The role of high mobility group of nucleosome binding proteins in stem cell biology and differentiation

By

Sylvia Patricia Garza Manero
MSc.

Submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy

Institute of Cancer Sciences
College of Medical, Veterinary and Life Sciences
University of Glasgow

March 2019

Abstract

The high mobility group of nucleosome binding proteins (HMGNs) are chromatin architectural proteins that bind specifically to nucleosomes and influence chromatin structure and DNA-dependent functions. However, the mechanisms underlying these events remain largely unknown. HMGN1 and HMGN2 are highly expressed by embryonic stem cells and are downregulated as differentiation proceeds. Nevertheless, embryonic and adult neural stem cells retain elevated levels of these proteins. Chromatin plasticity is essential for the pluri- and multipotency of stem cells and it is achieved by maintaining an open and dynamic chromatin conformation. Conversely, developmental potential seems to be restricted by chromatin condensation. The present work shows that loss of HMGN1 or HMGN2 in pluripotent embryonal carcinoma cells leads to increased spontaneous neuronal differentiation, which is accompanied by a reduction in pluripotency markers and higher gene expression of lineage-specific transcription factors. Inhibition of signalling pathways relevant for neurogenesis does not re-establish the phenotype observed in *Hmgn2*-knockout cells. Withdrawal of the factors sustaining pluripotency in embryonal carcinoma cells results in higher induction of pro-neural factors in cells lacking HMGN1 or HMGN2. Neural stem cells derived from *Hmgn*-knockout cells also display higher gene expression of pro-neural transcription factors and increased spontaneous neuronal differentiation. Loss of HMGN2 disrupts the active histone modification landscape, and therefore affects the chromatin structure at local and global levels. The proposition is that the local changes directly influence the transcription rates of pluripotency and lineage-specific transcription factors, while the global changes may restrict chromatin plasticity. The present data support a hypothesis whereby HMGNs contribute to the chromatin plasticity of stem cells by promoting an active histone modification landscape and open chromatin conformation, which are essential for preserving the self-renewal and developmental potential of stem cells.

Table of Contents

List of Tables

List of Figures

List of Accompanying Material

Appendix 1: Supplementary Figure Derivation of monoclonal lines from parental cells.

Appendix 2: Garza-Manero S., et. al. (2018) “Maintenance of active chromatin states by HMGN1 and HMGN2 is required for stem cell identity” (Submitted for publication).

Acknowledgement

I would like to dedicate this thesis to my life partner, Mauricio, who accepted this challenge with me and who has been in every step, good or bad, offering me all his love.

Infinite thanks to my family for always being beside me. Thank you both Mom and Dad for your eternal support, for believing in me, and for encouraging me to give the best, making you proud is one of my greatest motivations. Thanks Paulina and Raul for bringing happiness to my life and for your eternal company. Thanks to my grandmother Estela for all her love and to my grandmother Sylvia, who sadly could not see the end of this project, but who always trusted I could do this. Thanks to the rest of my family for all their love and always welcoming me with open arms.

Thanks to my friends; my beautiful ballerinas, my creative women from high school, and my bright friends from Uni. Thank you all for being with me beyond the distance. Thanks to my Mexican and non-Mexican family in Glasgow.

A special thanks to my supervisor Dr Katherine West for all her help, support, and time. Thank you for sharing your knowledge, for advising me, and for guiding me throughout this project. Thank you for your feedback and motivation. I would like to extend this thank to my lab members Abdul and Bnar; to Dr Andrew Hamilton, Sarah, and Dr Torsten; to Dr Adam West and his lab members Alejandro, Mauro, Laura, and Valentine; and to Dr Stuart Cobb and Dr Xu Huang. Thank you all for the scientific and non-scientific discussions. Thanks Sarah for being a wonderful lab partner and friend.

Thanks to the University of Glasgow for all the facilities and to the staff for all the support.

Finally, I thank CONACyT Mexico for funding my studies and giving me the opportunity to study abroad and to open my mind.

Author's Declaration

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

Definitions/Abbreviations

2i Two inhibitors (iMAP2K and iGSK3)

ACTB Actin beta

ANOVA Analysis of variance

AraC Cytosine arabinoside

ASCL1 Achaete-scute family bHLH transcription factor 1

ATP Adenosine triphosphate

AXIN2 Axis inhibition protein 2

bHLH Basic helix-loop-helix

BMP Bone morphogenetic protein

BSA Bovine serum albumin

Cas CRISPR associated protein

CREBBP CREB (cAMP response element-binding) binding protein

ChIP Chromatin immunoprecipitation

CRISPR Clustered regularly interspaced short palindromic repeats

Ct Threshold cycle

CTNNB1 Catenin beta 1

Dai Days after induction

DAPI 4, 6-Diamidino-2-phenylindole

DAPT Difluorophenacetyl-L-alanyl-S-phenylglycine t-butyl ester

DHS DNase I hypersensitivity site

DMSO Dimethyl sulfoxide

DNA Deoxyribonucleic acid

DTT Dithiothreitol

ECC Embryonal carcinoma cell

EDTA Ethylene diamine tetraacetic acid

EGF Epidermal growth factor

EGTA Ethylene glycol-2-aminoethylether tetraacetic acid

EP300 E1A binding protein p300

EpiSC Epiblast stem cell

ESC Embryonic stem cell

ESRRB Estrogen related receptor beta

FABP7 Fatty acid binding protein 7

FACS Fluorescence-activated cell sorting

FGF Fibroblast growth factor

GATA4 GATA binding protein 4

GFAP Glial fibrillary acidic protein

GFP Green fluorescent protein

GO Gene ontology

GPI1 Glucose phosphate isomerase 1

GSK3 Glycogen synthase kinase 3

gRNA Guide RNA

H1 Linker histone 1

H2 to 4 Histone 2 to 4

H3K122ac Acetylation on lysine 122 of histone 3

H3K27ac Acetylation on lysine 27 of histone 3

H3K27me3 Trimethylation on lysine 27 of histone 3

H3K4me3 Trimethylation on lysine 4 of histone 3

H3K9ac Acetylation on lysine 9 of histone 3

H3K9me3 Trimethylation on lysine 9 of histone 3

HAT Histone acetyl transferase

HDAC Histone deacetylase

HDM Histone demethylase

HES Hairy and enhancer of split family of bHLH transcription factors

hESC Human embryonic stem cell

HMG High mobility group

HMGA High mobility group AT hook binding proteins

HMGB High mobility group box domain proteins

HMGN High mobility group of nucleosome binding proteins

HMT Histone methyl transferase

HP1 Heterochromatin protein 1

HPRT Hypoxanthine phosphoribosyl transferase

HSC Haematopoietic stem cells

IEG Immediate early gene

IF Immunofluorescence

iPSC Induced pluripotent stem cell

JAK Janus associated kinase

KLF Kruppel like factor

LIF Leukaemia inhibitor factor interleukin 6 family cytokine

MAP2 Microtubule associated protein 2

MAP2K Mitogen-activated protein kinase kinase

MAPK Mitogen-activated protein kinase

MEF Mouse embryonic fibroblast

NBD Nucleosome binding domain

NES Nestin

NEUROD1 Neuronal differentiation 1

NEUROG1 Neurogenin 1

NICD NOTCH intracellular domain

NSC Neural stem cell

OLIG1/2 Oligodendrocyte transcription factor 1/2

OTX2 Orthodenticle homeobox 2

PcG Polycomb group of proteins

PCR Polymerase chain reaction

PIK Proteinase inhibitor cocktail

POU5F1 POU class V homeobox 1

PTM Post-translational modification

qPCR Quantitative polymerase chain reaction

qRT-PCR Reverse transcription linked to quantitative polymerase chain reaction

RA Retinoic acid

RNA Ribonucleic acid

SD Standard deviation

SDS Sodium dodecyl sulphate

SEM Standard error of the mean

siRNA Small interfering RNA

SOX2 SRY-box 2

SRY Sex determining region-Y

SSEA1 Stage-specific embryonic antigen 1

STAT3 Signal transducer and activator of transcription 3

SVZ Subventricular zone

TBX3 T-box 3

TBXT T-box

TrxG Trithorax group

TSS Transcription start site

TUBB Tubulin beta class

TUBB3 Tubulin beta class III

UTR Untranslated region

VZ Ventricular zone

WB Western blotting

WDR5 WD repeat-containing protein 5

WNT Wingless/Integrated

Chapter 1 Introduction

1.1 Chromatin biology

DNA is the molecule that carries heritable information encoding cellular structure and function (McCarty, 2003). In eukaryotic cells, it is packaged within the nucleus in a nucleoprotein complex called chromatin. The nucleosome is the building block of chromatin and consists of an octamer of histone proteins, two of each H2A, H2B, H3, and H4, wrapped by 147 base pairs (bp) of DNA (Felsenfeld & Groudine, 2003; Kornberg, 1974). Nucleosomes are connected by linker DNA, forming an array that folds into a compact fibre of 30 nm in the presence of linker histones (Ball, 2003; Felsenfeld & Groudine, 2003). The fibre is then looped and folded into higher order structures, such as the mitotic chromosomes. In interphase cells, however, chromatin is present in different states of condensation that are broadly divided in two types: highly compacted, repressive heterochromatin and unfolded, transcriptionally active euchromatin (Ball, 2003; Felsenfeld & Groudine, 2003).

All cellular activities that use the DNA as a template, such as transcription, replication, and repair, rely on mechanisms that modulate chromatin and DNA accessibility (Bell et al, 2011). In the context of a nucleosome, the linker DNA is accessible to catalytic enzymes, such as polymerases, and to transcription factors, whereas DNA wrapped tightly around the histone octamer is much less accessible. Therefore, nucleosome positioning and stability, and chromatin compaction are determining factors for the regulation of DNA-dependent functions (Bell et al, 2011).

Nucleosome positioning is largely reliant on ATP-dependent remodelling complexes, which use the energy produced by the hydrolysis of ATP to slide nucleosomes, exposing different DNA sequences (Becker & Hörz, 2002). This mechanism does not operate randomly throughout the genome, however. Nucleosome remodelling complexes are enriched in gene regulatory regions where they play roles in marking transcriptional direction and preventing inappropriate transcription (Whitehouse et al, 2007). Furthermore, they can be recruited by transcription factors to specific genomic regions (Hartley & Madhani, 2009).

Nucleosome stability is influenced by the nucleosome composition. The canonical histones can be replaced by different histone variants, which provide specific properties to nucleosomes (Talbert & Henikoff, 2010). Two variants of H3 have been widely studied; cenH3 is incorporated at the highly compact centromeres whereas H3.3 is associated with active transcription and nucleosome remodelling (Talbert & Henikoff, 2010). Interestingly, nucleosomes containing H3.3 and H2A.Z have been identified close to transcription start sites (TSS), where nucleosome eviction creates a nucleosome-free region, facilitating transcription initiation (Jin et al, 2009a).

Chromatin compaction is largely reliant on reversible and stable covalent modifications of the nucleosome. These includes DNA methylation, which can result in stable gene silencing, and post-translational histone modifications (PTMs), most of which are deposited on the histone tails that protrude from the nucleosome (Bannister & Kouzarides, 2011; Bird, 2002). These modifications constitute a code that can be read by effectors, influencing DNA-dependent activities in a variety of ways (Jenuwein & Allis, 2001).

For instance, acetylation of the histone tails is recognised by bromodomains, which are found in a variety of transcriptional co-activator and adaptor proteins. Interestingly, several protein complexes that contain bromodomains also possess histone acetyl transferase activity, thus magnifying the message of this particular modification (Dhalluin et al, 1999). Chromodomains are readers of histone lysine methylation; for example, the chromodomain of the heterochromatin protein 1 (HP1) specifically targets the trimethylation on lysine 9 of H3 (H3K9me3), an PTM that promotes chromatin compaction (Nakayama et al, 2001).

Histone methylation promotes both chromatin activation and repression, depending on the residue that is modified. H3K4me3 is enriched at actively transcribed genes, while H3K27me3 marks transcriptional silencing (Sims et al, 2003). These two modifications have been linked with developmental regulation by the polycomb group (PcG) of transcriptional repressors, and the trithorax group (TrxG) of transcriptional activators (Ringrose & Paro, 2004). Importantly, PcG and TrxG, in addition to DNA methylation, form the basis of epigenetic inheritance or cellular memory, enabling information concerning gene activation

or repression to be transmitted to a daughter cell, which in turn, ensures the maintenance of tissue-specific expression profiles and the unidirectionallity of developmental processes (Bird, 2002; Margueron & Reinberg, 2011; Ringrose & Paro, 2004; Weintraub et al, 1978).

Chromatin architectural proteins have been also shown to impact chromatin compaction. They are devoid of enzymatic activity and bind to chromatin without any specificity to the underlying DNA sequence (Postnikov & Bustin, 2016). There are two super families of chromatin architectural proteins, the high mobility group (HMG) and the linker histones (Bustin et al, 2005). The classification of the HMG proteins relates to their functional chromatin binding motifs, and includes HMGA, HMGB, and HMGN families (Bustin, 2001). HMGAs contain the AT hook motif, HMGBs bind and bend the minor groove of DNA, and HMGNs specifically target the nucleosome, with higher affinity than to free histones or purified DNA (Shirakawa et al, 2000a). In general, HMGs are associated with a relaxed chromatin conformation permissive for DNA-dependent functions such as transcription, replication, and repair, while linker histones promote chromatin compaction and higher order conformations (Postnikov & Bustin, 2016). The association of the architectural proteins with chromatin is transient and dynamic; they continuously exchange from one nucleosome to another (Phair & Misteli, 2000). FRAP studies have demonstrated that all HMGs compete with H1 for binding sites (Catez et al, 2002; Catez et al, 2004). The emerging picture is that the architectural proteins work in a dynamic network of interactions that ultimately modulates chromatin dynamics, structure, and function (Postnikov & Bustin, 2016).

1.2 HMGN family of chromatin architectural proteins

Pioneering studies identified the two founder members of the HMGN family of chromatin architectural proteins among the most abundant and ubiquitous non-histone proteins found in the nuclei of higher eukaryotes. Weintraub and Groudine first showed in 1976 that the chromatin conformation of active genes was particularly sensitive to DNase I digestion (Weintraub & Groudine, 1976). The DNase I hypersensitivity sites (DHSs) constitute nuclease-accessible regions located at transcriptional regulatory sequences, such as promoters and enhancers, and are considered to be a hallmark of genes poised or activated for

transcription (Martínez de Paz & Ausió, 2016). In a subsequent work, the authors eluted the chromatin of active genes with 0.35 M NaCl and noticed that it lost the DNase I hypersensitivity (Weisbrod & Weintraub, 1979). This property could be reconstituted with a fraction of the 0.35 M NaCl elution, which was highly enriched in two proteins that had high mobility on SDS polyacrylamide gels (Weisbrod & Weintraub, 1979). HMGN1 and HMGN2 were first described as proteins that bind the nucleosomes of actively transcribed genes (Weisbrod et al, 1980).

Expanding the repertory, three other members of the HMGN family were described in the subsequent years (Birger et al, 2001; Shirakawa et al, 2000b; West et al, 2001). HMGN1 to 5 share a general structure consisting of a bipartite nuclear localisation signal (Hock et al, 1998a), a variable negatively charged C-terminal or regulatory domain (Trieschmann et al, 1995), and a highly conserved nucleosome binding domain (NBD) that is the hallmark of the family (Ueda et al, 2008).

The genes encoding HMGN1 and HMGN2 have been identified in all vertebrates and consist of six exons and five introns. The first exon codes the 5´UTR and the first four amino acids of the proteins, the third and fourth exons code the NBD, and the last and largest exon codes for the complete 3´UTR, constituting 70% of the transcript (Bustin et al, 1990). Genes encoding HMGN3 and HMGN5 also exhibit the described structure, although the occurrence of *Hmgn3* has only been reported in mammals, birds, and amphibians, and that of *Hmgn5* exclusively in mammals (González-Romero et al, 2015). *Hmgn4*, however, represents an intronless gene that seems to be originated from an insertion of an *Hmgn2* retropseudogene next to an active promoter around 25 million years ago and is restricted to primates (Birger et al, 2001; González-Romero et al, 2015). The retropseudogene origin of *Hmgn4* is in line with the previous finding that the HMGN family is among the largest retropseudogene families (González-Romero et al, 2015). The structural similarities of the genes suggest that all five members evolved from a common ancestor that emerged after the divergence of vertebrates (González-Romero et al, 2015). Importantly, previous attempts to extract proteins from invertebrates homologous to HMGNs, carrying the NBD, have been unsuccessful (González-Romero et al, 2015).

Hmgn1 and *Hmgn2* transcripts are similar in size, from 1100 to 1250 bp long. The open reading frame represents approximately 25% of the sequence, the 5´UTRs are short and rich in C and G residues, while the 3´UTRs contain several A and U residues and constitute around 70% of the sequence (Bustin et al, 1990). The previous structural features are conserved in *Hmgn3*, *Hmgn4*, and *Hmgn5* mRNAs. In contrast, the sequences are distinctive of each variant. For instance, the open reading frame of *Hmgn3* shares 56 and 65% identity with those of *Hmgn1* and *Hmgn2*, respectively (West et al, 2001). *HMGN4* human transcript, however, is 86% identical to *HMGN2* but only 62% to *HMGN1*, which is consistent with its potential evolutionary origin (Birger et al, 2001).

Interestingly, the mRNA sequence of each variant is evolutionarily conserved between species, particularly that of the open reading frame. In this sense, the open reading frame of the *Hmgn2* has 87% sequence identity between chicken and human and 93% between mouse and human, which represent evolutionary distances of 350 and 96 million years, respectively (Bustin et al, 1990). In the case of *Hmgn3*, the open reading frame of the mouse mRNA is 91% identical to that of the human, while *Hmgn1* transcript is less conserved among species with sequence identities of 60-65% between chicken, mouse, and humans.

Analyses of the protein sequences agree with the conservation levels of the open reading frame of the mRNAs. The mouse, cow, and frog HMGN3 proteins share 91, 86, and 71% identity with human HMGN3, respectively (West et al, 2001). The amino acid sequence identity among HMGN2 proteins is higher than 90%, while among the HMGN1 varies between 50-95% (Bustin et al, 1990). Therefore, HMGN2 evolved slower than HMGN1.

Certain degree of amino acid sequence identity is observed between HMGN1 and HMGN2, but this is less than 50%; HMGN3 is 41 and 54% identical to HMGN1 and HMGN2 (Bustin et al, 1990; West et al, 2001). Importantly, the regions that are evolutionary conserved between all five variants encompass the functional domains characteristic of the HMGN family; the bipartite nuclear localisation signal encoded by exon I and exon V, and the nucleosome binding domain encoded by exons III and IV (Bustin et al, 1990; Malicet et al, 2011; West et al, 2001).

HMGN proteins are expressed in nearly all vertebrate cells and tissues. However, the expression patterns of the HMGN variants are different, raising the possibility that these proteins function in a tissue-specific manner. Their widespread occurrence and evolutionary conservation strongly suggest that they play a relevant role for the proper cellular function.

1.2.1 HMGN binding to chromatin

HMGNs bind to chromatin through their highly conserved NBD. Point mutations in the NBD abolish the structural and functional effects of HMGNs in chromatin (Birger et al, 2003; Catez et al, 2002; Deng et al, 2017; Lim et al, 2005; Prymakowska-Bosak et al, 2001; Rattner et al, 2009). Two molecules of the same variant bind simultaneously to one nucleosome (Postnikov et al, 1995). The core of the NBD contacts the acidic patch formed by H2A and H2B, while the NBD C-terminal region interacts with the DNA (Alfonso et al, 1994; Kato et al, 2011; Murphy et al, 2017; Ueda et al, 2008). The regulatory domain contacts and repositions the H3 and H4 tails disrupting their interaction with the linker DNA (Murphy et al, 2017; Trieschmann et al, 1998). These events result in chromatin decompaction via physical interactions.

Although recent studies have shown that HMGNs and H1 can simultaneously bind a nucleosome, the conformational changes induced after HMGN binding might affect H1 affinity, displacing it from chromatin and interfering with its architectural role in compaction (Murphy et al, 2017). Indeed, it has been reported that HMGNs counteract H1-mediated heterochromatin formation (Ding et al, 1997; Rochman et al, 2009). Furthermore, the interplay between the two major HMGN variants and H1 regulates gene expression relevant for oligodendrocyte development (Deng et al, 2017).

Since their discovery, the HMGNs have been related with active and unfolded chromatin (Bustin, 2001; Weisbrod et al, 1980). A major question is how these HMGNs are targeted to particular regions in the chromatin. The fact that the members of the HMGN family share a disordered structure that facilitates multiple protein to protein interactions raises the possibility that other proteins recruit HMGNs to specific genomic regions. Indeed, a high proportion of them

are found forming macromolecular complexes in the nuclei of cells (Hansen et al, 2006; Lim et al, 2002).

Considering that HMGNs induce conformational changes in the nucleosome and are capable of multiple protein to protein interactions, they are also proposed to mediate chromatin accessibility to transcription factors or chromatin modulators, or act as molecular bridges for epigenetic players and transcriptional activators/repressors. In this sense, an antagonistic role of HMGNs in ATP-dependent nucleosome remodelling has been described (Rattner et al, 2009). In addition, three members of the family have been shown to enhance the action of histone acetyl transferases (HATs), and consequently, *Hmgn1*-knockout cells display reduced levels of H3 acetylation (Barkess et al, 2012; Lim et al, 2005; Ueda et al, 2006). HMGN1 also modulates H3 phosphorylation (Lim et al, 2004). The influence of HMGNs on post-translational histone modifications may be a key mechanism underlying their roles in chromatin structure and function.

1.2.2 Biological functions of HMGN proteins

The generation of HMGN variant-specific knockout mice in recent years has offered considerable opportunities to gain insights into the biological function of HMGNs. Single knockout mice are available lacking HMGN1, HMGN2, HMGN3, or HMGN5 proteins, and additionally, *Hmgn1* and *Hmgn2* double-knockout mice have been generated. Surprisingly, all mice are viable and do not display severe or obvious phenotypes, considering the high degree of evolutionary conservation and widespread occurrence of HMGNs. Although in different levels, all variants are expressed in many tissues and cell types, which raises the possibility of functional compensation among them. In fact, HMGN2 binding to chromatin is increased in cells derived from the *Hmgn1*-knckout mice, and vice versa, without affecting the transcriptional or translational rates of the residual HMGN variants (Deng et al, 2015). Alternatively, other chromatin architectural proteins members of the HMG and H1 super families can compensate the loss of HMGNs, since they all work in a dynamic network of interactions (Postnikov & Bustin, 2016) that may be adjusted in order to maintain chromatin dynamics, structure, and function.

Nevertheless, HMGNs are not fully redundant. Variant-specific phenotypic alterations have been observed; *Hmgn1*-knockout mice exhibit increased tumorigenicity and impaired DNA damage response (Birger et al, 2005; Birger et al, 2003), the energy metabolism seem to be perturbed in *Hmgn2*-knockout mice (Deng et al, 2015), and *Hmgn3*-knockout mice are mildly diabetic (Ueda et al, 2009). Noteworthy, the phenotypes mentioned above suggest an inefficient response to stress or to metabolic conditions. In this regard, alterations in epigenetic regulators or chromatin architecture are not always evident in the steady state, but in the cellular response to extrinsic stimuli.

1.2.2.1 Establishment and maintenance of DNA regulatory regions

A major question is whether the HMGNs are randomly distributed throughout the genome in physiological conditions or whether they are positioned at specific sites. The development of genome-wide techniques facilitates the analysis of the genomic distribution of these proteins. For instance, HMGN1 and HMGN2 have shown to colocalise with DHSs of mammalian cells (Cuddapah et al, 2011; Deng et al, 2013; Deng et al, 2015; Zhang et al, 2016). Accordingly, the loss of the two major HMGN variants reduces the number and intensity of DHSs in living cells, suggesting that the chromatin is less accessible in regulatory regions (Deng et al, 2015; Zhang et al, 2016). Furthermore, cells derived from double-knockout mice show a remodelling of the DHS landscape, particularly at enhancer regions (Deng et al, 2015; Zhang et al, 2016). Importantly, the loss of the two major HMGN variants is required for the chromatin accessibility effects mentioned above, as in single-knockout mice the loss of HMGN1 is compensated by increased HMGN2 binding to chromatin, and vice versa, which ultimate maintains the most of the original DHSs.

Although there is evidence about a relationship between enhancers and HMGNs dating from when HMGNs were discovered (Martínez de Paz & Ausió, 2016), it remains to be resolved how HMGNs are targeted to enhancers and what mechanisms HMGNs utilise for the establishment and maintenance of the DHSs.

1.2.2.2 Regulation of gene expression

The role of the HMGNs in chromatin structure, and particularly, in the establishment and maintenance of DHSs, is expected to impact gene expression in vertebrate cells.

Microarray hybridisation and RNA-seq experiments performed in different tissues from single *Hmgn*-knockout mice suggest that the loss of an HMGN leads to both upregulation and downregulation of limited genes in a cell type- and variant-specific manner (Deng et al, 2015; Kugler et al, 2013). In other words, there is little overlap of the affected genes between the tissues after the loss of a particular HMGN variant, and additionally, the genes affected in a certain tissue are different after the loss of each variant. Several conclusions can be derived from the previous data. First, HMGNs do not act as general promoters or inhibitors of transcription, which contrasts with previous studies *in vivo* where displacing a major HMGN variant from chromatin arrests polymerase II dependent transcription (Hock et al, 1998b). Second, there is not a specific subset of genes regulated by each HMGN variant, which suggests that the HMGNs do not work as transcription factors and is in agreement with the lack of a consensus DNA sequence targeting HMGNs to certain genomic loci (Shirakawa et al, 2000a). Third, the HMGNs are not fully redundant, since the lack of a single variant affects gene expression in every tissue (Deng et al, 2015; Kugler et al, 2013). And fourth, the HMGNs do not modify the pre-existing expression profile of tissues, which is illustrated by principal component analysis where the different tissues cluster together regardless the presence or absence of an HMGN variant (Kugler et al, 2013).

The last conclusion is supported by transcriptional analysis during neuronal differentiation of embryonic stem cells derived from *Hmgn1*-knockout mice (Deng et al, 2013). In this study, the loss of HMGN1 did not interfere with the temporal regulation of gene expression, enabling the establishment of expression profiles dictated by the developmental stages. However, subtle differences in a subset of genes were observed in the three different stages evaluated (Deng et al, 2013). Furthermore, the transition from one stage to another is not identical; the directional changes of the genes are retained, i.e.

up and downregulation, yet the magnitude of change differs (Deng et al, 2013), suggesting that HMGNs optimise the regulation of gene expression.

Similar observations have been reported when studying the kinetics of immediate early genes (IEGs) activation following a stress stimulus. In mouse embryonic fibroblasts (MEFs) derived from the double-knockout mice lacking HMGN1 and HMGN2, the transcriptional response to treatment with a protein synthesis inhibitor and to heat shock displays different kinetics than that observed in MEFs derived from control littermates (Deng et al, 2015). Some of the evaluated IEGs are induced faster and others slower in the double-knockout mice (Deng et al, 2015), reinforcing the concept that HMGNs optimise gene expression processes rather than work as activators or repressors of transcription. Likewise, the loss of the two major members of the HMGN family dampens the magnitude of the defence transcriptional response of B lymphocytes when exposed to lipopolysaccharide (Zhang et al, 2016). Importantly, the proportion of genes showing alterations in the absence of HMGNs is greater when the cells are exposed to stimuli than in the rest state (Deng et al, 2015; Zhang et al, 2016). This may explain the impaired glucose tolerance of *Hmgn3*-knockout mice, since the levels of glucose and insulin in the serum of fasting animals are normal, while after feeding, sugar-stimulated insulin secretion is lower leading to higher blood sugar levels (Ueda et al, 2009). Similarly, the *Hmgn1*-knockout mice exhibit an impaired response to DNA damage (Birger et al, 2003).

Considering these recent experiments and the phenotypes of the knockout mice, the emergent picture is that HMGNs fine-tune an already established expression profile and ensure the appropriate cellular response to external cues such as stress, feeding, infections, and developmental programs. This in turn guarantees normal cellular and organism phenotypes.

1.2.2.3 HMGNs during embryonic development

The two major HMGN variants are developmentally regulated. They are expressed in oocytes and throughout all preimplantation stages (Deng et al, 2017; Deng et al, 2013; Mohamed et al, 2001). Interestingly, targeting both HMGN1 and HMGN2 with small interference RNAs or antibodies complementary

to the NBD delays initial cell divisions, suggesting that these proteins are essential for the proper timing of early developmental stages (Mohamed et al, 2001).

After implantation, HMGN1 and HMGN2 are progressively downregulated as differentiation proceeds (Furusawa et al, 2006; Lehtonen & Lehtonen, 2001; Lehtonen et al, 1998). Their decreasing levels have been observed during myogenesis, erythropoiesis, and chondrongenesis, and seems to be required since overexpression of HMGN1 impairs normal cellular differentiation (Crippa et al, 1991; Furusawa et al, 2006; Pash et al, 1990). Thus, downregulation of the two major members of HMGN family seems to be characteristic of and necessary for tissue differentiation.

It is worth mentioning that tissue-specific stem cells and transient amplifying precursors appear to retain high levels of HMGN1 and HMGN2 (Furusawa et al, 2006; Lehtonen & Lehtonen, 2001), suggesting that HMGNs are important for stem cell biology. Indeed, the Allen Developing Mouse Brain Atlas (Allen Institute for Brain Science, 2008) and the Allen Mouse Brain Atlas (Allen Institute for Brain Science, 2004) from the Allen Institute from Brain Sciences database reveal that *Hmgn1* and *Hmgn2* transcripts are highly abundant in the ventricular and subventricular zone (VZ and SVZ) of the developing and adult brain, in addition to the adult hippocampus. These regions are well-characterised neurogenic niches providing the adequate microenvironment sustaining neural stem cell (NSC) survival, self-renewal, and differentiation (Bjornsson et al, 2015; Gage, 2000; Temple, 2001). Immunofluorescence of coronal sections of the developing mouse brain demonstrates that HMGN1 and HMGN2 are strongly expressed in the VZ and SVZ, and largely colocalise with NSC markers (Nagao et al, 2014).

1.3 Modelling embryonic development *in vitro*

Embryonic development is the process through which the zygotic cell creates a complete organism, following cell division and differentiation programs. It is by definition epigenetic, since all cells in an organism contain the same genome, while expressing different genes. Therefore, it relies on the establishment of self-renewal or lineage-specific expression profiles according to a spatio-temporal patterning (see below).

In the 1970s, Stevens reported that pre- and post-implantation mouse embryos grafted into the testes of adult mice developed tumours of multilineage differentiated cells, and remarkably, that some of the embryonal cells remained undifferentiated retaining the capacity to form other teratocarcinomas when transplanted serially (Stevens, 1970). The observation that the undifferentiated cells were pluripotent and capable of proliferating indefinitely encouraged efforts to capture pluripotency *in vitro*, resulting in defined culture conditions suitable for derivation and propagation of embryonal carcinoma cell (ECC) lines (Martin, 1980).

1.3.1 Embryonal carcinoma cells

ECCs surprised researchers with their similarities to the cells that originate the developing mouse embryo; they share several markers corresponding to an embryonal identity and can contribute to all tissues in mouse chimaeras (Martin, 1980). Furthermore, when allowed to aggregate, ECCs form differentiating embryoid bodies that resemble early embryos (Martin & Evans, 1975).

P19 is an ECC line derived from a 7.5-day embryo that was transplanted into the testis of an adult mouse, and exhibits a normal male diploid karyotype (McBurney & Rogers, 1982). P19 cells does not spontaneously differentiate and can be indefinitely propagated in serum-supplemented media (McBurney, 1993). However, when stimulated with retinoic acid (RA) they differentiate into neurons and glia, while generating endodermal and mesodermal lineages, including cardiac and skeletal muscle, after exposure to dimethyl sulfoxide (DMSO) (McBurney, 1993). P19 cells can be genetically manipulated, providing a suitable system for studying mammalian embryonic development (McBurney, 1993).

1.3.2 Embryonic stem cells

Taking advantage of the derivation method for ECCs, embryonic stem cell (ESC) lines were established from mouse pre-implantation blastocysts (Evans & Kaufman, 1981). One of the main properties of ESCs is their pluripotency, defined as the ability to differentiate into all cell types present in an organism and tested as the colonisation of all embryonic tissues after injection into pre-

implantation mouse embryos (Smith, 2001). Importantly, unlike most of ECCs, ESCs present a stable diploid karyotype suitable for contribution into the germ line in chimeras, and therefore raise the possibility of transgenic mice derivation (Smith, 2001). Finally, ESCs are clonogenic, which means that a clonal population can be propagated from a single cell, demonstrating their indefinite self-renewal capacity (Smith, 2001).

Human embryonic stem cells (hESCs) originating from pre-implantation blastocysts have been also captured *in vitro* (Guo et al, 2016; Thomson et al, 1998). These cells are capable of multilineage differentiation tested *in vitro* or by teratocarcinoma formation, and are indefinitely propagated as undifferentiated cells (Guo et al, 2016; Thomson et al, 1998). Conventional hESC lines differ in several ways to their mouse counterparts, and interestingly, to human blastocysts, including signalling responses, epigenetic landscape, and expression profiles (Davidson et al, 2015). It is thought that these hESCs develop to post-implantation stages *in vitro*, which is consistent with the observation that human pluripotent cells are transient and rapidly proceed to differentiation (Rossant, 2015). Nevertheless, a recent study reported the derivation of hESC lines closely related to mouse ESCs and blastocysts in terms of transcriptome, epigenome, and metabolism, using a different formulation of the media (Guo et al, 2016). Although the coming years will provide further information about these hESCs, it seems likely that capturing human pre-implantation stages *in vitro* has been achieved (Guo et al, 2016).

1.3.3 Epiblast stem cells

Recently, pluripotent cell lines were derived from the mouse post-implantation epiblast, and therefore, are referred as epiblast stem cells (EpiSCs) (Brons et al, 2007; Tesar et al, 2007). Similar to ESCs, EpiSCs express pluripotency transcription factors, maintain their karyotype, and differentiate into all somatic cell types; however, they respond differently to signalling pathways involved in self-renewal and differentiation (Brons et al, 2007; Tesar et al, 2007). Furthermore, EpiSCs display distinct transcriptional profiles resembling late-gastrulation ectoderm, the layer that develops into neuroectoderm, suggesting that these cells are already primed for lineage specification (Kojima et al,

2014). Interestingly, mouse EpiSCs and hESCs share important features, such as signalling pathway responses and gene expression patterns (Tesar et al, 2007).

1.3.4 Signalling pathways regulating stemness

The culture conditions that supported the first successful derivation of mouse ESCs consisted of a co-culture with mitotically inactivated feeder cells in serum-supplemented media (Smith, 2001). Subsequent studies arrived at the conclusion that the contribution of feeder cells to ESC propagation relied on the secretion of leukaemia inhibitor factor (LIF), and therefore, the layer of feeders could be replaced by LIF (Smith et al, 1988). LIF activates the Janus associated kinases (JAKs), promoting signal transducer and activator of transcription 3 (STAT3)-mediated effects in ESC self-renewal (Niwa et al, 1998). However, LIF alone does not indefinitely sustain pluripotency and, after few cell passages, the cells differentiate towards the neural pathway (Ying et al, 2003b). Therefore, serum was suggested to contain inhibitors of neural fates. In this sense, bone morphogenetic proteins (BMPs) antagonise neural induction and can replace the serum in the culture media, allowing the development of chemically defined formulations (Ying et al, 2003a).

As previously mentioned, in the presence of LIF alone, ESCs proceed towards neural lineages (Ying et al, 2003b), and intriguingly, in the absence of LIF, BMPs promote non-neural specification (Ying et al, 2003a). Therefore, serum/LIF or BMP/LIF combinations sustain self-renewal by preventing differentiation, which, in chemically defined media, indicates that ESC differentiation is triggered by either intrinsic programs, or autocrine/paracrine cues (Martello & Smith, 2014). Further studies demonstrated that the autocrine stimulation of the fibroblast growth factor (FGF) receptor by FGF4 is required for ESC differentiation (Kunath et al, 2007; Stavridis et al, 2007). Ablation of FGF signalling in ESCs impedes neural and non-neural induction, indicating that FGF signalling acts upstream of the commitment to a certain lineage, and LIF/BMP block differentiation of committed cells (Kunath et al, 2007; Ying et al, 2003b). FGF signalling activates the mitogen-activated protein kinase (MAPK) cascade (Stavridis et al, 2007); accordingly, inhibition of the MAPK cascade sustains ESC self-renewal in the absence of LIF and BMPs (Ying et al, 2008).

It is worth mentioning that the FGF/MAPK cascade promotes proliferation in many cell types, including EpiSCs and hESCs (Brons et al, 2007; Tesar et al, 2007; Thomson et al, 1998). Therefore, independence from this signalling pathway is a particular characteristic of ESCs that changes during development (Martello & Smith, 2014; Smith, 2017).

The observation that the factors sustaining ESC self-renewal shield the cells from differentiation means that pluripotency can be regarded a stable cell state. Hence, ESCs will remain undifferentiated unless instructed otherwise (Ying et al, 2008). However, culturing ESC under MAPK cascade inhibition in chemically defined media leads to increased apoptosis (Ying et al, 2008). The inclusion of an antagonist of glycogen synthase kinase 3 (GSK3) in the culture media increases the survival of ESCs and, in conjunction with LIF, these two inhibitors (2i) support the propagation of a more homogeneous population in terms of morphology and pluripotency gene expression than that cultured in serum/LIF or BMP/LIF (Wray et al, 2010; Ying et al, 2008).

1.3.5 Transcription factors of stemness

The POU class V homeobox 1 (POU5F1) was the first described pluripotency transcription factor; it is highly expressed in P19 cells and its transcription rates drop upon RA induced neuronal differentiation (Okamoto et al, 1990). In mouse development, POU5F1 is detected in the oocyte, blastocyst, and epiblast; during gastrulation, however, POU5F1 is present exclusively in cells generating the germ line (Schöler et al, 1990). Furthermore, embryos lacking POU5F1 can proceed towards the blastocyst stage, but the pre-implantation epiblast loses pluripotency and differentiates into trophoectoderm (Nichols et al, 1998). Surprisingly, trophoectoderm differentiation is also observed after the loss of POU5F1 in ESCs, even when blastocysts does not differentiate into trophoectoderm (Niwa et al, 2000).

The sex-determining region-Y box 2 (SOX2) binds co-operatively with POU5F1 at enhancers of relevant pluripotent genes, including NANOG homeobox, *Pou5f1*, and *Sox2*, among others (Masui et al, 2007). SOX2 is essential in this positive feedback loop that stabilises *Pou5f1* gene expression (Masui et al, 2007). POU5F1 and SOX2 are included in the transcription factor cocktail that induces active

cellular reprogramming from fibroblasts into induced pluripotent stem cells (iPSCs), demonstrating their determinant role in pluripotency (Takahashi & Yamanaka, 2006).

The transcription factor NANOG can promote ESC self-renewal in serum-supplemented media in the absence of LIF (Chambers et al, 2003). Conversely, loss of NANOG reduced the clonogenic capacity of ESCs; however, it did not completely abolish self-renewal (Chambers et al, 2007). *Nanog* expression during development is more restricted than that of *Pou5f1* and *Sox2*; it starts in the morula stage and is downregulated upon implantation (Chambers et al, 2003). POU5F1, SOX2, and NANOG co-occupy several target genes (Boyer et al, 2005). In addition, they collaborate to form a pluripotency transcriptional network that is auto-regulated by positive feedback loops (Boyer et al, 2005; Chen et al, 2008).

Other elements of the pluripotency transcriptional network are KLF2, KLF4, ESRRB, TBX3, among others (Nichols & Smith, 2012). It has been shown that the expression of target genes regulated by this network is higher when four or more of the pluripotency transcription factors bind to the regulatory regions (Kim et al, 2008). NANOG, KLF2, KLF4, ESRRB, TBX3 are individually dispensable for ESC self-renewal, but collaborate in the maintenance of the pluripotency transcriptional network (Nichols & Smith, 2012). Interestingly, these are all uniformly expressed by ESCs cultured in 2i/LIF media, while displaying a mosaic gene expression in serum/LIF or BMP/LIF (Chambers et al, 2007; Kalmar et al, 2009; Marks et al, 2012; Wray et al, 2010). Furthermore, they are downregulated in the post-implantation epiblast and in EpiSCs, and their downregulation is accompanied by higher expression of lineage-specific genes (Chambers et al, 2007; Kalmar et al, 2009; Kinoshita & Smith, 2018; Nichols & Smith, 2009; Smith, 2017).

The previous observations led to the proposition of the existence of a naïve pluripotency state that is largely homogeneous in morphology and transcription factor gene expression, and a primed pluripotency state that represents a further developmental stage, displaying lineage-specific gene expression (Kinoshita & Smith, 2018; Nichols & Smith, 2009; Smith, 2017). In the mouse embryo, these pluripotency states correspond to pre- and post-implantation embryos, whereas in pluripotent cells they correspond to naïve ESCs and EpiSCs,

respectively (Kinoshita & Smith, 2018; Nichols & Smith, 2009; Smith, 2017). Naïve ESCs are uniformly propagated in 2i/LIF media and homogeneously express NANOG, KLF2, KLF4, ESRRB, TBX3, among other naïve pluripotency transcription factors (Marks et al, 2012; Wray et al, 2010; Ying et al, 2008). ESCs cultured in serum/LIF or BMP/LIF constitute a heterogeneous population of naïve, primed, and intermediate pluripotency states (Chambers et al, 2007; Kalmar et al, 2009; Kinoshita & Smith, 2018; Nichols & Smith, 2009; Smith, 2017).

It is worth mentioning that the expression profiles and morphological features of cells cultured in serum/LIF or BMP/LIF and 2i/LIF are interconvertible (Marks et al, 2012). Furthermore, cells expressing lower levels of naïve pluripotency factors and higher levels of lineage-specific genes are capable of self-renewal, and can originate cells with naïve expression profiles (Abranches et al, 2014; Chambers et al, 2007; Herberg et al, 2016; Kalmar et al, 2009).

In contrast to the loss of naïve pluripotency factors, POU5F1 and SOX2 remain expressed at high levels during early ESC differentiation (Kalkan et al, 2017; Trott & Martinez Arias, 2013). It is likely that these two factors contribute to cell fate decisions by interacting with lineage-specific transcription factors in the establishment of transitory expression profiles. For example, POU5F1 and OTX2 co-occupy a set of enhancers during early ESC differentiation that is different to the enhancers bound by POU5F1/SOX2 in pluripotent cells (Buecker et al, 2014). In addition, this developmental transition is accompanied by changes in the epigenetic landscape, highlighting the importance of chromatin in the establishment and maintenance of lineage-specific expression profiles (Buecker et al, 2014).

1.3.6 Epigenetic modulators of stemness

POU5F1, SOX2, NANOG, and other naïve pluripotency transcription factors collaborate to regulate the expression of target genes (Chen et al, 2008; Kim et al, 2008). Notable, their binding sites in ESCs strongly colocalise with those of the HATs CREBBP and EP300 (Chen et al, 2008; Fang et al, 2014; Göke et al, 2011). Additionally, POU5F1 physically interacts with WDR5, which is a member of TrxG, and their binding sites in ESCs extensively overlap (Ang et al, 2011). These observations suggest that pluripotency transcription factors recruit

chromatin modifiers to activate the expression of target genes. Furthermore, the overlapping sites include regulatory regions of the master regulators of pluripotency, POU5F1, SOX2, and NANOG, suggesting that chromatin modifiers participate in the positive feedback loops stabilising self-renewal of ESCs (Ang et al, 2011; Göke et al, 2011; Zhong & Jin, 2009).

Accordingly, the lower pluripotency transcription factor gene expression upon ESC differentiation is accompanied by a global reduction of histone acetylation, H3K4me3, and *Wdr5* expression (Ang et al, 2011; Krejčí et al, 2009; Lee et al, 2003). Furthermore, knockdown of WDR5 and double-knockdown of CREBBP and EP300 in mouse ESCs compromise self-renewal and facilitate differentiation, since the cells display lower alkaline phosphatase reactivity and pluripotency gene expression, in addition to higher lineage-specific transcription rates (Ang et al, 2011; Fang et al, 2014). Conversely, treatment with histone deacetylase (HDAC) inhibitors supports ESC self-renewal and delays induced differentiation (Qiao et al, 2015; Ware et al, 2009).

These observations suggest that an open chromatin conformation is essential for the maintenance of self-renewal and pluripotency of ESCs, and enhances the function of the pluripotency transcriptional network. In agreement, chromatin of ESCs present unique global features that distinguish it from that of differentiated cells, provide a homogeneously decondensed and dynamic structure, and confer great plasticity to the fibre (Mattout & Meshorer, 2010). Such features include numerous DHSs (Deng et al, 2013), enriched active marks (Efroni et al, 2008; Guenther et al, 2007; Krejčí et al, 2009), reduced DNA methylation and H3K9me3 (Bibikova et al, 2006; Efroni et al, 2008), fewer and less condensed heterochromatin foci (Efroni et al, 2008; Meshorer et al, 2006), and hyperdynamic binding of chromatin architectural proteins, such as H1 and HP1 (Meshorer et al, 2006), in addition to HMGNs (Deng et al, 2013).

The open chromatin conformation of ESCs favours global activation of the genome. Indeed, the total RNA content normalised to DNA is two fold higher in ESCs than in lineage-specific precursors (Efroni et al, 2008). ESCs transcribe genomic regions that are usually repressed in somatic cells, such as repetitive sequences and lineage-specific genes for other tissues, although at very low levels, denoting a noisy transcriptional hyperactivity, which is suggested to

constitute a hallmark of ESCs (Efroni et al, 2008). This transcriptional hyperactivity requires further control to avoid misinterpreted signals that may result in precocious differentiation (Chi & Bernstein, 2009). Chromatin modifiers have been shown to buffer gene expression; accordingly, HATs and HDACs are both targeted to active and silenced genes (Wang et al, 2009). Dynamic cycles of acetylation and deacetylation inhibit promiscuous initiation at active genes while also priming the repressed genes for future activation (Wang et al, 2009).

In ESCs, developmental transcription factors driving differentiation towards a specific lineage are repressed by PcG, displaying increased levels of H3K27me3 compared to differentiating cells of that lineage (Boyer et al, 2006). However, these genes also accumulate H3K4me3 at the promoter region (Azuara et al, 2006; Bernstein et al, 2006). These bivalent domains at developmental gene promoters constitute another hallmark of ESCs, which restrict gene expression while poising the genes for transcriptional activation (Azuara et al, 2006; Bernstein et al, 2006; Marks et al, 2012). Bivalent genes resolve into monovalency during differentiation, retaining either H3K4me3 or H3K27me3 marks, depending on whether the gene becomes active or silenced (Mikkelsen et al, 2007). In this sense, epigenetic pre-patterning has been shown to play an important role in cell fate decisions during development (Szutorisz et al, 2005; Xu et al, 2011).

Once a cell's fate has been specified, its expression profiles and cellular identity are highly stable, and modifiers of chromatin structure play an important role in cellular memory (Ringrose & Paro, 2004). Consequently, epigenetic cellular memory represents an enormous challenge for the reprogramming of somatic cells into iPSCs (Chen & Dent, 2013).

Considering all of the above, it is clear that chromatin structure and modifiers perform important roles in pluripotency maintenance and ESC fate decisions, and these roles are beyond that of stabilising the expression profile driven by transcription factors (Chen & Dent, 2013).

1.4 Neural fates

Neural induction of ESCs occurs in the gastrulation embryo under BMP/SMAD signalling inhibition (Wilson & Houart, 2004). This observation led to the proposition of a default model of neural induction, suggesting that ESCs naturally acquire a neural identity unless instructed otherwise (Hemmati-Brivanlou, 1997; Munoz-Sanjuan & Brivanlou, 2002; Wilson & Edlund, 2001). In agreement, ESCs differentiate into neural progenitors *in vitro* upon withdrawal of the factors that sustain pluripotency (Ying et al, 2003b). The neuroectoderm is the first neural structure and is comprised of neuroepithelial cells that amplify by symmetric divisions, constructing first the neural plate, and subsequently, the neural tube (Martynoga et al, 2012). At the onset of neurogenesis, these neuroepithelial cells specify into radial glia cells, expressing fatty acid binding protein 7 (FABP7) and attaching to the VZ, which function as NSCs during development and in the adult brain (Kriegstein & Alvarez-Buylla, 2009).

1.4.1 Neural stem cells: neurogenesis and gliogenesis

NSCs are multipotent tissue-specific stem cells capable of self-renewal, and they can differentiate into neurons and into the main macroglial lineages: astrocytes and oligodendrocytes (Gage, 2000; Kriegstein & Alvarez-Buylla, 2009; Temple, 2001).

Once the NSCs have specified into radial glia, they divide asymmetrically to generate another radial glial cell and a more differentiated cell, which leaves the VZ and initiates neuronal differentiation (Martynoga et al, 2012). In the particular case of the cerebral cortex, there are intermediate states mediated by intermediate progenitors that, after limited rounds of self-renewal at the SVZ, differentiate into early-born and late-born cortical neurons migrating towards deeper and upper cortical layers, respectively (Molyneaux et al, 2007). In agreement, NSCs isolated from early stages produce a majority of early-born neurons, while NSCs from later stages differentiate into late-born neurons (Shen et al, 2006). Interestingly, early-stage NSCs can differentiate into late-born neurons after exposure to the proper environmental cues, however, late-stage NSCs cannot produce early-born neurons even if they are embedded into an early

environment, which suggest that there is a temporal restriction in the developmental potential of NSCs (Shen et al, 2006).

In fact, the developmental potential of NSCs changes over time and a neurogenic to gliogenic switch is observed *in vitro* from isolated NSCs (Qian et al, 2000). After the neurogenic phase, most of the radial glia transform into astrocytes and the astrogenesis continues postnatally (Kriegstein & Alvarez-Buylla, 2009). Similarly, oligodendrocyte differentiation occurs in waves starting from the onset of astrogenesis to postnatal stages (Kessaris et al, 2005). The precise spatio-temporal regulating of the molecular switches defining the neurogenic to gliogenic transition is crucial for the proper size and morphology of neural tissues (Hirabayashi & Gotoh, 2005; Miller & Gauthier, 2007).

Finally, some NSCs persist in the adult brain at the SVZ and the dentate gyrus of the hippocampus, where they continuously generate neurons and glia (Gage, 2000; Kriegstein & Alvarez-Buylla, 2009; Temple, 2001).

1.4.2 Molecular mechanisms regulating neurogenesis

The basic helix-loop-helix (bHLH) pro-neural transcription factors perform pivotal roles in neuronal differentiation. Three of them, Neurogenin 1 and 2 (NEUROG1 and NEUROG2), and mammalian achaete-scute family bHLH transcription factor 1 (ASCL1), are expressed in cortical NSCs and their loss is associated with impaired neurogenesis and premature glial differentiation (Nieto et al, 2001). Furthermore, forced gene expression of a single one or a combination of these pro-neural transcription factors is sufficient to induce neurogenesis in P19, mouse ESCs, and human iPSCs (Farah et al, 2000; Reyes et al, 2008; Thoma et al, 2012; Velkey & O'Shea, 2013). These observations suggest that the pro-neural transcription factors orchestrate neuronal differentiation programs. In addition to its neurogenic role, NEUROG1 inhibits astrocyte differentiation of NSCs, contributing to a proper timing of the developmental switches (Sun et al, 2001).

Considering the above information, it is essential to maintain silencing of pro-neural genes in order to avoid precocious or extended neuronal differentiation. In this sense, NOTCH signalling has been proven to promote survival and self-

renewal of NSCs, and to promote gliogenesis following the neurogenic phase (Ohtsuka et al, 2001; Oishi et al, 2004). NOTCH signalling is mediated by the proteolytic cleavage of the cell membrane receptor upon ligand binding; subsequently, the NOTCH intracellular domain (NICD) translocates to the nucleus where it activates the transcription of the pathway effectors (Lai, 2004). The hairy and enhancer of split (HES) family of bHLH transcriptional repressors are among the main NOTCH effectors, regulating lineage-specific stem cell maintenance and fate decisions (Kageyama et al, 2007; Ohtsuka et al, 1999). In particular, NOTCH signalling promotes self-renewal of NSCs and gliogenesis, through a mechanism in which HES1 and HES5 antagonise the expression of pro-neural genes, and therefore, the commitment to neurogenesis (Kageyama et al, 2007; Shimojo et al, 2008).

In addition, FGF signalling potentiates self-renewal of NSCs in a cell-autonomous manner that impedes precocious differentiation into intermediate progenitors or neurons (Kang et al, 2009). Interestingly, WNT signalling performs dual roles in NSC fate decisions; at early stages, WNT stimulates self-renewal (Kuwahara et al, 2010), while instructing neurogenesis by the induction of *Neurog1* expression at later stages (Hirabayashi et al, 2004).

At the onset of the gliogenic phase, stable silencing is required to guarantee the neurogenic to gliogenic switch. The PcG repressor complex PRC2 is recruited to the *Neurog1* gene promoter where it deposits the repressive PTM H3K27me3, reducing the chromatin accessibility of this pro-neural gene (Hirabayashi et al, 2009). Furthermore, chromatin of NSCs experiences a progressive condensation, where the self-renewal and neurogenic potential of early-stage NSCs correlate with a higher DNase I hypersensitivity, while late-stage NSCs show a more compacted chromatin conformation (Kishi et al, 2012).

These observations provide evidence that NSC fate decisions are regulated by the interaction of extrinsic cues, intrinsic transcriptional programs, and epigenetic modulators.

1.4.3 HMGN proteins in neural lineages

As previously mentioned, NSCs express high levels of HMGN1 and HMGN2 (Nagao et al, 2014), and interestingly, Deng et al (2013) reported an increase in the expression of *Hmgn1* in the transition from ESC to NSC following a neuronal differentiation protocol. Although this result was out of the scope of the authors' work, it strongly suggests that HMGNs are relevant for NSC biology. In agreement, the SVZ of the adult *Hmgn1*-knockout mice harbor fewer cells for the NSC marker Nestin (NES) and displays lower *Nes* transcript levels (Deng et al, 2013).

The HMGNs are not only important for the maintenance of NSCs, but in addition, for NSC fate decisions. Nagao et al (2014) reported that HMGNs promote astrocyte differentiation at the expense of neurogenesis. Overexpression of either HMGN1 or HMGN2 in NSCs isolated from both early and late stage embryogenesis resulted in higher numbers of astrocytes compared to control, and conversely, siRNAs targeting *Hmgn1*, *Hmgn2*, or *Hmgn3* increased neurogenesis from these cells (Nagao et al, 2014). Similar outputs were observed when evaluating the brains of seven day old mice after introducing plasmids encoding HMGN1, HMGN2, HMGN3, or siRNAs targeting either of these variants into cortical neural precursor cells by *in utero* electroporation at the stage corresponding to gliogenesis onset (Nagao et al, 2014). Notably, *Hmgn*-expressing plasmids failed to increase gliogenesis when introduced to the neural precursor cells at earlier stages (Nagao et al, 2014), suggesting that a pre-requisite for these phenomena is that the neural precursor cells have acquired gliogenic potential.

Recent studies have shown that ESC differentiation towards neural lineages is altered when one or both of the major HMGN variants are lacking (Deng et al, 2017; Deng et al, 2013). Gene ontology analysis revealed that expression of genes in pathways related to neuronal development were altered in the knockout cells (Deng et al, 2017). Moreover, *in vitro* oligodendrocyte differentiation of double-knockout ESCs is significantly less efficient than control ESCs, which is consistent with the lower number of OLIG2-positive cells observed in the spinal cord of the double-knockout mice and the behavioural alterations of these animals (Deng et al, 2017).

Taken together, these studies support a role for HMGNs in the regulation of developmental processes along the neural lineage.

1.5 Hypothesis and specific aims

The general hypothesis of this project is that the two major members of the HMGN family of chromatin architectural proteins, HMGN1 and HMGN2, play an important role in the establishment and/or maintenance of the particular active chromatin conformation and histone modification landscape of pluripotent cells, which in turn are essential for the self-renewal and the preservation of the differentiation potential of these cells. Therefore, loss of HMGN1 or HMGN2 may interfere with the pluripotency of P19 embryonal carcinoma cells.

1.5.1 Specific aims:

1. To investigate whether the loss of HMGN1 and HMGN2 affects the protein and transcript levels of pluripotency and lineage-specific markers.

2. To investigate whether P19 cells retain their differentiation potential after the loss of a major HMGN variant.

3. To investigate whether NSCs can be derived from P19 cells lacking HMGN1 or HMGN2, as they can be derived from P19 WT cells and ESCs, and whether these cells present a normal phenotype.

4. To investigate whether the loss of a major HMGN variant modifies the histone modification landscape of P19 cells.

Chapter 2 Materials and Methods

2.1 Cell culture

Table 2.1 Cell culture product information

Product	Provider	Catalogue number
DMEM/F12	Thermo Fisher Scientific	12634-010
Neurobasal media	Thermo Fisher Scientific	21103-049
Recovery cell culture freezing media	Thermo Fisher Scientific	12648-010
New born calf serum	Thermo Fisher Scientific	26010-074
Fetal bovine serum	Sigma-Aldrich	F7624
GlutaMAX	Thermo Fisher Scientific	35050-038
N2 supplement	Thermo Fisher Scientific	17502-048
B27 supplement	Thermo Fisher Scientific	17504-044
BSA fraction V	Thermo Fisher Scientific	15260-037
TrypLE	Thermo Fisher Scientific	12604-013
LIF	ESGRO, Millipore	ESG1107
PD03259010	Stratech	S1036-SEL
CHIR99021	Stratech	G09-901B-SGC
FGF2	Peprotech	450-33
EGF	Peprotech	315-09
FGF8	Thermo Fisher Scientific	PHG0184
SU-5402	Sigma-Aldrich	SML0443
XAV-939	Stratech	S1180-SEL
DAPT	Sigma-Aldrich	D5942
RA	Sigma-Aldrich	R3255
Laminin	Sigma-Aldrich	L2020
Gelatine	Sigma-Aldrich	G1890

2.1.1 Embryonal carcinoma cells

P19 mouse ECCs (ATCC CRL-1825) were routinely propagated onto uncoated tissue culture flasks in DMEM/F12 basal media supplemented with 2 mM glutaMAX and 10% new born calf serum (P19 media) at 37°C in a 5% CO_2 incubator. Cells were passaged every two days using TrypLE at a constant 1:12 split ratio. For freezing, 2 x 10^6 cells were re-suspended in 1 ml of recovery cell culture freezing media, immediately transferred into a cryovial, and deposited into a cell freezing container (Biocision BCS-405) at -80°C. Then, cells were stored in liquid nitrogen. For thawing, a cryovial was partially thawed and cells were transferred drop by drop to 9 ml of pre-warmed P19 media. Cells were then centrifuged at 200 RFC, re-suspended in new P19 media and seeded onto an uncoated tissue culture flask.

2.1.2 Retroviral production of HMGN2

Retroviral production using the pBABE-hygro vector, and P19 cell infection were performed by Mariarca Bailo and Faika Laz Banti. Briefly, Phoenix-Eco retrovirus producer cells were obtained from ATCC and cultured in DMEM with 10% FBS. Protocols were based on those defined by the Nolan laboratory, Stanford University. Phoenix-Eco cells were transfected with plasmid pBABE-hygro (Addgene) or pBABE-hygro-HMGN2 using lipofectamine 2000 (Thermo Fisher Scientific 31985-062). pBABE-hygro-*HMGN2* contains the human *HMGN2* cDNA inserted in the EcoRI/Sal1 sites of the pBABE-hygro vector and was created by Brenna Flately. Retroviral suspension was passed through a 0.2 µm filter, concentrate with a centrifugal filter, then added to P19 cells in the presence of hexadimethrine bromide (Sigma-Aldrich H9268). Infected cells were selected with 400 µg/ml hygromycin, then maintained in 200 µg/ml hygromycin routinely.

2.1.3 Pathway inhibitor treatment

P19 WT and *Hmgn*-knockout cells were seeded at 2 x 10^5 cells/well in 6-well plates with 2 ml of P19 media. Cells were allowed to attach overnight at 37°C in a 5% CO_2 incubator. The next day, XAV-939 was used for inhibiting WNT, DAPT for NOTCH, and SU-5402 for FGF; all inhibitors were added to the culture media at a final concentration of 10 µM. After 24 h, cells were harvested according to the qRT-PCR protocol and samples were stored at -80°C until required. For the

present work, the experiment was performed in triplicated at consecutive cell passages.

2.1.4 Neural induction

Neuronal induction of P19 WT and *Hmgn*-knockout cells was performed according to the rapid and efficient method of neuronal differentiation previously reported (Nakayama et al, 2014). Cells were seeded at 1 x 10^5 cells/ml onto laminin-coated wells of 6-well plates in neural induction media. For the laminin-coating, laminin was added to sterile PBS at a final concentration of 0.5 µg/ml and the solution was filter-sterilised; then, enough volume to cover the surface was added to the dish and removed following 4 h of incubation at 37°C. Neural induction media consisted of DMEM/F12, 1X N2 supplement, 200 mM glutaMAX, 500 nM RA, 10 ng/ml FGF8, and 10 µM DAPT. The day of seeding was considered day 0. For the present work, the protocol was performed two independent times, including different cell lines each time, and therefore, the results were presented separately (see chapter 5).

2.1.5 Differentiation in N2 media

P19 WT and *Hmgn*-knockout cells were seeded at 1 x 10^5 cells/ml onto laminin-coated wells of 6-well plates (see above) in N2 media. N2 media consists of DMEM/F12, 1X N2 supplement, and 200 mM glutaMAX. The day of seeding was considered day 0.

2.1.6 Embryonic stem cells

E14 mouse ESCs (Genome Research Limited) were routinely propagated onto 0.1% gelatine-coated dishes in 2i media (Ying et al, 2008) at 37°C in a 5% CO_2 incubator. Cells were passaged every two days using TrypLE at a constant 1:10 split ratio (approximate seeding density of 1.5 x 10^4 cells/cm^2). For gelatine-coating, 0.1% w/v gelatine solution was prepared and filter-sterilised; then, enough volume to cover the surface of the dish was added and, after 1 h at 37°C, the solution was removed. For preparing 2i media 1 x 10^3 U/ml LIF, 1 µM PD03259010 (MAP2K inhibitor), and 3 µM CHIR99021 (GSK3 inhibitor) were added to N2B27 media (Table 2.2). For freezing, 2 x 10^6 cells were re-suspended in 1 ml of freshly prepared freezing media (60% N2B27 media, 30% fetal bovine serum,

and 10% DMSO), immediately transferred into a cryovial, and deposited into a cell freezing container (Biocision BCS-405) at -80°C. Then, the cells were stored in liquid nitrogen. For thawing, a cryovial was partially thawed and the cells were transferred drop by drop to 9 ml of pre-warmed 2i media. Cells were then centrifuged at 200 RFC, re-suspended in new 2i media, and seeded onto a gelatine-coated 10 cm dish.

Table 2.2 N2B27 media

Product	Final conc.
DMEM/F12	Mixture 1:1 ratio
Neurobasal	
N2 supplement	0.5X
B27 supplement	0.25X
GlutaMAX	2 mM
BSA fraction V	50 µg/ml
2-mercaptoethanol	0.1 mM

2.1.6.1 Differentiation

Neural induction and differentiation of E14 cells were performed according to a previous report (Ying et al, 2003b). Before starting, E14 cells were cultured at high density (1.5 x 10^5 cells/cm^2) onto gelatine-coated dishes in 2i media and usual conditions for 24 h. Then, cells were dissociated using TrypLE and seeded onto gelatine-coated wells of 6-well plates at low density (1.0 x 10^4 cells/cm^2) in N2B27 media (Table 2.2). Neural induction was therefore achieved by withdrawing the factors that sustain pluripotency of ESCs. The day of seeding in N2B27 media was considered day 0. The medium was renewed every day.

2.1.7 Neural stem cells

Neural induction of P19 cells and ESCs was achieved as previously described. At indicated time points, NSC derivation and differentiation were performed by following and adapting a previously reported method (Conti et al, 2005).

2.1.7.1 NSC derivation

E14 cells: at day 6, the cells were dissociated and 3 x 10^6 cells were seeded as suspension into an uncoated tissue culture flask in NSC media (10 ng/ml each, FGF and epidermal growth factor (EGF), were added to N2B27 media), allowing the formation of floating neurospheres. Three days after, the floating aggregates were collected and seeded onto gelatine-coated dishes in NSC media, this was considered passage 1. From this point, the cells were routinely propagated in gelatine-coated dishes and NSC media by passaging them every two days at 1:3-1:5 split ratio. Under these conditions, after few passages, the cells acquired a homogeneous characteristic bipolar morphology. Cells from passage 8 and onwards were used for analysis.

P19: at 3 days after induction, NSC media replaced the neural induction media. The day after, the cells were transferred to a gelatine-coated dish in NSC media, this was considered passage 1. From this point, the cells were routinely propagated in gelatine-coated dishes and NSC media by passaging them every two days at 1:3-1:5 split ratio. Under these conditions, after few passages, the cells acquired a homogeneous characteristic bipolar morphology. Cells from passage 8 and onwards were used for analysis.

2.1.7.2 NSC differentiation

Neuronal differentiation of E14- and P19-derived NSCs was triggered by withdrawing EGF and dropping the concentration of FGF2 to 5ng/ml in N2B27 media. The cells were seeded onto laminin-coated wells of 6-well plates or coverslips at a density of 1 x 10^5 cells/ml. The medium was replaced every two days and the cells were harvested for analysis one week after growth factor restriction.

For glial differentiation 2 x 10^5 cells/ml were seeded onto laminin-coated wells of 6-well plates and coverslips in N2B27 media supplemented with 1% of new born calf serum. The medium was replaced every two days and the cells were harvested for analysis one week after growth factor deprivation.

2.2 qRT-PCR

For RNA analysis of the undifferentiated cells, cells were seeded at 2 x 10^5 cells/well in 6-well plates with 2 ml of P19 media and cells were harvested after 48 h. For the present work, samples were collected from 3 to 10 independent cultures (different cell passages and/or frozen batches). In the cases of the treatment with inhibitors or the differentiation protocols, the time points in which the cells were harvested are either previously indicated in this chapter or mentioned in chapter 5. Table 2.3 contains relevant product information.

Table 2.3 qRT-PCR product information

Product	Provider	Catalogue number
RNeasy mini kit	Qiagen	74104
DNase set	Qiagen	79254
Superscript III	Thermo Fisher Scientific	18080044
RNase OUT	Thermo Fisher Scientific	10777019
FastStart universal SYBR green master (rox)	Sigma-Aldrich	04913914001

2.2.1 RNA isolation

Cells were harvested from one well of 6-well plates in 350 µl of RLT buffer (RNeasy mini kit) using a cell scraper and samples were stored at -80°C until required. RNA was isolated with RNeasy mini kit, following the manufacturer instructions. Briefly, cell lysates were mixed 1:1 with 70 % ethanol and applied to a column (to induce RNA binding to the silica membrane of the column). On-column DNase digestion was performed according to the manufacturer to remove genomic DNA traces. Then, the RNA was washed using two different buffer solutions, and eluted in 50 µl of RNase-free water. RNA concentration (A_{260}) and purity ($A_{260}/A_{280} > 2$) were determined using the Nanodrop and RNA integrity was tested on a 1% agarose gel.

2.2.2 cDNA synthesis

For the cDNA synthesis, a first mixture per reaction was prepared as follows: 300 ng of RNA, 0.5 µl of 100 µM oligo dT_{20}, 0.5 µl of 10 mM dNTPs, and RNase-free water up to 10 µl. Once prepared, the mixture was incubated for 5 min at 65°C and placed on ice for at least 1 min. During the incubation time, a second mixture per reaction consisted on: 3 µl of 5X first strand buffer, 1 µl of 0.1 M DTT, 0.5 µl of RNase OUT, and 0.5 µl of Superscript III. The RT negative control utilised RNase-free water instead of Superscript III. The second mixture was incubated for 50 min at 50°C, 5 min at 85°C, and ice for 1 min. Finally, 1 µl of 1:5 dilution of RNase H was added to each sample and the reaction was incubated for 20 min at 37°C to digest the RNA. Samples were stored shortly at -20°C.

2.2.3 Real time PCR

Real time or quantitative PCR (qPCR) reactions were performed in 96-well plates. Before starting, cDNA samples were diluted 1:5 with nuclease-free water. One qPCR reaction contained 5 µl of diluted cDNA, 12.5 µl of FastStart universal SYBR green master (with rox), 3 µl of forward primer, 3 µl of reverse primer, and 1.5 µl of nuclease-free water. The 96-well plate was centrifuged for 3 min at 1000 RFC. qPCRs were run in the Stratagen Mx 3000P thermocycler and the conditions were indicated in Table 2.4. Primer sequences and concentrations were provided in Table 2.5.

Table 2.4 qRT-PCR thermal cycling conditions

Segment	Cycles	Time	Temperature	Collection of data
Polymerase activation	1	10 min	95°C	
Denaturation	40	30 s	95°C	
Alignment-extension		60 s	60°C	In the extension
Dissociation curve analysis	1	60 s	95°C	
		30 s	55°C	In the ramp from 55 to 95°C
		30 s	95°C	

Table 2.5 qRT-PCR primer information

Primer	Sequence (5'-3')	Final conc. (nM)
Pou5f1 F	CGTTCTCTTTGGAAAGGTGTTCA	600
Pou5f1 R	GGTTCTCATTGTTGTCGGCTTC	600
Sox2 F	CGTTCATGTGCGCGTAGCTG	600
Sox2 R	GGAACAGCATGGCGAGCGG	600
Nanog F	ACCTGAGCTATAAGCAGGTTAAG	600
Nanog R	TCAGACCATTGCTAGTCTTC	600
Nes F	AAAGTTCCAGCTGGCTGT	300
Nes R	CACTTCCAGACTAAGGGACAT	300
Tubb3 F	AAGGTAGCCGTGTGTGACATC	300
Tubb3 R	ACCAGGTCATTCATGTTGCTC	300
Map2 F	TCTGCCTCTAGCAGCCGAAG	300
Map2 R	CACTGTGGCTGTTTGTTCTG	300
Gfap F	CAACCTGGCTGCGTATACCAG	600
Gfap R	TTAAGAACTGGATCTCCTCC	600
Neurog1 F	GGCTTCATGCATTATGGATCC	900
Neurog1 R	CTCCAGTCCAGTGCCTGAATAG	900
Ascl1 F	CCAACTGGTTCTGAGGACCTG	600
Ascl1 R	CTGCCATCCTGCTTCCAAA	600
Gata4 F	TAGTCTGGCAGTTGGCACAG	300
Gata4 R	ACGGGACACTACCTGTGCAA	300
Tbxt F	ATGCCAAAGAAAGAAACGAC	300
Tbxt R	AGAGGCTGTAGAACATGATT	300
Axin2 F	GAGAGTGAGCGGCAGAGC	300
Axin2 R	CGGCTGACTCGTTCTCCT	300
Fgf4 F	CGACCACAGGGACGCTGCTG	300
Fgf4 R	ACTCCGAAGATGCTCACCACG	300
Hes5 F	CACCAGCCCAACTCCAAGCT	300
Hes5 R	GGCGAAGGCTTTGCTGTGT	300
Hmgn1 F	AGAGACGGAAAACCAGAGTCCAG	300
Hmgn1 R	CGTGATGGATGCTTAGTCGGA	300
Hmgn2 F	AAAAGGCCCCTGCGAAGAA	300
Hmgn2 R	TGCCTGGTCTGTTTTGGCA	300
Gpi1 F	TCCGTGTCCCTTCTCACCAT	900
Gpi1 R	TGGCAGTTCCAGACCAGCTT	50

2.2.4 Analysis of data and statistics

Amplification plots and dissociation curves were observed and compared with the RT negative controls to determine the specificity and the performance of the qRT-PCR reactions. Dissociation curves should show the presence of a unique amplicon and in any other cases the product was considered as undetermined. Threshold values (Ct) were analysed by the $2^{-\Delta\Delta Ct}$ method, which determined the fold change between normalised values from the experimental condition in comparison with the control condition. In the present work, the control condition for P19 cells was represented by the average of normalised data (average ΔCt) displayed by undifferentiated parental cells. For E14 cells, the average ΔCt of undifferentiated E14 cells constituted the control condition. The normaliser was the house-keeping gene encoding for the enzyme glucose phosphate isomerase 1 (GPI1), unless indicated in the figure legend.

Statistical analysis was performed on undifferentiated P19 samples of *Hmgn*-knockout cells and control cells (from 3 to 10 independent cultures) using the GraphPad Prism 7 software. The input values corresponded to individual ΔCt per gene and cell line. Every gene was analysed separately, constituting an independent hypothesis, while all samples from each cell line were compared with all samples from parental cells. Statistical significance was determined by one-way analysis of variance (ANOVA) and the *post-hoc* test selected was Dunnett's multiple comparison test. Adjusted p-values were indicated in the figure legends.

2.3 Immunological techniques

Table 2.6 Product information related with immunological techniques

Product	Provider	Catalogue number
cOmplete mini EDTA-free protease inhibitor cocktail	Sigma-Aldrich	11836170001
Benzonase	Sigma-Aldrich	E1014
Pierce BCA protein assay kit	Thermo Fisher Scientific	23225
NuPAGE LSD sample buffer	Thermo Fisher Scientific	NP0007
NuPAGE sample reducing agent	Thermo Fisher Scientific	NP0004
NuPAGE 12% bis-tris protein gels	Thermo Fisher Scientific	NP0341, NP0342
NuPAGE antioxidant	Thermo Fisher Scientific	NP0005
NuPAGE MOPS-SDS running buffer	Thermo Fisher Scientific	NP0001
Benchmark protein ladder	Thermo Fisher Scientific	10747012
PVDF membranes	Bio-Rad	21703
SuperSignal West Duro Extended Duration Substrate	Thermo Fisher Scientific	34076
Prolong gold mounting media with DAPI	Thermo Fisher Scientific	P36931
Formaldehyde 16% solution	Thermo Fisher Scientific	28908
DYNA beads-Protein A	Thermo Fisher Scientific	10002D
Proteinase K	Sigma-Aldrich	1001954815
RNase A	Thermo Fisher Scientific	ENO531
Qiagen MinElute PCR pufirifiation kit	Qiagen	28004

2.3.1 Western blotting

2.3.1.1 Protein extraction

Whole cell extracts were prepared by scrapping the cells in 200 µl of cell lysis buffer (45 mM of Tris-HCl pH 7.5, 1 mM EDTA, 1% SDS, 10% glycerol, 2 mM DTT, 10 mM sodium butyrate, one Mini-tablet of protease inhibitor cocktail (PIK) per 10 ml of lysis buffer). Samples were stored at -20°C until required. Protein quantification was performed by analysing an immunoblot for the detection of tubulin beta class I (TUBB), using 10 µl of each sample.

2.3.1.2 Acid histone extraction

Cells were washed on plate twice with ice-cold PBS. Then, cells were scrapped in 3 ml PBS and counted. Subsequently, cells were centrifuged at 200 RFC and 4°C for 5 min. The cell pellet was re-suspended in triton extraction solution (4.5 µ/ml triton X-100, 0.02% w/v NaN_3, 4 mM sodium butyrate, all in PBS, and one Mini-tablet of PIK per 10 ml of triton extraction solution buffer) to a concentration of 10^7 cells/ml and was incubated on ice with gentle rocking. Cells were centrifuged at 300 RFC and 4°C for 10 minutes. The pellet was washed in half the previous volume of triton extraction solution, centrifuged in the same conditions, and re-suspended in 0.2 M HCl at a concentration of 4 x 10^6 cells/50 µl. Samples were incubated for 3 hours at 4°C and centrifuged at 300 RFC and 4°C for 10 minutes. The supernatant containing the histones was stored at -20°C until required. Protein quantification was performed by analysing an immunoblot for the detection of H3, using 1 µl of each sample.

2.3.1.3 Immunoblotting

Whole cell extracts and acid histone extracts were mixed with LSD buffer and reducing agent, and denaturing during 10 min at 70°C. Then, they were loaded in 12% bis-tris protein gels. Gels ran at 120 V for 20 min and 150 V for 1 h on MOPS-SDS running buffer with antioxidant agent. The separated proteins were transferred to PVDF membranes in the Trans-Blot Turbo transfer device (Bio-Rad). PVDF membranes were blocked with WB buffer (0.1% Tween-20 and 5% non-fat milk in PBS) for 1 h at RT on a shaker. After blocking, primary antibody dilutions prepared in WB buffer were added to the membranes and the antigen-

antibody reaction was incubated overnight at 4°C. The next day, the membranes were washed four times with 0.1% Tween-20/PBS for 10 min and incubated for 1 h at RT with the HRP-conjugated secondary antibody dilutions prepared in WB buffer. Finally, the antigen-antibody reaction was detected with chemiluminescence using a CCD camera imaging system (LAS 3000 Fujifilm). The antibody details are listed in Table 2.7. WB quantification of histone modifications was performed in Image J. The values were first normalised with total H3 and then expressed as relative to parental cells.

2.3.2 Immunofluorescence

Cells were fixed with 4% PFA for 30 mins at RT and stored in PBS at 4°C until required. Immunofluorescences (IF) were performed as follows: cells were permeabilised for 10 min with 0.1% PBS-triton X-100 and then blocked for 1 h in IF buffer (5% horse serum and 0.5% triton X-100 in PBS). Subsequently, cells were incubated for 2 h with primary antibody dilution prepared in IF buffer. Then, cells were washed thrice for 10 min with 0.1% PBS-triton X-100 and incubated for 1 h with the secondary antibody dilutions prepared in IF buffer, protecting from light. Finally, cells were washed thrice for 10 min with PBS and mounted onto slides, using mounting media containing DAPI to stain the nuclei. The complete protocol was carried at RT on a shaker. Details of the antibodies and dilutions are listed in Table 2.7. Slides were observed in an epifluorescence microscope (Olympus IX51).

2.3.3 Fluorescence-activated cell sorting

Cells were collected using TripLE and washed with PBS. Then, they were stained with zombie yellow (dead cells) for 15 min in the dark and gentle rotation. From this point onwards, all the steps were performed in the dark. Subsequently, the cells were washed with PBS by centrifugation at 200 RFC for 3 min and fixed during 30 min with 4% PFA. The 4% PFA was washed with 1 ml 10% FBS-PBS by higher centrifugation at 380 RFC for 3 min. For SSEA1 detection, cells were re-suspended in 200 µl of primary antibody dilution prepared in 2% BSA-PBS and incubated for 1.5 h in gentle rotation. The rest of the cells were permeabilised and blocked for 15 min in FACS buffer (5% horse serum and 0.5% tween 20 in PBS) and washed once with 1 ml 10% FBS-PBS by centrifugation at 380 RFC for 3

min. After fix-perm, cells were incubated for 1 h in gentle rotation with 200 µl of the primary antibody dilution prepared in FACS buffer. The primary antibody dilution was washed as previously described and cells were incubated for 45 min in 200 µl of the secondary antibody dilution prepared in of 2% BSA-PBS or FACS buffer, in gentle rotation. Finally, cells were washed twice with 2% PBS-BSA and re-suspended in 2% PBS-BSA. The IgG isotype controls were included in the staining protocol. NSCs derived from E14 ESCs were utilised as a negative control for the detection of pluripotency and endodermal markers. Complete details of the antibodies and dilutions are listed in Table 2.7. FACS analysis was carried using the Attune Focusing Cytometer (Applied Biosystems, Thermo Fisher Scientific). The yellow-stained cells were excluded from the data.

Three independent experiments were performed for quantitative FACS. Statistical analysis was carried in the GraphPad Prism 7 software. The median of the fluorescence intensity of each immunostaining was normalised dividing the value by the median of the fluorescence intensity of the negative control (E14 derived NSCs). The normalised data constituted the numeric input for the statistical tests. Each marker was analysed separately, constituting an independent hypothesis, while all samples from each cell line were compared with the samples from parental cells. Statistical significance was determined by one-way analysis of variance (ANOVA) and the *post-hoc* test selected was Dunnett's multiple comparison test. Adjusted p-values were indicated in the figure legends.

Table 2.7 Antibody information for WB, IF, FACS, and ChIP

Antibody	Provider	Catalogue number	IF/FACS dilution	WB dilution	ChIP vol/rxn
Rb anti POU5F1	Abcam	Ab19857	1:500		
Rb anti NANOG	Abcam	Ab80892	1:200		
Ms anti SSEA1	Abcam	Ab16285	1:200		
Rb anti GATA4	Abcam	Ab84593	1:500		
Ms anti NES	Abcam	Ab6142	1:200		
Rb anti FABP7	Abcam	Ab27171	1:1000		
Rb anti TUBB3	Abcam	Ab18207	1:2000		
Ms anti TUBB3	Millipore	MAB1637	1:500		
Ch anti MAP2	Abcam	Ab5392	1:2000		
Rb anti GFAP	Abcam	Ab7260	1:500		
Rb anti HMGN1	Home made		1:1000	1:1000	5 µl
Rb anti HMGN2	Home made		1:2000	1:1000	5 µl
Rb anti H3	Millipore	07-690		1:25000	2 µl
Rb anti H3.cs1	Active Motif	39574		1:1000	
Rb anti H3K4me3	Millipore	07-473		1:20000	5 µl
Rb anti H3K27me3	Millipore	07-449			7.5 µl
Rb anti H3K9ac	Millipore	07-352		1:10000	5 µl
Rb anti H3K27ac	Millipore	07-360		1:20000	5 µl
Rb anti H3K122ac	Abcam	Ab33309			5 µl
Ms anti TUBB	TFS	MA5-16308		1:3000	
rIgG	Sigma	I5006			7.5 µl
mIgG	Sigma	12-371	1:200		
Gt anti rb-HRP	TFS	32460		1:1000	
Dn anti rb-488	TFS	A21206	1:1000		
Dn anti rb-594	TFS	A11037	1:1000		
Gt anti ms-594	TFS	A11020	1:1000		
Gt anti ch-594	TFS	A11042	1:1000		

2.3.4 Chromatin immunoprecipitation-PCR

2.3.4.1 Chromatin preparation

P19 cells were seeded onto six 15 cm dishes. One plate was used for cell counting, and the others for chromatin collection. Just before starting, DMEM/F12 is added to replace the P19 media. The cells were cross-linked on a shaker for 5 min by adding 5% formaldehyde dilution (16% formaldehyde solution in formaldehyde dilution buffer: 50 mM HEPES pH 8.0, 0.1 M NaCl, 1 mM EDTA and 0.5 mM EGTA) directly to the DMEM/F12 media to reach a 0.5% final concentration. The cross-linking reactions were quenched with 1 ml of 1.5 M glycine, incubating for 5 min on the shaker. Cells were washed with 20 ml of chilled PBS and collected with a cell scraper in 5 ml of chilled PBS. The cell suspension was centrifuged for 5 min at 200 RFC and 4°C. Then, the pellet was washed with PBS by centrifugation and re-suspended in 5 ml of cell lysis buffer (10 mM Tris pH 8, 10 mM NaCl, 10 mM EDTA, 0.5 mM EGTA, 0.2% NP40, and one cOmplete Mini EDTA-free Cocktail tablet in 10 ml). Following an incubation of 10 min on ice, the cells were centrifuged in the same conditions, but this time the pellet was re-suspended in nuclei lysis buffer (50 mM Tris pH 8.0, 10 mM EDTA, 0.5% SDS at 2ml per 7.5 x 10^7 cells) and, after another incubation of 10 min on ice, samples were stored at -80°C until required.

Chromatin samples were then sonicated (Misonic sonicator 3000) for a total time of 7.5 min as follows: amplitude 3, amplitude power 10%, sonication ON for 10s, sonication OFF for 30s.

A 50 µl aliquot of sonicated chromatin was used for reverse cross-linking in order to verify in a 0.7% agarose gel that the DNA fragments after sonication were between 200-1000 bp. DNA was quantified in the Nanodrop.

Protein content of chromatin samples was determined by BCA assay, using a standard curve of BSA. The assay was read in the Nanodrop.

2.3.4.2 Chromatin immunoprecipitation

For each chromatin immunoprecipitation (ChIP) reaction, 50 µl of magnetic DYNA beads-Protein A were required. First, the magnetic beads were blocked by

washing them with blocking buffer (0.5% BSA and 0.5% tween-20 in PBS). Then, antibody conjugation of the magnetic beds was carried by adding 2-7.5 µl of antibody in 200 µl of blocking buffer and incubating in a rotation rocker for 1 h at RT. The supernatant was removed and 400 µl of chromatin diluted in RIPA buffer (10 mM Tris-HCl pH 8, 1 mM EDTA pH 8, 0.5 mM EGTA pH 8, 140 mM NaCl, 0.1% sodium deoxycholate, 0.1% SDS, 1% triton X-100 and 1X PIK) were added to the antibody-conjugated magnetic beads. The ChIP reaction was incubated in the rotor for 3 h at 4˚C. After the incubation, the beads were washed twice with 500 µl of RIPA buffer, twice with 500 µl of RIPA-500 (10 mM Tris-HCl pH 8, 1 mM EDTA pH 8, 0.5 mM EGTA pH 8, 500 mM NaCl, 0.1% sodium deoxycholate, 0.1% SDS, 1% triton X-100, 1X PIK), once with 500 µl of RIPA-LiCl (10 mM Tris-HCl pH 8, 1 mM EDTA pH 8, 0.5 mM EGTA pH 8, 250 mM LiCl, 0.1% SDS, 0.5% NP40, 1X PIK), and twice with 500 µl of TE pH 8 (10 mM Tris-HCl pH 8 and 1 mM EDTA). The beads and the input (10 µl of untreated chromatin) were incubated with 70 µl of elution buffer (10 mM Tris-HCl pH 8, 5mM EDTA, 0.5% SDS, 300 mM NaCl) and 4 µl of proteinase K (1µl of RNase A is added only to the input) for 1 h at 55˚C and overnight at 65˚C for reverse cross-linking. The next day, DNA was purified using the Qiagen MinElute PCR pufirifiation kit and eluted in 50 µl of Qiagen elution buffer. Details of the antibodies and dilutions are listed in Table 2.7.

2.3.4.3 Real time PCR

The eluted DNA from each ChIP reaction was diluted 1:5 with 10 mM Tris-HCl pH 8. Concentration of input DNA was adjusted to 0.1 ng/µl. qPCRs were performed in 96- well plates. One qPCR consisted of 5 µl of diluted cDNA or input DNA, 12.5 µl of FastStart universal SYBR green master (with rox), 5 µl of primer mix, and 2.5 µl of nuclease-free water. The 96-well plate was centrifuged for 3 min at 1000 RFC. qPCRs were run in the Stratagen Mx 3000P thermocycler and the conditions are indicated in Table 2.4. Primer sequences and concentrations are provided in Table 2.8.

Amplification plots and dissociation curves were observed to determine the specificity and the performance of the reactions. Ct values were used in a variation of the $2^{-\Delta\Delta Ct}$ method. The first ΔCt was calculated as the difference between the Ct of each gene of the chromatin sample and the Ct of the

corresponding input DNA from the same sample. Then, the resulting ΔCt was divided by the average ΔCt of H3 from the same chromatin sample. Finally, the fold change to the input was determined as follows: $2^{-\Delta Ct1/\Delta CtaveH3}$. The input DNA controlled for amplification particularities of each genomic region, while the H3 average for chromatin amount, acting as a normaliser. Two tailed Student T test was used to calculate the statistical significance of the observed differences between the qPCR triplicates, per gene and ChIP reaction, of the *Hmgn*-knockout chromatin and the parental.

Table 2.8 ChIP-PCR primer information

Primer	Sequence (5'-3')	Final conc. (nM)
Neurog1 F	GGTGAGGAAGCTGGACAGG	300
Neurog1 R	CCCTTTGGAGACCTGCATCT	300
Ascl1 F	CGTCTCCACCTTGCTCATCT	300
Ascl1 R	TTGGTCAACCTGGGTTTTGC	300
Gata4 F	CCAACAGGCAAAGTCCATGC	300
Gata4 R	CACTGAGGGCAGAACGGAG	300
Tbxt F	TGTAATCTTTGGGCTCCGCA	600
Tbxt R	CCTACCCAACAGCCACCTTC	600
Actb F	CGCCATGGATGACGATATCG	300
Actb R	CGAAGCCGGCTTTGCACATG	300
Pou5f1-1 F	GTGAGCATGACAGAGTGGAGGAA	300
Pou5f1-1 R	TCTCTGGCCCTCTCCATGAAT	900
Pou5f1-2 F	GTGGGTAAGCAAGAACTGAGGA	300
Pou5f1-2 R	TGGAGAGCCTAAAACATCCATT	900
Pou5f1-3 F	CAATGCCGTGAAGTTGGAGA	300
Pou5f1-3 R	TCACTTACCTCCTCGGGAGTTG	900
Nanog-1 F	GGAAGAACCACTCCTACCAATACTCA	300
Nanog-1 R	CGTAACATCTCCCATGTGAAGACTC	900
Nanog-2 F	TCTTTAGATCAGAGGATGCCCCCTAAGC	300
Nanog-2 R	AAGCCTCCTACCCTACCCACCCCCTAT	300
Nanog-3 F	TCAGCCCAGTACTCAGGCTTGT	300
Nanog-3 R	AGCCTAGCAGCCTCTTGGTTCT	300
Nanog-4 F	TAACTGGACCCTCTGACTGGCT	300
Nanog-4 R	CCCACCATCTTTTCTGCTAGTACAAG	300

Chapter 3 *Hmgn*-knockout embryonal carcinoma cells spontaneously differentiate at higher frequencies and exhibit increased levels of lineage-specific markers

3.1 Introduction

HMGN1 and HMGN2 share a spatio-temporal expression pattern; they are ubiquitously expressed in all adult cells, although their levels are higher in embryonic structures and decrease upon differentiation (Crippa et al, 1991; Deng et al, 2017; Deng et al, 2013; Furusawa et al, 2006; Lehtonen & Lehtonen, 2001; Lehtonen et al, 1998; Mohamed et al, 2001; Pash et al, 1990).

Modelling embryonic development *in vitro* has been achieved by culturing ECCs and ESCs. These cells have unique properties; they are capable of self-renewal and can differentiate into all cell types present in an organism. As previously reviewed, this great plasticity relies on a transcriptional program orchestrated by a network of pluripotency-related transcription factors and their interactions with chromatin components and modulators. Furthermore, the chromatin structure of stem cells differs from that of differentiated cells in a number of ways: it is less tightly compacted (Mattout & Meshorer, 2010), with abundant DHSs (Deng et al, 2013) and marks of active transcription (Efroni et al, 2008; Guenther et al, 2007; Krejčí et al, 2009; Lee et al, 2003), fewer and less condensed heterochromatin loci (Efroni et al, 2008; Meshorer et al, 2006), and hyperdynamic binding of major architectural proteins such as linker histones (Meshorer et al, 2006) and HMGNs (Deng et al, 2013).

HMGNs are not only highly abundant and hyperdynamic, but are related with other chromatin features of stem cells as these proteins positively influence histone acetylation (Lim et al, 2005) and play a role in the establishment and maintenance of the DHS landscape of cells (Cuddapah et al, 2011; Deng et al, 2017; Deng et al, 2013; Deng et al, 2015; Zhang et al, 2016).

Based on these studies, the hypothesis is that the HMGNs play an important role in maintaining the chromatin structure of stem cells, and therefore, in safeguarding self-renewal and pluripotency.

In order to test this hypothesis, the present work studies P19 monoclonal lines lacking either HMGN1 or HMGN2. As cited in the introduction, P19 cells were derived from post-implantation embryos (McBurney & Rogers, 1982), capturing a later developmental stage than mESCs. It has been suggested that P19 cells resemble EpiSCs (Kelly & Gatie, 2017; Mallanna et al, 2008), the primed state of pluripotency sharing several features with hESCs (Kinoshita & Smith, 2018; Nichols & Smith, 2009; Smith, 2017). Working with P19 offers a number of advantages as they are maintained in culture, directed to differentiate, and importantly, genetically manipulated with ease (McBurney, 1993).

The *Hmgn*-knockout monoclonal lines were generated by previous students in the laboratory employing the CRISPR/Cas system. CRISPR stands for clustered regularly interspaced short palindromic repeats and Cas for CRISPR associated proteins. The CRISPR/Cas system was first described in bacteria and archaea as a prokaryotic adaptive immune system that provides resistance against phage infections (Barrangou et al, 2007), and has been recently modified to induce mutagenesis in mammalian cells (Cong et al, 2013; Mali et al, 2013). Briefly, a guide RNA (gRNA), complementary in sequence to the gene of interest in the host cell, recruits the Cas9 nuclease to produce a targeted double DNA strand break. The DNA repair machinery of the host cell will recognise this double strand break and, while repairing it, might introduce mutations, such as insertions or deletions, leading to a change in the open reading frame of the gene. Hence, if transcription proceeds, aberrant transcripts will be synthesised from the mutated gene that either will not translate into proteins or will encode for peptides different in sequence and function.

The strategy used in the laboratory consisted of targeting mutations around the ATG start codon at exon 1 of the *Hmgn2* gene, with the aim of introducing a frameshift switch right at the start of the coding region (Sindi, 2017). P19 cells were transfected with three different plasmids; two containing gRNAs targeting the parallel and antiparallel DNA strands at the site of interest, and a third one encoding for the GFP reporter and the Cas9 nickase, a modified version of the Cas9 nuclease that only has a single active site. The concept of the CRISPR nickase is to introduce two single strand breaks close enough for the cells to still interpret these as a double strand break, while enhancing the specificity of the

system and decreasing off target effects (Ran et al, 2013). After transfection, cells were sorted by FACS and GFP-positive cells were plated at limiting dilution to generate clonal lines. Three genetically distinct lines lacking the HMGN2 protein (B8ΔN2, B19ΔN2, and B38ΔN2) were isolated, and the *Hmgn*2 locus was sequenced to confirm the gene disruption (Sindi, 2017). Additionally, a monoclonal line (B1), which was taken through all the procedures but was found to preserve HMGN2, was analysed to confirm the WT genotype and utilised as a control line (Sindi, 2017).

The second strategy employed the WT Cas9 nuclease, a gRNA directed to the ATG start codon at exon 1 of the *Hmgn1* gene, and an extra gRNA targeting the *Hprt* gene (Eden, West, unpublished). Hypoxanthine phosphoribosyl transferase (HPRT) metabolises 6-thioguanine into a toxic product leading to cell death. The *Hprt*-targeted cells are resistant to 6-thioguanine and can be selected, guaranteeing the presence of a functional operating CRISPR/Cas system, and thus, improving the probabilities of finding an *Hmgn1*-knockout cell (Liao et al, 2015). Two cell lines lacking the HMGN1 protein were derived by this method (N1-1ΔN1 and D3ΔN1), in parallel with a control line transfected with the *Hprt*-gRNA and the Cas9 nuclease (C2) (Eden, West, unpublished). It is worth mentioning that these three cell lines were generated later than the *Hmgn2*-knockout lines, and therefore not all the experiments shown here have included them.

3.2 *Hmgn1*- and *Hmgn2*- targeted embryonal carcinoma cells have lost HMGN1 and HMGN2 proteins, respectively

The first task of the present work was to confirm that the different P19 cell lines engineered by CRISPR/Cas were in fact expressing or lacking HMGN1 and/or HMGN2 proteins, respectively. RNA and protein extracts were prepared from multiple independent cultures. Additionally, cells were fixed and IF was performed for the detection of the two HMGN variants. Relevant information corresponding to these cell lines is summarised in Table 3.1.

Table 3.1 Nomenclature and description of the cell lines analysed in the present work

Cell line	Description
Parental	A P19 polypopulation of early passage number, from which all the following monoclonal lines originated.
B1	A control clonal line derived in parallel with the *Hmgn2*-knockout lines, in which the CRISPR/Cas9 nickase failed to induce mutagenesis as proven by sequencing.
C2	A clonal line in which *Hprt* was targeted with CRISPR/Cas9 nuclease and that was selected with 6-thioguanine.
N1-1ΔN1	*Hmgn1*-knockout monoclonal lines that were generated by targeting *Hmgn1* and *Hprt* genes with the CRISPR/Cas9 nuclease system and selecting with 6-thioguanine.
D3ΔN1	
B8ΔN2	*Hmgn2*-knockout monoclonal lines that resulted from successful mutagenesis in the *Hmgn2* gene accomplished by the CRISPR/Cas9 nickase.
B19ΔN2	
B38ΔN2	

As shown in Figure 3.1A, mutations in the first exon of *Hmgn1* significantly reduce the gene expression in N1-1ΔN1 and D3ΔN1 cells when compared with parental cells. However, there is still an *Hmgn1* transcript as determined by qRT-PCR and dissociation curve analysis of the amplicon. Similar results are observed with *Hmgn2* expression, which is three, five, and six fold less expressed in B8ΔN2, B19ΔN2, and B38ΔN2 cells, respectively, than in parental cells (Figure 3.1A).

Despite the fact that transcription of *Hmgn1* or *Hmgn2* is not fully blocked, neither HMGN1 protein in N1-1ΔN1 and D3ΔN1 cells, nor HMGN2 protein in B8ΔN2, B19ΔN2, and B38ΔN2 cells, are synthesised as demonstrated by WB in Figure 3.1B and C.

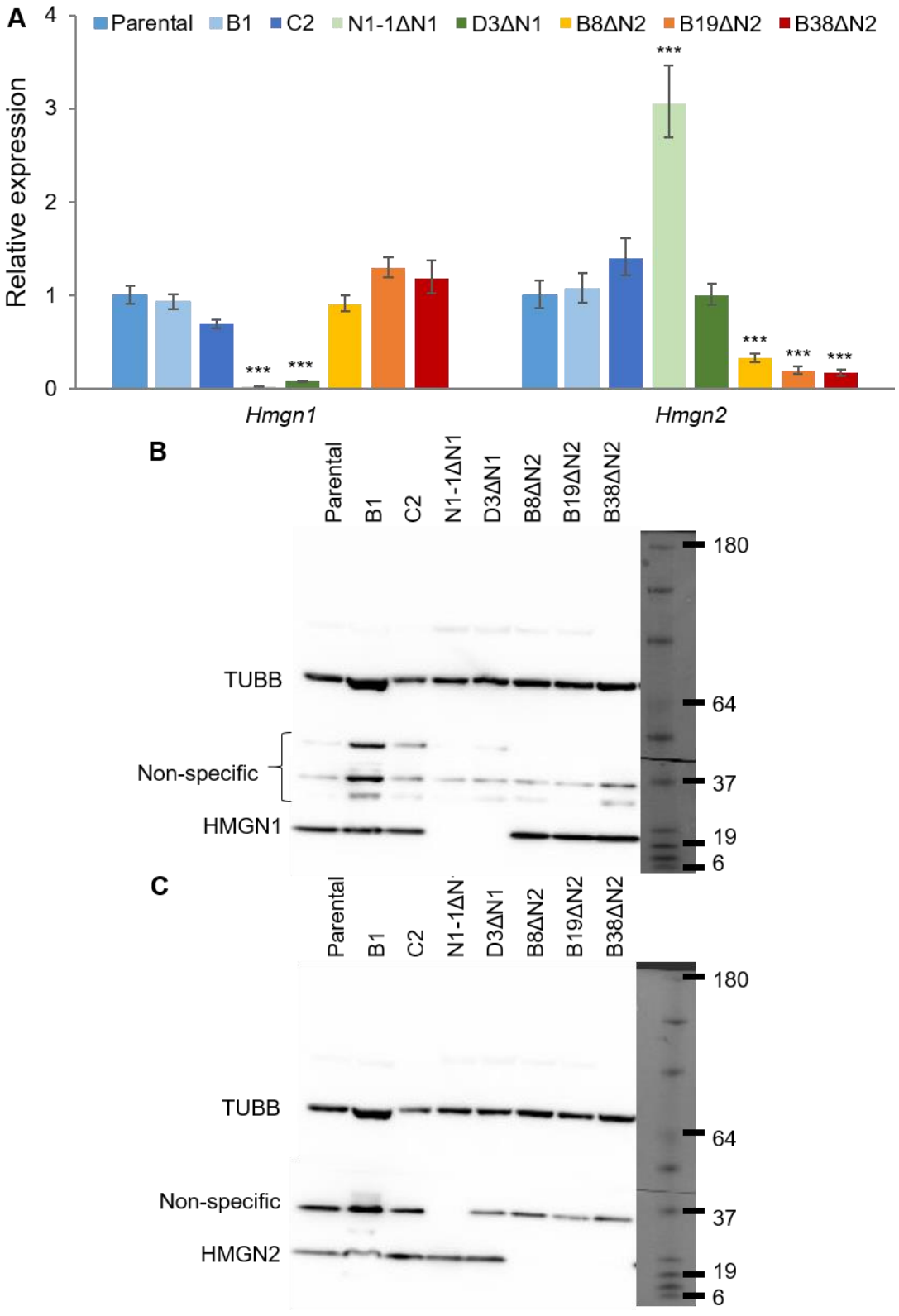


Figure 3.1 Gene expression and protein detection of the two major HMGN variants in *Hmgn1*- and *Hmgn2*-targeted ECCs
A) Relative expression of *Hmgn1* and *Hmgn2* as determined by qRT-PCR in control and *Hmgn*-knockout cells. The graph represents the fold change in comparison with parental cells, error bars symbolise the SEM from 3 to 10 independent cultures (with the cooperation of A. Sindi), the statistical significance was calculated by ANOVA and Dunnett's multiple comparison test (adjusted p value ***<0.0001). WB for the detection of (B) HMGN1 and (C) HMGN2 proteins, including TUBB as a loading control.

Interestingly, *Hmgn2*-knockout cells do not seem to compensate their loss by increasing transcription or translation of HMGN1 (Figure 3.1A and B). Neither do D3ΔN1 cells synthesise additional HMGN2 protein/mRNA. The expression of *Hmgn2* in N1-1ΔN1 cells, in contrast, is three times higher than in parental cells, although this is not translated into higher protein levels (Figure 3.1A and C). The fact that the overexpression of the residual HMGN variant is only observed in one cell line suggests a clonal variation in transcription rates, rather than a general mechanism of compensation. Indeed, HMGN1 and HMGN2 protein levels are unchanged in cells and tissues derived from *Hmgn2*- and *Hmgn1*-knockout mice, respectively (Birger et al, 2003; Deng et al, 2015).

HMGNs are chromatin architectural proteins that bind to nucleosomes, and therefore, are actively transported into the nuclei of the cells via an intrinsic bipartite nuclear localisation signal (Hock et al, 1998a). When HMGNs are detected by IF, the anti-HMGN2 antibody is highly specific resulting in a bright nuclear signal, easily distinguishable from the background (Figure 3.2). The nuclei of all cells in control and in N1-1ΔN1 and D3ΔN1 cultures contain HMGN2 (Figure 3.2). In contrast, there is no nuclear signal detected in B8ΔN2, B19ΔN2, and B38ΔN2 cells above the background (Figure 3.2).

The anti-HMGN1 antibody, however, produces a more dispersed signal and a higher cytoplasmic background. Nevertheless, the nuclei of parental, B1, C2, B8ΔN2, and B19ΔN2 cells are brighter than the cytoplasm (Figure 3.2). Conversely, N1-1ΔN1 and D3ΔN1 cells have completely lost the nuclear signal (Figure 3.2, magnifications) and merely the cytoplasmic background can be detected.

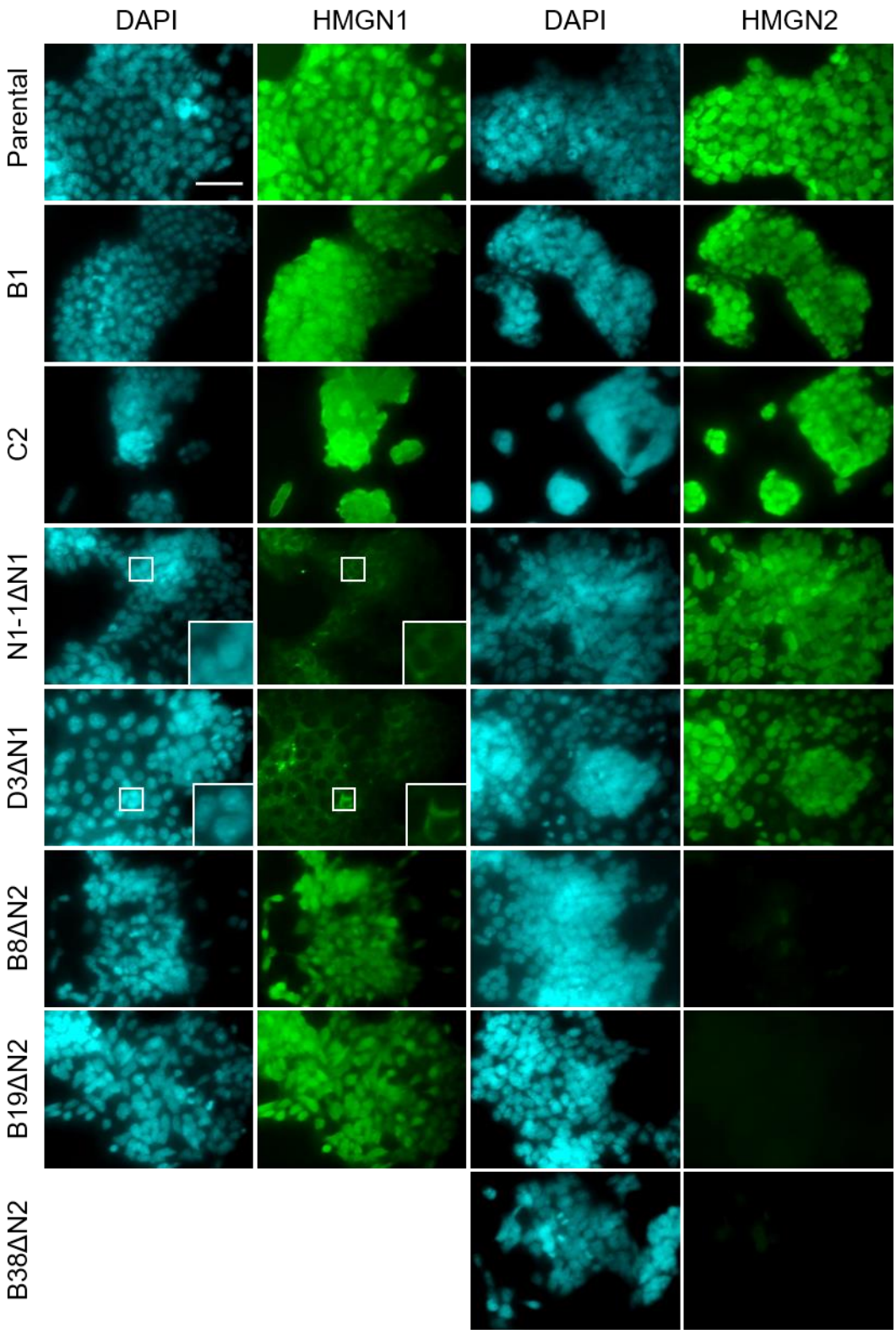


Figure 3.2 *Hmgn1*- and *Hmgn2*-targeted ECCs have lost HMGN1 or HMGN2 nuclear proteins
IF for the detection of HMGN1 and HMGN2 proteins (green). DAPI was used to stain the nuclei (cyan). Control cells contain both chromatin binding proteins. B8ΔN2, B19ΔN2, and B38ΔN2 cells lack HMGN2, but are positives for HMGN1; whereas N1-1ΔN1 and D3ΔN1 cells retain HMGN2. The 3X magnifications of the indicated cells highlight the loss of the nuclear signal corresponding to HMGN1 protein in N1-1ΔN1 and D3ΔN1 cells. Scale bar indicates 50 µm.

3.3 *Hmgn*-knockout cultures present reduced pluripotency markers and signs of spontaneous differentiation

As described in the introduction, ESCs can be propagated *in vitro* under feeder-free conditions and different formulation of the media. The requirement for feeder cells is replaced by the addition of LIF to serum-supplemented media (Smith et al, 1988); alternatively, ectopic BMPs can substitute for BMPs contained in the serum (Ying et al, 2003a). LIF and BMPs work in conjunction to support self-renewal by preventing differentiation into non-neural and neural lineages, respectively (Ying et al, 2003a). The notion that these factors shield ESC cells from differentiation derived from the development of a chemically defined media complimented with small molecule inhibitors of two kinases, MAP2K and GSK3. These small molecules directly restrict differentiation pathways and enhance pluripotency in all cells, as indicated by homogeneous morphology and pluripotency-related transcription factor levels (Wray et al, 2010; Ying et al, 2008).

P19 cells, in contrast, can self-renew indefinitely in the laboratory under feeder-free conditions in serum-supplemented media. They have lost the capacity to spontaneously differentiate, despite their teratocarcinoma origin, and can be propagated as mostly pure cultures of undifferentiated cells. As shown in Figure 3.3, parental cells are morphologically homogeneous and form colonies where only the cells at the edges show minor cytoplasmic protuberances (arrowheads). B1 cells also form colonies, although these seem to be more compact, and the morphology of the cells located at the centre is uniform, whereas the morphology of the cells at the edges or outside the colonies is heterogeneous (Figure 3.3).

The loss of a major HMGN variant, however, severely affects the cellular morphology and organisation. Fewer colonies are observed (Figure 3.3, white arrows) as a high proportion of *Hmgn*-knockout cells spread outside the colonies. These cells exhibit an extended cytoplasm and resemble differentiated cells, such as endothelial flat cells (Figure 3.3, orange arrow), and neural rosettes (Figure 3.3, yellow arrow); for examples of differentiated cells see (Trott & Martinez Arias, 2013). Interestingly, cells with cytoplasmic protuberances are

observed surrounding the round and compact ESC colonies cultured in serum/LIF media; these cells have lost alkaline phosphatase reactivity, which indicates they have exited the pluripotency state (Wray et al, 2010). Thus, it is possible to hypothesise that some of the *Hmgn*-knockout cells have progressed along differentiation programs. Nevertheless, a substantial proportion of cells in the *Hmgn*-knockout cultures can self-renew, as the cell lines can be propagated indefinitely, although at smaller split ratios than parental, B1, and C2 cells.

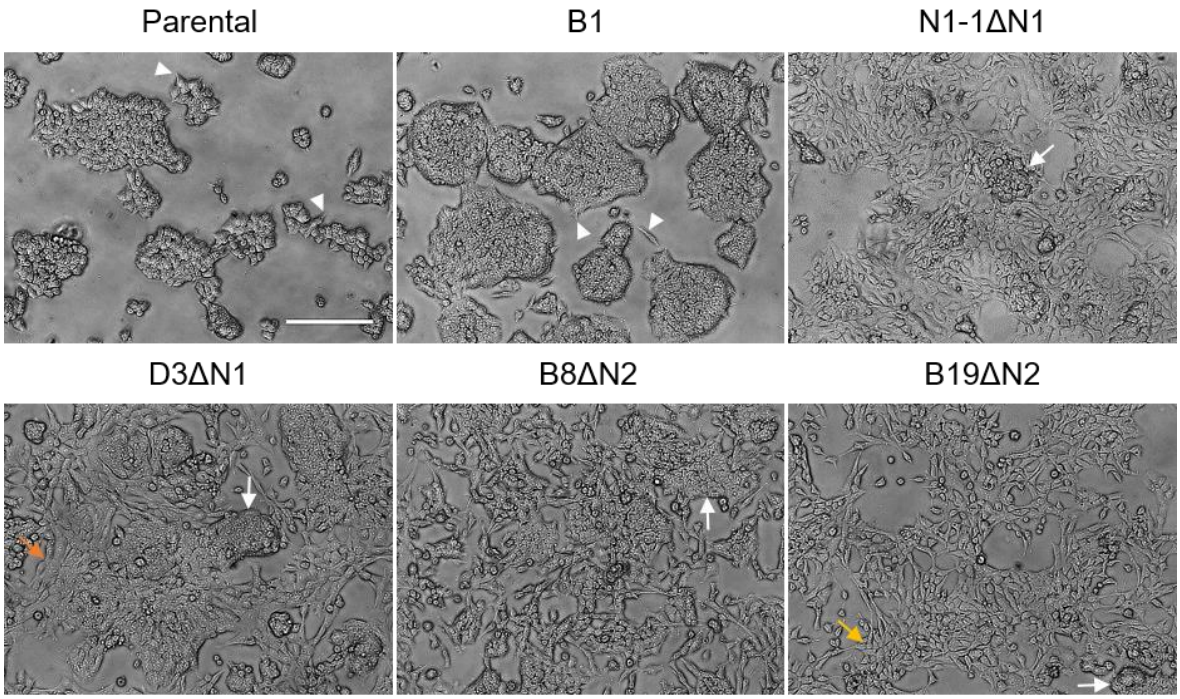


Figure 3.3 Colony morphology reveals substantial differences in cellular shape and organisation between control and *Hmgn*-knockout cells
Bright field pictures of control and *Hmgn*-knockout cultures. Parental and B1 cells are morphologically homogeneous and form colonies, where the cells at the edges show minor cytoplasmic protuberances (arrowheads). Fewer colonies can be observed in the *Hmgn*-knockout cultures (white arrows) as the cells grow spread outside the colonies. The morphology of *Hmgn*-knockout cells is heterogeneous and resembles endothelial cells (orange arrow) and neural rosettes (yellow arrow). Scale bar indicates 200 µm.

To evaluate whether the *Hmgn*-knockout cells are spontaneously differentiating, cells were fixed 24 and 48 h after seeding in normal P19 media, and IF was performed for the detection of pluripotency and lineage-specific markers. P19 cells are passaged every two days, and therefore, the two time points selected represent cultures of lower and higher confluency, respectively.

DAPI staining recapitulates what observed on bright field pictures. At 48 h, parental and B1 colonies have uniformly expanded as a monolayer, while *Hmgn*-knockout cells grow in three dimensional clusters (Figure 3.4A, arrows) surrounded by numerous cells that spread between these clusters (Figure 3.4A, arrowheads).

Stage-specific embryonic antigen 1 (SSEA1) plays a role in cell adhesion and migration in the pre-implantation embryo, is widely used as mouse ESC marker, and is also expressed by ECCs (Solter & Knowles, 1978; Zhao et al, 2012). As shown in Figure 3.4A, numerous *Hmgn*-knockout cells have lost this pluripotency marker, especially those spread between the clusters (arrowheads); some of the cells within these clusters, however, retain the signal (arrows), although it not as intense as in control cells.

FACS analysis shows that *Hmgn*-knockout cells contain variable levels of SSEA1, which is particularly evident in the extended D3ΔN1 and B8ΔN2 populations (Figure 3.4B). Importantly, most of D3ΔN1 cells are represented in regions of lower fluorescence intensity than parental cells (Figure 3.4B). Fluorescence intensity is directly proportional to the amount of SSEA1 in the cells, and quantification of the median fluorescence intensity of parental, B1, and *Hmgn*-knockout populations reveals that all monoclonal lines have reduced levels of the pluripotency marker (Figure 3.4C). *Hmgn1*-knockout cells display the lower SSEA1 protein levels, between 40 and 50% of those of the parental cells, while *Hmgn2*-knockout and B1 cells present a 25 to 30% reduction (Figure 3.4C). The numerous SSEA1-negative cells that are observed in the *Hmgn*-knockout cultures, especially in D3ΔN1 (Figure 3.4A), are expected to contribute to the reduction of the total fluorescence intensity of the population.

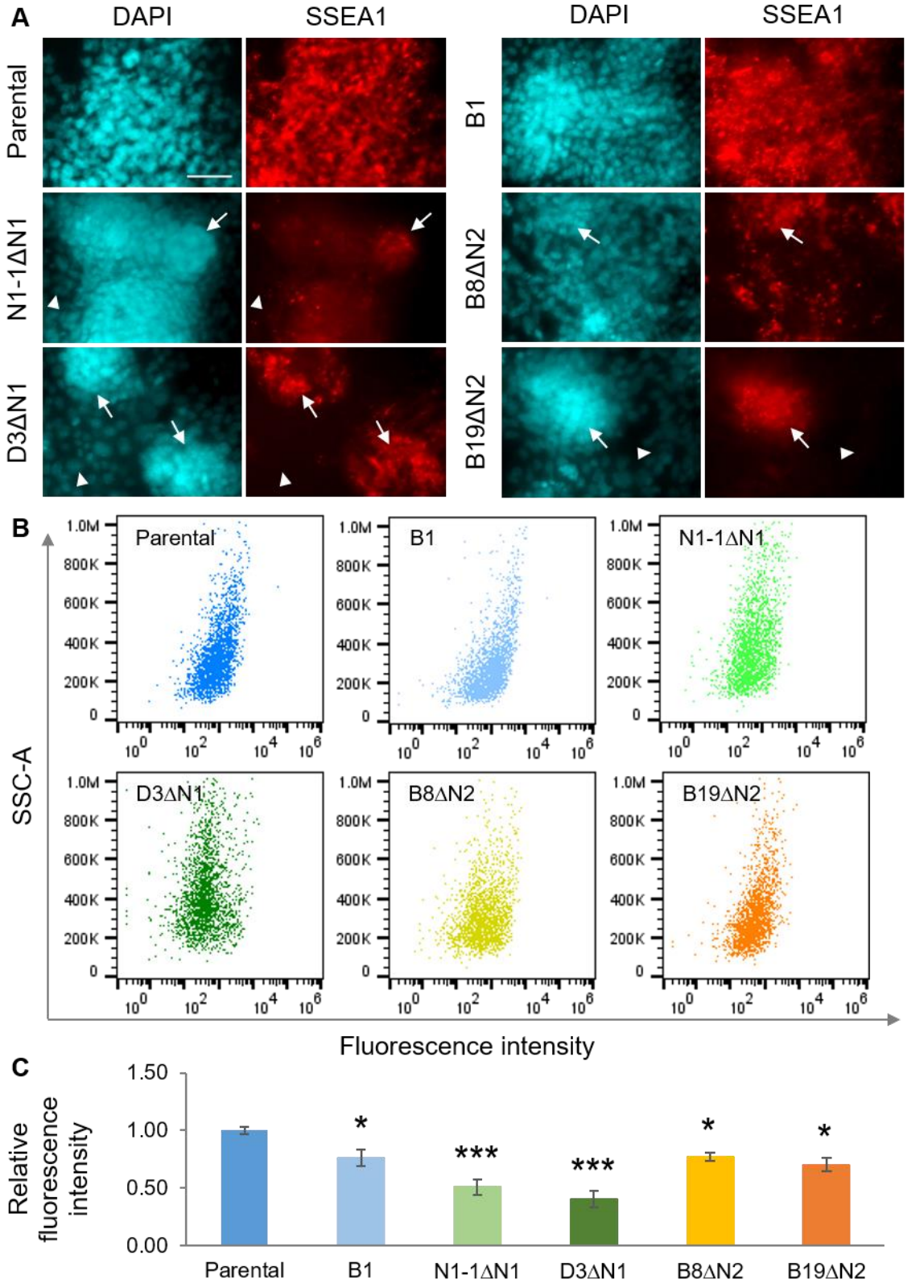


Figure 3.4 The pluripotency marker SSEA1 is lost in numerous *Hmgn*-knockout cells
Parental, B1, and *Hmgn*-knockout cells were fixed 48 h after seeding for the detection of the pluripotency marker SSEA1. A) IF show that clustered *Hmgn*-knockout cells retain higher levels of SSEA1 (arrows) than the cells that grow spread outside colonies (arrowheads). DAPI was used to stain the nuclei (cyan). Scale bar indicates 50 µm. B) Immunostaining of SSEA1, as determined by FACS. The dot plots show the fluorescence intensity displayed by parental, B1, and *Hmgn*-knockout cells. C) Quantification of the florescence intensity. The graph represents the relative fluorescence intensity (median) of each cell line, error bars symbolise the SEM from 3 independent experiments, and the statistical significance was calculated by ANOVA and Dunnett's multiple comparison test (adjusted p values *<0.05, **<0.001, ***<0.0001).

The naïve pluripotency transcription factor NANOG is heterogeneously expressed by parental cells as observed in Figure 3.5A, where some of the cells are brighter than others, mostly located at the edges of the colonies. FACS analysis are in line with the previous observations. Two subpopulations are clearly distinguished (Figure 3.5B), the larger displaying higher fluorescence intensity (arrow) than the shorter (arrowhead). This is in complete agreement with previous studies reporting that in pluripotent cell cultures around 20% of the population express lower NANOG levels, and suggesting that NANOG heterogeneity is related to diverse differentiation potentials within a population of pluripotent cells (Abranches et al, 2014; Chambers et al, 2007; Herberg et al, 2016; Kalmar et al, 2009).

The fluorescence intensity is 50 to 80% less in *Hmgn*-knockout cultures, and in B1 cells, indicating decreased NANOG levels in these cells (Figure 3.5C). However, in the IF it is clear that the cultures preserve some cells that have high NANOG protein levels (Figure 3.5A arrows) whereas others have completely lost the protein (Figure 3.5A arrowheads). The variability in the protein levels is consistent with the heterogeneity of this pluripotency transcription factor, while the reduction in the total protein levels suggests that pluripotency has moved closer to differentiation. As observed with SSEA1, most of the negatively-stained cells are those located outside the colonies (Figure 3.5A arrowheads).

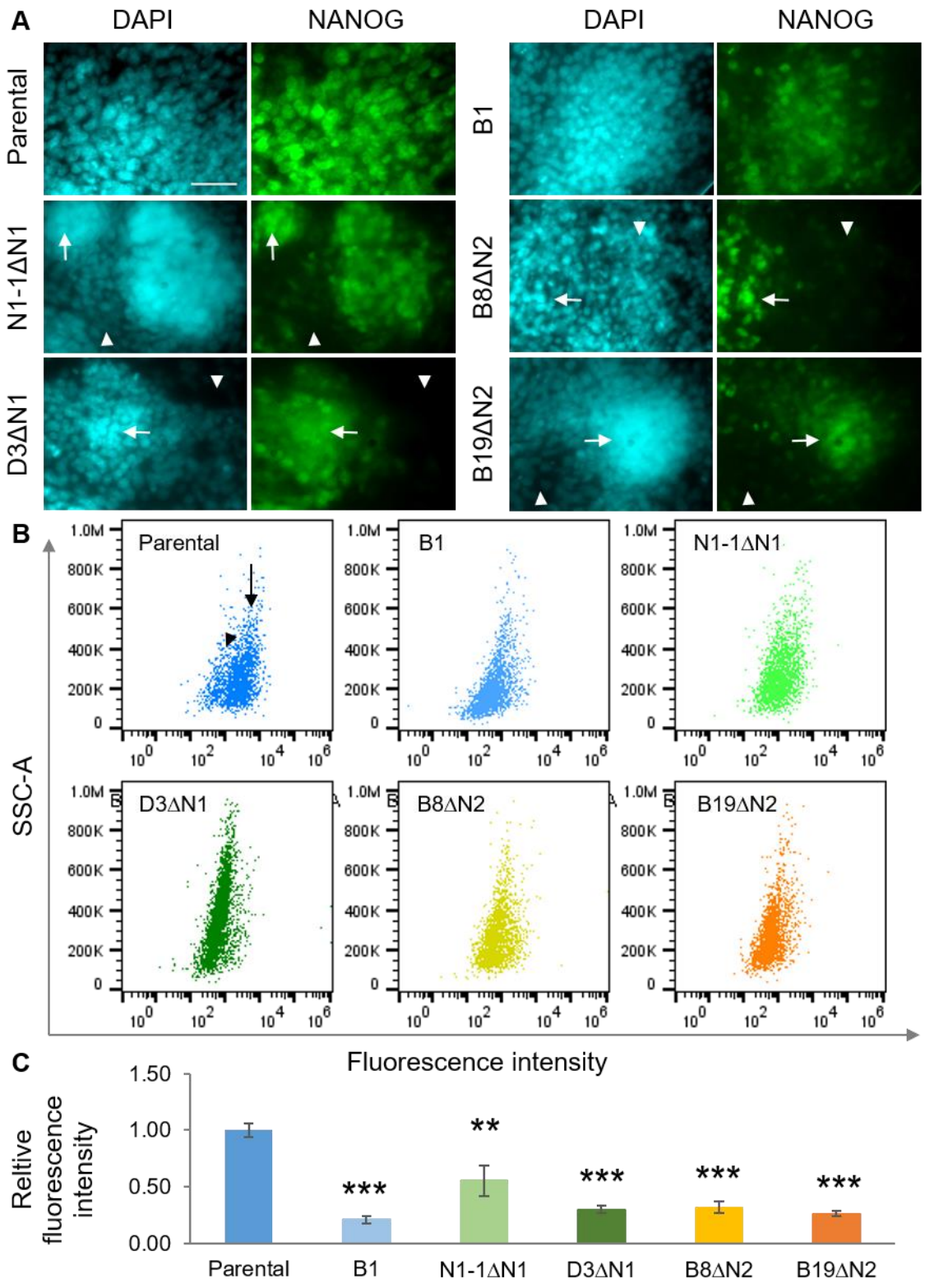


Figure 3.5 The naïve pluripotency transcription factor NANOG is heterogeneously expressed by P19 cells and its levels are reduced in *Hmgn*-knockout cultures
Parental, B1, and *Hmgn*-knockout cells were fixed 48 h after seeding for the detection of the naïve pluripotency transcription factor NANOG. A) IF show that NANOG levels are heterogeneous in all cell lines, as marked by a higher (arrows) and lower or none (arrowheads) fluorescence intensity. DAPI was used to stain the nuclei (cyan). Scale bar indicates 50 µm. B) Immunostaining of NANOG, as determined by FACS. The dot plots show the fluorescence intensity displayed by parental, B1, and *Hmgn*-knockout cells. C) Quantification of the florescence intensity. The graph represents the relative fluorescence intensity (median) of each cell line, error bars symbolise the SEM from 3 independent experiments, and the statistical significance was calculated by ANOVA and Dunnett's multiple comparison test (adjusted p values *<0.05, **<0.001, ***<0.0001).

SSEA1 and NANOG immunostaining suggest that pluripotency is compromised in *Hmgn*-knockout cultures. Nevertheless, IF and FACS analysis reveal that the master regulator of pluripotency in ESCs, POU5F1, is expressed at similar levels in most of the cell lines, and most of the cells display equivalent fluorescence intensity (Figure 3.6).

Some of the *Hmgn*-knockout cells, however, are completely devoid of POU5F1, which are especially abundant in D3ΔN1 cultures, but also can be identified in B8ΔN2 cultures (Figure 3.6A arrowheads). Interestingly, FACS analysis shows two subpopulations of D3ΔN1 and B8ΔN2 cells (Figure 3.6B). In the case of D3ΔN1, the larger subpopulation exhibits lower fluorescence intensity (Figure 3.6B arrowhead), which is reflected in a reduced median fluorescence intensity that represents approximately 56% of the parental (Figure 3.6C). In contrast, the larger subpopulation of B8ΔN2 cells expresses higher levels of POU5F1 (Figure 3.6B arrow), and therefore, the fluorescence intensity of the total population is similar to parental cells (Figure 3.6C).

In summary, a high proportion of D3ΔN1 cells and few B8ΔN2 cells have definitely abandoned pluripotency, while N1-1ΔN1 and B19ΔN2 cells contain POU5F1 at similar levels to parental cells. Recent findings have shown that POU5F1 is retained in the first stages of ESC differentiation (Kalkan et al, 2017; Thomson et al, 2011), therefore, the presence of POU5F1 does not exclude the possibility that some of the N1-1ΔN1 and B19ΔN2 cells have abandoned pluripotency to initiate differentiation programs.

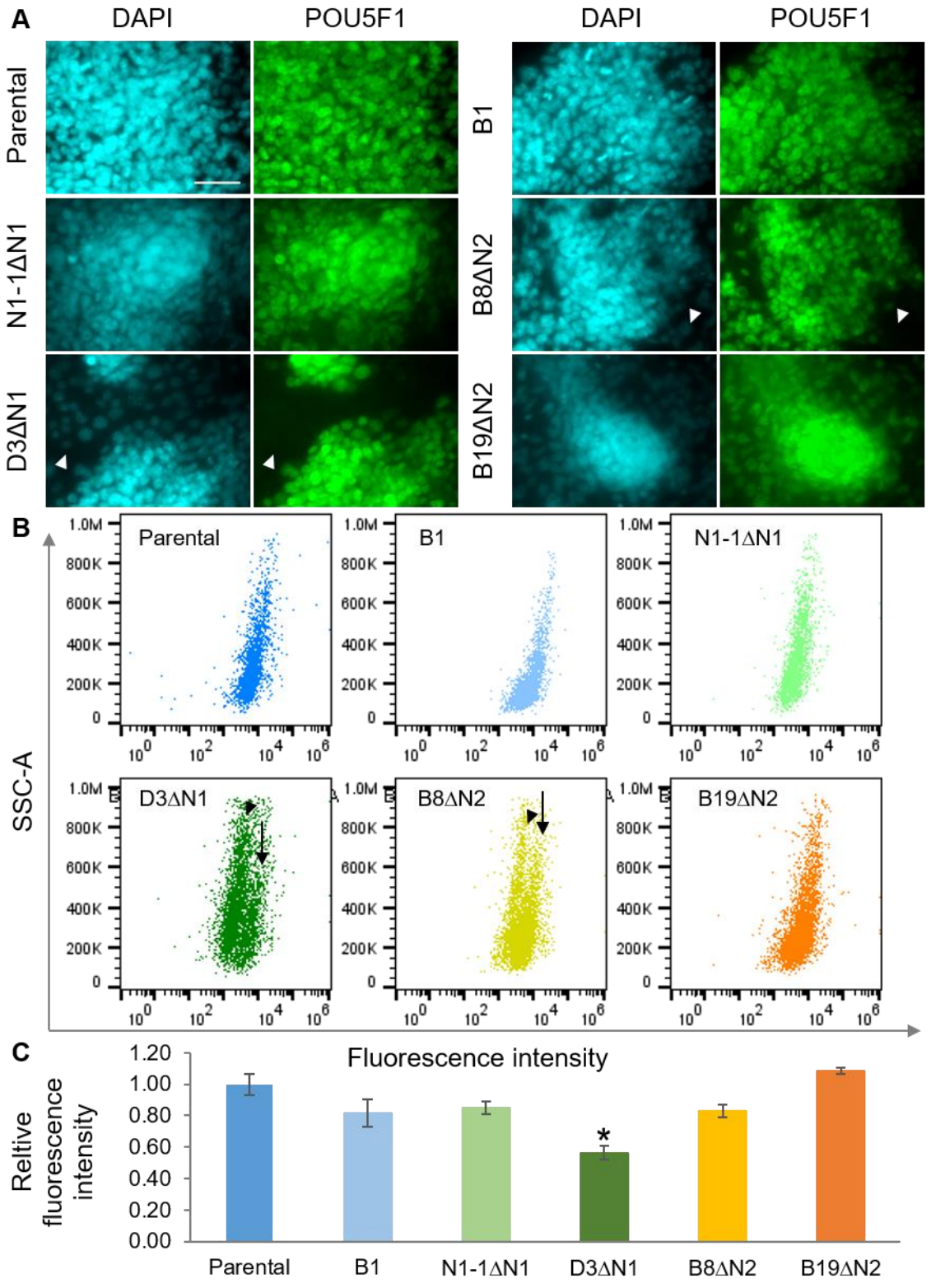


Figure 3.6 The pluripotency transcription factor POU5F1 is reduced in some *Hmgn*-knockout cultures

Parental, B1, and *Hmgn*-knockout cells were fixed 48 h after seeding for the detection of the pluripotency transcription factor POU5F1. A) IF show that most of the cells are positives for POU5F1, except some D3ΔN1 and B8ΔN2 cells (arrowheads). DAPI was used to stain the nuclei (cyan). Scale bar indicates 50 µm. B) Immunostaining of POU5F1, as determined by FACS. The dot plots show the fluorescence intensity displayed by parental, B1, and *Hmgn*-knockout cells. C) Quantification of the florescence intensity. The graph represents the relative fluorescence intensity (median) of each cell line, error bars symbolise the SEM from 3 independent experiments, and the statistical significance was calculated by ANOVA and Dunnett's multiple comparison test (adjusted p values *<0.05, **<0.001, ***<0.0001).

NES is an intermediate filament used as a classical marker of NSCs (Lendahl et al, 1990) that is also transiently expressed by other precursor and endothelial cells during development (Sejersen & Lendahl, 1993; Suzuki et al, 2010; Wroblewski et al, 1997). P19 undifferentiated cells express NES at basal levels (Jin et al, 2009b; Jin et al, 2006); however, as shown in Figure 3.7, NES levels seem to be higher in *Hmgn*-knockout cultures, particularly in cells in the three dimensional clusters (arrows). Furthermore, NES immunostaining evidences substantial morphological differences between control and *Hmgn*-knockout cells, as the cytoskeleton of the *Hmgn*-knockout cells is more extended (Figure 3.7, magnifications). Whether the NES-positive cells represent undifferentiated cells that have altered morphology or differentiation propensities, or whether they are neural or other precursor cells is not clear. Hence, the detection of lineage-specific markers is required to derivate accurate conclusions.

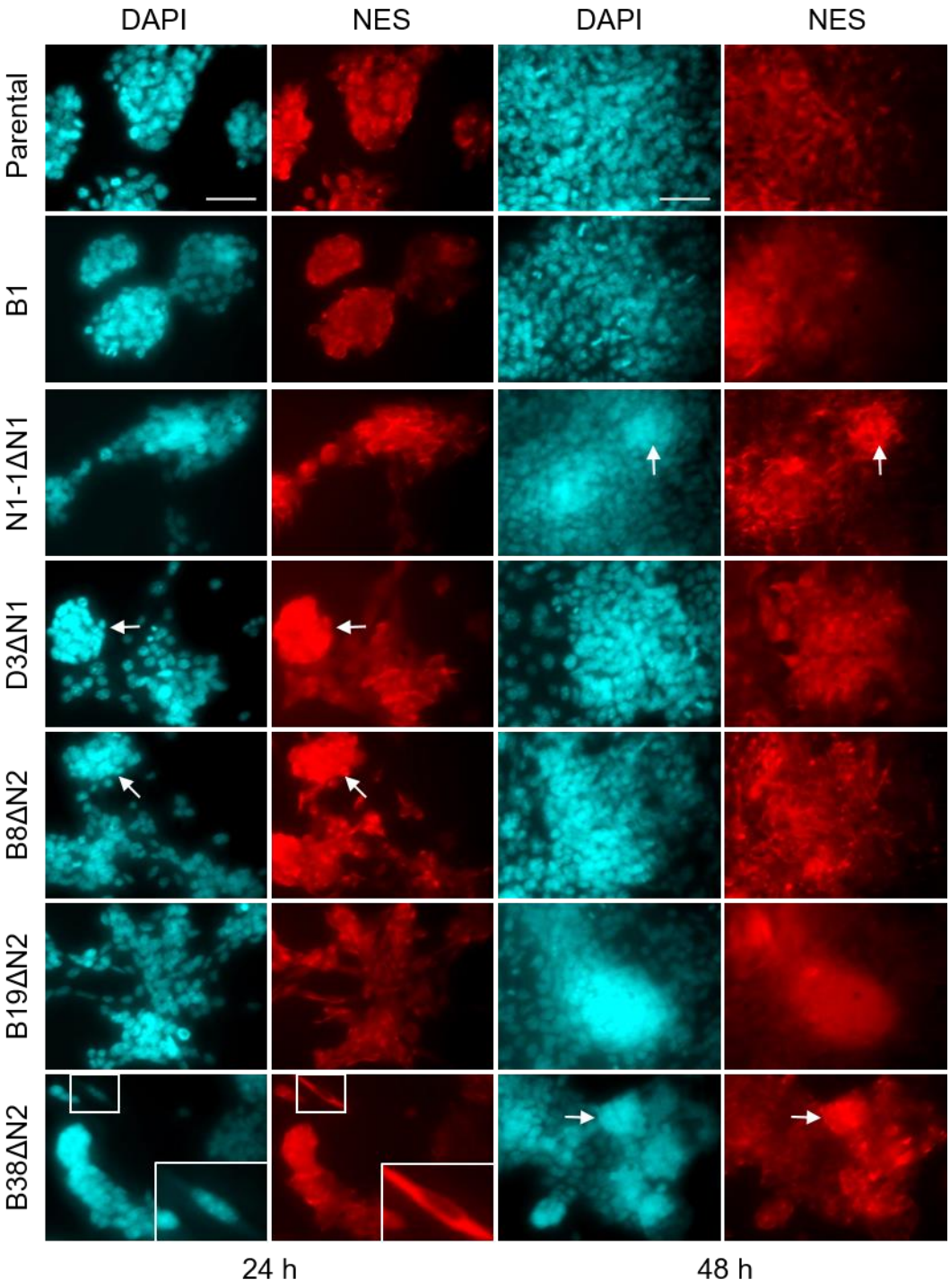


Figure 3.7 The intermediate filament NES is increased in *Hmgn*-knockout cultures
IF for the detection of the intermediate filament NES. Parental, B1, and *Hmgn*-knockout cells were fixed 24 and 48 h after seeding. C2 cells were just used 24 h after seeding. Fluorescence intensity corresponding to NES is higher in *Hmgn*-knockout cultures, especially in the three dimensional clusters (arrows). 2X magnifications of the specified area reveal cytoskeleton projections in *Hmgn*-knockout cells. DAPI was used to stain the nuclei (cyan). Scale bar indicates 50 µm.

IF for the neurofilament tubulin beta class III (TUBB3) reveals an increased frequency of spontaneous neuronal differentiation in *Hmgn*-knockout cultures as early as in 24 h (Figure 3.8). Cellular density and the incubation time impact the number of differentiated cells. At 48h, the *Hmgn*-knockout cultures display neuronal cells with complex cytoskeleton morphology, extending numerous neurites that may establish cell contacts. In contrast, few neurons are found in parental and B1 cultures at the same time point (Figure 3.8). In the *Hmgn*-knockout cultures, however, neurons are evident at both time points analysed (Figure 3.8), discarding the possibility that high confluence is the main factor driving the neuronal differentiation of *Hmgn*-knockout cells, and supporting the hypothesis that the loss of a major HMGN variant leads to spontaneous differentiation of ECCs.

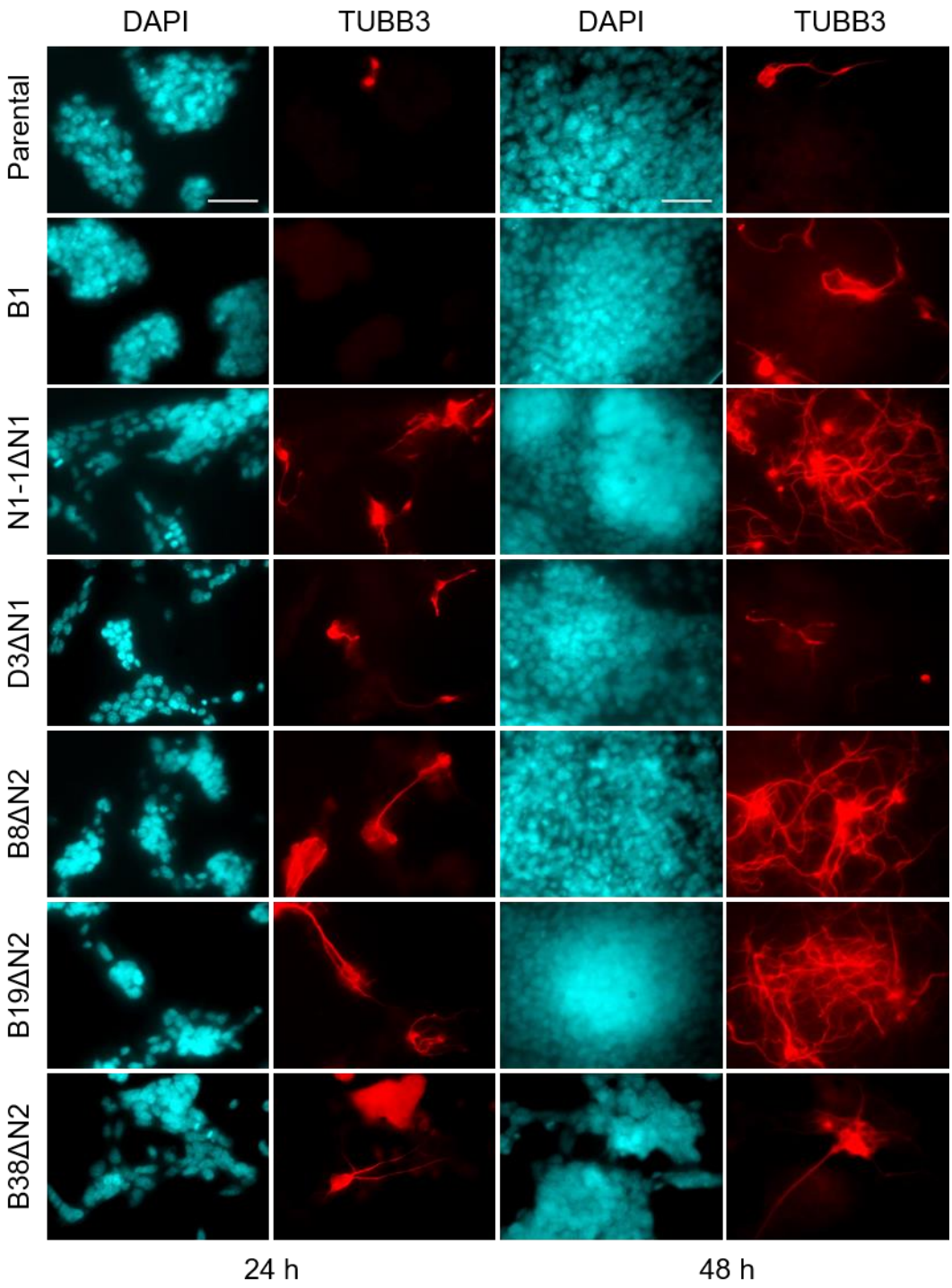


Figure 3.8 The neurofilament TUBB3 reveal the presence of numerous neurons and neurite outgrowth in *Hmgn*-knockout cultures
IF for the detection of the neurofilament TUBB3 (red). Parental, B1, and *Hmgn*-knockout cells were fixed 24 and 48 h after seeding. C2 cells were just used 24 h after seeding. DAPI was used to stain the nuclei (cyan). Scale bar indicates 50 µm.

To test whether *Hmgn*-knockout cells exclusively follow the neuronal differentiation program, or whether they alternatively progress towards other lineages, an anti-GATA4 antibody was acquired. GATA4 is expressed by the primitive streak, from where the definitive endoderm and the mesoderm arise, and also by the primitive endoderm that develops into extra-embryonic tissues (Arceci et al, 1993; Fujikura et al, 2002; Laverriere et al, 1994). Very few GATA4-positive cells are observed across all cultures and conditions (Figure 3.9A); the D3ΔN1 line displays the highest proportion of GATA4-positive cells, and these are preferentially spread between the clusters (Figure 3.9A arrows). These spread cells are highly abundant in D3ΔN1 cultures and are devoid of pluripotency markers (Figure 3.4A, Figure 3.5A, and Figure 3.6A), however, not all of them are stained with the GATA4 antibody (Figure 3.9A). Furthermore, D3ΔN1 cultures show fewer neurons than other *Hmgn*-knockout cultures (Figure 3.8), indicating that the commitment to certain lineages may differ between *Hmgn*-knockout lines.

FACS analysis are in line with the IF results, most of the cell lines display similar fluorescence intensity that may correspond to the background of the antibody, and are observed in the dot plots as compact populations (Figure 3.9B and C). Conversely, D3ΔN1 population is extended, which suggests variable GATA4 protein levels (Figure 3.9B). The previous observation indicates that not all the cells are GATA4-positives, as observed in the IF (Figure 3.9A); however, they are sufficient to shift the median of the fluorescence intensity 1.5 times higher than the parental value (Figure 3.9C).

Taken together the imaging analyses suggest that the loss of either HMGN1 or HMGN2 compromises pluripotency of ECCs, leading to the initiation of differentiation programs mainly, but not exclusively, towards neural lineages.

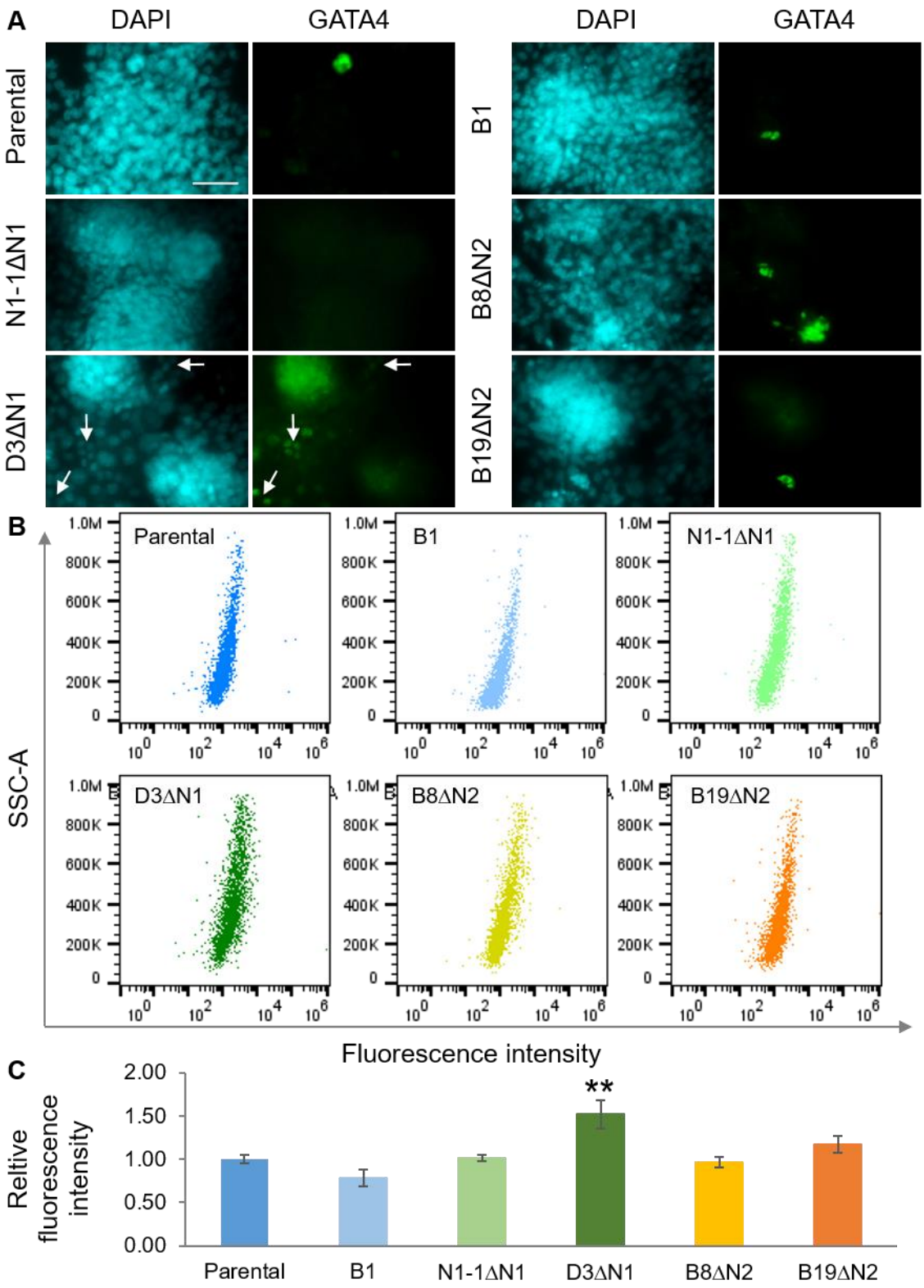


Figure 3.9 The protein levels of the endoderm marker GATA4 remain unchanged in most of the *Hmgn*-knockout cultures

Parental, B1, and *Hmgn*-knockout cells were fixed 48 h after seeding for the detection of the endoderm transcription factor GATA4. A) IF show similar levels of GATA4 in all the cultures, except in those of D3ΔN1 cells, where more GATA4-positive cells are identified (arrows). DAPI was used to stain the nuclei (cyan). Scale bar indicates 50 µm. B) Immunostaining of GATA4, as determined by FACS. The dot plots show the fluorescence intensity displayed by parental, B1, and *Hmgn*-knockout cells. C) Quantification of the florescence intensity. The graph represents the relative fluorescence intensity (median) of each cell line, error bars symbolise the SEM from 3 independent experiments, and the statistical significance was calculated by ANOVA and Dunnett's multiple comparison test (adjusted p values *<0.05, **<0.001, ***<0.0001).

3.4 *Hmgn*-knockout cells express higher levels of lineage-specific genes

To further investigate whether the loss of HMGN1 and HMGN2 compromises pluripotency of P19 cells, gene expression analyses were performed by isolating RNA 48 h after seeding the cells from several independent cultures.

Consistent with the immunostaining analyses where most of the cultured cells show uniform POU5F1 signal intensity, *Pou5f1* transcript levels are similar across all cultured lines, except in D3ΔN1 in which *Pou5f1* expression is decreased by 40% (adjusted p value of 0.007) (Figure 3.10A). It is noteworthy that a substantial proportion of D3ΔN1 cells have completely lost POU5F1 protein (Figure 3.6), raising the possibility that the rest of the cells express *Pou5f1* at equivalent levels to the parental cells, and that the observed reduction in mRNA levels derives from the average of *Pou5f1*-expressing and *Pou5f1*-negative cells.

Sox2 has a similar expression pattern to *Pou5f1* along the different pluripotency states and early ESC differentiation (Kalkan et al, 2017; Thomson et al, 2011). In agreement, *Sox2* is downregulated only in D3ΔN1 cultures (Figure 3.10A), where its levels are five times less than in parental cultures (adjusted p value <0.0001). Intriguingly, SOX2 is a marker of NSCs, and D3ΔN1 cells are not particularly neurogenic (Figure 3.8), which perhaps explains the further downregulation of *Sox2* in these cells. Furthermore, the highly neurogenic N1-1ΔN1 line shows higher *Sox2* transcript levels (Figure 3.10A, adjusted p value of 0.008).

In contrast to *Pou5f1* and *Sox2*, the other core pluripotency transcription factor NANOG is highly expressed in naïve pluripotency and downregulated upon lineage-priming (Kinoshita & Smith, 2018; Nichols & Smith, 2009; Smith, 2017). This may explain the 3 to 10 times lower *Nanog* transcript levels found in D3ΔN1, B8ΔN2 and B38ΔN2 cells (Figure 3.10A, adjusted p values <0.001). In N1-1ΔN1 and B19ΔN2 cells, however, *Nanog* expression is not altered (Figure 3.10A), suggesting that neither HMGN1 nor HMGN2 directly regulate the transcription of this gene. Importantly, in B1 cells the *Nanog* levels are half that observed in the parental cells (Figure 3.10A), which is in agreement with the observed reduction in the fluorescence intensity (Figure 3.5). As previously mentioned, NANOG is heterogeneously expressed in P19 cells, raising the possibility that when

generating the B1 monoclonal line, a low *Nanog*-expressing cell was selected and expanded.

The mRNAs levels of lineage-specific transcription factor were also assessed. NEUROG1 and ASCL1 are well-known pro-neural transcription factors (Bertrand et al, 2002), while GATA4 and TBXT (T-box) both play important roles in the primitive streak, and then in endoderm and mesoderm specification, respectively (Arceci et al, 1993; Fehling et al, 2003; Fujikura et al, 2002; Laverriere et al, 1994). Consistent with the increased spontaneous neuronal differentiation of *Hmgn*-knockout cells (Figure 3.8), *Neurog1* and *Ascl1* are upregulated in most of the *Hmgn*-knockout cells when compared with parental cells (Figure 3.10B). *Neurog1* transcript levels are increased by 5 to 25 fold in *Hmgn1*-knockout and *Hmgn2*-knockout lines (Figure 3.10B, adjusted p values <0.05), while *Ascl1* by 8 to 18 fold in *Hmgn2*-knockout cells and 67 fold in N1-1ΔN1 cells when compared with parental cells (Figure 3.10B, adjusted p values <0.0001). Consistent with previous reports that higher expression of lineage-specific genes is found in low *Nanog*-expressing cells (Abranches et al, 2014; Chambers et al, 2007; Herberg et al, 2016; Kalmar et al, 2009), *Ascl1* and *Neurog1* levels are four and five fold higher in B1 cells, respectively, than in parental P19 cells (Figure 3.10B, adjusted p values of 0.004 and 0.006, respectively).

In the case of *Gata4*, its expression is 2.5 to 5 times higher in N1-1ΔN1, D3ΔN1, and B8ΔN2 cells than in parental cells (Figure 3.10C, adjusted p values <0.01), suggesting that HMGN1 and HMGN2 are important for the preservation of stem cells rather than for the direct inhibition of differentiation into neural lineages. However, the cells do not prefer mesoderm lineages, since *Tbxt* expression is either unchanged or downregulated in *Hmgn*-knockout cells (Figure 3.10C).

As most of the *Hmgn*-knockout cells mainly prefer neural fates, the transcript levels of neural-specific markers were evaluated. NES, as previously mentioned, is a NSC marker that is also transiently expressed by precursors for other lineages and endothelial cells (Lendahl et al, 1990; Sejersen & Lendahl, 1993; Suzuki et al, 2010; Wroblewski et al, 1997). TUBB3 and MAP2 are neuronal-specific cytoskeleton or cytoskeleton-associated proteins (Kevenaar &

Hoogenraad, 2015). The neural markers are found at higher levels in most of the *Hmgn*-knockout cultures; however, the changes are not as dramatic as for the pro-neural transcription factors (Figure 3.10D). A possible explanation is that NES, TUBB3 and MAP2 are markers of already specified cells, whereas NEUROG1 and ASCL1 are required for initiating the neuronal differentiation program. Interestingly, D3ΔN1 cells only overexpress the promiscuous marker *Nes*, reinforcing the idea that these cells may prefer other fates than the neuronal.

In summary, the gene expression profiling demonstrates that the loss of either of the two HMGN major variants results in the upregulation of lineage-specific genes and, in some cases, the downregulation of the naïve pluripotency transcription factor NANOG.

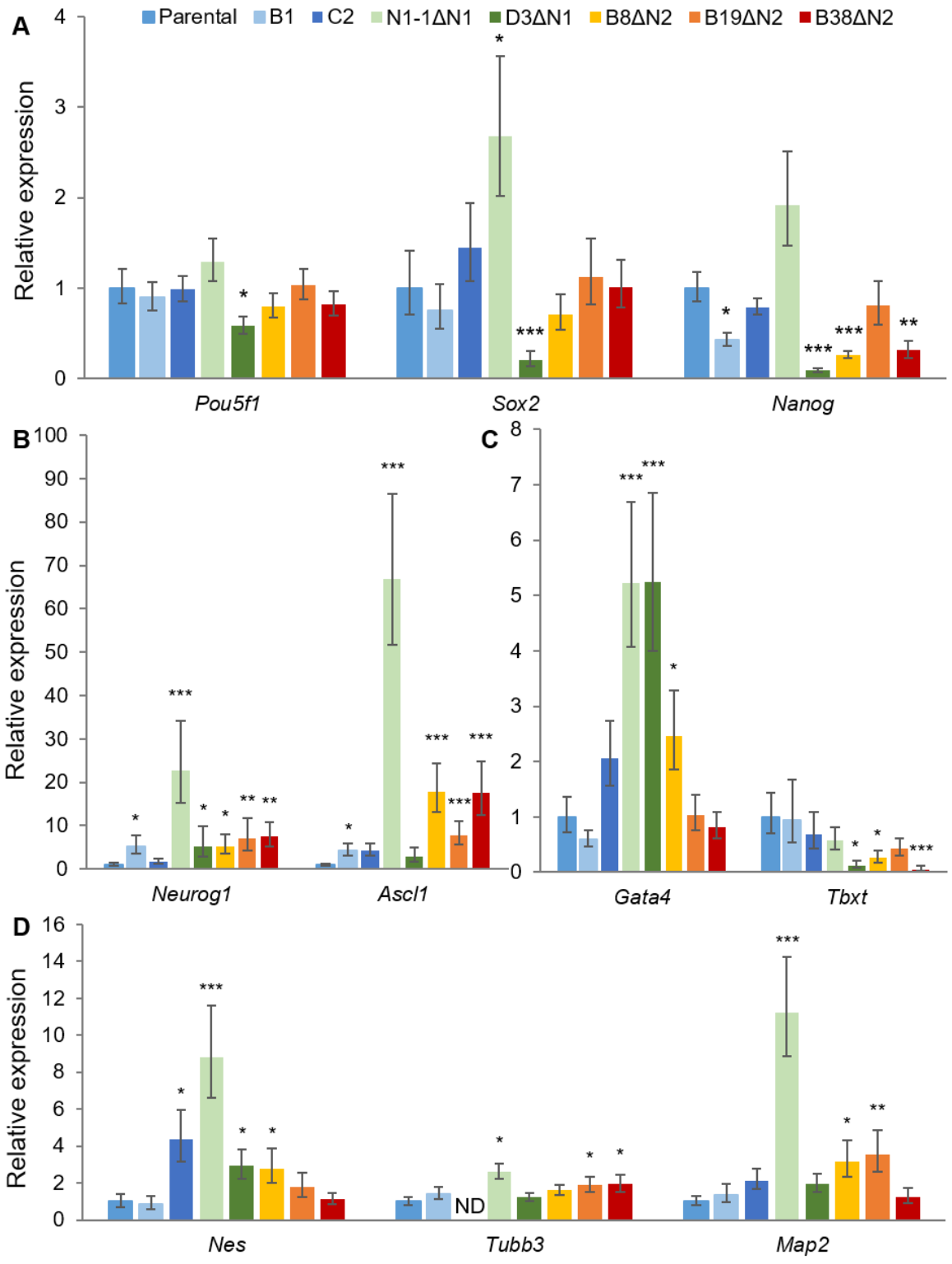


Figure 3.10 *Hmgn*-knockout cells express higher levels of lineage-specific genes
Relative expression of the (A) pluripotency transcription factors *Pou5f1*, *Sox2*, and *Nanog*; (B) lineage-specific transcription factors *Neurog1*, *Ascl1*, *Gata4*, and *Tbxt*; and (C) neural-specific markers *Nes*, *Tubb3*, and *Map2*, as determined by qRT-PCR. The graphs represent the fold change in comparison with parental cells, error bars symbolise the SEM from 3 to10 independent cultures (with the cooperation of A. Sindi), the statistical significance was calculated by ANOVA and Dunnett's multiple comparison test (adjusted p values *<0.05, **<0.001, ***<0.0001).

3.5 Re-expression of HMGN2 using a retroviral system

In order to test whether the previous observations were specifically due the loss of a major HMGN variant, HMGN2 was ectopically re-expressed in B8ΔN2 and B19ΔN2 cells using the second generation pBABE hygromycin retroviral vector system by the Master's students Mariarca Bailo and Faika Laz Banti. Briefly, the retroviral packing cell line Phoenix-Eco was transfected with the retroviral vectors pBABE-hygro or pBABE-hygro-*HMGN2* (Morgenstern & Land, 1990). The latter plasmid expresses the human *HMGN2* cDNA. In this system, retroviral particles are secreted into the media, which is then used to transduce P19 cells. Hygromycin was used to select for transduced P19 cells.

Re-expression of HMGN2 in B8ΔN2 and B19ΔN2 cells was confirmed by WB, and show that HMGN2 levels are comparable to those in parental cells (Figure 3.11A). Furthermore, HMGN2 nuclear localisation was confirmed by IF (Figure 3.11B). Intriguingly, the nuclei of the cells transfected with the pBABE vector, either the empty vector or carrying the *HMGN2* cDNA, seem to be heterogeneous in shape and size, and in particular, they are larger than the nuclei of B8ΔN2 and B19ΔN2 cells as evidenced by DAPI staining (Figure 3.2 and Figure 3.11B). Additionally, the cells grow completely spread out and do not form either compact colonies, or three dimensional clusters, as previously observed (Figure 3.4 and Figure 3.11B).

Gene expression analyses demonstrated that the methodology employed to re-express HMGN2 promotes major changes in the transduced cells. In general, cells carrying either the empty vector or the pBABE-*HMGN2* and selected with hygromycin show substantial differences in pluripotency and lineage-specific marker steady state mRNA levels compared to parental, B8ΔN2, and B19ΔN2 cells (Figure 3.11C). Little difference is observed between cells transduced with the empty pBABE plasmid and the pBABE-*HMGN2* plasmid. The pluripotency genes *Pou5f1*, *Sox2*, and *Nanog* are downregulated in all the transduced cells, as exemplified by *Nanog* expression (Figure 3.11C and data not shown). In contrast, the pro-neural genes *Neurog1* and *Ascl1*, and the neural-specific genes *Nes* and *Map2* are either unchanged or upregulated in parental cells and mostly reduced in B8ΔN2 and B19ΔN2 cells, as observed from *Ascl1* expression (Figure 3.11C and

data not shown). Additionally, *Gata4* is either unchanged or upregulated in transduced cells whereas *Tbxt* is downregulated (data not shown). This expression pattern resembles that of differentiating P19 cells.

The data indicate that cells transduced with the control vector have substantial changes in morphology and gene expression, which dwarf any effects arising from re-expression of HMGN2. It is possible that these effects are due to continued culture in the presence of hygromycin, so the cultures could be re-analysed after a several passages in the absence of hygromycin. Alternatively, the third generation pGIPZ lentiviral vector system could be used, as this lacks the truncated Gag sequence that is present in the pBABE vector that might be causing a problem (Morgenstern & Land, 1990). In conclusion, the retroviral methodology employed for the re-expression of HMGN2 failed to produce useful data that could support or refute the hypothesis that the major HMGN variants safeguard pluripotency in ECCs.

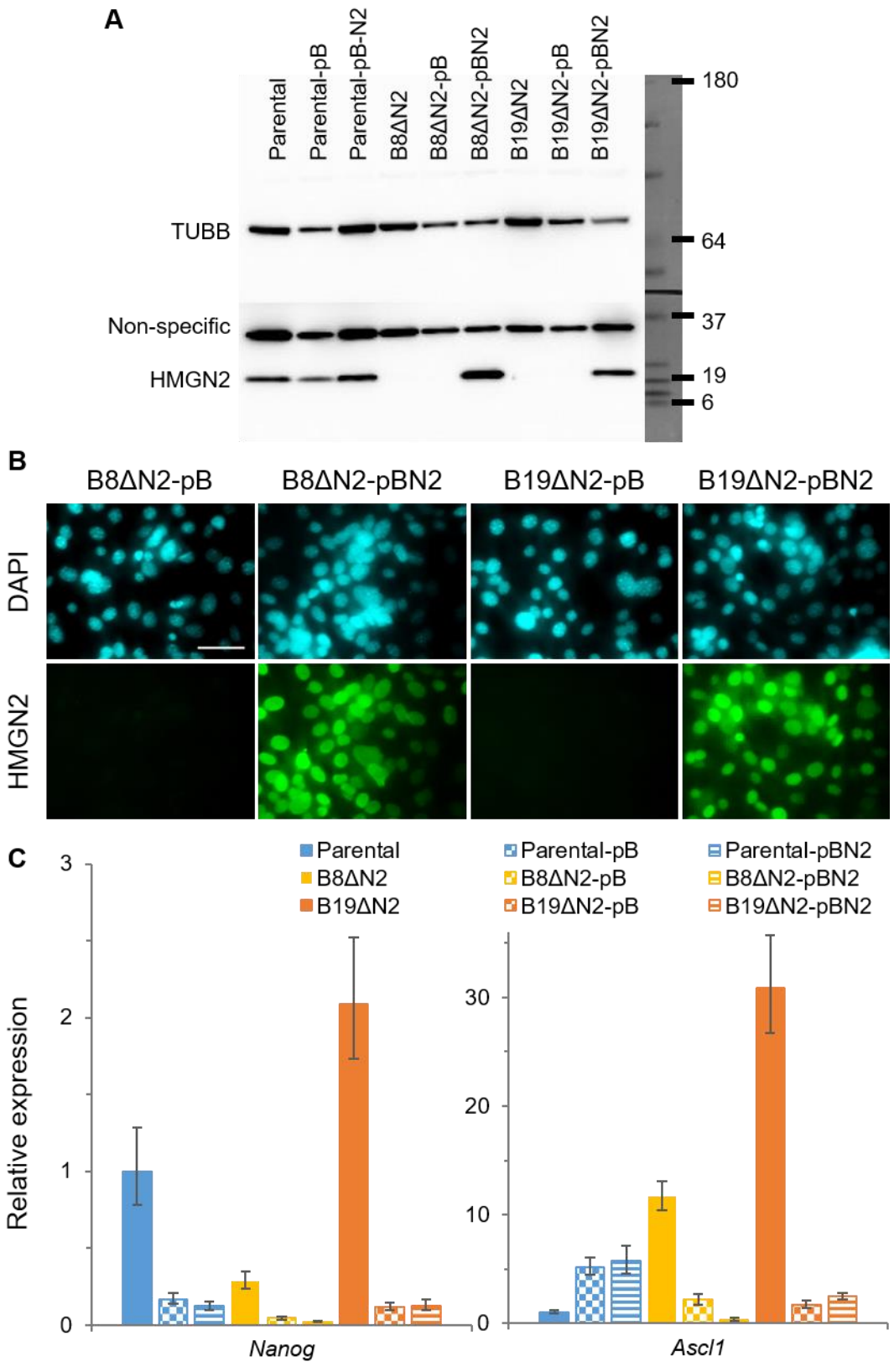


Figure 3.11 Re-expression of HMGN2 achieved with the pBABE-hygro retroviral vector
Detection of HMGN2 protein by (A) WB and (B) IF. DAPI was used to stain the nuclei. Scale bar indicates 50 µm. C) Relative expression of the pluripotency and pro-neural transcription factors *Nanog* and *Ascl1*, respectively, as determined by qRT-PCR. The graph represents the fold change in comparison with parental non-transduced cells and error bars symbolise the SEM from three independent cultures.

3.6 Discussion

This chapter studied the phenotypical alterations of P19 cells after the loss of HMGN1 and HMGN2. P19 cells are transformed pluripotent cells from embryonal origin that exhibit a distinct and mostly uniform morphology, express pluripotency markers, indefinitely self-renew, and do not differentiate unless stimulated (Jones-Villeneuve et al, 1982; McBurney, 1993; McBurney & Rogers, 1982). *Hmgn1*- and *Hmgn2*-knockout lines, in conjunction with control lines, were subjected to imaging and immunostaining analyses capturing the morphology and organisation of the cells, and detecting pluripotency and differentiation markers by IF. Additionally, steady state mRNA levels of pluripotency transcription factor and lineage-specific genes were assayed by qRT-PCR.

Hmgn-knockout cells are morphologically heterogeneous and grow in a more dispersed manner than the parental cells (Figure 3.3), resembling the differentiated cells in ESC cultures (Trott & Martinez Arias, 2013; Wray et al, 2010). Indeed, differentiating cells are observed in *Hmgn*-knockout cultures as evidenced by NES and TUBB3 immunostaining (Figure 3.7 and Figure 3.8). In addition, a high proportion of *Hmgn*-knockout cells is devoid of the pluripotency markers SSEA1 and NANOG (Figure 3.4 and Figure 3.5), while retaining POU5F1 (Figure 3.6). Unlike NANOG, POU5F1 and SOX2 are retained during the early stages of differentiation where they interact with different lineage-specific transcription factors to influence fate commitment (Kalkan et al, 2017; Thomson et al, 2011), raising the possibility that some of the observed POU5F1-positive cells have abandoned pluripotency and initiated differentiation programs.

Gene expression analyses are consistent with the previous observations. *Hmgn*-knockout cultures show higher levels of lineage-specific transcripts, in some cases accompanied by a reduction in *Nanog* expression (Figure 3.10). *Pou5f1* and *Sox2*, however, are comparable among control and *Hmgn*-knockout lines (Figure 3.10A). Interestingly, a set of pluripotency genes including *Nanog* is downregulated in the transition from the naïve pluripotency to the primed state for differentiation (Kinoshita & Smith, 2018; Nichols & Smith, 2009; Smith, 2017). Furthermore, NANOG is heterogeneously expressed in ESC and P19 cultures and the low-NANOG cells show higher lineage-specific gene expression

(Abranches et al, 2014; Chambers et al, 2007; Herberg et al, 2016; Kalmar et al, 2009). Although the low-NANOG cells are more prone to differentiate, they have not abandoned pluripotency as they indefinitely self-renew and even originate daughter cells with higher NANOG levels (Abranches et al, 2014; Chambers et al, 2007; Herberg et al, 2016; Kalmar et al, 2009).

In the cited works, the proportion of low-NANOG cells among the population varies around 20%. In fact, when deriving monoclonal lines from WT P19 cells, one out of six shows lower *Nanog* expression accompanied by higher levels of *Neurog1* and *Ascl1* transcripts (Supplementary Figure, Appendix). This may be also the case of B1 monoclonal line, as *Nanog* is downregulated whereas *Neurog1* and *Ascl1* are upregulated in these cells (Figure 3.10); nevertheless, B1 cells do not spontaneously differentiate like the *Hmgn*-knockout lines, which can be confirmed by IF for NES and TUBB3 (Figure 3.7 and Figure 3.8). Interestingly, this one-in-six proportion seems to be altered in the *Hmgn*-knockout lines as three out of five monoclones express *Nanog* in lower levels (Figure 3.10A). Whether the reduction in *Nanog* expression is the driver for the increased propensity to differentiate is not clear. However, these chromatin architectural proteins do not seem to directly regulate *Nanog* transcription as N1-1ΔN1 and B19ΔN2 cells display normal transcript levels, while expressing higher levels of lineage-specific markers (Figure 3.10) and presenting numerous differentiating cells (Figure 3.7 and Figure 3.8).

In fact, reports suggest that HMGNs do not act as specific regulators of gene expression, but fine-tune the existing gene expression profile and, in their absence, cumulative changes lead to altered phenotypes (Deng et al, 2017; Deng et al, 2013; Deng et al, 2015; Kugler et al, 2013). In the present work, a high proportion of ECCs exits pluripotency and initiates differentiation programs, mainly towards neural fates, in the absence of a major HMGN variant.

These observations, however, contrast with previous works performed in ESCs derived from *Hmgn1* single- and *Hmgn1* and *Hmgn2* double-knockout mice, where alkaline phosphatase reactivity and the cellular and colony morphology show no significant changes compared to ESCs derived from WT mice (Deng et al, 2017; Deng et al, 2013). Furthermore, the authors did not observe changes in

gene expression that suggest a higher frequency of spontaneous differentiation of *Hmgn1* single- and *Hmgn1* and *Hmgn2* double-knockout ESCs to differentiate, not even towards neural fates.

Several reasons can account for the exposed differences. Unlike P19 cells that indefinitely self-renew in serum-supplemented media and do not spontaneously differentiate, ESCs require further support of different factors for their propagation. As previously mentioned, feeder cells or LIF, serum or BMPs, and 2i sustain ESC self-renewal mainly by preventing differentiation (Smith et al, 1988; Wray et al, 2010; Ying et al, 2003a; Ying et al, 2008). Deng and collaborators (2013, 2017) cultured the cells with feeders or under feeder-free conditions including knockout serum replacement, LIF, and, in their most recent work, also 2i. It is possible to speculate that the loss of a major HMGN variant facilities the response of the P19 cells to inductive differentiation cues, and that this response is ultimately blocked in ESCs by LIF and 2i. Alternatively, the loss of HMGN1 or HMGN2 might also increase the spontaneous differentiation of ESC, however, this cannot be observed as in 2i media only naïve ESCs form colonies and their differentiating progeny is eliminated (Kalkan et al, 2017).

It is also worth considering the differences in the developmental stages from which ESCs and P19 cells were derived. ESCs capture the naïve state of pluripotency from pre-implantation embryos; once the embryo is implanted, the cells continue towards differentiation and the derived cell lines, such as EpiSCs, represent a pluripotency state primed for differentiation (Kinoshita & Smith, 2018; Martello & Smith, 2014; Nichols & Smith, 2009; Smith, 2017). Previous works support the idea that P19 cells resemble EpiSCs (Kelly & Gatie, 2017; Mallanna et al, 2008), and these cells were derived from post-implantation embryos, raising the possibility that they can immediately commit to differentiation, whereas ESCs first exit the naïve pluripotency to later respond to inductive cues (Kinoshita & Smith, 2018; Mulas et al, 2017; Smith, 2017). Therefore, HMGNs may not be relevant for the preservation of the naïve pluripotency, that *in vitro* is enhanced by culturing the cells in 2i; rather, these chromatin architectural proteins may be important for keeping the primed cells unresponsive to misplaced inductive differentiation cues.

Interestingly, ESCs start losing their identity upon withdrawal of the factors that maintain them in a pluripotency state and differentiate mostly into neural lineages (Ying et al, 2003b). This is in complete agreement with the default model of neural differentiation proposing that the cells commit to neural fates unless instructed otherwise (Hemmati-Brivanlou, 1997; Munoz-Sanjuan & Brivanlou, 2002; Wilson & Edlund, 2001). This may explain the preference of *Hmgn*-knockout cells towards neural lineages over others.

The following chapter investigates whether the proportion of differentiating cells in the *Hmgn*-knockout cultures is triggered by inductive differentiation cues.

Chapter 4 Inhibition of signalling pathways relevant for neurogenesis does not reverse the *Hmgn*-knockout phenotype

4.1 Introduction

In the previous chapter, it was suggested that the loss of a major HMGN variant increases the differentiation propensity of ECCs, leading to a higher expression of lineage-specific genes and increased spontaneous differentiation, mainly towards neural lineages. Additionally, it was speculated that *Hmgn*-knockout cells might respond more readily to inductive differentiation cues.

A major event in the differentiation of ESCs is the autocrine stimulation of FGF signalling through the activation of the MAPK cascade. As previously mentioned, POU5F1 and SOX2 seem to play a dual role controlling pluripotency and developmental programs. These two transcription factors promote *Fgf4* expression in ESCs (Yuan et al, 1995). Autocrine stimulation of the FGF receptor by FGF4 is required for the proper commitment of ESCs to neuronal differentiation after withdrawal of the factors that sustain pluripotency (Ying et al, 2003b). Interestingly, ESCs treated with pharmacological inhibitors of the FGF receptor and *Fgf4*-null ESCs resist neural and non-neural induction, suggesting that the FGF signalling initiates commitment into multiple lineages in ESCs before specification (Kunath et al, 2007; Ying et al, 2003b). Furthermore, pharmacological inhibition of the FGF receptor impairs MAPK activation, and an inhibitor of the MAP2K that activates MAPK also impedes neural and non-neural specification of ESCs, identifying MAPK as the downstream effector of the FGF signalling in ESC differentiation (Stavridis et al, 2007). Indeed, this MAP2K inhibitor is included in the 2i formulation to prevent ESC differentiation (Ying et al, 2008). Importantly, MAPK signalling is also upregulated upon differentiation in P19 cells and proper neural induction of these cells relies on the presence of FGF4 and FGF8 (Wang et al, 2006). These observations raise the possibility that HMGN1 or HMGN2 may be important for modulating events related to the activation of the FGF signalling.

WNT signalling is also relevant in ESC and P19 differentiation, although the scenario is complex. Canonical WNT signalling activation leads to catenin beta 1

(CTNNB1) stabilisation and translocation to the nuclei of the cells where it promotes transcription of target genes (MacDonald et al, 2009). When WNT signalling is inactive, CTNNB1 is sequestered and targeted for degradation in the cytoplasm by a complex containing, among others, GSK3 and AXIN2 (MacDonald et al, 2009). Self-renewal of ESCs is enhanced by pharmacological inhibition of GSK3 (Sato et al, 2004; Ying et al, 2008). Therefore, a GSK3 inhibitor is included in the 2i formulation (Ying et al, 2008), where it resists exit from naïve pluripotency through a mechanism that includes CTNNB1-mediated de-repression of relevant pluripotency genes (Wray, 2011). However, in the mouse embryo, WNT signalling is required during differentiation but not in the pre-implantation stages from which ESCs are derived (Huelsken et al, 2000), raising the possibility that the WNT requirement for ESC self-renewal is an artificial culture condition.

In fact, a recent study shows that WNT signalling is not very active in ESCs; however, a subpopulation express a WNT reporter and a target gene, *Axin2*, in combination with lower levels of pluripotency factors and higher differentiation gene expression (Trott & Martinez Arias, 2013). The results suggest that WNT signalling activation occurs in early stages of differentiation, when ESCs just exit pluripotency (Trott & Martinez Arias, 2013). In agreement, previous studies in P19 cells have shown that different *Wnt* genes are upregulated upon RA induction of neuronal differentiation, and that overexpression of WNT1 in the absence of RA promotes SSEA1 loss and increments in neuronal-specific markers and transcription rates pro-neural genes (Smolich & Papkoff, 1994; Tang et al, 2002). Canonical WNT signalling is likely to be responsible for these observations as inhibition of GSK3 and overexpression of CTNNB1 also initiate neuronal differentiation of P19 cells (Ding et al, 2003; Israsena et al, 2004), and of ESCs (Otero et al, 2004), through direct transcriptional regulation of *Neurog1* (Hirabayashi et al, 2004; Israsena et al, 2004).

Interestingly, CTNNB1 and HMGN2 have been shown to collaborate in the modulation of homeodomain target gene expression (Amen et al, 2008). In this study, the presence of CTNNB1 switches the homeodomain/HMGN2 complex from a transcriptional repressor to a ternary activator complex, providing a link between HMGNs and WNT signalling that, in conjunction with the roles of WNT

during pluripotent stem cell differentiation, potentially influences cell fate decisions, favouring neural specification.

Finally, NOTCH signalling activates the HES family of transcriptional repressors, regulating lineage-specific stem or progenitor cell maintenance and fate decisions (Kageyama et al, 2007). Particularly, NOTCH signalling promotes self-renewal of NSCs, through a mechanism in which HES1 and HES5 antagonise the expression of pro-neural genes, and therefore, the commitment to neurogenesis (Kageyama et al, 2007; Ohtsuka et al, 1999). However, NOTCH signalling is neither required for maintaining ESC identity, nor changes the phenotype of ESCs when constitutively active (Lowell et al, 2006). Nevertheless, there is substantial evidence that NOTCH signalling operates during early ESC differentiation. NOTCH components are expressed by ESCs, and constitutive expression of NICD promotes an accelerated and exclusive neuronal differentiation of ESCs upon withdrawal of self-renewal stimuli; conversely, its genetic or pharmacological interference inhibits neural fates (Kobayashi & Kageyama, 2010; Lowell et al, 2006). These observations suggest a role for NOTCH signalling in the early cell fate decisions of ESCs, which may also determine the neural fate of P19 cells after the loss of HMGN1 or HMGN2.

The aim of this chapter is to evaluate whether inductive cues trigger differentiation programs in P19 cells after the loss of a major HMGN variant.

4.2 Inhibition of WNT, FGF and NOTCH signalling reduces the transcription rates of reporter genes

The strategy used in the following experiments consists of culturing control and *Hmgn2*-knockout cells in normal conditions and under inhibition of WNT, FGF, and NOTCH signalling. The tankyrase inhibitor XAV-939 stabilises AXIN2 (Huang et al, 2009), enhancing CTNNB1 degradation, and therefore, impeding canonical WNT signalling and target gene expression. SU5402 is widely used as a competitive FGF receptor inhibitor (Mohammadi et al, 1997) that interferes with FGF signalling. DAPT selectively inhibits γ-secretase (Dovey et al, 2009), which is the enzyme responsible for NOTCH receptor cleavage, releasing the NICD to promote transcription of target genes.

In a first approach, parental P19 cells were exposed to the inhibitors, and additionally, to activators of two of the pathways, testing the response of the selected reporter genes. P19 cells were allowed to attach overnight in the usual media, and the day after, the three inhibitors and two activators were used independently. After 24 h of exposure to the different treatments, RNA was collected, and relative expression of selected target genes was determined by qRT-PCR in comparison with untreated cells.

Figure 4.1 demonstrates that the selected target genes effectively report the inhibition of the respective pathways. Activation of WNT signalling using a GSK3 inhibitor results in a 10 fold upregulation of *Axin2*, whereas inhibition of WNT signalling using the tankyrase inhibitor (iWNT) causes a 70% reduction in *Axin2* expression (Figure 4.1). *Fgf4* levels are increased by 60% after FGF8 exposure and reduced by 30% after FGF receptor blockade (iFGF; Figure 4.1). *Hes5* expression is decreased by 40% under DAPT treatment (iNOTCH; Figure 4.1).

It is worth noting that manipulation of WNT and FGF signalling modifies the expression of target genes from other pathways, whereas inhibition of NOTCH only affects *Hes5* expression (Figure 4.1). This is in complete agreement with previous works reporting that ESCs cultured in the undifferentiated state are sensitive to inhibitors of WNT and FGF signalling (Kunath et al, 2007; Stavridis et al, 2007; Wray et al, 2010; Ying et al, 2008), but not to NOTCH inhibitors (Lowell et al, 2006). Interestingly, WNT and FGF signalling seem to act in opposite directions in P19 cells (Figure 4.1), as previously reported in developmental studies (Dyer et al, 2014; Wurst & Bally-Cuif, 2001). Furthermore, stimulation of FGF signalling leads to an increase in *Hes5* (Figure 4.1), consistent with the idea that FGF signalling is required for proper ESC commitment into neural progenitors (Ying et al, 2003b), where HES5 regulates NSC fate decisions (Kageyama et al, 2007). In addition, previous findings show that NOTCH effectors can be modulated by different pathways in particular situations (Doetzlhofer et al, 2009; Dyer et al, 2014; Nakayama et al, 2008).

It can be concluded that the selected genes are capable of reporting changes in the activity of the WNT, FGF, and NOTCH signalling pathways in P19 cells.

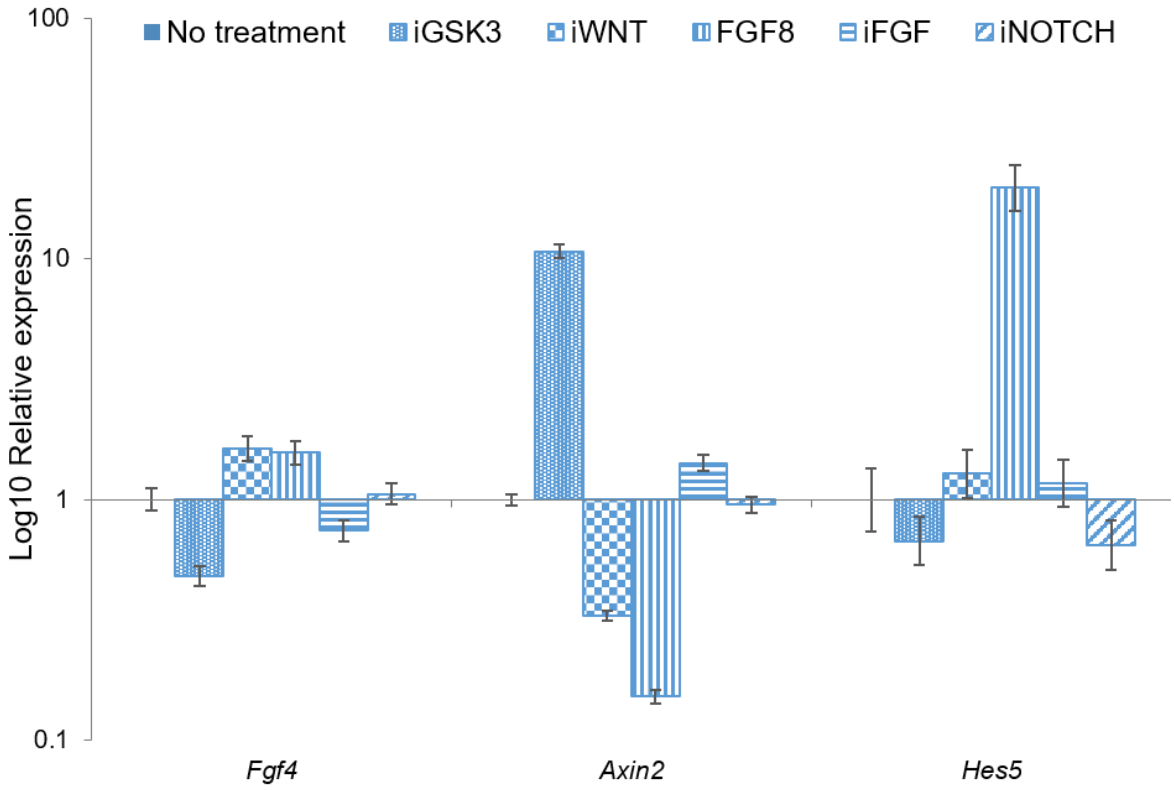


Figure 4.1 Target genes of WNT, FGF, and NOTCH signalling can report the inhibition or activation of the respective pathways
Relative expression of the reporter genes *Axin2*, *Fgf4*, and *Hes5*, which are targets of WNT, FGF, and NOTCH signalling pathways, respectively, as determined by qRT-PCR. P19 cells were allowed to attach overnight and exposed for 24 h to activators and inhibitors of the mentioned pathways. The graph represents the logarithm of the fold change in comparison with untreated cells and error bars symbolise the SD from technical triplicates.

4.3 Inhibition of WNT, FGF or NOTCH signalling does not reduce lineage-specific gene expression in *Hmgn2*-knockout cells

To investigate whether, after the loss of a major HMGN variant, P19 cells initiate differentiation programs triggered by inductive cues, FGF, WNT, and NOTCH signalling pathways were inhibited in control and *Hmgn2*-knockout cells using the conditions described above. The expression of the reporter genes was assayed in order to demonstrate effective inhibition of the signalling pathways.

The basal levels of *Axin2* are similar in all the cell lines, although slightly higher in B8ΔN2 and lower in B38ΔN2, suggesting that WNT signalling is not altered in a particular direction after the loss of HMGN2 (Figure 4.2A). When the cells are treated with the tankyrase inhibitor, *Axin2* is reduced by 60% in all the lines except for B38ΔN2, in which the basal levels are lower, and the reduction is around 40% (Figure 4.2A). Thus, iWNT indeed impedes WNT signalling.

Fgf4 basal expression is variable in the different lines tested, with a 60% reduction in control B1 and B19ΔN2 cells compared to the parental cells (Figure 4.2B). This suggests that autocrine stimulation of FGF signalling varies between clonal lines, regardless the presence of HMGN2. After inhibition of the FGF receptor, *Fgf4* levels are reduced by 40-50% in all cases (Figure 4.2B), indicating that iFGF successfully restricts FGF signalling.

The *Hes5* transcript is barely detectable in all the cell lines, which is consistent with previous work demonstrating that NOTCH signalling is required for proper ESC differentiation rather than for maintaining ESC identity (Kobayashi & Kageyama, 2010; Lowell et al, 2006). In B8ΔN2 and B19ΔN2 cells, *Hes5* is further reduced in comparison with parental cells (Figure 4.2C), providing a partial explanation for the upregulation of pro-neural genes observed in these cells (Figure 3.10B and Figure 4.3B and C), as *Hes1* and *Hes5* reduce the transcription rates of pro-neural transcription factors (Ohtsuka et al, 1999). iNOTCH treatment effectively interferes with NOTCH activity as demonstrated by the downregulation of *Hes5* by 40-70% in all lines (Figure 4.2C).

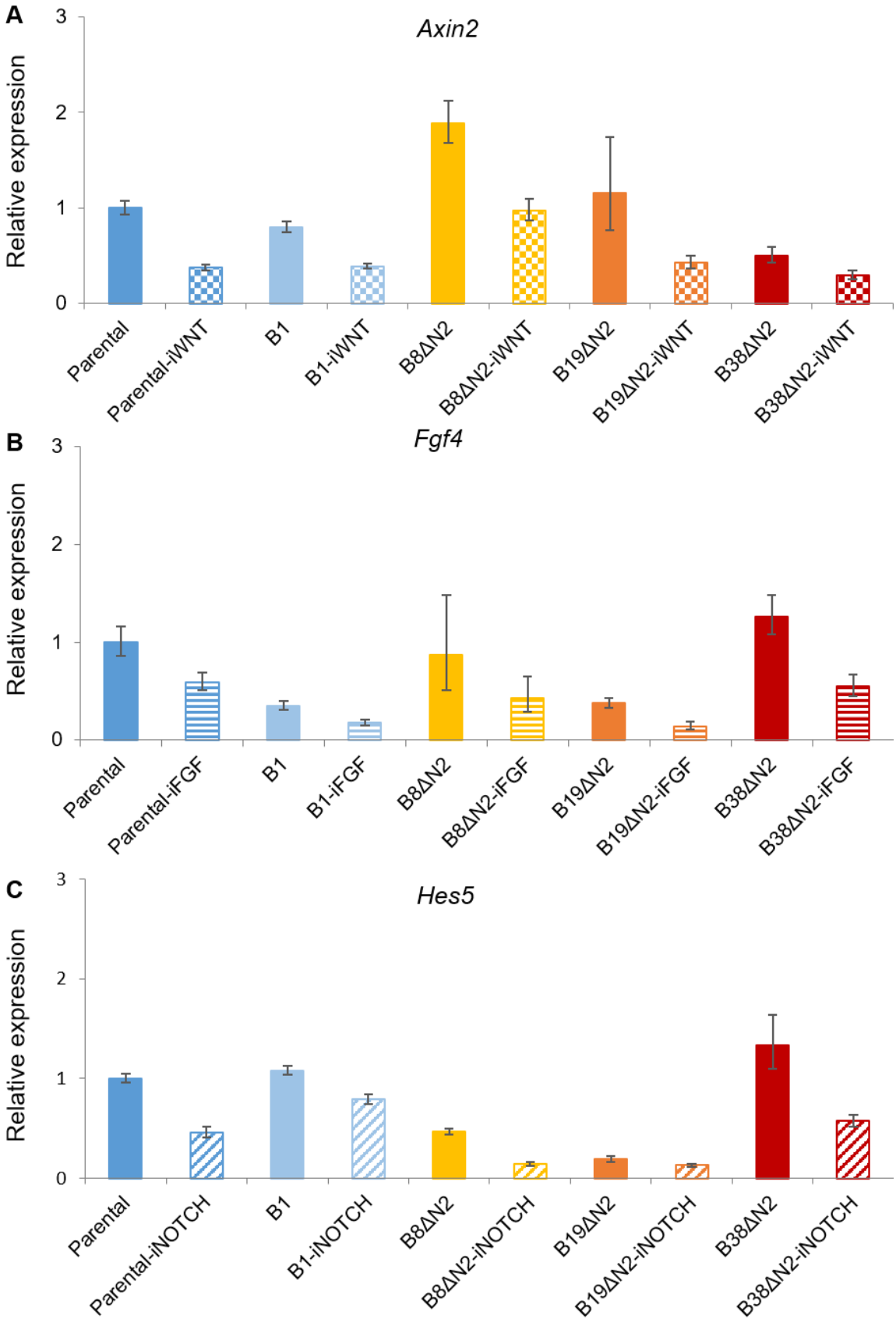


Figure 4.2 Interference of WNT, FGF, and NOTCH signalling in control and *Hmgn2*-knockout cells
Relative expression of the reporter genes *Fgf4*, *Axin2*, and *Hes5* under inhibition of FGF, WNT, and NOTCH signalling pathways, respectively, as determined by qRT-PCR. The graphs represent the fold change in comparison with untreated parental cells and error bars symbolise the SEM from 2 to 3 independent cultures.

Three of the genes previously shown to be altered in *Hmgn2*-knockout cells, *Nanog*, *Neurog1* and *Ascl1*, were assayed following signalling inhibition. The logic behind was that if the increased *Neurog1* and *Ascl1* expression was due to increased WNT, FGF, or NOTCH signalling in the *Hmgn2*-knockout cells, then inhibition of these pathways should reduce these genes back to the levels observed in the parental P19 cells.

As previously shown (Figure 3.10A), *Nanog* is significantly downregulated in B1, B8ΔN2, and B38ΔN2 cells. iWNT treatment reduces *Nanog* levels by around a half in all cell lines (Figure 4.3A), irrespective of the presence of HMGN2. This is in agreement with previous findings where activation of WNT signalling enhances the transcription rates of naïve pluripotency transcription factors (Wray et al, 2010; Ying et al, 2008). Conversely, iFGF roughly doubles *Nanog* expression in all cases (Figure 4.3A). This opposite effect to iWNT is in agreement with Figure 4.1 and previous developmental studies (Dyer et al, 2014; Wurst & Bally-Cuif, 2001). Furthermore, interfering with FGF signalling seems to re-establish *Nanog* levels in B1, B8ΔN2, and B38ΔN2 cells to that observed in untreated parental cells. However, consistent with the idea that inhibiting the MAPK cascade leads to upregulation of naïve pluripotency transcription factors (Kunath et al, 2007; Stavridis et al, 2007; Ying et al, 2008) and impedes RA-induced neuronal differentiation of P19 cells (Wang et al, 2006), iFGF also increases *Nanog* levels in parental cells, thus the differences between the cell lines are maintained. NOTCH signalling does not seem to be relevant for the transcriptional regulation of *Nanog* (Figure 4.3A), consistent with previous data that NOTCH is not required for the maintenance of the ESC identity (Lowell et al, 2006).

Conversely, *Neurog1* and *Ascl1* are upregulated in B1, B8ΔN2, B19ΔN2, and B38ΔN2 cells in comparison with parental cells (Figure 3.10B and Figure 4.3B and C). Interfering with WNT signalling has little effect on the transcription rates of pro-neural transcription factors, which contrasts previous findings where overexpression of WNT1 or CTNNB1 leads to neuronal differentiation of P19 cells by directly targeting the *Neurog1* promoter (Hirabayashi et al, 2004; Israsena et al, 2004; Smolich & Papkoff, 1994; Tang et al, 2002). It is possible to speculate that WNT signalling operates differently in undifferentiated P19 cells than during differentiation, and that WNT signalling is not important for transcriptional

regulation of *Neurog1* in the undifferentiated state. Indeed, previous studies show that WNT signalling becomes active after implantation in mouse embryos (Huelsken et al, 2000) and when ESCs exit pluripotency (Trott & Martinez Arias, 2013). Additionally, the canonical WNT signalling antagonist, AXIN2, reduces the formation of neurite-like structures in RA-induced neuronal differentiation of P19 cells, rather than impeding the initiation of differentiation (Lyu et al, 2003).

Blockade of FGF receptor activity increases the transcription rates of pro-neural transcription factor, as *Neurog1* and *Ascl1* levels are 7 to 13 and 3 to 5 fold higher, respectively, than in untreated cells (Figure 4.3B and C). This is an unexpected result, as iFGF upregulates *Nanog*, raising the possibility that FGF signalling operates differently in individual cells from control and *Hmgn2*-knockout cultures. Similarly to ESCs maintained in serum, P19 cells constitute a heterogeneous population in which different cells have different propensities to differentiate. This is evidenced by the heterogeneous *Nanog* levels and the reciprocal lineage-specific gene expression (Kalmar et al, 2009). Furthermore, these wide-ranging pluripotency states respond differently to FGF signalling, since in the two ends, naïve pluripotent cells proceed to EpiSCs whereas EpiSCs self-renew (Nichols & Smith, 2009). In such a scenario, inhibition of FGF signalling possibly enhances naïve pluripotency, and therefore, *Nanog* expression, while in EpiSCs compromises self-renewal, leading to differentiation and lineage-specific gene expression. An alternative explanation arises from the observation that stimulation of FGF signalling increase the levels of *Hes5* (Figure 4.1). Although *Hes5* does not seem to be relevant for embryonic pluripotent cells (see above), it is possible to speculate that FGF signalling influences the expression of other HES family members important for embryonic pluripotent cells, such as *Hes1* (Kobayashi & Kageyama, 2010; Kobayashi et al, 2009), indirectly impacting on the transcription rates of pro-neural transcription factors.

Neurog1 and *Ascl1* levels are also higher after iNOTCH treatment, since they are 2 to 4 and 1.5 to 2 fold the levels of untreated cells, respectively (Figure 4.3B and C). The magnitude of the changes is lower than after FGF inhibition and no differences are observed in parental cells. However, this is consistent with previous work demonstrating that NOTCH effectors antagonise the pro-neural

gene expression (Kageyama et al, 2007; Ohtsuka et al, 1999). It is worth mentioning that this work used NSCs as NOTCH does not seem to be relevant for ESCs (Lowell et al, 2006). A potential explanation is that NOTCH signalling becomes active in the cells with higher differentiation propensity and the differentiating P19 cells, and therefore, these cells gain responsiveness to NOTCH signalling. Nevertheless, the effects of iNOTCH in the pro-neural transcription factor mRNA levels do not support that NOTCH signalling triggers neural induction of *Hmgn*-knockout cells.

The previous results show that interfering with the main pathways related to neuronal differentiation does not revert lineage-specific gene expression in *Hmgn2*-knockout cells, suggesting that spontaneous neuronal differentiation after the loss of a major HMGN variant is not under the control of WNT, FGF, or NOTCH signalling.

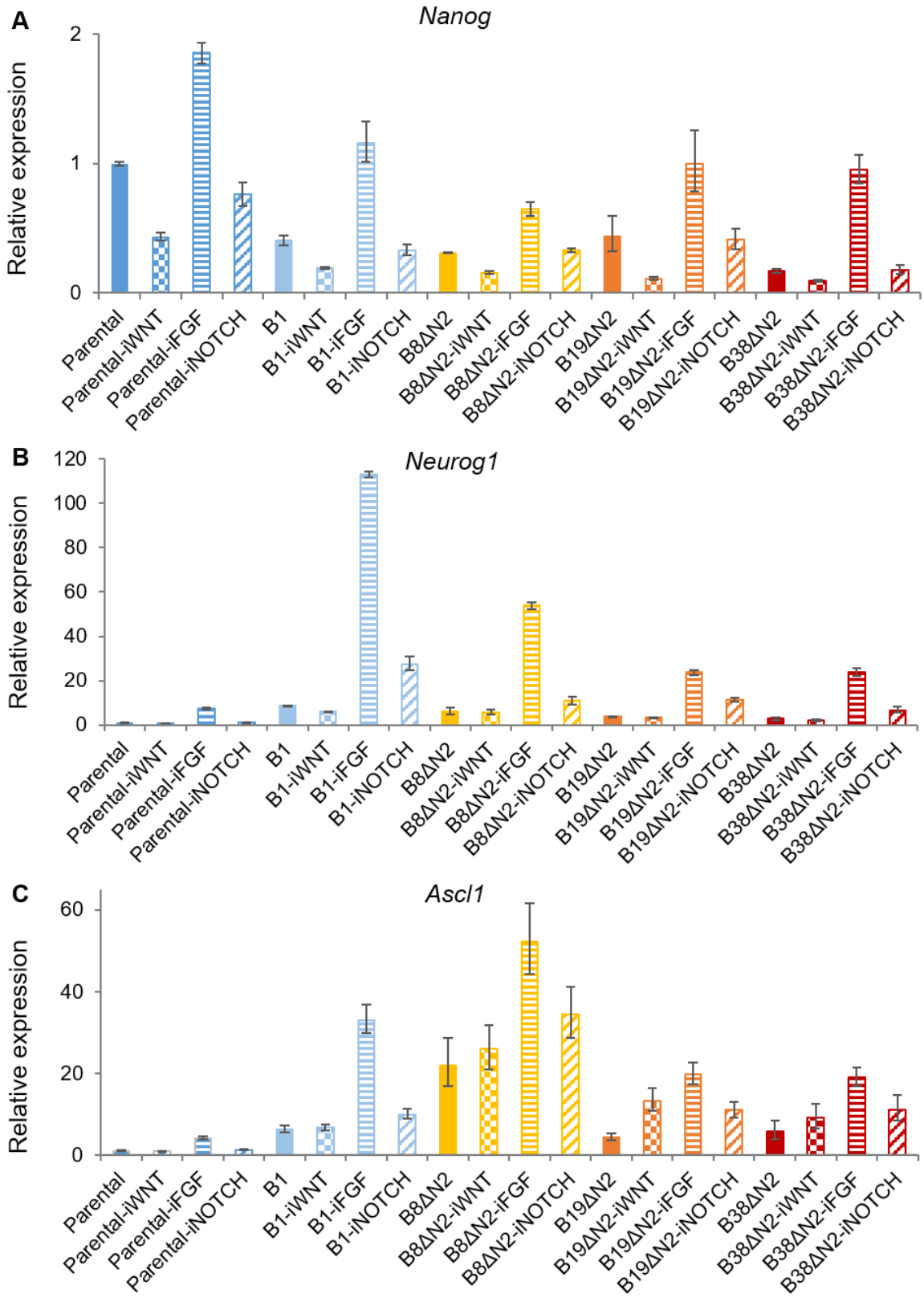


Figure 4.3 Interfering with WNT, FGF and NOTCH signalling does not reduce the lineage-specific gene expression in *Hmgn2*-knockout cells
Relative expression of (A) the pluripotency transcription factors *Nanog* and the pro-neural transcription factors (B) *Neurog1*, and (C) *Ascl1*, under inhibition of WNT, FGF, and NOTCH signalling pathways as determined by qRT-PCR. The graphs represent the fold change in comparison with untreated parental cells and error bars symbolise the SEM from 2 to 3 independent cultures.

4.4 Discussion

This chapter investigated whether the spontaneous differentiation, mainly towards neural lineages, observed in P19 cells after the loss of a major HMGN variant is triggered by inductive cues. WNT, FGF, and NOTCH signalling are involved in embryonic pluripotent cell differentiation events such as multilineage commitment and neural specification. Control and *Hmgn2*-knockout cells were treated with inhibitors of the three pathways, and the inhibition was confirmed using validated reporter genes prior to further analysis (Figure 4.2) by (Figure 4.1).

The loss of HMGN2 does not appear to either enhance or diminish the three signalling pathways, as expression of their target genes does not change in a consistent pattern in *Hmgn2*-knockout lines when compared with parental cells (Figure 4.2). Furthermore, there are differences in the levels of *Fgf4* between parental and B1 cells (Figure 4.2B), indicating clonal variability. FGF4 stimulates FGF signalling in an autocrine manner, allowing embryonic pluripotent cell commitment to differentiation (Kunath et al, 2007; Stavridis et al, 2007). Lineage-priming precedes the initiation of differentiation and is denoted by naïve pluripotency factor downregulation (Kalkan et al, 2017; Kinoshita & Smith, 2018; Smith, 2017). In agreement, *Nanog* is upregulated in all cell lines after blockade of FGF receptor (Figure 4.3A), suggesting that at least some of the cultured cells can revert to a naïve pluripotency state under FGF signalling inhibition (Kalkan et al, 2017; Marks et al, 2012; Ying et al, 2008).

Intriguingly, *Neurog1* and *Ascl1* are also increased by iFGF in all cell lines (Figure 4.3B and C), raising the possibility that at least a subset of P19 cells respond differently to FGF signalling compared to naïve ESCs, perhaps reflecting the different developmental stage from which P19 cells were derived. Indeed, EpiSCs and hESCs proliferate in the presence of FGF2 and differentiate upon FGF withdrawal, in contrast to mouse ESCs (Tesar et al, 2007), providing evidence that the independence of FGF signalling for self-renewal is a unique characteristic of naïve pluripotency and that the developmental stage influences the responsiveness to FGF signalling (Kinoshita & Smith, 2018; Smith, 2017). The fact that interfering with FGF signalling has similar effects in the different cell lines, including higher pro-neural gene expression, indicates that the

spontaneous neuronal differentiation of *Hmgn2*-knockout cells is not explained by changes in FGF signalling.

WNT signalling neither seem to be responsible for the neuronal differentiation of P19 cells after the loss of HMGN2, as iWNT has little effect on pro-neural gene expression (Figure 4.3B and C). This observation contrasts previous works showing that canonical WNT signalling promotes *Neurog1* expression (Hirabayashi et al, 2004; Israsena et al, 2004; Tang et al, 2002). However, similar to FGF signalling, canonical WNT signalling has shown to work differently in undifferentiated state than along differentiation.

In ESCs, the activation of canonical WNT signalling enhances self-renewal, through CTNNB1-mediated de-repression of pluripotency genes (Sato et al, 2004; Wray et al, 2010; Ying et al, 2008). In agreement, iWNT reduces *Nanog* levels (Figure 4.3A). Nevertheless, WNT signalling is not operating at the blastocyst stage from which ESCs are derived, rather it is activated at early differentiation stages (Huelsken et al, 2000). These observations suggest that the requirement of WNT signalling for ESC self-renewal is an artificial culture condition. When ESCs are allowed to differentiate by withdrawing the factors sustaining pluripotency, the cells gain WNT responsiveness and activates WNT signalling early after exiting pluripotency, and WNT supports differentiation (Kalkan et al, 2017; Trott & Martinez Arias, 2013). Then, it is possible that the neurogenic effect of WNT signalling starts in later developmental stages, and thus, WNT does not trigger differentiation of P19 cells.

Similarly, NOTCH is not relevant for pluripotency (Lowell et al, 2006), which explains the fact that *Nanog* expression is not influenced by iNOTCH in any of the cell lines, and that iNOTCH does not affect *Neurog1* and *Ascl1* transcript levels in parental cells (Figure 4.3). Interestingly, *Neurog1* and *Ascl1* are upregulated following NOTCH inhibition in all monoclonal lines, which have an increased differentiation propensity (Figure 4.3B and C). It has been previously shown that NOTCH components are upregulated upon neural induction (Kobayashi & Kageyama, 2010), and that the NOTCH effectors HES1 and HES5 act as transcriptional repressors of the pro-neural transcription factors in NSCs (Kageyama et al, 2007). It is possible that spontaneously differentiating P19 cells have increased levels of NOTCH signalling components, so that inhibition of

NOTCH leads to even greater increases in pro-neural gene expression. The direction of the change in the pro-neural gene expression, however, discards the possibility that precocious activation or responsiveness to NOTCH signalling in *Hmgn*-knockout cells triggers neuronal differentiation.

In conclusion, the data of this chapter indicate that none of the tested signalling pathways triggers differentiation of P19 cells after the loss of HMGN2, as inhibiting these pathways does not re-establish lineage-specific gene expression to the levels seen in parental levels. WNT and NOTCH signalling seem to act downstream of the ECC commitment to certain lineages, rather than initiating differentiation. FGF stimulation has different effects within the range of pluripotency states; however, these do not include directly promoting neuronal differentiation. These observations suggest that the loss of a major HMGN variant is permissive for the initiation of differentiation following an intrinsic and cell-autonomous program, in agreement with the default model of neural induction.

The next chapter studies the process of neural induction of P19 cells after the loss of a major HMGN variant and evaluates the potential consequences of a higher differentiation propensity when the cells are subject to neuronal differentiation protocols and converted to NSCs.

Chapter 5 Loss of HMGN1 or HMGN2 alters cell fate decisions of derived neural stem cells

5.1 Introduction

The previous chapters showed that loss of HMGN1 or HMGN2 increases the spontaneous neuronal differentiation of P19 cells, and additionally, that this is not mediated by signalling pathways relevant for embryonic neurogenesis. These observations lead to the proposition that P19 cells are more prone to differentiate after the loss of a major HMGN variant following an intrinsic and cell-autonomous program along mainly, but not exclusively, the neural pathway.

This chapter studies whether the higher differentiation propensity of *Hmgn*-knockout cells influences their performance in response to neuronal differentiation protocols. In addition, it evaluates whether HMGNs are required for proper cell fate decisions of other stem cells, specifically of NSCs, or whether this is a unique feature of embryonic pluripotent cells.

During development, neural induction of embryonic cells occurs during gastrulation under BMP/SMAD signalling inhibition when the cells apparently execute an intrinsic default program of differentiation (Hemmati-Brivanlou, 1997; Munoz-Sanjuan & Brivanlou, 2002; Wilson & Edlund, 2001). The neuroectoderm is the first neural structure and comprises cells that turn into NSCs (Kriegstein & Alvarez-Buylla, 2009). The NSCs are multipotent tissue-specific stem cells capable of self-renewal that can generate the main nervous system cell lineages: neurons and glia (Gage, 2000; Kriegstein & Alvarez-Buylla, 2009; Temple, 2001). Once the NSC population is established, the production of neurons and glia occurs in a spatio-temporal specific manner, which is crucial for the development of neural tissues with the correct size and morphology (Hirabayashi & Gotoh, 2005; Miller & Gauthier, 2007).

In the first neural development stage called the expansion phase, the population of NSCs increases in number by symmetric divisions. Subsequently in the neurogenic phase, NSCs undergo asymmetric divisions resulting in the production of neurons without extinguishing the NSC pool. Finally in the gliogenic phase, the NSCs proceed towards glial fates at expense of neuronal production (Qian et al,

2000). Astrocytes appear in late embryonic stages and their generation is a continuous process that peaks peri- and postnatally. In contrast, oligodendrocyte differentiation occurs in waves along the period of astrocyte differentiation (Kessaris et al, 2005). The precise timing of nervous system development involves changes in NSC features that limit their neurogenic potential (Qian et al, 2000). In fact, NSC isolated from early stages and late stages differ in their chromatin conformation (Kishi et al, 2012) and responsiveness to signalling pathways as, for example, NOTCH promotes self-renewal of early-stage NSCs, but favours glial versus neuronal differentiation in later stages (Morrison et al, 2000).

Neural development has been modelled *in vitro* using ESCs and ECCs. The first aim of this chapter is to reproduce previous neuronal differentiation protocols.

5.2 *In vitro* systems of neuronal differentiation of ESCs and ECCs

In ESCs, neural induction is achieved in chemically defined media by withdrawing the factors that maintain pluripotency (Ying et al, 2003b), consistent with the default model. ESC differentiation is asynchronous (Kalkan et al, 2017; Mulas et al, 2017; Ying et al, 2003b); however, after six days of factor withdrawal, most of the cells in the culture are positive for the NSC marker NES and show classical bipolar morphology (Figure 5.1). At this time point, early neurogenesis can be detected by immunostaining for the neurofilament TUBB3 (Figure 5.1). By day 12, the NSCs are arranged in structures that resemble the neural rosettes observed during neural development (Figure 5.1, box), in which the dividing cells occupy the centre (Figure 5.1, arrow), while the differentiated cells migrate to the edges. The production of neurons continues towards later stages without extinguishing the NSC pool and, mimicking the timing of neuronal development *in vivo*, the onset of gliogenesis occurs later, at around day 16 (Figure 5.1).

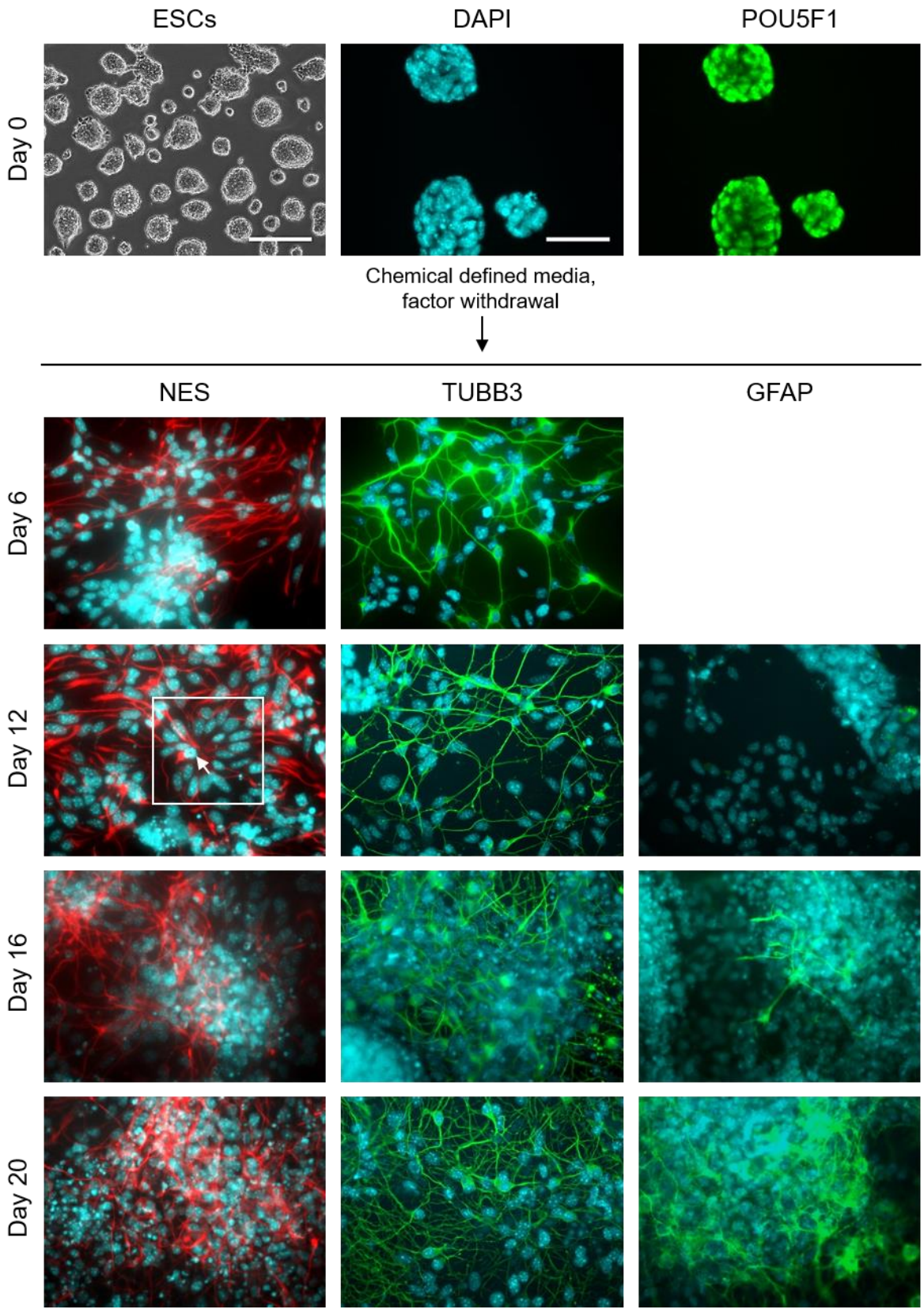


Figure 5.1 ESC differentiation along neural lineages
ESCs grow as round, compact colonies in 2i/LIF chemically defined media, and are positive for POU5F1 (green). Conversion to NSCs is initiated after withdrawal of 2i/LIF. IF at days 6, 12, 16, and 20 are presented for the detection of the NSC marker NES (red), the neurofilament TUBB3 (green), and the glial filament GFAP (green). DAPI was used to stain the nuclei (cyan). The box highlights a structure resembling neural rosettes and the arrow indicates a dividing cell located at the centre. Scale bars represent 200 µm in bright field pictures and 50 µm in IF.

P19 cells progress into NSCs in response to RA stimulation (Jones-Villeneuve et al, 1982). In this system, RA acts as a trigger of differentiation rather than as a neural inducer, as when P19 cells are exposed to RA and BMP, they do differentiate, but into non-neural fates (Hoodless & Hemmati-Brivanlou, 1997). Nevertheless, highly efficient protocols of neuronal differentiation from P19 cells have been developed and most of them include the addition of RA in combination with factors that facilitate neurogenesis. The protocol used here consists of seeding P19 cells in serum-free media containing RA, FGF8, and DAPT (Nakayama et al, 2014). As mentioned in previous chapters, serum contains BMPs that inhibit neurogenesis (Ying et al, 2003a), FGF8 is important for neuronal differentiation of P19 cells (Wang et al, 2006), and DAPT blocks NOTCH signalling (Dovey et al, 2009). NOTCH maintains the pool of NSCs in early stages by antagonising pro-neural transcription factors, and therefore, neurogenesis (Kageyama et al, 2007); thus, the addition of DAPT promotes neuronal differentiation of NSCs.

As shown in Figure 5.2, P19 cells undergo conversion to NSC by three days after induction (dai) as revealed by NES immunostaining. Similarly to ESCs, early neurogenesis is observed at this time point (Figure 5.2). The aim of this protocol is the rapid generation of mature neurons, and it therefore utilises AraC to kill diving cells between days four and six after induction. Hence, the NSC pool extinguishes towards later time points (Figure 5.2), and consequently, gliogenesis is abolished (Nakayama et al, 2014). Meanwhile, neurons are maturing in the cultures as shown by immunostaining for the mature neuron marker MAP2 (Figure 5.2).

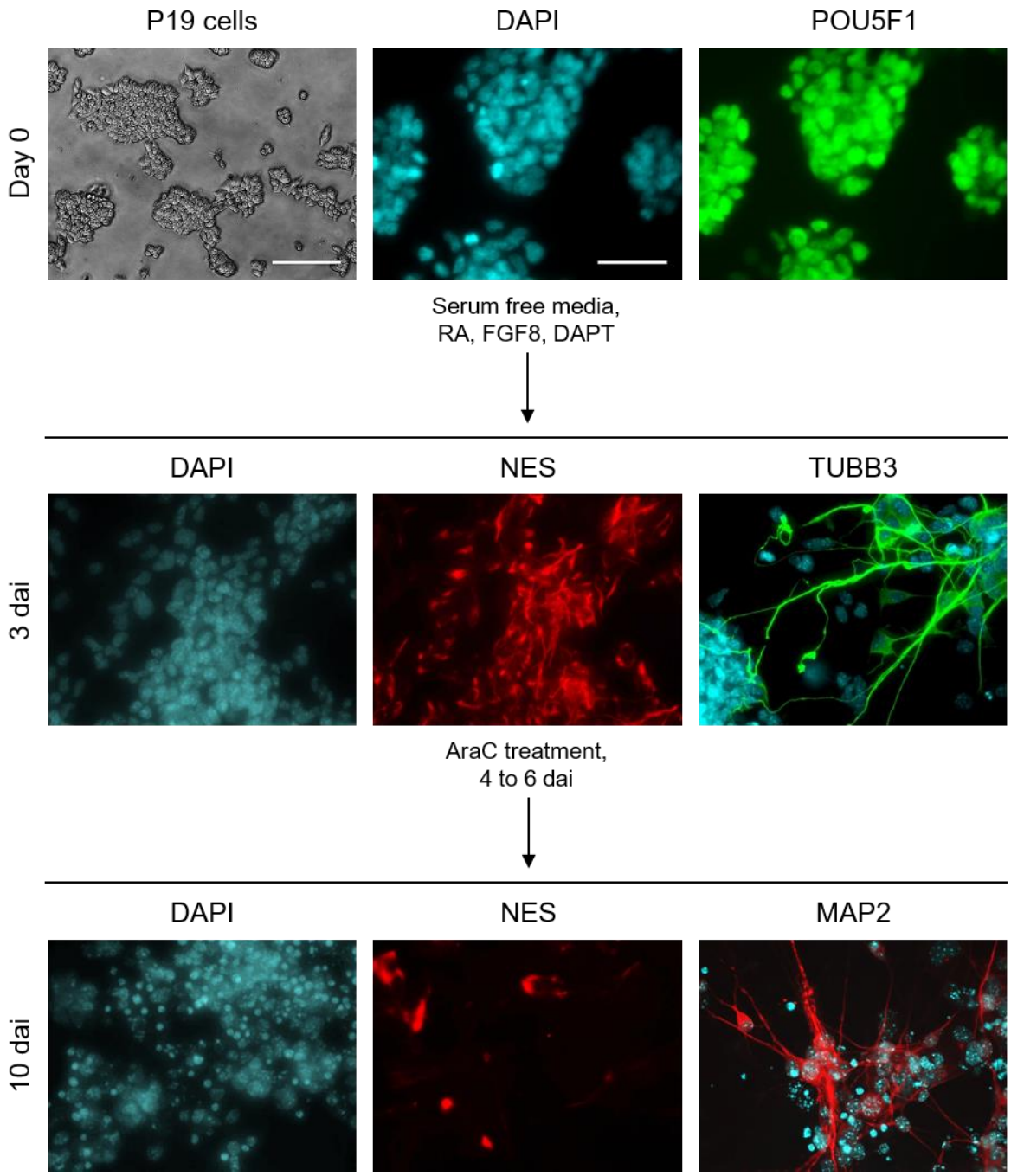


Figure 5.2 Rapid and efficient method of P19 neuronal differentiation
P19 cells grow as colonies with mostly defined edges in serum-supplemented media. They are all positive for the pluripotency marker POU5F1 (green). Neural induction is stimulated by RA, FGF8, and DAPT in serum-free media. AraC is added to the media from 4 to 6 dai. IF images are shown for the NSC marker NES (red), the neurofilament TUBB3 (green), and the neuronal-specific microtubule associated protein MAP2 (red) at 3 and 10 dai, respectively. DAPI was used to stain the nuclei (cyan). Scale bars represent 200 µm in bright field pictures and 50 µm in IF.

5.3 Neuronal differentiation of *Hmgn*-knockout cells

The next aim of this chapter is to determine whether the higher differentiation propensity of the *Hmgn*-knockout cells influences their performance in response to neuronal differentiation protocols. The method described above (Nakayama et al, 2014) was selected as it is rapid, efficient, and well characterised. The focus was on the initial steps just before the AraC treatment (days 1 to 4), as these comprise the period of neural induction and early neurogenesis. Parental, B1, and *Hmgn*-knockout cells were dissociated from one flask each and induced to differentiate by plating the cells in four different dishes, each one with a different end point at 1, 2, 3, or 4 dai. The entire procedure was performed twice and monitored by imaging and RNA analyses.

As shown in Figure 5.3, morphological changes are evident as early as 1 dai. Parental and B1 cells spread out from the colonies that are observed in the undifferentiated cultures (Figure 3.3) and develop cytoplasmic protuberances (Figure 5.3). *Hmgn*-knockout cells already displayed these features before neural induction (Figure 3.3), consistent with a more differentiated phenotype; however, they spread out even more at 1 dai (Figure 5.3).

By 4 dai, the cells seem to concentrate in clusters that are more compact in parental and B1 cells than in *Hmgn*-knockout cells (Figure 5.3). A network of neuronal projections connects these clusters (Figure 5.3). These observations suggest that the clusters of cells work as differentiation centres, perhaps mimicking the function of neural rosettes during neural development. Although no major changes are observed when comparing control and *Hmgn*-knockout neuronal differentiation, there are more cells growing outside the clusters in the *Hmgn*-knockout cultures, thus at least some part of the process is different after the loss of a major HMGN variant. Interestingly, neural induction appears less efficient in D3ΔN1 cells, as the arrangement of the cells at 4 dai is completely different and a layer of flattened cells can be observed below the undefined clusters of cells (Figure 5.3). This result is consistent with previous data suggesting that these cells are prone to differentiate towards non-neural lineages (see chapter 3); thus, this propensity seems to interfere with the neurogenic potential of some D3ΔN1 cells.

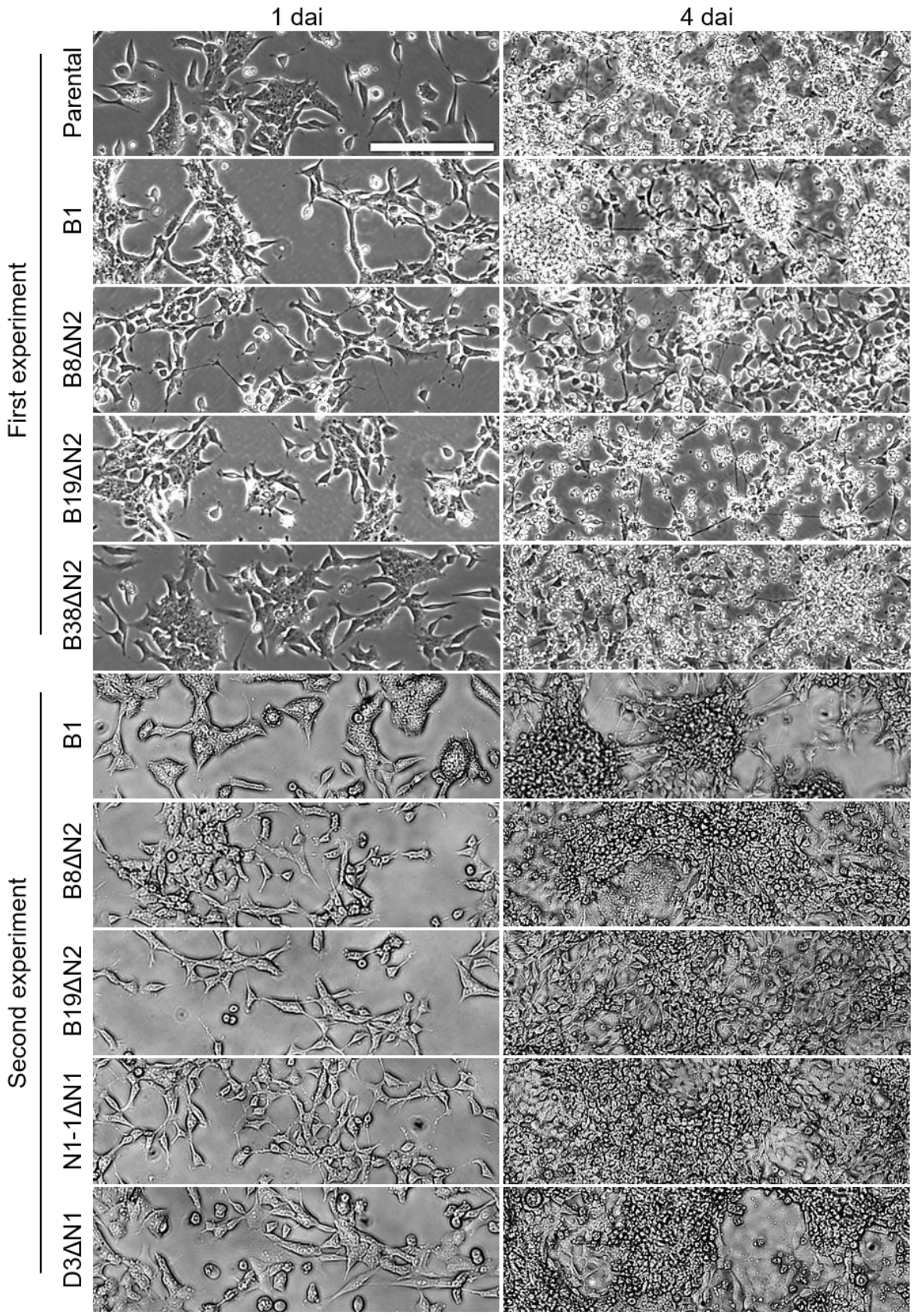


Figure 5.3 Morphological changes in control and *Hmgn*-knockout cells during a rapid and efficient method of neuronal differentiation
Bright field pictures at 1 and 4 dai of a rapid and efficient method of neuronal differentiation. Control and *Hmgn*-knockout cells were induced to differentiate in two independent experiments. Scale bar indicates 200 µm.

To further evaluate the neuronal differentiation process in control and *Hmgn*-knockout cells, RNA was collected at different ending points including 1, 2, 3, and 4 dai, and the expression of relevant genes was determined. It is worth mentioning that the system is not completely reproducible, and therefore, variations from one dish to another and from one experiment to another are observed. Taking this into consideration, the two independent experiments are presented separately, with data from the first experiment in Figure 5.4, and data from the second experiment in Figure 5.5.

The expression of the naïve pluripotency transcription factor *Nanog* is downregulated upon neural induction of parental cells; there is a 50% reduction by 1 dai that continues to decrease in 2 and 3 dai to approximately 100 and 1000 fold less than the initial levels, respectively (Figure 5.4). A similar expression pattern is observed in B1 cells (Figure 5.4 and Figure 5.5), although there is little change between 0 and 1 dai (Figure 3.10A). B8ΔN2, B38ΔN2, N1-1ΔN1 and D3ΔN1 cells have similar responses to B1 cells, while in B19ΔN2 cells *Nanog* is downregulated at 1 dai in a similar manner to parental cells (Figure 3.10A, Figure 5.4, and Figure 5.5). There are no consistent differences in *Nanog* expression when comparing the controls with the *Hmgn*-knockout samples, suggesting that all the cells abandon pluripotency with similar timing.

Neural induction of parental cells is also accompanied by changes in *Nes* expression; this NSC marker is upregulated in the first few days and tends to decrease again by 4 dai, although the timing of this is variable (Figure 5.4). *Nes* induction is also observed in B1 cells; however, the transcript levels are not as high as in parental cells and are comparable to the increased basal levels of B8ΔN2, B19ΔN2, N1-1ΔN1, and D3ΔN1 cells (Figure 3.10D, Figure 5.4, and Figure 5.5). Interestingly, *Nes* expression remains unchanged the first couple of days after neural induction of *Hmgn*-knockout cells, perhaps indicating that the basal transcript levels are sufficient for neural induction of these cells (Figure 5.4 and Figure 5.5). However, the downregulation of the NSC marker at 3 dai is observed in all cell lines, especially in the second experiment (Figure 5.5), suggesting that the process of neural induction is complete by this point and the NOTCH inhibitor DAPT is promoting neurogenesis from NSCs at the expense of self-renewal.

Neurog1 and *Ascl1* are strongly induced in all cases, even though their initial levels vary between the cell lines (Figure 5.4 and Figure 5.5). The pro-neural transcription factors reach their highest transcript levels at 2 to 3 dai, when *Neurog1* transcript is up to 2000 fold and *Ascl1* up to 600 fold more abundant than at basal levels (Figure 5.4 and Figure 5.5). Although both are required for neuronal differentiation, *Neurog1* induction is much higher than *Ascl1* in all cell lines during this protocol. Moreover, their induction is preceded by *Nes* expression and their levels are maintained while *Nes* is downregulated (Figure 5.4 and Figure 5.5). These observations are consistent with the developmental switch from the expansion to the neurogenic phase, which is enhanced by DAPT. Notably, the biological variability observed between experiments (Figure 5.4 and Figure 5.5) is higher than any potential influence of HMGN proteins.

The original neuronal differentiation protocol that was adapted for the current experiments aimed to maximise the rapid generation and maturation of neurons (Nakayama et al, 2014). At 1 dai, the cells are induced to convert into NSCs and by 3 dai a switch from expansion to neurogenic phase is achieved. Subsequently, the newborn neurons start maturing, which is evidenced by branching and neurite formation, involving complex cytoskeleton changes and projections. Neurofilament proteins are an important part of this process, and in agreement, *Tubb3* expression is induced after *Neurog1* and *Ascl1* (Figure 5.4). However, in the second experiment either the efficiency or timing of the differentiation process was not sufficient to detect a peak in *Tubb3* levels (Figure 5.5).

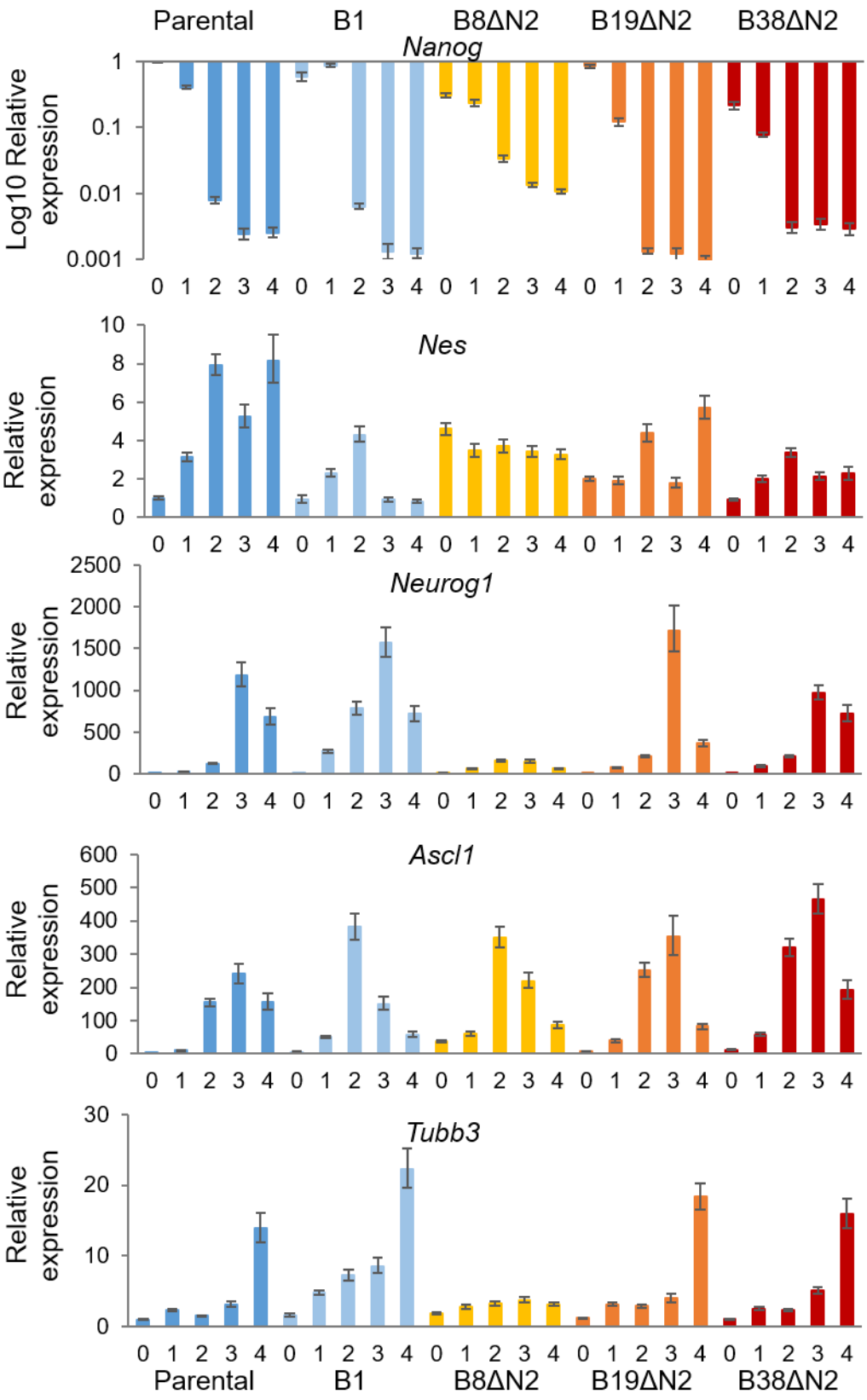


Figure 5.4 Gene expression analyses of control and *Hmgn*-knockout cultures along a rapid and efficient neuronal differentiation method: first experiment
Relative expression of the naïve pluripotency transcription factor *Nanog*, the NSC marker *Nes*, the pro-neural transcription factors *Neurog1* and *Ascl1*, and the neurofilament *Tubb3* along a rapid and efficient neuronal differentiation method as determined by qRT-PCR. The graphs represent the fold change in comparison with undifferentiated parental cultures and error bars symbolise the SD of a PCR triplicate.

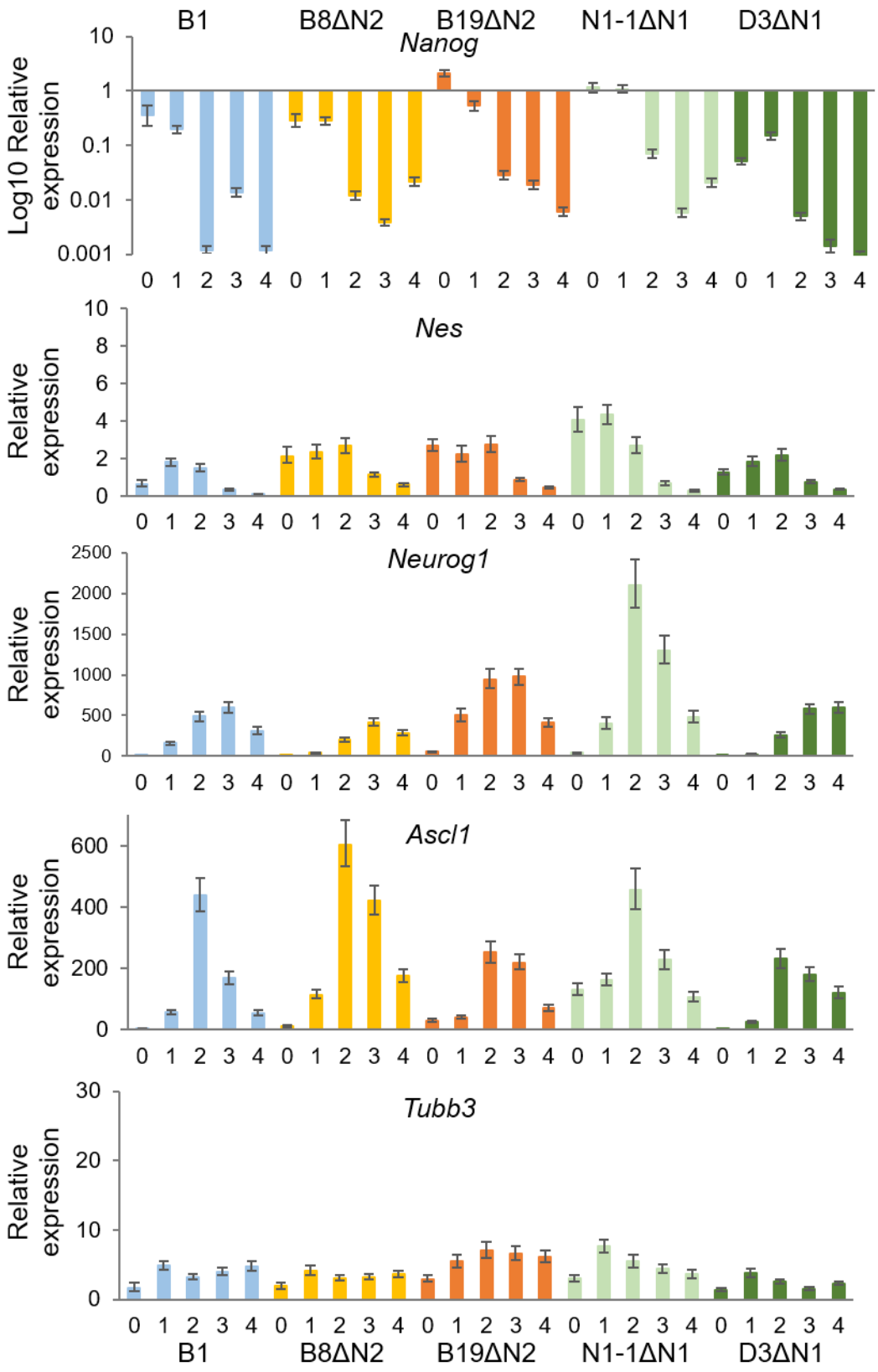


Figure 5.5 Gene expression analyses of control and *Hmgn*-knockout cultures along a rapid and efficient neuronal differentiation method: second experiment
Relative expression of the naïve pluripotency transcription factor *Nanog*, the NSC marker *Nes*, the pro-neural transcription factors *Neurog1* and *Ascl1*, and the neurofilament *Tubb3* along a rapid and efficient neuronal differentiation method as determined by qRT-PCR. The graphs represent the fold change in comparison with undifferentiated parental cultures and error bars symbolise the SD of a PCR triplicate.

Taken together, the bright field images and the RNA analyses reveal that the loss of HMGN1 or HMGN2 influence neither the efficiency of P19 neuronal differentiation, nor the timing and the sequence of the events. All the cell lines follow a process whereby first of all, the cells spread out of compact colonies, *Nanog* expression drops and *Nes* levels are sufficient for neural induction. Then, *Nes* is downregulated, while the pro-neural transcription factors are induced. Consequently, the cells concentrate in clusters functioning as differentiation centres that are connected by neuronal projections. Simultaneously, *Neurog1* and *Ascl1* are reduced and, in some cases, higher transcript levels of *Tubb3* are detected.

Further information is provided by IF performed at 4 dai. Figure 5.6 shows that the pluripotency marker SSEA1 is mostly lost by the end of the experiment in B1 and *Hmgn*-knockout cells. Although a few SSEA1-positive cells are still observed, most of the cells are positive for NES, evidencing the highly efficient ECC to NSC conversion (Figure 5.6). Importantly, NSCs with high NES protein levels are still numerous at 4 dai, when *Nes* transcript levels are already reduced (Figure 5.4 and Figure 5.5). Indeed, changes in gene expression precede both variations at the protein level and differentiation events (Kalkan et al, 2017).

Immunostaining of TUBB3 confirms that the induction of the pro-neural transcription factors and the conversion to NSCs progresses into neurogenesis (Figure 5.7). The protocol promotes the rapid generation and maturation of neurons, which involves branching, neurite formation, and synaptogenesis (Nakayama et al, 2014). In agreement, many projections are observed in B1, B8ΔN2, B19ΔN2, and N1-1ΔN1 cultures at 4 dai (Figure 5.7). However, the cytoskeleton complexity of D3ΔN1 cells is lower, suggesting alterations in either neurite formation or neurogenesis (Figure 5.7). This constitutes further evidence that the preference of D3ΔN1 cells is more likely towards non-neural lineages. Nevertheless, some of the D3ΔN1 cells are still capable of neuronal differentiation since TUBB3-positive cells are identified (Figure 5.7).

Consistent with the above observations and the bright field images showing different cell types in D3ΔN1 cultures (Figure 5.3), considerable numbers of GATA4-positive cells are identified in D3ΔN1 cultures at 4 dai, and interestingly, in B8ΔN2 cultures (Figure 5.8). These cells are located outside the large clusters

of cells where most neurogenesis is happening, and therefore, are spatially separated from NSCs and neurons. GATA4-positive cells are also detected in D3ΔN1 and B8ΔN2 undifferentiated cultures with higher frequencies than in parental (Figure 3.9). Furthermore, *Gata4* expression is upregulated in the undifferentiated state of these two *Hmgn*-knockout lines (Figure 3.10C). Hence, a subset of D3ΔN1 and B8ΔN2 cells are predisposed towards endodermal lineages, reinforcing the idea that HMGNs influence the differentiation propensity of P19 cells rather than inhibiting neurogenesis.

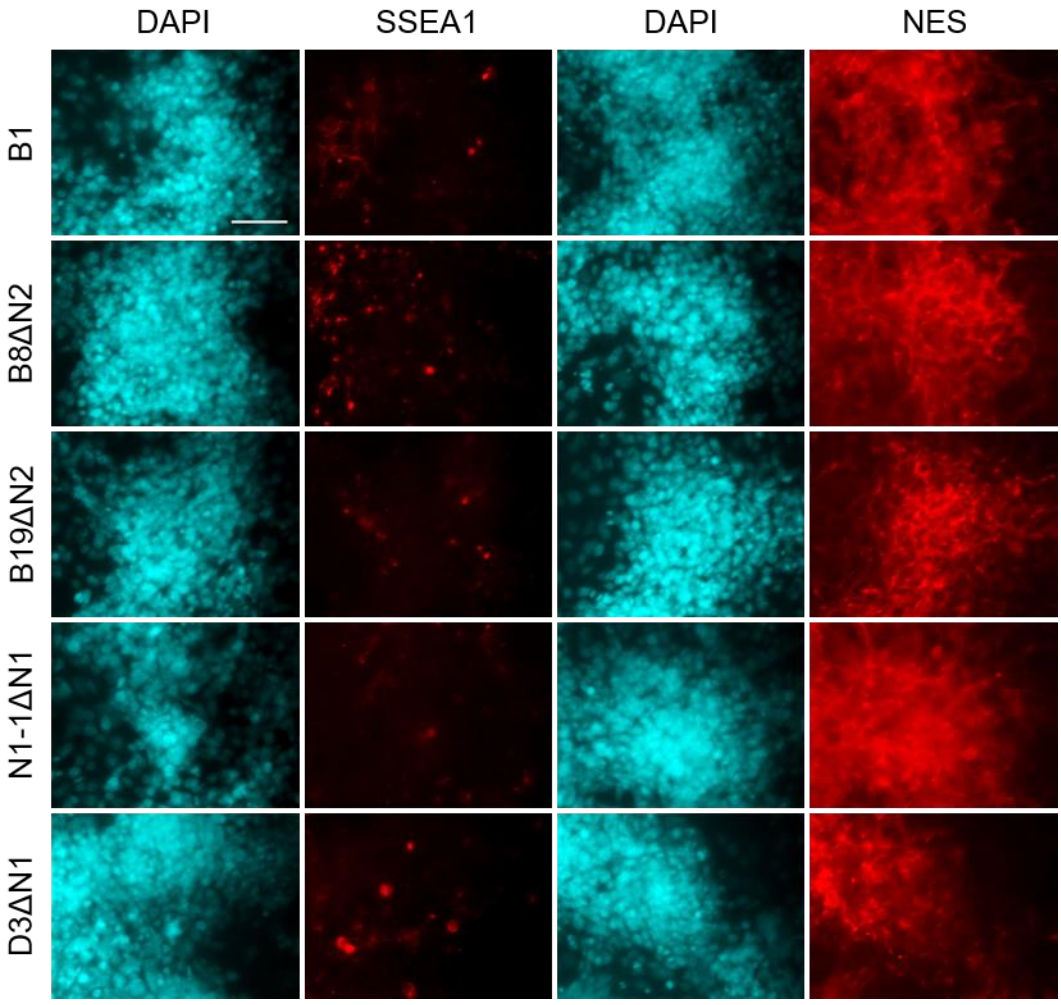


Figure 5.6 Neural induction of control and *Hmgn*-knockout cells results in loss of pluripotency and conversion into NSCs
IF for the detection of the pluripotency marker SSEA1 (left, red) and the NSC marker NES (right, red). B1 and *Hmgn*-knockout cells were fixed at 4 dai of a rapid and efficient neuronal differentiation method. DAPI was used to stain the nuclei (cyan). Scale bar indicates 50 µm.

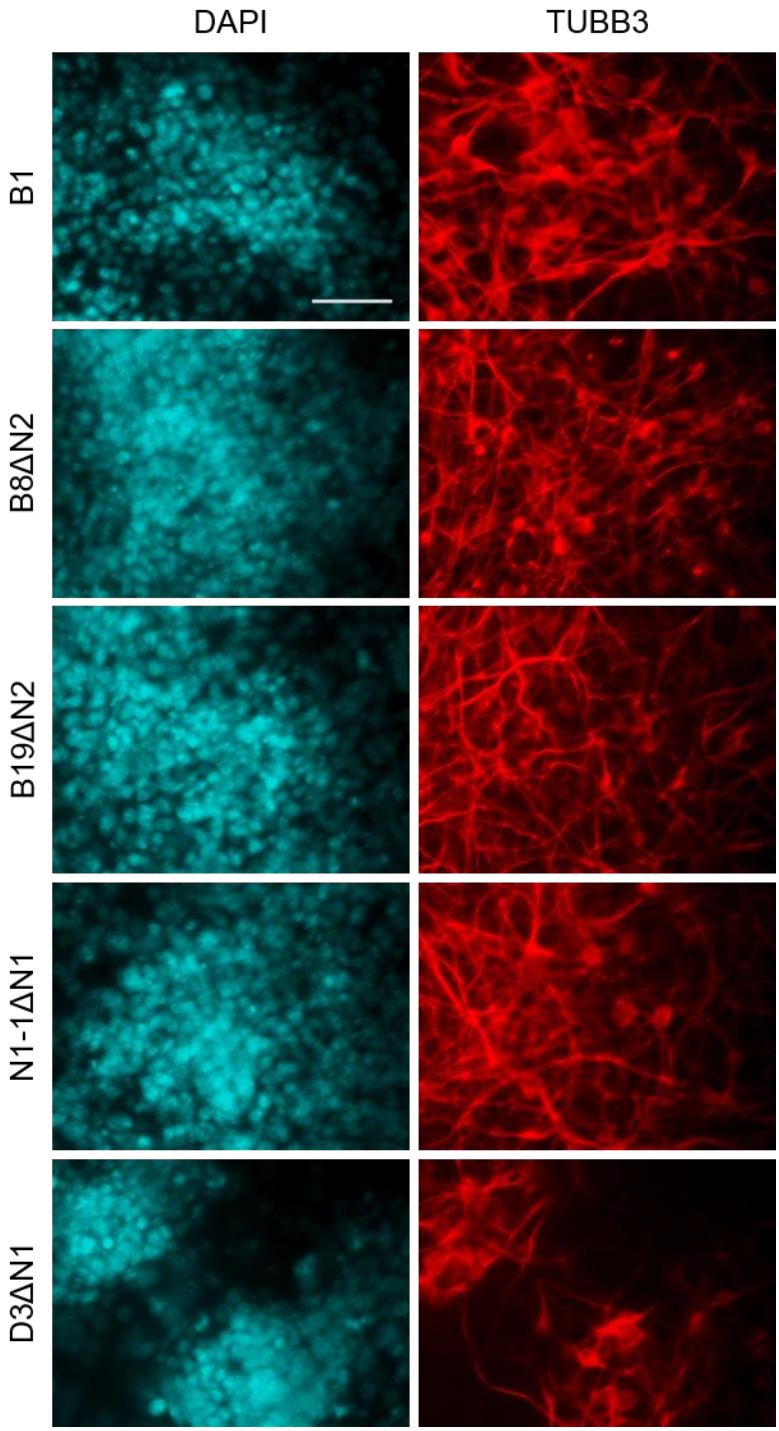


Figure 5.7 Neural induction of control and *Hmgn*-knockout cells generates neurons with complex cytoskeleton projections
IF for the detection of the neurofilament TUBB3 (red) reveals complex cytoskeleton architecture in all cell lines, except in D3ΔN1 cells. B1 and *Hmgn*-knockout cells were fixed at 4 dai of a rapid and efficient neuronal differentiation method. DAPI was used to stain the nuclei (cyan). Scale bar indicates 50 µm.

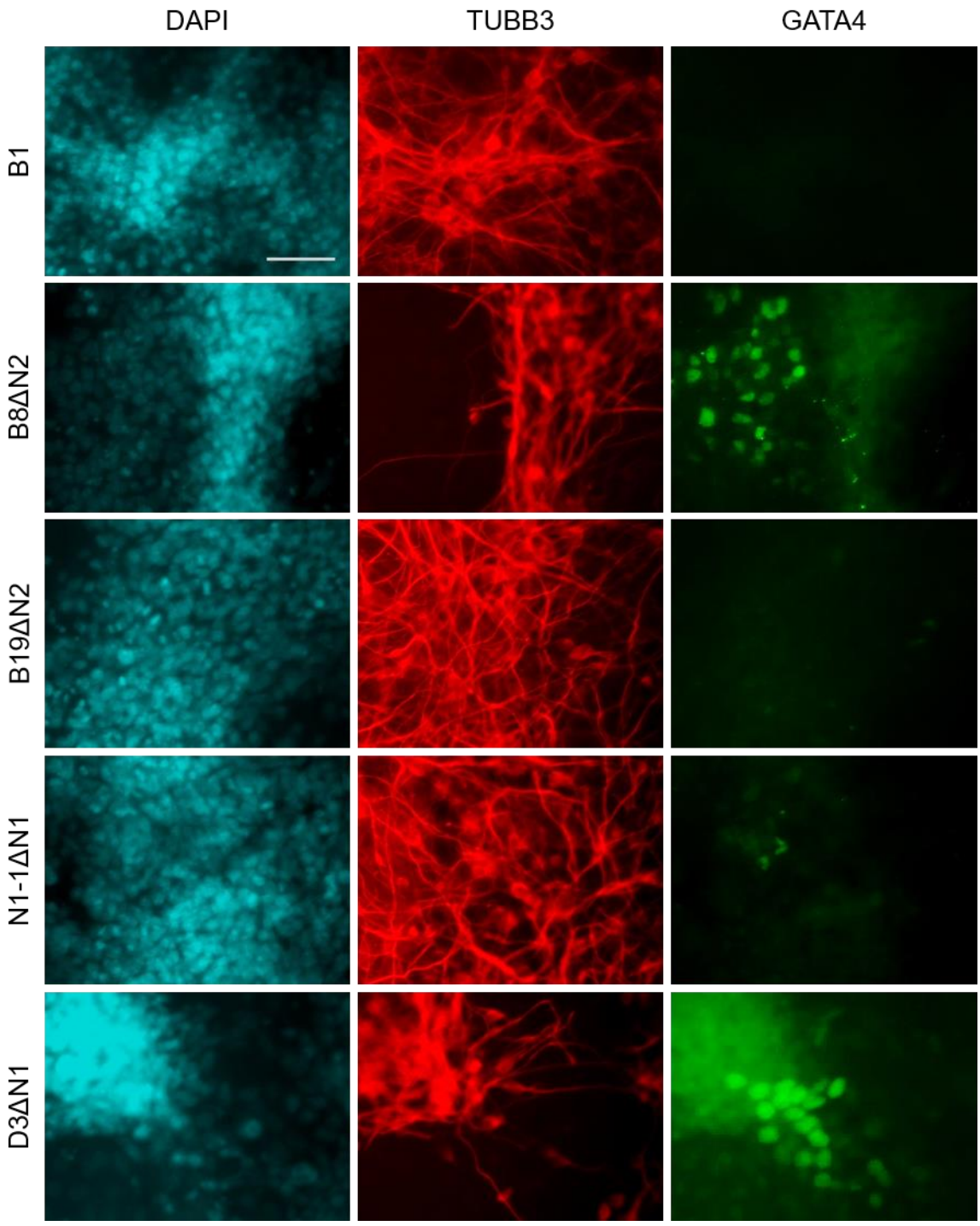


Figure 5.8 Identification of endodermal cells in *Hmgn*-knockout cultures induced to differentiate into neural fates
Double IF for the detection of the neurofilament TUBB3 (centre, red) and the endodermal transcription factor GATA4 (right, green). B1 and *Hmgn*-knockout cells were fixed at 4 dai of a rapid and efficient neuronal differentiation method. DAPI was used to stain the nuclei (cyan). Scale bar indicates 50 µm.

Important conclusions can be derived from these experiments. First, the higher neuronal differentiation propensity of *Hmgn*-knockout cells cultured under non-differentiating conditions does not enhance or accelerate the process of neural induction under active stimulation. Secondly, the propensity towards non-neural fates in the undifferentiated state may decrease the responsiveness to neural inductive cues. Finally, although there is a subset of cells differentiating towards non-neural lineages, there is another subset of cells that retains the potential to undergo neuronal differentiation. Overall, it can be concluded that the loss of a major HMGN variant does not limit pluripotency of embryonic cell cultures and does not seem to favour or inhibit a specific cell lineage.

5.4 HMGN1 and HMGN2 along neural induction of P19 cells

As reviewed in the introduction, the two major HMGN variants are progressively downregulated in the entire embryo during development; however, tissue-specific stem cells and other transiently amplifying precursor cells retain high levels of these proteins (Furusawa et al, 2006; Lehtonen & Lehtonen, 2001; Lehtonen et al, 1998).

The neuronal differentiation protocol presented above comprises all the stages of a differentiation process (see Figure 5.2, Figure 5.6, and Figure 5.8). At 1 dai, there are embryonic pluripotent cells committing to or even starting differentiation towards a certain lineage, in this case the neural lineage. Then by 2 to 3 dai, tissue-specific stem cells, here NSCs, appear in the cultures. Subsequently, differentiated cells, such as neurons, become highly abundant and are expected to mature fully over time. The sequential phases of the protocol facilitate the evaluation of HMGN1 and HMGN2 protein and mRNA levels during neuronal differentiation.

IF for the detection of HMGN1 and HMGN2 in B1, B8ΔN2, B19ΔN2 and B1, N1-1ΔN1, D3ΔN1 cells, respectively, reveals a mosaic expression pattern of the proteins at 4 dai (Figure 5.9A). These results differ from the homogenous distribution observed in undifferentiated cells (Figure 3.2) and are in complete agreement with a developmental regulation of the two major HMGN variants. There is a general trend that the cells in the middle of the differentiation

centres exhibit a brighter signal than the cells at the edges; however, numerous cells escape this trend (Figure 5.9A, see arrows for examples). Determining the identity of high- and low-expressing HMGN cells by IF is not simple considering the high cell density of the differentiation centres and the complex cytoskeleton architecture of the NSCs and neurons. Therefore, different experiments are required to derive accurate associations between the higher or lower HMGN levels and the specific cell types.

Gene expression analyses during neuronal differentiation show that the *Hmgn1* transcript is upregulated in B1, B8ΔN2, B19ΔN2 cells in early stages of the process, corresponding to the NSC conversion of embryonic pluripotent cells (Figure 5.9B). Subsequently, a progressive downregulation causes a 40 to 70% reduction in *Hmgn1* transcript levels by 4 dai, when fully differentiated cells represent a high proportion of the cultures (Figure 5.9B). Similarly, *Hmgn2* expression is induced at 1 dai in B1, N1-1ΔN1, D3ΔN1 cells, followed by a decline of 70 to 90% in comparison with the initial levels in undifferentiated cells (Figure 5.9C).

The previous results recapitulate what observed along embryonic development (Crippa et al, 1991; Furusawa et al, 2006; Lehtonen & Lehtonen, 2001; Lehtonen et al, 1998; Pash et al, 1990), as *Hmgn1* and *Hmgn2* transcript levels are higher in embryonic pluripotent cells (0 dai) than in cultures containing fully differentiated cells (4 dai). However, the NSCs seem to express higher levels of HMGNs than the undifferentiated cells (Figure 5.9B and C). Although this finding has been previously reported, it has not been highlighted or further investigated (Deng et al, 2013; Lehtonen & Lehtonen, 2001). Nevertheless, these observations suggest that HMGNs are important for tissue-specific stem cell biology, particularly for NSCs.

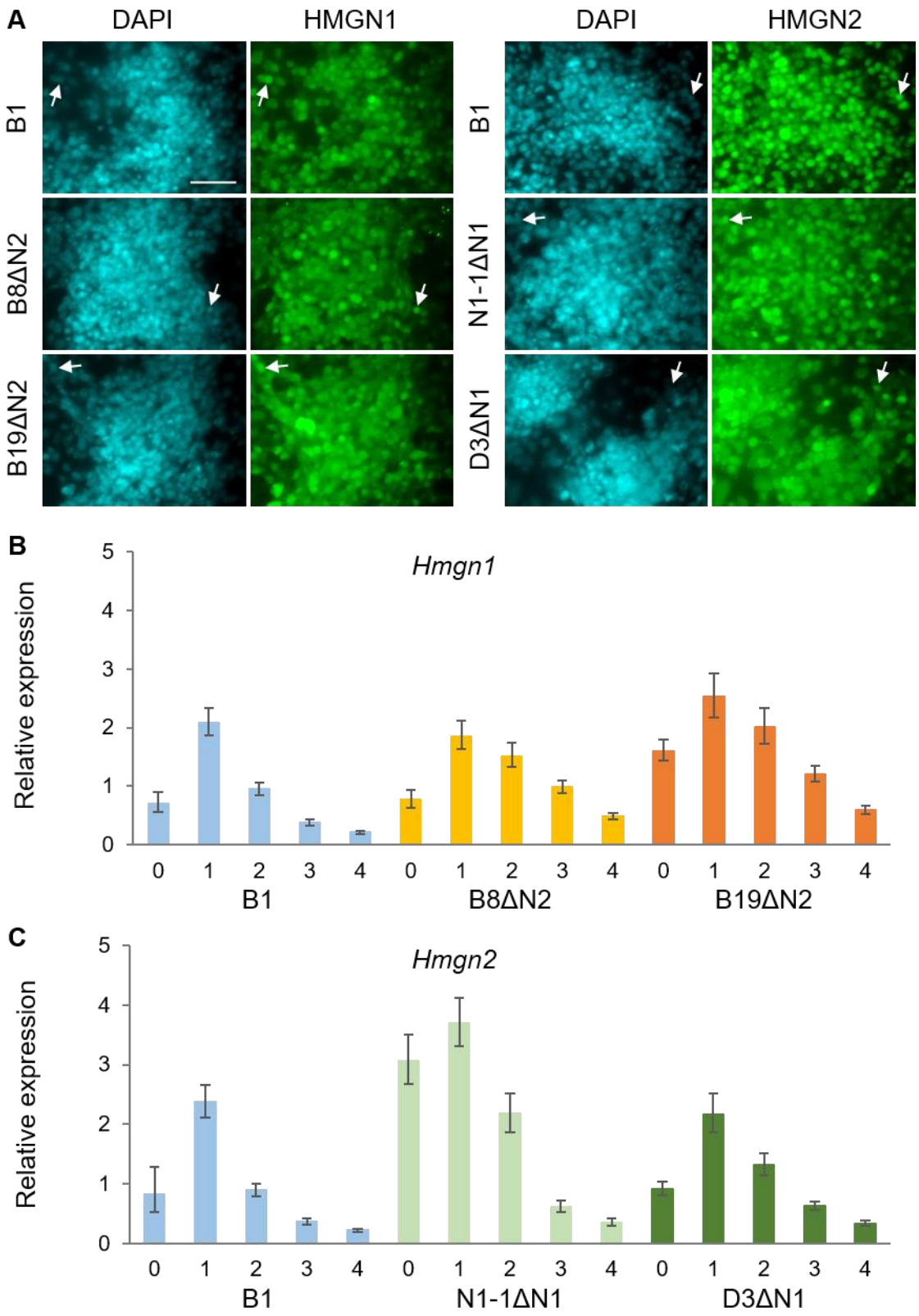


Figure 5.9 HMGN1 and HMGN2 protein and mRNA expression pattern along neuronal differentiation

B1 and *Hmgn*-knockout cells were induced to differentiate following a rapid and efficient method of neuronal differentiation. A) IF for the detection of HMGN1 and HMGN2 (green) at 4 dai are presented. DAPI was used to stain the nuclei (cyan). Arrows highlight cells at the edges of the clusters expressing high levels of HMGNs. Scale bar indicates 50 µm. Relative expression of *Hmgn1* (B) and *Hmgn2* (C) along neuronal differentiation as determined by qRT-PCR. The graphs represent the fold change in comparison with undifferentiated parental cultures and error bars symbolise the SD of a PCR triplicate.

5.5 Loss of a major HMGN variant influences P19 differentiation after serum withdrawal

It was previously shown that the higher neuronal differentiation propensity of *Hmgn*-knockout cells does not accelerate the process of neural induction following a rapid and efficient method of neuronal differentiation for P19 cells. This finding contrasts what was expected; however, it is possible to hypothesise that the active stimulation of differentiation results in an efficient process whatever the initial propensity of the cells. To investigate differentiation in the absence of active stimulation, the cells were subjected to a gentler process that is based on the protocol of ESC differentiation along neural lineages (Ying et al, 2003b). Briefly, P19 cells are routinely propagated in serum-supplemented media and is reasonable to believe that the factors sustaining self-renewal of these cells are contained in the serum. Hence, the serum was replaced by the N2-supplement containing chemically defined factors required for the survival of the cells without preventing them from differentiation. B1 and *Hmgn*-knockout cells were seeded in N2 chemically defined media; bright field images and RNA were collected after 48 h.

As shown in Figure 5.10A, the arrangement of B1 cells cultured in N2 chemically defined media differs from the round compact colonies characteristic of the undifferentiated state (Figure 3.3). Furthermore, some neural projections can be observed in the cells at the edges (Figure 5.10A, magnification), although most of the cells maintain a homogeneous morphology lacking cytoplasmic protuberances, especially those located at the centre. In contrast, most of the B19ΔN2 and N1-1ΔN1 cells have acquired a bipolar morphology (Figure 5.10A, magnifications). D3ΔN1 and, to a lesser degree B8ΔN2, cultures are heterogeneous; however, all the cells exhibit cytoplasmic protuberances as a feature of differentiation and cells with bipolar morphology are also identified (Figure 5.10A, magnifications).

At the gene expression level, *Nanog* is downregulated although in the range of tens of fold instead of hundreds, as was seen during induced neuronal differentiation protocol (Figure 5.5 and Figure 5.10B). This suggests either an inefficient or a delayed differentiation. This scenario might be useful for identifying different responses of the cells to serum withdrawal according to

their diverse differentiation propensities. The variations in *Nanog* levels among the cell lines in N2 chemically defined media may be explained by the variations in the levels of the undifferentiated state, which are lower in B1, B8ΔN2, and D3ΔN1 cells than in B19ΔN2 and N1-1ΔN1 (Figure 3.10A).

Nes transcript levels do not change substantially, similar to the induced differentiation protocol. However, differences among the cell lines that show spontaneous neuronal differentiation (B8ΔN2, B19ΔN2, and N1-1ΔN1) and those that does not (B1 and D3ΔN1) (Figure 3.8) are highlighted after 48 h in N2 chemically defined media (Figure 5.10B).

Neurog1 and *Ascl1* transcripts are strongly induced in all cell lines, although their levels do not reach the highest point observed during the induced protocol (Figure 5.5 and Figure 5.10B). This supports the idea of an inefficient or delayed differentiation that might be useful to identify particularities. Indeed, the gene expression of the pro-neural transcription factors is much higher in B19ΔN2 and N1-1ΔN1 than in other lines, in agreement with their increased neuronal differentiation propensity (Figure 5.10B). Furthermore, D3ΔN1 cells that prefer non-neural lineages show the lowest change in *Neurog1* and *Ascl1* levels (Figure 5.10B). Surprisingly, the pro-neural transcription factors are less induced in B8ΔN2 cells than in B1 (Figure 5.10B). The fact that at least a subpopulation of B8ΔN2 cells prefers endodermal lineages (Figure 5.7) may account for this observation.

Considering all of the above data, it can be concluded that the different cell lines initiate differentiation programs after serum withdrawal, and that the efficiency of neural specification seems to be in agreement with the respective basal neuronal differentiation propensity, which in turn, is influenced by the loss of a major HMGN variant.

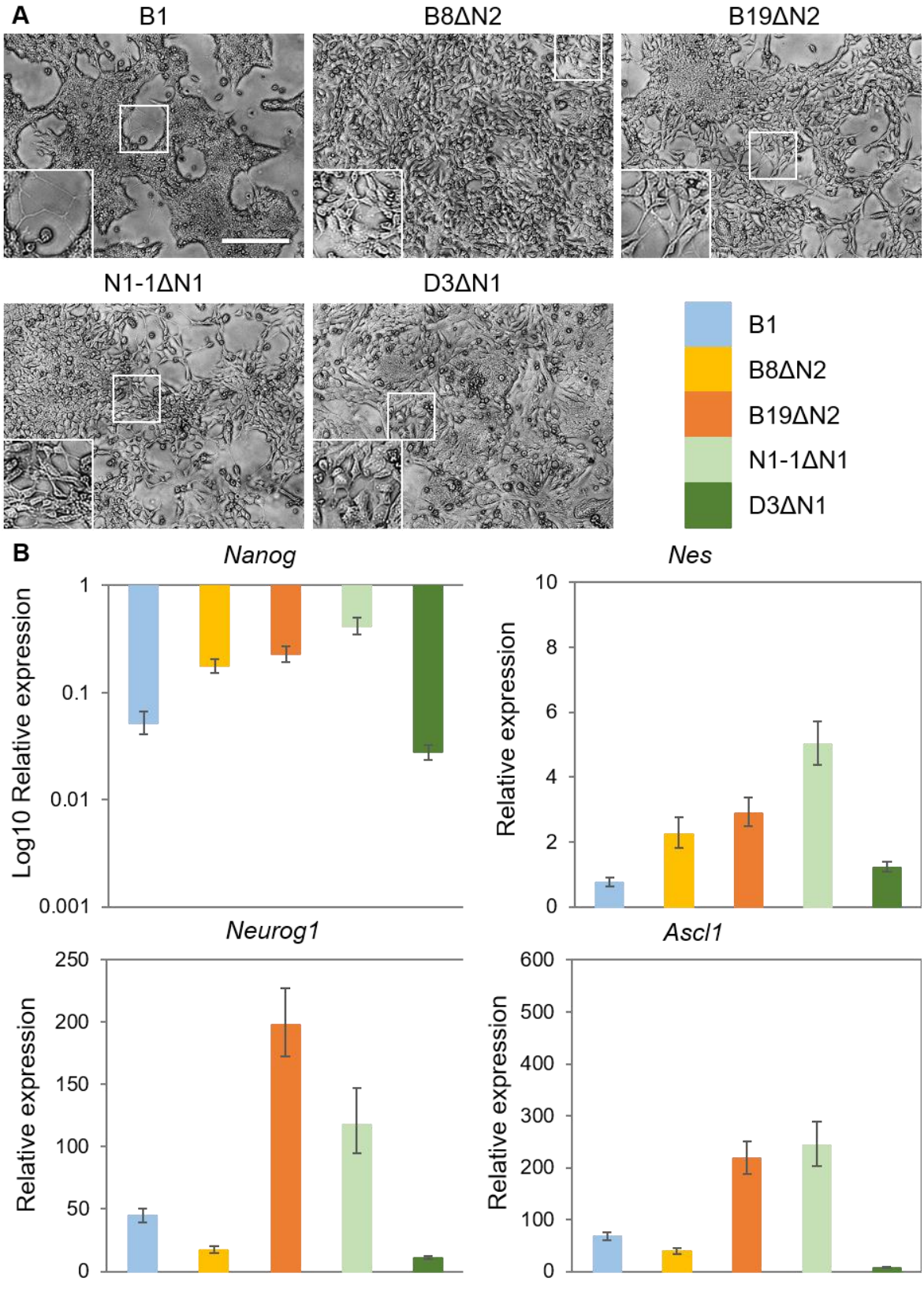


Figure 5.10 Loss of HMGN1 or HMGN2 influences P19 differentiation after serum withdrawal
B1 and *Hmgn*-knockout cells were allowed to differentiate by withdrawing the serum from the culture media. A) Bright field pictures of cultures at 48 h after serum withdrawal. 2X magnifications highlight neural morphology. Scale bar indicates 200 µm. B) Relative expression of the naïve pluripotency transcription factor *Nanog*, the NSC marker *Nes*, and the pro-neural transcription factors *Neurog1* and *Ascl1* as determined by qRT-PCR. The graphs represent the fold change in comparison with undifferentiated parental cells and error bars symbolise the SD of a PCR triplicate.

5.6 Neural stem cell derivation from embryonal carcinoma and embryonic stem cells

It was previously observed that *Hmgn1* and *Hmgn2* transcript levels are higher after conversion of ECCs to NSCs, and before fully differentiated cells represent a high proportion in the differentiating cultures (Figure 5.9). This finding leads to the proposition that HMGNs are important for NSC biology.

To gain further insights into this, a method of NSC derivation from P19 cells was developed in the laboratory based on an established protocol of NSC derivation from ESCs (Conti et al, 2005). Briefly, ESCs are allowed to differentiate in chemically defined media by withdrawing the factors that sustain pluripotency (Ying et al, 2003b). At day 6, when most of the population express the NSC marker NES (Figure 5.1), the cells are dissociated and seeded onto uncoated dishes in chemically defined media containing FGF2 and EGF, allowing the formation of neurospheres. These two growth factors ensure the propagation of NSCs in the laboratory (Gage, 2000). Then, the cells are plated onto gelatine-coated dishes and propagated routinely in the presence of FGF2 and EGF. After a few passages, differentiated progeny are removed from the population and a homogeneous NSC population is established (Conti et al, 2005).

Based on this method, P19 cells are induced to differentiate in serum-free media by the addition of RA, FGF8, and DAPT, following the rapid and efficient method of neuronal differentiation (Nakayama et al, 2014). At 3 dai, most of the cells are positive for the NSC marker NES (Figure 5.2), and therefore, this time point was selected for the switch to the chemically defined media containing FGF2 and EGF. The next day, the cells are dissociated and plated onto gelatine-coated dishes and propagated routinely in the presence of FGF2 and EGF.

NSCs derived from P19 cells and the ESC line, E14, are passaged 8 to 10 times before further evaluation. Bipolar cells are observed under the microscope, forming geometric figures with their cytoplasmic projections that extend from small growing centres (Figure 5.11A). The P19-derived NSCs present thicker cytoplasmic projections and higher cell density in the growing centres (Figure 5.11A, arrows). However, most of the cells are identified as NSC by double immunostaining for the detection of the NSC markers NES and FABP7 (Figure

5.11B). FACS analyses show a mostly homogeneous population expressing NES in P19- and E14-derived NSC cultures. Nevertheless, NSC derivation from P19 cells appears to be less efficient than from E14 cells, since the peak of the population is wider, revealing certain heterogeneity in NES levels. Furthermore, the peak overlaps with the isotype negative control, which indicates that not all the cells are NSCs (Figure 5.11C).

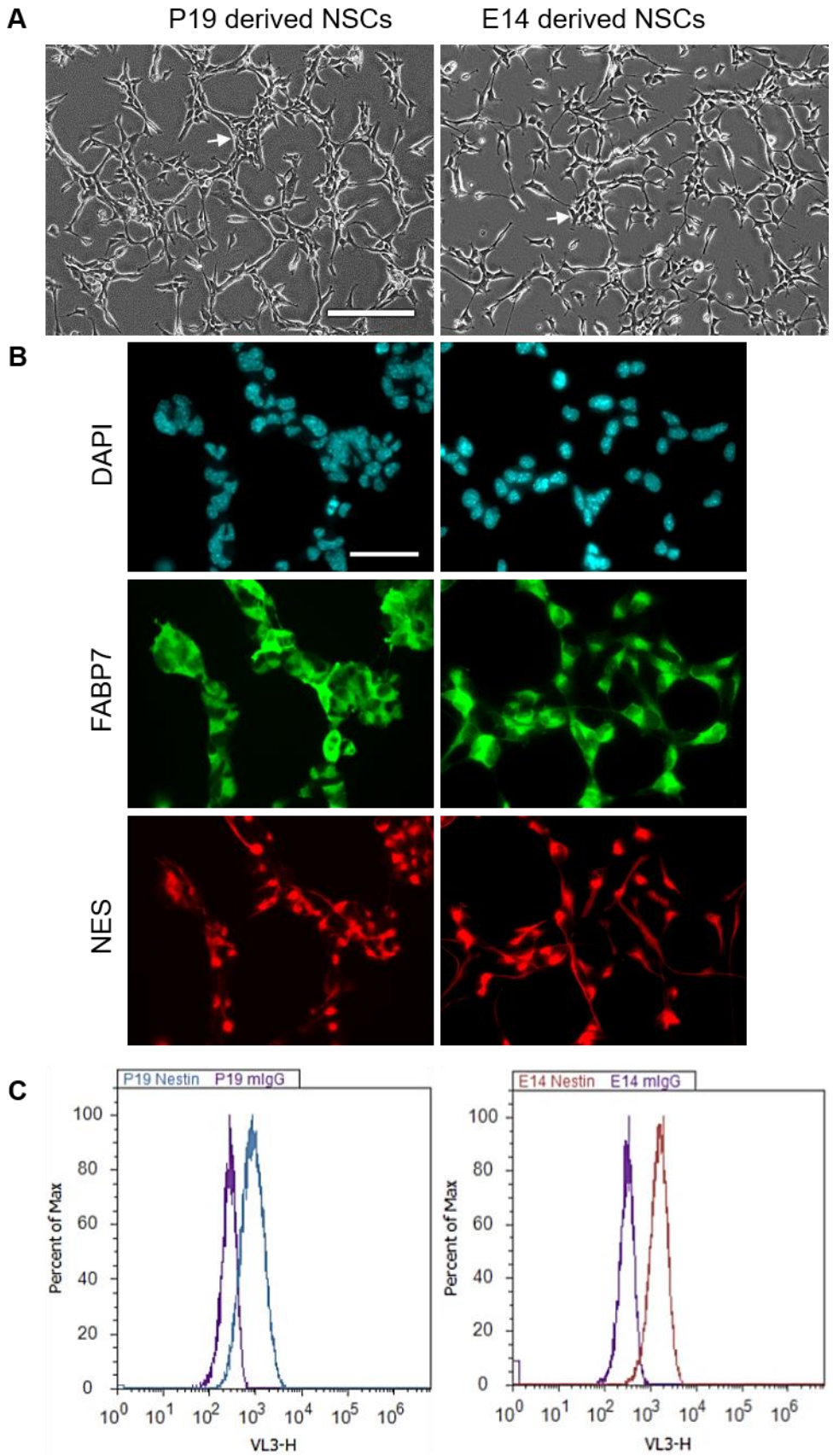


Figure 5.11 NSCs derived from P19 and E14 cells
A) Bright field pictures of NSCs derived from P19 (left) and E14 (right) cells. Arrows show growing centres. Scale bar indicates 200 µm. B) Double IF for the detection of the NSC markers FABP7 (green) and NES (red). DAPI was used to stain the nuclei (cyan). Scale bar indicates 50 µm. C) FACS analyses for the identification of the NSC marker NES in NSCs derived from P19 cells (left, blue) and from E14 cells (right, red), in comparison with the isotype negative control (purple).

E14-derived NSCs are capable of generating the main cell lineages of the nervous system, neurons and glia (Conti et al, 2005). To test whether P19-derived NSCs are also capable of differentiation, growth factors were withdrawn from P19 and E14-derived NSCs, and neurogenesis was observed by immunostaining for the neuron-specific markers TUBB3 and MAP2 (Figure 5.12A). In the presence of serum, the cells differentiate into glial cells with classical stellar morphology that, upon maturation, replaces the NES intermediate filament with the glial fibrillary acidic protein (GFAP) (Figure 5.12B). Interestingly, the differentiating P19 cultures exhibit higher proliferation rates, which results in the emergence of growing and differentiation centres (Figure 5.12A, arrows) and neural rosette-like structures (Figure 5.12B, box).

Gene expression analyses show that the pluripotency transcription factor *Pou5f1* is downregulated in the transition from embryonic pluripotent cell to NSCs in both cell lines, although greater downregulation in E14-derived NSCs is observed. *Pou5f1* transcript levels are 100,000 fold less in E14-NSCs in comparison with the undifferentiated state, while in P19-NSCs they are 1,000 fold less (Figure 5.12C). Upon growth factor withdrawal from the NSC cultures, E14 cultures maintain silenced *Pou5f1* expression, while P19 cultures show increased *Pou5f1* expression (Figure 5.12C). It is possible that *Pou5f1* expression is misregulated in NSCs and neuronal cells derived from P19 cells, due to their oncogenic mutations. Alternatively, it may be that residual undifferentiated pluripotent cells are present in the P19-NSC cultures, and that these cells continue to proliferate after the removal of growth factors, thus making up a larger proportion of the population. The latter hypothesis may explain the NES-negative cells detected by FACS (Figure 5.11).

Nes and *Map2* transcript levels are much higher in the NSC stage than in the undifferentiated cells, reflecting the self-renewal and neurogenic features of NSCs (Figure 5.12C). These two transcripts are at comparable levels in P19- and E14-derived NSCs, despite the higher basal levels detected in the undifferentiated ECCs. After growth factor withdrawal, *Nes* and *Map2* are downregulated exclusively in differentiating P19-NSCs (Figure 5.12C), although greater numbers of NSCs and neurons are observed in P19 cultures than in E14. It

is possible that other cell types, i.e. embryonal cells, might be present in P19 cultures that are decreasing the proportion of the neural lineage cells.

Gfap expression is negligible in the undifferentiated and the NSC stages. However, it is induced upon factor withdrawal in both cell lines (Figure 5.12C), which is in agreement with a gain of gliogenic potential and a reduction of neurogenic potential in the late stage NSCs.

Finally, *Hmgn2* levels are higher in NSCs derived from both P19 and E14 cells and drop upon NSC differentiation (Figure 5.12C). This is consistent with previous observations (Figure 5.9) and supports the hypothesis that HMGNs are important for NSC biology.

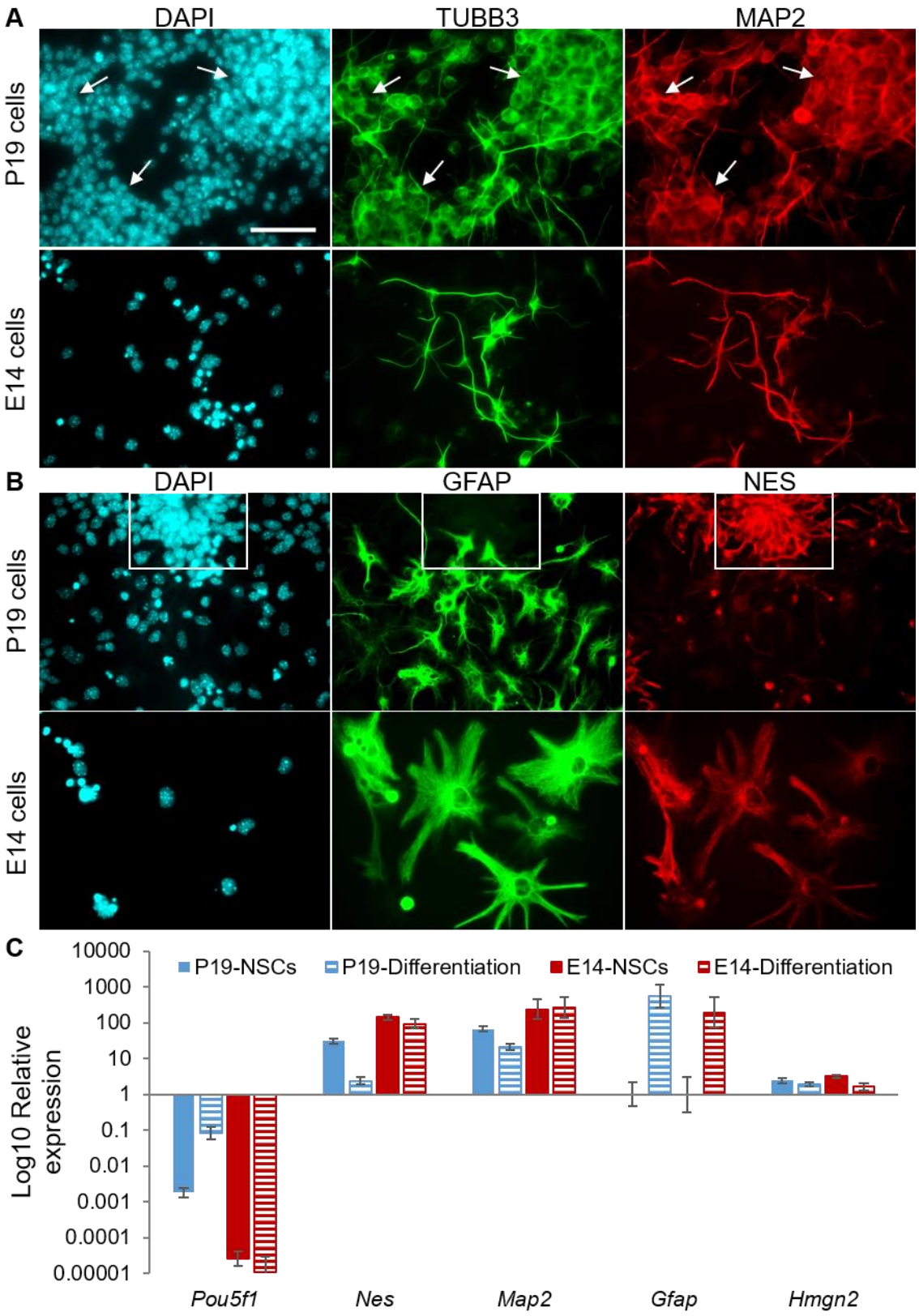


Figure 5.12 P19- and E14-derived NSCs possess neurogenic and gliogenic potential
NSCs derived from P19 and E14 were allowed to differentiate by growth factor withdrawal for six days (A and C), or induced to generate glia (B). Co-immunostaining was performed to detect (A) TUBB3 (green) and MAP2 (red), or (B) GFAP (green) and NES (red). DAPI was used to stain the nuclei (cyan). Arrows show growing and differentiation centres and a neural rosette-like structure is highlighted by the box. Scale bar indicates 50 µm. C) Relative expression of relevant genes as determined by qRT-PCR. The graph represents the log10 of the fold change in comparison with undifferentiated E14 or P19 cells, respectively. *Actb* was utilised as the reference gene. Error bars symbolise the SD of a PCR triplicate.

5.7 Neural stem cells lacking HMGN proteins differentiate with higher frequency and exhibit altered gene expression

In order to test the hypothesis that HMGNs are important for NSC biology, NSCs were derived from parental, B1, and *Hmgn*-knockout cells, following the method described above. After 8 to 10 passages, bipolar cells that are positive for the NSC marker NES represent the majority of the population in all lines (Figure 5.13). Although D3ΔN1 cells prefer non-neural fates, they are capable of generating NSCs (Figure 5.13), supporting the idea that the loss of a major HMGN variant does not limit pluripotency to a certain lineage in a population of P19 cells.

In B1 cells, the transition from the undifferentiated state to NSCs recapitulates what is observed in parental cells (Figure 5.12, blue columns and Figure 5.14A, light blue columns); *Pou5f1* expression decreases by approximately 1,000 fold, *Nes* expression is induced by 30 fold, and *Map2* transcript levels are around 65 fold higher than the undifferentiated state. Furthermore, *Sox2* transcript levels are similar to those of undifferentiated parental cells, as SOX2 is a marker of both ESCs and NSCs (Figure 5.14A, light blue columns). Conversely, *Neurog1* expression is downregulated (Figure 5.14A, light blue columns), possibly due to NOTCH signalling that maintains NSC self-renewal by antagonising the transcriptional activation of pro-neural transcription factors (Kageyama et al, 2007).

In *Hmgn*-knockout NSCs, *Sox2* transcript levels are comparable to those in B1-NCSs, even when the individual cultures prior to differentiation had varying steady state mRNA levels (Figure 5.14A). *Nes* and *Map2* are induced in all cases, reaching similar levels despite the initial variability (Figure 5.14A). The expression pattern of these three genes is consistent with the NSC identity of these cells. Nevertheless, important alterations are observed in the NSCs derived from *Hmgn*-knockout cells: *Pou5f1* is only slightly downregulated and its expression is not even 10 fold less than in the undifferentiated populations, while *Neurog1* expression remains 5 to 7.5 times higher than in undifferentiated parental cells (Figure 5.14A). Interestingly, *Pou5f1* and *Neurog1* levels of NSCs

lacking HMGN1 or HMGN2 are similar to those observed after growth factor withdrawal in parental-NSCs (Figure 5.14B).

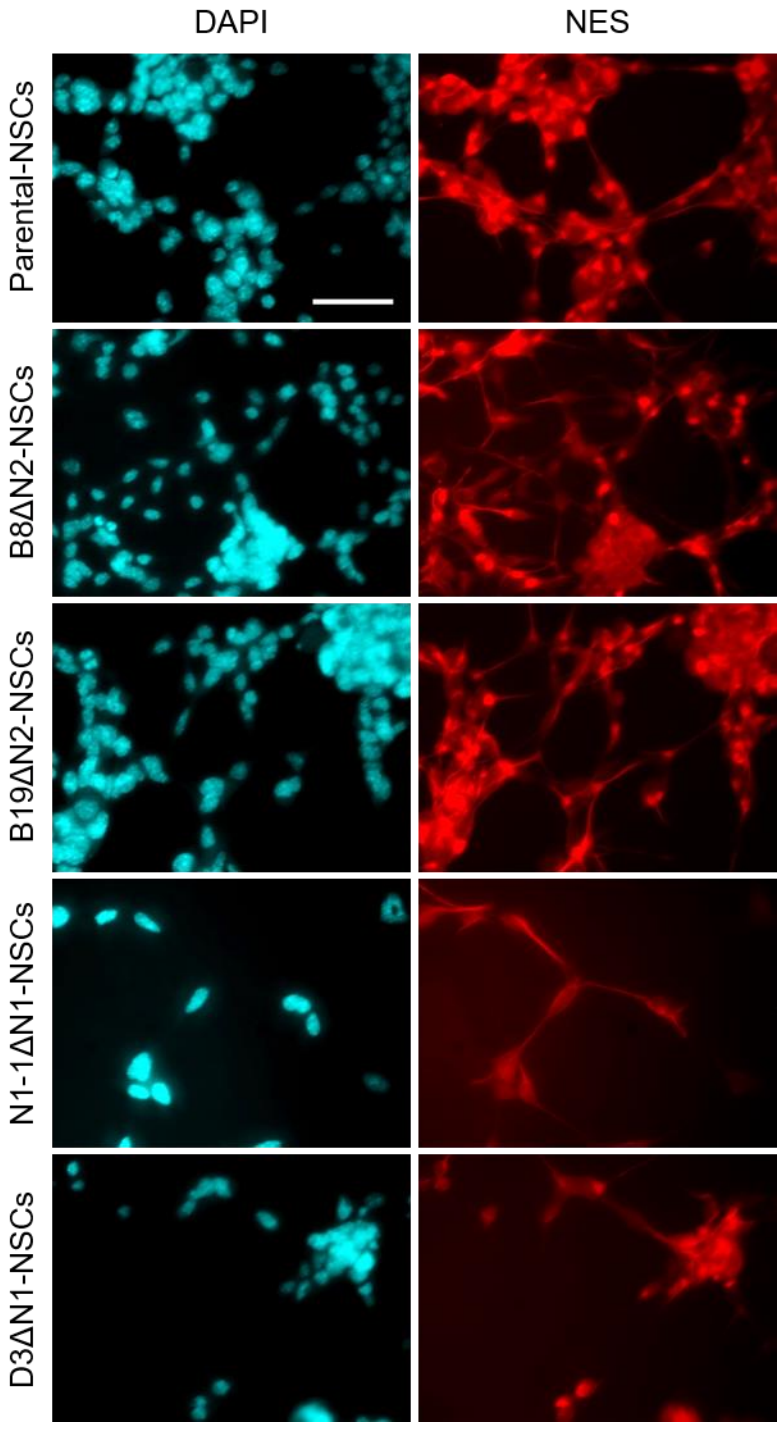


Figure 5.13 NSCs derived from parental and *Hmgn*-knockout cells
Parental and *Hmgn*-knockout cells were induced to differentiate following a rapid and efficient method of neuronal differentiation (Nakayama et al, 2014). At 3 dai, there is a switch to chemically defined media containing FGF2 and EGF. The next day, the cells were dissociated and seeded onto gelatine-coated dishes in the presence of FGF2 and EGF. After 8 to 10 passages in the mentioned conditions, the cells were fixed and IF were performed for the detection of the NSC marker NES (red). DAPI was used to stain the nuclei (cyan). Scale bar indicates 50 µm.

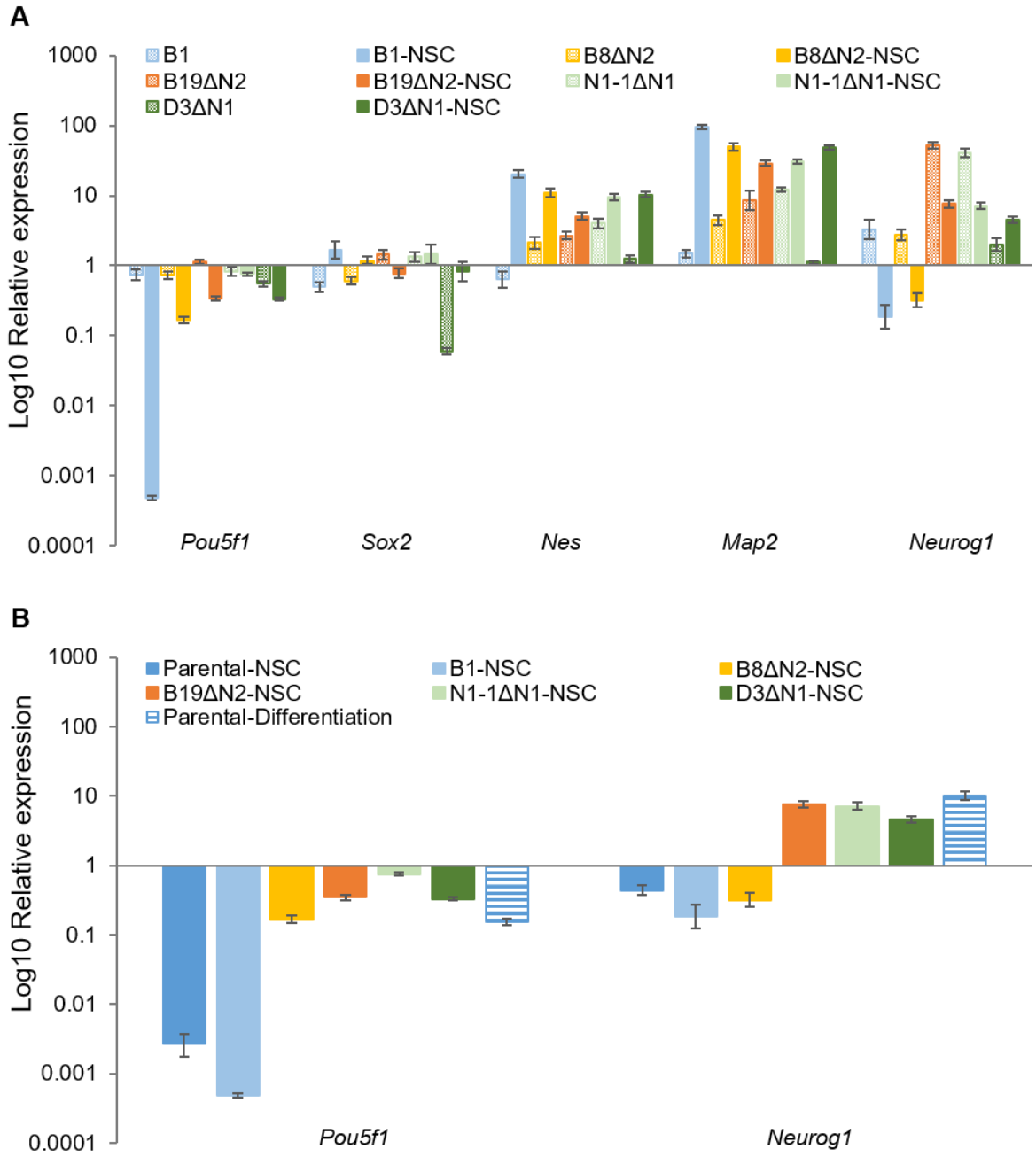


Figure 5.14 NSCs lacking a major HMGN variant exhibit altered gene expression
A) Relevant gene expression changes along the transition from the undifferentiated to the NSC state in B1 and *Hmgn*-knockout cells, as determined by qRT-PCR. B) Relative expression of the pluripotency transcription factor *Pou5f1* and the pro-neural transcription factor *Neurog1* in NSCs derived from parental, B1, and *Hmgn*-knockout cells in comparison to differentiating parental-NSCs, as determined by qRT-PCR. The graphs represent the log10 of the fold change in comparison with undifferentiated embryonal carcinoma parental cells and error bars symbolise the SD of a PCR triplicate.

To further evaluate the alterations presented by the NSCs lacking HMGN1 or HMGN2, IF was performed for POU5F1 and TUBB3. Notably, the growing and differentiation centres previously identified in P19-derived NSCs (Figure 5.11A,

arrows) contain cells expressing POU5F1 (Figure 5.15, arrows). Numerous clustered POU5F1-positive cells are also observed in NSCs derived from *Hmgn*-knockout cultures (Figure 5.15, arrows). Immunostaining for the intermediate neurofilament TUBB3 reveals the presence of increased numbers of neurons in the growing centres in all *Hmgn*-knockout cell lines (Figure 5.16).

The data show that the growing and differentiation centres observed in P19 and *Hmgn*-knockout derived NSCs appear to contain POU5F1-positive cells (Figure 5.15), NES-positive cells (Figure 5.13), and TUBB3-positive neurons (Figure 5.16). It seems likely that these cell clusters work as differentiation centres in a similar manner to the clusters observed during induced neuronal differentiation (Figure 1.7). However, it is not possible to say from the data whether these clusters comprise three different cell types growing side by side, or whether individual cells express all three genes. Dual-colour FACS analysis of NSC and differentiated cell populations would be required to investigate this further. Nevertheless, it is possible to hypothesise that undifferentiated P19 cells expressing POU5F1 are present in these clusters of NSC cells, and that after growth factor withdrawal and neuronal differentiation, these cells continue to proliferate, leading to a relative increase in overall *Pou5f1* transcript levels (Figure 5.12C). This would mean that P19 cells do not respond as efficiently as E14 cells to the conditions for deriving and maintaining NSCs.

Similarly, the increased *Pou5f1* transcript levels in NSCs derived from *Hmgn*-knockout lines could reflect an increase in the number of pluripotent undifferentiated cells in these populations. This leads to the proposition that after the loss of a major HMGN variant, the cells are even less responsive than parental cells to the growth factor conditions that selectively promote NSC maintenance and propagation. Residual undifferentiated cells lacking either HMGN1 or HMGN2 survive and proliferate in response to these growth factors at higher rates than parental cells as evidenced by the numerous POU5F1-positive cells (Figure 5.15), and the higher *Pou5f1* expression (Figure 5.14). Furthermore, the *Hmgn*-knockout cells that successfully undergo NSC conversion are not propagated with the same efficiency, as higher neuronal differentiation events are observed (Figure 5.16), consistent with the upregulation of the pro-neural

transcription factor NEUROG1 (Figure 5.14). Therefore, HMGNs seem to be required for the maintenance of the cellular identity of NSCs.

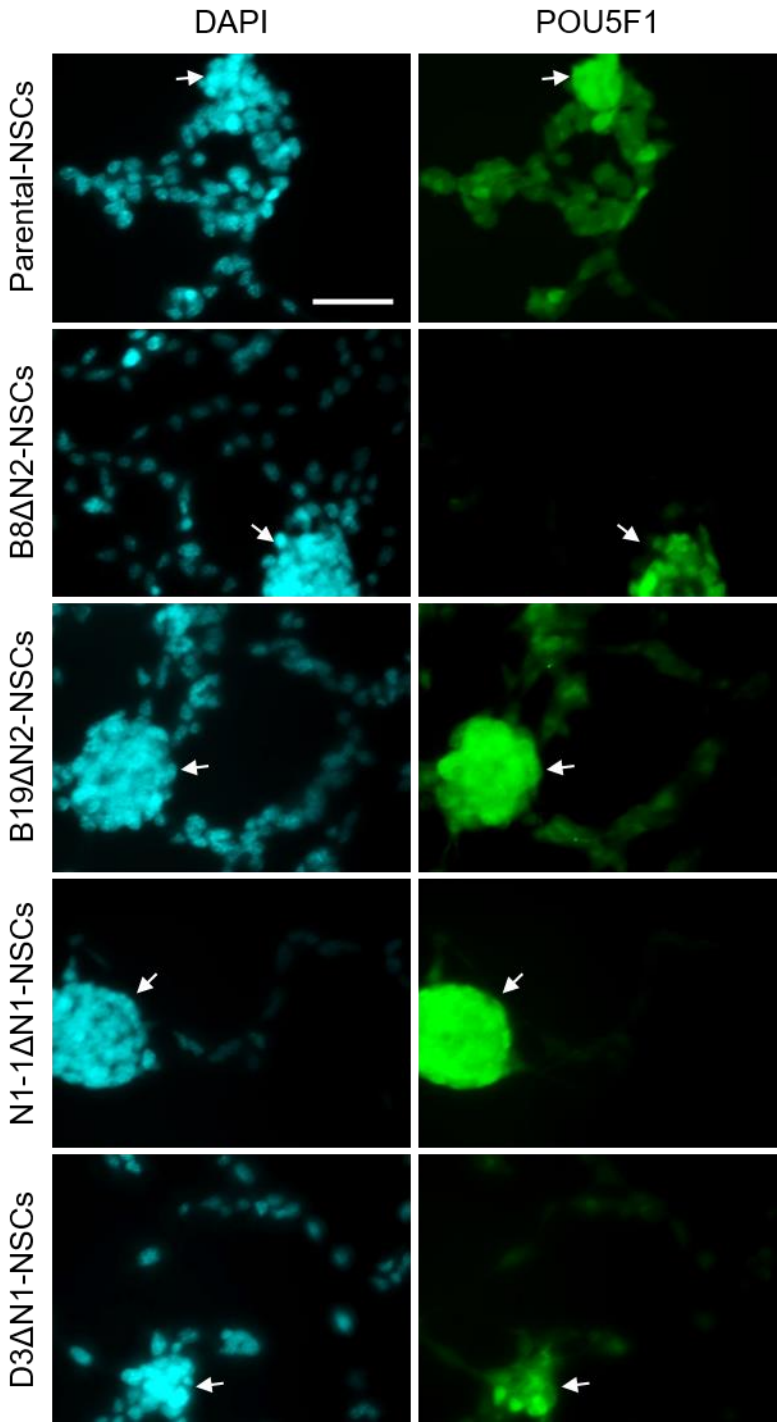


Figure 5.15 NSC cultures lacking a major HMGN variant contain numerous pluripotent cells NSCs derived from parental and *Hmgn*-knockout cells were fixed after 8 to 10 passages under NSC propagation conditions and IF were performed for the detection of the pluripotency transcription factor POU5F1 (green). DAPI was used to stain the nuclei (cyan). Arrows show the growing centres. Scale bar indicates 50 µm.

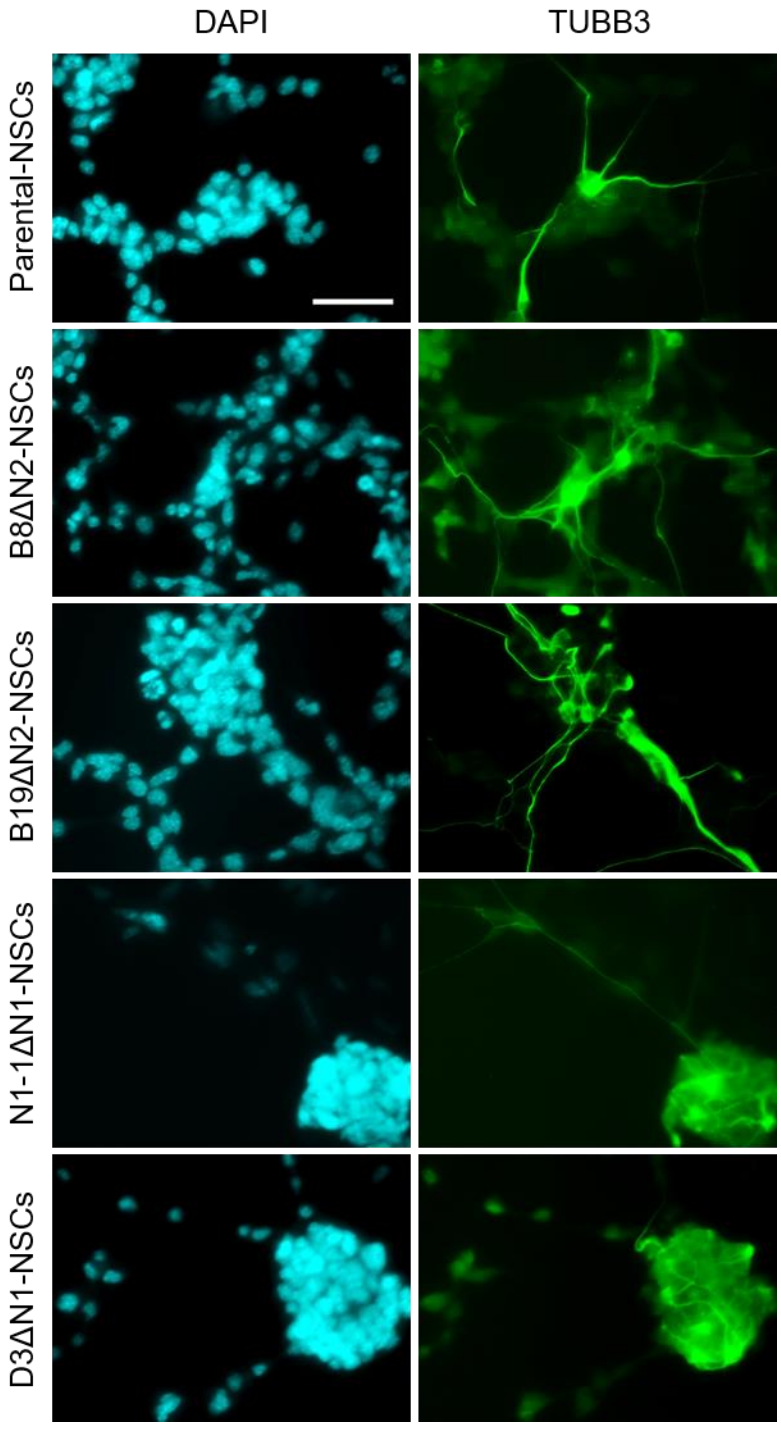


Figure 5.16 NSC cultures lacking a major HMGN variant contain numerous neurons
NSCs derived from parental and *Hmgn*-knockout cells were fixed after 8 to 10 passages under NSC propagation conditions and IF were performed for the detection of the neurofilament TUBB3 (green). DAPI was used to stain the nuclei (cyan). Scale bar indicates 50 µm.

5.8 Discussion

The loss of a major HMGN variant does not have a major influence on the neural induction of P19 cells after stimulation with factors relevant for neurogenesis (see Figure 5.3 to Figure 5.8). Nevertheless, the higher propensity of

unstimulated cells to differentiate towards neural fates following serum withdrawal (mainly observed in N1-1ΔN1 and B19ΔN2) increases the transcription rates of pro-neural transcription factors (Figure 5.10), which could result in the induction of neuronal differentiation. Moreover, the propensity of some *Hmgn*-knockout cells to differentiate into non-neural fates (mainly observed in D3ΔN1 cells) interferes with their neuronal differentiation upon stimulation. However, there are some D3ΔN1 cells that respond to these neurogenic cues, suggesting that the loss of HMGNs does not restrict the pluripotency of all P19 cells within the population.

Previous studies have tested the neurogenic potential of ESCs lacking a single or double HMGN variant. The first study implemented a standard protocol of dopaminergic neuron differentiation of ESCs derived from *Hmgn1*-knockout mice and their control littermates, which includes the addition of growth, selection, and specification factors (Deng et al, 2013). The authors did not observe either differences in the efficiencies of NSC and neuron production, or major gene expression changes in the NSCs and neurons derived from them (Deng et al, 2013). These findings are consistent with the results presented in Figure 5.3 to Figure 5.8, supporting that the loss of a major HMGN variant neither enhance, nor extinguish the neurogenic potential of embryonic pluripotent cell cultures.

In a more recent work, the authors tested the neurogenic potential of double-knockout ESCs lacking both HMGN1 and HMGN2 (Deng et al, 2017). This time the cells were allowed to differentiate by withdrawing the factors that sustain pluripotency, similar to the strategy utilised in Figure 5.10. Several genes showed altered expression along the protocol in the double-knockout cells. Gene ontology is a bioinformatic resource utilised for the identification of affected biological pathways considering a set of missregulated genes. The gene ontology analysis of the upregulated genes failed to recognise specific related pathways; however, it revealed alterations in neuron differentiation, neuron morphogenesis, neuron projection development, neuron projection morphogenesis, transmission of nerve impulse, synaptic transmission, and axogenesis, when considering the downregulated genes (Deng et al, 2017). These findings suggest that HMGNs are involved in neuron maturation processes, which are out of the scope of the present and the cited works. Nevertheless, D3ΔN1

cultures exhibit fewer projections at 4 dai (Figure 5.7 and Figure 5.8), and therefore, it would be interesting to further investigate whether HMGNs are indeed required for the processes identified by gene ontology analyses.

The fact that HMGN1 and HMGN2 are progressively downregulated along development (Crippa et al, 1991; Furusawa et al, 2006; Lehtonen et al, 1998; Pash et al, 1990), does not rule out the possibility that these proteins play an important role in fully differentiated cells as they are ubiquitously expressed in all adult tissues (Johns, 1982). However, it highlights their relevance in ESCs and other stem cells. As observed in Figure 5.9, *Hmgn1* and *Hmgn2* expression is induced in the ECC to NSC transition, consistent with higher transcription rates of *Hmgn1* in NSCs derived from control ESCs during dopaminergic neuron differentiation (Deng et al, 2013). Indeed, previous reports have emphasised the presence of HMGN proteins in progenitor cells committed to differentiate but still capable of self-renewal (Furusawa et al, 2006; Lehtonen & Lehtonen, 2001). However, these findings have not been further investigated.

NSCs can be derived in the laboratory from P19 cells following the induction of neuronal differentiation and selection with FGF2 and EGF (Figure 5.11). The presence of the NSC markers FABP7 and NES confirm the identity of the cells, further supported by higher *Nes* transcript levels and downregulation of pluripotency transcription factors (Figure 5.12C, solid blue). Moreover, the derived NSCs seem to activate NOTCH signalling as *Neurog1* expression is reduced (Figure 5.14, solid blue), and this pathway promotes self-renewal at expenses of neurogenesis. Nevertheless, they are capable of neuronal, and additionally, glial differentiation (Figure 5.12A and B), in line with their *in vivo* counterparts (Temple, 2001). The ability to derive NSCs from P19 cells allowed the evaluation of the potential role of HMGNs in these tissue-specific stem cells. NSC lacking a major HMGN variant show increased *Neurog1* and *Pou5f1* expression than control NSCs, resembling the differentiation process followed by parental cells after growth factor withdrawal (Figure 5.14B). Furthermore, numerous events of spontaneous neuronal differentiation are observed in NSCs derived from *Hmgn*-knockout cells (Figure 5.16), suggesting that the HMGNs are important for the balance between self-renewal and differentiation of tissue-specific stem cells. Indeed, limb bud mesenchymal progenitor cells derived from

Hmgn1-knockout mice differentiate faster into chondrocytes (Furusawa et al, 2006).

The proper timing of neurodevelopment is essential for the size, morphology and function of neural tissues, which is guaranteed by molecular switches between expansion, neurogenic, and gliogenic phase of NSC differentiation (Hirabayashi and Gotoh, 2005, Miller and Gauthier, 2007). It is possible to speculate that the spontaneous neuronal differentiation of NSC lacking a major HMGN variant leads to precocious neuronal differentiation *in vivo* that might compromise the pool of NSCs and the subsequent generation of glia. In line with this speculation, *Hmgn1*-knockout mice contain fewer NSCs in the SVZ (Deng et al, 2013), a widely known neurogenic niche where the developing and adult NSCs reside (Bjornsson et al, 2015). In addition, siRNA-mediated downregulation of *Hmgn1*, *Hmgn2*, or *Hmgn3* in mouse NSCs results in increased neurogenesis at expenses of astrocyte differentiation both *in vitro* and *in vivo* (Nagao et al, 2014), which may be explained by a reduced NSC pool at late stages when gliogenesis starts. Furthermore, the *Hmgn1* and *Hmgn2* double-knockout mice exhibit reduced numbers of oligodendrocytes in the spinal cord and related behavioural alterations (Deng et al, 2017). The mechanism proposed by the authors, however, involves the accumulation of a post-translational histone modification repressive for gene expression at the promoter regions of *Olig1* and *Olig2*, the master regulators of oligodendrocyte differentiation. It would be interesting to test whether the accumulation of repressive marks is related with a neurogenic fate of cells and with a precocious neuronal differentiation.

In conclusion, the present experiments suggest that HMGNs are necessary for the maintenance of the stem cell identity, which relies in mechanisms promoting self-renewal and inhibiting differentiation. As previously reviewed, developmental fate decisions rely on the interaction between extrinsic cues and intrinsic programs, and epigenetic events are important mediators of this interaction. The HMGNs are chromatin architectural proteins that modify local chromatin structure and mediate chromatin accessibility for other epigenetic players. This may explain the changes in gene expression and the altered cell fate decisions of *Hmgn*-knockout cells. The next chapter investigates the epigenetic landscape of these cells.

Chapter 6 HMGN1 and HMGN2 support an open chromatin conformation in embryonal carcinoma cells

6.1 Introduction

The data shown in the previous chapters indicate that HMGN proteins are important for safe warding the identity of stem cells. Specifically, the loss of a major HMGN variant increases the propensity of stem cells to differentiate into the lineages dictated by intrinsic programs, at expenses of their self-renewal.

Self-renewal of ESCs is influenced by chromatin modulators. As reviewed in the introduction, chromatin of ESCs exhibits a globally open conformation that is permissive for the particularly high transcriptional activity of these cells (Efroni et al, 2008). Accordingly, histone modifications associated with open and active chromatin, such as acetylation and H3K4me3, are enriched in ESCs (Azuara et al, 2006; Efroni et al, 2008; Krejčí et al, 2009). Interestingly, increasing histone acetylation through HDAC inhibition promotes self-renewal and prevents differentiation of ESCs (Lee et al, 2003; Qiao et al, 2015; Ware et al, 2009), while reducing HAT activity in ESCs leads to differentiation (Fang et al, 2014). Additionally, H3K4me3 is highly abundant in ESCs and its levels drop during differentiation; maintenance of this modification is required for ESC self-renewal, and optimises somatic cellular reprograming into iPSCs (Ang et al, 2011).

At the same time as maintaining a globally open chromatin conformation, the silencing of developmental genes is important for preventing the precocious lineage commitment of pluripotent stem cells (Boyer et al, 2006). The abundance of the transcriptional repressive mark H3K27me3 is lower in ESCs than in their differentiated progeny (Marks et al, 2012), consistent with a globally open chromatin conformation; however, it is enriched at promoters of developmental genes (Boyer et al, 2006). Interestingly, some of these promoters are also marked with H3K4me3, and are thought to be poised for transcriptional activation if required for initiating differentiation programs (Azuara et al, 2006; Bernstein et al, 2006). These bivalent domains tend to resolve into monovalent domains during differentiation (Bernstein et al, 2006; Mikkelsen et al, 2007).

The above mechanisms seem to be conserved in P19 cells. For instance, HDAC inhibitors decrease *Neurod1* expression (Fang et al, 2010), which is in agreement with previous findings that inhibition of HDAC prevents differentiation of ESCs (Lee et al, 2003). Furthermore, the enrichment of PcG proteins and their deposited histone modification H3K27me3 around the *Neurod1* TSS is higher in undifferentiated P19 cells than after RA-induced neuronal differentiation (Yellajoshyula et al, 2012). The presence of H3K4me3 at the *Neurod1* TSS in undifferentiated P19 cells suggests that *Neurod1* is bivalently marked (Yellajoshyula et al, 2012). Likewise, other genes relevant for neurogenesis show bivalent marks in undifferentiated P19 cells and their transcriptional activation during P19 neuronal differentiation involves loss of PcG-mediated repression and enrichment of H3K4me3 and H3K9ac (Yellajoshyula et al, 2012).

HMGN1 and HMGN2 proteins are known to influence histone modifications (Deng et al, 2017; Lim et al, 2005; Ueda et al, 2006). In particular, they are required to maintain proper cellular levels of H3 acetylation, in a mechanism that involves enhancing HAT function (Lim et al, 2005; Ueda et al, 2006). Furthermore, HMGNs indirectly prevent H3K27me3 accumulation at lineage-specific genes (Deng et al, 2017). The mechanism consists of antagonising the binding of linker histone H1 to chromatin; bound H1 recruits PcG to specific gene loci, increasing H3K27me3 local levels (Deng et al, 2017).

Considering the above information, the aim of this chapter is to investigate whether the loss of a major HMGN variant influences the histone modification landscape of P19 cells, in particular the levels of H3 acetylation, H3K4me3, and H3K27me3. The first objective focuses on histone modifications at specific gene loci, as changes in local chromatin architecture might explain the alterations in the expression of pluripotency and lineage-specific genes observed in *Hmgn*-knockout cells. The second objective investigates whether the loss of HMGN2 influences the global levels of certain histone modifications.

6.2 Chromatin preparation and immunoprecipitation controls

To achieve the first objective, quantitative ChIP-PCR for the detection of histone modifications at promoters and regulatory regions of pluripotency and lineage-specific genes was performed. Chromatin samples from parental, B19ΔN2, and B38ΔN2 cells were previously prepared by a Master student in the laboratory, Mariarca Bailo, and stored for further analysis. The present work utilised the stored samples. Briefly, chromatin was sonicated to produce DNA fragments of adequate sizes for PCR amplification (between 100 to 1000 bp). Figure 6.1 shows that the sizes of the DNA fragments are comparable between parental, B19ΔN2, and B38ΔN2 chromatin preparations and rely within the required range.

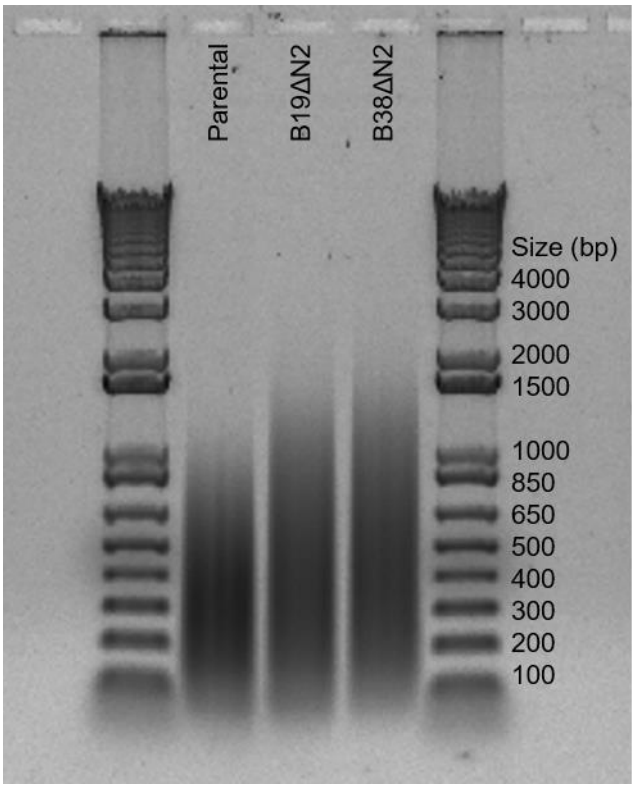


Figure 6.1 Sizes of DNA fragments after sonication of chromatin samples
Gel electrophoresis of DNA from parental, B19ΔN2, and B38ΔN2 chromatin samples. After sonication, chromatin samples were subjected to a reverse cross-linking reaction to release the DNA. Histones and chromatin bound proteins were enzymatically degraded and the DNA was purified in order to analyse the length of the fragments by gel electrophoresis. The DNA ladder flanks the samples and contains bands of the indicated sizes in bp.

DNA and protein concentrations of chromatin samples were determined by standard methods and results are provided in Table 6.1. Although there are some discrepancies between the two methods, both were considered for calculating the volume of each sample required to yield comparable results. The volumes per reaction indicated in Table 6.1 were maintained in all ChIP-PCR experiments.

Table 6.1 Chromatin amount loaded per ChIP reaction

Chromatin sample ID	DNA (ng/µl)	Protein (µg/µl)	Vol/rxn (µl)	DNA/rxn (µg)	Protein/rxn (µg)
Parental	172.6	2.8	30	5.2	84
B19ΔN2	108.4	2.3	40	4.3	92
B38ΔN2	101.1	2.2	40	4.0	88

Chromatin was immunoprecipitated with antibodies targeting H3K4me3, H3K27me3, H3K9ac, H3K27ac, and H3K122ac. In addition, antibodies recognising HMGN1 and HMGN2 proteins determined whether B19ΔN2 and B38ΔN2 chromatin is in fact devoid of HMGN2, and investigated whether there is functional compensation by HMGN1 in the *Hmgn2*-knockout cells (Deng et al, 2015). An antibody that binds the unmodified C-terminal tail of H3 was used to compare chromatin levels between samples and was subsequently used as a normaliser. Finally, since all antibodies were raised in rabbit, rabbit IgG (rIgG) was used as a negative control to detect non-specific binding during the ChIP protocol.

After ChIP, the DNA was purified and the enrichment relative to input DNA or H3 enrichment was determined for specific genomic regions by qPCR. Primer sets targeting these genomic regions were designed considering mouse ES Bruce4 information deposited in the UCSC genome browser and generated by the Ren lab as part of the ENCODE/LICR project (http://genome.ucsc.edu) (Consortium, 2012; Kent et al, 2002). Several locations along the pluripotency genes *Pou5f1* and *Nanog* were investigated, especially those corresponding to H3K4me3 peaks, which mark active promoters, and to sites of colocalisation of both H3K4me1 and H3K27ac, which is characteristic of active enhancers. Specifically, *Pou5f1*-1 primer set targets a potential enhancer marked by H3K4me1 and H3K27ac, while *Pou5f1*-2 and *Pou5f1*-3 amplify regions upstream and downstream of the first exon, respectively, both highly enriched in H3K4me3 (Figure 6.2). Similarly, *Nanog*-1 primer set is located in a region enriched in H3K4me1, *Nanog*-2 and *Nanog*-3 flank the first exon occupied by H3K4me3, and *Nanog*-4 targets an

intronic region that is also enriched in H3K4me3 (Figure 6.2). Primer sets for the lineage-specific genes *Neurog1*, *Ascl1*, *Gata4*, and *Tbxt* target potential bivalent domains near each TSS that are enriched in both H3K4me3 and H3K27me3 (Figure 6.2). In addition, *Actb* serves as a control for active transcription.

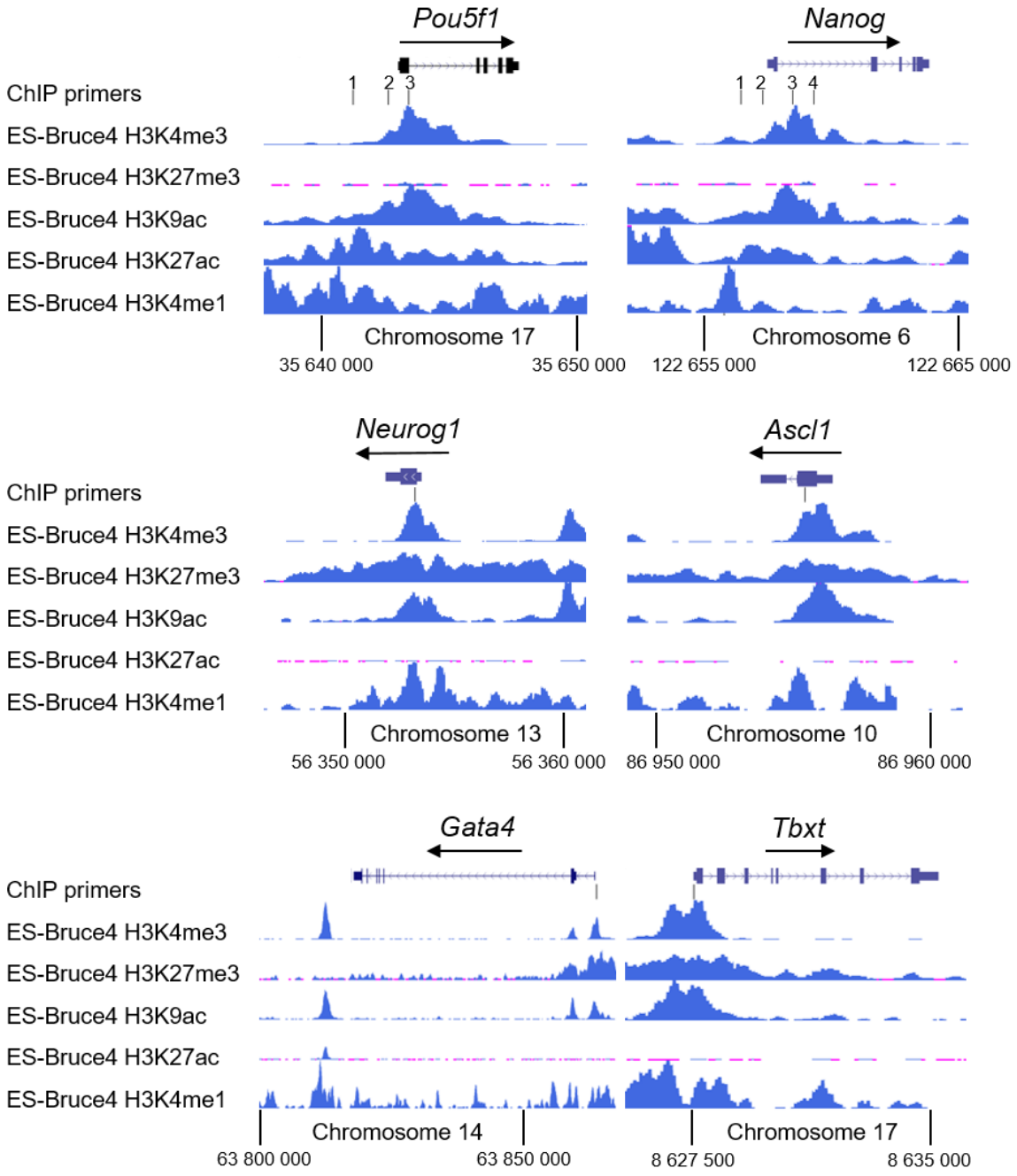


Figure 6.2 ChIP-PCR primer positioning along pluripotency and lineage-specific genes
ChIP-PCR primer positions are indicated along the pluripotency genes, *Pou5f1* and *Nanog*, and at the lineage-specific genes *Neurog1*, *Ascl1*, *Gata4*, and *Tbxt*. Sequences were aligned to the mm9 mouse genome. Data for mouse ES Bruce4 cells were obtained from the UCSC genome browser and was generated by the Ren lab as part of the ENCODE/LICR project (http://genome.ucsc.edu) (Consortium, 2012; Kent et al, 2002). Y-axis maxima are 30 for H3K4me3, 4 for H3K27me3, 15 for H3K9ac, 25 for H3K27ac, and 3 for H3K4me1.

H3 ChIP-PCR shows that the amount of chromatin loaded per reaction is comparable between the three different preparations from parental, B19ΔN2, and B38ΔN2 cells; however, it is not identical (Figure 6.3A). The average enrichment of H3 in the parental, B19ΔN2, and B38ΔN2 chromatin samples was 2.65, 1.72 and 2.1 respectively. These observations indicate that less amount of chromatin from *Hmgn2*-knockout lines was utilised, which is particularly evident at the actively transcribed genes *Actb*, *Pou5f1*, and *Nanog* (Figure 6.3A). In order to control for these differences, the average H3 enrichment across all primer sets tested was used to normalise the data from different chromatin samples in subsequent experiments.

The isotype control is a negative control that reveals the background levels of non-specific immunoprecipitation at the genomic locations tested. The data show that PCR signal from rIgG is less than 1% of that for H3 at most of the primer sets, although in some cases the signal approximates to 10% of the H3 signal (*i.e.* relative enrichment of less than 0.35, Figure 6.3B). In these reactions, however, dissociation curve analysis reveals that the signal does not derive from the correct amplicon (not shown). The previous data demonstrate that enrichment above these background levels corresponds to selective immunoprecipitation by the antibodies tested, and confirm primer specificity.

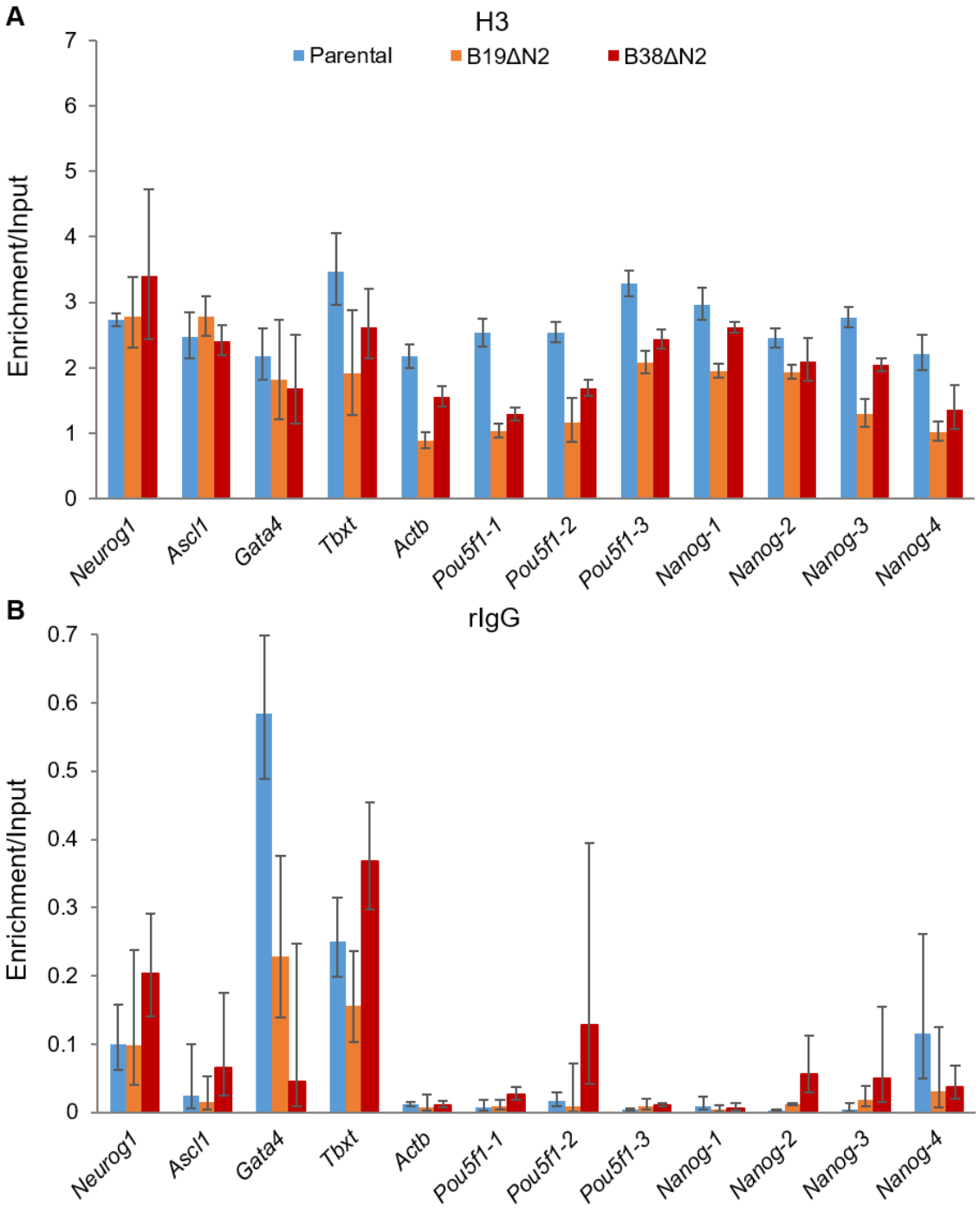


Figure 6.3 ChIP-PCR controls for chromatin loading and antibody/primer specificity
Chromatin samples were prepared from parental, B19ΔN2, and B38ΔN2 cells and immunoprecipitated with an antibody targeting (A) the C-terminal tail of H3 or (B) using the isotype control rIgG. The resulting DNA was purified and assayed by qPCR using primers corresponding to the promoter and/or regulatory regions of pluripotency and lineage-specific genes. The graphs represent the enrichment in the immunoprecipitated sample compared to that of input data. Error bars symbolise (A) the SEM from three independent experiments or (B) the SD from a PCR triplicate.

6.3 HMGN1 and HMGN2, chromatin architectural proteins

Recent ChIP-seq studies suggest an enrichment of HMGN1 and HMGN2 at DHSs, the hallmark of regulatory regions such as promoters and enhancers (Martínez de Paz & Ausió, 2016), in mammalian cells (Deng et al, 2013; Deng et al, 2015; Zhang et al, 2016). In addition, a more recent paper shows that HMGN1 and HMGN2 colocalise with histone modifications related to decondensed and active chromatin such as H3K4me3, H3K9ac, H3K27ac, and H3K4me1, while the proteins are nearly depleted from condensed and transcriptionally silent chromatin regions marked by H3K9me3 and H3K27me3 in three different cell types derived from mouse (He et al, 2018).

In this work, ChIP-PCR detected HMGN1 and HMGN2 bound at pluripotency and lineage-specific genes (Figure 6.4). Importantly, all the PCR primer sets encompass regions enriched in histone modifications characteristics to active chromatin (Figure 6.2), and therefore, the presence of the proteins was expected. In parental cells, HMGN1 and HMGN2 are bound at similar levels at all the *Nanog* and *Pou5f1* gene primer sets (Figure 6.4). Interestingly, HMGN binding is around a third higher at the actively transcribed genes (*Actb*, *Pou5f1*, and *Nanog*) than at the poised lineage-specific genes *Neurog1*, *Gata4* and *Ascl1* (Figure 6.4). Although the profile of HMGN1 distribution mimics that of HMGN2 (Figure 6.4A and B), the HMGN1 enrichment is approximately 70% lower. This suggests that HMGN2 binding is higher than HMGN1, or alternatively, that the anti-HMGN2 antibody is more efficient at immunoprecipitating chromatin.

Figure 6.4A shows that in the *Hmgn2*-knockout lines, B19ΔN2 and B38ΔN2, the levels of HMGN2 bound to chromatin are reduced to approximately one third of that observed in parental cells. This confirms that the HMGN2 binding detected in parental cells is specific. The residual binding in the *Hmgn2*-knockout cells is likely to be due to non-specific chromatin binding to the antibody or the beads, as western blotting shows that HMGN2 protein synthesis is completely abolished in B19ΔN2 and B38ΔN2 cells (Figure 3.1 and Figure 3.2).

Previous studies found that, although gene expression and protein levels of HMGN1 in *Hmgn2*-knockout cells and tissues remain the same as in control littermates, there is some functional compensation, and an anti-HMGN1 antibody

precipitates twice as much chromatin from *Hmgn2*-knockout cells than from control cells (Deng et al, 2015). Figure 6.4B shows that the HMGN1 binding to chromatin increases in B19ΔN2 and B38ΔN2 cells, specifically at the genes that are actively transcribed in pluripotent cells (*Actb*, *Pou5f1*, and *Nanog*) and at the pro-neural transcription factors that are also expressed in *Hmgn*-knockout cells. Hence, it seems that increased HMGN1 binding partially compensates for the loss of HMGN2, particularly at actively transcribed genes.

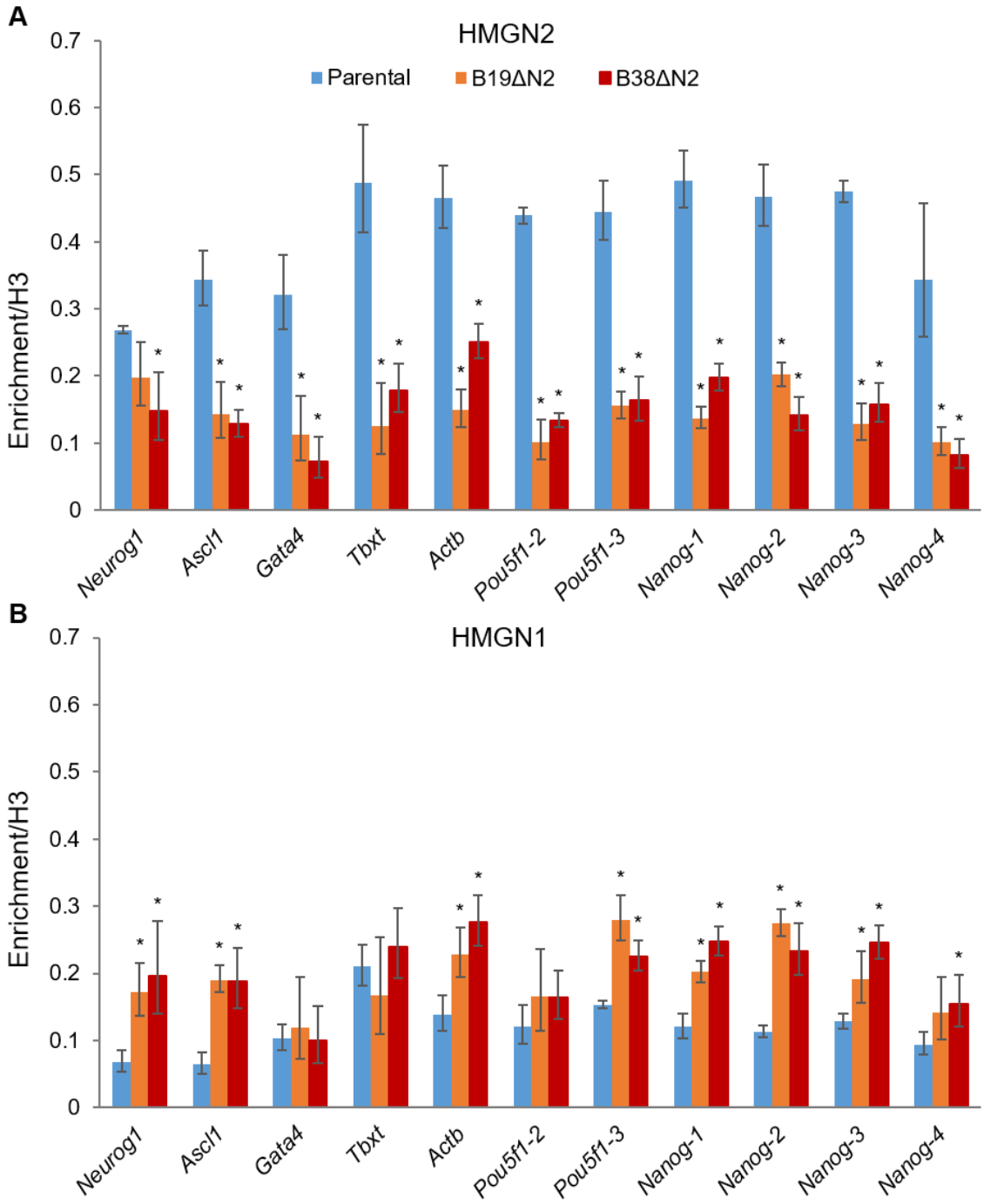


Figure 6.4 HMGN1 and HMGN2 binding to chromatin at pluripotency and lineage-specific genes in parental cells and after the loss of HMGN2
Chromatin samples were prepared from parental, B19ΔN2, and B38ΔN2 cells and immunoprecipitated with antibodies for (A) HMGN2 and (B) HMGN1. The resulting DNA was purified and tested for promoter and regulatory regions of pluripotency and lineage-specific genes by qPCR. The graphs represent the fold change in comparison with H3, error bars symbolise the SD from a PCR triplicate, and the statistical significance was calculated by two-tailed Student T test and Bonferroni correction for multiple comparisons, considering two comparisons per gene and immunoprecipitation (p value *<0.025).

6.4 H3K4me3 and H3K27me3, marks of transcriptional activation and repression, respectively, and bivalent domains

H3K4me3 is a modification associated with active transcription (Sims et al, 2003). Conversely, H3K27me3 is preferentially deposited at genes that require transcriptional silencing (Sims et al, 2003). In addition, both modifications are found in bivalent domains, which ensures transcriptional repression while poising the genes for transcriptional activation in later developmental stages (Azuara et al, 2006; Bernstein et al, 2006; Mikkelsen et al, 2007).

Most of the primer sets utilised in the present study target regions enriched in H3K4me3 in ESCs, indicating either active transcription or bivalent domains (Figure 6.2). Accordingly, H3K4me3 is enriched by up to 25 fold relative to H3 at the genomic regions tested (Figure 6.5A). In parental cells, the distribution of H3K4me3 along *Nanog* is in agreement with the classical pattern observed in the laboratory (West, unpublished) and in ESCs (Figure 1.2), where the mark is low at the potential upstream enhancer (*Nanog*-1), higher in the proximal promoter (*Nanog*-2), highest just after the first exon (*Nanog*-3), and lower towards the centre of the first intron (*Nanog*-4) (Figure 6.5A). Data from ESCs also support a peak of H3K4me3 at the end of the first exon of *Pou5f1* (*Pou5f1*-3, Figure 6.2). The enrichment of H3K4me3 is lower at *Nanog* than at *Pou5f1* and lineage-specific genes (Figure 6.5A), which is not related with transcription rates since *Nanog* transcript levels are considerable higher than those of lineage-specific genes in parental cells, although *Nanog* gene expression is indeed lower than that of *Pou5f1*.

Loss of HMGN2 influences the accumulation of H3K4me3, since some or most of the chromatin regions tested in B19ΔN2 and B38ΔN2 cells, respectively, exhibit reduced levels of this modification compared to parental cells (Figure 6.5A). The lower H3K4me3 signal is not a simple consequence of reduced gene expression, as *Pou5f1* transcript levels are similar in *Hmgn2*-knockout and parental cells (Figure 3.10A). However, there is some correlation between gene expression and H3K4me3 at the other genes tested. For example, in B38ΔN2 cells, the levels of *Nanog* mRNA and H3K4me3 along *Nanog* gene are both reduced, whereas in B19ΔN2 cells, *Nanog* mRNA levels are unchanged and the reduction in H3K4me3

is restricted to the region targeted by *Nanog*-3 primer set (Figure 3.10A and Figure 6.5A). Similarly, *Neurog1* and *Ascl1* are highly expressed in B19ΔN2 cells (Figure 3.10B) and higher accumulation of H3K4me3 is detected (Figure 6.5A), while *Gata4* and especially *Tbxt* transcripts are found at lower levels (Figure 3.10C), correlating with lower H3K4me3 (Figure 6.5A).

H3K27me3 levels are low in undifferentiated ESCs and increase during differentiation (Marks et al, 2012). Consistent with the repressive nature of this modification, its accumulation at the active pluripotency genes in parental P19 cells is negligible (Figure 6.5B). In fact, dissociation curve analysis fails to identify the desired amplicon when using *Pou5f1* and *Nanog* primer sets, except in the case of *Nanog*-4. In contrast, H3K27me3 is detected at lineage-specific genes (Figure 6.5B).

As mentioned above, H3K27me3 is part of the mechanism to silence lineage-specific genes in ESCs (Boyer et al, 2006). The pro-neural genes are highly expressed in B19ΔN2 and B38ΔN2 cells (Figure 3.10B), although changes in the pattern of H3K27me3 do not seem to be the reason, since it is not affected by the loss of HMGN2 (Figure 6.5B). Previous studies on ESCs lacking both HMGN1 and HMGN2 report a substantial accumulation of H3K27me3 at lineage-specific genes as a result of increased nucleosome binding by H1 (Deng et al, 2017). In the *Hmgn2*-knockout P19 cells, the compensatory activity of HMGN1 (Figure 6.4) may be sufficient to compete with H1, providing a possible explanation for this contrasting result.

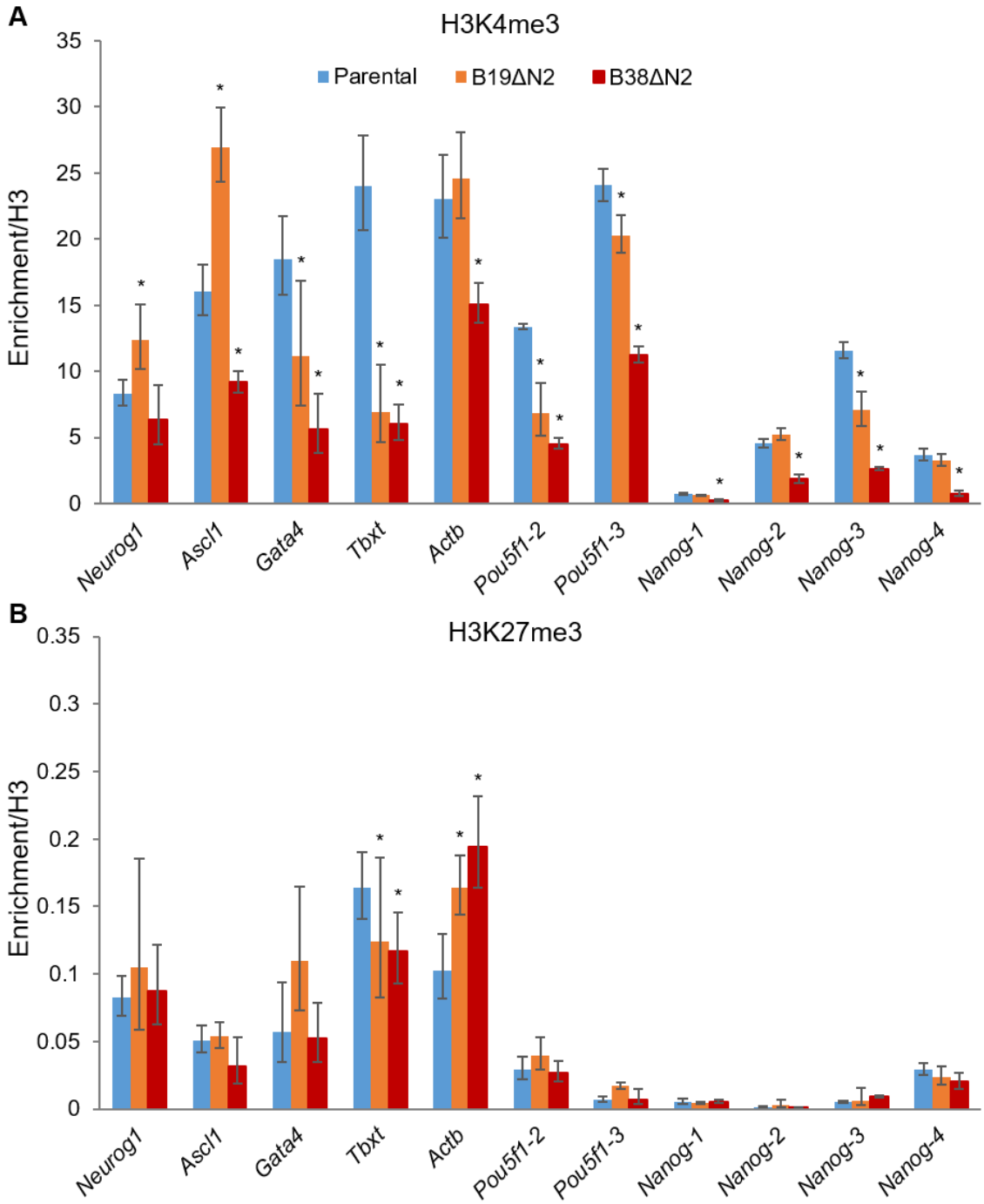


Figure 6.5 H3K4me3 and H3K27me3 accumulation at pluripotency and lineage-specific genes of parental cells and after the loss of HMGN2
Chromatin samples were prepared from parental, B19ΔN2, and B38ΔN2 cells and immunoprecipitated with antibodies specifically targeting (A) H3K4me3 and (B) H3K27me3. The resulting DNA was purified and tested for promoter and regulatory regions of pluripotency and lineage-specific genes by qPCR. The graphs represent the fold change in comparison with H3, error bars symbolise the SD from a PCR triplicate, and the statistical significance was calculated by two-tailed Student T test and Bonferroni correction for multiple comparisons, considering two comparisons per gene and histone modification (p value *<0.025).

6.5 Histone acetylation, characteristic of open and active chromatin

HMGN1 and HMGN2 proteins enhance histone acetylation (Ueda et al, 2006) and loss of HMGN1 is associated with a marked reduction of this modification in cells and tissues (Lim et al, 2005). ESC self-renewal is compromised when HAT activity is abolished (Fang et al, 2014), while self-renewal is supported by inhibition of histone deacetylation (Qiao et al, 2015; Ware et al, 2009). To evaluate whether histone acetylation is reduced after the loss of HMGN2, which may provide a partial explanation for the changes in self-renewal of *Hmgn2*-knockout cells, ChIP-PCR was performed for the detection of H3K9ac, H3K27ac, and H3K122ac.

Acetylation of H3K9 is prevalent in ESCs and drops dramatically upon differentiation (Efroni et al, 2008; Krejčí et al, 2009; Lee et al, 2003; Qiao et al, 2015). Figure 6.6A shows that H3K9ac is highly enriched at the actively transcribed genes *Actb* and *Pou5f1* of parental P19 cells, with the highest enrichment at the primer set downstream of the *Pou5f1* TSS (*Pou5f1*-3). Indeed, previous work reports that accumulation of H3K9ac tends to be particularly high at around 1 kb downstream of gene TSSs (Qiao et al, 2015). Likewise, the region targeted by *Nanog*-3 primer set shows the highest H3K9ac accumulation at *Nanog* gene locus (Figure 6.6A), even though H3K9ac levels are lower at this gene, in line with the H3K4me3 data (Figure 6.5A).

Importantly, B19ΔN2 and B38ΔN2 cells display lower H3K9ac levels in comparison with parental cells (Figure 6.6A). The reduction varies between 50 and 70% at most of the regions evaluated in this study, suggesting a global effect following the loss of HMGN2 (Figure 6.6A). In line with the H3K4me3 data (Figure 6.5A), this excludes the pro-neural genes (Figure 6.6A) that are highly expressed in B19ΔN2 and B38ΔN2 cells (Figure 3.10B). However, at all other genomic regions investigated, the lower H3K9ac accumulation does not seem to correlate with changes in gene expression, as most of the genes are expressed at similar levels among the three cell lines (Figure 3.10). These observations suggest that loss of HMGN2 interferes with H3K9ac deposition or maintenance, which in turn might destabilise the pluripotency transcriptional network. In fact, a growing body of evidence suggest that the master regulators of the pluripotency transcriptional network recruit HATs as coactivators of their target genes (Fang et al, 2014;

Göke et al, 2011; Zhong & Jin, 2009). Importantly, this mechanism is also part of the positive feedback loop that potentiates the transcription rates of the master regulators (Zhong & Jin, 2009).

POU5F1, SOX2, and NANOG strongly colocalise with the HATs CREBBP/EP300 at enhancers (Göke et al, 2011). Active enhancers are usually marked by the colocalisation of H3K4me1 and H3K27ac, and recently, of H3K122ac (Pradeepa et al, 2016). In agreement, the *Pou5f1*-1 primer set reports the highest accumulation of both H3K27ac (Figure 6.6B) and H3K122ac in parental P19 cells (Figure 6.6C). Interestingly, similar enrichments are detected by *Pou5f1*-2 primer set (Figure 6.6B and Figure 6.6C) and may be explained by interactions between enhancer and promoter that propagate the modifications (Fang et al, 2014). In contrast with H3K27ac, H3K122ac is also detected in promoter and genic regions (Tropberger et al, 2013), which may explain the presence of H3K122ac at lineage-specific genes poised for transcriptional activation, while H3K27ac is barely detected (Figure 6.6B and Figure 6.6C).

H3K27ac and H3K122ac are substantially reduced in B19ΔN2 and B38ΔN2 cells at most of the evaluated genomic regions (Figure 6.6 and Figure 6.6C). *Nanog* shows the lowest accumulation of these modifications in the *Hmgn2*-knockout cells, which correlates with the reduced transcript levels in B38ΔN2 cells (Figure 3.10A). However, the reduced acetylation at the *Pou5f1* and *Gata4* genes does not correlate with gene expression (Figure 3.10). Importantly, *Tbxt* displays less histone acetylation and H3K4me3 in the *Hmgn2*-knockout cells, while retaining H3K27me3 (Figure 6.5 and Figure 6.6). This may result in a more repressive chromatin state that is consistent with the reduced *Tbxt* transcript levels (Figure 3.10C). Conversely, *Neurog1* and *Ascl1* retain normal levels of modifications associated with active transcription (Figure 6.5 and Figure 6.6), and therefore, the chromatin structure at these genes may be permissive for transcription in B19ΔN2 and B38ΔN2 cells (Figure 3.10B).

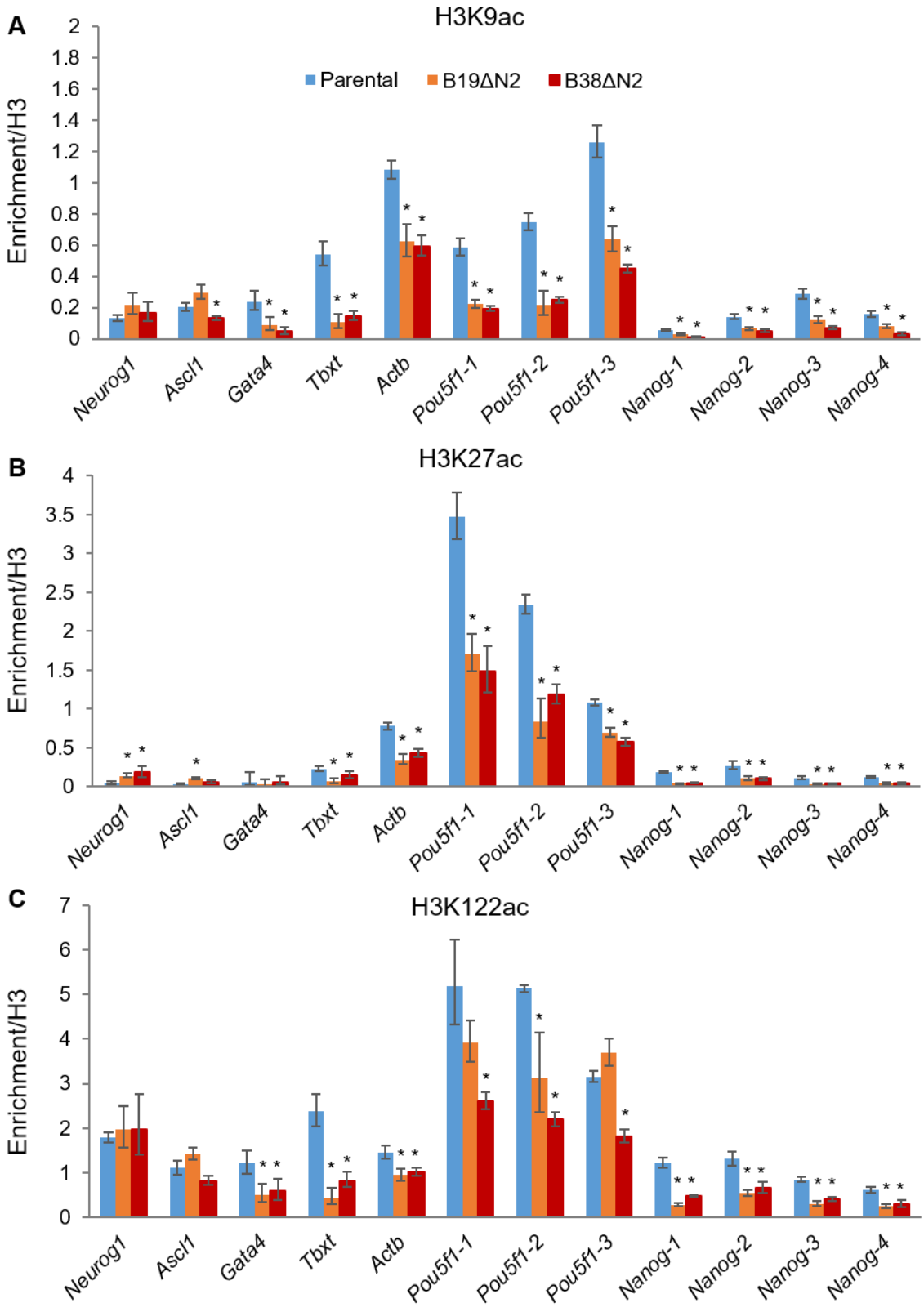


Figure 6.6 H3 acetylation in pluripotency and lineage-specific genes of parental cells and after the loss of HMGN2

Chromatin samples were prepared from parental, B19ΔN2, and B38ΔN2 cells and immunoprecipitated with antibodies specifically targeting (A) H3K9ac, (B) H3K27ac, and (C) H3K122ac. The resulting DNA was purified and tested for promoter and regulatory regions of pluripotency and lineage-specific genes by qPCR. The graphs represent the fold change in comparison with H3, error bars symbolise the SD from a PCR triplicate, and the statistical significance was calculated by two-tailed Student T test and Bonferroni correction for multiple comparisons, considering two comparisons per gene and histone modification (p value *<0.025).

Taken together, the ChIP-PCR experiments suggest that HMGN2 promotes open chromatin conformations in P19 cells at all genomic regions tested, as its loss results in lower accumulation of modifications related with active transcription, such as histone acetylation and H3K4me3. The pattern of histone modifications does not always correlate with gene expression, suggesting that there are other factors regulating transcription. Furthermore, histone acetylation and methylation does not exclusively rely on the presence of HMGN2, since the genes encoding pro-neural transcription factors display normal levels of these modifications.

6.6 Loss of a major HMGN variant leads to global reduction of H3 acetylation

The maintenance and potentiation of a global open chromatin conformation seems to be crucial for pluripotency (Ang et al, 2011; Qiao et al, 2015; Ware et al, 2009). Therefore, in order to extend the observations made at individual genomic locations, the global levels of several histone modifications were investigated. Acid histone extracts were prepared from control and *Hmgn*-knockout cells and immunoblotting for the detection of H3K4me3, H3K9ac, and H3K27ac was performed.

Figure 6.7 shows that the total levels of H3K4me3 are not affected after the loss of HMGN1, since they are comparable in N1-1ΔN1, D3ΔN1 and B1 control cells. In addition, B8ΔN2 cells display similar total levels of this modification (Figure 6.7). However, a large reduction of 75% in total H3K4me3 is observed in B19ΔN2 cells (Figure 6.7), which is consistent with the lower accumulation at pluripotency and lineage-specific genes detected by ChIP-PCR (Figure 6.5).

In the case of H3K9ac, the levels exhibited by B1, N1-1ΔN1, and D3ΔN1 cells correspond to approximately half of the levels in parental cells (Figure 6.7). Further reduction is observed after the loss of HMGN2, since total H3K9ac in B8ΔN2 and B19ΔN2 cells represent around 35% and 25% of parental, respectively (Figure 6.7). This is in complete agreement with previous reports showing that the effect of HMGN2 on histone acetylation is greater than that of HMGN1 (Ueda et al, 2006). However, the fact that B1 cells also contain reduced total levels of H3K9ac (Figure 6.7) suggest that the lack of a major HMGN variant is not the

only factor regulating this modification. Chapter 3 shows that B1 cells display lower *Nanog* and higher pro-neural transcription factor mRNA levels without resulting in spontaneous differentiation. Considering the previous data, it is possible to speculate that B1 cells represent a stable pluripotency state that is positioned one step further towards differentiation in comparison with parental cells. Indeed, loss of H3K9ac is an early event in ESC differentiation (Krejčí et al, 2009; Qiao et al, 2015). Perhaps there is a lower threshold in the level of histone acetylation that correlates with the initiation of differentiation programs, and the loss of a major HMGN variant may contribute to reaching that threshold (see below).

H3K27ac preferentially colocalises with active enhancers. Total H3K27ac is substantially reduced in *Hmgn*-knockout cells (Figure 6.7). Although B1 cells also display lower levels of this modification than parental cells, the reduction is greater after the loss of a major HMGN variant (Figure 6.7). Previous ChIP-seq studies show that HMGN1 and HMGN2 are enriched at DHSs particularly marking enhancers (Deng, 2015). This may explain the greater changes in total H3K27ac than in H3K9ac observed in *Hmgn*-knockout cells (Figure 6.7).

As previously mentioned, loss of histone acetylation correlates with the initiation of differentiation programs in ESCs (Krejčí et al, 2009; Qiao et al, 2015). Another major epigenetic event during ESC differentiation is the clipping of histone H3, which may represent an efficient mechanism for removing histone modifications (Duncan et al, 2008; Vossaert et al, 2014). In mouse, proteolysis of the N-terminal tail of H3, mainly at the residue 21, is developmentally regulated (Duncan et al, 2008). Accordingly, the abundance of this H3 cleavage product (H3.cs1) is slightly higher in *Hmgn*-knockout cells than in parental and B1 (Figure 6.7). Importantly, H3 clipping does not explain the global reduction in histone acetylation, as the acetylation of the globular domain of H3 at K122 is conserved in all cleavage products and is also reduced after the loss of a major HMGN variant (Figure 6.6). Furthermore, H3K27 lies after the main cleavage site at position 21 (Duncan et al, 2008). Finally, the small increase in H3 clipping does not correlate with the large reductions in global H3 acetylation (Figure 6.7).

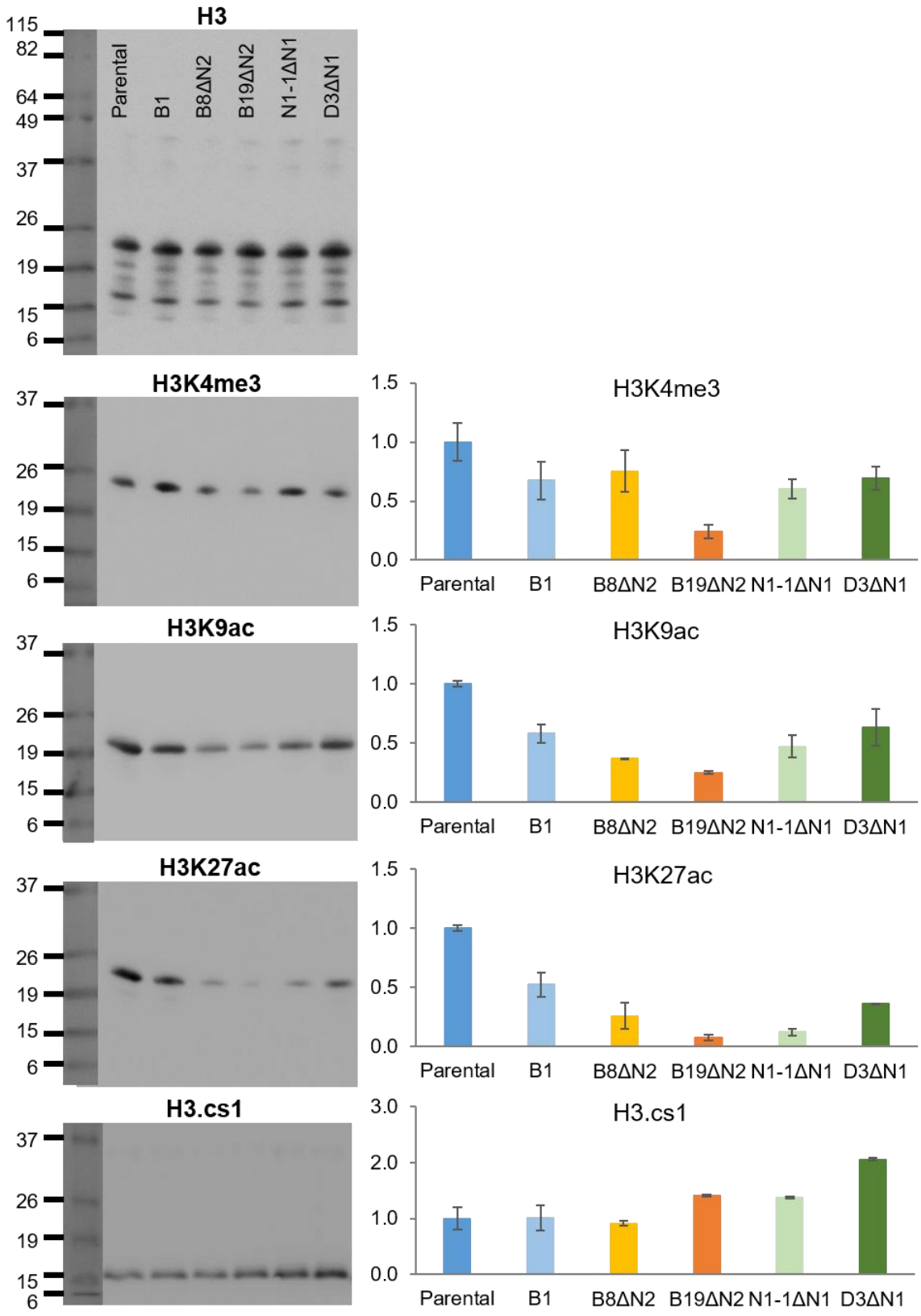


Figure 6.7 Global reduction of histone acetylation in *Hmgn*-knockout cells
WB for the detection of histone modifications related to active transcription, H3K4me3, H3K9ac, H3K27ac, and the proteolytic product after clipping at residue 21 of H3 (H3.cs1). Acid histone extracts were prepared in duplicate from two different cell passages of parental, B1, and *Hmgn*-knockout cells and separated by SDS-PAGE. WB for each histone modification and sample set was performed twice or three times. Representative images are displayed (left). The graphs represent the average result from the densitometry relative to parental cells and error bars symbolises the SEM of two independent samples.

6.7 Discussion

The present chapter investigated the epigenetic landscape in parental and *Hmgn2*-knockout cells at pluripotency genes, silenced lineage-specific genes, and lineage-specific genes that are activated after the loss of a major HMGN variant. This includes the accumulation of histone acetylation, H3K4me3, H3K27me3, and the binding of HMGN1 and HMGN2. In addition, it evaluated global levels of histone modifications associated with active transcription in ECCs lacking HMGN2.

Pou5f1 and *Nanog* are actively transcribed in pluripotent cells, and accordingly, they are devoid of the repressive mark H3K27me3, and enriched in H3K4me3 and histone acetylation (Figure 6.5 and Figure 6.6). In addition, the binding of HMGN1 and HMGN2 is slightly higher at these genes than at the poised lineage-specific genes in parental cells (Figure 6.4). In agreement, previous ChIP-seq data of NSCs show higher enrichment of HMGN1 at the top 40 expressed genes in comparison with the bottom 40 (Deng et al, 2013).

Loss of HMGN2 leads to increased HMGN1 binding along *Pou5f1* and *Nanog* (Figure 6.4), consistent with functional compensation between the two major HMGN variants (Deng et al, 2015). Nevertheless, these proteins are not fully redundant, and changes in the epigenetic landscape of pluripotency genes are observed in *Hmgn2*-knockout cells. In particular, histone modifications related to transcriptional activation, such as H3K4me3, H3K9ac, H3K27ac, and H3K122ac are reduced (Figure 6.5 and Figure 6.6). The lower accumulation of these modifications does not influence *Pou5f1* expression but might relate to *Nanog* downregulation in B38ΔN2 cells (Figure 3.10A). Unlike *Pou5f1*, *Nanog* is transcribed in bursts (Abranches et al, 2014; Kalmar et al, 2009), which may increase the dependence of its transcription on histone modifications that promote an open local chromatin conformation. In fact, *Nanog* is downregulated when the HAT/co-activator protein CREBBP/EP300 is knocked down, while *Pou5f1* remains unchanged (Qiao et al, 2015). Furthermore, EP300 regulates *Nanog* expression, but not that of *Pou5f1* (Zhong & Jin, 2009). These observations lead to the proposition that HMGNs support an open chromatin conformation that enhances transcription at pluripotency genes, particularly at genes that rely on their histone acetylation status.

The endodermal and mesodermal transcription factors, *Gata4* and *Tbxt*, are expressed at low levels in parental cells, which correlates with the bivalent modifications found at both promoters, *i.e.* H3K4me3 and H3K9ac for activation, and H3K27me3 for repression (Figure 6.2 and Figure 6.5). As studied in chapter 3, *Gata4* transcript and protein levels remain unchanged in B19ΔN2 and B38ΔN2 cells, while *Tbxt* expression is lower than in parental cells (Figure 3.9 and Figure 3.10C). Importantly, the accumulation of H3K27me3 at promoters of both genes remains comparable before and after the loss of HMGN2, while H3K4me3, H3K9ac, and H3K122ac are less abundant (Figure 6.5 and Figure 6.6). The prevalence of H3K27me3 may result in a repressive chromatin environment, explaining the reduced *Tbxt* transcript levels. This resembles what was previously described for pluripotency genes, suggesting that HMGN2 enhances modifications related to active transcription without influencing H3K27me3, resulting in an open chromatin conformation at several genomic regions. These observations contrast with recent work reporting higher enrichment of H3K27me3 at lineage-specific genes in *Hmgn1* and *Hmgn2* double-knockout mice; however, the presence of HMGN1 may be sufficient to antagonise the deposition of H3K27me3 (Figure 6.4).

A different scenario is observed at the genes encoding pro-neural transcription factors. Similar to endodermal and mesodermal transcription factors, *Neurog1* and *Ascl1* transcripts are scarce in parental cells and their gene promoters are bivalently marked (Figure 6.2 and Figure 6.5). However, loss of HMGN2 does not modify this bivalent signature, as enrichment of H3K27me3, H3K4me3, H3K9ac, and H3K122ac is maintained in B19ΔN2 and B38ΔN2 cells at parental levels (Figure 6.5 and Figure 6.6). Interestingly, HMGN1 binding increases at *Neurog1* and *Ascl1* genes in B19ΔN2 and B38ΔN2 cells, similar to what was observed at the actively transcribed genes (Figure 6.4). Perhaps HMGN1 and other epigenetic players maintain an open chromatin conformation at pro-neural genes in the absence of HMGN2. Since the bivalent domains silence genes, while poising them for activation (Azuara et al, 2006; Bernstein et al, 2006), other mechanisms should induce *Neurog1* and *Ascl1* expression in B19ΔN2 and B38ΔN2 (Figure 3.10B). As these mechanisms operate in the absence of HMGN2, this protein may indirectly regulate lineage-specific gene expression in P19 cells (see below).

Although the pro-neural genes escape the rule, the loss of HMGN2 leads to lower accumulation of H3K4me3 and histone acetylation at several genomic regions in P19 cells (Figure 6.5 and Figure 6.6). WB of acid histone extracts from parental, B1, and *Hmgn*-knockout cells confirms a global reduction of H3K9ac and H3K27ac in cells lacking a major HMGN variant (Figure 6.7). Lower total levels of these modifications are also observed in B1 cells without perturbing self-renewal in these cells. However, the changes in *Hmgn*-knockout lines are greater (Figure 6.7), perhaps reaching a threshold in which pluripotency is no longer stable. It is known that the pluripotency transcription factors POU5F1, SOX2, and NANOG promote gene expression by recruiting CREBBP/EP300 to target genes in ESCs, especially at enhancers (Fang et al, 2014). It is possible that in P19 cells, HMGNs optimise HAT activity, as previously reported for other systems (Lim et al, 2005; Ueda et al, 2006). The loss of HMGN proteins would then lead to lower H3K27ac at enhancers and lower H3K9ac at promoters and gene bodies of POU5F1, SOX2, and NANOG targets (Figure 6.6 and Figure 6.7), which consequently destabilises pluripotency gene expression. In this regard, the inhibition of HDACs has been shown to increase both H3K9ac and pluripotency gene expression (Qiao et al, 2015). In addition, there is growing body of evidence suggesting that global histone acetylation and open chromatin conformation support ESC self-renewal (Mattout & Meshorer, 2010; Niwa, 2007).

In conclusion, HMGN proteins promote an open chromatin conformation by influencing the accumulation of histone modifications related to active transcription, such as H3K4me3, H3k9ac, H3K27ac, and H3K122ac. Loss of a major HMGN variant leads to changes at local and global levels of the mentioned modifications, and therefore, of chromatin structure. At local levels, those changes directly influence gene expression, while at a global level they seem to perturb the pluripotency transcriptional network. Importantly, intrinsic cellular programs compensate or potentiate the effects triggered by the loss of HMGN1 or HMGN2, which ultimately results in the perturbation of the stem cell identity.

Chapter 7 Conclusions and future work

7.1 Summary and conclusions

The present work studied the phenotypical alterations of P19 cells after the loss of HMGN1 or HMGN2. P19 cells represent a widely used model of embryonal pluripotent cells that express pluripotency markers, consistent with their embryonal identity. These cells indefinitely self-renew and do not differentiate unless stimulated, and accordingly, they present a mostly uniform morphology (Jones-Villeneuve et al, 1982; McBurney, 1993; McBurney & Rogers, 1982). In contrast, P19 cells lacking HMGN1 or HMGN2 are morphologically heterogeneous and grow in a more dispersed manner than control cells, resembling the differentiated cells in ESC cultures (Trott & Martinez Arias, 2013; Wray et al, 2010). Indeed, a proportion of *Hmgn*-knockout cells spontaneously differentiate into mainly, but not exclusively, neural lineages. Furthermore, *Hmgn*-knockout cultures display lower levels of pluripotency markers and higher lineage-specific gene expression, particularly of pro-neural transcription factors, which suggest a higher propensity towards differentiation. These observations support a relevant role for HMGNs in maintaining ECC identity.

Inhibition of the main signalling pathways involved in neuronal differentiation, such as WNT, FGF, and NOTCH (Martynoga et al, 2012), does not re-establish pluripotency and pro-neural gene expression of *Hmgn*-knockout cells, suggesting that the cells that initiate neuronal differentiation follow an intrinsic and cell-autonomous program that is consistent with the default model of neural induction (Hemmati-Brivanlou, 1997; Munoz-Sanjuan & Brivanlou, 2002; Wilson & Edlund, 2001).

Loss of a major HMGN variant does not have a major influence on neuronal differentiation of P19 cells after active stimulation of neurogenesis (Nakayama et al, 2014). However, the higher propensity of unstimulated cells to differentiate towards neural fates following serum withdrawal increases the induction of pro-neural transcription factors, which could result in increased neuronal differentiation (Busskamp et al, 2014; Farah et al, 2000; Reyes et al, 2008; Thoma et al, 2012; Velkey & O'Shea, 2013). Furthermore, NSCs derived from *Hmgn*-knockout cells display higher pro-neural gene expression and

spontaneous neuronal differentiation, suggesting that HMGNs are important for proper cell fate decisions of all the stem cells, including ECCs and NSCs, with their presence favouring the preservation of the stem cell identity.

A growing body of evidence supports an essential role for chromatin structure and modifiers during developmental decisions and transitions, that is beyond simply stabilising the expression profile driven by transcription factors (Chen & Dent, 2013). Indeed, the precise control of gene expression in cell fate determination requires collaborative interactions, and the master regulators of pluripotency recruit chromatin modifiers to achieve both target gene expression and feedback loops (Ang et al, 2011; Chen et al, 2008; Fang et al, 2014; Göke et al, 2011; Zhong & Jin, 2009). Furthermore, chromatin structure has been shown to act as a buffer where the chromatin modifiers fine-tune gene expression (Wang et al, 2009). This function becomes particularly relevant in pluripotent cells that have inherent transcriptional noise (Efroni et al, 2008). The chromatin of pluripotent cells has a remarkable plasticity that becomes more limited in differentiated cells (Deng et al, 2013; Efroni et al, 2008; Guenther et al, 2007; Krejčí et al, 2009; Meshorer et al, 2006). Chromatin plasticity and developmental potential have been directly correlated (Kishi et al, 2012), and seem to be supported by mechanisms that maintain an open chromatin state (Mattout & Meshorer, 2010; Niwa, 2007).

In this sense, HMGNs have been shown to fine-tune gene expression (Deng et al, 2013; Deng et al, 2015; Kugler et al, 2013; Zhang et al, 2016), to dampen the magnitude of the transcriptional response (Zhang et al, 2016), and to promote an open chromatin conformation at particular genomic regions (Cuddapah et al, 2011; Deng et al, 2017; Deng et al, 2013; Deng et al, 2015; Zhang et al, 2016). However, little is known about the mechanisms underlying these observations.

In the present work, HMGN proteins influence the accumulation of histone modifications related to active transcription, such as H3K4me3, H3K9ac, H3K27ac, and H3K122ac. Loss of a major HMGN variant disrupts the active histone modification landscape, and therefore, chromatin structure at local and global levels. Locally, those changes directly influence gene expression of pluripotency and lineage-specific transcription factors, while globally they may restrict chromatin plasticity. Importantly, intrinsic cellular factors compensate

or potentiate the effects triggered by the loss of HMGN1 or HMGN2, thus the HMGNs are not the only proteins modulating chromatin structure. The present data support a role for HMGNs in contributing to chromatin plasticity of stem cells by promoting an active histone modification landscape and an open chromatin conformation that are essential for preserving the self-renewal and developmental potential of stem cells.

The hypothesis of this project was that the two major members of the HMGN family, HMGN1 and HMGN2, play an important role in the establishment and/or maintenance of the active chromatin conformation and histone modification landscape of pluripotent cells, which in turn are essential for the preservation of the self-renewal and the differentiation potential of these cells. The results confirm the hypothesis since loss of HMGN1 or HMGN2 in ECCs, and additionally in NSCs, leads to increased spontaneous differentiation, mainly towards neural fates, to higher lineage-specific gene expression, and to reduced accumulation of histone modifications associated with active and decondensed chromatin, which ultimately implies that, at least, a high proportion of cells loses their stem cell identity.

7.2 Future work

A major issue during the development of the present work was the clonal variation. The derivation of *Hmgn1*- and *Hmgn2*-knockout lines, previously performed in the laboratory, included the selection and expansion of monoclonal lines lacking either HMGN1 (Eden, West, unpublished) or HMGN2 (Sindi, 2017). Two different methods were used after the transfection of the P19 population with the CRISPR/Cas system; selection with 6-thioguanine facilitated the expansion of exclusively *Hprt*- and *Hmgn1*-knockout single-cell colonies (Eden, West, unpublished), while plating at limiting dilution allowed the expansion of single-cell colonies that subsequently were screened by IF and WB, selecting those lacking HMGN2 (Sindi, 2017). As a result, seven monoclonal lines including two *Hmgn1*-knockout, three *Hmgn2*-knockout, and two controls, constituted the object of study of this work, in addition to the parental cells.

Inconsistencies in the prevalence and abundance of pluripotency and lineage-specific markers are observed among B8ΔN2, B19ΔN2, and B38ΔN2 cells, and

strikingly, between N1-1ΔN1 and D3ΔN1 cells (see chapter 3). Furthermore, the accumulation of H3K4me3 after the loss of HMGN2 differs in B19ΔN2 and B38ΔN2 cells (Figure 6.5). These inconsistencies are attributed to clonal variations for the reasons presented below. First, the pluripotency of P19 cells is not homogeneous, which is evidenced by the cytoplasmic protuberances of the cells at the edges of the colonies (Figure 3.3, arrowheads) that are absent in the cells at the centre. In addition, P19 cells contain different levels of the naïve pluripotency transcription factor NANOG, and therefore, the reciprocal rates of transcription of lineage-specific genes within a population (Kalmar et al, 2009). Second, the cloning process required to derivate monoclonal lines involves phenotypical adaptations that allow the colony expansion from individual cells, regardless of the presence of other cells and of cell-to-cell communication. These adaptations are conserved and propagated in the monoclonal line. Then, it is possible to speculate that D3ΔN1 cells originated from a pluripotent cell with a transcriptional profile biased to endodermal lineages, while N1-1ΔN1 cells were poised to neuronal differentiation before CRISPR editing. In such a scenario, the loss of HMGNs disturbs pluripotency and the cells differentiate into the lineages dictated by their pre-existing transcription profile. Alternatively, the cloning process could apply a selection force to an *Hmgn2*-knockout cell that resulted in B19ΔN2 cells with *Nanog* transcript levels similar to parental cells, rather than to B8ΔN2 and B38ΔN2 cells (Figure 3.10A).

To overcome these inconsistencies that introduce noise to the data and challenge its credibility, it is worth to consider other systems for the derivation of *Hmgn1*- and *Hmgn2*-knockout cells. Huge improvements have been achieved from the start date of the present project in the application of the CRISPR/Cas system for genome editing and regulation, developing strategies that do not require monoclonal derivation and screening. Cas9-induced homology directed repair has been successfully used to tag endogenous genes with GFP (Leonetti et al, 2016) among other tags, raising the possibility of using GFP as a donor DNA that replaces the target gene. In this sense, *Hmgn1*- and *Hmgn2*-knockout cells could be sorted by FACS after transfection and immediately propagated and characterised as polypopulations, overcoming the clonal variability. Moreover, *Hmgn1* and *Hmgn2* double-knockout cells could be generated and selected with ease by using two different fluorescent reporters, which would extend the

knowledge generated in this work by circumventing the functional redundancy between the two major HMGN variants (Deng et al, 2015).

Another strategy would be the application of the CRISPR interference (CRISPRi) system, which utilises a deactivated version of the Cas9 lacking endonuclease activity that is targeted by a gRNA to the gene of interest to repress its expression either allosterically or by recruiting transcriptional repressors (Dominguez et al, 2015). Therefore, P19 cells could be transfected with the deactivated Cas9 and gRNAs targeting the gene promoters of *Hmgn1* and *Hmgn2*, resulting in their transcriptional silencing. The phenotypical alterations could be directly evaluated in the polypopulation, which circumvents the clonal variability. Interestingly, numerous gRNAs targeting the promoters of the genes coding for all HMGN variants would generate the first cells depleted of all HMGNs. Furthermore, CRISPRi represents a reversible system, since the deactivated Cas9 allow the manipulation of gene expression without altering the DNA sequence (Dominguez et al, 2015), which is suitable for the analysis of the role of HMGNs during differentiation processes. In this regard, NSCs can be derived from P19 cells and their identity and homogeneity can be ensured before the application of CRISPRi targeting HMGNs. The previous strategy would guarantee that the phenotypical alterations observed in the targeted NSCs are exclusively caused by the loss of HMGNs, rather than by previous phenotypes in the pluripotent state or by problems in the neural induction.

Beyond the clonal variability, the five *Hmgn*-knockout monoclonal lines present higher transcriptional rates of lineage-specific genes, especially of pro-neural transcription factors, in comparison with control lines (Figure 3.10). In addition, numerous cells have lost pluripotency markers and are positive for the neuronal marker TUBB3 (Figure 3.8). The previous data support a relevant role of HMGN1 and HMGN2 in the maintenance of the ECC identity, at least in a proportion of ECCs. Nevertheless, there is not enough evidence to conclude that the loss of HMGNs increases the propensity of all ECCs to differentiate, since the present work evaluated the expression of limited genes and few differentiation markers.

Broader methods such as RNA-seq would elucidate whether the loss of HMGNs in P19 cells switches their pluripotency state towards another closer to differentiation. CRISPRi can target *Hmgn1*, *Hmgn2*, and both genes in different

P19 cell cultures. A pilot experiment would inform about the timing required to deplete HMGN proteins from P19 cells and to observe the first differentiated cells. Collection of RNA samples would be performed just before the appearance of the first differentiated cells that can alter the expression profile of the remaining ECC population. RNA-seq data from these cultures can be used in GO enrichment analysis to reveal whether *Hmgn*-knockout cells moved from the parental state of pluripotency to a poised pluripotency biased to certain cellular lineages. Single-cell RNA-seq carried would complement the previous experiments by showing the frequencies in which a cell prefers one lineage or another, perhaps identifying functional differences between HMGN1 and HMGN2. Considering the observations of the present work, it would be expected that most of the *Hmgn*-knockout cells are found in a pluripotency state poised to neuronal differentiation. However, single-cell RNA-seq results may explain the phenotypical differences of D3ΔN1 cells in comparison with the other *Hmgn*-knockout lines studied in this work.

Developmental specification is accompanied by epigenetic restriction and chromatin condensation (Kishi et al, 2012). ChIP-PCR carried out in B19ΔN2, B38ΔN2, and parental cells shows that, there is a reduction in H3 acetylation and H3K4me3 accumulation at different putative regulatory regions of pluripotency and lineage-specific genes after the loss of HMGNs (Figure 6.5 and Figure 6.6). Chromatin can be collected from P19 cultures in which *Hmgn1*, *Hmgn2*, and both genes are repressed by CRISPRi at the same time point of the collection of RNA for RNA-seq experiments. The working hypothesis is that HMGNs are relevant for the ECC identity, and therefore, their loss switches the pluripotency state of P19 cells to a poised pluripotency biased to neural lineages. In addition to expression profiles that correspond to a transition state between pluripotency and early neuronal differentiation potentially identified by RNA-seq, ChIP-seq data would reveal chromatin condensation in *Hmgn*-knockout cells in comparison with parental. It is expected a reduced accumulation of H3K4me3, H3K9ac, H3K27ac, and H3K4me1 in several regulatory regions, especially in those of pluripotency genes, which may destabilise the pluripotency network lowering the threshold in which *Hmgn*-knockout cells respond to differentiation. Conversely, pro-neural genes, that are expected to orchestrate neuronal differentiation, would retain the active histone modifications at their regulatory regions.

In summary, alternative knockdown or knockout approaches in combination with broader systems for the analysis of gene expression profiles and histone modification landscapes would reinforce the findings and conclusions of the present work. In such scenario, the interpretation would be that HMGNs maintain P19 cell identity by promoting an active chromatin organisation that optimises the pluripotency network and protect the cells from precocious differentiation. Interestingly, a very recent paper from Bustin lab shows that the binding of HMGN to cell-specific enhancers stabilises cell identity (He et al, 2018). The authors observed that loss of HMGNs enhances the rate of reprogramming of MEFs into iPSCs following the ectopic expression of POU5F1, SOX2, KLF4, and MYC (He et al, 2018). Beyond the application that this finding can have in cellular reprogramming, the paper demonstrates that HMGNs are required to maintain a chromatin organisation that respond adequately to biological cues (He et al, 2018), which is complete agreement with the propositions of the present work. The emerging picture is that HMGNs increase the threshold required for cells to respond to signalling events inducing different cellular fates by providing an active histone modification landscape and open chromatin conformation at regulatory regions of cell-type specific genes, stabilising the transcriptional networks.

Although it seems to be concordance in the conclusions of other works and the present work, there are certain discrepancies in the data. Bustin lab has not observed neither important differences in the expression profiles of ESCs derived from *Hmgn1* and *Hmgn2* double-knockout and WT mice, nor changes in the steady state levels of H3K27ac and H3K4me3 (Deng et al, 2017; Deng et al, 2013; Deng et al, 2015; He et al, 2018). However, the culturing conditions for ESCs are designed to sustain pluripotency and prevent differentiation. Then, it is possible to speculate that the extrinsic stimulation is sufficient to stabilise the pluripotency network regardless the presence of HMGNs. In this regard, P19 cells are closer to the primed state of pluripotency, EpiSCs and hESCs, rather than to the naïve pluripotency achieved by culturing ESCs in 2i. Perhaps the threshold of P19 to undergo differentiation is lower than that of ESCs, and therefore, the cells are more sensitive to epigenetic perturbations.

ECCs constituted the first capture of pluripotency *in vitro*, however, their transformed nature questions their suitability as a stem cell model. As previously discussed, the P19 system was permissive for the identification of a phenotype after the loss of HMGNs that has not been identified in highly stable systems. Importantly, the experiments carried in P19 cells derived in similar conclusions of a recent work (He et al, 2018), which supports the suitability of P19 cells as a model to investigate the function of HMGNs in mammalian cells. Therefore, P19 cells can be further exploited to investigate the molecular mechanisms by which these proteins stabilise an active chromatin conformation, and therefore, the pre-existing transcriptional networks, impeding chromatin remodelling that result in precocious cellular conversion.

Nevertheless, P19 cells might not be the model of choice to conduct experiments aiming to understand the role of HMGNs in an entire organism. In this sense, the findings of the present work contrast the apparently normal development of *Hmgn1*-knockout, *Hmgn2*-knockout, and double-knockout mice. As previously reviewed, these mice are viable and do not present strong or obvious phenotypes (Birger et al, 2003; Deng et al, 2015). However, the pre- and post-implantation mouse epiblast is remarkably flexible. For instance, it is capable to rapidly adjust increases or decreases in cell number, which allow embryological perturbation and transplantation, and additionally, it shows cell fate re-specification by heterotopic grafting (Martello & Smith, 2014; Smith, 2017). Hence, the observations reported in the present work of increased differentiation and compromised self-renewal and pluripotency of P19 cells after the loss of HMGN1 or HMGN2 are not relevant for the early development of an entire organism that can adjust to variations in cell number in the case of the precocious differentiation of a few cells.

In contrast to the early embryo, the size and morphology of neural tissues relies on the precise spatio-temporal regulation of cellular fates and on the molecular switches between the expansion, neurogenic, and gliogenic phases (Hirabayashi & Gotoh, 2005; Miller & Gauthier, 2007). Interestingly, a previous study suggested that HMGN proteins promote astrocyte differentiation of NSCs at the onset of the astrogenic phase, but not in the neurogenic, during brain development (Nagao et al, 2014). Overexpression of HMGN1, HMGN2, and HMGN3

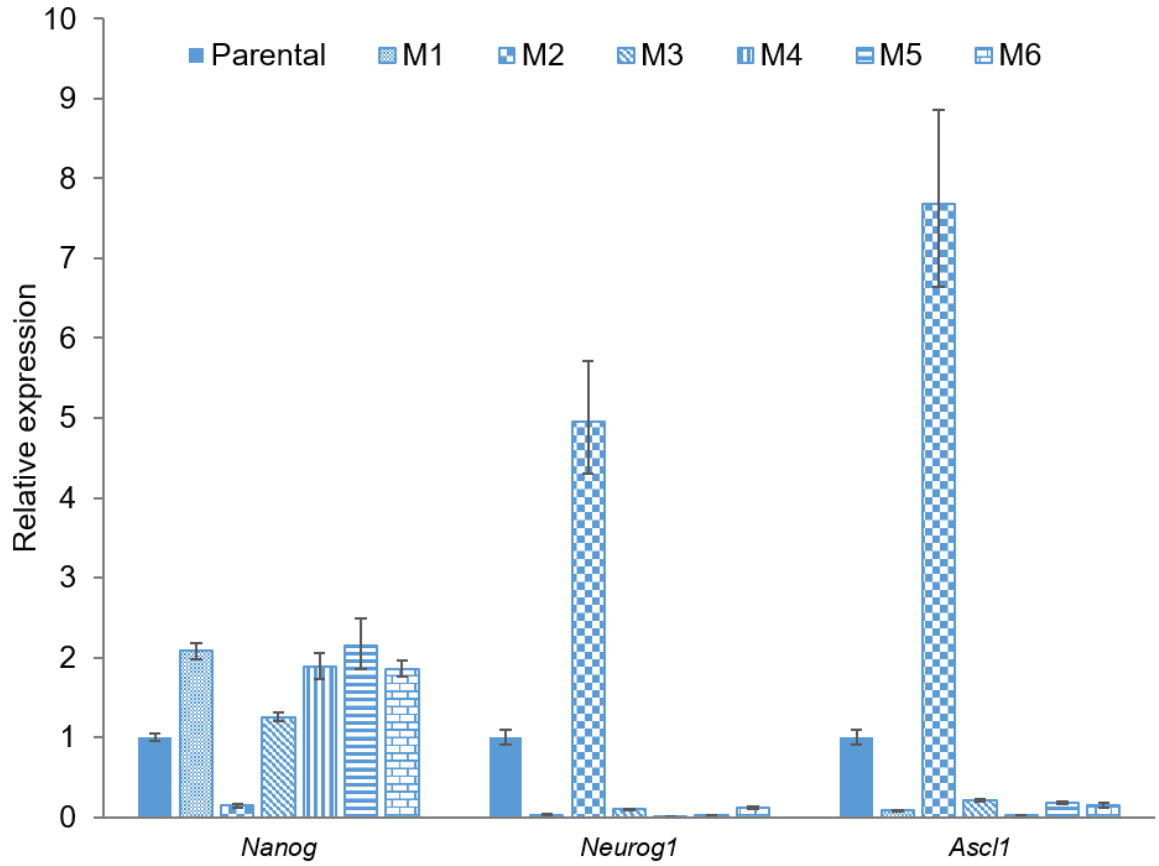
at the gliogenesis onset resulted in increased number of astrocytes at the expenses of neurons in the cerebral cortex of one-week mice, however, when the overexpression of the proteins occurred in the neurogenic phase, no effects were observed (Nagao et al, 2014). Conversely, knockdown of HMGN1, HMGN2, and HMGN3 at bird, markedly reduced the number of GFAP-positive astrocytes in the cortex of one-week old mice, while increasing the number of upper layer neurons (Nagao et al, 2014). These results suggest that HMGNs stabilise the neurogenic to gliogenic transition of NSCs, highlighting the relevance of the maintenance of the cell identity by HMGNs *in vivo*.

In addition, the present work shows that HMGNs maintain NSC identity, as *Hmgn*-knockout NSCs display higher mRNA levels of pro-neural factors and spontaneous neuronal differentiation. This could translate into precocious neurogenesis during nervous system development and could explain the reduction of NSCs at the SVZ of two-week old *Hmgn1*-knockout mice (Deng et al, 2013). It would be interesting to extend this finding to the neurogenic niches, SVZ and dentate gyrus of the hippocampus, of *Hmgn2*-knockout and double-knockout mice. Subsequently, to evaluate whether the lower number of NSCs leads to impaired adult neurogenesis when stimulated with exercise or learning and memory tasks. Finally, to investigate whether these neurogenic niches remain with lower NSCs numbers during all adulthood, or whether the pool is completely extinguished or re-established at certain age. These experiments are of special interest, since adult NSCs contribute to brain homeostasis, and adult neurogenesis confers repair and regeneration capacity, among other plasticity functions (Winner et al, 2011).

HMGNs are ubiquitous proteins with high degree of evolutionary conservation from their appearance in vertebrates. It is surprising that their loss does not result in stronger phenotypes. The present and previous studies have identified that the loss of HMGNs particularly influences processes where chromatin remodelling is required, ensuring proper cellular responses to stimuli. This is the case of cellular differentiation and reprogramming (He et al, 2018), B-cell activation by lipopolysaccharide (Zhang et al, 2016), glucose metabolism after feeding (Ueda et al, 2009), and DNA repair after damage (Birger et al, 2003). Therefore, it is possible to speculate that stress conditions reveal stronger

phenotypes in mice. Likewise, aging could trigger robust alterations. The *Hmgn1* and *Hmgn2* double-knockout mouse was recently generated (Deng et al, 2015) and there is not available information of aging processes in these or single-knockout animals. Observation and analysis of aging *Hmgn*-knockout mice would have a high value, particularly considering the growing body of evidence that associates aging with tissue-specific stem cell function (Goodell & Rando, 2015), cell types in which HMGNs are highly expressed and have important roles in their stability.

Appendix 1



Supplementary Figure Derivation of monoclonal lines from parental cells
P19 parental cells were plated at limited dilution in 96-well plates and six monoclonal lines were expanded from single cells (M1-M6). Relative expression of the pluripotency transcription factor *Nanog* and the pro-neural transcription factors *Neurog1* and *Ascl1* was determined by qRT-PCR. The graph represents the fold change in comparison with parental cells and error bars symbolise the SD of a PCR triplicate. One out of six clones shows lower *Nanog* expression. The transcription rates of *Nanog* and pro-neural transcription factor are inversely correlated. Data provided by A. Sindi (unpublished).

Appendix 2

Maintenance of active chromatin states by HMGN1 and HMGN2 is required for stem cell identity

Sylvia Garza-Manero[1,2], Abdulmajeed Abdulghani A. Sindi[1,2,6], Gokula Mohan[2,5], Ohoud Rehbini[2], Valentine H. M. Jeantet[2], Mariarca Bailo[2], Faeezah Abdul Latif[2], Maureen P. West[2], Ross Gurden[2], Lauren Finlayson[2], Silvija Svambaryte[2] , Adam G. West[2], Katherine L. West[2,3*]

1: equal contributions. 2: Institute of Cancer Sciences, 3: School of Life Sciences, College of Medical Veterinary and Life Sciences, University of Glasgow, University Avenue, Glasgow, UK, G12 8QQ. 5: Current address: Institute of Biological Sciences, Faculty of Science, University of Malaya, Kuala Lumpur. 6: Current address: Department of Basic Medical Sciences, Faculty of Applied Medical sciences, Albaha University, Albaha-Alaqiq, Saudi Arabia.

* corresponding author.

email addresses:

sylgzmn@hotmail.com, asindi_10@hotmail.com, g.mohan@um.edu.my, ohoud.rehbini@temple.edu, maureenwest91@gmail.com, v.jeantet.1@research.gla.ac.uk, Mariarca.Bailo@uws.ac.uk, faeezah8@live.com.my, r.gurden.1@research.gla.ac.uk, Lauren.Finlayson@ed.ac.uk, ssvambaryte@gmail.com, adam.west@glasgow.ac.uk, katherine.west@glasgow.ac.uk

Abstract

Background: Chromatin plasticity is thought to be fundamental to the pluripotency of embryonic stem cells. Here, we investigate whether HMGN proteins play a role in maintaining the epigenetic landscape in stem cells. HMGN family members modulate chromatin structure and influence epigenetic modifications. They are highly expressed during early development and in the neural stem/progenitor cells of the developing and adult brain.

Results: We show that functional knockout of *Hmgn1* or *Hmgn2* in pluripotent embryonal carcinoma cells leads to increased levels of spontaneous neuronal differentiation. This is accompanied by the loss of pluripotency markers and increases in expression of the pro-neural transcription factors *Neurog1* and *Ascl1*. Neural stem cells derived from these *Hmgn*-knockout lines also show increased levels of spontaneous neuronal differentiation and *Neurog1* expression. The loss of HMGN2 is associated with the disruption of active chromatin states at specific classes of gene. The levels of H3K4me3, H3K9ac, H3K27ac and H3K122ac are considerably reduced at the pluripotency genes *Nanog* and *Pou5f1*, which impacts transcription. At endodermal/mesodermal lineage-specific genes, the loss of HMGN2 leads to a switch from a bivalent state to a repressive chromatin configuration. However, at the neuronal lineage genes *Neurog1* and *Ascl1*, no epigenetic changes are observed and their bivalent states are retained.

Conclusions: We conclude that HMGN proteins play important roles in maintaining chromatin plasticity in stem cells, and are essential for maintaining stem cell identity and pluripotency.

Keywords: HMGN, chromatin, epigenetics, stem cells, embryonal carcinoma cells, differentiation, neuronal

List of References

Abranches, E., Guedes, A. M., Moravec, M., Maamar, H., Svoboda, P., Raj, A. & Henrique, D. (2014) Stochastic NANOG Fluctuations allow Mouse Embryonic Stem Cells to Explore Pluripotency. Development, 141(14), 2770-2779.

Alfonso, P. J., Crippa, M. P., Hayes, J. J. & Bustin, M. (1994) The Footprint of Chromosomal Proteins HMG-14 and HMG-17 on Chromatin Subunits. Journal of Molecular Biology, 236(1), 189-198.

Allen Institute for Brain Science (2004) Allen Mouse Brain Atlas.

Allen Institute for Brain Science (2008) Allen Developing Mouse Brain Atlas.

Amen, M., Espinoza, H. M., Cox, C., Liang, X., Wang, J., Link, T. M., Brennan, R. G., Martin, J. F. & Amendt, B. A. (2008) Chromatin-associated HMG-17 is a Major Regulator of Homeodomain Transcription Factor Activity Modulated by Wnt/beta-catenin Signaling. Nucleic Acids Research, 36(2), 462-476.

Ang, Y. S., Tsai, S. Y., Lee, D. F., Monk, J., Su, J., Ratnakumar, K., Ding, J., Ge, Y., Darr, H., Chang, B., Wang, J., Rendl, M., Bernstein, E., Schaniel, C. & Lemischka, I. R. (2011) Wdr5 Mediates Self-Renewal and Reprogramming via the Embryonic Stem Cell Core Transcriptional Network. Cell, 145(2), 183-197.

Arceci, R. J., King, A. A. J., Simon, C., Orkin, S. & Wilson, D. (1993) Mouse GATA-4: a Retinoic Acid-Inducible GATA-Binding Transcription Factor Expressed in Endodermally Derived Tissues and Heart. Molecular and Cellular Biology, 13(4), 2235-2246.

Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jorgensen, H. F., John, R. M., Gouti, M., Casanova, M., Warnes, G., Merkenschlager, M. & Fisher, A. G. (2006) Chromatin Signatures of Pluripotent Cell Lines. Nature Cell Biology, 8(5), 532-538.

Ball, P. (2003) Portrait of a Molecule. Nature, 421, 421.

Bannister, A. J. & Kouzarides, T. (2011) Regulation of Chromatin by Histone Modifications. Cell Research, 21, 381-395.

Barkess, G., Postnikov, Y., Campos, C. D., Mishra, S., Mohan, G., Verma, S., Bustin, M. & West, K. L. (2012) The Chromatin-binding Protein HMGN3 Stimulates Histone Acetylation and Transcription across the Glyt1 Gene. The Biochemical Journal, 442(3), 495-505.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A. & Horvath, P. (2007) CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. Science, 315(5819), 1709-1712.

Becker, P. B. & Hörz, W. (2002) ATP-Dependent Nucleosome Remodeling. Annual Reviews of Biochemistry, 71(1), 247-273.

Bell, O., Tiwari, V. K., Thomä, N. H. & Schübeler, D. (2011) Determinants and Dynamics of Genome Accessibility. Nature Reviews Genetics, 12, 554.

Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., Jaenisch, R., Wagschal, A., Feil, R., Schreiber, S. L. & Lander, E. S. (2006) A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. Cell, 125(2), 315-326.

Bertrand, N., Castro, D. S. & Guillemot, F. (2002) Proneural Genes and the Specification of Neural Cell Types. Nature Review Neurosciences, 3(7), 517-530.

Bibikova, M., Chudin, E., Wu, B., Zhou, L., Garcia, E., Liu, Y., Shin, S., Plaia, T. W., Auerbach, J. M., Arking, D. E., Gonzalez, R., Crook, J., Davidson, B., Schulz, T. C., Robins, A., Khanna, A., Sartipy, P., Hyllner, J., Vanguri, P., Savant-Bhonsale, S., Smith, A. K., Chakravarti, A., Maitra, A., Rao, M., Barker, D. L., Loring, J. F. & Fan, J. B. (2006) Human Embryonic Stem Cells Have a Unique Epigenetic Signature. Genome Research, 16(9), 1075-1083.

Bird, A. (2002) DNA Methylation Patterns and Epigenetic Memory. Genes and Development, 16(1), 6-21.

Birger, Y., Catez, F., Furusawa, T., Lim, J. H., Prymakowska-Bosak, M., West, K. L., Postnikov, Y. V., Haines, D. C. & Bustin, M. (2005) Increased Tumorigenicity and Sensitivity to Ionizing Radiation upon Loss of Chromosomal Protein HMGN1. Cancer Research, 65(15), 6711-6718.

Birger, Y., Ito, Y., West, K. L., Landsman, D. & Bustin, M. (2001) HMGN4, a Newly Discovered Nucleosome-Binding Protein Encoded by an Intronless Gene. DNA and Cell Biology, 20(5), 256-263.

Birger, Y., West, K. L., Postnikov, Y. V., Lim, J. H., Furusawa, T., Wagner, J. P., Laufer, C. S. & Kraemer, K. H. (2003) Chromosomal Protein HMGN1 Enhances the Rate of DNA Repair in Chromatin. EMBO, 22(7), 1665-1675.

Bjornsson, C. S., Apostolopoulou, M., Tian, Y. & Temple, S. (2015) It Takes a Village: Constructing the Neurogenic Niche. Developmental Cell, 32(4), 435-446.

Boyer, L. A., Lee, T. I., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., Guenther, M. G., Kumar, R. M., Murray, H. L., Jenner, R. G., Gifford, D. K., Melton, D. A., Jaenisch, R. & Young, R. A. (2005) Core Transcriptional Regulatory Circuitry in Human Embryonic Stem Cells. Cell, 122(6), 947-956.

Boyer, L. A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L. A., Lee, T. I., Levine, S. S., Wernig, M., Tajonar, A., Ray, M. K., Bell, G. W., Otte, A. P., Vidal, M., Gifford, D. K., Young, R. A. & Jaenisch, R. (2006) Polycomb Complexes Repress Developmental Regulators in Murine Embryonic Stem Cells. Nature, 441(7091), 349-353.

Brons, I. G., Smithers, L. E., Trotter, M. W., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes, S. M., Howlett, S. K., Clarkson, A., Ahrlund-Richter, L., Pedersen, R. A. & Vallier, L. (2007) Derivation of Pluripotent Epiblast Stem Cells from Mammalian Embryos. Nature, 448(7150), 191-195.

Buecker, C., Srinivasan, R., Wu, Z., Calo, E., Acampora, D., Faial, T., Simeone, A., Tan, M., Swigut, T. & Wysocka, J. (2014) Reorganization of Enhancer Patterns in Transition from Naive to Primed Pluripotency. Cell Stem Cell, 14(6), 838-853.

Busskamp, V., Lewis, N. E., Guye, P., Ng, A. H. M., Shipman, S. L., Byrne, S. M., Sanjana, N. E., Murn, J., Li, Y., Li, S., Stadler, M., Weiss, R. & Church, G. M. (2014) Rapid Neurogenesis through Transcriptional Activation in Human Stem Cells. Molecular Systems Biology, 10(11), 1-21.

Bustin, M. (2001) Chromatin Unfolding and Activation by HMGN Chromosomal Proteins. Trends in Biochemical Sciences, 26(7), 431-437.

Bustin, M., A. Lehn, D. & Landsman, D. (1990) Structural Features of the HMG Chromosomal Proteins and their Genes. Biochimica et Biophysica Acta, 1049(3), 231-243.

Bustin, M., Catez, F. & Lim, J. H. (2005) The Dynamics of Histone H1 Function in Chromatin. Molecular Cell, 17(5), 617-620.

Catez, F., Brown, D. T., Misteli, T. & Bustin, M. (2002) Competition Between Histone H1 and HMGNs for Chromatin Binding Sites. EMBO, 3(8), 760-766.

Catez, F., Yang, H., Tracey, K. J., Reeves, R., Misteli, T. & Bustin, M. (2004) Network of Dynamic Interactions between Histone H1 and High-Mobility-Group Proteins in Chromatin. Molecular and Cellular Biology, 24(10), 4321-4328.

Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S. & Smith, A. (2003) Functional Expression Cloning of Nanog, a Pluripotency Sustaining Factor in Embryonic Stem Cells. Cell, 113(5), 643-655.

Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L. & Smith, A. (2007) Nanog Safeguards Pluripotency and Mediates Germline Development. Nature, 450(7173), 1230-1234.

Chen, T. & Dent, S. Y. R. (2013) Chromatin Modifiers and Remodellers: Regulators of Cellular Differentiation. Nature Reviews Genetics, 15, 93.

Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V. B., Wong, E., Orlov, Y. L., Zhang, W., Jiang, J., Loh, Y. H., Yeo, H. C., Yeo, Z. X., Narang, V., Govindarajan, K. R., Leong, B., Shahab, A., Ruan, Y., Bourque, G., Sung, W. K., Clarke, N. D., Wei, C. L. & Ng, H. H. (2008) Integration of External Signaling Pathways with the Core Transcriptional Network in Embryonic Stem Cells. Cell, 133(6), 1106-1117.

Chi, A. S. & Bernstein, B. E. (2009) Pluripotent Chromatin State. Science, 323(5911), 220-221.

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A. & Zhang, F. (2013) Multiplex Genome Engineering Using CRISPR/Cas Systems. Science, 339(6121), 819-823.

Consortium, E. P. (2012) An Integrated Encyclopedia of DNA Elements in the Human Genome. Nature, 489, 57.

Conti, L., Pollard, S. M., Gorba, T., Reitano, E., Toselli, M., Biella, G., Sun, Y., Sanzone, S., Ying, Q. L., Cattaneo, E. & Smith, A. (2005) Niche-independent Symmetrical Self-renewal of a Mammalian Tissue Stem Cell. Plos Biology, 3(9), 1594-1606.

Crippa, M. P., Nickol, J. M. & Bustin, M. (1991) Developmental Changes in the Expression of High Mobility Group Chromosomal Proteins. Journal of Biological Chemistry, 266(5), 2712-2714.

Cuddapah, S., Schones, D. E., Cui, K., Roh, T. Y., Barski, A., Wei, G., Rochman, M., Bustin, M. & Zhao, K. (2011) Genomic Profiling of HMGN1 Reveals an Association with Chromatin at Regulatory Regions. Molecular and Cellular Biology, 31(4), 700-709.

Davidson, K. C., Mason, E. A. & Pera, M. F. (2015) The Pluripotent State in Mouse and Human. Development, 142(18), 3090-3099.

Deng, T., Postnikov, Y., Zhang, S., Garrett, L., Becker, L., Racz, I., Holter, S. M., Wurst, W., Fuchs, H., Gailus-Durner, V., de Angelis, M. H. & Bustin, M. (2017) Interplay Between H1 and HMGN Epigenetically Regulates OLIG1&2 Expression and Oligodendrocyte Differentiation. Nucleic Acids Research, 45(6), 3031-3045.

Deng, T., Zhu, I., Zhang, S., Leng, F., Cherukuri, S., Hansen, L., Mariño-Ramírez, L., Meshorer, E., Landsman, D. & Bustin, M. (2013) HMGN1 Modulates Nucleosome Occupancy and DNase I Hypersensitivity at the CpG Island Promoters of Embryonic Stem Cells. Molecular and Cellular Biology, 33(16), 3377-3389.

Deng, T., Zhu, Z. I., Zhang, S., Postnikov, Y., Huang, D., Horsch, M., Furusawa, T., Beckers, J., Rozman, J., Klingenspor, M., Amarie, O., Graw, J., Rathkolb, B., Wolf, E., Adler, T., Busch, D. H., Gailus-Durner, V., Fuchs, H., Hrabe de Angelis, M., van der Velde, A., Tessarollo, L., Ovcherenko, I., Landsman, D. & Bustin, M. (2015) Functional Compensation Among HMGN Variants Modulates the DNase I Hypersensitive Sites at Enhancers. Genome Research, 25(9), 1295-1308.

Dhalluin, C., Carlson, J. E., Zeng, L., He, C., Aggarwal, A. K. & Zhou, M. M. (1999) Structure and Ligand of a Histone Acetyltransferase Bromodomain. Nature, 399, 491-496.

Ding, H. F., Bustin, M. & Hansen, U. (1997) Alleviation of Histone H1-Mediated Transcriptional Repression and Chromatin Compaction by the Acidic Activation Region in Chromosomal Protein HMG-14. Molecular and Cellular Biology, 17(10), 5843-5855.

Ding, S., Wu, T. Y. H., Brinker, A., Peters, E. C., Hur, W., Gray, N. S. & Schultz, P. G. (2003) Synthetic Small Molecules that Control Stem Cell Fate. Proceedings of the National Academy of Sciences, 100(13), 7632-7637.

Doetzlhofer, A., Basch, M. L., Ohyama, T., Gessler, M., Groves, A. K. & Segil, N. (2009) Hey2 Regulation by FGF Provides a Notch-Independent Mechanism for Maintaining Pillar Cell Fate in the Organ of Corti. Developmental Cell, 16(1), 58-69.

Dominguez, A. A., Lim, W. A. & Qi, L. S. (2015) Beyond Editing: Repurposing CRISPR–Cas9 for Precision Genome Regulation and Interrogation. Nature Reviews Molecular Cell Biology, 17, 5.

Dovey, H. F., John, V., Anderson, J. P., Chen, L. Z., De Saint Andrieu, P., Fang, L. Y., Freedman, S. B., Folmer, B., Goldbach, E., Holsztynska, E. J., Hu, K. L., Johnson-Wood, K. L., Kennedy, S. L., Kholodenko, D., Knops, J. E., Latimer, L. H., Lee, M., Liao, Z., Lieberburg, I. M., Motter, R. N., Mutter, L. C., Nietz, J., Quinn, K. P., Sacchi, K. L., Seubert, P. A., Shopp, G. M., Thorsett, E. D., Tung, J. S., Wu, J., Yang, S., Yin, C. T., Schenk, D. B., May, P. C., Altstiel, L. D., Bender, M. H., Boggs, L. N., Britton, T. C., Clemens, J. C., Czilli, D. L., Dieckman-McGinty, D. K., Droste, J. J., Fuson, K. S., Gitter, B. D., Hyslop, P. A., Johnstone, E. M., Li, W. Y., Little, S. P., Mabry, T. E., Miller, F. D., Ni, B., Nissen, J. S., Porter, W. J., Potts, B. D., Reel, J. K., Stephenson, D., Su, Y., Shipley, L. A., Whitesitt, C. A., Yin, T. & Audia, J. E. (2009) Functional Gamma-secretase Inhibitors Reduce Beta-amyloid Peptide Levels in Brain. Journal of Neurochemistry, 76(1), 173-181.

Duncan, E. M., Muratore-Schroeder, T. L., Cook, R. G., Garcia, B. A., Shabanowitz, J., Hunt, D. F. & Allis, C. D. (2008) Cathepsin L Proteolytically

Processes Histone H3 During Mouse Embryonic Stem Cell Differentiation. Cell, 135(2), 284-294.

Dyer, C., Blanc, E., Hanisch, A., Roehl, H., Otto, G. W., Yu, T., Basson, M. A. & Knight, R. (2014) A Bi-modal Function of Wnt Signalling Directs an FGF Activity Gradient to Spatially Regulate Neuronal Differentiation in the Midbrain. Development, 141(1), 63-72.

Efroni, S., Duttagupta, R., Cheng, J., Dehghani, H., Hoeppner, D. J., Dash, C., Bazett-Jones, D. P., Le Grice, S., McKay, R. D. G., Buetow, K. H., Gingeras, T. R., Misteli, T. & Meshorer, E. (2008) Global Transcription in Pluripotent Embryonic Stem Cells. Cell Stem Cell, 2(5), 437-447.

Evans, M. J. & Kaufman, M. H. (1981) Establishment in Culture of Pluripotential Cells from Mouse Embryos. Nature, 292, 154-156.

Fang, F., Xu, Y., Chew, K. K., Chen, X., Ng, H. H. & Matsudaira, P. (2014) Coactivators p300 and CBP Maintain the Identity of Mouse Embryonic Stem Cells by Mediating Long-range Chromatin Structure. Stem Cells, 32(7), 1805-1816.

Fang, H. B., Mi, Y., Zhang, Y., Wu, N. H. & Shen, Y. F. (2010) HDAC3 Augments the Autoregulation of NeuroD Gene in P19 Cells. NeuroReport, 21(1), 19-23.

Farah, M. H., Olson, J. M., Sucic, H. B., Hume, R. I., Tapscott, S. J. & Turner, D. L. (2000) Generation of Neurons by Transient Expression of Neural bHLH Proteins in Mammalian Cells. Development, 127(4), 693-702.

Fehling, H. J., Lacaud, G., Kubo, A., Kennedy, M., Robertson, S., Keller, G. & Kouskoff, V. (2003) Tracking Mesoderm Induction and its Specification to the Hemangioblast during Embryonic Stem Cell Differentiation. Development, 130(17), 4217-4227.

Felsenfeld, G. & Groudine, M. (2003) Controlling the Double Helix. Nature, 421, 448-453.

Fujikura, J., Yamato, E., Yonemura, S., Hosoda, K., Masui, S., Nakao, K., Miyazaki, J. & Niwa, H. (2002) Differentiation of Embryonic Stem Cells is Induced by GATA Factors. Genes and Development, 16(7), 784-789.

Furusawa, T., Lim, J. H., Catez, F., Birger, Y., Mackem, S. & Bustin, M. (2006) Down-regulation of Nucleosomal Binding Protein HMGN1 Expression during Embryogenesis Modulates Sox9 Expression in Chondrocytes. Molecular and Cellular Biology, 26(2), 592-604.

Gage, F. H. (2000) Mammalian Neural Stem Cells. Science, 287(5457), 1433-1438.

Göke, J., Jung, M., Behrens, S., Chavez, L., O'Keeffe, S., Timmermann, B., Lehrach, H., Adjaye, J. & Vingron, M. (2011) Combinatorial Binding in Human and Mouse Embryonic Stem Cells Identifies Conserved Enhancers Active in Early Embryonic Development. Plos Computational Biology, 7(12), e1002304.

González-Romero, R., Eirín-López, J. M. & Ausió, J. (2015) Evolution of High Mobility Group Nucleosome-Binding Proteins and Its Implications for Vertebrate Chromatin Specialization. Molecular Biology and Evolution, 32(1), 121-131.

Goodell, M. A. & Rando, T. A. (2015) Stem Cells and Healthy Aging. Science, 350(6265), 1199-1204.

Guenther, M. G., Levine, S. S., Boyer, L. A., Jaenisch, R. & Young, R. A. (2007) A Chromatin Landmark and Transcription Initiation at Most Promoters in Human Cells. Cell, 130(1), 77-88.

Guo, G., von Meyenn, F., Santos, F., Chen, Y., Reik, W., Bertone, P., Smith, A. & Nichols, J. (2016) Naive Pluripotent Stem Cells Derived Directly from Isolated Cells of the Human Inner Cell Mass. Stem Cell Reports, 6(4), 437-446.

Hansen, J. C., Lu, X., Ross, E. D. & Woody, R. W. (2006) Intrinsic Protein Disorder, Amino Acid Composition, and Histone Terminal Domains. Journal of Biological Chemistry, 281(4), 1853-1856.

Hartley, P. D. & Madhani, H. D. (2009) Mechanisms that Specify Promoter Nucleosome Location and Identity. Cell, 137(3), 445-458.

He, B., Deng, T., Zhu, I., Furusawa, T., Zhang, S., Tang, W., Postnikov, Y., Ambs, S., Li, C. C., Livak, F., Landsman, D. & Bustin, M. (2018) Binding of HMGN Proteins to Cell Specific Enhancers Stabilizes Cell Identity. Nature Communications, 9(1), 5240.

Hemmati-Brivanlou, A. (1997) Vertebrate Neural Induction. Annual Review of Neuroscience, 20, 43-60.

Herberg, M., Glauche, I., Zerjatke, T., Winzi, M., Buchholz, F. & Roeder, I. (2016) Dissecting Mechanisms of Mouse Embryonic Stem Cells Heterogeneity Through a Model-based Analysis of Transcription Factor Dynamics. Journal of Royal Society Interface, 13(117), 1-10.

Hirabayashi, Y. & Gotoh, Y. (2005) Stage-dependent Fate Determination of Neural Precursor Cells in Mouse Forebrain. Neuroscience Research, 51(4), 331-336.

Hirabayashi, Y., Itoh, Y., Tabata, H., Nakajima, K., Akiyama, T., Masuyama, N. & Gotoh, Y. (2004) The Wnt/beta-catenin Pathway Directs Neuronal Differentiation of Cortical Neural Precursor Cells. Development, 131(12), 2791-2801.

Hirabayashi, Y., Suzki, N., Tsuboi, M., Endo, T. A., Toyoda, T., Shinga, J., Koseki, H., Vidal, M. & Gotoh, Y. (2009) Polycomb Limits the Neurogenic Competence of Neural Precursor Cells to Promote Astrogenic Fate Transition. Neuron, 63(5), 600-613.

Hock, R., Scheer, U. & Bustin, M. (1998a) Chromosomal Proteins HMG-14 and HMG-17 Are Released from Mitotic Chromosomes and Imported into the Nucleus by Active Transport. The Journal of Cell Biology, 143(6), 1427-1436.

Hock, R., Wilde, F., Scheer, U. & Bustin, M. (1998b) Dynamic Relocation of Chromosomal Protein HMG-17 in the Nucleus is Dependent on Transcriptional Activity. EMBO, 17(23), 6992-7001.

Hoodless, P. A. & Hemmati-Brivanlou, A. (1997) Inhibitory Control of Neural Differentiation in Mammalian Cells. Development Genes and Evolution, 207(1), 19-28.

Huang, S. M., Mishina, Y. M., Liu, S., Cheung, A., Stegmeier, F., Michaud, G. A., Charlat, O., Wiellette, E., Zhang, Y., Wiessner, S., Hild, M., Shi, X., Wilson, C. J., Mickanin, C., Myer, V., Fazal, A., Tomlinson, R., Serluca, F., Shao, W., Cheng, H., Shultz, M., Rau, C., Schirle, M., Schlegl, J., Ghidelli, S., Fawell, S., Lu, C., Curtis, D., Kirschner, M. W., Lengauer, C., Finan, P. M., Tallarico, J. A., Bouwmeester, T., Porter, J. A., Bauer, A. & Cong, F. (2009) Tankyrase

Inhibition Stabilizes Axin and Antagonizes Wnt Signalling. Nature, 461(7264), 614-620.

Huelsken, J., Vogel, R., Brinkmann, V., Erdmann, B., Birchmeier, C. & Birchmeier, W. (2000) Requirement for ß-Catenin in Anterior-Posterior Axis Formation in Mice. The Journal of Cell Biology, 148(3), 567-578.

Israsena, N., Hu, M., Fu, W., Kan, L. & Kessler, J. A. (2004) The Presence of FGF2 Signaling Determines Whether Beta-catenin Exerts Effects on Proliferation or Neuronal Differentiation of Neural Stem Cells. Developmental Biology, 268(1), 220-231.

Jenuwein, T. & Allis, C. D. (2001) Translating the Histone Code. Science, 293(5532), 1074-1080.

Jin, C., Zang, C., Wei, G., Cui, K., Peng, W., Zhao, K. & Felsenfeld, G. (2009a) H3.3/H2A.Z Double Variant–containing Nucleosomes Mark "Nucleosome-free Regions" of Active Promoters and other Regulatory Regions. Nature Genetics, 41, 941-945.

Jin, Z., Liu, L., Bian, W., Chen, Y., Xu, G., Cheng, L. & Jing, N. (2009b) Different Transcription Factors Regulate nestin Gene Expression during P19 Cell Neural Differentiation and Central Nervous System Development. The Journal of Biological Chemistry, 284(12), 8160-8173.

Jin, Z., Liu, L., Zhong, H., Zhang, K., Chen, Y., Bian, W., Cheng, L. & Jing, N. (2006) Second Intron of Mouse Nestin Gene Directs its Expression in Pluripotent Embryonic Carcinoma Cells through POU Factor Binding Site. Acta Biochimica et Biophysica, 38(3), 207-212.

Johns, E. W. (1982) The HMG Chromosomal ProteinsAcademic Press.

Jones-Villeneuve, E. M., McBurney, M. W., Rogers, K. A. & Kalnins, V. I. (1982) Retinoic Acid Induces Embryonal Carcinoma Cells to Differentiate into Neurons and Glial Cells. The Journal of Cell Biology, 94(2), 253-262.

Kageyama, R., Ohtsuka, T. & Kobayashi, T. (2007) The Hes Gene Family: Repressors and Oscillators that Orchestrate Embryogenesis. Development, 134(7), 1243-1251.

Kalkan, T., Olova, N., Roode, M., Mulas, C., Lee, H., Nett, I., Marks, H., Walker, R., Stunnenberg, H. G., Lilley, K. S., Nichols, J., Reik, W., Bertone, P. & Smith, A. (2017) Tracking the Embryonic Stem Cell Transition from Ground State Pluripotency. Development, 144(7), 1221-1234.

Kalmar, T., Lim, C., Hayward, P., Munoz-Descalzo, S., Nichols, J., Garcia-Ojalvo, J. & Martinez Arias, A. (2009) Regulated Fluctuations in Nanog Expression Mediate Cell Fate Decisions in Embryonic Stem Cells. Plos Biology, 7(7), e1000149.

Kang, W., Wong, L. C., Shi, S. H. & Hébert, J. M. (2009) The Transition from Radial Glial to Intermediate Progenitor Cell Is Inhibited by FGF Signaling during Corticogenesis. The Journal of Neuroscience, 29(46), 14571-14580.

Kato, H., van Ingen, H., Zhou, B. R., Feng, H., Bustin, M., Kay, L. E. & Bai, Y. (2011) Architecture of the High Mobility Group Nucleosomal Protein 2-Nucleosome Complex as Revealed by Methyl-based NMR. Proceedings of the National Academy of Sciences, 108(30), 12283-12288.

Kelly, G. M. & Gatie, M. I. (2017) Mechanisms Regulating Stemness and Differentiation in Embryonal Carcinoma Cells. Stem Cells International, 2017, 3684178.

Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M. & Haussler, D. (2002) The Human Genome Browser at UCSC. Genome Research, 12(6), 996-1006.

Kessaris, N., Fogarty, M., Iannarelli, P., Grist, M., Wegner, M. & Richardson, W. (2005) Competing Waves of Oligodendrocytes in the Forebrain and Postnatal Elimination of an Embryonic Lineage. Nature Neuroscience, 9, 173-179.

Kevenaar, J. T. & Hoogenraad, C. C. (2015) The Axonal Cytoskeleton: from Organization to Function. Frontiers in Molecular Neuroscience, 8, 44.

Kim, J., Chu, J., Shen, X., Wang, J. & Orkin, S. H. (2008) An Extended Transcriptional Network for Pluripotency of Embryonic Stem Cells. Cell, 132(6), 1049-1061.

Kinoshita, M. & Smith, A. (2018) Pluripotency Deconstructed. Development, Growth and Differentiation, 60(1), 44-52.

Kishi, Y., Fujii, Y., Hirabayashi, Y. & Gotoh, Y. (2012) HMGA Regulates the Global Chromatin State and Neurogenic Potential in Neocortical Precursor Cells. Nature Neurosci, 15(8), 1127-1133.

Kobayashi, T. & Kageyama, R. (2010) Hes1 Regulates Embryonic Stem Cell Differentiation by Suppressing Notch Signaling. Genes to Cells, 15(7), 689-698.

Kobayashi, T., Mizuno, H., Imayoshi, I., Furusawa, C., Shirahige, K. & Kageyama, R. (2009) The Cyclic Gene Hes1 Contributes to Diverse Differentiation Responses of Embryonic Stem Cells. Genes and Development, 23(16), 1870-1875.

Kojima, Y., Kaufman-Francis, K., Studdert, Joshua B., Steiner, Kirsten A., Power, Melinda D., Loebel, David A. F., Jones, V., Hor, A., de Alencastro, G., Logan, Grant J., Teber, Erdahl T., Tam, Oliver H., Stutz, Michael D., Alexander, Ian E., Pickett, Hilda A. & Tam, Patrick P. L. (2014) The Transcriptional and Functional Properties of Mouse Epiblast Stem Cells Resemble the Anterior Primitive Streak. Cell Stem Cell, 14(1), 107-120.

Kornberg, R. D. (1974) Chromatin Structure: A Repeating Unit of Histones and DNA. Science, 184(4139), 868-871.

Krejčí, J., Uhlířová, R., Galiová, G., Kozubek, S., Šmigová, J. & Bártová, E. (2009) Genome-wide Reduction in H3K9 Acetylation during Human Embryonic Stem Cell Differentiation. Journal of Cellular Physiology, 219(3), 677-687.

Kriegstein, A. & Alvarez-Buylla, A. (2009) The Glial Nature of Embryonic and Adult Neural Stem Cells. Annual Review of Neuroscience, 32(1), 149-184.

Kugler, J. E., Horsch, M., Huang, D., Furusawa, T., Rochman, M., Garrett, L., Becker, L., Bohla, A., Holter, S. M., Prehn, C., Rathkolb, B., Racz, I., Aguilar-Pimentel, J. A., Adler, T., Adamski, J., Beckers, J., Busch, D. H., Eickelberg, O., Klopstock, T., Ollert, M., Stoger, T., Wolf, E., Wurst, W., Yildirim, A. O., Zimmer, A., Gailus-Durner, V., Fuchs, H., Hrabe de Angelis, M., Garfinkel, B., Orly, J., Ovcharenko, I. & Bustin, M. (2013) High Mobility Group N Proteins Modulate the Fidelity of the Cellular Transcriptional Profile in a Tissue- and Variant-specific Manner. The Journal of Biological Chemistry, 288(23), 16690-16703.

Kunath, T., Saba-El-Leil, M. K., Almousailleakh, M., Wray, J., Meloche, S. & Smith, A. (2007) FGF Stimulation of the Erk1/2 Signalling Cascade Triggers Transition of Pluripotent Embryonic Stem Cells from Self-renewal to Lineage Commitment. Development, 134(16), 2895-2902.

Kuwahara, A., Hirabayashi, Y., Knoepfler, P. S., Taketo, M. M., Sakai, J., Kodama, T. & Gotoh, Y. (2010) Wnt Signaling and its Downstream Target N-myc Regulate Basal Progenitors in the Developing Neocortex. Development, 137(7), 1035-1044.

Lai, E. C. (2004) Notch Signaling: Sontrol of Cell Communication and Cell Fate. Development, 131(5), 965-973.

Laverriere, A. C., MacNeill, C., Mueller, C., Poelmann, R., Burch, J. B. E. & Evans, T. (1994) GATA-4/5/6, a Subfamily of Three Transcription Factors Transcribed in Developing Heart and Gut. The American Society for Biochemistry and Molecular Biology, 269(37), 23177-23184.

Lee, J. H., Hart, S. R. L. & Skalnik, D. G. (2003) Histone Deacetylase Activity is Required for Embryonic Stem Cell Differentiation. Genesis, 38(1), 32-38.

Lehtonen, S. & Lehtonen, E. (2001) HMG-17 is an Early Marker of Inductive Interactions in the Developing Mouse Kidney. Differentiation, 67, 154-163.

Lehtonen, S., Olkkonen, V., Stapleton, M., Zerial, M. & Lehtonen, E. (1998) HMG-17, a Chromosomal Non-Histone Protein, Shows Developmental Regulation during Organogenesis. International Journal of Developmental Biology, 42, 775-782.

Lendahl, U., Zimmerman, L. B. & McKay, R. D. G. (1990) CNS Stem Cells Express a New Class of Intermediate Filament Protein. Cell, 60(4), 585-595.

Leonetti, M. D., Sekine, S., Kamiyama, D., Weissman, J. S. & Huang, B. (2016) A Scalable Strategy for High-Throughput GFP Tagging of Endogenous Human Proteins. Proceedings of the National Academy of Sciences, 113(25), E3501-E3508.

Liao, S., Tammaro, M. & Yan, H. (2015) Enriching CRISPR-Cas9 Targeted Cells by Co-Targeting the HPRT Gene. Nucleic Acids Research, 43(20), e134.

Lim, J. H., Bustin, M., Ogryzko, V. V. & Postnikov, Y. V. (2002) Metastable Macromolecular Complexes Containing High Mobility Group Nucleosome-binding Chromosomal Proteins in HeLa Nuclei. Journal of Biological Chemistry, 277(23), 20774-20782.

Lim, J. H., Catez, F., Birger, Y., West, K. L., Prymakowska-Bosak, M., Postnikov, Y. V. & Bustin, M. (2004) Chromosomal Protein HMGN1 Modulates Histone H3 Phosphorylation. Molecular and Cellular Biology, 15(4), 573-584.

Lim, J. H., West, K. L., Rubinstein, Y., Bergel, M., Postnikov, Y. V. & Bustin, M. (2005) Chromosomal Protein HMGN1 Enhances the Acetylation of Lysine 14 in Histone H3. EMBO, 24(17), 3038-3048.

Lowell, S., Benchoua, A., Heavey, B. & Smith, A. (2006) Notch Promotes Neural Lineage Entry by Pluripotent Embryonic Stem Cells. Plos Biology, 4(5), e121.

Lyu, J., Costantini, F., Jho, E. H. & Joo, C. K. (2003) Ectopic Expression of Axin Blocks Neuronal Differentiation of Embryonic Carcinoma P19 Cells. The Journal of Biological Chemistry, 278(15), 13487-13495.

MacDonald, B. T., Tamai, K. & He, X. (2009) Wnt/ß-catenin Signaling: Components, Mechanisms, and Diseases. Developmental cell, 17(1), 9-26.

Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E. & Church, G. M. (2013) RNA-Guided Human Genome Engineering via Cas9. Science, 339(6121), 823-826.

Malicet, C., Rochman, M., Postnikov, Y. & Bustin, M. (2011) Distinct Properties of Human HMGN5 Reveal a Rapidly Evolving but Functionally Conserved Nucleosome Binding Protein. Molecular and Cellular Biology, 31(13), 2742-2755.

Mallanna, S. K., Boer, B., Desler, M. & Rizzino, A. (2008) Differential Regulation of the Oct-3/4 Gene in Cell Culture Model Systems that Parallel Different Stages of Mammalian Development. Molecular Reproduction and Development, 75(8), 1247-1257.

Margueron, R. & Reinberg, D. (2011) The Polycomb Complex PRC2 and its Mark in Life. Nature, 469, 343-349.

Marks, H., Kalkan, T., Menafra, R., Denissov, S., Jones, K., Hofemeister, H., Nichols, J., Kranz, A., Francis S., Smith, A. & Stunnenberg, H. G. (2012) The Transcriptional and Epigenomic Foundations of Ground State Pluripotency. Cell, 149(3), 590-604.

Martello, G. & Smith, A. (2014) The Nature of Embryonic Stem Cells. Annual Review of Cell and Developmental Biology, 30, 647-675.

Martin, G. R. (1980) Teratocarcinomas and Mammalian Embryogenesis. Science, 209(4458), 768-776.

Martin, G. R. & Evans, M. J. (1975) Differentiation of Clonal Lines of Teratocarcinoma Cells: Formation of Embryoid Bodies in vitro. Proceedings of the National Academy of Sciences of the United States of America, 72(4), 1441-1445.

Martínez de Paz, A. & Ausió, J. (2016) HMGNs: The Enhancer Charmers. BioEssays, 38(3), 226-231.

Martynoga, B., Drechsel, D. & Guillemot, F. (2012) Molecular Control of Neurogenesis: A View from the Mammalian Cerebral Cortex. Cold Spring Harbor Perspectives in Biology, 4(10), a008359.

Masui, S., Nakatake, Y., Toyooka, Y., Shimosato, D., Yagi, R., Takahashi, K., Okochi, H., Okuda, A., Matoba, R., Sharov, A. A., Ko, M. S. H. & Niwa, H. (2007) Pluripotency Governed by Sox2 via Regulation of Oct3/4 Expression in Mouse Embryonic Stem Cells. Nature Cell Biology, 9, 625-635.

Mattout, A. & Meshorer, E. (2010) Chromatin Plasticity and Genome Organization in Pluripotent Embryonic Stem Cells. Current Opinion in Cell Biology, 22(3), 334-341.

McBurney, M. W. (1993) P19 Embryonal Carcinoma Cells. International Journal of Developmental Biology, 37, 135-140.

McBurney, M. W. & Rogers, B. J. (1982) Isolation of Male Embryonal Carcinoma Cells and their Chromosome Replication Patterns. Developmental Biology, 89(2), 503-508.

McCarty, M. (2003) Discovering Genes are Made of DNA. Nature, 421, 406.

Meshorer, E., Yellajoshula, D., George, E., Scambler, P. J., Brown, D. T. & Misteli, T. (2006) Hyperdynamic Plasticity of Chromatin Proteins in Pluripotent Embryonic Stem Cells. Developmental Cell, 10(1), 105-116.

Mikkelsen, T. S., Ku, M., Jaffe, D. B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T. K., Koche, R. P., Lee, W., Mendenhall, E., O'Donovan, A., Presser, A., Russ, C., Xie, X., Meissner, A., Wernig, M., Jaenisch, R., Nusbaum, C., Lander, E. S. & Bernstein, B. E. (2007) Genome-Wide Maps of Chromatin State in Pluripotent and Lineage-Committed Cells. Nature, 448(7153), 553-560.

Miller, F. D. & Gauthier, A. S. (2007) Timing is Everything: Making Neurons versus Glia in the Developing Cortex. Neuron, 54(3), 357-369.

Mohamed, O. A., Bustin, M. & Clarke, H. J. (2001) High-Mobility Group Proteins 14 and 17 Maintain the Timing of Early Embryonic Development in the Mouse. Developmental Biology, 229(1), 237-249.

Mohammadi, M., McMahon, G., Sun, L., Tang, C., Hirth, P., Yeh, B. K., Hubbard, S. R. & Schlessinger, J. (1997) Structures of the Tyrosine Kinase Domain of Fibroblast Growth Factor Receptor in Complex with Inhibitors. Science, 276(5314), 955-960.

Molyneaux, B. J., Arlotta, P., Menezes, J. R. L. & Macklis, J. D. (2007) Neuronal Subtype Specification in the Cerebral Cortex. Nature Reviews Neuroscience, 8, 427-437.

Morgenstern, J. P. & Land, H. (1990) Advanced Mammalian Gene Transfer: High Titre Retroviral Vectors with Multiple Drug Selection Markers and a Complementary Helper-free Packaging Cell Line. Nucleic Acids Research, 18(12), 3587-3596.

Morrison, S. J., Perez, S. E., Qiao, Z., Verdi, J. M., Hicks, C., Weinmaster, G. & Anderson, D. J. (2000) Transient Notch Activation Initiates an Irreversible Switch from Neurogenesis to Gliogenesis by Neural Crest Stem Cells. Cell, 101(5), 499-510.

Mulas, C., Kalkan, T. & Smith, A. (2017) NODAL Secures Pluripotency upon Embryonic Stem Cell Progression from the Ground State. Stem Cell Reports, 9(1), 77-91.

Munoz-Sanjuan, I. & Brivanlou, A. H. (2002) Neural Induction, the Default Model and Embryonic Stem Cells. Nature Review Neurosciences, 3(4), 271-280.

Murphy, K. J., Cutter, A. R., Fang, H. B., Postnikov, Y. V., Bustin, M. & Hayes, J. J. (2017) HMGN1 and 2 Remodel Core and Linker Histone Tail Domains within Chromatin. Nucleic Acids Research, 45(17), 9917-9930.

Nagao, M., Lanjakornsiripan, D., Itoh, Y., Kishi, Y., Ogata, T. & Gotoh, Y. (2014) High Mobility Group Nucleosome-Binding Family Proteins Promote Astrocyte Differentiation of Neural Precursor Cells. Stem Cells, 32(11), 2983-2997.

Nakayama, J., Rice, J. C., Strahl, B. D., Allis, C. D. & Grewal, S. I. S. (2001) Role of Histone H3 Lysine 9 Methylation in Epigenetic Control of Heterochromatin Assembly. Science, 292(5514), 110-113.

Nakayama, K., Satoh, T., Igari, A., Kageyama, R. & Nishida, E. (2008) FGF Induces Oscillations of Hes1 Expression and Ras/ERK Activation. Current Biology, 18(8), R332-R334.

Nakayama, Y., Wada, A., Inoue, R., Terasawa, K., Kimura, I., Nakamura, N. & Kurosaka, A. (2014) A Rapid and Efficient Method for Neuronal Induction of the P19 Embryonic Carcinoma Cell Line. Journal of Neurosciences Methods, 227, 100-106.

Nichols, J. & Smith, A. (2009) Naive and Primed Pluripotent States. Cell Stem Cell, 4(6), 487-492.

Nichols, J. & Smith, A. (2012) Pluripotency in the Embryo and in Culture. Cold Spring Harbor Perspectives in Biology, 4(8), a008128.

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Schöler, H. R. & Smith, A. (1998) Formation of Pluripotent Stem Cells in the Mammalian Embryo Depends on the POU Transcription Factor Oct4. Cell, 95(3), 379-391.

Nieto, M., Schuurmans, C., Britz, O. & Guillemot, F. (2001) Neural bHLH Genes Control the Neuronal versus Glial Fate Decision in Cortical Progenitors. Neuron, 29(2), 401-413.

Niwa, H. (2007) Open Conformation Chromatin and Pluripotency. Genes and Development, 21(21), 2671-2676.

Niwa, H., Burdon, T., Chambers, I. & Smith, A. (1998) Self-renewal of Pluripotent Embryonic Stem Cells is Mediated via Activation of STAT3. Genes and Development, 12(13), 2048-2060.

Niwa, H., Miyazaki, J. & Smith, A. (2000) Quantitative Expression of Oct-3/4 Defines Differentiation, Dedifferentiation or Self-renewal of ES Cells. Nature Genetics, 24, 372-376.

Ohtsuka, T., Ishibashi, M., Gradwohl, G., Nakanishi, S., Guillemot, F. & Kageyama, R. (1999) Hes1 and Hes5 as Notch Effectors in Mammalian Neuronal Differentiation. EMBO, 18(8), 2196-2207.

Ohtsuka, T., Sakamoto, M., Guillemot, F. & Kageyama, R. (2001) Roles of the Basic Helix-Loop-Helix Genes Hes1 and Hes5 in Expansion of Neural Stem Cells of the Developing Brain. Journal of Biological Chemistry, 276(32), 30467-30474.

Oishi, K., Kamakura, S., Isazawa, Y., Yoshimatsu, T., Kuida, K., Nakafuku, M., Masuyama, N. & Gotoh, Y. (2004) Notch Promotes Survival of Neural Precursor Cells via Mechanisms Distinct from those Regulating Neurogenesis. Developmental Biology, 276(1), 172-184.

Okamoto, K., Okazawa, H., Okuda, A., Sakai, M., Muramatsu, M. & Hamada, H. (1990) A Novel Octamer Binding Transcription Factor is Differentially Expressed in Mouse Embryonic Cells. Cell, 60(3), 461-472.

Otero, J. J., Fu, W., Kan, L., Cuadra, A. E. & Kessler, J. A. (2004) Beta-Catenin Signaling is Required for Neural Differentiation of Embryonic Stem Cells. Development, 131(15), 3545-3557.

Pash, J. M., Bhorjee, J. S., Patterson, B. M. & Bustin, M. (1990) Persistence of Chromosomal Proteins HMG- 14/- 17 in Myotubes following dependent Reduction of HMG mRNA. Journal of Biological Chemistry, 265(8), 4197-4199.

Phair, R. & Misteli, T. (2000) High Mobility of Proteins in the Mammalian Cell Nucleus. Nature, 404, 604-609.

Postnikov, Y. V. & Bustin, M. (2016) Functional Interplay Between Histone H1 and HMG Proteins in Chromatin. Acta Biochimica et Biophysica, 1859(3), 462-467.

Postnikov, Y. V., Trieschmann, L., Rickers, A. & Bustin, M. (1995) Homodimers of Chromosomal Proteins HMG-14 and HMG-17 in Nucleosome Cores. Journal of Molecular Biology, 252(4), 423-432.

Pradeepa, M. M., Grimes, G. R., Kumar, Y., Olley, G., Taylor, G. C., Schneider, R. & Bickmore, W. A. (2016) Histone H3 Globular Domain Acetylation Identifies a New Class of Enhancers. Nature Genetics, 48(6), 681-686.

Prymakowska-Bosak, M., Misteli, T., Herrera, J. E., Shirakawa, H., Birger, Y., Garfield, S. & Bustin, M. (2001) Mitotic Phosphorylation Prevents the Binding of HMGN Proteins to Chromatin. Molecular and Cellular Biology, 21(15), 5169-5178.

Qian, X., Shen, Q., Goderie, S. K., He, W., Capela, A., Davis, A. A. & Temple, S. (2000) Timing of CNS Cell Generation: A Programmed Sequence of Neuron and Glial Cell Production from Isolated Murine Cortical Stem Cells. Neuron, 28(1), 69-80.

Qiao, Y., Wang, R., Yang, X., Tang, K. & Jing, N. (2015) Dual Roles of Histone H3 Lysine 9 Acetylation in Human Embryonic Stem Cell Pluripotency and Neural Differentiation. Journal of Biological Chemistry, 290(4), 2508-2520.

Ran, F. A., Hsu, P. D., Lin, C. Y., Gootenberg, J. S., Konermann, S., Trevino, A. E., Scott, D. A., Inoue, A., Matoba, S., Zhang, Y. & Zhang, F. (2013) Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. Cell, 154(6), 1380-1389.

Rattner, B. P., Yusufzai, T. & Kadonaga, J. T. (2009) HMGN Proteins Act in Opposition to ATP-dependent Chromatin Remodeling Factors to Restrict Nucleosome Mobility. Molecular Cell, 34(5), 620-626.

Reyes, J. H., Shea, K. S., Wys, N. L., Velkey, J. M., Prieskorn, D. M., Wesolowski, K., Miller, J. M. & Altschuler, R. A. (2008) Glutamatergic Neuronal Differentiation of Mouse Embryonic Stem Cells after Transient Expression of Neurogenin 1 and Treatment with BDNF and GDNF: In Vitro and In Vivo Studies. Journal of Neuroscience, 28(48), 12622-12631.

Ringrose, L. & Paro, R. (2004) Epigenetic Regulation of Cellular Memory by the Polycomb and Trithorax Group Proteins. Annual Review of Genetics, 38(1), 413-443.

Rochman, M., Postnikov, Y., Correll, S., Malicet, C., Wincovitch, S., Karpova, T. S., McNally, J. G., Wu, X., Bubunenko, N. A., Grigoryev, S. & Bustin, M. (2009) The Interaction of NSBP1/HMGN5 with Nucleosomes in Euchromatin Counteracts Linker Histone-Mediated Chromatin Compaction and Modulates Transcription. Molecular Cell, 35(5), 642-656.

Rossant, J. (2015) Mouse and Human Blastocyst-Derived Stem Cells: Vive les Differences. Development, 142(1), 9-12.

Sato, N., Meijer, L., Skaltsounis, L., Greengard, P. & Brivanlou, A. H. (2004) Maintenance of Pluripotency in Human and Mouse Embryonic Stem Cells through Activation of Wnt Signaling by a Pharmacological GSK-3-Specific Inhibitor. Nature Medicine, 10(1), 55-63.

Schöler, H. R., Dressler, G. R., Balling, R., Rohdewohld, H. & Gruss, P. (1990) Oct-4: a Germline-Specific Transcription Factor Mapping to the Mouse T-Complex. EMBO, 9(7), 2185-2195.

Sejersen, T. & Lendahl, U. (1993) Transient Expression of the Intermediate Filament Nestin during Skeletal Muscle Development. Journal of Cell Science, 106(4), 1291-1300.

Shen, Q., Wang, Y., Dimos, J. T., Fasano, C. A., Phoenix, T. N., Lemischka, I. R., Ivanova, N. B., Stifani, S., Morrisey, E. E. & Temple, S. (2006) The Timing of Cortical Neurogenesis is Encoded within Lineages of Individual Progenitor Cells. Nature Neuroscience, 9, 743-751.

Shimojo, H., Ohtsuka, T. & Kageyama, R. (2008) Oscillations in Notch Signaling Regulate Maintenance of Neural Progenitors. Neuron, 58(1), 52-64.

Shirakawa, H., Herrera, J. E., Bustin, M. & Postnikov, Y. (2000a) Targeting of High Mobility Group-14/-17 Proteins in Chromatin Is Independent of DNA Sequence. Journal of Biological Chemistry, 275(48), 37937-37944.

Shirakawa, H., Landsman, D., Postnikov, Y. V. & Bustin, M. (2000b) NBP-45, a Novel Nucleosomal Binding Protein with a Tissue-specific and Developmentally Regulated Expression. Journal of Biological Chemistry, 275(9), 6368-6374.

Sims, R. J., Nishioka, K. & Reinberg, D. (2003) Histone Lysine Methylation: a Signature for Chromatin Function. Trends in Genetics, 19(11), 629-639.

Sindi, A. (2017) Investigating the Role of HMGN2 in the Self-Renewal and Neuronal Differentiation of ECCs using the CRISPR-Cas9 Knockout System. Doctor of Philosophy University of Glasgow.

Smith, A. (2001) Embryo-Derived Stem Cells: Of Mice and Men. Annual Review of Cell and Developmental Biology, 17(1), 435-462.

Smith, A. (2017) Formative Pluripotency: the Executive Phase in a Developmental Continuum. Development, 144(3), 365-373.

Smith, A., Heath, J. K., Donaldson, D. D., Wong, G. G., Moreau, J., Stahl, M. & Rogers, D. (1988) Inhibition of Pluripotential Embryonic Stem Cell Differentiation by Purified Polypeptides. Nature, 336(6200), 688-690.

Smolich, B. D. & Papkoff, J. (1994) Regulated Expression of Wnt Family Members during Neuroectodermal Differentiation of P19 Embryonal Carcinoma Cells: Overexpression of Wnt-1 Perturbs Normal Differentiation-Specific Properties. Developmental Biology, 166(1), 300-310.

Solter, D. & Knowles, B. B. (1978) Monoclonal Antibody Defining a Stage-Specific Mouse Embryonic Antigen (SSEA-1). Proceedings of the National Academy of Sciences of the United States of America, 75(11), 5565-5569.

Stavridis, M. P., Lunn, J. S., Collins, B. J. & Storey, K. G. (2007) A Discrete Period of FGF-Induced Erk1/2 Signalling is Required for Vertebrate Neural Specification. Development, 134(16), 2889-2894.

Stevens, L. C. (1970) The Development of Transplantable Teratocarcinomas from Intratesticular Grafts of Pre- and Postimplantation Mouse Embryos. Developmental Biology, 21(3), 364-382.

Sun, Y., Nadal-Vicens, M., Misono, S., Lin, M. Z., Zubiaga, A., Hua, X., Fan, G. & Greenberg, M. E. (2001) Neurogenin Promotes Neurogenesis and Inhibits Glial Differentiation by Independent Mechanisms. Cell, 104(3), 365-376.

Suzuki, S., Namiki, J., Shibata, S., Mastuzaki, Y. & Okano, H. (2010) The Neural Stem/Progenitor Cell Marker Nestin Is Expressed in Proliferative Endothelial Cells, but Not in Mature Vasculature. Journal of Histochemistry and Cytochemistry, 58(8), 721-730.

Szutorisz, H., Canzonetta, C., Georgiou, A., Chow, C. M., Tora, L. & Dillon, N. (2005) Formation of an Active Tissue-Specific Chromatin Domain Initiated by Epigenetic Marking at the Embryonic Stem Cell Stage. Molecular and Cellular Biology, 25(5), 1804-1820.

Takahashi, K. & Yamanaka, S. (2006) Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. Cell, 126(4), 663-676.

Talbert, P. B. & Henikoff, S. (2010) Histone Variants — Ancient Wrap Artists of the Epigenome. Nature Reviews Molecular Cell Biology, 11, 264-275.

Tang, K., Yang, J., Gao, X., Wang, C., Liu, L., Kitani, H., Atsumi, T. & Jing, N. (2002) Wnt-1 Promotes Neuronal Differentiation and Inhibits Gliogenesis in P19 Cells. Biochemical and Biophysical Research Communications, 293(1), 167-173.

Temple, S. (2001) The Development of Neural Stem Cells. Nature, 414, 112-117.

Tesar, P. J., Chenoweth, J. G., Brook, F. A., Davies, T. J., Evans, E. P., Mack, D. L., Gardner, R. L. & McKay, R. D. (2007) New Cell Lines from Mouse Epiblast share Defining Features with Human Embryonic Stem Cells. Nature, 448(7150), 196-199.

Thoma, E. C., Wischmeyer, E., Offen, N., Maurus, K., Sirén, A. L., Schartl, M. & Wagner, T. U. (2012) Ectopic Expression of Neurogenin 2 Alone is Sufficient to Induce Differentiation of Embryonic Stem Cells into Mature Neurons. Plos One, 7(6), e38651.

Thomson, J. A., Itskovitz-Eldor, J., Shapiro, S. S., Waknitz, M. A., Swiergiel, J. J., Marshall, V. S. & Jones, J. M. (1998) Embryonic Stem Cell Lines Derived from Human Blastocysts. Science, 282(5391), 1145-1147.

Thomson, M., Liu, S. J., Zou, L. N., Smith, Z., Meissner, A. & Ramanathan, S. (2011) Pluripotency Factors in Embryonic Stem Cells Regulate Differentiation into Germ Layers. Cell, 145(6), 875-889.

Trieschmann, L., Martin, B. & Bustin, M. (1998) The Chromatin Unfolding Domain of Chromosomal Protein HMG-14 Targets the N-terminal Tail of Histone H3 in Nucleosomes. Proceedings of the National Academy of Sciences of the United States of America, 95(10), 5468-5473.

Trieschmann, L., Postnikov, Y. V., Rickers, A. & Bustin, M. (1995) Modular Structure of Chromosomal Proteins HMG-14 and HMG-17: Definition of a Transcriptional Enhancement Domain Distinct from the Nucleosomal Binding Domain. Molecular and Cellular Biology, 15(12), 6663-6669.

Tropberger, P., Pott, S., Keller, C., Kamieniarz-Gdula, K., Caron, M., Richter, F., Li, G., Mittler, G., Liu, E. T., Bühler, M., Margueron, R. & Schneider, R. (2013) Regulation of Transcription through Acetylation of H3K122 on the Lateral Surface of the Histone Octamer. Cell, 152(4), 859-872.

Trott, J. & Martinez Arias, A. (2013) Single Cell Lineage Analysis of Mouse Embryonic Stem Cells at the Exit from Pluripotency. Biology Open, 2(10), 1049-1056.

Ueda, T., Catez, F., Gerlitz, G. & Bustin, M. (2008) Delineation of the Protein Module That Anchors HMGN Proteins to Nucleosomes in the Chromatin of Living Cells. Molecular and Cellular Biology, 28(9), 2872-2883.

Ueda, T., Furusawa, T., Kurahashi, T., Tessarollo, L. & Bustin, M. (2009) The Nucleosome Binding Protein HMGN3 Modulates the Transcription Profile of Pancreatic ß Cells and Affects Insulin Secretion. Molecular and Cellular Biology, 29(19), 5264-5276.

Ueda, T., Postnikov, Y. V. & Bustin, M. (2006) Distinct Domains in High Mobility Group N Variants Modulate Specific Chromatin Modifications. The Journal of Biological Chemistry, 281(15), 10182-10187.

Velkey, J. M. & O'Shea, K. S. (2013) Expression of Neurogenin 1 in Mouse Embryonic Stem Cells Directs the Differentiation of Neuronal Precursors and Identifies Unique Patterns of Down-stream Gene Expression. Developmental Dynamics, 242(3), 230-253.

Vossaert, L., Meert, P., Scheerlinck, E., Glibert, P., Van Roy, N., Heindryckx, B., De Sutter, P., Dhaenens, M. & Deforce, D. (2014) Identification of Histone H3 Clipping Activity in Human Embryonic Stem Cells. Stem Cell Research, 13(1), 123-134.

Wang, C., Xia, C., Bian, W., Liu, L., Lin, W., Ang, S. L. & Jing, N. (2006) Cell Aggregation-induced FGF8 Elevation Is Essential for P19 Cell Neural Differentiation. Molecular Biology of the Cell, 17, 3075-3084.

Wang, Z., Zang, C., Cui, K., Schones, D. E., Barski, A., Peng, W. & Zhao, K. (2009) Genome-wide Mapping of HATs and HDACs Reveals Distinct Functions in Active and Inactive Genes. Cell, 138(5), 1019-1031.

Ware, C. B., Wang, L., Mecham, B. H., Shen, L., Nelson, A. M., Bar, M., Lamba, D. A., Dauphin, D. S., Buckingham, B., Askari, B., Lim, R., Tewari, M., Gartler, S. M., Issa, J. P., Pavlidis, P., Duan, Z. & Blau, C. A. (2009) Histone Deacetylase Inhibition Elicits an Evolutionarily Conserved Self-Renewal Program in Embryonic Stem Cells. Cell Stem Cell, 4(4), 359-369.

Weintraub, H., Flint, S. J., Leffak, I. M., Groudine, M. & Grainger, R. M. (1978) The Generation and Propagation of Variegated Chromosome Structures. Cold Spring Harbor Symposia on Quantitative Biology, 42, 401-407.

Weintraub, H. & Groudine, M. (1976) Chromosomal Subunits in Active Genes Have an Altered Conformation. Science, 193(4256), 848-856.

Weisbrod, S., Groudine, M. & Weintraub, H. (1980) Interaction of HMG 14 and 17 with Actively Transcribed Genes. Cell, 19(1), 289-301.

Weisbrod, S. & Weintraub, H. (1979) Isolation of a Subclass of Nuclear Proteins Responsible for Conferring a DNase I-Sensitive Structure on Globin Chromatin. Proceedings of the National Academy of Sciences of the United States of America, 76(2), 630-634.

West, K. L., Ito, Y., Birger, Y., Postnikov, Y., Shirakawa, H. & Bustin, M. (2001) HMGN3a and HMGN3b, Two Protein Isoforms with a Tissue-specific Expression Pattern, Expand the Cellular Repertoire of Nucleosome-Binding Proteins. Journal of Biological Chemistry, 276(28), 25959-25969.

Whitehouse, I., Rando, O. J., Delrow, J. & Tsukiyama, T. (2007) Chromatin Remodelling at Promoters Suppresses Antisense Transcription. Nature, 450, 1031-1035.

Wilson, S. I. & Edlund, T. (2001) Neural Induction: toward a Unifying Mechanism. Nature Neuroscience, 4 Suppl, 1161-1168.

Wilson, S. W. & Houart, C. (2004) Early Steps in the Development of the Forebrain. Developmental Cell, 6(2), 167-181.

Winner, B., Kohl, Z. & Gage, F. H. (2011) Neurodegenerative Disease and Adult Neurogenesis. European Journal of Neuroscience, 33(6), 1139-1151.

Wray, J., Kalkan, T. & Smith, A. (2010) The Ground State of Pluripotency. Biochemical Society Transactions, 38(4), 1027-1032.

Wroblewski, J., Engström, M., Edwall-Arvidsson, C., Sjöberg, G., Sejersen, T. & Lendahl, U. (1997) Distribution of Nestin in the Developing Mouse Limb Bud In Vivo and in Micro-Mass Cultures of Cells Isolated from Limb Buds. Differentiation, 61(3), 151-159.

Wurst, W. & Bally-Cuif, L. (2001) Neural Plate Patterning: Upstream and Downstream of the Isthmic Organizer. Nature Reviews Neuroscience, 2, 99-108.

Xu, C. R., Cole, P. A., Meyers, D. J., Kormish, J., Dent, S. & Zaret, K. S. (2011) Chromatin "Prepattern" and Histone Modifiers in a Fate Choice for Liver and Pancreas. Science, 332(6032), 963-966.

Yellajoshyula, D., Lim, J., Thompson, D. M., Witt, J. S., Patterson, E. S. & Kroll, K. L. (2012) Geminin Regulates the Transcriptional and Epigenetic Status of Neuronal Fate-Promoting Genes during Mammalian Neurogenesis. Molecular and Cellular Biology, 32(22), 4549-4560.

Ying, Q. L., Nichols, J., Chambers, I. & Smith, A. (2003a) BMP Induction of Id Proteins Suppresses Differentiation and Sustains Embryonic Stem Cell Self-Renewal in Collaboration with STAT3. Cell, 115(3), 281-292.

Ying, Q. L., Stavridis, M., Griffiths, D., Li, M. & Smith, A. (2003b) Conversion of Embryonic Stem Cells into Neuroectodermal Precursors in Adherent Monoculture. Nature Biotechnology, 21(2), 183-186.

Ying, Q. L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P. & Smith, A. (2008) The Ground State of Embryonic Stem Cell Self-Renewal. Nature, 453(7194), 519-523.

Yuan, H., Corbi, N., Basilico, C. & Dailey, L. (1995) Developmental-Specific Activity of the FGF-4 Enhancer Requires the Synergistic Action of Sox2 and Oct-3. Genes and Development, 9, 2635-2645.

Zhang, S., Zhu, I., Deng, T., Furusawa, T., Rochman, M., Vacchio, M. S., Bosselut, R., Yamane, A., Casellas, R., Landsman, D. & Bustin, M. (2016) HMGN Proteins Modulate Chromatin Regulatory Sites and Gene Expression during Activation of Naive B Cells. Nucleic Acids Research, 44(15), 7144-7158.

Zhao, W., Ji, X., Zhang, F., Li, L. & Ma, L. (2012) Embryonic Stem Cell Markers. Molecules, 17(6), 6196-6236.

Zhong, X. & Jin, Y. (2009) Critical Roles of Coactivator p300 in Mouse Embryonic Stem Cell Differentiation and Nanog Expression. Journal of Biological Chemistry, 284(14), 9168-9175.