

Epigenetic writers can add histone modification such as acetyl and methyltransferase, kinase and ubiquitinase, whereas eraser enzymes remove those histone modifications, such as deacetylase, phosphatase, demethylase, and deubiquitinase. In addition to writer and eraser enzymes, readers are domain regions that can identify histone modifications such as bromo- and chromo- domains, PHD fingers, and malignant brain tumour domain (Arrowsmith et. al., 2012, Marmorstain and Zhou 2014, Alexander Tarakhovsky 2010 and Blakey and Lltt 2016) (Figure 1.3).

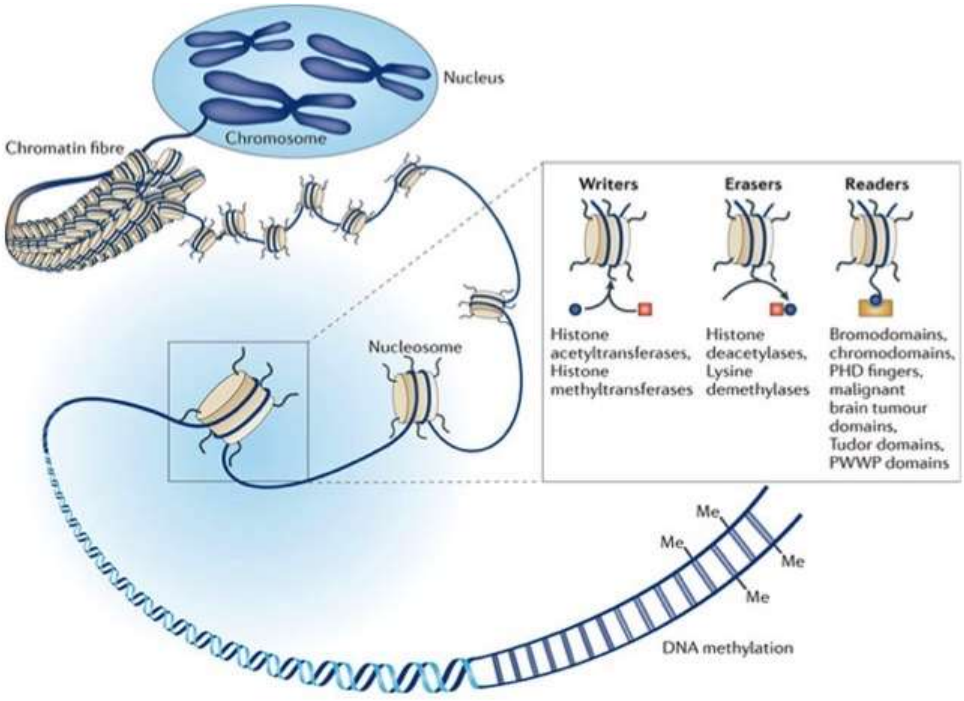


Figure 1.4 : the tools of epigenetics

Writers, such as histone acetylase, kinases DNA and histone methyltransferase and ubiquitin ligases, present histone marks (orange circle) and can enhance the PTMs on either DNA or protein. Erasers, such as histone deacetylase and histone demethylase, can remove histone marks from chromatin organization and prepare these histones for other modifications. Readers include proteins with domain such as bromo-, chromo-, Tudor-, MBT-, PWWP-, and PHD-domains. These proteins can bind to specific modifications and it can interpret a particular form of histone

modification and impose the change in chromatin structure adapted from (Hojfeldt et. al., 2013).

1.3.1 Histone acetylation

The most common posttranslational modifications of histones are the acetylation and methylation of lysine located at the histone tails (Jenuwein and Allis, 2001). Acetylation defines an interaction between organic compounds when it introduced an acetyl group into it. Either histones or non-histones proteins can be acetylated. Histone becomes acetylated by histone acetyltransferase (HATs) that derived the chromatin to be more open in favor of gene activation whereas decondensed the chromatin can be recruited using histone deacetylase (HDACs) (Eberharter and Becker, 2002 and Shen and Casaccia-Bonnefil, 2008) (Figure 1.6).

Histones hyperacetylation implicated to be related with active gene expression profiles. Moreover, some studies revealed that the acetylation could influence disruption of cellular processes such as DNA repair and replications (Trygve Tollefsbol, 2011). Moreover, in some modification acetylation in lysine residue 16 at histone H4 reduced the chromatin compaction and increases the gene expression in vivo and in vitro. In addition to H4K16ac, H3K155 and H3K122 are a novel acetylation sites on the lateral surface of the histone called dyad axis. This region is covered by one fiber of DNA where the histones bind to DNA. It was found that H3K122ac region has less nucleosome, but it enriched in transcription starting sites (TSS). Nevertheless, it shares the binding sites with other active gene markers such as H3K4me3, H2A.Z, and H3.3 and other enhancers such as H3K4me1 and H3K27ac. Another acetylated mark is H3K56ac, which may be active by p300, or GCN5 in human cells. H3K56ac occurs by interaction between histone chaperons (Asf1) and Rtt109. Moreover, it degraded during H3 synthesis and during G2 of cell cycle process. Therefore, it could have role as a response for DNA damage or in chromatin modification after DNA repair (Lawrence et. al., 2015).

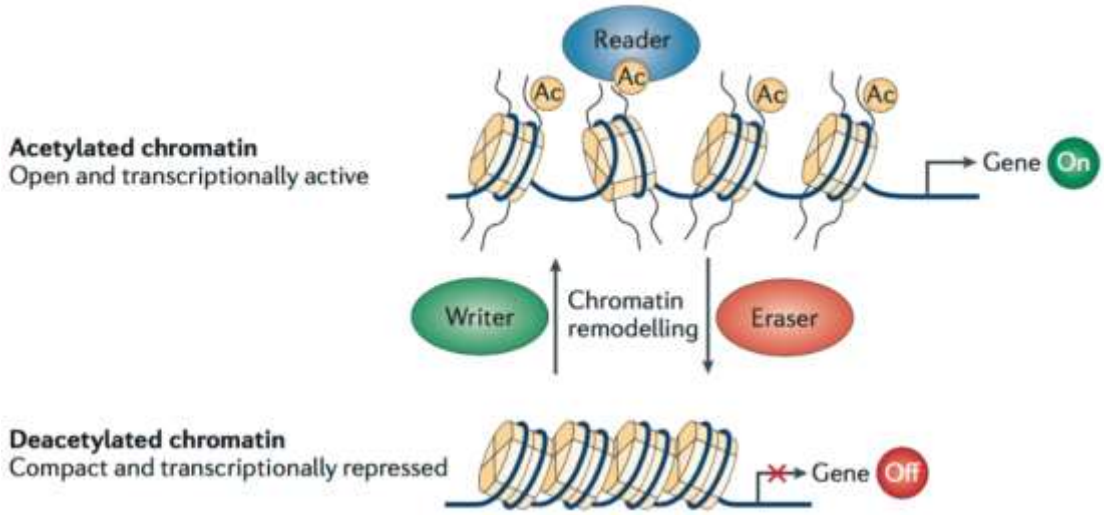


Figure 1.5 schematic representation of histone acetylation in case of chromatin modification and transcriptional process.
Lysine residues at the N-terminal tail of histone region of core histone proteins showed to be acetylated in H2B and H4 histones using histone acetyltransferase (writer) enzyme. This action helps to open the chromatin and promotes the transcriptional mechanism by binding to acetyl reader such as H3K9ac. On the other hand, the chromatin can fold to tertiary and secondary structure by histone deacetylase enzyme (HDACs) (eraser), which prevents the transcriptional process adapted from (Verdin and Ott, 2015).

1.3.2 Histone methylation

Methylated histone tails are found in arginine, lysines, and histidine. Monomethylated or dimethylated histones are found in arginine whereas lysines can have one, two, or three, methyl groups. Three enzymes, lysine-specific SET domain containing histone methyltransferase (HMT), non-SET domain containing lysine methyltransferases and arginine methyltransferases, are responsible for the methylation of lysine and arginine regions and each of them involves the location of lysine (Trygve Tollefsbol, 2011). In contrast, histone demethylase (HDMs) can derive to the chromatin condensation using H3K9me3 and H3K27me3 based on target gene. It was found that this mechanism could mediate the inflammatory and metabolic disorders (Duygu et. al., 2013) (Figure 1.7).

Methylation of the histone has a role in transcriptional regulation that could be related to activation, elongation, and repression. For instance, the lysine mono-, di- and a tri- methyl group of histone H3 (H3K4me, me2, and me3)

were found to be inactive promoters and it linked to transcription initiation and elongation. H3K4me3 was studied in mammalian cells showed the sharp peak of H3K4me3 at the transcription starting sites (TSS) of many genes besides it can positively regulate transcriptions through recruiting nucleosome remodelling enzymes and histone acetylates (Bernstein et. al., 2006 and Zhou et. al., 2011). In embryonic stem cells (ESCs) and Epiblast stem cells (EpiSCs), chromatin studies revealed that H3k4me3 has been identified in the broad range of high CpG promoter content despite the expression profile. On the other hand, H3K27me3 correlate with repressed promoters, which determine the transcriptional silencing mode in mammalian (Zhou et. al., 2011 and Hemberger et. al., 2009). Moreover, the main Polycomb repressive complexes PRC1 and PRC2 target a large part of high CpG island content promoters. However, in mouse embryonic stem cells (ESCs) and Epiblast stem cells (EpiSCs), 20% of high CpG island content promoters bind to Polycomb repressive complex 2 (PRC2) and marked by H3k27me3 (Hemberger et. al., 2009 and Zhou et al., 2011)(Figure 1.6).

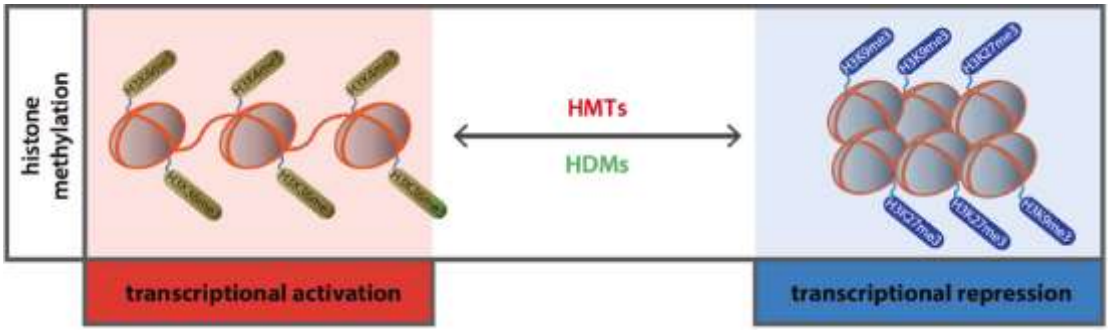


Figure 1.6 diagram of the role of histone methylation in transcriptional process
Histone methylation can stimulate either activation or repression of target gene. it is regulated by histone methyltransferase (HMTs) and histone demethylase (HDMs). HMTs derive the chromatin decondensation and gene activation using H3K4me3 whereas HDMs catalysed by H3K9me3 and H3K27me3 toward chromatin condensation and transcriptional inhibition. Adapter from (Duygu et. al., 2013)

1.3.3 Histone phosphorylation

Histone phosphorylation occurs during cellular response to DNA damage repair, transcription regulation, and chromatin condensation in apoptosis and mitosis and meiosis cell cycle (Rossetto et. al., 2012). For example, in mouse, H3 phosphorylation on T11 and T6 have involved in transcription regulation when they induced to androgen and to DNA-damage (Rossetto et. al., 2012). Another example, in chromosome compaction during mitosis, Histone phosphorylation regulates the serine/threonine kinase Aurora-B affecting H3 at serine location 10 and 28 (Figure 1.8) (Minarovits et. al., 2016). Phosphorylated serine 10 at H3 recruits members of 14-3-3 protein families to open chromatin and transcriptional activation especially when there is K9 or K14 lysine close to serine 10 acetylated. This double modification confirms full transcriptional activation of Hdac1 and p21 genes (Minarovits et. al., 2016) and MAPK and p38 signalling in interphase stage at cell cycle analysis (Cheung et. al., 2000) (Figure 1.8).

Phosphorylation has a high turnover rate around 30 min in 2-3 hours using isotopic pulse labelling method. It has been implied that Histone phosphorylation can regulate the nucleosome dynamic through changing the histone charge. For example, when the serine, threonine and tyrosine residues become phosphorylated, they obtain negative charge, which could cause repulsion against DNA. In contrast, adding nuclei residue to unphosphorylated histone tail protect the DNA from DNase I enzyme more than phosphorylated histone tail. This result suggests that when the repulsion charged on histone tails, it opens the nuclear DNA sequence (Blakey and Litt, 2016). Phosphorylated Histone H3 regulates induction of many genes by driving the extracellular signals in different cell types and possibility plays roles in transient derepression of promoters silenced by histone H3K9me3 (Minarovits et. al., 2016) (Figure 1.7)

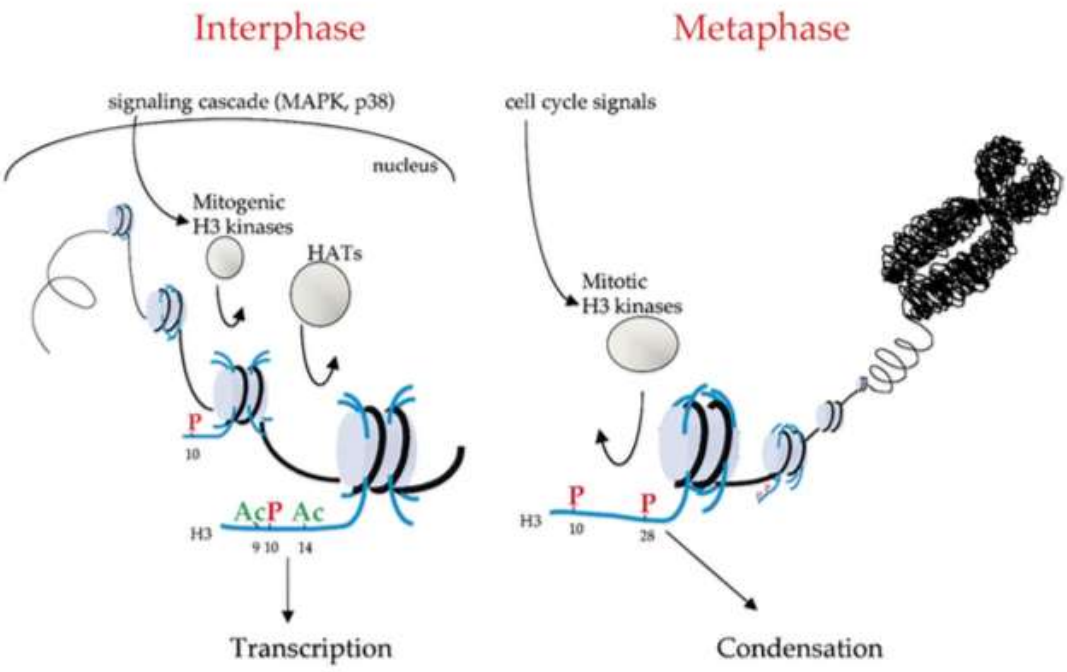


Figure 1.7 phosphorylation of histone H3 in interphase and metaphase cell cycle

The diagram represents the mechanism of histone phosphorylation at chromatin status in cell cycle analysis. The activation of MAP kinase and p38 pathway (right diagram) showed activation in H3 kinase of mitogen such as Rsk1 or Msk1. Coincidently, this process occur when there is an activation of acetylation by histone acetyltransferase enzyme at lysine 9 and 14 which is located close to phosphorylated serine 10 at histone H3 tail region. Both actions derived the transcriptional activation in the interphase region in the cells. On the other hand, in the metaphase (right diagram) region, where the nucleosomes condensed into chromosome, another signalling pathway involved into chromatin condensation through H3 Kinase inhibition. (Adapted from Cheung et. al., 2000).

1.3.4 Histone ubiquitination and Histone SUMOylation

Ubiquitin is defined as a protein with the addition of 8.5-kDa polypeptide of 78 amino acids using ubiquitin activation, conjugation and ligation enzymes. These enzymes are histone ubiquitin ligase and histone deubiquitinating enzyme. Histone ubiquitin ligase identified polycomb group protein RING1B that is responsible for H2A monoubiquitination at lysine 119. Moreover, in breast cancer gene BRCA1, ubiquitin ligase enzyme was found to cooperate with BRCA1 to increase the reaction of monoubiquitination H2A/H2A.X. On

the other hand, histone deubiquitinating enzyme acts as an eraser for ubiquitin modification and it involved in: silencing HOX gene, inactivating X-chromosome and progression in cell cycles mediated by H2Aub. For example BRCA1 associated protein 1 (BAP1) was found to be H2A-specific histone C-terminal hydrolase and it can remove monoubiquitin from H2A (Cao and Yan, 2012). Histone ubiquitination involves in several process in nucleus including transcription, DNA repair, X-chromosome inactivation process, stem cells maintenance and differentiation (Wadosky and Willis, 2012).

Another important ubiquitination of the histone is H2Bub1. This modification was found to regulate the transcriptional elongation stage, DNA replication and cell cycle. H2Bub1 occurred at lysine 123 in yeast and at lysine 120 in human and was found at low level in vivo. A group of researchers found that H2B ubiquitination regulated the H3K4me and H3K79me using Rad6 gene (ubiquitin-conjugating enzyme E2 in yeast) through trans-tail mechanism. Therefore, the results suggested that H2Bub1 has role in the regulation of transcription and gene silencing in heterochromatin region in yeast. Another study used chromatin immuniprecipitation analyses (ChIP) found that Rad6 cooperatively with H2Bub1 may identify the ORF of target gene. This cooperation relay on other complex called PAF (it was found in human and yeast), which was found to be critical for activation and interaction with Rad6 complex. This direct relation between PAF and Rad6 complex could promote the elongation step of Pol II to generate H2Bub1 and the subsequent of H3K4me and H3K79me in the target gene (Laribee et. al., 2007) (Figure 1.8).

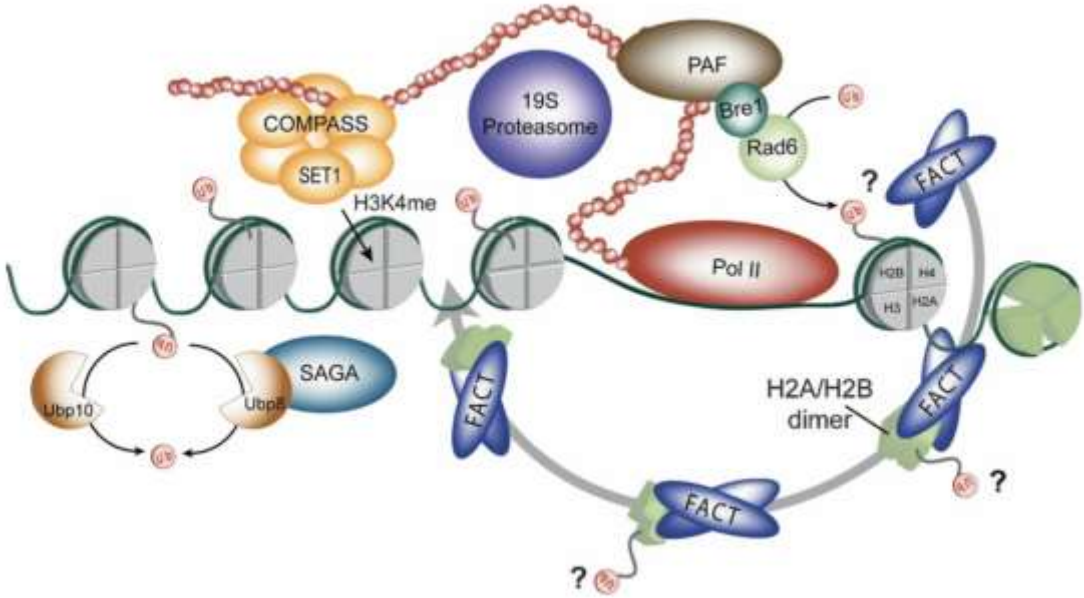


Figure 1.8 the function of ubiquitination in histone H2B
A diagram represents the dynamic of ubiquitination in chromatin. Ubiquitin was first added to histone H2B by BRE1/Rad6 complex. This action enhances the promoter POL II through PAF transcriptional elongation. However, in the other side of PAF complex, another complexes called COMPASS with SET1 recruits H3K4me. When the ubiquitination existed in H2B, other complexes become activated such as 19S proteasome. During the transcription elongation, H2A/H3B dimer is removed by FACT (facilitate chromatin transcription) and enables Pol II elongation via chromatin. H2Bub1 can be removed by SAGA that linked with Ubp10 or Ubp8 for fixing the methylated histone. Question mark on ubiquitin at the H2A/H2B suggests that FACT may recruit the ubiquitination in this dimer during nucleosome disruption or reassembly. This disruption associated with proteasome and may help to establish methylated histone. Adapted from (Laribee et. al., 2007).

Sumoylation is very similar to ubiquitination. It contains 101 amino acids polypeptide, approximately 10 kDa, with 18% sequence homology and different surface charge. PolySUMOylation is associated with stress response unlike monoSUMOylation is linked to regulating gene expression, nuclear transport, and protein-protein interaction regulation (Wadosky and Willis, 2012) SUMOylation and ubiquitination share the same enzymes performance inactivation, conjugation and ligation (Wadosky and Willis, 2012 and Minarovits et. al. 2016). Histone SUMOylation connects with transcriptional suppression and it occurs in all four histone cores. For example, histone H3 in Hela cells was found to be SUMOylated at lysine 19 (K19) (Hendriks et.

al., 2014). Another role of SUMOylation may affect the nucleosomal dynamics. For example, the interference between chromatin condensation and transcriptional suppression may be mediated when H4K2 is modified by SUMO-3. This could recommend the SUMOylation interfered with promoter and enhancers with direct interaction. On the other hand, histone SUMOylation may interact with many proteins indirectly in favour of chromatin regulation (Blakey and Litt, 2016).

A novel SUMOylation was found in Wiskott-Aldrich syndrome (WAS) which is an X-linked immunodeficiency and autoimmunity arising from mutation in *WAS* protein. This mutation affects the regulation of chromatin in T_{helper} (T_H) cells. Previous research found that there is promoter as coactivator of nuclear-WASp for transcriptional regulation in T_H1 gene. The results showed WASp is linked with SUMOylation pathway throughout SUMO1. This relation was due to a SUMO motif in WASp in lysine 66, 76, 144, 147, and 235. Applying mutation at in WASp hindered NF-kB signalling in T_H 1 cells. Furthermore, a mutation SUMOylation in WASp helped COMMD1 function as an inhibitor for NF-kB signalling to remove the NF-kB from DNA, activated HDAC6 and reduced the acetylation in H3K14ac at p300 promoter of NF-kB genes in T_H 1 cells (Sarkar et. al., 2015) (Figure 1.9).

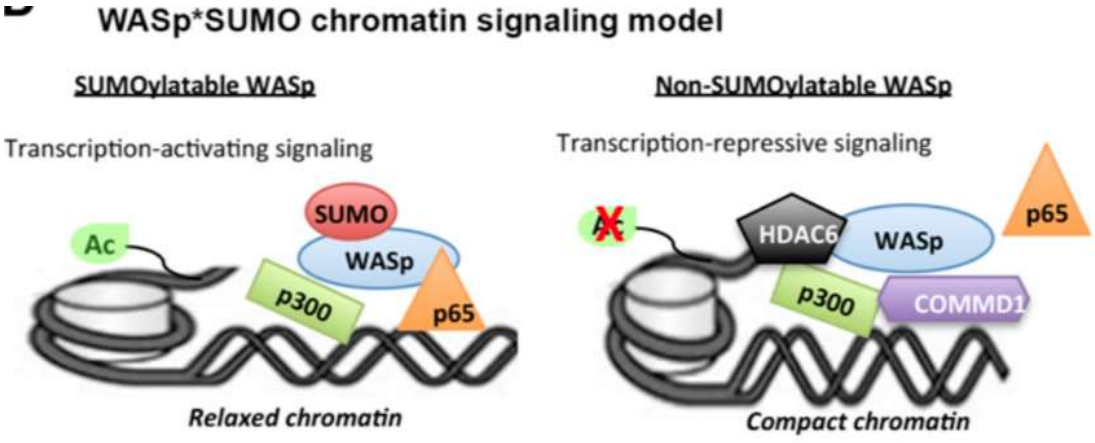


Figure 1.9 the roles of SUMOylation deficient in nuclear-WASp chromatin signalling

In normal case (left diagram) SUMOylated WASp helps to bind with p65 and target the promoter p300 in NF-kB signalling under acetylated H3K4 to derive the transcription process. In mutated SUMOylated WASp, there is less acetylated H3K14 and no p65 effects. In this case, HDAC6 activates the

WASp binding with COMMD1 and p300 to derive the chromatin compaction and repress the NF-kB signalling Adapted from (Sarkar et. al., 2015).

1.3.5 Other large modifications

Using NAD+ as a substrate, ADP-ribosylation is formed by the addition of an ADP-ribose modified residues on protein; thus, it imparts a negative charge. ADP-ribosyltransferase (ART) modifies Glutamate residues histone. There are three forms of this protein named mono-, poly- ribosylation (PARP) and O-acetyl- ADP-ribosylation. It was found that mono-ADP-ribosyltransferase is the main of all the core histones and linker histone H1. This cooperation can involve in cell cycle stage or response to genotoxic stress, proliferation activity and differentiation termination (Minarovits et. al., 2016 and Trygve Tollefsbol, 2011). For example, a study was found in single strand break, PARP1 seems to have some roles in histone through poly-ADP-ribosylation the linker histone H1, but PARP2 prefer core histone. Moreover, they observed when PARP1 and PARP2 become poly-ADP-ribosylation, C- and N-terminal at the histone H1 and H2B cause relaxation in chromatin for access the single strand break repair. (Trygve Tollefsbol, 2011 cited from Kreimeyer et. al. 1984).

Deminiation is named to remove methylated arginine forming citrulline, which is not possible to be methylated. This irreversible reaction derived by peptidyl arginine deiminase 4 (PADI4) causing a structural modification in histone H3 and H4. Therefore, citrullination may involve in chromatin structure and gene regulation (Minarovits et. al., 2016).

Histone proline isomerize are a post-translational non-covalent modification of histone tails and it can naturally interconvert between cis- and trans backbone conformation. In yeast, Isomerization is slow from minute to hours to fold the protein (Monneau et. al., 2013). Moreover, proline isomerases were shown to be facilitated the interconversion between cis- and trans-. This function may affect the stability of histone tails leading to effect on gene expression. For example, Nelson and colleagues (2006) identify Fpr4 binding to the N-terminal tail of histone H3 and H4 through nucleolin-like

domain and found two substrates of Fpr4 called H3P30 and H3P38 that are close to Fpr4 binding sites. Modified H3 tail was found methylated in lysine K36 by set2 histone methyltransferase. This indicates that there is communication between H3P28 substrate and methylated H3K36 (Nelson et. al., 2006 and Bannister and Kouzarides 2011).

Histone modification O-linked B-N-acetyl glucosamine, O-GlcNAc, is an unchanged acetylated hexosamine sugar linked through glycosyl linkage. This histone was found in cytosolic, nuclear and mitochondrial proteins and mark histone cores. Therefore, it affects major cellular signalling mechanisms (Bond and Hanover, 2015 and Minarovits et. al., 2016).

Histone lysine crotonylation (KCR) is a novel histone modification that has an active role in post-meiotic testis. It marks a group of X/Y genes that is active in haploid spermatids (Montellier et. al., 2011). Moreover, it was found that p300, which is histone acetyltransferase, derived to histone crotonylation is a strong transcriptional activator than histone acetylation in Hela cells based on the nature of intracellular of crotonyl-CoA. This could have major effects between cellular metabolism and gene expression (sabari et. al., 2015)

1.3.6 Methylation and acetylation of non-histone proteins

Non-histone proteins are important members of chromatin architectures that maintain chromatin condensation causing to form high-order structure and decompaction causing more chromatin accessibility leading to more gene expression. Some proteins are responsible for chromatin high order structure such as HMG, HP1, MeCP2, polycomb group proteins, and PARPs. Many studies focused on the methylation of non-histone proteins such as histone lysine demethylases (PKDMs) and histone methyltransferase (PKMTs) due to their influences in many cellular functions such as cell signalling transduction and tumorigenesis (Carr et. al., 2015). Non-Histone acetylation was reported in few studies although it involves in cell signalling, transcription and protein stability, cancer, apoptosis and cell cycle. P53 protein can be acetylated in lysine 120, 164, 370, 372, 373, 381, 382, and 386. For example, when p53 become acetylated lysine at position 120, it

enhances the expression of a gene that is involved in DNA damage-induced apoptosis using hMOF and tip6 (Singh et al., 2010).

1.4 Chromatin in embryonic stem cells

Embryonic stem cells (ESCs) originated from inner cell mass before the implantation stage (see section 1.2). Genomic DNA arranged into chromatin, which influences the transcription, translational and other regulation processes. Besides the organization of chromatin, studying chromatin in stem cell is considering a significant challenge due to stem cell existence in pluripotent and some somatic cells and the difference in chromatin organization between embryonic stem cells and differentiated cells. Chromatin found in pluripotent stem cells with less condensed and more hyperdynamic related proteins. These chromatin features play important roles in stem cell maintenance and self-renewal (Mattout and Meshorer, 2010).

Chromatin features in stem cells were observed by genome-wide study in mouse and human stem cells when researchers found that methylation and acetylation of histone H3 at lysine 4 was more dominant than DNA methylation. This action caused an increased in embryonic stem cells activities throughout the genome (Chen and Dent, 2014). Coincidently, some embryonic stem cells genes deriving the cell differentiation enrich with high conserved noncoding elements (HCNEs). One this modification was found to have a large region of H3K27me3 hindering a small area of H3K4me3 termed “bivalent chromatin.” in pluripotent stem cells lineage, Oct4, Nanog, and the Sox2 binding domain was mapped using genome analysis. It was found around 50% of bivalent domain occurred with at least one binding site of these transcription factors such as Oct4 and Nanog interaction. Moreover, it was found that more transcription factors committed to bivalent domain tend to be in the poised state whereas many of bivalent domains have a correlation with Oct4, Sox2, and Nanog. However, in differentiation state, transcription factors show to be active or repressive. In fact, the bivalent domain associated with Polycomb response elements and Trithorax-group

proteins that were found to provide epigenetic memory for cell committed lineage (Bernstein et. al., 2006). Nevertheless, current research found that some bivalent sites in ESCs are modified based on cell culture condition. For example, the standard serum medium for mouse stem cells contains pluripotent cells in addition to some differentiated cells showing cells heterogeneity in morphology and gene expression. on the other hand mouse stem cells cultured in no serum medium with two inhibitors, GSK3, and MEK, tend to show high enrichment in H3K4me3, H3ac, and H4ac, but low level in H3K27me3 and almost no heterogeneity exist (Chen and Dent, 2014) (Figure 1.10).

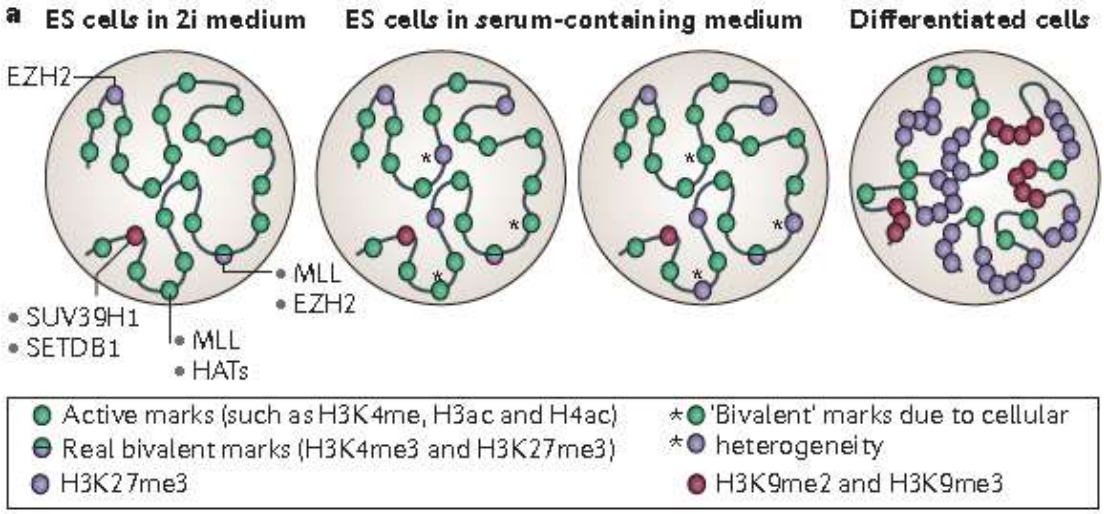


Figure 1.10 chromatin organization in pluripotent and differentiated stem cells Stem cells in 2i medium (serum free media) show more active marks (H3K4me, H3ac, and H4ac) than repressive mark (H3K27me3) whereas stem cells in medium with serum show cell heterogeneity, bivalent markers causing by heterogeneity, more repressive marks, and active mark. The differentiated cells, stem cells committed to differentiation contain more repressive mark than active marks and more repressive marks, H3K9me2 and H3k9me3. Adapted from (Chen and Dent, 2014).

1.4.1 Chromatin structure in stem cells

This project uses embryonic carcinoma cells, which are a malignant cell similar to embryonic stem cells. Therefore, it is important to consider whether the chromatin structure in embryonic stem cells is different to that

in somatic cells and cell lines. Embryonic Stem cells and differentiated stem cells such as neural or cardiac stem cells contain signals that help to distribute epigenetic modification through the changes in chromatin structure and transcriptional process. A study of human embryonic stem cells (H1) examined higher-order of chromatin structure from the lower position of the chromosome using chromosome-span haplotypes. It found significant changes in chromatin structure at these posts. Moreover, the study highlighted chromatin structure varieties among homologs chromosome. They found high similar to genome folding sections (A/B compartment) whereas only 0.6-2.3% of genome differences indicated between alleles in each lineage. These small changes in genome influenced chromatin structure between alleles regions in homologs chromosome (Dixon, et. al., 2015) (figure 1.11)

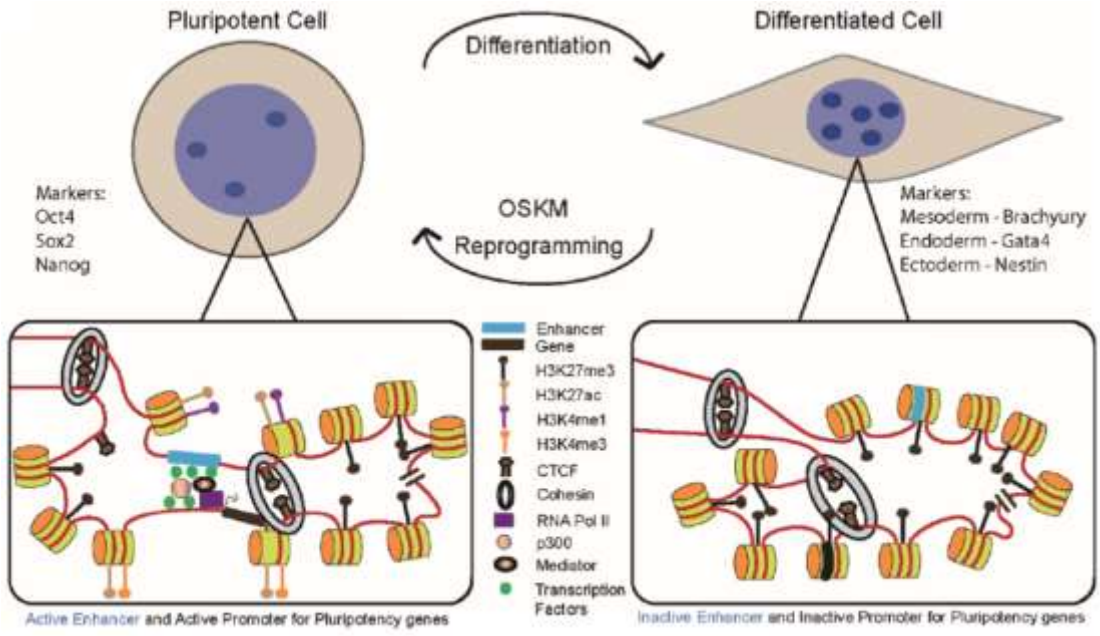


Figure 1.11 the diagram for chromatin accessibility and modification in pluripotent, differentiation, and reprogramming stem cells.
This chart represents the difference between chromatin organization from pluripotent stem cells to differentiated stem cells and vice versa. Chromatin in pluripotent cells shows genes, such as Oct4, Sox2, and Nanog, are enhanced by downstream regulatory factors such as transcription factors, RNA Pol II, p300, mediator and active enhancer. While downstream factors increased gene transcriptional process, the same target gene promoted by upstream factors such as active CTCF, and histone modification such as H3K4me3 and H3K27ac. Differentiated stem cells for the same gene undergo to silence events. However, The committed cells derive active

differentiated genes such as Brachyury in mesoderm, Gata4 in endoderm, and Nestin in the ectoderm. In this case, inactive enhancer and target pluripotent gene lied on silenced chromatin under repressive marker H3K27me3. In inducing pluripotent stem cells (cell reprogramming), the differentiated cells reactive the silenced pluripotent markers, Oct4, Sox2, Klf4 and Myc to generate embryonic stem cell like. Adapted from (Shchuka et. al., 2015).

1.4.2 Chromatin accessibility in stem cells

Besides chromatin structure modification, accessible chromatin can be affected by histone modifications and ATP-dependent chromatin remodeling (Delgado-Olguin and Recillas-Targa, 2011). Moreover, genome-folding section (A/B compartment) may affect the chromatin accessibility (Dixon et. al., 2015). Besides, chromatin accessibility can influence the structure of chromatin through the transcriptional process by changing the regulatory proteins in distinct gene loci (Meshorer and Misteli, 2006). For example, the stem cells genome-wide study showed that BRG1 (ATPase subunit of SWI/SNF-BRG1 chromatin remodeling complex) was found to share the location with other pluripotent markers and the loss of BRG1 caused termination in self-renew and cell reprogramming. These finding suggesting that the complexes of SWI/SNF-BRG1 are critical for pluripotent cell process. However, there is another factor that supports the self-renewal called LIF (leukemia inhibitory factor) throughout activating the STAT3 signaling pathway. Another study showed that SWI/SNF-BRG1 complex maintain open-chromatin status throughout STAT3 signal when it antagonists with Polycomb group proteins complex (Chen and Dent, 2014). Histone modification, H3K9me3 is linked with heterochromatin sites in pluripotent and reprogramming cells. These regions organized in 10 nm chromatin fibers similar to the open chromatin region. However, some of closed chromatin regions in stem cells such as, iPSCs, have been derived from unique transcriptional activities such as Oct4, Sox2, Klf4, and Myc, which able them to access chromatin, change the enhancers and permit the binding of other factors and vice verse (Figure 1.13) (Shchuka et. al., 2015).

1.4.3 Histone modification in embryonic stem cells

Monitored gene expression is important to maintain and establish cell identity in stem cells. One of the methods to control gene expression is by chromatin modification, which was found to control gene expression. Histone modification factors and chromatin remodeling can change the access of regulatory protein to DNA sequence. Therefore, the four histone-codes can be read in trans-acting factors to translate the chromatin organization and gene activity by becoming acetylated, methylated, phosphorylated, ubiquitinated, SUMOylated, or any other modification changes (Lunyak and Rosenfeld, 2008 and Schmittwolf et. Al., 2005).

1.4.3.1 Histone acetylation and methylation in stem cells

A study found that histone acetylation was found stimulating neuronal stem cells (NSC) to differentiate to hematopoietic stem cells in trans. Previously, it was known The resistance activity of histone acetylation transferase (HATs) and deacetylases (HDACs) control the level of histone acetylation. However, much microbial such as Streptomyces can inhibit the HDACS and One of these inhibitors is trichostatin A (TSA) that can indirectly stimulate the acetylation of histones, cell cycle arrest and apoptosis. Moreover, Inhibitor 5-aza-2'-deoxy-cytidine (AzaC) can bind to DNA covalently and prevent to retain methyltransferase (Dnmt1), which allow an increase in demethylation (Schmittwolf et. al., 2005). In this study, they treated neurosphere cells with TSA inhibitor for acetylation combined with AzaC inhibitor for methylation before transplantation. They found that treatment elevates the level of H4 acetylation and reduce the level of methylation at CpG Island in protein 2 (MeCP2). They conclude that the treatment increased the histone acetylation and decreased cytidine methylation (Schmittwolf et. al., 2005). Another study found that modification in epigenetic may allow generating neuronal from non-neuronal stem cells such as bone marrow stromal stem cells (MSCs). This study demonstrated the addition of histone deacetylase inhibitor such as Valproic acid can turn

the bone marrow stromal cells to neuronal stem cells (NSCs) (Woodbury et. al., 2002).

Activation the transcription was one of the factors that associated with histone acetylation and methylation. Tight marks of histone H3 on lysine 4 acetylation and methylation are more likely to contribute to establishing an active epigenetic profile, such as λ5-VpreB1 locus in stem cells, not by generalized histone acetylation and methylation of histone H3 lysine 4. To confirm this hypothesis, this study used Mouse lambda5-VpreB1 locus marked by histone acetylation (H3k4ac) and histone lysine 4 methylation (H3K4me) at individual sites in ESCs. These epigenetic marks distribute from modification sites to genes at the later stage of B-cell development. Next, the large active chromatin was found to establish pre-B-cell when the genes are fully expressed. The recommendation in this study was that epigenetic mark is crucial for achieving transcriptional accomplishment for lambda5 and VpreB1 genes at the early stage of pluripotent ESC stage to establish the region for bivalency (Szutorisz et. al., 2005). Another study was on Neocortex extracted from E18 rat embryo to study the epigenetic regulatory mechanism of gene transcription and differentiation such as FGF2, CNTF, STAT and GFAP. In this experiment, ChIP experiment was performed against H3K4me and H3K4ac. They found that ChIP for anti-H3K4ac on the STAT-binding site of GFAP promoter was high due to CNTF with irreversible action by FGF2 treatment whereas ChIP for anti- H3K4me increased in FGF2 binding and suppression in H3K9me in GEAP promoter (Song and Ghosh, 2004).

Another evidence showed that acetylated histone is associated with active transcription such as Klf4 and C-myc. Klf (Kruppel-like factor) family are transcription factors involved in cell differentiation and proliferation. P300/CBP, the protein with the binding protein, was found to be associated with histone acetylation by being recruited to the specific region of DNA using other transcription factors such as p53 and KLF6 interaction with p300. Besides KLF6, in human cancer cells, KLF4 was found to be acetylated by p300 in, which is important for transactivation. In addition, Klf4 inhibition by mutating p300 or overexpression of Klf4 was shown to be

directly interacting with Klf4 and it suggested the KLF4 regulate histone acetylation differently at the promoter of target genes (Evans et. al., 2007). Same results were observed in C-myc interact with histone acetyltransferase (HAT) including p300 and CBP and it may induce histone acetylation which permits Oct4 and Sox2 to bind to their target loci in inducing pluripotent stem cells (Takahashi and Yamanaka, 2006).

Histone methylation was identified as a bivalent modification, which contains large sites of histone H3 at lysine 27 and small sites of histone H3 at lysine 4 methylation in mouse embryonic stem cells. These modified histones are catalysed by trithorax- and polycomb-group proteins. Methylation of lysine 27 of histone H3 (H3K27me) is associated with silence chromatin and transcriptional repression whereas Histone H3 methylated lysine 4 (H3K4me) is defined as active chromatin associated with transcription factor genes that have roles in embryonic development and lineage specification such as HOX genes (Bernstein et. al., 2006). One of polycomb-group proteins in drosophila embryo called Polycomb repressive complex (PRC2) have roles in maintaining the structure of repressive chromatin through the function of chromatin modifiers such as enhancers of zeste (EZh2), embryonic ectoderm development protein (Eed) and suppressor of zeste 12 (Suz12), histone methyltransferases responsible for depositing H3K27me3 mark on chromatin. However, the study did not explain fully embryonic stem cells model deficiency for PRC2 component. Mutant ESCs showed growth lost of H3K27me but remain able for self-renewal and maintain normal morphology (Lunyak and Rosenfeld, 2008). Moreover, the factors of DNA sequence-specific can provide a site for catalysed specific enzymes that can remove or add the modification and substrate of the enzyme. For example, leaving the self-renewal state is accomplished via altering the covalent modification of the histone such as increasing silencing-associated histone H3k9me2 and H3K9me3 which marks on chromatin and removed H3K27me (Lunyak and Rosenfeld, 2008). Moreover, ESC transcription factor Oc4 was found to regulated H3K9me2 and H3K9me3 demethylase genes, Jmjd2c and Jmjd1a, Together those genes Jmjd2c and Jmjd1a in knockdown condition, it led to stem cell

differentiation with a reduction of lineage marker genes (Lunyak and Rosenfeld, 2008).

1.4.3.2 Histone phosphorylation in stem cells

Phosphorylation on histones has frequently occurred at specific sites throughout cell division. It was found binding at serine, threonine or tyrosine residues and catalysed by an amount of protein kinase and dephosphorylation by phosphatases (Ding and Wu, 2015). Previously, it was mentioned that phosphorylation involved in chromatin condensation. One of the mechanisms of signalling for maintaining self-renew is Jak2 signalling and for chromatin modulation is ERK pathway. For example, in mESCs, LIF-independent role for Jak signalling has been demonstrated in the phosphorylation of histone H3 on tyrosine 41. This circumstance caused a decrease in heterochromatin protein 1 alpha (HP1α) binding site on pluripotent genes (Fagnocchi et al., 2015). Mutated Jak2 can possibly phosphorylate and impede Protein arginine N-methyltransferase 5 (PRMT5) preventing histone arginine methylation. Consequently, it favours for uncontrolled haematopoietic progenitor cell expansion (Fagnocchi et al., 2015). However, MAP kinase signalling supports differentiation of stem cells during JNK-mediated H3 serine 10 phosphorylation of its target genes (Fagnocchi et al., 2015). Other mechanisms of singling called ERK pathway. It can regulate PRC2 deposition at developmental genes, by phosphorylating the RNA polymerase II at serine 5 and establishing poised domain (Fagnocchi et al., 2015).

1.4.3.3 Histone ubiquitination and SUMOylation in stem cells

Ubiquitination in embryonic stem cells was reported to affect the self-renewal and differentiation. For example, BMI1, proto-Oncogene polycomb ring finger, is important for self-renewal and maintenance of hematopoietic and neural stem cells (Cao and Yan, 2012). Moreover, in mouse stem cells, ubiquitinated H2A has an ability to stop RNA polymerase II during regulator

gene development and RING1A and RING1B subunit loss cause a release in silencing RNA polymerase II and other genes (Cao and Yan, 2012). During cell development and epigenetic regulation, histone ubiquitination (H2A) was found to play important roles in chromatin compaction at target site through Polycomb-group proteins PRC1 and PRC2 in ESCs (Endoh et. al., 2012).

Besides ubiquitination, histone SUMOylation may have roles in DNA damage repair and possibly in telomere activity in eukaryotic cells and probably in human stem cells, reproductive cells and cancer cells. Telomerase that maintain telomere integrity in those cells, function by adding TTAGGG repeats on telomeres. However, there is an alternative mechanism in cancer for retain the telomere integrity (ALT) which depends on the association with DNA repair mechanism. Similarly, some DNA repair protein was found to be affiliated with SUMOylation substrates such as Blm and Wrn. Nevertheless, sumoylation activity on DNA damage repair protein from telomere side in eukaryotic whereas in yeast SUMO and PIAS deletion caused an increase in telomere length (Bischof and Dejean, 2007).

1.5 High mobility proteins family (HMG)

Another nonhistone is high mobility group proteins (HMG) is a superfamily and it contains HMGA, HMGB and HMGN proteins existing in all vertebrate cells. Each protein has subgroups of proteins and other DNA or chromatin binding motifs. The functional motifs of HMG proteins associated with DNA and nucleosome to modify chromatin and gene expression (Postnikov and Bustin, 2015). For example when HMGB1 and HMGB2 become acetylated at lysine 2 through CBP, it caused an increase in DNA binding ability (Figure 1.8) (Stros,, 2010). Each member of HMG proteins interacts with chromatin differently. This means, that all HMGs promote toward open chromatin and form a change in transcription, replication and DNA repair mechanism. Moreover, they can interact with other regulatory factors and bind directly to DNA to alter the transcriptional process (Postnikov and Bustin, 2015) (FIGURE 1.12).

Figure 1.12 the localization of HMGN, HMGB, and HMGA in the nucleosome
HMGNs (red circle) are DNA-protein binding sites. They bind to nucleosome core domain. HMGBs (yellow circle) bind to linker DNA. HMGA (green circle) bind to miner groove in AT rich DNA.

Further roles of HMG proteins are highly regulated by their post-translational modification including: acetylation, methylation, phosphorylation, and ADP-ribosylation etc. These biochemical modifications are dynamic and rapidly responsive to intra and extra cellular signaling event since different families of HMG proteins have distinct modification profiles, hence different biological implication. HMGA, HMGB and HMGN share phosphorylation, acetylation, histones interaction and roles in cancer (Zhang and Wang 2009).

1.5.1 HMGA and HMGB

High mobility group A (HMGA) non-histone chromatin proteins is characterized by short repeat located in the minor groove of DNA called "AT-hook" motif (Ozturk et. al., 2014) and it is high positive charge except high negative charge in C-terminal region (Sgarra et. al., 2004). It contains HMGA1a, HMGA1b, HMGA1c and HMGA2 (Ozturk et. al., 2014). HMGA1 proteins came from alternative splicing of HMGA1 gene, which was found to be important for development and diagnostic of cancer. For example, human breast cancer cell lines showed high expression level of HMGA1 in tumorigenesis rather than benign cell lines (Maurizio et. al., 2016). Moreover, HMGA can be modified based on their roles in chromatin formation, transcription and embryonic development (Ozturk et al., 2014).

At the chromatin structure level, HMGA acetylation can help to decondensed chromatin structure and gene expression by moving the histone H1 whereas it can help to condense the chromatin and inhibit the transcription using histone H1. At promoter and enhancer level, many transcription factors and HMGA bind to DNA resulting either facilitating transcription factors binding to DNA or HMGA interact with DNA and these factors to form a complex. HMGA interaction with transcription factor can be described as high or low affinity interaction to explain the interaction with enhancers or promoters in favour of DNA binding (Sgarra et. al., 2004).

High mobility group box family (HMGB) is non-histone chromosomal proteins in mammals. HMGB is highly conserved and contains two L-shape DNA binding domains named HMG boxes A and B, which highly binding to DNA independently, and 30 amino acids of the C-terminal tail (Ueda and Yoshida, 2010). The L-shape binds to minor groove of DNA non-specifically (Reeves, 2010). It is the most abundant proteins in mammals and composed of: HMGB1, HMGB2, HMGB3, and HMGB4. The period of binding HMGB with chromatin is very short (Figure 1.8) (Ueda and Yoshida, 2010).

Beside HMGB binding to DNA in high affinity, it involves in many nuclear process such as transcription and replication. HMGB interact with DNA at entry/exit region of nucleosome similar to histone H1. However, HMGB binding help to assemble the chromatin remodeling proteins such as ACF/CHARC to derive the nucleosome sliding. Moreover, HMGB has role in as transcriptional repressor. It can bind to TATA binding protein to form complexes, which can prevent forming other complexes at preinitiation stage of transcription mechanism. Furthermore, HMGB tight binding to chromatin has roles in apoptotic cells, which can inhibit the protein from releasing from dying cells and affect the immune response (Reeves, 2010).

1.6 HMGN

HMGN is a subgroup of the high mobility group proteins, but it targets the nucleosomal binding domain (Figure 1.8). HMGNs family contains five

members: HMGN1, HMGN2, HMGN3a, HMGN3b, HMGN4, and HMGN5. All HMGNs share five features: nuclear localization signal, conservative, nucleosomal binding domain positive change, same nucleosomal core particle sequence (RRSARLSA), and C-terminal regulatory domain region (Postnikov and Bustin, 2015) (Figure 1.9).

They are ubiquitously expressed in all vertebrate. At chromatin level, DNAase hypersensitivity sites were the preferable binding location for HMGNs proteins. Therefore, HMGNs were found to change the global and the local chromatin structure. Besides chromatin remodeling, histone modification, transcription mechanism and cell cycle are affected by HMGNs proteins. Histone H1 domain associated with DNA close to nucleosome dyad axis and beside linker DNA. This global domain of H1 helped to stabilize the wrapping DNA around histone variants. In addition, histone H1 General structure of Nucleosome core particles can be identified by all HMGNs proteins. Therefore, it was found that HMGNs bind to nucleosome, which interact with two the same HMGNs molecules (Figure 1.8). Moreover, HMGN was found to be located near linker DNA. This suggest that there may be overlapping between HMGNs and H1 domain at linker DNA which may affect the chromatin modification. One of histone H1 role is to help and promote the chromatin condensation whereas HMGNs may help to access chromatin throughout the binding site of histone H1 to nucleosome and connecting HMGNs with acidic patch in H2A-H2B dimer besides the binding sites of the N-terminal of H3 (Postnikov and Bustin, 2015).

1.6.1 The structure and the functional domain of HMGN proteins

HMGNs in vertebrates bind to nucleosomes core particles (CP) and histones. Nucleosomal binding domain (NBD) sequence in all HMGN contains RRSARLSA residues (Figure 1.9), which are the core-binding domain that is the building block of chromatin fibre. HMGN proteins except HMGN5 consist of four parts: two nuclear localization signals (NLSs) at the N-terminal domain, regulatory domain (RD) at the C-terminal, and nucleosomal binding domain (NBD). Regulatory domain (RD) can facilitate chromatin decompaction and it

can influence the HMGN activities on histone posttranslational modification (Kugler et. al., 2012). On the other hand, HMGN5 contains long C-terminal domain about 300 amino acids in mouse and 200 amino acids in human. This domain interacts with linker histone H1 in positive charge (Bustin et al, 1995) (Figure 1.13).

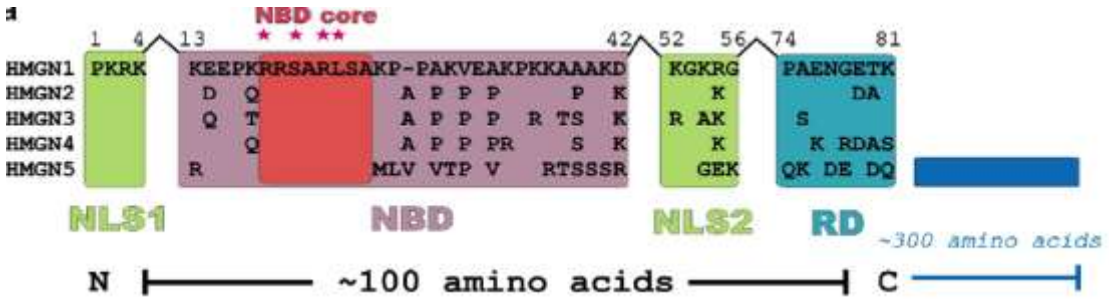


Figure 1.13 the protein sequences of HMGN1, HMGN2, HMGN3, HMGN4 and HMGN5
All HMGN proteins contain around 100 amino acids except HMGN5 contains around 300 amino acids. HMGN proteins contain nuclear localization signal (NLS1 and NLS2), nucleosomal binding proteins (NBD) and regulatory domain (RD). The N-terminal region has nuclear localization signal (NLS1, bright green) very conservative throughout the all HMGN proteins and it contains 4 amino acids. Nucleosome binding proteins (NBD, light purple) contains very conserved sequence called nucleosomal core binding domain which contains eight amino acids, RRSARLSA (bright red), and the red star above some of them indicates the specific binding sites to nucleosome. The rest of NBD are variable among HMGN proteins. The C-terminal regions contain nuclear localization signal (NLS2, bight green) and regulatory domain (RD, cyan) and both are different in protein sequences. HMGN5 is the only HMGN proteins member that contains extra sequence around 300 amino acids at the C-terminal region. However, all the mentioned domains are functional in HMGN proteins. Adapted from (Postnikov and Bustin, 2010).

Several studies identified HMGNs amino acids residues that interact with chromatin and regulates transcriptional process. To begin with, the nucleosome core binding of the HMGN2 plays an autonomous functional domain. Since nucleosomal core binding domain was found between residues 15 to 40 amino acids in HMGN1 and HMGN2, Crippa and colleagues (1992) used gel retardation assays to identify the Hmgn1 and Hmgn2

interactions with core particles comparing with recombinant HMGN2 from bacteria. The Nucleosome core binding sites of HMGN2 (HMG-17) plays an autonomous functional domain. Mobility shift application was shown that single nucleosomal core particle contains two binding sites either HMGN1 or HMGN2. These proteins bind near the gate points of nucleosomal DNA of the core. They found the location between 17 and 47 amino acids residues in HMGN1 and HMGN2 encoded by two exons and it is very conserved during the evolution (Crippa et. al., 1992)

In addition, the regions between 17 to 47 amino acids residues were specific for binding HMGN2 core particles to DNA. Study on this region suggested that conformational change occurs in the core particles. Moreover, this region maintains most of HMGN2 properties. On the other hand, HMGN1 gel shift assays and DNase I digestion found an increase in radius of gyration and temperature of nucleosomal core particles. Removing the histones tails of HMGN1/2 using proteolytic digestion caused an elimination and separation between proteins and peptides. This indicates that there is interaction between core particles domain and histone tails. Therefore, there are a similarity between The structure of HMGN2/1 and many transcriptional activators, which many of them have modular organization, may facilitate multiple cooperation, interactions, increasing the specificity and flexibility of multicomponent structure such as functional transcription complex (Crippa et. al., 1992).

Another study highlighted the functional domain in HMGN1/2 proteins and examined the binding condition between proteins and core particles domain. Postnikov and colleagues (1994) used gel shift assays to study the interaction between mutant HMGN1/2 from human and nucleosomal core particles from chicken erythrocytes. They found that non-cooperative binding occurred at low ionic binding between HMGN1/2 proteins and nucleosome core particles whereas cooperative binding occurred at high ionic binding between HMGN1/2 and nucleosome core particles. Low ionic strength indicates that the binding of HMGN1 or HMGN2 with nucleosome core particles generates two extra lower bands reflect the concept of one or two HMGN molecules bind to one nucleosome core particles. In contrast,

only two molecules of HMGN binds to one nucleosome core particle. Under mutation condition, gel shift assay for non-cooperative binding showed no effect in the binding to nucleosome core particles. However, cooperative binding showed some changes in single amino acid residues by replacing the alanine with proline at position 21 and change lysine with cysteine at position 26 in the nucleosomal binding domain of HMGN1. These changes caused a decrease the binding affinity between protein and core particle domain (Postnikov et. al., 1994). The deletion Region from 17-47 bp corresponded to the entire nucleosomal binding for HMGN2 and it incorporated into chromatin but failed to enhance the transcription (Trieschmann et. al., 1995). Then, the use of Cross-linking experiments revealed that HMGN2 was found to be near histone H2A located between histones octamer and DNA (Crippa et. al., 1992 and Trieschmann et. al., 1995).

The amount of HMGNs proteins binding to nucleosome in chromatin was identified by gel shift assay in pervious studies. The organization behavior of HMGNs such as HMGN1 and HMGN2 in cellular chromatin found to be very similar and to act as homodimers binding to nucleosomes that contains two molecules of either HMGN1 or HMGN2 proteins in vitro and in vivo. However, this binding found to have no influence in histone modification in vivo and in vitro. In fact, HMGN1 and HMGN2 organization throughout the chromatin fiber not fully understood. Possibly, HMGN1 and HMGN2 distribute in chromatin in two ways: one of them is random distribution and the other is clustered distribution, which can be either mixture of heterogeneous HMGN1 and HMGN2 homodimers or homogeneous separation of HMGN1 and HMGN2 homodimers. This organization was studied on mouse thymus or chicken erythrocytes using micrococcal nuclease digestion. This study found that HMGN2 is distributed into specific regions that have nucleosome assembly lacking HMGN1, enrich in core histones, and hold six attached HMGN1 to nucleosome. this results could be agreed with the connection of hmgn2 to nucleosome that could control the 30 nm of chromatin fiber (Postnikov et. al., 1997).

Although many studies attempted to determine the protein nucleosomes complexes using X-ray crystallography such as proteins called LANA and RCC1, The structural basis of HMGNs-nucleosome complexes is unknown. Therefore, Kato and colleague (2011) used nuclear magnetic resonance spectroscopy (methyl-TROSY) for macromolecules to manage the histone methyl groups in Ile, Leu and Val (ILV) residues in the nucleosome and to determine the HMGN2 binding sites. They found that Leu6482 in H2A and Val45γ2 and Leu103δ2 in H2B showed a large chemical shift changes (Figure 1.10). Moreover, the N-terminal region 19-30 residues of the NBD was found to influence the same methyl groups which are very close to negative charged residues in an acidic patch of the H2A and H2B dimer surface. Another experiment was performed to measure the distance between HMGN2-nucleosome complexes called paramagnetic relaxation enhancement (PRE). This experiment showed, that N-terminal region at amino acids residues proline (PRO28 and PRO44) in HMGN2 affect the methyl group in the C-terminal part of H2A. On the other hand, the N-terminal part of H3 indicates that C-terminal region of the binding domain is close to where DNA entry and exit the region of the lysine residues in the nucleosome. Therefore, either HMGN2 fully bound or only from NBD sites to nucleosome binding showed more opposition to thermal denaturation unlike the HMGN2 in free state. In addition, the NBD location at entry/exit region may act as antagonist to ATP-dependent chromatin remodelling factors by tightening the DNA movements in the nucleosome or preventing the remodelling factors from binding to nucleosomes. These results agreed with other results that showed HMGNs reduce the chromatin compaction induced by linker histone H1 in vitro, but at the same time H1 competes for binding with chromatin in vivo (Kato et. al., 2011). Furthermore, the acidic region (C-terminal region) of the HMGN1 proteins can reduce the down regulation of transcription of the histone H1 and chromatin condensation (Ding et. al., 1997) (Figure 1.14).

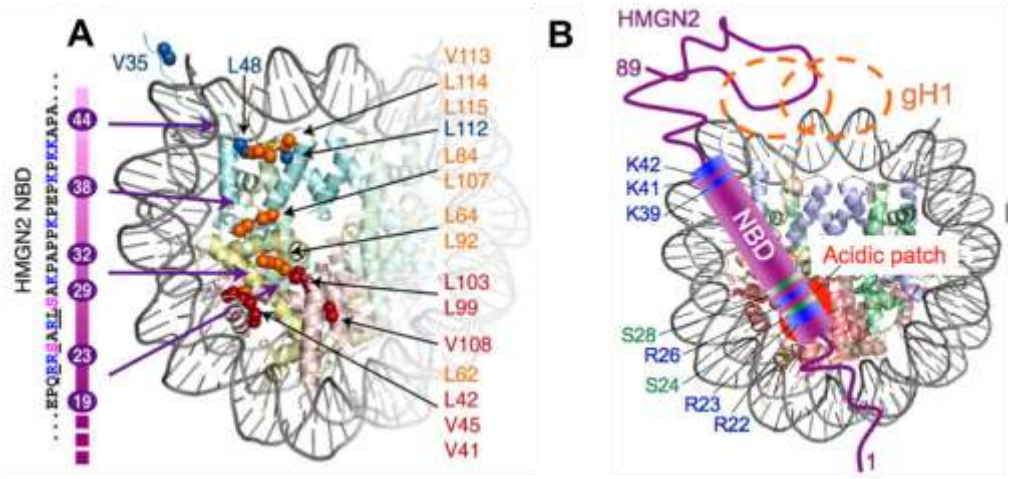


Figure 1.14 the interaction map of nucleosome-binding proteins (NBD) of HMGN2
A. Histone modifications, H2A, H2B, H3, and H4, from mouse mapped for nucleosome binding interaction regions. HMGN2 NBD region has around 27 amino acids and some of them specifically bind to nucleosome. Some of these amino acids residues are in blue colors represent Arg/Lys residues whereas the amino acids in pink colors represent Ser24 and Ser28. In nucleosome, the core Histones H2A (orang), H3B (red), H3(blue) and H4(green). These histones showed methylated group in nucleosome where HMGN2 NBD interact with them. B. HMGN2 interactions show NBD (cylinder shape) region interact with histones H2A (light yellow), H4 (light green) , H3 (light blue), H2B (light pink) and the red oval indicates the acidic patch region. Different colors of HMGN2 represent different amino acids residues. The long tail represents C-terminal region of HMGN2 and around this region there are globular domain of H1 (gH1) (dashed orange oval). This region is found to be around DNA entry/exit point. Adapted from (Kato et. al., 2011)/

1.6.2 The role of HMGN proteins in vertebrate

Several evidence showed that the role of HMGNs affect the architecture of chromatin fibre in several aspects. First, HMGNs proteins can affect the chromatin dynamic by changing the chromatin remodelling factors. Second, HMGNs proteins can induce the chromatin decondensation by modulating the interaction with histone H1. Third, HMGNs proteins can affect the chromatin fibre by influencing several histone Post-translational modifications (Reeves, 2015). As consequence of chromatin remodelling influencing, HMGNs proteins seem to act as transcriptional coactivators and/or transcriptional regulators, DNA replication, DNA repair and possibly nucleosomal spacing in vitro and in vivo.

1.6.3 HMGN proteins and chromatin architecture

To begin with, it was found that HMGN presents in nuclear of all vertebrate cells. It involve in chromatin architecture and dynamics. Several studies applied minichromsome from several organisms such as Xenopus egg extract, Hela nuclear extract, and Drosophila oocyte extract. All these studies found that HMGN proteins enhance the rate of transcription using either Pol II or Pol III promoters and replication using SV40 origin of replication (Bustin, 2001).

For example, a combined HMGN1 and HMGN2 into native nucleosome caused functional increased in the transcription of active chromatin and chromatin remodeling in Xenopus laevis egg extract. The study used a minichromosome contains an extract of 5S rRNA genes from Xenopus and these two elements were combined in the egg with HMGN1 and with no HMGN1. They found that active RNA polymerase II and likely increase the HMGN1 and HMGN2 caused an accessibility of active 5S rRNA gene, which was previously found to be affected by addition of histones and TFIIIA (TBP factor). Consequently, the chromatin assembly mixture formed after an increase in 5S rRNA activity. Moreover, the presences of HMGN1/2 increase the open-chromatin regions; thus more proteins facilitate the RNA polymerase accessibility (Trieschmann et. al., 1995).

However, combining the HMGN1 and 2 into native nucleosome increase the function of transcription for active chromatin and chromatin remodeling. It was found that 30% of the HMGN1/2 sequence is conserved. Both proteins were found to have similar cellular function and both have to incorporate into chromatin using transcriptional enhancement assay and micrococcal nuclease digestion was used for this aspect. Moreover, the influence of nuclease digests of chromatin brought by HMG binding to preassembled chromatin opposite to the effect of HMG incorporation into nascent chromatin. When both hmgn1/2 incorporate into nascent chromatin, the rate of digestion by nuclease enzyme increased using micrococcal nuclease digest to identify kinetics of proteins digestion instead of nucleosomal spacing. On the other hands, when inhibition of transcriptional activity occur using nucleosomal binding domain of HMGN1/2 (Region from 17-47bp), the nucleosomal binding domain alone did not enhance the transcription potential of chromatin (Trieschmann et al., 1995).

This cooperation binding of HMGN1/2 into nucleosome cores were found to produce homodimer complexes using antibodies against purified proteins and analyzed by ELISA, western blots and radioimmunoassay. However, under non-cooperative condition the full length of HMGN1/2 are required for keeping the homodimeric mode of binding to core particles using low ionic strength solution (Postnikov et. al., 1995).

Another example used minichromsome of Xenpus eggs and micrococcal nuclease showed that HMGN2 involved in chromatin assembly and nucleosome spacing. To elaborate, the period of removing HMGN2 from chromatin fiber during DNA replication step to perform the chromatin assembly is critical to obtain active chromatin. On the contrary the addition of HMGN2 after chromatin assembly had no effect on transcription process. Furthermore, chromatin during the replication required deposition of H3-H4 tetramer and assembly of two H2A-H2B dimers to form nucleosome spacing. Like Xenopus, mammalians when the chromatin assembled in the absence of Mg2+-ATP, they digest faster using micrococcal nuclease than in the presence of Mg2+-ATP. Nevertheless, the presence of HMGN2 has major effects on the structure of chromatin assembled when MG2+-ATP is absence.

In this case, it showed well-defined nucleosomal spacing and the nucleosome ladder that was obtained from chromatin assembled with HMGN2 displayed extreme separation and diminished backgrounds between bands. Nucleosomal spacing occurs due to the stabilization of core particles by HMGN2 (Crippa et. al., 1993). However, micrococcal nuclease assays showed that hmgn2 proteins helped to define the boundaries of nucleosomal core particles beside improve the nucleosomal spacing, but not necessarily the HMGN2 acts as nucleosomal spacing factor (Crippa et. al., 1993). The association between HMGN2 and the dynamic of chromatin was confirmed in tissue culture cells (Hock et. al., 1998).

From Protein Data Bank, it revealed that HMGN proteins could have high developmental disorder and diseases relation such as cancer. However, HMGNs are different from each other based on amino acid composition. Therefore, HMGNs proteins have roles in the regulation of cellular transcriptional profile. The expression profile in cell specific manner was tested on mouse embryonic fibroblast (MEFs) using retrovirus to knockout or overexpressed HMGNs. The structure of HMGN1, HMGN2 and HMGN3a are similar in size unlike HMGN5 whereas HMGN3b contains a lacks of 21 C-terminal residues. The knockout was performed on HMGN1, HMGN3 and HMGN5 unlike HMGN2 knockout that was embryonic lethal in MEF cells. Overexpression experiment was performed on HMGN1/2/3a and HMGN5 on MEF cells. The knockout results showed that no overlap was identified in knockout, but changes in gene expression involve up and down regulation in transcriptional level, which influence the chromatin structure and regulate and correct the transcriptional process (Rochman et. al., 2011)

1.6.4 HMGN proteins and linker histone H1

The association between linker histone H1 and HMGN family affects the chromatin dynamic and nucleosome. It was known that HMGNs beside histone H1 cooperate in nucleosome and either positively or negatively stimulate the chromatin modification activities using enzymes. Histone H1 secured the chromatin structure from decompaction, which prevent the

activity of HATs enzyme and nucleosome remodeling complexes. However, HMGNs proteins enhance the chromatin decompaction, more nucleosomal accessibility and in some cases promote the transcription and replication (Catez et. al., 2003). Previously, it was shown that HMGN and histone H1 overlap between each other close to dyad axis of the nucleosome. This location between HMGN and H1 may be important for their interaction and chromatin remodeling. Possibly the C-terminal domain has more negative charge amino acids located near to linker DNA and it may interact with H1 at the mentioned site. Moreover, due to the role of H1 as promoter for chromatin compaction opposed to the role of HMGN as promoter for chromatin decompaction, this may inhibit the binding of H1 to nucleosome. However, HMGN role as a promoter toward chromatin decompaction can interfere with acidic patch of H2A-H2B dimer and with N-terminal domain of H3 to facilitate the internucleosomal interaction activities (Postnikov and Bustin, 2015). For example, in vitro study investigated the chromatin assembly that showed histone H1 was able to induce toward chromatin remodeling using ATP dependent SWI/SNF. On the other hand, HMGNs enhance open-chromatin and it is possible to promote the chromatin remodeling (Postnikov and Bustin, 2015).

1.6.5 HMGN proteins and transcriptional activation

HMGNs proteins affect transcriptional regulations, DNA repair and embryonic development due to the chromatin modification. For example, Paranjape and colleagues (1995) found that HMGN2 (HMG17) has a role as “transcriptional coactivator” because they found that either in Hela cells or drosophila embryo when HMGN2 was added to chromatin, it increases the GAL4-VP16, which contains amino-terminal 147 amino acid residues of yeast GAL4 proteins fused to a fragment of the herpesvirus VP16 protein that has potent transcriptional activation region, activated transcription. On the contrary when GAL-VP16 is absence, the HMGN2 has no stimulation for general transcription factors by RNA polymerase II unlike Trieschmann and colleagues (1995) results. Moreover, in Xenopus system, it was shown that

HMGN2 did not increase the transcription initiation by RNA polymerase III. However, same system revealed that using RNA polymerase II could increase the efficiency of transcription initiation in the presence of HMGN2 (Paranjape et. al., 1995).

1.6.6 HMGN proteins and DNA repair

HMGNs proteins control the DNA repair efficiency throughout cellular repair mechanisms. Two Studies investigated nucleotide excision repair (UV-induced DNA damage) (NER) knockout, heterogeneous knockout and wild type of HMGN1 in mice and mouse embryonic fibroblast stem cells (MEF). The cells lacking of HMGN1 became more sensitive to UV light four times than heterogeneous knockout mice (Reeves, 2015). Beside nucleotide excision repair (NER), double-strand break (DSB) raised by many DNA cuts. This mechanism is recognized by all cells to ensure the genomic stability. Hmgn1 knockout in mice and MEF cells showed high sensitivity to ionizing radiation (Reeves, 2015). Another experiment performed by Briger and colleagues (2003) showed the influence of HMGN1 in TCR and the ability of HMGN1 in helping cells to survive after these cells experienced IR. Following this damage the cells DNA repaired itself by the association of NBD of HMGN1 to nucleosome and facilitate the chromatin accessibility from secondary and tertiary structure. Moreover, in MEF experiment due to the hmgn1 knockout encountered to IR, the damaged cells did not managed to terminate the cell cycle check point G2/M. therefore, it is possible that hmgn1 has function in the mechanism of ATM (Birger et. al., 2003). The third DNA repair mechanism is base excision repair (BER) pathway, which remove the base leaving an apurinic/apyrimidinic (AP) site using enzyme lesion-specific DNA glycosylases. Knockout hmgn1 in MEF cells showed the regulation of HMGN1 in nuclear enzyme poly (ADP-ribose) polymerase 1(PARP-1) which was found high in BER reaction, AP sites and DNA strand breaks. However, the knockout hmgn1 with untreated PARP-1 was found lower than wild-type hmgn1 in MEF cells. This possible interaction between

HMGN1 and PARP-1 is regulated by protein-protein interaction using immunofluorescence-imaging experiments in vivo (Reeves, 2015).

1.6.7 The role of HMGN1 and HMGN2 in mouse embryonic stem cells (ESCs) and Epiblast stem cells (EpiSCs)

All HMGN proteins show high expressions during the embryonic stage, then gradual down-regulation. The down-regulation may cause changes in chromatin organization and in the cellular transcription. HMGN expression showed major alterations during the developmental program (Hock et. al., 2006). In mouse early embryonic development gene expression and microarray profiling was used to examine the number of genes that were expressed in oocytes, two cells, eight cells and blastocyst stages. Comparison performed between oocytes and blastocysts identified major differences in some genes “oocyte specific genes” including HMGN2. Gene transcriptional profiles suggested that these genes may have possible implication in early bovine development and embryonic genome activation. Genes were expressed in the first week of embryonic development maternal genes, shared regulation of transcription and cell cycle. This regulation showed rapid degradation following fertilization such as NLRP9, FIGLA and ZP2/3/4 and gradually degrading transcription such as CCNB1, HMGN2, and BTG4. NLRP9 and HMGN2 showed similar expression profile in the microarray data during bovine early embryonic development in vitro (Vallee et al., 2009). Previously it was shown that knockdown of expression of HMGN1 and/or HMGN2 in mouse was found to cause delay in cell cleavage during the preimplantation development (Mohamed et al., 2001 and Hock et. al., 2006). On the other hand in Xenopus embryo overexpression of HMGN1 and/or HMGN2 resulted in malfunction in closed blastopore and head structure, distorted body axis and loss of mesodermal competence of animal cap (Korner et. al., 2003 and Hock et. al., 2006).

Recently, Deng and colleagues (2015) examine the role of HMGN1 and HMGN2 in mouse embryonic fibroblast (MEFs). Although HMGN1 and HMGN2 are widely expressed in vertebrate cells particularly in embryonic stem

cells, it was argued that they are functional redundant due to ubiquitous expression and competition in NBD region. Therefore, an investigation was made using HMGN1 and HMGN2 and double HMGN1/2 knockout mice from HMGN1 and HMGN2 knockout mice. MEFs cells were collected from three knockout mice to examine the expression and chromatin profiles with lack of HMGN1 and/or HMGN2. They found that DNase I hypersensitive sites (DHSs) were reduced due to the loss of HMGN proteins, which affect the chromatin remodeling sites. The majority of DHSs are located in promoter and enhancers, moreover in the majority of HMGN1/2 knockout MEFs cells, DHSs were found in enhancer regions. Furthermore, DNase I digestion in knockout HMGN1/2 was found to be more accessible to regions that have reduction of DNase I hypersensitive sites; however at the same level new DHSs sites were found in knockout cells that were not found in wild type cells. Unlike new DHSs sites in knockout cells, most of the regulatory sites that were formed became less accessible to DNase enzyme. Furthermore, histones modification, H3K4me1, H3K4me3 and H3K27ac did not change in knockout HMGN1/2. This suggested that HMGN-mediated chromatin could act as a mediator for other regulatory factors to interact with chromatin such as PCNA and ATP-dependent nucleosome remodeling complexes. In addition, the loss of HMGN1/2 showed enhanced in H1 binding leading to reduce the DNA accessibility at chromatin regions using DNase I digestion. Therefore, HMGN1/2 may act as modulator for other regulatory factors and chromatin modifications; thus they play important roles in maintaining the DHS profiles (Deng et. al., 2015).

1.7 Pluripotent stem cells of mouse

Embryonic stem cells are pluripotent cells that have the ability for self-renewal and differentiation to any type of cells. At the early stage of blastocyst, mammalian embryo developed to two different types of cell populations: trophectoderm (TE) and inner cell mass (ICM). Trophectoderm cells develop to trophoblast giant cells and placenta from trophoblast stem cells. This is controlled under the activation of transcription factor Tead4,

as well as down regulation of pluripotent genes Oct4, Sox2 and Nanog in the cleavage-stage embryo.

Inner cell mass (ICM) cells form at embryo day E3-E3.75. These inner cell mass (ICM) activate Oct4 as main transcription factor, followed by other transcription factors Sox2, Nanog, Esrrb, Tbx3, and FGF4. These cells give raise to primitive endoderm (PrEn) and epiblast naïve stem cells. Primitive endoderm (PrEn) cells are extra embryonic lineage and form the yolk sac of the embryo. Transcription factors Gata6/4, Sox17/7 and Pdgfra become up regulated in these cells. Epiblast Naïve stem cells derived from ICM of blastocysts express pluripotency associated factors such as Oct4, Sox2, Nanog, Tbx3, Klf2/4/5 and Esrrb (Nichols and Smith, 2012; Patra et al., 2011 and Plusa and Hadjantonakis 2014). Naïve cells also are also controlled by the levels of negative regulators such as transcriptional repressors, transcription factors, the regulators of mRNA stability and translation and regulators of nuclear transport (Kalkan and Smith, 2014). After that epiblast naïve stem cells undergo to implantation process. Then the primitive ectoderm epiblast form in two stages: formative (early post implantation at E5.5) and primed (late post implantation E6.5). At the early post implantation a subset of transcription factors shows over expression such as FGF5, Oct6, Otx2 (Kalkan and Smith 2014) Pitx2, Sox17, Nodal, Foxa2 (Boroviak et al., 2014 and Tesar et al 2007) These cells are the signs of the development of epiblast cells until it extends to gastrulation (Ohtsuka and Dalton, 2008 and Prelle et al, 2002). Pluripotent cells developed into three germ layers: ectoderm, endoderm and mesoderm (Ohtsuka and Dalton, 2008). Several studies identified that pluripotent stem cells are found in two distinct states: naïve ICM exists in mouse and human stem cells and primed post-implantation epiblast was found in epiblast stem cells in human (Nichols and Smith 2009 Welling and Geijsen, 2013) (Figure 4.1) .

Martin and Evans (1975) is the first study, which succeeded to isolate pluripotent cells from teratocarcinoma stem cells. These cells were polyclonal derivatives from undifferentiated cells and primordial germ cells similar to embryonic stem cell at the early stage. For this reason they were called embryonic carcinoma (EC) cells (Wobus et al., 1984, Prelle et al,

2002 and Welling and Geijsen, 2013). However, other researchers thought that these cells are neither ICM/Epiblasts cells nor primodial germ cells due to spontaneous re-implantation and phenotypic similarity with embryonic stem cells (Peter Andrew 2002). Many studies have used EC cells to establish experimental foundation in culturing embryonic stem cells. The reason for this was that EC cells if they are able to transplant into ectopic sites, often influence blastocysts developing normal tissues in chimeric mice (Wobus et al., 1984, Prelle et al, 2002 and Welling and Geijsen, 2013). Moreover, these cells, that are similar to ESCs in morphology, gene expression profiles and functional ability when injected to form chimera, can generate three germ layers in vitro and in vivo.

1.7.1 The formation of mammalian embryo and gene expression profile

Similar to EC cells are murine ES cell lines derived from inner cell mass (ICM) of blastocysts, 8-cell embryos, or from pluripotent cells that were generated from ectoderm of implantation-delayed blastocysts (Prelle et al, 2002). To begin with, germ cells derived from pluripotent primed epiblast cells called primodial germ cells (PGCs). These cells develop oocytes in females and spermatozoa in male. Zygotes (totipotent) form after fertilization forming renewable and consistence mammalian life cycle. These cells are the early stage of Inner cell Mass (ICM) cells.

Murine ESCs encode genes that are important in the development of germ cells. These germ cells can be only from post-implantation cell type, which continually express the pluripotent genes, Oct4, Sox2 and Nanog (Welling and Geijsen 2013) (Figure 1.15).

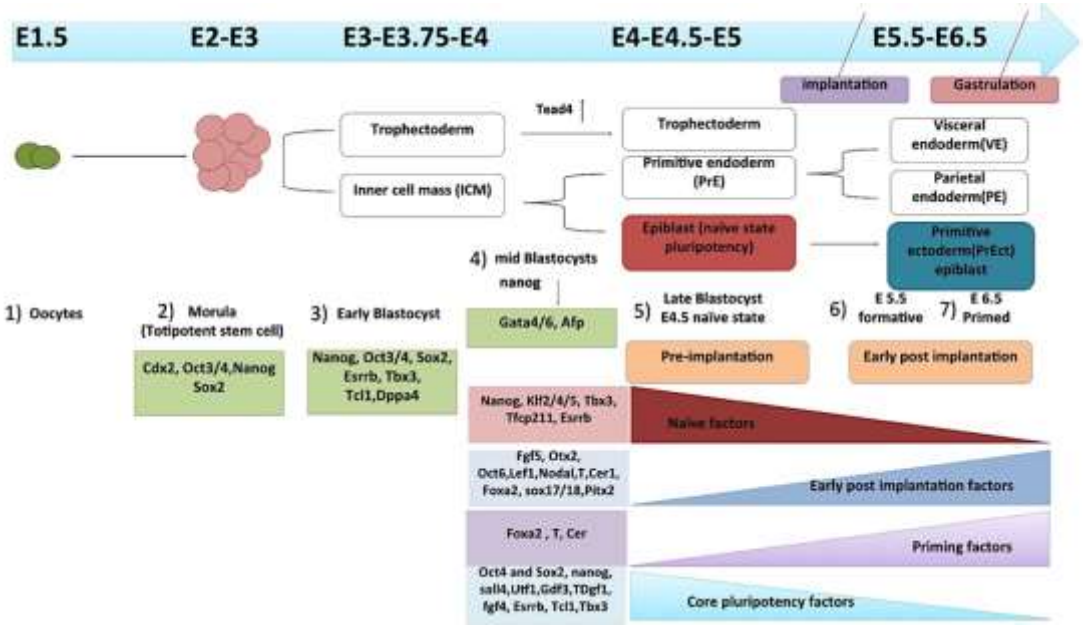


Figure 1.15 Summary of Pluripotent lineages profile of mouse embryo.
Schematic representation in the peri-implantation mouse development and gene expression (1) Oocytes: around the embryonic day 1.5 (E1.5), oocytes has two cells undergo cleavage process that generate embryo. (2) Morula; around the embryonic day 2.5 (E2.5), the first pluripotent cells in the embryo shows 8 cells of pre-inner cell mass (ICM) which express some genes. (3) Early blastocyst; around the embryonic day E3, trophectoderm and ICM cells were generated from Morula. Subsets of different genes, which are highly expressed in this stage (green box), indicate the mouse ES cell developmental identity. (4) Mid blastocyst; around the embryo day 3.75, ICM still at the pre implantation stage before epiblasts development; however, some genes were activated in this stage such as GATA4/6 whereas Nanog gene showed down regulation. (5) Late blastocyst stage appears around embryonic day E4-4.5. Trophectoderm cells develop placenta under the activation of Tead4 gene and ICM cells give rise to primitive endoderm (PrE) and Epiblast (Naïve state-pluripotency). These cells develop before the implantation process and activate some of Naïve genes such as Klf2/4/5, nanog, Tbx3, Esrrb, Gata6/4,Sox7/17, and Tfcp112. (6) early post implantation, which is around embryonic day E5.5. This stage called formative stage that lineage of specific cells emerge. Moreover, list of genes start to become activated such as fgf5, otx2, sox17 and oct6. (7) Primed stage or late post implantation at the embryonic stage E6.5 before specialized cells lineages emerge (gastrulation) by activating new subset of genes such as Foxa2 and Cer1. Core pluripotency factors become activated from E4; then they become downregulated thorough cell development to primordial germ cells and somatic cells. Adapted from (Niwa Hitoshi, 2007; Patra et al., 2010; Patra et. al., 2011; Ohtsuka and Dlton 2007; Welling and Gijsen, 2013; Plusa and Hadjantonakis, 2014; Boroviak et al., 2014; Kalkan and Smith, 2014; Nichols and Smith, 2012).

1.7.2 Embryonal carcinoma cells

Teratomas are tumours, which occur in germ cells in the ovary or testis. Generally these tumours are benign when found in the ovary (benign ovarian cysts) whereas similar tumours found in the testis are mostly malignant (Teratocarcinoma). These tumours derived from germ cell tumours (GCTs) mostly cause testicular cancers (Andrews Peters, 2002). P19 is an embryonic carcinoma stem cell line derived from murine embryonic teratocarcinoma stem cells in C3H/He mice from the time period of embryo day 7 to the testes of an adult male mouse (Choi et al 2012 and Chen et al, 2014).

The first stable pluripotent stem cells were isolated from teratocarcinoma. Moreover, these cells contain benign tumour cells and embryonic carcinoma cells. These cells contain random of differentiated cells, which found during the early stages of embryo development. These cells have the ability to re-transplant to another host maintaining an undifferentiated state (Peter Andrews, 2002).

The P19 EC stem cell line has the ability to differentiate into all three germ layers and the molecular pathways that may explain the self-renewal pathway and differentiation have been established. Moreover, the P19 cells are an appropriate model for EC and ES because the cells can grow in monolayers and without feeder layers. Moreover, they are easy to study expression of ectopic genes that are important for self-renewal and maintaining an undifferentiated state. (Choi et. al., 2012 and Chen et. al, 2014).

The characterization of embryonic carcinoma (EC) cells is very similar to normal embryonic stem cells (ESCs). EC cells derived from primary tumours possess euploid male karyotype (40:XY) (Angello et al., 1997 and Heyden and Defize 2002). P19 is defined as embryonic carcinoma stem cells that have normal mice euploid male karyotype (40:XY). These cells were isolated from 7.5 days embryos that stem from C3H/He female mice. Some mice can form Teratocarcinoma from early embryos. P19 EC cell lines are a preferable in-vitro model to investigate the molecular mechanisms of early mammalian development. P19 can differentiate to normal cardiac myocytes

with embryonal bodies under dimethylsulfoxide (DMSO) (Skerjanc Ilona 1999 and Qin et al., 2013), (Angello et al., 1997 and Heyden and Defize 2002), and normal neuronal cell under Retinoic acid inducers. This is because EC cells can differentiate to all three germ layers: endoderm, mesoderm and ectoderm (Bain et. al., 1994).

Another reason for P19 as preferable model is that establish heterozygous for x-linked alleles (Jones-Villeneuve et al, 1982 and Heyden and Defize, 2003). P19 can grow as an undifferentiated primary cancer stem cell and small amount of endoderm with no factors or feeder cells in vitro (Jones-Villeneuve et al, 1982. Recently P19 embryonal carcinoma cells were investigated by Riego and her colleagues (1995) who examined the interferons (IFNs) that generate heterogeneous family of cytokines in P19 EC. Moreover, it was shown that IFNs and IRFs responsible in cell growth and differentiation in embryonic carcinoma (EC) cells. They found that IFN-alpha genes were identified in early embryos and P19 EC cells. However, IFN-alpha expressed in P19 EC cells unlike IFN-beta. IRF-2 expression was found in early stage of P19 cells (Riego et. al., 1995).

1.7.3 Epiblast Stem cells

Epiblast stem cells are pluripotent cells that generate during the cell division within inner cell mass. These cells give rise to three germ layers, endoderm, ectoderm and mesoderm. They express the pluripotent marker sox2, oct4, and nanog at the early epiblast cells stage; however, they do not express Zfp42 and they less express Gbx2 unlike ICM. In the late epiblast stem cells FGF5 and Nodal were expressed after the implantation unlike mouse ES cells (Brons et al., 2007). In mouse embryos, it revealed that hsp 68, IRF1, IFN-alpha are highly expressed in two cells embryos. However, in P19 cells RNA expression levels of IFN-alpha, hsp68, and APP genes highly express from post cleavage stage of the embryos whereas IRF1 expresses later stage of embryo (Yeom et. al., 1996).

EC P19 cells are considered as epiblasts cells of the early post-implantation blastocyst development. The derivation of the cell lines affects the

pluripotent gene expression such as Oct4. Embryonic stem cells, embryonal carcinoma stem cells and embryonic germ cells were used to test the expression level of Oct4 gene. Embryonic stem cell (ESCs) derived from blastocysts, which generate high production of germline chimeras. However, Embryonic carcinoma (EC) P19 cells were generated from early to mid-gastrulation embryos. It has similar properties of the epiblasts, but this type of cells has different developmental stages and only some of epiblasts cells derived from totipotent cells. This was suggested that Oct4 expression has different regulation roles in these cells (Yeom et. al., 1996) (Figure 16.1).

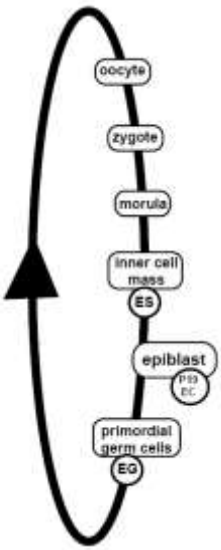


Figure 1.16 mouse embryonic stem cell development based on Oct4 expression
The box represents the stem cells express Oct4 gene at different developmental stages whereas the circle arrows represent indicates different stage of stem cell development. P19 EC cell is not on the black arrow line to show that not all the epiblast stem cells are totipotent stem cells that are precursor of primodal germ cells.

1.7.4 Teratocarcinoma stem cells (P19) and neuronal development

Previously, it was mentioned that P19 cells could differentiate to neuronal stem cells using retinoic acids induction; consequently, neuroectoderm, neurons and glia is developed from neuroepithelial-like germinal cells close to mammalian central nervous system. At functional level, these differentiated cells can express neuronal markers and neurofilament and it can exhibit neuronal mechanisms. Under retinoic acids (RA) exposure twice

or three times, P19 cells can generate only 15% of neuronal cells whereas the rest of differentiated cells are unknown or lack of neuronal properties (Nye et. al., 1994). Therefore, P19 has been used as a model for neuronal differentiation and studying signalling molecules such as Wnt, Notch, BMP, and Sonic Hedgehog (Marikawa et. al., 2009). Furthermore, P19 can be used as model to investigate possible neurodegenerative treatment such as Parkinson (Noh et. al., 2012).

For instance, a study used activation of *Notch1* gene to investigate neurogenesis regulation in P19 cells using overexpression of mNotch proteins in differentiated P19 cells. it was found that mNotch overexpression in differentiated P19 cells caused suppression in neurogenesis and myogenesis. Moreover, Undifferentiated P19 cells can expression embryonic cell surface marker (SSEA1) which is a glycolipid antigen found in preimplantation embryo at the inner cell mass and primitive ectoderm. When P19 cells induced by RA, SSEA1 reduced and filament protein (Nestin) increased after 4 days of RA induction. Then another population called neurons is raised after Nestin reduction. Nestin expression is an indication of neuroepithelial and myotomal cells (Nye et. al., 1994). However, it loses its expression during the neural precursors development. Mature neuron is developed by neurofilament and NeuN expression. During P19 cell developed to glia cells after 12 days of RA induction. The overexpression of *Notch* gene in P19 cells did not affect the development to glia cells and GFAP marker showed increase level in the same cells. This showed that *Notch* gene may not affect the differentiated state, but it may have role in developing neuroepithelial cells to a neuron (Nye et. al., 1994).

Several cell culture protocols have been developed to reduce the contamination of non-neuronal cells. For example, Monzo and colleagues (2012) developed low serum medium method to overcome low neuronal cells rate. Performing the P19 neuronal differentiation initiate by plating the P19 cells as suspension cells and induced them with RA to form embryoid bodies (EBs) then neuronal precursor cells. However, this method did suspend generation of proliferative non-neuronal cells and only 2% was found to be mature neuron after 12 day of culture (Monzo et. al., 2012).

Another method was performed by Nakayama and colleagues (2014) using no serum medium and no formation of suspension culture. This method showed that increase in neurogenesis about 50% with no glia formation within 6 days after the RA induction (Nakayama et. al., 2014).

1.8 Approaches to identify the function of a protein

Targeting the gene using homologous recombination techniques (loss of function) aids understanding of the biological roles of the gene and its protein by disrupting the gene open reading frame in mouse. One of the most common methods is knockout gene in mice. This method involves a drug resistance marker replacement of genetic coding region (Hall et al., 2009). However, some knockout genes can be lethal to the mice although conditional knockout mice were developed to overcome this issue. On the other hand, studying RNA of gene function in vitro can be unclear. Interference RNA interference (RNAi) provides a novel tool for studying loss of gene function using endogenous pathway. The use of knockdown methods in vitro and in vivo can generate long-term specific gene silencing in the target model. Moreover, it can eliminate the expression of genes in brain and hematopoietic stem cells while reducing the issue of a lethal model in vivo or in vitro (Tiscornia et al., 2002).

RNA interference (RNAi) has expanded the understanding of loss of function for some genes in development and disease. It is an innate defense mechanism in mammalian systems that protects the cells against viruses and other activities caused by transposable elements. However, RNAi has been used as an alternative pathway to the gain of function or deletion of genes using homologous recombination (HR). RNAi technologies use a natural gene-silencing phenomenon as an advantage to knockdown the expression or block mutant genes (Fellmann and Lowe, 2014, and Rytlewski and Beronja, 2015).

1.8.1 RNA Interference pathway

RNAi is endogenous mechanisms that can mediate the gene silencing. This can be used as a powerful tool to study the biological role of specific genes. This tool can be used in cell culture using delivery system of chemically synthesized siRNAs or shRNAs. Constitutive gene silencing of shRNA is derivn by RNA polymerase III (U6 or H1) promoters. This system showed a sufficient knockdown of target gene. However, the lack of sufficient the level of knockdown can cause limitation during the analysis of genes that are important for cell survival, cell cycle analysis and development. Moreover, the long period of knockdown of certain genes may cause nonphysiological responses (Gupta et al., 2004).

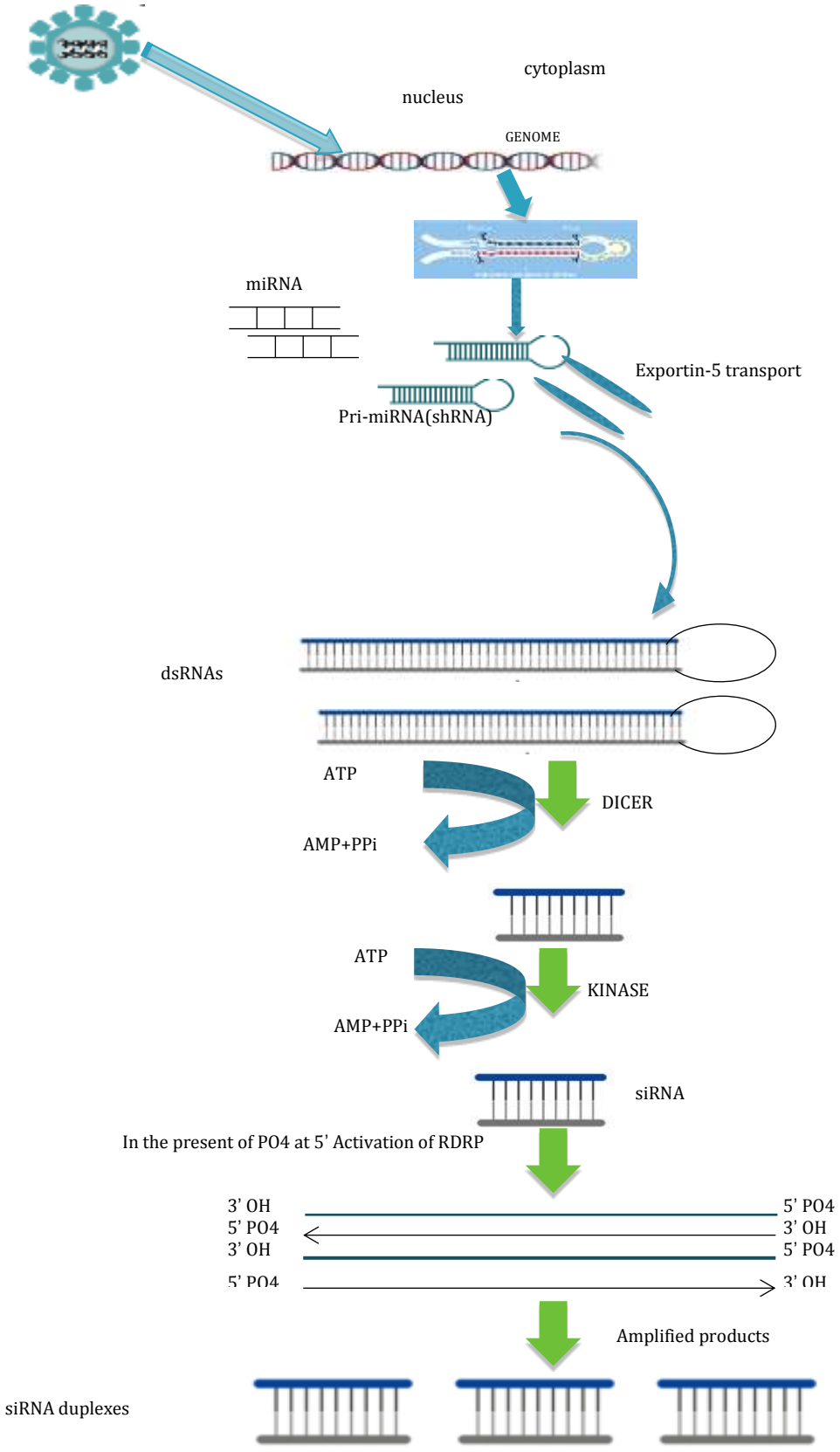
RNAi mechanisms involve nucleases and ATP dependent process in the cytoplasm and can be divided into two steps. The first step is the formation of double stranded siRNA. RNase III enzymes Dicer and Dorsha cleave the 3' end of shRNA that contain miRNA and the 5' end loop (Aagaard and Rossi, 2007). Drosha forms a microprocessor complex, which contains an RNase III family enzyme, DiGe orge syndrome critical region gene 8 (DGCR8). DGCR8 consists of two double stranded RNA-binding domain (dsRBDs). DGCR8 can determine the pri-miRNA and single strand RNA sites. Drosha recognizes the cleavage region of around 11bp leaving around 65-70 nt of precursor miRNA (pre-miRNA) (Wilson and Doudna, 2013). The long sequence double strand RNAs form after pre-miRNA is exported to the cytoplasm. They have a specific siRNA of 21-23 base pair long for particular genes at the 3' end and there are around 2-5 nucleotide that allow to be recognized by Dorsha and Dicer (Aagaard and Rossi, 2007 and Rychahou et al, 2006). This dsRNAs mimics the innate immune response from antivirus through increasing the interferon-linked pathway. One of the important interferon-based defence mechanisms is the cellular dsRNA-dependent protein kinase (PKR) which can cause inhibition of translation (Rana, 2007 and Garcı´a et. al., 2007).

The Second step is the use of the effector RNA induced silencing complexes (RISC). Either synthetic or endogenous siRNA, which are formed as a small double-stranded RNA from the first step, are incorporated into the RISC protein complex (Hammond, 2005 and Rychahou et al, 2006). RISC is a

complex that contains a splicing protein Argonaute that can recognize siRNA. When RISC binds to the siRNA sequence, a destruction of RNA transcription occurs by activating RISC, leading to knockdown of the mRNA of the gene of interest. This process results in translational repression (Figure 1.12) (Aagaard and Rossi, 2007 and Nowotny and Yang, 2009).

The endogenous mechanism of RNAi

A. **Pre-miRNA**



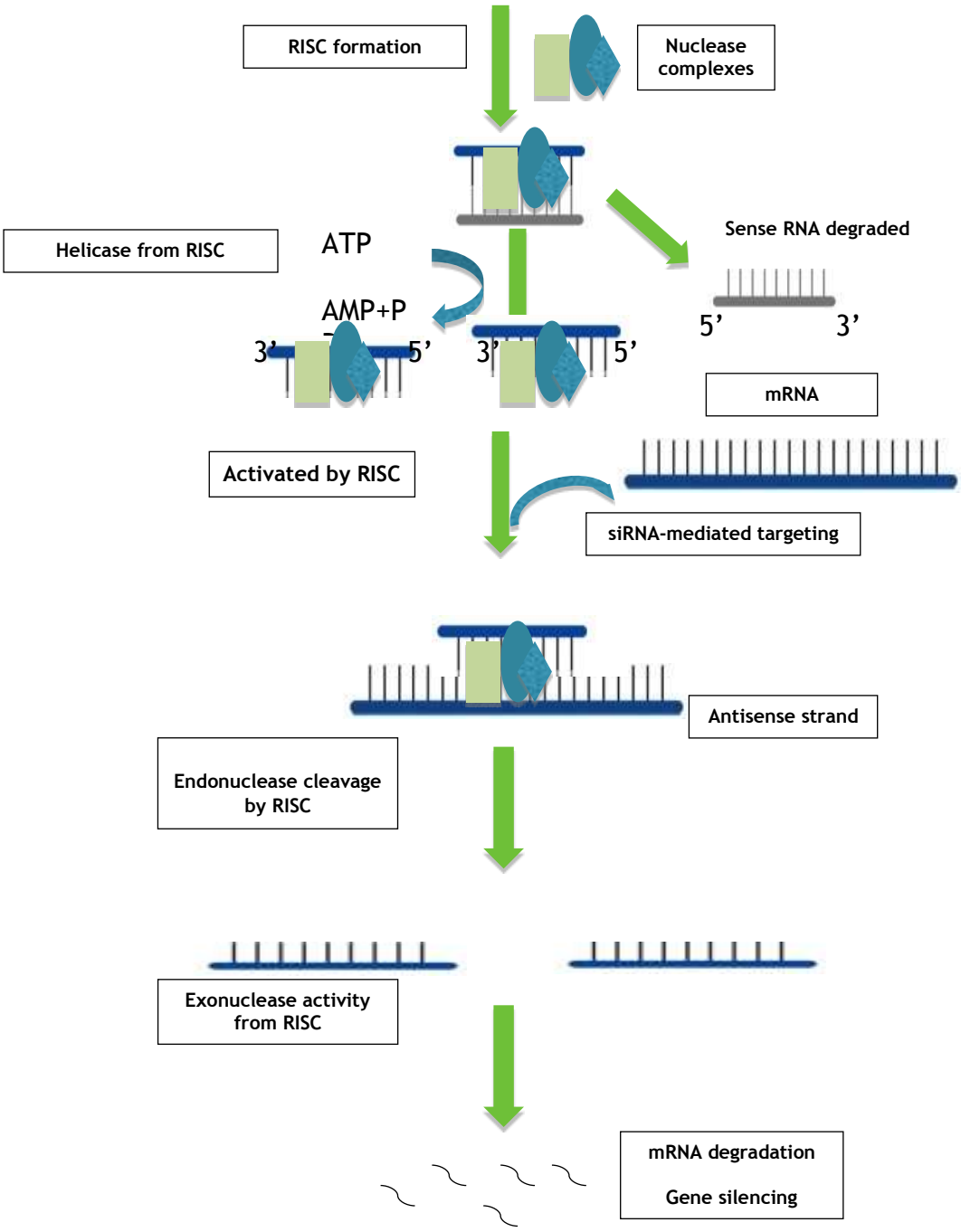


Figure 1.17 RNAi mechanism and the formation of siRNA

A. the formation of double stranded pre-miRNA with stem loop structure; the pre-miRNAs require enzymatic activity and ATP release to generate double helix siRNA. B. RNA-induced silencing complex (RISC) activation process. Unwinding the 5' end of the siRNA duplexes using piRNA guide strand at the PIWI domain in the Arganote-2 region with the help of thermodynamic activity help to generate the siRNA with Arganote-2 recognition site. Arganote-2 helps to identify the complementary sequence at the endogenous mRNA and binds with siRNA. This process causes mRNA degradation.

1.8.2 RNAi in the nucleus: (shRNAs and siRNA)

Two different pathways can trigger the RNAi mechanism. Firstly, 21 base pair duplexes of siRNAs can be directly cleaved by dicer enzyme and the effects are transient. Secondly, nuclear synthesis of shRNAs can be used based on miRNAs synthesized from DNA vectors (Aagaard and Rossi, 2007). Both types of molecules can control gene expression by RNAi process. It was shown that siRNA contains a hairpin structure and it can be expressed in cells using DNA plasmid construct derived by RNA Polymerase III (Pol III) promoter. Although siRNA and shRNA use a similar process, there are some differences in terms of their actions, off target effects and applications (Rao et. al., 2009).

shRNA is synthesized in the nucleus of the cells. Once it formed as pre-miRNA, it integrated in to the genome. ShRNA can transfer to the cytoplasm throughout viral plasmid or deliverer to involve into RISC activity (Figure 1.12). It was mentioned that shRNAs use the endogenous miRNA pathway, in particular miR-30 based shRNAs (Rao et. al., 2009). The advantages of having a DNA based strategy is in the easiness of delivery and the production of shRNAs. The expression vectors can generate shRNAs that have sense, and antisense within the 21 mer. In between them there is a short 5-6 residues terminator sequence. When the transgene expression cassette is transcribed in vivo, the shRNA fold back to form a hairpin structure using cellular nucleases, and the shRNA is cleaved to generate siRNAs (Aagaard and Rossi, 2007).

ShRNAmir can be designed to better targeting the siRNAs process and give high level of knockdown efficiency. The design of shRNAmir uses the RNAi natural enzymatic process by RNase III Drosha. This permits more Dicer recognition and specificity. ShRNAmir can use the RNAi pathway that is processed in the nucleus by Drosha and cytoplasm by Dicer and RISC complex. This causes more siRNAs to be produced and specific mRNA degradation. Unlike siRNA, shRNAmir showed limited rate of enzymatic process in living cells and less off target effects (Fewell and Schmitt, 2006). The idea of adding miR-30 as part of shRNA construct is that miR-30 can

provide efficient knockdown through RNA polymerase II promoter. The main reason of choosing miR30 was due to the well characterization of pre-miRNA transcription and functional process toward miRNA in vitro and in vivo (Cullen 2004 and Stegmeier et al., 2005).

1.8.3 Constitutive versus inducible RNAi systems

Vectors used in this project are bicistronic vectors, which permit high transfection efficiency and enhanced report gene expression for recognizing the transduced cells from non transduced cells such as PGFP and PRFP. These vectors were designed based on HIV-1 which is considered to be the most extensively utilized in gene therapy applications using gene transfer vector systems.

The use of stable long-term expression is essential to study gene behaviour in vitro and in vivo for therapeutic reasons such as neurodegenerative diseases (Eleftheriadou and Mazarakis, 2015). Using constitutive systems can be preferable in some studies that show high transfection rates and it can spread to more mature cells types such as in mouse midbrain (Bensadoun et al., 2000). The disadvantages of long constitutive gene knockdown using shRNA can generate a secondary effect which is inappropriate to study important gene function that highlight the growth of cancer cells (Wiederschain et al 2009).

On the other hand, the use of inducible knockdown gene using shRNA is preferable to study some genes. For example, demonstrating the function of tumour suppressor genes and oncogenes in epithelial cancer cell lines (law et al, 2012) was challenging using stable gene expression methods. Therefore, developing an inducible system to modulate the gene expression in spatially and temporally controlled manners is important. An Inducing agent regulates inducible gene silencing. Moreover, inducible shRNAs need several clone selection to produce stable cell lines (Wiederschain et al 2009 and Shuen et al., 2015).

1.9 Vectors for RNAi systems

1.9.1 The characteristic of single lentivector for inducible knockdown vector (pSLIK-CD4)

pSLIK stands for single lentivector for inducible knockdown. The design of pSLIK inducible system is slightly different than the pTRIPZ. Ubi-c promoter drives a different cell surface selectable marker, CD4, whereas in pTRIPZ, puromycin is a drug selection marker . Moreover, pSLIK uses TRE to express GFP and miR30-shRNA. Furthermore, it has constitutive part derived by Ubc-I promoter to express reverse tetracycline transactivator (rtTA) and it associated with bicistronic elements IRES that express cell surface marker. These elements allow pSLIK to be tracktable and inducible of one gene or multiple genes knockdown system.

pSLIK-CD4 was design to limit the issue of leaking expression and low transgene expression level in vivo. Furthermore, this system supports the constitutive expression using the GFP to recognize the transduced cells and CD4 for cell sorting. The TRE driven by rtTA3 expression can obtain high expression level in the present of Dox (Shin et. al., 2006).

1.9.2 The characteristic of pGIPZ and pTRIPZ

Open-bio system manufacture design two different plasmids a constitutive plasmid named pGIPZ and inducible system named pTRIPZ (see chapter 2 material and methods, figure 2.3).

pGIPZ and pTRIPZ lentivirus construct was designed based on Puromycin drug selection. Streptomyces Alboniger bacterium makes puromycin. It contains PAC gene that encodes a puromycin ZV-acetyl transferase (PAC). S. Alboniger cells contain N-acetyl puromycin. Exposing living cells to Puromycin showed an active N-acetylate-puromycin since the ribosomes are very sensitive to puromycin in vitro. Activating the N-acetylation of antibiotic inhibits the protein synthesis NH2 group that can stimulate puromycin. The puromycin mechanism occurs during the translation by inhibiting the peptide chain elongation in living cells (Azzam and Algranati, 1973 and Vara et. al., 1985). Thus, cloning and expressing the PAC gene in

living cells has advantage as selectable marker for puromycin resistance (Vara et. al., 1985 and Luna et. al., 1987). PAC gene is encoded in pGIPZ (constitutive system) and pTRIPZ (inducible system) vector backbones (open-biosystem). Therefore, affected cells become puromycin resistant. Thus stable cell lines were studied under certain amount of puromycin which were tested on wild type cells.

pGIPZ (Figure 2.3, A) and pTRIPZ (Figure 2.3, B) lentiviral vectors use MicroRNA (miRNA) based on hairpin developed by Dr Greg Hannon and Dr Steve Elledge (thermo scientific open biosystem pGIPZ, and pTRIPZ shRNAmir). miRNAs have been used in backbone of shRNA sequences. This design creates a configurative recognition as a natural substrate of the RNAi pathway. Moreover, the use of miRNA in shRNA backbone allows a stable and regulative expression from RNA polymerase II (Pol II) promoters. Furthermore, miRNA based on shRNA design enables the "polycistronic tandem" shRNA vectors and linked to reporter genes such as green florescent protein (Fellmann and Lowe, 2014). In addition, the natural occurring of miRNA to generate the desire shRNA can have a potent gene silencing without toxicities (Manjunath et. al., 2009 and Chang et al., 2006). Zeng and et al (2002) use artificial and endogenous human mir-30 to show the mRNA expression for selected target genes in human cells in vivo. They transfect both artificial and endogenous mir-30 with plasmid pCMV-mir30 in 293T/17 cell. They found that artificial mir-30 similar to endogenous mir-30 under RNAi construct can act as siRNA to knockdown the target gene and it can stabilize the gene silencing in human cells. This stable inhibition of target gene can be expressed using viral vectors, which permit the miRNA production in mammalian cells. Furthermore, the expression of inhibitor miRNA works effectively as a part of miRNA (2002). Next, miRNA can reduce the effects of RNAi by interacting with miRNA processing machinery (Aagaard and Rossi, 2007).

pTRIPZ (inducible system) Tetracycline based systems (Tet-On system) can drive the expression of shRNAmir in pTRIPZ system in the presence of Doxycycline (manual technical pTRIPZ). Tet-on system contains two elemen transgenes that work together: the transgenic transcription factors (rtTA)

,and the ic tetracycline-inducible promoter (TRE) (Buchholz, 2012 and TRIPZ technique manual). pTRIPZ was designed after pSLIK. Like pSLIK vector, pTRIPZ uses TRE promoter to drive the RFP turbo following the miR30-shRNA, and the Ubc-I promoter to drive the rtTA-IRES-PURO cassette. pTRIPZ vector gives tighten regulation of shRNA expression in many cell types (Shin et al., 2006 and Hu and Luo, 2012) . Doxycycline (dox) is a widely used antibiotic. It can pass through the placenta barriers and brain tissues. Moreover, it has less cellular toxicity and it can bind to all kinds of tetracycline receptors (TetR) efficiently (Corbel and Rossi, 2002). In the presence of Doxycycline, the structure of rtTA changes. This change allows rtTA to bind to DNA and drive the expression of shRNAmir from the TRE promoter. On the other hand, when there is no Doxycycline, rtTA cannot bind to DNA or regulate the expression of shRNAmir (Buchholz, 2012) (Figure 1.12).

However, inducible systems have some disadvantages. First, the leaking expression can raise an issue in the absence of inducer such as rtTA. Second, the drug toxicity from TRE and/or inducer (rtTA) can affect the cell phenotype. Third, expressing other genes than the target genes by using rtTA may produce off target effects (Meerbrey et al., 2011).

To maintain siRNAs activities in mammalian cells, constitutive vectors were designed based on shRNA. These constitutive vectors contain promoters (Pol II or Pol III) that can drive Transgene expressions in vivo and in vitro (Leung and Whittaker, 2005).

pGIPZ was designed from Hannon-Elledge libraries. They used Lentiviral vectors to obtain high transduction efficiency, and they can transduced non-dividing cells such as neurons. pGIPZ vector (open biosystems, Huntsville, USA) was a modified version from a pervious vector called pPRIME vector from the same library. It uses CMV promoter to derive the expression cassette turboGFP-IRES-puro following by miR30-shRNA. Cells with pGIPZ depend on transduction status not on the expression level of shRNA due to the independent expression of reporter and shRNA

Internal ribosome entry site (IRES) sequence was found in picornavirus RNAs by Jang et. al. ,and Pelletier and Sonenberg in 1988. Two mechanisms

promote the translation in eukayotic cells: first, the association of eukaryotic initiation factors called elfs to 5' end Cap structure of the mRNA; second, the limited elfs association to 5' untranslated region (UTR) of mRNA called IRES. This translational process can be achieved by some metabolic conditions (Amorim et al., 2014). In eukaryotic cells, IRES allow the initiation of translation of RNA, which can permit the expression of more than one protein from a polycistronic transcription unit. Bicistronic application vectors became more preferable for expressing two proteins. One protein is used as a marker such as reporter gene, GFP/RFP, or cell surface marker, CD4, whereas the other protein can express the target gene. The advantages of using bicistronic vectors in vivo or ex-vivo is the high selection of transgene by marker expression. Nonetheless, the first gene in a bicistronic expression vector stimulates the cap-dependent translation while the second gene is activated in an IRES depended manner. This means that the positive selection determines the transcriptional efficiency; Moreover, IRES expression for the second translational protein shows no interference with first translational protein (Martinez-Salas, 1999, and Hellen and Sarnow 2001).

1.9.3 Delivery of RNAi triggers (shRNAmir30)

Lentiviral system delivered RNAi has been shown to be more efficient for integration in to the human genome (Stewart et al., 2003) and mouse cells (Kuhn et al., 2007). Moreover, due to the mutation in U3 region of LTR, the transcriptional process in lentivirus is inactive (Stewart et al., 2003). Furthermore, lentivirus showed better transduction of cycling haematopoietic or cancer cells (Kuhn et al., 2007).

The natural function of retrovirus and lentivirus is to integrate into the host genome and provide a long expression of transgenes through viral genomic backbone. Therefore, the main usage of applying viruses is to generate a stable transgene expression for gene therapy purposes. The advantages of using lentivirus over retrovirus is that lentivirus can transduce dividing as well as non dividing cells.It can be used for long-term expression of a transgene as well as to acommodate large transgene sequences in the viral

plasmid (Manjunath et. al., 2009)

1.9.4 Lentiviral transduction systems for gene delivery

The common use of lentivirus is HIV-based Lentiviral vector. Furthermore, in clinical trial, it was found that the use of envelop HIV-based Lentiviral vector is safe and no sign of mutagenesis or any recombination virus side effects. However, it highlighted the fact that lentivirus gene therapy still at the early stage and there is positional risk of biosafety from using HIV as a vehicle (Manjunath et al., 2009).

Minimizing the viral genome component is achieved by introducing the different components of the viral genome in 3-4 plasmids to produce a transduced virus. These plasmids contain: the packaging cassette that may contain one or two plasmids, the vector expression cassette and the envelop cassette. The packaging elements are *gag and pol* genes which encode the structural and enzymatic proteins that is important to form the viral particles called tat transactivator element, *tat*, that can drive the full expression of RNA vector. Packaging vector also contains *rev* gene that express the Rev protein. This protein is essential for nucleocytoplasmic transport of full length of RNA and viral RNA. RRE is called rev responsive elements. It is a *cis* elements that binds to Rev to regulate nucleocytoplasmic export. These elements pertinent to post-transcriptional mRNA processing events are retained and the full length is achieved using a heterologous constitutive promoter (CMV) and polyadenylation signals (SV40).

The packaging vector started with the first generation with all HIV-1 accessory genes, which raised a concern about bio-safety. The second generation eliminates all the accessory genes except the Rev and Tat elements. The third generation splits the viral vectors into three vectors (Pol/gag, Rev and VSV-g), which are made safer to use without the Tat gene (Cockrell and Kafri, 2007). The last generation has been used by pSLIK-shRNAmir-CD4. Vector expression cassette is a plasmid that contains target shRNA/ shRNAmir. It contains self-inactivating vector (SIN), which showed significant improvement in vector safety. SIN deletion of the

enhancer/promoter elements in the U3 at the 3'LTR increase the vector safety profile. The safety improvement was the reduction of the expression of packaging vector, which reduce the chance for vector mobilization through the infection with HIV virus (Cockrell and Kafri, 2007 and Manjunath et al., 2009 and Zufferey et al., 1998). Improved the transgene expression and transduction efficiency of the lentivirus vector was performed through the incorporation with two elements: the woodchuck hepatitis virus post-transcriptional regulatory element (WPRE) and the central polypurine tract (cPPT) (Cockrell and Kafri, 2007). The Envelop cassette contains vesicular stomatitis virus G protein (VSV-G). VSV-G protein has three main advantages, which are considered to be one of the major developments in lentivirus. First, VSV-G protein can generate pseudotypes with envelope replacement, such as rabies virus G protein and Ross river virus glycoprotein, in various species and cell types. Second, it can be applied to many host cells and tissues. Third, it can maintain the structure of viral particles during the ultra-centrifugation (Petersen et al, 2014 and Cockrell and Kafri, 2007).

Open Biosystem uses five Lentiviral packaging vectors offering more safety profiles (Figure 1.13). The expression of gag-pol and tat-rev are driven by TER promoters, which can only apply when there is tTA expressed in the cells. pTLA1-ENZ vector contains RT-IN (pol) which is packaged by in-frame fusion with accessory vector called virion-associated protein Vpr. This vector is derived from HIV-2 LTR promoter, which is trans-activated by the Tat protein. VSV-g and SV40 elements are part of pTLA1-ENV derived by CMV promoter. pTLA1-TOFF has the tTA element regulated by CMV promoter. This element is critical for the gag-pol and Tat-rev expression vectors. pTLA-1-TAT/REV is controlled by TRE promoter. This vector has IRES, which is a bicistronic element in between TAT and REV.

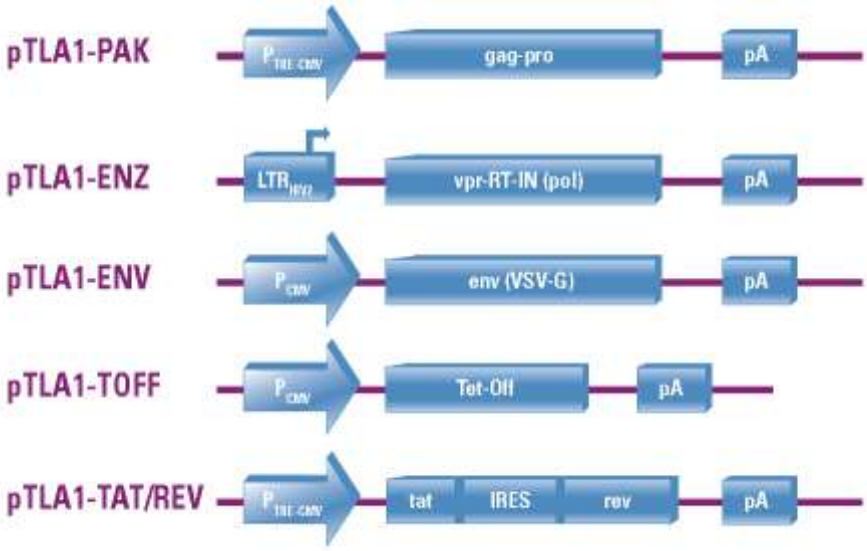


Figure 1.18 open biosystem Trans-Lentiviral packaging vectors (technical manual open bio-systems). Each plasmid is explained in the text.

1.9.5 Measuring the MOI for tittering the Lentiviral particles

Measuring the Titre of lentivirus can predict the viral efficiency of the target cells by calculating the Multiplicity of infection (MOI) (Zhang et al., 2004).

TU/ml= number of Green cells dilution factors* the amount of virus with medium.*

MOI= (TU/ml) / the number of total cells/ml

MOI represents the ratio of viral titration to the number of cells for transduction. Lentivirus vectors mediated gene transfer can be hindered by low gene transfer efficiency or other inhibitory molecules (Zhang et. al., 2004 and Shearer and Saunders 2015). Other factors that also influence the transfection efficiency of the viral vector are high Lentiviral transduction into target cell, which may cause insertion mutagenesis of genomic sequences which give an off target effects. These effects occur when there are several integrations. To overcome these issues, it is favourable to determine the number of infected units of viral supernatant. Serial dilution of the virus particles can obtain an acceptable range of transduction efficiency. Furthermore, focus on single integration (copy number) per cell

can reduce the off target effect (Zhang et. al, 2004, Shearer and Saunders 2015 and Connolly 2002).

In stem cell research, the range of effective MOI in hematopoietic stem cell and gene therapy remains unclear. Moreover, the period of time for the target genes to affect on the results of gene expression is unknown. Therefore, two points were highlighted: the MOI in vitro is not consistent with clinical application and the common use of reporter gene, GFP, in preclinical evaluation (Wahlers et. al., 2001).

The generation of safe, stable and high_titer Lentiviral has been shown promises in gene therapy. The expression of target gene depends on the three packaging vectors to produce Lentiviral particles. To generate stable producer cell line is challenging. However, it will help generating large batches of vectors to apply in vivo and in vitro. Another limitation is the amount of vector titration can be variable from cell to cell depend on VSV-g viral vector particles (Klages et al, 2000).

1.10 Background to this project

Previously, siRNAs were used to transiently knock down Hmgn1 and/or Hmgn2 in the P19 cell line (Mohan 2012). Hmgn1/2 knockdown in undifferentiated P19 cellos resulted in downregulation of the pluripotent markers Oct4, Sox2 and Nanog. Repeated siRNA transfections allowed the knockdown of Hmgn1/2 in P19 cells differentiated into neuronal cells via embryoid bodies. Day 3 neuronal cells showed a reduction of Hmgn1 and Hmgn2 mRNA expression in the double knockdown. However, at the protein level only around half of the protein was lost. Rest, Zfp521, Nse, and GlyT2 were downregulated into the Hmgn1/2 double knockdown whereas Nestin, Map2, NF-160, Nmda receptor subunit 2, GlyT1a expression showed upregulation (Mohan, 2012). These experiments were difficult to reproduce due to variability in siRNA transfection levels, cytotoxicity following transfection, and variability in the efficiency of differentiation

1.11 Aim

This aim of this project was to develop a lentiviral system to knockdown Hmgn2 expression in embryonal carcinoma cells. This system would be used to determine the role of Hmgn2 in pluripotent stem cells, with particular focus on regulation of key pluripotency factors.

1.12 Hypothesis

These experiments aim to test the hypothesis that Hmgn2 modulates chromatin structure and regulates the transcription of key transcription factors associated with pluripotency.

1.13 Project Outline

The work presented here uses lentiviral vectors to knock down Hmgn2 expression, thus avoiding repetitive transfection and cytotoxicity. Constitutive (pGIPZ) and doxycycline-inducible (pSLIK and pTRIPZ) shRNA vectors were compared for their ability to achieve a high level of knockdown in P19 EC cells. FACS was used to monitor the efficiency of viral

transduction, and changes in gene expression were assayed using qRT-PCR, western blotting, FACS and immunofluorescence. ChIP was using to assay changes in chromatin structure following Hmgn2 knockdown.

1.14 Objectives

The main objectives can be summarised as follows:

1. To optimise conditions for lentiviral transduction of P19 cells
2. To compare the pSLIK, pTRIPZ and pGIPZ lentiviral systems for their ability to knockdown Hmgn2 expression and choose the best system for further experiments.
3. To investigate whether Hmgn2 knockdown affects the expression and chromatin binding of various transcription factors associated with pluripotency
4. To investigate whether Hmgn2 knockdown affects Hmgn1 binding, H3K4me3 and H3K27me3 at the *Oct4* and *Nanog* gene loci.

Chapter 2 Materials and methods

2.1 Materials

2.1.1 Gateway cloning system

Reagents and samples

TE buffer PH 8.0 (10mM Tris-HCL, PH 8.0, 1mM EDTA)

shRNA sequences for HMGN1 and HMGN2 were designed at Qiagen company

Names (entry vector)	Catalogue	Stock Concentration
HMGN1-3 in PEN-TGm1	C09879B	4ug
HMGN2-1 in PEN-TGm1	C09879C	4ug
HMGN1-2 in PEN-TGm1	C09879A	4ug
Scramble in PEN-TGm1	C09879D	4ug

Destination vectors	Catalogue	Stock concentration
pSLIK-CD4	L70DDLUTICXA	250 ug/ul
pSLIK-HYGRO	L55DDLUTIHX	202.91 ug/ul

LR Recombination Reaction

Gateway LR Clonase II plus enzyme Mix catalogue number: 12538-120

LR recombination component	Reaction sample	Negative control	Positive control
Entry clone	1ul	1ul	-
Destination clone	1ul	1ul	1ul
TE buffer	2ul	3ul	3ul
Clonase II enzyme	1ul	-	1ul
Total	5ul	5ul	5ul

2.1.2 PSLIK and pGIPZ, pTRIPZ vectors and packaging vectors

PSLIK-CD4, pSLIK-hygromycin, and pENtgMRC3

PSLIK vectors on bacterial strap were purchased from Add-gene Company without shRNA sequences

Construct barcode	Construct name	Vector details	Vector size	Digest fragment	Bacterial selection	Mammalian selection	Ecoli-propagation-strain
L55DDLUTIHXA	pSLIK-Hygro	shRNA Lentiviral Expression; Ubi-c promoter driven rtTA3-IRES-Hygro	13918	KpnI: 10173; 2403; 1342	Ampicillin	Hygromycin	DB3.1
L70DDLUTICXA	pSLIK-CD4	shRNA Lentiviral Expression; Ubi-c promoter driven rtTA3-IRES-delta CD4 (human)	13770	KpnI: 10173; 2321; 1276	Ampicillin	CD4	DB3.1
P30ETREMIRAG	pEN_TmiRc3	shRNA expression; miR30-based topology; TRE Pmin promoter; Entry vector backbone; +ccdB in parent; Multiple unique sites d/s of TRE	4587	BsrG1: 1889, 1478, 773, 447	Gentamicin	--	DB3.1

PGIPZ and pTRIPZ

PGIPZ was purchased from thermo scientific with shRNA sequences

Vectors names	Barcode name	Vector size	Bacterial selection	Mammalian selection
pGIPZ-shHMGN2	V3LMM-480381	14345	Zeocin and ampicillin	Puromycin
pGIPZ-shHMGN2	V3LMM-516968	14575	Zeocin and ampicillin	Puromycine
pGIPZ-shHMGN2	V2LMM-55514	14656	Zeocin and ampicillin	Puromycin
pGIPZ-shGAPDH	RH54371	14766	Zeocin and ampicillin	Puromycin
pGIPZ-shNon-silencing	RH54346	14676	Zeocin and ampicillin	Puromycin
pGIPZ-shEG5	RH54480	13565	Zeocin and ampicillin	Puromycin

PTRIPZ empty vector was gift from Professor Peter D. Adam, Beatson Institute, Glasgow G61 1BD. shHMGN2 were transformed following the protocol from thermo scientific and it was performed by Dr Katherine west.

Packaging vectors

Vectors name	Barcode name	Vector size	Manufacture
PSLIK-CD4 and pSLIK-HYGRO (3rd generation packaging vectors)			
pCMV-VSV-G	8454	6363bp	Add-gene
pRSV-REV	12253	4174bp	Add-gene
pMDLg/pRRE	12251	8895bp	Add-gene
pGIPZ and pTRIPZ (trans-packaging mix)			
Trans-lenti shRNA Packaging kit	TLP5913	Five Mix-packaging vectors	Thermo scientific

2.1.3 Lentiviral reagents

Reagents	Manufacture	Catalogue number	Lentiviral system
Lipofectamin 2000	Invitrogen	11668-019	pSLIK
OptIMEM	Gibco	31985-062	pSLIK
ULTRA cultural	Lonza	g1234	pSLIK , pGIPZ and pTRIPZ
CacL	Thermo scientific	TLP5913	pGIPZ and pTRIPZ
HBSS buffer 10x	Thermo scientific	TLP5913	pGIPZ and pTRIPZ

2.1.4 Lentiviral transduction kits

Kit name	manufacture	Catalogue number	System
Amicon ultra-15 centrifugal filter device	Millipore	UFC910024	Fast filtration with the capability for high concentration factors and easy concentrate recovery from dilute and complex sample matrices
Lenti-x geoStix	Clontech	631243	confirming the presence of lentivirus in packaging cell supernatants
Transdux	System Bioscience	LV850A-1	Enables high transduction rates of virus into most cells, even those that are resistant to infections.
Polyberen	Sigma	107689	The presence of polybrene Can efficiently serve as the method of transduction for many Cell types, some cells are more difficult to transduce.
Lenti-X™ Concentrator	Clontech	PT4421-2	concentrating lentiviral stocks
Lenti-X™ p24 Rapid Titer Kit	Clontech	632200	quickly determine the titre of any HIV-1-based lentiviral supernatant using standard ELISA methods

2.1.5 Plasmid purification, transformation, and electrophoresis

DNA Plasmid purification

QIAprep spin Miniprep Kit catalogue number 27104

Qiagen plasmid plus midi kit catalogue number 12943

Transformation

Solid Lysogeny broth (LB) in 1L

Reagents	Company	Catalogue number	Concentration
Lysogeny broth (LB) bacteria medium	Sigma-Aldrich	L3022-1KG	20g/l
NaCL			5 g/l
Agar			10g/l
To be complete up to 1000ml with dH2O			

Liquid Lysogeny broth (LB) in 1L

Reagents	Company	Catalogue number	Concentration
Lysogeny broth (LB) bacteria medium	Sigma-Aldrich	L3022-1KG	20g/l
NaCl			5g/l
To be complete up to 1000ml with dH2O			

Subcloning Efficiency™ DH5α™ Competent Cells from Invitrogen catalogue number 18265-017

Screening using restriction digest

Reagents	Catalogue number	Concentration at 1x
Buffer 1	B70015 10X conc.	4ul
BSA	0121005 10mg/UL	1ul
DNA	-----	1ug
Kpn1 (NEB)	1425 10000U/ml	0.5 ul
DH2O		Up to 40ul

Gel electrophoresis

To prepare Buffer and dyes

50x TAE

242g TRIS BASE

57.1ml Acetic acid

18.69 EDTA 1hr for 1L

Product	Manufacturer	Product Code
P19 CELL LINES with alphaMEM		
Medium (MEM)-alpha containing deoxyribonucleotides, ribonucleosides and ultraglutamine 1	Lonza	BE02-002F
New born calf serum 7.5%	Source Bioscience,	7.03Hi
Fetal bovine serum 2.5%	Sigma	F9665
P19 cell lines with DMEM/F12		
Advanced DMEM/F12	GIBCO	12634-010
New born calf serum 10%		
Glutamax 100x	Gibco	
HEPES (transduced cells only) 100x	Gibco	
Hepa cell lines		
DMEM with glutamine	Gibco	10566-016
FBS 10%	Sigma	F9665
E14 cell serum free		
NDiff 227 500ml	Stem Cell Sciences	SCS-SF-NB-02
LIF-recombinant	ESGRO, Millipore,	10^7 U/ml, 500 U/ml final
MEK inhibitor	SelleckCHem	1 µM final, PD0325901 Cat. S1036) 50 ul of 10 mM stock in DMSO
GSK3-B inhibitor	SelleckCHem	3 µM final, CHIR99031 Cat. S1263) ,150 ul of 10 mM stock in DMSO
HEK392/17		
DMEM GLUTAMAX	Gibco	10566-016
FBS 10%	Sigma	F9665

BPB 300bp

1kb loading ladder

75ul H2O

5ul ladder

20ul LB (loading buffer).

Xylene cyanol dye 4kb

25mg xylene cyanol in 10ml

6 X of leading buffer in 30% glycerol

DNA ladder

Reagents	Manufacture	Product number
100 bp DNA ladder	NEB	N32315
1 Kb DNA ladder	Invitrogen	10787-018

2.1.6 Cell lines

Cell lines	Reference	Source
P19 cell lines Passage 14	(McBurney and Rogers, 1982)	P19 EC cell line was provided by Dr. Andrew Hamilton, Institute of Cancer Sciences, University of Glasgow
Mouse hepatoma line Hepa-1	(West et. al., 2004)	From ATCC,
E14 cell lines	(Hooper et. al., 1987)	From Cambridge
HEK293T/17	(Pear et. al., 1993)	From Adam west lab

2.1.7 Drug reagents

Drug names	Manufacture	Catalogue number
Puromycine	Sigma	P7255-25MG
Hygromycin b	Sigma	H3274-50MG
Doxycycline	Sigma	D9891-1G

2.1.8 Molecular biology reagents

RNA extraction:

RNeasy mini kit catalogue 74104 from qiagen

CDNA formation:

Reagent	Manufacturer	Product Number
Oligo DT 100UM	Bioscience	
100 mM dNTP set	Invitrogen	10297018
FastStart Universal SYBR Green Master (Rox)	Roche	04913914001
Super script III	INVITROGEN	18080-044
Proteinase K	Sigma	P5568
RNase A	Fermentas	EN0531
RNase OUT	Invitrogen	160000840
RNase H	NEB	M0297

Oligonucleotides

Oligonucleotides were ordered from IDT or MWG Eurofins used as primer for RT-PCR

Gene	Primer	Sequence	Primer Efficiency
Hmgn1	Forward	AGAGACGGAAAACCAGAGTCCAG	1.98
	Reverse	CGTGATGGATGCTTAGTCGGA	
Hmgn2	Forward	AAAAGGCCCCTGCGAAGAA	1.96
	Reverse	TGCCTGGTCTGTTTTGGCA	
Oct4	Forward	CGTTCTCTTTGGAAAGGTGTTCA	1.98
	Reverse	GGTTCTCATTGTTGTCGGCTTC	
Nanog	Forward	ACCTGAGCTATAAGCAGGTTAAG	2.0
	Reverse	TCAGACCATTGCTAGTCTTC	
Sox2	Forward	GGAACAGCATGGCGAGCGG	1.96
	Reverse	CGTTCATGTGCGCGTAGCTG	
α-tubulin	Forward	ACCCACGGTCATCGATGAAGTT	1.98
	Reverse	TCCTTGCCAATGGTGTAGTGGC	
β-Actin	Forward	GTGAAAAGATGACCCAGATC	1.97
	Reverse	GTGTGGGTGACCCCGTCTCC	
C-myc		CTGTTTGAAGGCTGGATTTCC	1.92
Klf4		AACCTATACCAAGAGTTCTCAT	2.40

Oligonucleotides for QPCR-ChIP

Oligonucleotides were ordered from Eurofins for Oct4 and Nanog genes primer design around the TSS

Gene	Primer	Sequence	Primer position
Oct1	Forward	GTGAGCATGACAGAGTGGAGGAA	-1781
	Reverse	TCTCTGGCCCTCTCCATGAAT	
Oct2	Forward	GTGGGTAAGCAAGAACTGAGGA	-399
	Reverse	TGGAGAGCCTAAAACATCCATT	
Oct3	Forward	CAATGCCGTGAAGTTGGAGA	+410
	Reverse	TCACTTACCTCCTCGGGAGTTG	
Oct4	Forward	TCAAGGCTTCTCACCTCCAGA	+2070
	Reverse	ACAGAGAAAGCAAGGCAGCAG	
Oct5	Forward	GCTGCCTCACTCACTCTGTTTG	+3342
	Reverse	TGGCTGAACACCTTTCCTGAA	
Nanog1	Forward	GGAAGAACCACTCCTACCAATACTCA	-1081
	Reverse	CGTAACATCTCCCATGTGAAGACTC	
Nanog2	Forward	TCTTTAGATCAGAGGATGCCCCCTAAG	-219
	Reverse	AAGCCTCCTACCCTACCCACCCCCTAT	
Nanog3	Forward	TCAGCCCAGTACTCAGGCTTGT	+929
	Reverse	AGCCTAGCAGCCTCTTGGTTCT	

Nanog4	Forward Reverse	TAACTGGACCCTCTGACTGGCT CCCACCATCTTTTCTGCTAGTACAAG	+1740
Nanog5	Forward Reverse	TGTAGATGCGGTCACGGTTC CCAATCCCTTCTCCCTGCT	+2636
ß-actin	Forward Reverse	CGCCATGGATGACGATATCG CGAAGCCGGCTTTGCACATG	
MSat	Forward Reverse	GACGACTTGAAAAATGACGAAATC CATATTCCAGGTCCTTCAGTGTGC	

2.1.9 Antibodies

Antibodies for IF (immune florescent)

1st Antibody	Cat Number	Concentration	Species	2nd Antibody	Cat Number	Concentration	Colour
Hmgn1	Homemade	1/500	Rabbit	Donkey anti-rabbit	A21206	1/1000	Green 488
Hmgn2	Homemade	1/500	Rabbit	Donkey anti-rabbit	A21206	1/1000	Green 488
Hmgn3 (EP062146)	Homemade	1/5000	Rabbit	Donkey anti-rabbit	A21206	1/1000	Green 488
SSEA1	AB16285	1/200	MOUSE	Donkey anti-mouse	AB150111	2ug/ml	Red (647)
Nanog	AB80892	1/1000	RABBIT	Donkey anti-rabbit	A21206	1/1000	Green 488
Sox2	Ab97959	1/1000	Rabbit	Donkey anti-rabbit	A21206	1/1000	Green 488
Map2	Ab5392	1/1000	Chicken	Goat anti-chicken	Ab150175	1/500	Red 647
Oct3/4 santa	SC8628	1/500	GOAT	RABBIT ANTI-GOAT	A11078	1/1000	Green 488
Oct4 AB	AB19857	1/500	RABBIT	Donkey anti-rabbit	A21206	1/1000	Green 488
NeuroD1	Ab60704	1/1000	Mouse	Donkey anti-mouse	A21206	2ug/ml	Red (647)
TBP		1/400	Mouse	Goat anti-mouse	A21206	1/200	Yellow

Antibodies for western blot on whole protein and chromatin

1st Antibody	Cat number	Expected size	Concentration	Species	2nd Antibody	Cat number	Concentration
Hmgn1	Homemade	14 kDa	1/1000	Rabbit	Goat anti rabbit HRP	2346	1/2000
Hmgn2	Homemade	17 kDa	1/2000	Rabbit	Goat anti rabbit HRP	2346	1/2000
TBP	Ab51841	34-40 kDa	1/3000	Mouse	GOAT ANTI MOUSE HRP	4567	1/5000

Oct4	Ab19857	50 KDa	1/500	Rabbit	Goat anti rabbit HRP	2346	1/2000
H2AZ Millipore	07-594	13 KDa	1/1000	RABBIT	Goat anti rabbit HRP	2346	1/2000
P300	Sc584	300 KDa	1/1000	RABBIT	Goat anti rabbit HRP	2346	1/10,000
P53 (SC)	Sc-99	50 KDa	1/500	Mouse	GOAT ANTI MOUSE HRP	4567	1/2000
Nanog	Ab80892	35KDa	1/300	Rabbit	Goat anti rabbit HRP	2346	1/2000
Sox2	Ab97959	34 KDa	1/1000	Rabbit	Goat anti rabbit HRP	2346	1/5000
H3.3		17 KDa	1/100,000	RABBIT	Goat anti rabbit HRP	2346	1/2000
B-tubulin (thermo)	12430-1	43KDa	1/3000	Mouse	GOAT ANTI MOUSE HRP	4567	1/10,000
HMGN3b			1/3000	Rabbit	Goat anti rabbit HRP	2346	1/2000
H3K4me3 (Millipore)	07-473	14 KDa	1/5000	Rabbit	Goat anti rabbit HRP	2346	1/2000

2.1.10 Protein extraction and analysis

SDS-PAGE using Invitrogen protocols:

Reagents and buffers	Manufacture	Catalogue number
Nupage 4-12% bis-tris gel	Invitrogen (novex)	NP0322BOX
Nupage 12%	Invitrogen	NP0321BOX
Reducing agents 10%	Invitrogen	NP0004
Antioxidant	Invitrogen	NP0005
Nupage Mops	Invitrogen	NP0001
Running buffer 20x	Invitrogen	
Transferring buffer 10x	Invitrogen	NP0006
Nupage LDS sample Buffer 4X	Invitrogen	NP0007

Methanol	Thermo scientific	
Milk (SEMI-SKIMMED)	Marvel	MARVEL
TWEEN 20	Thermo scientific	
TRIS PH7.5		
NaCL 5M		

BIO-RAD protocol

Reagents and machine	Catalogue number
BLOTING TRANS-BLOT TURBO TRANFER SYSTEM PACK	1704155
MINI TRANSFER PACKS	1704156

BUFFERS
TBST buffer 1x (antibody blocking buffer)
20mM tris ph 7.5 20ml 1M TRIPZ PH7.5
150mM NaCl 30ml 5M NaCl
Up to 1L of dH2O + TWEEN 20

2.1.11 Chromatin immunoprecipitation reagents and buffers

5% Formaldehyde Solution (can be stored at 4C for 1 week)
(5% HCHO, 50 mM HEPES pH8.0, 0.1 M NaCl, 1 mM EDTA, 0.5 mM EGTA)

Formaldehyde dilution buffer
50 mls
1M HEPES pH 8.0 2.5 ml
5M NaCl 1.0 ml
0.5M EDTA pH 8.0 0.1 ml
0.5M EGTA pH 8.0
Water

Use single use vials of 16% (?) formaldehyde.
Dilute to 5% in formaldehyde dilution buffer.

1 M Glycine Solution

	50 ml	100 ml
Glycine	3.76g	7.51g

2x Cell lysis buffer
(1x is 10 mM Tris pH8.0, 10 mM NaCl, 10 mM EDTA, 0.5 mM EGTA, 0.2%NP40 and protease inhibitors)

	25 mls
1M Tris pH8.0	500 µl
5M NaCl	100 µl
10% NP40 or IPEGAL (w/v)	1 ml

0.5M EDTA
0.5M EGTA
Water

10x stock of mini protease inhibitor tablet – 1 tab in 1 ml water.

6 ml of 1x cell lysis buffer: 3 ml of 2x, 600 ul protease inhib stock, 2.4 ml water.

2x Nuclei Lysis Buffer
1x is 50 mM Tris pH8.0, 10 mM EDTA, 0.5% SDS

	25 mls
1M Tris pH8.0	2.5 ml
0.5M EDTA pH8.0	1 ml
10% SDS	2.5 ml

3 ml 1x nuclei lysis buffer: 1.5 ml of 2x, 300 ul of protease inhibitor stock, 1.2 ml water.

IP Dilution Buffer
(12.5 mM Tris pH8.0, 0.19 M NaCl, 0.01% SDS, 1.2% Triton X-100)

	50 ml	100ml
1M Tris pH8.0	625 μl	1.25 ml
5M NaCl	1.9 ml	3.8 ml
10% SDS	50 μl	100 μl
10% Triton X-100	6 ml	12 ml
Water	41.42 ml	

IP Wash I
(20 mM Tris pH8.0, 0.15 M NaCl, 2 mM EDTA, 0.1% SDS, 1% Triton X-100)

	50ml
1M Tris pH8	1 ml
5M NaCl	1.5 ml
0.5M EDTA	0.2 ml
10% SDS	0.5 ml
10% Triton X-100	5 ml
Water	41.8 ml

High Salt Buffer
(20 mM Tris pH8.0, 0.5 M NaCl, 2 mM EDTA, 0.1% SDS, 1% Triton X-100)

	50 ml
1M Tris pH8	1 ml
5M NaCl	5 ml
0.5M EDTA	0.2 ml
10% SDS	0.5 ml
10% Triton X-100	5 ml
Water	38.3 ml

IP Wash II
(10 mM Tris pH8.0, 0.25 M LiCl, 1 mM EDTA, 1% Nonidet P40, 0.5% sodium deoxycholate)

	50 ml
1M Tris pH8	0.5 ml
5M LiCl	2.5 ml
0.5M EDTA	0.1 ml
10% Nonidet P40	5 ml
10% sodium deoxycholate	2.5 ml
Water	39.4 ml

TE Buffer
(10 mM Tris pH8.0, 1 mM EDTA)

	50 ml
1M Tris pH8	0.5 ml
0.5M EDTA	0.1 ml
Water	49.4 ml

Elution Buffer: prepare fresh (some recommend heating to 65C before use)
(1% SDS, 0.1 M $NaHCO_3$)

	10ml
1% SDS	10 ml
$NaHCO_3$	0.08 g

(Easiest: weigh an aliquot of NaHCo3 and add the correct amount of 1% SDS solution.)

2.1.12 Co-Immunoprecipitation (Co-IP) for chromatin and whole protein cell extract

Buffer1
TBS
0.5% NP40
15mM MgCl2

Buffer2 (benzonase buffer)
50mM Tris pH 7.5
300mM NaCl
0.5% NP40
2.5 mM MgCl2

Buffer3 (Dilution buffer)
50mM Tris pH 7.5
300mM NaCL
0.5% NP40
15mM EDTA

Buffer4 (wash buffer1)
TBS (TBS has 150mM NaCl)

0.5% NP40
5mM EDTA

Buffer5 (wash buffer2)
50mM Tris
5mM EDTA
0.5% NP40
300mM Nacl

2x sample Buffer
4x sample buffer
10x reducing agent
Sterile dh2o

2.2 Methods

2.2.1 ShRNA commercial sequences

siRNA triggers obtained from Qiagen. The manufacture had siRNA triggers for Hmgn1 and Hmgn2 and negative control. It was tested which triggers performed better in knockdown. Each trigger was identified by two points: the position in the genome and N-scramble (based on HMGN1_3 siRNA target sequence) targeted the Hmgn2. For selecting the shRNA we use a web tool developed by lab of Greg Hannon. The output was filtered selecting for sequence feature experimentally shown to increase the efficiency of RNAi processing (Mittal , 2004) as well as for any potential off-target effects using a BLAST search for the human genome leaving four putative targets shown in table .

Names	siRNA targets	shRNA sequences
HMGN1_2	5'CACTGGAACA AGTTCAAAGAA 3'	5'AGCGCACTGGAACAAGTTCAAAGAATGATATG TGCATTCTTTGAACTTGTTCCAGTG -3' 3'GTGACCTTGTTCAAGTTTCTTACTATACACGT AAGAAACTTGAACAAGGTCACACGG -5'
HMGN1_3	5'TTGTGATAAT GTGCTGTGAAA 3'	5'- AGCGTTGTGATAATGTGCTGTGAAATGATATGT GCATTTCACAGCACATTATCACAA -3' 3'AACACTATTACACGACACTTTACTATACACGTA AAGTGTCGTGTAATAGTGTTACGG -5'
HMGN2_1	5'CAGATTGATA ATTCTGCCTAA 3'	5'AGCGCAGATTGATAATTCTGCCTAATGATATG TGCATTAGGCAGAATTATCAATCTG -3' 3'GTCTAACTATTAAGACGGATTACTATACACGT AATCCGTCTTAATAGTTAGACACGG -5'

HMGN scramble	5'GGTATAGTAC GGTTTAGATAT 3'	5'AGCGGGTATAGTACGGTTTAGATATTGATATG TGCAATATCTAAACCGTACTATACC -3' 3'CCATATCATGCCAAATCTATAACTATACACGTT ATAGATTTGGCATGATATGGACGG -5'

Table 2.1 the list of selected target sequence
The number of each target ID indicate the position of the first nucleotide with respect to the Hmgn1 and Hmgn2 coding sequence from Qiagen manufacture.

Plasmid pGIPZ is constitutive Lentiviral system that can knockdown gene for long time (Figure 3.20A). Five Lentiviral pGIPZ plasmids: Three triggers for Hmgn2 gene, one negative control (non-silencing), and two positive control GAPDH and EG2 were performed in E14, P19 and Hepa cell lines (Figure 3.20 B). Moreover, the potential off target effect were performed using BLAST search of mouse RNA-seq database leaving putative targets of Hmgn2 and some pseudogenes. The selected RNA prediction gene targets were selected based on E value 0.039(the number of hits for particular size the lower E-value the more significant a match to the database sequence and very small probability to have match by chance) coverage 100%, and identification 100%.

target ID	shRNA sequence (19 nt)	RNA-seq targets	TARGET	shRNAmir type
V3LMM_480381 Mature Antisense	TAGCAGACAACCTCGCAGA	predicted gene 10357 (Gm10357), misc_RNA predicted gene 6750 (Gm6750), misc_RNA predicted pseudogenes 7931 (Gm7931), misc_RNA predicted gene 10357 (Gm10357), misc_RNA high mobility group nucleosomal binding domain 2 (Hmgn2), mRNA	ORF	GIPZ/TRIPZ
V3LMM_516968 Mature Antisense	GTTACTGACACATGATCCT	predicted gene 10357 (Gm10357), misc_RNA predicted pseudogenes 7931 (Gm7931), misc_RNA predicted gene 10357 (Gm10357), misc_RNA	3'UTR	GIPZ/TRIPZ

		high mobility group nucleosomal binding domain 2 (Hmgn2), mRNA		
V2LMM_55514 Mature Antisense	ACAATCAT GGTATTTC AGG	predicted pseudogenes 7931 (Gm7931), misc_RNA high mobility group nucleosomal binding domain 2 (Hmgn2), mRNA	3'UTR	GIPZ/TRI PZ
N2-1 mature antisense (Qiegen)[1]	5'CAGATT GATAATTC TGCCTAA 3'	Mus musculus predicted gene 10357 (Gm10357) Mus musculus predicted pseudogene 7931 (Gm7931) Mus musculus predicted gene 10357 (Gm10357) Mus musculus high mobility group nucleosomal binding domain 2 (Hmgn2)	3'UTR	TRIPZ only

Table 2.2 the list of selected target sequence
The number of each target ID indicate the position of the first nucleotide with respect to mRNA of Hmgn2 coding sequence in pGIPZ and pTRIPZ

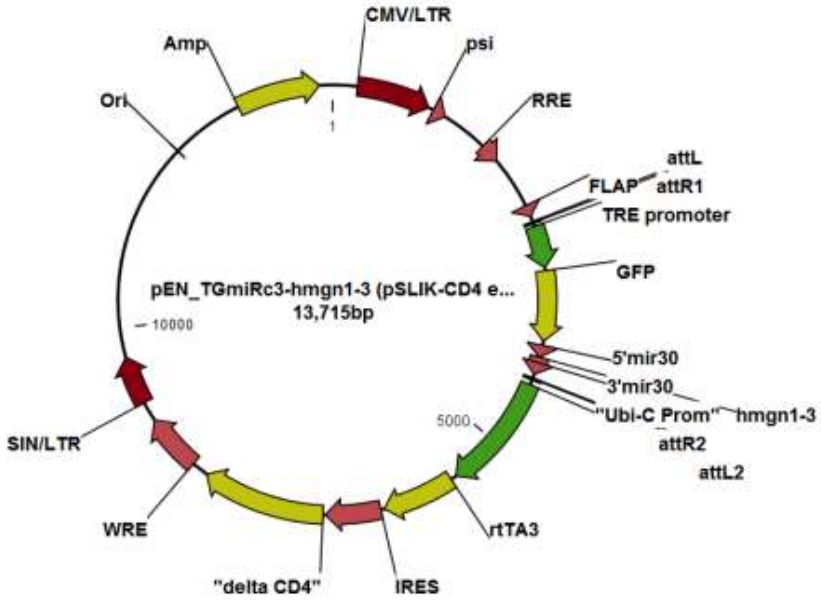
2.2.2 Plasmids construct (pSLIK, pGIPZ and pTRIPZ)

2.2.2.1 The construct of pEN-shRNA-Hmgn_(pSLIK-CD4)

Schematic Construct of pEN-TGRmiRc3-Hmgn2-1-(pSLIK-CD4-expression vector) from pENTR-TGmiRC3 and pSLIK-CD4 after gateway system was performed to generate Lentiviral knockdown gene using RNAi system (shRNAmir30). pSLIK-CD4 vector is tet-regulatable expression system of shHmgn2miR-30. it carries GFP and shRNAmir under the control of TRE promoter which can be activated under Doxycycline binding to rtTA3 (Tet-on/off system). Ubi-C promoter regulates CD4 , rtTA3 , and IRES which allows multisictronic transcript and drive the constitutive expression of the vector. Adding the human miR30 sequence between shHmgn ensured the expression specificity of Hmgn. pSLIK has another feature which can be constitutive system instead of inducible by changing the selection marker (CD4). 5=-long terminal repeat; FLAP, DNA flap region consisting of a central polypurine tract and a central termination sequence; TRE, tetracycline responsive element; Ubi-c, ubiquitin C promoter; rtTA3, reverse

[1] Target ID N2-1 sequence trigger was designed by quiegen company

tetracycline transactivator 3; Hygromycin, hygromycin-resistance gene; WRE, woodchuck hepatitis virus response element; 3=SIN LTR, 3=-self-inactivating long-terminal repeat.



Vector Elements	Terminology	Function	References
Expression elements			
TRE promoter	Tetracycline responsive element	Inducible Promoter controls the tTA transactivator that mediates the transgene expression.	(Pan et. al., 2008 and Hioki et. al., 2009)
GFP reporter gene	Green fluorescent protein	Reporter gene for visualized the transfected cells.	(Hioki et al, 2009)
5' and 3' miR30	Micro-RNA 30	Increasing the RNAi knockdown efficiently at protein level	(Stegmeier et al., 2005)
shRNA	Short hairpin RNA	RNAi technology for gene silencing.	(Song and Yang, 2010)
UBC-I promoter	Ubiquitin C promoter	Allow the high expression of transgene	(Schorpp et al, 1996)
rtTA3	Reverse tetracycline transactivator	Fusion protein interacts with Dox deriving the gene expression.	(Sisson et al., 2006)
IRES	Internal ribosome entry site	Translational initiation from an internal region of the mRNA	(Ibrahimi et al, 2009)
Delta CD4	Cell surface marker	Cell surface protein able to detect the transfected cells without affecting the cell function	(Cooray et al, 2012)
Lentiviral elements			
WRE	Woodchuck	Post-transcriptional regulatory	(Zufferey et

	hepatitis virus response element	element stimulates the expression of transgene via increase the nuclear export.	al., 1998)
SIN/LTR	3′-self-inactivating long-terminal repeat	Help insertion mutagenesis, integration site selection, and the potency of transgene expression	(Xu et al., 2011)
CMV/LTR	5′LTR, 5′-long terminal repeat	Drive a high level of expression	(Miyoshi et al.,1998)
PSI	Retroviral Psi packaging element	Serve as the site of initiation of positive-strand DNA synthesis during reverse transcription	(Warnock et al.,2011)
RRE	HIV rev responsive elements	Essential for viral replication and serve as RNA scaffold that help to assemble the RNP complex to regulate the nuclear export of viral massage	(Fernandes et al., 2011)
FLAP	FLAP, DNA flap region consisting of a central polypurine tract and a central termination sequence	Enhance the HIV-1 infection by assessing nuclear import of proviral DNA and it can protect the viral DNA.	(Kankia and Musier-Forsyth, 2006)
Cloning elements			
Ori	Open reading frame	Encode no more than 15 proteins that involved in the viral life cycle including structural and regulatory proteins	(Tiscornia et al.,2007)
Amp	Drug resistance ampicillin	Antibiotic bacterial selection	(shin et la.,2006)

Figure 2.1 lentivirus plasmid map within PSLIK system as destination vector flanked by att recombination sites

PSLIK as backbone contains CMV/LTR (polymerase II) promoter, GFP a reporter gene to mark shRNAmir expression, rtTA3 tetracycline antiduction (TET system) or tetracycline responsive gene expression to allow construction of stable helper cell lines that produce pseudotyping VSV-G. A. pSLIK-CD4/mmHmgn plasmid map the production of gateway cloning system between pSLIK and pEN vectors. B. table describe the elements and their function in pSLIK-CD4 plasmid map. Schematic Construct of pEN-TGRmiRc3-Hmgn2-1-(pSLIK-Hygromycin-expression vector) from pENTR-TGmiRC3 and pSLIK-Hygromycin after gateway system was performed to generate Lentiviral knockdown gene using RNAi system (shRNAmir30). TRE, tetracycline responsive element; Ubi-c, ubiquitin C promoter; rtTA3, reverse tetracycline transactivator 3; Hygromycin, hygromycin-resistance

gene; WRE, woodchuck hepatitis virus response element; 3=SIN LTR, 3=-self-inactivating long-terminal repeat.

2.2.2.2 The construct of constitutive pGIPZ Hmgn2 knockdown in P19 and Hepa cells lines

Experimental process to generate viral particles containing Hmgn2 miRNAs, were performed using open biosystem procedures protocols for pGIPZ viral production. The viral particles were generated in the packaging HEK293T clone 17 cell lines (Figure 3.21). The plasmids were received from Open-biosystem and they were cloned into competent cells (DH5alpha); then midi DNA purification was performed.

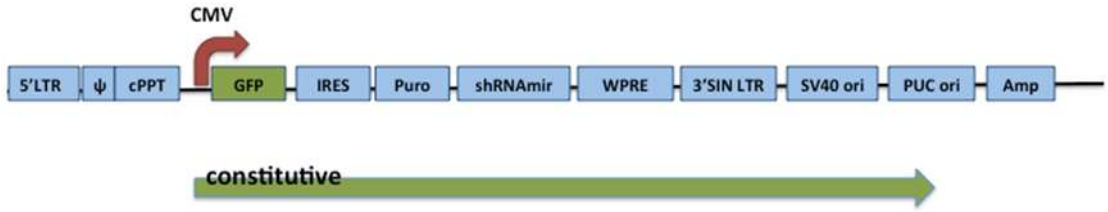


Figure 2.2 schematic diagram illustrating the expression cassette encoded by pGIPZ lentiviral vector
Bi-cistronic element (IRES) transcript encoding green florescent proteins (GFP) and puromycin drug selection and shRNAmir driven by CMV promoter. Abbreviation: IRES, internal ribosome entry elements; WRE, woodchuck response element; LTR, long terminal repeats; puro, puromycin.

2.2.2.3 Construction of inducible Hmgn2 knockdown in P19 and Hepa cell lines

Based on preliminary studies, Hmgn2 knockout in mice embryo is lethal. This indicates that constitutive system could cause cell toxicity and high rate of cell death before and after drug selection. Therefore, it was thought that inducible system could reduce the cell toxicity and could gain high Hmgn2 knockdown efficiency.

This section describes the steps that undertaken in order to generate P19 cell lines with Hmgn2 knockdown using inducible system called TRIPZ from open bio-system. Four TRIPZ inducible shRNAs sequence triggers in addition to two each negative and positive controls were used in this experiment.

The same sequence triggers were used in GIPZ and TRIPZ experiments except N2-1 was used on only TRIPZ system. GIPZ plasmids were received from Open-biosystem and they were cloned into competent cells (DH5alpha); then midi DNA purification was performed. shRNAmir sequence triggers of GIPZ were cloned into empty TRIPZ vector. TRIPZ empty vector was gifted from Dr. Peter Adam. Moreover, the same process was performed on pSLIK shRNAmir sequence triggers cloned into TRIPZ system.

Experimental processes to generate viral particles were generated in the packaging HEK293T clone 17 cell lines (Figure 3.11).

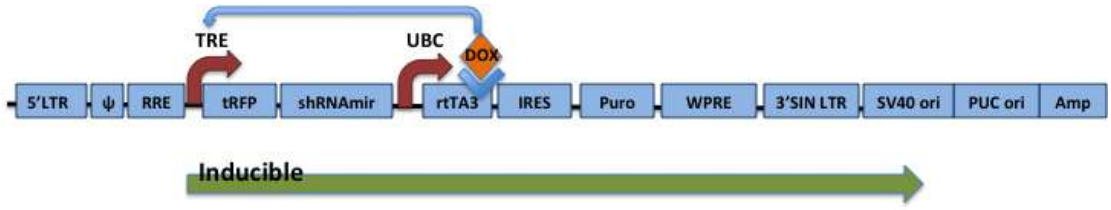


Figure 2.3 schematic diagram illustrating the expression cassette encoded by pTRIPZ lentiviral vector.
TRIPZ inducible system contains two parts. First part is constitutive system derived by UBC promoter. This constitutive part consists of Bi-cistronic element (IRES) and rtTA3 element following Puromycin drug selection. The second part if inducible, which only can be activated by supplying the medium with Doxycycline, contains RFP and shRNAmir derived by TRE promoter. Abbreviation: IRES, internal ribosome entry elements; WRE, woodchuck response element; LTR, long terminal repeats; puro, puromycin.

The lentiviral miRNA constructs targeted against Hmgn2 and Hmgn1 were purchased from Open Biosystems in the constitutive pGIPZ vector system, and cloned into the inducible pTRIPZ vector system. (pTRPZ kindly provided by Dr. Peter Adams' lab). in addition, the shRNAs that were used in previous vector pSLIK –CD4 were cloned into empty pTRIPZ vector. The control virus used was pGIPZ and pTRIPZ containing a shRNA toward turboGFP and turboRFP (open biosystems).

2.2.3 Cloning reagents and purification kits

2.2.3.1 Gateway cloning technology for pSLIK

2.2.3.1.1 Introduction to Gateway cloning technology

In order to study the structure and the function of genes, having those gene cloned into expression vector is essential. Gateway cloning reactions are based on recombination reaction, which depends on the integration and excision of lambda bacteriophages. Integration is reversible reaction and it involves two enzymes: BP clonase and LR clonase. BP stands for attB (PCR product or expression clone) × attP (donor vector) recombination whereas LR stands for attL (entry clone) × attR (destination vector) recombination (Figure 3.2)

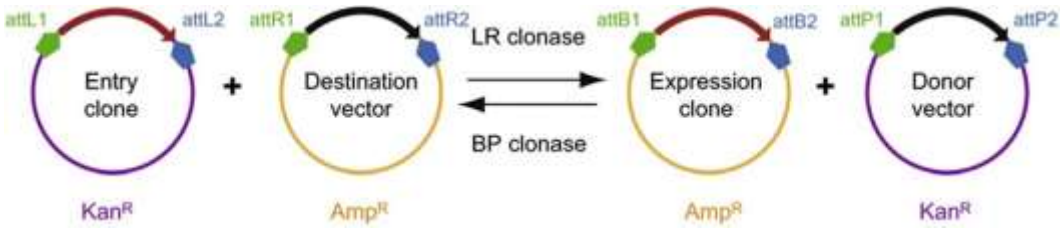


Figure 2.4 the Invitrogen Gateway recombination cloning technology

Briefly, BP clonase reaction is mediated by phage protein integrase (Int) and bacterial protein integration host factor (IHF) (Walhout et al., 2000). This reaction uses the lysogenic pathway of lambda bacteriophages to create the attL and attR reaction. On the other hand, LR clonase reaction, Invitrogen kit (cat no, 11828-029), uses the lytic pathway of lambda bacteriophages. They have been used for the cloning of recombinant DNA which can be activated using enzymes binding to specific sequence called (att sites) (Figure 3.1) . The lytic pathway can be activated by recombination pathway of bacteriophage lambda Int, excisionase (Xis) protein, integration host factor (IHF) protein and protein integrase (life technology manual sheet version 2011 and Walhout et al, 2000).

The advantages of applying the Gateway technology is that it is very flexible due to the fast and easy way to transfer the gene of interest from one

vector to another. The reaction is Very fast reaction and under time management for less than an hour at room temperature. Moreover, this reaction does not required restriction enzyme, ligation, gel purification or re-sequencing. Furthermore, the reaction maintains the open reading frame of the E-coli genome. It uses site-specific recombinase to recombine the DNA plasmids. Gateway technology can clone multi-site to assemble many reporter transgenes (Walhout et al, 2000 and Spitzer et al, 2013).

2.2.3.1.2 Gateway LR clonase reaction process

LR ClonaseTM II enzyme (Cat. No. 11791-020) mix is supplied as a 5X solution and 100 ng (2 µl) of pENTRTM-gus was used as control for the reaction. (50-150 ng) from Entry clone and 150 ng/µl from destination vector were added to 1 µl of TE buffer under pH of 8.0. Then LR clonaseTM II enzyme was thawed and vortex on ice for two minutes. Next each mixture of destination and entry vector, a 2 µl of LR clonaseTM II enzyme was added and mixed with reaction by briefly vortexing or microcentrifuge for two times. After using LR clonaseTM II enzyme mix, turning the LR clonaseTM II enzyme mix to -20° or -80° degree would maintain the efficiency of the enzyme. The reaction should be incubated at 25° degree for 1 hour following an addition of 1 µl of the proteinase K solution to each sample to stop the reaction. Finally, vortex the reaction briefly would ensure the homogeneity of the reaction following an incubation sample for 10 minutes at 37°C.

2.2.3.1.3 Transformation

After LR clonase reaction occurs between entry and destination vectors, transforming 1 µl this reaction is applied through 50 µl of phage resistant cells (One Shot® OmniMAXTM 2 T1 Phage-Resistant Cells (Catalog no. C8540-03)). The reaction with phage cells were Incubated on ice for 30 minutes; then, Heat-shock used for the cells by incubating at 42°C for 30 seconds. 250 µl of S.O.C. Medium was added into the cells and they were incubate at 37°C for 1 hour under shaking machine. 20 µl and 100 µl of each transformation were plated into selective plates.

It is important to note that any competent cells with a transformation efficiency of >1.0 × 10^8 transformants/µg may be used for clone screening. Moreover, 1 µl of pUC19 DNA (10 ng/ml) was transformed into 50 µl of One Shot® OmniMAXTM 2 T1 Phage-Resistant Cells as transformation control as described above. 20 µl and 100 µl were plated on LB plates containing 100 µg/ml ampicillin. In order to generate pSLIK-shRNAmir30-CD4 and pSLIK-shRNAmir30-hygromycin, the experimental strategy from life technology manufacture explained the process (Figure 3.6, B).

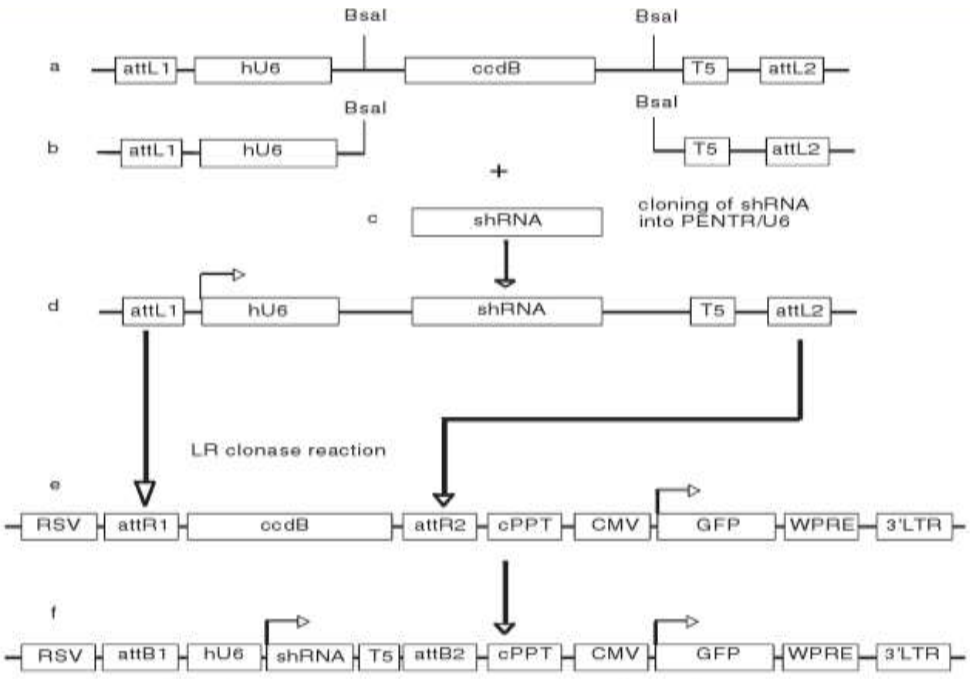


Figure 2.5 Diagram shows from a-f the development of the expression clone system using LR reaction from gateway technology

A. Is a portion of destination vector contains ccdB which required for LR gateway reaction. This destination vectors contains two genes ccdB and chloramphenicol (CmR) flanked by attL sequences. Expression ccdB gene is lethal to cells that carry this vector. The reason is that ccdB gene encodes for a toxin targeting DNA gyrase. B is shRNA cloning into ccdB site using restriction enzyme BsaI. C is ligation and D. PCR amplification and check sequences are preformed to generate a destination vector contains shRNA and attL sequences for gateway cloning reaction. E. LR clonase reaction F. generating the entry vector with shRNA using LR clonase reaction.

2.2.3.1.4 Bacterial transformation

Transformation reactions use bacterial cells to amplify the DNA of interest and the competent cells (DH5alpha) which are able to take up foreign DNA

and acquire the new genetic information. DH5alpha is an ordinary E. coli but it requires treatment with calcium ions to be made into a competent cell. This cell has a very fragile cell wall and must be handled gently. During the transformation reaction the competent cells can be combined with ligation product and incubated on ice. DNA sticks to the outer cell walls, the heat shocked membranes become more porous and allow DNA to enter but not all the competent cells will take in the DNA, so the frequency will be determined. Then, the cells are incubated in LB broth at 37 degree centigrade overnight and this will allow transcription and translation of ampicillin resistance genes. PUC19, which is used as positive control, is a genetically engineered plasmid. It has a 2.7 kb and circular non-genomic DNA. The source of PUC19 plasmid is PBR322, but PUC19 differs from PBR322 by two point mutations. The chromosomal genotype of PUC19 is: fhua2Δ (argF-lacZ) U169phoAglnV44Q80Δ (lacZ) M15gyrA96recA1relA1endA1thi-1hsdR17 (figure 8).

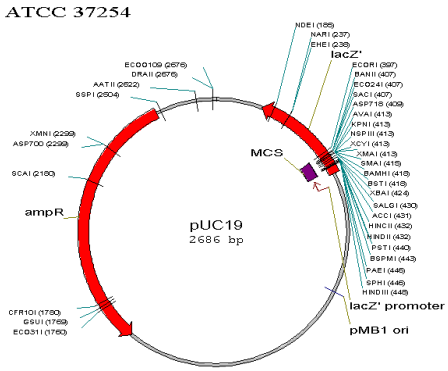


Figure 2.6 the PUC19 vector map
Adapted from:
http://www.atcc.org/ATCCAdvancedCatalogSearch/ProductDetails/tabid/452/Default.aspx?ATCCNum=37254&Template=vectors

PUC19 contains: (1) PMB1 replicon responsible for replication of the plasmid. The high copy number of PUC plasmid is a result of the lack of the rop gene and a single point mutation in rep of PMB1; (2) bla gene, coding for beta lactamase that confers resistance to ampicillin; (3) region of E-coli

operon lac containing CAP protein binding site, promoter place, lac repressor binding site and 5'-terminal part of the lacZ gene encoding the N-terminal fragment of beta-galactosidase. The LacZ gene is a part of the lac operon, which encodes the beta-galactosidase enzyme that breaks down lactose into glucose and galactose. The polylinker cloning site (MCS) within the lacz gene contains a recognition sequence for several restriction enzymes and inactivates the N-terminal fragments of beta-galactosidase and abolishes alpha-complementation. The genotype of PUC19 showed the most useful mutation. These mutations were:

- EndA1 mutation which can inactivate an intracellular endonuclease that degrades plasmid of DNA in many miniprep methods (plasmid purification DNA).
- HsdR17 mutation that can eliminate the restriction endonuclease of the ECOR1 restriction modification system so that methylated DNA from LacZ will not be degraded.
- Δ(LacZ)M15 is an alpha-acceptor allele which means that proteins with the alpha acceptor inactived dimer are much less stable as measured by their ability to act as alpha acceptor after heat treatment. Therefore, it is needed for the blue-white screening for recombinant cells;
- RecA1 can eliminate homologous recombination. This can make the strain sticky but reduces deletion formation and plasmid mutimerization; in other words, it reduces homologous recombination for a more stable insert;
- GlnV44 is an amber suppressor. This means that tRNA suppressor is often caused by changes in the anti-codon loop of tRNA to allow recognition of a stop codon. Due to the wobble rules, the amber suppressor is specific for amber (UAG) codons.

The possible products, which may result from the transformation reaction are: plasmid with insert but no functional LacZ; plasmid without insert but with functional LacZ, or no plasmid as well as no LacZ gene. After the plates were left overnight in a 37 degree Centigrade incubator, the calculation transformation efficiency was performed following this equation provided by

using the Subcloning efficiency DH5alpha kit:

Transformation efficiency (# transformants/ug DNA) = # of colonies / pg PUC19 DNA * 10(6) PG/ug * volume of transformants/ Xul plated * dilution factor

2.2.3.1.5 Plasmid purification

Plasmid DNA purification from E. coli is a core technique for molecular cloning.Small scale purification (Miniprep) from less than 5 ml of bacterial culture is a quick way for clone verification or DNA isolation, followed by further enzymatic reactions (polymerase chain reaction and restriction enzyme digestion). Here, we video recorded the general procedures of Miniprep through the QIAGEN QIAprep 8 Miniprep Kit, aiming to introducing this highly efficient technique to general beginners for molecular biology techniques. The whole procedure is based on alkaline lysis of E. coli cells followed by adsorption of DNA onto silica in the presence of high salt. It consists of three steps: 1) preparation and clearing of a bacterial lysate, 2) adsorption of DNA onto the QIAprep membrane and 3) washing and elution of plasmid DNA. All steps are performed without the use of phenol, chloroform, caesium chloride, ethidium bromide, and without alcohol precipitation. It usually takes less than 2 hours to finish the entire procedure, screening the purified DNA using restriction digest enzymes and using the plasmid editor to show the size of expected band for each plasmid.

2.2.3.1.6 Checking the plasmid clone using Electrophoresis reagents and restriction enzymes

Gel electrophoresis experiment is very important step to check the plasmid restriction site. The samples were prepared as following: First, the total digest each sample is 40 µl divided into: 1 µl of enzyme (Kpn1 and ECOR1), 4 µl of the expected buffer, 1 µl BSA (depend on the enzyme reaction), 1-2 µg of DNA, and the rest of 40 µl amount was filled with distilled water. 1% gel in 120 ml buffer TAE 1x was used to prepare the gel (agarose gel 1.2 g in

TAE buffer 1x =120 ml). The gel was poured into 100 ml plate and 2 µl of dye was added to the gel to visuals the expected band. In addition, the sample was prepared by adding 4 µl of bromophenol (BPB), which can digest lower than 300 bp.. Loading the samples beside 1 kb and 100 bp of DNA ladder.

Note:

For restriction sites:

Use APE plasmid editors to estimate the digest size.

Then, you add buffer 1 = 4ul , BSA= 1ul , DNA x (1ug-2ug) , KpN1 or/ ECORI =1ul, H2O= x in 40ul total amount.

2.2.4 Lentivirus protocol

2.2.4.1 The Production of lentivirus vectors for pSLIK

For viral transfection, 293T/17 cell line was used and it was plated in 10cm dish. The transfection divided into three parts:

The first part is that plating the 293T/17 cell line at passage 2 into 10cm dish as passage 0 before the day of transfection. Lipofectamine 2000 (invitrogen; cat no, 11668-019) was used as transfection reagent, which is a Cationic lipid, and diluted in Opti-MEM® I Reduced Serum Medium (Cat. no. 31985-062). Following the lipofectmine protocol and the amount of packaging vectors (pMCL/pRRE (7.5ug), pRSV-REV (7.5ug), and pCMV-VSV-g (5ug) from the literature (shin et al., 2006) as well as 10ug of recombinant vector pSLIK-shRNA-HMGN-miR30. Then the mixture of media, DNA and packaging vectors were added into 293T/17 cells with ultraculture media (Lonza cat no g1234) and it had been left for 48 hours as protocol recommend (figure: 4, part1).

The second part is the freezing and transducing the target cells. The viral supernatant within the media was centrifuged for 10 minutes at 13,000 rpm to remove any cell from the media. Then, the supernatant was filtered via minisart syringe filter. Finally, aliquot 1ml of the lentivirus particles in eppendorf tube and apply a snap freezing by throwing the tubes into liquid

nitrogen and freeze them down in -80 or transduced target cells with fresh lentivirus particles.

The third part of the lentivirus process is optimizing the viral particles dilution factor. Making a dilution of viral particles on 96 well plates then the viral particles diluted was transferred to 24 well plates. These 24 well plate contained target cells (P19 cell lines) and another 24 well plate with 293T/17 cell to compare the viral particles amounts. The main purpose of making this step is to measure the MOI (multiplicity of infection) of the target cells. As it showed on the scheme after 4-hour incubation with only the virus particles the 24 wells was completed with normal DMEM for 293T/17 cell lines and alpha-MEM for P19 cell lines media. Then after 24 hr., the cells were observed under the florescent microscope to detect the transduced cells (green cells) from non-transduced cells.

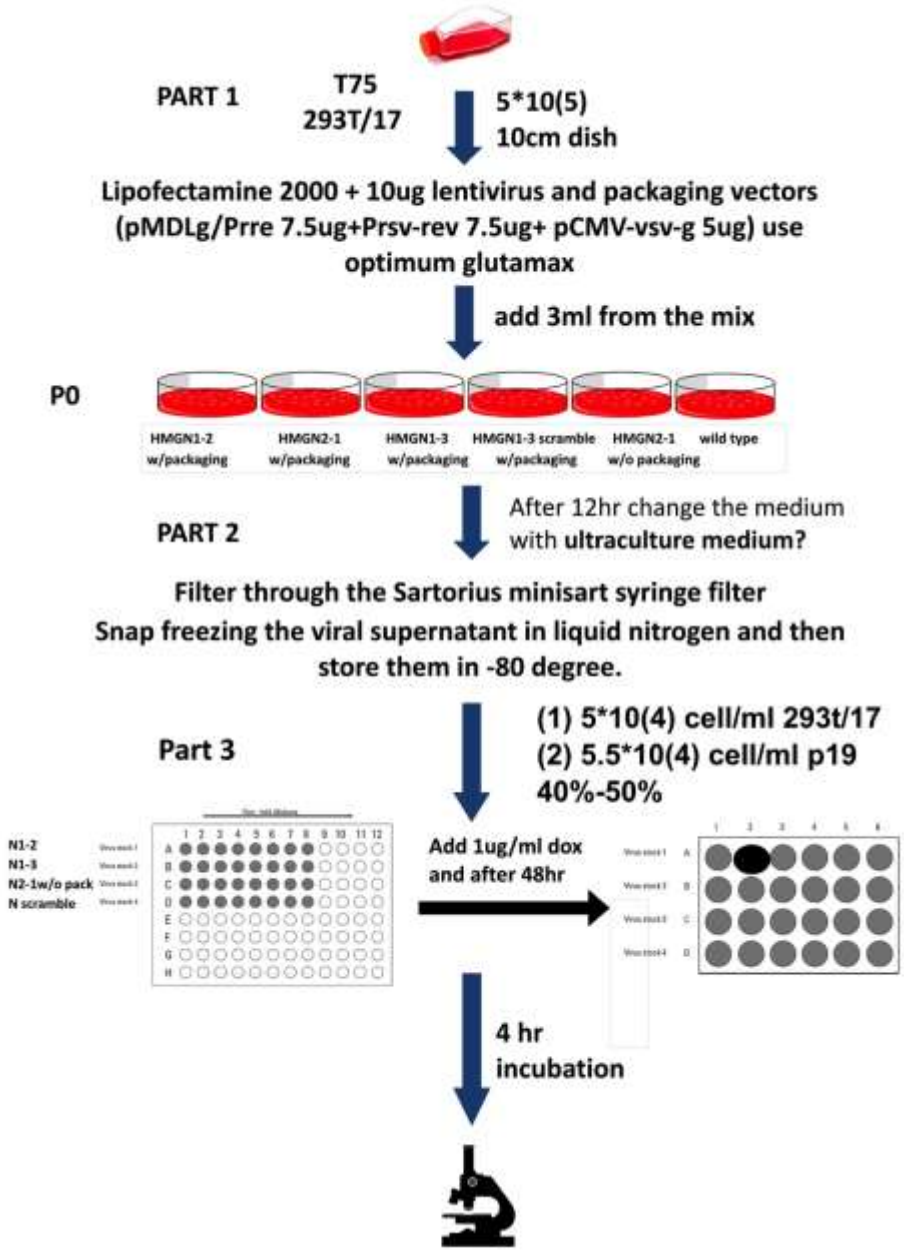


Figure 2.7 generating the lentiviral particles

Part1 shows the viral production preparation and the amount of packaging and recombinant vectors. Part2 shows the generating the viral particles and freezing down the viral particles.

2.2.4.2 The Production of Lentivirus vectors for pGIPZ and pTRIPZ

2.2.4.2.1 Structure of the viral vectors

pGIPZ: The short hairpin RNA constructs are expressed as human microRNA-30 (miR30) primary transcripts. All vectors contain the turboGFP coding sequence located in the middle of the lentiviral vector. miRNA and tGFP sequences are driven by human cytomegalovirus (hCMV) promoter and terminates using the 5` long terminal repeat (5`LTR). Downstream of tGFP is woodchuck hepatitis virus regulatory element (WPRE) that enhances the expression of transgene. pGIPZ also expressed the gene for puromycin resistance.

pTRIPZ: The human ubiquitin C promoter (UBC) drives constitutive expression of reverse tetracycline-transactivator 3 (rtTA3). Addition of doxycycline activates rtTA3, which activates the tetracycline-responsive element (TRE) promoter that drives miRNA and turboRFP expression. The puromycin resistance gene is consitutively expressed.

2.2.4.2.2 Virus production

The recombinant Lentivirus was produced by cotransfection of vectors and calcium phosphate (TLP5913) reagents in HEK293T/17 cells. The component of the translentiviral pGIPZ and pTRIPZ packaging systems was designed to enhance the biosafety and increase the level of shRNA expression. This packaging system contains pTLA1-Pak (packaging protein), pTLA1-enz (viral particles), pTLA1-Env (envelop protein), pTLA1-Rev (transporter protein) and pTLA1-TOFF (transactivator protein binds to TRE promoter).

The first Lentivirus experiment was done on 9/05/2012 using open biosystems protocol on 6 well plates.

2.2.4.2.3 Producing the Lentiviral particles in HEK293T/17 cells

The experiment begins using the human embryonic kidney 293 cells from clone 17 (HEK293T/17). These cells contain the SV40 T antigen which is important for increasing the efficiency of Lentivirus production. When Calcium chloride combined with DNA and co-precipitated by adding the

phosphate buffer, it can increase the uptake of DNA in transformed, 293T/17 cells.

A. Cell plating:

1.2*10(6) cell/well 293T/17 cells were prepared in one well of each 6-well plate with 2ml full media with serum for each transfer vector to be packaged into Lentiviral particles. Then, the cells were Incubated at 37°C with 5% CO2 overnight.

B. Transfection

CaCl and 2X HBSS were brought into a room temperature after thawing them for transfection step.

Each well of a 6-well plate was already indicated the quantity of DNA plasmid and Trans-Lentiviral packaging mix in 15 ml polystyrene tube. Sterile water was used to bring DNA mix to the indicated total volume as it showed in the table.

	Lentivirus Transfer Vector DNA (shRNA or ORF)	Trans-Lentiviral Packaging Mix	Total Volume (With sterile water)
One well of a 6-well plate	8µg	4.3µL	To 135µL

The volume of CaCl was indicated for each 6 well plate by adding 15ul into the DNA assembled above. The tubes were vortexed at a speed sufficient to thoroughly mix reagents without spillover. While running the sample into the vortex, a drop wise of indicated 2X HBSS was added 150 µL in One well of a 6-well plate. The tubes were incubated at room temperature for 3 minutes. A chalky precipitate should appear with time. The total volume is 300 µL of transfection mix was added drop-wise into the cells and the cell was checked with medium once you add a drop of the Lentivirus mix.

Cells were incubated at 37 C with 5% CO2 for 10-16 hours. After 16 hours of incubation, examine the cells microscopically for the presence of a fluorescent reporter protein, such as TurboGFP.

The media was changed to ultracultural- serum free media from lonza (BE12-725F). The cells with serum free media were left from 36-48 hrs. The supernatant was collected to proceed the Lentivirus concentration using Millipore (Amicon ultra-15 centrifugal filter device).

C. Viral Particle Collection and Concentration

After 48 hours incubation time with ultraculture medium the supernatant was taken from the host cell (293T/17). The supernatant was centrifuged at 8000 rpm for 15 min. Then, the supernatant was filtered via 0.22 or 0.45 um ministrant filter (low protein binding). To obtain high transfection efficiency for the target cells (P19 EC cells) several experiments were performed:

Amicon ultra-15 centrifugal filter device

For concentration, the filtered supernatant was poured into an Amicon ultra-15 centrifugal filter device (Millipore).The device was centrifuged at 40,000 rpm for 20min at 4 degree.

Lenti-x Geostix

Following the manufacture protocol, 20 µl was taken from prepared lentiviral supernatant and was applied to the sample well of the GoStix cassette. Next, 3 drops of Chase Buffer was added to the sample well. Then a few moments was spent in incubation for the bands to develop. a test band called (T) will start to appear within 5 min and reach maximum intensity at 10 min if the target sample contains sufficient lentivirus.

Transdux and Polyberen

To enhance the transfection reagents, transfection reagents called polyberen and transdux was used on P19, E14 and 293T/17 cells that will be transduced with the pGIPZ- GAPDH (positive control) lentivirus system and pSLIK-hygromicine (N-scramble) (negative control). Figure 2 showed a lentivirus serial dilution on P19, E14 and 293t/17 cells with set of polyberen and transdux on each well. Sadly there were no GFP expressed on neither 293T/19 nor P19 and E14. Moreover, this experiments was used the pTRIPZ

(inducible system) after 72hrs doxycycline treatment but did not show any expression of RFP cells.

Following the manufacture protocol from system bioscience after viral particle collection day and before the transduction day, 50,000 cells per well were plated in a 24 well plate in cell culture medium. The next day Cells should be between 50 to 70% confluent. 7*10(4) cell/well 293t/17 and 8*10(4) cell/well p19. Medium was aspirated from cells. culture medium with TransDux or Polyberen were combined to a 1X final concentration.for example, 2.5 µl of TransDux was added to 500 µl culture medium and then transfer to each well. The virus was added to each well and swirl to mix.Optional step, increasing amounts of virus was added to different wells at varying MOIs (5, 10 and 20, etc.) to optimize the transduction.

Applying the Transdux protocol from system bioscience on p19 and 293t/17 failed to work using neither GIPZ nor TRIPZ with transfection reagents. However, 293t/17 worked well after 24 hr transfection using GIPZ (LV) in both transfection reagents whereas the TRIPZ (LV) with transfection reagents on 293t/17 did not work even after 72 hr with Dox 1ug/ml treatments.

Lenti-X concentrator

L.V containing supernatant pool similar stocks was harvested and Centrifuged briefly for 500x for 10 min or filter via 0.45 um filter. Then, clarified supernatant was transferred to a sterile container and combines 1 volume of Lenti-x concentrator with 3 volumes of clarified supernatant. Next, these supernatant was gently mixed by inversion and larger volumes may be accommodated via the use of larger centrifuge tubes. the mix was incubated at 4 C for 1 hr (but if the amount >100 ml make it O/N) and it was Centrifuged at 1500 X g for 45 min at 4 degree. After the centrifugation, the sample will be visible as an off white pellet. Then, supernatant was removed carefully without disruption the pellet. The supernatant can be removed by either pipette tip or by brief centrifugation at 1500x g. the pellet was re-suspended in 1/10 to 1/100 of the original volume using complete DMEM or PBS or TNE (note: Recommended to be in ultraculture (serum free media). Titre immediately the sample or store at -

70 degree in single use aliquots. pGIPZ and pTRIPZ are the target cells were plated into 12 well plate. A serial dilution 10x was made and was added into the 12 plates from column 2 to 4 the first column was added nude virus 500ul for 24 hr and then adds to normal media with full serum. You add 500ul normal media and 500ul from LV supernatant. 1/10 – 1/100 – 1/1000 – 1/10,000.

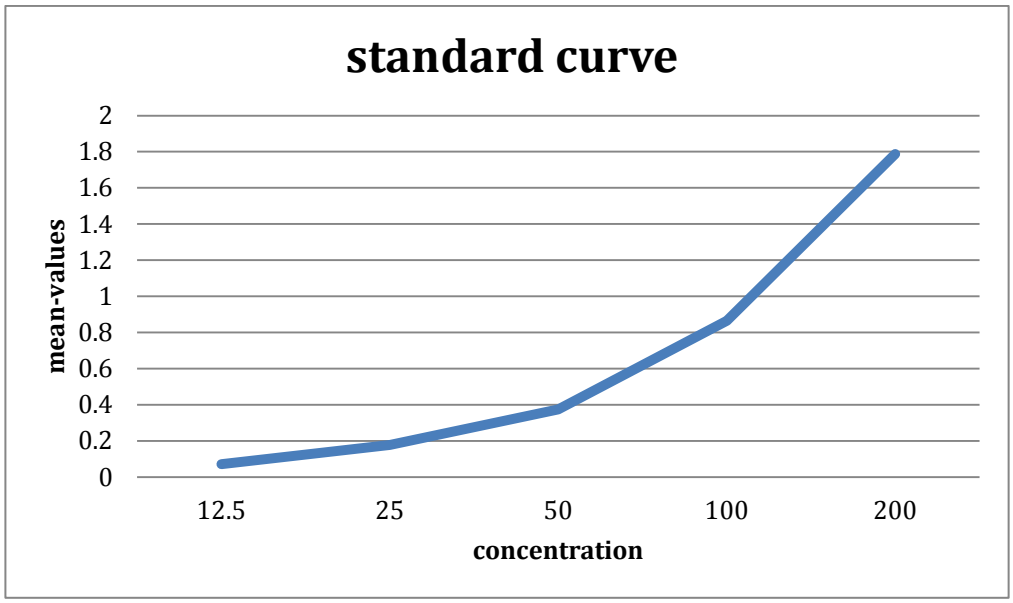
Lenti-X™ p24 Rapid Titer Kit

To ensure that the lentivirus have a high titration and can be efficient for any cells types. A titration kit from clone tech was used to measure the titration of lentivirus particles using ELISA machine. Following the manufacture protocol, first step is to Preparing Dilutions for the p24 Standard Curve:

Samples were tested quantitatively and determined accurate virus titers, the test sample was needed to prepare a p24 standard curve (0–200 pg/ml). A working strength p24 positive control stock solution was prepared by diluting 20 µl of the p24 Control (10 ng/1. ml) into 980 µl of fresh complete tissue culture medium (e.g. DMEM containing 10% FBS), for a 1:50 dilution. 200 pg/ml was generated as stock solution. 200 pg/ml stock was used and tissue culture medium was completed as the diluent. A series of four additional standard dilutions of 100, 50, 25, and 12.5 pg/ml was made and 500 µl of media was dispensed into each of four labelled tubes. 500 µl of the 200 pg/ml stock was added into the 100 pg/ml tube, mix, and using a fresh pipet tip, 500 µl of this 100 pg/ml solution was transferred into the 50 pg/ml tube and mix. Similar transfers were prepared similarly for the 25 and 12.5 pg/ml tubes.

Sample data for standard curve

Standard (pg/ml)	Absorbance (450nm)				
	A	B	Mean value	Standard dev.	CV%
12.5	0.082	0.061	0.072	0.015	20.6
25	0.154	0.201	0.178	0.033	18.5
50	0.395	0.354	0.374	0.030	7.9
100	0.817	0.913	0.865	0.069	7.9
200	1.429	2.146	1.787	0.507	28.4



Second step of the protocol is to Assaying the Lentiviral Supernatants:

All reagents were allowed to reach room temperature (18-25°C). A sufficient number of 8-well strips were selected to accommodate all standards, test specimens, controls, and complete culture medium blanks (negative controls) in duplicate. 20 µl of lysis was dispensed buffer into each well. 200 µl of each standard curve dilution was dispensed with supernatant sample, and culture medium into appropriately labeled duplicate wells. The samples were 37(±1) °C for 60 (±5) minutes. sample contents of the wells was removed , and washed the microtiter plate 100 µl of Anti-p24 (Biotin conjugate) detector antibody was distributed into each well. The sample was Incubated at 37(±1) °C for 60 (±5) minutes. The detector antibody was removed from the wells, and washed the microtiter plate. 100 µl of Streptavidin-HRP conjugate was added into each well. The samples was Incubated at room temperature (18-25°C) for 30 (±5) minutes. The conjugate was removed from the wells, and washed the microtiter plate. 100 µl of Substrate Solution was immediately distributed into each well with multichannel pipette. Plate was protected from direct sunlight, and incubated at room temperature (18-25°C) for 20 (±2) minutes. The reaction was stopped by adding 100 µl of Stop Solution to each well including the culture medium blanks. Solution color will change to a uniform yellow color without air bubbles and the content dry. After adding the Stop

Solution, the absorbance values were detected at 450 nm using a microtiter plate reader blanked on the negative control well.

Table 1 showed lentivirus particles that were used and results of each sample at wavelength 450 as protocol suggested. Some of the samples were concentrated first before measuring the titration. The results, which were identified with ELISA machine, did not show any low titration values, but it gave no results because the values were higher than the wavelength level.

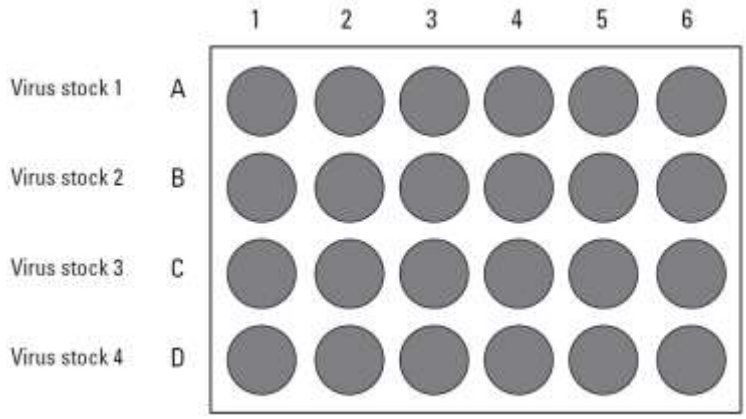
pTRIPZ , pSLIK-hygromycine and pGIPZ titration kit					
Sample	Values	Results	Mean-results	Stand. Dev.	CV%
pTRIPZ-GAPDH	3.544	390.437	390.437	0	0
pGIPZ-GAPDH	3.544	390.437	390.437	0	0
pTRIPZ-N2-1	3.544	390.437	390.437	0	0
pGIPZ-N2	3.544	390.437	390.437	0	0
pTRIPZ-Nscram	0.419	54.505	54.505	0	0
pTRIPZ-Non-sil	3.544	390.437	390.437	0	0
pSLIK-HYGRO-Nscram	3.544	390.437	390.437	0	0
pGIPZ-EG1	3.544	390.437	390.437	0	0

Table 2.3 measuring the titration of Lentivirus virus using ELISA machine

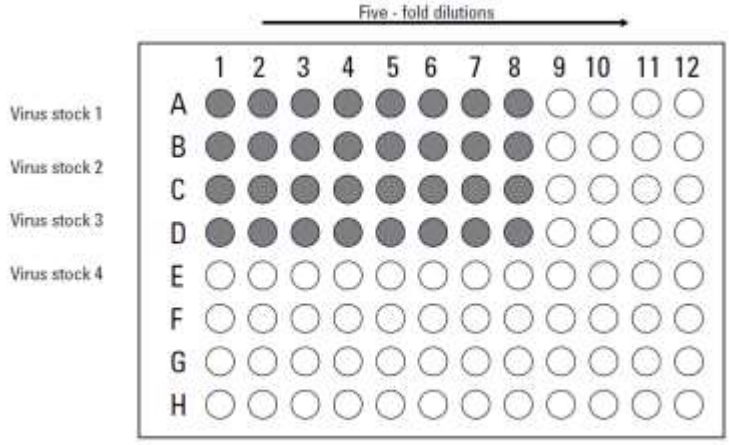
Some of The samples were concentrated using the concentrator from clone-tech.

D. Determining the titer of Lentivirus particles stocks and transduced target cells

24-well tissue culture plate of HEK 293T cells at 5 x 104 cells per well in DMEM (10% FBS, 1% pen-strep) and another 24 well plate of P19 at 3*10(4) cells/well in full media with serum were seeded the day before the transduction under cell density from 40-50% confluent.



The dilutions of the viral stock in a round bottom 96-well plate using serum-free media were made at the same time of viral collection. A series of 5-fold dilutions was made to reach to the final dilution of 390625-fold.



Each well an 80 µL was added with serum-free media. 20 µL of thawed viral particles stock was added to each corresponding well in column 1 (nude Lentivirus without dilution) the contents was mixed up and down 10-15 times. Pipette tip was discarded. With new pipette tips, 20 µL was transferred from each well of column 2 to the corresponding well in column 3 then from 3 to 4 . Mixed 10-15 times and discard pipette tips each time you move to new well. At the 24 well plate for P19 and 293T/17 Remove culture media from the cells in the 24-well plate. 225 µL was added with serum-free media to each well. The target cells P19 and 293T/17 cells were transduced by adding 25 µL of diluted viral particles from the original 96-well plate to a well on the 24-well destination plate containing the cells. Transduced cultures were incubated at 37 °C for 4 hours. The transduction mix was removed from the culture and 1 mL of full serum media was added.

Cells were left in the incubation for 48 hours. Then the GFP was detecting using the florescent microscope.

2.2.5 Cell lines, cell medium and reagents

P19 cells were cultured in alpha-MEM with UltraGlutamine™ I, deoxyribonucleosides and ribonucleosides (Lonza BE02-002F), supplemented with 7.5% newborn calf serum (PAA B15-102) and 2.5% tetracycline-negative fetal bovine serum (A15-109) 293T/17 and Hepa cells were cultured in DMEM with glutamax and 10% fetal bovine serum.

Cell lines used in this study were a mouse embryonic carcinoma cell line, P19 and a derivative from human embryonic kidney cell line, 293T/17. P19 was maintained at 37°C in Minimum Essential Medium Eagle - Alpha Modification (Alpha MEM) with Nucleosides (Alpha MEM Eagle; w/o L-Glutamine, Deoxyribonucleosides or Ribonucleosides; Quantity: 500mL) containing 2,5% Fetal Bovine Serum (FBS) from Sigma-Aldrich and 7,5% new born calf bovine serum (NBCBS) from Lonza in a 5% CO2 incubator. Another cell type, HEK293T/17, was maintained at 37°C in Dulbecco's Modified Eagle Medium (DMEM) containing 2 mM glutamine, 10% Fetal Bovine Serum (FBS) similarly at 5% CO2.

2.2.6 Drug reagents

2.2.6.1 Puromycin kill curve

Puromycin (sigma cat number P7255) was added into wild type P19 EC cells for 5 days to detect the amount of puromycin needed to remove all the wild type cells. On day 0, P19 EC cells were plated around 5 - 8 x 10^4 cells per well in all 24-well plates which are sufficient for adding several concentration of puromycin in each well. Next days cell medium was prepared and added into falcon tube supplemented with various amount of puromycin ranging from 0 - 22 µg/ml of 10 mg/ml puromycin stock. Then from day 1 to day 5 a medium was changed with new puromycin supplement and some cells needs to split during these days.

Therefore, Monitor the cells daily and observe the percentage of surviving cells is critical to detect the right puromycin concentration. Thus, The minimum antibiotic concentration to use is the lowest concentration that kills 100% of the cells in 1 - 5 days from the start of antibiotic selection.

2.2.6.2 Hygromycin B kill curve

Hygromycin B (sigma cat number H3274) was used for pSLIK-hygro plasmids. The kill curve process is similar to puromycin. In day 0, 2 x 10^5 cells per well was plated in six 10-cm tissue culture dishes containing 10 ml of the appropriate complete medium. Next day the medium was changed and supplemented with hygromcyin B concentration (0, 50, 100, 200, 400, and 800 µg/ml). Then. the cells Incubated for 10–14 days, replacing the selective medium every four days (or more often if necessary). During this period, the cells were examined for viable cells every two days.

2.2.6.3 Toxilight bioassay kit from Lonza

Doxycycline amount was examined for cell cytotoxicity in wild type cells although 1 µg/ml of Doxycycline is the optimal concentration for most of the cell lines and the doxycycline half-life in culture is 24 hours as previous studies suggest. Therefore, it is advisable to add doxycycline amount in the daily bases to maintain the doxycycline effects constant and avoid the fluctuation of shRNA expression during the knockdown process.
Following the protocol from Lonza Vialight plus kit LT27-221 500 tests on P19 EC cells in wild type and knockdown cell lines. The protocol is for adherent cells Luminescence incompatible plate 96 well formats. To begin with, all reagents were prepared at the room temperature before using. Then, the Reconstitute the AMR plus in assay buffer was left for 15 minutes at room temperature to ensure complete rehydration. Next, the culture plate was removed from the incubator and allowed it to cool to room temperature for at least 5 minutes. The luminometer program took a 1 second integrated reading of each appropriate well. Moreover, 50 µl of cell lysis reagent was added to each well and waiting for at least 10 minutes.

Furthermore, 100 µl of cell lysate was transferred to a white walled luminometer plate. 100 µl of AMR plus was added to each appropriate well and the plate was incubated for 2 minutes at room temperature. Finally, The Plate was placed in luminometer and initiate the program. (Luminoskan Ascent 2.5)

2.2.7 Molecular biology reagents

2.2.7.1 Trizol extraction

Trizol reagent (cat number: 15596-018) from Ambion was used to extract the RNA. The cells were suspended and centrifuge, then the supernatant was removed. After that, a 1 ml of the Trizol reagent was added in each cell pellet (approximately $5x10^6$ cell). To homogenize the cells with trizol reagent should be Pipetting up and down the mixture.

A tube, which contained 1 ml of trizol, was used as negative control to ensure no contamination was detected during the extraction procedure. The samples were incubated at room temperature for 5 min then they were stored at -70 or -20 c for 1-2 hours. 100 µl BCP (Sigma B9673-200 mL) was added per ml of trizol reagent in each sample. The sample's lid was held down and the tube was shaken for 30-40 second without vortex. The sample was incubated at room temperature for 5-15 min. Then it was centrifuged at 12000 g for 15 min at 4°C. A fresh set of labelled eppendorf tubes were labelled and 500 µl of the isopropanol (19516-500 ml) from sigma were dispensed in each tube. The colourless solution was transferred to labelled tube with isopropanol. Next, The samples were vortex and were incubated at room temperature for 5-10 min and the samples were incubated at -20 c for 20 minute. Then the samples were centrifuged at 12,000 g for 15 min at 4°C.

The supernatant was completely removed and touched with clean tissue paper to clean the remaining supernatant. 1 ml of 75% of ethanol was added into each sample's pellet and it was centrifuged for 5 min at 7500 g at 4°C. Then the supernatant was removed. This step was repeated twice. The sample was briefly centrifuged at 12,000 g to remove all the remaining

ethanol from the sample pellet without disrupting the pallet. The tubes were left the lid open and upside down to the tissue paper to allow the pellet to dry for 3-5 min and then it was heated at 30 °C on heat block for 3 min. The pellet was suspended and in 50 µl of RNA free water and it was left at RT for 2 min.

2.2.7.2 RNeasy extraction

RNA was extracted using the Qiagen Rneasy kit with a QIAshredder column and on-column Dnase digestion, following the manufacturers' instructions. RLT buffer was added to cells in a 60 mm dish and RNA sample was extracted from transduced cells using Qiagen RNase extraction. The cells were harvested from 60 mm dish, the media was removed and wash with DPBs. Next, Cells were disrupted by adding Buffer RLT and mixed properly to ensure the lysis homogenized with the cells. To homogenize the lysate, a step was followed using QIAshreddere spin Colom. Lysate was added directly into a QIAshreddere spin column, and it was centrifuged for 2 min at full speed. 1 volume of 70% ethanol was added to the homogenized lysate, and mixed well by pipetting only.

Up to 700 µl of the sample was transferred including any precipitate that the sample may have formed, the sample then was placed to an RNeasy spin column in a 2 ml collection tube. The lid was closed gently, and centrifuged for 15 s at _8000 x g. The flow-through was discarded and the collection tubes were reused If the sample volume exceeds 700 µl. The successive aliquots were centrifuge in the same RNeasy spin column. The flow-through was discarded after each centrifugation. Furthermore, 700 µl Buffer RW1 was added to the RNeasy spin column and the lid was closed gently, and centrifuged for 15 s at _8000 x g (_10,000 rpm) to wash the spin column membrane. The flow-through was discarded. 500 µl Buffer RPE was added to the RNeasy spin column and the lid was closed gently, and centrifuged for 15 s at _8000 x g (_10,000 rpm) to wash the spin column membrane and the flow-through was discarded. The collection tube was reused if it is necessary. 500 µl of Buffer RPE was added to the RNeasy spin column and

the lid was closed gently, and centrifuged for 2 min at _8000 x g (_10,000 rpm) to wash the spin column membrane.

The long centrifugation dries the spin column membrane, ensuring that no ethanol was carried over during RNA elution because residual ethanol may interfere with downstream reactions. The RNeasy spin column was placed in a new 1.5 ml collection tube and 30–50 µl RNase-free water was added directly to the spin column membrane. The lid was closed gently, and centrifuged for 1 min at _8000 x g (_10,000 rpm) to elute the RNA.

2.2.7.3 cDNA and Real time RT- PCR

Total RNA was extracted using methods described in section 2.6.1. The RNA was used in a first strand cDNA synthesis reactions using SuperScript™ III kit (Invitrogen). cDNA synthesis was carried out using oligo(dT)20 primer in order to reverse transcribed polyadenylated transcripts only. The following steps were employed:

Denaturation/ Annealing steps

The following reaction was set up and incubated at 65^{0}C for 5 minutes and then rapidly cooled on ice for 5 minutes.

Reagents	Final volume (µl)
RNA (300 ng in 10 ul dH20)	10
oligo(dT)20 (100 µM)	0.5
dNTPs (50 µM)	0.5
Total	11

Reverse transcription

cDNA synthesis master mix was prepared as below and added to the RNA/primer/dNTPs mix. The mix was incubated first at 25^{0}C for 10 minutes and then at 50^{0}C for 50 minutes. The reaction was heat inactivated at 85^{0}C for 5 minutes.

Reagents	Final volume (µl)
RNA (300 ng in 10 ul dH20)	10
oligo(dT)20 (100 µM)	0.5
dNTPs (50 µM)	0.5
Total	11

RNA digestion

The RT reaction was cooled on ice for 30 seconds. The RNA component was degraded by adding 1 ul of RNase H. The reaction were incubated at 37^{0}C for 20 minutes and then either used for Real-time PCR amplification or stored at -20^{0}C.

2.2.7.4 Real-time PCR amplification

The gene expression profiles studies shown in this project were conducted using SYBR®- green assays. This assay is based on detection of double stranded PCR products formed in the reaction using oligonucleotide primers. The primers for specific genes were design using either Primer Express or Primer 3 software. The basic criteria for primer design included designing primers over exon-intron-junctions, melting temperatures of 55-60° and an amplicon of 100-150 bp were applied. The primer pair sequences were compared against the entire mouse genome using BLAST alignment tool. The concentration of each primer was optimised, using dissociation curves to detect primer dimer formation and standard curves to check PCR efficiency. The list of primer sets used is shown in table 2.3

The cDNA used in real-time PCR amplification was diluted 1:5, and 5 µl of this used in all PCR reactions. The PCR reactions was set up using transparent 96-well PCR plates (Abgene) in 25 µl reactions, typically in triplicates for each sample, by mixing all the reagents shown below. Fast start universal SYBR green master mix (Rox) from Roche was used (cat. no. 04 913 914 001)

SYBR®-green assay	Final volume (µl)
SYBR PCR mix	12
Forward primer (2.5-7.5 µM)	3
Reverse primer (2.5-7.5 µM)	3
dH20	1.5
Template DNA	5

PCR was performed on an MxPro 3000P (4 filter set plate) (Stratagene) using the thermal cycling conditions shown below. All SYBR green readings were normalised to Rox dye fluorescence. A dissociation curve (double stranded PCR product melting curve) was carried out to allow screening for non-specific products amplified.

Segment	Number of cycles	Thermal cycling condition	
1	1	10 minutes	95^{0}C
2	40	15 seconds	95^{0}C
1 minute		60^{0}C	
3	1	1 minute	95^{0}C
30 seconds		55^{0}C	
30 seconds		95^{0}C	

The data was analysed using excel file and the statistical analysis was summarize as following:

CT (cycle threshold) average is used to detect the accumulation of a fluorescent signal. Then, STDEVP (CT) the standard deviation of delta Ct value measures the technical triplicates. The triplicates are valid when the SD is smaller than 0.25. Subtract the Sample average from housekeeping average (avct-hk avct) is applied to Compare the Ct difference for calibrator (GOI minus HK). Then calculate the error is important ERROR1 =SQRT (^{power} (SD-SAMPLE, 2)+^{power} (SD-HK,2)) to check for errors and variation among the

triplicate sample. Delta-delta CT value of ((SAMPLEAVCT-HKAVCT) DDCT-internal (wt)) is used to measure the amplification efficiencies of the reference control gene and the target gene of interest between treated and untreated sample.

Measuring the error bar (ERROR 2 (SQRT (^{power} (ERROR$1,2)+^{power} (ERROR1,2))) represent the range of possible relative of expression (delta-delta CT values) and it defined by the standard error of the delta Ct's. the fold change (POWER(2,-DDCT)) represent the difference in gene expression which means gene expression normalized to an endogenous reference of the gene and relative to the untreated control. Finally ERROR UP (POWER (2, -(DDCT-ERROR1)-FOLD CHANGE)) and ERROR DOWN (FOLD CHANGE-POWER (2, -(DDCT+ERROR1))) are measured to the standard error of delta-delta CTs values.

2.2.8 Protein analysis reagents and kits

2.2.8.1 Protein extraction

2.2.8.1.1 Whole cell extraction

Cells were collected from 10 cm dish and washed twice with cold PBS. 3 ml of cell lysis buffer (CLB) was added and scraped off cells with plastic scraper. The solution was mixed by pipetting into 15 cm tube. cells were freezed and thawed from -20°C. the lysate cells were Heated at 65 °C for 20 min or pass through 25 gaug needle 10 times or sonicate to beak up the chromatin and Spin for 15 min . finally supernatant was added into new tube.

Cell lysis buffer (CLB)

Ingredients	Concentration
4.5ml of 0.5M Tris-HCL pH 7.5 or 8	45mM
100ul of 0.5M EDTA	1mM
5ml of 10% SDS	1%
5ml glycerol	10%
Tiny bit of bromophenol blue	0.01%
Water up to	50ml

10 ml of the following reagents was added to the cells and lysis solution was Freeze at -20°C for 3 month. Adding the 10 ml CLB from 2 mM stock 20 ul of 1M DTT and 400 ul of protease inhibitor cocktail.

2.2.8.1.2 Nuclear cell extraction

Two buffers are used in the nuclear cell extraction

Buffer A	Buffer C
10mM HEPES PH 7.9	20mM HEPES PH7.9
10mM KCL	0.4M NaCL
0.1mM EDTA	1mM EDTA
0.1mM EGTA	1mM EGTA
1mM DTT	1mM DTT
0.5mM AE-BSF	1mM AE-BSF
Total distal water 50ml	Total distal water 50ml

One million cells should yield about 100 ug of protein at about 2 ug/ul at the end of following process:

To begin with cells were washed once with TBS and resuspended in 1 ml of TBS and transfer to eppendorf tube; pellet by quick-spin. Supernatant was aspirated and 400 ul of cold Buffer A was added to the solution and resuspend gently. Supernatant was left on ice for 15 minute to allow cells to swell. And 25 ul of 10% NP-40 was added to the cells. Vortex for 10 second was applied and supernatant used the Centrifugation in the cold at top speed for 30 second. Then supernatant (cytoplasm + RNA) was removed and the pellet (nucleus) was separated and resuspended in 50 ul of ice-cold Buffer C. then the supernatant was on tube for rock or shake in the cold for 15 minutes to solubilize pellet. The supernatant was Centrifuge for 5 minutes in the cold at top speed and removed the supernatant that contain nuclear extract at about 2ug/ul to freeze aliquots at -70 degree.

2.2.8.1.3 Cytoplasmic cell extraction

The Extraction of monolayer cells:

Culture medium was removed from the cells. As an Optional step, If components of the culture media (i.e. phenol red) are inhibitory to protein

analysis, cells should be washed once with PBS (PBS; 43 mM Na2HPO4, 15 mM KH2PO4, 137 mM NaCl, 27 mM KCl, pH 7.4) or Hanks' Buffered Salts Solution (HBSS) prior to Cyto-Buster addition. The recommended amount of Cyto-Buster (from Table 1) was added and allowed to extract at room temperature for 5 min. cell debris were removed using a cell scraper or rubber policeman and the plate was oriented so all debris is pooled in the Cyto-Buster Protein Extraction Reagent. the extraction was removed to a suitably sized tube and centrifuge for 5 min at 16,000 Å~ g (4°C). it is important to note that the Longer centrifugation at higher speeds may be desired for certain applications. Then, the supernatant was removed to a new tube and proceed with analysis. It is important to note that the Extracts prepared with Cyto-Buster can be used immediately or may be stored for extended periods of time at –20 °C or –80 °C.

2.2.8.2 Protein quantification

2.2.8.2.1 Protein quantification using commassie stain

SimplyBlueTM SafeStain

- 1 L (Cat. no. LC6060)

The microwave procedure is fast, takes just 12 minutes, and yields results with sensitivity as low as 5 ng with an additional incubation with a salt solution. The procedure is for 1.0 mm mini-gels. For 1.5 mm mini-gels, use the values in parentheses.

Caution: Use caution while using the stain in a microwave oven. Do not overheat the staining solutions.

After electrophoresis, place the gel in 100 mL of ultrapure water in a loosely covered container and microwave on High (950 to 1,100 watts) for 1 minute until the solution almost boils. Then the gel was on an orbital shaker for 1 minute (2 minutes). The water was discarded and their previous steps were repeated for 2 more times.

After the last wash, 20 mL (30 mL) of SimplyBlueTM SafeStain was added and microwave was applied on High for 45 seconds to 1 minute (1.5 minutes) until the solution almost boils. The gel was an orbital shaker for 5 minutes (10 minutes). Detection limit: 20 ng BSA. The gel was washed in 100

mL of ultrapure water for 10 minutes on a shaker. And the Detection limit: 10 ng BSA. 20 mL of 20% NaCl was added for at least 5 minutes. Detection limit: 5 ng BSA. The gel can be stored for several weeks in the salt solution.

2.2.8.2.2 Protein quantification using using Nano-orang kit from Invitrogen

Protein preparation:

Protein was harvested from the cell culture using lysis buffer as following :

4.5 ml tris-base 0.5M

5ml glycerol

5ml SDS 10%

100ul EDTA 0.5M

0.01% tiny from bromophenol blue

Up to 50 ml d.H2O

400 ul of protease inhibitors cocktail (diluted one tables in 1 ml H2O) and 1 M of DTT were added Into 10 ml of the lysis buffer . Then 3 ml of the mixture was added into each sample. The protein was sonicated at 4.5 W for 3 times in 5 second each and it Quantified using the Nano Orange kit from Invitrogen cat number N6666.

Following the manufacturers' instructions, the standard curve was made Figure 1.5 page 19 beside the samples at each run. Then the qRT-PCR detects the fluorescence when the sample was dropped to the room temperature after heated for 10 min at 95 c.

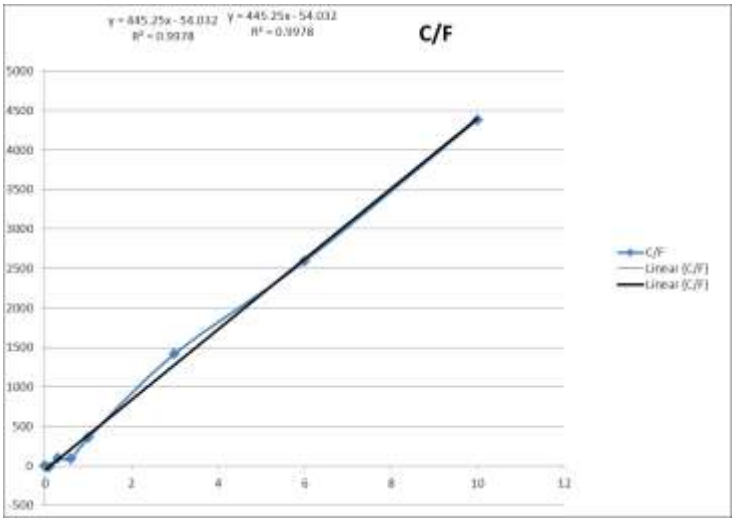


Figure 2.8 Standard curve of the P19 cells transduced with pTRIPZ Lentivirus
The standard curve was made by bovine serum albumin (BSA) as reference standard. The chart showed that BSA at different concentration and it showed the equation of the line and R square amount.

2.2.8.3 Western blot assay

2.2.8.3.1 Western blot assay using manual SDS page from Invitrogen

After protein quantification was assessed, the blotting assay can be performed.

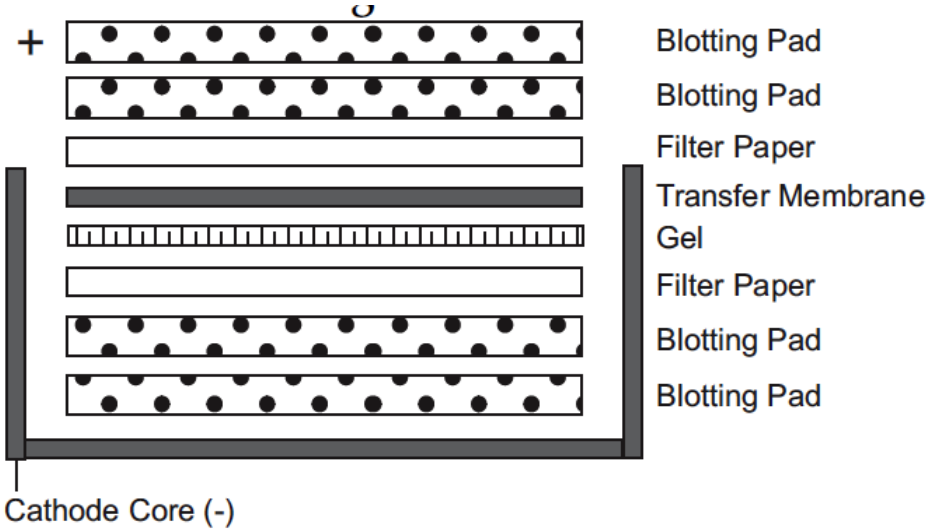
Following the manufacture instruction of NuPAGE electrophoresis system from Invitrogen. The manufacture explained the PVDF transfer membrane method.

To begin with, load the protein after quantification and heated at 70 c for 10 min.

then add reducing agent , antioxidant and loading buffer. run the gel (NuPage 4-12% Bis-triz gel) at 55minute at 120 v and mA 115. During the electrophoresis run prepare the antibody blocking assay buffer (TBST).

Tbs-TEEN 1X.
20 Mm, Triz PH 7.5 20ml 1M Tris ph 7.5
150 Mm. Nacl 30ml 5M Nacl
In 1L WATER + 1ml TWEEN20

The antibody that was used mHmgn2 1/2000 and actin 1/2000 as a control the membrane, and watman paper were cut and socked into the methanol then to DH2O and lastly to the transfer buffer set up the gel with membrane into the following order



Then the gel with membrans was run at 1:30min on 25V. After the transfering the gel into the membarn, blocking the membrane was made with TBST and 5% milk and left it for 1 hr into orbital shaker.

First antibody Hmgn2 was added into the new TBST +5% MILK at dilution 1:2000 in boch bag over night. At 4c. The next day the member was washed 4 times with TBST. 15 minute each washing. Then the secondary antibody was added which is IgG 1:4000 dilution within TBST AND 5% buffer.

And left in orbital shaker for 1hr at RT. Finaly the membran was washed 4 times for 15min with TBST.

The illumino peroxidase was added into the membrane for viewing the target band using 3000pixl picture programme.

2.2.8.3.2 Western blot assay using Bio-Rad assay kit

The precast gel was used 12% non-gradient gel from Novex Invitrogen and the buffer was used MOPS 20X. the transferring membrane was from Bio-rad and the transferring membrane was used called TRUBO cat number 21703.

The blocking buffer was 1X PBS with 1ml tween 20 and 5% non-fat Milk and the washing step was with 1x PBS with tween 20 for 3 times in 5 min each.

2.2.9 Flow cytometry analysis

2.2.9.1 Cell cycle analysis (propidium iodide)

In the First day 10^6 of the cells (quantity of DNA same number of cells) was used to apply propidium iodide experiment. The cells were washed with PBS. Then the cell Pellet was in the 2.5ml PBS solution and 5 ml of 100% ice cold ethanol

The cell pellet was vortex 7.5ml in 50ml falcon tube. Immediately, the solution was added in -20°C overnight

The Next day, the cells were centrifuge to removed the alcohol and the rest of solution was removed by flipping the tube on tissues and leave it to dry out. (50ug/ml) of propidium iodide was used with Rnase A (0.5 mg/ml) in fresh 1 ml of PBS. 500 ul was taken from the mixture and mixed with pellet. The cells were incubated at RT. Then the cell will be ready to read out through Attune.

22.9.2 Cell compensation using FACS analysis

Cell compensation experiment was performed on FACS analysis. Before applying the target sample, P19 EC wild type controls need for setting the experiment. In this protocol, three controls were used: negative wild type P19 EC cells, P19 stained with green fluorescents antibody and P19 stained with red fluorescent antibody. Then experiment will be performed using another control for the staining and transfection beside the target sample.

To begin with, Cell compensation was optimized using wild type P19 EC cells as control. GFP (green) P19 EC cells control was performed using transient transfection for GFP plasmid in wild type P19 EC cells in 48 hours incubation period. Non-stained P19 EC cells were used to gate the four quadrants. Red fluorescent SSEA1 was stained in wild type P19 EC cells for 2 hours without cell fixation. The analysis was collected twice live cells with no fixation on the last day (day 25).

2.2.10 Immunofluoresent reagents

Immunofluoresent assay was developed by Dr. Torstein Stain lab at the University of Glasgow. To being with, 4% paraformaldehyde was prepared in warm smaller volume of PBS to 60°C ;then PFA was added to warm PBS and mixed . 1M of NaOH was added as dropwise until the PBS+PFA dissolved. Similar volume of 1 M HCl was added to neutralise the solution using pH paper to check pH. Take volume up to final volume with H_2O, filter sterilise. 4% PFA solutions mix needs to be warmed up every time there is immunoflorecent application. Therefore, this solution can be used either as fresh or as frozen.

To Fix cells, the medium was removed from the plate (wells) and 100 µl 4% PFA drop-wise was added to each well. All the volumes in 24 wells are 100 µl, although it can use 200 µl to be sure. The plate or well were Incubated for 20 min at RT. During this incubation period, methanol was chilled at -20°C. PFA was removed and 100ul of chilled methanol was added per well. Then the plate was Incubated at -20 degree for 1 min. the plate was Aspirated and the cells were washed for one time in PBS (slides can be stored in PBS at 4C at this stage). To Permeabilise cells, PBS was added with 0.1% triton x-100 and Incubated at room temperature for 10 min. then, Washing once with PBS and 20 mM glycine for 10 minutes to remove any cell debris. IF buffer was prepared as following: 100 ul per well (PBS+ 0.3% triton x-100 , 2.5% horse serum).

Then, the cells were blocked with IF buffer for 45 mins - 1hr at RT. (NB: prepare excess block to dilute primary Ab). The Primary Ab was used as the following: dilute primary Ab (1:200-1:1000) in IF buffer and Add 100ul/ well and incubate at 1 hr at RT. It is important to note that, it is possible to combine several 1' antibodies together. The Best way is actually to place 30 ul of solution on nescofilm and place coverslip up side down on the droplet. Cover to prevent evaporation. Then, primary Ab was removed and rinsed for 3 times for 10 min washes with PBS and 0.3% triton.

Secondary Ab was used as following step: (NB. Keep slides in the dark after secondary is added) Dilute the Ab (1:500-1:1000) in IF buffer for 1 hr at RT. It is important to note that it can combine several 2^{nd} antibodies at this

step. The plate was protected from light and covered to prevent evaporation. The Best way is to place 30 µl solution on nescofilm and place coverslip up side down on the droplet. Next, 3 times of washing step with PBS for the plates in 10 minutes happened after blotting with first and second antibodies; then Washing one time with dH_2O.

The plate was Mounted with prolong gold mounting medium with contain DAPI to counterstain the nuclei. If prolong gold doesn't contain DAPI already, add DAPI to 1 ug/ml final concentration just before use. The slides were left to set for 24 hours in the dark, covered. The slide was stored at RT short tem, or -20 degree for long term and thaw before you use. For chamber slides, cover with PBS and leave at 4°C. Allow reaching RT before imaging.

2.2.11 Chromatin Immunoprecipitation reagents and buffers

Protein and DNA cross-linked with formaldehyde:

P19 cells were cultured under advanced DMEM/F12, 1X glutamine and 10% NBCS . the culture was expanded in 10X of 15 mm dish. Each plate should be 50-60% confluent. Each chromatin prep use $7.5*10^7$ cells. $7.5*10^7$ cells were collected from P19 cells and $6*10^7$ cells were collected from E14 cells. Count one plate of cells and collect the total number of the cell and how many plates is needed to reach to $7.5*10^7$ .The cells were trypsinized and collected under the same medium but without serum (the serum reduce the formaldehyde efficiency).

Reagents preparation:

Before DNA pre-clearing step PBS , 2X cell lysis buffer, 2X nuclei buffer were pre-chilled in ice. In addition 1 ml of 10x nuclei buffer, 5% formaldehyde, and 1.5 ml of 1M of glycine were prepared for ready use.
5% Formaldehyde was added into the final concentration of 0.5% (20 ml of medium so 2.2 ml from 5% formaldehyde added into cells with medium).The plates were rotated into orbital shaker for 5 min each plate at room temperature. Then 1M of glycine to stop cross link reaction. The cells were

washed, scrapped from the plate using cold PBS. And centrifuged for 5 min. Cell lysis buffer with protease inhibitor were added into the cell pellet to remove the cell membrane. Then the nuclei lysis buffer was added into the cell pellet at the end DNA and protein was only left in the solution. The chromatin (DNA+ protein) was sonicated using Ultrasonic energy to disrupt the chromatin into smaller fragments. the process was done using machine SONICATOR 3000 at amplitude 3 for 5 min 10 sec on and 30 sec off. Then the sonicated chromatin was run in electrophoresis to check the size of the sample.

Pre-clearing chromatin

To eliminate the Chip background, pre-clearing was performed to ensure that almost no signal was obtained with the IgG control precipitation, which was done with each experimental set. Each IP has amount of chromatin, which range from 5-20 ug. Chromatin amount is equal to $1*10^6$ - $6*10^6$ cells. In P19 the total cells in 1 ml was $3.75*10^7$ whereas E14 a 1 ml from chromatin $3*10^7$ cells. The total cell was used in each IP was 1*10(6) cell/900ul IP with IgG antibody. Using protein A sepharose beads is necessary for chip to reduce the background level that caused by IgG antibody.

Immune-precipitation of chromatin

After pre-clearing different antibodies of the interest were added into each 900 ul of pre-clearing chromatin include one IgG antibody as positive control and 260 ul was for input control.

The antibodies was used as following

Antibody	Cat number (Millipore)	Concentration
H3 CT PAN	07690	2ul
H3K4me3	07473	2ul
H3K14ac	07353	2ul
Hmgn2	Homemade rabbit	2ul
Hmgn1	Homemade rabbit	1ul
H3K27ac		2ul

Another sephorose beads were added into chromatin with antibodies solution after centrifugation and washing steps. After several washes, wash with buffer I twice, then wash with high salt buffer to increase the chance that chromatin will be prevented from interacting with the beads. Next wash with IP wash buffer II, finally with 50-degree pre warm TE buffer. To elute the chromatin with elution buffer; then to maintain the protein-DNA crosslinked 17 ul from 5M NaCl and 1ul of Rnase A in the input only. Finally the sample was kept in at 65 degree overnight.

Recover the DNA:

To remove the DNA, Qiagen minlute column using PB buffer was obtained. The chromatin was eluted in 50 ul Qiagen elution buffer.

Western blot for chromatin extraction

This is the procedure for the benzonase digest for chromatin western preparation 100 ul of prepared chromatin was applied in this experiment. Then 10 ul of 10X TBS (200 mM Tris 7.5 PH and 1.5 M NaCl) was added into chromatin. in addition, 10 ul of 10% Triton X-100 was and 1.2 ul of MgCl2 100 mM were added into chromatin. It is important to note that this mixture is to give a final concentration of 1X of MgCl2, which is required for benzonase activation.

Make aliquots of 60 ul of the digest mixture. For example, 2 aliquots of the same sample was needed for the experiment. Benzonase (25 Unites) enzyme was measured to detect the amount of benzonase is needed in 100 ul of chromatin. it is important to note that The benzonase is highly concentrated. Therefore, ~250 U/ul is needed to dilute it at least 1 in 10 and then take 30 unites. 3 ul of diluted benzonase was used to obtain 1 in 10 of digested chromatin and it kept for further use. Benzonase with chromatin was left to digest for 15 minutes at room temperature. After digestion step, samples was prepared for western blotting assay as following: 14 ul chromatin + 5 ul of 4X Nupage loading buffer (LDS) + 1 ul of sample reducing reagent DTT. Then the sample mix was heated at 75 C for 10 minutes and loaded on pre-cast gel.

NOTE: if chromatin were not highly concentrated, a start up with a bigger amount of chromatin rather than 100 ul would be feasible. For example, 300 ul starting chromatin amount and the volumes of 10X TBS (30 ul), Triton X-100 (30 ul) and $MgCl_2$ (3.5 ul) was adjusted based on chromatin amount.

2.2.12 Co-Immunoprecipitation (Co-IP) for chromatin and whole protein cell extract

Cell harvesting/prep (for T75/20 mL, 80-95% confluence, 200000000 cells approximately) were collected in falcon tube and centrifuge 500 g/5 minutes at RT . Supernatants were removed and the cells were washed one time with PBS following centrifuge 500 g/5 minutes. Cell lysis and benzonase digested the resuspended cells in 1 of 2 ml buffer to 4 of dH_2O (1:4) + protease inhibitors. The cells were incubated on ice for 10 minutes and resuspend occasionally; Then the cells were Centrifuge 300 g/3 minutes at 4°C to collect nuclei and to save aliquots of supernatant for checking the fractionation efficiency. Resuspend the pellet was applied in 500 µl benzonase buffer and protease inhibitors; next, 25u/100 µl benzonase was added in to the pellet and incubated on ice for 30 minutes following resuspention occasionally. Then, the pellet Centrifuged in speed 300 g for 3 minutes at 4°C. to Recover the supernatant, the pellet was dissolved in sample buffer to check chromatin prep and solubilisation efficiency. Adding 500 µl dilution buffer and inhibitors diluted 2 fold of supernatant. This step is important to deactivates benzonase with (EDTA). The process is continue as following to prepare the sample for western blot essay using CO-IP.

Bead preparation begins with, 400 µl protein G (should be A for rabbit Ab) beads was prepared by washing twice with 1ml PBS and 0.5% BSA (adjust volumes according to requirements); then suspend beads step started by centrifuge the beads at speed 800 g for 1 minutes at RT, and supernatant was removed. This Resuspend pelleted beads became 1 to 1 with buffer 1. Pre-clear step started by Adding 50 µl beads and 10 µl rabbit IgG to lysate and Incubate with rotation 30 minutes at 4 degree C. then, the beads were Centrifuge at 800 g for 5 minutes under 4 degree c. Dilution step used dilute lysate ¼ in dilution buffer and protease inhibitors for approximately $5x10^6$

cells/ml and save 1-2 aliquots of 5% input (i.e 5% volume to be used in each sample).

IP step started with sample was in 1.5 ml eppendorf and 1 ml pre-cleared lysate and Ab (30ug Hmgn2 or 10mg of IgG – 10ul) were added. This step was repeated three times; then the sample was incubated under rotation from 2 hours to overnight at 4 degree. The final step used 50 ul of prepared proteins A/G beads was added to the samples and incubated under rotation for 3 hours at 4 degree. The sample was washed two times with buffer 4 and 2 times with buffer 5. Then the supernatant was removed for final wash and two times of 30 ul of sample buffer was added to the sample. The sample was heated for 10 minutes at 68-70°C. finally, the samples were ready to be run on western blot.

Chapter 3 The characterisation of lentiviral-mediated gene silencing of Hmgn2 in embryonic carcinoma EC P19 stem cells

3.1 Introduction

From previous study on *Xenopus laevis* the expression and the location of Hmgn1 and Hmgn2 are parallel. These two proteins found to be developmentally regulated such as down regulation of Hmgn1 and Hmgn2 affect the development of mouse embryo at pre-implantation stage. These results showed that losing the expression of Hmgn1 and Hmgn2 could terminate mouse embryo development (Postnikov and Bustin, 2009). Moreover, a study on Xenopus laevis in Hmgn1 and Hmgn2 found to be highly expressed in mesoderm and neuroectoderm region (Korner et. al., 2003). In addition, Hmgn1 and Hmgn2 expression showed at the early stage of embryo development and Hmgn2 downregulation in bovine embryo showed the embryo failing to develop to blastocysts due to the change in chromatin remodeling in histone H3 hyperacetylation at lysine 14 (Furusawa and Cherukuri, 2010). From these previous studies, it indicates that Hmgn1 and Hmgn2 may involve in maintaining pluripotency.

To confirm this hypothesis, loss of function experiment was carried out by siRNA and shRNA against Hmgn1 AND Hmgn2 in mouse teratocarcinoma stem cells (P19 EC cells). Lentiviral system is the best system because shRNA based on microRNA encoded to DNA vector to facilitate the ectopic mRNA expression. Moreover, lentiviral system can be constitutive or inducible to derive the expression and each shRNA based on microRNA gave arise to single siRNA for more gene target specificity (Hannon and Rossi, 2004).

In this chapter pSLIK-CD4 and pTRIPZ inducible lentivirus knockdown were used in P19 EC cells. both systems are tet-inducible TRE2 and they are only inducible systems. However, the reporter genes that track shRNA expression in pSLIK is GFP whereas pTRIPZ the reporter gene is turboRFP. PSLIK-CD4 design has TRE to encode GFP and miR-shRNA and it has separate Ubc promoter to express the reverse tetracycline transactivator (rtTA).

Moreover, this promoter express the drug resistance marker by IRES and it can silence many genes (Figure 3.1).

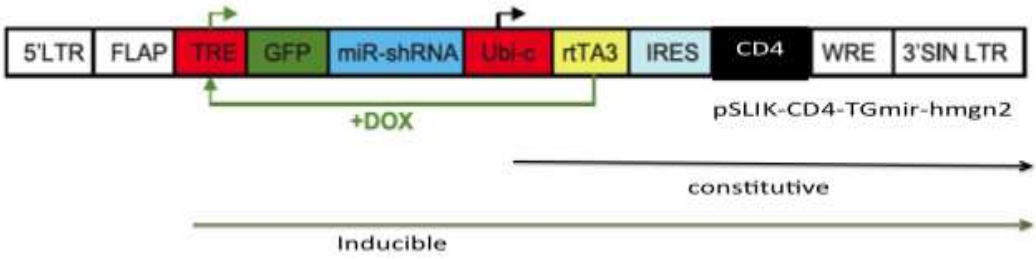


Figure 3.1 Schematic showing the insertion of miR-shRNA and GFP transgene in pSLIK-CD4
lentivirus encoding the Hmgn2 miR-shRNA. The schematic was adapted from (Shin et al., 2006). pSLIK-CD4 and pSLIK-hygro is the destination vector, which consists of four parts. First part is attR sequence elements are required for LR gateway reaction to flanked ccdB gene from destination vector. The insert TRE-GFP-shRNAmir is performing with rtTA3 to generate inducible knockdown system. The second part is the constitutive system of pSLIK using FLAP recognitions sites derived by CMV promoter. This part is optional to choose the constitutive or inducible system. The third part is for activating the Lentiviral recombination system (SIN/LTR, WRE, and SV40). The forth part is for cloning the vector into the competent cells (Ori and amp) adapted from (shin et. al., 2006).

Similar to pSLIK-CD4, pTRIPZ has TRE promoter, which can derive the expression of turbo RFP besides miRNA-shRNA. Furthermore, it had UBC promoter to derive the expression of rtTA, IRES, and Puromycine. PTRIPZ consider being more tightly regulated shRNA expression in many cell lines (Hu and Luo, 2012). Therefore, these lentiviral systems were optimizing to knockdown Hmgn1 and Hmgn2 in P19 EC cells (Figure 3.2)

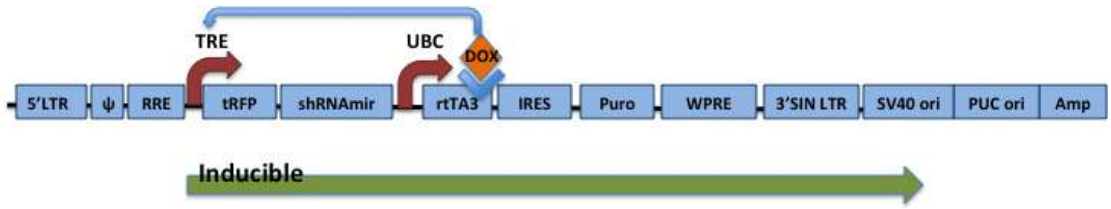


Figure 3.2 schematic diagram illustrating the expression cassette encoded by pTRIPZ lentiviral vector

TRIPZ inducible system contains two parts. First part is constitutive system derived by UBC promoter. This constitutive part consists of Bi-cistronic element (IRES) and rtTA3 element following Puromycin drug selection. The second part if inducible, which only can be activated by supplying the medium with Doxycycline, contains RFP and shRNAmir derived by TRE promoter. Abbreviation: IRES, internal ribosome entry elements; WRE, woodchuck response element; LTR, long terminal repeats; puro, puromycin (adapted from open-biosystem)

3.2 Aims and objectives

The aim of this part of the work was to investigate whether lentiviral shRNAmir systems could be used in P19 cells for the inducible knockdown of Hmgn1 and/or Hmgn2.

The specific objectives of the experimental work described in this chapter were:

1. To test for detrimental effects of doxycycline on P19 cells
2. To optimise conditions for generating P19 cells stably transduced with pSLIK or pTRIPZ lentiviral particles.
3. To investigate whether P19 cells transduced with pSLIK or pTRIPZ shRNAmir viral particles have reduced expression of Hmgn2.

3.3 Results

3.3.1 Testing for detrimental effects of doxycycline on P19 EC cells

Doxycycline is a common reagent used for regulating inducible expression, and it is typically used at 1-2 µg/ml. In order to investigate whether growing P19 cells in doxycycline could be detrimental (Chang et. al., 2014 and Souza et. al., 2015), several different tests were carried out. The effect on cell proliferation was assayed using a luminescent 96-well plate assay for cellular ATP (Figure 3.3).

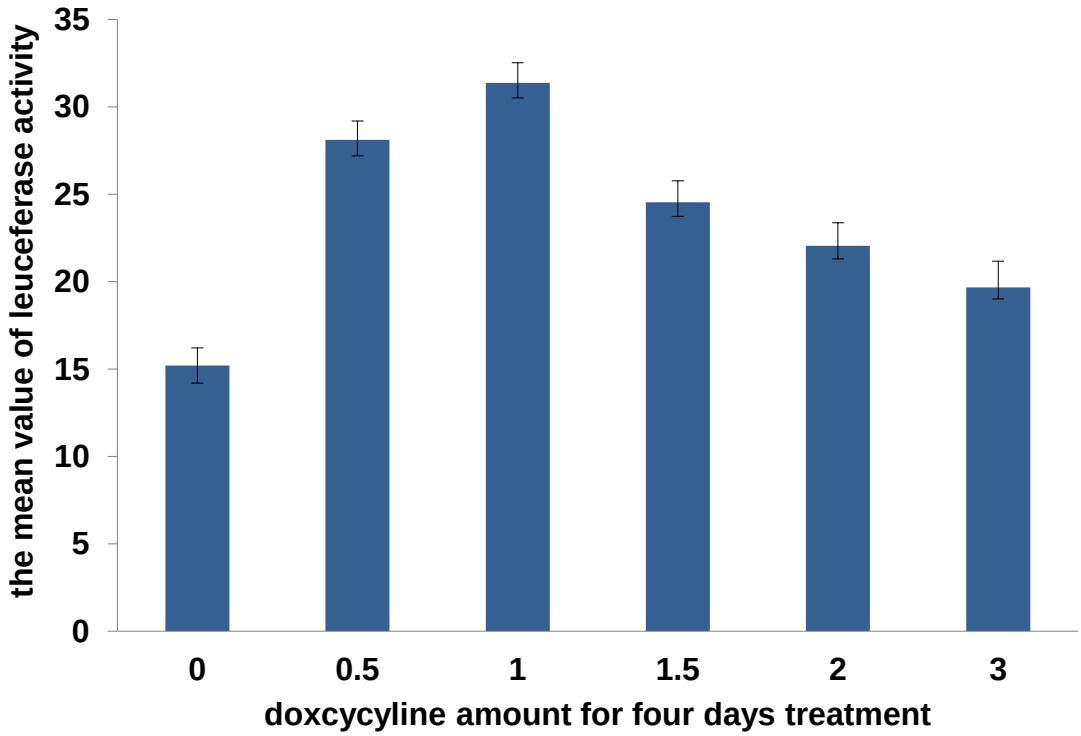
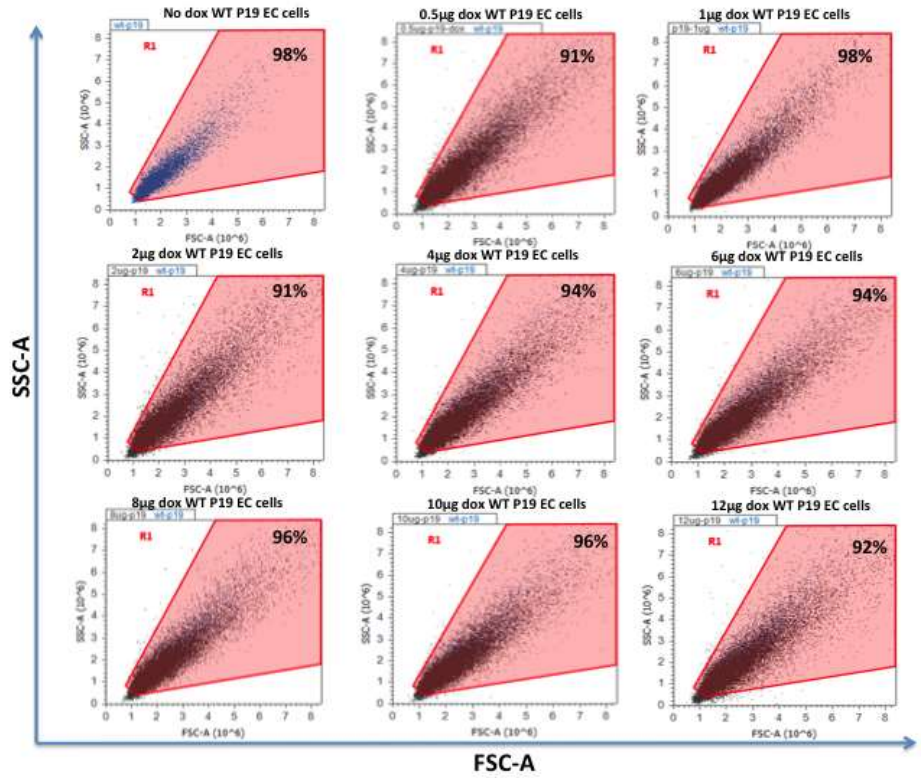


Figure 3.3 The effect of Doxycycline treatment on the growth and viability of P19 cells. Cells were incubated with the indicated concentrations of doxycycline (µg/ml) and grown for four days. Cell viability was assayed using the Vialight assay according to the manufacturer's instructions (Lonza). Error bars were calculated based on the concentration to the number of repeat.

In order to test whether doxycycline causes gross defects in cell morphology, cells were assayed by FACS after 3 and 5 days of treatment with different concentrations of doxycycline (Figure 3.4). The forward and

side scatter measurements would reveal whether cells are significantly altered in terms of their size, shape or granularity, which might indicate changes in cell health or identity. No significant changes were observed in any of the samples.

A)



B)

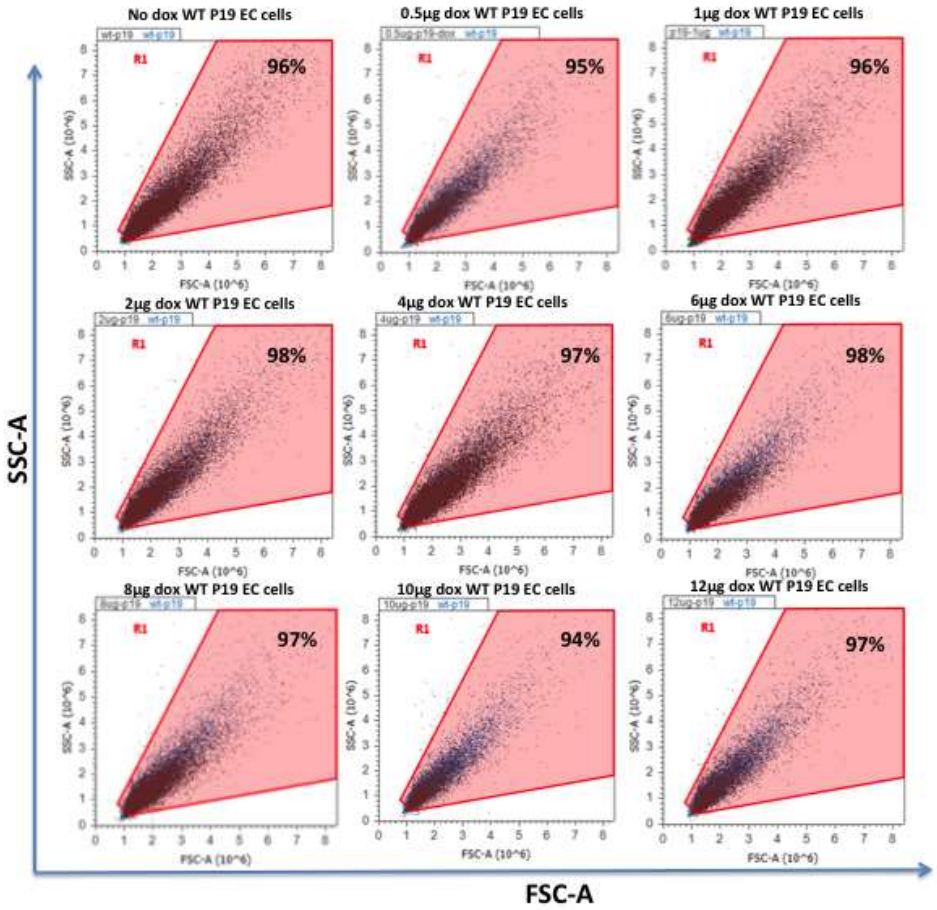


Figure 3.4 The Effects of doxycycline on P19 EC cells
Flow cytometer analyses (FACS) of wild type EC P19 cells after 3 (A) and 5 (B) days of Dox treatment with the indicated concentration of doxycycline (0.5, 1, 2, 4, 6, 8, 10, and 12 µg/ml). Media was replaced every 24 hours, and cells were passaged every 48 hours. The gate for healthy cells was drawn based on wild type P19 cells without Dox treatment. FSC-A (forward scatter) corresponds to a relative size for the cell whereas SSC-A (side scatter) corresponds to granularity. The percentage of cells falling within the gate is indicated.

In a further attempt to investigate whether doxycycline might have a major impact on this investigation, the mRNA expression level of Hmgn2 and Oct4 were assayed following five days of treatment with varying levels of doxycycline (Figure 3.5). It can be seen that there is quite wide variability in the expression levels between samples, and there is no consistent trend with increasing doxycycline concentration. In particular, at 1 µg/ml

doxycycline, the levels of Hmgn2 and Oct4 mRNA are very similar to those of untreated cells.

These investigations indicated that 1 µg/ml of doxycycline treatment for up to 5 days does not have a major detrimental on P19 cells. However, doxycycline may be causing other changes in gene expression, so it is important to include dox-treated control cells with each experiment.

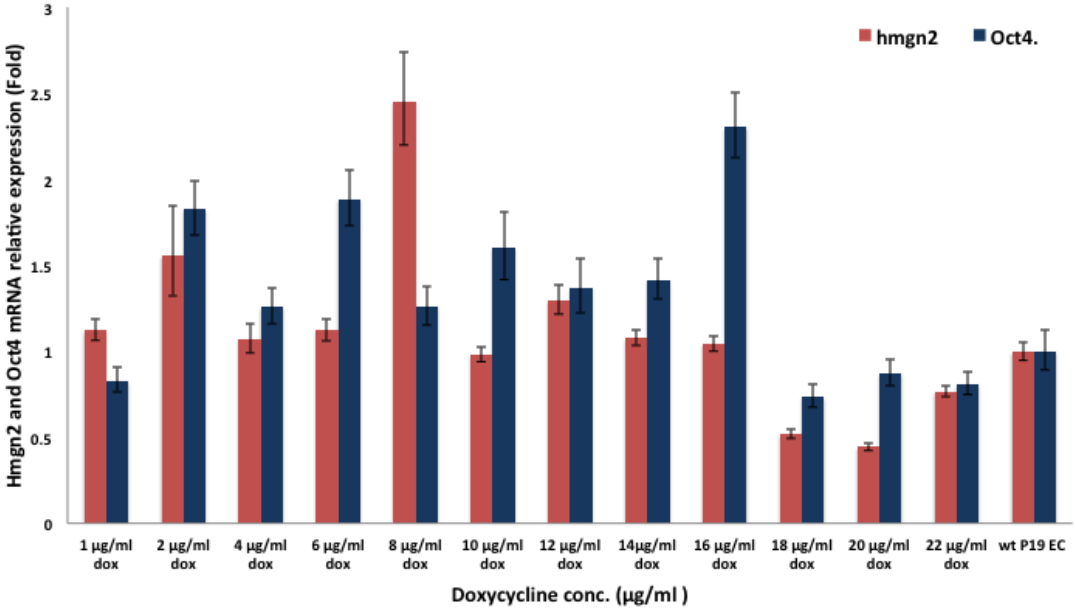


Figure 3.5 the influence of doxycycline on Hmgn2 and Oct4 expression in P19 EC cells
Cells were incubated with the indicated concentration of doxycycline for five days. Medium was replaced every 24 hours. RNA was prepared, and qRT-PCR carried out to quantify the relative expression of Hmgn2 and Oct4. Gene expression was normalised to ß-Actin as housekeeping gene and is calculated relative to expression in untreated wt P19 EC cells. Error bars represent the standard deviation from QPCR triplicates.

3.3.2 The Construction of pSLIK-CD4-shRNAmir30

The pSLIK lentiviral system was initially chosen for expression of shRNA. This system enables doxycycline-induced expression of miRNA containing the shRNA for RNA-interference. The Ubi-C promoter drives constitutive expression of rtTA3 and delta CD4, which are separated by an internal ribosome entry site (IRES). Doxycycline causes rtTA3 to bind the TRE promoter, driving expression of shRNAmir and GFP. Mohan (2012) had previously used siRNA for successful knockdown of HMGN1 and HMGN2 in P19 cells, so the same trigger sequences were used in the design of the lentiviral

shRNAs: N1-2 and N1-3 target Hmgn1, N2-1 targets Hmgn2, and sc is a negative control containing a scrambled version of the N2-1 trigger sequence.

The cloning strategy is shown in Figure 3.6. shRNA sequences were synthesised and cloned into the entry vector pENTR-TGmiRC3 by the company Shinegene (Figure 3.7).

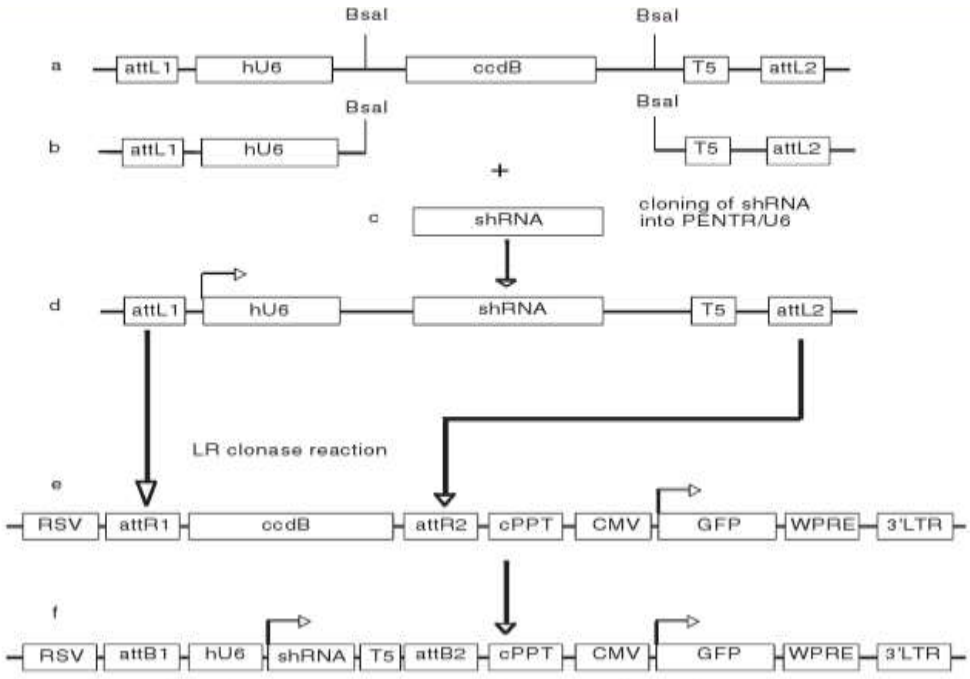


Figure 3.6 Diagram shows from a-f the construction of the pSLIK lentiviral shRNA expression vector

Part a: the entry vector (pENTR-TGmiRC3) for receiving the shRNA contains the ccdB gene, which is toxic when expressed in standard E.coli strains. Insertion of the shRNA into BsaI sites removes the ccdB gene, to generate the vector in part d. Gateway recombination is used to move the shRNA from the entry vector into the destination vector shown in part e (pSLIK-CD4). The final vector is shown in part f (pSLIK-shHmgn-CD4).

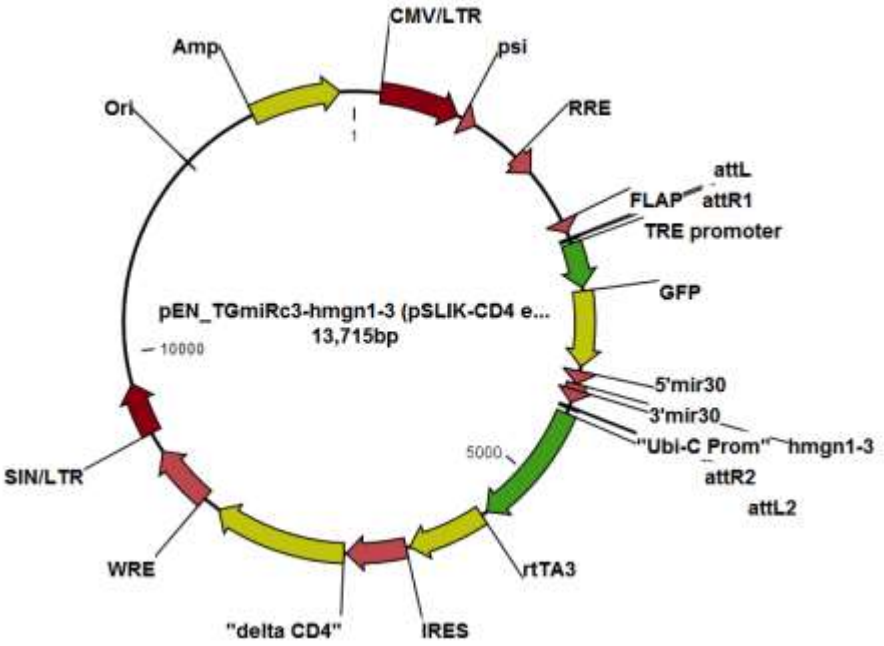


Figure 3.7 : Elements in the pSLIK-shHmgn-CD4 lentiviral vector Plasmid map of the final lentiviral expression vector

3.3.2.1 Cloning of pSLIK-shHmgn-CD4 vectors

Gateway recombination technology was used to insert the shRNAmir30 sequences from the entry vectors into the pSLIK-CD4 destination vector. The final vectors were checked by restriction enzyme digestion using EcoRI (Figure 3.8) and Kpn1 (not shown), and then confirmed by DNA sequencing.

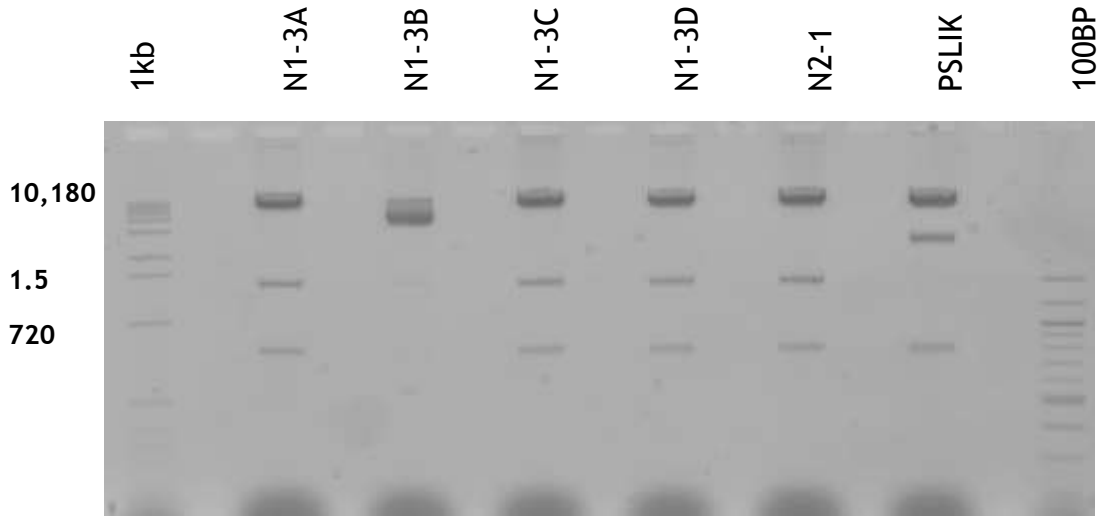


Figure 3.8 Example restriction digest of plasmid clones resulting from the LR-recombination reaction

LR recombination reactions were performed using pSLIK-CD4 and the entry vectors containing N1-3 or N2-1. Colonies were grown in culture and DNA extracted using a miniprep kit (Qiagen). DNA (1.5 ug) was digested with EcoRI-high fidelity and run on a 0.8% agarose gel in TAE buffer. D. The expected bands for unrecombined pSLIK-CD4 are 10.18 kb, 2054 bp, 780 bp. The expected bands of recombined pSLIK-shHmgn-CD4 plasmids are 10.18 kb, 1.5 kb, 780 bp. In this gel, clone N1-3B is not correct, but N1-3A, N1-3C, N1-3D and N2-1 show the expected bands.

3.3.2.2 The Production of pSLIK-shHmgn-CD4 viral particles

PSLIK-shHmgn-CD4 vectors were transfected into the cell line Hek293/T17 along with three third generation packaging vectors (see chapter 2; materials packaging vectors) (Shin and et al., 2006).

Supernatant containing viral particles was harvested as described in chapter 2, diluted to varying degrees with serum-free ultracultural medium, and then added to exponentially growing P19 cells. After three days, 1 μg/ml doxycycline was added, and the induction of GFP expression was assayed by microscopy three days later (Figure 3.9).

Figure 3.9 shows the results from a typical experiment. Strong GFP expression can be seen in many cells transduced with undiluted virus, or with virus diluted by 2 or 4 fold. Fewer cells express GFP following transduction with virus used at the highest dilution (16x). However, even at the high viral concentration, less than 50% of cells are expressing GFP.

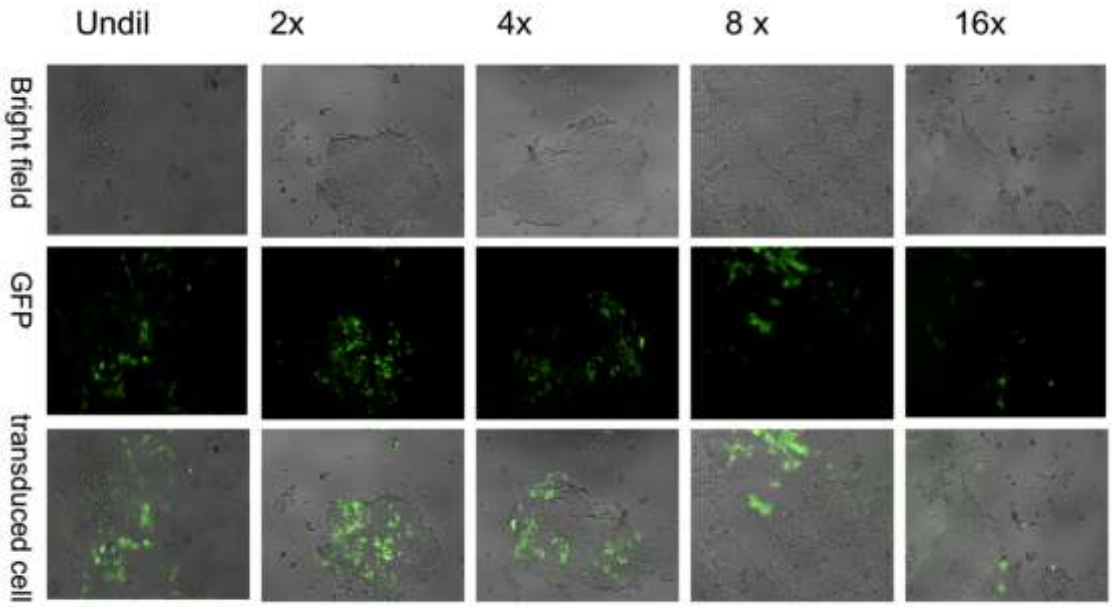


Figure 3.9 Fluorescent microscopy shows serial dilution of pSLIK-Hmgn1-2-CD4 viral particles
The green cells indicate the transduced cell with viral particles. These cells derived by TRE promoter, which is activated by rtTA3 gene under three days doxycycline presences. The image divided into three parts: the bright field, the GFP and transduced cells (merge). Each image contains 2 fold dilutions the first image is undiluted Lentiviral particles.

3.3.2.3 The Optimisation of pSLIK-shHmgn-CD4 viral transduction

In order to optimise viral transduction of P19 cells, several parameters were tested: A range of doxycycline concentrations was tested to ensure maximum induction. There was little difference between 1 µg/ml and 8 µg/ml doxycycline, and it was decided to use 2 µg/ml in all further experiments.

Several batches of fetal bovine serum were tested to ensure there were no trace levels of tetracycline that could induce GFP expression in the absence of exogenous doxycycline. Finally, a range of viral supernatant dilutions in different culture vessels was tested, and it was decided to use 3 ml of nude viral supernatant in a 25 cm^2 flask for 24 hrs.

3.3.2.4 The Analysis of pSLIK-transduced P19 cells

Given that it was not possible to transduce more than 50% of P19 cells with the lentiviral particles, we attempted to use magnetic cell sorting (MACS) to enrich for transduced cells. However, this was not successful in our hands, and colleagues using the same system in our department also found this was unsuccessful. It appears that the CD4 fragment is not expressed from this vector as expected. Therefore, we used FACS sorting for GFP following doxycycline treatment to enrich for transduced cells.

To determine the Hmgn1 and Hmgn2 expression in transduced P19 cell lines before and after FACS sorting, we used real-time quantitative RT-PCR. Expression of Hmgn1 and Hmgn2 in unsorted cells is shown in Figure 3.10. with and without DOX (Figure 3.10). None of the transduced cell lines showed a reduction in Hmgn2 or Hmgn1 expression below the levels in wt cells. The N1-2 cells did show a reduction in Hmgn1 expression in dox-treated cells compared to no-dox cells. However, a similar reduction in Hmgn2 expression was observed in these same samples, so the effect on Hmgn1 expression was not specific. In Knockdown P19 cell lines did not show reduction neither Hmgn2 nor Hmgn1. We concluded that knockdown P19 cell lines still a mixture population of transduced and normal cells (Figure 3.10).

Expression of Hmgn1 and Hmgn2 in sorted cells is shown in Figure 3.11. In the N1-2 cells, the expression of Hmgn1 was reduced in dox-treated cells compared to wt cells. In N2-1 cells, expression of Hmgn2 was reduced in dox-treated cells compared to untreated cells, but the level was not significantly reduced compared to wt cells. Furthermore, a similar reduction in Hmgn1 expression was observed in dox-treated cells compared to untreated cells.

Western blotting was used to investigate whether the protein levels of Hmgn1 or Hmgn2 were reduced in the transduced cells. No reduction in either protein compared to the wt P19 cells was observed (data not shown – western not good enough).

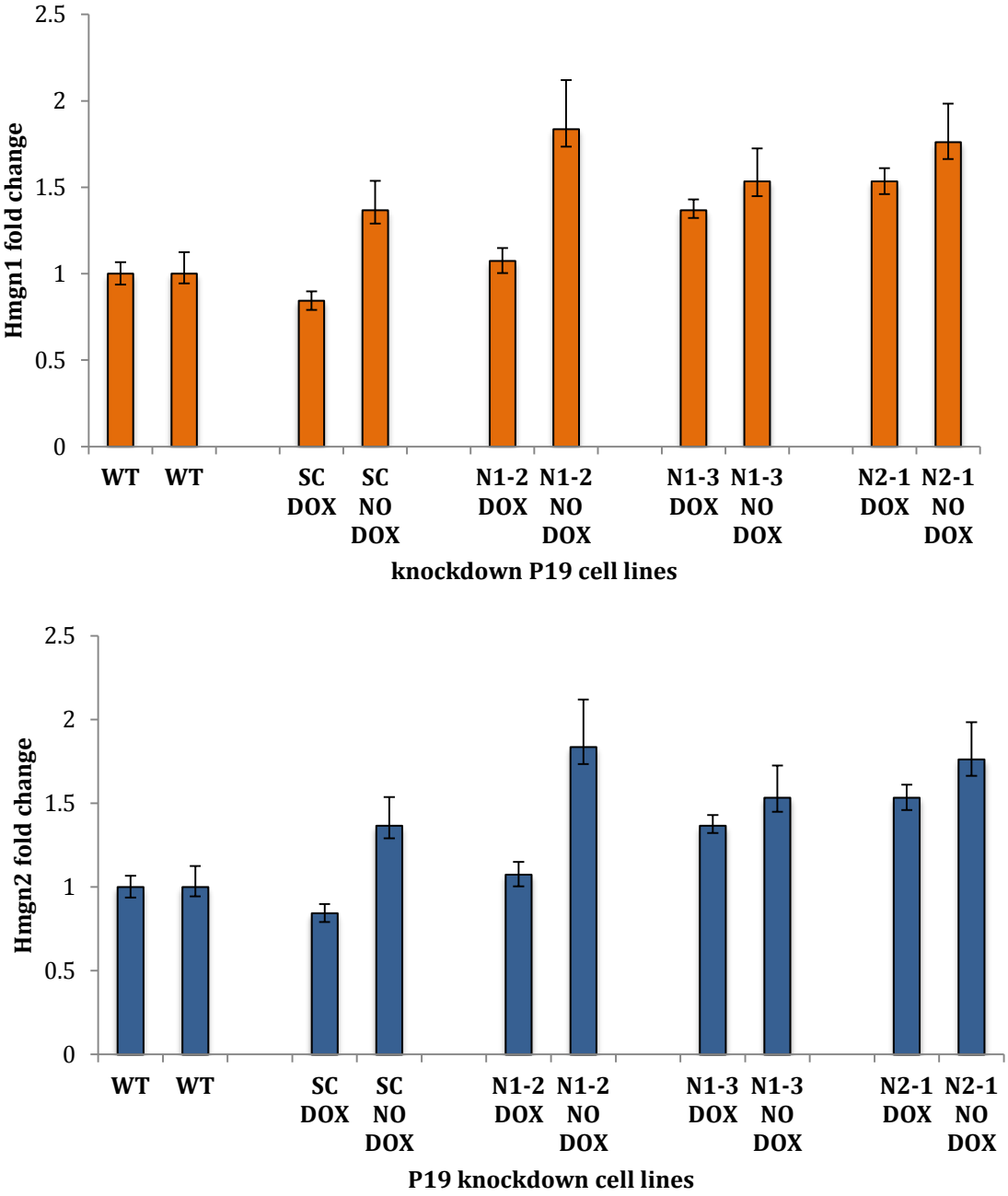
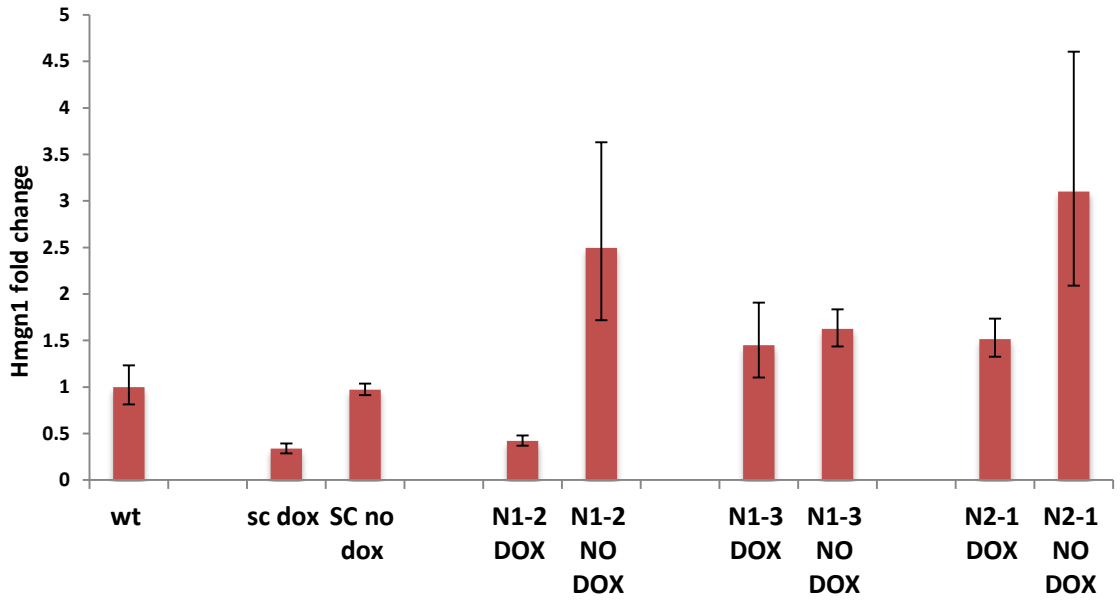


Figure 3.10 A. Hmgn1 and B. Hmgn2 RNA level in pSLIK-transduced P19 polyclonal cell lines

These cell lines were grown with or without 2 ug/ml Dox for five days. The expression level are normalized against control gene and expressed as fraction of expression in wtP19 cells. Error bars reflect standard deviation between technical PCR replicates. Knockdown P19 polyclonal cell lines: sc (Hmgn-scramble), N1-2 (Hmgn1-2), N1-3 (Hmgn1-3) and N2-1 (Hmgn2-1).

A.



B.

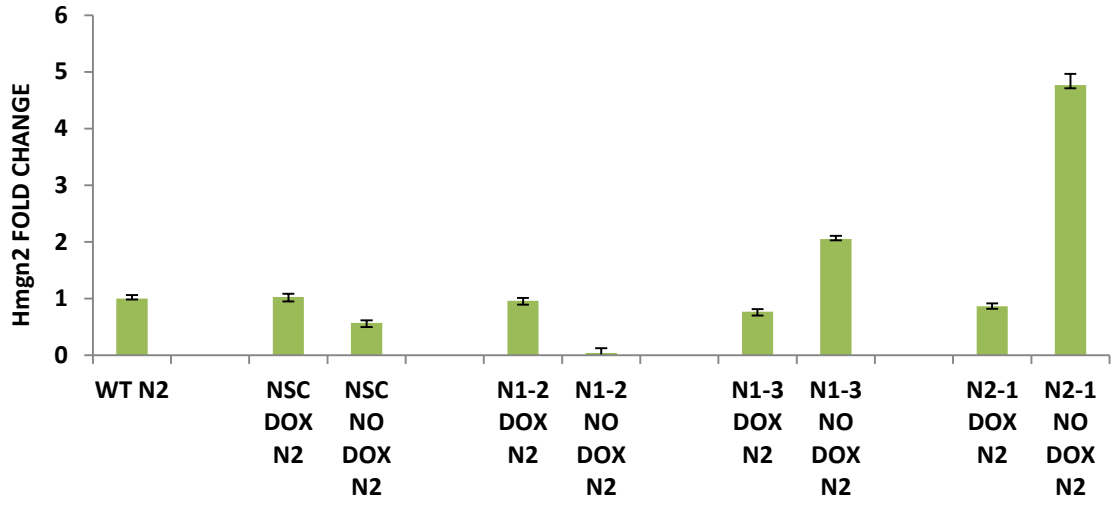


Figure 3.11 A. Hmgn1 and B. Hmgn2 gene expression in pSLIK-transduced P19 cells enriched for high GFP expression using FACS sorting

The sorted cells were cultured with or without 2 ug/ml Dox treatment for 30 days. Expression level was normalized against selected control gene and is plotted as a fraction of wt P19 levels. Error bars reflect the standard deviation between technical PCR triplicates. Knockdown P19 monoclonal cell lines: sc (Hmgn-scramble), N1-2 (Hmgn1-2), N1-3 (Hmgn1-3) and N2-1 (Hmgn2-1).

3.3.2.5 The Conclusions from pSLIK experiments

From these experiments, we concluded that the performance of the pSLIK system was not good enough for our requirements. Levels of P19 cell transduction were low, and the absence of a drug-selection marker or CD4 expression meant that FACS sorting of dox-treated cells is the only way to enrich for transduced cells. Even following sorting, evidence for knockdown of Hmgn1 or Hmgn2 was inconclusive. For these reasons, we decided to try a different inducible lentiviral vector: pTRIPZ.

3.3.3 pTRIPZ as an inducible lentiviral vector system

The pTRIPZ lentiviral vector (Thermo-scientific) is an alternative system for the production of lentivirus particles that express shRNAs in response to doxycycline. Greg Hannon reference A schematic of the expression cassette is shown in Figure 3.12. The UBC promoter drives constitutive expression of rtTA3 and the puromycin resistance gene, separated by an IRES element. In the presence of doxycycline, the rtTA3 transactivator binds the TRE promoter and drives expression of turboRFP and the shRNAmir. (Rytlewski and Beronja, 2014; Dull et. al., 1998).

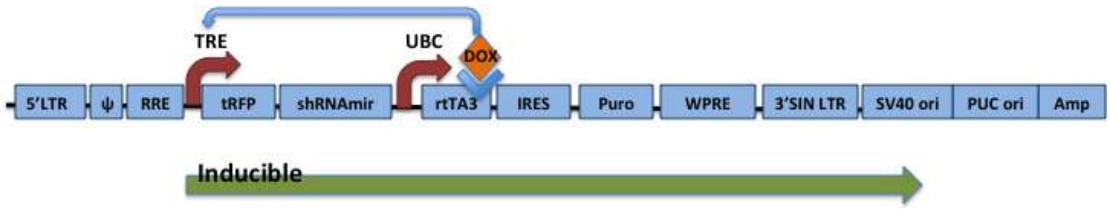


Figure 3.12 Schematic diagram illustrating the expression cassette encoded by pTRIPZ lentiviral vector
Abbreviation: IRES, internal ribosome entry elements; WRE, woodchuck response element; LTR, long terminal repeats; puro, puromycin. TRE, RRE, tRFP, amp, ori

The main advantage of pTRIPZ over pSLIK is that it carries the gene for puromycin resistance, so that transduced cells can be selected. A puromycin kill curve was carried out with each new puromycin vial in order to determine the optimum puromycin concentration to use for P19 cells. The

optimum concentration was typically 1.5 µg/ml. No difference in sensitivity was observed whether the cells were grown in DMEM/F12 or alpha-MEM media.

3.3.3.1 The Experimental design of Inducible pTRIPZ for Hmgn2 in EC P19 cells

Various shRNAmir triggers were cloned from pSLIK and pGIPZ vectors (see chapter 4) into pTRIPZ (provided by K. West):

shHMGN2-381:	targets Hmgn2
shHMGN2-514:	targets Hmgn2
shHMGN2-968:	targets Hmgn2
shHMGN2-N2-1:	targets Hmgn2
shGAPDH:	targets Gapdh
shN-scramble	negative control
sh-346:	non-silencing

3.3.3.2 Transfection of packaging cell line 293T/17 with inducible lentiviral vector

Viral particles were produced by transfecting HEK293T/17 cells with the pTRIPZ vector and packaging mix (Thermo Scientific), as described in chapter 2.

Media supernatant containing the viral particles was collected. P19 cells were cultured in ultracultural medium for 24 hours, then the medium was changed to alpha-MEM supplemented with 10% serum. Cells were selected with 1.6 ug/ml puromycin, and then 1 ug/ml doxycycline added for four days. FACS was performed to quantify the percentage of cell expressing RFP:

shHMGN2-381:	6%
shHMGN2-514:	2%
shHMGN2-968:	35%
shHMGN2-N2-1:	43%
shGAPDH:	2%
sh-346:	59%

It can be seen that the percentage of cells carrying the vector varied from 2% - 59%, despite all the cells being resistant to puromycin. Fluorescent microscopy confirmed the general trend of these results (Figure 3.13). However, there was concern that the cells did not appear to be healthy when cultured in the presence of both puromycin and doxycycline. Specifically, cell morphology was altered (Figure 3.13) and the proliferation rate was decreased, even when only a small percentage of cells were expressing RFP.

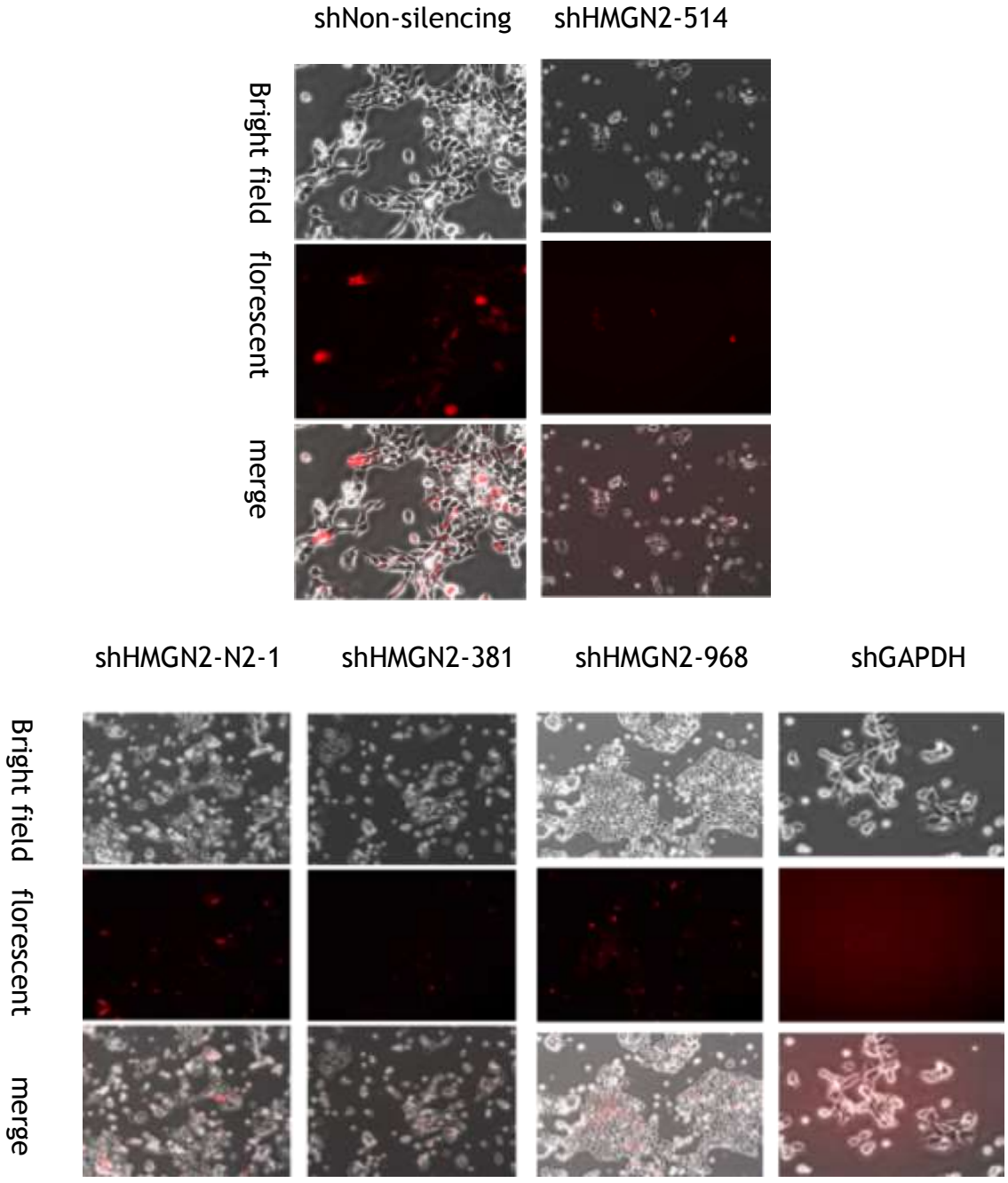


Figure 3.13 Fluorescent microscopy of P19 cell lines transduced with pTRIPZ (1:1) lentiviral particles

Non-concentrated particles were diluted 1:1 with ultracultural serum free media and added to P19 cells. After 24 hours, the medium was changed to alpha-MEM plus 10% serum. Cells were selected with 1.6 ug/ml puromycin, then 1 ug/ml doxycycline added for five days.

RT-PCR expression of Hmgn2 knockdown P19 cells was performed to determine the knockdown efficiency (Figure 3.14). It was found that only shHMGN2-968 cells showed a modest knockdown of Hmgn2 in the presence of doxycycline compared to wt P19 cells by 20%.

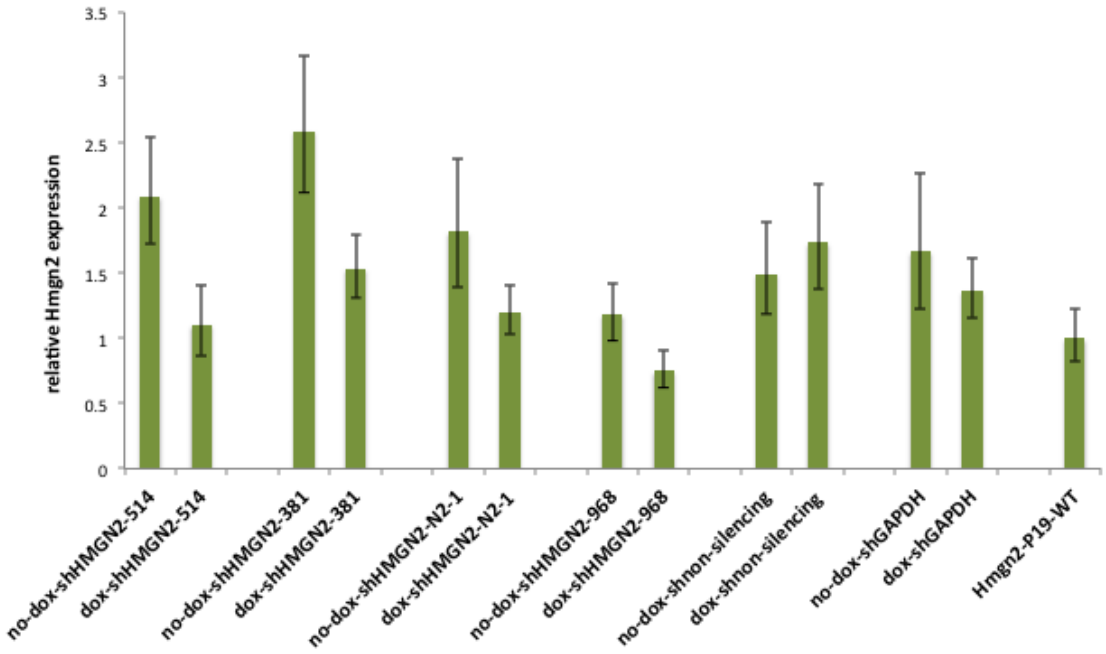


Figure 3.14 RT-PCR of Hmgn2 knockdown in P19 cell lines transduced with TRIPZ lentiviral vectors

Cells were selected with puromycin and induced with 1ug/ml Doxycycline daily for five days. Two controls are 16.1(364) (negative control), and GAPD (positive controls) and four shRNAmir-Hmgn2 sequence triggers are 4.1(381),8.1 (968), 2.1 (514) and N2-1. The y-axis presents the fold change from delta CT value.

3.3.3.3 The Optimisation of pTRIPZ transduction of P19 cells

The above experiments demonstrated that it is possible to transduce P19 cells with pTRIPZ lentiviral particles and detect RFP. However, the percentage of transduced cells was low, the cells did not look healthy, and it was unclear whether knockdown of Hmgn2 was occurring. Further experiments revealed that transduced cells were much healthier when grown in advanced DMEM-F12 instead of alpha-MEM (data not shown), and from this point onwards, P19 cells were always grown in advanced DMEM-F12.

To increase the lentivirus titer, two different methods of concentrating viral particles were tested. A polyethylene glycol precipitation method was unsuccessful, but concentration by centrifugation through a filter unit increased the efficiency of transduction significantly (kit – as described in Chapter 2- section 2.2.4.2.3: PART C PAGE 121). The multiplicity of infection was also increased by reducing the density of P19 cells to $5x10^5$ cells per well of a six-well plate.

3.3.3.4 The Generation of pTRIPZ-transduced P19 lines

The concentrated viral particles were diluted 1:10, 1:50, 1:100 and 1:500 in ultracultural media in four separate transductions for each pTRIPZ vector, in order to determine the best dilution ratio.

Transduced cells were selected with puromycin for one week as described earlier, and treated with doxycycline for up to 12 days to induce expression of RFP and shRNA. FACS for RFP was performed at 0, 2, 4, 5, 8 and 12 days of doxycycline treatment. Fluorescent microscopy was performed at day 4, and RNA samples and protein extracts were collected at the same time point (Figure 3.15).

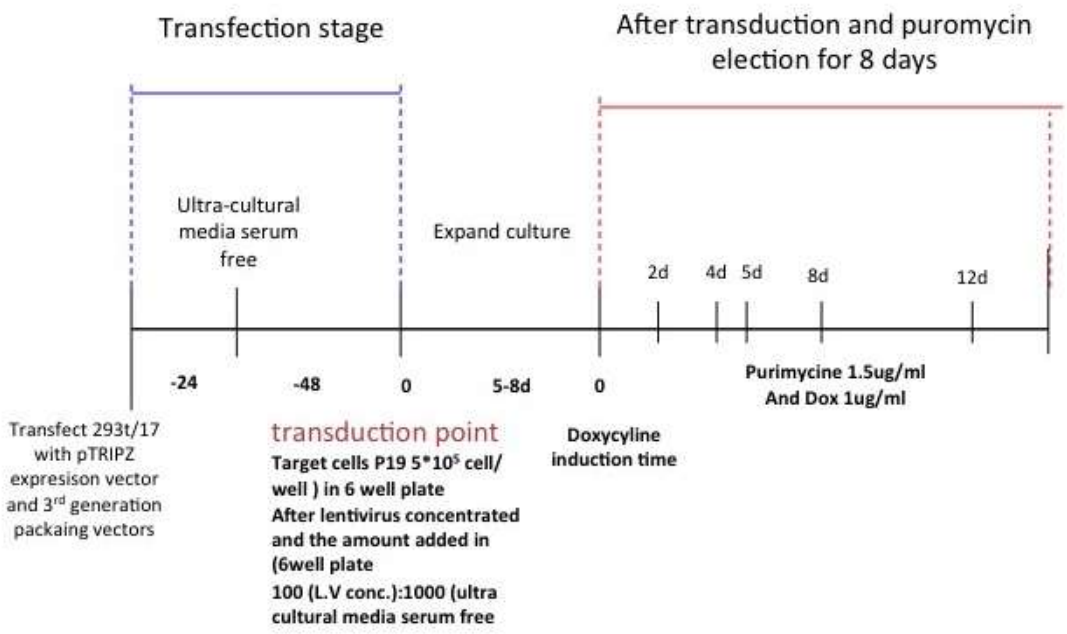
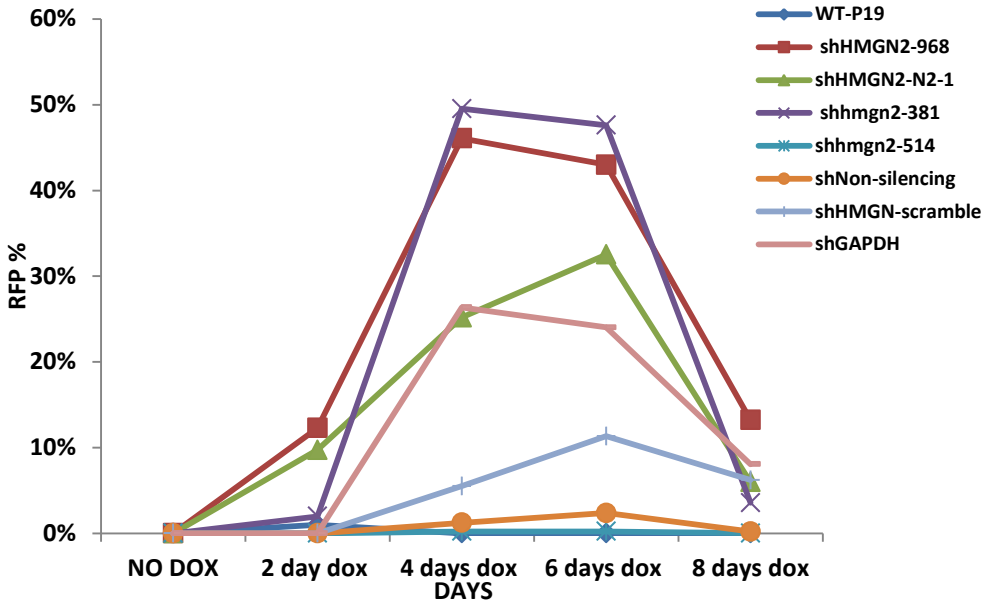


Figure 3.15 Experimental design for transducing P19 cells with pTRIPZ lentivirus
P19 cells were transduced with concentrated pTRIPZ viral particles and selected with 1.5 µg/ml puromycin for five to eight days. Doxycycline (1 µg/ml) was added at day 0 and samples collected for FACS at 0, 2, 4, 6 and 8 days.

Figure 3.16a and b show the FACS data from the experiments using 1:10 and 1:500 viral dilutions, respectively. It can be seen transduction efficiency correlates with viral titre, as transduction with 1:10 dilution of virus (Figure 3.16a) yielded up to 50% of RFP positive cells, whereas transduction with a 1:500 dilution of virus yielded a maximum of 5% RFP positive cells. Not all pTRIPZ vectors were equally effective. For example, shHMGN2-968 yielded 45% and 4.5% RFP-positive cells with the 1:10 and 1:500 viral dilutions respectively, whereas shHMGN2-514 yielded 3% and 0.5% RFP-positive cells in the same experiments. After day 6, the level of RFP in the 1:10 experiment gradually declined, and the experiment was terminated after day 8 due to the loss of RFP Given the significant difference in the percentage of cells expressing RFP, only cells from the 1:10 dilution experiment were analysed in subsequent assays.

A)



B)

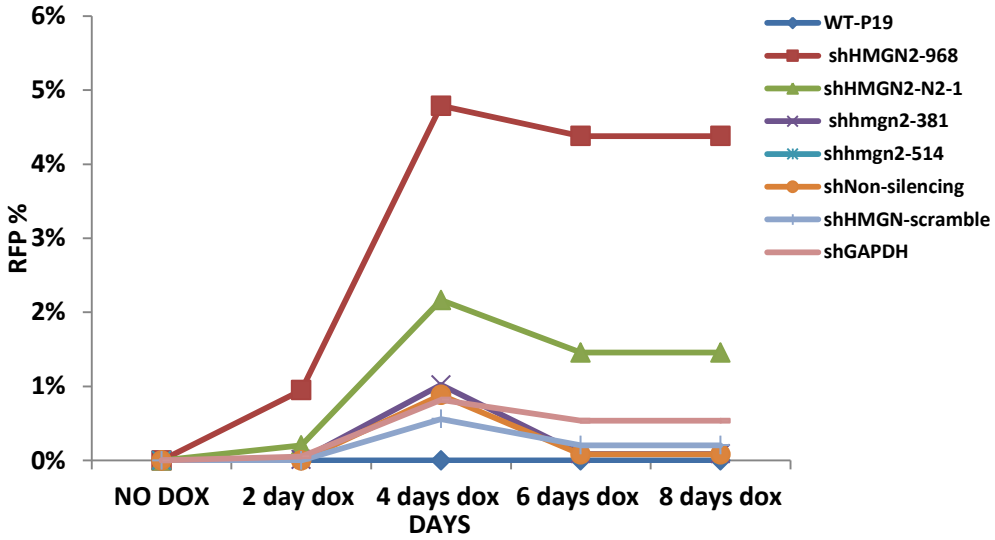
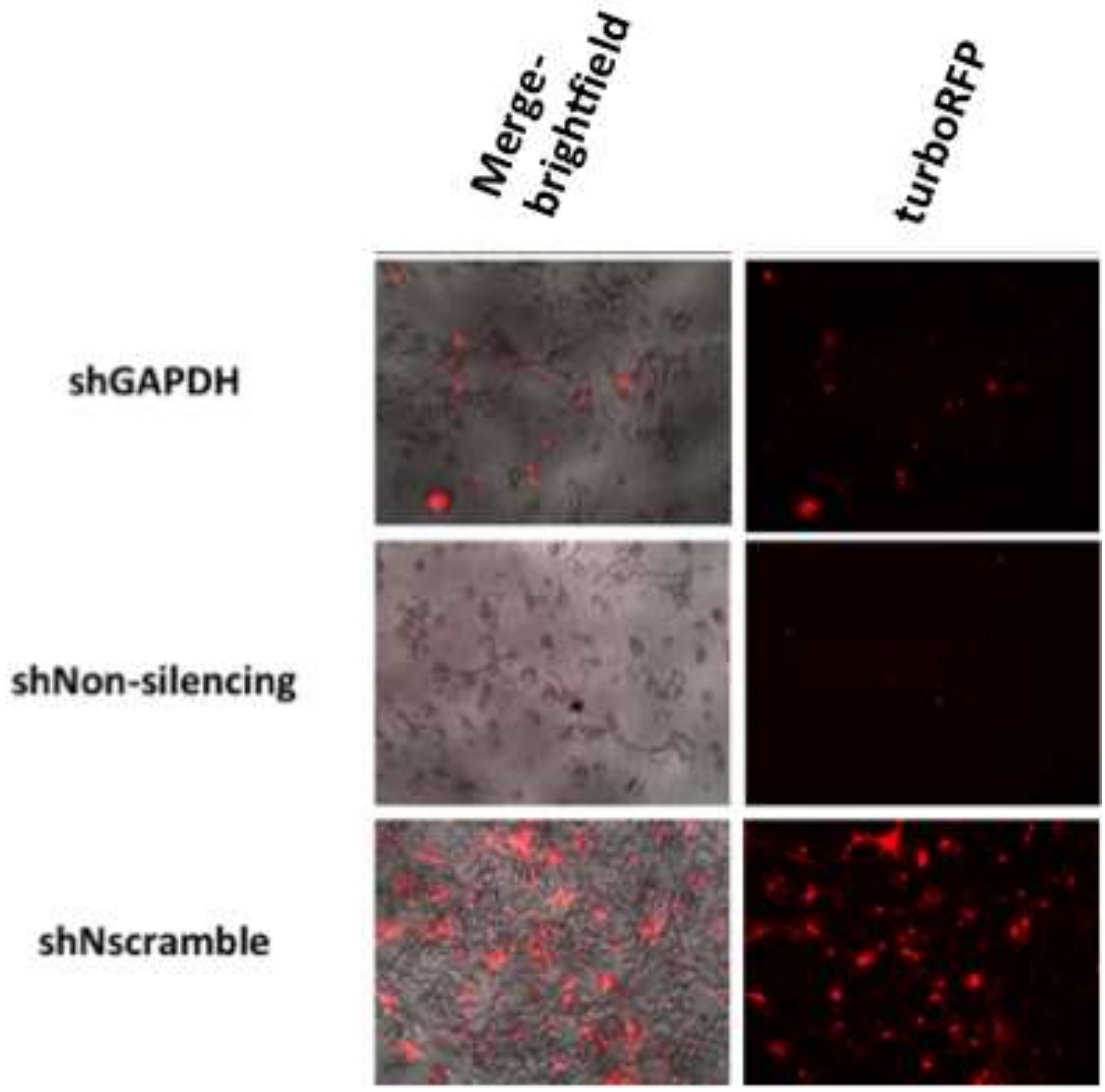


Figure 3.16 Flow cytometer (FACS) of RFP expression in P19 EC cells transduced with pTRIPZ lentivirus

There are four shHMGN2 (shHMGN2-968, shHMGN2-N2-1, shHMGN2-381 and shHMGN2-514) and four controls, WT P19, sh-346 (non-silencing) and shN-scramble (negative controls) and shGAPDH (positive control). A) 1:10 diluted viral particles B) 1:500 diluted viral particles. Cells were selected with 1.5 µg/ml of puromycin for 8 days, then treated with 1 µg/ml doxycycline for the indicated times. Media was canged every 24 hours.

Fluorescent microscopy was performed at day 4 to check the cell morphology and health and RFP level (Figure 3.17). The results were similar to FACS analysis in that approximately 50% of cells transduced with the shHMGN2-381 vector were RFP0% of cells. On the other hand, the shHMGN2-N2-1 cell line had less than 10% RFP-positive cells, whereas it was 30% in FACS analysis. Negative control shN-scramble had around 50% RFP-positive cells by fluorescent microscopy, while it was 10% in FACS analysis (Figure 3.17).



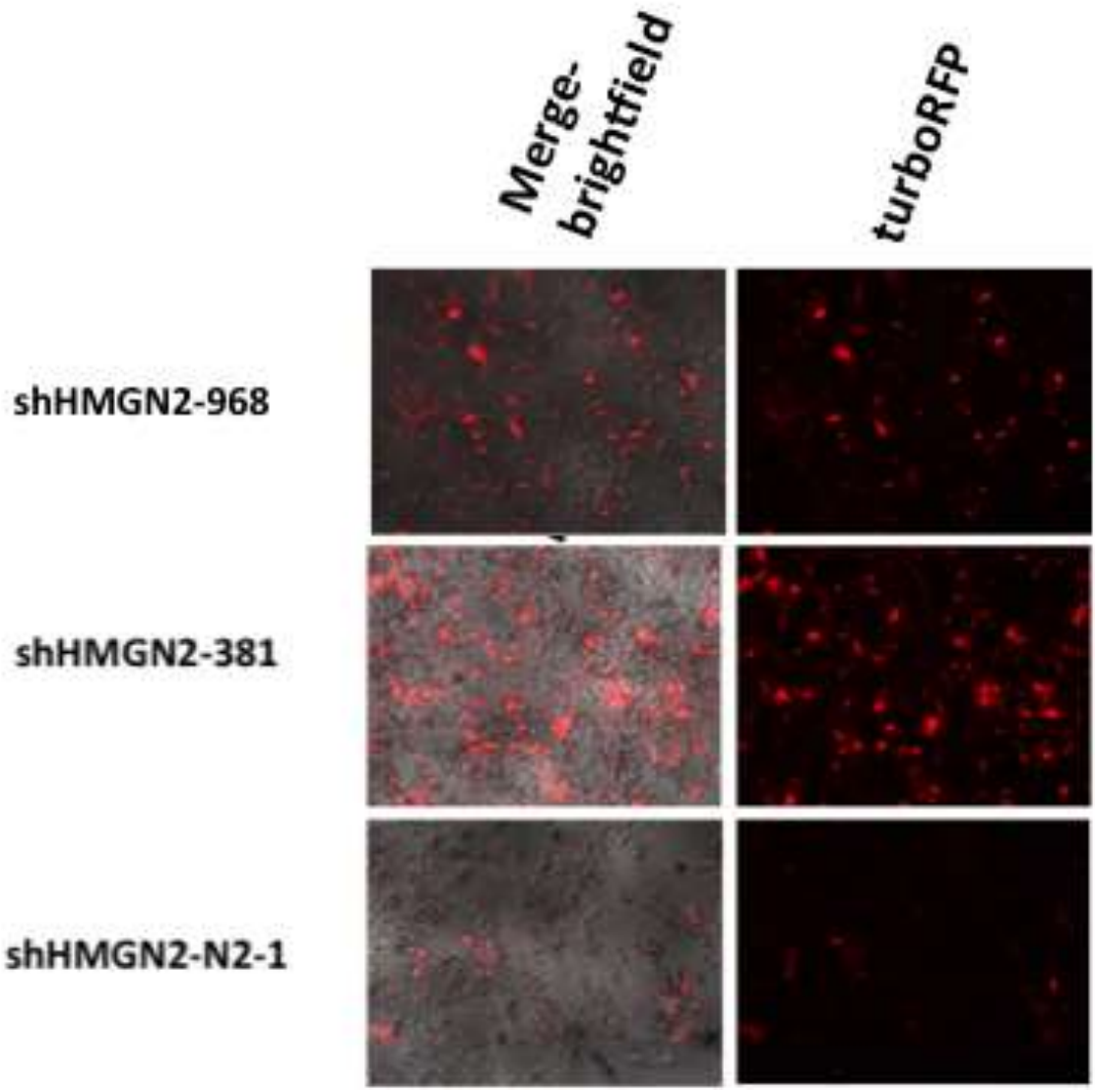


Figure 3.17 Fluorescent microscopy of P19 cells transduced with 1:10 diluted pTRIPZ after 4 days doxycycline treatment

P19 EC cells were transduced with 1:10 diluted viral particles, selected with puromycin and treated with doxycycline for four days. Shown are the TurboRFP (fluorescent) image and the merge between brightfield and fluorescent field for each shHmgn2 (shHMGN2-381, shHMGN2-N2-1 and shHMGN2-968) and controls (shGAPDH, sh-346 and shN-scramble), using the 10x objective.

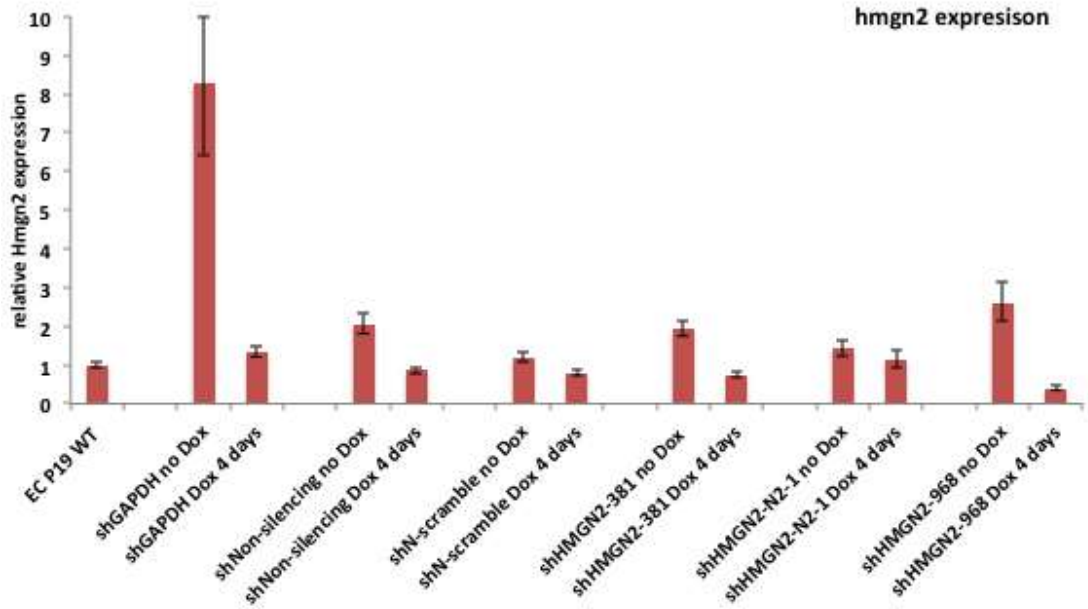
3.3.3.5 Examine Hmgn2 expression using reverse transcriptase-coupled quantitative real-time (PCR) and western blot assay

QRT-PCR was employed to detect the changes in Hmgn2 mRNA expression after four days of doxycycline treatment (Figure 3.18a). The results showed that Hmgn2 expression in shHMGN2-968 and shHMGN2-381 cells induced with Dox was downregulated by 80% and 30% respectively compared with wild type P19 EC cells. In contrast, shHMGN2-N2-1 cells failed to show any knockdown of Hmgn2, even though this same siRNA sequence was previously used successfully in transient siRNA transfections (Mohan, 2012). These observations may be related to the percentage of cells expressing RFP: shHMGN2-968 and shHMGN2-381 cells were around 45% RFP-positive at day 4, whereas shHMGN2-2-1 cells were 25% RFP-positive at this time point (Figure 3.16A).

Negative Controls (shN-scramble and shNon-silencing) with Dox showed a slight reduction in Hmgn2 expression. None of the cell lines showed a reduction in Hmgn2 expression in the absence of doxycycline. However, in two of the lines, Hmgn2 expression was more than two fold higher than wt in the absence of doxycycline.

Hmgn1 expression was assayed in these cell lines in order to check for non-specific knockdown by the shRNA, and to look for possible upregulation of Hmgn1 to compensate for a loss of Hmgn2 (Figure 4.18B). Hmgn1 expression was not significantly altered in shHMGN2-968 cells with or without doxycycline. However, Hmgn1 expression was reduced by 80% in shHMGN2-2-1 cells following doxycycline treatment. Furthermore, four of the samples showed Hmgn1 expression more than two-fold higher than wt cells. Without further replicates, it is not possible to conclude whether this wide variability in Hmgn1 expression is due to the pTRIPZ transductions, or is due to other factors such as cell confluency or genetic drift due to time in culture.

A)



B)

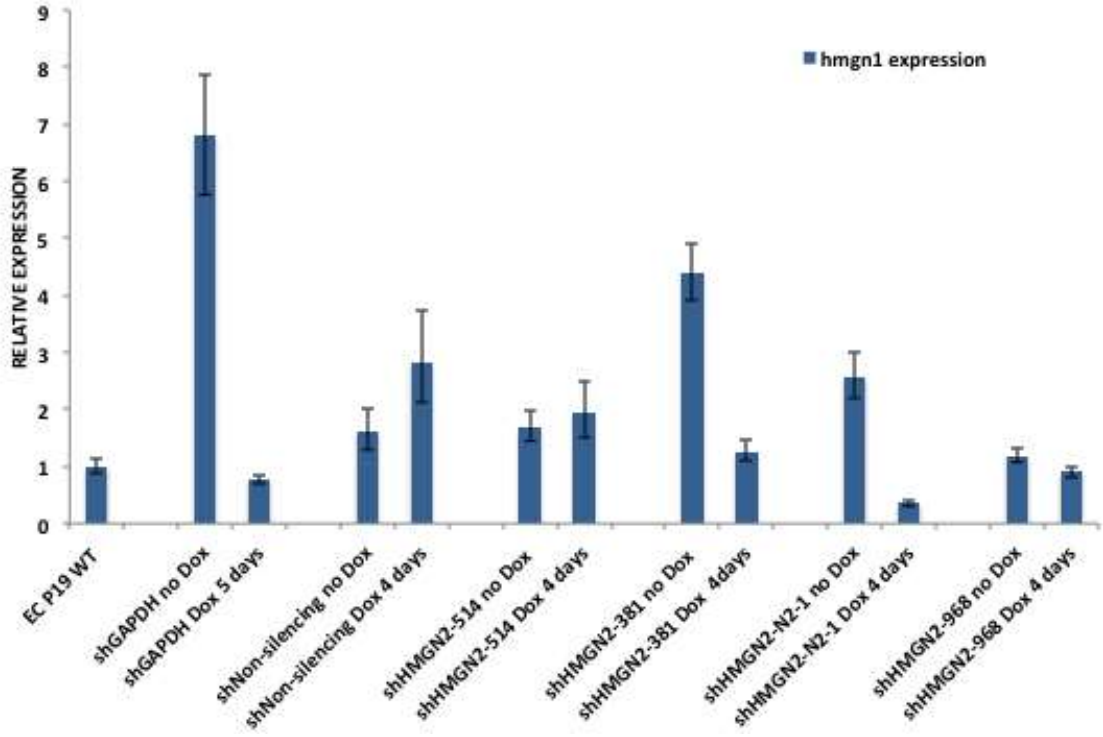


Figure 3.18 Expression of Hmgn2 and Hmgn1 in EC P19 cells transduced with pTRIPZ viral particles

EC P19 cells were transduced with a 1/10 dilution of concentrated pTRIPZ viral particles, selected with puromycin, and treated with doxycycline for four days. RNA was prepared and Hmgn2 (A) and Hmgn1 (B) expression

quantified using qRT-PCR. Data was normalized to GAPDH, and is plotted relative to expression in wt P19 cells. Mean and s.d from PCR triplicates are shown.

Western blotting was performed in order to investigate whether Hmgn2 protein levels were reduced in pTRIPZ-transduced cells (Figure 3.19). Coomassie blue staining of the gel prior to transfer, and western blotting using an antibody against β-actin were used as loading controls. Hmgn2 expression was significantly reduced in shHMGN2-381 and shHMGN2-N2-1 cells (lanes 9 and 10 respectively) compared to WT P19 cells (lane 1). shHMGN2-968 cell lines did not show any loss of Hmgn2 protein, in contrast to the qRT-PCR analysis.

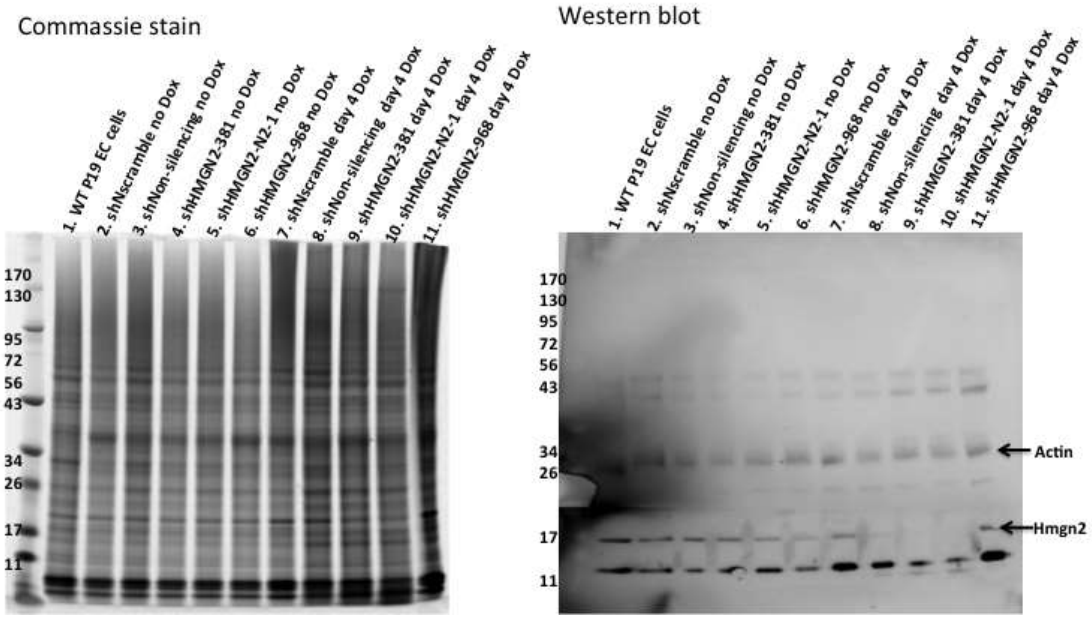


Figure 3.19 Western blot analysis of Hmgn2 protein levels
EC P19 cells were transduced with a 1/10 dilution of concentrated pTRIPZ viral particles, selected with puromycin, and treated with doxycycline for four days. A: coomassie blue-stained protein gel. B: western blot cut in half and probed with antibodies against actin (upper blot) or Hmgn2 (lower blot). Wild type (Lane 1); shN-scramble (lanes 2 and 7); shNon-silencing (lanes 3 and 8); shHMGN2-381 (Lanes 4 and 9); shHMGN2-N2-1 (lanes 5 and 10); shHMGN2-968 (lanes 6 and 11) respectively. No dox: lanes 1-6. 4 days dox: lanes 7-11.

3.3.3.6 Checking the expression of HMGN2 in knockdown EC P19 cells using fluorescent immunostaining

The FACS analysis in Figure 3.16 shows that less than 50% of cells were transduced with pTRIPZ lentivirus. This means that methods such as western blotting and QRT-PCR are less accurate for reporting whether the different vectors are capable of knocking down Hmgn2 expression. Immunofluorescence was therefore used in order to analyse knockdown on a cell-by-cell basis (Figure 3.20).

The top row of images in Figure 3.20 shows RFP expression in red, which represents cells transduced by the virus. The second row shows Hmgn2 immunofluorescence in green. The third row shows a merge of the DAPI signal and Hmgn2 signal for the two lines with shRNAmir against Hmgn2 – N2-1 and 381. This shows the total number of cells in the field. The bottom row shows a merge between the red and green fields.

In the shGAPDH and shN-scramble controls, it is clear that the intensity of green Hmgn2 staining is similar in all cells, whether they are expressing RFP or not. In shNon-silencing cells, the overall level of Hmgn3 expression is similar to the shGAPDH and shNscramble controls. However, some shNon-silencing cells with high RFP expression appeared to have low Hmgn2 expression (Figure 3.20).

In doxycycline-treated shHMGN2-381 and shHMGN2-N2-1 cells, the overall level of green Hmgn2 staining appears to be reduced in over 80% and 60% of cells respectively (compared to the controls). This is particularly clear in the cells highlighted by arrows – these express high levels of RFP and have lower Hmgn2 than other cells in the same field. Cells not treated with doxycycline expressed Hmgn2 at similar levels to the controls (Figure 3.17).

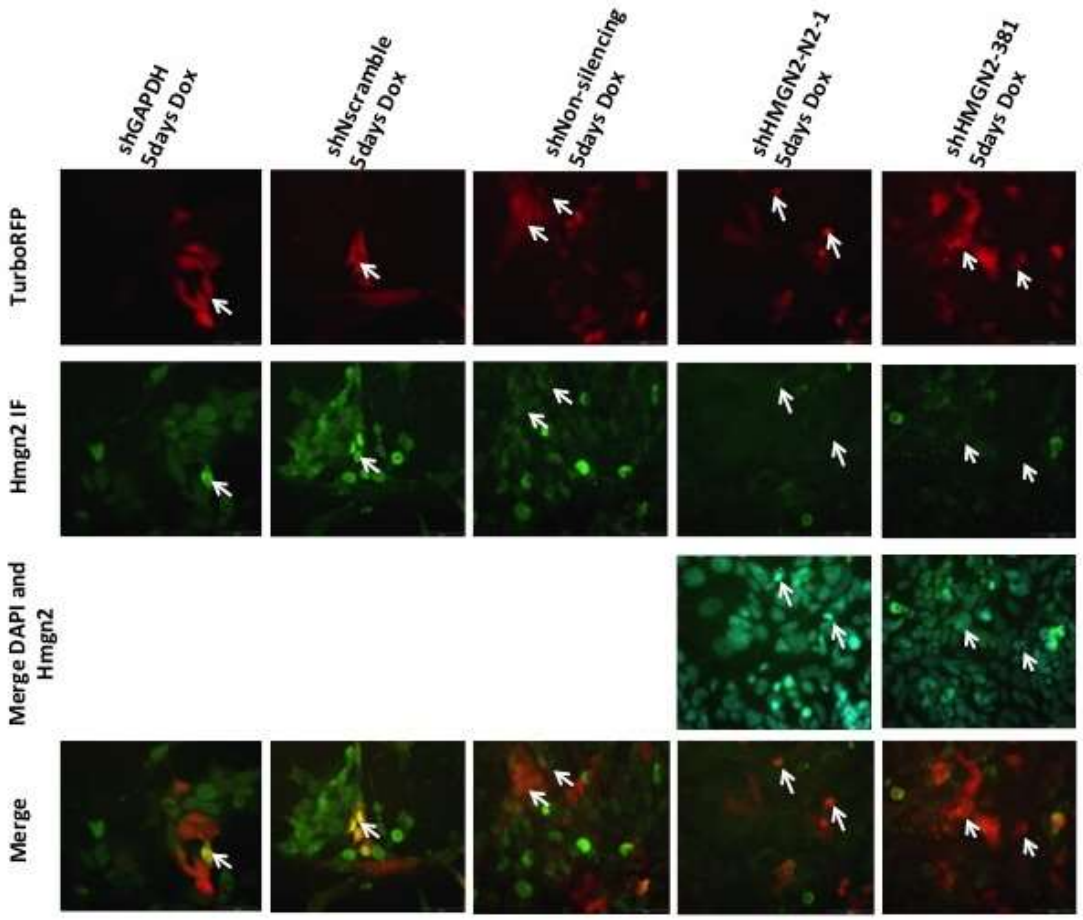


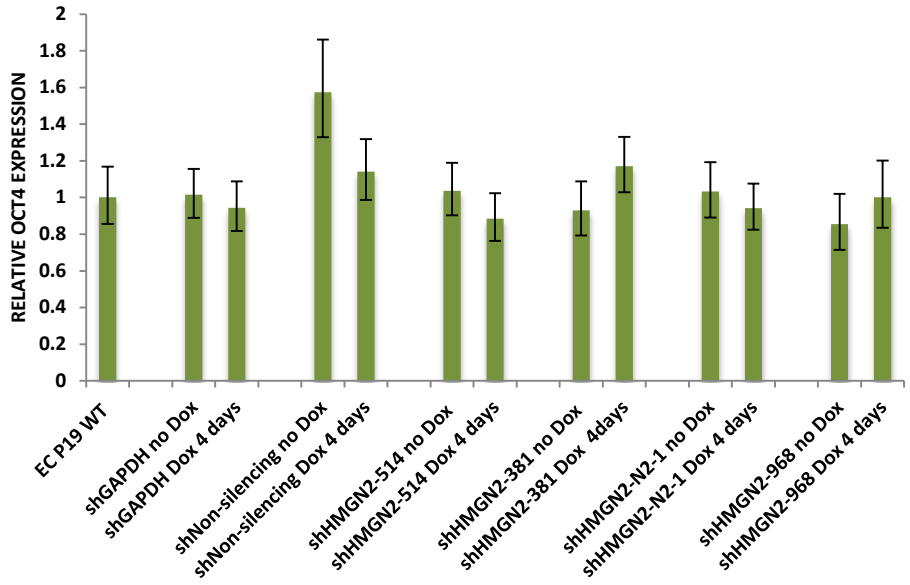
Figure 3.20 Immunofluorescence of Hmgn2 knockdown P19 EC cells
Immunofluorescence analysis of Hmgn2 expression in P19 cells transduced with pTRIPZ lentiviral particles. Cells were treated with 1 ug/ml dox for five days(× 1000 magnification). A. RFP expression from the pTRIPZ vector. B. Hmgn2 immunofluorescence staining. C. DAPI and Hmgn2 merge image D. Merge image between RFP (red) and Hmgn2 (green).

3.3.3.7 Investigating the influence of HMGN2 knockdown in the expression of stem cells pluripotent markers

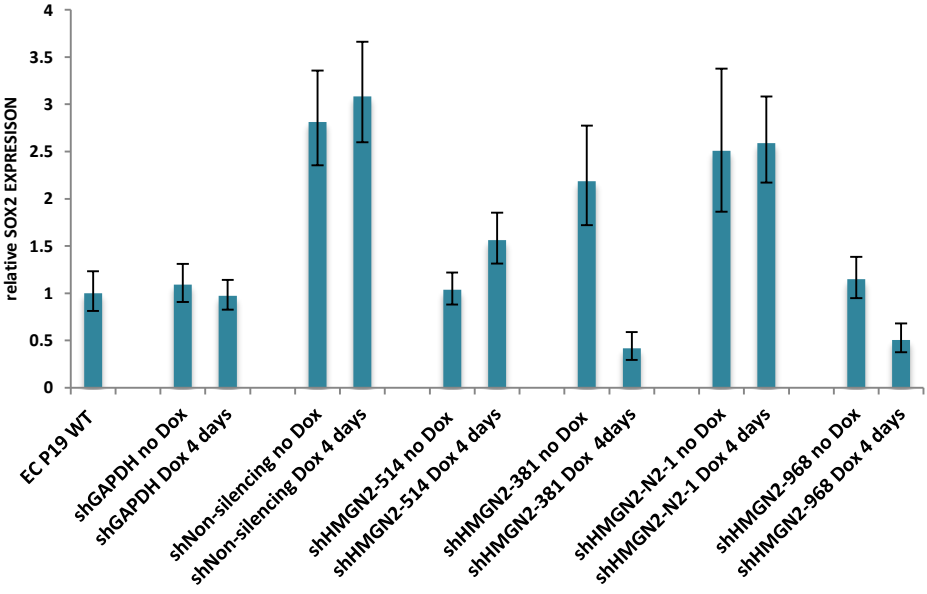
Previously, knockdown of Hmgn2 in P19 EC cells using siRNA was shown to reduce the expression level of Oct4, Sox2 and Nanog, as shown by qRT-PCR and western blot analysis (Mohan, 2012). In order to investigate whether a similar phenomenon was observed in cells transduced with Hmgn2-knockdown pTRIPZ vectors, Oct4, Sox2 and Nanog levels were quantified by qRT-PCR. The results showed the expression levels of Oct4 and Nanog were relatively unchanged compared to wild type P19 EC cells after 4 days

(Figure 3.21a and c) or 6 days (not shown) of Dox induction. Sox2 expression was reduced to 50% of wt in shHMGN2-968 cells, and reduced to 20% of wt in shHMGN2-381 cells (Figure 3.21B).

A. Oct4 expression in Hmgn2 knockdown P19 EC cells with pTRIPZ under doxycycline for 4 days.



B. Sox2 expression in Hmgn2 knockdown P19 EC cells with pTRIPZ under Doxycycline for 4 days



C. Nanog expression in Hmgn2 knockdown P19 EC cells with pTRIPZ under doxycycline for 4 days

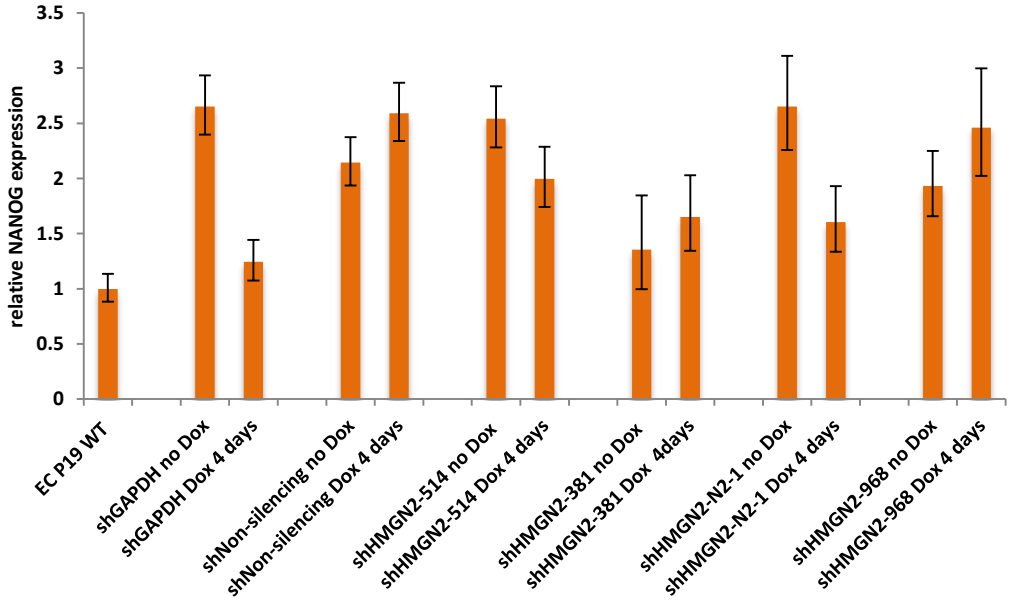


Figure 3.21 Expression of Oct4, Nanog and Sox2 in pTRIPZ-transduced cells
EC P19 cells were transduced with pTRIPZ virus, selected with puromycin and treated with doxycycline for 4 days. RNA was harvested and qRT-PCR analysis for Oct4 (A), Sox2 (B) and Nanog (c) expression was carried out. Data was normalized to Gapdh and relative expression in wt EC P19 cells was set to 1. Mean and s.d from PCR triplicates are shown.

3.4 Discussion

The aim of this chapter was to test two different doxycycline-inducible lentiviral systems for the ability to knock down Hmgn2 in mouse epiblast carcinoma stem cells (P19 EC cells). The pSLIK vector system was used initially. However, the lack of a constitutively expressed selection marker on this vector made it hard to select transduced cells. FACS sorting for the doxycycline-induced expression of GFP was not particularly helpful, and although a dox-induced reduction in *Hmgn2* expression was observed in cells expressing the N2-1 shRNA, a comparable reduction in *Hmgn1* expression

was also observed. This suggests that the *Hmgn2* knockdown was not specific in this system.

In an attempt to improve on the pSLIK system, the pTRIPZ lentiviral vector was tested. The key difference is that pTRIPZ constitutively expresses the gene for puromycin resistance, so transduced cells can be selected. Three constructs containing different shRNAmirs were chosen to knockdown Hmgn2: shHMGN2-381, shHMGN2-968 and shHMGN2-N2-1. It was found that concentration of the viral supernatant using a centrifugal filter unit enabled more efficient transduction. Selection of transduced cells with puromycin followed by 4 – 6 days of doxycycline treatment resulted in up to 50% of cells expressing the RFP lentiviral marker. However, there was considerable variation in transduction efficiency between vectors, which reduces the utility of this system. It is possible that the puromycin selection was not strong enough, or that transduction with a higher viral MOI would improve efficiency. Attempts to repeat these transductions, both by this researcher and by others, were not successful.

It was observed that expression of RFP was maximal with 4 – 6 days of doxycycline treatment, but decreased significantly by day 8. Furthermore, freezing transduced cells and subsequent recovery resulted in much lower levels of RFP expression, further reducing the utility of this system. Several studies have demonstrated epigenetic silencing of lentiviral transgenes in P19 cells. Specifically, DNA methylation has been shown to silence CMV, EF1α and SFFV promoters in less than 17 days, whereas the A2UCOE promoter was not epigenetically silenced, even after 44 days (He at al, 2005, Zhang et al, 2010). The doxycycline-responsive TRE promoter found in pTRIPZ and pSLIK includes the CMV promoter, so it is possible that epigenetic silencing caused the reduction in RFP expression observed after 8 days of doxycycline treatment, and the loss of RFP expression following cryopreservation.

Hmgn2 expression was assayed at the mRNA and protein level to determine whether shRNA-driven knockdown was apparent in pTRIPZ-transduced cells. QRT-PCR indicated some reduction in *Hmgn2* expression in dox-treated shHMGN2-381 and shHMGN2-986 cells. Western blotting suggested

knockdown of Hmgn2 protein in dox-treated shHMGN2-381 and shHMGN2-N2-1 cells but not shHMGN2-968 cells. Immunofluorescence indicated downregulation of Hmgn2 protein in RFP-positive shHMGN2-381 and shHMGN2-N2-1 cells.

Previous work has suggested that knockdown of Hmgn2 in P19 cells using siRNA leads to a reduction in the expression of the pluripotent markers, *Oct4*, *Sox2* and *Nanog* (Mohan, 2012). In order to explore this further, expression of these three genes was assayed in the pTRIPZ cells. The data revealed that Oct4 and Nanog expression was not affected in the different cell lines, but that expression of Sox2 was downregulated in dox-treated shHMGN2-968 and shHMGN2-381 cells.

3.5 Conclusions

In summary, although the pTRIPZ system is superior to the pSLIK system in that transduced cells can be selected more easily, the overall percentage of pTRIPZ cells expressing the RFP marker did not exceed 50%. Western blots and immunofluorescence indicated significant knockdown of Hmgn2 in shHMGN-N2-1 and shHMGN2-381 cells, but this was not consistent with mRNA expression data from qRT-PCR assays. The inability to freeze and recover pTRIPZ-transduced cells proved to be a major obstacle to further analyses of these cells. In the next chapter, I present data using an alternative lentiviral vector, pGIPZ, which constitutively expresses shRNAmir and GFP.

Chapter 4 The optimization of pGIPZ constitutive lentiviral knockdown Hmgn2 in P19 EC and Hepa cells

4.1 Introduction: the pGIPZ lentiviral vector

In parallel with the pTRIPZ work described in the previous chapter, pGIPZ plasmids that constitutively express shRNAmir were also tested. PGIPZ expresses the reporter gene GFP, the puromycin resistance gene, and shRNAmir from a constitutive CMV promoter (Figure 4.1). PGIPZ used lentivirus shRNA vector have high transduction effciency and it can transduced non-dividing cells such as neurons besides other primary cells (Hu and Luo, 2011).

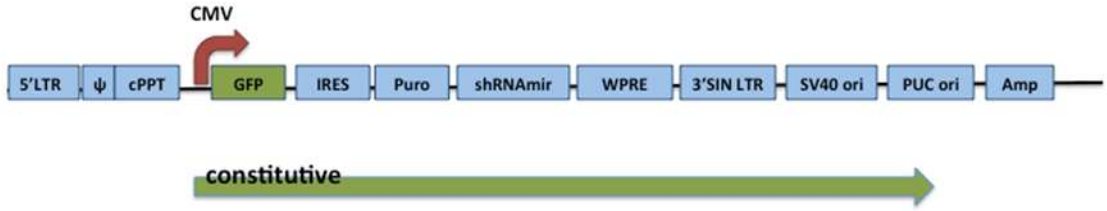


Figure 4.1 schematic diagram illustrating the expression cassette encoded by pGIPZ lentiviral vector
Bi-cistronic element (IRES) transcript encoding green florescent proteins (GFP) and puromycin drug selection and shRNAmir driven by CMV promoter. Abbreviation: IRES, internal ribosome entry elements; WRE, woodchuck response element; LTR, long terminal repeats; puro, puromycin.

From previous results, pTRIPZ and pSLIK showed to have a lack of stability under doxycycline influence and FACS and mRNA detected low transduction efficiency. Another issue was reported that UBC promoter is a type of Pol III promoter that showed to have less lentiviral transduction efficiency over CMV promoter (Hu and Luo, 2011 and Shin et. al., 2006). Therefore, due to the difficulties in knockdown P19 EC cells, a decision was made to knockdown P19 EC cells using constitutive system pGIPZ as this may provide better knockdown over pTRIPZ and pSLIK.

4.2 aims

The principle aims of this chapter are to:

1. Applying constitutive lentiviral system to knockdown Hmgn2 in P19 EC and Hepa cells.
2. Comparing the pGIPZ performance between cancer stem cells (P19 EC cells) and hepatoma cells (Hepa cells)
3. Identify the quick silenced transgene after the transduction due to the methylated promoter

4.3 Results

PGIPZ vectors used Five Lentiviral pGIPZ plasmids were obtained from Open Biosystems/Thermo Fisher Scientific. They were transformed into competent cells (DH5alpha), and then midi DNA purification was performed.

shHMGN2-381	targets Hmgn2
shHMGN2-514	targets Hmgn2
shHMGN2-968	targets Hmgn2
shGAPDH-391	targets Gapdh, positive control
shEG2-480	targets EG2, positive control
sh-346	non-silencing, control

4.3.1 pGIPZ transduction of Hepa cells

PGIPZ plasmids were transfected into HEK293T/17 cells along with packaging mix in order to generate viral particles. Cell supernatant containing viral particles was harvested after 2 days and added to P19 cells in various dilutions. Puromycin was then added in order to select for transduced cells.

However, initial attempts to transduce pGIPZ lentiviral particles into P19 cells had a very low success rate, with no cells surviving after puromycin selection. In order to investigate whether the problem was due to poor lentiviral production, or if P19 cells are particularly hard to transduce, the

same lentiviral supernatant was used to transduce the murine Hepa cell line.

Transduction of Hepa cells by pGIPZ was much more efficient than the P19 cells. Constitutive GFP expression from the pGIPZ vector was monitored by fluorescence microscopy (not shown) and FACS analysis (Figure 4.2 and 4.3). Even before puromycin selection was applied, five out of the six vectors tested showed transduction of more than 60% of cells (Figure 4.2). After 2 – 4 days of selection, the percentage of cells expressing GFP increased to more than 80%.

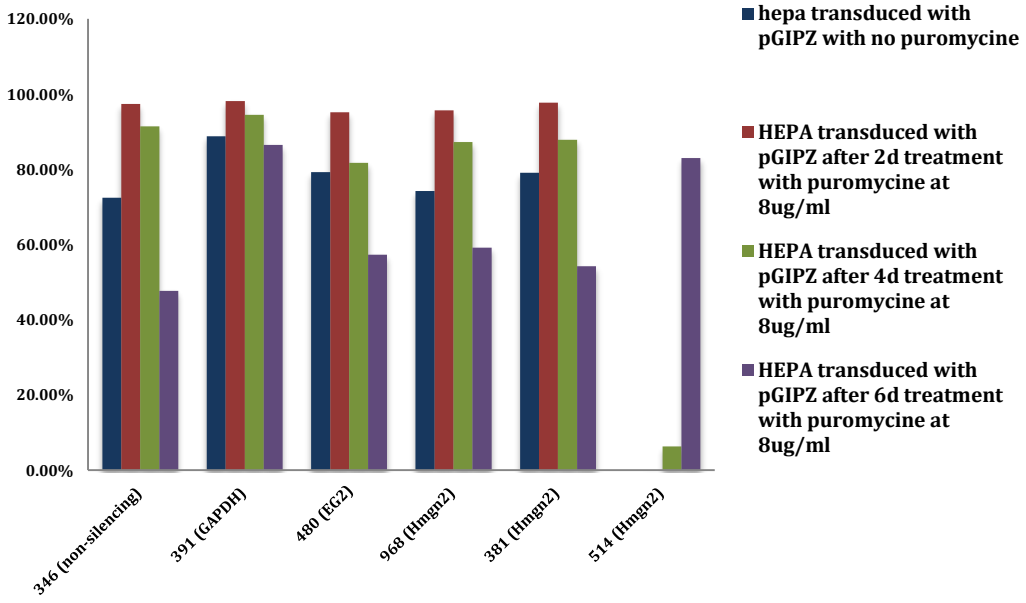
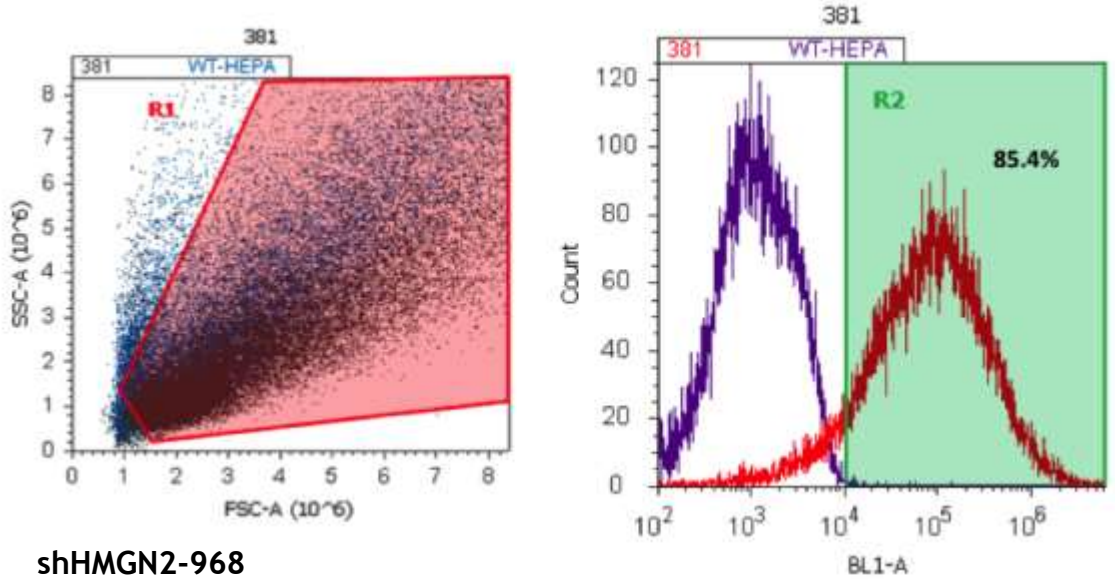


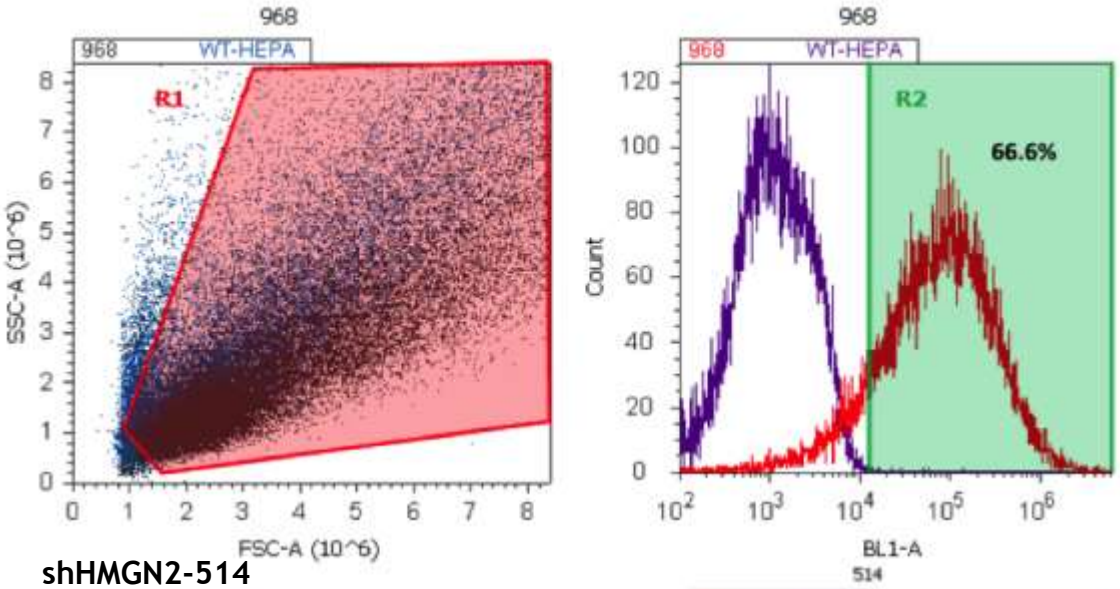
Figure 4.2 Percentage of Hepa cells expressing green fluorescent protein (GFP) following pGIPZ transduction and selection

Hepa cells were transduced with viral supernatant, expanded, then puromycin added at 8 µg/ml for 0, 2, 4 or 6 days. The percentage of cells expressing GFP was quantified by FACS analysis. 346 is negative control, 391 and 480 is positive control and shRNAmir-Hmgn2 sequence triggers are 968, 381 and 514.

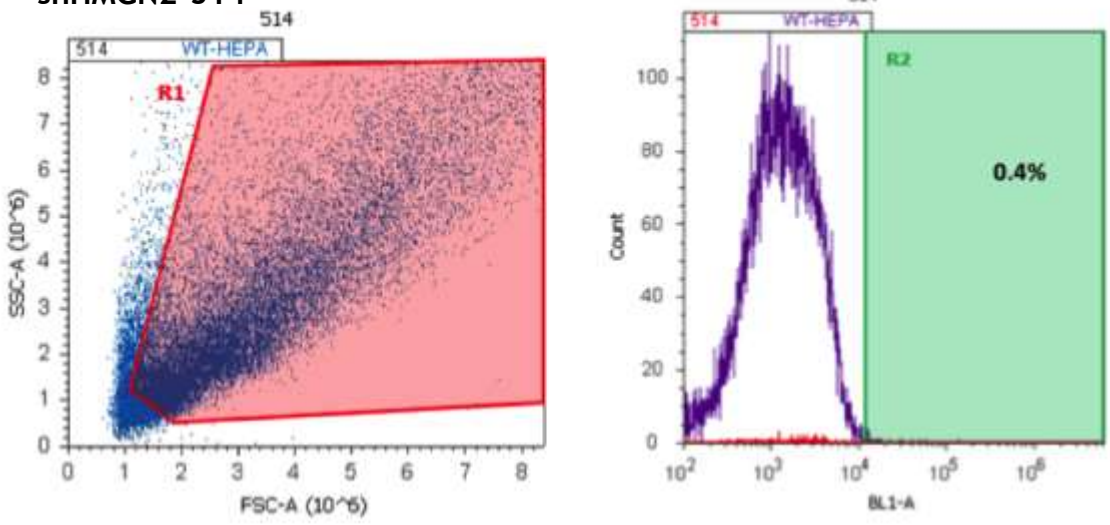
shHMGN2-381



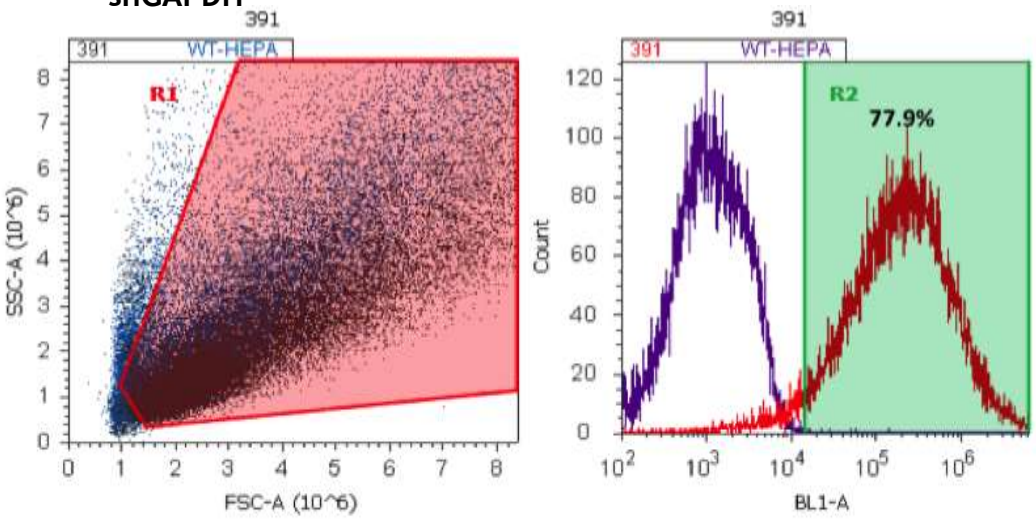
shHMGN2-968



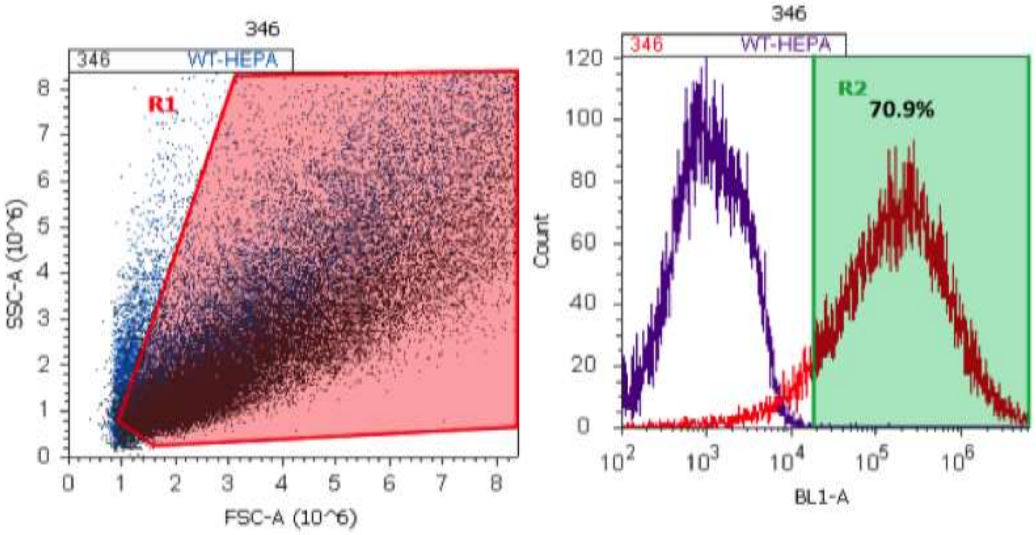
shHMGN2-514



shGAPDH



shNon-silencing



shEG2

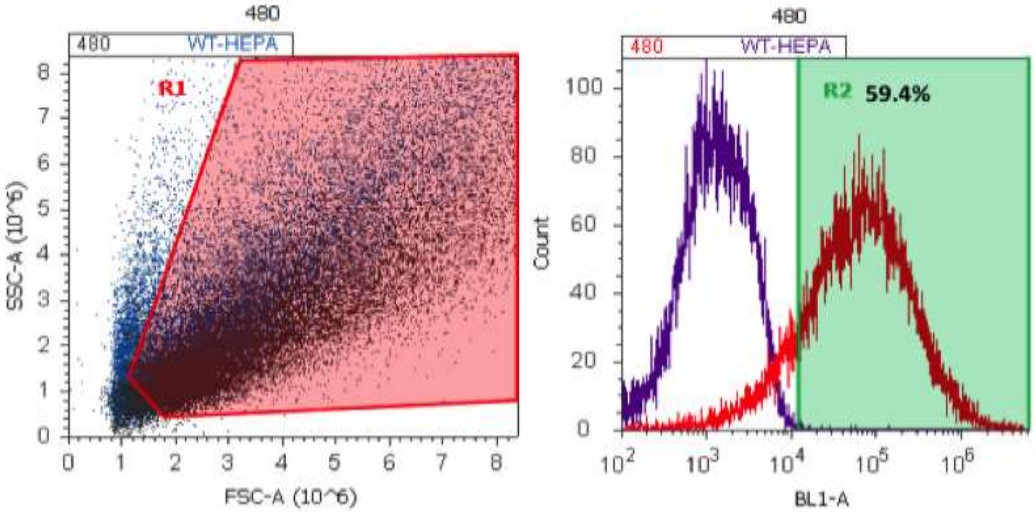


Figure 4.3 FACS analysis of pGIPZ-transduced Hepa cells

Transduced cells were selected with 8 μg/ml puromycin for two days prior to analysis. Each pair of graphs corresponds to a different pGIPZ vector. ShRNAmir-Hmgn2 vectors are 381, 968 and 514, whereas 391 and 480 are positive controls and 346 is the negative control. The dot plots on the left of each pair represent the health and size of the transduced cells (black dots) comparing with wt Hepa cell lines (blue dots). The x axis shows the forward scatter (shape) of the cells and the y-axis shows the side scatter (size) of the cells. The histogram plot on the right shows GFP expression in transduced cells (red line) comparing with WT P19 cells (purple line) in the R1 gate. The x axis represents the green channel from 10^2 to 10^6 and the y axis represents the number of cells.

4.3.2 pGIPZ transduction of P19 cells

The observation that the pGIPZ lentiviral particles could transduce Hepa cells efficiently implied that a much higher multiplicity of infection is required to transducer P19 cells. Consequently, and as described in Chapter 3 for the pTRIPZ particles, 12ml of pGIPZ viral particles from 10 cm dish were concentrated with a centrifugal filter by 350ul.

P19 cells were transduced with concentrated virus, expanded, and selected for up to 6 days with 2.5 μg/ml puromycin. GFP expression was assayed every two days (Figure 4.4 and 4.5) FACS results showed 30-90% GFP expression for all vectors except for pGIPZ-shHMGN2-514. Indeed, pGIPZ-shHMGN2-514 cells showed low survival and low GFP expression in both Hepa and P19 cells (Figures 4.2, 4.3, 4.4 and 4.5), so these cells were not analysed further.

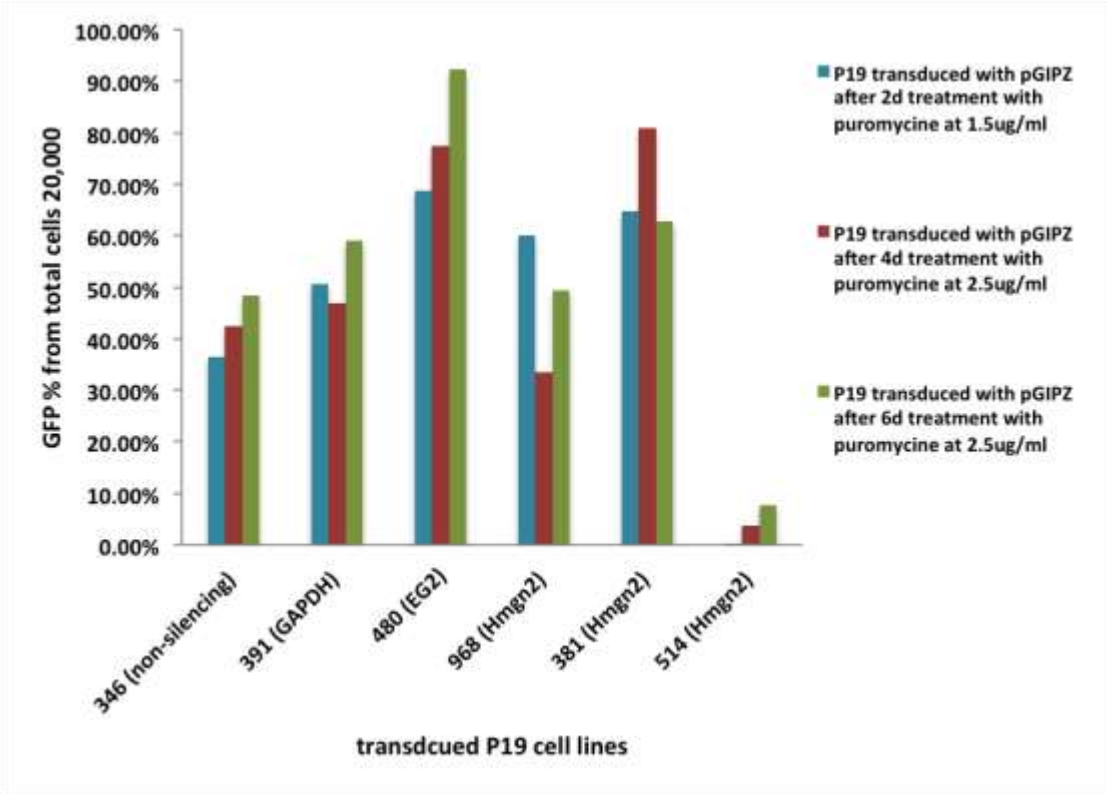
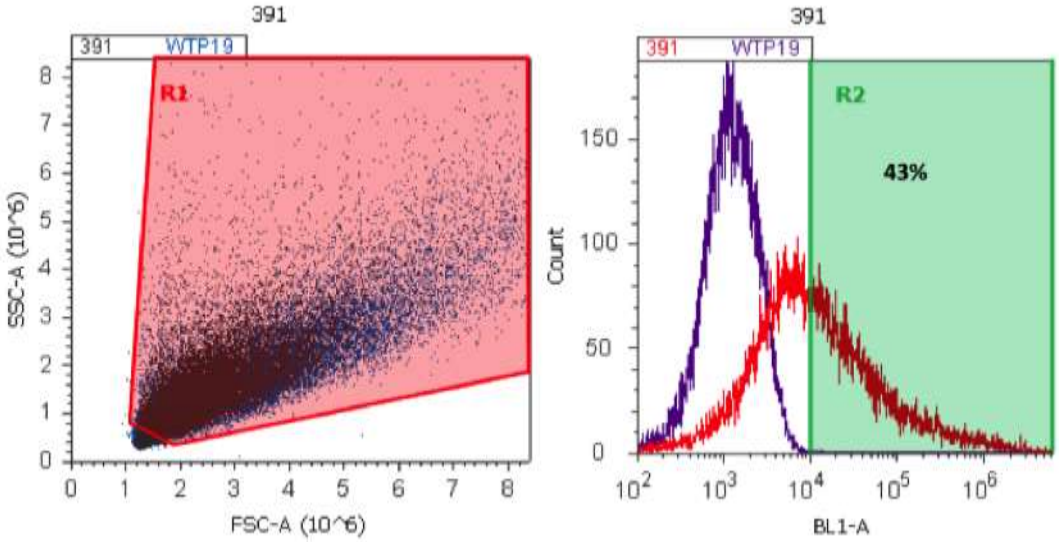


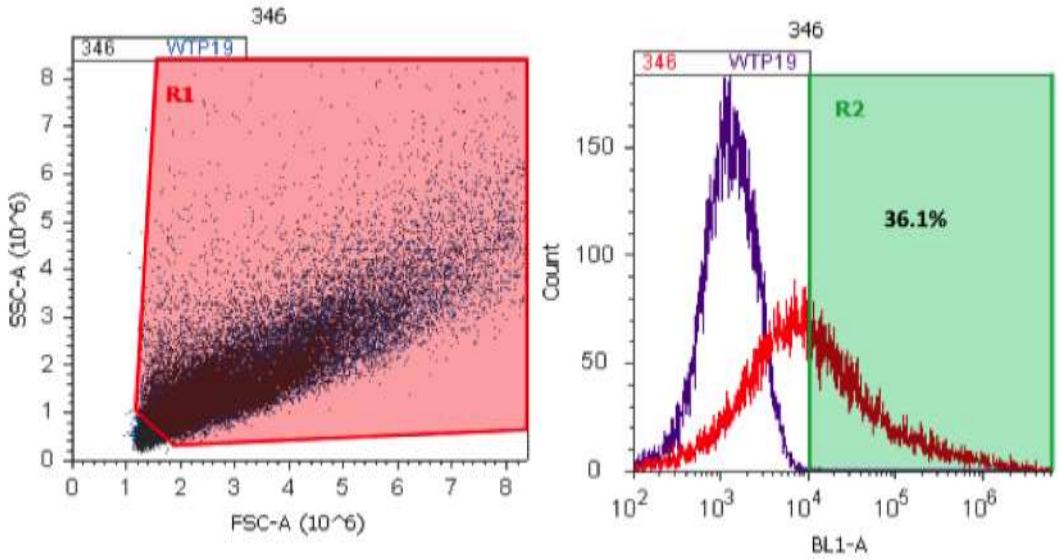
Figure 4.4 Percentage of P19 cells expressing green fluorescent protein (GFP) following pGIPZ transduction and selection

P19 cells were transduced with viral supernatant, expanded, then puromycin added at 2.5 µg/ml for 0, 2, 4 or 6 days. The percentage of cells expressing GFP was quantified by FACS analysis. 346 is negative control, 391 and 480 is positive control and shRNAmir-Hmgn2 sequence triggers are 968, 381 and 514.

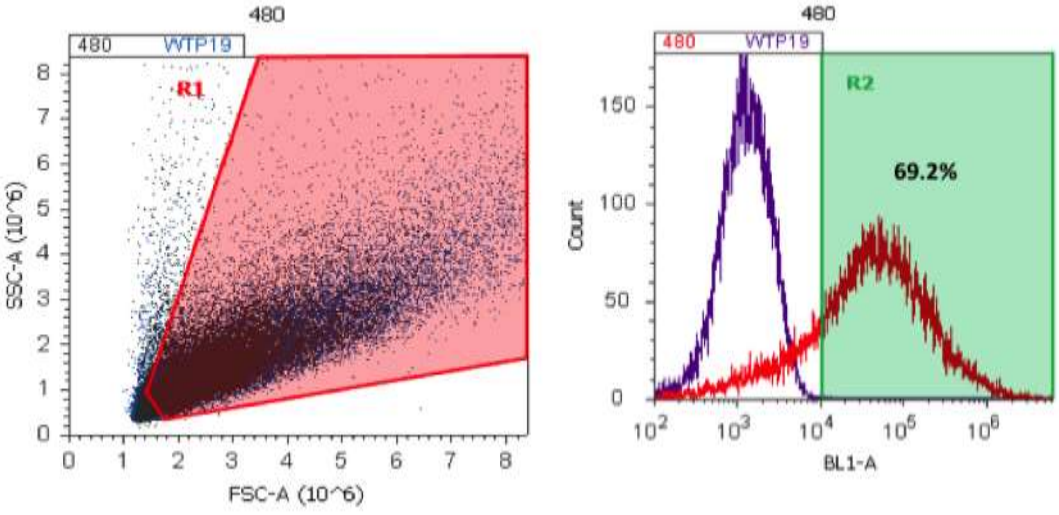
shGAPDH



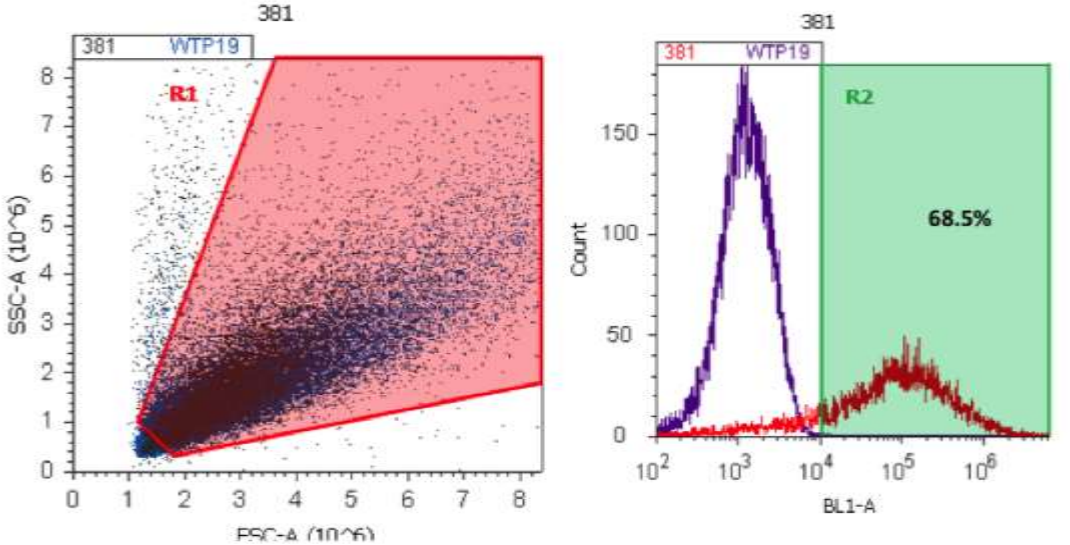
Shnon-silencing



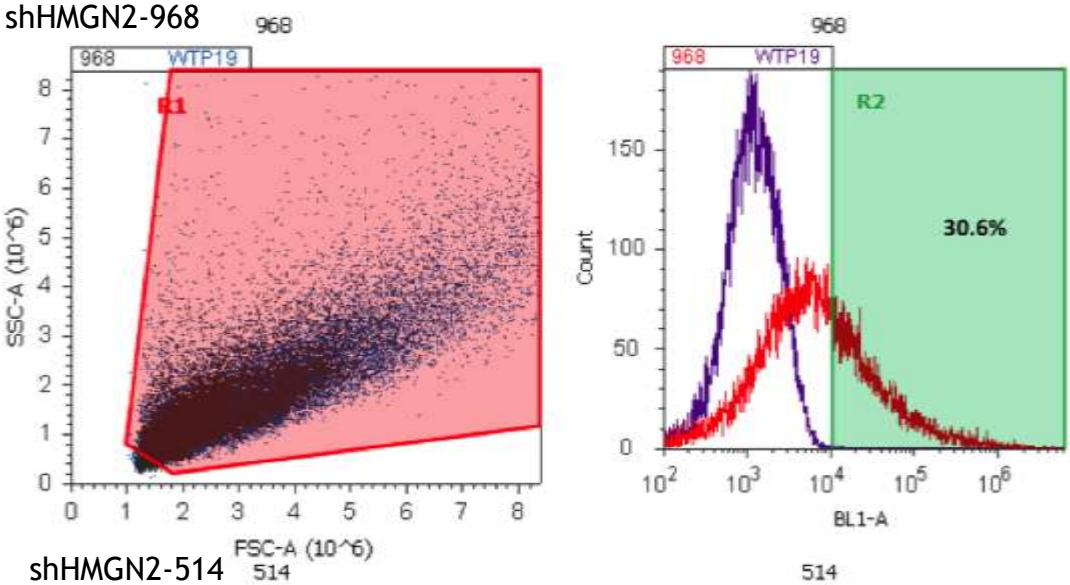
shEG2



shHMGN2-381



shHMGN2-968



shHMGN2-514

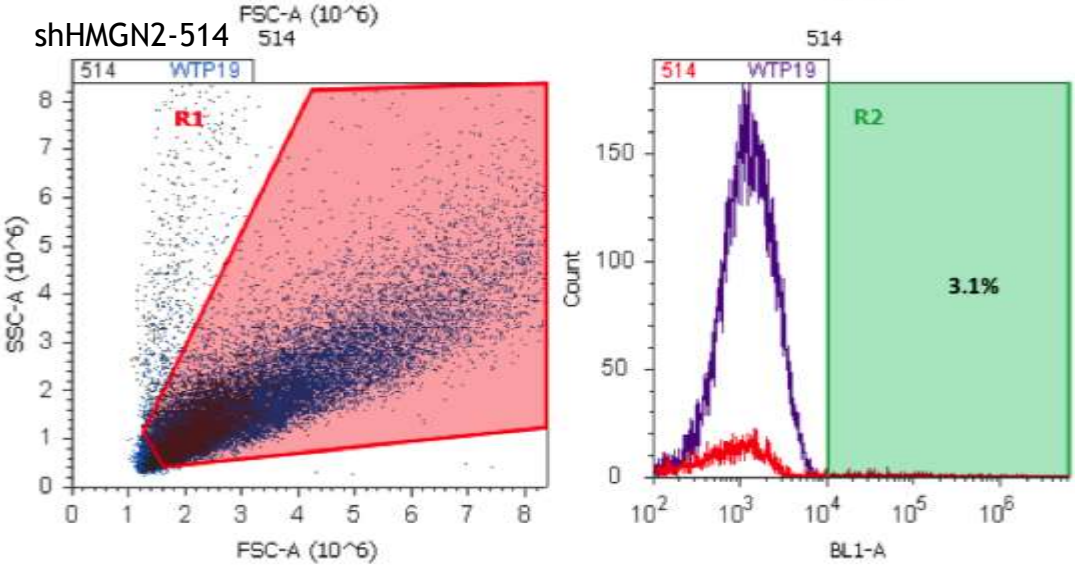


Figure 4.5 FACS analysis of pGIPZ-transduced P19 cells

Transduced cells were selected with 2.5 µg/ml puromycin for four days prior to analysis. Each pair of graphs corresponds to a different pGIPZ vector. shRNAmir-Hmgn2 vectors are 381, 968 and 514, whereas 391 and 480 are positive controls and 346 is the negative control. The dot plots on the left of each pair represent the health and size of the transduced cells (black dots) comparing with wt Hepa cell lines (blue dots). The x axis shows the forward scatter (shape) of the cells and the y-axis shows the side scatter (size) of the cells. The histogram plot on the right shows GFP expression in transduced cells (red line) comparing with WT P19 cells (purple line) in the R1 gate. The x axis represents the green channel from 10^2 to 10^6 and the y axis represents the number of cells.

The ability of the different pGIPZ vectors to knockdown Hmgn2 expression was assayed using qRT-PCR (Figure 4.6). It can be seen that expression of Hmgn2 in shHMGN2-381 cells was reduced by over 70%, whereas expression in the positive and negative controls was unaltered. shHMGN2-968 did not cause a knockdown of Hmgn2 expression.

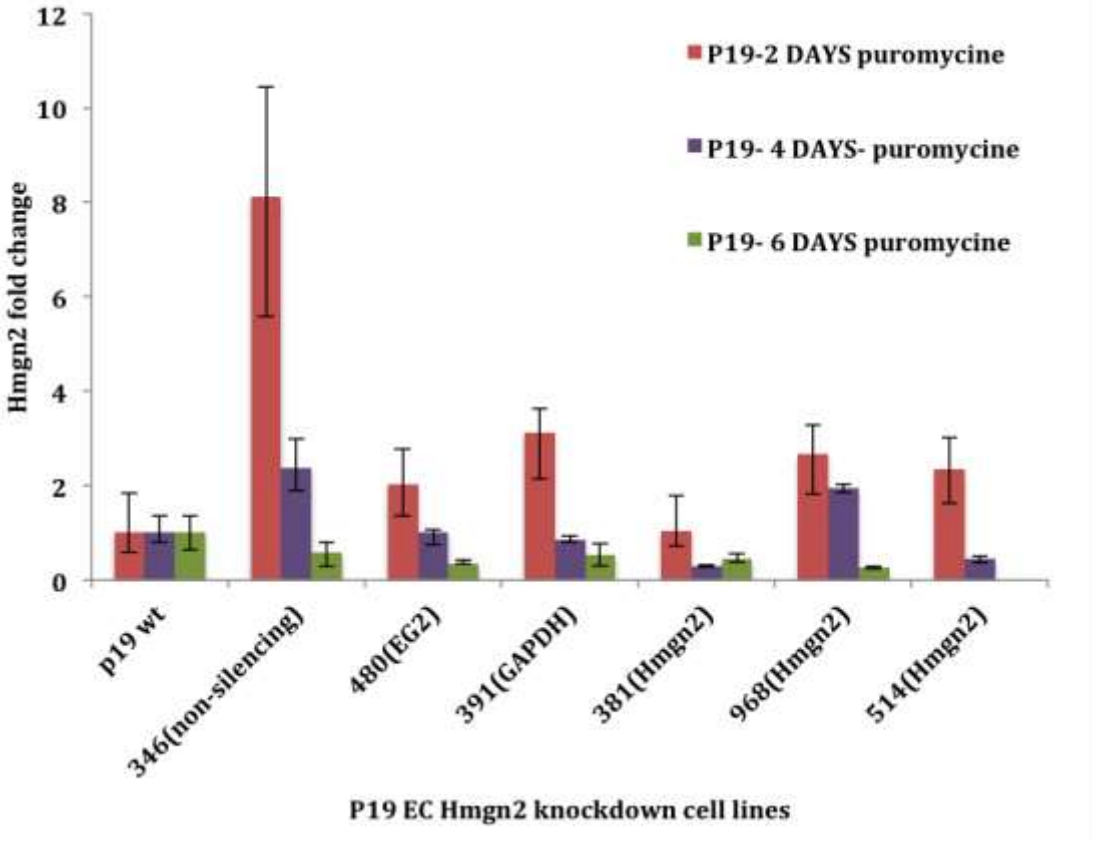


Figure 4.6 Hmgn2 expressions in transduced P19 cell lines with PGIPZ lentiviral constitutive system

The transduced cells were normalized based on WT P19 cells. Three controls are 346 (negative control), 391 and 480 (positive controls) and three shRNAmir-Hmgn2 sequence triggers are 381, 968 and 514. The y-axis presents the fold change from delta CT value. Each RT-analysis represents the day of Collection after the Puromycin selection. Day 2 (A) shows no knockdown in Hmgn2 in P19 cells whereas Day4 (B) and Day6 (C) shows knockdown in 318 around 60% to 65%. For 968 the knockdown did not appear until day 6 for around 90% whereas 514 shows 60% knockdown in day 4 but it failed to survive until day 6.

Figure 4.6 Hmgn2 expressions in P19 cell lines transduced with PGIPZ lentiviral vectors. Cells were selected for four days with 2.5 μg/ml puromycin. Hmgn2 expression was assayed using real time RT-PCR and normalised to actin. Relative expression compared to wt cells is plotted on the y-axis. Data represent the mean and standard deviations of PCR triplicate reactions. Controls are 346 (negative control), 391 and 480 (positive controls) and three shRNAmir-Hmgn2 sequence triggers are 381, 968 and 514.

We observed on several occasions that when P19 cells carrying pTRIPZ or pGIPZ vectors were frozen then thawed; expression of the reporter gene was greatly reduced. This is illustrated in Figure 4.7 Before the cryopreservation, FACS analysis showed greater than 50% GFP-positive cells in various pGIPZ cell lines. After thawing and puromycin selection for four days, the fraction of GFP positive cells was less than 2%. To investigate whether the GFP reduction was due to promoter silencing, we treated the cells with 5-Aza-2'-deoycytidine. We found that two days of treatment with this drug increased the percentage of GFP-positive cells to 15-20%, and five days of treatment further increased the percentage to 20-50% (Figure 4.7). This suggests that the CMV promoter at pGIPZ expression vector can be silenced by DNA methylation during cryopreservation.

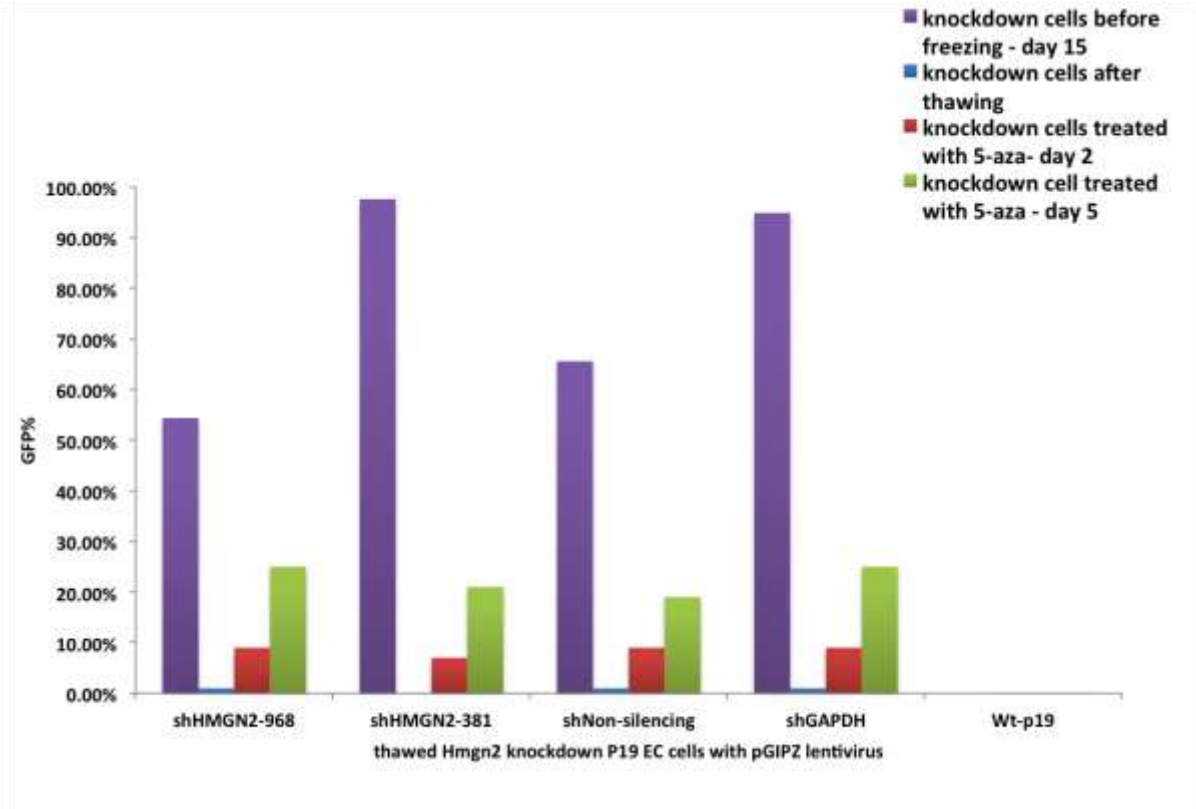


Figure 4.7 The effect of 5-Aza-2-deoxycytidine on the percentage of GFP-positive cells following cryopreservation.
pGIPZ-transduced cells were grown in 2.5 µg/ml puromycin for 15 days, then cryopreserved. After thawing, cells were grown for 3 days, before addition of 2.5 µg/ml puromycin for 4 days. 5-aza-2-deoxycytidine (8 µM) (Hofmann et. al., 2005) was added for 2 days (red) and 5 days (green), or not at all (blue).

4.4 Discussion

In this chapter, the lentiviral vector system pGIPZ was tested. After optimization of media, transfection conditions, drug selection and concentration of viral particles, high efficiencies of P19 cell transduction were obtained. In contrast, Hepa cells were much more easily transduced, and did not require concentrated viral particles.

Cribbs and his colleagues (2013) compared the concentrated with unconcentrated Lentiviral particles. They demonstrated that concentrated Lentiviral particles decreased 10 fold in viral collection. This is indicated that there was a loss of virus during the ultracentrifugation process.

However, they found increase of target gene called NGFP (nerve growth factor receptor) expression when they transduced their primary T cells under concentrated lentivirus. They suggested that concentrator eliminates several conditions from the viral particles supernatant. Consequently the ultracentrifugation with a reduction of centrifuge speeds from 90,000 g to 20,000 g can increase the transduction efficiency and stability. Therefore, optimizing the centrifugation speed is critical to produce high viral titer with high transfection stability (2013).

Comparison of the pSLIK, pTRIPZ and pGIPZ vector systems in chapter 3 and 4 highlights the importance of choosing an appropriate lentiviral system for the cell type being studied. Key properties include:: (i) shRNA expressed in the form of shRNAmir, so a polymerase II promoter can be used to drive expression (ii) a strong, or regulatable, polymerase promoter to drive shRNAmir expression (iii) a single vector containing all doxycycline-regulatory machinery, with no need for specific cell line (iv) a reporter gene (GFP or RFP) to be used as an indication of transduced cells (Meerbrey et. al., 2011) (v) a selectable marker such as the puromycin resistance gene.

In chapter 3, it was discussed how epigenetic silencing of the TRE promoter in pSLIK and pTRIPZ vectors might be contributing to the loss of reporter gene expression over time and during cryopreservation. Loss of reporter gene expression was also observed during cryopreservation of cells transduced with pGIPZ. Treatment of these cells with 5-Aza-2'-deoycytidine reactivated GFP expression, suggesting that the loss of GFP expression was due to DNA methylation of the CMV promoter, as has been previously observed in P19 cells (He et al 2005, Zhang et al 2010). One way to address this problem in future would be to replace the CMV promoter in the pGIPZ vector with a promoter that is resistant to epigenetic silencing, such as A2UCOE or CAG (Zhang et al 2010)

Both pTRIPZ and pGIPZ use an IRES (internal ribosomal element) to allow bicistronic expression of the puromycin resistance gene. In pTRIPZ, the

construct is *rtTA3*-IRES-*puro* and in pGIPZ, the construct is *GFP*-IRES-*puro*. Ibrahimi and his colleagues (2009) found that IRES is less efficient and more cell type-dependent than the self cleaving peptides T2A and P2A at allowing production of two proteins from one transcript. Poor functioning of the IRES element in P19 cells could result in low expression of the puromycin resistance gene. This would mean the multiple copies of the viral vector per cell are required to generate resistance, and so a higher concentration of virus is needed to obtain efficient transduction. Changing the IRES element to P2A or T2A in the pTRIPZ and pGIPZ vectors could improve transduction efficiency in future experiments.

4.5 Conclusion remark

The purpose of this chapter was to optimize the lentivirus system and to address the difficulties in transduced P19 cell lines. Due to the difficulties to transduced stem cells, Choosing the efficient Lentiviral system for stem cells depend on the following properties: (i) the backbone of the vector is miR-30 and these miRNA expressed from strong promoter such as polymerase II promoter (ii) using single vector with no need for specific cell line (iii) in inducible Lentiviral system, the basal transcription profiles of shRNA should be high in polyclonal population (iv) the reporter gene (GFP or RFP) is used as an indication of transduced cells (Meerbrey et. al., 2011) (v) puromycin consistency depend on bicistronic vectors from IRES can be problematic with a replacement of T2A or P2A can increases the gene expression efficiency (Ibrahimi et. al.,2009). However, creating a stable robust transduced cell with lentivirus can be challenging for many reasons: (i) constitutive system can generate gene silencing during cell culture for more than one month (ii) inducible system can cause leaking expression due to the rtTA random positions in the cell lines (iii) freezing and thawing the transduced cell in constitutive system pGIPZ can cause the gene silencing. Therefore, a further experiment need to preform to modify the plasmids and transduced P19 EC cells.

Chapter 5 The functional role of Hmgn2 in embryonic carcinoma stem cells using a constitutive knockdown

5.1 Introduction

Chapters 3 and 4 have described the testing of pSLIK, pGIPZ and pTRIPZ lentiviral vectors for the knockdown of Hmgn1 and/or Hmgn2 in P19 embryonal carcinoma cells. To summarise: we were unable to get high levels of pSLIK transduction, mainly due to the lack of a selection marker on the lentiviral vector. Although pTRIPZ carries a puromycin selectable marker, we were again unable to obtain cell populations with more than 50% of cells expressing the RFP reporter. In contrast, with pGIPZ we were able to obtain 80-90% of cells expressing the GFP reporter. Knockdown of Hmgn2 was observed with vectors carrying shRNAmir against Hmgn2 in several instances. However, the data was not always consistent, possibly due to variable infection levels. The difference between pGIPZ and pTRIPZ is likely to be the strength of the promoter driving the reporter-shRNAmir construct: in pGIPZ, the promoter is CMV, whereas in pTRIPZ it is the doxycycline-regulated TRE promoter.

With both pTRIPZ and pGIPZ, we found that expression of the reporter was greatly reduced by freezing and recovering the cells. This expression could be rescued to certain extent by treating the cells with the DNA methylation inhibitor, 5-azacytidine, indicating that the lentiviral vectors may be epigenetically silenced during the freeze-thaw process. This phenomenon significantly impacted on our ability to analyse the phenotype of the knockdown cells, as we could not go back to previous cultures to collect more samples.

At this stage, a decision was made to focus on the pGIPZ system, as this was giving the highest levels of transduction. This chapter describes how viral particles were made from control and experimental vectors in parallel, and P19 cells were transduced, selected and expanded to allow the collection of FACS data, RNA, protein and chromatin samples at different time points.

5.2 Aims

1. To establish a constitutive microRNA-mediated Hmgn2 knockdown system in EC P19 cells using the pGIPZ lentiviral approach.
2. Investigate the cell surface marker expression in knockdown Hmgn2
3. Investigate the expression of putative Hmgn2 target genes upon knockdown Hmgn2.
4. Further examine other pluripotent marker in Hmgn2 knockdown cells.

5.3 Results

5.3.1 Generation of P19 EC cells transduced with pGIPZ lentiviral vectors

Figure 5.1 shows the experimental timeline for generating P19 EC cells transduced with the pGIPZ constitutive shRNAmir30 system. The plan is divided into five parts: transfection, transduction, expansion, puromycin selection and sample collection. Each stage was optimised to achieve a high transduction level.

Transfection: Five pGIPZ vectors were used for constitutive expression of GFP-shRNAmir.

shHMGN2-381:	targets Hmgn2
shHMGN2-514:	targets Hmgn2
shHMGN2-968:	targets Hmgn2
shGAPDH:	targets Gapdh, positive control
sh-346:	non-silencing, negative control

Lentiviral expression vectors (pGIPZ) and packaging vector mix were transfected into HEK293/17 cells, and medium containing viral particles was collected 48 hours later.

Viral particles were concentrated 30 fold using centrifugal concentrators. P19 EC cells were transduced with these lentiviral particles for 36 hours in

ultracultural media without serum, before transferring to standard P19 media with serum. Cultures were then expanded for seven days to generate enough cells for the selection process.

Puromycin was added at 1.5 µg/ml on day 0. On day 5, the puromycin concentration was increased to 2 µg/ml, and on day 15 it was increased to 2.5 µg/ml. This gradual increase in selection pressure is intended to eliminate cells with low levels of lentiviral gene expression, thus enriching for cells with high expression. GFP expression was monitored by FACS every five days, and protein and RNA samples were collected at the same time points. At the final time point, day 25, chromatin was prepared from the cells (Figure 5.1)

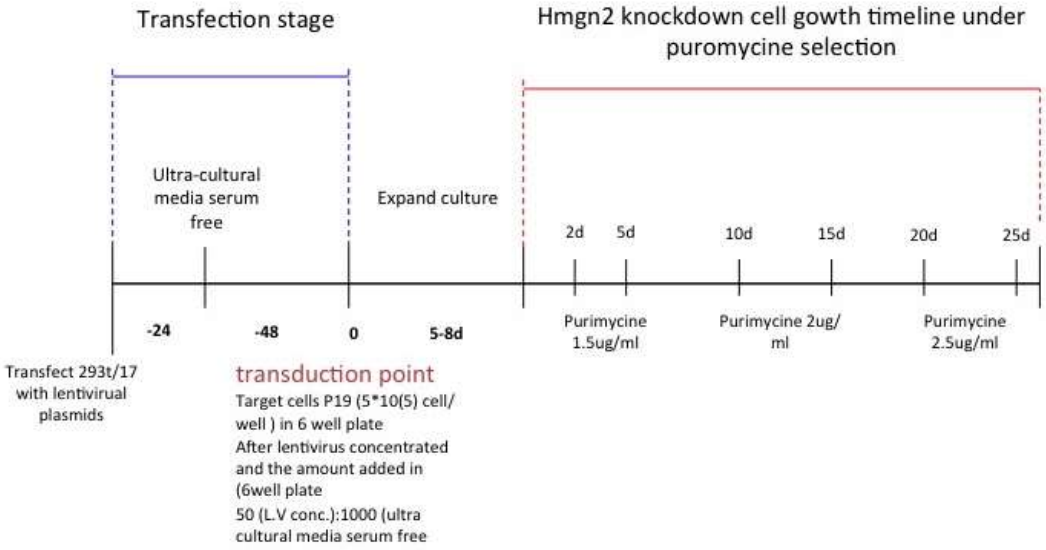


Figure 5.1 Experimental model for generating the P19 knockdown cell lines using the pGIPZ lentiviral constitutive system

Lentiviral particles were generated and collected from host HEK293T/17 cell lines. Concentrated viral particles were transduced into P19 cells, and the culture expanded for 7 days. Puromycin was added at 1.5 µg/ml on day 0. On day 5, the puromycin concentration was increased to 2 µg/ml, and on day 15 it was increased to 2.5 µg/ml. Every five days a FACS analysis, protein and RNA were extracted and collected for further analysis.

5.3.2 FACS analysis on transduced P19 EC cell line

FACS was used to assay for GFP expression over the time course of the experiments (Figure 5.2). It can be seen that the percentage of cells expressing GFP rises as selection pressure is increased, leading to a maximum of 97% GFP in the GAPDH and 381 cells. 968 and 381 had 72% and 53% GFP respectively, whereas the 514 cells failed to survive.

Healthy cells were gated based on their forward and side scatter, in order to distinguish them from dying or dead cells (Figure 5.2). It can be seen that after day 2, the percentage of healthy cells was greater than 60% in most samples. This indicates that increasing the selection pressure did not have a detrimental effect on the overall health of the cell cultures.

Sample ID	**Abbrev-iation**	**%GFP**	**% healthy cells**	**%GFP**	**% healthy cells**	**%GFP**	**% healthy cells**	**%GFP**	**% healthy cells**	**%GFP**	**% healthy cells**	**%GFP**	**% healthy cells**	**%GFP**	**% healthy cells**
		No puromycin		**2 days puromycin 1.5 ug**		**5 days puromycin 1.5ug/ml**		**10 days puromycin 2ug/ml**		**15 days puromycin 2ug/ml**		**20 days puromycin 2.5ug/ml**		**25 days puromycin 2.5ug/ml**	
GAPDH	8	1.3%	46%	43%	69%	65%	49%	87%	74%	95%	76%	99%	88%	97%	87%
Non-silencing	10	0.9%	45%	19%	63%	35%	64%	70%	74%	66%	79%	38%	68%	72%	68%
968 (Hmgn2)	11	0.6%	46%	14%	60%	26%	35%	37%	67%	54%	79%	42%	68%	53%	65%
381 (hmgn2)	12	0.4%	44%	42%	75%	53%	65%	Not enough events to be counted		98%	73%	99%	84%	98%	81%
514 (Hmgn2)	13	0.1%	47%	0.1%	66%	0.4%	57%	Dead cells		0%	0%	0%	0%	0%	0%

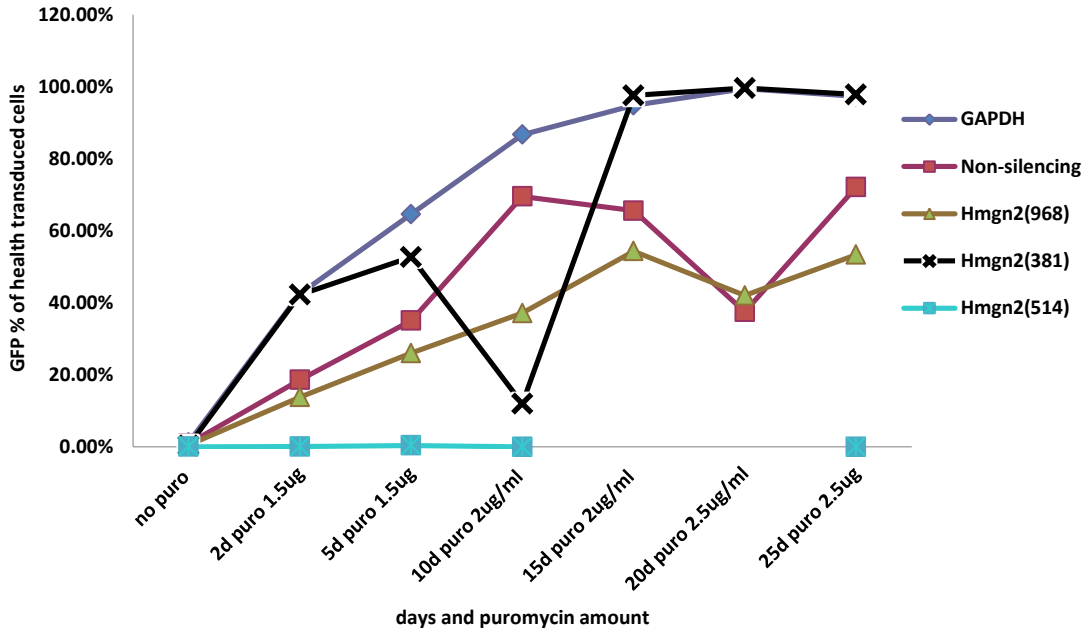


Figure 5.2 Flow cytometer analysis of pGIPZ-transduced cells
A: Percentage of healthy cells within the population, and the percentage of GFP-positive cells within the healthy cell population. B: GFP percentage from part A is plotted against time for each cell line.

5.3.3 Screening the positive transduced knockdown cells using fluorescent microscopy

Transduced P19 EC cell lines were analyzed at day 27 for any changes in gross morphology due to loss of Hmgn2. Fluorescent microscope pictures were taken on live cells on the same day of performing FACs analysis for the cell surface marker SSEA1. Figure 5.3 shows the GFP expression and cell morphology of the transduced cell lines. Cell morphology was not significantly altered in the Hmgn2 knockdown P19 EC cell lines (381 and 968) compared with control shRNA-non-silencing or shGAPDH P19 EC cell lines.

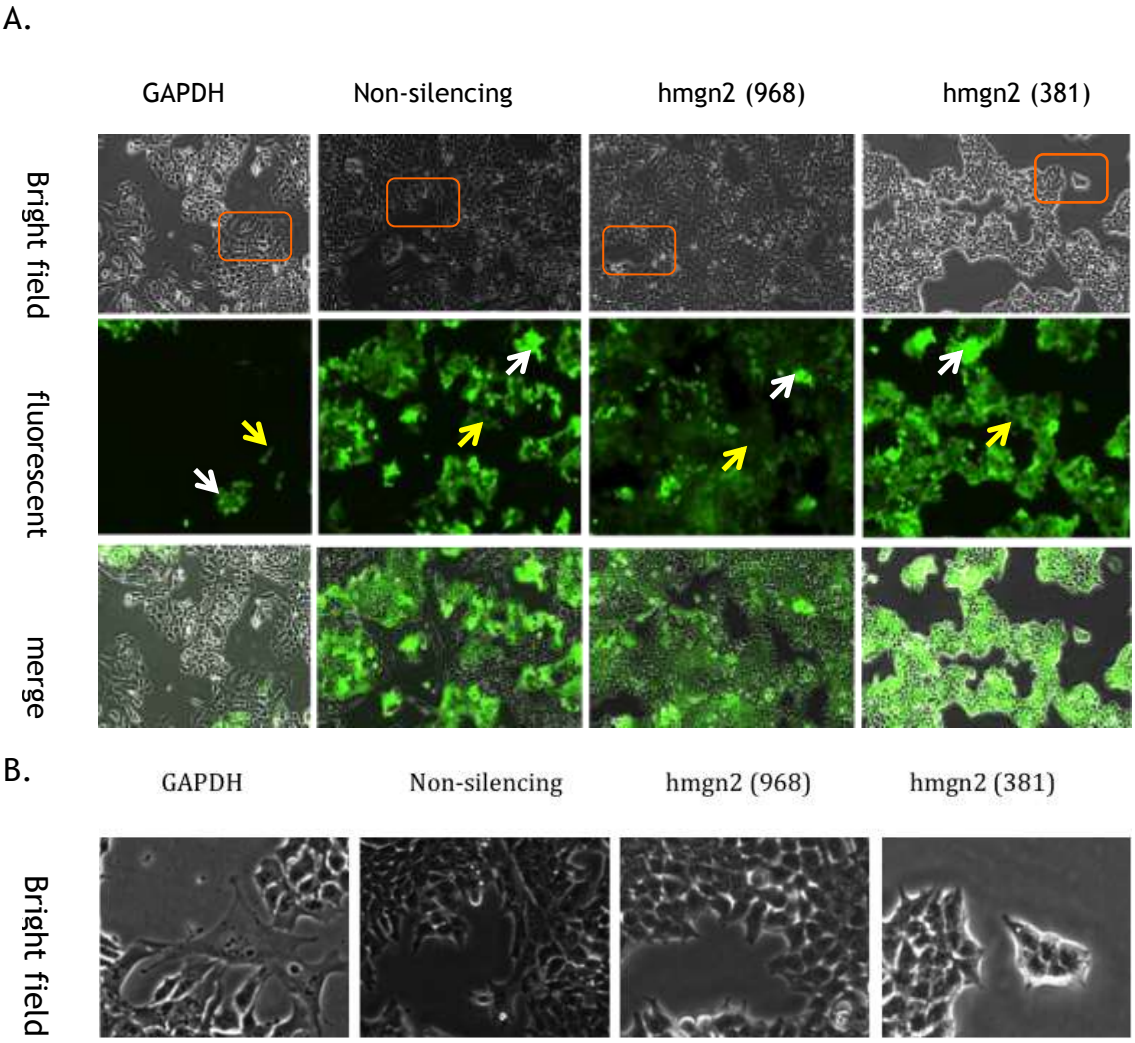


Figure 5.3 Microscopy examination of HMGN2 knockdown P19 cell lines
GAPDH (positive control), non-silencing (negative control), HMGN2 (968) and (381) at x20 magnification. A: White and yellow arrows point to regions of high and low GFP, respectively. B: The regions within orange boxes on the bright field are shown magnified in the lower panels to illustrate cell morphology.

5.3.4 Screening the Hmgn2 knockdown P19 cell lines for possible spontaneous cell differentiation

The pluripotent status of pGIPZ-transduced P19 EC cell lines was assessed using FACS analysis for SSEA-1 pluripotency marker, which is located on the cell surface.

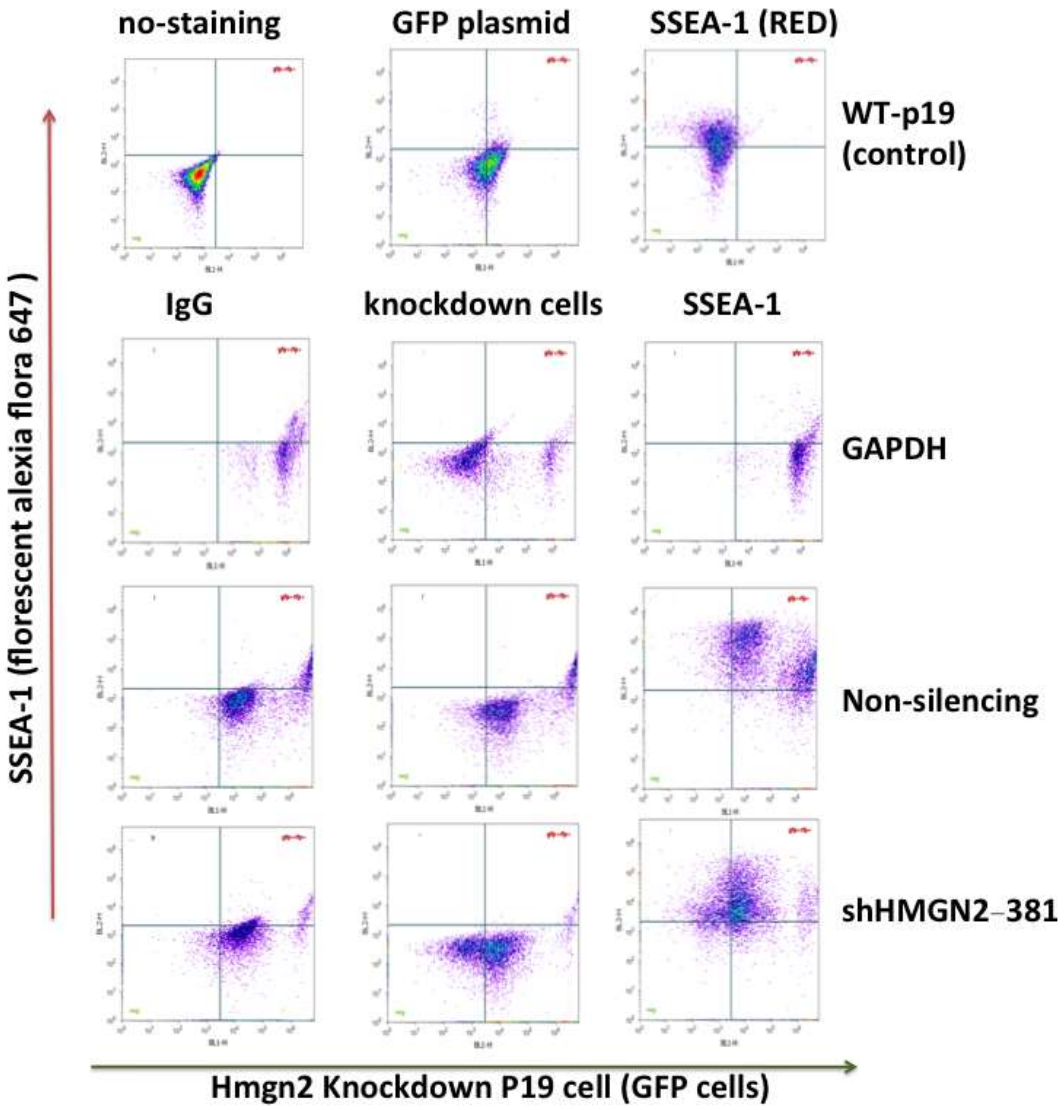
As expected, the majority of GAPDH, non-silencing and shHMGN2-381 cells

expressed GFP. This is shown by most of the cells lying in the right half of each plot, compared with the untreated wt P19 control, where all the cells are in the left half of the plot. Interestingly, the GFP-expressing cells fall into two discrete groups, of cells with low and high GFP. It can be seen that the high intensity green signal of the second group appears to "leak" into the red channel, giving some red signal even in the absence of a red fluorescent antibody (middle column).

The negative control IgG antibody did not change the overall pattern of red and green staining for the shNon-silencing and shHMGN2-381 - transduced cells (compare the left with the middle column). However, most of the shGAPDH cells in the lower left quadrant of the middle plot are absent from the IgG control and the SSEA-1 plot. This phenomenon occurred in both replicates. It could be due to non-specific staining by the IgG and SSEA-1 antibodies giving signal in the green channel, or it could be due to loss of these cells during the washing steps (Figure 5.4A).

Figure 5.4B shows that 74% of wt P19 cells are SSEA-1 positive. Strikingly, in pGIPZ-shNon-silencing and pGIPZ-shHMGN-381 cells that express GFP (ie are carrying the lentivirus), the proportion of SSEA-1-expressing cells is very similar, at 73% and 76% respectively. This suggests that the pluripotency of these cells has not be significantly affected by the lentiviral vector.
In contrast, cells carrying the pGIPZ-shGAPDH vector and expressing GFP are only 4% SSEA-1 positive. This could represent a loss of pluripotency in cells that have a knockdown in GAPDH expression. However, given the concerns about the IgG control in the shGAPDH cells, it is not possible to draw any firm conclusions.

A.



B.

Sample name	%SSEA1 positive cells	%GFP cells	%GFP cells that are SSEA1 positive
WT-P19 (CONTROL)	74%	n.a	n.a
GAPDH(Positive control)	5%	99%	4%
Non-silencing (Negative control)	78%	79%	73%

shHMGN2-381	88%	59%	76%

Figure 5.4 FACS analysis for GFP and SSEA-1 expression in pGIPZ-knockdown P19 cells

A: Dot plots, where the x-axis represents the (BL1) green channel and corresponds to GFP expression. The y-axis represents the (BL2) red channel and corresponds to SSEA-1 expression. Each horizontal row corresponds one cell type: WT-P19, GAPDH , Non-silencing and shHMGN2-381. In the top row, the left hand plot is unstained cells, the middle plot is cells transiently transfected with a plasmid expressing GFP, and the right plot is cells incubated with anti-SSEA-1. These plots were used to set the gates for the four quadrants. In the other three rows, plots on the left column show cells incubated with the IgG control. The middle plots are unstained cells, and the right plots are cells incubated with anti-SSEA-1. The analysis was performed twice on unfixed, live cells on the day 25. B: the left column is the total percentage of SSEA-1-positive cells, the middle column is the total percentage of GFP-positive cells, and the right column is the percentage of GFP cells that are also SSEA-1 positive.

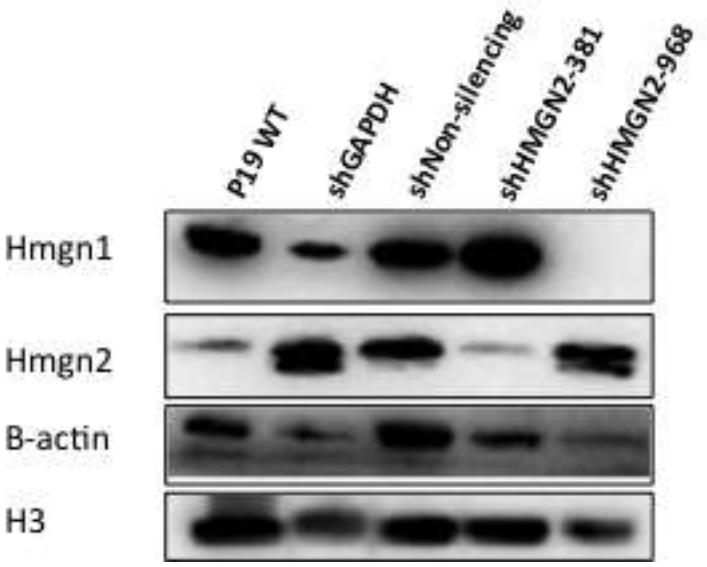
5.3.5 HMGN2 knockdown in undifferentiated mouse embryonic carcinoma EC P19 cells

Western blotting (Figure 5.5A and B) and real time RT-PCR (Figure 5.5C) were performed to examine HMGN1 and HMGN2 expression in pGIPZ-transduced P19 EC cells. Using whole cell extracts, western blotting was performed and normalized based on histone H3 (Figure 5.5A and B). In cells carrying pGIPZ-shHMGN2-381, expression of Hmgn2 was reduced at both the RNA and the protein level, indicating successful knockdown. Expression of Hmgn1 in this line was similar to wt. levels.

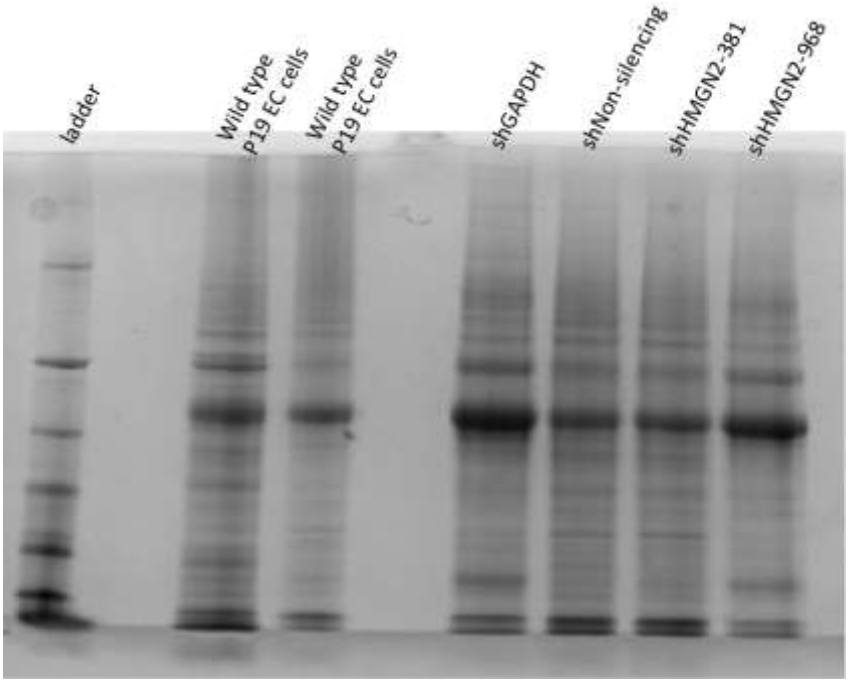
Hmgn2 expression was not reduced in pGIPZ-shHMGN2-968 cells, even though this vector also targets Hmgn2. Hmgn1 protein levels were reduced in these cells, although Hmgn1 mRNA was unaffected.

In pGIPZ-shNon-silencing cells, Hmgn1 and Hmgn2 mRNA was reduced, but protein levels were unaffected. Similarly, in pGIPZ-shGAPDH cells, Hmgn2 mRNA was reduced but HMGN2 protein was unaffected.

A.


P19 WT
shGAPDH
shNon-silencing
shHMGN2-381
shHMGN2-968
Hmgn1
Hmgn2
B-actin
H3

B.


ladder
Wild type P19 EC cells
Wild type P19 EC cells
shGAPDH
shNon-silencing
shHMGN2-381
shHMGN2-968

C.

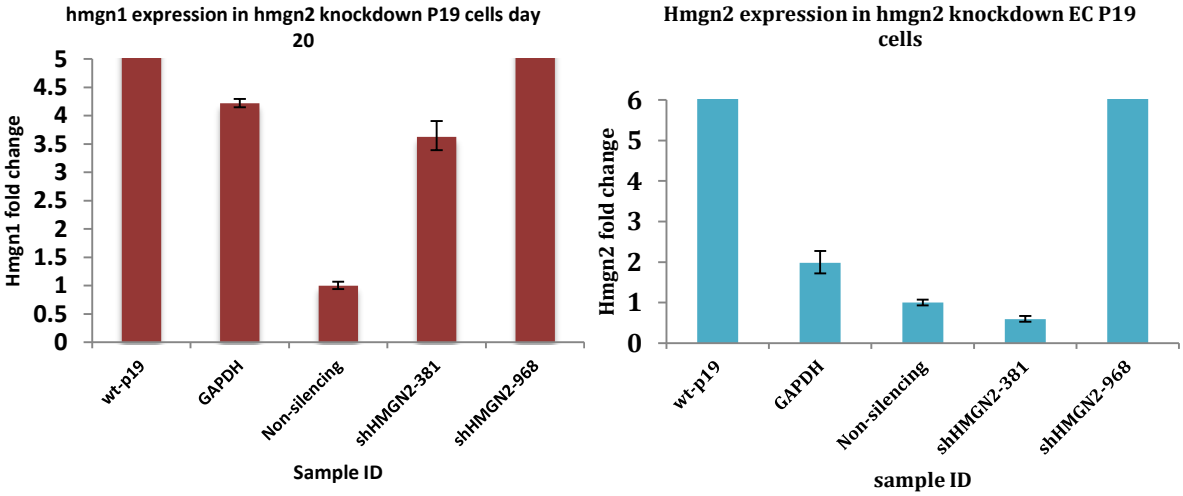


Figure 5.5 Hmgn1 and Hmgn2 expression in pGIPZ-transduced P19 cells after 20 days of selection

A. Western blotting was performed for Hmgn1 and Hmgn2, with H3 as the loading control. B. Coomassie blue gel of normalised whole cell extracts. C. Hmgn1 and Hmgn2 mRNA expression as quantified using real time rt-PCR, with α-tubulin as the control and normalised to the shNon-silencing cells. The error bars represent standard deviations from three qPCR reaction replicates.

5.3.6 HMGN2 knockdown influence the expression of key pluripotency genes

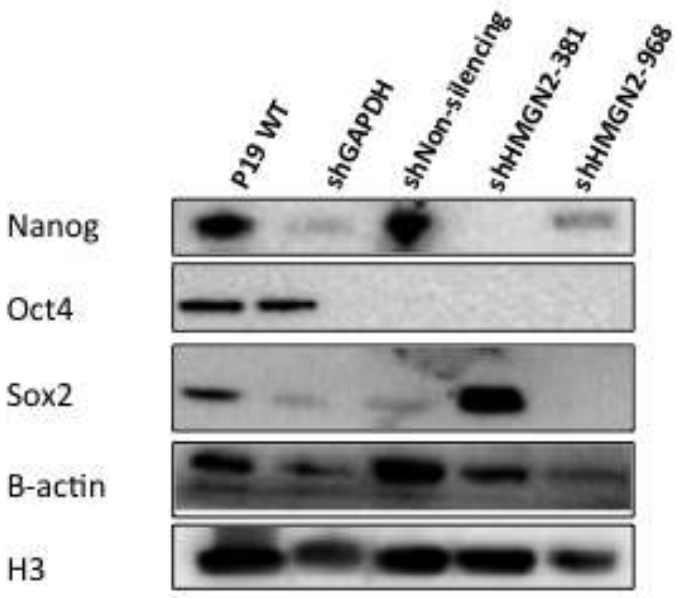
Previous data from the lab had indicated that expression of Nanog, Oct4 and Sox2 was reduced in Hmgn2-knockdown cells. Figure 5.6A and B show gene expression and protein analysis were performed in Hmgn2 knockdown P19 EC cells for Nanog, Oct4 and Sox2.

It can be seen that Oct4 protein and mRNA is reduced in the shNon-silencing cells, as well as in shHMGN2-381 and shHMGN2-968 cells.

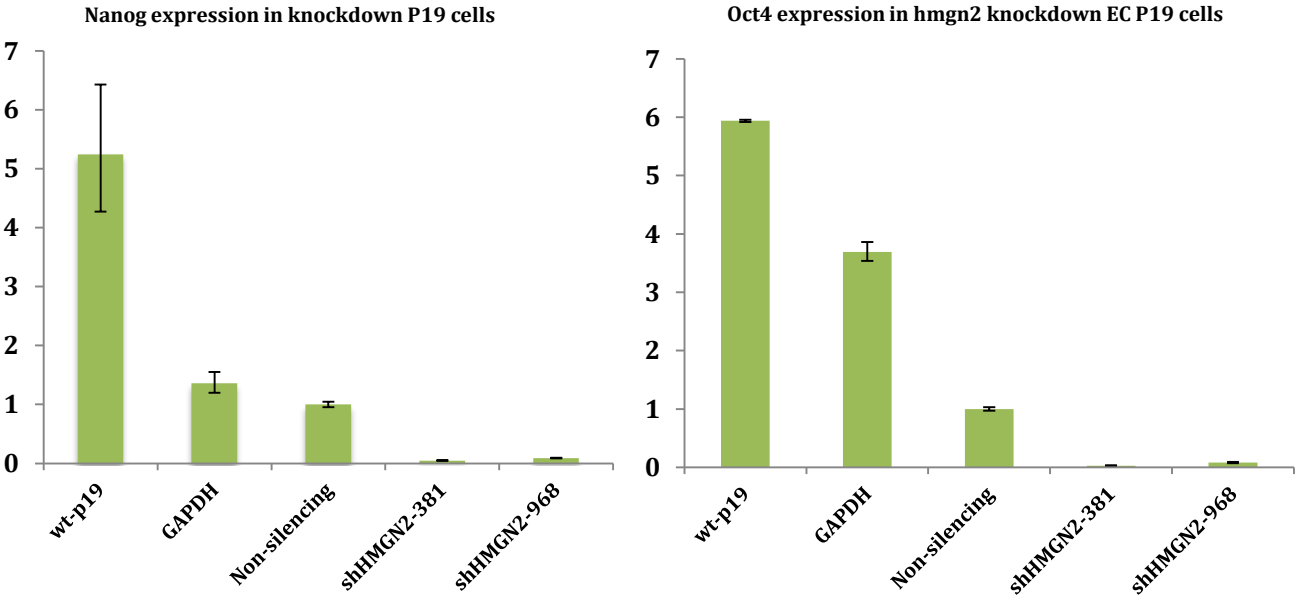
Nanog mRNA and protein is reduced in shGAPDH, shHMGN2-381 and shHMGN2-968 cells, and the mRNA is also reduced in shNon-silencing cells.

Sox2 mRNA and protein is reduced in shNon-silencing and shHMGN2-968 cells, and the proteins are also reduced in shGAPDH cells.

A.


P19 WT
shGAPDH
shNon-silencing
shHMGN2-381
shHMGN2-968
Nanog
Oct4
Sox2
B-actin
H3

B.


Nanog expression in knockdown P19 cells
7
6
5
4
3
2
1
0
wt-p19
GAPDH
Non-silencing
shHMGN2-381
shHMGN2-968

Oct4 expression in hmgn2 knockdown EC P19 cells
7
6
5
4
3
2
1
0
wt-p19
GAPDH
Non-silencing
shHMGN2-381
shHMGN2-968

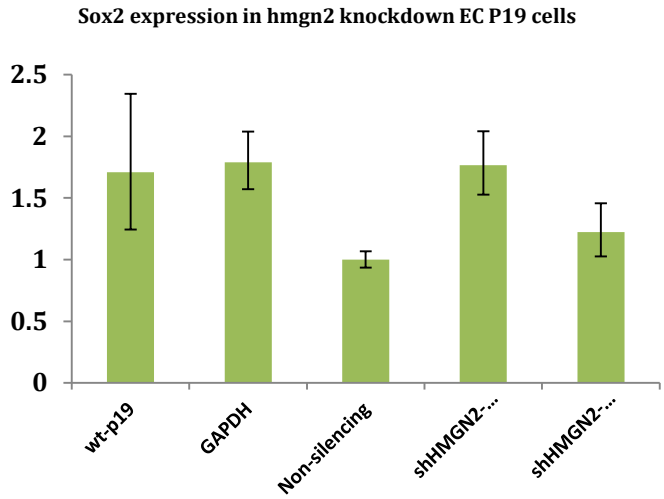

Sox2 expression in hmgn2 knockdown EC P19 cells
2.5
2
1.5
1
0.5
0
wt-p19
GAPDH
Non-silencing
shHMGN2-...
shHMGN2-...

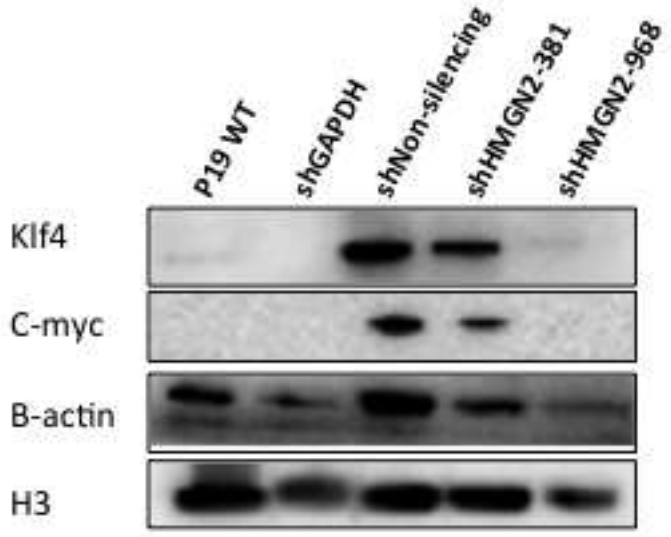
Figure 5.6 Nanog, Oct4 and Sox2 expression in pGIPZ-transduced P19 cells after 20 days of selection

A. Western blotting was performed for Nanog, Oct4 and Sox2, with H3 as the loading control. B. Coomassie blue gel of normalised whole cell extracts. C. Nanog, Oct4 and Sox2 mRNA expression as quantified using real time rt-PCR, with α-tubulin as the control and normalised to the shNon-silencing cells. The error bars represent standard deviations from three qPCR reaction replicates.

5.3.7 Knockdown HMGN2 in EC P19 cells influence other genes expressions that involved in induce pluripotent stem cells formation.

Two additional pluripotency markers were also studied: c-myc and KLF4. Both genes showed increased protein levels in shNon-silencing and shHMGN2-381 cells compared to the other lines. However, levels of c-myc mRNA were not significantly altered, and Klf4 mRNA was increased in shGAPDH and shNon-silencing samples (FIGURE 5.7A and B).

A.



B.

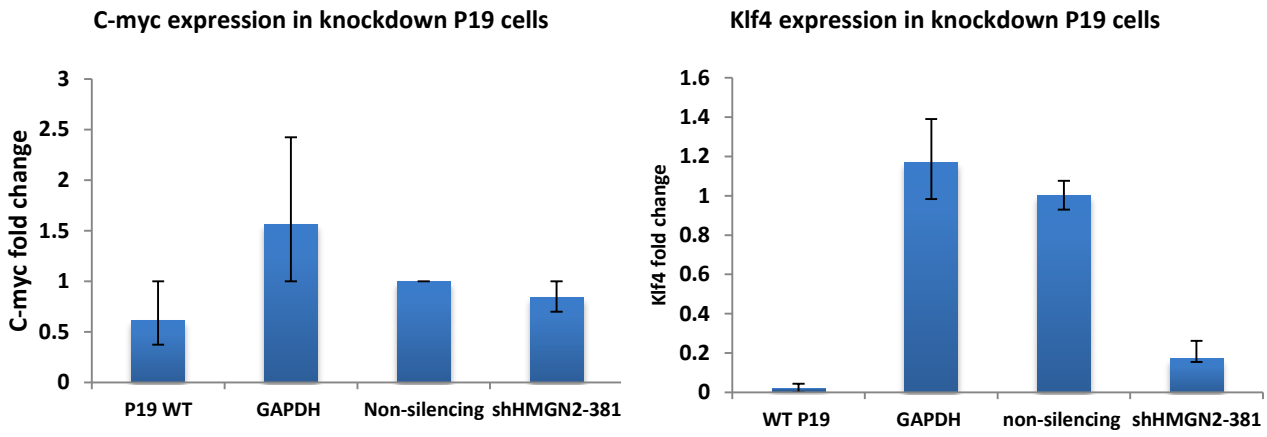


Figure 5.7 Klf4 and c-myc expression in pGIPZ-transduced P19 cells after 20 days of selection

A. Western blotting was performed for Klf4 and c-myc, with H3 as the loading control. B. Coomassie blue gel of normalised whole cell extracts. C. Klf4 and c-myc mRNA expression as quantified using real time rt-PCR, with α-tubulin as the control and normalised to the shNon-silencing cells. The error bars represent standard deviations from three qPCR reaction replicates.

The above data do not indicate specific changes in the expression of Oct4, Sox2, Nanog, c-myc or Klf4 that correlate with the loss of Hmgn2 expression. Changes in gene expression that occur in both Hmgn2-knockdown and control cells may be a result of non-specific effects of shRNA expression, either through off target miRNA-like effects on protein translation, or through activation of the interferon pathway.

Interpretation of the data is complicated by the fact that the mRNA and protein results are not always consistent. This could be due to biological effects on protein translation or stability that are not seen at the RNA level. The differences could also be technical artefacts arising from variable protein extraction efficiency, RNA contamination, or protein extract degradation.

5.4 Discussion

The aim of this chapter was to gain further insight into the effect of Hmgn2 loss in pluripotent and induce pluripotent markers in epiblast P19 EC cells. This was accomplished by applying a constitutive knockdown lentiviral shRNA based on microRNA system.

Constitutive long period of Hmgn2 knockdown in undifferentiated epiblast P19 EC cells revealed three findings: i) long term knockdown for Hmgn2 in epiblast P19 EC cell protocol process, ii) two population bright green and less green of Hmgn2 knockdown P19 EC cells using fluorescent microscope and SSEA1 cell surface staining, iii) the loss of Oct4 and Nanog and gain of Sox2, Klf4, C-myc and Hmgn1 expression using western blot and qRT-PCR. Several experiments attempted to generate Hmgn2 knockdown P19 cell lines (Chapter 3 and Chapter 4). Three stages used to process to achieve the knockdown of Hmgn2 in EC P19 cells: generating viral particles by host cells (HEK293/17 cell lines), P19 EC cell transduced with lentiviral particles and puromycin selection. These three stages depend on four factors: lentiviral plasmid promoters, the health of the HEK293T/17 and EC P19 cell lines, lentiviral efficiency, and puromycin concentration influence. However, after short period of puromycin selection (day 2 and 5), Hmgn2 did not show any evidence of GFP cells. PGIPZ plasmid expression represents by the expression of GFP report gene in transduced cells (green or GFP cells) . GFP cells were observed in day 10, 15 and 20 under increase of puromycin selection start from 1.5, 2 and 2.5 ug/ml. The last day of collection (day25) GFP cells showed a slight reduction in GFP cells in transduced P19 EC cells. This result suggests two points: one is the chance of gaining high Hmgn2 knockdown P19 cell population by increasing the puromycin concentration and one CMV promoter silencing in pGIPZ lentiviral plasmid due to methylation on promoter site in constitutive knockdown system.

Hmgn2 knockdown process revealed two populations after purimycine selection using Cell compensation and fluorescent microscope. Bright and light GFP population can be seen clearly on cell compensation using SSEA1 staining for cell pluripotency. This polyclonal population may affect on the

level of Hmgn2 knockdown and other affected gene expressions. These results suggest that although using puromycin drug selection for increasing the chance for collecting high positive Hmgn2 knockdown, the level of Hmgn2 knockdown still variant due to the differences of viral particles concentration in some cells.

The Knockdown of Hmgn2 may affect significantly by overexpression the Hmgn1 in RNA and protein. However, still unclear if the knockdown of Hmgn1 will gain the same effect on Hmgn2 expression using the same constitutive system. In addition, the knockdown of Hmgn2 did not change the expression of Hmgn3a RNA (data did not shown). These results suggest that the loss of Hmgn2 may be compensated by increase expression of Hmgn1.

The knockdown of Hmgn2 may leads to reduction in the expression of some pluripotent genes Oct4 and Nanog whereas Sox2 was over expressed in Hmgn2 knockdown P19 EC cells in RNA and Proteins. In the induce pluripotent markers, Hmgn2 knockdown P19 EC cells leads to expression of C-myc and Klf4. In protein analysis, it showed that there is slight reduction of Klf4 comparing with negative control non-silencing while in C-myc, it showed that more reduction. In gene expression, it revealed that C-myc and Klf4 had a significant down regulated. Interestingly, wild type P19 EC cells showed less or no expression of C-myc and Klf4 in protein and RNA which can bring skeptical thoughts that wild type epiblast P19 EC cells may not have high expression of Klf4 and C-myc comparing with embryonic stem cells. However, there were some contradictions in the data from western blot and gene expression analysis due to variability levels of protease activity in protein extract (the protein lysis buffer was lab made) and variability level of transduced cells in RNA extract (see appendix Figure S1). on the other hand, These results may recommend two points: i) Hmgn2 may be act as mediator of key pluripotent genes and ii) Hmgn2 can possibly induce the epiblast stem cells to be more embryonic stem cells

In earlier study shows that Furusawa and colleagues (2006) found that Hmgn1 expression was high in epiblast mouse embryogenesis from day E7.5-

E9.5 and fairly strong on E10.5 unlike ectoderm embryonic region. they found that knockdown Hmgn1 in mice enhance the expression of Hmgn2 expression due to the lack of phenotype raising the possibility of homeostatic mechanisms. The level of hmgn2 in knockout Hmgn1 was the same as the level of Hmgn1 over expression mice. In siRNA study Knockdown Hmgn2 in EC P19 cells did not affect the expression of Hmgn family members (Mohan, 2012). On the other side, Furusawa and colleagues finding was match with the presented results here, Hmgn1 was more expressed than Hmgn2 in P19 EC cells in protein analysis. However, in gene expression both Hmgn1 and Hmgn2 wild type P19 EC cells give high level of expression. in Hmgn2 knockdown P19 EC cells the normalization was accomplished by using negative control (non-silencing). In this case non-silencing protein analysis in Hmgn1 and Hmgn2 were relatively high, but it seems that there is more expression of Hmgn1 than Hmgn2 whereas Hmgn2 knockdown in shHMGN2-381 clearly showed over expression of Hmgn1. Two possible hypotheses can be used to explain the data. First, P19 EC cell lines have endogenous high amount of Hmgn1 than Hmgn2 which may affect on the Hmgn2 knockdown P19 EC cells. Second, Hmgn1 compensate for the loss of Hmgn2 to either drive the cells to become more specific or pluripotent cells using novel pathway. Third, technical difficulties by using constitutive lentiviral system based on drug selection may affect on the P19 cells.

to investigate whether the Hmgn2 knockdown cells that have lost the key pluripotent markers ,Oct4 and Nanog, may maintain the cell pluripotency. In presented study here, Protein and RNA analysis showed that knockdown of Hmgn2 affect the Oct4 and Nanog expression while Sox2 expression were up regulated in the same cells by 2 fold change. Moreover, cell cycle analysis was performed on thawed Hmgn2 knockdown P19 EC cell lines using pripeidoum iodide (PI) showed more cell arrested in G1 and cell proliferation rate was higher than wild type P19 EC cells (data did not show). These results suggested two hypothetical questions: one if more cells arrested in the S-G2 phase of the cell cycle, and one whether the Hmgn2 Knockdown lead to cells (epiblast cells are on late blastocysts) commitment to next (early post-implantation) or previous cell state (early blastocysts).

To investigate the first hypothesis, SSEA1 (CD15) were expressed in Hmgn2 knockdown P19 EC cells by almost half of expression comparing with non-silencing

During culturing Hmgn2 knockdown P19 EC cells, there were several points were raised from observing Hmgn2 knockdown cell under the cell microscope and noticing cell proliferation rate. First, the knockdown P19 EC cells seems to grow as colonies but not as monolayers as P19 EC cells. second, the cells were growing very fast as wild type P19 EC cells. these two observation gain indication that the knockdown cells may reserve its pluripotency instead of committed to differentiation state. Klf4 and C-myc are common transcriptional factors to be used in induced pluripotent stem cells beside Oct4 and Sox2. In this study here, Hmgn2 knockdown showed downregulation of C-myc and Klf4 normalized either negative control (non-silencing) or wild type P19 EC cells. in western blot, Klf4 and C-myc have slight reduction comparing with non-silencing.

Nevertheless, these results showed four issues: i) pGIPZ a constitutive plasmid contain CMV promoter, which can be methylated less than a month, ii) western blot data from whole protein extract may have experienced some proteinase activity which showed some degradation in the quality of protein, iii) RNA was extracted from polyclonal population that had variability in GFP level, and iv) from public mouse RNA sequences data showed that shRNA sequences for Hmgn2, 381, 968, and 514 contain pseudogenes which may also explain the variability in gene expression. Therefore, it will be challenging to repeat just the protein extract without repeating the whole experiment. Consequently, analysing the protein content of crosslinked chromatins preparation may be very useful for this scenario.

5.5 Conclusion remark

The results of this chapter show that Hmgn2 may play novel roles in pluripotency and reprogramming of epiblast P19 EC cells. Hmgn2 long knockdown specifically affect the expression of pluripotent and induce pluripotent markers. Further experiments need to carry out before deriving any further conclusion.

Chapter 6 Profiling the endogenous Hmgn2 and Hmgn2 knockdown events across the Oct4 and Nanog region in mouse epiblast P19 EC

6.1 Introduction

Hmgn proteins act by binding to nucleosomes and modulating chromatin structure. They have been shown to alter histone modifications, compete with linker histones and influence enhancer structure (Deng et. Al. 2015 and Subramanian et. al., 2009). A major aim of this project was to investigate whether chromatin structure was altered in P19 cells that have a knockdown in Hmgn2.

Previous work in the lab had indicated that the expression of Oct4, Sox2 and Nanog was decreased in cells with a transient knockdown of Hmgn2 (Mohan 2012). ChIP assays had shown that histone H3 lysine 4 methylation was reduced and H3K27me3 was increased, at the Oct4 and Nanog loci in Hmgn2-knockdown cells (K. West, unpublished). H3K4me3 is typically associated with active promoters, and H3K27me3 is often associated with genes that become silenced, for example during development (Bernstein et. Al., 2006).

The first aim of this chapter was to investigate whether the pattern of histone modifications at the Oct4 and Nanog loci is similar in P19 cells compared to mouse embryonic stem cells. Embryonic stem cells (E14) were chosen as a control, as ENCODE research teams have characterised histone modifications, chromatin features and the binding profiles of many transcription factors across the mouse E14 genome. These ChIP assays should indicate whether the chromatin features in the teratocarcinoma P19 cells are significantly different to the normal mouse ES cells.

Before performing ChIP at the Oct4 and Nanog loci in Hmgn2 knockdown cells, it is important to know whether the expression of these genes is

actually altered. The data presented in Chapter 5 were inconclusive as to whether the expression of these genes was altered in Hmgn2 knockdown cells. There was concern as to whether the whole cell extract method adequately extracts chromatin-bound factors, or whether proteolytic activity had degraded the protein samples. The fact that frozen stocks of these cells cannot be regrown meant that new protein extracts could not be obtained from these cells. In an alternative approach, this chapter presents western blotting of chromatin in which cross links have been reversed. This method reveals the levels of several chromatin-bound proteins in pGIPZ-transduced P19 cells.

Following on from the western blotting of chromatin, ChIP assays were performed to investigate Hmgn binding and histone modifications at the Oct4 and Nanog loci in Hmgn2 knockdown cells.

6.2 Aims

5. Compare the global levels of histone modifications and chromatin-bound Hmgn1 and Hmgn2 in P19 and mES E14 cells using western blotting of crosslinked chromatin.
6. Compare the histone modifications at the Oct4 and Nanog loci in EC P19 cells and ES E14 cells using crosslinking ChIP assay. Examine the global changes in chromatin-bound factors in Hmgn2 knockdown P19 EC cells using western blot.
7. Investigate changes in histone modifications at the Oct4 and Nanog genes in Hmgn2 knockdown cells using ChIP.
8. Investigate the migration of TATA-binding protein using CO-IP in chromatin and whole cell protein extract.

6.3 Results

6.3.1 Preparation of cross-linked chromatin from wild type P19 and E14 cell lines.

An optimised ChIP protocol routinely used in the lab was followed for all ChIP experiments (Barkass et. al., 2012). Crosslinked chromatin was prepared from P19 EC and mES E14 cells, and sonicated to an average fragment size of 300-100 bp ready for immunoprecipitation (Figure 6.1).

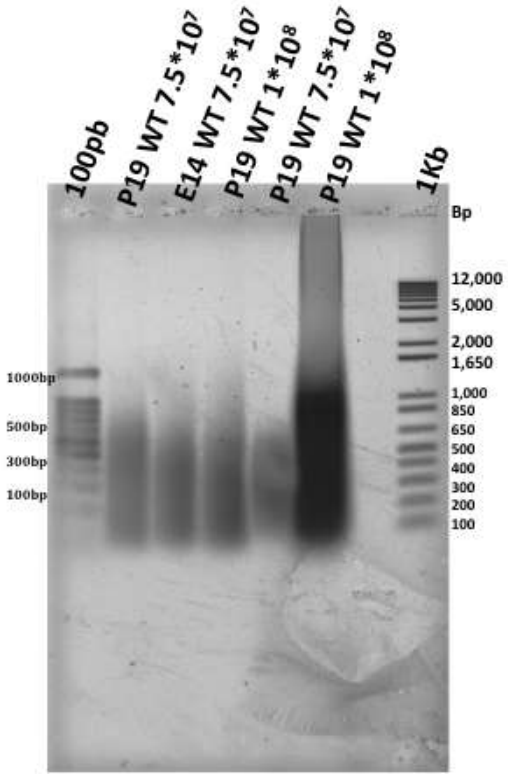


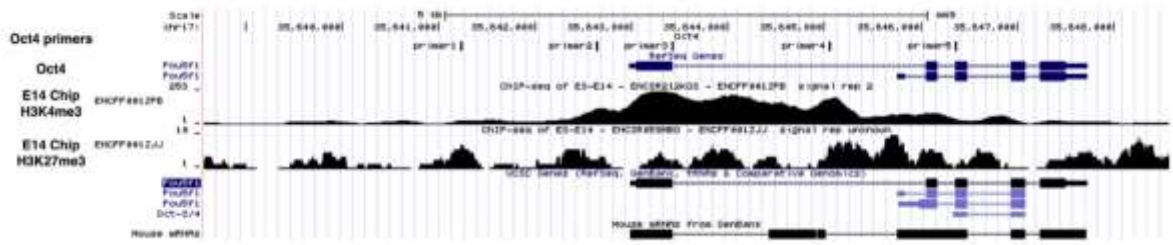
Figure 6.1 Shearing of wild type P19 and E14 chromatin.
Agarose gel electrophoresis of sonicated P19 and E14 chromatin after reversal of crosslinks and purification of DNA.

6.3.2 ChIP primer design

ChIP real time PCR primers were designed against the promoter and transcribed region of the Oct4 and Nanog genes, using ENCODE data as a guide (Figure 6.2) . This data revealed the enrichment of the active promoter mark H3K4me3 active over the transcription starting sites (TSS) of the Oct4 and Nanog genes. On the other hand, the same promoter

regions have no enrichment of the repressive mark H3K27me3 (Figure 6.3 A and B).

A.



B.

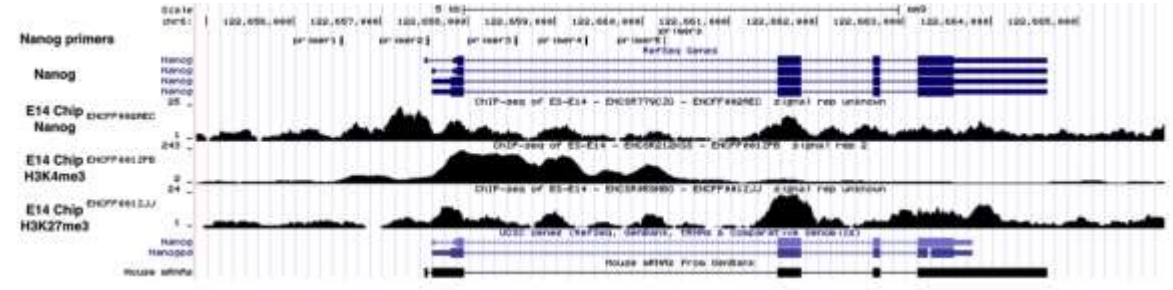


Figure 6.2 UCSC genome browser view of mouse Oct4 (A) and Nanog (B) genes on chromosome 17 and 6.
The data includes the primer sets that were used in QPCR-ChIP experiment. Moreover, the data presents E14 ChIP-seq data for histone (H3K4me3 and H3K27me3) at Oct4 and Nanog regions. Oct4 and Nanog both contain enrichment of H3k4me3 at the transcription starting sites (TSS). (https://www.encodeproject.org).

Primers were designed to encompass the H3K4me3 peak and surrounding region (Table 6.1). A primer set that detects the major satellite (Msat) repeat served as a control for heterochromatin, and a primer for the housekeeping gene (ß-actin) was also used.

Oct4 primer set	Distance from Oct4 TSS	Nanog primer set	Distance from Nanog TSS
Oct1	-1781bp	N1	-1081bp
Oct2	-399bp	N2	-219bp
Oct3	+410	N3	+929bp
Oct4	+2070	N4	+1740 bp
Oct5	+3342	N5	+2636bp

Table 6.1 Location of Oct4 and Nanog loci primer sets

6.3.3 HMGN proteins in E14 and P19 cells

After the chromatin extraction step, western blotting can be performed to ensure the quality of chromatin and detect the bound proteins. Benzonase digestion was used to digest native or heat-denatured DNA and RNA followed by reversal of crosslinks to separate any conjugated proteins. Then samples were ready for western blot application. Figure 6.3 shows western blotting for Hmgn1 and Hmgn2 proteins in wt P19 and E14 cells. The results show that more Hmgn2 is bound to chromatin in E14 cells than in P19 cells, whereas Hmgn1 is present at similar levels in both cell types.

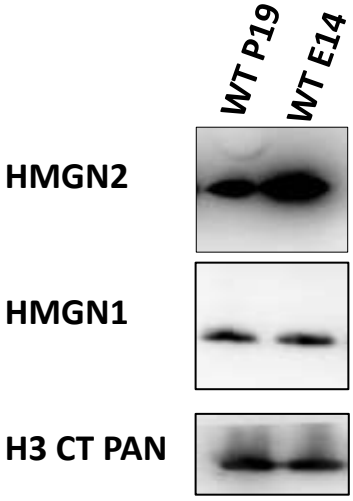


Figure 6.3 Western blotting for Hmgn2 and Hmgn1 in chromatin isolated from P19 EC and E14 cells
The chromatin was treated with benzonase enzyme to remove the DNA, then crosslinks were reversed prior to electrophoresis. H3 is the loading control.

ChIP assays were performed as described previously (Barkess et al 2012) . ChIPs were performed in duplicate and the final samples were combined before analysis. Real time PCR was used to quantify DNA that was immunoprecipitated. Ct values were normalised to input and then to the MSat primer set using the □□Ct method.

ChIP was performed using an antibody to histone H3 as a control (Figure 6.4). This antibody should detect all modified forms of H3, as it recognises an epitope in the C-terminus. In Figure 6.4 it can be seen that enrichment

of H3 is very even across both the Oct4 and Nanog loci in P19 cells. In E14 cells, the enrichment of H3 is slightly lower over the Nanog gene

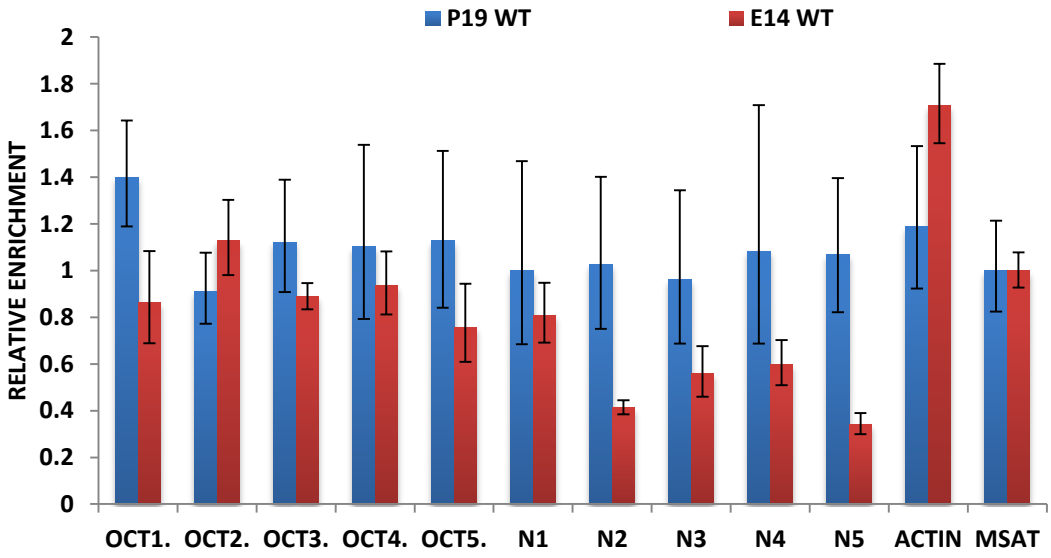


Figure 6.4 H3 enrichment at the Oct4 and Nanog gene loci in P19 and E14 cells

CHIP-QPCR analysis of H3 enrichment at the Oct4 and Nanog loci. Data was normalised to input and the MSat primer set. Error bars represent the errors from PCR triplicates.

Hmgn1 and Hmgn2 binding at the Oct4 and Nanog pluripotent genes is shown in Figure 6.5. On the Nanog gene, Hmgn2 binding is enriched by up to 6-fold compare to the MSat control. In both P19 and E14 cells, the highest level of Hmgn2 binding is at the N2 primer set, which is the one nearest to the TSS (-291 bp).

Similar levels of Hmgn2 binding are seen at the Oct4 locus. Binding is slightly higher at the two primer sets nearest the TSS: OCT2 (-399 bp) and OCT3 (+410bp), but the trend is less clear than it is at the Nanog locus.

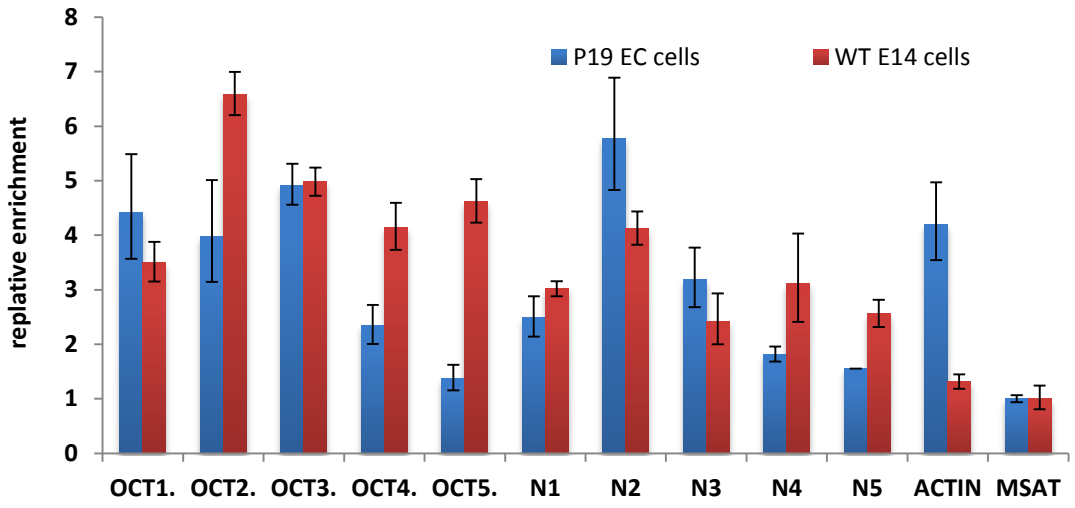


Figure 6.5 Hmgn2 binding at the Oct4 and Nanog gene loci in P19 and E14 cells.

CHIP-QPCR analysis of Hmgn2 enrichment at the Oct4 and Nanog loci. Data was normalised to input and the MSat primer set. Error bars represent the errors from PCR triplicates.

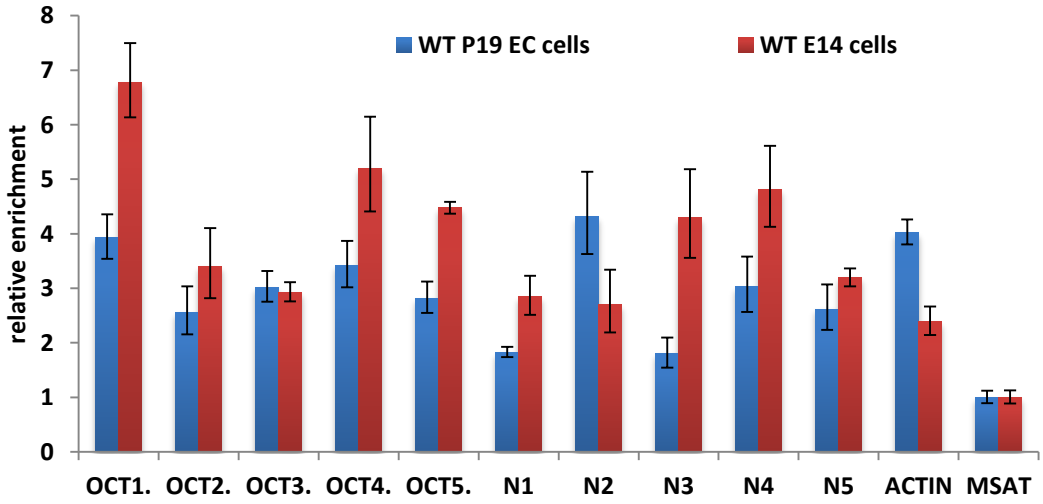


Figure 6.6 Hmgn1 binding at the Oct4 and Nanog gene loci in P19 and E14 cells.

CHIP-QPCR analysis of Hmgn1 enrichment at the Oct4 and Nanog loci. Data was normalised to input and the MSat primer set. Error bars represent the errors from PCR triplicates.

Hmgn1 binding at the Oct4 and Nanog loci also showed up to 7 fold enrichment compared to the MSat control (Figure 6.6). However, there did not appear to be increased enrichment near the TSS of either gene in either P19 or E14 cells.

6.3.4 Active and repressive Histone modifications in P19 and E14 cells

H3K4me3 and H3K27me3 are key histone modifications that often mark active promoters and repressed genes, respectively (Bernstein et. al., 2006). In order to investigate whether the global levels of these modifications are different in P19 and E14 cells, western blotting on chromatin samples was performed (Figure 6.7). TBP was used as a loading control. The levels of both H3K4me3 and H3K27me3 were similar in P19 cells compared to E14 cells (Figure 6.7).

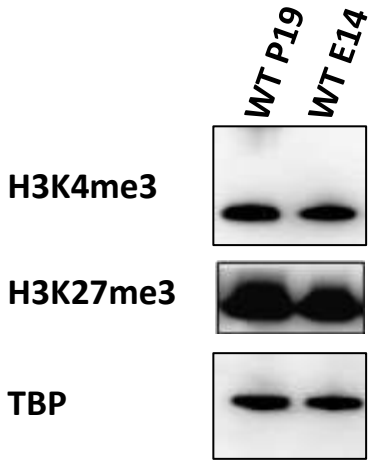


Figure 6.7 Western blotting for H3K4me3 and H3K27me3 chromatin from EC P19 and E14 ES cells
Chromatin was treated with Benzonase enzyme to remove the DNA, then crosslinks were reversed prior to electrophoresis. TBP is the loading control.

In E14 cells. H3K4me3 was enriched by 30 -35 fold at Oct4 primer sets 2, 3 and 4 (Figure 6.8). These primer sets correspond to the regions of highest H3K4me3 enrichment observed in the publically available ChIP-seq data from the ENCODE project (Figure 6.2). H3K4me3 was also highly enriched (20-40 fold) at Nanog primers sets 3, 4 and 5 (Figure 6.8), again corresponding to regions with high H3K4me3 in Figure 6.2.

In P19 cells, relative levels of H3K4me3 at the Oct4 and Nanog genes were only up to six fold higher than those observed at MSat (Figure 6.9).

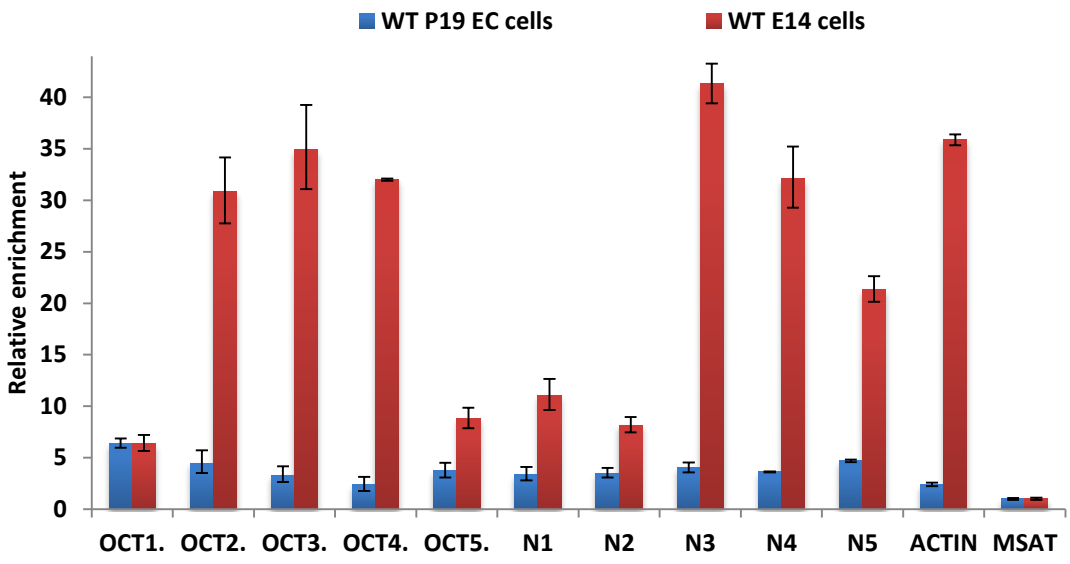


Figure 6.8 H3K4me3 enrichment at the Oct4 and Nanog gene loci in P19 and E14 cells.

CHIP-QPCR analysis of H3K4me3 enrichment at the Oct4 and Nanog loci. Data was normalised to input and the MSat primer set. Error bars represent the errors from PCR triplicates.

ChIP for H3K27me3 is shown in Figure 6.9. As expected, H3K27me3 levels at the Oct4 and Nanog genes were low, as these genes are not repressed in P19 or E14 cells.

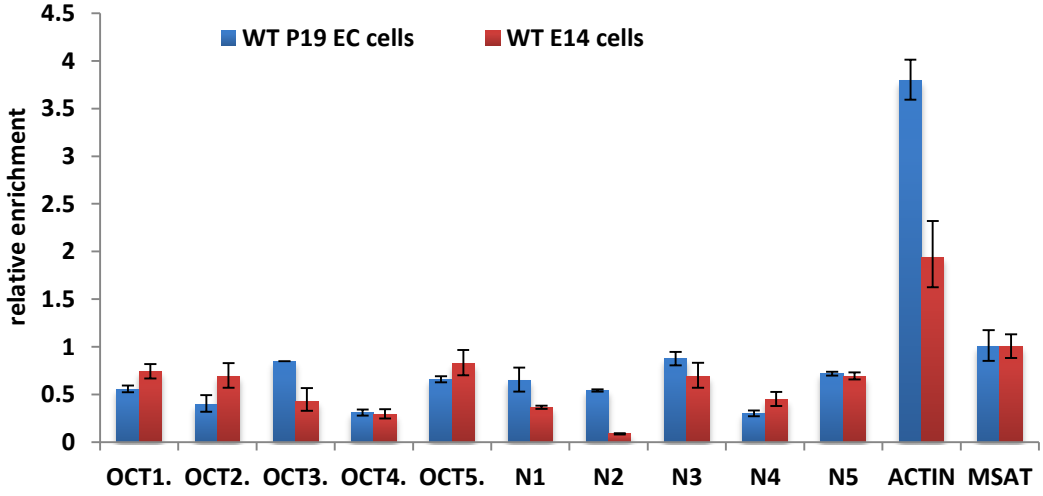


Figure 6.9 H3K27me3 enrichment at the Oct4 and Nanog gene loci in P19 and E14 cells

CHIP-QPCR analysis of H3K27me3 enrichment at the Oct4 and Nanog loci. Data was normalised to input and the MSat primer set. Error bars represent the errors from PCR triplicates.

6.3.5 Genome-wide analysis of Hmgn2 binding

The ChIP data presented above suggests that Hmgn1 and Hmgn2 do not bind specifically at promoters, but are instead bound at fairly similar levels at all the genomic loci tested. Similar conclusions have been reached from other studies in the lab (Barkess et. al., 2012). However, it is possible that peaks of Hmgn binding have been missed due to the locations of the PCR primer sets. To gain a clearer understanding of where Hmgn2 is bound across the genome, ChIP-seq was used to profile the binding of HMGN2 across the mouse P19 genome (performed by Dr. K. West).

Figure 6.10 shows the ChIP-seq data of Hmgn2 binding at the Nanog gene locus in P19 cells. In addition, ChIP-seq data from ENCODE project representing the histone modification (H3K4me3 and H3k4me1) from mouse embryonic stem cells. are shown. The Hmgn2 ChIP-seq track shows well-defined peaks of enrichment. These peaks are not present in the Histone H3 control track, showing that these genomic sequence elements might be specifically enriched by immunoprecipitation of HMGN2 protein. Moreover, histones H3K4me1 ChIP-seq from ENCODE project showed a peak of enrichment that correlates with one of the Hmgn2 peaks, upstream of the Nanog gene. These peaks seem to be located at the enhancer region of Nanog genes marked by red arrows, whereas histone H3K4me3 is enriched at the 5’ end of transcription starting sites (TSS).

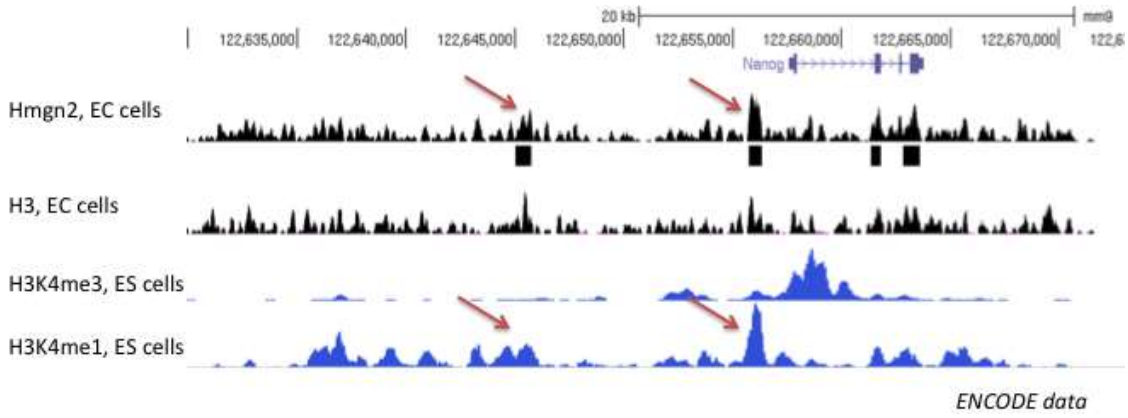


Figure 6.10 ChIP-seq of HMGN2 and H3 in undifferentiated EC cells
ChIP-seq localization of Hmgn2 and H3 in P19 EC cells, and ENCODE project views of H3K4me3 and H3K4me1 ChIP-seq profiles from ES cells. ChIP-seq peaks and summits identified by Sicer analysis with H3 as control are shown in black below the ChIP-seq track. UCSC annotated genes are shown in blue and the direction of the transcription is indicated by arrows (above track). The enriched location is identified by black track for Hmgn2 comparing with H3 as control and histone modifications (H3K4me3 and H3K4me1) in ES cells. The enriched location between HMGN2 and histone H3K4me1 are shown in red arrow.

6.3.6 Preparation of cross-linked chromatin from Hmgn2 knockdown P19 cell lines.

As described in Chapter 5, Hmgn2 knockdown cells were generated using shRNAmir expressed using the pGIPZ constitutive lentiviral system. These cell lines were analyzed using FACS, RT-PCR, and western blotting (Chapter 5). Chromatin was prepared as described in Chapter 2 (Barkess et. al., 2012), and sonicated to reach to an average fragment length of 500-100bp ready for immunoprecipitation (Figure 6.11).

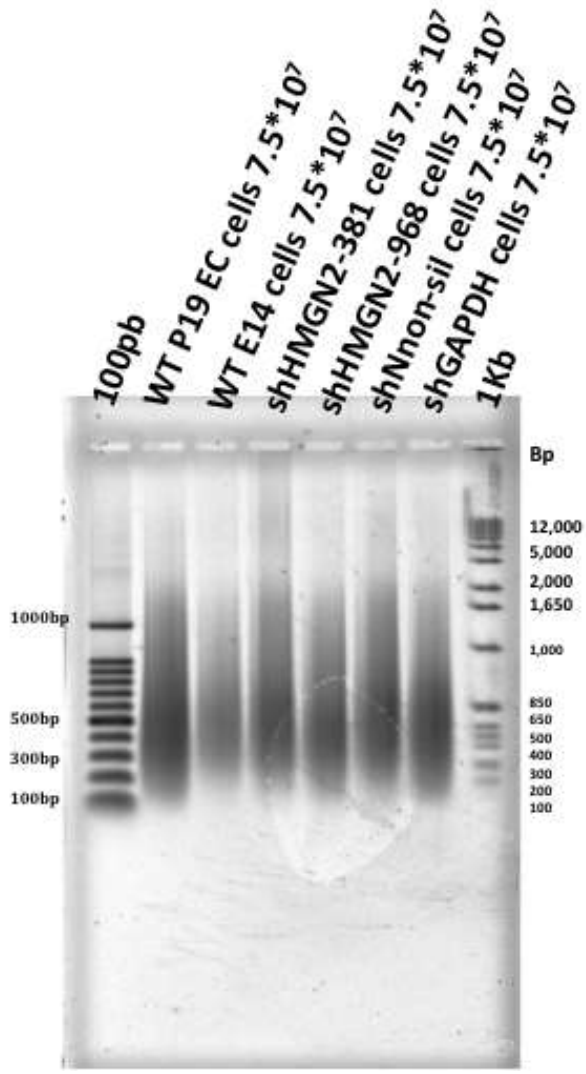


Figure 6.11 Shearing of wild type P19 and E14 and HMGN2 KNOCKDOWN P19 chromatin.
Agarose gel electrophoresis of P19 EC and E14 chromatin after sonication to an average fragment size of 200 – 1000 bp for use in ChIP.

6.3.7 Hmgn2 is undetectable in chromatin from pGIPZ-shHMGN2-381 – transduced P19 cells.

In order to investigate the level of chromatin-bound Hmgn1 and Hmgn2 in pGIPZ-transduced P19 cells, western blotting was used (Figure 6.12). As described for Figure 6.3 chromatin was treated with Benzonase to digest the nucleic acid, and crosslinks were reversed prior to electrophoresis. H3 was used as the loading control. Strikingly, Hmgn2 protein was undetectable in chromatin from shHMGNG-381 cells, whereas it was present in cells carrying

the negative control, shNon-silencing, vector. No reduction in Hmgn2 protein was observed in cells carrying the shHMGN2-968 vector, even though this is intended to target Hmgn2. Conversely, the amount of chromatin-bound Hmgn1 was increased in shHMGN2-381 cells, compared to wt P19 cells (Figure 6.12).

These results are similar to those observed using whole cell extracts (WCE) (Figure 5.5). In shHMGN2-381 cells, both WCE and chromatin show a large reduction in Hmgn2 levels, and an increase in Hmgn1 levels. In shHMGN2-968 cells, no reduction in Hmgn2 expression is observed in either WCE or chromatin, indicating that this shRNA trigger is unsuccessful in knocking down Hmgn2. There are some inconsistencies between the WCE and chromatin data, however. For example, the WCE blots show reduced Hmgn2 levels in the wt P19 sample, and Hmgn1 is absent from the shHMGN2-968 sample (Figure 5.5).

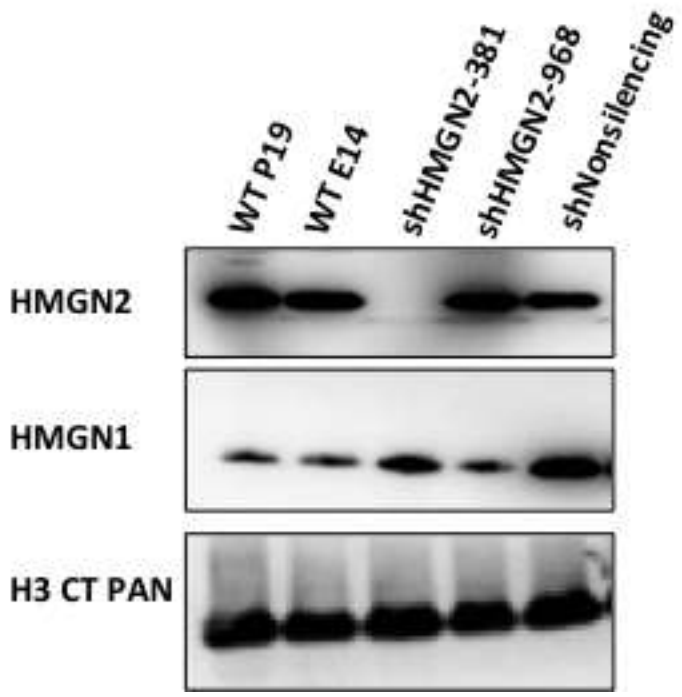


Figure 6.12 Analysis of Hmgn2 knockdown in pGIPZ-transduced P19 cells
The chromatin was treated with benzonase enzyme to remove the DNA, and then crosslinks were reversed prior to electrophoresis. H3 is the loading control.

6.3.8 Analysis of chromatin-bound factors in pGIPZ-transduced cells

Western blotting was unable to detect Oct4 in chromatin from shHMGN2-381 cells (Figure 6.13), whereas there was no reduction in shHMGN2-968 cells. Oct4 was significantly reduced, but not absent, in shNon-silencing cells.
TBP was assayed initially as a loading control. However, in shHMGN2-381 and shNon-silencing cells, the normal TBP is absent or reduced, and a higher mobility band was detected by the antibody (Figure 6.13).

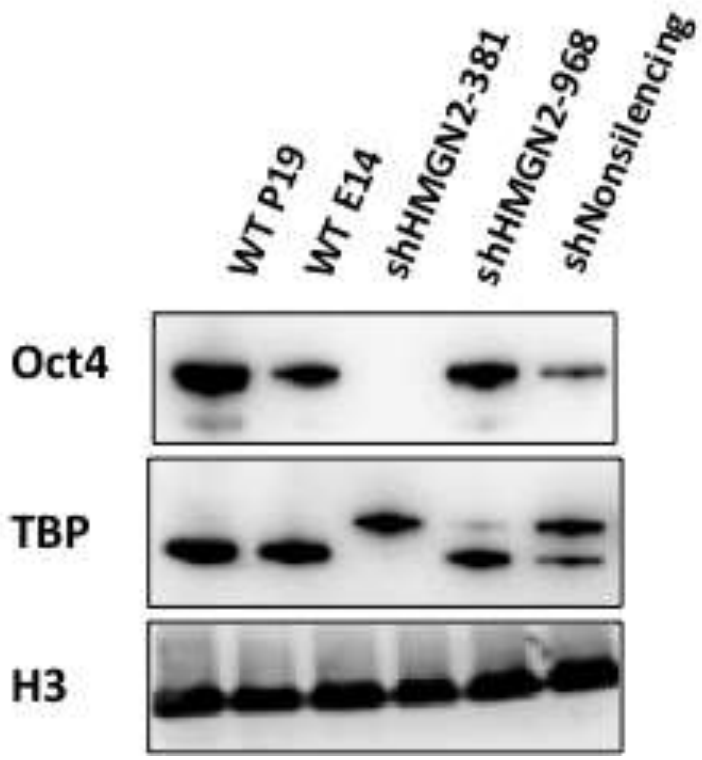


Figure 6.13 Oct4, Nanog and TBP immunoblot from chromatin (input) after 25 days of HMGN2 knockdown in EC P19 cells.

Other pluripotent and induce pluripotent factors were conducted using western blot for Input. Sox2 and C-myc was lost in Input of shHMGN2-381 comparing with non-silencing while Klf4 showed a gain of expression in shHMGN2-381 comparing with non-silencing (Figure 6.16). These results contradict with western blot for whole protein extract which showed UP-

regulation in Sox2 and slight down regulation in C-myc and Klf4 for the same shHMGN2-381 cell line (Chapter 5).

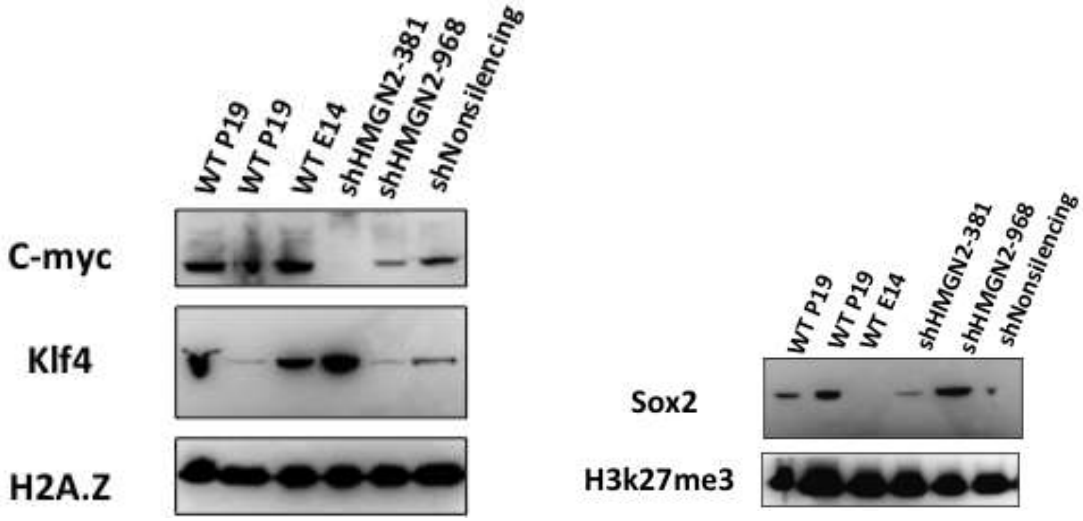


Figure 6.14 C-myc, Klf4 and Sox2 Immunoblot from chromatin (input) after 25 days of HMGN2 knockdown in EC P19 cells

6.3.9 Chromatin immunoprecipitation of Hmgn proteins in Hmgn2 knockdown cells

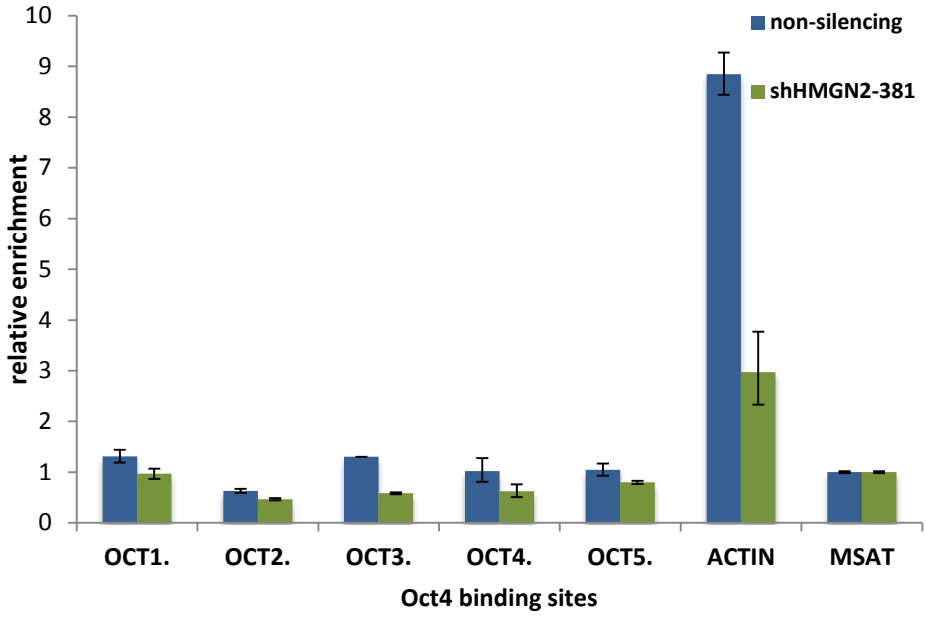
As shown above, western blotting revealed that Hmgn2 and Oct4 were undetectable in chromatin from shHMGN2 cells. To investigate whether the binding of Hmgn2 to the Oct4 gene locus was reduced, ChIP assays were performed using chromatin from shHMGN2-381 cells, and the control, shNon-silencing cells.

Control ChIPs using non-immune IgG and histone H3 were performed in order to compare the chromatin from the two cell lines (Figures 6.15 and 6.16). IgG is a control for background levels of non-specific immunoprecipitation, and reveals primer sets that might have abnormal results. H3 controls for differences in the ability of the chromatin preparations to be immunoprecipitated. For example, significant differences in crosslinking efficiency, size differences due to varying sonication efficiency, or loss of

protein epitopes due to over-sonication could alter the profile of chromatin-bound proteins or modifications. All enrichments are calculated relative to that for the Msat primer set, which normalizes for differences in the total amount of chromatin that is immunoprecipitated.

Similar levels of IgG immunoprecipitation were observed in chromatin from shNon-silencing and shHMGN2-381 cells (Figure 6.15). Enrichment was higher at the actin promoter compared with the Oct4 and Nanog primers, particularly with the shNon-silencing chromatin (9 fold higher than Msat), indicating that care must be taken when interpreting data from this primer set.

Levels of H3 enrichment were similar in both cell types, although slightly higher in chromatin from shNon-silencing cells (Figure 6.16). Enrichment at the actin promoter was up to 2.5 fold higher than Msat in shNon-silencing cells, which is less than the 9 fold observed with the IgG control.



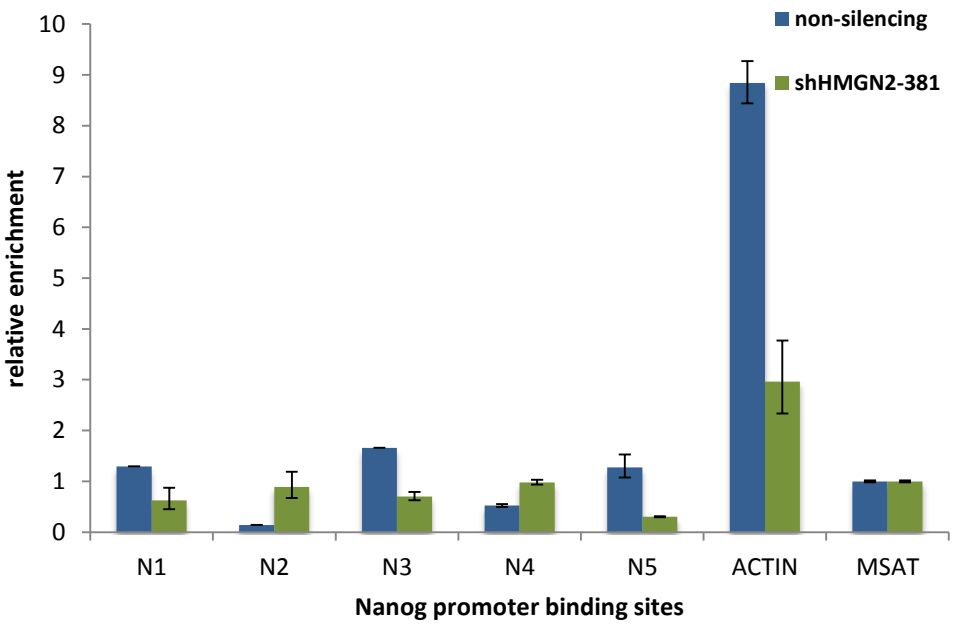


Figure 6.15 ChIP using non-immune IgG in Hmgn2 knockdown cells.

Chromatin was prepared from P19 cells transduced with pGIPZ-ShNon-silencing (negative control) or pGIPZ-shHMGN2-381 (Hmgn2 knockdown). ChIP was performed using rabbit IgG, and qPCR carried out with primers to the Oct4 and Nanog gene loci. Data was normalised to input and the Msat primer set. Error bars represent the errors from PCR triplicates.

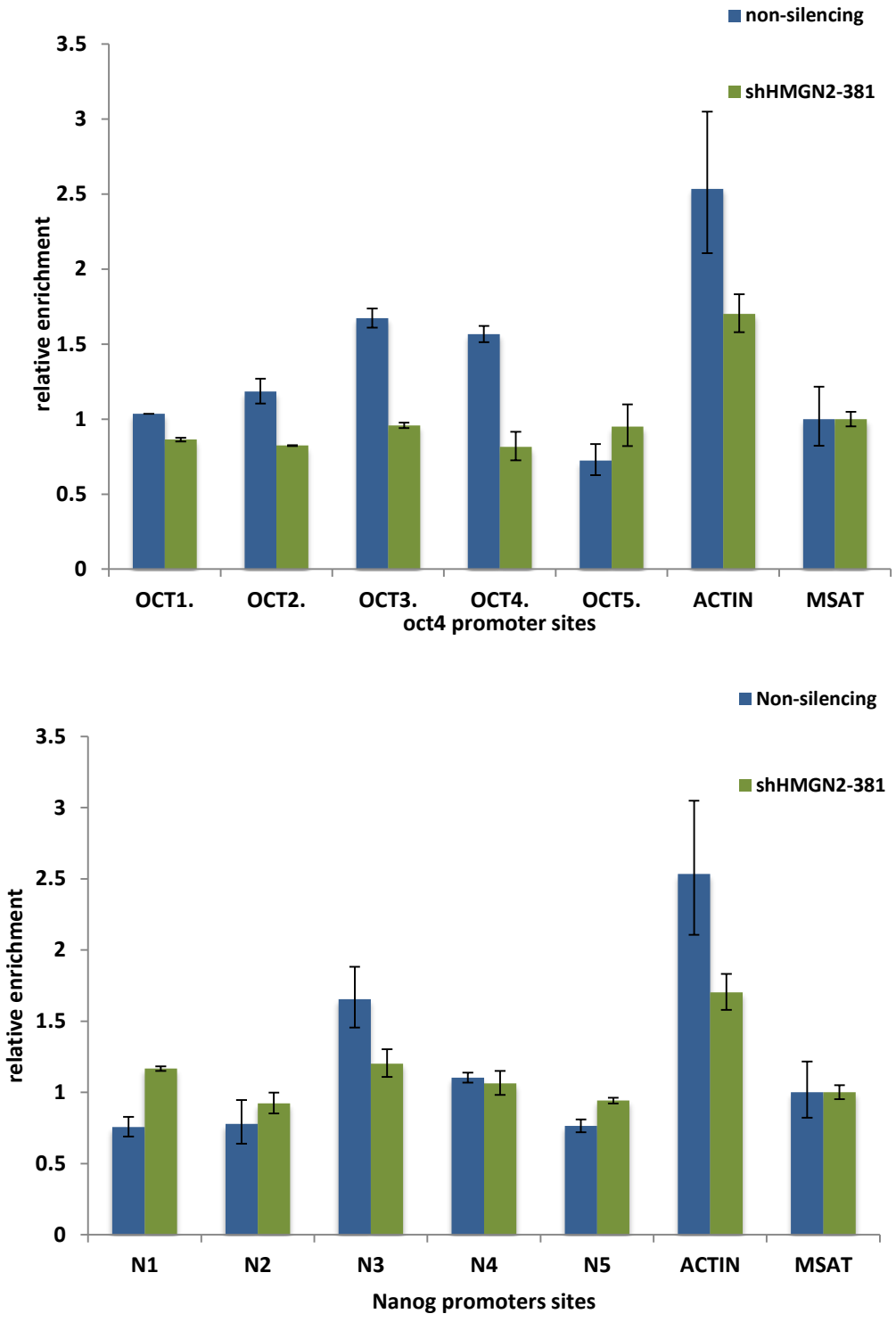
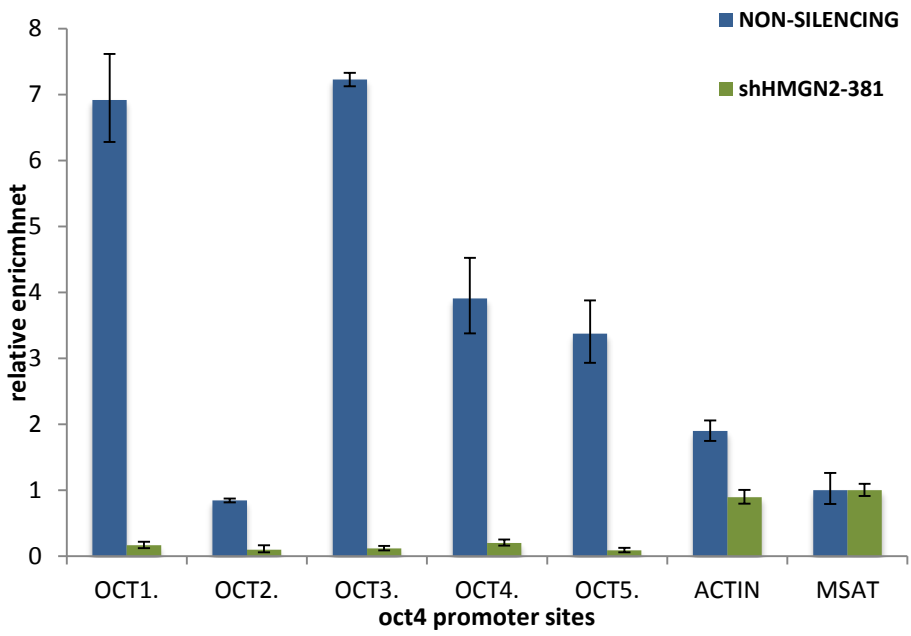


Figure 6.16 Enrichment of H3 around the Oct4 and Nanog TSS in knockdown P19 EC cells

Chromatin was prepared from P19 cells transduced with pGIPZ-ShNon-silencing (negative control) or pGIPZ- shHMGN2-381 (Hmgn2 knockdown). ChIP was performed using an antibody to the C-terminal region of H3, and qPCR carried out with primers to the Oct4 and Nanog gene loci. Data was normalised to input and the Msat primer set. Error bars represent the errors from PCR triplicates.

ChIP for Hmgn2 showed that most of the Hmgn2 binding at the Oct4, Nanog and Actin loci is absent in shHMGN-381 cells, in contrast with up to 7 fold enrichment at these sites in shNon-silencing cells (Figure 6.17). This is consistent with the global loss of Hmgn2 protein observed in the western blots of the chromatin (Figure 6.12).

Notably, binding of Hmgn1 is increased at the Oct4, Nanog and Actin loci in shHMGN2-381 cells compared to shNon-silencing cells (Figure 6.18) This is consistent with the increase in chromatin-bound Hmgn1 observed by western blotting in Figure 6.12.



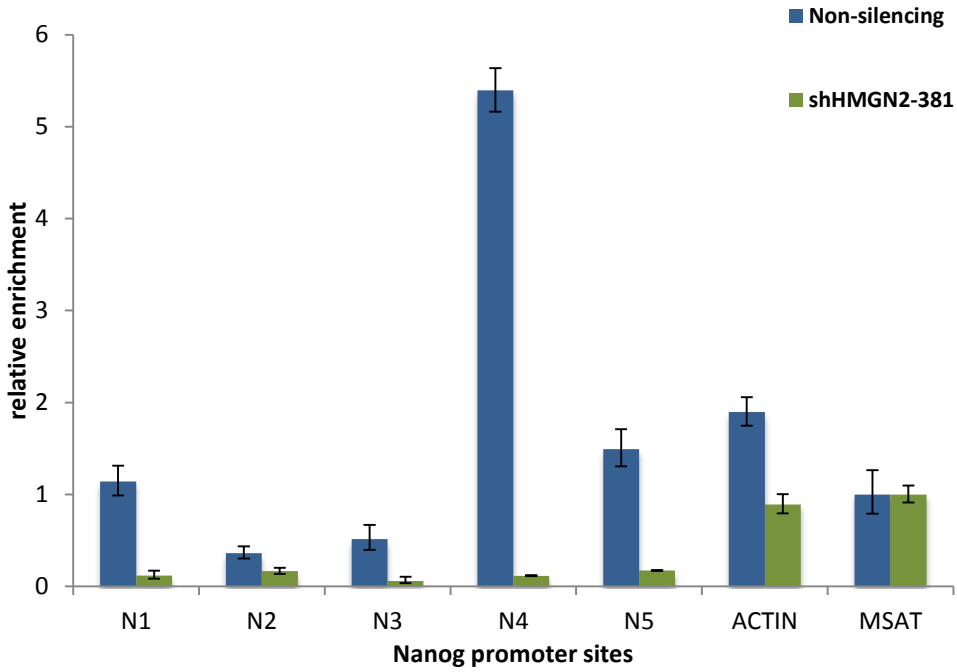


Figure 6.17 Global loss of Hmgn2 around Oct4 and Nanog TSS regions in knockdown P19 EC cell lines.

Chromatin was prepared from P19 cells transduced with pGIPZ-ShNon-silencing (negative control) or pGIPZ- shHMGN2-381 (Hmgn2 knockdown). ChIP was performed using an antibody to the C-terminal region of Hmgn2, and qPCR carried out with primers to the Oct4 and Nanog gene loci. Data was normalised to input and the Msat primer set. Error bars represent the errors from PCR triplicates.

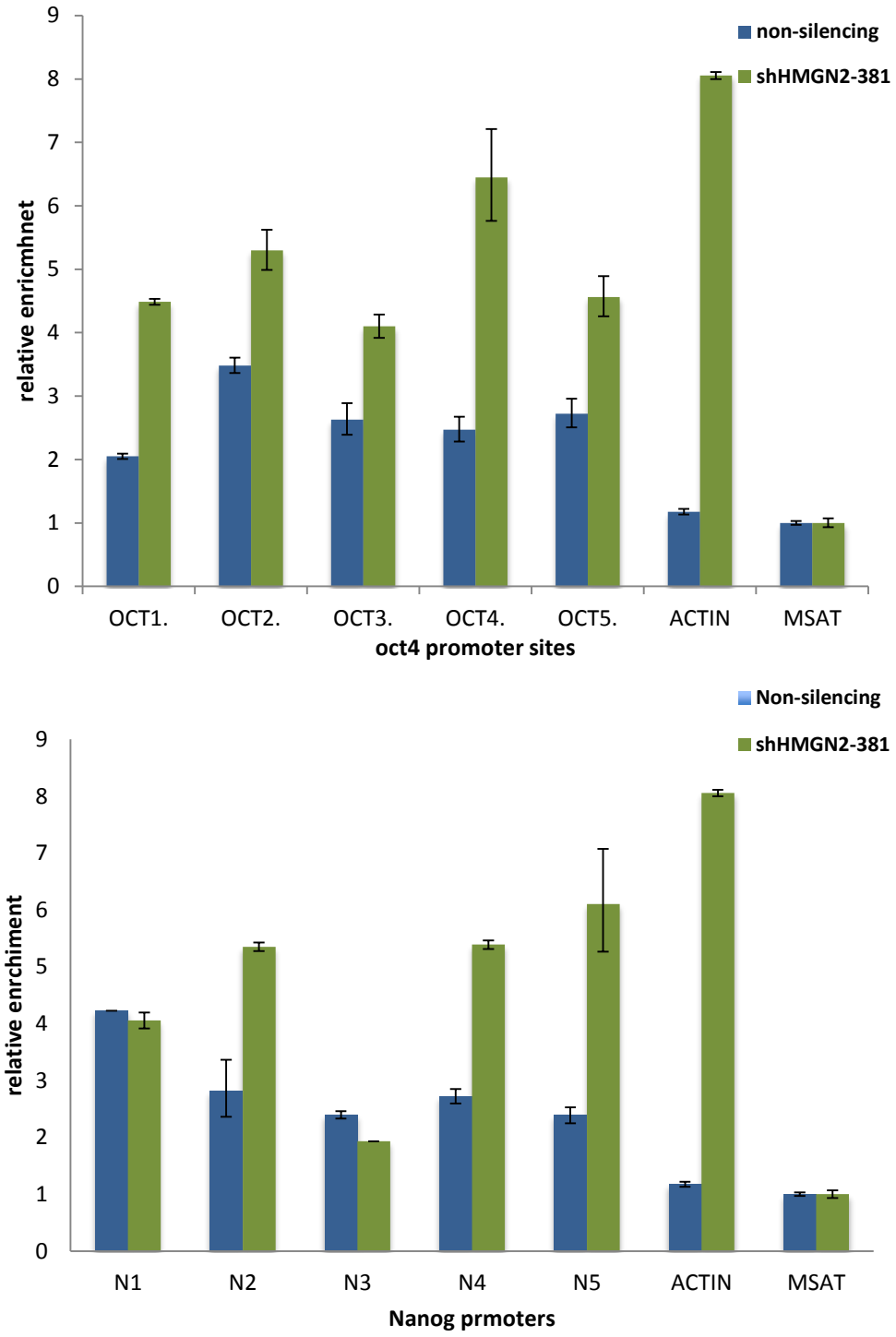


Figure 6.18 The enrichment of Hmgn1 around Oct4 and Nanog TSS in knockdown EC P19 cells.

Chromatin was prepared from P19 cells transduced with pGIPZ-ShNon-silencing (negative control) or pGIPZ- shHMGN2-381 (Hmgn2 knockdown). ChIP was performed using an antibody to the C-terminal region of Hmgn1, and qPCR carried out with primers to the Oct4 and Nanog gene loci. Data was normalised to input and the Msat primer set. Error bars represent the errors from PCR triplicates.

6.3.10 Changes in histone modifications at the Oct4 and Nanog genes in Hmgn2 knockdown cells

Western blotting of chromatin was performed to investigate if there were any global changes in the levels of H3K4me3 or H3K27me3 in Hmng2 knockdown cells (Figure 6.19). The results showed no significant changes in the levels of these modifications. The histone variant H2A.Z was also assayed, but this did not change in shHMGN”-381 cells either (Figure 6.19).

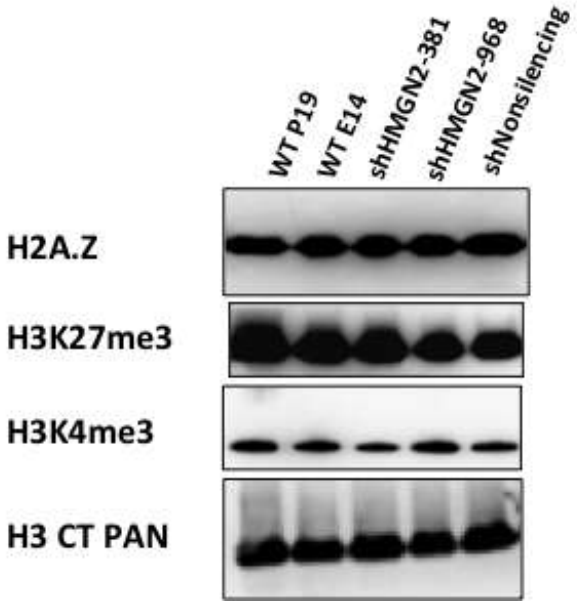


Figure 6.19 analysis of HMGN2 knockdown cells EC P19 cells.
Western blot from chromatin INPUT on knockdown cells comparing with WT EC P19 and ES E14 cells shows the histones variants. The knockdown binding sites of HMGN2 (shHMGN2-381 and shHMGN2-968), a negative control region (shNon-silencing), and positive control region (shGAPDH).

Chromatin immunoprecipitation was performed to investigate whether level of histone modifications at the Oct4 and Nanog loci were altered in Hmgn2 knockdown cells (Figure 6.20 and 6.21). The results show that H3K4me4 was significantly reduced at most Oct4, Nanog and Actin primer sets in Hmgn2 knockdown P19 cells (shHMGN2-381) compared to control (shNon-silencing) cells (Figure 6.20).Levels of H3K27me3 across these loci were also reduced in Hmgn2 knockdown cells (Figure 6.21)

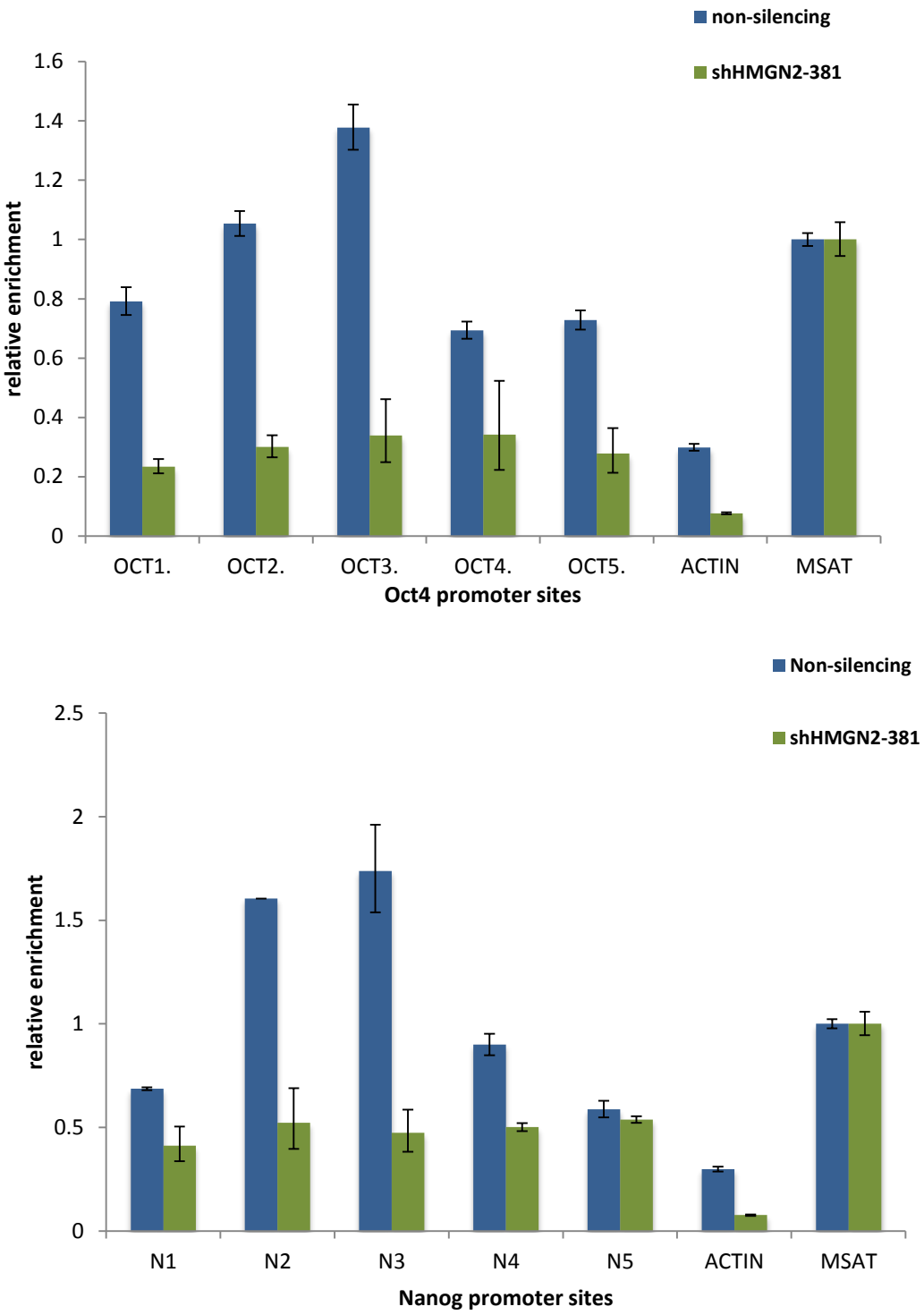


Figure 6.20 The enrichment of H3K4me3 around Oct4 and Nanog TSS in knockdown EC P19 cells

Chromatin was prepared from P19 cells transduced with pGIPZ-ShNon-silencing (negative control) or pGIPZ- shHMGN2-381 (Hmgn2 knockdown). ChIP was performed using an antibody to H3K4me3, and qPCR carried out with primers to the Oct4 and Nanog gene loci. Data was normalised to input and the Msat primer set. Error bars represent the errors from PCR triplicates.

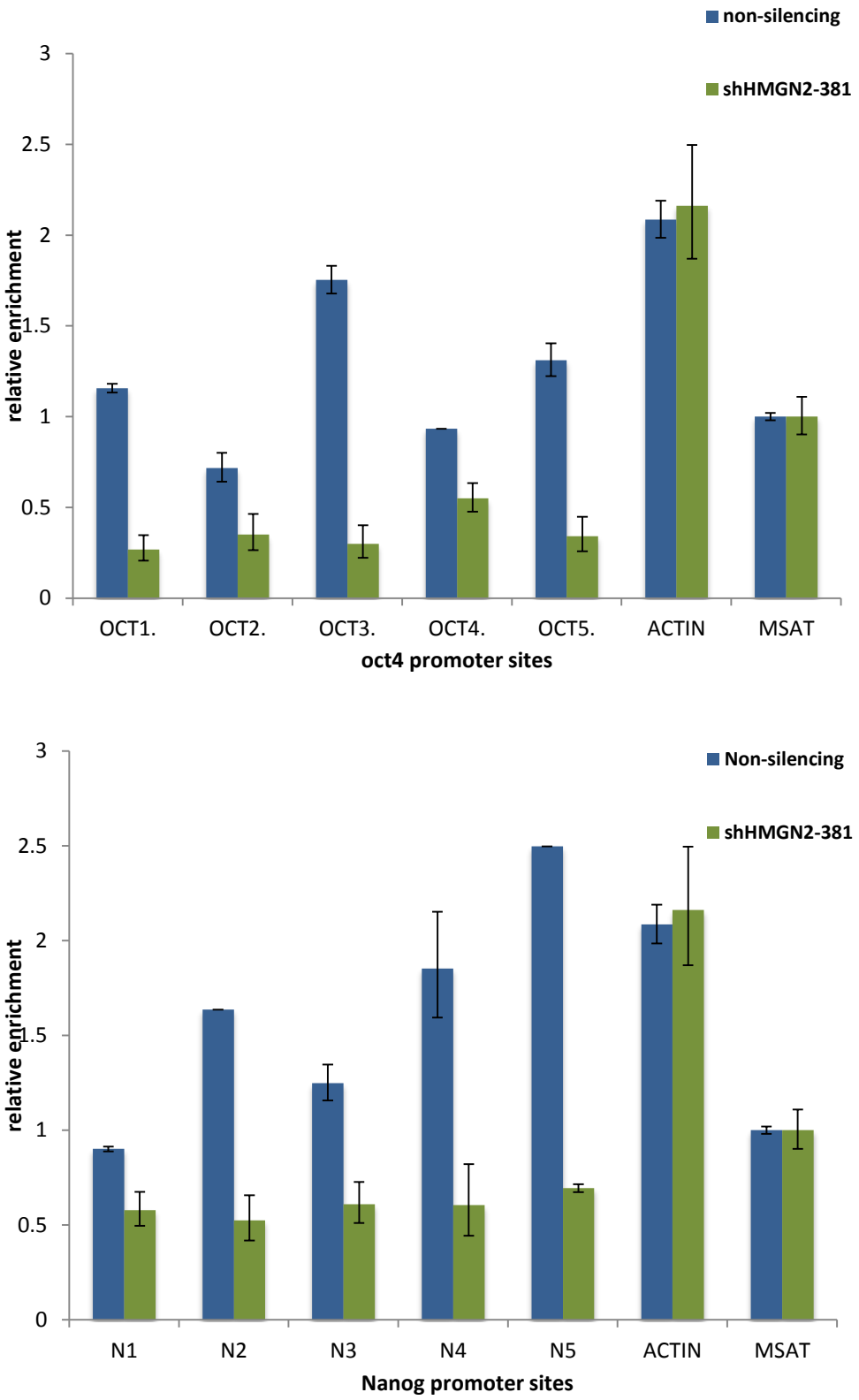


Figure 6.21 The enrichment of H3K27me3 around Oct4 and Nanog TSS in knockdown EC P19 cells

Chromatin was prepared from P19 cells transduced with pGIPZ-ShNon-silencing (negative control) or pGIPZ- shHMGN2-381 (Hmgn2 knockdown). ChIP was performed using an antibody to H3K27me3, and qPCR carried out with primers to the Oct4 and Nanog gene loci. Data was normalised to input and the Msat primer set. Error bars represent the errors from PCR triplicates.

6.3.11 The investigation of TBP migration using CO-IP against ubiquitination and/or acetylation

As previously shown in Figure 6.13, anti-TBP recognised a protein approximately 10-15 kDa larger than typical TBP in chromatin from some of the GIPZ-transduced cell lines. In order to investigate these further, westerns for TBP in whole cell extracts (WCE) were also carried out (Figure 6.22). The higher Mw band was also observed in WCE for some pGIPZ_transduced cell lines, although this phenomenon was not specific to the Hmgn2 knockdown cells (ShHMGN2-381) (Figure 6.22)

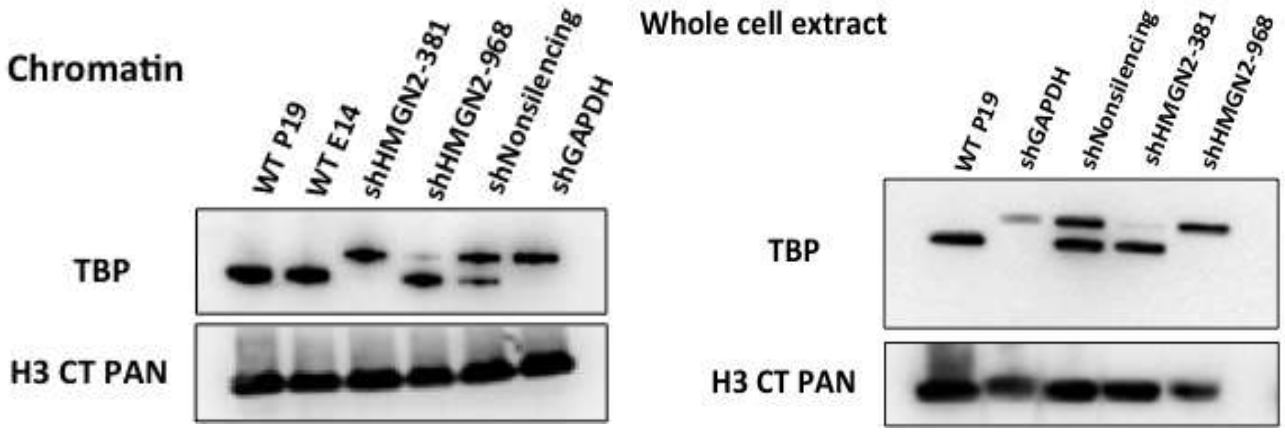
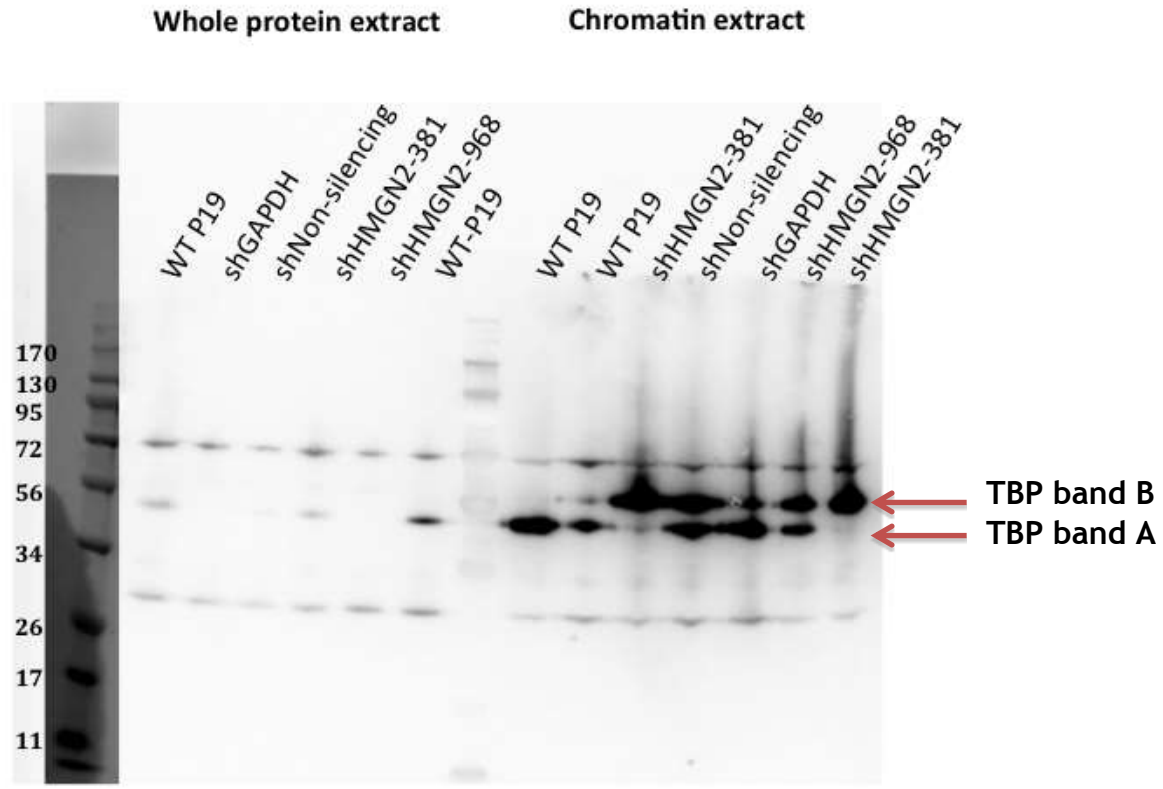


Figure 6.22 TBP migration immunoblot after 25 days HMGN2 knockdown whole-cell extract and chromatin

We reasoned that the higher Mw band could be an alternative splice form of TBP, a phosphorylated form, or an ubiquitinated form. Multiple phosphorylation and ubiquination sites on TBP have been demonstrated (follow from www.phosphosite.org). In order to test whether ubiquitination was involved, a co-immunoprecipitation assay was performed.

Co-immunoprecipitation (CO-IP) was carried out using an antibody for ubiquitin (Figure 6.23A) or non-immune IgG (Figure 6.23B). Chromatin samples were treated with benzonase and crosslinks reversed prior to the Co-IP assay. After CO-IP, samples were western blotted, and membranes probed with the antibody to TBP.

A)



B)

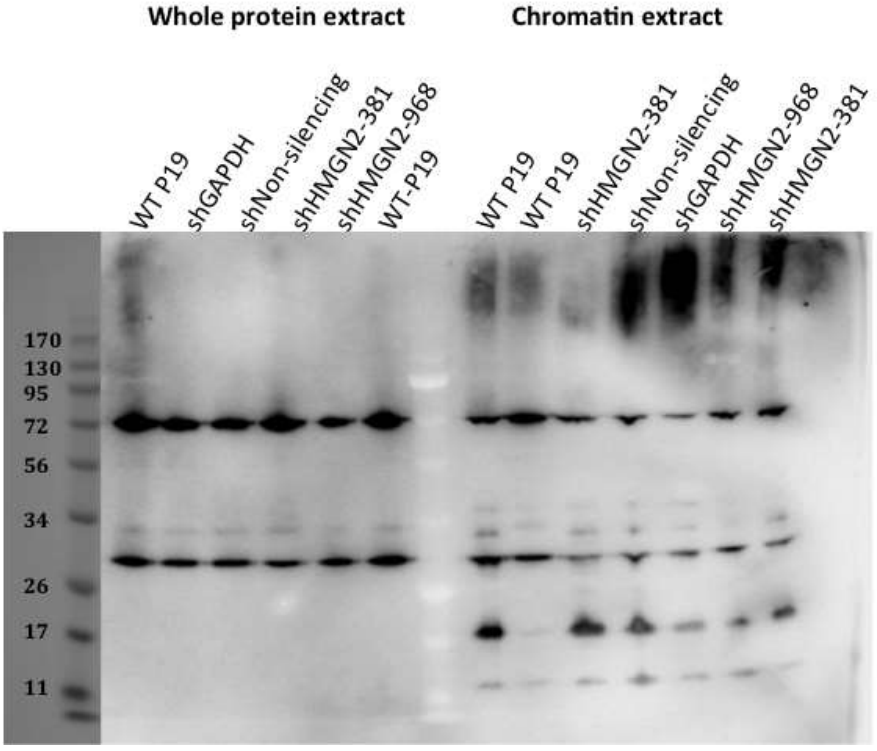


Figure 6.23 TPB may interact with di- and mono- ubiquitination.
Immunoprecipitation with ubiqutinylation antibody FK2 (A) or non-immune IgG (B). Either whole cell extracts (left) or chromatin samples (right) were used. Chromatin samples were treated with benzonase and crosslinks reversed prior to the Co-IP assay. Western blots of Co-IP samples were probed with anti-TBP (expected size 37 kDa).

Co-IP using IgG did not pull down any bands corresponding to the expected 37kDa band, or higher mobility forms (Figure 6.23B). However, co-IP using the ubiquitin antibody in chromatin extracts pulled down two bands A and B, from pGIPZ-transduced cells. Co-IP from WCE pulled down the lower band A in wt P19 and shHMGN2-381 samples only. Close inspection of these blots indicated that the lower band A has an apparent Mw of ~50 kDa, and the upper band B has an apparent mobility of ~60 kDa, by comparison with the 56 Kda marker. The calculated Mw of TBP is 37 kDa, although the apparent Mw in cell extracts varies between 40 and 45 kDa (Abcam manufacture), and a single ubiquitin group has a Mw of 8.5 kDa (Li et. al., 2015). It is possible that the lower band A corresponds to TBP with a single ubiquitin mark conjugated to it, and the upper band B has two ubiquitin marks. However, further experiments are required in order to investigate this further (see appendix Figure x3 and x4 the difference between TBP western blot in PGIPZ and PTRIPZ).

6.4 Discussion

The aim of this chapter was to use ChIP technology to identify Hmgn2-DNA interactions in the P19 EC and E14 cell line and compare with the loss of Hmgn2 in P19 EC cell lines. Moreover, the use of western blot for input and ChIP-qPCR was the first stage to validate the loss of Hmgn2 on pluripotent genes such as Oct4 and Nanog.

6.4.1 The effect of Hmgn2 knockdown in pluripotent stem cells

Many studies from different tissues showed the remodelling the chromatin structure level could affect the core and linked histones, the regulatory factor of DNA assembly and epigenetic modification. These modifications regulate the transition period from an inactive to active genome during embryonic development. One of these modifications was conducted in Hmgn proteins knockout mouse, which caused a delay early embryonic development (Bastos et. al., 2008). From previous study, Knockdown of Hmgn2 and/or Hmgn1 showed reduction in Oct4, Sox2, and Nanog genes in P19 EC cells using siRNA (Mohan, 2012). This may indicate that Hmgn1 and Hmgn2 binding site enhance the

regulatory region of these pluripotent genes. Therefore, Hmgn1 and Hmgn2 may act as mediator to maintain the pluripotent markers by interacting with other transcription factors. Hmgn2 knockdown with shRNA showed downregulation of Oct4 and Nanog at the protein and RNA level.

6.4.2 Hmgn2 and H3K4me3 are enriched in Oct4 and Nanog region of P19 EC cells and E14 stem cells

ENCODE data of H3K4me3 ChIP-seq from E14 cells showed enrichment in Oct4 and Nanog at the transcription starting sites (TSS) of ESC cells. Hence the enrichment of H3K4me3 at the TSS of Oct4 and Nanog was investigated in mouse stem cells, Hmgn2 may affect at the regulatory region of Oct4 and Nanog. Therefore, ChIP followed by QPCR was performed to confirm any changes in regulatory region of Oct4 and Nanog in only P19 EC cells. Normal ChIP was applied in wild type P19 EC cells and E14 cells. The results show more of Hmgn2 than Hmgn1 by 2-3 relative enrichment in E14 and P19 EC cells. Moreover, H3K4me3 is highly enriched in E14 cells at Oct4 and Nanog regions whereas H3K4me3 was found P19 EC cells less enrichment. However, in H3K27me3 there is more or similar level of enrichment in E14 and P19 EC cells. This may indicates that due to P19 EC is embryo carcinoma stem cells, more genes may be silenced. Since transgene silencing in P19 EC cells showed that P19 EC cells play roles in normal preimplantation embryos and other type of precursor cells. Therefore, it was concluded that both P19 EC and ES cells might have similar gene silencing mechanism (He et. al., 2005).

6.4.3 Western blots from Hmgn2 knockdown input downregulated Oct4, Sox2, and C-myc and up-regulated Klf4

Previously, Hmgn2 knockdown in P19 EC cells was conducted to confirm the down regulation of pluripotent genes such as Oct4, Sox2, and Nanog and induce pluripotent genes such as Klf4 and C-myc by RNAi technology (Chapter 5). ShHMGN2-381 and non-silencing (negative control) knockdown cells were selected to perform ChIP assay. After collecting chromatin samples, western blot for input was applied to examine the pluripotent genes expressions profile in these cells without adding antibodies. The results show similarity and variability between the chromatin and whole protein western blot in the same genes. Western blot results show similarity in Hmgn1 and Hmgn2. Hmgn2 knockdown

and over expression of Hmgn1 in shHMGN2-381 was confirmed in both protein and chromatin western blot. Nanog and Oct4 did not show protein or chromatin expression.

6.4.4 QPCR for ChIP revealed a loss of Hmgn2 and H3K4me3 enriched binding and slightly enrichment of H3K27me3 in Oct4 and Nanog binding regions

Beside western blot for input, QPCR-ChIP results on Hmgn2 knockdown P19 EC cells demonstrate three points. First, Hmgn2 showed low enrichment in Oct4 and nanog region whereas Hmgn1 showed high enrichment in Oct4 at position, O1 (-1781) and O2 (-399) and Nanog at position N1 (-1081), and N2 (-219) in Hmgn2 knockdown cell. On the other hand, H3K4me3 and H3K27me3 at the same region (TSS) in Oct4 and nanog showed loss of enrichment in Hmgn2 knockdown P19 EC cells.

6.4.5 TBP western blot showed migrated bind leading to suspect that TBP may be experienced ubiquitination

The protein and input western blot results of Hmgn2 knockdown P19 EC cells demonstrate there is TBP migration occurs. CO-IP of TBP was performed against phosphorylation (data did not show) and ubiquitination. However, the results fail to explain the changes in TBP protein. In addition, it is important to note that ChIP-QPCR was from one biological replicated only. Moreover, there is a chance of background raised from normal method of ChIP which was shown by measuring the delta CT value to input sample. Furthermore, there were difficulties in sonicated chromatin of Hmgn2 knockdown.

6.5 Conclusion remark

The results of this chapter showed ChIP technology for P19 EC, E14 and Hmgn2 knockdown in only P19 EC cells. This finding may exhibit novel roles of Hmgn2 as safeguard for pluripotency and reprogramming mechanism due to its interaction with Oct4, Sox2, Nanog, Klf4 and C-myc. Hmgn2 knockdown ChIP-QPCR showed H3k4me3 beside Hmgn2 lost enrichment in Oct4 and Nanog region whereas Hmgn1 gained the enrichment at the same region. Further experiments need to carry out before deriving any further conclusion.

Chapter 7 Summary and Future work

7.1 Summary of work presented in this thesis

Gene transcription, genomic stability, cellular differentiation and development are affected by chromatin remodelling. A group of architectural proteins called High mobility group proteins (HMGN) are associated with nucleosome involving in chromatin remodelling architecture and modifying gene expression. One of the main critical functions of these proteins was found in the development of mouse, human and chicken. However, investigating the role of HMGN2 proteins in embryonic stem cells remains unclear. A pilot project in Dr. Katherine West lab previously found that Hmgn2 knockdown with siRNA technology caused downregulation of Sox2, Oct4 and Nanog in P19 EC cells (Mohan, 2012).

The principle aims of this thesis were to confirm the previous results of siRNA knockdown tool in P19 EC cells. Using lentivirus based on shRNAmir technology. Hmgn2 expression in undifferentiated P19 EC cells throughout endogenous RNA interference process was preformed to achieve a stable, long term, and high level of Hmgn2 knockdown to study the role of Hmgn2 in mouse cancer stem cells. Oct4 and Nanog binding sites across the mouse genome during mouse genome of P19 EC cells and ES cells were identified by western blot of chromatin, qPCR-ChIP and ENCODE protein of ES cells. Moreover, western blot of chromatin and qPCR-ChIP was used to investigate the Hmgn2 knockdown peaks throughout Oct4 and Nanog binding sites besides investigating the enrichment of Hmgn1 and histone (H3, H3K4me3, and H3K27me3) at Oct4 and Nanog binding sites.

7.2 The optimization of lentiviral pSLIK-CD4, pGIPZ, and pTRIPZ knockdown in P19 EC, Hepa and E14 cells

P19 EC cells are well-characterized cell line to study cancer stem cell and stem cells induced to neuron, muscles and cardiac differentiation. Chapter 3 describes the optimization condition of the undifferentiated P19 EC cells which were applied as main target knockdown Hmgn1 and Hmgn2 using three lentiviral based on shRNAmir technology. Prior to generating lentiviral particles, P19 EC cell was examined under puromycin and doxycycline drug concentration.

Doxycycline drug was used for inducible lentiviral system to induce shRNAmir expression. The optimal concentration for Doxycycline amount was 1 µg/ml using luminescent 96-well plate assay for cellular ATP and gene expression of Hmgn2 and Oct4. Puromycin drug was tested using puromycin kill curve and Flow cytometer analysis (FACS) on day 3 and 5 to identify the amount of puromycin concentration and check the health of the cells.

Four shRNAmir-30 (Hmgn2-1, Hmgn1-2, Hmgn1-3, and Hmgn-scramble) were designed with ccdB gene and received inside pENTER-TGmiRC3. Then, shRNA was moved to pSLIK-CD4 (lentiviral vector) throughout gateway cloning recombination system. Finally pSLIK-CD4-shHmgn was created for all four Hmgn triggers. Generating knockdown of mmHmgn1 or mmHmgn2 proteins failed beyond expectation due to the low level of transduction of Hmgn2 expression in knockdown cells. Alternatively, pTRIPZ is another inducible lentiviral knockdown system was developed by open biosystem. The difference between pSLIK and pTRIPZ is that shRNAmir-30 is under the control of TRE promoter in pSLIK-CD4 whereas shRNAmir-30 and red florescent protein reporters are under the control of TRE promoter in pTRIPZ. Moreover, The reverse tetracycline transactivator (rtTA3) and reporter gene are under the control of constitutively active promoter (UBC-i) in pSLIK system. However, the reverse tetracycline transactivator and drug selection are under the control of constitutive promoter (UBC) in pTRIPZ. PSLIK system contains inactive CMV promoter can be used to switch the plasmid to constitutive system rather than doxycycline inducible whereas pTRIPZ does not contain CMV promoter. Four shRNA-Hmgn from pGIPZ and shRNA-Hmgn2-1 from pSLIK-CD4 were cloned to pTRIPZ. After generating knockdown of Hmgn2 using pTRIPZ, the level of Hmgn2 knockdown in mRNA and FACS analysis was not high enough to affect the pluripotent gene (Oct4, Sox2 and Nanog) expressions.

Therefore, Chapter 4 represents the constitutive lentiviral system (pGIPZ) from open bio-system was applied to achieve high, efficient, and stable knockdown for Hmgn2 in P19 EC and Hepa cell lines. Hmgn2 knockdown using pGIPZ in P19 EC and Hepa cells was very efficient. A significant reduction of Hmgn2 was observed by mRNA analysis. However, it was noticed that knockdown Hmgn2 in Hepa cells were significantly efficient and easier than P19 EC cells.

7.3 Characterization and identification of Hmgn2 knockdown and Hmgn2 binding sites in undifferentiated P19 EC cells (Chapter 4 and 5)

Chapter 5 describe the establishment of Hmgn2 knockdown using lentiviral pGIPZ in undifferentiated P19 EC cells. Five shRNAs (381, 968, 514, non-silencing and GAPDH) generated knockdown of Hmgn2 proteins by almost 80%-90%. Hmgn2 knockdown cell lines showed significant upregulation of Hmgn1. This result indicate that Hmgn2 compensated with Hmgn1 when there is a loss of Hmgn2. Moreover, these knockdown cells were used to investigate in the expression of key pluripotent and induce pluripotent genes. Hmgn2 knockdown cells showed dramatic downregulation of pluripotent genes, Oct4 and Nanog whereas Sox2 showed a clear upregulation. These results indicate that Hmgn2 play a role in the regulation of pluripotent genes, Oct4 and Nanog. On the other hand, the upregulation of Sox2 in knockdown Hmgn2 cell lines. These results showed that Hmgn2 may require for the regulation of Oct4 and Nanog while Sox2 overexpression in knockdown Hmgn2 cells may be critical for maintaining the cancer stem cells and consequence preventing the cell differentiation. Moreover, knockdown the Hmgn2 in P19 EC cells showed slight reduction in induce pluripotent stem cell genes, Klf4 and C-myc. This result was challenging to demonstrate due to protease activity in protein.

On the other hand, Chapter 6 presents the changes of Hmgn2 knockdown at the chromatin level. Western blot for chromatin analysis used to ensure the quality of chromatin and check the protein effects in knockdown Hmgn2 in P19 EC cells. The results showed that pluripotent markers, Oct4 and Sox2, were significantly reduced in Hmgn2 knockdown P19 EC cells. However, induced pluripotent marker Klf4 showed high expression while C-myc showed significant reduction. These results do not match whole protein analysis from Chapter 5. Therefore, These results may indicate that some of these proteins have not been affected at chromatin level.

Moreover, chapter 6 describes and compares the chromatin from the knockdown of Hmgn2 in P19 EC cells using pGIPZ with mouse embryonic stem cell's public data from the ENCODE project. ChIP data from ENCODE project highlighted the Oct4 and Nanog transcription starting sites because they contain an enrichment peak of histone for active promoter (H3K4me3) in ES cells. To confirm this data experimentally, qPCR for ChIP analysis applied the enrichment peak of H3K4me3

in Oct4 and Nanog binding site in E14 and P19 EC cells. The results showed high enrichment of H3K4me3 besides Hmgn2 at the same region. Knockdown Hmgn2 in P19 EC cells used one target shHmgn2-381 and one negative control (non-silencing). ShHmgn2-381 found to have significant reduction of Hmgn2 from protein and mRNA analysis (Chapter 5) and from western blot for chromatin (Chapter 6). Therefore, qPCR-ChIP used this target to investigate the effect of Hmgn2 reduction in Oct4 and Nanog binding sites. QPCR-ChIP for knockdown Hmgn2 P19 EC cells confirmed a loss of Hmgn2 peak in Oct4 and Nanog binding sites while there is clear high enrichment peak of Hmgn1 in the same region. Histone for active promoter (H3K4me3) showed a loss in Oct4 and Nanog binding sites whereas histone for Polycomb repression (H3K27me3) showed slight enrichment peak at the same Oct4 and Nanog region.

Furthermore, chapter 6 attempts to investigate the western blot analysis from whole protein and chromatin of the general transcription factor binding specifically to DNA (TATA binding protein or TBP). The results showed the migration of TBP in knockdown Hmgn2 in P19 EC cells. an expected phosphorylated or ubiquitinated TBP preformed using CO-IP, but further investigation need to address.

7.4 Future work

Work described in this thesis aimed to understand biological roles played by a highly conserved non-histone chromosomal protein Hmgn2 and identify the novel roles in the mouse cancer stem cells. One of the main achievements of this work is that it allowed several new avenues for further research, some of which are briefly outlined below.

7.4.1 Identify the difference between embryonic stem cells and epiblast stem cells

P19 EC cells are considered as epiblast stem cells (EpiSCs). These epiblast stem cells showed two stage of embryo development: early epiblast and late epiblast. These epiblast stem cells are different from embryonic stem cells (ES cells) in epigenetic profiles and signalling responses, but they share some similarities with human embryonic stem cells (hESC) in gene expression and signalling pathways. A combine between chromatin immunoprecipitation and microarray

(ChIP-on-Chip) helped to understand the correlation between chromatin domain in epiblast stem cells and pluripotent makers such as Oct4, Sox2 and Nanog. The result showed Oct4 expression has two regulatory sites: distal enhancers were found to express stem cell from inner cell mass (ICM) whereas proximal enhancers were found to express epiblast stem cell from post-implantation. These enhancers have the ability to direct the Oct4 expression in those cells. This difference was observed in chimaera mice when stem cells introduced to pre-implantation embryo, epiblast cells did not survive as single cells and no committing to pre-implantation embryo was detected throughout morula aggregation. Moreover, epiblast stem cells can form teratomas that include different type of cells (Tesar et. al., 2007). P19 EC studies did not highlight if P19 EC cells are raised from early epiblast or late epiblast. However, it was mentioned that P19 EC cells similar to inner cell mass (ICM) cells. However, study identified the Oct4 expression is activated by proximal enhancer in P19 EC cells. Therefore, P19 EC cells are raised from post-implantation (late epiblast) (Yeom el. al., 1996).

Defining gene expressions that involve in ES and Epiblast stem cells is important for better cell differentiation. Both cells express high Oct4 and Sox2, but less Nanog. Genes are associated with ICM stage such as Pecam1, Tbx3 and Gbx2 express in mouse stem cells. Genes such as Otx2, Eomes, Foxa2, Brachyury(T), Gata6, Sox17, and Cer1 highly express in Epiblast stem cells, human ES cells, and early germ layers. However, Rex1 and Stella found to be down regulated in naïve epiblast stem cells. ChIP-on-Chip applied in three genes Stella, Nanog, Otx2 in mouse ES, Epiblast and human ES cells to detect the histone modification for active and repressive promoter (H3K4me3 and H3K27me3) at the transcription-starting site (TSS). The results indicate that epiblast and human stem cells share similar epigenetic regulation of transcription. Moreover, same experiment was performed in Oct3/4 in the same cells showed clear difference of Oct4 transcriptional network to maintain pluripotent epiblast stem cells (Tesar et. al., 2007).

Signaling pathways is very important to detect specific genes for cellular and molecular process. Epiblast stem cells are able to maintain pluripotent state through ActivinA (Inbb)/Nodal and SMAD2/3 signaling whereas ES pluripotency is controlled by LIF signaling. Therefore, preventing the phosphorylation of STAT3 at tyrosine residue 705 in JAK inhibitor showed no ES cell differentiation (Tesar

et. al., 2007). Moreover, growth factors FGF/MAPK and TGF-B signaling was detected in Epiblast post-implantation (Loh et. al., 2015). It would be helpful to identify and characterize epiblast stem cell genes in P19 EC cells using qRT-PCR for gene expression and signaling facto

7.4.2 Developing pGIPZ, pSLIK and/or pTRIPZ tools to study the Hmgn2 and/or Hmgn1

There were several issues were observed in all the Lentiviral systems. First, Chapter 4 showed that P19 EC knockdown cell lines with constitutive lentivirus system (pGIPZ) driven by CMV promoter for 20 days became methylated. Similar results were found in inducible lentiviral system (pTRIPZ and pSLIK-CD4) (Chapter 3) that is derived by Ubc-I promoter. Therefore, replace the promoters from pGIPZ and pTRIPZ with GAC or A2UCOE promoter could help to solve methylation issue and stabilize the transgene expression. Second, inconsistence amount of puromycine from pGIPZ and pTRIPZ results suggested that bicistronic vectors from (IRES) could be an issue. Therefore, it was recommended that picornaviridae virus family (T2A and P2A) could be preferable choice due to its high eGFP expression and gene activity. Third, the optimization of doxycycline in inducible lentiviral system (pTRIPZ and/or pSLIK-CD4) using shRNA for empty vector of pTRIPZ and/or pSLIK beside other targeting vectors could help determine the Hmgn2 knockdown that caused by endogenous loss of gene or other factors or genes. Equally, pH level in DMEM/F12 medium during culturing P19 EC knockdown cell lines changed to be more acidic medium. Thus, HEBES had helped to maintain pH level under 7.4. Moreover, lentiviral transduction process was preformed with two protocols: pSLIK protocol used lipofectamine 2000 and pTRIPZ and pGIPZ protocols used Ca_2PO_4 transfection reagent. Both methods obtained low multiplicity of infection (MOI) transfection number for P19 comparing with Hepa cells. Although Ca_2PO_4 showed low cytotoxicity in host cell (293T/17 cell line), the transfection reagents did not achieve the high transfection level in P19 EC cells. Therefore, it was critical to concentrate the viral particles to transduce P19 EC cells to obtain high MOI level, but the expression of shRNA derived by weak promoter should be changed at the first stage. Furthermore, freezing and thawing the transduced P19 EC cells with either pGIPZ or pTRIPZ failed to restore high and efficient knockdown due to

gene silencing. Finally, repeating pGIPZ and pTRIPZ lentivirus system showed a variable results in each time depend on the shRNA integration into the genome.

7.4.3 Further investigation the knockdown P19 EC and E14 cells

In chapter 5, a single lentiviral platform was employed for constitutive knock down of Hmgn2 Although more than 80% reduction in Hmgn2 protein levels was observed, still more investigation need to ensure that knockdown affect pluripotent and epiblast genes. Hence it is unclear whether the achieved Hmgn2 knockdown levels are sufficient to mediate effect on expression of its pluripotent and epiblast genes from one single experiment. It is also not possible to distinguish between the effects of Hmgn2 activity and redundancy mediated by other TFs. Further studies addressing Hmgn2 function would require an improve of the current tool to address the above described limitation.

Previously, it was suggested that one of the possible approaches could be to improve the efficiency of current RNAi mediated knockdown of Hmgn2. For instance, replace the CMV promoter to GAC or A2UCOE promoters and replace IRES to T2A or P2A. In addition, Hmgn2 mRNA could be targeted by two or more different shRNA. Therefore, cloning the shRNA for Hmgn2 (HMGN2-1) from pSLIK vector would be helpful. Then further investigation would be in gene expression and protein analysis to confirm gene knockdown. However, ensuring the knockdown of Hmgn2 is real, it can be examined by rescue experiment using retrovirus or vector to overexpress Hmgn2 and examine the change in pluripotent and epiblast stem cells by comparing them with pluripotent stem cells that changes in knockdown Hmgn2. Rescue experiment was previously used in Hmgn1 knockout mice and mouse embryonic fibroblast (MEFs) to reintroduce Hmgn1 to ATM and p53 activation. Results showed that the loss of Hmgn1 reduce the ionizing radiation (IR) phosphorylation of SMC1, CHK1 and CHK2 which are ATM targets and reduces the level of p53 and MDM2. When the Hmgn1 re-introdcued to ATM and p53, the expression of Hmgn1 increased in ATM and p53 phosphorylation and stabilization. This help to form multi-protein the factors at the DSB sites which support the ATM activation (Kim et. al., 2008).

Flow cytometer (FACS) analysis was performed in Chapter 5 for knockdown Hmgn2 in P19 EC cells. Moreover, culturing Hmgn2 knockdown P19 EC showed increase in cell proliferation (data did not show). Therefore, it was suggested that cell proliferative assay could help to identify if Hmgn2 knockdown EC cells

roles which presumably due to additional genetic events. Further investigation would be in cell cycles analysis. This was examined before on knockdown Hmgn2 P19 EC cells. However, the tested cells where from thawed cells, which had some inactive Hmgn2 knockdown vector. The results showed that the knockdown Hmgn2 EC cells accumulated in the G1 phase of the cell cycle (data not shown). Moreover, western blot was used to detect p53, but the western blot did not work (data not shown). Nevertheless, these results cannot be valid due to thawing and methylated promoter issues. Thus, cell-cycle analysis could be important experiment because Knockdown EC cells did not show substantial apoptosis.

Quantitative reverse transcriptase PCR (qRT-PCR) applied in knockdown Hmgn2 P19 EC cells (chapter 5) and confirmed the reduction of Oct4 and Nanog and Klf4. It would be helpful to distinguish changes in gene expression occur between negative control (non-silencing) and knockdown Hmgn2 in epiblast and epiblast/ES genes. Then microarray gene expression analysis of knockdown cells, which could highlight the significant gene expression differences in genes responsible for self-renewal, cell cycle, cell migration, DNA repair, apoptosis and cell differentiation.

7.4.3.1 Pluripotent markers regulation under the loss of Hmgn2

It was known that HMGN1 and HMGN2 are the only non-histone protein binding specifically to nucleosome and it is highly dynamic proteins, which are associated to transcription activity, cell cycle process, and DNA repair mechanism. Chapter 5 showed that knockdown of Hmgn2 affect the expression of pluripotent genes and induce pluripotent genes. However, mouse fibroblast stem cells Hmgn2 knockout did not detect phenotypes changes such as eye, energy metabolism, haematology, immunology, and clinical chemistry. Moreover, 19 genes were up regulated and 29 genes were down regulated in Hmgn2 knockout in mouse brain whereas in mouse thymus 173 were up regulated and 310 were down-regulated (Deng et. al., 2015). On the other hand, this study found that phenotype and gene expression changes in Hmgn1 and Hmgn2 double knockout gene in mouse (Deng et. al., 2015). Another study found that knockout Hmgn1 in mice had neither phenotypic nor transcriptional effects on mouse embryo development. In addition, this study used two different models:

nonchrondrogenic MEFs and E10.5 limb bud cells to detect the interaction between Sox9 and Hmgn1 at chromatin level. Consequently, the results showed that Hmgn1 expression regulates Sox9 chromatin in limb bud (in vivo), but failed to detect in MEFs (in vitro) (Furasawa et. al., 2006). However, both studies highlighted that whether the loss of either Hmgn1 and/or Hmgn2 is associated with chromatin modification. Thus, these studies focused on the region of active genes "Dnase hypersensitive sites" that regulates its transcription.

Active genes in embryonic stem cells and induce pluripotent stem cells were tested and discussed in chapter 5 and 6. Further experiments to confirm and address the functional discrepancy between in vitro and in vivo in mouse Hmgn2 and the adaptive time for cell to obtain stable knockdown or knockout mouse model. Hmgn2 knockdown experiments in chapter 5 showed that transduced cells took five-seven days to adapt with the loss of Hmgn2 in P19 EC cells comparing with previous studies. The significant of this study need more investigation in normal mouse embryonic stem cells such as E14 cells.

Analysis of Hmgn1 expression following Hmgn2 knockdown in P19 EC cells showed significant up regulation in Chapter 5. This possibly may explain the homeostatic mechanisms between Hmgn1 and Hmgn2. Gene targeting of Sox9 under the loss of Hmgn1 in mice produce same results although there is a lack of phenotype. The level of HMGN2 expression was similar to HMGN1 in wild type cell (Furusawa et. al., 2006). Further research is required to establish a precision mechanism between Hmgn1 and Hmgn2 in Sox9 location and whether there are other locations recruit this compensation.

7.4.4 Investigating the neuronal development in P19 EC and E14 cells

Epiblast stem cell (E 6.5) is specific to develop to three types of neurons by day E 7.5. Neuronal progenitor cells are the first neural committed cells and it is induced by signal from mesoderm and endoderm (Hitoshi et. al., 2004). Both embryonic stem cells (ESCs) and epiblast stem cells (EpiSC) undergo to different cell culture and neuronal differentiation process (Tesar et. al., 2007).

In vitro, epiblast first neural tissue marker genes are Sox1 and Nestin at day E 7.0- E 8.0. Then fibroblast growth factor (FGF2) is recruited epidermal growth factor (EGF) and it becomes detectable at day E 8.5. Moreover, bone morphogenetic proteins (BMPs) inhibit the epiblast stem cells once the neural tube form. (Hitoshi et. al., 2004). Embryonic Stem cells derived from E 5.5-E

7.5, LIF-dependent primitive neural stem cells presents in the epiblast and neuroectodermal of E 5.5-E 7.5 mouse embryo. These cells expressed GATA 4, FGF5, and Sox2/1, but these cells could not detect Brachyury (mesodermal marker), endodermal marker (HNF4), epidermal marker (cytokeratin17) at day E 7.5. However, downregulation of GATA 4 and FGF5 expression detect at day E 6.5. Based on these results it showed that epiblast LIF-dependent and ES cells are raised proliferative neural stem cells. P19 EC cells similar to ES cells gives arise to three germ layers: endoderm, mesoderm and ectoderm (Hitoshi et. al., 2004). Nerve cells derive from ectoderm layer and it forms from neuroepithelial germinal cells similar in function and morphology to cells in mammalian central nerve system. Previously P19 EC neural differentiation protocol used serum medium. Although this method can form neuronal population such as oligodendrocytes, glia cells and mature neuron, many proliferative cells (non-neuronal cells) such as astrocytes, fibroblast and skeletal cells remain in the culture. Moreover, the previous protocol based on suspension cells form embryo bodies (Ebs) (Manzo et. al., 2012).

A novel method allows generating more neural cells under defined medium. Investigating for more neural markers such as Calretinin, Calbindin, synapsin I, and NeuN. Moreover, this method helps to reduce the non-neuronal cells that reduce the neuronal number in long-term culture. In addition, more neuronal precursor cells increase which permit more investigation and characterization. The method starts with retinoic acid (RA) induction in monolayer culture. After the induction cells, the neuronal progenitor cells form under medium contains neuronal basal A medium (NBA) with supplement N2 and glutamax for five days. Then, neurite outgrowth cells form under NBA medium supplemented with B27, AraC and 2dCTD. Post-mitotic neurons become committed between culture day 10-15 UNDER NBA supplemented with B27 and synaptic integration cells (mature neurons) form by day 15-20 under the same medium. During neural forming, genes expressions are responsible for neuronal developments were detected such as Gfap, NeuN, Calretinin, Calbindin, Synapsin I, and Nestin (Monzo et. al., 2012).

Another method for P19 neuronal differentiation had improved the previous method by Monzo and colleagues (2012) in cell culture medium and less culture time point. P19 uses RA to induce the neuronal differentiation for only two days under DMEM/F12 supplemented with N2, FGF8 and DAPT in adherent culture.

Then exogenesis form for two days under the same medium with no RA. By day 4 synaptogenesis forms under medium NBA supplemented with B27 and AraC until day 6. Next, synapse maturation cells form under NBA supplemented with B27. During this neural differentiation several genes are detected such as Oct4, E-cadherin, Nestin, Laminin, MAP2c, NF-L, and synapsin I in addition to pervious genes (Nakayama et. al., 2014). However, still proliferative non-neuronal cells remain an issue.

A current explanation by Hamada-Kanazawa and colleagues (2016) suggested that RA induction caused an increase of BMP-4 expression while Sox6 inhibit the BMP-4 expression which cause more wild type P19 cells remains in the culture. However, an alternative method is to overexpress the Sox6 to increase the expression of neuronal differentiation lineage without the need to RA (Hamada-Kanazawa et. al., 2016) using viral vector.

7.4.5 Investigating P19 EC Knockdown during neuronal differentiation under serum free medium

As discussed in chapter 5, Hmgn2 knock down in P19 EC cells resulted in downregulation of Oct4 and Nanog, and Klf4 but upregulation of Sox2 and C-myc. It was knockdown that Sox2 overexpression could lead to cell committed toward differentiation. However, drawing a conclusion based on current result presents significant limitation, as Hmgn2 knockdown could be vary between biological triplicates. Similarly, Hmgn2 knockdown in P19 during neuronal differentiation could be challenging due to the limitation in current knockdown system and differentiation protocol. Previous study showed the P19 neural development was investigated at day 3, 6, and 18 and knockdown during these days using siRNA (Mohan, 2012). The later could be knockdown P19 at embryonic stage then inducing these cells using RA beside inducing wild type P19 cells to neural development using serum free media and growth factors as supplement. Then confirming the key neuronal development markers would apply detection by immunofluorescence (IF) against Hmgn2 and neural gene makers besides qRT-PCR analysis.

7.4.6 Genome wide approaches to study Hmgn2 regulatory network

Unfortunately there is no public ENCODE region for HMGN2 binding site in mouse genome. However, Public ENCODE data for Oct4, Sox2 and Nanog binding sites from mouse embryonic stem cells is available. Therefore, investigate pluripotent genes using qRT-PCR and western blot for chromatin to ensure the quality of data are still highly demanding following Chip-seq and bioinformatics analysis, which ENCODE regions represents a large portion of mouse genome. This would permit manual data analysis, which is the first attempting of qPCR-ChIP as well as being readily affordable. In addition, Oct4, Sox2 and Nanog are well characterized in mouse stem cells.

As discussed in Chapter 6, Hmgn2 knockdown qPCR-ChIP highlighted in two regions that is enriched with H3K4me3 in Oct4 and Nanog transcription-starting sites (TSS). QPCR-ChIP primers were designed before TSS, within TSS and after TSS. Moreover, shRNA for Hmgn2-381 found to be within TSS region of Oct4 and Nanog. Hmgn2 knockdown results revealed global loss of Hmgn2 besides loss of H3K4me3 across Oct4 and Nanog TSS. These results contradict with Deng and colleagues (2015), which showed the level of H3K4me3 and H3K4me1 from ChIP-seq data did not changed due to either the knockout of Hmgn2 or double knockout of Hmgn1 and Hmgn2 although they mentioned that the loss of Hmgn1 and Hmgn2 affects Dnase hypersensitive sites (DHSs) at the enhancer region (Deng et. al., 2015). In addition, enrichment peak of Hmgn1 besides modest enrichment peak of H3k27me3 identified at the Oct4 and Nanog TSS. These results have some contradictions with other studies. The enrichment of Hmgn1 at the active genes regions such as Oct4 and Nanog due to the loss of Hmgn2 found in the study of investigating the role of Hmgn1 in Sox9 expression. ChIP-seq analysis found that the loss of Hmgn1 in limb bud associated with Sox9 chromatin and significant enrichment of Hmgn2 at the same Sox9 region. This data suggests that during the loss of Hmgn1 in limb bud, but not in MEFs, HMGN2 expression level is similar to normal HMGN1 (Furusawa et. al., 2006). Further results need to establish to confirm three points before applying comprehensive analysis of Hmgn2 function by combining Hmgn2 ChIP-seq analysis with RNA-seq data or gene expression data.

First, further research is required to identify the epiblast stem cell gene expression under Hmgn2 knockdown cells. This will allow studying the effect of Hmgn2 during cell development since no study focused on the relationship

between Hmgn2 and chromatin of active genes that are important for epiblast identity including Oct4, Sox2 and Nanog.

Second, identify epiblast stem cell population that can commit to neuronal precursor cells using cell surface markers that distinguish epiblast from embryonic stem cells. This step is important because epiblast stem cells have two options: reverse to embryonic stem cells or commit to neuronal stem cells and it will help to avoid proliferative cells during neuronal differentiation.

Then, Hmgn2 knockdown in P19 EC cells besides E14 cells using modified pGIPZ would help to detect the loss Hmgn2 effects on epiblast and embryonic stem cell genes. Furthermore, qPCR-ChIP experiment is still a valid analysis to ensure the loss of Hmgn2 in pluripotent gene regions.

However, it would be feasible to apply qPCR-ChIP to investigate promoter and super-enhancer at Dnase hypersensitive regions. In general, promoters is a DNA cis-regulatory elements that found in proximity of transcription starting sites to confirm the initiation of the transcription process while enhancers can stimulate the transcription level of its target gene. The interaction between promoters and enhancers may happen during “looping” and “tracking”: looping means that DNA location between enhancer and promoter form loop to increase the interaction and tracking occur when enhancer stimulate the RNA polymerase II to DNA promoter to initiate the transcription mechanism. Moreover, both promoter and enhancer showed high sensitivity to Dnase I digest which indicate the alteration of nucleosome organization leading to increase chromatin accessibility. In addition to promoter and enhancers, Dnase I hypersensitive sites include other regulatory gene complexes such as silencer and insulator (Martinez de Paz and Ausio, 2016). Therefore, they are highly dynamic during cells committing to differentiation. Previously, Deng and colleges (2015) shed the light toward the roles of Hmgn1 and Hmgn2 at Dnase hypersensitive sites (DHSs). Therefore, it will be helpful to identify the common promoters and super-enhancers based on epiblast stem cell genes. QPCR for ChIP besides mRNA analysis would be the initial step to identify these genes.

According to recent study epiblast stem cells promoters are different from embryonic stem cells based on signalling pathways that control the self-renewal and De novo DNA methylation such as methyltransferase 3a, 3b, and 3l using RNA-seq. Each promoter should be classified based on methylation level in epiblast stem cells. Epiblast hyper methylation promoters such as Dppa3 (Stella),

Zfp42 (Rex1), Tbx3, which found to be part of gene regulations expressed different in naïve from primed pluripotent cells. In addition, these promoters found to be linked to genes responsible for molecular transport, metabolism, signalling pathway and neural development. On the other hand, epiblast hypomethylation promoters such as Abcb1a, Aebp1, Chrna3, Cspg4, Daam2, and Gfra3 contribute with bivalent promoter in ESCs. These promoters play roles in membrane transportation, nervous system and muscle development. Besides all these genes, gene expression of epiblast stem cells such as Esrrb, Klf4 (from early epiblast stem cells) Nodal, Fgf5, Brachyury and Cer1 (late epiblast stem cells) should be examine including Oct4, Sox2, and Nanog (Velliard et. al., 2014).

Recent studies highlight on super-enhancer in stem cells. Super-enhancer (high level of mediators) defined as large cluster of transcriptional enhancers that are dynamically involve in regulatory elements deriving the gene expression responsible for cell identity (Whyte et. al., 2013 and Ding et. al., 2015). May be most of super-enhancer study focused on embryonic stem cells genes such as Sox2, Oct4 and Nanog and other genes such as Klf4, Essrb, C-myc that are involve in embryonic and early epiblast stem cells development. Therefore, it is challenging to reply on these studies albeit it is very important to understand gene selection via ChIP-seq data. However, Oct4, Sox2 and Nanog are express in both epiblast and embryonic stem cells, which can be used to compare with P19 EC ChIP-seq. furthermore, nucleosome with the histone modification H3K27ac and H3K4me1 are enriched with active enhancers in ESCs. Another acetylated nucleosome found in enhancer and promoter called acetylated histone at lysine 9 and phosphorylated at serine 10 (H3K9acS10ph) which play important roles in enhancer function and transcription process of drosophila (Martinez da Paz and Ausio, 2016). This can be used to compare P19 EC ChIP-seq data from Sox2, Oct4, Nanog, and Otx2, mediator, Histone modifications and Dnase I hypersensitive sites to segregate super-enhancers from normal enhancers.

7.4.7 Preforming the Hmgn2 knockout using CRISPR-CAS9 technology as an alternative effective method

As result of the gene targeting performed in mouse stem cells, Hmgn2 mouse embryonic knockout become available (Deng et al., 2015). These provide a tool for investigating Hmgn2 function, albeit accompanied with several limitations. In

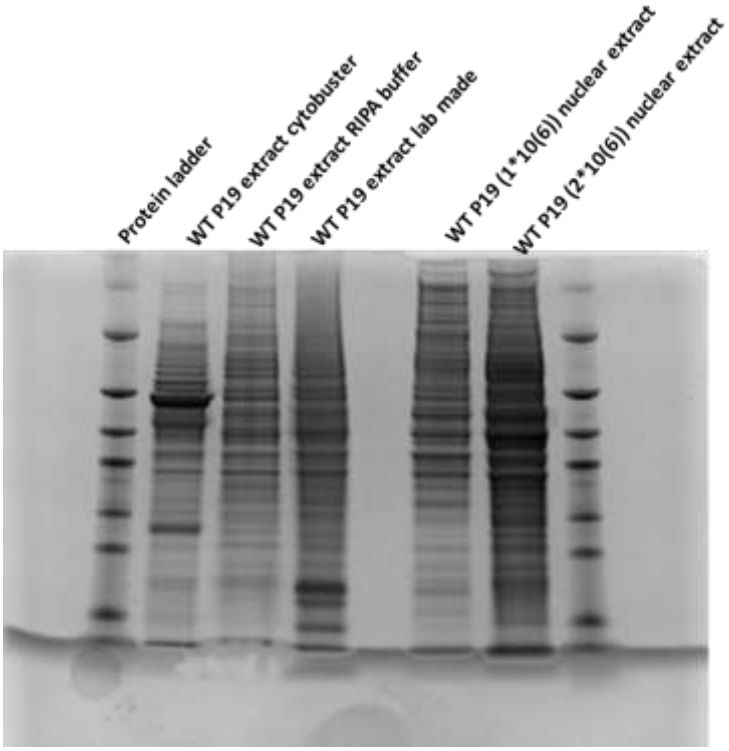
particular it may not be possible to investigate the role of Hmgn2 in neural gene regulation and development due to the long mouse breed and the possibility that Hmgn2 replaced with other modulators or protein complexes that are involve in Dnase I hypersensitive sites to maintain cell pluripotency and identity. Recent studies have successfully employed genome engineered clustered regularly interspaced palindromic repeat associated Cas9 nuclease (CRISPR-Cas9) in order to perform stable gene manipulation experiment in mammalian cells (Doudna and Charpentier, 2014). It may therefore be possible to design CRISPR-Cas9 targeting the full length of HMGN2 coding sequences in order to generate stable HMGN2 knockout in mouse stem cells (P19 EC cell and E14 cells used in this study).

The CRISPR-Cas9 approach to knockout Hmgn2 can clarify if the hmgn2 present following the Chapter 5 and 6 described RNAi knockdown is responsible for changes pluripotent stem cell expression of some Hmgn2 target genes. Flowing the successful application of CRISPR-Cas9 described approaches; the global effects of disrupting Hmgn2 function can be assessed by performing whole genome gene expression analysis (RNA-seq).

Appendix

Protein cell extract using different lysis buffer

A.



B.

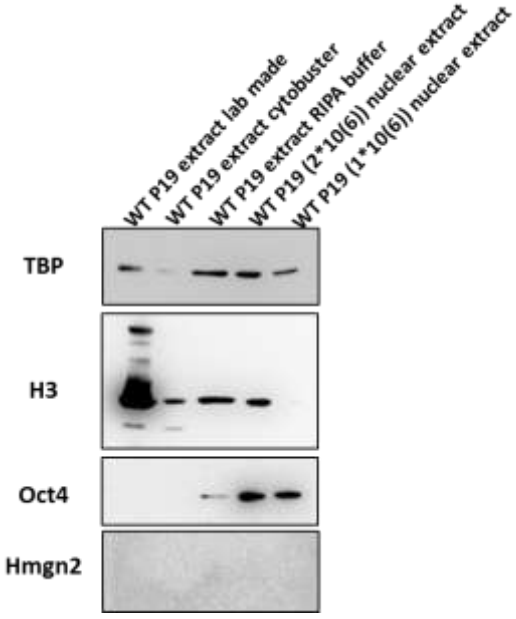


Figure X 1 western blot and Coomassie blue for wild type P19 EC cells extracted with different cell extraction buffer
A. Coomassie blue showed wild type P19 EC cells with different lysis buffer: cytobuster (Millipore), RIPA buffer (Millipore), lab made (provided by Dr. Katherine west lab), and nuclear extract (provided by Dr. Adam west lab) with different cell number extraction. B. western blots showed some proteins expression; TBP, H3, Oct4 and Hmgn2 under different lysis buffer for Wild type P19 EC cells. TBP and H3 were used as control for Oct4 and Hmgn2.

ShHMGN2-381 sequence alignment from UCSC browser

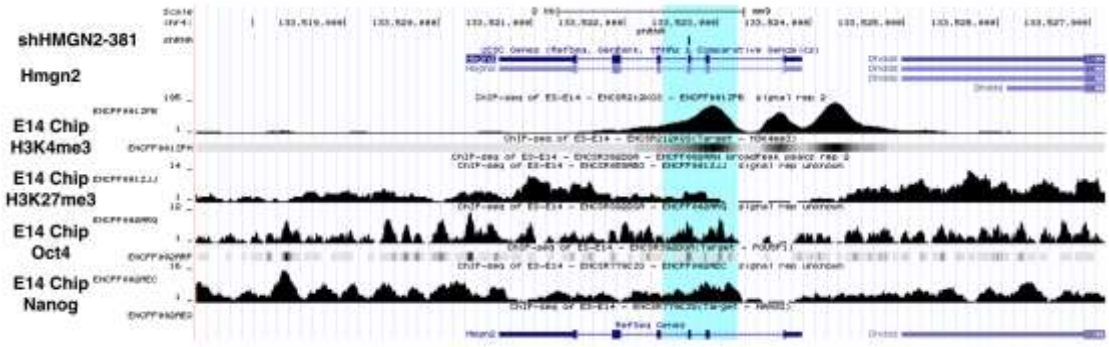


Figure X 2 UCSC genome browser of mouse Hmgn2 region in chromosome 4
The data include shHMGN2-381 sequence targeting the nucleosome binding domain region. the sequence was mapped with E14 chip-seq data for histones (H3K4me3 and H3K27me3) and pluripotent genes (Oct4 and Nanog) (https://www.encodeproject.org). ShHMGN2-381 region was located at enrichment site of H3k4me3 of E14 cells. E14 Chip-seq experiments came from three experiments were downloaded from ENCODE project.

Western blot for active histone promoter H3K4me3 in Hmgn2 knockdown P19 EC cells

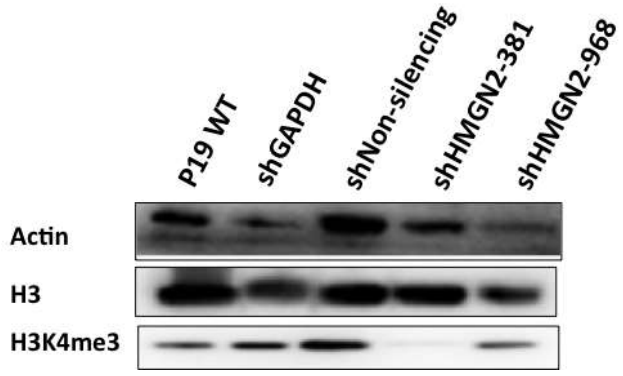


Figure X 3 Hmgn2 knockdown on P19 EC cells using lentiviral constitutive system
Western blot analysis H3K4me3 on shRNA (shGAPDH, shNon-silencing, shHMGN2-381 and shHMGN2-968) was performed on day 20. Anti-H3K4me3 antibodies were

detected against H3K4me3 in Hmgn2 knockdown P19 EC cells whole protein extract. All the knockdown cells showed normal expression comparing with H3 and ACTIN (controls) except shHMGN2-381 did not show any H3k4me3 signal on day 20.

Western blot for TPB protein in Hmgn2 knockdown P19 EC cell with PTRIPZ after Dox treatment for 5 days

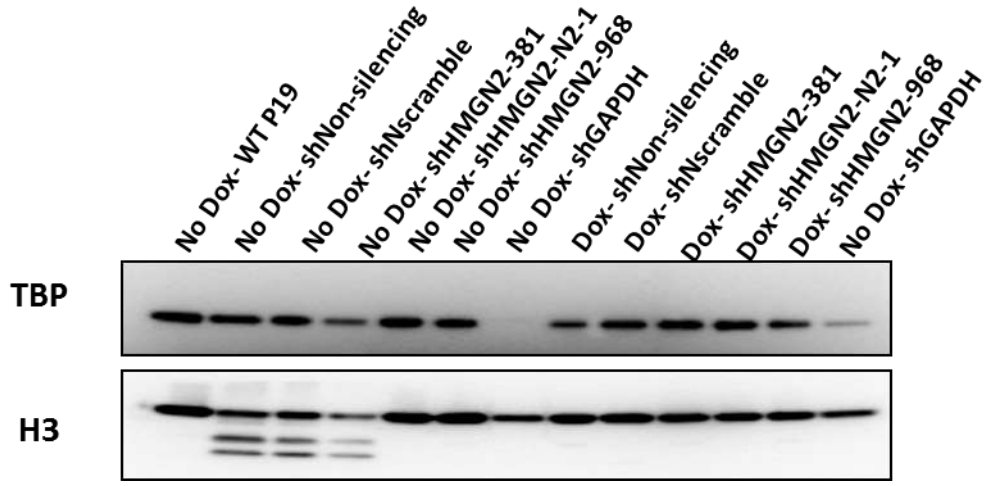


Figure X 4 TBP and H3 (control) western blot in Hmgn2 knockdown P19 EC cells with pTRIPZ after 5 days dox treatment
(Notice that PTRIPZ did not gain high hmgn2 knockdown in FACS analysis and RT-PCR).

Western blot for TBP protein in Hmgn2 knockdown P19 EC cells using PGIPZ

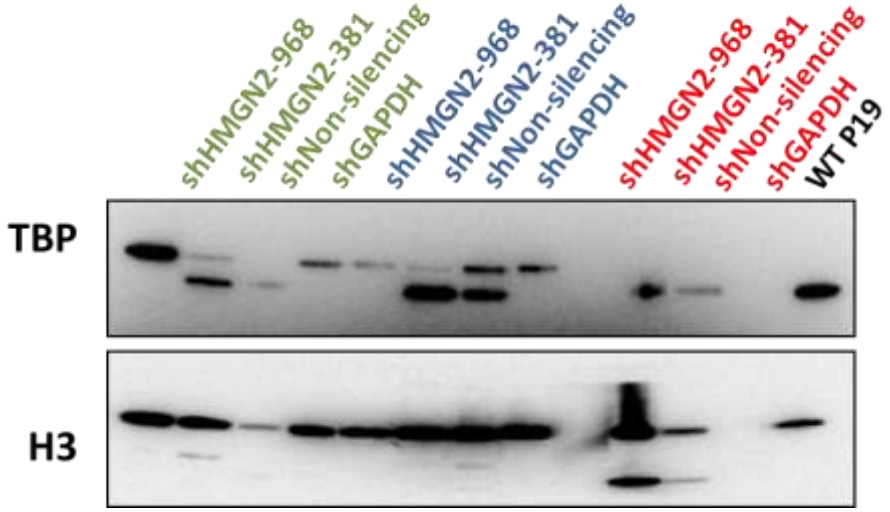


Figure X 5 TBP and H3(control) western blot in Hmgn2 knockdown P19 EC cells with PGIPZ lentivirus at day 10 (red), day 20 (blue) and day 25 (green)

List of Reference

1. Aagaard, L. and J. J. Rossi (2007). "RNAi Therapeutics: Principles, Prospects and Challenges." Advanced drug delivery reviews **59**(2-3): 75-86.

2. Abu-Remaileh, M., et al. (2010). "Oct-3/4 regulates stem cell identity and cell fate decisions by modulating Wnt/beta-catenin signalling." EMBO J **29**: 3236 - 3248.

3. Ambrosetti, D., et al. (1997). "Synergistic activation of the fibroblast growth factor 4 enhancer by Sox2 and Oct-3 depends on protein-protein interactions facilitated by a specific spatial arrangement of factor binding sites." Mol Cell Biol **17**: 6321 - 6329.

4. Amorim, R., et al. (2014). "HIV-1 Transcripts Use IRES-Initiation under Conditions Where Cap-Dependent Translation Is Restricted by Poliovirus 2A Protease." PLoS ONE **9**(2): e88619.

5. Andrews, P. W. (2002). "From teratocarcinomas to embryonic stem cells." Philosophical Transactions of the Royal Society B: Biological Sciences **357**(1420): 405-417.

6. Angello, J. C., et al. (1997). "P19 Embryonal Carcinoma Cells: A Model System for Studying Neural Tube Induction of Skeletal Myogenesis." Developmental biology **192**(1): 93-98.

7. Annunziato, A. ((2008)). DNA Packaging: Nucleosomes and Chromatin. . Nature Education. **1:** 26

8. Ansaloni, S., et al. (2010). "A streamlined sub-cloning procedure to transfer shRNA from a pSM2 vector to a pGIPZ lentiviral vector." Journal of RNAi and Gene Silencing : An International Journal of RNA and Gene Targeting Research **6**(2): 411-415.

9. Ansorge, S., et al. (2010). "Recent progress in lentiviral vector mass production." Biochemical Engineering Journal **48**(3): 362-377.

10. Arnaudo, A. M. and B. A. Garcia (2013). "Proteomic characterization of novel histone post-translational modifications." Epigenetics & Chromatin **6**(1): 1-7.

11. Arrowsmith, C. H., et al. (2012). "Epigenetic protein families: a new frontier for drug discovery." Nat Rev Drug Discov **11**(5): 384-400.

12. Artus, J., et al. (2011). "The primitive endoderm lineage of the mouse blastocyst: Sequential transcription factor activation and regulation of differentiation by Sox17." Developmental biology **350**(2): 393-404.

13. Atwood, B., et al. (2015). Contributors. Epigenetic Gene Expression and Regulation. Oxford, Academic Press.

14. Avilion, A. A., et al. (2003). "Multipotent cell lineages in early mouse development depend on SOX2 function." Genes & Development **17**(1): 126-140.

15. Azzam, M. E. and I. D. Algranati (1973). "Mechanism of Puromycin Action: Fate of Ribosomes after Release of Nascent Protein Chains from Polysomes." Proceedings of the National Academy of Sciences of the United States of America **70**(12 Pt 1-2): 3866-3869.

16. Bai, L. and A. V. Morozov (2010). "Gene regulation by nucleosome positioning." Trends in Genetics **26**(11): 476-483.

17. Bain, G., et al. (1994). "From embryonal carcinoma cells to neurons: The P19 pathway." BioEssays **16**(5): 343-348.

18. Bain, G., et al. (1994). "From embryonal carcinoma cells to neurons: The P19 pathway." BioEssays **16**(5): 343-348.

19. Ballaré, C., et al. (2013). "Nucleosome-Driven Transcription Factor Binding and Gene Regulation." Molecular Cell **49**(1): 67-79.
20. Bandeira, V., et al. (2012). "Downstream Processing of Lentiviral Vectors: Releasing Bottlenecks." Human Gene Therapy Methods **23**(4): 255-263.

21. Bannister, A. J. and T. Kouzarides (2011). "Regulation of chromatin by histone modifications." Cell Research **21**(3): 381-395.

22. Barkess, G., et al. (2012). "The chromatin-binding protein HMGN3 stimulates histone acetylation and transcription across the Glyt1 gene." The Biochemical journal **442**(3): 495-505.

23. Bastos, G. M., et al. (2008). "Immunolocalization of the High-Mobility Group N2 protein and acetylated histone H3K14 in early developing parthenogenetic bovine embryos derived from oocytes of high and low developmental competence." Molecular Reproduction and Development **75**(2): 282-290.

24. Baumann, V. and J. Winkler (2014). "miRNA-based therapies: strategies and delivery platforms for oligonucleotide and non-oligonucleotide agents." Future Medicinal Chemistry **6**(17): 1967-1984.

25. Bell, J. C. and A. F. Straight (2015). "Condensing chromosome condensation." Nat Cell Biol **17**(8): 964-965.

26. Bensadoun, J.-C., et al. (2000). "Lentiviral Vectors as a Gene Delivery System in the Mouse Midbrain: Cellular and Behavioral Improvements in a 6-OHDA Model of Parkinson's Disease Using GDNF." Experimental Neurology **164**(1): 15-24.

27. Ben-Shushan, E., et al. (1995). "A dynamic balance between ARP-1/COUP-TFII, EAR-3/COUP-TFI, and retinoic acid receptor:retinoid x receptor heterodimers regulates Oct-3/4 expression in embryonal carcinoma cells."

Mol Cell Biol **15**: 1034 - 1048.

28. Benyoucef, A. and M. Brand (2015). Chapter 7 - Epigenetic gene regulation and stem cell function A2 - Blakey, Suming HuangMichael D. LittC. Ann. Epigenetic Gene Expression and Regulation. Oxford, Academic Press**:** 149-181.

29. Bernstein, B. E., et al. (2006). "A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells." Cell **125**(2): 315-326.

30. Birger, Y., et al. (2003). "Chromosomal protein HMGN1 enhances the rate of DNA repair in chromatin." The EMBO Journal **22**(7): 1665-1675.

31. Birger, Y., et al. (2005). "Increased Tumorigenicity and Sensitivity to Ionizing Radiation upon Loss of Chromosomal Protein HMGN1." Cancer Research **65**(15): 6711-6718.

32. Bischof, O. and A. Dejean (2007). "SUMO is growing senescent." Cell cycle (Georgetown, Tex.) **6**(6): 677-681.

33. Blakey, C. A. and M. D. Litt (2015). "Histone modifications—models and mechanisms." Epigenetic Gene Expression and Regulation: 21.

34. Bond, M. R. and J. A. Hanover (2015). "A little sugar goes a long way: the cell biology of O-GlcNAc." The Journal of Cell Biology **208**(7): 869-880.

35. Boroviak, T., et al. (2014). "The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification." Nat Cell Biol **16**(6): 513-525.

36. Boyer, L. A., et al. (2005). "Core Transcriptional Regulatory Circuitry in Human Embryonic Stem Cells." Cell **122**(6): 947-956.

37. Brant, J. O. and T. P. Yang (2015). Chapter 15 - Epigenetic effects of environment and diet A2 - Blakey, Suming HuangMichael D. LittC. Ann. Epigenetic Gene Expression and Regulation. Oxford, Academic Press**:** 339-365.

38. Brehm, A., et al. (1997). "The carboxy-terminal transactivation domain of Oct-4 acquires cell specificity through the POU domain." Mol Cell Biol **17**: 154 - 162.

39. Brons, I. G. M., et al. (2007). "Derivation of pluripotent epiblast stem cells from mammalian embryos." Nature **448**(7150): 191-195.

40. Buchholz, D. (2012). Tet-On Binary Systems for Tissue-Specific and Inducible Transgene Expression. Xenopus Protocols. S. Hoppler and P. D. Vize, Humana Press. **917:** 265-275.

41. Buck, M. J. and J. D. Lieb (2006). "A chromatin-mediated mechanism for specification of conditional transcription factor targets." Nat Genet **38**(12): 1446-1451.

42. Bustin, M. (2001). "Chromatin unfolding and activation by HMGN* chromosomal proteins." Trends in Biochemical Sciences **26**(7): 431-437.

43. Bustin, M., et al. (1995). "The HMG-14/-17 chromosomal protein family: architectural elements that enhance transcription from chromatin templates." Seminars in Cell Biology **6**(4): 247-255.

44. Cantone, I. and A. G. Fisher (2013). "Epigenetic programming and reprogramming during development." Nat Struct Mol Biol **20**(3): 282-289.

45. Cao, J. and Q. Yan (2012). "Histone Ubiquitination and Deubiquitination in Transcription, DNA Damage Response, and Cancer." Frontiers in Oncology **2**: 26.

46. Card, D., et al. (2008). "Oct4/Sox2-regulated miR-302 targets cyclin D1 in human embryonic stem cells." Mol Cell Biol **28**: 6426 - 6438.

47. Carmo, M., et al. (2009). "Stabilization of gammaretroviral and lentiviral vectors: from production to gene transfer." The Journal of Gene Medicine **11**(8): 670-678.

48. Carr, S. M., et al. (2015). "Post-translational control of transcription factors: methylation ranks highly." FEBS Journal **282**(23): 4450-4465.

49. Cartwright, P., et al. (2005). "LIF/STAT3 controls ES cell self-renewal and pluripotency by a Myc-dependent mechanism." Development **132**(5): 885-896.

50. Catez, F., et al. (2003). "HMGN dynamics and chromatin function." Biochemistry and Cell Biology **81**(3): 113-122.

51. Chambers, I. and S. Tomlinson (2009). "The transcriptional foundation of pluripotency." Development **136**: 2311 - 2322.

52. Chambers, I., et al. (2007). "Nanog safeguards pluripotency and mediates germline development." Nature **450**(7173): 1230-1234.

53. Chang, G., et al. (2010). "Linking Incomplete Reprogramming to the Improved Pluripotency of Murine Embryonal Carcinoma Cell-Derived Pluripotent Stem Cells." PLoS ONE **5**(4): e10320.

54. Chang, K., et al. (2006). "Lessons from Nature: microRNA-based shRNA libraries." Nat Meth **3**(9): 707-714.

55. Chang, M.-Y., et al. "Doxycycline Enhances Survival and Self-Renewal of Human Pluripotent Stem Cells." Stem Cell Reports **3**(2): 353-364.

56. Chavez, L., et al. (2009). "In silico identification of a core regulatory network of OCT4 in human embryonic stem cells using an integrated approach." BMC Genomics **10**: 314.

57. Chen, T. and S. Y. R. Dent (2014). "Chromatin modifiers and remodellers: regulators of cellular differentiation." Nat Rev Genet **15**(2): 93-106.

58. Chen, T., et al. (2014). "Foxa1 contributes to the repression of Nanog expression by recruiting Grg3 during the differentiation of pluripotent P19 embryonal carcinoma cells." Experimental Cell Research **326**(2): 326-335.

59. Chen, X., et al. (2008). "Integration of external signaling pathways with the core transcriptional network in embryonic stem cells." Cell **133**: 1106 - 1117.

60. Cheung, P., et al. (2000). "Synergistic Coupling of Histone H3 Phosphorylation and Acetylation in Response to Epidermal Growth Factor Stimulation." Molecular Cell **5**(6): 905-915.

61. Chew, J., et al. (2005). "Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells." Mol Cell Biol **25**: 6031 - 6046.

62. Chibazakura, T., et al. (1997). "Phosphorylation of Human General Transcription Factors TATA-Binding Protein and Transcription Factor IIB by DNA-Dependent Protein Kinase." European Journal of Biochemistry **247**(3): 1166-1173.

63. Choi, S.-C., et al. (2012). "Nanog regulates molecules involved in stemness and cell cycle-signaling pathway for maintenance of pluripotency of P19 embryonal carcinoma stem cells." Journal of Cellular Physiology **227**(11): 3678-3692.

64. Cockrell, A. and T. Kafri (2007). "Gene delivery by lentivirus vectors." Molecular Biotechnology **36**(3): 184-204.

65. Connolly, J. B. (2002). "Lentiviruses in gene therapy clinical research." Gene therapy **9**(24): 1730-1734.

66. Cooray, S., et al. (2012). Chapter three - Retrovirus and Lentivirus Vector Design and Methods of Cell Conditioning. Methods in Enzymology. F. Theodore, Academic Press. **Volume 507:** 29-57.

67. Corbel, S. Y. and F. M. V. Rossi (2002). "Latest developments and in vivo use of the Tet system: ex vivo and in vivo delivery of tetracycline-regulated genes." Current Opinion in Biotechnology **13**(5): 448-452.

68. Couto, L. B. and K. A. High (2010). "Viral vector-mediated RNA interference." Current Opinion in Pharmacology **10**(5): 534-542.

69. Cribbs, A., et al. (2013). "Simplified production and concentration of lentiviral vectors to achieve high transduction in primary human T cells." BMC Biotechnology **13**(1): 98.

70. Cribbs, A., et al. (2013). "Simplified production and concentration of lentiviral vectors to achieve high transduction in primary human T cells." BMC Biotechnology **13**(1): 98.

71. Crippa, M. P., et al. (1992). "Nucleosome core binding region of chromosomal protein HMG-17 acts as an independent functional domain." Journal of Molecular Biology **228**(2): 442-449.

72. Crippa, M. P., et al. (1993). "Deposition of chromosomal protein HMG-17 during replication affects the nucleosomal ladder and transcriptional potential of nascent chromatin." The EMBO Journal **12**(10): 3855-3864.

73. Cullen, B. R. (2004). "Transcription and Processing of Human microRNA Precursors." Molecular Cell **16**(6): 861-865.

74. Cusanovich, D. A., et al. (2015). "Multiplex single-cell profiling of chromatin accessibility by combinatorial cellular indexing." Science **348**(6237): 910-914.

75. Cutter, A. R. and J. J. Hayes (2015). "A brief review of nucleosome structure." FEBS Letters **589**(20, Part A): 2914-2922.

76. Dailey, L. (2015). "High throughput technologies for the functional discovery of mammalian enhancers: New approaches for understanding transcriptional regulatory network dynamics." Genomics **106**(3): 151-158.

77. de las Mercedes Segura, M., et al. (2006). "Downstream processing of oncoretroviral and lentiviral gene therapy vectors." Biotechnology Advances **24**(3): 321-337.

78. Delgado-Olguín, P. and F. Recillas-Targa (2011). "Chromatin structure of pluripotent stem cells and induced pluripotent stem cells." Briefings in Functional Genomics **10**(1): 37-49.

79. Deng, C., et al. (2015). Chapter 5 - Chromatin dynamics and genome organization in development and disease A2 - Blakey, Suming HuangMichael D. LittC. Ann. Epigenetic Gene Expression and Regulation. Oxford, Academic Press**:** 95-115.

80. Deng, L.-X., et al. (2012). "The Chromosomal Protein HMGN2 Mediates the LPS-Induced Expression of β-Defensins in Mice." Inflammation **35**(2): 456-473.

81. Deng, T., et al. (2013). "HMGN1 Modulates Nucleosome Occupancy and DNase I Hypersensitivity at the CpG Island Promoters of Embryonic Stem Cells." Molecular and Cellular Biology **33**(16): 3377-3389.

82. Deng, T., et al. (2015). "Functional compensation among HMGN variants modulates the DNase I hypersensitive sites at enhancers." Genome Research.

83. Deshpande, A., et al. (2009). "Cdk2ap1 is required for epigenetic silencing of Oct4 during murine embryonic stem cell differentiation." J Biol Chem **284**: 6043 - 6047.

84. Ding, J., et al. (2015). "Tex10 Coordinates Epigenetic Control of Super-Enhancer Activity in Pluripotency and Reprogramming." Cell Stem Cell

16(6): 653-668.

85. Ding, L., et al. (2009). "A genome-scale RNAi screen for Oct4 modulators defines a role of the Paf1 complex for embryonic stem cell identity." Cell Stem Cell **4**: 403 - 415.

86. Ding, L., et al. (2015). "Systems Analyses Reveal Shared and Diverse Attributes of Oct4 Regulation in Pluripotent Cells." Cell Systems **1**(2): 141-151.

87. Dixon, J. R., et al. (2015). "Chromatin architecture reorganization during stem cell differentiation." Nature **518**(7539): 331-336.

88. Donohoe, M., et al. (2009). "The pluripotency factor Oct4 interacts with Ctcf and also controls X-chromosome pairing and counting." Nature **460**: 128 - 132.

89. Doudna, J. A. and E. Charpentier (2014). "The new frontier of genome engineering with CRISPR-Cas9." Science **346**(6213).

90. Duygu, B., et al. (2013). "Genetics and epigenetics of arrhythmia and heart failure." Frontiers in Genetics **4**: 219.

91. Dykxhoorn, D. M., et al. (2003). "Killing the messenger: short RNAs that silence gene expression." Nat Rev Mol Cell Biol **4**(6): 457-467.

92. Eberharter, A. and P. B. Becker (2002). "Histone acetylation: a switch between repressive and permissive chromatin: Second in review series on chromatin dynamics." EMBO Reports **3**(3): 224-229.

93. Eleftheriadou, I. and N. Mazarakis (2015). Lentiviral Vectors for Gene Delivery to the Nervous System. Gene Delivery and Therapy for Neurological Disorders. X. Bo and J. Verhaagen, Springer New York. **98:** 23-66.

94. Endoh, M., et al. (2012). "Histone H2A Mono-Ubiquitination Is a Crucial Step to Mediate PRC1-Dependent Repression of Developmental Genes to Maintain ES Cell Identity." PLoS Genet **8**(7): e1002774.

95. Engel, N. (2015). Chapter 3 - Genomic imprinting in mammals—memories of generations past A2 - Blakey, Suming HuangMichael D. LittC. Ann. Epigenetic Gene Expression and Regulation. Oxford, Academic Press**:** 43-61.

96. Fagnocchi, L., et al. (2016). "Integration of Signaling Pathways with the Epigenetic Machinery in the Maintenance of Stem Cells." Stem Cells International **2016**: 8652748.

97. Fan, A. X., et al. (2015). Chapter 11 - Regulation of erythroid cell differentiation by transcription factors, chromatin structure alterations, and noncoding RNA A2 - Blakey, Suming HuangMichael D. LittC. Ann. Epigenetic Gene Expression and Regulation. Oxford, Academic Press**:** 237-

264.

98. Fellmann, C. and S. W. Lowe (2014). "Stable RNA interference rules for silencing." Nat Cell Biol **16**(1): 10-18.

99. Felsenfeld, G. (2015). Preface A2 - Blakey, Suming HuangMichael D. LittC. Ann. Epigenetic Gene Expression and Regulation. Oxford, Academic Press**:** xv-xviii.

100. Feng, B., et al. (2009). "Molecules that promote or enhance reprogramming of somatic cells to induced pluripotent stem cells." Cell Stem Cell **4**: 301 - 312.

101. Fernandes, J., et al. (2012). "The HIV-1 Rev response element: An RNA scaffold that directs the cooperative assembly of a homo-oligomeric ribonucleoprotein complex." RNA Biology **9**(1): 6-11.

102. Fewell, G. D. and K. Schmitt (2006). "Vector-based RNAi approaches for stable, inducible and genome-wide screens." Drug Discovery Today **11**(21–22): 975-982.

103. Fiorillo, A. A., et al. (2011). "HMGN2 Inducibly Binds a Novel Transactivation Domain in Nuclear PRLr to Coordinate Stat5a-Mediated Transcription." Molecular Endocrinology **25**(9): 1550-1564.

104. Fryer, R. M., et al. (2002). "Global Analysis of Gene Expression: Methods, Interpretation, and Pitfalls." Nephron Experimental Nephrology **10**(2): 64-74.

105. Fuhrmann, G., et al. (2001). "Mouse germline restriction of Oct4 expression by germ cell nuclear factor." Dev Cell **1**: 377 - 387.

106. Furusawa, T. and S. Cherukuri (2010). "Developmental function of HMGN Proteins." Biochimica et Biophysica Acta **1799**(1-2): 69.

107. Furusawa, T., et al. (2006). "Down-Regulation of Nucleosomal Binding Protein HMGN1 Expression during Embryogenesis Modulates Sox9 Expression in Chondrocytes." Molecular and Cellular Biology **26**(2): 592-604.

108. Fussner, E., et al. (2012). "Open and closed domains in the mouse genome are configured as 10-nm chromatin fibres." EMBO Reports **13**(11): 992-996.

109. Gagliardi, A., et al. (2013). "A direct physical interaction between Nanog and Sox2 regulates embryonic stem cell self-renewal." The EMBO Journal **32**(16): 2231-2247.

110. García, M. A., et al. (2007). "The dsRNA protein kinase PKR: Virus and cell control." Biochimie **89**(6–7): 799-811.

111. Gerlitz, G., et al. (2009). "The dynamics of HMG protein–chromatin

interactions in living cells()." Biochemistry and cell biology = Biochimie et biologie cellulaire **87**(1): 127-137.

112. Ghosh, S. and Z. Zhou (2015). Chapter 14 - Epigenetics of physiological and premature aging A2 - Blakey, Suming HuangMichael D. LittC. Ann. Epigenetic Gene Expression and Regulation. Oxford, Academic Press**:** 313-338.

113. Giles, K. E., et al. (2015). Chapter 6 - ncRNA function in chromatin organization A2 - Blakey, Suming HuangMichael D. LittC. Ann. Epigenetic Gene Expression and Regulation. Oxford, Academic Press**:** 117-148.

114. Gorman, C. M., et al. (1983). "Expression of recombinant plasmids in mammalian cells is enhanced by sodium butyrate." Nucleic Acids Research **11**(21): 7631-7648.

115. Greber, B., et al. (2010). "Conserved and Divergent Roles of FGF Signaling in Mouse Epiblast Stem Cells and Human Embryonic Stem Cells." Cell Stem Cell **6**(3): 215-226.

116. Gu, P., et al. (2005). "Orphan nuclear receptor LRH-1 is required to maintain Oct4 expression at the epiblast stage of embryonic development." Mol Cell Biol **25**: 3492 - 3505.

117. Guo, G., et al. (2009). "Klf4 reverts developmentally programmed restriction of ground state pluripotency." Development **136**(7): 1063-1069.

118. Guo, Y., et al. (2002). "The embryonic stem cell transcription factors Oct-4 and FoxD3 interact to regulate endodermal-specific promoter expression." Proc Natl Acad Sci USA **99**: 3663 - 3667.

119. Gupta, P., et al. (2008). "Retinoic acid-stimulated sequential phosphorylation, PML recruitment, and SUMOylation of nuclear receptor TR2 to suppress Oct4 expression." Proc Natl Acad Sci USA **105**: 11424 - 11429.

120. Gupta, S., et al. (2004). "Inducible, reversible, and stable RNA interference in mammalian cells." Proceedings of the National Academy of Sciences of the United States of America **101**(7): 1927-1932.

121. Guzman-Ayala, M., et al. (2015). "Chd1 is essential for the high transcriptional output and rapid growth of the mouse epiblast." Development **142**(1): 118-127.

122. Hales, B. F., et al. (2011). "Epigenetic programming: From gametes to blastocyst." Birth Defects Research Part A: Clinical and Molecular Teratology **91**(8): 652-665.

123. Hall, B., et al. (2009). "Overview: Generation of Gene Knockout Mice." Current protocols in cell biology / editorial board, Juan S. Bonifacino ... [et al.] **CHAPTER**: Unit-19.1217.

124. Hammond, S. M. (2005). "Dicing and slicing: The core machinery of

the RNA interference pathway." FEBS Letters **579**(26): 5822-5829.

125. Han, D. W., et al. (2011). "Direct reprogramming of fibroblasts into epiblast stem cells." Nat Cell Biol **13**(1): 66-71.

126. Hanai, J. i., et al. (2013). "ATP citrate lyase knockdown impacts cancer stem cells in vitro." Cell Death & Disease **4**(6): e696.

127. Hansen, J. C. and A. P. Wolffe (1992). "Influence of chromatin folding on transcription initiation and elongation by RNA polymerase III." Biochemistry **31**(34): 7977-7988.

128. He, J., et al. (2005). "Dynamic DNA Methylation and Histone Modifications Contribute to Lentiviral Transgene Silencing in Murine Embryonic Carcinoma Cells." Journal of Virology **79**(21): 13497-13508.

129. Hellen, C. U. T. and P. Sarnow (2001). "Internal ribosome entry sites in eukaryotic mRNA molecules." Genes & Development **15**(13): 1593-1612.

130. Hemberger, M., et al. (2009). "Epigenetic dynamics of stem cells and cell lineage commitment: digging Waddington's canal." Nat Rev Mol Cell Biol **10**(8): 526-537.

131. Hendriks, Ivo A., et al. (2015). "SUMO-2 Orchestrates Chromatin Modifiers in Response to DNA Damage." Cell Reports **10**(10): 1778-1791.

132. Heng, J., et al. (2010). "The nuclear receptor Nr5a2 can replace Oct4 in the reprogramming of murine somatic cells to pluripotent cells." Cell Stem Cell **6**: 167 - 174.

133. Herr, W. and M. Cleary (1995). "The POU domain: versatility in transcriptional regulation by a flexible two-in-one DNA-binding domain." Genes Dev **9**: 1679 - 1693.

134. Hioki, H., et al. (2009). "High-level transgene expression in neurons by lentivirus with Tet-Off system." Neuroscience Research **63**(2): 149-154.

135. Hirano, T. (2015). "Chromosome Dynamics during Mitosis." Cold Spring Harbor Perspectives in Biology.

136. Hirasaki, M., et al. (2013). "Striking Similarity in the Gene Expression Levels of Individual Myc Module Members among ESCs, EpiSCs, and Partial iPSCs." PLoS ONE **8**(12): e83769.

137. Hobbs, S., et al. (1998). "Development of a Bicistronic Vector Driven by the Human Polypeptide Chain Elongation Factor 1α Promoter for Creation of Stable Mammalian Cell Lines That Express Very High Levels of Recombinant Proteins." Biochemical and Biophysical Research Communications **252**(2): 368-372.

138. Hock, R., et al. (1998). "Chromosomal Proteins HMG-14 and HMG-17 Are Released from Mitotic Chromosomes and Imported into the Nucleus by Active Transport." The Journal of Cell Biology **143**(6): 1427-1436.

139. Hock, R., et al. (1998). "Dynamic relocation of chromosomal protein HMG-17 in the nucleus is dependent on transcriptional activity." The EMBO Journal **17**(23): 6992-7001.

140. Hock, R., et al. (2007). "HMG chromosomal proteins in development and disease." Trends in Cell Biology **17**(2): 72-79.

141. Hojfeldt, J. W., et al. (2013). "Histone lysine demethylases as targets for anticancer therapy." Nat Rev Drug Discov **12**(12): 917-930.

142. Hu, G. and J. Luo (2012). "A primer on using pooled shRNA libraries for functional genomic screens." Acta Biochimica et Biophysica Sinica **44**(2): 103-112.

143. Hu, G. and K. Zhao (2015). Chapter 10 - Identification of intergenic long noncoding RNA by deep sequencing A2 - Blakey, Suming HuangMichael D. LittC. Ann. Epigenetic Gene Expression and Regulation. Oxford, Academic Press**:** 223-235.

144. Hu, X., et al. (2015). Chapter 17 - Structure, regulation, and function of TET family proteins A2 - Blakey, Suming HuangMichael D. LittC. Ann. Epigenetic Gene Expression and Regulation. Oxford, Academic Press**:** 379-395.

145. Ibrahimi, A., et al. (2009). "Highly Efficient Multicistronic Lentiviral Vectors with Peptide 2A Sequences." Human Gene Therapy **20**(8): 845-860.

146. Ipsaro, J. J. and L. Joshua-Tor (2015). "From guide to target: molecular insights into eukaryotic RNA-interference machinery." Nat Struct Mol Biol **22**(1): 20-28.

147. Ivanova, N., et al. (2006). "Dissecting self-renewal in stem cells with RNA interference." Nature **442**(7102): 533-538.

148. Jenuwein, T. and C. D. Allis (2001). "Translating the Histone Code." Science **293**(5532): 1074-1080.

149. Jeter, C. R., et al. (2009). "Functional Evidence that the Self-Renewal Gene NANOG Regulates Human Tumor Development." Stem Cells (Dayton, Ohio) **27**(5): 993-1005.

150. Ji, X., et al. (2015). "Chromatin proteomic profiling reveals novel proteins associated with histone-marked genomic regions." Proceedings of the National Academy of Sciences of the United States of America **112**(12): 3841-3846.

151. Jonesvilleneuve, E. M. V., et al. (1982). "Retinoic Acid Induces Embryonal Carcinoma-Cells to Differentiate into Neurons and Glial-Cells."

Journal of Cell Biology **94**(2): 253-262.

152. Kabe, Y., et al. (2005). "NF-Y Is Essential for the Recruitment of RNA Polymerase II and Inducible Transcription of Several CCAAT Box-Containing Genes." Molecular and Cellular Biology **25**(1): 512-522.

153. Kalkan, T. and A. Smith (2014). Mapping the route from naive pluripotency to lineage specification.

154. Kankia, B. I. and K. Musier-Forsyth (2007). "The HIV-1 central DNA flap region contains a "flapping" third strand." Biophysical Chemistry **127**(1–2): 64-68.

155. Karwacki-Neisius, V., et al. (2013). "Reduced Oct4 Expression Directs a Robust Pluripotent State with Distinct Signaling Activity and Increased Enhancer Occupancy by Oct4 and Nanog." Cell Stem Cell **12**(5): 531-545.

156. Kashyap, V., et al. (2009). "Regulation of stem cell pluripotency and differentiation involves a mutual regulatory circuit of the NANOG, OCT4, and SOX2 pluripotency transcription factors with polycomb repressive complexes and stem cell microRNAs." Stem Cells Dev **18**: 1093 - 1108.

157. Kato, H., et al. (2011). "Architecture of the high mobility group nucleosomal protein 2-nucleosome complex as revealed by methyl-based NMR." Proceedings of the National Academy of Sciences of the United States of America **108**(30): 12283-12288.

158. Kim, J., et al. (2008). "An extended transcriptional network for pluripotency of embryonic stem cells." Cell **132**: 1049 - 1061.

159. Kim, J., et al. (2009). "Direct reprogramming of human neural stem cells by OCT4." Nature **461**: 649 - 653.

160. Kim, J., et al. (2009). "Oct4-induced pluripotency in adult neural stem cells." Cell **136**: 411 - 419.

161. Klages, N., et al. (2000). "A Stable System for the High-Titer Production of Multiply Attenuated Lentiviral Vectors." Molecular Therapy **2**(2): 170-176.

162. Körner, U., et al. (2003). "Developmental role of HMGN proteins in Xenopus laevis." Mechanisms of Development **120**(10): 1177-1192.

163. Kugler, J. E., et al. (2012). "The HMGN Family of Chromatin-Binding Proteins: Dynamic Modulators of Epigenetic Processes." Biochimica et Biophysica Acta **1819**(7): 652-656.

164. Kugler, J. E., et al. (2012). "The HMGN family of chromatin-binding proteins: Dynamic modulators of epigenetic processes." Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms **1819**(7): 652-656.

165. Kugler, J. E., et al. (2013). "High Mobility Group N Proteins Modulate the Fidelity of the Cellular Transcriptional Profile in a Tissue- and Variant-specific Manner." Journal of Biological Chemistry **288**(23): 16690-16703.

166. Kulkeaw, K., et al. (2012). "Ectopic expression of Hmgn2 antagonizes mouse erythroid differentiation in vitro." Cell Biology International **36**(2): 195-202.

167. Kuroda, T., et al. (2005). "Octamer and Sox elements are required for transcriptional cis regulation of Nanog gene expression." Mol Cell Biol **25**: 2475 - 2485.

168. Lane, A. A., et al. (2014). "Triplication of a 21q22 region contributes to B cell transformation through HMGN1 overexpression and loss of histone H3 lysine 27 trimethylation." Nature genetics **46**(6): 618-623.

169. Lanza, A. M., et al. (2013). "Evaluating the influence of selection markers on obtaining selected pools and stable cell lines in human cells." Biotechnology Journal **8**(7): 811-821.

170. Laribee, R. N., et al. (2007). "H2B ubiquitylation in transcriptional control: a FACT-finding mission." Genes & Development **21**(7): 737-743.

171. Lau, P. N. I. and C. W. E. So (2015). Chapter 4 - Polycomb and Trithorax factors in transcriptional and epigenetic regulation A2 - Blakey, Suming HuangMichael D. LittC. Ann. Epigenetic Gene Expression and Regulation. Oxford, Academic Press**:** 63-94.

172. Law, E. W. L., et al. (2012). "Anti-angiogenic and tumor-suppressive roles of candidate tumor-suppressor gene, Fibulin-2, in nasopharyngeal carcinoma." Oncogene **31**(6): 728-738.

173. Lawrence, K. S., et al. (2015). "DNA Damage Response and Spindle Assembly Checkpoint Function throughout the Cell Cycle to Ensure Genomic Integrity." PLoS Genetics **11**(4): e1005150.

174. Lee, Y., et al. (2004). "MicroRNA genes are transcribed by RNA polymerase II." The EMBO Journal **23**(20): 4051-4060.

175. Lemischka, I. (2010). "Hooking up with Oct4." Cell Stem Cell **6**: 291 - 292.

176. Lengner, C., et al. (2007). "Oct4 expression is not required for mouse somatic stem cell self-renewal." Cell Stem Cell **1**: 403 - 415.

177. Leung, R. K. M. and P. A. Whittaker (2005). "RNA interference: From gene silencing to gene-specific therapeutics." Pharmacology & Therapeutics **107**(2): 222-239.

178. Li, J., et al. (2007). "Synergistic function of DNA methyltransferases Dnmt3a and Dnmt3b in the methylation of Oct4 and Nanog." Mol Cell Biol **27**: 8748 - 8759.

179. Li, L., et al. (2010). "Stk40 links the pluripotency factor Oct4 to the Erk/MAPK pathway and controls extraembryonic endoderm differentiation." Proc Natl Acad Sci USA **107**: 1402 - 1407.

180. Li, L., et al. (2015). "A specific E3 ligase/deubiquitinase pair modulates TBP protein levels during muscle differentiation." eLife **4**: e08536.

181. Li, R., et al. (2010). "A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts." Cell Stem Cell **7**: 51 - 63.

182. Li, X., et al. (2014). "A model for the molecular underpinnings of tooth defects in Axenfeld–Rieger syndrome." Human Molecular Genetics **23**(1): 194-208.

183. Li, Y. (2015). Chapter 9 - Transgenerational epigenetic regulation by environmental factors in human diseases A2 - Blakey, Suming HuangMichael D. LittC. Ann. Epigenetic Gene Expression and Regulation. Oxford, Academic Press**:** 209-222.

184. Liang, J., et al. (2008). "Nanog and Oct4 associate with unique transcriptional repression complexes in embryonic stem cells." Nat Cell Biol **10**: 731 - 739.

185. Liao, B. and Y. Jin (2010). "Wwp2 mediates Oct4 ubiquitination and its own auto-ubiquitination in a dosage-dependent manner." Cell Res **20**: 332 - 344.

186. Liu, Y. P. and B. Berkhout (2011). "miRNA cassettes in viral vectors: Problems and solutions." Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms **1809**(11–12): 732-745.

187. Liu, Y. P., et al. (2010). "Titers of lentiviral vectors encoding shRNAs and miRNAs are reduced by different mechanisms that require distinct repair strategies." RNA **16**(7): 1328-1339.

188. Loh, Y., et al. (2006). "The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells." Nat Genet **38**: 431 - 440.

189. Luger, K. (2006). "Dynamic nucleosomes." Chromosome Research **14**(1): 5-16.

190. Luger, K. and J. C. Hansen (2005). "Nucleosome and chromatin fiber dynamics." Current Opinion in Structural Biology **15**(2): 188-196.

191. Luna, S. d. l., et al. (1988). "Efficient transformation of mammalian cells with constructs containing a puromycin-resistance marker." Gene **62**(1): 121-126.

192. Lunyak, V. V. and M. G. Rosenfeld (2008). "Epigenetic regulation of stem cell fate." Human Molecular Genetics **17**(R1): R28-R36.

193. Lyu, J., et al. (2003). "Ectopic Expression of Axin Blocks Neuronal Differentiation of Embryonic Carcinoma P19 Cells." Journal of Biological Chemistry **278**(15): 13487-13495.

194. Ma, Y., et al. (2003). "High-Level Sustained Transgene Expression in Human Embryonic Stem Cells Using Lentiviral Vectors." STEM CELLS **21**(1): 111-117.

195. Manjunath, N., et al. (2009). "Lentiviral delivery of short hairpin RNAs." Adv Drug Deliv Rev **61**(9): 732-745.

196. Marikawa, Y., et al. (2009). "Aggregated P19 Mouse Embryonal Carcinoma Cells as a Simple In Vitro Model to Study the Molecular Regulations of Mesoderm Formation and Axial Elongation Morphogenesis." Genesis (New York, N.Y. : 2000) **47**(2): 93-106.

197. Mariño-Ramírez, L., et al. (2005). "Histone structure and nucleosome stability." Expert Review of Proteomics **2**(5): 719-729.

198. Marmorstein, R. and M.-M. Zhou (2014). "Writers and Readers of Histone Acetylation: Structure, Mechanism, and Inhibition." Cold Spring Harbor Perspectives in Biology **6**(7).

199. Martínez de Paz, A. and J. Ausió (2016). "HMGNs: The enhancer charmers." BioEssays **38**(3): 226-231.

200. Martínez-Salas, E. (1999). "Internal ribosome entry site biology and its use in expression vectors." Current Opinion in Biotechnology **10**(5): 458-464.

201. Matoba, R., et al. (2006). "Dissecting Oct3/4-regulated gene networks in embryonic stem cells by expression profiling." PLoS ONE **1**: e26.

202. Mattout, A. and E. Meshorer (2010). "Chromatin plasticity and genome organization in pluripotent embryonic stem cells." Current Opinion in Cell Biology **22**(3): 334-341.

203. Maurizio, E., et al. (2016). "Translating Proteomic Into Functional Data: An High Mobility Group A1 (HMGA1) Proteomic Signature Has Prognostic Value in Breast Cancer." Molecular & Cellular Proteomics **15**(1): 109-123.

204. Meerbrey, K. L., et al. (2011). "The pINDUCER lentiviral toolkit for inducible RNA interference in vitro and in vivo." Proceedings of the

National Academy of Sciences of the United States of America **108**(9): 3665-3670.

205. Meissner, A. (2010). "Epigenetic modifications in pluripotent and differentiated cells." Nat Biotech **28**(10): 1079-1088.

206. Melcer, S., et al. (2012). "Histone modifications and lamin A regulate chromatin protein dynamics in early embryonic stem cell differentiation." Nat Commun **3**: 910.

207. Merrill, B. J. (2012). "Wnt Pathway Regulation of Embryonic Stem Cell Self-Renewal." Cold Spring Harbor Perspectives in Biology **4**(9): a007971.

208. Meshorer, E. and T. Misteli (2006). "Chromatin in pluripotent embryonic stem cells and differentiation." Nat Rev Mol Cell Biol **7**(7): 540-546.

209. Meshorer, E., et al. (2006). "Hyperdynamic Plasticity of Chromatin Proteins in Pluripotent Embryonic Stem Cells." Developmental cell **10**(1): 105-116.

210. Mills, B. B., et al. (2015). Chapter 8 - Epigenetic inheritance A2 - Blakey, Suming HuangMichael D. LittC. Ann. Epigenetic Gene Expression and Regulation. Oxford, Academic Press**:** 183-208.

211. Minarovits, J., et al. (2016). "Epigenetic Regulation." Advances in experimental medicine and biology **879**: 1-25.

212. Mitsui, K., et al. (2003). "The Homeoprotein Nanog Is Required for Maintenance of Pluripotency in Mouse Epiblast and ES Cells." Cell **113**(5): 631-642.

213. Miyoshi, H., et al. (1998). "Development of a Self-Inactivating Lentivirus Vector." Journal of Virology **72**(10): 8150-8157.

214. Mizuguchi, H. (2000). "IRES-Dependent Second Gene Expression Is Significantly Lower than Cap-Dependent First Gene Expression in a Bicistronic Vector." Mol. Ther. **1**: 376-382.

215. Mohamed, O. A., et al. (2001). "High-Mobility Group Proteins 14 and 17 Maintain the Timing of Early Embryonic Development in the Mouse." Developmental biology **229**(1): 237-249.

216. Mohan, G. (2012). Chromatin-binding HMGN proteins and the neuronal differentiation of enbryonal carcinoma cells in vitro, University of Glasgow.

217. Monneau, Y. R., et al. (2013). "Structure and Activity of the Peptidyl-Prolyl Isomerase Domain from the Histone Chaperone Fpr4 toward Histone H3 Proline Isomerization." The Journal of Biological Chemistry **288**(36): 25826-25837.

218. Montellier, E., et al. (2012). "Histone crotonylation specifically

marks the haploid male germ cell gene expression program." BioEssays **34**(3): 187-193.

219. Monzo, H. J., et al. (2012). "A method for generating high-yield enriched neuronal cultures from P19 embryonal carcinoma cells." Journal of Neuroscience Methods **204**(1): 87-103.

220. Mountford, P. S. and A. G. Smith (1995). "Internal ribosome entry sites and dicistronic RNAs in mammalian transgenesis." Trends in Genetics **11**(5): 179-184.

221. Nakatake, Y., et al. (2006). "Klf4 Cooperates with Oct3/4 and Sox2 To Activate the Lefty1 Core Promoter in Embryonic Stem Cells." Molecular and Cellular Biology **26**(20): 7772-7782.

222. Nakayama, Y., et al. (2014). "A rapid and efficient method for neuronal induction of the P19 embryonic carcinoma cell line." Journal of Neuroscience Methods **227**: 100-106.

223. Naughton, C., et al. (2013). "Transcription forms and remodels supercoiling domains unfolding large-scale chromatin structures." Nature structural & molecular biology **20**(3): 387-395.

224. Nelson, C. J., et al. (2006). "Proline Isomerization of Histone H3 Regulates Lysine Methylation and Gene Expression." Cell **126**(5): 905-916.

225. Nichols, J. and A. Smith (2009). "Naive and Primed Pluripotent States." Cell Stem Cell **4**(6): 487-492.

226. Nichols, J. and A. Smith (2012). "Pluripotency in the Embryo and in Culture." Cold Spring Harbor Perspectives in Biology **4**(8): a008128.

227. Nichols, J., et al. (1998). "Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4." Cell **95**: 379 - 391.

228. Nishimoto, M., et al. (1999). "The gene for the embryonic stem cell coactivator UTF1 carries a regulatory element which selectively interacts with a complex composed of Oct-3/4 and Sox-2." Mol Cell Biol **19**: 5453 - 5465.

229. Nishino, Y., et al. (2012). "Human mitotic chromosomes consist predominantly of irregularly folded nucleosome fibres without a 30-nm chromatin structure." The EMBO Journal **31**(7): 1644-1653.

230. Niwa, H. (2007). "How is pluripotency determined and maintained?" Development **134**(4): 635-646.

231. Niwa, H. (2014). Chapter 7 - Mechanisms of Stem Cell Self-Renewal. Essentials of Stem Cell Biology (Third Edition). R. Lanza and A. Atala. Boston, Academic Press**:** 81-94.

232. Niwa, H., et al. (2000). "Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells." Nat Genet **24**: 372 - 376.

233. Niwa, H., et al. (2005). "Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation." Cell **123**: 917 - 929.

234. Noh, Y. H., et al. (2013). "Recovery from Parkinsonism with N-acetylcysteine-differentiated neurons." Molecular Biology **47**(4): 538-543.

235. Nordhoff, V., et al. (2001). "Comparative analysis of human, bovine, and murine Oct-4 upstream promoter sequences." Mamm Genome **12**: 309 - 317.

236. Nowotny, M. and W. Yang (2009). "Structural and functional modules in RNA interference." Current Opinion in Structural Biology **19**(3): 286-293.

237. Nye, J. S., et al. (1994). "An activated Notch suppresses neurogenesis and myogenesis but not gliogenesis in mammalian cells." Development **120**(9): 2421-2430.

238. Ohtsuka, S. and S. Dalton (2007). "Molecular and biological properties of pluripotent embryonic stem cells." Gene therapy **15**(2): 74-81.

239. Ohtsuka, S. and S. Dalton (2007). "Molecular and biological properties of pluripotent embryonic stem cells." Gene therapy **15**(2): 74-81.

240. Okamoto, K., et al. (1990). "A novel octamer binding transcription factor is differentially expressed in mouse embryonic cells." Cell **60**: 461 - 472.

241. Ozturk, N., et al. (2014). "HMGA proteins as modulators of chromatin structure during transcriptional activation." Frontiers in Cell and Developmental Biology **2**.

242. Paddison, P. J., et al. (2002). "Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells." Genes & Development **16**(8): 948-958.

243. Pan, G., et al. (2002). "Stem cell pluripotency and transcription factor Oct4." Cell Res **12**: 321 - 329.

244. Pan, H., et al. (2008). "Dual-promoter lentiviral system allows inducible expression of noxious proteins in macrophages." Journal of Immunological Methods **329**(1–2): 31-44.

245. Paranjape, S. M., et al. (1995). "HMG17 is a chromatin-specific transcriptional coactivator that increases the efficiency of transcription initiation." Genes & Development **9**(16): 1978-1991.

246. Pardo, M., et al. (2010). "An expanded Oct4 interaction network: implications for stem cell biology, development, and disease." Cell Stem Cell **6**: 382 - 395.

247. Park, S., et al. (2007). "SUMOylation of Tr2 orphan receptor involves Pml and fine-tunes Oct4 expression in stem cells." Nat Struct Mol Biol **14**: 68 - 75.

248. Patra, S. K., et al. (2011). "Molecular marks for epigenetic identification of developmental and cancer stem cells." Clinical Epigenetics **2**(1): 27-53.

249. Pesce, M. and H. Scholer (2001). "Oct-4: gatekeeper in the beginnings of mammalian development." STEM CELLS **19**: 271 - 278.

250. Petersen, G. F., et al. (2014). "Efficient transduction of equine adipose-derived mesenchymal stem cells by VSV-G pseudotyped lentiviral vectors." Research in Veterinary Science **97**(3): 616-622.

251. Pfeifer, A., et al. (2002). "Transgenesis by lentiviral vectors: Lack of gene silencing in mammalian embryonic stem cells and preimplantation embryos." Proceedings of the National Academy of Sciences **99**(4): 2140-2145.

252. Pink, R. C., et al. (2011). "Pseudogenes: Pseudo-functional or key regulators in health and disease?" RNA **17**(5): 792-798.

253. Plusa, B. and A.-K. Hadjantonakis (2014). "Embryonic stem cell identity grounded in the embryo." Nat Cell Biol **16**(6): 502-504.

254. Pogna, E. A., et al. (2010). "Signalling to chromatin through post-translational modifications of HMGN." Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms **1799**(1–2): 93-100.

255. Postnikov, Y. and M. Bustin (2010). "Regulation of chromatin structure and function By HMGN proteins." Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms **1799**(1–2): 62-68.

256. Postnikov, Y. V. and M. Bustin "Functional interplay between histone H1 and HMG proteins in chromatin." Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms.

257. Postnikov, Y. V., et al. (1994). "The cooperative binding of chromosomal protein HMG-14 to nucleosome cores is reduced by single point mutations in the nucleosomal binding domain." Nucleic Acids Research **22**(21): 4520-4526.

258. Postnikov, Y. V., et al. (1995). "Homodimers of Chromosomal Proteins HMG-14 and HMG-17 in Nucleosome Cores." Journal of Molecular Biology **252**(4): 423-432.

259. Postnikov, Y. V., et al. (1997). "Clusters of nucleosomes containing

chromosomal protein HMG-17 in chromatin1." Journal of Molecular Biology **274**(4): 454-465.

260. Postnikov, Y. V., et al. (2014). "Loss of the nucleosome-binding protein HMGN1 affects the rate of N-nitrosodiethylamine induced hepatocarcinogenesis in mice." Molecular cancer research : MCR **12**(1): 82-90.

261. Prelle, K., et al. (2002). "Pluripotent Stem Cells – Model of Embryonic Development, Tool for Gene Targeting, and Basis of Cell Therapy." Anatomia, Histologia, Embryologia **31**(3): 169-186.

262. Pugh, B. F. (2010). "A preoccupied position on nucleosomes." Nat Struct Mol Biol **17**(8): 923-923.

263. Qin, D.-N., et al. (2013). "Effects of miR-19b Overexpression on Proliferation, Differentiation, Apoptosis and Wnt/β-Catenin Signaling Pathway in P19 Cell Model of Cardiac Differentiation In Vitro." Cell Biochemistry and Biophysics **66**(3): 709-722.

264. Qiu, C., et al. (2010). "Lin28-mediated post-transcriptional regulation of Oct4 expression in human embryonic stem cells." Nucleic Acids Res **38**: 1240 - 1248.

265. Quina, A. S., et al. (2006). "Chromatin structure and epigenetics." Biochemical Pharmacology **72**(11): 1563-1569.

266. Radman-Livaja, M. and O. J. Rando (2010). "Nucleosome positioning: how is it established, and why does it matter?" Developmental biology **339**(2): 258-266.

267. Ramirez, T., et al. (2008). "Sodium arsenite modulates histone acetylation, histone deacetylase activity and HMGN protein dynamics in human cells." Chromosoma **117**(2): 147-157.

268. Rana, T. M. (2007). "Illuminating the silence: understanding the structure and function of small RNAs." Nat Rev Mol Cell Biol **8**(1): 23-36.

269. Rao, D. D., et al. (2009). "siRNA vs. shRNA: Similarities and differences." Advanced drug delivery reviews **61**(9): 746-759.

270. Rathert, P., et al. (2015). "Transcriptional plasticity promotes primary and acquired resistance to BET inhibition." Nature **525**(7570): 543-547.

271. Recillas-Targa, F. (2015). Chapter 13 - Long noncoding RNAs and carcinogenesis A2 - Blakey, Suming HuangMichael D. LittC. Ann. Epigenetic Gene Expression and Regulation. Oxford, Academic Press**:** 291-312.

272. Reeves, R. (2015). "High mobility group (HMG) proteins: Modulators of chromatin structure and DNA repair in mammalian cells." DNA Repair

36: 122-136.

273. Ricci, Maria A., et al. (2015). "Chromatin Fibers Are Formed by Heterogeneous Groups of Nucleosomes In Vivo." Cell **160**(6): 1145-1158.

274. Rochman, M., et al. (2011). "Effects of HMGN variants on the cellular transcription profile." Nucleic Acids Research **39**(10): 4076-4087.

275. Rodda, D., et al. (2005). "Transcriptional regulation of nanog by OCT4 and SOX2." J Biol Chem **280**: 24731 - 24737.

276. Rosner, M., et al. (1990). "A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo." Nature **345**: 686 - 692.

277. Rossetto, D., et al. (2012). "Histone phosphorylation: A chromatin modification involved in diverse nuclear events." Epigenetics **7**(10): 1098-1108.

278. Rugg-Gunn, Peter J., et al. (2012). "Cell-Surface Proteomics Identifies Lineage-Specific Markers of Embryo-Derived Stem Cells." Developmental cell **22**(4): 887-901.

279. Rust, M. J., et al. (2006). "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)." Nat Meth **3**(10): 793-796.

280. Rychahou, P. G., et al. (2006). "RNA interference: Mechanisms of action and therapeutic consideration." Surgery **140**(5): 719-725.

281. Rytlewski, J. A. and S. Beronja (2015). "RNAi in the mouse: rapid and affordable gene function studies in a vertebrate system." Wiley Interdisciplinary Reviews: Developmental Biology **4**(1): 45-57.

282. Sabari, Benjamin R., et al. (2015). "Intracellular Crotonyl-CoA Stimulates Transcription through p300-Catalyzed Histone Crotonylation." Molecular Cell **58**(2): 203-215.

283. Saitou, M., et al. (2012). "Epigenetic reprogramming in mouse pre-implantation development and primordial germ cells." Development **139**(1): 15-31.

284. Saiz, N. and B. Plusa (2013). "Early cell fate decisions in the mouse embryo." Reproduction **145**(3): R65-R80.

285. Saladi, S. V. and I. L. de la Serna (2010). "ATP Dependent Chromatin Remodeling Enzymes in Embryonic Stem Cells." Stem cell reviews **6**(1): 62-73.

286. Sarkar, K., et al. (2015). "SUMOylation-disrupting WAS mutation converts WASp from a transcriptional activator to a repressor of NF-κB response genes in T cells." Blood **126**(14): 1670-1682.

287. Scholer, H., et al. (1990). "New type of POU domain in germ line-specific protein Oct-4." Nature **344**: 435 - 439.

288. Schones, D. E., et al. (2008). "Dynamic Regulation of Nucleosome Positioning in the Human Genome." Cell **132**(5): 887-898.

289. Schorpp, M., et al. (1996). "The Human Ubiquitin C Promoter Directs High Ubiquitous Expression of Transgenes in Mice." Nucleic Acids Research **24**(9): 1787-1788.

290. Sgarra, R., et al. (2004). "Nuclear phosphoproteins HMGA and their relationship with chromatin structure and cancer." FEBS Letters **574**(1–3): 1-8.

291. Shakiba, N., et al. (2015). "CD24 tracks divergent pluripotent states in mouse and human cells." Nat Commun **6**.

292. Shalem, O., et al. (2015). "High-throughput functional genomics using CRISPR-Cas9." Nat Rev Genet **16**(5): 299-311.

293. Sharov, A., et al. (2008). "Identification of Pou5f1, Sox2, and Nanog downstream target genes with statistical confidence by applying a novel algorithm to time course microarray and genome-wide chromatin immunoprecipitation data." BMC Genomics **9**: 269.

294. Shchuka, V., et al. (2015). "Chromatin Dynamics in Lineage Commitment and Cellular Reprogramming." Genes **6**(3): 641.

295. Shearer, R. F. and D. N. Saunders (2015). "Experimental design for stable genetic manipulation in mammalian cell lines: lentivirus and alternatives." Genes to Cells **20**(1): 1-10.

296. Sheik Mohamed, J., et al. (2010). "Conserved long noncoding RNAs transcriptionally regulated by Oct4 and Nanog modulate pluripotency in mouse embryonic stem cells." RNA **16**: 324 - 337.

297. Shen, S. and P. Casaccia-Bonnefil (2008). "Post-Translational Modifications of Nucleosomal Histones in Oligodendrocyte Lineage Cells in Development and Disease." Journal of molecular neuroscience : MN **35**(1): 13-22.

298. Shi, G. and Y. Jin (2010). "Role of Oct4 in maintaining and regaining stem cell pluripotency." Stem Cell Research & Therapy **1**(5): 39.

299. Shi, L. and J. Wu (2009). "Epigenetic regulation in mammalian preimplantation embryo development." Reproductive Biology and Endocrinology : RB&E **7**: 59-59.

300. Shi, W., et al. (2006). "Regulation of the Pluripotency Marker Rex-1 by Nanog and Sox2." Journal of Biological Chemistry **281**(33): 23319-

23325.

301. Shi, Y., et al. (2008). "A combined chemical and genetic approach for the generation of induced pluripotent stem cells." Cell Stem Cell **2**: 525 - 528.

302. Shin, K.-J., et al. (2006). "A single lentiviral vector platform for microRNA-based conditional RNA interference and coordinated transgene expression." Proceedings of the National Academy of Sciences of the United States of America **103**(37): 13759-13764.

303. Shuen, W. H., et al. (2015). "Novel lentiviral-inducible transgene expression systems and versatile single-plasmid reporters for in vitro and in vivo cancer biology studies." Cancer Gene Ther **22**(4): 207-214.

304. Singh, A. M. and S. Dalton (2009). "The Cell Cycle and Myc Intersect with Mechanisms that Regulate Pluripotency and Reprogramming." Cell Stem Cell **5**(2): 141-149.

305. Singh, B. N., et al. (2010). "Nonhistone protein acetylation as cancer therapy targets." Expert Review of Anticancer Therapy **10**(6): 935-954.

306. Sisson, T. H., et al. (2006). "Expression of the Reverse Tetracycline-Transactivator Gene Causes Emphysema-Like Changes in Mice." American Journal of Respiratory Cell and Molecular Biology **34**(5): 552-560.

307. Skerjanc, I. S. (1999). "Cardiac and Skeletal Muscle Development in P19 Embryonal Carcinoma Cells." Trends in Cardiovascular Medicine **9**(5): 139-143.

308. Sommer, C. A., et al. (2009). "Induced Pluripotent Stem Cell Generation Using a Single Lentiviral Stem Cell Cassette." STEM CELLS **27**(3): 543-549.

309. Song, H. and P.-C. Yang (2010). "Construction of shRNA lentiviral vector." North American Journal of Medical Sciences **2**(12): 598-601.

310. Song, M. and A. Ghosh (2004). "FGF2-induced chromatin remodeling regulates CNTF-mediated gene expression and astrocyte differentiation." Nature neuroscience **7**(3): 229.

311. Soufi, A., et al. (2012). "Facilitators and Impediments of the Pluripotency Reprogramming Factors' Initial Engagement with the Genome." Cell **151**(5): 994-1004.

312. Spitzer, J., et al. (2013). Chapter Eight - Rapid Creation of Stable Mammalian Cell Lines for Regulated Expression of Proteins Using the Gateway® Recombination Cloning Technology and Flp-In T-REx® Lines. Methods in Enzymology. L. Jon, Academic Press. **Volume 529:** 99-124.

313. Sprengel, R. and M. T. Hasan (2007). Tetracycline-Controlled

Genetic Switches. Conditional Mutagenesis: An Approach to Disease Models. R. Feil and D. Metzger, Springer Berlin Heidelberg. **178:** 49-72.

314. Sridharan, R., et al. (2009). "Role of the murine reprogramming factors in the induction of pluripotency." Cell **136**: 364 - 377.

315. Stegmeier, F., et al. (2005). "A lentiviral microRNA-based system for single-copy polymerase II-regulated RNA interference in mammalian cells." Proceedings of the National Academy of Sciences of the United States of America **102**(37): 13212-13217.

316. Stewart, S. A., et al. (2003). "Lentivirus-delivered stable gene silencing by RNAi in primary cells." RNA **9**(4): 493-501.

317. Štros, M. (2010). "HMGB proteins: Interactions with DNA and chromatin." Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms **1799**(1–2): 101-113.

318. Subramanian, M., et al. (2009). "The nucleosome-binding protein HMGN2 modulates global genome repair." FEBS Journal **276**(22): 6646-6657.

319. Suh, L. H., et al. (2007). "Cryopreservation and Lentiviral-Mediated Genetic Modification of Human Primary Cultured Corneal Endothelial Cells." Investigative Ophthalmology & Visual Science **48**(7): 3056-3061.

320. Sutani, T., et al. (2015). "Condensin targets and reduces unwound DNA structures associated with transcription in mitotic chromosome condensation." Nat Commun **6**.

321. Szutorisz, H., et al. (2005). "Formation of an Active Tissue-Specific Chromatin Domain Initiated by Epigenetic Marking at the Embryonic Stem Cell Stage." Molecular and Cellular Biology **25**(5): 1804-1820.

322. Szymczak, A. L., et al. (2004). "Correction of multi-gene deficiency in vivo using a single 'self-cleaving' 2A peptide-based retroviral vector." Nat Biotech **22**(5): 589-594.

323. Tagami, H., et al. (2004). "Histone H3.1 and H3.3 Complexes Mediate Nucleosome Assembly Pathways Dependent or Independent of DNA Synthesis." Cell **116**(1): 51-61.

324. Takahashi, K. and S. Yamanaka (2006). "Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors." Cell **126**: 663 - 676.

325. Tang, M., et al. (2015). Chapter 12 - Genetically altered cancer epigenome A2 - Blakey, Suming HuangMichael D. LittC. Ann. Epigenetic Gene Expression and Regulation. Oxford, Academic Press**:** 265-289.

326. Tang, Y., et al. (2015). "Optimization of lentiviral vector production using polyethylenimine-mediated transfection." Oncology

Letters **9**(1): 55-62.

327. Tarakhovsky, A. (2010). "Tools and landscapes of epigenetics." Nat Immunol **11**(7): 565-568.

328. Tay, Y., et al. (2008). "MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation." Nature **455**: 1124 - 1128.

329. Teif, V. B., et al. (2012). "Genome-wide nucleosome positioning during embryonic stem cell development." Nat Struct Mol Biol **19**(11): 1185-1192.

330. Tesar, P. J., et al. (2007). "New cell lines from mouse epiblast share defining features with human embryonic stem cells." Nature **448**(7150): 196-199.

331. Tirosh, I. and N. Barkai (2008). "Two strategies for gene regulation by promoter nucleosomes." Genome Research **18**(7): 1084-1091.

332. Tiscornia, G., et al. (2003). "A general method for gene knockdown in mice by using lentiviral vectors expressing small interfering RNA." Proceedings of the National Academy of Sciences of the United States of America **100**(4): 1844-1848.

333. Tiscornia, G., et al. (2006). "Design and cloning of lentiviral vectors expressing small interfering RNAs." Nat. Protocols **1**(1): 234-240.

334. Tiscornia, G., et al. (2007). "3 Development of Lentiviral Vectors Expressing siRNA."

335. Tollefsbol, T. (2011). Handbook of epigenetics: the new molecular and medical genetic, elsevierdirect.

336. Trieschmann, L., et al. (1995). "Incorporation of chromosomal proteins HMG-14/HMG-17 into nascent nucleosomes induces an extended chromatin conformation and enhances the utilization of active transcription complexes." The EMBO Journal **14**(7): 1478-1489.

337. Trieschmann, L., et al. (1995). "Modular structure of chromosomal proteins HMG-14 and HMG-17: definition of a transcriptional enhancement domain distinct from the nucleosomal binding domain." Molecular and Cellular Biology **15**(12): 6663-6669.

338. Turner, B. M. (2005). "Reading signals on the nucleosome with a new nomenclature for modified histones." Nat Struct Mol Biol **12**(2): 110-112.

339. Ueda, T. and M. Yoshida (2010). "HMGB proteins and transcriptional regulation." Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms **1799**(1–2): 114-118.

340. Ueda, T., et al. (2006). "Distinct Domains in High Mobility Group N Variants Modulate Specific Chromatin Modifications." Journal of Biological Chemistry **281**(15): 10182-10187.

341. Vallée, M., et al. (2009). "Revealing the bovine embryo transcript profiles during early in vivo embryonic development." Reproduction **138**(1): 95-105.

342. van den Berg, D. L. C., et al. (2010). "An Oct4-Centered Protein Interaction Network in Embryonic Stem Cells." Cell Stem Cell **6**(4): 369-381.

343. van den Berg, D., et al. (2008). "Estrogen-related receptor beta interacts with Oct4 to positively regulate Nanog gene expression." Mol Cell Biol **28**: 5986 - 5995.

344. van der Heyden, M. A. G. and L. H. K. Defize (2003). "Twenty one years of P19 cells: what an embryonal carcinoma cell line taught us about cardiomyocyte differentiation." Cardiovascular Research **58**(2): 292-302.

345. Vara, J., et al. (1985). "Cloning and expression of a puromycin N-acetyl transferase gene from Streptomyces alboniger in Streptomyces lividans and Escherichia coli." Gene **33**(2): 197-206.

346. Varlakhanova, N. V., et al. (2010). "myc maintains embryonic stem cell pluripotency and self-renewal." Differentiation; research in biological diversity **80**(1): 9-19.

347. Veillard, A.-C., et al. (2014). "Stable methylation at promoters distinguishes epiblast stem cells from embryonic stem cells and the in vivo epiblasts." Stem cells and development **23**(17): 2014-2029.

348. Verdin, E. and M. Ott (2015). "50 years of protein acetylation: from gene regulation to epigenetics, metabolism and beyond." Nat Rev Mol Cell Biol **16**(4): 258-264.

349. Vestner, B., et al. (1998). "Stimulation of Replication Efficiency of a Chromatin Template by Chromosomal Protein HMG-17." Journal of Biological Chemistry **273**(16): 9409-9414.

350. Voss, T. C. and G. L. Hager (2014). "Dynamic regulation of transcriptional states by chromatin and transcription factors." Nat Rev Genet **15**(2): 69-81.

351. Waddington, C. (1957). New York: Macmillan.

352. Wadosky, K. M. and M. S. Willis (2012). "The story so far: post-translational regulation of peroxisome proliferator-activated receptors by ubiquitination and SUMOylation." American Journal of Physiology - Heart and Circulatory Physiology **302**(3): H515-H526.

353. Wahlers, A., et al. (2001). "Influence of multiplicity of infection and protein stability on retroviral vector-mediated gene expression in hematopoietic cells." Gene therapy **8**(6): 477-486.

354. Walhout, A. J. M., et al. (2000). [34] GATEWAY recombinational cloning: Application to the cloning of large numbers of open reading frames or ORFeomes. Methods in Enzymology. S. D. E. Jeremy Thorner and N. A. John, Academic Press. **Volume 328:** 575-IN577.

355. Wang, J., et al. (2006). "A protein interaction network for pluripotency of embryonic stem cells." Nature **444**: 364 - 368.

356. Wang, X. and J. Dai (2010). "Concise review: isoforms of OCT4 contribute to the confusing diversity in stem cell biology." STEM CELLS **28**: 885 - 893.

357. Wang, Y.-C., et al. (2014). "Protein post-translational modifications and regulation of pluripotency in human stem cells." Cell Research **24**(2): 143-160.

358. Wang, Z. and D. J. Patel (2011). "Combinatorial Readout of Dual Histone Modifications by Paired Chromatin-associated Modules." Journal of Biological Chemistry **286**(21): 18363-18368.

359. Wang, Z., et al. (2007). "Oct4 and Sox2 directly regulate expression of another pluripotency transcription factor, Zfp206, in embryonic stem cells." J Biol Chem **282**: 12822 - 12830.

360. Warnock, J., et al. (2011). Introduction to Viral Vectors. Viral Vectors for Gene Therapy. O.-W. Merten and M. Al-Rubeai, Humana Press. **737:** 1-25.

361. Webster, D., et al. (2009). "O-GlcNAc modifications regulate cell survival and epiboly during zebrafish development." BMC Dev Biol **9**: 28.

362. Wei, F., et al. (2007). "Sumoylation of Oct4 enhances its stability, DNA binding, and transactivation." J Biol Chem **282**: 21551 - 21560.

363. Wei, F., et al. (2014). "The Alarmin HMGN1 Contributes to Antitumor Immunity and Is a Potent Immunoadjuvant." Cancer Research **74**(21): 5989-5998.

364. Weigmann, N., et al. (1997). "Enhancement of the transcription potential of nascent chromatin by chromosomal proteins HMG-14/-17 is coupled to nucleosome assembly and not DNA synthesis." DNA and cell biology **16**(10): 1207-1216.

365. Welling, M. and N. Geijsen (2013). "Uncovering the true identity of naïve pluripotent stem cells." Trends in Cell Biology **23**(9): 442-448.

366. Wernig, M., et al. (2007). "In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state." Nature **448**: 318 - 324.

367. West, K. L. (2004). "HMGN proteins play roles in DNA repair and gene expression in mammalian cells." Biochemical Society Transactions **32**(6): 918-919.

368. West, K. L., et al. (2001). "HMGN3a and HMGN3b, Two Protein Isoforms with a Tissue-specific Expression Pattern, Expand the Cellular Repertoire of Nucleosome-binding Proteins." Journal of Biological Chemistry **276**(28): 25959-25969.

369. Whyte, Warren A., et al. "Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes." Cell **153**(2): 307-319.

370. Wiederschain, D., et al. (2009). "Single-vector inducible lentiviral RNAi system for oncology target validation." Cell Cycle **8**(3): 498-504.

371. Wilson, R. C. and J. A. Doudna (2013). "Molecular Mechanisms of RNA Interference." Annual Review of Biophysics **42**(1): 217-239.

372. Wobus, A. M., et al. (1984). "Characterization of a pluripotent stem cell line derived from a mouse embryo." Experimental Cell Research **152**(1): 212-219.

373. Wolffe, A. P. and R. H. Morse (1990). "The transcription complex of the Xenopus somatic 5 S RNA gene. A functional analysis of protein-DNA interactions outside of the internal control region." Journal of Biological Chemistry **265**(8): 4592-4599.

374. Wong, F. C. K., et al. (2016). Chapter 10 - SOX2-Dependent Regulation of Pluripotent Stem Cells. Sox2. H. K. Lovell-Badge. Boston, Academic Press**:** 163-185.

375. Wong, M. D., et al. (2005). "Molecular Mechanisms of Germline Stem Cell Regulation." Annual Review of Genetics **39**(1): 173-195.

376. Woodbury, D., et al. (2002). "Adult bone marrow stromal stem cells express germline, ectodermal, endodermal, and mesodermal genes prior to neurogenesis." Journal of Neuroscience Research **69**(6): 908-917.

377. Xia, C., et al. (2007). "Induction of a high population of neural stem cells with anterior neuroectoderm characters from epiblast-like P19 embryonic carcinoma cells." Differentiation **75**(10): 912-927.

378. Xu, H., et al. (2004). "Wwp2, an E3 ubiquitin ligase that targets transcription factor Oct-4 for ubiquitination." J Biol Chem **279**: 23495 - 23503.

379. Xu, H., et al. (2009). "WWP2 promotes degradation of transcription factor OCT4 in human embryonic stem cells." Cell Res **19**: 561 - 573.

380. Xu, N., et al. (2009). "MicroRNA-145 regulates OCT4, SOX2, and

KLF4 and represses pluripotency in human embryonic stem cells." Cell **137**: 647 - 658.

381. Xu, W., et al. (2012). "Evaluation of Residual Promoter Activity in [gamma]-Retroviral Self-inactivating (SIN) Vectors." Mol Ther **20**(1): 84-90.

382. Yan, B., et al. (2015). Chapter 18 - Epigenetic drugs for cancer therapy A2 - Blakey, Suming HuangMichael D. LittC. Ann. Epigenetic Gene Expression and Regulation. Oxford, Academic Press**:** 397-423.

383. Yang, H., et al. (2007). "Transcriptional regulation of human Oct4 by steroidogenic factor-1." J Cell Biochem **101**: 1198 - 1209.

384. Yang, J., et al. (2008). "Genome-wide analysis reveals Sall4 to be a major regulator of pluripotency in murine-embryonic stem cells." Proc Natl Acad Sci USA **105**: 19756 - 19761.

385. Ye, Z., et al. (2008). Lentiviral Gene Transduction of Mouse and Human Stem Cells. Hematopoietic Stem Cell Protocols. K. Bunting, Humana Press. **430:** 243-253.

386. Yeom, Y. I., et al. (1996). "Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells." Development **122**(3): 881-894.

387. Young, N. L., et al. (2010). "The significance, development and progress of high-throughput combinatorial histone code analysis." Cellular and Molecular Life Sciences **67**(23): 3983-4000.

388. Yu, H., et al. (2009). "Zfp206, Oct4, and Sox2 are integrated components of a transcriptional regulatory network in embryonic stem cells." J Biol Chem **284**: 31327 - 31335.

389. Yu, J., et al. (2007). "Induced pluripotent stem cell lines derived from human somatic cells." Science **318**: 1917 - 1920.

390. Yuan, H., et al. (1995). "Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3." Genes & Development **9**(21): 2635-2645.

391. Yuan, P., et al. (2009). "Eset partners with Oct4 to restrict extraembryonic trophoblast lineage potential in embryonic stem cells." Genes Dev **23**: 2507 - 2520.

392. Zafarana, G., et al. (2009). "Specific knockdown of OCT4 in human embryonic stem cells by inducible short hairpin RNA interference." STEM CELLS **27**: 776 - 782.

393. Zaratiegui, M. (2015). "Molecular biology: RNA interference hangs by a thread." Nature **520**(7546): 162-164.

394. Zeineddine, D., et al. (2006). "Oct-3/4 dose dependently regulates specification of embryonic stem cells toward a cardiac lineage and early

heart development." Dev Cell **11**: 535 - 546.

395. Zeng, Y., et al. (2005). Use of RNA Polymerase II to Transcribe Artificial MicroRNAs. Methods in Enzymology. R. E. David and J. R. John, Academic Press. **Volume 392:** 371-380.

396. Zhang, B., et al. (2004). "The significance of controlled conditions in lentiviral vector titration and in the use of multiplicity of infection (MOI) for predicting gene transfer events." Genetic Vaccines and Therapy **2**: 6-6.

397. Zhang, F., et al. (2010). "A Ubiquitous Chromatin Opening Element (UCOE) Confers Resistance to DNA Methylation–mediated Silencing of Lentiviral Vectors." Molecular Therapy **18**(9): 1640-1649.

398. Zhang, J., et al. (2006). "Sall4 modulates embryonic stem cell pluripotency and early embryonic development by the transcriptional regulation of Pou5f1." Nat Cell Biol **8**: 1114 - 1123.

399. Zhang, P., et al. (2010). "Kruppel-like Factor 4 (Klf4) Prevents Embryonic Stem (ES) Cell Differentiation by Regulating Nanog Gene Expression." The Journal of Biological Chemistry **285**(12): 9180-9189.

400. Zhang, P., et al. (2010). "Kruppel-like Factor 4 (Klf4) Prevents Embryonic Stem (ES) Cell Differentiation by Regulating Nanog Gene Expression." Journal of Biological Chemistry **285**(12): 9180-9189.

401. Zhang, Q. and Y. Wang (2008). "High mobility group proteins and their post-translational modifications." Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics **1784**(9): 1159-1166.

402. Zhang, Q. and Y. Wang (2010). "HMG modifications and nuclear function." Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms **1799**(1–2): 28-36.

403. Zhang, X., et al. (2008). "Esrrb activates Oct4 transcription and sustains self-renewal and pluripotency in embryonic stem cells." J Biol Chem **283**: 35825 - 35833.

404. Zhang, Z., et al. (2007). "Post-translational modification of POU domain transcription factor Oct-4 by SUMO-1." FASEB J **21**: 3042 - 3051.

405. Zhou, B.-R., et al. (2015). "Structural Mechanisms of Nucleosome Recognition by Linker Histones." Molecular Cell **59**(4): 628-638.

406. Zhou, L. (2015). Chapter 16 - Epigenetic control of stress-induced apoptosis A2 - Blakey, Suming HuangMichael D. LittC. Ann. Epigenetic Gene Expression and Regulation. Oxford, Academic Press**:** 367-377.

407. Zhou, V. W., et al. (2011). "Charting histone modifications and the functional organization of mammalian genomes." Nat Rev Genet **12**(1): 7-18.

408. Zlatanova, J., et al. (1998). "Chromatin Fiber Structure: Morphology, Molecular Determinants, Structural Transitions." Biophysical Journal **74**(5): 2554-2566.

409. Zlatanova, J., et al. (1999). "Chromatin structure revisited." Critical reviews in eukaryotic gene expression **9**(3-4): 245-255.

410. Zufferey, R., et al. (1998). "Self-Inactivating Lentivirus Vector for Safe and Efficient In Vivo Gene Delivery." Journal of Virology **72**(12): 9873-9880.

Index