Characterisation of H3.3 mutants in terms of pluripotency

Thesis submitted in partial fulfilment of the requirements for the degree of
Master of Science in Engineering at the University of Applied Sciences
Technikum Wien - Degree Program Tissue Engineering and Regenerative Medicine

By: Shona Robinson
Student Number: 1410692032

Supervisor 1: Dr Thomas Machacek
Supervisor 2: Dr Justin Brumbaugh

Boston, January 2016-August 2016

FACHHOCHSCHULE
TECHNIKUM WIEN

Declaration

„I confirm that this thesis is entirely my own work. All sources and quotations have been fully acknowledged in the appropriate places with adequate footnotes and citations. Quotations have been properly acknowledged and marked with appropriate punctuation. The works consulted are listed in the bibliography. This paper has not been submitted to another examination panel in the same or a similar form, and has not been published. I declare that the present paper is identical to the version uploaded."

Boston, August 30th 2016

Place, Date

Signature

Abstract

A high percentage of chondroblastomas, giant cell tumour of bone, and paediatric gliomas (namely diffuse intrinsic pontine gliomas (DIPGs) and supratentorial glioblastoma multiforme (GBM) have been found to express mutations in histone variant H3.3. Mutations in the *H3F3A* or *H3F3B* genes (the two genes encoding the histone H3.3 protein in humans) have been shown to result in amino acid substitutions at critical positions within the histone tail of Lysines 27, 36 or 9 to Methionine, or Glycine 34 to Arginine or Valine[1–3]. The Lysine-to-Methionine substitutions were shown to cause a global eradication of the di- and tri-methylation of the respective H3K residues, and a simultaneous increase in the acetylation marks. Lewis *et al* demonstrate that this eradication in methylation of Lysine residues within the H3 tail after substitution to Methionine (and also Isoleucine) was caused by the mutant H3.3 peptide causing a conformational change within the PRC2 complex via interaction with the EZH2 active site, preventing the complex from disengaging with the peptide and moving to its downstream targets of methylation, resulting in a global loss of methylation of H3K residues[2].

This prevalence of these specific H3.3 Lysine residue substitutions in the cancer types previously mentioned has led to a large number of studies dissecting the mechanisms causing these mutations and the subsequent tumorigenesis.

Due to the similarities between pluripotent cells and cancer cells, including their potential for self-renewal and high proliferation rates, we hypothesised that these H3.3 mutants may have some form of effect on pluripotency and also reprogramming of somatic cells to iPSCs.

In order to investigate this hypothesis, mESCs expressing the histone H3.3 transgenes H3.3K36 to I/M and H3.3K9 to I/M were used to perform a self-renewal assay to observe the effects of these mutations in differing cell culture conditions, such as the presence or absence of LIF, on these mESCs grown at a clonal density. Alkaline phosphatase staining was used to quantify the number of pluripotent cells remaining after 6 days of culture. Embryoid body assays were then performed in order to analyse the effect of these mutations on the differentiation potential of the mESCs at a genetic level, via RT-qPCR.

It was found through AP staining that the mutations are not holding these cells in a pluripotent state, and may in fact be directing cells towards differentiation, or even apoptosis (in the case of H3.3K9M). Embryoid bodies expressing these histone H3.3 transgenes showed a phenotype of smaller EBs, which could also suggest the exit from pluripotency into differentiation.

Contents

1 Introduction 5

1.1 Histones and Disease 11

1.2 Pluripotency and Cancer 13

1.3 Context of this investigation 14

2 Materials and Methods 16

2.1 Cell Lines 16

2.1.1 Table of Cell Lines 16

2.2 Mycoplasma 17

2.3 Mice 17

2.3.1 Table of Mouse Strains 17

2.4 Cell Culture 17

2.5 Derivation of Murine Embryonic Fibroblasts (MEFs) 18

2.6 DNA Extraction 19

2.7 Genotyping and Mycoplasma Testing 19

2.7.1 PCR Reaction Mixes 19

2.7.2 Primers 20

2.7.3 PCR Programmes 20

2.7.4 Agarose Gel Electrophoresis 21

2.7.5 Band Sizes 21

2.8 SDS-PAGE and Western Blot Analysis 22

2.9 Self-Renewal Assay and Alkaline Phosphatase Staining 23

2.10 Annexin V Staining Apoptosis Assay 24

2.11 Embryoid Body Assays 24

2.11.1 Embryoid Body Culture 24

2.11.2 Reverse Transcription Quantitative PCR 25

2.12 Reprogramming Assay 27

3 Results 29

3.1 Expression of H3.3 mutants K36I/M and K9I/M causes a global reduction in H3.3K36 and H3.3K9 tri-methylation 29

3.2 Alkaline Phosphatase staining shows H3.3 mutations do not hold mESCs in a pluripotent state 30

3.3 Annexin V staining of H3.3K9M showed no effect on the amount of apoptosis occurring 33
3.4 Transgenic H3.3-expressing embryoid bodies show a size phenotype 36
3.5 It is unclear whether the EBs are differentiating more or less as a result of the H3.3 transgenes 40

4 Discussion 43
4.1.1 Expression of transgenic histone H3.3 reduces the tri-methylation of lysine residues 36 and 9 43
4.1.2 Self-renewal ability is not maintained by the expression of transgenic histone H3.3 44
4.1.3 H3.3K9M does not seem to have a cytotoxic effect on mESCs 45
4.1.4 qPCR analysis did not produce any concrete findings 46
4.1.5 Reprogramming Assay 47
4.1.6 Findings from this investigation in the context of cancer 47

5 Bibliography 48

6 List of Figures 56

7 List of Tables 57

8 List of Abbreviations 58

1 Introduction

The term ‘epigenetics’ was first coined by Conrad Waddington in 1942, later defined as differences in phenotype without differences in genotype, as an explanation for certain developmental processes for which there was little understanding of the mechanisms[4–6]. Epigenetic mechanisms are now much better characterised and understood as the alteration of gene expression patterns by the adaptation of chromatin, without the alteration of the underlying DNA sequence[6]. The importance of epigenetic control lies in its ability to widely change gene expression programmes and the direction of cell-type identitites[6].

A cell’s vast amount of genetic information is stored within DNA molecules, which can reach lengths of up to 2 meters (total extended length of human DNA)[7]. In order to fit this large amount of material into the cell’s nucleus, which has a diameter of 5 to 10 μm[7], the DNA molecules are compacted up to $2x10^{-5}$ times smaller in length with the aid of proteins known as histones to form chromatin[8]. Chromatin is the compacted form of DNA comprising a complex of proteins (histones) and DNA that forms chromosomes in the nuclei of eukaryotic cells[8].

Histones are highly conserved proteins in eukaryotes which form multi-unit complexes around which DNA wraps to form nucleosomes. Nucleosomes are the fundamental units which allow for the compacting and organisation of DNA into chromatin, whilst also facilitating critical cellular processes such as gene transcription, DNA replication and damage-induced DNA repair[9]. The core nucleosome particle consists of a hetero-octamer of histones comprising two copies of each of the core histones H2A, H2B, H3 and H4, with approximately 146 base pairs (bp) of DNA are wrapped around the octamer [9,10]. Nucleosomes are separated by DNA of variable length known as linker DNA, to which the histone H1 binds, allowing further compaction of chromatin[11,12]. Various in vitro and in vivo investigations revealed that the core nucleosome particle are assembled in two steps. A tetramer of both old and new histones H3 and H4 $(H3\text{-}H4)_2$ is deposited onto the DNA first. This allows the organisation of the central 120 base pairs of DNA. Two H2A-H2B dimers subsequently associate with the H3-H4 tetramer, completing the structure of the nucleosome core particle, and allowing the organisation of the 146 base pairs of DNA to complete the nucleosome[13].

The core histones are characterised by two components; a C-terminal globular domain, around which the DNA wraps, and an N-terminal tail[14,15]. The N-terminal tail of histones vary in length and extend outside of the nucleosome core[8,16]. Near the C-terminus of every core histone is a structurally conserved motif known as the histone fold domain (HFD) which consists of three 11-residue long α-helices[16] (α1, α2, and α3), connected by two unstructured ‘short-strap’ loops (L1 and L2)[16]. These loops allow heterodimeric interactions, known as the handshake motif, to occur between core histones[8,16,17]. A-helices are secondary structures of proteins which are formed by the arrangement of peptides into a right-handed helical

conformation[18]. These helices assemble into heterodimers through the interleaving of the helices and subsequent juxtaposition of the loops into parallel β-bridges[19,20].

The HFDs assemble together in antiparallel pairs: H3 with H4, and H2A with H2B[10], forming the dimers which then allow the subsequent assembly of the octameric nucleosome core.

The HFD has also been observed in archaeal histones and in some eukaryotic transcription factors[16]. Since all three of the major archaeal branches have been found to express histones, it can therefore be concluded that histones were present in the common ancestor of eukaryotes and archaea[21,22].

Regulation of chromatin involves dynamic mechanisms which underlie the vast majority (if not all) of biological processes, including DNA methylation, chromatin remodelling by ATP-dependent complexes, such as the SWI/SNF related enzymes[23], histone variants, and covalent histone modifications[24].

The N-terminal tails of the core histones are subject to extensive post-translational modifications (PTMs), which play a fundamental role in the epigenetic regulation of gene expression through transcriptional activation, silencing, assembly of chromatin and DNA replication[8]. These PTMs act as switches for gene expression regulation by affecting the binding of proteins to the histone tails, therefore regulating the type of protein complexes that associate with certain regions of chromatin[25]. PTMs are chemical modifications of the histones that do not cause any changes to the DNA sequence. Histone PTMs are being widely studied; at least eight different classes of covalent modifications of histones, involving upwards of sixty distinct sites of modification within the core histones have been characterised so far; acetylation of lysine, methylation of lysine and arginine, phosphorylation of serine and threonine, ubiquitylation of lysine, poly-ADP ribosylation of arginine, sumoylation of lysine, deimination of arginine, and isomerisation of proline[15,24] (see Table 1).

Chromatin Modifications	Residues Modified	Functions
Acetylation	K-ac	Transcription, Repair, Replication, Condensation
Methylation: Lysine Arginine	K-me1/ me2/ me3 R-me1/ me2a/ me2s	Transcription, Repair
Phosphorylation: Serine Threonine	S-ph T-ph	Transcription, Repair, Condensation
Ubiquitylation	K-ub	Transcription, Repair
Poly-ADP ribosylation	E-ar	Transcription
Sumoylation	K-su	Transcription
Deimination (citrullination)	R > Citrulline	Transcription
Proline isomerisation	P-cis > P-trans	Transcription

Table 1. The eight classes of post-translational modifications identified on histones. The residues modified and the associated functions of the modifications are outlined[15].

The histone code hypothesis proposes that the complex pattern of PTMs, resulting in several layers of regulation in the process of interpretation and transcription of the genome, renders distinct biological consequences[11,26,27]. This hypothesis generally states that there are two, non-mutually exclusive mechanisms through which post-translational mechanisms can act. *Cis* mechanisms regulate access to the DNA by transcription factors, for example, by structurally changing the chromatin fibre through inter-nucleosomal contacts. *Trans* mechanisms initiate distinct biological processes by acting as docking sites for effector molecules, such as WDR5 or heterochromatin protein 1 (HP1)[11]; specific proteins associating with certain modifications on histone tails. These proteins have the ability to act as activators or inhibitors of transcription, or to act in the maintenance of chromatin structures[25].

It is highly likely that multiple chromatin states are regulated and maintained in a tissue-specific manner, allowing the transcription machinery access to the DNA at specific time points at precise locations[8,28–31]. The major histone encoding genes which produce the bulk of the cellular histones, hereby referred to as canonical histones, are primarily expressed during the S phase (synthesis phase) of the cell cycle[30,31]. These histones are used for the packing of newly synthesised DNA via nucleosome assembly[8]. Histone variants, hereby referred to as non-canonical histones, which differ from their canonical paralogues in primary amino acid sequence are not limited in their expression to the S phase[22]. These non-canonical histones have been found to be expressed throughout the cell cycle, and have roles in various processes (described in more detail further into this chapter)[22,34,35].

With the exception of H4, multiple histone variants exist which have specialised functions in DNA packaging, further contributing another level of epigenetic complexity and regulation. These histone variants differ at the level of their primary sequence ranging from a few amino

acid changes to large domains[9,36]. These differences in primary sequence confer changes upon the physical properties of nucleosomes and nucleosome dynamics.

Histone H3.1 and H3.2 differ by only a single amino acid (Ser96 in H3.2), and H3.3 is distinguished by four additional amino acid substitutions (Ser31, Ala87, Ile89, Gly90)[37].

Genes encoding canonical histones are organised in tandem multi-copy repeat clusters[36]. The transcription of these canonical histone-encoding genes is tightly coupled to DNA replication. Conversely, non-canonical histone variant-encoding genes are typically found singly in the genome and are constitutively expressed[22] and are therefore also known as 'replacement' histones[36].

These non-canonical histone variants are involved in a wider range of processes than their canonical counterparts. For example, histone variants were shown to be critical for DNA repair, meiotic recombination, chromosome segregation, sex chromosome condensation, sperm chromatin packaging, and initiation and termination of transcription[22].

As a result of the combination of PTMs and histone variants, a wide diversity of nucleosomes can be created in vivo. This variation is important in determining the chromatin fibre properties at a local and regional level with respect to replication, transcription[38], heterochromatin or euchromatin formation, repair, or kinetochore formation (for review see 'Epigenetic and replacement roles of histone variant H3.3 in reproduction and development'[36]). There are two varieties of chromatin that reflect the activity level of a cell; euchromatin and heterochromatin. Euchromatin is the lightly packed, most active form of chromatin, and is predominantly actively transcribed. Heterochromatin is a highly condensed and tightly packed form of DNA that comes in two forms; constitutive and facultative heterochromatin. Constitutive heterochromatin, found at the pericentromeric regions and at telomeres, is generally the most static form of chromatin, and does not contain many genes. Facultative heterochromatin forms at various regions within the chromosome, and often contains genes which are generally silenced upon developmental cues. It is thought that the tight packing of the DNA in heterochromatin renders the DNA inaccessible to polymerases, therefore preventing transcription (for review see [6,39]).

As previously mentioned, canonical histone encoding genes are usually found clustered in repeated arrays, do not contain introns, and replicative histone mRNAs are not polyadenylated[36]. In place of polyadenylation, the binding of Stem Loop Binding Protein (SLBP) and U7 snRNP to the 3' end of the histone RNAs tightly regulates the translation of these canonical histones[36,40,41]. This unusual genomic organisation and transcriptional regulation mechanism allows the translations of high numbers of canonical histones at the beginning of the S phase, prior to DNA replication. This ensures a massive supply for histone incorporation events during replication in a manner coupled to DNA synthesis, whilst also synthesising stoichiometric quantities of each protein[36,42].

Canonical histone H3 (referring to H3.1 and H3.2 in higher eukaryotes) can be incorporated into chromatin outside of S phase, but is still coupled to DNA synthesis during DNA repair after damage by ultraviolet irradiation, which involves deposition of the H3.1 by histone chaperone CAF-1 (chromatin assembly factor-1)[13,43,44]. The histone H3.3 variant is deposited via replication-independent nucleosome assembly by the histone chaperones histone regulator A (HIRA), and death-associated protein 6 (DAXX)[13,45].

Different H3 variants have been identified in mammals, including two canonical variants, H3.2 and mammalian-specific H3.1, and three replacement variants, H3.3, the centromere-specific variant CenH3 (CENP-A in mammals)[9,46,47] and the testis-specific histone H3t[48]. Two primate-specific H3 variants, H3.X and H3.Y, have also been characterised[9,49].

Histone H3.3 is one of the most conserved proteins in eukaryotes. It is encoded by two genes *H3f3a* and *H3f3b* in mice, and *H3F3A* and *H3F3B* in humans, which produce the exact same H3.3 protein despite containing distinct untranslated regions (UTRs)[50–52]. This is an interesting result of histone H3.3 being unique amongst the histone proteins, being the only variant whose genes contain introns and whose mRNA is polyadenylated. This presence of introns and the polyadenylation of the mRNA allows another level of control over the expression of the H3.3 variant. For example, the transcripts produced after transcription of the genes have varying polyadenylation sites, resulting in differing lengths of transcripts. These different polyadenylation sites do not seem to be random, and have been found to display some tissue specificity[53].

The H3 variants also differ from their canonical counterparts in that they are constitutively expressed, enabling histone deposition and exchange which is independent of DNA synthesis and the cell cycle. This constitutive expression provides a continuous source of the H3 variant in all instances, converse to the availability of the canonical H3 being limited to the S phase[54]. Both genes (*H3F3A* and *H3F3B*) are expressed during quiescence and are enriched in various stages of differentiation[9,51,52,54,55]. For instance, it has been demonstrated in *Drosophila melanogaster* that there is a genome wide incorporation of the H3.3 variant within the male pro-nucleus at the point of fertilisation, before the first round of DNA replication[13,56,57]. The levels of histone H3.3 reach up to 90% of the total H3 content in mature cortical neurons as a result of this cell cycle-independent expression[58,59].

Canonical histone H3 is slowly replaced by histone H3.3 as differentiating cells exit the cell cycle[60]. Various means have been suggested for the deposition of H3.3 independently of transcription, since as well as being highly enriched within genes undergoing active transcription[61–63], H3.3 is also found to be involved in sperm reprogramming[56,64,65], and also accumulates at telomeric and pericentric heterochromatin[44,45,66,67].

This specific histone variant deposition is most likely to be directed by histone chaperones, such as the HIRA chaperone complex, as this chaperone complex promotes histone deposition independently of DNA synthesis[57]. This complex which is comprised of HIRA,

UBN1 (Ubinuclein-1), and CABIN1 (Calcineurin Binding Protein 1) has been implicated in this deposition of H3.3 variants due to having been found with H3.3 in soluble complexes[13,44]. HIRA has been found to hold a critical role in the accumulation of H3.3 at active genes, promoters and certain regulatory elements, unrelated to other chaperones (outlined below)(reviewed by Ray-Gallet *et al*. 2011[44]).

Other potential pathways which contribute to the distinct localised enrichment of H3.3 are other H3.3 chaperones; DEK (Drosophila Eph kinase)[68], and a complex formed by the alpha-thalassemia/ mental retardation X-linked syndrome protein (ATRX) and DAXX[45,67,69,70]. The DAXX/ATRX complex have been implicated in enrichment of H3.3 at telomeres[45].

Histone 3 is widely modified by methylation at lysine and arginine residues, and acetylation at arginine residues. Histone acetylation is the enzymatic transfer of acetyl groups ($COCH_3$) from acetyl coenzyme A (AcetylCoA) to residues on the N-terminal tail of a histone by histone acetyltransferases (HAT). Acetylation causes the overall charge of the histone tail to change from positive to neutral, subsequently weakening DNA-histone contacts due to the reduced affinity for the now neutral histone tails to the negative DNA. This results in greater access of transcription factors to promotors in the chromatin[71–73]. It has also been suggested that histone acetylation may prevent nucleosomal arrays from folding into more complex structures which would have had inhibitory functions towards transcription[74,75].

Histone methylation is the transfer of methyl groups (CH_3) from S-adenosyl-L-methionine to lysine, or arginine residues in the histone's N-terminal tail by a histone methyltransferases (HMT). Lysines can be mono-, di-, or tri-methylated on their ε amine group[76–79], and arginines can be mono-, symmetrically dimethylated or asymmetrically dimethylated on their guanidyl (NH_3^+) group[76,80,81]. The methylation of lysine residues in the histone tail can be associated with either activation or repression, depending on the site of methylation[82], as shown in the overview of the possible modifications of histone H3 in Figure 1. Methylation of histone H3 lysine 9 (H3K9) is generally found in correlation with the repression of transcription[83], and serves as a binding site for HP1[25]. In contrast, methylation of histone H3 lysine 36 (H3K36) has been associated with transcriptional activation[15,84].

The histone variant H3.3 is found in transcriptionally active or "poised" domains which are commonly enriched for H3 lysine 4 trimethylation (H3K4me3) or possessing both lysine 27 trimethylation (H3.3K27me3) and H3K4me3[85]. It has been shown that only the methylation of lysine 9 is present prior to the incorporation of the histone variant into chromatin, all other residues are methylated post-incorporation into chromatin[86].

The histone modifying enzymes Set2 (Sc/Sp SET2), NSD1, and SYMD2 are involved in the methylation of lysine 36 of histone H3[87]. Several histone-modifying enzymes are capable of methylating H3K9, including SUV39H1, SUV39H2, G9a[15,22]

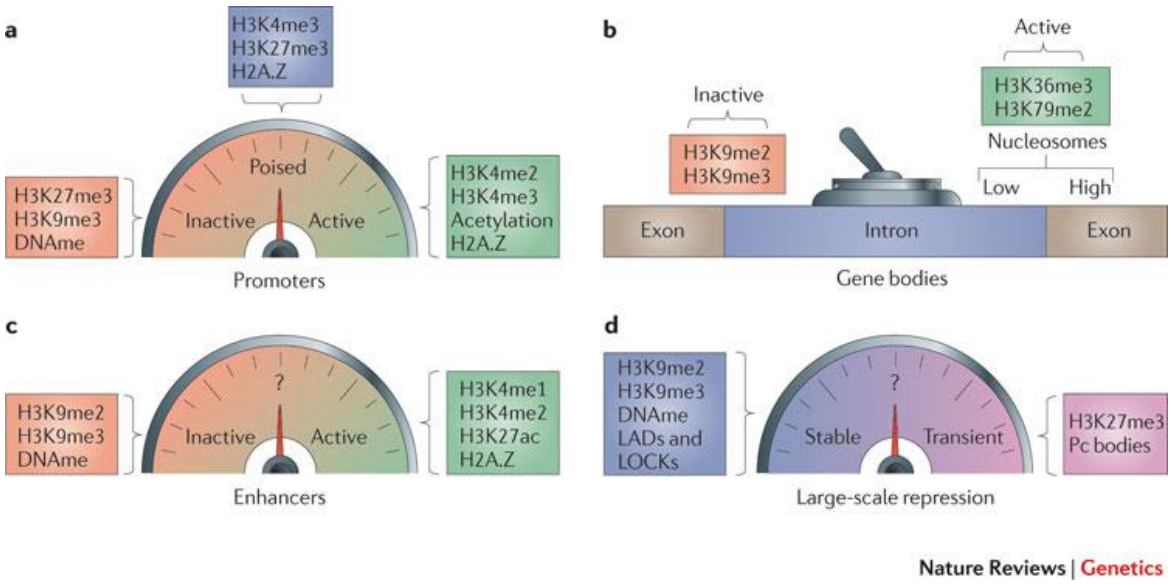

Figure 1. Overview of the various possible histone H3 modifications, and their associated functions: A) Contribution to fine tuning expression levels at promoters. B) Discrimination between active and inactive conformations at gene bodies. More histone 3 lysine 36 trimethylation (H3.3K36me3) and H3K79me2-modified histones are found in exons in active genes than introns, due to the higher nucleosome occupancy on active exons. C) Histone marks correlate with levels of enhancer activity at distal sites. D) Possible conferral of repression of varying stability and association with different genomic features on a global scale.[88]

1.1 Histones and Disease

Due to the complexity and multiple levels of control involved in the regulation of histone deposition, replacement and modification, there are many ways in which the regulation or the regulators of any of these processes may go wrong. Many diseases have been found to be caused as a result of mutations affecting the regulation of these processes. For example, α-thalassemia mental retardation X-linked (ATRX) syndrome has been found to be caused by a downregulation of α-globin gene expression as a result of mutations which may cause subtle changes in the localisation of histone H3.3 upon loss of ATRX[89]. 43% of pancreatic neuroendocrine tumours (PanNETs) have been found to express a mutation in either the ATRX or DAXX histone chaperones[90].

The most common type of disease linked to histone variants and their associated factors is cancer. Colorectal cancer has been found to be linked to the upregulation and mis-targeting of CENPA (centromere protein A), a histone H3-like nucleosomal protein found in nucleosomes located in the centromere, which may play a role in aneuploidy[91,92]. Breast cancers have also been linked to histone variants and their factors; the upregulation of Holliday Junction Recognition Protein (HJRP), a protein involved in the incorporation and maintenance of CENPA at centromeres, has been linked to a poor prognostic outcome[93]. Also the overexpression of the histone variant H2A.Z was found to facilitate the activation of genes responsive to oestrogen, and has been associated with metastasis[94]. Conversely, reduced expression of histone variant H2A.Z found in colon cancer was found to be associated with genomic instability[95,96].

The histone variant pertinent to this investigation is the H3.3, which has been found to be linked to various cancers.

Mutations in the histone variant H3.3 have been found in a high percentage of chondroblastomas, giant cell tumour of bone, and paediatric gliomas (diffuse intrinsic pontine gliomas (DIPGs) and supratentorial glioblastoma multiforme (GBM))[1–3]. More specifically, Schwartzentruber *et al.*[3] observed that 31% of GBM tumours contained mutations in the *H3F3A* gene, which resulted in amino acid substitutions at critical positions within the histone tail of Lysine 27 to Methionine (K27M), or Glycine 34 to Arginine or Valine (G34R/V).

Behjati *et al.*[1] observed that in 95% of the chondroblastoma cases studied, there was a mutation in the *H3F3B* gene resulting in a substitution of lysine 36 to methionine (K36M) in the histone tail, and that 92% of giant cell tumours of bone displayed histone H3.3 variants, exclusive to *H3F3A,* which were G34W (and G34L in one case).

Lewis *et al.*[2] discovered that 60% of glioma patients studied exhibited one of two mutations in either *H3F3A*, or *HIST1H3B*, one of the several H3.1 encoding genes. This mutation resulted in a substitution of K27M in 80% of diffuse intrinsic pontine gliomas (DIPGs) and 22% of non-brain stem gliomas studied[97]. It was found that this K27M substitution resulted in both the global decrease of di- and tri-methylation of H3.3K27 residues, and a moderate increase in acetylation of H3.3K27. These global changes in histone post-translational modifications were specific to H3.3K27, since two histone modifications related to transcription (H3K4me3 and H3.3K36me3) were similar in DIPG samples in all genotypes.

Together, these results suggest that the mechanism behind the formation of tumours in the case of the K27M mutation is different to the mechanism behind those caused by H3.3G34R/V (two common mutations found in other cancers), since the H3.3G34R/V mutations had no impact on the methylation or acetylation levels of any of the H3 Lysine residues. Lewis *et al* found that the mutant K27M peptides inhibit the Polycomb Repressive Complex 2 (PRC2) via interaction with the Enhancer of Zeste Homolog 2 (EZH2) active site[2]. The mechanism behind this inhibition involves the PRC2 complex being unable to detach from the K27M peptide, because EZH2 cannot hydrolyse S-Adenosyl Methionine (SAM) in order to methylate the residue. Upon binding both SAM and K27M, allosteric changes within the complex strengthen the association of PRC2 with K27M, preventing the complex from being ejected, since SAM is not converted to *S*-Adenosyl homocysteine (SAH).

Lewis *et al* also expressed every amino acid substitution for H3.3K27, and discovered through immunoblotting that only the K27I and K27M substitutions had the same effect of globally reducing tri-methylation of H3.3K27.

Subsequently, the group investigated whether highly conserved active sites of other SET-domain proteins are sensitive to methionine substitution, by creating H3.3 transgenes containing methionine at K4, K9, and K36. Interestingly, H3.3 transgenes containing K9M and K36M were also found to decrease di- and tri- methylation of K9 and K36 respectively.

Mutations in SETD2, a histone H3 lysine tri-methyltransferase, resulting in a decrease in the amount of H3K36me3, has also been implicated in metastatic renal cell carcinoma high resolution profiling of histone h3 lysine 36 in metastatic renal cell carcinoma

1.2 Pluripotency and Cancer

Because of the similarities between stem cells and cancer cells, for example their unlimited proliferative capacity, self-renewal ability, a block in differentiation, telomerase activity, and telomere lengths there may be a link between these K-to-M/I substitutions in histone H3.3 and pluripotency.

Pluripotent stem cells show two distinct characteristics: self-renewal, the ability to maintain proliferation without causing damage or changes to their cellular properties and characteristics, and the capacity to produce any type of cell found in adult organisms[98]. There have been a variety of important transcription factors that have been identified as being key regulatory factors in the control of pluripotency in embryonic stem cells (ESCs). These factors are Oct-4 (Octamer-binding transcription factor 4)[99], Sox2 (SRY-box2)[100], and Nanog[101].

Yamanaka *et al* showed that through the introduction of four transcription factors, this allowed the reprogramming of somatic cells back into ES cell-like-cells known as induced pluripotent stem cells (iPSCs). The four transcription factors they identified for use in this process are Oct4, Sox2, Klf4 (Kruppel-like factor 4), and Myc, and are now known as the Yamanaka factors[102,103]. This discovery underlines the importance and relevance to cell fate control and the pluripotent state of transcriptional regulatory mechanisms. This has resulted in the production of a powerful tool for the study of cell fate choice during differentiation; lineage reprogramming[104].

It has been suggested that the similarities seen between stem cells and cancer cells might be a result of shared regulation of gene expression patterns, potentially associated with the 'embryonic' state. Studies involving somatic cell reprogramming have further highlighted the similarities between cancer cells and iPSCs[98].

The polycomb-group proteins, via the generation of repressive histone marks, play an important role in contributing to the repression of critical developmental or lineage specific regulators. This means that their role is essential early on in development, and in ESCs. Polycomb-mediated gene silencing lies in the regulation of chromatin structure, partially through the post-translational modification of histones, as previously described (See earlier in this Introduction)[105]. These polycomb-repressive complex (PRC) proteins have been found to be involved in the reprogramming of somatic cells[106,107].

Myc, a highly studied oncogene, has been found to be expressed in up to 70% of cancers in humans. Due to this high occurrence, it has been suggested that the expression of Myc is required for tumorigenesis[108]. One of the many cellular functions in which Myc is involved is

the control of self-renewal in ES cells (reviewed by Meyer *et al* [109]). It has also been found that Myc's role in the maintenance of pluripotency in ESCs is different from that of the core factor interaction network[110,111] or the PRC network. The histone acetyltransferase complex proteins Tip60-p400 has been shown to interact with Myc in ES cells, to play a crucial role in ESC identity[112,113]. Myc can be activated via a variety of mechanisms, such as amplification, translocation, or enhanced translation or protein stability (for review, see Nilsson *et al*[108]).

Characteristics that have been found to be shared by ESCs and cancer cells have been investigated on the level of gene expression patterns in order to identify the mechanisms behind these shared characteristics. Inactivation of the p53 pathway, which is seen to occur in a majority of cancers, causes an increase in the efficiency of the reprogramming process[102].

Oct4, Nanog and Sox2 have been shown to function as repressors of differentiation in human ESCS (hESCs), and prevent ESCs from diverging into the wrong cell type too soon, as well as maintaining self-renewal and pluripotency in ESCs. Each transcription factor regulates a specific cell fate; Oct4 regulated extraembryonic and epiblast- derived cell fates, embryonic ectoderm differentiation is repressed by Nanog, but Nanog has little effect on other lineages. Sox2 (and Sox3) repress mesendoderm differentiation[114].

These transcription factors are found to be expressed in cancer stem cell-like-cells (CSCLCs), and it has been shown multiple times that the overexpression of these three genes is linked to or results tumorigenicity, tumour transformation and metastasis[115].

These three transcription factors have been found to be co-up-regulated in a large variety of human cancers[115–117].

Even though it has been shown that ESCs and cancer cells display some similar properties, cancer cells are not truly pluripotent like ESCs.

1.3 Context of this investigation

The aim of this investigation was to assess whether the histone H3.3 substitution mutations of lysine 36 and lysine 9 to methionine or isoleucine has any kind of effect upon the maintenance of the pluripotent state of mouse embryonic stem cells (mESCs), and whether these mutations have an effect, whether negative or positive, on the reprogramming efficiency of mouse embryonic fibroblasts back to the pluripotent state. The relevance of this investigation is the fact that multiple cancers (outlined in Histones and Disease section) have been found to express substitution mutations at critical points within the histone tail of the variant H3.3[1–3]. These mutations cause a global loss in the methylation of the residues which had been substituted, which is causing the formation of tumours. It has not yet been investigated as to whether there is a link between these mutations and the pluripotent state. The rationale behind this investigation is that there are multiple similarities between pluripotent stem cells and cancer cells, including the gene expression patterns, self-renewal ability, and telomerase activity, amongst others outlined previously in this introduction. If the mESCs expressing these transgenic H3.3 variants are indeed being held in the pluripotent

state, this could lead to tumour formation due to the fact that pluripotent stem cells are immortal, and can continue proliferating for as long as feasible. This persistence of mitotic activity of stem cells increases the potential of these cells becoming a reservoir for the accumulation of oncogenic mutations[118]. This is due to the fact that the self-renewal mechanisms that allow stem cells to maintain this state involve proto-oncogenic pathways; Wnt, Shh, and Notch[118]. Another potential hypothesis of the cause of these cancers as a result of these substitutions in the histone H3.3 tail is that the loss of methylation causes a loss in repressive marks on tumorigenic genes, or conversely a gain in active marks on tumorigenic genes.

2 Materials and Methods

All general consumables were obtained from Corning (USA) and Thermo Fisher Scientific (USA). Nikon microscopes were used for observation and counting of cells in culture, a Nikon dissection microscope was used during the derivation of MEFs, and an Olympus IX81 microscope was used in the imaging of the Embryoid Bodies.

2.1 Cell Lines

Mouse embryonic stem cell (mESCs) lines were engineered as previously described[119] under a doxycycline (Dox) inducible promoter, and obtained from Dr. Justin Brumbaugh. The two components of the Dox inducible system are the reverse tetracycline controllable trans-activator (rtTA), and the tetracycline operator minimal promoter (tetOP), which drive expression of the gene of interest[119]. In the endogenous bacterial system, transcription occurs in the presence of tetracycline, or one of its analogues such as doxycycline, as the Tet repressor (tetR) will bind to the tetracycline and not the Tet response element (TRE).
A reverse Tet repressor (rTetR) was developed by Gossen *et al*[120] which required the presence of tetracycline (or an analogue) for induction of transcription. This new trans-activator was named rtTA (reverse tetracycline controllable trans-activator), and is capable of binding to the operator only if bound to tetracycline, or in this case doxycycline. Therefore, the addition of Tet or Dox initiated the transcription of the gene of interest[119,120].
The mESCs were engineered to contain either one of the histone H3.3 transgene variants (K36M, K36I, K9M or K9I), overexpression of the wild type histone H3.3, or overexpression of Nanog (obtained from Dr Konrad Hochedlinger).

2.1.1 Table of Cell Lines

H3.3K36I	Substitution of H3.3 Lysine 36 by Isoleucine
H3.3K36M	Substitution of H3.3 Lysine 36 by Methionine
H3.3K9I	Substitution of H3.3 Lysine 9 by Isoleucine
H3.3K9M	Substitution of H3.3 Lysine 36 by Methionine
H3.3 OE	Overexpression of wild-type histone variant H3.3. This cell line was used as a negative control in order to show that the increased expression of the histone variant H3.3 does not have any impact on the methylation state of H3.3K36 or H3.3K9.
Nanog OE	Overexpression of Nanog. This cell line was selected for use as a positive control in the self-renewal assay. See section 2.9.

Table 2 Overview of Cell Lines used.

2.2 Mycoplasma

Mycoplasmas are a highly common bacterial contaminant of cells in culture, and are the simplest and smallest self-replicating bacteria. Mycoplasma infection of cultured cells can cause alterations to many processes within host cells, including interference with membrane receptors and transport mechanisms[121], oxidative damage to the cell membrane[122], and hydrolysis of phospholipids on the host cell membrane[123]. These could in turn could lead to the triggering of various signalling cascades[124].

Unfortunately Mycoplasmas are often resistant to the antibiotics commonly used in cell culture, and also do not often cause visible cell damage or produce turbid growth[15,125]. This makes testing via methods such as PCR necessary for accurate detection of contamination. To avoid the use of cells contaminated with Mycoplasma for these investigations, all cells in culture were tested for Mycoplasma every 2-3 passages (see Genotyping and Mycoplasma testing section 2.7). Murine Embryonic Fibroblasts (MEFs) were tested after derivation at passage 0, and before freezing.

2.3 Mice

All protocols which involved the use of mice during this investigation were performed as per the Institutional Animal Care and Use Committee (IACUC) and the Massachusetts General Hospital Centre for Comparative Medicine's (CCM) regulations for the humane treatment of laboratory animals, to ensure the adherence to all pertinent laws and regulations.

2.3.1 Table of Mouse Strains

H3.3K36M/ rtTA	Heterozygous/ heterozygous
OKSM/ rtTA	Heterozygous/ heterozygous

Table 3 Overview of mouse strains used for producing MEFs for reprogramming assay.

2.4 Cell Culture

Cell culture plates were prepared for mESC culture by incubation with 0.1% gelatine solution for a minimum of 10 minutes (min) at 37°C. After removing the gelatine solution from the culture plate, irradiated Murine Embryonic Fibroblasts (MEFs) were subsequently seeded on to the gelatinised plates as feeder cells in MEF media (DMEM (Dulbecco's Modified Eagle Medium, Gibco™ ThermoFisher), 10% filtered FBS (Foetal Bovine Serum, Sigma), 1X NEAA (Non-Essential Amino Acids, Gibco™ ThermoFisher), 1X GlutaMAX™ (Gibco™ ThermoFisher), 55 µM β-Mercaptoethanol (Gibco™ ThermoFisher)). The feeder cells were plated and incubated overnight at 37°C, allowing the MEFs to completely cover the growth surface prior to the plating of the mESCs.

The mESCs were plated onto the prepared feeder cells in ESC media (Gibco™ Knockout-DMEM (ThermoFisher), 15% filtered FBS (Hyclone), 1x Non-Essential Amino Acids (NEAA), 1x GlutaMAX™, 55 µM β-Mercaptoethanol, 1000 U/mL Leukaemia Inhibitory Factor (LIF, produced in-house)). They were incubated at 37°C, in 5% CO_2.

The ESC media was changed daily, and the mESCs were passaged every 3-4 days depending on the confluency of the cells. The cells were passaged by first removing the old media from each well. The cells were then dissociated by adding Trypsin-EDTA (Gibco™ ThermoFisher), and incubating the plates for approximately 3-5 min at 37°C, 5% CO_2, until the cells had detached from the plate. The Trypsin was then quenched with fresh ESC media (1-2 times the volume of the Trypsin used). The cells were gently re-suspended by pipetting the media up and down, and subsequently the required volume of cell suspension was transferred to a new well with pre-plated feeder cells and fresh ESC media.

2.5 Derivation of Murine Embryonic Fibroblasts (MEFs)

MEFs were derived and expanded in culture to be used as feeder cells for the culturing of mESCs.

To time the pregnancy, the female mice were checked in the mornings for a vaginal plug to determine the day 0 point. The foetuses were harvested at between 13 and 14 d.p.c (day post-coitum) by first sacrificing the mother via cervical dislocation, and immediately disinfecting the female by spraying the animal with 70% (v/v) ethanol. With the dam on her back, the fur covering the belly was removed and the peritoneum was carefully cut open to access the abdomen. The uterine horns were dissected out, and placed into a dish containing cold DPBS (-Ca^{2+}, -Mg^{2+}, Gibco™ ThermoFisher). Each embryo was separated from its placenta and embryonic sac before being moved to a new dish with fresh cold DPBS. The head, limbs, tail, internal organs and spinal cord were all removed from the foetus, and the remaining tissue was placed into a 6 well plate (one foetus per well) in DPBS.

Under a tissue culture hood in sterile conditions, the tissue was transferred to a new, sterile 6 well plate and finely minced in 200 µl of Trypsin EDTA by using two sterile scalpels. When the tissue was minced finely enough to be possible to pipette, 500 µl of trypsin was added to each well and incubated for 10 min at 37°C, 5% CO_2. 4 mL of MEF media with 1X sterile filtered Pen-Strep (Penicillin Streptomycin, Sigma-Aldrich) were used to re-suspend the cells as thoroughly as possible by pipetting the media up and down multiple times. The cell suspension was transferred to a T75 cell culture flask containing 6 mL MEF media with 1X Pen-Strep, and incubated at 37°C in low oxygen.

2 mL of MEF media (without Pen-Strep) were added to the wells of the 6 well plates where the tissue was minced to expand out any remaining MEFs. These cells were then used for Mycoplasma contamination testing via Polymerase Chain Reaction (PCR).

The MEFs were expanded in a low oxygen incubator until they became confluent (usually after 2 days of culture), and split 1:3 into 3 x T175 flasks (the splitting was performed using the same procedure as described in section 2.4, Cell Culture). The MEFs were expanded in this fashion until passage 3 or 4.

Prior to use as a feeder layer for mESCs, MEFs must be mitotically inactivated via irradiation[126]. Before being irradiated, the MEFs were harvested by using the same method as that used to passage cells (described in section 2.4, Cell Culture), counted using a haemocytometer, and collected into 50 mL conical flasks on ice. A caesium source irradiator was used to irradiate the MEFs at a dose of 30 gray.

After the MEFs had been irradiated, the cells were pelleted by using a Beckman Coulter Allegra X-12R centrifuge at 1000 rpm for 5 min, and the media was removed. To re-suspend the MEFs for freezing, half of the desired final volume of MEF media (depending on the desired cell concentration of the frozen cell stocks) was added to the cells, and the suspension was subsequently made up to the desired final volume by the addition of freezing media (20% DMSO (Dimethyl Sulfoxide, Sigma-Aldrich) in MEF media). The freezing media was added to the cell suspension dropwise with slight shaking until the suspension was homogenised. The cells were pipetted into cryovials, which were then transferred to a Mr. Frosty™ Freezing Container (Thermo Fisher Scientific) filled with 100% Isopropanol, and placed into a - 80°C freezer to achieve a rate of cooling of -1°C per min.
A small sample of cells was taken for Mycoplasma testing of the MEFs before freezing (outlined in section 2.6).

2.6 DNA Extraction

Tissue for the genotyping of the mice was obtained by cutting and collecting approximately 2 mm of the tip of the tail from mouse pups of around 7-10 days old and placed into a PCR tube.
In order to collect DNA from cells for either genotyping or Mycoplasma contamination testing, the cells were harvested using the procedure described in section 2.4, Cell Culture. The harvested cells were pelleted (1000 rpm for 5 min) and the media was removed. The pellet was washed twice with DPBS.
75 µL of Alkaline extraction buffer (5 mL H_2O, 11.65 µL 30% NaOH, 0.7 M EDTA) were added to the tail tissue or cells (the lysed cell suspension was then transferred to PCR tubes). The tubes containing the tail tissue/ cells were incubated in a thermocycler (BioRad PTC-200 Thermocycler) at 95°C for 30 min, followed by 4°C for 15 min. The reaction was immediately quenched via the addition of 75 µL of 40 mM Tris HCl, pH 5.5.
The DNA was then used for PCR reactions for genotyping or Mycoplasma testing.

2.7 Genotyping and Mycoplasma Testing

2.7.1 PCR Reaction Mixes

2 µL of DNA (extracted from mouse tail tissue or cultured cells) were added to a PCR tube, along with 10 µL GoTaq® G2 Hot Start Green Master Mix (Promega), 1 µL of 10 mM forward and reverse primer mix, and 7 µL H_2O.
For the Mycoplasma testing, positive controls were used to ensure that the PCR worked. These positive controls were DNA derived from cells known to be contaminated with Mycoplasma.

2.7.2 Primers

Primers for Mouse Genotyping

Col1a1 (locus of H3.3K36M and OKSM insert):

Col/frt-B	CCCTCCATGTGTGACCAAGG	Reverse
Col/frtA1	GCACAGCATTGCGGACATGC	Forward
Col/frtC1	GCAGAAGCGCGGCCGTCTGG	Targeting vector

ROSA26 (locus of rtTA insert):

Rosa1	AAA GTC GCT CTG AGT TGT TAT
Rosa2	GCG AAG AGT TTG TCC TCA ACC
Rosa3	GGA GCG GGA GAA ATG GAT ATG

Mycoplasma Primers

Myco 5/6:

Myco 5	GGT GTG GGT GAG TTA TTA CAA ART CAA TT
Myco 6	GGA GTG AGT GGA TCC ATA AAT TGT GA

Myco 2/11:

Myco 2	CTT CWT CGA CTT YCA GAC CCA AGG CAT
Myco 11	ACA CCA TGG GAG YTG GTA AT

2.7.3 PCR Programmes

The tubes containing the PCR reaction mixes were capped and placed into the thermocycler, and the following programmes were used for the PCRs:

Mouse Genotyping

Col1A1 –

1. 95°C for 1 min
2. 94°C for 30 s
3. 70°C for 45 s
4. Repeat from step 2 for 1 cycle
5. 94°C for 30 s
6. 68°C for 45 s
7. Repeat from step 5 for 4 cycles
8. 94°C for 20 s
9. 66°C for 1 min
10. Repeat from step 8 for 29 cycles
11. 72°C for 5 min

ROSA26 –

1. 94°C for 30 s
2. 60°C for 30 s
3. 72°C for 60 s
4. Repeat all steps for 30 cycles

Mycoplasma Testing

Myco 5/6 –

1. 94°C for 1 min
2. 94°C for 30 s
3. 50°C for 1 min
4. 72°C for 1 min
5. Repeat from step 2 for 2 cycles
6. 94°C for 30 s
7. 60°C for 1 min
8. 72°C for 1 min
9. Repeat from step 6, 34 cycles
10. 72°C for 5 min
11. 10°C forever

Myco 2/11 –

1. 95°C for 2 min
2. 95°C for 45 s
3. 55°C for 45 s
4. 72°C for 1 min
5. Repeat from step 2, 31 cycles
6. 72°C for 10 min
7. 10°C forever

2.7.4 Agarose Gel Electrophoresis

The PCR products were separated by size by agarose gel electrophoresis for analysis. The gel comprised of 2% Agarose in Tris-acetate-EDTA (TAE) Buffer. The PCR products were pipetted into the wells of the agarose gel, which was submerged in TAE buffer in an electrophoresis chamber. An electrical current of 100 Volts (V) was then passed through the gel for approximately 30 min, or until the dye front reached the end of the gel (depending on how similar the size of the products to be separated were; the gel was run longer if the products were close in size, to allow for an adequate degree of separation of the bands in the gel).

2.7.5 Band Sizes

Col1A1 locus:

Wild type fragment: 331 bp	Primers B + A1
Mutant fragment*: 551 bp	Primers B + C1

ROSA26 locus:

Wild type fragment: 600 bp
Mutant fragment*: 350 bp

Mycoplasma:

Myco 5/6: 250 bp
Myco 2/11: 450 bp

*The ‘mutant fragment’ refers to the band produced if there is an insert at that locus. In the case of Col1A1, the inserts were H3.3K36M and OKSM (The Yamanaka Factors Oct4, Klf4, Sox2 and c-Myc[127], to be used for the reprogramming of the MEFs). The insert at ROSA26 was rtTA, which is required for the induction of the expression of the Yamanaka Factors and the H3.3K36M.

The wild type fragment refers to the band produced when there is no insert at that locus.

2.8 SDS-PAGE and Western Blot Analysis

mESCs (H3.3 mutants and H3.3 overexpression cells, see 2.1.1, Table of Cell Lines) were plated onto gelatine only without feeder cells, in ES media without LIF, to stress the cells towards differentiation. Dox (1000 U/mL) was used to induce the expression of the mutation, and a control was cultured without Dox.

The samples were prepared by first removing the media and detaching the cells as previously described (see section 2.4, Cell Culture). The cells were washed with 1X DPBS and centrifuged at 1000 rpm for 5 min to pellet the cells. The DPBS was removed and the cells were re-suspended in 500 µl of Buffer NB-A (85 mM KCl, 5.5% Sucrose, 10 mM Tris-HCl pH 7.5, 0.2 mM EDTA). Subsequently, 500µl of Buffer NB-B (85 mM KCl, 5.5% Sucrose, 10 mM Tris-HCl pH 7.5, 0.2 mM EDTA, 0.1% IGEPAL) were added and the tube was inverted gently to mix. The suspension was centrifuged at 900 x g for 3 min, and the supernatant was then removed. The pellet was re-suspended in 1 mL of Buffer NB-R (85 mM KCl, 5.5% Sucrose, 10 mM Tris-HCl pH 7.5, 1.5 mM $CaCl_2$, 3 mM $MgCl_2$) and centrifuged at 300 x g for 3 min. The supernatant was removed and the pellet was re-suspended in 50 µl Buffer NB-R. Laemnli Sample Buffer (5X stock: 0.35 M SDS, 1 M Tris pH 6.8, 1.5 mM Bromophenol blue, 0.5 M DTT (dithiothreitol), 50% Glycerol), was added to samples at a 1:5 concentration. The samples were sonicated for 5 min in a bath sonicator at 4°C, were boiled for 5 min to re-dissolve precipitated SDS, and to prepare the samples for loading.

The samples were loaded into a 4-20% polyacrylamide gel (Biorad Mini-PROTEAN TGX Stain-Free™ pre-cast gel), which was mounted in a Biorad Mini-PROTEAN® Tetra Vertical Electrophoresis Cell containing 1X TGS (Tris-Glycine SDS) buffer. Biorad Precision Plus Protein™ Dual Colour Standards protein ladder was loaded alongside the samples. A 200 V current was passed through the cell for ~20 min, until the dye front reached the end of the gel. The gel was placed into a Western Blot transfer sandwich with nitrocellulose membrane in 1X TG (Tris-Glycine) Buffer with 20% methanol (v/v), and transferred at 100v for 1 hour at 4°C. The membrane was then blocked using 5% (w/v) milk (non-fat milk powder in Tris-Buffered Saline with 1:1000 Tween (TBST)) overnight at 4°C on a shaker. The milk was removed and the membrane was incubated with primary antibody (Ab) (rabbit Ab against K36me3 (1:2000), K9me3 (1:2000) or H3 (1:10000)) and 5% (w/v) milk in TBST for 1 hour at room temperature on a shaker. The samples stained for H3 were used as the loading control. The primary antibody was removed and the membrane was washed 3 times for 5 min each in TBST, with shaking. The secondary antibody (goat α rabbit) in 5% milk TBST was added to membrane and incubated for 1 hour at room temp with shaking.

The membrane was washed 3 times for 5 minutes each in TBST. Biorad Clarity™ ECL Substrate (Enhanced chemiluminescence) used to detect the proteins bound on the

nitrocellulose membrane, and to develop them onto photographic film. The proteins were able to be visualised onto photographic film due to luminescence produced as a result of the reaction between the ECL Substrate and the Horseradish peroxidase (HRP) which was conjugated to the secondary antibodies. HRP catalyses the oxidation of luminol into a reagent which emits light upon decaying. The size and location of the bands which were revealed upon development of the photographic film directly correlates with the amount and location of protein on the membrane. This is due to the luminol oxidation being catalysed by the HRP, which is complexed with the protein of interest on the membrane. The more light produced means the more luminol being oxidised due to a larger concentration of HRP, as a result of more protein being present on the membrane.

2.9 Self-Renewal Assay and Alkaline Phosphatase Staining

mESCs containing one of either the H3.3 mutants (K36M, K36I, K9M, K9I), H3.3, or Nanog, were plated at a clonal density of 1000 cells per well of a 6 well plate (in triplicate). The assay was repeated 3 times for each cell line.

The cells were cultured for 6 days in standard ES media with or without LIF (1000 U/mL), and with or without Dox (1000 U/mL). The media was changed every day. The ESCs were plated onto wells coated with 0.1% gelatine only (gelatine pre-plated overnight), with no feeder cells in order to further stress the ESCs towards differentiation. The reason for stressing the cells towards differentiation is to highlight that the mutations are capable of preventing the ESCs from exiting the pluripotent state, if this in indeed the effect that the transgenes are having.

The Nanog over-expression cell line was used as a positive control showing the ideal situation of cells being held in a pluripotent state in the absence of LIF, and to ensure that the assay was working. Nanog was selected as it is known to be a key mediator of the acquisition of pluripotency; embryonic and also induced pluripotency[128], has been shown to be sufficient in driving self-renewal of undifferentiated ESCs in a cytokine-independent manner[129], and Nanog is also able to promote the transfer of pluripotency after ESC fusion[130].

The over-expression of wild type H3.3 line was used as a negative control in order to show that the increased amount of histone H3.3 does not have any impact on the methylation of H3.3K36 or H3.3K9.

After 6 days of culture, the media was removed, the cells were washed in 1X PBS and stained using VECTOR Red Alkaline Phosphatase (Red AP) Substrate kit (in solution with 125 mM Tris HCl pH8.2, Vector Laboratories). Alkaline Phosphatase (AP), a hydrolase enzyme responsible for the de-phosphorylation of nucleotides and proteins under alkaline conditions, is a marker for pluripotency in all types of pluripotent stem cells; ESCs, embryonic germ cells, and induced pluripotent stem cells[131]. A high level of AP expression is characteristic of pluripotent stem cells, and is present on the cell membrane[132,133].

The plates were then scanned into a computer. The number of AP positive colonies in each well was counted using Fiji (Fiji Is Just Image J) software (parameters: threshold 100, pixel size 9-∞, circularity 0-1). The mean of the number of AP positive colonies was taken for each

technical replicate, for each repetition of the assay, and the subsequent mean was taken for the three repetitions of the assay (n=9).

2.10 Annexin V Staining Apoptosis Assay

In order to assess whether the H3.3K9M mutation was causing apoptosis of the mESCs, an apoptosis assay was performed via Annexin V staining (BD Pharmingen) and analysed by Flow Cytometry using MACSquant® Analyzer (Miltenyl Biotec). Annexin V-PE (hereby referred to as PE) was used to stain cells undergoing apoptosis, and DAPI was used to stain dead cells. The un-induced (-Dox) H3OE cells were used for the following controls: unstained cells (no DAPI or PE), PE only, DAPI only, and a positive control for which the H3OE cells were microwaved for 15 s in order to kill the cells, and stained with both PE and DAPI. The H3OE negative controls were the H3OE cells cultured with or without Dox.

mESCs containing the H3.3K9M transgene, or the H3OE were cultured in the same way as for the self-renewal assay (section 2.9). After 6 days of culture, the cells were trypsinised, and washed twice with cold PBS. The cells were then re-suspended at a concentration of 1 x 10^6 cells/mL in 1X Binding Buffer (10X Binding Buffer: 0.1 M HEPES pH 7.4, 1.4 M NaCl, 25 mM $CaCl_2$. Diluted to 1X prior to use). 100 µl of the suspension (~1 x 10^5 cells) were transferred to a 5 mL culture tube. The Annexin V and Vital Dye were added as follows: 5 µl of PE and DAPI (diluted to 1:1000 prior to use) were added to the relevant samples, and gently mixed. The samples were then incubated for 15 min at room temperature in the dark. 400 µl of 1X Binding Buffer were added to each tube, and analysed by flow cytometry within 1 hour. FlowJo software was used for the analysis of the flow cytometry data. All debris and cell doublets were removed from the data collected by selectively gating only the single cell populations, before the population sizes were measured for analysis. The single cells were selected based on their forward side scatter profiles. Single cells were plotted for DAPI against Annexin to distinguish apoptotic cells from dead cells.

2.11 Embryoid Body Assays

2.11.1 Embryoid Body Culture

mESCs (H3.3 mutants and H3.3 overexpression cells, see 2.1.1, Table of Cell Lines) were expanded in 1 well of a 6-well plate on MEFs in ES media until confluent (approximately 3 days). The mESCs were pre-plated 30 min before EB preparation in order to separate the mESCs from the MEFs. The MEFs would settle and adhere to the culture plate surface within 30 min, leaving the detached mESCs in the supernatant. To pre-plate, the mESCs were trypsinised and re-suspended in a small volume of embryoid body (EB) media (Iscove's Modified Dulbecco's Medium (IMDM, Gibco™ ThermoFisher), FBS (Hyclone), GlutaMAX (Gibco™ ThermoFisher), Monothioglycerol (Sigma-Aldrich), Iron-Saturated transferrin, Ascorbic Acid (Gibco™ ThermoFisher), Non-essential amino acids, Sodium Pyruvate) and incubated for 30 min at 37°C onto plastic only.

The supernatant containing the non-adherent mESCs was collected and the cells were counted. The cells were centrifuged at 1000 rpm for 5 min and re-suspended at a density of 400 cells/ 30 µl in EB media. The cell suspension was divided in half and 1000 U/mL of Dox

were added to the media of one of the suspensions (two media conditions; with Dox and without Dox). An 8 well multichannel pipette was used to make 30 µl drops onto the up-turned lid of a low attachment 15 cm dish. The dish was replaced onto the lid and the dish was carefully flipped over to create hanging drops. For each culture condition (with or without Dox) two dishes were set up for the day 3 time point (to ensure there was enough material to produce a sufficient yield of RNA), one for day 6, and two for day 9 (one for reverse transcription quantitative PCR (RT qPCR) analysis, one for Western Blot analysis).
The hanging drops were incubated at 37°C, 5% CO_2 for 3 days. After 3 days, the EBs were harvested with 10 mL of 1X DPBS per plate using a 25 mL pipette to reduce shear stress on the EBs, and collected into one 50mL conical flask per plate. The EBs were then centrifuged at 200 rpm for 5 min. The centrifugation speed was lower than that used in regular cell pelleting to reduce the risk of breaking up the EBs. The EBs were re-suspended gently in 10 mL of EB media and transferred to non-adherent 10 cm dishes (except for those harvested from two plates for the Day 3 time point).
The dishes were incubated at 37°C and 5% CO_2 on a shaker at low speed (~ 100 rpm) for 6 days. The media was changed every other day, with Dox being included in the media for the +Dox EBs for the entire period to ensure continuous expression of the mutations.

Photographs of the EBs were taken The EBs were collected at four different time points for RT qPCR analysis; at Day 0 the remaining ESCs from the hanging drop preparation were collected, Day 3 (two plates of EBs were harvested), Day 6, and Day 9. Two plates were collected separately for the Day 9 time point; one for Western Blotting, and one for qPCR). The EBs were collected by using a pipette to transfer the media containing the EBs to a conical and centrifuging at 200 rpm to pellet the EBs. The media was removed and the EBs were washed two times with 1X DPBS. The EBs were pelleted as before, and the PBS removed. The EBs were stored at - 80°C.

2.11.2 Reverse Transcription Quantitative PCR

Reverse Transcription Quantitative PCR (RT qPCR) was performed on RNA isolated from the EBs in order to identify the markers expressed at different time points during the EB assay.

RNA Isolation and cDNA Synthesis

RNA was isolated from the EBs using the QIAGEN RNeasy Kit 250 as per the manufacturer's instructions. All reagents named (unless otherwise stated) were included pre-prepared in the RNeasy kit.
Prior to using the kit, all work surfaces, pipettes and tube racks were cleaned with ethanol and RNase*Zap*® Decontamination Solution (ThermoFisher Scientific), to remove any RNases.
The harvested EBs (see section 2.11.1) were re-suspended in 350 µl of Buffer RLT (with added β-mercaptoethanol). The EBs were then syringed up and down several times using a 1 mL syringe and 11.5 gauge needle, until the EBs had been lysed. The lysate was then transferred into a QIAshredder spin column placed in a 2 mL collection tube, and centrifuged for 2 min at full speed.

1 volume of 70% ethanol (v/v diluted with UltraPure™ Dnase/Rnase-Free Distilled Water) was added to the lysate in the collection tube to precipitate the nucleic acids, and pipetted to mix. The sample, and any precipitate formed, was transferred to an RNeasy MinElute spin column placed in a new 2 mL collection tube. The lid was carefully closed and the column was centrifuged for 15 s at ≥10,000 rpm. The flow-through was discarded. 350 µl of Buffer RW1 was added to the column and centrifuged for 15 s at ≥10,000 rpm. The flow through was discarded. This step was repeated once more. The flow through and collection tube were discarded. The column was placed into a new 2 mL collection tubed, 500 µl of Buffer RPE was added to the column, and the column was centrifuged for 15 s at ≥10,000 rpm. The flow through was discarded. 500 µl of 80% ethanol was added to the column, centrifuged for 2 min at ≥10,000 rpm, and the flow through discarded. With the lid open, the column was centrifuged at full speed for 5 min. The column was placed into a new 1.5 mL collection tube, 30 µl of RNase-free water was added directly to the centre of the spin column membrane, and centrifuged for 1 min at full speed to elute the RNA.
The yield of RNA was measured using a NanoDrop spectrophotometer.

cDNA was synthesised from the RNA collected using the SuperScript IV kit (Thermo Fisher) for use in Quantitative Real Time PCR. 1 µg of RNA was used per reaction, and combined in a PCR tube with 1 µl of 50 µM Oligo $d(T)_{20}$ primer, 10 mM dNTP mix (10 mM each) and the reaction mix was made up to 13 µl with nuclease free water. The RNA-primer mix was heated to 65°C in a thermocycler for 5 min and then incubated on ice for 1 min to anneal the primer to the template RNA. The reverse transcription mix, consisting of 4 µl of 5X SSIV Buffer, 1 µl of 100 mM DTT, 1 µl of RNaseOUT™ Recombinant RNase Inhibitor and 1 µl of SuperScript® IV Reverse Transcriptase (200 U/µl), was added to the reaction mix. The cDNA was synthesised by incubating the combined mix at 52°C for 10 minutes in a thermocycler, and subsequently inactivating the reaction by incubating it at 80°C for 10 min. Any remaining RNAs were removed by adding 1 µl of *E. coli* RNase H and incubating the mix at 37°C for 20 min. The cDNA was stored at - 20°C until required for use in qPCR.

Quantitative Real Time PCR (qPCR)

Quantitative real time PCR (qPCR) is a technique developed by Heid *et al*[134] used to quantify the amount of gene expression in real time, by measuring accumulation of PCR products by using a fluorescent DNA binding dye such as SYBR® green, which emits light as a result of being broken down by Taq polymerase. As the PCR progresses, the SYBR dye is incorporated into the amplified DNA, causing an increase in the amount of fluorescence. The increase in the amount of fluorescence is measured throughout the cycle, allowing the calculation of the amount of DNA that has been amplified in real time. A threshold in the amount of fluorescence measured is achieved once enough target DNA has been amplified. The C_t value is the cycle number at which this threshold is reached. This value is used in the calculation of the relative expression level of the genes under investigation. The relative expression of a gene is measured by using C_t value of a housekeeping gene (genes that are involved in the maintenance of basic cellular functions and are thus expressed at a constant level in all cells under normal and pathophysiological conditions) and the C_t value of the target gene.

For this investigation, Brilliant III SYBR® Master Mix (Agilent Technologies) was used in a reaction mix consisting of 1µl of cDNA (100 ng/µl) and 10 mM of forward and reverse primers. The reaction mix was made up to the desired reaction volume with nuclease free water.
All samples were loaded in triplicate into a 384 well plate, sealed with an optical film (Bio-Rad) and centrifuged at 1000 rpm for 30 s. The plate was then placed into a LightCycler® 480 (Roche Life Science) machine.
The qPCR programme used was:

1. 95°C for 5 min
2. 95°C for 20 min
3. 60°C for 20 min
4. 72°C for 20 min
5. Repeat from step 2 for 40 cycles

The melting curve selected for the programme was 65 - 95°C, with increments of 0.5°C.

Primer Efficiencies

Primer efficiencies were calculated by testing primers at different dilutions (1:10, 1:100, 1:1000, and 1:10000) and the C_T values were collected. The C_T values were plotted onto a graph and the slope of the linear regression model of the data points was used in the calculation of the primer efficiency:

$$E = 10^{\frac{1}{\Delta CT}}$$

$\Delta C_T = C_T$ reference gene - C_T target gene

Relative Expression Calculation

The relative expression levels of the genes investigated were calculated using a modified version of the Livak and Schmittgen *et al* $2^{-\Delta\Delta CT}$ method[135], using Actin as the reference housekeeping gene:

$$Relative\ target\ yield = \ 2 - \Delta\Delta CT$$

C_T = cycle number at which the fluorescence threshold was reached
$\Delta\Delta$ = difference between C_T values of reference gene and target gene
$\Delta\Delta C_T = \Delta C_T$ target gene - ΔC_T reference gene

2.12 Reprogramming Assay

A cross between the mouse strains H3.3K36M/rtTA and OKSM/rtTA (see section 2.3, Table 3) was set up with the aim of producing MEFs containing both H3.3K36M and OKSM, along with rtTA. The resulting MEFs were to be used in a reprogramming assay in order to analyse the effect of the H3.3K36M mutation on the reprogramming of somatic cells containing the transgenic H3.3.

The MEFs were derived and expanded in culture as described in section 2.5 (Derivation of Murine Embryonic Fibroblasts (MEFs)) and genotyped and tested for mycoplasma contamination as described in section 2.7 (Genotyping and Mycoplasma Testing).

3 Results

3.1 Expression of H3.3 mutants K36I/M and K9I/M causes a global reduction in H3.3K36 and H3.3K9 tri-methylation

SDS-PAGE and Western Blot analysis were performed on transgenic mESCs (see 2.1.1 Table 2) which had been cultured for 4 days in standard ESC culture conditions in the absence of LIF, and with or without Doxycycline (Dox). Immunostaining of the Western Blots (Figure 2) was used in order to determine that the expression of the substitution mutations H3.3K36I/M and H3.3K9I/M in our mESCs re-produced the effect of wiping out the H3.3K36 and H3.3K9 tri-methylation as previously demonstrated by Lewis *et al*[2], in the case of the H3.3K27I/M substitutions and the resulting loss of the tri-methylation on H3.3K27.

As shown in Figure 2, the overexpression of histone variant H3.3 (H3.3OE) has no effect on the methylation state of either H3.3K36me3 or H3.3K9me3. However, the substitution of Histone H3 Lysine 9 (H3.3K9) or Histone H3 Lysine 36 (H3.3K36) with Isoleucine or Methionine is shown to cause a global eradication of the tri-methylation of H3.3K9 or H3.3K36, respectively.

As well as this loss of K36me3 in cells with the K36I/M substitution, and K9me3 in K9I/M substituted cells, the immunoblots also show a slight reduction in K9me3 in the cells expressing H3.3K36I/M, and the same reduction of K36me3 in the cells expressing H3.3K9I/M.

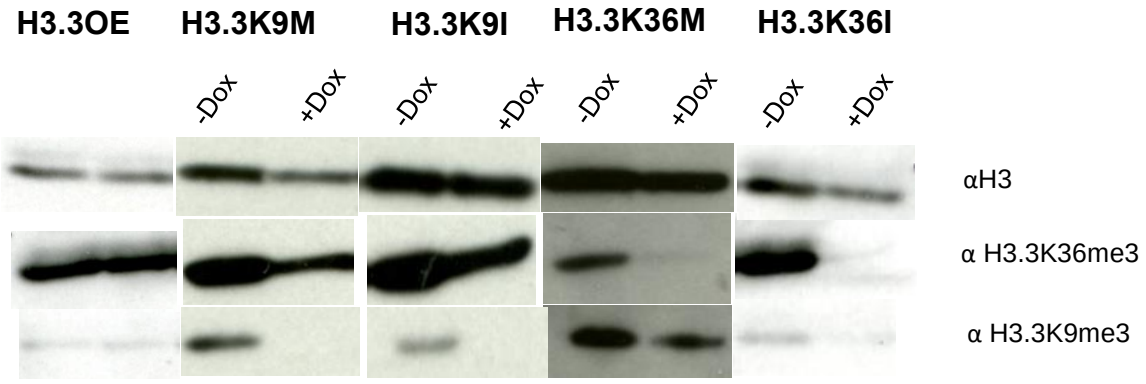

Figure 2. Transgenic H3.3 mESCs cultured with or without Dox for 4 days in the absence of LIF. Immunostained for H3.3K36me3, H3.3K9me3, and for H3 as the loading control.

3.2 Alkaline Phosphatase staining shows H3.3 mutations do not hold mESCs in a pluripotent state

In order to observe the effect of these Histone H3.3 mutations on the self-renewal ability of mESCs, the cells were cultured at a clonal density on a gelatine coated 6 well plate, in four different media conditions: +LIF +Dox, +LIF –Dox, -LIF +Dox, and –LIF –Dox. The cells were then stained with Alkaline Phosphatase after 6 days of culture, in order to determine the number of cells which remained in the pluripotent state. The number of AP positive cells was measured by using FIJI (FIJI is just Image J) software to count the red, AP positive cells. The significance of the results was then determined through using Graphpad Prism to perform t-tests in order to compare the cells with and without induction of the mutation for both culture conditions (with and without LIF). A p value less than 0.05 ($p < 0.05$) is considered to be significant.

Alkaline phosphatase (AP) staining of the cells after 6 days of culture revealed that the H3.3 mutations did not cause the mESCs to be held in the pluripotent state.

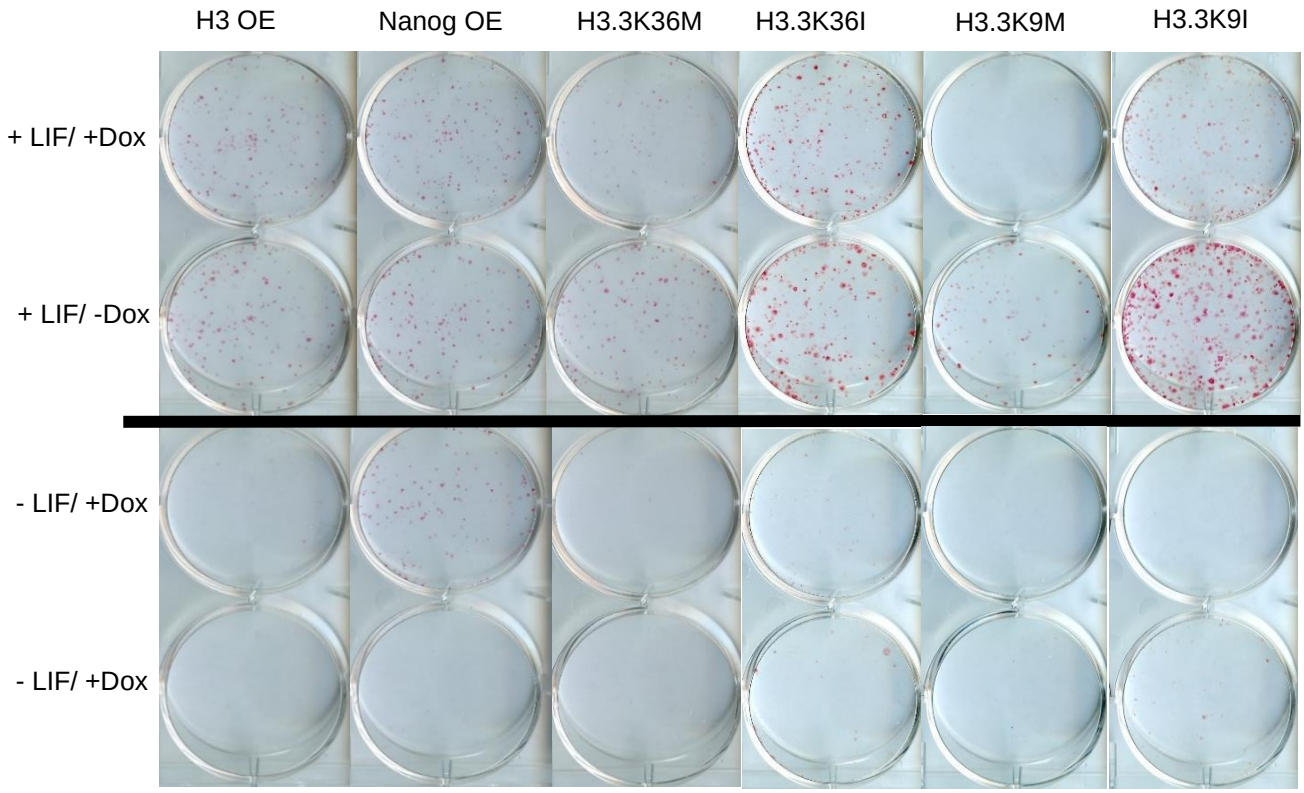

Figure 3 Alkaline phosphatase staining of mESC clones expressing the Histone H3.3 mutations (H3.3K36I/M, H3.3K9I/M) in 4 ES media culture conditions.

The H3OE line (used as a negative control) was used to demonstrate that the removal of LIF results in the differentiation of the mESCs. These results show that the overexpression of the histone variant H3 does not have an effect on the differentiation behaviour of mESCs. This can be seen in Figure 3, which shows no difference in the number of AP positive colonies between those cultured with or without Dox in the presence of LIF, and no difference between those cultured with or without Dox in the absence of LIF. This is shown quantitatively in Figure 4, where the bar graphs show that the means and standard deviation between the replicates were almost identical. A t-test performed to measure the difference between the number of AP positive colonies between the induced and un-induced transgenes within each culture condition (with or without LIF) produced a p value of > 0.05 for both conditions. This confirms that there is no significant difference in the number of AP positive cells between the two conditions.

These assays show that only the mESCs in which Nanog was overexpressed were held in a pluripotent state in the absence of LIF. This is can be seen clearly in Figure 3, and is shown quantitatively in Figure 4 B. The t-test performed produced a p value of > 0.05 between the induced and un-induced cells cultured with LIF, showing that there was no significant difference in the number of AP positive colonies. A p value of < 0.05 ($p < 0.0001$) showed a significant difference in the number of AP positive colonies between the induced and un-induced cells cultured in the absence of LIF.

Figure 4 C- F shows fewer AP positive mESCs in both conditions (+LIF, -LIF) for all transgenic H3.3 cell lines where the mutations were induced (+Dox), demonstrating that they are not being held in a pluripotent state by the expression of the H3.3 transgenes.
The minor exception to this are the cells expressing H3.3K36I, which do not show a perceivable difference in the number of AP positive colonies between those cultured with and without Dox in the presence of LIF (Figure 4 D). The t-test produced using the mean number of AP positive colonies taken for each technical replicate, and subsequently for each repeat, resulted in a p value of > 0.05 for the H3.3K36I cells cultured in LIF, confirming that there is no significant difference between the number of AP positive colonies between the un-induced (-Dox) and induced (+Dox) H3.3K36I lines. However, more variation in the number of colonies remaining pluripotent was seen between the repeats of the assay in the condition +LIF +Dox.

In the absence of LIF, marginally more AP positive colonies can be seen in the induced H3.3K36I line (Figure 4 D). The t-test produced a p value of > 0.05, which showed that this difference between the un-induced and induced H3.3K36I lines was not significant.

The t-test performed for both transgenic H3.3K9 lines showed no significant difference in the number of AP positive cells between the induced and un-induced cells cultured with LIF, with a p value of > 0.05. There was found to be a significant difference between the induced and un-induced cells lines cultured without LIF for both H3.3K9I and H3.3K9M, with a p value of < 0.05.

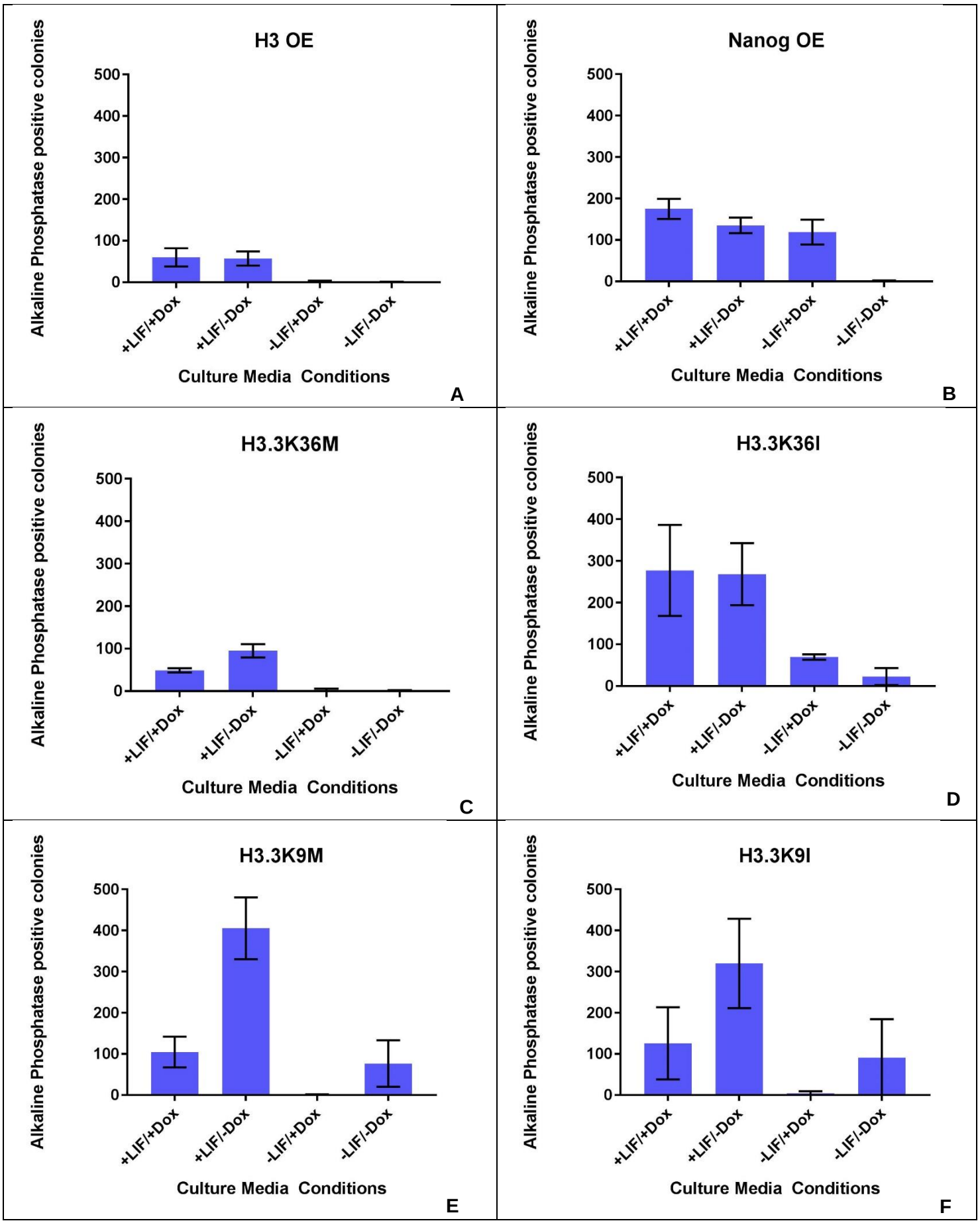

Figure 4 AP staining of mESC clones expressing the Histone H3.3 mutations H3.3K36I/M, H3.3K9I/M in 4 ES media culture conditions (n=9).

3.3 Annexin V staining of H3.3K9M showed no effect on the amount of apoptosis occurring

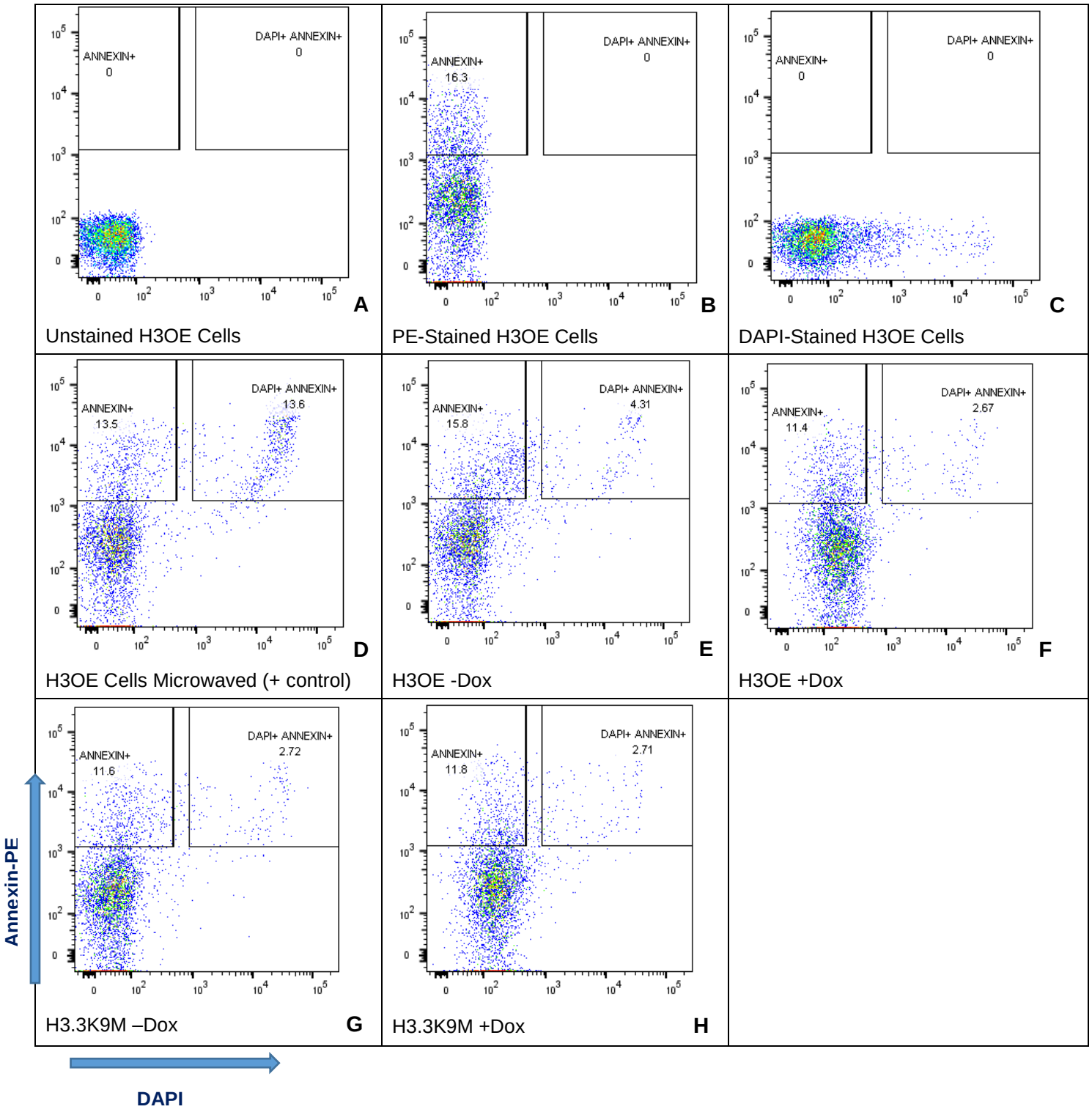

Figure 5 Flow cytometry analysis of H3.3K9M apoptosis assay using FlowJo software. Annexin V PE dye stains apoptotic cells, DAPI stains dead cells. H3OE cells were used for the controls.

Annexin V staining was used to perform an apoptosis assay to determine whether the H3.3K9M mutation is fatal to mESCs. H3.3K9M or H3OE expressing mESCs were cultured for 6 days –LIF, on gelatin only. The cells were then stained with Annexin-PE and DAPI, and the results were analysed via flow cytometry. The single cells were selected based on their forward side scatter profiles. Single cells were plotted for DAPI against Annexin to distinguish apoptotic cells from dead cells. This was done using FlowJo software.
Apoptotic cells are positive for PE and negative for DAPI, since DAPI can only stain dead cells due to their membranes not being intact, and the cell membranes in apoptotic cells is still intact.
The H3OE cell line was used for the following controls; unstained cells, PE stain only, DAPI stain only, the positive control (cells killed by microwaving, and stained for both DAPI and PE), and the negative controls which were H3OE cells cultured either with or without Dox, to be the reference point for the H3.3K9M +/- Dox conditions.
DAPI staining was used to differentiate cells undergoing apoptosis from dead cells. PE positive and DAPI negative cell populations are those undergoing apoptosis. This apoptotic population can be seen in the top left corner of the graphs, labelled Annexin+. The DAPI positive and PE negative population, seen in the bottom right area of the graphs, contains the population of dead cells. Cells which are double positive for DAPI and PE are dead, and are seen in the top right area of the graphs, labelled DAPI+ Annexin+. The population which is negative for both PE and DAPI are cells which are alive, seen in the bottom left area of the graphs.

The flow cytometry analysis of the results, shown in Figure 5, reveal that all of the controls showed the expected results; the unstained H3OE control (Figure 5 A) only has one population, which is double negative for both PE and DAPI. The PE only control (Figure 5 B) showed no cells which were positive for DAPI, and 16.3% of the total cell number stained positive for PE. The DAPI only control (Figure 5 C) showed no PE positive cells, and a small number of DAPI positive cells. This shows that there were very few dead cells, however the difference between dead and alive cells can be seen clearly. The positive control of microwaved H3OE cells stained for both PE and DAPI (Figure 5 D) showed a population of 13.5% of cells which were positive for PE. This population is undergoing apoptosis. 13.6% of cells were positive for both DAPI and PE, which means that this population consists of dead cells. There were no cells which were positive for DAPI and negative for PE.
The negative controls for which H3OE cells were cultured without and with Dox (Figure 5 E and F, respectively) showed very little difference between the size of the population of PE positive, apoptotic cells between the un-induced (-Dox) and induced (+Dox) overexpression of histone H3. The un-induced H3OE line showed an apoptotic population size of 15.8% (of the total cell number), compared with a population of 11.4% in the induced H3OE line. The population of dead cells, which were positive for both DAPI and PE, were also of similar size between the un-induced and induced H3 overexpressing cells, with a population size of 4.31% in the un-induced line, and 2.67% in the induced line.
The H3.3K9M cells showed no real difference between the sizes of the population of apoptotic cells (PE positive), or the population of dead cells (DAPI+ and PE+) between the un-induced and induced cells (Figure 5 G and H, respectively). The population sizes of

apoptotic cells were 11.6% in the un-induced H3.3K9M cells, and 11.8% in the induced H3.3K9M cells. The sizes of the population of dead cells were 2.72% in the un-induced cells, and 2.72% in the induced cells.

No significant differences can be seen between the population sizes of either the apoptotic cells (PE positive) or the dead cells (DAPI and PE positive) between the H3OE negative control cells and the H3.3K9M cells, for either the un-induced (-Dox) or the induced (+Dox) groups. The un-induced H3OE cells do however show marginally larger populations of apoptotic, and dead cells than the other 3 groups (H3OE +Dox, H3.3K9M –Dox and H3.3K9M +Dox).

When induced by Dox, the entire cell populations appear to shift to the right of the graph, resulting in the populations appearing one log higher on the DAPI axis than the un-induced populations (-Dox), for both cell lines, H3OE and H3.3K9M.

There also seems to be an upward shift in PE negative cells in the samples which were stained with the Annexin-PE. This can be seen upon comparison of samples stained with Annexin-PE with the unstained cells (Figure 5 A, C), which shows the unstained cell population to be lower on the Annexin-PE scale than the negative population on the graphs in which PE had been added to the samples (Figure 5 B, D, E, F, G and H)

3.4 Transgenic H3.3-expressing embryoid bodies show a size phenotype

In order to quantify the effect of the transgenic H3.3 expression on differentiation, the mESCs containing the H3.3K36I/M or H3.3K9I/M mutation were used to produce embryoid bodies (EBs) by using the hanging drop formation method (see Materials and Methods chapter 2.11.1 Embryoid Body Culture). After 3 days of hanging drop culture, the EBs were transferred to non-adherent culture dishes and incubated on a shaker for 6 more days to allow the EBs to differentiate.

The EBs were cultured in EB media with or without Dox from the beginning (the day of the hanging drop preparation), in order to ensure continual expression of the transgenes throughout both the formation and culture period of the EBs.

The diameter of the EBs was measured by using photographs of the EBs and ImageJ software. The mean values were used in the t-test calculation to compare the sizes between the induced (+Dox) and un-induced (-Dox) EBs for each condition. A p value of < 0.05 is accepted as being significant.

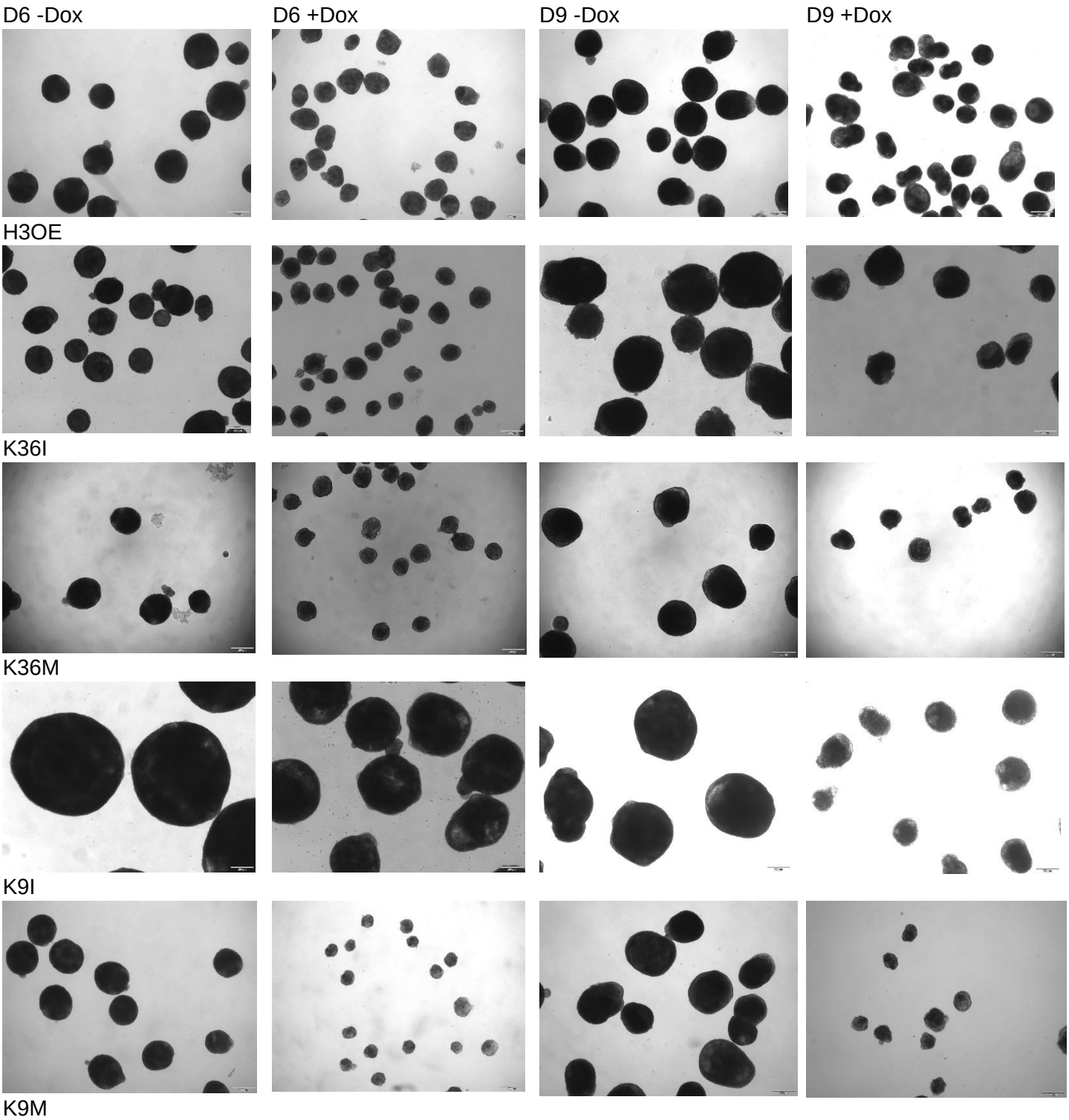

Figure 6 Embryoid bodies formed from mESCs expressing H3.3 transgenes H3.3K36I/M or H3.3K9I/M cultured in embryoid body media +/- Dox. Representative images taken at time points Day 6 and Day 9, scale bar = 200µm.

Figure 6 depicts 4X magnified photographs of EBs at the day 6 and day 9 time points, with and without Dox. The H3OE EBs (negative control) display a slight, but not significant ($p > 0.05$), difference in size between the EBS at both time points, with and without Dox.

All of the transgenic H3.3-expressing EBs show a phenotype of smaller EBs when induced (throughout the culture period) with Dox (Figure 6).

Both H3.3K36M and H3.3K36I show significantly smaller EBs when the mutations are induced, with p values of < 0.05 for both day 6 and day 9 with H3.3K36M, and $p < 0.05$ at day 9 with H3.3K36I (Figure 7). The difference was not significant between the un-induced and induced H3.3K36I EBs at day 6, with a p value of > 0.05.

The biggest difference in size can be seen in both of the K9 mutants (K9I and K9M); the H3.3K9M EBs are significantly smaller than their non-induced control (Figure 7) with analysis via t-test showing a p value of < 0.05 for the EBs at day 6, but the p value at day 9 ($p > 0.05$) showed that the size difference at day 9 was not significant.

The H3.3K9I EBs show a significant difference in size, with a p value of < 0.05 on day 6. The K9I EBs are all significantly larger than all other EBs of each genotype. The K9I EBs without Dox are much larger than all other EBs in the absence of Dox, whereas all other un-induced transgenic EBs are approximately the same size as each other and the H3OE (Wild Type) EB control cells.

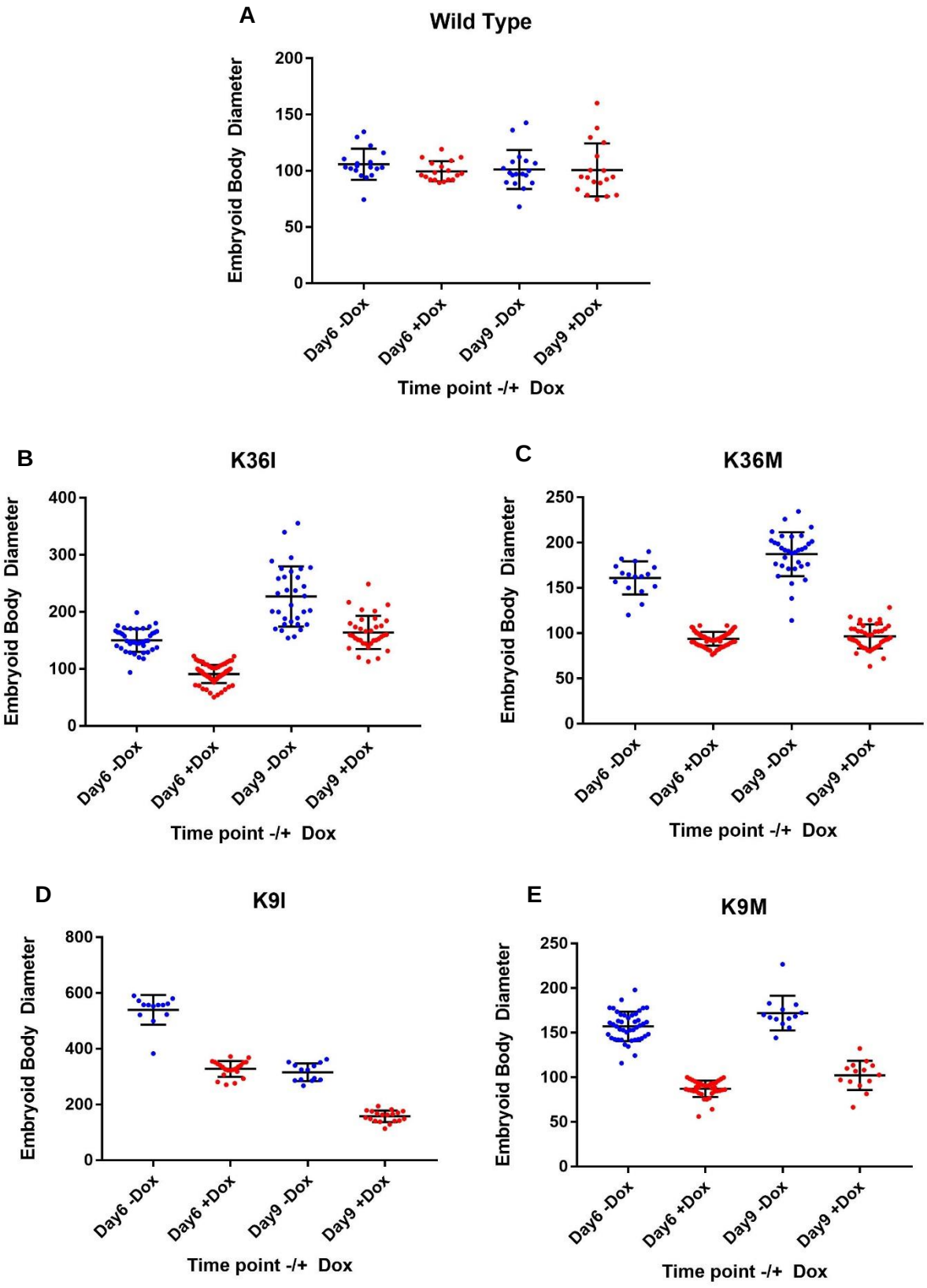

Figure 7 Size analysis of embryoid bodies. Diameters of EBs measured at day 6 and day 9 +/- Dox using ImageJ

3.5 It is unclear whether the EBs are differentiating more or less as a result of the H3.3 transgenes

In order to identify and compare the markers expressed within the embryoid bodies at different time points throughout the EB assay, RNA was isolated from the EBs at 4 different time points; day 0, day 3, day 6, and day 9. The RNA was then used to synthesise the cDNA, via reverse transcription, to be used in the quantitative real time PCR. (RT-qPCR).
The RT-qPCR allowed the analysis of the gene expression profiles of each transgenic H3.3 line (H3OE, H3.3K36I/M, and H3.3K9I/M) both with and without induction of the mutations in order to analyse the effect the induction of the transgenic H3.3 genes had upon the differentiation ability of the EBs, by measuring the expression of markers of pluripotency over the period of the assay.
The qPCR was performed by using the fluorescent DNA binding dye SYBR green, in order to use the changes in the amount of fluorescence produced to calculate the relative expression of the pluripotency markers Oct4, Sox2 and Nanog in comparison to the house keeping gene Actin. The relative expression levels of the genes investigated were calculated using the $2^{-\Delta\Delta CT}$ method, using the C_T values of the gene of interest and the house keeping gene. The relative expression of all genes was normalised to day 0, since that is the time point at which the cells are the least differentiated, and are still in the pluripotent state (i.e. still mESCs).The assay was repeated 3 times.

Figure 8 shows the relative expressions of the pluripotency markers Oct4, Sox2 and Nanog normalised to Actin, for each transgenic H3.3 line, over the course of 9 days. Figure 8 A shows the negative control; the H3.3 overexpression line both with and without Dox. The graph shows what appears to be an initial three-fold increase in the relative expression of Oct4, and a two-fold increase in Nanog expression for both the induced and un-induced lines at day 3, and a slight decrease in the relative expression of Sox2 in the un-induced line. The relative expression levels of all genes can then be seen to drop by day 6 in both the induced and un-induced lines, with a marginally smaller decrease in expression of the 3 genes in the induced line. By day 9, the expression levels of all of the genes appears to increase again, more so in the un-induced line.
Figure 8 B shows the H3.3K36I line. At day 3, a slight increase in the relative expression of Oct4 can be seen in both the induced and un-induced lines, whereas the expression of both Sox2 and Nanog appear to drop to half of the starting expression level, in both conditions. Day 6 –Dox appears to show a complete loss in expression of the pluripotency markers, and day 6 +Dox shows only a marginal reduction in the relative expression for all 3. The expression levels appear to increase again at day 9 for Sox2 in the un-induced line, and for all 3 markers in the induced line.
Figure 8 C shows the H3.3K36M line, with a general trend of a large reduction in the expression of all three markers by day 6, and a subsequent smaller increase after day 6. At day 3, the relative expression of Oct4 increases by 0.5 fold in the un-induced line, and 2.5 fold in the induced line. The relative expression for the other two markers decrease slightly in both lines. Day 6 shows the loss of pluripotency in the un-induced line, a small decrease in the expression levels of Sox2 and Nanog, and a large decrease in the expression of Oct4.

At day 9, the relative expression of all three markers increases; a negligible increase in the un-induced line, and a larger increase in the induced line, almost returning to the level seen at day 0 for Sox2 and Nanog. Oct4 expression marginally increases.

Figure 8 D shows the H3.3K9I line. The general trend in the relative expression for all three markers is a decrease in the expression levels until day 6, and a slight increase in the expression of Sox2. Day 3 shows a large reduction in the relative expression of Sox2 in the un-induced line, a smaller reduction in the Sox2 expression can be seen in the induced line. The expression of Oct4 and Nanog both drop in both conditions; the largest decrease is in the level of Nanog expression in the induced line. There is an overall larger reduction in relative expression of all three markers in the induced line at day 3. Day 6 shows a loss of expression for all three markers in the un-induced line, and a marginally higher level of expression of the three markers in the induced line. Day 9 shows a slight increase in the relative expression of Sox2 in the un-induced line, and a larger increase in the induced line.

Figure 8 E shows the H3.3K9M line. The general trend in the relative expression is the initial loss in expression of all three pluripotency markers by day 6, and a subsequent increase in relative expression after day 6. Day 3 shows a very slight drop in Oct4 expression in the un-induced line, and a reduction of ~0.75 fold in Sox2 and Nanog when un-induced. ~0.75 fold decrease in Oct4 and Sox2 expression is seen at day 3 in the induced line, and the largest reduction in expression is Nanog in the induced line. Day 6 shows an almost complete loss in all three markers in the un-induced line, a small drop in Oct4 expression in the induced line, and a marginal increase in relative expression of Sox2 and Nanog in the induced line. There is an increase in all three markers after day 9 in both the induced and un-induced line.

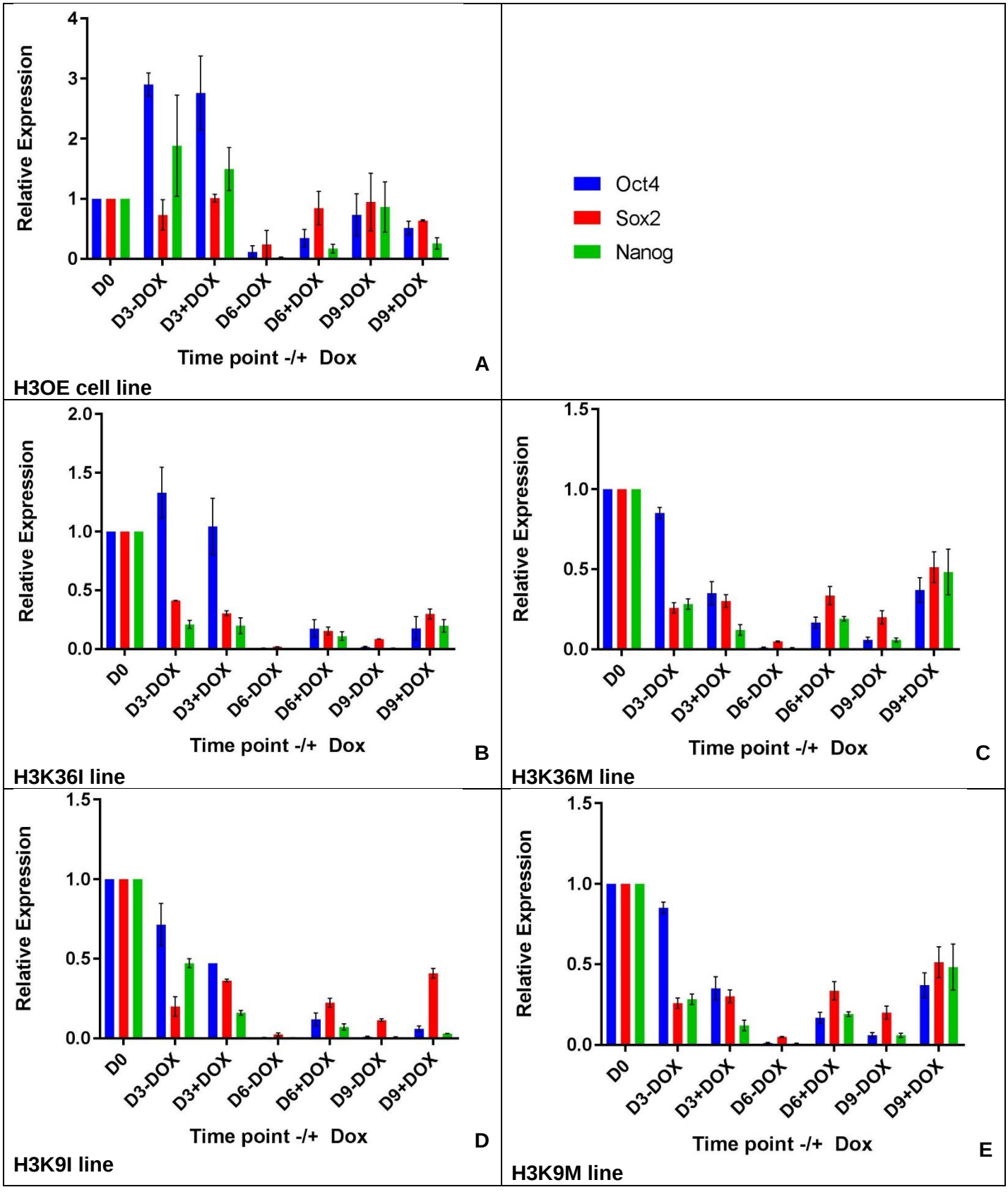

Figure 8 qPCR analysis of the relative expression of the pluripotency markers Oct4, Sox2 and Nanog in EBs cultured +/- Dox, over a period of 9 days.

4 Discussion

The aim of this investigation was to determine the effect on pluripotency of the substitution of amino acids at critical positions within the tail of histone H3.3. The lysine 9 (K9) or lysine 36 (K36) in the tail of the histone variant H3.3 were substituted in murine embryonic stem cells for either methionine or isoleucine, producing 4 different cell lines; H3.3K36I, H3.3K36M, H3.3K9I, and H3.3K9M (see Table 2 section 2.1 of Materials and Methods).

As discussed in the introduction of this paper, a high percentage of certain cancers, such as chondroblastomas, giant cell tumour of bone, and paediatric gliomas (DIPGs and GBM), were found to express mutations in the genes encoding the histone variant H3.3. These mutations resulted in amino acid substitutions at critical locations in the histone tail[1–3].

These mutations were shown to cause a global reduction in di- and tri-methylation of both the residues which had been mutated, and of the non-mutated wild-type residues throughout the cells, along with a moderate increase in the acetylation of these same residues.

Due to the similarities in characteristics between cancerous cells and pluripotent cells (outlined in the introduction, section 1.2), this investigation was undertaken to assess whether there may be a link between these K-to-I/M substitutions in histone H3.3 and pluripotency.

In order to analyse the effect of these substitutions on pluripotency, a variety of experiments were performed using mESCs expressing the different transgenic histone H3.3; self-renewal assays were used to assess whether or not the transgenic mESCs were held in the pluripotent state after being cultured in conditions designed to induce differentiation. Differentiation potential of the transgenic cells was measured by growing embryoid bodies (EBs) from the cells, and analysing their gene expression profiles over various time points by using RT-qPCR.

4.1.1 Expression of transgenic histone H3.3 reduces the tri-methylation of lysine residues 36 and 9

Immunostaining of Western Blots performed on the mESCs expressing one of the transgenic histone H3.3 (H3.3K36I, H3.3K36M, H3.3K9I, or H3.3K9M) showed that these mutations are having the same effect upon the tri-methylation of H3.3K36 and H3.3K9, respectively, as previously demonstrated by Lewis *et al* in their study of the substitution of H3.3K27 with isoleucine or methionine[2].

As shown in Figure 2 (results section 2.8), the substitution mutations H3.3K36I/M and H3.3K9I/M globally wipe out the tri-methylation marks on H3.3K36 and H3.3K9, respectively. This loss of tri-methylation suggests a similar mechanism as outlined in the case of the H3.3K27M. It was shown that upon interaction of the K27M peptides with the EZH2 active site of the PRC2 complex, a conformational change occurred within the complex, thus preventing its detachment from the mutated peptide, and subsequently preventing the methylation of downstream H3.3K27 residues[2]. Lewis *et al* also demonstrated that histone H3.3K9M peptides inhibited recombinant histone H3.3K9 methyltransferases SUV39H1 and G9a[2], and argued that SET-domain inhibition is the likely cause of the reduction of di- and tri-methylation of H3.3K9 in H3.3K9M expressing cells.

Figure 2 also shows that there seems to be some cross talk between both the mutation of both residues (K36 and K9) and the methylation (or rather the subsequent loss in methylation) on the other residue, along with the expected loss of methylation on the substituted residue. This is shown by the immunoblot for the H3.3K36I and M expressing cells, where a slight reduction in H3.3K9me3 can be seen as well as the total loss of the H3.3K36me3. Additionally, this effect is also shown by the immunoblot for H3.3K9I and M expressing cells, where a reduction in H3.3K36me3 is seen as well as the loss of H3.3K9me3.

This result contrasts the data obtained from Chao Lu *et al*[136], whose work demonstrates that the expression of the H3.3K36M transgene also appeared to cause an increase in H3.3K27me2/3 in various cell types. H3K27me3 and H3K36me2/me3 are rarely seen to co-exist on the same histone H3 polypeptide. H3 histones that are un-methylated at K36 are usually methylated at K27. This suggests that there is crosstalk between the methylation of H3K27 and H3K36 *in vivo*[137]*.* Pre-existing H3K36me inhibits H3K27me via PRC2 *in vivo.* The suggested cause of this concomitant increase in H3.3K27me2/3 is that the methylation of H3K36 and H3K27 negatively regulate each other[137]. Thus the K36M-mediated loss of H3.3K36 methylation could result in the lack of negative regulation upon the methylation of H3K27, resulting in an increase in H3K27me upon the loss of H3K36me. Also, Caho Lu *et al* suggest that the loss of methylation on H3K36 provides new nucleosomal substrates for PRC2, since H3K36me2/me3-containing nucleosomes are poor substrates for PRC2[136].

4.1.2 Self-renewal ability is not maintained by the expression of transgenic histone H3.3

Self-renewal ability of mESCs expressing the transgenic H3.3 was assessed by performing a self-renewal assay on cells, under various culture conditions. The culture conditions were designed to stress the mESCs towards differentiation by growing the cells at a clonal density, without MEFs, and without LIF. Alkaline phosphatase staining was used to identify the number of cells which had remained in the pluripotent state after several days of culture.

The H3 overexpressing (H3 OE) mESCs were used as a negative control in order to ascertain that it is not the overexpression of the Histone H3 which is causing any potential effects on the cell's behaviour. It has previously been shown in the yeast Saccharomyces cerevisiae that excess levels of histones can cause severe side effects within a cell, such as excessive chromosome loss, resulting in genomic instability, higher sensitivity to DNA damaging agents, and cytotoxicity[138]. Other investigations found that the overexpression of histone H3 interferes with the deposition of the histone H3 variant CENP-A, which is centromere specific, causing a loss in chromosome-phenotype[139,140]. In this investigation, the H3 OE cells did not seem to cause any adverse effects upon the cells when the overexpression was induced; Figure 4 A showed no significant difference in the number of alkaline phosphatase (AP) positive cells in the presence of LIF, and with or without Dox. Upon the removal of LIF, there were very few colonies which were stained positive for AP both with and without Dox. This clearly demonstrates that the overexpression of Histone H3 is having no effect upon the cells in terms of differentiation, and so any differences in cell

behaviour when the histone H3.3 K36 or K9 transgenes are expressed can be accredited to the effect of the substitution of the K36 or K9 with I or M.

Nanog overexpressing (Nanog OE) mESCs were used as a positive control. Nanog was selected for this purpose as it is known that Nanog expression is crucial to the pluripotent state. The prediction was that the overexpression of Nanog would be sufficient in order to hold the cells in the pluripotent state. Thus Nanog was selected to ensure that the assay was working. Figure 3 shows that the Nanog OE indeed has a strong effect, as it is the only line which holds the mESCs in a pluripotent state after the removal of LIF from the culture media.

Figure 4 B shows a higher number of AP positive cells in the presence of LIF and Dox (when Nanog is over expressed) in comparison to those without Dox (without Nanog overexpression). A significant difference can be seen in the absence of LIF; the overexpression of Nanog holds almost as many cells in the pluripotent state as seen in the presence of LIF without the Nanog OE (+LIF –Dox). In the absence of both LIF and Dox, no colonies are positive for AP, which shows that the removal of LIF causes mESCs to differentiate.

mESC colonies expressing the H3.3 transgene did not remain in a pluripotent state. Very few AP colonies can be observed in the absence of LIF in the mutant cell lines, as shown in Figure 4 C-F. In fact, the mutations seem to be having the opposite effect, since it can be seen that there are fewer AP positive colonies in the lines in which the mutations had been induced (+Dox), even in the presence of LIF.

This could suggest that the K-to-I/M substitutions are driving more rapid differentiation in these cells. However, it could be that the transgenes are potentially effecting the cell cycle in some way, such as by reducing the speed of the cell cycle and proliferation of the transgene-expressing cells.

One additional interesting observation is that during the course of the assay, the colonies formed by the K9M cells in the presence of Dox appeared flatter and sparser than the other colonies. The amount of dead cells floating in the media each day appeared to be larger than that of the other cell lines, however this was not quantified. These results could suggest that the K9M substitution has a cytotoxic effect upon the mESCs expressing the mutation, rather than causing an increase in the differentiation.

4.1.3 H3.3K9M does not seem to have a cytotoxic effect on mESCs

Due to the observed unusual physical appearance of the H3.3K9M colonies produced during the self-renewal assay, and the possible larger number of dead cells present floating in the media, and apoptosis assay was performed on this cell line. In order to assess whether the H3.3K9M substitution was having a cytotoxic effect on these cells, Annexin V staining was used to measure the number of apoptotic cells after 6 days of culture under the same conditions as the self-renewal assay. The results of the Annexin V staining were analysed by flow cytometry (Figure 5).

The results of the Annexin V stain revealed no difference in the number of apoptotic or dead cells in the induced H3.3K9M line, in comparison to the un-induced H3.3K9M line and the negative controls of un-induced and induced H3OE lines. The populations of apoptotic (PE positive) and dead (DAPI positive and PE positive) cells were all the same size for both cell lines, whether induced or un-induced. However, the experiment was only performed one time, and would need to be repeated at least two more times in order to confirm these results.

There were some abnormalities seen in the results; when induced by Dox, the entire cell populations for both cell lines, H3OE and H3.3K9M, appear to shift to the right of the graph, resulting in the populations appearing one log higher on the DAPI axis than the un-induced populations (-Dox). The reason for this shift is unclear and is most likely a result of a technical error at some stage of the experiment or data processing.
There also seems to be an upward shift in PE negative cells in the samples which were stained with the Annexin-PE. This can be seen upon comparison of samples stained with Annexin-PE with the unstained cells (Figure 5 A, C), which shows the unstained cell population to be lower on the Annexin-PE scale than the negative population on the graphs in which PE had been added to the samples (Figure 5 B, D, E, F, G and H). This most likely means that the Annexin is binding non-specifically.

4.1.4 qPCR analysis did not produce any concrete findings

Analysis of the expression of the pluripotency markers Oct4, Sox2, and Nanog in EBs over four time points did not answer the question of whether the EBs were differentiating more or less when expressing the transgenic H3.3 variants. The qPCR data (Figure 8) does show a general decrease in the relative expression of all three markers until day 6, suggesting the differentiation of the EBs. On average, it appears that across all of the transgenic EBs, there is marginally less differentiation occurring in the EBs in which the transgenes are expressed (+Dox). However, the data does not appear to be consistent or reliable, considering the expression levels appear to increase again after day 6. After comparing the trends of the changing relative expression to the C_T values of the house keeping gene Actin, which were used in the calculation of the relative expression of the pluripotency markers, there is an explanation for the unusual appearance of the data trend. The C_T values for Actin expression are inconsistent across the time-points in every line, with the values initially increasing, and then decreasing. Some of the values reached above 27, which should not be the case. Since the C_T values for Actin were going up and down, this would explain the unexpected decrease and increase in relative expression of the genes of interest. Since the same results was obtained when GAPDH was used as the reference gene, the most likely explanation for the unusual changes in the expression levels of all of the genes is not that the relative expression levels of Sox2, Oct4 and Nanog are going, rather that there was low sample mRNA material. A low amount of mRNA would cause an increase in the C_T value, as it would take more PCR cycles to reach the C_T threshold. Since the pluripotency markers were all normalised to Actin, it would appear that the markers were coming up in expression leve.ls again. One thing to note in relation to the re-increase in Sox2 expression levels is that it was expected that Sox2 expression would come up again upon differentiation of the EBs, since Sox2 is expressed in neural progenitor populations[141].
It would definitely be necessary to repeat all of the EB assays and the qPCR, to obtain better sample material, and also select more housekeeping genes, such as HMBS (hydroxymethylbilane synthase), or B-ACT (β-actin), since some may change upon differentiation of the cells.

4.1.5 Reprogramming Assay

Unfortunately, none of the MEFs derived from the OKSM/H3K36M/rtTA cross had the desired genotype of containing all three inserts, so the reprogramming assay could not be performed.

4.1.6 Findings from this investigation in the context of cancer

Although the acetylation state of the residues studied in this investigation were not assessed, it can be assumed, due to findings throughout the relevant literature, that the reduction of the methylation of H3.3K36 and H3.3K9 resulted in a concurrent increase in the acetylation of the same residues. It has been observed that histone acetylation might prevent the folding of nucleosomal arrays into more complex structures that would otherwise be inhibitory[71,74,75]. It could thus be deduced that the decrease in methylation as a results of these substitution mutations, and subsequent increase in acetylation, could cause a removal of complex nucleosomal arrays which control the repression of tumorigenic genes, resulting in the formation of the tumours observed in previous investigations into the mutation of histone variant H3.3.

5 Bibliography

1. Behjati, S. *et al.* Distinct H3F3A and H3F3B driver mutations define chondroblastoma and giant cell tumor of bone. *Nat. Genet.* **45,** 1479–82 (2013).

2. Lewis, P. W. *et al.* Inhibition of PRC2 activity by a gain-of-function H3 mutation found in pediatric glioblastoma. *Science* **340,** 857–61 (2013).

3. Schwartzentruber, J. *et al.* Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. *Nature* **482,** 226–31 (2012).

4. Waddington, C. H. The epigenotype. 1942. *Int. J. Epidemiol.* **41,** 10–3 (2012).

5. WADDINGTON, C. H. Canalization of Development and the Inheritance of Acquired Characters. *Nature* **150,** 563–565 (1942).

6. Allis, C. D. & Jenuwein, T. The molecular hallmarks of epigenetic control. *Nat. Rev. Genet.* **17,** 487–500 (2016).

7. Cooper, G. M. Chromosomes and Chromatin. (2000).

8. Mariño-Ramírez, L., Kann, M. G., Shoemaker, B. A. & Landsman, D. Histone structure and nucleosome stability. *Expert Rev. Proteomics* **2,** 719–29 (2005).

9. Szenker, E., Ray-Gallet, D. & Almouzni, G. The double face of the histone variant H3.3. *Cell Res.* **21,** 421–34 (2011).

10. Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389,** 251–60 (1997).

11. Bernstein, E. & Hake, S. B. The nucleosome: a little variation goes a long way. *Biochem. Cell Biol.* **84,** 505–17 (2006).

12. Happel, N. & Doenecke, D. Histone H1 and its isoforms: Contribution to chromatin structure and function. *Gene* **431,** 1–12 (2009).

13. Tagami, H. *et al.* Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell* **116,** 51–61 (2004).

14. Tropberger, P. & Schneider, R. Scratching the (lateral) surface of chromatin regulation by histone modifications. *Nat. Struct. Mol. Biol.* **20,** 657–661 (2013).

15. Kouzarides, T. *et al.* Chromatin modifications and their function. *Cell* **128,** 693–705 (2007).

16. Arents, G. & Moudrianakis, E. N. The histone fold: a ubiquitous architectural motif utilized in DNA compaction and protein dimerization. *Proc. Natl. Acad. Sci. U. S. A.* **92,** 11170–4 (1995).

17. Baxevanis, A. D., Arents, G., Moudrianakis, E. N. & Landsman, D. A variety of DNA-binding and multimeric proteins contain the histone fold motif. *Nucleic Acids Res.* **23,** 2685–91 (1995).

18. PAULING, L., COREY, R. B. & BRANSON, H. R. The structure of proteins; two hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Natl. Acad. Sci. U. S. A.* **37,** 205–11 (1951).

19. Alva, V., Ammelburg, M., Söding, J. & Lupas, A. N. On the origin of the histone fold. *BMC Struct. Biol.* **7,** 17 (2007).

20. Arents, G., Burlingame, R. W., Wangt, B.-C., Love, W. E. & Moudrianakis, E. N. The nucleosomal core histone octamer at 3.1 A resolution: A tripartite protein assembly and a left-handed superhelix. *Biochemistry* **88,** 10148–10152 (1991).

21. Sandman, K. & Reeve, J. N. Archaeal histones and the origin of the histone fold. *Curr. Opin. Microbiol.* **9,** 520–525 (2006).

22. Talbert, P. B. & Henikoff, S. Histone variants--ancient wrap artists of the epigenome. *Nat. Rev. Mol. Cell Biol.* **11,** 264–75 (2010).

23. Tang, L., Nogales, E. & Ciferri, C. Structure and function of SWI/SNF chromatin remodeling complexes and mechanistic implications for transcription. *Prog. Biophys. Mol. Biol.* **102,** 122–8 (2010).

24. Wang, G. G. *et al.* Chromatin remodeling and cancer, Part I: Covalent histone modifications. *Trends Mol. Med.* **13,** 363–72 (2007).

25. Stewart, M. D., Li, J. & Wong, J. Relationship between histone H3 lysine 9 methylation, transcription repression, and heterochromatin protein 1 recruitment. *Mol. Cell. Biol.* **25,** 2525–38 (2005).

26. Chi, P., Allis, C. D. & Wang, G. G. Covalent histone modifications — miswritten, misinterpreted and mis-erased in human cancers. *Nat. Rev. Cancer* **10,** 457–469 (2010).

27. Strahl, B. D. & Allis, C. D. The language of covalent histone modifications. *Nature* **403,** 41–5 (2000).

28. De Gobbi, M. *et al.* Tissue-specific histone modification and transcription factor binding in ␣ globin gene expression. doi:10.1182/blood-2007

29. Dion, M. F., Altschuler, S. J., Wu, L. F. & Rando, O. J. Genomic characterization reveals a simple histone H4 acetylation code. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 5501–6 (2005).

30. Roh, T., Ngau, W. C., Cui, K., Landsman, D. & Zhao, K. High-resolution genome-wide mapping of histone modifications. *Nat. Biotechnol.* **22,** 1013–1016 (2004).

31. Roh, T.-Y., Cuddapah, S. & Zhao, K. Active chromatin domains are defined by acetylation islands revealed by genome-wide mapping. *Genes Dev.* **19,** 542–52 (2005).

32. Kamakaka, R. T. & Biggins, S. Histone variants: deviants? *Genes Dev.* **19,** 295–310 (2005).

33. Carrino, J. J., Kueng, V., Braun, R. & Laffler, T. G. Distinct replication-independent and -dependent phases of histone gene expression during the Physarum cell cycle. *Mol. Cell. Biol.* **7,** 1933–7 (1987).

34. Henikoff, S. & Ahmad, K. ASSEMBLY OF VARIANT HISTONES INTO CHROMATIN. *Annu. Rev. Cell Dev. Biol.* **21,** 133–153 (2005).

35. Sarma, K. & Reinberg, D. Histone variants meet their match. *Nat. Rev. Mol. Cell Biol.* **6,** 139–149 (2005).

36. Orsi, G. A., Couble, P. & Loppin, B. Epigenetic and replacement roles of histone variant H3.3 in reproduction and development. *Int. J. Dev. Biol.* **53,** 231–43 (2009).

37. Franklin, S. G. & Zweidler, A. Non-allelic variants of histones 2a, 2b and 3 in mammals. *Nature* **266,** 273–5 (1977).

38. Weber, C. M. & Henikoff, S. Histone variants: dynamic punctuation in transcription. *Genes Dev.* **28,** 672–82 (2014).

39. Saksouk, N., Simboeck, E. & Déjardin, J. Constitutive heterochromatin formation and transcription in mammals. *Epigenetics Chromatin* **8,** 3 (2015).

40. Jaeger, S., Barends, S., Giegé, R., Eriani, G. & Martin, F. Expression of metazoan replication-dependent histone genes. *Biochimie* **87,** 827–834 (2005).

41. Marzluff, W. F., Gongidi, P., Woods, K. R., Jin, J. & Maltais, L. J. The Human and Mouse Replication-Dependent Histone Genes. *Genomics* **80,** 487–498 (2002).

42. Osley, M. A. The regulation of histone synthesis in the cell cycle. *Annu. Rev. Biochem.* **60,** 827–61 (1991).

43. Polo, S. E. *et al.* New histone incorporation marks sites of UV repair in human cells. *Cell* **127,** 481–93 (2006).

44. Ray-Gallet, D. *et al.* Dynamics of Histone H3 Deposition In Vivo Reveal a Nucleosome Gap-Filling Mechanism for H3.3 to Maintain Chromatin Integrity. *Mol. Cell* **44,** 928–941 (2011).

45. Goldberg, A. D. *et al.* Distinct Factors Control Histone Variant H3.3 Localization at Specific Genomic Regions. *Cell* **140,** 678–691 (2010).

46. Earnshaw, W. C. & Migeon, B. R. Three related centromere proteins are absent from the inactive centromere of a stable isodicentric chromosome. *Chromosoma* **92,** 290–6 (1985).

47. Buchwitz, B. J., Ahmad, K., Moore, L. L., Roth, M. B. & Henikoff, S. A histone-H3-like protein in C. elegans. *Nature* **401,** 547–8 (1999).

48. Witt, O., Albig, W. & Doenecke, D. Testis-Specific Expression of a Novel Human H3 Histone Gene. *Exp. Cell Res.* **229,** 301–306 (1996).

49. Wiedemann, S. M. *et al.* Identification and characterization of two novel primate-specific histone H3 variants, H3.X and H3.Y. *J. Cell Biol.* **190,** 777–791 (2010).

50. Akhmanova, A. S. *et al.* Structure and expression of histone H3.3 genes in Drosophila melanogaster and Drosophila hydei. *Genome / Natl. Res. Counc. Canada = Génome / Cons. Natl. Rech. Canada* **38,** 586–600 (1995).

51. Frank, D., Doenecke, D. & Albig, W. Differential expression of human replacement and cell cycle dependent H3 histone genes. *Gene* **312,** 135–143 (2003).

52. Krimer, D. B., Cheng, G. & Skoultchi, A. I. Induction of H3.3 replacement histone mRNAs during the precommitment period of murine erythroleukemia cell differentiation. *Nucleic Acids Res.* **21,** 2873–9 (1993).

53. Feng, R. *et al.* Regulation of the expression of histone H3.3 by differential polyadenylation. *Genome* **48,** 503–510 (2005).

54. Wu, R. S. *et al.* Patterns of histone variant synthesis can distinguish G0 from G1 cells. *Cell* **31,** 367–74 (1982).

55. Castiglia, D., Cestelli, A., Scaturro, M., Nastasi, T. & Di Liegro, I. H1(0) and H3.3B mRNA levels in developing rat brain. *Neurochem. Res.* **19,** 1531–7 (1994).

56. Loppin, B. *et al.* The histone H3.3 chaperone HIRA is essential for chromatin assembly in the male pronucleus. *Nature* **437,** 1386–1390 (2005).

57. Ray-Gallet, D. *et al.* HIRA is critical for a nucleosome assembly pathway independent of DNA synthesis. *Mol. Cell* **9,** 1091–100 (2002).

58. Skene, P. J. & Henikoff, S. Histone variants in pluripotency and disease. *Development* **140,** 2513–2524 (2013).

59. Piña, B. & Suau, P. Changes in histones H2A and H3 variant composition in differentiating and mature rat brain cortical neurons. *Dev. Biol.* **123,** 51–8 (1987).

60. Ahmad, K. & Henikoff, S. The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol. Cell* **9,** 1191–200 (2002).

61. Jin, C. *et al.* H3.3/H2A.Z double variant–containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. *Nat. Genet.* **41,** 941–945 (2009).

62. Mito, Y., Henikoff, J. G. & Henikoff, S. Genome-scale profiling of histone H3.3 replacement patterns. *Nat. Genet.* **37,** 1090–1097 (2005).

63. Wirbelauer, C. Variant histone H3.3 is deposited at sites of nucleosomal displacement throughout transcribed genes while active histone modifications show a promoter-proximal bias. *Genes Dev.* **19,** 1761–1766 (2005).

64. Torres-Padilla, M.-E., Bannister, A. J., Hurd, P. J., Kouzarides, T. & Zernicka-Goetz, M. Dynamic distribution of the replacement histone variant H3.3 in the mouse oocyte and preimplantation embryos. *Int. J. Dev. Biol.* **50,** (2006).

65. van der Heijden, G. W. *et al.* Asymmetry in Histone H3 variants and lysine methylation between paternal and maternal chromatin of the early mouse zygote. *Mech. Dev.* **122,** 1008–1022 (2005).

66. Wong, L. H. *et al.* Histone H3.3 incorporation provides a unique and functionally essential telomeric chromatin in embryonic stem cells. *Genome Res.* **19,** 404–414 (2008).

67. Drane, P., Ouararhni, K., Depaux, A., Shuaib, M. & Hamiche, A. The death-associated protein DAXX is a novel histone chaperone involved in the replication-independent deposition of H3.3. *Genes Dev.* **24,** 1253–1265 (2010).

68. Sawatsubashi, S. *et al.* A histone chaperone, DEK, transcriptionally coactivates a nuclear receptor. *Genes Dev.* **24,** 159–170 (2010).

69. Lewis, P. W., Elsaesser, S. J., Noh, K.-M., Stadler, S. C. & Allis, C. D. Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 14075–80 (2010).

70. Wong, L. H. *et al.* ATRX interacts with H3.3 in maintaining telomere structural integrity in pluripotent embryonic stem cells. *Genome Res.* **20,** 351–60 (2010).

71. Kurdistani, S. K. & Grunstein, M. Histone acetylation and deacetylation in yeast. *Nat. Rev. Mol. Cell Biol.* **4,** 276–284 (2003).

72. Struhl, K. Histone acetylation and transcriptional regulatory mechanisms. *Genes Dev.* **12,** 599–606 (1998).

73. Hong, L., Schroth, G. P., Matthews, H. R., Yau, P. & Bradbury, E. M. Studies of the DNA binding properties of histone H4 amino terminus. Thermal denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 "tail" to DNA. *J. Biol. Chem.* **268,** 305–14 (1993).

74. Horn, P. J. & Peterson, C. L. Molecular biology. Chromatin higher order folding--wrapping up transcription. *Science* **297,** 1824–7 (2002).

75. Tse, C., Sera, T., Wolffe, A. P. & Hansen, J. C. Disruption of higher-order folding by core histone acetylation dramatically enhances transcription of nucleosomal arrays by RNA polymerase III. *Mol. Cell. Biol.* **18,** 4629–38 (1998).

76. Greer, E. L. & Shi, Y. Histone methylation: a dynamic mark in health, disease and inheritance. *Nat. Rev. Genet.* **13,** 343–57 (2012).

77. MURRAY, K. THE OCCURRENCE OF EPSILON-N-METHYL LYSINE IN HISTONES. *Biochemistry* **3,** 10–5 (1964).

78. Paik, W. K. & Kim, S. E-N-dimethyllysine in histones. *Biochem. Biophys. Res. Commun.* **27,** 479–483 (1967).

79. Hempel, K., Lange, H. W. & Birkofer, L. [Epsilon-N-trimethyllysine, a new amino acid in histones]. *Naturwissenschaften* **55,** 37 (1968).

80. Byvoet, P., Shepherd, G. R., Hardin, J. M. & Noland, B. J. The distribution and turnover of labeled methyl groups in histone fractions of cultured mammalian cells. *Arch. Biochem. Biophys.* **148,** 558–67 (1972).

81. Borun, T. W., Pearson, D. & Paik, W. K. Studies of histone methylation during the HeLa S-3 cell cycle. *J. Biol. Chem.* **247,** 4288–98 (1972).

82. Zhang, Y. & Reinberg, D. Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes Dev.* **15,** 2343–60 (2001).

83. Vakoc, C. R., Mandat, S. A., Olenchock, B. A. & Blobel, G. A. Histone H3 Lysine 9 Methylation and HP1γ Are Associated with Transcription Elongation through Mammalian Chromatin. *Mol. Cell* **19,** 381–391 (2005).

84. Gerber, M. & Shilatifard, A. Transcriptional Elongation by RNA Polymerase II and Histone Methylation. *J. Biol. Chem.* **278,** 26303–26306 (2003).

85. Delbarre, E. *et al.* Chromatin environment of histone variant H3.3 revealed by quantitative imaging and genome-scale chromatin and DNA immunoprecipitation. *Mol. Biol. Cell* **21,** 1872–84 (2010).

86. Loyola, A. *et al.* PTMs on H3 variants before chromatin assembly potentiate their final epigenetic state. *Mol. Cell* **24,** 309–16 (2006).

87. Krogan, N. J. *et al.* Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. *Mol. Cell. Biol.* **23,** 4207–18 (2003).

88. Zhou, V. W., Goren, A. & Bernstein, B. E. Charting histone modifications and the functional organization of mammalian genomes. *Nat. Rev. Genet.* **12,** 7–18 (2011).

89. Law, M. J. *et al.* ATR-X Syndrome Protein Targets Tandem Repeats and Influences Allele-Specific Expression in a Size-Dependent Manner. *Cell* **143,** 367–378 (2010).

90. Jiao, Y. *et al.* DAXX/ATRX, MEN1, and mTOR Pathway Genes Are Frequently Altered in Pancreatic Neuroendocrine Tumors. *Science (80-.).* **331,** 1199–1203 (2011).

91. Tomonaga, T. *et al.* Overexpression and mistargeting of centromere protein-A in human primary colorectal cancer. *Cancer Res.* **63,** 3511–6 (2003).

92. Li, Y. *et al.* ShRNA-Targeted Centromere Protein A Inhibits Hepatocellular Carcinoma

Growth. *PLoS One* **6,** e17794 (2011).

93. Hu, Z. *et al.* The expression level of HJURP has an independent prognostic impact and predicts the sensitivity to radiotherapy in breast cancer. *Breast Cancer Res.* **12,** R18 (2010).

94. Rangasamy, D. Histone variant H2A.Z can serve as a new target for breast cancer therapy. *Curr. Med. Chem.* **17,** 3155–61 (2010).

95. Dunican, D. S., McWilliam, P., Tighe, O., Parle-McDermott, A. & Croke, D. T. Gene expression differences between the microsatellite instability (MIN) and chromosomal instability (CIN) phenotypes in colorectal cancer revealed by high-density cDNA array hybridization. *Oncogene* **21,** 3253–3257 (2002).

96. Rangasamy, D., Greaves, I. & Tremethick, D. J. RNA interference demonstrates a novel role for H2A.Z in chromosome segregation. *Nat. Struct. Mol. Biol.* **11,** 650–655 (2004).

97. Wu, G. *et al.* Somatic histone H3 alterations in pediatric diffuse intrinsic pontine gliomas and non-brainstem glioblastomas. *Nat. Genet.* **44,** 251–3 (2012).

98. Kim, J. & Orkin, S. H. Embryonic stem cell-specific signatures in cancer: insights into genomic regulatory networks and implications for medicine. *Genome Med.* **3,** 75 (2011).

99. Nichols, J. *et al.* Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* **95,** 379–91 (1998).

100. Avilion, A. A. *et al.* Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev.* **17,** 126–40 (2003).

101. Mitsui, K. *et al.* The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113,** 631–42 (2003).

102. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126,** 663–76 (2006).

103. Takahashi, K. *et al.* Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell* **131,** 861–872 (2007).

104. Graf, T. & Enver, T. Forcing cells to change lineages. *Nature* **462,** 587–94 (2009).

105. Margueron, R. & Reinberg, D. The Polycomb complex PRC2 and its mark in life. *Nature* **469,** 343–9 (2011).

106. Zhang, Z. *et al.* PRC2 Complexes with JARID2, MTF2, and esPRC2p48 in ES Cells to Modulate ES Cell Pluripotency and Somatic Cell Reprograming. *Stem Cells* **29,** 229–240 (2011).

107. Pereira, C. F. *et al.* ESCs require PRC2 to direct the successful reprogramming of differentiated cells toward pluripotency. *Cell Stem Cell* **6,** 547–56 (2010).

108. Nilsson, J. A. & Cleveland, J. L. Myc pathways provoking cell suicide and cancer. *Oncogene* **22,** 9007–9021 (2003).

109. Meyer, N. & Penn, L. Z. Reflecting on 25 years with MYC. *Nat. Rev. Cancer* **8,** 976–990 (2008).

110. Kim, J., Chu, J., Shen, X., Wang, J. & Orkin, S. H. An Extended Transcriptional Network for Pluripotency of Embryonic Stem Cells. *Cell* **132,** 1049–1061 (2008).

111. Chen, X. *et al.* Integration of External Signaling Pathways with the Core Transcriptional Network in Embryonic Stem Cells. *Cell* **133,** 1106–1117 (2008).

112. Kim, J. *et al.* A Myc Network Accounts for Similarities between Embryonic Stem and Cancer Cell Transcription Programs. *Cell* **143,** 313–324 (2010).

113. Fazzio, T. G., Huff, J. T. & Panning, B. An RNAi screen of chromatin proteins identifies Tip60-p400 as a regulator of embryonic stem cell identity. *Cell* **134,** 162–74 (2008).

114. Liu, A., Yu, X. & Liu, S. Pluripotency transcription factors and cancer stem cells: small genes make a big difference. *Chin. J. Cancer* **32,** 483–7 (2013).

115. Ben-Porath, I. *et al.* An embryonic stem cell–like gene expression signature in poorly differentiated aggressive human tumors. *Nat. Genet.* **40,** 499–507 (2008).

116. Chiou, S.-H. *et al.* Positive Correlations of Oct-4 and Nanog in Oral Cancer Stem-Like Cells and High-Grade Oral Squamous Cell Carcinoma. *Clin. Cancer Res.* **14,** 4085–4095 (2008).

117. Gu, G., Yuan, J., Wills, M. & Kasper, S. Prostate Cancer Cells with Stem Cell Characteristics Reconstitute the Original Human Tumor In vivo. *Cancer Res.* **67,** 4807–4815 (2007).

118. Reya, T., Morrison, S. J., Clarke, M. F. & Weissman, I. L. Stem cells, cancer, and cancer stem cells. *Nature* **414,** 105–111 (2001).

119. Beard, C., Hochedlinger, K., Plath, K., Wutz, A. & Jaenisch, R. Efficient method to generate single-copy transgenic mice by site-specific integration in embryonic stem cells. *Genesis* **44,** 23–8 (2006).

120. Gossen, M. *et al.* Transcriptional activation by tetracyclines in mammalian cells. *Science* **268,** 1766–9 (1995).

121. DeBey, M. C. & Ross, R. F. Ciliostasis and loss of cilia induced by Mycoplasma hyopneumoniae in porcine tracheal organ cultures. *Infect. Immun.* **62,** 5312–8 (1994).

122. Almagor, M., Kahane, I., Gilon, C. & Yatziv, S. Protective effects of the glutathione redox cycle and vitamin E on cultured fibroblasts infected by Mycoplasma pneumoniae. *Infect. Immun.* **52,** 240–4 (1986).

123. Shibata, K., Sasaki, T. & Watanabe, T. AIDS-associated mycoplasmas possess phospholipases C in the membrane. *Infect. Immun.* **63,** 4174–7 (1995).

124. Rosenshine, I. & Finlay, B. B. Exploitation of host signal transduction pathways and cytoskeletal functions by invasive bacteria. *Bioessays* **15,** 17–24 (1993).

125. Rottem, S. & Naot, Y. Subversion and exploitation of host cells by mycoplasmas. *Trends Microbiol.* **6,** 436–40 (1998).

126. Conner, D. A. Mouse embryo fibroblast (MEF) feeder cell preparation. *Curr. Protoc. Mol. Biol.* **Chapter 23,** Unit 23.2 (2001).

127. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126,** 663–76 (2006).

128. Silva, J. *et al.* Nanog is the gateway to the pluripotent ground state. *Cell* **138,** 722–37 (2009).

129. Chambers, I. *et al.* Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113,** 643–55 (2003).

130. Silva, J., Chambers, I., Pollard, S. & Smith, A. Nanog promotes transfer of pluripotency after cell fusion. *Nature* **441,** 997–1001 (2006).

131. Brambrink, T. *et al.* Sequential Expression of Pluripotency Markers during Direct Reprogramming of Mouse Somatic Cells. *Cell Stem Cell* **2,** 151–159 (2008).

132. Štefková, K. *et al.* Alkaline Phosphatase in Stem Cells. *Stem Cells Int.* **2015,** 1–11 (2015).

133. O'Connor, M. D. *et al.* Alkaline Phosphatase-Positive Colony Formation Is a Sensitive, Specific, and Quantitative Indicator of Undifferentiated Human Embryonic Stem Cells. *Stem Cells* **26,** 1109–1116 (2008).

134. Heid, C. A., Stevens, J., Livak, K. J. & Williams, P. M. Real time quantitative PCR. *Genome Res.* **6,** 986–94 (1996).

135. Livak, K. J. & Schmittgen, T. D. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods* **25,** 402–408 (2001).

136. Lu, C. *et al.* Histone H3K36 mutations promote sarcomagenesis through altered histone methylation landscape. *Science* **352,** 844–9 (2016).

137. Yuan, W. *et al.* H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. *J. Biol. Chem.* **286,** 7983–9 (2011).

138. Singh, R. K. *et al.* Excess histone levels mediate cytotoxicity via multiple mechanisms. *Cell Cycle* **9,** 4236–44 (2010).

139. Glowczewski, L., Yang, P., Kalashnikova, T., Santisteban, M. S. & Smith, M. M. Histone-histone interactions and centromere function. *Mol. Cell. Biol.* **20,** 5700–11 (2000).

140. Au, W.-C., Crisp, M. J., DeLuca, S. Z., Rando, O. J. & Basrai, M. A. Altered dosage and mislocalization of histone H3 and Cse4p lead to chromosome loss in Saccharomyces cerevisiae. *Genetics* **179,** 263–75 (2008).

141. Hutton, S. R. & Pevny, L. H. SOX2 expression levels distinguish between neural progenitor populations of the developing dorsal telencephalon. *Dev. Biol.* **352,** 40–47 (2011).

6 List of Figures

Figure 1. Overview of the various possible histone H3 modifications, and their associated functions: A) Contribution to fine tuning expression levels at promoters. B) Discrimination between active and inactive conformations at gene bodies. More histone 3 lysine 36 trimethylation (H3.3K36me3) and H3K79me2-modified histones are found in exons in active genes than introns, due to the higher nucleosome occupancy on active exons. C) Histone marks correlate with levels of enhancer activity at distal sites. D) Possible conferral of repression of varying stability and association with different genomic features on a global scale.[88] 11
Figure 2. Transgenic H3.3 mESCs cultured with or without Dox for 4 days in the absence of LIF. Immunostained for H3.3K36me3, H3.3K9me3, and for H3 as the loading control. 29
Figure 3 Alkaline phosphatase staining of mESC clones expressing the Histone H3.3 mutations (H3.3K36I/M, H3.3K9I/M) in 4 ES media culture conditions. 30
Figure 4 AP staining of mESC clones expressing the Histone H3.3 mutations H3.3K36I/M, H3.3K9I/M in 4 ES media culture conditions (n=9). 32
Figure 5 Flow cytometry analysis of H3.3K9M apoptosis assay using FlowJo software. Annexin V PE dye stains apoptotic cells, DAPI stains dead cells. H3OE cells were used for the controls. 33
Figure 6 Embryoid bodies formed from mESCs expressing H3.3 transgenes H3.3K36I/M or H3.3K9I/M cultured in embryoid body media +/- Dox. Representative images taken at time points Day 6 and Day 9, scale bar = 200µm. 37
Figure 7 Size analysis of embryoid bodies. Diameters of EBs measured at day 6 and day 9 +/- Dox using ImageJ 39
Figure 8 qPCR analysis of the relative expression of the pluripotency markers Oct4, Sox2 and Nanog in EBs cultured +/- Dox, over a period of 9 days. 42

7 List of Tables

Table 1. The eight classes of post-translational modifications identified on histones. The residues modified and the associated functions of the modifications are outlined[15]. 7
Table 2 Overview of Cell Lines used. 16
Table 3 Overview of mouse strains used for producing MEFs for reprogramming assay…17

8 List of Abbreviations

Ab	Antibody
Acetyl CoA	Acetyl Coenzyme A
ADP	Adenosine Diphosphate
AP	Alkaline Phosphatase
ATP	Adenosine Triphosphate
Bp	Base pairs
CABIN1	Calcineurin Binding Protein 1
CAF-1	Chromatin assembly factor-1
CCM	Centre for Comparative Medicine
cDNA	Complementary DNA
CENPA	Centromere protein A
CO_2	Carbon dioxide
DAXX	Death-associated protein 6
DEK	Drosophila Eph kinase
D.P.C	Day post-coitum
DIPG	Diffuse intrinsic pontine glioma
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotides
Dox	Doxycycline
DPBS	Dulbecco's Phosphate-Buffered Saline
DTT	Dithiothreitol
EB	Embryoid body
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
ESCs	Embryonic stem cells
EZH2	Enhancer of Zeste Homolog 2
FBS	Foetal Bovine Serum
GBM	Glioblastoma multiforme
HAT	Histone acetyltransferase
HFD	Histone Fold Domain
HIRA	Histone regulator A
HJURP	Holliday junction recognition protein
HMT	Histone methyltransferase
HP1	Heterochromatin protein 1
IACUC	Institutional Animal Care and Use Committee
IMDM	Iscove's Modified Dulbecco's Medium
LIF	Leukaemia Inhibitory Factor
MEFs	Murine embryonic fibroblasts
mESCs	Murine embryonic stem cells
Min	Minutes

mRNA	Messenger Ribonucleic acid
NEAA	Non-essential amino acids
OKSM	Yamanaka Factors Oct4, Klf4, Sox2, c-Myc
PCR	Polymerase Chain Reaction
PRC2	Polycomb Repressive Complex 2
PTM	Post-translational modification
RNA	Ribonucleic acid
Rpm	Rotations per minute
RT qPCR	Reverse Transcription Quantitative Real Time Polymerase Chain Reaction
rTetR	Reverse Tet repressor
rtTA	Reverse tetracycline-controlled trans activator
SAH	S-Adenosylhomocysteine
SAM	S-Adenosyl Methionine
SDS	Sodium dodecyl sulphate
SLBP	Stem Loop Binding Protein
SWI/SNF	SWItch/Sucrose Non-Fermentable
TAE	Tris-acetate-EDTA
TBST	Tris-buffered saline with Tween
tetOP	Tetracycline operator minimal promoter
TG	Tris-glycine
TGS	Tris-glycine-SDS
TRE	Tet response element
UBN1	Ubinuclein-1
UTRs	Untranslated regions
v/v	Volume/volume
w/v	Weight/volume