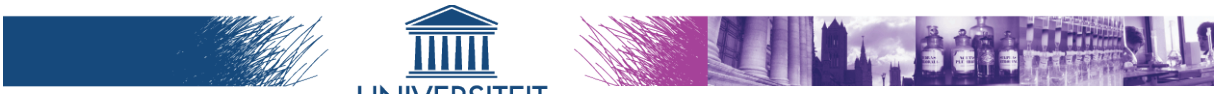


HISTONE H3 CLIPPING IN HUMAN EMBRYONIC STEM CELLS: IN PURSUIT OF AN EPIGENETIC OUTCAST

LIESBETH VOSSAERT

Laboratory for Pharmaceutical Biotechnology
Faculty of Pharmaceutical Sciences
Ghent University

Promotor: prof. dr. apr. Dieter Deforce
Co-promotor: prof. dr. Petra De Sutter and dr. Maarten Dhaenens

Thesis submitted in the fulfillment of the requirements for the degree of Doctor in Pharmaceutical Sciences
Academic year 2014 – 2015

Members of the reading committee:

prof. dr. apr. Christophe Stove, chairman
Laboratory of Toxicology, Faculty of Pharmaceutical Sciences, Ghent University

prof. dr. apr. Dieter Deforce
Laboratory for Pharmaceutical Biotechnology, Faculty of Pharmaceutical Sciences, Ghent University

prof. dr. Björn Heindryckx
Department for Reproductive Medicine, Ghent University Hospital

prof. dr. ir. Jo Vandesompele
Center for Medical Genetics, Ghent University

prof. dr. Wim Vanden Berghe
Department of Biomedical Sciences, Faculty of Pharmaceutical, Biomedical and Veterinary Sciences, University of Antwerp

Members of the examination committee:

prof. dr. Petra De Sutter
Department for Reproductive Medicine, Ghent University Hospital

dr. Maarten Dhaenens
Laboratory for Pharmaceutical Biotechnology, Faculty of Pharmaceutical Sciences, Ghent University

prof. dr. Paul Coucke
Center for Medical Genetics, Ghent University Hospital

prof. dr. Sebastien Carpentier
Department of Biosystems, Faculty of Bioscience Engineering, University of Leuven

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Ab	Antibody
ACTB	*β-Actin*
AEBSF	4-(2-Aminoethyl)benzenesulfonyl fluoride hydrochloride
ANOVA	Analysis of variance
B2M	*β-2-microglobulin*
bFGF / FGF-2	Basic fibroblast growth factor
BMP	Bone morphogenetic protein
BSA	Bovine serum albumin
CAPS	*N*-cyclohexyl-3-aminopropanesulfonic acid
cDNA	Complementary DNA
CENP-A	Histone H3-like centromeric protein A
cH3	Clipped histone H3
CpG	Cytosine-phosphate-Guanine
Cq / Ct	Quantification cycle / cycle threshold
ddPCR	Droplet digital PCR
DMEM	Dulbecco's modified Eagle medium
DNMT	DNA methyltransferase
dsDNA	Double stranded DNA
E8	Essential 8
EB	Embryonic body
EDTA	Ethylenediaminetetraacetic acid
eGFP	Enhanced green fluorescent protein
ESC	Embryonic stem cells
ESI	Electrospray ionization
FAM	Fluorescein amidite
FRET	Fluorescence resonance energy transfer
GAPDH	*Glyceraldehyde-3-phosphate dehydrogenase*
gDNA	Genomic DNA
GOI	Gene of interest
H3.cs1 Ab	Histone H3 cleavage site 1 Ab (antigen alanine 21)
H3K4me3	Histone H3 lysine 4 trimethylation
H3K27me3	Histone H3 lysine 27 trimethylation
HA	Hydroxylamine
HEPES	4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid
hESC	Human embryonic stem cells
hiPSC	Human iPSC
HMT	Histone methyltransferase
HP1	Heterochromatin protein 1
HRP	Horseradish peroxidase
ICM	Inner cell mass
iPSC	Induced Pluripotent stem cells
IVF	*In vitro* fertilization

KLF4	Kruppel-like factor 4
LC	Liquid chromatography
LIF	Leukemia inhibitory factor
lncRNA	Long non-coding RNA
MEF	Mouse embryonic fibroblasts
mESC	Mouse embryonic stem cells
MIQE	Minimum information for publication of quantitative Real-Time PCR experiments
miRNA	Micro RNA
mRNA	Messenger RNA
MS	Mass spectrometry
MS/MS	Tandem MS
MW	Molecular weight
m/z	Mass over charge
ncRNA	Non-coding RNA
NE	Nuclear extract
NET	Neutrophil extracellular traps
Oct4	Octamer-binding transcription factor 4
PBS	Phosphate buffered saline
PcG	Polycomb Group proteins
PCR	Polymerase chain reaction
PI	Propidium iodide
PIC	Protease inhibitor cocktail
PICMI	Post-inner cell mass intermediate
PMA	Phorbol 12-myristate 13-acetate
POU5F1	*POU class 5 homeobox 1*
PPIA	*Peptidylprolyl isomerase A*
PRC1, 2	Polycomb repressive complex
PTM	Posttranslational histone modification
PVDF	Polyvinylidene fluoride
Q-TOF	Quadrupole – Time-of-flight
RA	Retinoic acid
RNAPII	RNA polymerase II
ROX	6-carbonyl-X-rhodamine
RPL13A	*Ribosomal protein L13A*
RPLC	Reversed phase LC
RPMI	Roswell Park Memorial Institute
RQI	RNA quality index
rRNA	Ribosomal RNA
RT-qPCR	Reverse transcription – quantitative PCR
SDS	Sodium dodecyl sulphate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
siRNA	Short interfering RNA
Sox2	(Sex determining region Y)-Box 2

SSEA1, 3, 4	Stage-specific embryonic antigen 1, 3, 4
SUMO	Small ubiquitin-like modifier
TAMRA	Tetramethylrhodamine
Taq	*Thermus aquaticus*
TEB	Triton-X extraction buffer
TET	Ten-eleven translocation methylcytosine dioxygenase
Tm	Melting temperature
Tris	Tris(hydroxymethyl)aminomethane
trxG	Trithorax Group proteins
UTR	Untranslated region
5caC	5-carbonylcytosine
5fmC	5-formylcytosine
5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine

PREFACE:

AIMS AND OVERVIEW

PREFACE

This thesis merges two fascinating, relatively new research areas: **human embryonic stem cell biology** and **epigenetics**. Despite the thousands of reports that have been published up to now on both these subjects, numerous questions and research opportunities still remain. One such question is addressed in this dissertation: Is histone proteolysis an important part of stem cell biology?

Embryonic stem cells (ESC) have proven to serve as an excellent model for studying fundamental biological processes involved in the early development of (human) organisms. Equally important, they served as a template for the generation of induced pluripotent stem cells (iPSC) and will contribute to a better (clinical) application of these cells in the future. The awarding of the 2012 Nobel prize in Physiology or Medicine to John B. Gurdon and Shinya Yamanaka illustrates the importance of this research and its place in the scientific community. These two scientists were honored for discovering the potential of specialized cells to be reprogrammed to immature cells capable of differentiating towards all cell types of the body [1]. Despite the hurdles in the reprogramming process, the iPSC safety issues and the overall question on ESC and iPSC *in vivo* relevance, stem cells have yet proven useful in regenerative medicine, modeling diseases and candidate-drug screenings, and will continue to play an important role as such [2].

Also in the field of epigenetics, which studies gene expression regulatory mechanisms that are not based on DNA sequence alterations, an extending number of challenges and applications lies ahead. Aberrant recruitment of epigenetic modifiers amongst other factors, is involved in a number of pathologies, primarily cancer. Because of the potential reversibility of epigenetic marks such as histone modifications, they are promising therapeutic targets when dealing with human pathologies. This is illustrated by the approval of several so-called epi-drugs by the US Food and Drugs Administration and the European Medicines Agency for cancer treatment and the various other candidate drugs that are undergoing clinical trials for both mono- or combination therapy [3].

The aim of this thesis was to dig deeper into the field of **regulated histone proteolysis** as a posttranslational histone modification (PTM) in human ESC (hESC), thus combining the two aforementioned themes.

The first section of **Chapter 1** gives a broad introduction on hESC biology, culture methods and differentiation possibilities. Subsequently the epigenetics part of this thesis is launched by generally introducing the term, defining the different levels of epigenetic regulation and finally submitting this back to a hESC setting. Upon hESC differentiation, the epigenetic signature shifts drastically, leading to equally drastic changes in transcription. One of the possible techniques to monitor such changes is

reverse transcription-quantitative PCR (RT-qPCR). In **Chapter 2**, an optimization for RT-qPCR analysis of samples obtained from differentiating hESC is described. More specifically, a panel of suitable reference loci is provided for correct normalization of RT-qPCR data.

Chapter 3 and **Chapter 4** describe in more detail our findings on the presence of histone H3 clipping in hESC. While this event has been reported in many distinct biological settings, the true role of this modification however remains unclear. Notwithstanding its potential epigenetic function as a PTM that sweeps away *en masse* all other modifications, histone clipping is not picked up by the broader scientific community, although nowadays over 5000 papers are published on histones every year. In **Chapter 5** we thus discuss its current position in biology, while also reviewing the different reports on histone proteolysis in the past. This thesis adds one more piece to the puzzle by demonstrating for the first time that histone clipping also occurs in hESC, in which basic processes as self-renewal, pluripotency and lineage programming are highly epigenetically determined.

CHAPTER 1:

INTRODUCTION:

HUMAN EMBRYONIC STEM CELLS &

EPIGENETICS INTERTWINED

1. INTRODUCTION

1.1. Human embryonic stem cells

1.1.1. Establishment and characterization of embryonic stem cells

Human embryonic stem cells (hESC) are hallmarked by their prolonged proliferation potential and their capacity to differentiate towards all cell types of the human body.

Human ESC lines are derived from the blastocyst's inner cell mass (ICM), usually from 'spare' blastocyst-stage embryos that were in first instance intended for clinical use. In case of infertility issues, patients can opt for assisted reproductive technologies such as *in vitro* fertilization (IVF), intra uterine insemination and intracytoplasmic sperm injection. During the IVF procedure, the embryos are kept in culture for several days until they reach the blastocyst-stage. On day 5-6, the quality of the blastocysts is scored (e.g. according to the Stephenson grading system [4]), thus determining whether an embryo is suitable for being transferred back to the patient or not. Good quality embryos that are left unused after successful implantation and that do not have to be cryopreserved for any future pregnancy attempts, can be donated for research with informed consent from the patient. Also the embryos of poor quality, insufficient for re-implantation, can serve as a source for the derivation of new hESC lines.

Once the blastocyst-stage has been reached, five days after fertilization, the embryo consists of an ICM, surrounded by trophectoderm cells and the blastocoel cavity. *In vivo*, the ICM (or embryoblast) gives rise to the eventual embryo, whilst the trophectoderm develops into extra-embryonic tissues. For cell line derivation, the blastocyst as a whole or the isolated ICM can be plated on a mouse embryonic fibroblast (MEF) feeder layer or another adherent substrate for culturing. Via a post-ICM intermediate (PICMI) outgrowth formed within 7 days after plating, the ICM evolves towards an established ESC line (Figure 1.1) [5]. As such, the first human ESC line was derived in 1998 by J. Thomson [6].

The quality of a hESC line can be monitored via a range of characteristics. Morphologically, undifferentiated hESC colonies in culture appear round and flat-growing, and they have well-defined edges (Figure 1.2A-B). The individual cells have a high nucleus-to-cytoplasm ratio (Figure 1.2C). Pluripotent hESC also display an abbreviated cell cycle (16 h) compared to differentiated cells, due to a shortened G1 phase and thus have a high proliferation rate [7, 8].

For more detailed investigation of their molecular signature, the expression of cell surface markers such as the glycolipid antigens stage-specific embryonic antigen (SSEA)-3 and -4, and the keratin sulfate antigens TRA-1-60 and TRA-1-81 is often evaluated by immunofluorescence (microscopy and/or flow cytometry) as a marker for undifferentiated hESC [9–11]. The specific transcription factors

Oct4 (encoded by *POU5F1*), Nanog (*NANOG*) and Sox2 (*SOX2*) amongst others, are essential in early development and in propagation of the undifferentiated state, and share many target genes. The expression of these pluripotency factors can be quantitatively monitored via RT-qPCR or microarray assays [9, 10, 12] (*optimization for RT-qPCR analysis described in Chapter 2* [13]). Also the elevated activity of telomerase and alkaline phosphatase is a characteristic feature [6, 9, 10, 14]. Besides the evaluation of specific markers and expression profiling, also a cytogenetic analysis needs to be carried out in order to monitor the genetic stability over time. By evaluating the karyotype, usually obtained via G-banding, chromosomal abnormalities can be revealed (Figure 1.2E) [11]. Additionally, the typical epigenetic profile of hESC can be assessed too, which is described in more detail further in this chapter (*Section 1.3*).

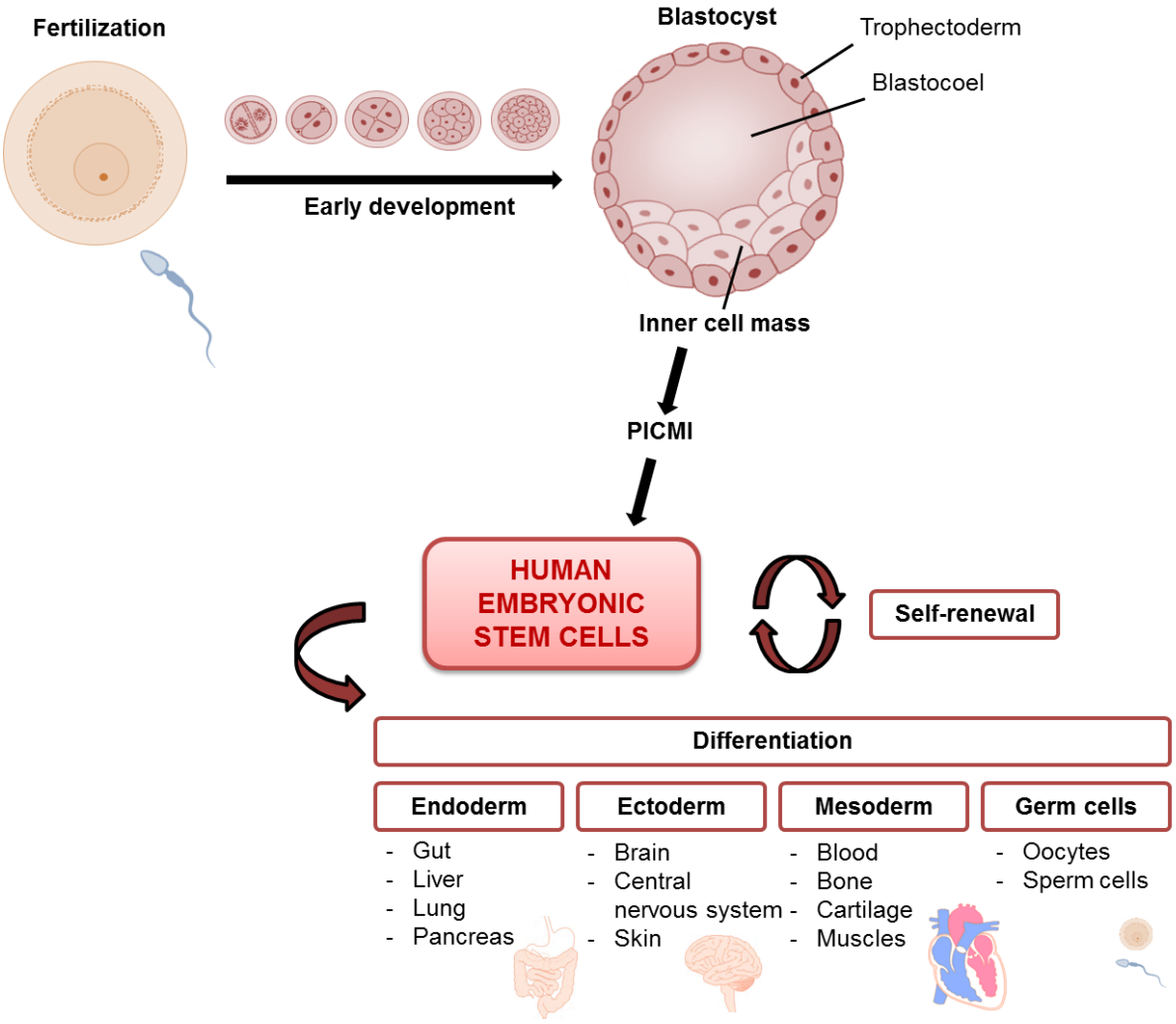


Figure 1.1. Development of a blastocyst and derivation of human embryonic stem cells. After *in vitro* blastocyst maturation, hESC are derived from the blastocyst's inner cell mass. They are characterized by the potential for self-renewal and their capacity to differentiate towards all cell types of the three germ layers, their respective progeny and germ cells.

Besides the aforementioned features, also the differentiation potential of hESC can be evaluated via different tests. The capability to form embryonic bodies *in vitro*, spherical clusters spontaneously formed by hESC in suspension, is used as a criterion for pluripotency and thus hESC line quality [10, 11, 15]. Additionally, a range of *in vitro* differentiation protocols are readily available to direct the hESC towards a particular lineage. Via expression-analysis of cell type-specific markers the generation of cells from all three germ layers can be followed [10, 11]. Furthermore, also the *in vivo* differentiation capacity can be assessed. As the formation of chimeras is obviously not possible for human cells, teratoma formation in immunocompromised animals is considered to be the most stringent method for measuring pluripotency [10, 11, 15]. In order to maintain the self-renewing hESC status and inhibit spontaneous differentiation, specific and adjusted culture methods are required as outlined below.

1.1.2. Human embryonic stem cell culture systems

HESC can be maintained in culture both on a feeder layer or in feeder-free conditions. The most commonly used feeder layer consists of MEF, although also human feeders have been applied [16–18]. To preserve the hESC from being overgrown by the feeder cells, the latter are mitotically inactivated with e.g. mitomycin C, a DNA-crosslinker which inhibits DNA synthesis and cell division. Feeder cell density, the time span of culturing and the MEF exposure time to mitomycin C are determining for the feeder's ability to support hESC growth [19, 20]. The feeders sustain the self-renewing, pluripotent hESC status by secretion of factors such as basic fibroblast growth factor (bFGF or FGF-2), transforming growth factor β, activin A and antagonists of the bone morphogenetic protein (BMP) signaling pathway, all of which have proven to be involved in hESC self-renewal. The downside is that the exact composition of that secretome is often unknown and batch-to-batch variability is thus easily induced. On top of that, working with feeders makes the culturing protocol rather time-consuming and labor-intensive.

As we want to move away from these undefined and variable circumstances, culture protocols have been adapted during the last couple of years towards better determined conditions, preferably feeder-free and definitely xeno-free [15]. The exclusion of xenogenic contaminants is unquestionably crucial when moving on to implementation of hESC and hiPSC in the clinic e.g. for transplantation purposes and regenerative medicine [21, 22]. The use of Matrigel as hESC support was a first step in adjusting culture methods [23], although it is still a mixture of extracellular matrix constituents of mouse origin. By taking a closer look at the responsible components supporting the hESC, several recombinant proteins were tested as adherent substrate; as such e.g. vitronectin proved to be an optimal substrate for maintaining the undifferentiated hESC state [15, 24] (illustrated in Figure 1.2B).

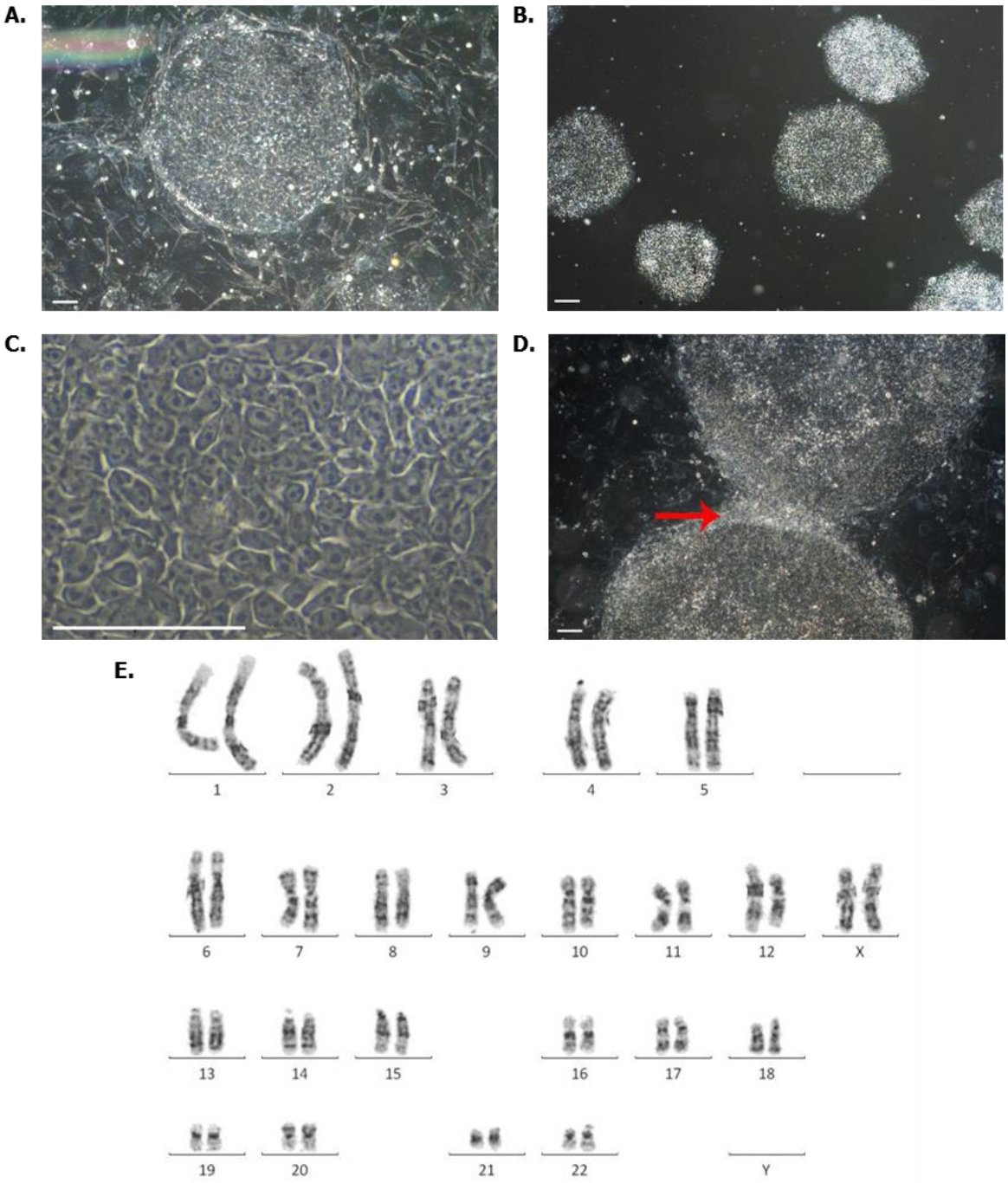


Figure 1.2. Human embryonic stem cells in culture. Undifferentiated hESC grow in round, flat colonies with well-defined borders, as can be seen in **A:** a culture on a MEF feeder layer or in **B:** feeder-free culture (with vitronectin coating). **C:** The high nucleus-to-cytoplasm ratio of the individual cells becomes clear when studying a hESC colony in more detail. **D:** Differentiation sets in at the contact area of different colonies (indicated with red arrow), where subsequently cells will pile-up and acquire a more lengthened shape (All scale bars = 100 µm). **E:** Cytogenetic analysis reveals a normal karyotype for the UGENT2 cell line. All images shown were gathered throughout the research for this dissertation.

The concomitant culture media originally contained fetal bovine serum, which suffers from batch-to-batch variability and also has the disadvantage of being of animal origin. Serum products were

switched to serum-replacements over time, but recently even less complex and more specialized media have been introduced, e.g. Essential 8 (E8) medium accompanying vitronectin-based feeder-free culture [25, 26]. The component that has remained constant during the fine tuning of cell culture media was the addition of bFGF, a molecule part of the FGF/extracellular signal-regulated kinase signaling pathway which promotes the self-renewing state. This factor is considered as an indispensable component for keeping hESC undifferentiated, although a wide range of applied concentrations is found in literature [27].

ESC culture is characterized by the need to refresh the feeder layer or feeder-free substrate on at least weekly basis, in order to sustain pluripotency and prevent overgrowing colonies with subsequent spontaneous *in vitro* differentiation (Figure 1.2D) [11, 15]. Depending on the culture methodology, a different passage technique is employed. Originally, hESC were mechanically passaged by dissociating the mature colonies into small cell clumps with a sharpened needle [15]. As this is however not a feasible method when scaling up cultures, enzymatic dissociation has been implemented as an alternative approach. Frequently used agents are collagenase IV, dispase or other proteolytic enzymes and combinations thereof [11]. In feeder-free cultures, non-enzymatic detachment is achieved by chelation of calcium and magnesium with EDTA [15]. Of note, cell survival after dissociation into single cells is very low and might also cause aneuploidies; hence, irrespective of the passaging method applied, it is essential that the hESC remain in small clumps during the splitting procedure [11, 15].

Additionally, oxygen exposure appears to be an important parameter to be controlled for in hESC culture: the capacity to maintain the undifferentiated hESC state is believed to be enhanced under 5 % oxygen growth conditions compared to atmospherical oxygen pressure (20 %) [28–30].

1.1.3. ESC differentiation

The level of differentiation potential can be classified in different potency-types [31]. The totipotent/omnipotent zygote is capable of forming both embryonic and extra-embryonic tissues, and develops further into the blastocyst. The ICM and the embryonic stem cells derived thereof are pluripotent, thus they can give rise to all cell types of the three germ layers but no longer to extra-embryonic tissues (*see Section 1.1.1*). As development continues, the differentiation capacity gets progressively more restricted. The subsequent stem cells are multipotent and are able to develop into the various cells of a few lineages. The probably best studied example in this category is the hematopoietic stem cell. Further specialization generates oligopotent stem cells or progenitors, only competent to differentiate into a few cell types of a particular lineage. The last step before reaching the differentiation-terminus consists of specific precursor cells (unipotent).

Besides the spontaneous differentiation in hESC culture [30, 32], a variety of protocols and differentiation agents are available to push development in general as well as towards a specific cell lineage (non-exhaustive overview in Figure 1.3) [33]. As such, ascorbic acid is e.g. applied to direct hESC differentiation into cardiac myocytes [34]. Chen and colleagues succeeded in pushing hESC into the pancreatic lineage via indolactam V addition [35]. BMP4-supplementation after hESC derivation in the presence of Activin A enhances the competence to differentiate towards primordial germ cell-like cells [36]. Another widely applied differentiating agent, also in the research conducted for this thesis, is all-trans retinoic acid (RA) [37], of which the resulting cell types are concentration- and exposure time-dependent [38].

A first feature of differentiation is the morphological change that takes place. Although the final outcome is dependent on the differentiation protocol applied, differentiation starts in general with the loss of the round shape of the hESC colonies and their well-defined borders. The morphology of the individual cells alters drastically and cells begin to pile up at the colony periphery where different colonies get into contact. Also in the colony center, cells start to accumulate.

The hESC cell cycle lengthens upon lineage specialization as the cells acquire an elongated G1 phase [39, 40], a phenomenon also readily seen in the differentiation experiments conducted for this thesis (Figure 1.4). When comparing pluripotency-promoting culture conditions on one hand and induced differentiation on the other, flow cytometry analysis with propidium iodide (PI) shows that the proportion of cells present in the G1 phase clearly increases in differentiation-allowing conditions and this occurs earlier in forced than in spontaneous differentiation.

Besides morphological and cell cycle changes, all characteristic features of unspecialized hESC described in Section 1.1.1 also change considerably. The surface antigens SSEA-3, SSEA-4, TRA-1-60 and TRA-1-81 become downregulated, whilst differentiation markers such as SSEA-1 are upregulated [41]. Alkaline phosphatase expression diminishes [41] and the pluripotency factors' expression levels, e.g. of Oct4, decrease [42]. Of note, the formation of Oct4-positive "islands" within each colony has been detected during RA-induced differentiation of hESC cultured on MEF [43, 44]. However, the biology behind this colony heterogeneity remains to be elucidated. Furthermore, also from an epigenetic point of view radical changes occur (*described in more detail in Section 1.3*).

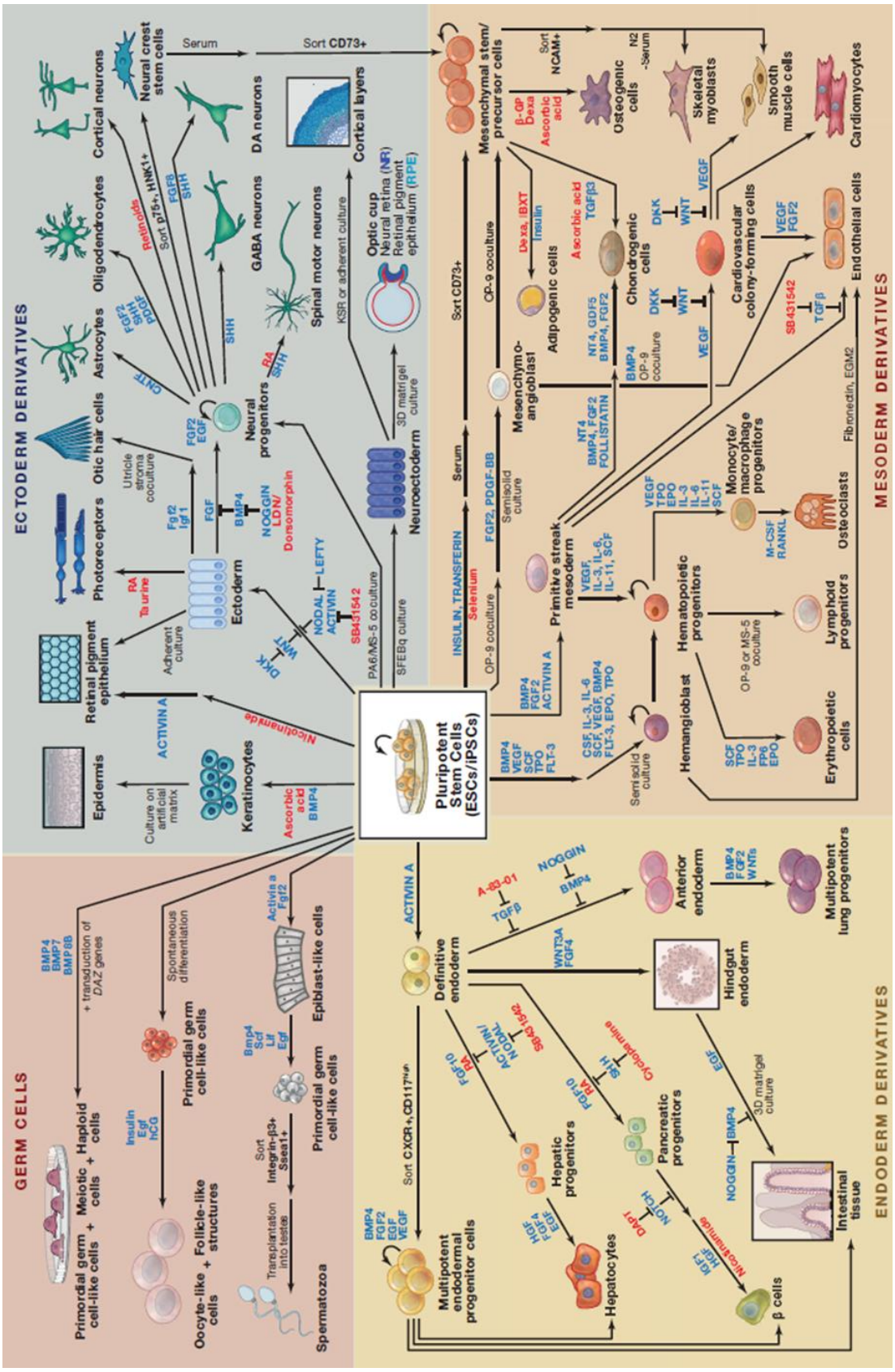


Figure 1.3. Directed differentiation of pluripotent stem cells. This figure illustrates the broad spectrum of differentiation agents that can be applied for induction of lineage specification (image taken from [233]).

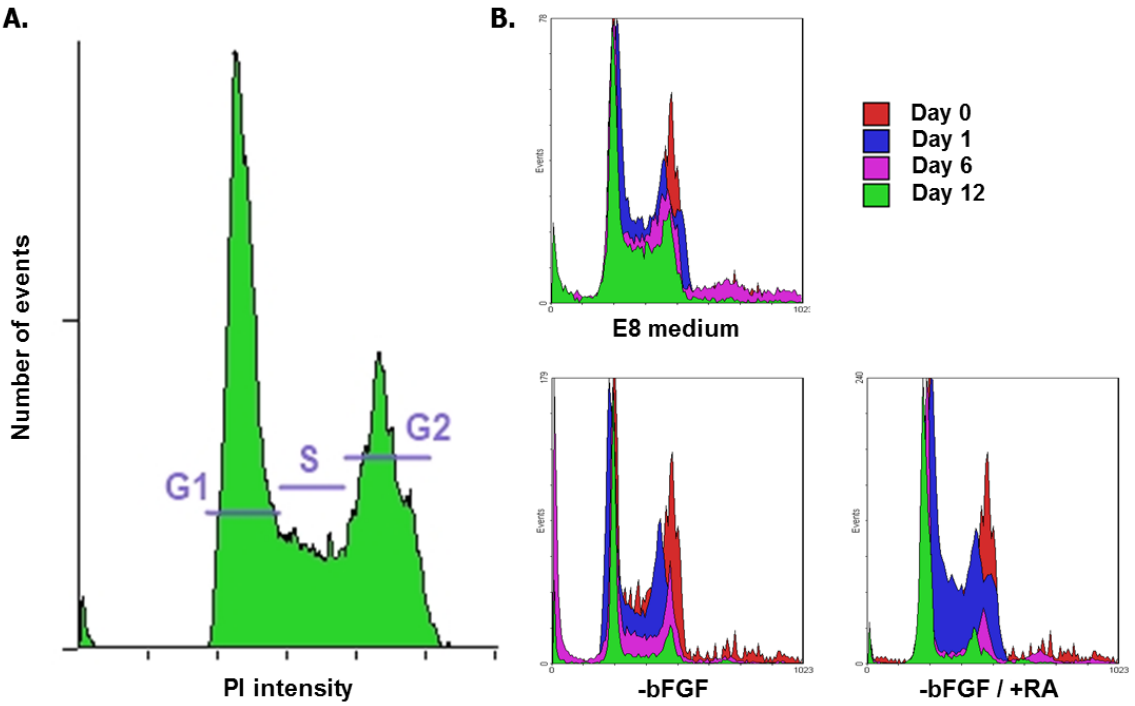


Figure 1.4. Cell cycle analysis of hESC. A: Flow cytometric analysis after PI staining displays the cell population's distribution over the different cell cycle phases. The PI intensity is directly correlated to the amount of DNA present in the cell; hence, the intensity of the signal shifts from G2 > S > G1. **B:** Illustration of the lengthening cell cycle due to an increasing number of cells in the G1 phase. When comparing three different cultures (E8 medium = pluripotent cells; -bFGF = spontaneous differentiation; -bFGF / +RA = forced differentiation) the relative proportion of the S and G2 phase evidently decreases when differentiation is launched in comparison to cells maintained in the pluripotency-preserving E8 medium. Samples were taken on day 0 and day 1, 6 and 12 after onset of differentiation.

1.1.4. Human versus mouse: Primed versus naïve

Whether hESC and mESC really represent the true pluripotent state is a question still open for debate. Recent findings suggest that hESC resemble more mouse epiblast stem cells than mESC, and that they thus represent a more developed, more heterogeneous, primed pluripotent state [45, 46]. Mouse ESC are considered to be more naïve and ground state pluripotent.

Naïve pluripotent stem cells are more homogeneous in terms of gene expression than primed populations (both within and amongst different cell lines) and are less 'biased' towards a certain lineage specification. Hence, they are believed to be better controllable during differentiation experiments. Moreover, they can be passaged as single cells and thus allow better for bulk culture.

The first mESC derivations date back to 1981 [47, 48]. They have several features in common with hESC. Mouse ESC are also rapidly proliferating and they do display a similar abbreviated cell cycle comparable to hESC, which also lengthens upon differentiation [49]. However, the observed differences indicate that caution needs to be taken when comparing data from both cell types [50].

Already in terms of morphology, mESC tend to grow more dome-shaped in comparison to the flat hESC colonies [45]. Unlike hESC, which are bFGF-dependent, mESC rely on Leukemia inhibitory factor (LIF)-supplementation in the culture medium to sustain their pluripotent, self-renewing status and they thus maintain a slightly different signaling network [51]. Although they rely on the same core pluripotency transcription factor network (Oct4, Nanog and Sox2), the expression of several surface antigens differs: mESC do express SSEA-1 in pluripotent state, but lack SSEA-3, SSEA-4, TRA-1-60 and TRA-1-81 [50]. The same holds for the levels of other specific differentiation markers [50]. Finally, the epigenetic signature of hESC and mESC is dissimilar; e.g. female mESC lines have two active X chromosomes, whilst in hESC one of the two is epigenetically silenced [46].

Only recently several labs succeeded in also deriving naïve hESC [52–54]. This naïve pluripotency, which is believed to resemble more the status of the ICM cells [46], obviously holds many research opportunities for the future.

1.1.5. Inducing pluripotency

The study of naïve pluripotency is increasingly being complemented by the research on induced pluripotent stem cells. In 2006, Yamanaka and co-workers established the first iPSC from mouse fibroblasts, by means of a four factor-mix [55]. The introduction of *OCT4*, *SOX2*, *c-MYC* and *KLF4* into somatic cells generated cells that exhibited the same morphology, growth properties and marker genes as ESC. In 2007, they applied the same cocktail to create the first human iPSC, starting from adult dermal fibroblasts [56]. Simultaneously, the Thomson lab published the production of hiPSC with a combination of *OCT4*, *SOX2*, *NANOG* and *LIN28* [57]. Since then, several attempts have been made to improve the protocol and enhance reprogramming efficiency e.g. by the inclusion of small molecules. As such, hiPSC have yet been generated from an assortment of different somatic cell types [58–60].

Human ESC are grateful subjects for basic developmental research and hold great potential for (future) applications in regenerative medicine. However, they inevitably bring along ethical and scientific issues, to which iPSC could provide a solution. The use of human embryos, even very early ones, is still a touchy topic for discussion but could be made irrelevant when applying iPSC. Another advantage of iPSC is that, for disease modeling and regenerative purposes (e.g. [61–64]), the starting material for reprogramming can be directly taken from the patient, thus (theoretically) overcoming the rejection issues that accompany transplantation.

However, the potential application of iPSC clearly still has to overcome many hurdles before they can be put to use in an actual clinical setting [65]. Human iPSC lines apparently are more susceptible to

DNA duplications, mutations and rearrangements than hESC lines. Perhaps even more important is the tumorigenicity of iPSC. Of the four Yamanaka factors used for reprogramming, *c-MYC* enhances the still rather low reprogramming efficiency substantially but is also a well-known oncogene. Although efforts have been made to improve the composition of the reprogramming cocktail and to establish iPSC without chromosomal integration, many publications still describe the administration of the original transcription factor combination, thus still carrying the added risk of tumor formation [65, 66].

Thus, whether iPSC could functionally replace ESC remains a subject of discussion. There are plenty of reports both accentuating the differences between those two cell types or stating that iPSC could replace ESC research [67, 68]. Although the latter would be eagerly welcomed out of an ethical point of view, even Yamanaka himself pointed out that ESC are still the best standard to which iPSC developments should be compared [67].

1.2. Epigenetics

1.2.1. The term "Epigenetics"

The central dogma in molecular biology states that DNA is transcribed into RNA, which in turn is further translated to proteins. All cells within one individual comprise the same DNA with >20.000 protein coding genes, but epigenetic mechanisms regulate their transcription and thus sort these cells into >200 cell types [69]. "Epigenetics" (literally "on top of genetics") was termed for the first time in the 1940's [70, 71], and serves as the connecting link between "genotype" and "phenotype". The resulting "epigenotype" is thus defined as a network of developmental processes where both genetic and environmental changes are brought together, towards phenotypic development. Epigenetic regulatory instructions on top of the genomic information are found on both the DNA, the RNA and the histone level (Figure 1.5) [72]. Epigenomics thus explores in depth the mitotically and/or meiotically inherited variation of gene expression based on changes in histone code, DNA methylation and non-coding RNA mechanisms, instead of changes in the primary DNA sequence [73, 74].

Nucleosomes, the basic DNA packaging elements, capture (environmental) stimuli via histone modifications and other mechanisms, which collaborate to regulate gene expression [74]. As such, two chromatin forms can be distinguished: euchromatin and heterochromatin, each accompanied by their own characteristic modifications. Heterochromatin is the highly compacted form where only low levels of transcriptional activity are found, whereas in euchromatin the nucleosomes are much less tightly compacted and display clear traits of high gene expression.

1.2.1. Non-histone epigenetics

This thesis focusses exclusively on histone epigenetics, though for completeness, also the other mechanisms are briefly described in the following paragraphs.

DNA methylation

Methylation at the 5' end of cytosine (5mC) is mostly associated with transcriptional repression and is established via members of the DNA methyltransferase (DNMT) family [72]. These DNMTs do not only establish *de novo* marks but also maintain methylation, thus making DNA methylation a long-term mark. DNA methylation-mediated silencing, which is established via chromatin remodeling complexes or by direct blocking of transcription factor binding, is involved in cellular programming events, X chromosome inactivation and genomic imprinting [72, 75, 76]. The 5mC mark mostly occurs at CpG-dinucleotides in mammalian genomes and 70-80 % of those CpGs are methylated. Exceptional are the so-called CpG-islands, areas with 50 % CG-content or higher, which are mostly unmethylated and are often found at active gene promotors [69].

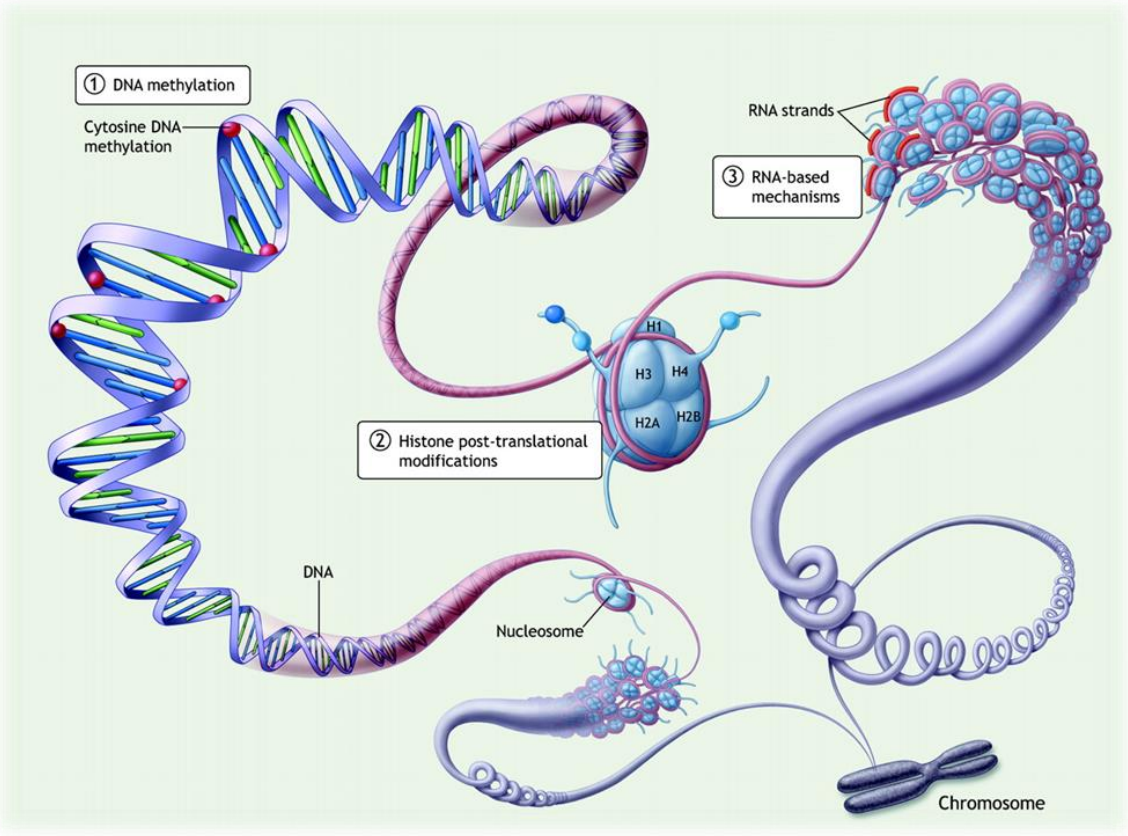


Figure 1.5. Overview of epigenetic mechanisms. DNA methylation, posttranslational histone modifications and RNA-based mechanisms constitute the epigenetic network (Image taken from [77]).

Although for a long time considered to be irreversible, recently several mechanisms for both passive and active demethylation have been proposed [72, 75]. 5mC can be modified by TET enzymes to 5-hydroxymethylcytosine (5hmC) and possibly further oxidized to 5-formylcytosine (5fmC) and 5-carbonylcytosine (5caC). Via base excision repair mechanisms these residues can be re-converted to an unmodified cytosine. Next to these active demethylation processes, the methyl marks can also dilute passively over several cell divisions, upon loss of DNMT function or (sustained) lack of their recruitment.

RNA-based epigenetic mechanisms

Although only 1-2 % of the total human genome actually encodes for proteins, a major part of it is transcribed anyway thus leading to a large pool of non-coding RNAs (ncRNAs), among which ribosomal RNAs (rRNAs), transfer RNAs and regulatory RNAs such as microRNAs (miRNAs) and long ncRNAs (lncRNAs) [72, 78]. MiRNAs, Piwi-interacting RNAs (piRNAs) and short interfering RNAs (siRNAs) are all short ncRNAs, but differ in their biogenesis and target regulatory mechanism [78]. MiRNAs are small RNAs (about 21 nucleotides long), that help in regulating messenger RNA (mRNA) stability and thus protein abundance, primarily in inhibitory mode [72, 78]. They are initially transcribed

as single stranded RNAs and subsequently fold into stem-loop structures, which are then processed by the RNase III-endoribonuclease Dicer. One of the resulting single stranded miRNAs interacts with the RNA-induced silencing complex (RISC), which then binds the specific target mRNA containing a sequence complementary to the miRNA loaded on RISC. This consequently induces the degradation of this mRNA or blocks its translation [78]. PiRNAs interact with Piwi proteins for chromatin remodeling and transposon silencing [78]. Also lncRNAs (>200 base pairs) exert functions at several stages of gene expression, from targeting epigenetic modifications to mRNA stability and translation modulation e.g. in the regulation of X chromosome inactivation [78, 79]. These lncRNAs can act as scaffolds for the generation of ribonucleoprotein complexes and networks containing diverse chromatin modifiers and remodelers, to modulate chromatin states (mostly silencing) by functioning at the histone – genome interface ([80] and Supplementary Material of [81]). Alternatively, lncRNAs operate as a protein decoy by sequestering proteins such as DNMT1 from their sites of action thus preventing the protein from exerting its function [80]. In summary, both short and long regulatory ncRNAs are thus involved in epigenetic regulation e.g. via RNA interference [78] and also fulfill specific roles in ESC (*Section 1.3*).

Chromatin remodeling complexes

ATP-dependent chromatin remodeling complexes modulate the interaction of DNA and histone octamers, leading to a change in chromatin architecture [72]. These enzymes transiently disrupt this interaction which allows for nucleosome sliding, translocation, eviction and replacement on one hand and compositional nucleosome changes due to histone variant exchange on the other. The numerous remodeling complexes are divided into several families, based on activity and composition, e.g. SWI/SNF, NuRD or CHD1 complexes [72, 82]. In ESC, some of these are closely involved in pluripotency-regulation.

1.2.2. Epigenetics on the histone level

Nucleosomes, the building blocks of chromatin [83, 84], are composed of 146 bp of DNA wrapped around two copies of four core histones (H2A, H2B, H3 and H4) kept together via electrostatic interactions and hydrogen bonds (Figure 1.6) [85]. Histones mutually interact via a characteristic "handshake" motif [83]. Two H2A-H2B dimers combine with an H3-H4 tetramer and thus altogether form the histone octamer. The interaction of consecutive nucleosomes creates a "beads on a string" formation, which further tightens in combination with linker H1 histones [86], linker DNA and other, non-histone proteins. This hierarchical compaction leads to local chromatin loops, then chromatin domains and ultimately chromosome territories [87]. Hence, via both protein-protein and protein-DNA interactions a more compact, higher-order structure is formed, termed the "30 nm chromatin fiber" in

its most condensed form [88]. The approximately two meter-long strand of DNA in a human diploid cell thus fits perfectly into the cell's nucleus [89].

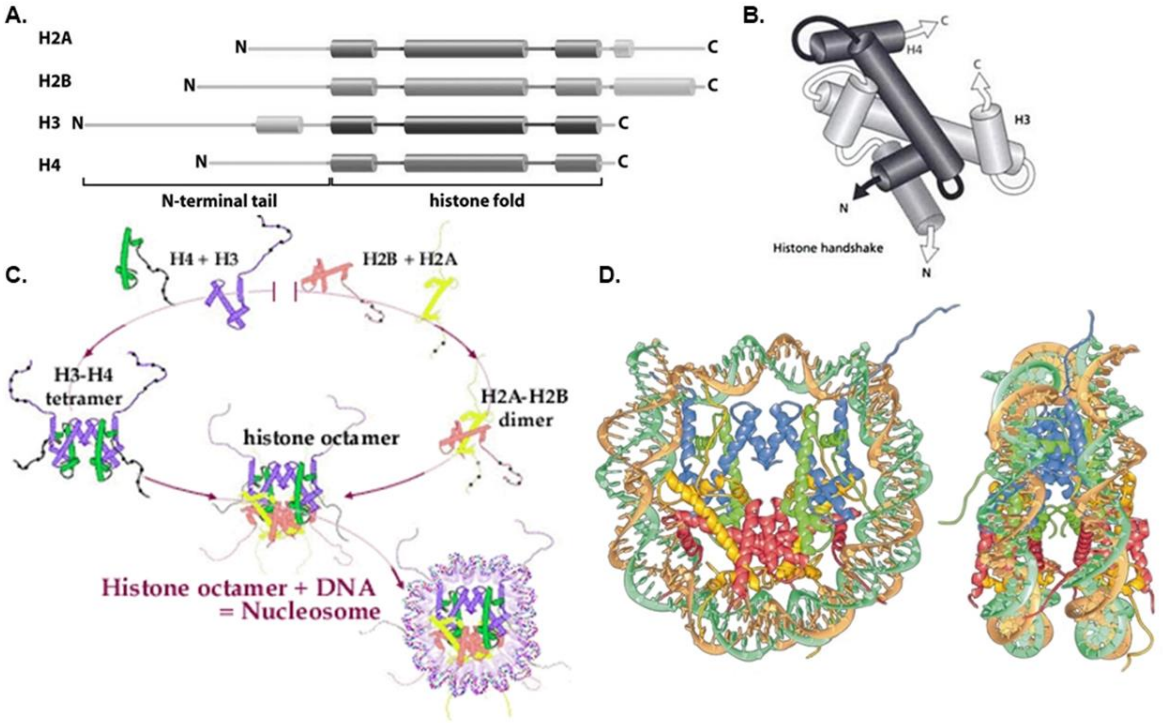


Figure 1.6. The nucleosome. A: The structural organization of the core histones (Image adapted from [90]). **B:** The histone 'handshake' motif illustrated for H3 and H4 (Image taken from [91]). **C:** The assembly of a histone octamer (Image adapted from [92]). **D:** A nucleosome particle, with protruding histone tails (with H2A = yellow - H2B = red - H3 = blue - H4 = green) (Image taken from [84]).

Core histones are small proteins, basic in nature and found in almost all eukaryotic cells [93]. Genes encoding for these histones are usually clustered in repeat arrays and their transcription is tightly connected with the cell cycle and DNA replication [94]. Their amino acid sequence is highly conserved throughout evolution and also their 3D-structure seems comparable between different eukaryotic species [95]. The globular domains of the core histones share the similar structure of a histone fold domain, which consists of three α-helices and two loops in between [94]. As such, about half of the histone octamer is arranged in α-helices while only a small fraction is organized in β-strands and the remaining part adopts a random-coil configuration [83]. In addition to its structural and packaging role, this higher-order organization participates in regulating the accessibility of the genomic material and as such in gene expression management amongst other functions [83, 87].

Posttranslational histone modifications

The N-terminal histone tails and the C-tail of H2A are protruding out of the nucleosome and are a well-known target for posttranslationally modifying enzymes [94]. Despite some conflicting

publications, it has become clear that the N-terminal tails with their modifications do play important roles in stabilizing both the higher-order structure of chromatin as well as the nucleosome itself [96, 97]. The first reports of PTMs date from the early 1960's [98] and described histone acetylation. Over the years, a whole spectrum of (reversible) modifications was revealed, together with their accompanying installing and removing enzymes (termed "writers" and "erasers" respectively) [99–102]. Gradually it became clear that not only the histone tails but also the globular domains can be modified (Figure 1.7) [103, 104].

The core histone tails are particularly rich in lysine and arginine residues (about 20 % of their amino acid composition) and consequently the N-termini compose 32 % of all positive charges of the histone complex [97]. These lysines and arginines underlie the histones' basic nature and are frequently targeted for modification, e.g. by acetyl- and methyltransferase enzymes.

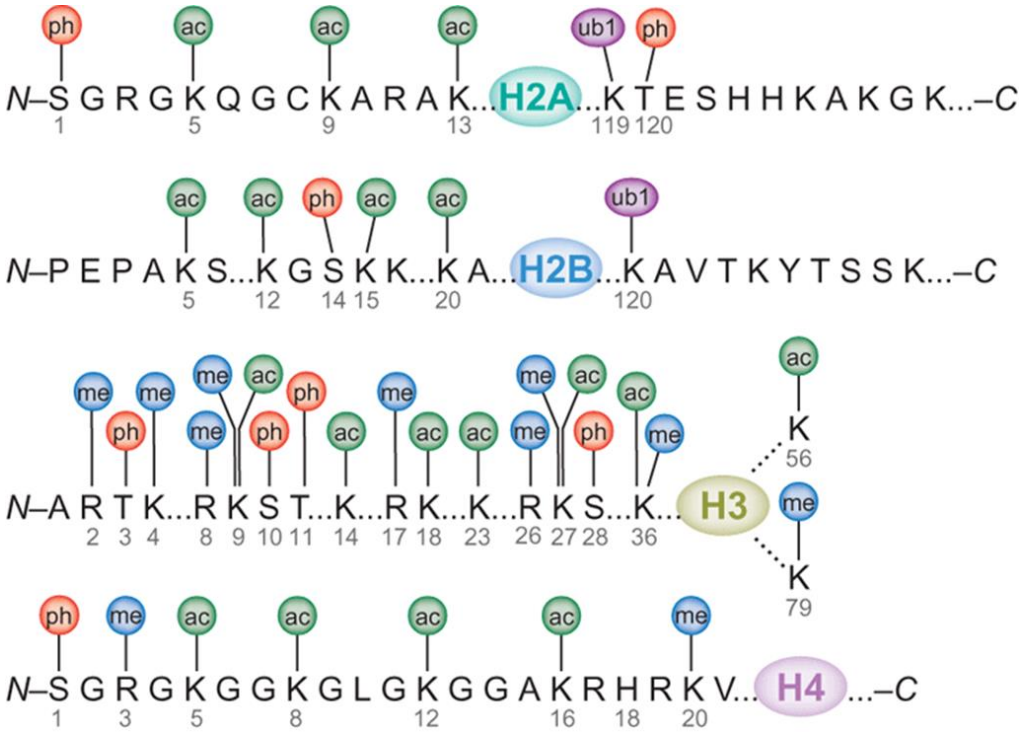


Figure 1.7. Histone modifications. A non-exhaustive list of possible histone PTMs of the four core histones. Abbreviations: ph = phosphorylation - ac = acetylation - ub1 = mono-ubiquitination - me = methylation. The ovals represent the histones' globular domains (Image taken from [103]).

Several different types of histone PTMs can be induced: the addition of a functional group or another protein to a specific amino acid, modification of the chemical nature of the residue itself or more drastic structural changes of the histone protein can be brought about. Well-studied examples in which a group or protein is added are lysine methylation, arginine methylation, lysine acetylation and ubiquitination amongst others, while citrullination and proline isomerization convert the amino acid

itself. An overview of this kind of modifications is given in Table 1.1. Histone clipping on the other hand, leads to a more radical change, as a fragment of the histone tail is cleaved off. However, histone clipping surprisingly still remains an outcast in the epigenetic landscape. Our findings on histone clipping are elaborately described in Chapter 3 and 4 [44], followed by a more in-depth discussion in Chapter 5.

All modifications bring along their own effect due to the changes in chromatin structure they directly induce or due to the recruitment of other transcriptional regulators [99, 105]. Chromatin structure is directly modulated by modifications such as acetylation, phosphorylation and ubiquitination. Lysine acetylation induces a change by neutralizing the positive histone charge, thereby loosening the histone-DNA interaction and creating a more open chromatin organization [103]. Phosphorylation might similarly have an impact via charge changes [99]. H2B ubiquitination is believed to physically separate chromatin fibers, hence making the DNA more accessible [105]. Nevertheless, most biological consequences are accommodated under the second group and are mediated via "reader" proteins [105]. For example, lysine or arginine methylation does not induce any net change in charge or steric hindrance and thus exerts its influence indirectly via the recruitment of other non-histone proteins [81]. To this end the modifications serve as an anchor site and are recognized via specific binding domains.

These binding pockets can be categorized according to their architecture and are able to distinguish a particular sequence surrounding a PTM, allowing for very specific chromatin targeting by their host proteins to ensure the proper biological outcome. The major binding domain types, as displayed in Figure 1.8A [99], can be found in various histone modifiers and chromatin remodeling complexes or components that recruit these (extensively reviewed in [81]). As such, specific domains for methyl-lysine, methyl-arginine, acetyl- and phospho-sites have already been identified. Unmodified and methylated lysines can be recognized by different domain families such as Tudor, Chromodomain, Plant Homeodomain (PHD) finger, Proline-Tryptophan-Tryptophan-Proline (PWWP) and MBT (Malignant Brain Tumor) amongst others. For example, heterochromatin protein 1 (HP1) and Polycomb proteins bind their targets via a chromodomain. Analogously, some of these families' members are also involved in arginine and methyl-arginine readout. Bromodomains are one of the best studied acetylation recognition sites and are a known component of several histone acetyl transferases and remodeling complexes. Phosphorylated residues are recognized by 14-3-3 proteins, which are involved in several regulatory processes such as signal transduction and chromosome condensation [81, 106]. In any case, the different types of binding modules are often simultaneously incorporated in reader complexes, to cooperatively recruit the necessary nuclear signaling network components to chromatin [81, 106].

The binding of such reader proteins might go via several modes (Figure 1.8B-C-D). First, multisite recognition occurs when the same mark is recognized on several locations, on the same or different histone tails. Second, for combinatorial PTM readout a single reader engages in reading several different marks simultaneously. The modifications may mutually antagonize or enhance binding. Finally, in multivalent binding a protein complex that comprises several binding sites interacts with coinciding modifications at the same time.

In terms of the resulting functional effect, it is now clear that the modification's context is one of the determining factors. As such, different modifications often cooperate to achieve the aimed effect via mutual communication or so-called "histone crosstalk" within the same or among different histone tails (Figure 1.8E) [99, 107]. Such communication can occur at several levels [99]. **(I)** Lysine residues are very susceptible targets for a whole set of modifications. Obviously these marks cannot be placed on a specific residue at the same time, implicating that some PTMs are mutually exclusive and thus regulate each other's deposition. **(II)** As described above, one mark can interfere with and disrupt the binding of another non-histone protein on another modification (Figure 1.8C). The binding of specific enzymes may also be influenced: **(III)** a modification can compromise the enzyme's recognition site or **(IV)** conversely, the placement of an adjacent mark is able to enhance the enzyme recognition of an earlier deposited modification.

Thus, whether they directly alter the biophysical properties of nucleosomes or act as a histone code that is decoded by specific effector proteins, histone PTMs participate both directly and indirectly in dictating the global character of the chromatin environment [99, 108]. Euchromatin (*Section 1.2.1*) displays high levels of acetylated lysines and trimethylation on H3K4, H3K36 and H3K79 residues. Heterochromatin on the contrary, is associated with low levels of acetylation and phosphorylation and high levels of certain methylations (H3K9, H3K27, H4K20) together with DNA methylation and siRNAs. The silent heterochromatin state is thus installed and maintained through e.g. the recruitment of HP1 via the H3K9me mark. Heterochromatin can be further subdivided into constitutive and facultative heterochromatin, which is respectively a stable form and a reversible chromatin state. A rather exceptional third form of chromatin organization, is the so-called "bivalent domain", where both activating and repressive histone modifications are combined into one with resulting silent genes that are yet poised for transcriptional activation (*discussed in more detail in Section 1.3*). Initially they were considered as a unique ESC feature, but also in other cells evidence for these domains has been found [72, 109].

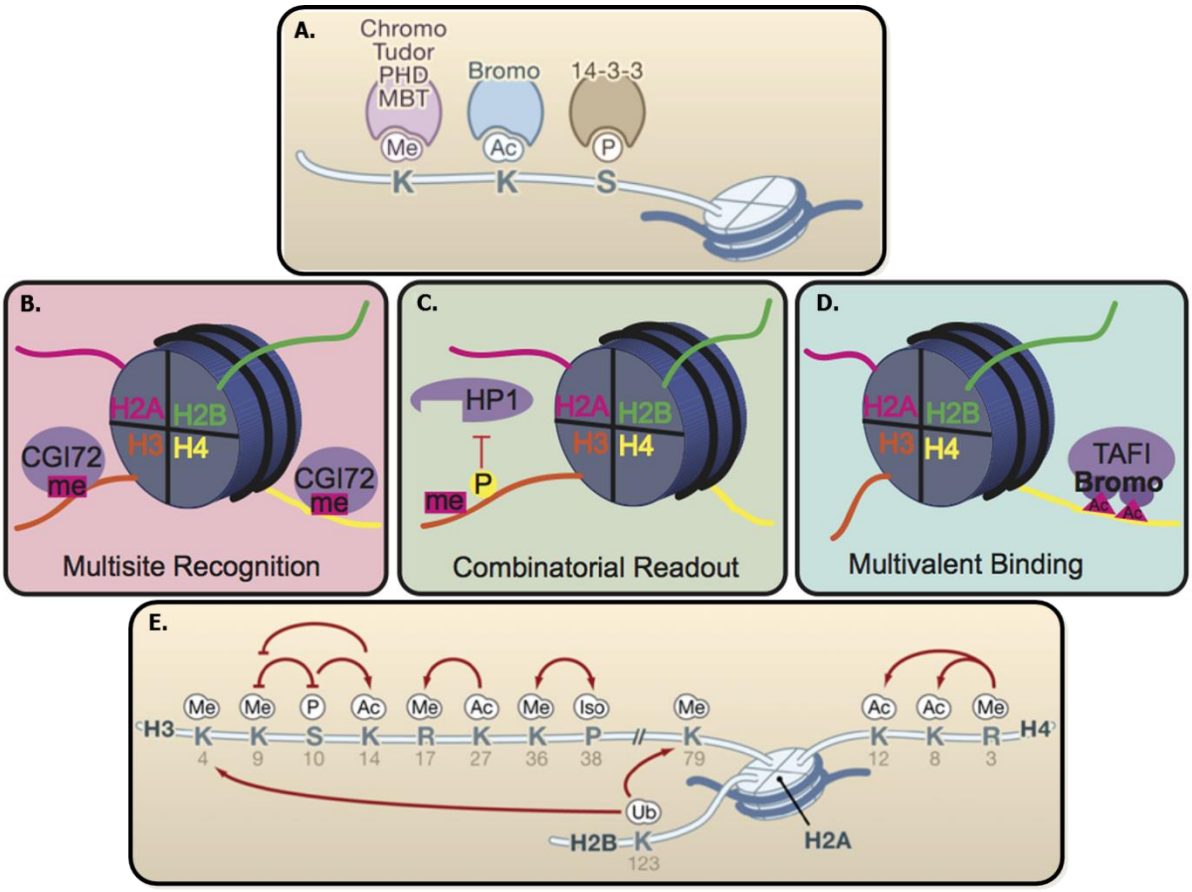


Figure 1.8. Binding of histone readers. A: Possible binding domains interacting with histone PTMs. **B:** Multisite recognition: e.g. the reader CGI72 can bind both H3K4me1 and H4K20me1. **C:** Combinatorial readout of H3S10ph and H3K9me3 by HP1. The phosphorylation prevents HP1 from binding to the neighboring methyl mark. **D:** Multivalent binding of the TAF1, a component of the TFIID transcription factor, has two bromodomains that can bind acetylated lysines on H4 via their bromo-domains. **E:** Histone PTM crosstalk illustrated (positive influence = arrow, negative = dish-line) (Images taken and adapted from [99] and [105]).

Table 1.1. Overview of histone PTMs and their function.

Modification	Residue	Established by	Reversible?	Involved in
Acetylation [99, 103, 124, 234]	Lysine **[K-ac]**	Histone acetyltransferases (HATs)	Histone deacetylases (HDACs)	- Generally associated with transcriptional activation (e.g. H2BK5ac, H3K9ac, H4K8ac), although association with inactive genes also has been reported (H3K56ac) - Genomic stability - Aiding in DNA repair - Chromatin decondensation
Methylation [99, 103, 108, 234]	Lysine **[K-me1 / K-me2 / K-me3]**	Histone lysine methyltransferases (HMTs)	Histone demethylases	- Residue-dependent transcription regulation: - Activating: H3K4me1/2/3, H3K9me1, H3K36me3, H3K79me1/2/3, H4K20me1 - Repressive: H3K9me2/3, H3K27me2/3, H4K20me3 - DNA repair (anchor for recruitment and stabilization of other regulatory factors)
	Arginine **[R-me1 / R-me2s / R-me2a]**	Histone arginine methyltransferases	Histone demethylases	Gene activation or repression, depending on context and methyltransferase involved – e.g. activating: H3R17me2 and H3R26me2
Phosphorylation [99, 103, 234]	Serine / Threonine / Tyrosine **[S/T/Y-ph]**	Kinases	Phosphorylases	- Transcription: gene activation - DNA repair (anchor for remodeling complexes) - Cell cycle regulation (chromatin (de)condensation)
Citrullination (Deimination) [99, 125, 235]	Arginine **[R → Cit]**	Peptidyl arginine deiminase type IV	*No evidence*	Promoting transcription: e.g. H3cit leads to a more open chromatin structure by expelling HP1 from the chromatin. Potentially antagonizing arginine methylation.
(mono-) Ubiquitination [99, 103, 234]	Lysine **[K-ub1]**	Ubiquitin-ligase complex (an E1 enzyme for activation, E2 for conjugation, E3 for ligation)	Deubiquitinases	- Context-dependent effect on gene expression: e.g. H2BK120ub1 is activating, H2AK119ub1 repressive - DNA repair

Table 1.1. Overview of histone PTMs and their function – continued.

Modification	Residue	Established by	Reversible?	Involved in
SUMOylation [99, 234]	Lysine **[K-su]**	SUMO-ligase complex (E1 – E2 – E3)	SUMO proteases	Repressive influence on transcription (antagonizing active marks such as acetylation)
ADP-ribosylation [99, 234]	Glutamic acid **[E-ar]**	Mono-ADP-ribosyltransferases or Poly-ADP-ribose-polymerases	*No evidence*	Transcription regulation and DNA repair: ADP-ribosylated histones are removed, leading to an open chromatin state
Proline isomerization [99, 234, 236]	Proline **[Cis → Trans]**	Proline isomerases	Proline isomerases	Transcription: distortion of the polypeptide backbone by isomerization influences regulation of other PTMs as methylation
β-N-acetylglucosaminylation [101, 104, 237]	Serine / Threonine **[S/T-O-GlcNAc]**	O-GlcNAc transferase	β-N-acetylglucosaminidase (O-GlcNAcase)	Predominantly associated with condensed chromatin, hint for a role in chromatin remodeling
Propionylation [104, 238, 239]	Lysine **[K-prop]**	Acetyltransferases (substrate propionyl-CoA) and/or novel enzymes	Histone deacetylases and/or novel enzymes	*Specific role not yet defined*
Butyrylation [104, 238]	Lysine **[K-bu]**	Acetyltransferases (substrate butyryl-CoA) and/or novel enzymes	Histone deacetylases and/or novel enzymes	*Specific role not yet defined*
Biotinylation [240, 241]	Lysine **[K-biot]**	Biotinidase	Uncertain, perhaps also biotinidase (dependent on context and/or alternate splicing)	- Gene silencing - DNA repair
N-Formylation [104, 242]	Lysine **[K-fo]**	Reaction with formylphosphate, induced by oxidative stress	*No evidence*	Specific role not fully defined; can interfere with formation and stabilization of binding sites for regulatory proteins

Table 1.1. Overview of histone PTMs and their function – continued (2).

Modification	Residue	Established by	Reversible?	Involved in
Crotonylation [100, 104]	Lysine **[K-cr]**	*Unknown effectors*	*Unknown effectors*	Specific role not fully defined; largely associated with active chromatin
Succinylation [102, 104]	Lysine **[K-succ]**	Unknown effectors; probably enzymatic, via succinyl-CoA	*No evidence*	Specific role not fully defined; interference of interaction between histones and DNA
Malonylation [102, 104]	Lysine **[K-mal]**	Unknown effectors; probably enzymatic, via malonyl-CoA	*No evidence*	Specific role not fully defined; induces diminished interaction between histones and DNA
Hydroxylation [100, 104]	Tyrosine **[Y-oh]**	Tyrosine hydroxylase	*No evidence*	Specific role not fully defined; probable involvement in chromatin structure regulation as identified sites are located near the H2B-H4 contact

Histone variants

Another level of variation in epigenetics is introduced by the replacement of canonical histones by histone variants (Table 1.2) [72]. They differ (slightly) in primary sequence, hence they alter chromatin structure and have a different modification susceptibility, leading to a different epigenetic regulation and altered chromatin dynamics [72, 94]. In contrast to their canonical counterparts, their expression (generally originating from a single gene) is often not replication-coupled; variants can be expressed throughout the whole cell cycle. Usually they get help from their own specific chaperone proteins for their implementation [94].

Table 1.2. Overview of the known histone variants of core histone H2A and H3 and linker histone H1 in ESC. Based on [72].

H2A	H3	H1
H2A.X	H3.1	H1.0
H2A.Z	H3.2	H1.1
macroH2A	H3.3	H1.2
H2A-BDB	CENP-A	H1.3
		H1.4
		H1.5
		H1X

Histone variants are involved in transcription, DNA repair, meiotic recombination, sperm chromatin packaging, etc. [94, 110]. H3.1, H3.2 and H3.3 are the main histone H3 forms and only differ by one to five amino acids. They show a different PTM pattern with H3.1 carrying both repressive and activating marks, H3.2 being enriched with inhibitory modifications and H3.3 mainly bearing active modifications, which might imply distinct roles for each [111–113]. As such, H3.3 is associated with transcriptional elongation and is deposited at active genes and promotors or at least at genes that are poised for activation. CENP-A, the human centromeric histone H3 variant, is essential for the kinetochore-assembly during mitosis. H2A.Z is involved in both gene activation and silencing, DNA repair, nucleosome turnover, heterochromatin, antagonizing DNA methylation in plants, etc. These often contradictory roles can at least be partially explained by the different possible PTMs found on H2A.Z (e.g. acetylation versus ubiquitination). The H2A.X variant is participating in DNA repair amongst other roles.

Apart from these aforementioned universal eukaryotic histone variants, several specific variants also exist, such as macroH2A isoforms, mostly implicated in gene silencing e.g. on the X chromosome, or testis-specific variants, which in combination with protamines tightly pack DNA in sperm cells.

In summary, the PTM pattern introduced at a specific locus is dependent on the specific chromatin assembly pathway that is chosen (determining the histone variant and modifying enzymes before histone deposition) and on the local neighborhood (locally acting modifiers after histone installation) [111].

1.2.3. Epigenetic memory

The stable passing on of a change in gene expression induced by developmental or environmental signals, i.e. the epigenetic memory, is established by multiple mechanisms which often include chromatin-based changes such as posttranslational histone modifications [114]. There are at least three types of epigenetic memory that can be defined: **(I)** cellular memory, **(II)** transcriptional memory and **(III)** transgenerational memory.

(I) Cellular memory is instituted during development in response to developmental (environmental) stimuli. The induced changes outline the cell's identity and its differentiation potential and are maintained through subsequent cell divisions, even if the initial stimuli are abrogated. For example, Polycomb Group (PcG) proteins bookmark genes that need to remain repressed during mitosis by means of H3K27 methylation [114].

(II) Transcriptional memory involves the changes that are achieved by cells and organisms in response to changes in their environment in order to adapt and survive in that habitat. As such, the reactivation of interferon gamma (IFNγ)-dependent *HLA-DRA* in HeLa cells is stronger and more rapidly achieved if cells have been exposed to IFNγ before. This example of transcriptional memory occurs through the propagation of H3K4 dimethylation [114].

(III) Previous experiences can lead to epigenetically changed gene expression in offspring via transgenerational memory. Maternal habits (e.g. in context of nutrition, stress, etc.) and environment as well as paternal behavior and environment can impact the next generation's physiology and behavior substantially, via epigenetic mechanisms such as genomic imprinting. The resulting phenotype can be maintained to adulthood and even in subsequent generations [114].

To what extent and how exactly the epigenetic memory is established still remains to be elucidated [114]. The DNA methylation pattern on a newly synthesized strand can be directly copied from the template strand. However, the mechanisms by which histone modifications get passed on are not fully understood [115]. During DNA replication, also the histone content needs to double to provide each daughter cell with a complete chromatin structure [111]. These histones are recycled from the parent cell or synthesized *de novo*, produced from replication-dependent canonical histone genes and/or replacement histone variant genes whose expression is not restricted to the S phase. One hypothesis

to maintain the existing PTMs and variants in the progeny is that the H3-H4 tetramer evenly splits into two dimers, which are partitioned to the daughter cells and serve as a template for copying the respective PTMs. This is however not totally supported by the findings of Xu et al. [115]. Their results indicate that most canonical H3-H4 tetramers remain intact, but some variant H3-H4 tetramers however, can indeed be split. Whether this splitting is unique for specific loci remains uncertain. They show that silent PTMs are maintained by literally copying from preexisting neighboring nucleosomes, but for active marks it is not so clear. Since the H3.3 variant is predominantly associated with transcriptionally active regions, splitting of H3.3-containing tetramers is proposed as one way for partitioning and maintaining this euchromatic feature.

Alternatively, Greer and Shi propose that histones on the newly synthesized DNA immediately become modified after binding, based on evidence that PcG proteins can remain bound during replication [116]. Another idea states that histones get marked before deposition. However, this would also imply that not all types of methylation marks are inherited since it is impossible for the cell to keep all possible modification combinations in stock. They too confirm that the splitting of tetramers is rather unlikely, but state that the segregation of entire octamers to each of the daughter strands is possible. Other models involve the collaboration of other epigenetic marks. For example, DNA methylation might help in maintaining histone methylation by recruiting the correct modifiers. However, it is not yet certain that the correlation between DNA and histone methylation always holds. It is also suggested that lncRNAs take care of the modifiers' recruitment or that siRNAs lend a hand in preserving heterochromatin regions.

In summary, it is clear that epigenetic variation and plasticity are crucial to how organisms develop, adapt and interact with their environment. Importantly, the impact of epigenetic changes can both be beneficial or harmful, depending on the context [114]. It is increasingly becoming clear that our genome is not the only determining factor for development, evolution, the occurrence of diseases, etc. Anand and colleagues boldly state that cancer development is predominantly environmentally regulated and consequently epigenetics is turning into a major focus in the search for molecular changes underlying disease [74, 117].

1.3. Epigenetic reorganization upon ESC differentiation

Waddington conceptuated an epigenetic landscape model [118] in which a progenitor cell is illustrated as a rolling ball in a landscape divided into several valleys or tracks, representing the various possible phenotypes in which the cell can differentiate (Figure 1.9). The correct track is chosen once the cell is committed towards a certain cell lineage. The development towards a specific phenotype is however canalized, which means that minor variations do not influence the one definite lineage commitment, which is illustrated by the high ridges that enclose the phenotypic valleys [119]. Nowadays, the emergence of cell programming possibilities supplements this model [109, 120]. Somatic cells can be pushed back up *in vitro* to generate induced pluripotent stem cells (*Section 1.1.5*). Alternatively, during direct conversion or transdifferentiation, a cell can directly jump over a hill to convert into another related tissue-specific cell or into a cell type of another germ layer.

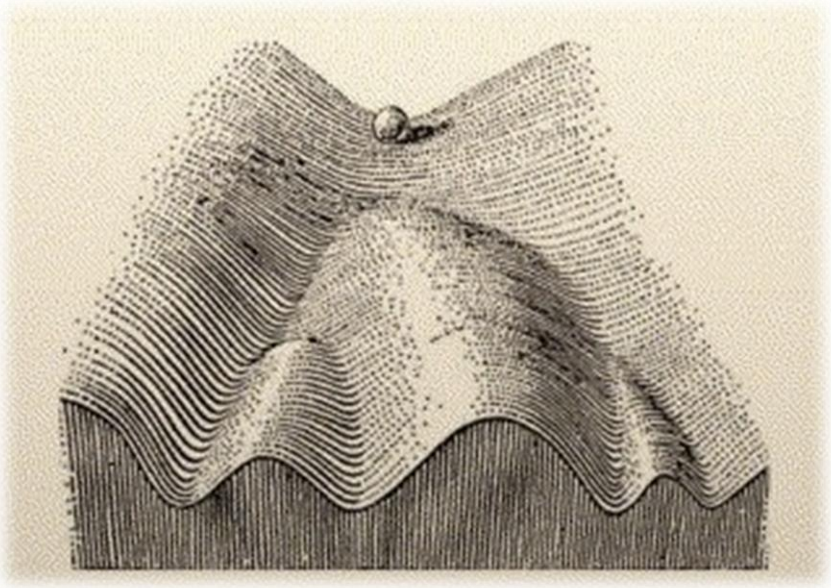


Figure 1.9. The original epigenetic landscape model of Waddington. A progenitor cell rolls into a specific valley after lineage commitment (Image taken from [121]).

The specific epigenetic characteristics of hESC are now increasingly being taken into account for their characterization as a hallmark of pluripotency. In general, ESC have a loose chromatin structure with accompanying high levels of gene activity. Heterochromatin markers are barely concentrated and highly dynamic in ESC and become more focused at specific loci upon differentiation [109].

At the DNA level, genes involved in the control of cell differentiation in mammalian germ cells are methylated and thus inactive. However, during embryonic development the paternal as well as the maternal genome gets to be heavily demethylated via both active and passive removal processes. After this genome-wide reorganization, promotors of pluripotency-associated genes such as *POU5F1*

encoding Oct4, thus become hypomethylated in ESC. E.g. DNA hydroxymethylation, often associated with transcriptional activation, helps in this demethylation process in mESC [72]. Upon differentiation however, progressive gene silencing occurs through DNA hypermethylation [109]. Also, TET levels decline and hydroxymethylation levels at pluripotency gene promotors decrease, which in combination with upregulated methylation at these sites leads to gene silencing [72].

Non-coding RNAs such as miRNAs also play an important role in both mESC differentiation and the self-renewal program [72, 109]. They can promote differentiation via reduction of the pluripotency factors' protein levels [109]. Oct4, Nanog and Sox2 have been found to occupy the promotors of several miRNAs in mESC to regulate the expression thereof [72].

Epigenetic changes on histones play an equally important role during stem cell differentiation. The most extensively studied histone PTMs in this context are acetylation and methylation [109]. Histone methylation, as described above, can have an inhibitory or an activating effect on transcription and chromatin structure. The main mediators for altering methyl marks on histone tails during the differentiation process are HMT proteins belonging to either the PcG or trithorax Group (trxG). These respectively install repressing or activating methylation marks on histones positioned at e.g. the developmentally important Homeobox genes amongst numerous other genes [112]. As such, in pluripotent hESC so-called bivalent domains are established at up to 16 % of the gene content, preferably at key developmental genes [122, 123]. In these domains, activating trimethylation on lysine 4 of H3 (H3K4me3) (trxG) and repressing trimethylation on lysine 27 (H3K27me3) marks (PcG) are combined [109, 123], resulting in silent but poised genes ready for rapid transcriptional activation. As differentiation progresses, these marks will resolve in either the activating mark or the repressing, depending on the lineage specification of that cell type. Also in multipotent stem cells of mouse and human origin, neural progenitors and terminal neurons, evidence for bivalency has been found, where it is believed to help with gene plasticity [72, 109]. The low levels of repressive H3K9 trimethylation and high overall acetylation levels found in mESC correspond with their general open chromatin structure and high transcriptional activity, that diminish largely upon differentiation [109]. The acetylation pattern on H3K56, which correlates positively with Oct4, Nanog and Sox2 binding along promotors in pluripotent hESC, undergoes upon differentiation a substantial redistribution from pluripotency-related genes towards developmental genes [124]. Recently, also citrullination of histone H1, H3 and H4 was added to the list of ESC PTMs as it leads to chromatin decondensation favoring the expression of mESC pluripotency markers [125].

The achievement of substantial PTM changes during differentiation such as the resolution of bivalent domains can be effectuated through several mechanisms [44, 126, 127]. First and most obviously, the

installing enzymes are counteracted by modification-removing enzymes, as happens e.g. during enzymatic lysine demethylation (Table 1.1). Second, histone exchange and the deposition of histone variants also notably contributes to a changed epigenetic pattern. In one such example, Duncan and colleagues found that the ratio of (H3.2+H3.3)/(H3.1) increased after RA induction of mESC due to an increase in H3.2 and/or H3.3 or a decrease in H3.1 [112]. Finally, the removal of a histone tail by enzymatic histone proteolysis provides another mechanism by which multiple modifications can be erased at once in both hESC and mESC ([44, 127]; *further elucidated in Chapter 3 - 5*). Despite the seemingly drastic nature of this latter modification, surprisingly little publications reporting this phenomenon attempt to figure out the full mechanism behind it. In any case, it is clear that different factors collaborate to specifically bring about the necessary changes in the epigenetic ESC landscape.

CHAPTER 2:

THE APPLICATION OF RT-QPCR IN HUMAN EMBRYONIC STEM CELL RESEARCH

Adapted from "**Reference loci for RT-qPCR analysis of human embryonic stem cells**"

Liesbeth Vossaert, Thomas O'Leary, Christophe Van Neste, Björn Heindryckx, Jo Vandesompele, Petra De Sutter and Dieter Deforce

BMC Molecular Biology (2013) 14:21

2. APPLICATION OF RT-QPCR IN HUMAN EMBRYONIC STEM CELL RESEARCH

As previously outlined in the introduction, the epigenetic signature of hESC undergoes a considerable shift upon differentiation. This heralds an accordingly drastic change in transcription. One targeted technique to follow up changes in the expression of specific genes is reverse transcription-quantitative PCR (RT-qPCR). Following a general introduction on the technical aspects of this method, this chapter describes the optimization for adequate RT-qPCR data analysis of samples obtained from differentiating hESC.

2.1. An introduction to RT-qPCR

RT-qPCR combines the amplification of small complementary DNA (cDNA) sequences by PCR with their quantification via fluorescence measurement and is widely distributed to monitor gene expression levels, even in high throughput.

PCR was introduced in the 1980's by Kary Mullis [128]. During the PCR process, a specific piece of DNA is exponentially amplified to generate a large number of identical copies of this sequence. The PCR cycle starts with a denaturation step, during which double stranded DNA is separated by heat (95 °C) to two single stranded molecules. In the following annealing step, specific primers bind to the DNA (around 60 °C) and are subsequently elongated by DNA polymerase (at 72 °C) (Figure 2.1). Depending on the protocol, the PCR cycle can be reduced to two steps by combining the annealing and elongation step at 60 °C.

By combining the PCR technology with fluorescence detection of the formed products during the course of the PCR cycle (instead of endpoint detection), gene expression levels can be accurately quantified. The complete RT-qPCR workflow is displayed in Figure 2.2.

2.1.1. Setting up an RT-qPCR experiment

The starting material for RT-qPCR is RNA. Cells or tissue samples are collected and stored in TRIzol reagent for cell lysis. After RNA isolation and DNase treatment to remove genomic DNA (gDNA), cDNA is synthesized via reverse transcription. This cDNA is then further on used for amplification, after removal of the RNA template by RNase.

Before moving on to the actual reverse transcription, amplification and expression quantification, the quality and quantity of the input RNA needs to be assessed as it is obviously determining for the downstream results. A SPUD assay, a 5'/3' ratio mRNA integrity assay and the determination of the RNA quality index (RQI) and 28 S/18 S rRNA ratio can all be employed to this end. The latter reflects the ratio of the 28 S and the 18 S rRNA species, which are both derived from a common transcription

unit, thus providing the exact same number of both molecules in a cell. Since they are approximately 5 kb and 2 kb in size respectively, their theoretical ratio equals 2.7, but in practice a ratio of 2 is considered as high quality RNA. The RQI translates this ratio into a value from 1 (degraded RNA) to 10 (good quality). In a 5′/3′ ratio mRNA integrity assay, cDNA is synthesized from the RNA template with the use of oligo(dT) primers that anneal specifically to the mRNA 3′ polyA-tail. In a following PCR step, two sets of primers are included in order to amplify both 5′ and 3′ fragments. The ratio of these amplified fragments provides a direct indication for RNA integrity and equals 1 for good quality RNA. As degradation often starts at the 5′ end, amplified fragments of the latter will be outnumbered compared to 3′ fragments, resulting in a lower ratio. A SPUD assay can be included to detect possible PCR inhibitors present in a sample. For this assay, an artificial template is added to the samples, together with the corresponding primers, which will render a specific quantification cycle (Cq) value. If there are inhibiting compounds present, this Cq will shift to a higher value compared to the control reaction.

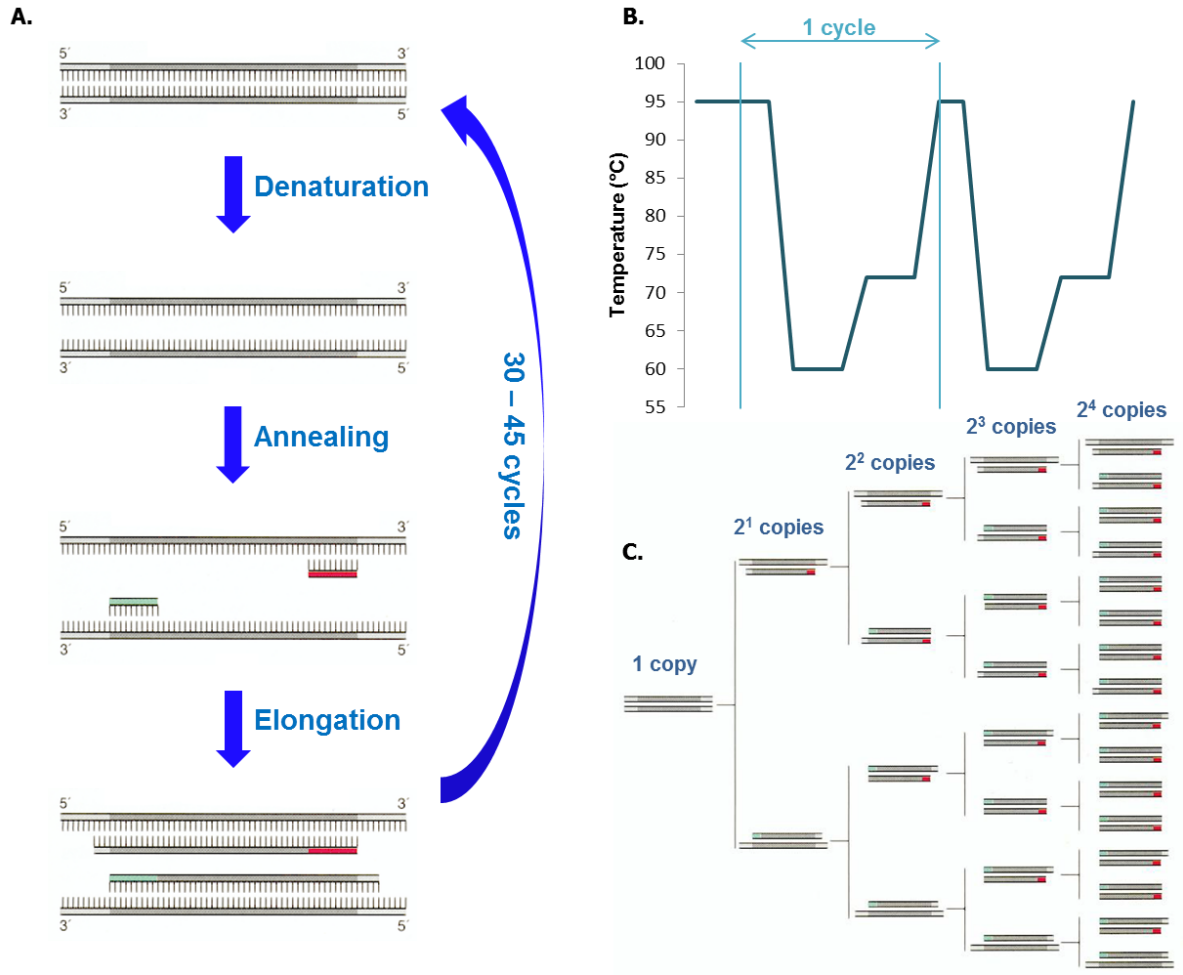


Figure 2.1. Polymerase chain reaction. A: PCR consists of a repeated cycle of denaturation, annealing and elongation. **B:** The according thermal profile. **C:** The amount of DNA doubles in each cycle, thus leading to an exponential increase in nucleotide material (Images taken and adapted from [129]).

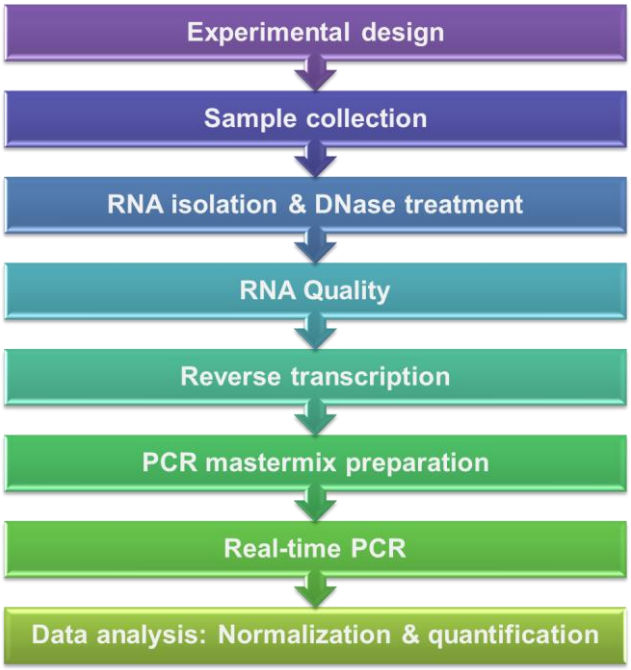


Figure 2.2. Overview of the RT-qPCR workflow.

In the reverse transcription reaction, random or oligo(dT) primers or a combination thereof can be used. Oligo(dT) primers is only bind specifically to a polyA-tail which is present on mRNA. However, the use of random primers is more desirable when working with partially degraded samples, as these primers do not require a specific sequence such as the polyA-tail and thus can enable a higher coverage. Alternatively, specific primers can be applied to bind directly to the gene of interest. In this way, the reverse transcription and amplification can be combined into one. After reverse transcription, the cDNA can be quantified by fluorescent assays such as Quant-iT OliGreen kit (Life Technologies), to provide the same starting amount for each sample.

Theoretically the DNA amount doubles each cycle when PCR-efficiency is 100 %. In practice this efficiency cannot be taken for granted and thus a preceding step-by-step optimization is required. The startup of a PCR experiment begins with an adequate primer design [130, 131]. In general the optimal primer length is about 18 to 30 bp long, with a melting temperature (Tm) of about 60-65 °C which is favorably similar for the two primers of the primer pair. The Tm or temperature where an oligonucleotide duplex is 50/50 double-stranded/single-stranded, is determined by the GC-content of the oligonucleotide (optimal 40-60 %). This parameter is decisive for the choice of annealing temperature in the PCR cycle, which will be 55-60 °C. Mismatches, complementarities within a primer and between the primers need to be avoided. To prevent co-amplification of possibly contaminating gDNA, it is best to design intron-spanning primers, as cDNA only consists of exon sequences in contrast to gDNA (with the exception of retrotransposed pseudogenes, which also lack introns).

Alternatively, intron-flanking primers can be applied, although these still possibly allow for gDNA amplification and are only advantageous if the resulting amplicons are sufficiently longer than the target amplicon. The length of the amplicon of interest is generally set between 80 to 150 bp and also here possible secondary structures need to be avoided. Several software tools are available to assist in primer design and assure the specificity and uniqueness of the primer pair. Some examples are Primer-BLAST from NCBI [132], RTPrimerDB [133], Primer3 [134] or applications designed by qPCR product vendors (ABI, IDT, Eurogentec,…).

The additional mastermix components necessary to bring the PCR reaction to completion are the dNTPs (the four nucleotide triphosphates, alias the building blocks for synthesizing new DNA strands), a thermo-stable Taq polymerase, $MgCl_2$ as a buffer component and co-factor for the DNA polymerase and a fluorophore (probe or intercalating dye) used for quantification. Often a HotStart Taq DNA polymerase is used to reduce aspecific binding during the startup of the reaction. A preceding heating step at 95 °C for several minutes is included to abrogate the activity-block (antibody-based or chemical blocking) of the Taq enzyme.

In quantitative PCR the amount of the product formed is monitored during the reaction by fluorescence [135]. The value used for quantification is determined once the fluorescence intensity reaches a certain threshold i.e. the Cq or threshold cycle (Ct) value (Figure 2.3).

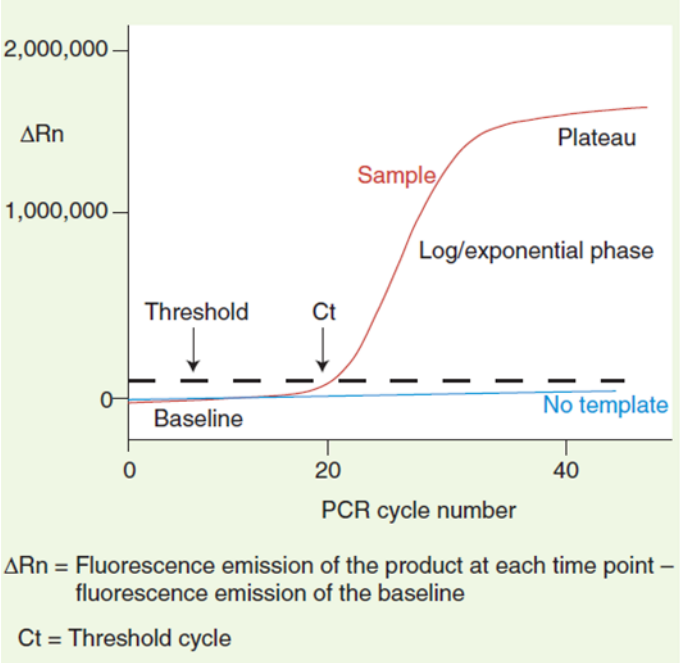


Figure 2.3. Determination of the Cq or Ct value (Image taken from [130]).

To this end different detection-chemistries are possible. In general, two groups can be distinguished: DNA-intercalating fluorescent dyes on one hand and specific fluorescent probes on the other – each with its own advantages and shortcomings [130, 131, 136].

Of the general DNA-binding dyes (Figure 2.4A), SYBR Green is one of the most commonly used reagents. This non-sequence-specific dye binds in the minor groove of double-stranded DNA (dsDNA). It only emits a strong fluorescent signal when bound to dsDNA, which increases during the elongation step and decreases again during denaturation. Alternatively, High Resolution Melting dyes such as EvaGreen can be used.

Because of the reversible nature of this fluorescent signal, it is possible to generate melting curves. For this purpose, at the end of the PCR run an additional heating step with a temperature ranging from 60 to 95 °C is added and fluorescence is measured in function of that temperature gradient. The fluorescence intensity drops upon denaturation at the PCR product's melting point (breakpoint in the melting curve). By transforming this curve to its negative first derivative, the melting point is determined as the maximum of the melting peak. In case of specific PCR products, only one tight peak will be visible; if additional products such as primer-dimers are also formed, the peak will be blurred or extra peaks appear. The fact that non-specific amplification can be visualized in this way adds a considerable advantage to the use of intercalating dyes. Also, as multiple dye molecules can bind to the amplified material, sensitivity is increased. In addition, they are the cheaper option and easy in use, as they do not require a designing step as is the case for specific fluorescent probes. The downside of their aspecificity is that multiplexing is impossible and that they might generate false positive signals.

The latter is not a problem when using fluorescent probes, as in case of aspecific amplification due to e.g. mispriming or primer-dimers, these side-products will not be detected since specific hybridization between probe and template is necessary for fluorescence emission [130]. Useful application of probes does call for adequate design of the probe specifically for the amplicon of interest. In general the same rules as for primer design apply: length of 18-30 bp, no (self-)complementarity, analogous GC-content.

Different types of labeled probes and primers can be applied, both with a double-dye as well as with fluorophores that change their fluorescence properties upon DNA-binding [130, 136]. One of the most frequently used double-dye probes is a so-called TaqMan probe (Figure 2.4B). This hydrolysis probe has both a fluorescent reporter and a quencher covalently attached. After excitation, the fluorophore passes on its energy via Fluorescence Resonance Energy Transfer (FRET) to the quencher, because of their vicinity. During the elongation process the bound probe is cleaved due to the 5'-3' exonuclease activity of the DNA polymerase, so that both dyes are separately released in solution and FRET is

abolished. The emitted reporter signal is detected and fluorescence intensity gradually increases over time. An often used combination is Fluorescein amidite (FAM) as the reporter dye and Tetramethylrhodamine (TAMRA) as the quencher. The diverse possibilities in fluorophore choice allow for multiplex reactions. Since the fluorescent signal when working with probes is irreversible, melting curves and peaks cannot be obtained, but are also less necessary as the probes do not bind aspecific primer products. But, despite their general advantage of high specificity for the target of interest, labeled probes require suitable design and are more expensive in comparison to aspecific DNA-binding dyes.

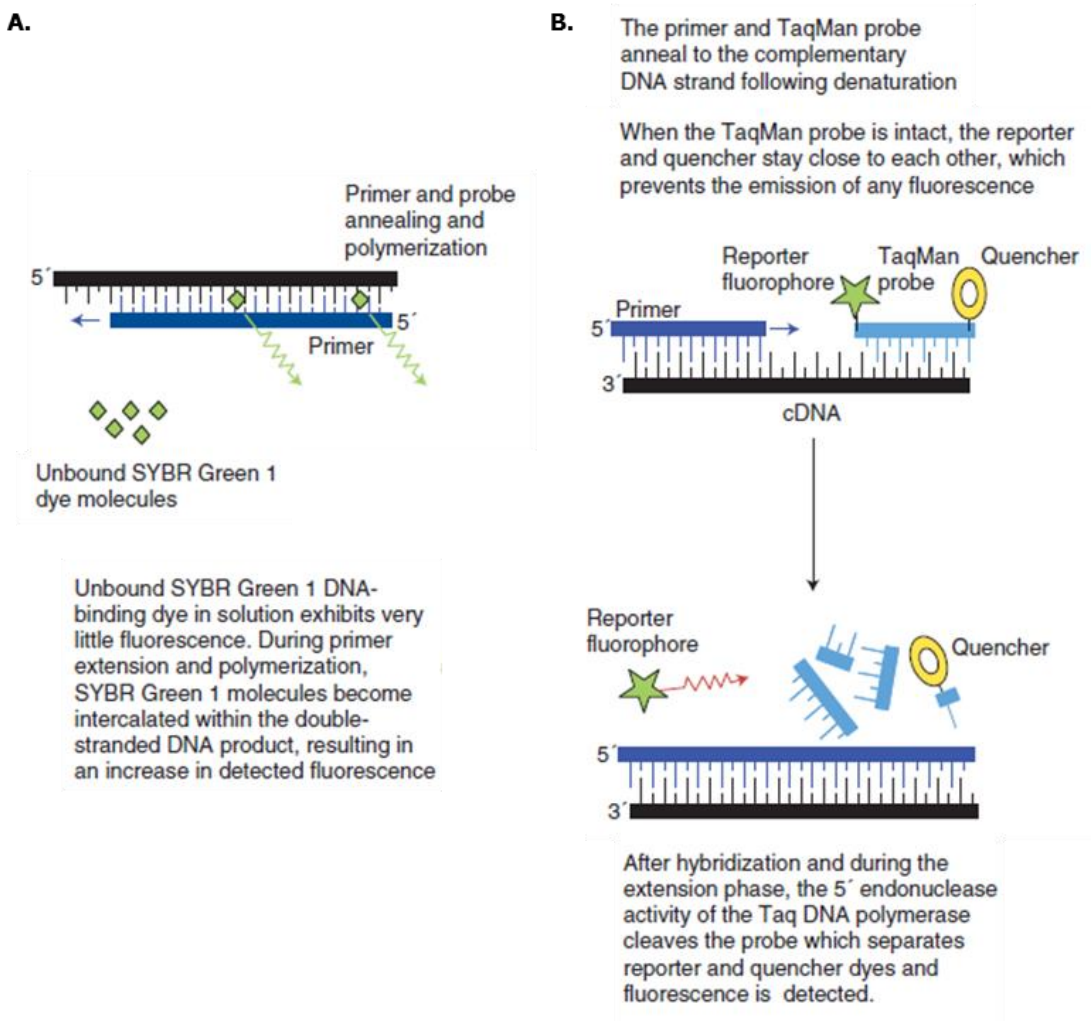


Figure 2.4. Fluorescence reporter mechanisms. A: Non-sequence-specific DNA-binding dyes e.g. SYBR Green, only emit a fluorescent signal when bound to dsDNA. **B:** Hydrolysis probes such as TaqMan probes (Images adapted from [130]).

Most mastermixes also contain a reference dye such as ROX. This dye serves to provide a stable baseline signal and normalizes for non-PCR-related fluctuations in fluorescence.

Before proceeding to the actual quantification, the quality of the qPCR run needs to be evaluated [131]. Negative controls e.g. a no-template, a no-amplification and a no-RT control can be included in the run and provide information on unwanted co-amplification, contamination issues and possible probe degradation. Also positive controls are useful as they give an idea about the quality of the employed reagents. The implementation of those controls, the type of thermocycler available and the normalization strategy all need to be taken into account when developing a PCR plate layout. Sample maximization is preferred over gene maximization, i.e. it is better to load all samples for one gene on one plate instead of running all genes for one sample in one plate. Sample maximization avoids interrun variation, although this is not always possible (in case of prospective studies or too large sample numbers). For that matter, a number of samples can be run repeatedly on each plate, the so-called interrun calibrators and aid to correct for this variability later on during data analysis.

2.1.2. Data analysis

The measured fluorescence is directly correlated with the amount of DNA present in the sample. As such, this can be accurately quantified, both relatively and absolutely. In absolute quantification the exact copy number of a gene of interest (GOI) is calculated by means of a standard curve. In case of relative quantification, as also applied in this thesis, the fold change in expression of a GOI is compared between different samples, whether or not after rescaling to a reference sample (= calibrator) and to a (set of) reference gene(s). The determined Cq values are put in a calculation based on the delta delta Ct (ΔΔCt) method:

$$\Delta Ct = Ct(GOI) - Ct(reference\ gene)$$

$$\Delta\Delta Ct = [\Delta Ct]sample - [\Delta Ct]calibrator$$

$$Fold\ change = 2^{-\Delta\Delta Ct}$$

A fold change value of > 1 indicates an upregulation of the GOI towards the reference, a value < 1 signifies a GOI downregulation.

This calculation however assumes that the PCR efficiency equals 100 %. If there is more than 10 % deviation from this ideal situation, the true efficiency needs to be taken into account. In this case "2" in the latter formula is replaced by the value calculated as follows:

$$Value = \frac{\left[E(GOI)^{\Delta Ct\ GOI(calibrator-sample)}\right]}{\left[E(reference)^{\Delta Ct\ reference\ (calibrator-sample)}\right]}$$

where E is the PCR efficiency. The efficiency itself can be determined by running a sample dilution series and evaluating the slope of the resulting standard curve (see equation below). A 2-fold dilution

gives a ΔCt of 1, a 10-fold dilution gives a difference of about 3.2. As stated above, an efficiency of 90-110 % is acceptable.

$$E = \left(10^{-\frac{1}{slope}} - 1\right) x\ 100$$

Although other normalization methods are available (s*ee Section 2.2*), the most widespread is the implementation of reference genes. Normalization to reference genes provides correction of any technical variability resulting from sample handling, variable input and differing enzymatic efficiencies, brought along in the procedure from sample preparation to the PCR reaction itself. The prerequisite for those references is that they remain stably expressed throughout all samples included in the RT-qPCR analysis. Reference gene stability can be evaluated via several available tools such as geNorm [137], BestKeeper [138] and Normfinder [139]. Multiple reference genes can be taken into account simultaneously and normalization is then based on the geometrical averaging thereof [137, 140]. In fact, the inclusion of multiple references is even preferred, as the use of only a single reference gene results in more erroneous normalization due to the inherent oscillations in expression for each gene. Hence, normalization based on more than one reference gene averages these variations and allows for more accurate measurement of expression levels [137]. Reference gene sets might not be ubiquitously applicable and thus need to be optimized for each experimental set-up. A reference gene optimization for RT-qPCR analysis of hESC samples is described in the second part of this chapter [13], proposing *β-2-microglobulin* (*B2M*), *ribosomal protein L13A* (*RPL13A*) and Alu repeats as suitable references.

It is generally accepted that all relevant information to reproduce experiments should be fully disclosed when publishing PCR data. The Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) guidelines help in providing the essential info [141, 142], although they are not always consequently applied [143].

The newest trend for absolute quantification is digital PCR [144], which is very well suited for detection of low abundant targets. This method provides an absolute measure of nucleic acids and does not even require a standard series. The first step in droplet digital PCR (ddPCR) is partitioning the sample into hundreds or thousands of droplets, representing a large number of reaction chambers. Hereby the input molecules are randomly distributed so that some droplets do contain one or more copies and others do not. The use of fluorescent probes during the following PCR reaction delivers a binary signal for each droplet: "0" for no-template-containing and "1" where a GOI is amplified. By registering the number of positive signals, the original amount can be calculated based on Poisson statistics. Accordingly, also for ddPCR MIQE guidelines have been established [145].

2.2. Reference loci for RT-qPCR analysis of human embryonic stem cells

2.2.1. Abstract

Background: Selecting stably expressed reference genes is essential for proper RT-qPCR gene expression analysis. However, this choice is not always straightforward. In the case of differentiating hESC, differentiation itself introduces changes whereby reference gene stability may be influenced.

Results: In this study, we evaluated the stability of various references during RA-induced (2 µM) differentiation of hESC. Out of 12 candidate references, *β-2-microglobulin* (*B2M*), *ribosomal protein L13A* (*RPL13A*) and Alu repeats are found to be the most stable for this experimental set-up.

Conclusions: Our results show that some of the commonly used reference genes are actually not amongst the most stable loci during hESC differentiation promoted by RA. Moreover, a novel normalization strategy based on expressed Alu repeats is validated for use in hESC experiments.

2.2.2. Background

Human ESC, derived from the inner cell mass of a blastocyst stage embryo, are able to differentiate into all cell types of the adult body and have the potential for unlimited growth [6, 10, 146, 147]. As a consequence of their pluripotency and self-renewal capacity, hESC are ideal for investigating the basic mechanisms of development and cell differentiation. In addition, they may be a source of differentiated cells of a particular cell type, to be used in toxicity screening, cell replacement therapies and many other applications [10, 148].

Pluripotent cells are characterized by several features, such as the expression of pluripotency factors (Oct4 (encoded by *POU5F1*), Nanog (*NANOG*) and Sox2 (*SOX2*)), the presence of specific cell surface antigens (e.g. SSEA-3, SSEA-4, TRA-1-60 and TRA-1-81), and distinct chromatin signatures [9, 10, 12, 82, 149].

To date, the molecular basic mechanisms of (spontaneous) hESC differentiation and development are largely unknown [42, 147]. Differentiation can be induced *in vitro* under specific culture conditions, such as the addition of RA, a morphogen commonly used for multilineage differentiation of ESC in general and for specific development along the neural lineage [37, 41, 150]. Amongst other techniques, RT-qPCR is very well suited for monitoring pluripotency and differentiation, as it allows accurate messenger RNA quantification of numerous samples at the same time [137, 151]. In the context of hESC characterization, RT-qPCR is applied for evaluating the expression of the transcription factors Oct4 and Nanog, since the expression of these core pluripotency circuitry members [12] decreases drastically within a few days after onset of differentiation [41, 42].

For proper RT-qPCR data evaluation, several variables need to be taken into account. These include sample handling and storage, starting material quantity and quality, efficiency of different enzymatic reaction steps and overall transcriptional activity differences between cells [137, 152].

To correct for these variables, different normalization methods have been reported. Gene expression levels can be standardized to cell number, however, it is not always possible to obtain an accurate enumeration of cells. In addition, this strategy does not consider possible insufficient enzymatic reaction efficiencies [137, 153]. Alternatively, data are normalized for RNA mass quantity, although this is not always representative for the mRNA content. Ribosomal RNA molecules make up the major part of the total RNA mass and may be regulated, thus resulting in a variable rRNA/mRNA ratio [137, 152–154]. The most frequently utilized strategy is the inclusion of one or preferably more reference genes as an internal standard. The expression of these references should ideally not vary between cells of interest or as a consequence of experimental handling [137, 152, 153, 155].

Selecting stable reference genes is critical for correct interpretation of RT-qPCR data. However, when studying differentiating hESC, proper reference gene selection is not straightforward. Differentiation does not only include various morphological changes, but also major alterations in gene expression levels of numerous genes. The regulation of some reference genes may be associated with these cellular changes, hence the stability of the used references has to be evaluated. The available differentiation protocols may induce distinct gene expression variability, which impedes finding stably expressed reference genes over the different samples and making protocol-dependent optimization required [42, 137, 152, 153, 155, 156].

In this study, we emphasize the importance of determining suitable reference genes by performing an expression stability analysis for RA induced differentiating hESC, using the geNorm algorithm in the qbase+ software (Biogazelle) [137, 140]. The possibility of co-regulation was reduced by opting for 11 candidate reference genes from different functional categories [137]. In addition to these candidates, a new normalization strategy was applied, based on the measurement of expressed Alu repeats.

Alu insertions are repetitive DNA sequences, approximately 300 bp long and occurring generally at high copy number in introns, 3′ untranslated regions (UTR) of genes and intergenic genomic regions [157]. These short interspersed mobile elements are not equally spread throughout the human genome, since they preferentially accumulate in gene-rich regions [152, 157]. In total, Alu elements comprise more than 10 % of the genome mass, thus being the most abundant of all mobile elements and they are divided in several well-conserved subfamilies (e.g. AluSq, AluSx, AluY) [152, 157, 158]. Alu repeats, named after a recognition site for the restriction enzyme Alu I, are thought to be amplified by retrotransposition, a process in which the Alu element is transcribed by RNA polymerase III, followed by reverse transcription and incorporation into the genome [152, 157–159].

Because of their genome-wide distribution, including in the 3′ UTR of protein-coding genes, individual gene expression variability in the cells of interest will not substantially influence total Alu element expression. This feature makes the Alu repeats a valuable and interesting strategy for RT-qPCR normalization for biological systems such as differentiating stem cells [152, 157, 160].

2.2.3. Methods

Human embryonic stem cell cultures and sample preparation

Human ESC (UGENT1 (XY) and UGENT2 (XX) cell line) were generated in-house [161]. The cells were cultured in 6-well dishes or flasks on a nearly confluent layer of Mitomycin C (Sigma-Aldrich, Steinheim, Germany, #M4287) treated MEF (passage 3), in knock-out DMEM (#41965-039) supplemented with knock-out serum replacement (#10828-010), antibiotics (PenStrep, #15140-122), L-

glutamine (#25030-024), bFGF (#13256-059), non-essential amino acids (#11140-035, all culture medium products purchased from Invitrogen, Carlsbad, CA, USA) and β-mercaptoethanol (Sigma-Aldrich, #M7522) and incubated at 37 °C, 5 % CO_2, 5 % O_2.

Differentiation of hESC was induced by adding 2 μM RA (Certa, Braine-l'Alleud, Belgium, #640327 T) to the culture medium, and eliminating bFGF. Cells were harvested using 0.25 % trypsin-EDTA (Invitrogen, #15596-026) and glass beads (Sigma-Aldrich, #Z265926-1EA). For Experiment 1 hESC (passage 30) were isolated every 24 hours during 8 days after onset of differentiation, plus an extra sample on day 12. In Experiment 2 (passage 43), samples were also collected every 24 hours, during 6 days. For Experiment 3 (passage 32), samples were collected every 4 hours during day 3, 4 and 5 after differentiation was induced. For each time point, approximately 2×10^5 cells were isolated.

The hESC were split two days before the start of a differentiation experiment, using 1 % collagenase (Type IV, Invitrogen, #17104-019) and glass beads.

Microscopy

Phase contrast and bright field images of the hESC culture were acquired with an Axiovert 25 light microscope (Carl Zeiss, Munich, Germany) (objective magnification 5×) and a Sony Alpha 100 camera.

RNA isolation and RNA quality assessment

After isolation, the cells were immediately resuspended in 1 ml of TRIzol (Invitrogen, #15596-026) and stored at −80 °C. For RNA isolation, 200 μL chloroform (Sigma-Aldrich, #C2432) was added to the thawed samples, with subsequent phase separation and purification using an RNeasy Mini kit (Qiagen, Valencia, CA, USA, #74104). After DNase treatment (Qiagen, #79254) and a washing step, RNA was eluted and concentrated using Vivacon 500 spin columns (Sartorius Stedim Biotech, Aubagne Cedex, France, #VN01H32).

RNA quality was assessed for a representative set of samples (6 extra samples taken during Experiment 1) by means of microfluidic capillary electrophoresis. An RNA HighSens Chip (Experion, Bio-Rad Laboratories, Hercules, CA, USA, #7007105) was used to determine a 28 S/18 S rRNA ratio and an RQI, after determining RNA concentrations using a Quant-iT RiboGreen RNA kit (Invitrogen, #R11490). The 28 S/18 S rRNA ratios ranged from 1.59 to 1.78 and the RQIs from 8.7 to 9.9, which indicate good quality samples.

cDNA synthesis

Complementary DNA (cDNA) was synthesized using a Superscript II kit with oligo(dT) primers (Invitrogen, #11904-018). The cDNA concentration was determined with a Quant-iT OliGreen ssDNA

Assay kit (Invitrogen, #O11492), using a spectrophotometer (Tecan, Männedorf, Switzerland). Samples were stored afterwards at −20 °C.

Reverse transcription quantitative PCR

Two different devices were used for RT-qPCR. When using the LightCycler 480 (Roche, Basel, Switzerland), for each reaction 2 µl of cDNA (2.5 ng/µl) was mixed with 3 µl of mastermix, in a 384-well plate. Utilising the ABI Prism 7000 Sequence Detection System (Applied Biosystems, Foster City, CA, USA), each reaction consisted of 5 µl of cDNA (2 ng/µl) and 20 µl of mastermix, in a 96-well plate. The same thermocycling conditions were applied for both systems: 2 min at 95 °C before 45 cycles of 15 sec at 95 °C followed by 1 min at 60 °C. When applicable, an additional heating step from 60 °C to 95 °C was added to obtain melting curves.

The mastermix comprises the primers, iTaq Supermix with ROX (Bio-Rad Laboratories, #1725855) and water. Depending on the locus of interest, probes were included or iTaq Supermix containing SYBR Green (#1725851) was used.

The primers for *ACTB* (Forward: AGAAAATCTGGCACCACACC; Reverse: TAGCACAGCCTGGATAGCAA, SYBR Green detection), and primers and 6-FAM probes for *PPIA* (F: CAAATGCTGGACCCAATACAAA; R: GCCATCCAACCCCTCAGTCT; Probe: TGTTCCCAGTGTTTCATCTGCACTGCC) and *GAPDH* (F: AGCCTCAAGATCAGCAATG; R: ATGGACTGTGGTCATGAGTCCTT; Probe: CCAACTGCTTAGCACCCCTGGCC) were designed and validated in-house (obtained from Applied Biosystems). These primers were applied at a concentration of 300 nM. The primer sequences for the remaining references (final concentration 250 nM, all detected with SYBR Green) are available in the RTPrimerDB database (http://www.rtprimerdb.org) [133]: *B2M* (RTPrimerDB ID #2), *HMBS* (#4), *HPRT1* (#5), *RPL13A* (#6), *SDHA* (#7), *UBC* (#8), *YWHAZ* (#9). The primer sequences for *TBP* are described in [32]. The sequence of the Alu repeats primers is CATGGTGAAACCCCGTCTCTA for the forward primer and GCCTCAGCCTCCCGAGTAG for the reverse primer. TaqMan assays (Applied Biosystems) were used for the analysis of *POU5F1* (Hs01895061_u1) and *NANOG* (Hs02387400_g1). All reactions were performed in duplo and no template controls were included for all genes. All primer efficiencies lie within the range of 90 % to 110 %.

Data analysis

Stability analysis of the different references was performed using the geNorm application in the qbase+ software version 2.0 (Biogazelle) [137, 140]. Relative quantification of the pluripotency markers data (Oct4 and Nanog) was calculated using the qbase+ software version 2.0. Each sample was rescaled to a calibrator, in this case undifferentiated hESC (day 0), and was normalized for three

reference loci; for *GAPDH*, *ACTB* and *PPIA* on the one hand, or *B2M*, AluSq and *RPL13A* on the other hand.

The relative quantification data for both normalization strategies were statistically analyzed performing an analysis of variance (ANOVA) in R (version 2.13.1).

2.2.4. Results

Human ESC were induced to differentiate for several days by addition of RA to the culture medium. The fading undifferentiated state of hESC was assessed morphologically, using light microscopy. Differentiation was initially visible at the colony periphery where cells start to pile-up and in comparison to undifferentiated cells, differentiating colonies lost their round shape with well-defined borders, as illustrated in Figure 2.5.

To confirm differentiation, the expression levels of pluripotency genes *POU5F1* and *NANOG* can be followed over time using RT-qPCR. For this purpose, the most stable normalization references were determined by means of a stability analysis using the geNorm application in qbase+. The stability of 12 candidate references (see Table 2.1 for gene names and function) was analyzed for a first series of samples, which were isolated every 24 hours during 8 days of differentiation (= Experiment 1). In the stability ranking, *B2M*, *RPL13A* and Alu repeats (AluSq) were found to be the most stable reference loci for this experimental set-up (reference stability ranking and accompanying M values displayed in Table 2.2). The stability measure of these three references varied from 0.277 to 0.290, which indicates a high stability in comparison to the other candidates, as M values up to 0.500 are considered acceptably stable for homogenous samples [140]. In the determination of the optimal reference number, two appeared to be sufficient for RT-qPCR normalization, as the pairwise variation (V value) was approximately 0.09 for upgrading from two to three reference loci (V2/3). V values less than 0.15 indicate that increasing the number of references would not add any more significant value to the assay.

Table 2.1. Reference loci and target genes included in the experiments.

Reference	Name	NCBI RefSeq	Function	Amplicon size (bp)	Exon location
ACTB*	β-Actin	NM_001101	Cytoskeletal protein	173	7p22
B2M[§]	β-2-microglobulin	NM_004048	Major histocompatibility complex I component	86	15q21-q22
GAPDH*	Glyceraldehyde-3-phosphate dehydrogenase	NM_002046	Oxidoreductase in glycolysis and gluconeogenesis	111	12p13
HMBS*	Hydroxymethyl-bilane synthase	NM_000190	Porphyrin metabolism and heme synthesis	64	11q23
HPRT1*	Hypoxanthine phosphoribosyl-transferase 1	NM_000194	Purine synthesis in salvage pathway	94	Xq26
PPIA*	Peptidylprolyl isomerase A (Cyclophilin A)	NM_021130	Catalyzation of cis-trans isomerization of proline imidic peptide bonds	71	7p13
RPL13A*	Ribosomal protein L13A	NM_012423	60S ribosomal subunit structural component	126	19q13
SDHA*	Succinate dehydrogenase complex, subunit A	NM_004168	Electron transporter in citric acid cycle and respiratory chain	86	5p15
TBP*	TATA box binding protein	NM_003194	General transcription factor for RNA polymerase II	89	6q27
UBC*	Ubiquitin C	NM_021009	Involved in protein degradation	133	12q24
YWHAZ[§]	Tyrosine 3-monooxygenase / tryptophan 5-monooxygenase activation protein, zeta polypeptide	NM_003406	Signal transduction; binds to phosphorylated serine residues on several signaling molecules	94	2p25
AluSq	Alu repeats, subfamily Sq				
POU5F1[§]	POU class 5 homeobox 1, Oct4	NM_002701.4	Marker for embryonic stem cell pluripotency	130	6p21
NANOG*	Nanog homeobox	NM_024865.2	Marker for embryonic stem cell pluripotency	109	12p13

* exon-spanning primer pair; [§] forward and reverse primer localized in the same exon

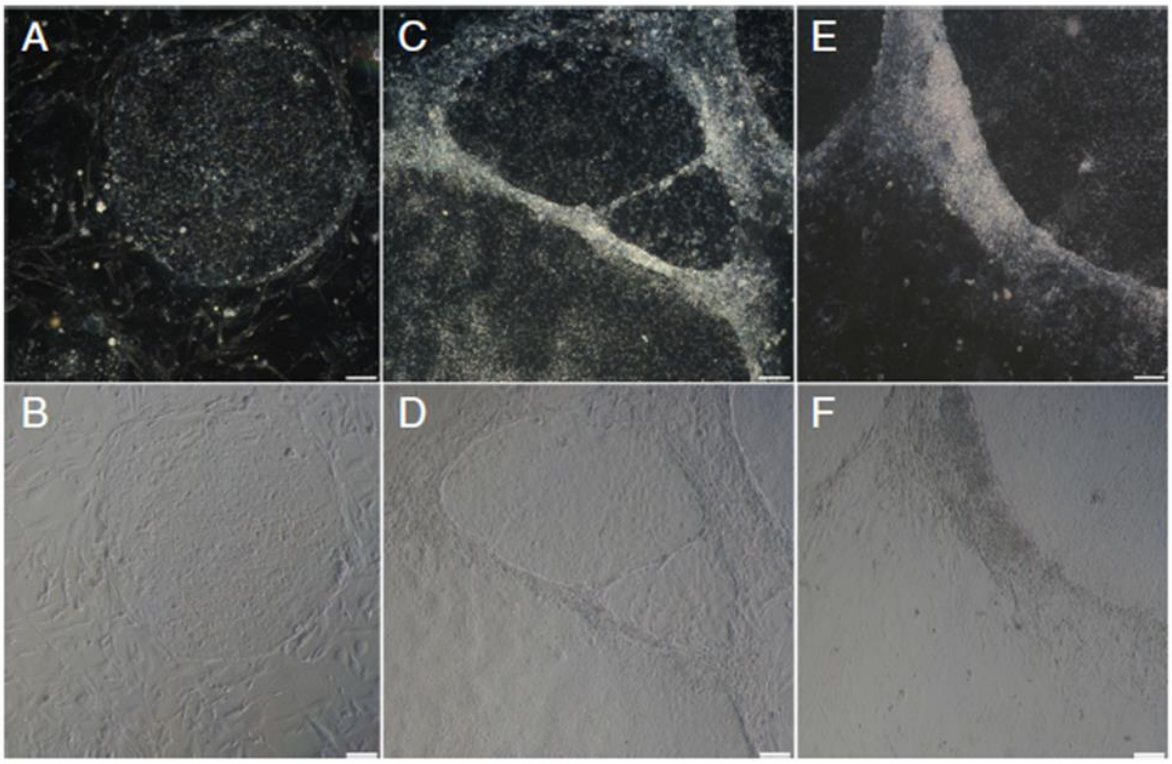


Figure 2.5. Morphological evaluation of human embryonic stem cell culture. Comparison of an undifferentiated colony (**A:** phase contrast – **B:** bright field image) with differentiated hESC on day 3 (**C**, **D**) and day 5 (**E**, **F**) after RA induction (scale bars = 100 µm).

Table 2.2. Reference gene stability analysis.

	Experiment 1		Experiment 2		Experiment 3	
Ranking	*Reference*	*M value*	*Reference*	*M value*	*Reference*	*M value*
1	**B2M**	0,277	**B2M**	0,203	**B2M**	0,378
2	**AluSq**	0,284	**AluSq**	0,213	**RPL13A**	0,378
3	**RPL13A**	0,290	**RPL13A**	0,221	**AluSq**	0,386
4	HPRT1	0,472	**GAPDH**	0,267	**PPIA**	0,480
5	YWHAZ	0,593	SDHA	0,360	**GAPDH**	0,609
6	TBP	0,696	TBP	0,409	HPRT1	0,672
7	HMBS	0,886	YWHAZ	0,451	TBP	0,813
8	**ACTB**	1,151	**ACTB**	0,480	YWHAZ	0,903
9	**GAPDH**	1,350	HPRT1	0,506	HMBS	1,088
10	**PPIA**	1,476	**PPIA**	0,524	SDHA	1,216
11	SDHA	1,598	HMBS	0,550	**ACTB**	1,347
12	UBC	1,697	UBC	0,579	UBC	1,484

Reference gene stability analysis was analyzed for three independent hESC differentiation experiments, applying the geNorm algorithm. Results are displayed with decreasing stability from top to bottom.

As a confirmation for this first experiment, a reference stability analysis was performed during two more hESC differentiation experiments. For experiment 2, samples were collected every 24 hours during 6 days, and a third series was collected every 4 hours during day 3, 4 and 5 after onset of differentiation (= Experiment 3). Again, the same three reference loci were found to be the most stable in both experiments (Table 2.2). The M values for *B2M*, *RPL13A* and AluSq varied from 0.203 to 0.221, and from 0.378 to 0.386 for Experiment 2 and 3 respectively (all M values can be found in Table 2.2). Also in both cases, two references were shown to be enough for normalization, considering the low V2/3 values.

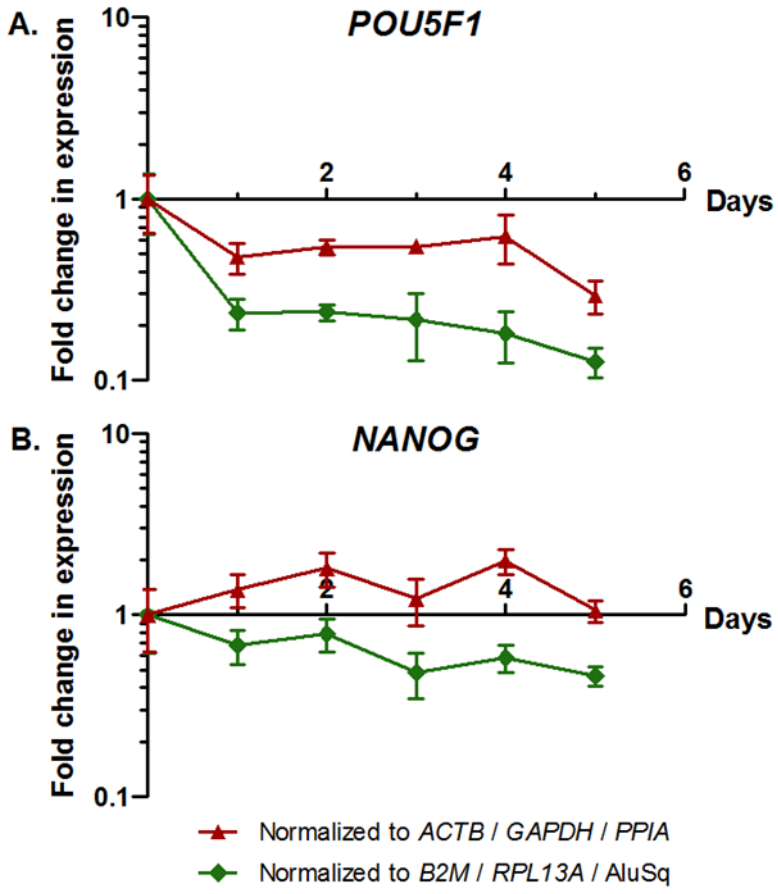


Figure 2.6. Relative quantification of *POU5F1* and *NANOG*. The importance of proper reference gene selection is illustrated by comparing normalized data of the pluripotency factors Oct4 and Nanog to different reference sets. The Cq values for the pluripotency genes *POU5F1* and *NANOG* were both normalized following the ΔΔCt-method (applying qbase+ software), against the geometric average of two different reference sets: on the one hand for *ACTB*, *GAPDH* and *PPIA* (red triangle); on the other hand for *B2M*, *RPL13A* and Alu repeats (green rhomb). The resulting values represent the fold change in expression levels of the pluripotency factors, and are displayed for *POU5F1* and *NANOG* for both normalization methods in respectively panel **A** and **B**. Data shown are obtained from differentiation Experiment 2 (sample isolation every 24 h during 6 days of differentiation) (mean and standard deviation of 2 replicates).

Subsequently, the expression data of *POU5F1* and *NANOG* were normalized applying two different reference sets: relative quantification using three commonly used genes (*GAPDH*, *ACTB* and *PPIA*) versus the three most stable references determined in the analyses described above (*B2M*, *RPL13A* and AluSq). As can be expected, this comparison revealed a substantial difference in the change of expression levels of *POU5F1* and *NANOG*, emphasizing the importance of proper reference gene selection. As illustrated for Experiment 2 in Figure 2.6, the decrease in expression of these pluripotency factors is significantly less pronounced using the 'traditional' reference genes (*GAPDH*, *ACTB* and *PPIA*) than with the most stable reference loci as defined in this study (*B2M*, *RPL13A* and AluSq) (ANOVA, p-value = 1.30e-05). Also for the other two experiments a significant difference was seen (p-value = 0.0034 and = 0.0215 for Experiment 1 and 3 respectively).

2.2.5. Discussion

Human ESC provide a unique opportunity to study early development and may hold great potential for regenerative medicine [10, 41, 148, 149]. The transcription factors Oct4 (*POU5F1*) and Nanog (*NANOG*) contribute to self-renewal and are required for maintaining the pluripotent state of hESC [12, 82]. Therefore, the expression of these factors is commonly assessed with RT-qPCR, as mRNA levels decrease substantially within a few days after inducing differentiation [41, 42]. Morphological evaluation of the differentiating cells shows that hESC start to accumulate at the edges of the colonies and that individual cells adopt a more lengthened, neuronal-like phenotype during early hESC differentiation (Figure 2.5) [33].

Adequate RT-qPCR normalization is essential for valid data interpretation. However, *in vitro* hESC differentiation entails massive gene expression alterations in general and specifically due to the differentiation agent itself, whereby the expression of individual reference genes may vary, thus making it difficult to select the most suitable and stable references [42, 151]. Synnergren and colleagues (2007) already made note of a unique reference gene expression pattern when differentiating hESC spontaneously [156]. And as Willems et al. (2006) also showed, normalization results vary substantially depending on the reference used [155].

In this study, the suitability of 12 different references was evaluated using the geNorm algorithm [137, 140]. From our results, it can be concluded that *B2M*, *RPL13A* and Alu repeats (AluSq) are the most stable reference loci for this specific differentiation protocol. The use of two references was shown to be sufficient for accurate normalization of RT-qPCR data, though generally the use of three stable references is recommended in literature [137].

The importance of selecting the most stable and suitable references is illustrated by normalization of gene expression levels of pluripotency factors Oct4 and Nanog. Comparison of *B2M*, *RPL13A* and AluSq with more 'traditional' reference genes, resulted in a significantly different normalization, indicating that classic reference genes such as *GAPDH* are not always appropriate for a given set-up. In the field of stem cell differentiation, optimization is required for each specific differentiation protocol.

Despite the comparability of different algorithms for determining reference stability (Normfinder, Bestkeeper, Comparative Delta-Ct method), their application may result in a slightly different stability order in comparison with the geNorm applet. This may indicate though that different reference sets are applicable, resulting in analogous normalization data. When applying those algorithms to the results of Experiment 1, Alu repeats and *RPL13A* remain among the Top 3 stable reference loci (data not shown). *B2M* deteriorates a few ranks when using Normfinder and Bestkeeper, and is replaced by *TBP* as a more stable reference. Nevertheless, the comparison of normalization data for *B2M*, *RPL13A* & AluSq and *TBP*, *RPL13A* & AluSq, gave no significant difference (p-value > 0.5). In conclusion, *B2M*, *RPL13A* and AluSq are suitable reference loci for this experimental set-up of RA induced hESC differentiation.

β-2-microglobulin (*B2M*) is a component of major histocompatibility complex I, hence being expressed in every nucleated cell, and has already been applied before as a normalization scalar in different set-ups [162–164]. Ribosomal protein L13A (*RPL13A*) is involved in the process of transcript translation, and has also been widely included as a reference gene for RT-qPCR analyses [165–168], in spite of possible presence of pseudogenes [169]. However, pseudogene detection does not necessarily imply that a specific reference is not usable. The main point of interest is the reference stability, which in this case is clearly maintained as supported by the data described above. In our hands both *B2M* and *RPL13A* repeatedly do come forward as a stable reference, supporting their use as normalization genes.

The fact that Alu repeats were one of the best scoring reference loci in the described analyses is not surprising. Because of their genome-wide distribution, they can be considered as a measure for the total amount of mRNA, and the overall Alu element expression will not be influenced by a variation in expression of individual genes [152, 157, 170]. For this reference, primer specificity is of minor importance. The more sequences are detected by the assay, the less the impact of individual expression variations on the total Alu content. Hence, Alu repeats provide a new strategy for reliable normalization of RT-qPCR data, in particular in experiments where dramatic changes are expected. An additional advantage when working with limited amounts of starting material, is that the input can be lowered, since Alu repeats are highly expressed and thus lead to low Cq values [152].

2.2.6. Conclusions

This study shows that some of the commonly used reference genes cannot always be included as a stable normalization scalar. Selection of suitable references is highly dependent on the experimental set-up, as is illustrated here for early hESC differentiation induced by RA. Furthermore, a new normalization strategy based on Alu repeat expression is proposed and validated for hESC (RA induced) differentiation experiments.

CHAPTER 3:

HISTONE CLIPPING

IN HUMAN EMBRYONIC STEM CELLS

Adapted from "**Identification of histone H3 clipping activity in human embryonic stem cells**"

Liesbeth Vossaert*, Paulien Meert*, Ellen Scheerlinck, Pieter Glibert, Nadine Van Roy, Björn Heindryckx, Petra De Sutter, Maarten Dhaenens and Dieter Deforce

Stem Cell Research (2014) 13 (1), 123-134

* Equal contribution

3. HISTONE H3-CLIPPING ACTIVITY IN HUMAN EMBRYONIC STEM CELLS

3.1. Abstract

Histone PTMs are essential features in epigenetic regulatory networks. One of these modifications has remained largely understudied: regulated histone proteolysis. In analogy to the histone H3 clipping during early mouse ESC differentiation, we report for the first time that also in human ESC this phenomenon takes place in the two different analyzed cell lines. Employing complementary techniques, different cleavage sites could be identified, namely A21, R26 and residue 31. The enzyme responsible for this cleavage is found to be a serine protease. The formation of cleaved H3 follows a considerably variable pattern, depending on the timeframe, culture conditions and culture media applied. Contrary to earlier findings on H3 clipping, our results disconnect the link between declining Oct4 expression and H3 cleavage.

3.2. Background

In all eukaryotes, DNA is tightly associated with histone proteins in order to form chromatin, of which the fundamental subunit is the nucleosome [84]. Each nucleosome consists of four different core histone types (H2A, H2B, H3 and H4), which have been very well evolutionarily conserved. Chromatin structure is essential for compaction of genomic DNA but also represents a physical barrier to control DNA accessibility and gene expression.

In ESC the delicate balance of self-renewal and differentiation into specific lineages is determined by many lineage-restricted promoters that are associated with highly combinatorial histone PTM patterns which may determine their selective priming of gene expression during lineage commitment. Together with DNA methylation, ATP-dependent chromatin remodeling, RNA interference, non-coding RNA and incorporation of histone variants, these properties form the "epigenetic signature" [72].

Not only ESC differentiation but also other biological contexts are characterized by a continuous interplay of installation and removal of histone PTMs. To accomplish the latter, several mechanisms can be at play. Apart from enzymatic elimination of modifications [72, 99] and histone exchange [72, 171, 172], also regulated proteolytic histone cleavage has been suggested to play such role [127, 173].

Duncan et al. showed that H3 is proteolytically cleaved at its N-terminus during early differentiation of mouse ESC (mESC) and they provide evidence for the regulatory capacity of covalent modifications herein [127]. Cathepsin L was found to cleave histone H3, with alanine 21 being the primary site of cleavage [127, 174]. This truncated H3 form is detected during the first days of both monolayer differentiation (with and without RA induction) and embryonic body (EB) formation. Similar clipping events of H3 associated with other cellular processes including viral infection [175, 176], aging [177, 178] and sporulation [179] have also been reported. Additionally, H3 protease activity was also found in chicken liver and *Tetrahymena* micronuclei [180–183]. Although the molecular consequences of any histone clipping event are yet to be defined, these data seem to suggest an evolutionary conserved process.

Here, we show for the first time that histone H3 clipping also occurs in human ESC lines (hESC) in addition to mESC. Several cleavage sites were assigned and the clipping enzyme was characterized as a serine protease. The manifestation of this proteolytic event can theoretically have an impact on several levels such as pluripotency and differentiation. Our results indicate that H3 cleavage can indeed be accompanied by a loss of Oct4 expression but does not exclusively occur in this condition. Hence, a direct correlation between these two processes cannot be claimed.

3.3. Methods

Cell culture

Two human ES cell lines were used: the UGENT2 cell line (XX), created in-house and the WA01 Oct4-eGFP knock-in reporter cell line (XY), obtained from WiCell [184]. Both cell lines were cultured (5 % O_2 and 5 % CO_2 at 37 °C) on a feeder layer of Mitomycin C inactivated MEF. Cells were passaged every 4 to 6 days, using 1 % collagenase type IV and glass beads. Culture medium consisted of knock-out DMEM supplemented with 2 mM L-glutamine, 1 % non-essential amino acids, 20 % knock-out serum replacement, 4 ng/ml bFGF, 100 U/ml penicillin and 100 µg/ml streptomycin.

Feeder-free cultures were maintained on a Vitronectin XF coating (Primorigen), in combination with Essential 8 medium (E8, Life Technologies). Cultures were split every 2 to 3 days by means of EDTA-passaging.

Differentiation was induced by omitting bFGF from the medium and by adding RA to the culture medium at a final concentration of 2 µM.

Fluorescence microscopy

Fluorescence microscopy images of eGFP expression (ex. 485 nm, em. 515 nm, exposure time 5000 ms) were acquired on an Axiovert 200M inverted fluorescence microscope equipped with the Axiovision multichannel fluorescence module and an AxioCam MRM camera (Carl Zeiss). Colonies (without auto-fluorescent medium) were screened at 10x magnification using a Carl Zeiss short distance Plan-Acromat objective and visualized using Zeiss filter set n° 38 (BP 470/40, FT 495, BP 525/50). For larger colonies, several images were stitched by use of Photoshop CS4 on TIFF images (Adobe).

Acid histone extraction

Cells were isolated using 0.25 % trypsin-EDTA. After a washing step, cells were resuspended in Triton extraction buffer (TEB; 0.5 % Triton X 100 and 0.02 % NaN_3 in PBS) at a cell density of $1x10^7$ cells/ml, and incubated for 10 min on ice. Subsequently the cells were centrifuged at 1500 rpm for 10 min at 4 °C, and resuspended in half the volume of TEB. After centrifuging again, the cells were resuspended in 0.2 N HCl at a cell density of $4x10^7$ cells/ml and incubated overnight at 4 °C, after which the supernatant containing the histones, was isolated. Protease inhibitors (Roche, 11836170001) were freshly added to each buffer for every experiment, unless stated otherwise. Protein concentration of the extracts was measured by Bradford/Coomassie assay.

Direct boiling in Laemmli buffer

Direct boiling in SDS-loading dye for protein isolation was performed. Human ESC were directly boiled (3 min at 100 °C) after harvesting with trypsin, in 2X Laemmli buffer (4 % SDS, 20 % glycerol and 10 % β-mercaptoethanol in 50 mM Tris (pH 6.8)).

Gel electrophoresis and Western blotting

Per sample the required amount of histone extract was vacuum-dried and dissolved in 2X Laemmli buffer, separated on a 15 % Tris-HCl gel (BioRad Laboratories) and tank-blotted on a nitrocellulose or a PVDF membrane. The membrane was incubated overnight at room temperature with the appropriate antibody. The following antibodies were purchased from commercial vendors: a C-terminus directed histone H3 Ab (1:1000; Abcam, ab10799), an N-terminus directed histone H3 Ab (1:1000; Merck Millipore, 05-499) and an antibody specific for the H3 N-terminus when cleaved after A21 (H3.cs1 Ab; 1:1000; Active Motif, 39573). All blots were detected by chemiluminescence using a Versadoc imaging system (BioRad Laboratories). Biotinylated histone H3 (as described below) was specifically visualized using HRP-conjugated avidin (45 min incubation; 1:100000; eBioscience, 18-4100-94). When needed, blots were stripped by incubating at 50 °C for 4 to 6 hours in stripping buffer containing 2 % SDS, 0.1 M β-mercaptoethanol, 0.05 M Tris (pH 6.8).

In-gel digest of propionylated histones

Vacuum-dried histones were propionylated as described [185] and separated by gel electrophoresis. Sypro Ruby (Invitrogen) stained gel pieces were cut out and subsequent in-gel digestion was performed as described before [186], with only a slight modification to the protocol: 25 mM ammonium bicarbonate buffer was used and the alkylating agent applied was iodoacetamide (100 mM). After peptide extraction out of the gel pieces, samples were vacuum-dried and a second round of propionylation was completed to propionylate the newly generated N-termini.

Flow cytometry

Oct4-eGFP levels were monitored after cell isolation and resuspension in flow buffer (1 % BSA, 0.1 % NaN_3 in PBS) with an FC500 (Beckman Coulter) using the CXP analysis software.

Reverse transcription – quantitative PCR

RNA isolation, cDNA preparation and RT-qPCR analysis were performed as described before [13]. In short, cells were suspended in Trizol after isolation and stored at -80 °C. After reverse transcription (SuperScript II kit, Invitrogen), qPCR analysis was performed for pluripotency gene *POU5F1* (Taqman assay Hs01895061_u1) with normalization to an optimized pool of 3 reference loci, namely *B2M*

(RTPrimerDB ID #2), *RPL13A* (#6) and AluSq (Forward: CATGGTGAAACCCCGTCTCTA – Reverse: GCCTCAGCCTCCCGAGTAG) (all SYBR Green assays).

Histone biotinylation

One mg of purified histone H3, isolated from calf thymus (Roche), was biotinylated using the EZ-Link Sulfo-NHS-Biotinylation kit (Thermo Scientific, 21425), according to the manufacturer's conditions. After the biotinylation reaction, the excess of biotin-label was removed by addition of 5 % hydroxylamine (HA)-solution in a concentration of 6 µl HA per 100 µl sample.

N-terminal sequence analysis

A histone extract taken one day after induction of differentiation and known to contain cH3 as validated by Western blotting, was blotted onto PVDF and proteins were visualized using Ponceau S staining. Subsequently the cleaved area was cut and subjected to Edman degradation, which was outsourced to Eurosequence (Groningen, The Netherlands) and performed according to standard procedures.

Mass spectrometry method

Propionylated peptides were dissolved in 0.1 % formic acid in water (buffer A) and separated on a PepMap 100 (C18) column (I.D. 75 µm, length 25 cm, particle size 5 µm) by use of a U3000 LC-system (Dionex) at a flow rate of 300 nl/min. Elution was performed with 80 % acetonitrile/0.1 % formic acid (buffer B) using a gradient of 10 % to 60 % buffer B in 60 min. A Q-TOF Premier mass spectrometer (Waters) was operated in the data-dependent mode, using a nano-ESI source with resolution of 10.000. Survey MS scans were acquired (*m/z* 425-1300) and up to 7 precursors (*m/z* 50-2300) with charge state 2+, 3+ or 4+ exceeding the signal threshold were isolated for fragmentation by collision induced dissociation, using the collision energy profile as suggested by the manufacturer. An inclusion list contained precursor *m/z* value 574.3 since preliminary experiments suggested that this is a cleavage site fragment.

Mass spectrometry data analysis

Database searching was performed against a custom-made database containing human histone sequences obtained from the National Center for Biotechnology (NCBI) database, using a Mascot 2.3 in house server (Matrix Science). Mass error tolerances for the precursor ions and its fragment ions were set at 0.35 Da and 0.45 Da respectively. Enzyme semi-specificity was set to Arg-C, allowing for up to two missed cleavage sites. Variable modifications included acetylation, dimethylation and propionylation on lysine (K), methylation and dimethylation on arginine (R) and oxidation of

methionine. Lysine monomethylation was searched as the sum of propionylation and methylation since monomethylated lysine residues can still be propionylated. N-terminal propionylation and carbamidomethylation of cysteine residues were set as fixed modifications. Redundant peptides were filtered including only the highest scoring match under the highest scoring protein containing that match. Low confidence identifications were excluded using an expectancy value cut-off of 0.05. Nevertheless all of the used spectra were additionally manually validated using Mascot Distiller software (Matrix Science).

Nuclear extraction and protease inhibitor experiments

Nuclear extracts were prepared using the commercially available Episeeker Nuclear Extraction Kit (Abcam, ab113474), omitting protease inhibitors during extraction. Calf H3 (1 µg) and nuclear extract (0.5 µg) were incubated together in buffer (150 mM NaCl, 2 mM β-mercaptoethanol, 0.1 mM EDTA, 10 % glycerol and 25 mM Tris-HCl (pH 7.5)) at 37 °C for 1 hour, after which the assay was inactivated by incubation at 99 °C for 10 min. All samples were analyzed by Western blotting.

For the inhibition assay the following commercially available inhibitors were used (Figure 3.8B): 4-(2-Aminoethyl)-benzenesulfonyl fluoride hydrochloride (AEBSF, Sigma-Aldrich, 76307), E64 (Sigma-Aldrich, E3132), EDTA (Sigma-Aldrich, E5134), Bestatin (Sigma-Aldrich, B8385), Pepstatin A (Sigma-Aldrich, P5318), protease inhibitor cocktail tablets (PIC, Roche, 11836170001) and three specific cathepsin L inhibitors (I (219421), III (219427) and CAA0225 (219502), purchased from Calbiochem. Inhibitors were pre-incubated with the nuclear extract for 15 min at 37 °C prior to addition of the H3 substrate. Their inhibitory capacity was validated and concentrations yielding aspecific inhibition were excluded from the assay.

3.4. Results

3.4.1. Feeder-Free cultured hESC show continuous histone H3 cleavage during differentiation

To verify whether the histone H3 N-tail is cleaved during human embryonic stem cell differentiation in analogy to mESC [127], two hESC lines were monitored: an Oct4-eGFP reporter hESC cell line that expresses eGFP under the control of the Oct4 promoter and a non-reporter cell line UGENT2. Both cell lines were cultured in feeder-free conditions during 5 days after induction of differentiation by addition of 2 µM RA to the culture medium, in the absence of bFGF. Cells were isolated every 24h and histone extracts were prepared.

The differentiation status during the experiments was validated for both cell lines. Morphologically each cell line displayed clear traits of differentiation, resulting in more lengthened cells and loss of

round colony shape (data not shown). To further validate the efficiency of differentiation, Oct4-eGFP reporter cells were directly monitored by both flow cytometry and fluorescence microscopy [43]. Flow data showed a clear drop in eGFP signal through time (Figure 3.1A), and microscopy imaging visualized a sustained but definite reduction in eGFP throughout the hESC colonies (Figure 3.1B). To monitor loss of pluripotency in the conventional non-reporter stem cell line (UGENT2), RT-qPCR analysis of the pluripotency gene *POU5F1* encoding for the transcription factor Oct4, coordinately confirmed differentiation in these cells (Figure 3.1C).

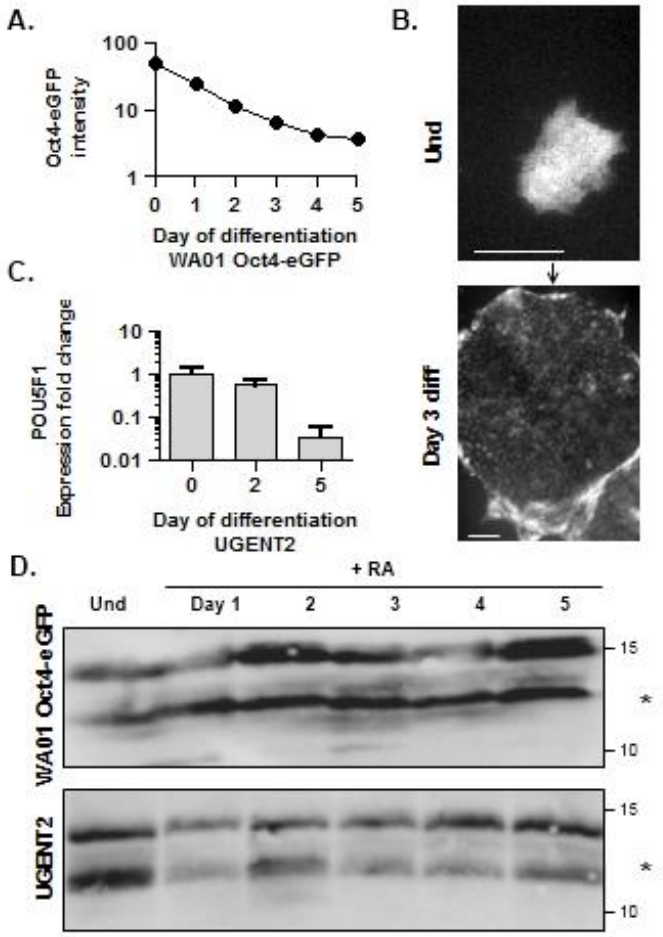


Figure 3.1. Histone H3 cleavage in human embryonic stem cells under feeder-free culture conditions. Human ESC samples were collected during five days after differentiation induction with RA. Differentiation was confirmed by **A:** flow cytometry Oct4-eGFP analysis for the Oct4-eGFP cell line. In addition, **B:** fluorescence microscopy data affirmed a drastic drop in Oct4-eGFP signal when comparing an undifferentiated colony (Und) and a colony after three days of differentiation (scale bars represent 200 µm). Also **C:** RT-qPCR analysis illustrated the decrease in *POU5F1* expression (encoding for Oct4) over time for the UGENT2 cell line (mean and standard deviation of 2 replicates). **D:** Western blot analysis (2 µg histone extract per sample) using a C-terminal H3 antibody reveals N-terminal H3 clipping (cleaved H3 indicated with an asterisk) in both the Oct4-eGFP reporter (upper image) and UGENT2 cell line (lower image).

Samples were monitored for histone H3 cleavage by means of Western blotting analysis, using an H3 C-terminally directed antibody. For both cell lines, lower molecular weight bands of histone H3 (indicated with an asterisk) were visualized at each time point throughout the differentiation experiment, indicating N-terminal cleavage of H3 (Figure 3.1D).

3.4.2. Continuous histone H3 cleavage in the Oct4-eGFP reporter hESC line is not related to the Oct4 expression level

The results above differ substantially from the data reported for differentiating mESC, which display a pattern with upcoming cleaved H3 (cH3), reaching a maximum intensity after 4 days, and decreasing again towards the end of the experiment [127]. The continuous H3 clipping found in feeder-free cultured hESC could be best explained by assuming that the timeframe was not extended enough to demonstrate the disappearance of cH3. The cH3 pattern might also be influenced by the speed and heterogeneity of differentiation, as also shown by Duncan et al., where spontaneous and RA induced differentiation and EB formation all lead to different cleavage patterns in terms of timescale. Therefore we repeated the experiment with the Oct4-eGFP reporter cell line, monitoring the hESC in feeder-free culture for a longer period of time (14 days), and subjected them to three different methods of differentiation to vary the speed at which they lose stemness throughout this timeframe: (i) in their undifferentiated status by culturing the cells in E8 medium, (ii) inducing spontaneous differentiation by omitting bFGF from the culture medium and (iii) enforcing directed differentiation by addition of RA. Morphological changes were observed using light microscopy (data not shown), global Oct4 expression and differentiation were monitored using flow cytometry. According to the flow plots shown in Figure 3.2A, the three used culture conditions indeed differ greatly in their (non)maintenance of stemness throughout the experiment. When analyzing the hESC, it is clear that E8 medium promoted the undifferentiated status in contrast to both differentiation methods which led to a decreased Oct4 expression. While spontaneous differentiation divided the culture system into two groups with loss of the Gaussian distribution in the flow histogram, RA induced a collective decrease in Oct4 expression of all cells resulting in a preserved Gaussian distribution shifting towards lower fluorescence. Also, differentiation induced by RA occurred much faster and a minimal Oct4 expression was reached after 12 days of differentiation while spontaneous differentiation continued at a slower pace and did not reach a minimum limit in Oct4 expression after 14 days of differentiation.

Despite the clear differences in Oct4 expression as described above, all time points in each culture condition appeared to display clear N-terminal H3 clipping when monitored with Western blotting (Figure 3.2B).

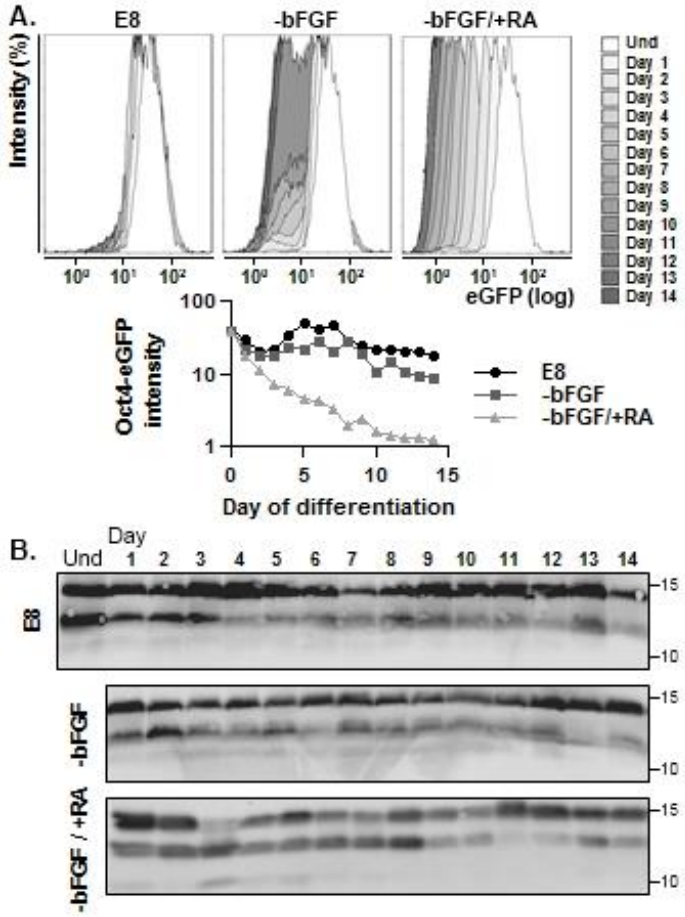


Figure 3.2. H3 cleavage in feeder-free cultured hESC under application of different culture media within an extended time frame. The Oct4-eGFP reporter cell line was cultured for 14 days in different conditions: (i) maintaining the pluripotent state in E8 medium, (ii) differentiating spontaneously (in the absence of bFGF) and (iii) differentiating after RA induction. **A:** Flow cytometry analysis displays a definite reduction in Oct4-eGFP levels under differentiation-allowing circumstances. The flow histogram plots show a variable shift in Oct4-eGFP signal under differentiating conditions: whilst the signal in RA-stimulated cells decreases collectively for the total population, the spontaneously differentiating cells are divided into two groups with different paces of differentiation. All data shown on the flow histogram depict data from the undifferentiated stage until day 14 of differentiation, with the most anterior graph representing the Oct4-eGFP signal from undifferentiated cells, and the graph most at the back illustrating day 14. **B:** Western blot images (2 µg samples, C-terminal H3 antibody) show continuous cH3 formation in hESC, regardless of the hESC culture condition.

3.4.3. Histone H3 clipping in hESC is not an *in vitro* artifact

To exclude the possibility that this continuous histone H3 cleavage in a feeder-free culturing system is actually an *in vitro* artifact induced during extraction, we spiked biotinylated calf histone H3 during hESC histone extraction at day 1 of differentiation for subsequent avidin-HRP detection on Western blot. We initially confirmed that indeed the biotinylated histone H3 still is susceptible to proteolytical degradation (Figure 3.3A).

After spiking biotinylated calf H3 during extraction, a lower band was detected on the Western blot from the histone extracts only when using the C-terminal H3 antibody. This band was absent when immunoblotting was done using avidin-HRP, which specifically highlights the biotinylated fraction (Figure 3.3B). This implies that the histone H3 cleavage was indeed already present during culturing and that the truncated form was not created during extraction. Moreover, when the extraction was performed both with and without protease inhibitors (PIC), no difference between these samples was found, confirming that no additional cleavage was induced during extraction.

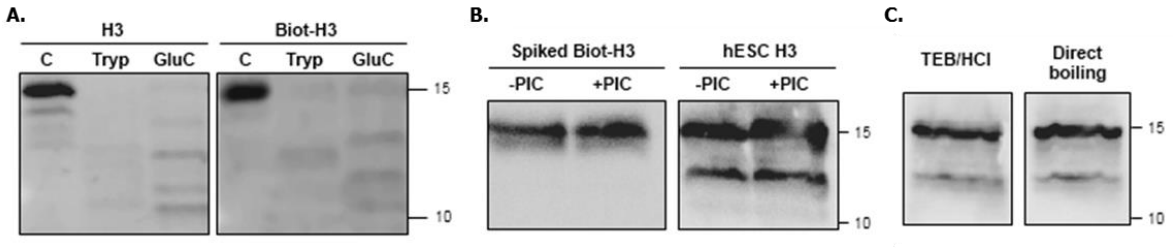


Figure 3.3. Histone H3 cleavage is not an *in vitro* artifact. A: The digestion of purified calf H3 and biotinylated calf H3 with both trypsin or GluC was compared, and Sypro Ruby staining after SDS-PAGE illustrated the analogous susceptibility of both samples. **B:** Comparison of spiked biotinylated H3 (Biot-H3) with the H3 content extracted from hESC shows that only the latter undergoes truncation. As both immunoblotting images demonstrate, including the protease inhibitor cocktail (-PIC versus +PIC) during the extraction procedure does not influence the outcome. Per lane 2 µg sample was loaded. **C:** Illustrates the comparison of two different extraction protocols: TEB/HCl extraction (left, 2 µg sample) versus direct boiling in Laemmli buffer (right, 5 µl loaded). Western blotting analysis shows no difference in cH3 content between those protocols.

In addition, the TEB/HCl extraction protocol was compared with direct boiling of the hESC after harvesting. Directly boiling in Laemmli buffer diminishes the steps of the extraction procedure and thus the steps in which cH3 can be artificially formed. For this experimental set-up, Oct4-eGFP reporter hESC were differentiated with 2 µM RA for 24h. After this the hESC were harvested and split into two to carry out both protocols simultaneously on the same starting material. As can be seen in Figure 3.3C, both protocols result in the same Western blotting image containing cH3.

3.4.4. Similar histone H3 cleavage sites are found in mouse and human ESC

To verify if this H3 truncation event has any parallel to the clipping event reported in mESC, we next set out to identify the cleavage site(s).

Applying the antibody directed against the C-terminus of H3, cH3 is visualized, in some cases as multiple bands. As opposed to the C-terminal antibody, immunoblotting with an antibody directed against the N-terminal end of histone H3 detected no cH3 for both the Oct4-eGFP reporter and UGENT2 cell line (Figure 3.4A and 3.4B). Only the band of intact histone H3 was seen, indicating N-terminal cleavage. When using the C-terminal antibody, the distance of the most intense cH3 band to

the intact H3 form, implies a loss of approximately 3 kDa, which roughly corresponds to about 30 amino acids. Using the histone H3.cs1 antibody that was developed to specifically detect H3 truncated after A21 [127], a weak signal was detected for samples of the Oct4-eGFP cell line, suggesting A21 as a possible target site also in hESC.

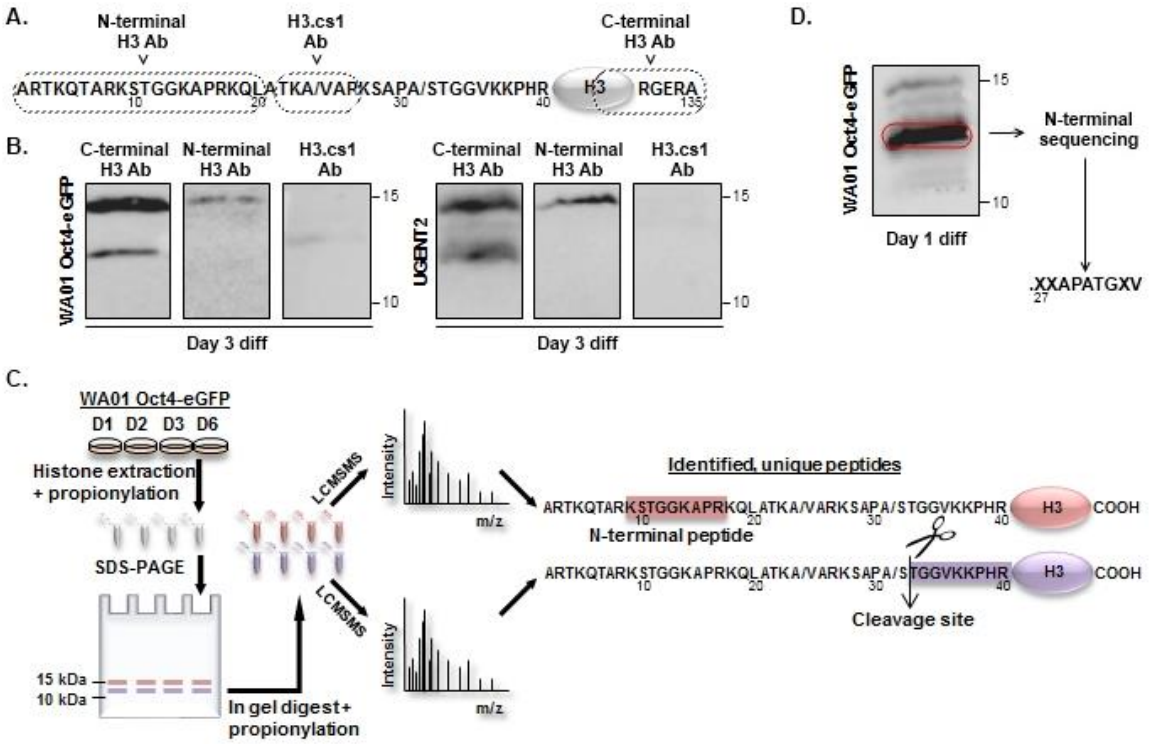


Figure 3.4. H3 cleavage in feeder-free cultured hESC under application of different culture media within an extended time frame. The application of differentially directed antibodies for immunoblotting (antibody epitopes displayed in **A**) illustrates that the cleavage is situated N-terminally, and proposes alanine 21 as one of the possible cleavage sites (**B**). **C:** Mass spectrometry analysis after propionylation of different hESC samples (Day 1, day 2, day 3 and day 6 of differentiation) also reveals residue 31 as a site of cleavage. **D:** An additional site, arginine 26, was annotated by Edman degradation.

Apart from the most prominent A21, Duncan et al. found additional histone H3 cleavage sites in mESC, so we pursued further investigation of the presence of any other possible cleavage site(s) in hESC. A specific mass spectrometry (MS) approach for histone analysis was therefore optimized which is subsequently described in more detail. A regular bottom-up approach using trypsin would cleave after each lysine (K) or arginine (R) (except when followed by a proline), resulting in unidentifiably small peptides because of the abundant presence of these basic amino acids in histones. We thus propionylated the histones prior to digestion, which modifies all free primary amine groups (the N-termini and the ε-amino group of unmodified and monomethylated K). Of note, when the N-terminus or a lysine residue is covered by an endogenous modification other than monomethylation, this residue cannot be propionylated but is also blocked from proteolytical digestion by that modification

itself. Consequently, trypsin now only cleaves C-terminal to arginine residues mimicking an Arg-C digestion for histone H3, resulting in larger, identifiable peptides [187]. By post-digestion propionylation all newly generated N-termini are subsequently being propionylated as well. The specific cleavage sites present in cultured hESC can then be detected as peptides containing a non-arginine C-terminus or which start N-terminally with an amino acid not neighboring an arginine. Spectra corresponding to such so-called semi-Arg-C peptides were additionally manually validated.

Hence we separated propionylated histone extractions of the feeder-free cultured Oct4-eGFP reporter cell line at four different time points after RA induction (D1, D2, D3 and D6) on SDS-PAGE (Figure 3.4C). This way, multiple molecular weight bands could be cut out for subsequent in-gel digestion using trypsin. After a second round of propionylation, samples were analyzed by RPLC and ESI-MSMS. Of all identified histone H3 peptides, 100 % ended C-terminally with R and 94.4 % had an N-terminus adjoining an R, confirming successful propionylation.

First, to confirm the N-terminal histone H3 cleavage by MS we monitored the presence or absence of an N-terminal histone H3 peptide in the intact H3 and the cH3 gel bands. The peptide $K_{[9]}$STGGKAPR starting after $R_{[8]}$ covers the N-terminus of histone H3 and was only identified in the H3 bands and not the cH3 (Figure 3.4C), in contrast to the globular histone H3 peptides and the C-terminal $V_{[117]}$TIMPKDIQLAR which were present in both. Two differentially modified semi-Arg-C peptides both started at residue 32 which N-terminally does not flank an R: $T_{[32]}GGVK_{prop}K_{prop}PHR$ and $T_{[32]}GGVK_{prop}K_{me\text{-}prop}PHR$ (Figure 3.5). Since these spectra correspond to the peptide generated by cleavage after amino acid 31 and this residue itself is not represented in the spectrum, it is impossible to define whether amino acid 31 is an alanine or serine, coming from the H3.1/H3.2 or the H3.3 isoform respectively. At each of the four different time points $A_{[31]}$ or $S_{[31]}$ clipping was identified in the gel pieces containing cH3, but not intact H3. Of note, using this mass spectrometry method no cleavage at an arginine residue can be detected, since the protocol itself introduces clipping after arginine during digestion. Indeed, by applying Edman degradation, we also assigned R26 as a cleavage site (Figure 3.4D). Since the amino acid at position 31 is identified by Edman degradation as an alanine this cleaved form is derived from the H3.1 and/or H3.2 isoforms.

In conclusion, both Western blotting and mass spectrometry confirmed N-terminal cleavage with an intact C-terminus, and the latter assigned amino acid 31 as a cleavage site for histone H3 without isoform specification. N-terminal sequence analysis added R26 as a cleavage site, at least for the H3.1 and/or H3.2 isoforms.

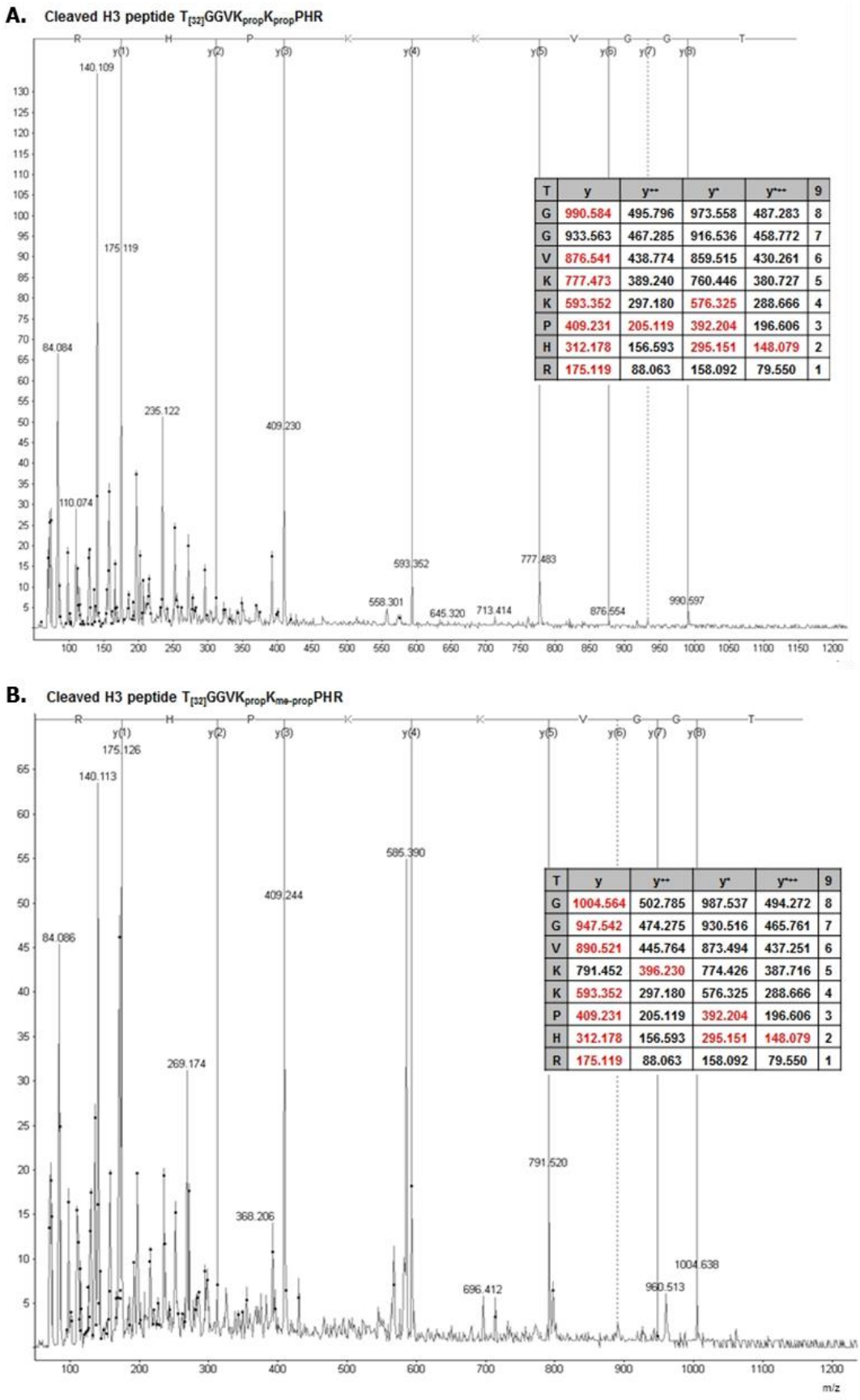


Figure 3.5. Manual validation of cleaved H3 peptides. Validation of the MSMS spectrum representing **A:** $T_{[32]}GGVK_{prop}K_{prop}PHR$ and **B:** $T_{[32]}GGVK_{prop}K_{me\text{-}prop}PHR$ respectively. Matching fragments are displayed in red.

3.4.5. When grown on MEF hESC cleave H3 in an identical temporal window compared to mESC

Since mESC differentiation showed various temporal clipping patterns depending on the differentiation protocol applied, we further extended the different culture conditions to also include the more classical culture system on a MEF feeder layer. Although these MEF are mitotically inactivated prior to cell culturing, cH3 formation derived from MEF cannot be excluded. Thus, RA was first added to a culture consisting only of a confluent layer of MEF, without hESC present. No cleavage was detected by immunoblotting (Figure 3.6).

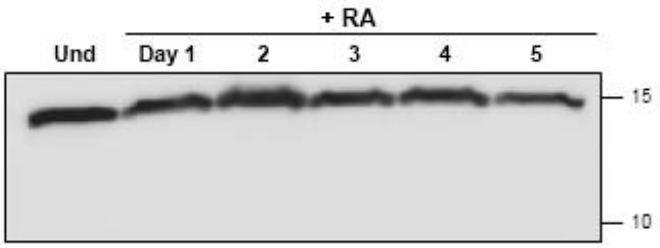


Figure 3.6. Mouse embryonic fibroblasts display no histone H3 cleavage. When incubating a culture consisting only of MEF with RA for 5 days, no cH3 could be detected.

When RA was added to hESC cultured on MEF, both morphological assessment of the differentiated status, and flow cytometry as well as RT-qPCR confirmed a decrease in Oct4 level as differentiation proceeded for the Oct4-eGFP reporter and UGENT2 cell line respectively (Figure 3.7A and 3.7B). Of note, as seen by fluorescence microscopy within one colony patches or 'islets' of undifferentiated cells became apparent, indicating heterogeneous stemness within colonies (Figure 3.7C). Remarkably, this Oct4 expression pattern is clearly different from what was seen when growing hESC feeder-free, which showed a diffuse expression pattern in the abovementioned experiments.

Surprisingly, when analyzing these histone extracts with Western blotting, a more 'mouse-like' pattern of H3 cleavage becomes visible: on day 2 of differentiation cH3 appears, attains a maximum on day 4 and fades considerably on the last day of the experiment. This pattern was visualized for both cell lines, again confirming the cleavage capability of both the reporter and non-reporter cell line (Figure 3.7D).

Also here the N-terminal nature of the H3 cleavage was double-checked using the N-terminally directed antibody, which visualized only the band of intact histone H3. In addition, a clear signal was detected using the H3.cs1 antibody for the UGENT2 sample, once more pointing out A21 as a possible cleavage site (Figure 3.7E).

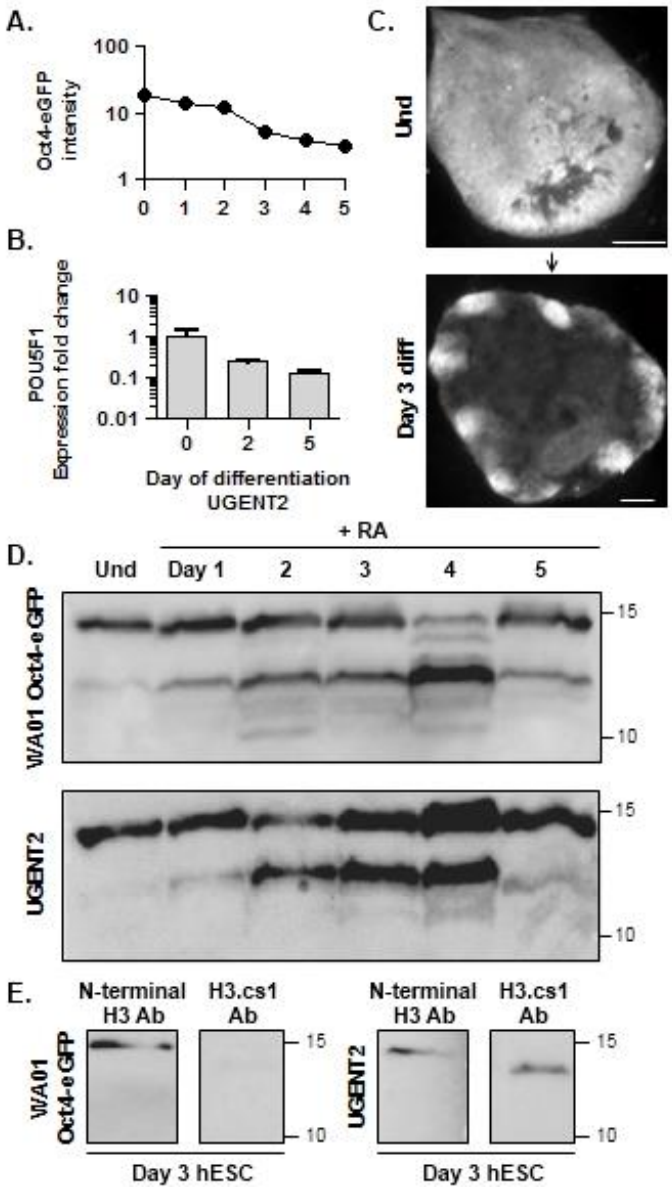


Figure 3.7. A different H3 cleavage pattern is seen in MEF cultured hESC. Aside from the application of feeder-free culture, additional differentiation experiments with hESC cultured on a MEF feeder layer were conducted. The differentiated status is again confirmed by **A:** flow cytometry and **B:** RT-qPCR (mean and standard deviation of 2 replicates). **C:** Fluorescence microscopy analysis indicates differentiation heterogeneity among colonies, as within each colony several 'islets' of concentrated Oct4-eGFP expression were seen (scale bars represent 200 µm). **D:** Western blotting analysis reveals that when hESC are cultured on MEF, a cleavage pattern is shown that resembles more to the mESC results. **E:** Also, the N-terminal orientation of the cleavage and A21 as a possible cleavage site is confirmed by application of the N-terminal and H3.cs1 antibody respectively.

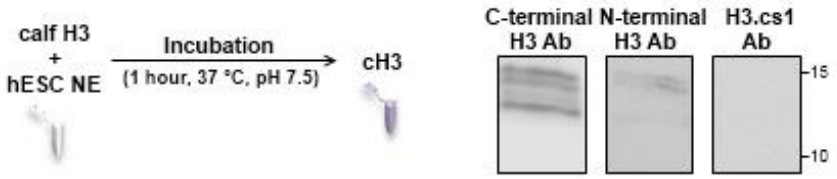
3.4.6 Histone H3 clipping activity in hESC is exerted by a serine protease

In order to identify the clipping enzyme, we conducted different calf H3 incubation assays, whether or not with the application of various protease inhibitors. To this end, we prepared a nuclear extract (NE) from Oct4-eGFP hESC for which histone H3 cleavage was confirmed in the corresponding histone

extract, and validated it to be capable of calf H3 proteolysis. Endogenous H3 was not detectable in the amount of NE used in these experiments.

In a first experiment we compared the incubation of H3 with hESC NE after immunoblotting with the three different antibodies described before. This confirmed that the cleavage pattern induced by the NE is similar to what was seen in the histone extract of Oct4-eGFP hESC: the cleavage is exclusively of N-terminal nature and there is no detection of A21 as a possible cleavage site (Figure 3.8A).

A.



B.

Inhibitor	Target	Concentrations applied		Inhibition
		Min (*)	Max (**)	
AEBSF	Serine proteases	0.1 mM	1 mM	+
E64	Cysteine proteases	1 µM	10 µM	-
PIC	Serine and cysteine proteases	1 tablet/10 ml buffer	1 tablet/2 ml buffer	+
EDTA	Metalloproteases	1 µM	10 µM	-
Bestatin	Aminopeptidases	1 µM	10 µM	-
Pepstatin A	Aspartyl proteases	5 µM	10 µM	-
Cath L inh I	Cathepsin L	1 µM	10 µM	-
Cath L inh III	Cathepsin L	10 µM	30 µM	-
Cath L inh CAA0225	Cathepsin L	5 µM	20 µM	-

C.

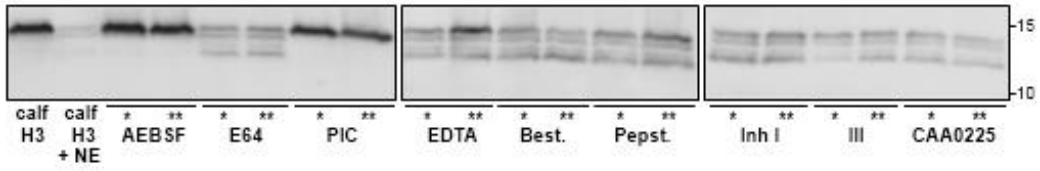


Figure 3.8. The histone H3 clipping activity in hESC is exerted by a serine protease. A: In order to investigate the responsible cleavage enzyme, incubation assays were set up in which 1 µg calf histone H3 was incubated with 0.5 µg hESC NE. The nuclear extract was prepared from the Oct4-eGFP cell line and contains N-terminal histone H3 cleavage activity, as was confirmed in the corresponding histone extracts (data not shown). Immunoblotting analysis of the incubation product displays several clipped H3 (cH3) fragments, with confirmed N-terminal origin. A21 could not be confirmed as a cleavage site (H3.cs1 Ab). The list of all inhibitors included, with their targets, the applied concentrations and their outcome in the inhibition assay is displayed in panel **B**. **C:** Western blotting analysis points out that only when using a serine protease inhibitor the cleavage activity of the nuclear extract can be inhibited. The latter is confirmed by the additional inhibitory effect of the protease inhibitor cocktail, targeting mainly serine and cysteine proteases.

In the following inhibition assay, several inhibitors were included (Figure 3.8B), comprising both general inhibitors for a certain protease enzyme class and specific inhibitors for cathepsin L, in parallel to the mESC results, where cathepsin L was identified as the responsible clipping enzyme. Western blotting analysis pointed out that only AEBSF, a serine protease inhibitor, could establish inhibition of the cleavage, whereas all other inhibitors could not demonstrate any notable effect on the incubation.

Also the PIC was able to inhibit the proteolytic activity on histone H3. The specific inhibitors for cathepsin L were not able to diminish the cleavage (Figure 3.8C).

3.5. Discussion

Histone clipping was reported for the first time even before these proteins received their current nomenclature [188, 189]. Surprisingly however, these truncation events are still greatly understudied. Yet, from a practical point of view, techniques such as chromatin-immunoprecipitation run the risk of not detecting the substantial influence that histone clipping might have on their outcome. All techniques employing specific antibodies to the H3 N-terminus in general or to modifications thereof in particular, will fail in case of H3 proteolytic cleavage. Also mass spectrometry-based (quantitative) comparisons of e.g. H3K4 and H3K27 methylations devaluate when dealing with samples where clipped histones are present.

Histone clipping has been reported in several distinct biological systems [175–178, 180–183], entailing different biological settings and thus potential roles of this PTM. However, only recently it was admitted into the epigenetic landscape with the discovery of transient histone H3 clipping in differentiating mESC [127] and sporulating yeast [179]. The evolutionarily conserved sequence of histone proteins has been suggested to underlie the surprising occurrence of A21 clipping in both mouse and yeast, and we thus set out to verify whether the H3 clipping that accompanies mESC differentiation also takes place during hESC differentiation.

Duncan and colleagues monitored cH3 levels under several differentiation protocols in mESC, namely spontaneous monolayer differentiation after withdrawal of LIF, induced differentiation with RA and EB formation. Depending on the protocol applied, a different cH3 pattern was visualized. Where RA induction leads to an undulating pattern of upcoming and decreasing H3 clipping centered around day 2 and 3 of differentiation, EB formation displayed a faster migrating H3 band which peaked between day 8 and 12 but did not disappear completely after 14 days.

Here, we report that indeed histone H3 N-terminal clipping occurs in differentiating hESC and that its temporal appearance is equally influenced by the culture conditions. In feeder-free conditions, the two hESC lines tested (UGENT2 and WA01 Oct4-eGFP) show a continuous cH3 pattern after RA induced differentiation, resembling the results of EB formation in mESC. In contrast, when switching to culturing hESC on a feeder layer of MEF, both these cell lines obtain a pattern of upcoming cleavage appearing on day 1, reaching a maximum intensity at day 4 and fading again at the last day of the experiment, similar to the temporal pattern described for RA induced mESC differentiation. Together, this suggests that within both human and mouse, the specific control over histone H3 cleavage during

differentiation is profoundly influenced by experimental culture conditions applied. Although the cleavage event in mESC and hESC is appreciably similar, caution should be taken when functionally comparing these events, not in the least because hESC resemble more mouse epiblast stem cells than regular mESC. The latter are considered to represent a more homogeneous and naïve pluripotent state compared to hESC, which are designated to be in a heterogeneous and primed pluripotent condition and thus probably reflect a more developed state [5, 46, 52]. In contrast to Duncan et al., we could not find any clear correlation between the progression of differentiation and histone H3 clipping. The pluripotency status was monitored in this report by following Oct4 levels, i.e. specifically with an Oct4-eGFP reporter cell line. When monitoring three different differentiation methodologies, H3 cleavage seemed to be unaffected, while the Oct4 status was found to be influenced in the predicted way: a decrease in expression following (induced) differentiation. Considering the open and highly dynamic state of pluripotent hESC chromatin, this might not be such a surprise [72]. In line with this, on a Sypro Ruby stained SDS-PAGE gel we also observed some reduction in H4 band intensity at some time points (data not shown). Whether this indeed is a clipping event remains to be determined. Of note, when Duncan et al. inhibited cathepsin L, H3 clipping was abrogated while Oct4 expression still decreased, adding yet another argument against a direct link between Oct4 expression levels and histone clipping.

Histone H3 cleavage is not *in vitro* generated during the experiment as visualized here by the use of biotinylated histone H3 during the extraction procedure. Also, protease inhibitors did not seem to influence the cH3 intensity, whether or not they are present. Moreover, both TEB/HCl extraction and direct boiling in SDS-PAGE Laemmli yields the same results, further supporting the endogenous formation of cH3. The fact that MEF themselves do not display any clipping, but can completely change the temporal appearance of cH3 in differentiating hESC, further argues in favor of a biologically regulated process. Of note, we also verified the occurrence of H3 clipping in a human, terminally differentiated cell line (Burkitt's lymphoma Raji cells) and these cells did not form any clipped histone H3 (data not shown).

Many different histone H3 cleavage sites have been reported in the past, even within one study. In differentiating mESC [127], A21 is the primary site of cleavage, though multiple other sites were also found at T22, K23, A24, R26 and K27. By the use of Western blotting, Edman degradation and mass spectrometry respectively, we confirmed two of those cleavage sites in hESC, namely A21 and R26, and assigned residue 31 as an additional new cleavage site in hESC. Unfortunately, with the techniques applied, no definite distinction could be made between the H3.1/H3.2 or H3.3 isoform cleavage

respectively. Nevertheless, we expect H3.1 or H3.2 to be the cleaved isoform, since A31 is followed by threonine just as A21, which can thus be suspected to be susceptible to a similar enzyme activity.

It is worth noting that despite the many hurdles that need to be overcome when using mass spectrometry for the analysis of histone H3, several arguments add up to the likelihood of this newly identified cleavage site at residue 31. First, the annotated peptide was N-terminally propionylated, indicating no in-source decay causing this fragment to appear. Second, elution time patterns of the precursor and its cleaved form are distinct, further arguing against in-source formation of the cleaved fragment out of the intact precursor. Third, by first separating the histones with SDS-PAGE, the location of annotated histone H3 and its cleaved form in the gel allows to project the peptide data back onto the precursor proteins: the N-terminally cleaved fragment could only be annotated in lower MW (cH3) gel bands, in contrast to the N-terminal peptide itself which was found only in the highest MW (intact H3) fraction. Finally, all spectra corresponding to a cleaved peptide were manually validated by an expert before taken into account. On the other hand, more cleavage sites are expected to be present which cannot be identified by MS, as seen for the A21 and R26 sites detected by a specific antibody and Edman degradation respectively. This can be explained by the limited possibilities of MS to annotate small peptides with high reliability. When a cleavage site is found to be close to an arginine, which is cleaved during digestion, the resulting peptides will consist of too little amino acids to be identified reliably. As a result not all possible cleavage sites can theoretically be identified by MS but their existence should not be disregarded.

Only few of the published histone clipping reports actually also categorize the protease responsible for this event (reviewed in [190]). Duncan and colleagues (2008) assigned the cleavage to cathepsin L, a lysosomal cysteine protease, whilst glutamate dehydrogenase has been brought forward by Mandal et al. (2013). Santos-Rosa et al. (2009) who originally published the yeast clipping event, could only categorize the enzyme as a serine protease. It was just recently that Xue and colleagues [191] found that the vacuolar protein Prb1 (Cerevisin) is required for the N-terminal H3 clipping in *Saccharomyces cerevisiae*. Finally, the Foot-and-Mouth disease virus expresses the so-called protease 3C, a cysteine protease, in the host cells which mediates clipping of host histone H3 at leucine 20 [175, 176].

The incubation assays performed here with the application of different inhibitors point out that the clipping enzyme is a serine protease, since apart from AEBSF none of the inhibitors for other protease classes could establish any effect on the incubation. This was confirmed by the fact that the PIC, which inhibits mainly serine and cysteine proteases, was also able to inactivate the clipping activity. We also included three specific inhibitors for cathepsin L, in analogy to the clipping in mESC, but no decrease in cleavage was seen due to these inhibitors. However, caution should be taken in the search for the

responsible enzyme. As also others already suggested [127, 179], enzyme redundancy and overlapping functions could impede its identification.

Despite the great epigenetic promise of such radical PTM, histone clipping and more specifically its biological potential and the mechanisms by which it could exert its transcriptional effects, remain surprisingly understudied. Since the two landmark discoveries of H3 clipping in mouse and yeast in 2008 and 2009 respectively, four major mechanisms have been formulated by which histone proteolysis can influence gene expression programs [173].

First, in yeast, a direct regulatory role of gene expression has been attributed to the removal of the N-tail and its repressive marks at promoter regions [179]. As such, histone H3 cleavage clears repressive marks massively, hence allowing for gene expression activation. Second, cH3 might provide a new binding site for protein complexes that could not be bound before clipping, thus fulfilling an active role in protein recruitment. On the other hand, also a passive regulatory role should be considered if other proteins are no longer able to bind the shortened histone H3 [173, 179]. Third, Santos-Rosa and colleagues [179] also proposed nucleosome eviction and histone replacement as another framework in which clipping might regulate gene expression, as they report that H3 clipping precedes nucleosome eviction and subsequent gene induction. Finally, the N-terminal peptide itself might establish its translation regulation by binding its own mRNA [173]. Which of these mechanisms, if any, is at play in (differentiating) hESC remains to be elucidated.

In short, based on previous findings and theoretical background, H3 clipping could be involved in several processes linked to gene expression control and differentiation. But the question still remains as to whether H3 proteolysis correlates directly with or causes such processes. Further studies will hopefully help elucidate the role of this new epigenetic mark in hESC and their differentiation process.

3.6. Conclusions

As a member of the epigenetic network, regulated histone proteolysis has been occasionally described earlier in diverse biological settings, yet being largely understudied. Histone H3 clipping and histone proteolysis in general might skew experimental findings substantially, both from a biological and a technical point of view. In this report we show for the first time that this posttranslational modification is also present in human ESC, and is mediated by a serine protease. The temporal pattern of cleaved H3 is highly dependent on the culture protocol applied, as seen for both cell lines used in this report. Although in first instance we too detected the clipping upon early stem cell differentiation, we found that the clipping process is not necessarily accompanied by a decrease in Oct4 expression, as also

undifferentiated hESC can contain cleaved H3 fragments. Thus, more research is needed to fully elucidate the potential biological role(s) of histone H3 cleavage.

Chapter 4:

Supplementary findings on H3 clipping in hESC

4. SUPPLEMENTARY FINDINGS

4.1. Abstract

H3 clipping is present as a histone PTM in hESC, but a clear framework of its biological function is still lacking. Preliminary data obtained from MEF cultured hESC initially led to the hypothesis that this event is cell cycle related. However, the clipping data collected from other, synchronized cell lines (THP-I and Raji cells) in combination with data retrieved from feeder-free hESC experiments contradict this hypothesis. Initially also a link with ongoing differentiation was proposed, but as discussed in Chapter 3, this seems implausible. After conducting two more control experiments to exclude *in vitro* H3 clipping, we also ruled out two of the best studied clipping proteases as potential mediators of the clipping event: cathepsin L/V nor neutrophil elastase are probable candidates for the responsible clipping enzyme.

4.2. Background

Several epigenetic mechanisms cooperate in regulating stem cell biology, but the elucidation of this complex network still has a long way to go. Histone clipping illustrates perfectly how much more research is needed to complete current insights. As clipping is still an under-studied PTM despite its epigenetic potential, it is hard to attribute the correct biological context to new experimental findings. On top of the results on hESC H3 clipping described in the previous chapter, some additional unpublished data are outlined here. They helped in phrasing the discussion in the manuscript in Chapter 3, but were not included as results because of their unclear biology. Nevertheless, they can contribute to build the correct framework for histone clipping in hESC.

4.3. Methods

Cell cultures and differentiation

Human ESC culture and their differentiation were performed as described in Methods 3.3. The following cell lines were applied: UGENT1 (G1 (XY)), UGENT2 (G2 (XX)) and a WA01 Oct4-eGFP knock-in cell line.

THP-I cells were maintained in suspension culture in completed RPMI consisting of RPMI medium supplemented with 2 mM L-glutamine, 100 U/ml penicillin, 100 µg/ml streptomycin and 10 % fetal bovine serum (37 °C, 5 % CO_2). Medium was replenished or refreshed totally every other day. Differentiation was induced by addition of 1 µM RA.

Raji cells (suspension) were cultured in completed DMEM consisting of DMEM medium supplemented with 2 mM L-glutamine, 100 U/ml penicillin, 100 µg/ml streptomycin and 10 % fetal bovine serum (37 °C, 5 % CO_2). Medium was replenished or refreshed totally every other day.

Cell cycle synchronization

Raji and THP-I cells were synchronized by means of a double thymidine block. A first blocking was reached by exposure to an excessive amount of thymidine (2 mM) overnight. After this incubation period the cells were released again for 7 to 8 hours by washing off the thymidine. Subsequently a second overnight thymidine block was applied, after which the cells were synchronized at the early S phase.

PI staining and flow cytometry analysis

The results of the cell cycle synchronization can be monitored with flow cytometry. To this end, the cells were stained with PI (ex. 488 nm, em. 617 nm). Prior to staining, the cells were fixated in ethanol at 4 °C during at least 2 hours. After washing the cells with PBS, staining was performed in the dark

during 30 min (50 µg/ml PI and 10 µg/ml RNase in PBS). The flow cytometric analysis afterwards was carried out with an FC500 (Beckman Coulter) using the CXP analysis software.

Protein sample preparation

Histone extraction from the different cell types and the nuclear extraction from Raji cells were performed as described in Methods 3.3.

The protocol applied for isolating the nuclear fraction of THP-I cells used in the first step a hypotonic lysis buffer (10 mM Tris (pH 8), 1 mM KCl, 1.5 mM $MgCl_2$, 1 mM dithiothreitol, PIC), in which the cells were incubated for 30 min at 4 °C. After centrifugation (10 min, 4 °C, 2000 rpm), the isolated nuclei were resuspended in R1-buffer (40 mM Tris, supplemented with tributylphosphine, phosphatase inhibitors, PIC and endonuclease) and sonicated for 10 min.

SDS-PAGE and western blotting

Gel electrophoresis, Sypro Ruby staining and blotting analysis were performed as described in Methods 3.3. The applied primary antibodies for the additional experiments were: anti-cathepsin L catalytic domain antibody (1:1000; ab49984), anti-cathepsin L mature domain antibody (1:500; ab58991), anti-cathepsin V antibody (1:1000; ab49982), anti-neutrophil elastase antibody (1:500; ab21595), all purchased from Abcam.

RT-qPCR

RNA isolation, cDNA preparation and RT-qPCR analysis were performed as described in Methods 3.3. The additional TaqMan assays used were: cathepsin L (Hs00377632_m1), cathepsin V (H100426731_m1), neutrophil elastase (Hs00975992_g1 and Hs00975994_g1), all purchased from Life Technologies.

In vitro enzymatic incubation assays

Recombinant active cathepsin L was purchased from Abcam (ab81780). Recombinant pro-cathepsin V was obtained from Enzo Life Sciences (BML-SE554-0010) and could be activated by pre-incubating the proenzyme in assay buffer for 5 to 30 min. The pH 5.5 buffer consisted of 10 mM HEPES, 10 mM KCl, 1.5 mM $MgCl_2$, 0.34 M sucrose, 10 % glycerol and 5 mM β-mercaptoethanol. The pH 7.5 buffer was composed of 25 mM Tris-HCl, 150 mM NaCl, 0.1 mM EDTA, 10 % glycerol and 2 mM β-mercaptoethanol. Incubations were performed at 37 °C for 1 hour.

4.4. Results

4.4.1. Histone H3 clipping in differentiating MEF-cultured hESC follows an undulating pattern

The initial differentiation experiments carried out on MEF-cultured hESC were also repeated both in a more detailed time frame and over a more extended time period. To acquire a sufficient number of cells in the least possible number of passages, two hESC lines (G1 and G2) were combined in these experiments. Although this may have yielded a different clipping pattern from what might have been obtained when using a single cell line, both G1 and G2 display a pattern of upcoming and decreasing H3 cleavage (Figure 4.1).

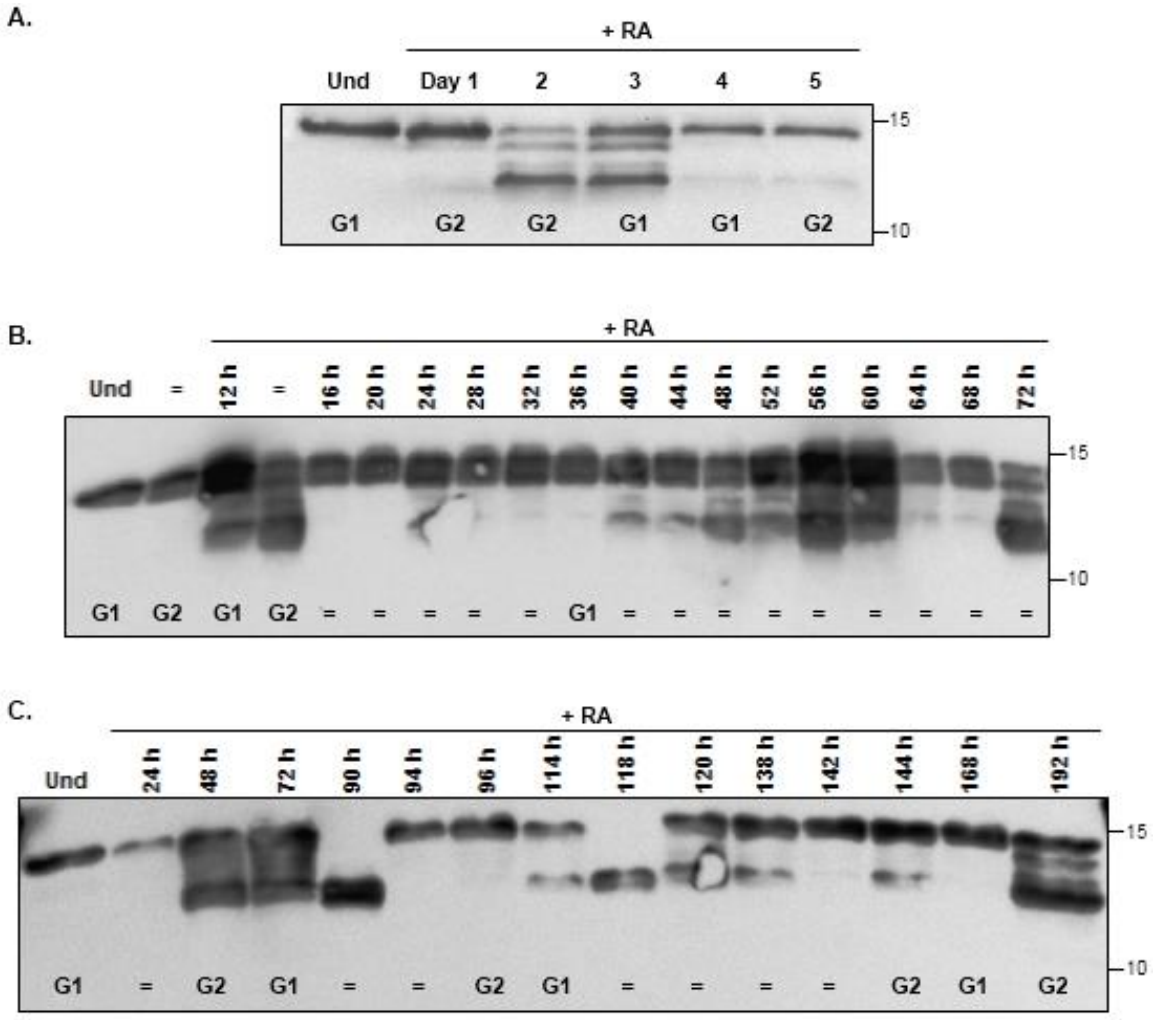


Figure 4.1. Undulating pattern of H3 clipping in MEF cultured hESC. A: The initial differentiation experiment showed cleavage for both the G1 and the G2 cell line, with upcoming cH3 on day 2 and fading again on day 4 of differentiation. **B:** When increasing the time resolution on day 3, 4 and 5 of differentiation, western blotting analysis shows a 'wave' pattern of upcoming and decreasing clipping. **C:** The same result is obtained on a larger time scale when the experiment was extended up to 12 days. Also for **B** and **C** samples from both cell lines were used, as indicated on the blot image. Respectively 3 µg (**B**) and 2 µg (**C**) histone extract were loaded per lane. The time points of sample collection are indicated on all images ("x hours/days" after onset of differentiation).

This 'wave' pattern is in stark contrast with the continuous cleavage found in feeder-free cultured hESC and emphasizes again the crucial role that is played by the culture conditions applied. However, we interpreted this undulating pattern to be possibly cell cycle-related, as the addition of RA to the cell culture most probably also has a somewhat synchronizing effect on the stem cell cycle, as they are primed for differentiation. One possibility is then that the clipping occurs during the S phase, in which the DNA is duplicated and the histone content is replenished accordingly. As the DNA structure is loosened upon replication, it becomes more susceptible for clipping proteases. This also implies it is more accessible to aspecific and/or cytoplasmic enzymatic activity. *In vitro* enzyme activity does thus become an important experimental concern, although this does not necessarily exclude any biological functionality for histone clipping.

4.4.2. Artificial *in vitro* clipping

If the clipping is indeed merely induced during cell lysis, this too could generate an undulating pattern, assuming that the clipping primarily occurs when DNA structure is less condense. Apart from the experiments reported in Results 3.4.3, we also performed a final experiment that circumvented the need for biotinylated H3, although we validated that biotinylated histone H3 displayed sustained enzymatic susceptibility (Figure 3.3A). In this experiment, we added an excessive amount of calf histone H3 to the buffers during the extraction procedure. This calf H3 did not carry a biotin- or any other label and was thus *in se* not distinguishable from the endogenous hESC H3 content by means of western blotting analysis. The comparison of the H3 intensity from hESC samples with and without extra H3 added showed a diminished cH3 intensity in the former samples (Figure 4.2A). Although we cannot exclude that this was due to an inhibitory effect of the excess calf H3 towards the clipping of the endogenous H3, this result at least suggests that the added calf H3 was probably not clipped to a notable amount during extraction and did not contribute to the cH3 content in the hESC samples.

Besides the experiments to exclude possible unwanted *in vitro* clipping activity during the histone extraction protocol, we also briefly looked into whether or not the other core histones are processed during our hESC experiments. In the mass spectrometry data obtained in search for the clipping sites (Results 3.4.4), we could not find any evidence for other clipped histone forms. Additionally, we analyzed several samples from differentiating hESC by means of Sypro Ruby staining after SDS-PAGE to visualize possible additional protein bands indicative of other degradation products. Again, no obvious changes in histone patterns were found, except for that of histone H3 (Figure 4.2B). Although this is evidently not stringent enough to affirm their absence, it does add to the evidence for the specificity of this clipping event.

Taking these data together, we conclude that the detected H3 fragments must have been generated prior to the extraction procedure in our experiments. Nevertheless, undisputable confirmation hereof could only be attained by detecting the specific fragment in intact cells, e.g. by live-cell immunofluorescence staining with the H3.cs1 antibody. Yet, we consider the evidence outlined here adequate to imply a biological origin of the cH3 fragments in differentiating hESC.

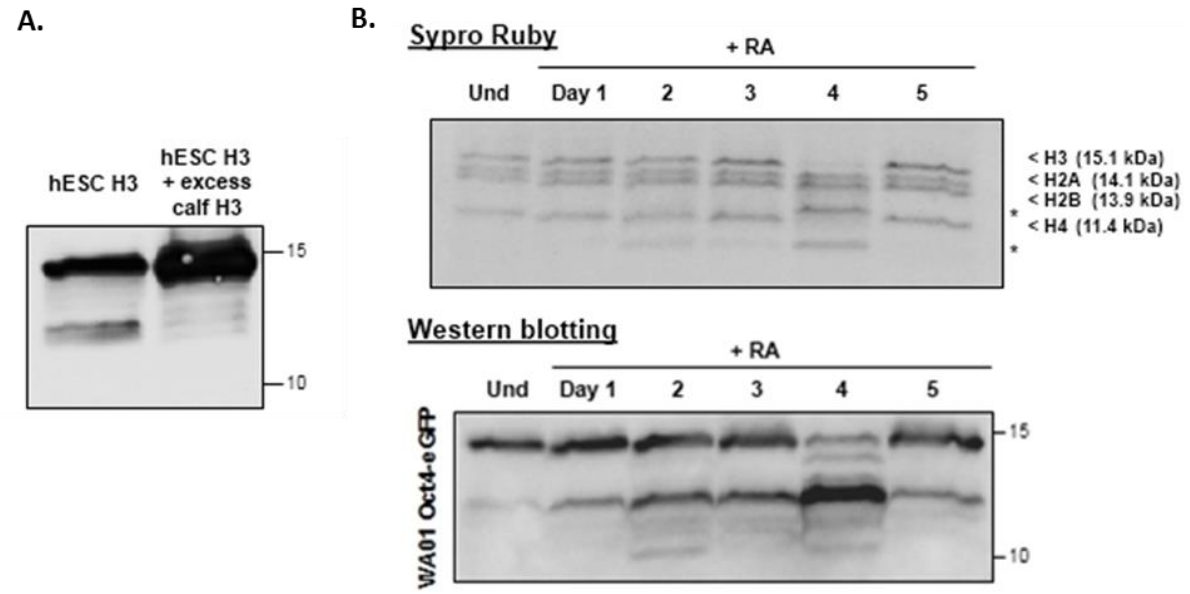


Figure 4.2. No *in vitro* clipping in hESC. A: Western blotting analysis showed that the cH3 intensity decreased in samples where an excessive amount of calf H3 was added during the extraction procedure, at least suggesting that the cH3 content originated from before extraction and was already present in the hESC. For each lane 2 µg histone extract was loaded of samples obtained from the Oct4-eGFP reporter cell line (feeder-free culture) after 24 hours of RA exposure. **B:** Sypro Ruby staining after SDS-PAGE (upper panel) shows no notable changes in any other core histone content other than histone H3. Specific western blotting analysis (lower panel) with the C-terminal H3 antibody clearly showed the cH3 content. Samples were obtained from a differentiation (2 µM RA) experiment performed on Oct4-eGFP reporter cells (MEF culture) during 5 days. The input for each lane was 3 µg histone extract.

4.4.3. Is clipping linked to the cell cycle? Is differentiation involved in its regulation?

If not generated *in vitro*, this undulating clipping pattern still suggests cell cycle involvement. To investigate this hypothesis, a cell cycle synchronization experiment was performed with Raji cells (Burkitt's lymphoma). Synchronization was most optimally induced by subjecting the cells to a double thymidine block (stop in early S phase) and could be visualized by flow cytometric analysis after PI staining (as shown in Figure 4.3A). As already mentioned in the discussion of 3.5, no cH3 was seen in Raji cells, disregarding whether they are synchronized or not (Figure 4.3B). Next, also the progenitor THP-I cell line (acute monocytic leukemia) was included, which displays H2A clipping activity when stimulated with RA or phorbol 12-myristate 13-acetate (PMA) [192–194]. Samples were isolated on

several time points after onset of differentiation (1 µM RA). Although the obtained THP-I results were not always consistent, they do show that these cells are capable of H3 clipping (Figure 4.3B). However, cell cycle synchronization did not alter the constant pattern of clipping found in these cells.

Since these cell lines do not have an equal differentiation capacity, the involvement of differentiation in clipping could be monitored simultaneously. Raji cells are terminally differentiated, while the THP-I's are pro-monocyte cells. Yet, differentiating the latter with RA did not induce any changes in the abundance of the H3 clipping (Figure 4.3B); this in support of our hESC results in Section 3.4, which indicate that differentiation as such (monitored via Oct4 expression) does not seem to be a major biological process at work in histone clipping.

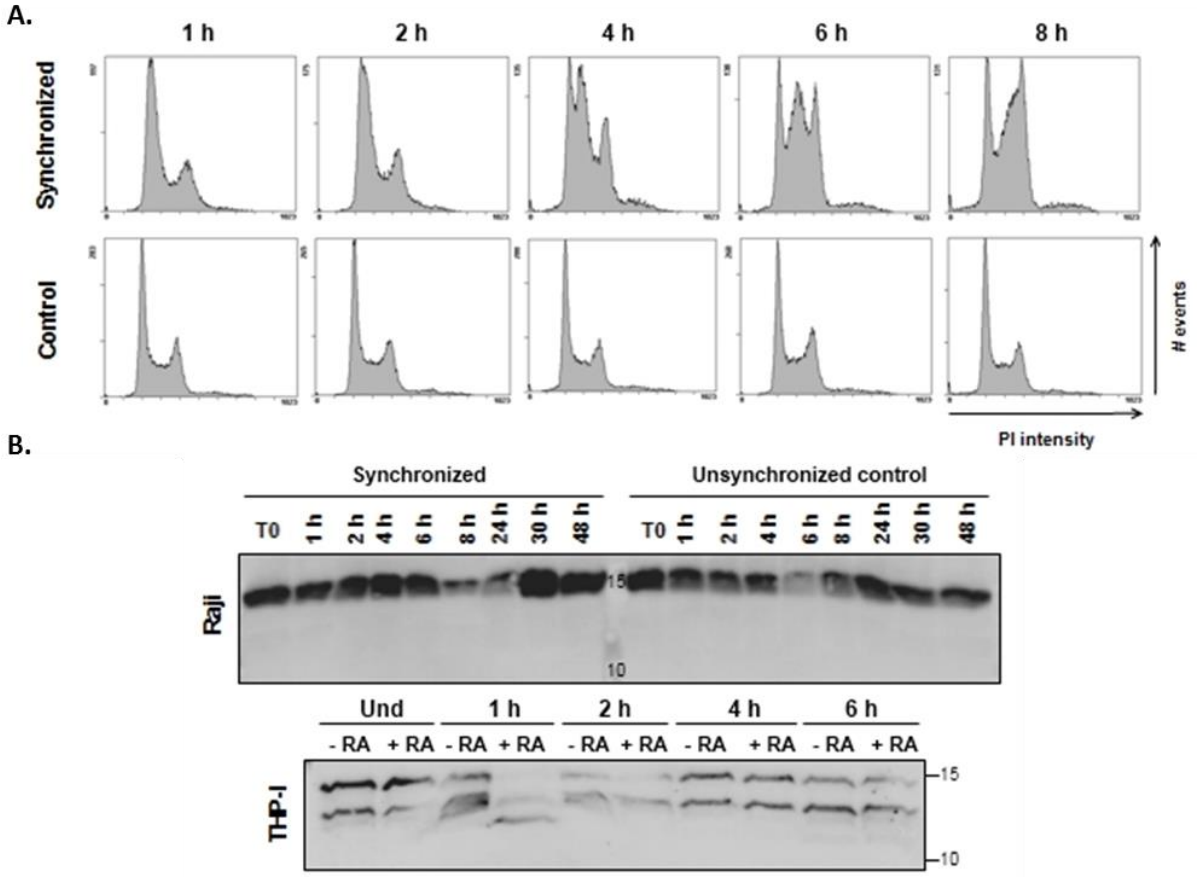


Figure 4.3. Histone H3 clipping in other cell lines. A: Validation of cell cycle synchronization. Flow plots after PI staining show a clear shift in population distribution after release from the thymidine block, compared to the steady profile of the control samples. **B:** No cH3 was observed in Raji cells in western blotting analysis, neither in unsynchronized or synchronized conditions (T0 = starting point right after release of second thymidine block, other time points indicate number of hours after release). THP-I cells on the other hand are capable of H3 clipping, but the addition of RA does not influence the clipping pattern (Time points given are the number of hours after onset of differentiation).

In summary, these findings suggest that the clipping is not directly correlated to either cell cycle or differentiation but nevertheless seem to imply that some degree of differentiation capacity is required

for it to occur. As ESC in general have a more loosened chromatin structure compared to terminally differentiated cells, it is plausible that their DNA is more susceptible for clipping protease activity.

4.4.4. The clipping enzyme

Besides the inhibition experiments described in Section 3.4.6, we also applied a targeted approach towards the cysteine protease cathepsin L, the cleavage enzyme in mESC and thus an obvious candidate for H3 clipping in hESC. Importantly, there are two cathepsin L isoforms present in human cells, namely cathepsin L (or L1) and cathepsin V (or L2). We first verified whether cathepsin L and V were capable of generating H3 fragments *in vitro*, by comparing their proteolytic activity on calf H3 in both pH 5.5 and pH 7.5. They both appeared to be capable of doing so, as visualized by western blotting. In contrast to the enzymatic activity of the NE from hESC however, they were more active in acid environment than at neutral pH (Figure 4.4A).

Next, we checked the expression of cathepsin L and V in hESC. RT-qPCR analysis showed that there was indeed expression of both cathepsins but no obvious quantitative correlation was found with the H3 clipping pattern (Figure 4.4B). Western blotting analysis only yielded inconclusive data about whether the active form of cathepsin L or V was present in the hESC nuclear extract (data not shown). Similarly, we verified the expression of both proteases in THP-I cells. Also these cells showed expression, albeit very low according to the RT-qPCR results (data not shown).

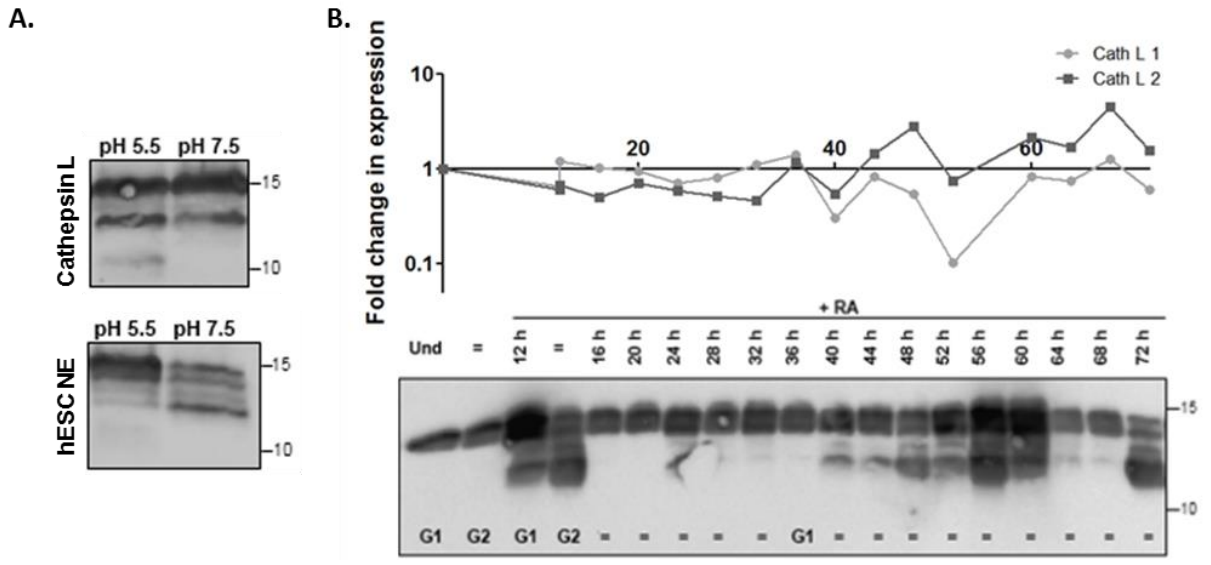


Figure 4.4. The clipping enzyme in hESC. A: The enzymatic activity of the hESC NE was compared with that of cathepsin L in two different reaction buffers. The latter enzyme was more active in acid environment, while the NE showed higher clipping activity at neutral pH. **B:** The expression of cathepsin L does not show any quantitative correlation with the cH3 pattern.

Nevertheless, the experiment outlined in Results 3.4.6 pointed out that the enzymatic activity in hESC is mediated by a serine protease. Though the option is still open that the aforementioned cathepsins might mediate histone clipping in hESC, the inhibition assay contradicts that the main responsibility lies with these cysteine proteases. Of note, the cH3 found in THP-I seemed to have a slightly higher MW than the fragment(s) found in hESC. The involvement of another enzyme in these cells should thus be taken into consideration for future experiments.

As such, a serine protease seems responsible, at least partially, for the clipping activity in hESC. In parallel with the study described here, our lab discovered that the serine protease neutrophil elastase was responsible for the H2A clipping in THP-I cells [195]. Therefore we also checked this enzyme's expression in hESC. However, no expression was seen with RT-qPCR analysis and simultaneous western blotting analysis also suggested that this enzyme is not present in hESC, or at least not in a considerable amount (data not shown). Taken together these results seem to suggest that instead of cathepsin L/V a serine protease, but not neutrophil elastase, executes the H3 clipping in hESC.

4.5. Conclusions

The lack of a biological framework hampers the research on histone cleavage and many open questions still remain. Although the clipping activity was thus far only found in cell lines that are not terminally differentiated, any data functionally linking clipping to the process of differentiation is missing. Chapter 5 provides a general discussion wherein the current literature on histone proteolytic events is classified for the first time, to give, in combination with our own results, a first impulse to build a biological framework for histone clipping.

CHAPTER 5:

DISCUSSION

Partially based on "**Histone proteolysis: A proposal for categorization into 'Clipping' and 'Degradation'**"

Maarten Dhaenens, Pieter Glibert, Paulien Meert, Liesbeth Vossaert and Dieter Deforce

BioEssays (2014), Epub ahead of print

5. DISCUSSION

Human ESC still serve as the gold standard in studying early development. Epigenetic mechanisms play an important role in hESC biology. Epigenetics is mediated via histone PTMs, histone variants, DNA methylation and RNA-based mechanisms, which all cooperate in a far from completely defined, complex network to regulate essential processes such as e.g. maintenance of pluripotency and controlling differentiation in hESC. Histone proteolysis or the enzymatic removal of several amino residues of a histone's tail, is assumed to be a rather new, emerging candidate to participate in that network. It is surprising that despite the numerous reports that propose such a potential epigenetic function, only few actually tried to attribute a biological significance and that this process does not have a broader scientific basis.

Our own findings on H3 and H2A clipping [44, 195] could as such not easily be submitted back to compare with earlier findings and hypotheses. Hence, in order to provide a general framework to build on further research and interpret new findings, we propose a first categorization of histone proteolysis into "histone degradation" and "histone clipping" [196] (further divided into several (biologically relevant) functional classes). Of note, for convenience we mainly focus here on reports on mammalian cells, though histone proteolysis also has been described in other eukaryotes such as amphibians and *Tetrahymena*. After summarizing this suggested categorization, we here frame our own results in this new classification.

5.1. Classifying histone proteolysis

5.1.1. Histone degradation

The earliest reports on histone proteolysis already date from over 50 years ago. Its theoretical implications in a living cell suggest themselves when considering the histones' nuclear functions, but conclusive evidence of transcriptional consequences is only now gradually emerging. Reports on specific enzymatic interactions with an epigenetic connotation are all classified under "histone clipping". However, these histones might have been degraded instead of being truncated inside the nucleosome and these events are subsequently classified as "histone degradation". These two categories are often hard to distinguish both experimentally and biologically, e.g. because of the fact that some proteolytic enzymes proved to be able to mediate both. And although the focus for this thesis is on histone clipping, also histone degradation provided us with some interesting insights nevertheless and is discussed here first to illustrate the complexity of histone proteolysis.

Biologically unclassified degradation

Although degradation might seem rather random at first sight, it is an important biological process. Early reports from the 1970's already described a "neutral protease" activity, mainly degrading histone H1 and H3. This enzymatic activity was attributed to a serine protease with trypsin-like activity and was primarily found in rat liver and calf thymus, but a biological function could not convincingly be attributed (e.g. [197, 198]).

Histone degradation has also been induced *in vitro* in enzymatic incubation assays conducted for structural purposes. These showed that histone proteins retain their globular structure even after the tails are removed [199], but also that these histone tails are essential for correct formation of higher-order chromatin structures [96, 97]. Upon enzymatic incubation, histone H1 appears to be the most susceptible substrate, followed by histone H3 and subsequently H2A > H4 > H2B [200]. This is not surprising as the linker histone H1 is not part of the histone octamer and is more accessible, and histone H3 has the largest protruding N-tail out of a nucleosome. Similarly, incubation experiments with crude protein extracts from different tissues can deliver plausible candidates for those proteolytic events in which the responsible enzyme has not been elucidated yet and can indicate which tissues are thus most prone to histone degradation.

Biologically classified degradation

Histone degradation in a biological context can be further divided in continuous degradation and degradation coupled to specific biological processes. At least two general mechanisms are responsible for the continuous degradation of histones: lysosomal and proteasome processing. Lysosomes are small membrane-bound cell organelles, filled with (proteolytic) enzymes in an acidic environment. For example, in senescent cells, the histone content in cytoplasmic chromatin fragments budded off the nucleus is processed by the lysosomal pathway [201]. Of note, cathepsin L – the mESC H3 cleavage enzyme – is also a lysosomal component and thus helps in mediating this histone processing. The proteasome processes proteins after they have been tagged for degradation with polyubiquitination marks. In the case of histones, also acetylation seems to play a role herein [202].

Degradation can also occur associated with specific transitions in development. For example, histones are removed during spermatogenesis when they need to be replaced with protamines [203]. Further on, upon early embryogenesis, these protamines and the remaining sperm histones again need to be degraded. Of note, this latter degradation process was recently attributed to a nuclear cathepsin L [204].

Also in the immune system histone proteolysis can fulfill specific roles [205, 206]. Neutrophil extracellular traps (NET) are composed of decondensed chromatin and antimicrobial proteins, and have evolved in neutrophils, amongst other cell types, as a way to eliminate invading pathogens. Neutrophil elastase plays an important role in these NET. During T-cell and natural killer cell granule-mediated apoptosis, histone digestion by GranzymeA arranges the unfolding of compacted chromatin and facilitates endogenous DNase access to DNA [206].

5.1.2. Histone clipping

Histone clipping represents the group of the more specific, epigenetically important proteolytic events and also the findings reported in this dissertation are dealt with as histone clipping.

The direct expression of truncated histone forms *in vivo* illustrates the potential molecular impact of a clipping event. Although the tails are dispensable for viability as was found for yeast [207–210] and human embryonic kidney cells [211, 212], the truncation of C- and N-termini has an important influence on transcription and chromatin compaction. This also illustrates the caveats inherent to studying histone clipping: as histone tails can be discarded genome-wide without affecting viability, while they can also be modified by an extensive mix of PTMs (as outlined in Table 1.1), cells seem to survive through redundancy of regulation. This redundancy might hamper the interpretation of experimental results.

Up to now, only H2A C-tail and H3 N-tail clipping have been described in considerable detail and these are outlined below more elaborately (also partially reviewed in [190]).

H2A clipping

The first reports in this context date from the 1970's and 80's [213–217]. The group of Moudrianakis described for the first time a chromatin-bound proteolytic activity in calf thymus with unique specificity for H2A. A "H2A-specific protease" was able to clip H2A after valine 114, under high ionic strength. A similar cleavage pattern was described in several myeloid and lymphatic leukemia cell types shortly after [218, 219]. However, it took another two decades for an additional confirmation of these results. The research group of Ohkawa found that an analogous clipping event and an additional N-terminal truncation took place in THP-I cells, an acute monocytic leukemia cell line [192, 193] after stimulation with RA or PMA. Only recently, we were able to define the so-called "H2A-specific protease" responsible for the V114-proteolytic activity as neutrophil elastase [194, 195].

H2A clipping has also been implicated in host defense. Enzymatically generated fragments of the H2A N-terminus can function as antimicrobial peptides called buforins. Such pepsin-mediated activity has also been found in the gastric fluid from human, bovine and pig species [220].

H3 clipping

In 2008, Duncan et al. published their findings on the N-terminal proteolytic processing of histone H3 in differentiating mESC [112, 127]. They found that cathepsin L cleaves H3, with alanine 21 as the primary clipping site, independently of the differentiation protocol applied. This manuscript inspired us to investigate whether hESC have a similar mechanism and define this clipping event in a broader context. Simultaneously another report described H3 truncation after alanine 21 in yeast during the stationary phase or sporulation by a serine protease [179]. Recently vacuolar proteinase B1 (PRB1) was proposed to be at least partially responsible for this clipping activity in yeast [191]. Glutamate dehydrogenase appears to be the responsible enzyme for H3 cleavage in chicken liver cells [182, 183]. In parallel, also in quail liver and oviduct H3 clipping appears under circumstances of aging and steroid induction [178, 221]. Legumain activity from colorectal cancer cells yields a 12 kDa H3 form, analogous to the one produced by cathepsin L in mESC [222]. The latest publications report histone H3 clipping activity in mammary glands upon involution and adipocyte differentiation [223, 224]. Two labs described this proteolytic event independently from each other, but they attributed the enzymatic activity to two different enzymes, namely calpain 1 and cathepsin D.

Most of the aforementioned clipping events share cleavage sites within the same H3 sequence stretch, namely amino acids 20 to 30 on the N-terminus, as was also found here for hESC. The fact that the same clipping site can be targeted by different enzymes and that multiple clipping sites can be generated by one proteolytic enzyme again alludes on redundancy, in line with the high sequence conservation of histones along eukaryotic evolution. E.g. Xue et al. [191] found that there was still a slight residual proteolytic activity in nuclear extracts of PRB1 mutants, a similar finding to what we have seen for histone H2A in neutrophil elastase knock-out mice [195]. Also the description of histone clipping during mammary gland involution and its attribution to two different enzymes, is an excellent example of the enzyme redundancy that might be at play. Furthermore, another indirect indication for redundancy was deduced from the fact that cathepsin L knock-out mice are still viable and fertile [127].

5.2. Histone H3 clipping in hESC

In this dissertation, we show for the first time that also in hESC histone H3 clipping takes place [44]. At first sight, this proteolytic event seems very similar to the one reported for mESC, but in fact turned out to be very different. Nevertheless, as this publication was the starting point for our own investigation, we first shortly summarize this report below.

Histone H3 can be N-terminally processed by cathepsin L upon mESC differentiation, whether occurring spontaneously, upon EB formation or induced by RA [127]. The H3.2 variant in particular appeared to be preferentially cleaved over H3.1, but H3.3 was no substrate for this proteolytic activity.

Since the H3 N-terminus in general is highly prone to PTMs, Duncan et al. also investigated the clipping's PTM environment. Both active and repressive histone marks were present and had the ability to modulate the cathepsin L activity. As such, dimethylation of H3K27 strongly increased the clipping activity, as did acetylation on H3K18. H3K23 acetylation on the contrary abrogated the cleavage almost completely, disregarding whether or not it was deposited in combination with K18 acetylation. A potential functional effect of this clipping, was that cleavage after alanine 21 greatly reduced the binding affinity of the Polycomb protein Chromobox homolog 7 (CBX7).

Although we also found the N-terminal processing of histone H3 in (whether or not differentiating) hESC, the latter appears to have a very different molecular regulation in comparison to mouse.

Culture conditions dictate H3 clipping activity in hESC

One of our most important conclusions is that the ESC culture conditions substantially influence the pattern of this proteolytic activity. Duncan et al. described a temporally different cleavage pattern in mESC depending on the differentiation protocol applied. After RA induction, the truncated H3 form is obvious in samples taken on day 2 and 3 of differentiation. In contrast, when differentiation is not forced by RA, the clipping event is delayed in time and does not fade as abruptly (in spontaneous differentiation) or does not disappear at all (after EB formation). Nevertheless, they seem to associate clipping with the differentiation process itself, disregarding the differentiation protocol applied and consider the RA-induced clipping as sufficiently representative to complete their experiments. In contrast, their following observations describe that the inhibition of cathepsin L does not influence Oct4 levels and does not have an influence on the maintenance of pluripotency or the capacity for differentiation, although the authors initially seemed to suggest so.

Our results in Chapter 3 and 4 illustrate that for hESC cultured in feeder-free conditions the cleavage pattern differs from that one seen in MEF-cultured hESC. Feeder-free hESC display more or less continuous H3 clipping, under both pluripotent and differentiating conditions (Figure 3.2), while feeder-culture hESC yield a similar pattern to that one of mESC after RA induction (Figure 3.7). Studying differentiation in MEF hESC cultures in a more extended and detailed time frame, even visualized a sort of undulating pattern of upcoming and decreasing H3 cleavage (Figure 4.1). The difference in clipping outcome between culture conditions might be explained by the difference in mechanical cues these culture systems induce. It is known that processes such as lineage specification and stem cell differentiation are also determined by the mechanical and geometrical properties of the cells' environment (e.g. 2D MEF culture versus 3D EB culture) [225]. In the case of ESC H3 clipping, the differential response to different culture conditions might be expressed in the form of a distinct

clipping pattern, as mechanical stress and forces acting on a cell may alter chromatin organization and gene expression.

Of note, one of the most important questions popping up in this context is whether there is an artificial *in vitro* contribution during the acid histone extraction procedure as the cells' compartmentalization is lost and their content is released. Based on the results of the different experiments outlined in Chapter 3 and 4, we believe that there is no *in vitro* clipping of histone H3. Similarly, we could not find any evidence of other core histones being cleaved and we interpret this as an additional argument for the specificity of this hESC H3 clipping event.

H3 clipping in hESC is not primarily determined by the cell cycle

We initially interpreted the undulating pattern of cH3 as potential evidence for cell cycle involvement. Also Duncan et al. state that the H3 clipping can be cell cycle-dependently regulated. In mESC, the H3.2 variant was preferably cleaved while contrastingly, H3.3 was no substrate for histone cleavage, which consequently was interpreted as an indication for being S phase- or replication-dependent [112]. The preferred clipping of H3.2, whose expression is tightly replication-coupled, over H3.3, which on the other hand can be expressed throughout the whole cell cycle, might indicate a preference of the cleavage event towards the S phase. Of note, mESC divide rapidly and have a relatively high proportion of cells residing in the S phase. In addition, they refer to the processing of the CDP/Cux transcription factor by cathepsin L, which also occurs during the G1/S phase transition [226]. The same reasoning might hold for hESC, as our data also indirectly indicate that H3.1 and/or H3.2 are the preferred cleaved variants in our set-up. However, our results from synchronization experiments on other cell lines and the vast influence of the culture conditions applied argue that the cell cycle is not the primary factor determining histone clipping.

The hESC clipping process does not seem differentiation-coupled

Analogously to H3 clipping in mESC, our first experiments were carried out only on differentiating hESC and thus initially supported such a correlation. Nevertheless, when introducing hESC maintained under pluripotent conditions (E8 medium) in a feeder-free culture, our results showed that also then histone H3 is continuously clipped. As no direct link to the Oct4-expression became apparent, a direct correlation to differentiation seemed rather unlikely (Figure 3.2). Although the meta-analysis of several previous reports on H3 seems to suggest that clipping is a plausible common feature of differentiation, it could not be confirmed in hESC that clipping is a specific differentiation-feature.

Nevertheless, it appears that cells require a certain degree of multipotency for the clipping to occur, illustrated by the fact that terminally differentiated Raji cells are incapable of H3 clipping in contrast to

oligopotent THP-I cells. In general, the more pluripotent the cells are, the more their chromatin structure is loosened – as such this might explain their higher susceptibility to histone clipping. Importantly, from a substrate's point of view, it might also be that the histone H3 content present in cells unable to form cH3 does not contain the 'correct epigenetic signature' or is not accessible enough to allow for clipping. Alternatively, it is possible that these cells fail to activate the responsible clipping enzyme(s), if they even express such protease(s).

The hESC clipping enzyme is a serine protease

Besides the characterization of the clipping event itself, we also tried to define the clipping enzyme responsible in hESC. We applied both a targeted and a general approach to deal with this question.

The targeted method focused mainly on cathepsin L, the responsible clipping enzyme in mESC and thus a very likely candidate for the clipping event in hESC [127]. In contrast to mouse cells, two variants of this cysteine protease can be present in human cells, namely cathepsin L (L1) and cathepsin V (L2) [227]. In terms of amino acid composition, cathepsin V exerts a higher sequence homology to the mouse variant than does cathepsin L but at the other hand its expression is much less ubiquitous. We took both variants into account for our experiments. Despite their ability to clip H3 *in vitro*, neither of them seems a plausible candidate for exerting H3 clipping in hESC, as none of the specific inhibitors for these enzymes or the general cysteine protease inhibitors could abrogate the cleavage event when histones were incubated with hESC NE that displayed clipping capability. Although the cleavage sites found in hESC are very similar to those in mESC and the clipping process itself seems highly conserved, these findings caution for blindly extrapolating results from one system to another.

In analogy, we also briefly investigated the expression of neutrophil elastase in hESC, as we just identified this to be the H2A-specific protease, the serine protease which clips histone H2A after valine 114 [195]. However, we could not find conclusive evidence for such presence in hESC.

In parallel, we also incubated purified calf H3 with hESC NE supplemented with a whole range of protease inhibitors, as outlined in Section 3.4.6. As the serine protease inhibitor AEBSF was the only inhibitor able to abrogate the processing of histone H3, we concluded that a serine protease is responsible (at least in part) for the clipping activity in hESC. As mentioned earlier (*Section 5.1.2*), possible enzymatic redundancy as seen by Xue et al. [191], is still an important consideration.

Whether the clipping enzyme - assumed to be a serine protease - is solely located in the nucleus also needs to be explored in the future. It might very well be that the cleavage is catalyzed by e.g. a cytoplasmic enzyme that traffics to the nucleus due to a certain trigger. Several examples of nuclear

active enzymes usually located elsewhere have already been reported, e.g. cathepsin L in mESC and NIH3T3 cells [127, 226], and cathepsin B in thyroid carcinoma cells [228].

The potential (biological) outcomes of histone clipping

The histone H3 N-terminus protrudes out of the nucleosome and is thus very susceptible for a whole range of PTMs (as illustrated in Table 1.1), among which also histone clipping. One exemplary case of the involvement of these PTMs in epigenetic regulation, is the interaction of the modifications deposited by the PcG and trxG proteins. The group of PcG proteins can be further subdivided into several Polycomb repressive complexes (PRC1 and PRC2), which respectively mediate H2A ubiquitination and H3 methylation. In this hierarchical recruitment model, PRC2 first installs a trimethyl group on H3K27, which in turn serves as an anchor site for the PRC1 complex that subsequently deposes a ubiquitin on lysine 119 of H2A. The trxG proteins on the other hand trimethylate H3K4 and have an activating effect on transcription. When occurring together in bivalent domains, these modifications counteract to result in silent genes that are poised for activation. Obviously, histone tail clipping can severely disturb this mechanism (and its downstream effects) or any other PTM cascade. Also other arginine and lysine methylations, lysine acetylation, citrullination and phosphorylation reside on the histone H3 tail, each bringing along their own effect. The latter three types of modification usually have an activating effect towards transcription, whilst for methylation the situation is more context- and residue-dependent and can lead both to activation or repression of the transcription process.

As most PTMs exert their effect via mediator proteins, histone clipping will not only impair the deposition of other PTMs but also the binding of their according effector proteins, as e.g. Psathas et al. show that the H3 N-tail removal interferes with intratail regulatory mechanisms [229]. Additionally, also Duncan found that the binding of the Polycomb protein CBX7 decreased significantly when the N-tail was cleaved off. On the other hand, it might also provide (a) new binding site(s) for other regulatory proteins that could not bind before due to e.g. sterical hindrance or the presence of other inhibitory modifications.

When considering the more global level of chromatin organization, it is plausible that clipping can tag nucleosomes that need to be evicted and replaced by new canonical and/or variant histones. As preferably H3.1 and 3.2 are cleaved over H3.3, this might represent a mechanism for H3.3 to be included after removal of the other two variants. Also the higher-order compaction of the chromatin fiber will be affected by histone clipping, as histone N-termini have proven to play an important role in the organization of the general chromatin structure.

For interpreting H3 clipping in mESC, it is noteworthy that cells lacking cathepsin L do have an abnormal distribution of euchromatic and heterochromatic marks and factors across the genome, suggesting a role for the cathepsin L activity in the epigenetic regulation of gene expression [107]. Another article worth mentioning in this context is that of Pauli et al. [230], who encountered an H3 clipping event in protein extracts from Drosophila salivary glands. Their experiments served to investigate the transcriptional changes caused by depletion of cohesin, a multisubunit complex essential for sister chromatid cohesion and mitotic chromosome segregation. The depletion of one of the subunits of cohesion, Rad21, induced some drastic (posttranslational) alterations, e.g. upregulation of actin protein levels and histone H3 clipping. Surprisingly, coincidentally with the H3 clipping there appeared a presumably heavily posttranslationally modified form of RNA polymerase II (RNAPII). Personal communication with the authors learned that they assumed it to be a phosphorylated or even a ubiquitinated form of RNAPII, probably emerging as a stress-response, but they did not further pursue this matter. Phosphorylation of the serine and threonine residues in its carboxy-terminal domain aids in RNAPII functioning and is essential for correct 3' end processing, 5' capping, transcription termination, etc. [231]. These phosphorylated forms are enriched at active and/or poised genes [232]. As such, this RNAPII modification and histone clipping might be possibly co-operating to promote transcription at certain regions.

The regulation of this enzymatic activity also remains elusive. Duncan et al. already showed that the neighboring PTMs on the H3 tail in mESC do have a substantial influence on the clipping activity, as outlined above. This effect is most probably due to its influence on the level of histone folding and interaction, as differential modification of histone peptide substrates does not substantially impact cathepsin L activity [174]. Alternatively, the presence of natural enzyme inhibitors in the nucleus might provide another level of clipping regulation. Ceru et al. already showed the interaction of stefin B (cystatin B), an endogenous cysteine protease inhibitor, with histones and cathepsin L in the nucleus of T98G human glioblastoma cells amongst other cell lines [227]. Increased levels of stefin B will thus influence the cathepsin L activity on e.g. the CDP/Cux transcription factor and help in regulating cell cycle progression into the S phase. The interaction with histone H2A.Z, H2B and H3 *in vitro* even seemed to enhance the inhibitory activity of stefin B on cathepsin L.

Considering all of above arguments, although only delivering indirect evidence, clipping is definitely a plausible candidate to be involved in the complex network regulating transcription, also in hESC.

Technical implications

On top of the biological consequences, we also should take some technical implications of histone clipping into account. All techniques that are focusing on a specific modification within the H3 N-

terminus or other histone tails such as e.g. chromatin immunoprecipitation, risk being blind for clipping events and hence create biased results. From our western blotting analyses it is clear that the amount of cH3 can be up to 50 % or more of the total histone content at certain time points in our hESC differentiation experiments, illustrating the extent of the potential impact. The lack of detection can be minimized by applying different antibodies to monitor both the N- and C-terminus of the core histones, as we did for histone H3 in hESC. In addition, also techniques such as Top Down mass spectrometry could provide an applicable platform, as they take all proteoforms into consideration. Nevertheless, only by circumventing the need for isolating histones all together, the presence of and the correlation between histone clipping and degradation can be studied in a cellular context, which could for example be attained by directly staining truncated histone forms in fixed cells.

5.3. Conclusions

We have shown for the first time that histone H3 clipping needs to be considered as an important player in the histone landscape of hESC, adding another layer of complexity to the already extended network of PTMs. We conclude that the applied culture conditions exert a substantial influence on the clipping pattern and that the underlying mechanism is at least more determining the clipping activity, than does the process of differentiation or cell cycle organization. Nevertheless, a certain degree of differentiation capacity seems required for the clipping to occur, as terminally differentiated cell lines are incapable of forming cH3. The real regulating mechanisms behind histone clipping remains thus elusive.

The responsible clipping enzyme in this context is most probably a serine protease. Cathepsin L/V and neutrophil elastase, previously shown to be the clipping enzyme in mESC (H3) and myeloid cell lines (H2A) respectively, were excluded as potential clipping candidates in hESC.

Finally, we propose a categorization for histone proteolysis into “histone clipping” and “histone degradation”, although we also emphasize that these classes are hard to distinguish. Despite numerous previous reports and the piece to the puzzle that was added in this dissertation, histone clipping remains an epigenetic outcast.

Summary

Samenvatting

SUMMARY

Epigenetic mechanisms help in regulating gene expression and variations thereof without altering the primary DNA sequence. Instead, these mechanisms are based on posttranslational histone modifications (PTM), incorporation of histone variants, DNA methylations and RNA-based processes. One very relevant model to study these modifications are human embryonic stem cells (hESC). These cells are hallmarked by their indefinite capacity for self-renewal and their ability to differentiate towards all cell types of the adult body. The biology of hESC is more comprehensively described in the first part of **Chapter 1**. Subsequently this chapter also elaborates on the domain of epigenetics, the involved mechanisms and the levels on which they manifest. Submitting this back in a stem cell context, epigenetics is shown to be the key to coordinate the development of all different cell types within an organism, as all cells of that organism contain the same genetic material.

When differentiation is induced, the epigenetic signature of hESC alters substantially, which as a consequence will also affect their transcriptional activity. Hence, e.g. the expression of the core pluripotency factors such as Oct4 decreases upon the differentiation process. These changes in expression levels can be monitored via a number of techniques. One such, often employed method is relative quantification via reverse transcription-quantitative PCR (RT-qPCR). To make RT-qPCR analysis successful several possible (technical) variations need to be taken into account, such as sample handling, sample quality and enzymatic (in)efficiencies, which can be normalized for during data processing. This normalization is based on e.g. on one or more reference genes amongst other possible strategies. These reference genes should be equally expressed throughout all samples and thus need to be carefully selected for each experimental set-up. Thus **Chapter 2** summarizes suchlike optimization for differentiating hESC. By use of the geNorm algorithm (qbase+ software), we found that the genes *β-2-microglobulin* (*B2M*), *ribosomal protein L13A* (*RPL13A*) and Alu repeats were the most stable reference loci in the setting we applied (i.e. hESC differentiation induced by retinoic acid). Moreover, their combination appeared to be much better than applying the 'classic' references such as *glyceraldehyde-3-phosphate dehydrogenase* (*GAPDH*) or *β-actin* (*ACTB*). As stated before, this method can thus conveniently be applied to validate if hESC are keeping up in culture or to monitor how they evolve during differentiation experiments.

Getting back to the epigenetics part, the main question addressed in this dissertation is whether histone clipping also plays an important role in hESC. Histones are highly conserved proteins that compact DNA into the nucleus of a cell and help regulating its transcription. The four core histones compact together in duplo to form the histone octamer, around which circa 146 bp of DNA is wrapped, forming altogether the nucleosome, alias the building block of chromatin. The tails of these

small basic proteins protrude out of the nucleosome and are thus highly susceptible for posttranslationally modifying enzymes such as methyltransferases and demethylases, kinases, acetyltransferases and deacetylases, etc. These modifications can participate in gene expression regulation via direct genome accessibility management and/or recruitment of other effector proteins involved in the transcription process.

One such modification of particular interest to us is histone clipping, a process in which a series of amino residues gets to be enzymatically cleaved off of a histone tail. Histone proteolysis in general can be roughly subdivided in "histone degradation" and "histone clipping", but only in the latter category the processing of histones is more considered as an epigenetically important PTM. This proteolytic event has been described in several cell types, tissues and organisms under different circumstances and for different histones. For example, H2A clipping has been described in myeloid cells catalyzed by neutrophil elastase. H3 clipping has been investigated in more detail and has been reported in yeast, chicken and quail liver, colorectal cells and mammary glands. A notably interesting report discusses the clipping of histone H3 primarily after alanine 21 by the cysteine protease cathepsin L in differentiating mouse ESC (mESC). The authors found that this event could have a substantial influence on transcription by e.g. impacting the binding affinity of effector proteins such as the Polycomb protein Chromobox homolog 7, an enzyme involved in transcriptional silencing.

In analogy, we analyzed whether this N-terminal H3 cleavage is also present in hESC, in which basic processes such as pluripotency and self-renewal are equally highly epigenetically determined. Indeed, we were able to show for the first time that this clipping event is actually taking place in hESC. Our findings are outlined in **Chapter 3** and **4** and further discussed in **Chapter 5**. We studied several hESC lines, in feeder-free culture as well as maintained on a feeder layer of mouse embryonic fibroblasts (MEF), during various differentiation experiments. The clipping of histone H3 -analyzed by western blotting- appeared in both culture systems, albeit in a different pattern which highlights the influence culture conditions can exert. Feeder-free culture hESC displayed continuous clipping during the course of our experiments, whilst more of an undulating pattern with fluctuating clipped H3 (cH3) intensity was seen in MEF cultured hESC. When elaborating on the feeder-free hESC, we saw that they continued to show this clipping even after extension of the studied time frame to two weeks. Moreover, whether they were being forced into differentiation or not (by application of retinoic acid) did not seem to have an influence on the character of the cleavage. Since we could not find an obvious direct correlation with Oct4 expression, we suspect that this clipping event in hESC might not be related the process of differentiation, although contrastingly previous cH3 reports pointed in that direction and yet some differentiation capacity is required. However, this clipping process might very

well have different functionalities depending on the setting. Along the course of our experiments, we could also exclude the possible *in vitro* origin of this truncated H3 form during the histone extraction procedure. Furthermore, we were able to validate three clipping sites in hESC by means of complementary techniques: alanine 21 (western blotting), arginine 26 (Edman degradation) and residue 31 (mass spectrometric analysis). Of note, these cleavage sites are very similar to the ones found in mESC.

In the pursuit of finding the responsible clipping enzyme, we set up an incubation assay in which purified calf H3 was incubated with the nuclear protein fraction isolated from hESC. By applying several inhibitors for the distinct protease classes, we found that not cathepsin L but a serine protease instead is processing histone H3 in hESC.

Despite the reports yet published on histone clipping, this topic has remained a surprisingly understudied modification. Though more research is needed to elucidate its exact role in the epigenetic scenery, we nevertheless showed that histone cleavage is also occurring in (differentiating) hESC.

Considering its potential effects, it might represent an epigenetically important process. Histone tails seem dispensable for a cell's viability but do play an essential role in the organization of the higher-order structure of chromatin. Proteolytic processing of these tails could thus evidently have a substantial influence on that organization. Clipping could also abrogate the binding of other effector proteins on the histone tails by removing the site of interaction or alternatively tag histones to be excised and replaced by e.g. a histone variant and thus as such provide new binding sites.

Additionally to its potential biological influence, histone proteolysis should also be taken into account from a technical point of view. All studies using techniques based on the specific detection of particular modifications within the first part of the H3 N-terminus (e.g. chromatin immunoprecipitation) may generate biased results because of the lack of those modifications' detection. Instead, methods that monitor the total protein with its corresponding proteoforms, such as Top Down mass spectrometry, are better suited for this purpose.

In conclusion, more research will be needed to fully uncover the purpose of histone clipping but we can yet conclude that this proteolytic event definitely claims its spot, not only in hESC but also in other biological systems.

SAMENVATTING

Epigenetica bestudeert de mechanismen die instaan voor de regulatie van genexpressie waarbij geen veranderingen aan de primaire DNA-sequentie aangebracht worden. In plaats daarvan zijn deze mechanismen gebaseerd op posttranslationele histonmodificaties, het incorporeren van histonvarianten, DNA-methylatie en RNA-gebaseerde processen. Een dankbaar model om deze modificaties te bestuderen zijn humane embryonale stamcellen (hESC). Deze worden gekenmerkt door hun capaciteit om zichzelf *in vitro* oneindig in stand te houden en door hun vermogen om te differentiëren naar eender welk celtype van het volwassen lichaam. Aldus wordt in het eerste deel van **Hoofdstuk 1** een overzicht van hESC biologie gegeven. Vervolgens gaat dit hoofdstuk dieper in op epigenetica en de mechanismen die daarbij betrokken zijn. Epigenetische processen vervullen een sleutelrol binnen de ontwikkeling van alle verschillende celtypes in een organisme, gezien het feit dat alle cellen van dat organisme toch hetzelfde genetisch materiaal bevatten.

De epigenetische handtekening van hESC verandert sterk wanneer differentiatie geïnduceerd wordt en bijgevolg zal ook hun transcriptionele activiteit wijzigen. Zo zal de expressie van bepaalde specifieke pluripotentie-transcriptiefactoren zoals Oct4 sterk dalen tijdens het differentiatie-proces. Deze veranderingen in expressie kunnen met behulp van verschillende technieken gevolgd worden, waaronder relatieve kwantificatie via reverse transcriptie-kwantitatieve PCR (RT-qPCR). Om RT-qPCR correct toe te passen, moeten tijdens de data-analyse verschillende (technische) variabelen in rekening gebracht worden zoals de kwaliteit van de stalen en de efficiëntie van de verschillende enzymatische reacties. Om zulk een normalisatie-strategie op te stellen kan onder meer gebruikt gemaakt worden van een calibrator-staal en één of meerdere referentie-genen. Zulke referentie-genen worden geacht een stabiel expressie-patroon te vertonen over de stalen heen en moeten zorgvuldig gekozen worden per nieuwe experimentele set-up. In **Hoofdstuk 2** wordt een dergelijke optimalisatie van de meest gepaste referentie-genen voor differentiërende hESC toegelicht. Analyse met geNorm wees uit dat in deze setting van retinoïnezuur-geïnduceerde hESC *β-2-microglobuline* (*B2M*), *ribosomaal proteïne L13A* (*RPL13A*) en Alu repeats de meest stabiele referentie-genen zijn. Bovendien bleek de toepassing van hun combinatie beter dan die van de klassieke en veelgebruikte *glyceraldehyde-3-fosfaat dehydrogenase* (*GAPDH*) en *β-actine* (*ACTB*). Zodoende kan deze geoptimaliseerde methode gebruikt worden om de pluripotentie van hESC in cultuur te monitoren of om hun evolutie na te gaan tijdens een differentiatie-experiment.

De belangrijkste onderzoeksvraag binnen deze thesis was of histonklieving ook plaatsvindt in een hESC setting, wat ons terug brengt naar de epigenetica. Histonen zijn sterk geconserveerde eiwitten die zorgen voor de compacte verpakking van DNA in een celkern en die helpen in de regulatie van het

transcriptie proces. De vier kernhistonen worden in duplo gecombineerd tot een histon-octameer, waarrond vervolgens ongeveer 146 bp DNA gewikkeld worden en zo de nucleosomen, ofwel de basis-units van het chromatine, vormen. De uiteinden of 'staarten' van deze kleine, basische eiwitten steken uit het nucleosoom en zijn dus zeer vatbaar voor posttranslationeel modificerende enzymen zoals methyltransferases en demethylases, kinases, acetyltransferases en deacetylases, enz. Deze modificaties spelen een belangrijke rol in de regulatie van genexpressie doordat zij rechtstreeks de toegankelijkheid van het DNA kunnen beïnvloeden of meewerken aan de rekrutering van andere effector-proteïnen betrokken in transcriptie.

Een dergelijke modificatie is histonklieving, waarbij een reeks aminozuren van een histonstaart enzymatisch verwijderd wordt. Algemeen kan histonproteolyse onderverdeeld worden in enerzijds histondegradatie en anderszijds histonklieving, waarbij het proces enkel in de laatste categorie als een epigenetisch belangrijke histonmodificatie beschouwd wordt. Histonklieving is reeds beschreven voor verschillende celtypes en organismen, voor verschillende biologische settings én voor verschillende histonen. Bijvoorbeeld H2A-klieving, gekatalyseerd door neutrofiel elastase, werd reeds beschreven in myeloïde cellijnen en H3-klieving werd al gerapporteerd in gistcellen, kip- en kwartel-levercellen, colorectale cellen en melkklierweefsel. Een specifiek belangrijke publicatie handelt over H3-klieving, voornamelijk na alanine 21, gekatalyseerd door het cysteïne protease cathepsine L in differentiërende muis ESC (mESC). De auteurs vonden dat deze klieving wel degelijk een beduidende invloed kan hebben op transcriptie, bijvoorbeeld door het beïnvloeden van de bindingsaffiniteit van Polycomb Group eiwitten zoals Chromobox homolog 7, een enzym betrokken in transcriptie-respressie.

Analoog trachtten wij hier het voorkomen van deze N-terminale H3 klieving na te gaan in humane ESC, aangezien ook in deze cellen basis-processen zoals zelf-vernieuwing en pluripotentie grotendeels epigenetisch geregeld worden. Als dusdanig konden wij voor het eerst aantonen dat deze modificatie ook in hESC wel degelijk plaatsvindt. **Hoofdstuk 3** en **4** geven onze belangrijkste onderzoeksresultaten weer, die vervolgens verder besproken worden in **Hoofdstuk 5** met de bijhorende conclusies. In onze opzet werden verschillende hESC lijnen opgenomen, zowel gekweekt op een feeder layer van muis embryonale fibroblasten als in feeder-vrije culturen, en geanalyseerd via western blotting. Ongeacht de cellijn, werd de klieving onder de verschillende cultuuromstandigheden teruggevonden, weliswaar in een ander patroon, wat de mogelijke invloed van de cultuurcondities sterk benadrukt. Feeder-vrij gecultiveerde hESC vertoonden een continu klievingspatroon binnen de differentiatie-experimenten, terwijl er bij de op MEF gehouden hESC eerder sprake is van een soort golfpatroon met een fluctuerende intensiteit van de geknipte H3-vorm. Bij het uitbreiden van de feeder-vrije experimenten naar een totale duur van twee weken bleef de klieving aanhouden. Uit deze

experimenten kon bovendien afgeleid worden dat het wel of niet toepassen van differentiatie (door toevoeging van retinoïnezuur aan het cultuurmedium) geen verschil uitmaakt voor het bekomen resultaat. Aangezien we geen directe link konden vinden tussen de klieving en de expressie van Oct4, lijkt de beschouwing van de klieving als een algemeen differentiatie-mechanisme niet plausibel – hoewel vorige publicaties wel in die richting deden vermoeden. Niettemin is het wel mogelijk dat de klieving een verschillende functionele rol speelt naargelang de biologische set-up waarin het voorkomt. Doorheen deze experimenten konden we tevens mogelijke *in vitro*-klieving tijdens de histonextractie uitsluiten. Verder konden we ook drie verschillende klievingsplaatsen identificeren met behulp van complementaire technieken, namelijk alanine 21 (western blotting), arginine 26 (Edman degradatie) en aminozuur 31 (massa spectrometrie). Niet onbelangrijk is dat deze sites zeer gelijk(aardig) zijn aan degene teruggevonden in mESC.

In de context van het klievingsenzym, ontwikkelden we een incubatie assay waarin opgezuiverd kalf H3 geïncubeerd werd met de nucleaire eiwitfractie geïsoleerd uit hESC. Door hierin meerdere inhibitoren voor verschillende protease-klassen te includeren, vonden we dat niet cathepsine L maar wel een serine protease verantwoordelijk is voor de H3-klieving in hESC.

Niettegenstaande er al verschillende publicaties over histonklieving verschenen zijn, blijft het een relatief weinig bestudeerde modificatie. En hoewel er meer onderzoek nodig is om het totaalbeeld op te helderen, konden wij hier aantonen dat deze klieving ook voorkomt in (differentiërende) hESC. De potentiële effecten ervan beschouwend, kunnen we stellen dat histonklieving mogelijk een epigenetisch belangrijk regulatiemechanisme is. Histonstaarten lijken niet essentieel voor de viabiliteit van een cel maar spelen wel een cruciale rol in de organisatie van de chromatine-structuur. Het knippen van de histonstaarten kan dus een belangrijke invloed hebben op die organisatie. Ook kan de binding van andere effector-proteïnen door klieving verhinderd worden doordat de interactiesite verwijderd is of kunnen daar tegenover gesteld bestaande histonen als het ware 'gelabeld' worden om vervolgens verwijderd en vervangen te worden door bijvoorbeeld een andere histonvariant en zo nieuwe bindingssites voor andere effectors te creëren.

Ook vanuit technisch oogpunt moet dit proces in rekening gebracht worden. Alle technieken die gefocust zijn op een specifieke modificatie op het eerste deel van de H3 N-terminus (zoals chromatine immunoprecipitatie) kunnen een bias aan hun resultaten meegeven doordat ze dergelijke modificaties kunnen missen in geval van histonklieving. In plaats hiervan zouden technieken moeten toegepast worden die het eiwit in zijn geheel met bijhorende proteovormen, kan monitoren, zoals bijvoorbeeld Top Down massaspectrometrie.

Kort samengevat kan gesteld worden dat er nog meer onderzoek nodig is om het volledige karakter van de klieving en zijn (biologische) rol op te helderen, maar niettemin eist histonklieving zijn plaatsje op, niet enkel in hESC maar ook in andere biologische settings.

REFERENCES

1. **The 2012 Nobel Prize in Physiology or Medicine - Press Release** [http://www.nobelprize.org/nobel_prizes/medicine/laureates/2012/press.html]

2. Yamanaka S: **Induced pluripotent stem cells: past, present, and future.** *Cell Stem Cell* 2012, **10**:678–84.

3. Nebbioso A, Carafa V, Benedetti R, Altucci L: **Trials with "epigenetic" drugs: an update.** *Mol Oncol* 2012, **6**:657–82.

4. Stephenson EL, Braude PR, Mason C: **Proposal for a universal minimum information convention for the reporting on the derivation of human embryonic stem cell lines.** *Regen Med* 2006, **1**:739–50.

5. O'Leary T, Heindryckx B, Lierman S, van Bruggen D, Goeman JJ, Vandewoestyne M, Deforce D, de Sousa Lopes SMC, De Sutter P: **Tracking the progression of the human inner cell mass during embryonic stem cell derivation.** *Nat Biotechnol* 2012, **30**:278–82.

6. Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, Jones JM: **Embryonic stem cell lines derived from human blastocysts**. *Science (80-)* 1998, **282**:1145–1147.

7. Becker KA, Ghule PN, Therrien JA, Lian JB, Stein JL, van Wijnen AJ, Stein GS: **Self-renewal of human embryonic stem cells is supported by a shortened G1 cell cycle phase**. *J Cell Physiol* 2006, **209**:883–893.

8. Kapinas K, Grandy R, Ghule P, Medina R, Becker K, Pardee A, Zaidi SK, Lian J, Stein J, van Wijnen A, Stein G: **The abbreviated pluripotent cell cycle.** *J Cell Physiol* 2013, **228**:9–20.

9. Adewumi O, Aflatoonian B, Ahrlund-Richter L, Amit M, Andrews PW, Beighton G, Bello PA, Benvenisty N, Berry LS, Bevan S, Blum B, Brooking J, Chen KG, Choo AB, Churchill GA, Corbel M, Damjanov I, Draper JS, Dvorak P, Emanuelsson K, Fleck RA, Ford A, Gertow K, Gertsenstein M, Gokhale PJ, Hamilton RS, Hampl A, Healy LE, Hovatta O, Hyllner J, et al.: **Characterization of human embryonic stem cell lines by the International Stem Cell Initiative**. *Nat Biotechnol* 2007, **25**:803–816.

10. Smith KP, Luong MX, Stein GS: **Pluripotency: toward a gold standard for human ES and iPS cells**. *J Cell Physiol* 2009, **220**:21–29.

11. Hoffman LM, Carpenter MK: **Characterization and culture of human embryonic stem cells.** *Nat Biotechnol* 2005, **23**:699–708.

12. Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG, Gifford DK, Melton DA, Jaenisch R, Young RA: **Core transcriptional regulatory circuitry in human embryonic stem cells**. *Cell* 2005, **122**:947–956.

13. Vossaert L, O Leary T, Van Neste C, Heindryckx B, Vandesompele J, De Sutter P, Deforce D: **Reference loci for RT-qPCR analysis of differentiating human embryonic stem cells.** *BMC Mol Biol* 2013, **14**:21.

14. Hiyama E, Hiyama K: **Telomere and telomerase in stem cells**. *Br J Cancer* 2007, **96**:1020–1024.

15. Hasegawa K, Pomeroy JE, Pera MF: **Current technology for the derivation of pluripotent stem cell lines from human embryos.** *Cell Stem Cell* 2010, **6**:521–31.

16. Stacey GN, Cobo F, Nieto A, Talavera P, Healy L, Concha A: **The development of "feeder" cells for the preparation of clinical grade hES cell lines: challenges and solutions.** *J Biotechnol* 2006, **125**:583–8.

17. Richards M, Fong C-Y, Chan W-K, Wong P-C, Bongso A: **Human feeders support prolonged undifferentiated growth of human inner cell masses and embryonic stem cells.** *Nat Biotechnol* 2002, **20**:933–6.

18. Crook JM, Peura TT, Kravets L, Bosman AG, Buzzard JJ, Horne R, Hentze H, Dunn NR, Zweigerdt R, Chua F, Upshall A, Colman A: **The Generation of Six Clinical-Grade Human Embryonic Stem Cell Lines.** *Cell Stem Cell* 2007, **1**:490–494.

19. Villa-Diaz LG, Pacut C, Slawny NA, Ding J, O'Shea KS, Smith GD: **Analysis of the factors that limit the ability of feeder cells to maintain the undifferentiated state of human embryonic stem cells**. *Stem Cells Dev* 2009, **18**:641–651.

20. Zhou D, Liu T, Zhou X, Lu G: **Three key variables involved in feeder preparation for the maintenance of human embryonic stem cells**. *Cell Biol Int* 2009, **33**:796–800.

21. Tannenbaum SE, Turetsky TT, Singer O, Aizenman E, Kirshberg S, Ilouz N, Gil Y, Berman-Zaken Y, Perlman TS, Geva N, Levy O, Arbell D, Simon A, Ben-Meir A, Shufaro Y, Laufer N, Reubinoff BE: **Derivation of xeno-free and GMP-grade human embryonic stem cells - platforms for future clinical applications.** *PLoS One* 2012, **7**:e35325.

22. Nakagawa M, Taniguchi Y, Senda S, Takizawa N, Ichisaka T, Asano K, Morizane A, Doi D, Takahashi J, Nishizawa M, Yoshida Y, Toyoda T, Osafune K, Sekiguchi K, Yamanaka S: **A novel efficient feeder-free culture system for the derivation of human induced pluripotent stem cells.** *Sci Rep* 2014, **4**:3594.

23. Xu C, Inokuma MS, Denham J, Golds K, Kundu P, Gold JD, Carpenter MK: **Feeder-free growth of undifferentiated human embryonic stem cells**. *Nat Biotechnol* 2001, **19**:971–974.

24. Braam SR, Zeinstra L, Litjens S, Ward-van Oostwaard D, van den Brink S, van Laake L, Lebrin F, Kats P, Hochstenbach R, Passier R, Sonnenberg A, Mummery CL: **Recombinant vitronectin is a functionally defined substrate that supports human embryonic stem cell self-renewal via alphavbeta5 integrin.** *Stem Cells* 2008, **26**:2257–65.

25. **Essential 8™ Medium - Life Technologies**
[http://www.lifetechnologies.com/order/catalog/product/A1517001]

26. Chen G, Gulbranson DR, Hou Z, Bolin JM, Ruotti V, Probasco MD, Smuga-Otto K, Howden SE, Diol NR, Propson NE, Wagner R, Lee GO, Antosiewicz-Bourget J, Teng JMC, Thomson JA: **Chemically defined conditions for human iPSC derivation and culture.** *Nat Methods* 2011, **8**:424–9.

27. Fernandes AM, Meletti T, Guimarães R, Stelling MP, Marinho PAN, Valladão AS, Rehen SK: **Worldwide survey of published procedures to culture human embryonic stem cells.** *Cell Transplant* 2010, **19**:509–23.

28. Zachar V, Prasad SM, Weli SC, Gabrielsen A, Petersen K, Petersen MB, Fink T: **The effect of human embryonic stem cells (hESCs) long-term normoxic and hypoxic cultures on the maintenance of pluripotency**. *Vitr Cell Dev Biol Anim* 2010, **46**:276–283.

29. Lengner CJ, Gimelbrant AA, Erwin JA, Cheng AW, Guenther MG, Welstead GG, Alagappan R, Frampton GM, Xu P, Muffat J, Santagata S, Powers D, Barrett CB, Young RA, Lee JT, Jaenisch R, Mitalipova M: **Derivation of pre-X inactivation human embryonic stem cells under physiological oxygen concentrations.** *Cell* 2010, **141**:872–83.

30. Ezashi T, Das P, Roberts RM: **Low O2 tensions and the prevention of differentiation of hES cells.** *Proc Natl Acad Sci U S A* 2005, **102**:4783–8.

31. Singh RK, Gaikwad SM, Chatterjee S, Ray P: **Stem Cells: The Holy Grail of Regenerative Medicine**. In *Eng Transl Med*. Edited by Cai W. London: Springer London; 2014:19–69.

32. Reubinoff BE, Pera MF, Fong CY, Trounson A, Bongso A: **Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro.** *Nat Biotechnol* 2000, **18**:399–404.

33. Trounson A: **The production and directed differentiation of human embryonic stem cells**. *Endocr Rev* 2006, **27**:208–219.

34. Takahashi T, Lord B, Schulze PC, Fryer RM, Sarang SS, Gullans SR, Lee RT: **Ascorbic acid enhances differentiation of embryonic stem cells into cardiac myocytes.** *Circulation* 2003, **107**:1912–6.

35. Chen S, Borowiak M, Fox JL, Maehr R, Osafune K, Davidow L, Lam K, Peng LF, Schreiber SL, Rubin LL, Melton D: **A small molecule that directs differentiation of human ESCs into the pancreatic lineage.** *Nat Chem Biol* 2009, **5**:258–65.

36. Duggal G, Heindryckx B, Warrier S, O'Leary T, Van der Jeught M, Lierman S, Vossaert L, Deroo T, Deforce D, Chuva de Sousa Lopes SM, De Sutter P: **Influence of activin a supplementation during human embryonic stem cell derivation on germ cell differentiation potential.** *Stem Cells Dev* 2013, **22**:3141–55.

37. Tonge PD, Andrews PW: **Retinoic acid directs neuronal differentiation of human pluripotent stem cell lines in a non-cell-autonomous manner**. *Differentiation* 2010, **80**:20–30.

38. Rohwedel J, Guan K, Wobus AM: **Induction of cellular differentiation by retinoic acid in vitro**. *Cells Tissues Organs* 1999, **165**:190–202.

39. Becker KA, Stein JL, Lian JB, van Wijnen AJ, Stein GS: **Human embryonic stem cells are pre-mitotically committed to self-renewal and acquire a lengthened G1 phase upon lineage programming**. *J Cell Physiol* 2009, **222**:103–110.

40. Calder A, Roth-Albin I, Bhatia S, Pilquil C, Lee JH, Bhatia M, Levadoux-Martin M, McNicol J, Russell J, Collins T, Draper JS: **Lengthened G1 phase indicates differentiation status in human embryonic stem cells.** *Stem Cells Dev* 2013, **22**:279–95.

41. Draper JS, Pigott C, Thomson JA, Andrews PW: **Surface antigens of human embryonic stem cells: changes upon differentiation in culture**. *J Anat* 2002, **200**(Pt 3):249–258.

42. Noaksson K, Zoric N, Zeng X, Rao MS, Hyllner J, Semb H, Kubista M, Sartipy P: **Monitoring differentiation of human embryonic stem cells using real-time PCR**. *Stem Cells* 2005, **23**:1460–1467.

43. Scheerlinck E, Van Steendam K, Vandewoestyne M, Lepez T, Gobin V, Meert P, Vossaert L, Van Nieuwerburgh F, Van Soom A, Peelman L, Heindryckx B, De Sutter P, Dhaenens M, Deforce D: **Detailed method description for noninvasive monitoring of differentiation status of human embryonic stem cells.** *Anal Biochem* 2014, **461**:60–66.

44. Vossaert L, Meert P, Scheerlinck E, Glibert P, Van Roy N, Heindryckx B, De Sutter P, Dhaenens M, Deforce D: **Identification of histone H3 clipping activity in human embryonic stem cells**. *Stem Cell Res* 2014, **13**:123–134.

45. Welling M, Geijsen N: **Uncovering the true identity of naïve pluripotent stem cells.** *Trends Cell Biol* 2013, **23**:442–8.

46. Nichols J, Smith A: **Naive and primed pluripotent states.** *Cell Stem Cell* 2009, **4**:487–92.

47. Evans MJ, Kaufman MH: **Establishment in culture of pluripotential cells from mouse embryos.** *Nature* 1981, **292**:154–6.

48. Martin GR: **Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells.** *Proc Natl Acad Sci U S A* 1981, **78**:7634–8.

49. Coronado D, Godet M, Bourillot P-Y, Tapponnier Y, Bernat A, Petit M, Afanassieff M, Markossian S, Malashicheva A, Iacone R, Anastassiadis K, Savatier P: **A short G1 phase is an intrinsic determinant of naïve embryonic stem cell pluripotency.** *Stem Cell Res* 2013, **10**:118–31.

50. Ginis I, Luo Y, Miura T, Thies S, Brandenberger R, Gerecht-Nir S, Amit M, Hoke A, Carpenter MK, Itskovitz-Eldor J, Rao MS: **Differences between human and mouse embryonic stem cells.** *Dev Biol* 2004, **269**:360–80.

51. Dahéron L, Opitz SL, Zaehres H, Lensch MW, Lensch WM, Andrews PW, Itskovitz-Eldor J, Daley GQ: **LIF/STAT3 signaling fails to maintain self-renewal of human embryonic stem cells.** *Stem Cells* 2004, **22**:770–8.

52. Gafni O, Weinberger L, Mansour AA, Manor YS, Chomsky E, Ben-Yosef D, Kalma Y, Viukov S, Maza I, Zviran A, Rais Y, Shipony Z, Mukamel Z, Krupalnik V, Zerbib M, Geula S, Caspi I, Schneir D, Shwartz T, Gilad S, Amann-Zalcenstein D, Benjamin S, Amit I, Tanay A, Massarwa R, Novershtern N, Hanna JH: **Derivation of novel human ground state naive pluripotent stem cells.** *Nature* 2013, **504**:282–286.

53. Ware CB, Nelson AM, Mecham B, Hesson J, Zhou W, Jonlin EC, Jimenez-Caliani AJ, Deng X, Cavanaugh C, Cook S, Tesar PJ, Okada J, Margaretha L, Sperber H, Choi M, Blau CA, Treuting PM, Hawkins RD, Cirulli V, Ruohola-Baker H: **Derivation of naive human embryonic stem cells.** *Proc Natl Acad Sci U S A* 2014, **111**:4484–9.

54. Duggal G, Ghimire S, Broekaert D, Deroo T, Van der Jeught M, Warrier S, Lierman S, Peelman A, Van Soom A, Cornelissen R, Menten B, Mestdagh P, Vandesompele J, Roost M, Slieker RC, Heijmans BT,

Deforce D, De Sutter P, Chuva De Sousa Lopes S, Heindryckx B: **Novel culture conditions inducing naïve pluripotency in human embryonic stem cells**. *Nat Commun - Manuscr Submitt* .

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

56. Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S: **Induction of pluripotent stem cells from adult human fibroblasts by defined factors**. *Cell* 2007, **131**:861–872.

57. Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, Nie J, Jonsdottir GA, Ruotti V, Stewart R, Slukvin II, Thomson JA: **Induced pluripotent stem cell lines derived from human somatic cells**. *Science (80-)* 2007, **318**:1917–1920.

58. Yu J, Chau KF, Vodyanik MA, Jiang J, Jiang Y: **Efficient feeder-free episomal reprogramming with small molecules.** *PLoS One* 2011, **6**:e17557.

59. Zhu S, Li W, Zhou H, Wei W, Ambasudhan R, Lin T, Kim J, Zhang K, Ding S: **Reprogramming of human primary somatic cells by OCT4 and chemical compounds.** *Cell Stem Cell* 2010, **7**:651–5.

60. Drews K, Jozefczuk J, Prigione A, Adjaye J: **Human induced pluripotent stem cells--from mechanisms to clinical applications.** *J Mol Med (Berl)* 2012, **90**:735–45.

61. Imaizumi Y, Okano H: **Modeling human neurological disorders with induced pluripotent stem cells.** *J Neurochem* 2014, **129**:388–99.

62. Mohamet L, Miazga NJ, Ward CM: **Familial Alzheimer's disease modelling using induced pluripotent stem cell technology.** *World J Stem Cells* 2014, **6**:239–247.

63. Fujita J, Fukuda K: **Future Prospects for Regenerated Heart Using Induced Pluripotent Stem Cells.** *J Pharmacol Sci* 2014, **125**:1–5.

64. Holditch SJ, Terzic A, Ikeda Y: **Concise Review: Pluripotent Stem Cell-Based Regenerative Applications for Failing β-Cell Function.** *Stem Cells Transl Med* 2014, **3**:653–61.

65. Isobe K-I, Cheng Z, Nishio N, Suganya T, Tanaka Y, Ito S: **iPSCs, aging and age-related diseases.** *N Biotechnol* 2014, **31**:411–421.

66. Condic ML, Rao M: **Alternative sources of pluripotent stem cells: ethical and scientific issues revisited.** *Stem Cells Dev* 2010, **19**:1121–9.

67. Rossant J: **Reprogramming to pluripotency: from frogs to stem cells**. *Cell* 2009, **138**:1047–1050.

68. Puri MC, Nagy A: **Concise review: Embryonic stem cells versus induced pluripotent stem cells: the game is on.** *Stem Cells* 2012, **30**:10–4.

69. Flemish Training Network Life Sciences: **Workshop: The Epigenetics Revolution: development and disease biology revisited - Gabriella Ficz - Regulation of epigenetic reprogramming by signalling pathways; Ground state pluripotency and beyond**. Leuven, BE; 2013.

70. Waddington CH: *An Introduction to Modern Genetics*. New York: The Macmillan Company; 1939.

71. Waddington CH: **The epigenotype. 1942.** *Int J Epidemiol* 2012, **41**:10–3.

72. Tollervey J, Lunyak V V: **Epigenetics: Judge, jury and executioner of stem cell fate**. *Epigenetics* 2012, **7**:823–40.

73. Wu Ct, Morris JR: **Genes, genetics, and epigenetics: a correspondence.** *Science* 2001, **293**:1103–5.

74. Flemish Training Network Life Sciences: **Workshop: The Epigenetics Revolution: development and disease biology revisited - Wim Vanden Berghe - The seven plagues of nutritional epigenomics**. Leuven, BE; 2013.

75. Kohli RM, Zhang Y: **TET enzymes, TDG and the dynamics of DNA demethylation.** *Nature* 2013, **502**:472–9.

76. Cheng X, Hashimoto H, Horton J, Zhang X: **Mechanisms of DNA methylation, Methyl-CpG recognition and demethylation in mammals**. In *Handb Epigenetics new Mol Med Genet*. Edited by Tollefsbol T. Academic Press; 2011:9–24.

77. Matouk CC, Marsden PA: **Epigenetic Regulation of Vascular Endothelial Gene Expression**. *Circ Res* 2008, **102**:873–887.

78. Collins L, Schönfeld B, Chen XS: **The epigenetics of non-coding RNA**. In *Handb Epigenetics new Mol Med Genet*. Edited by Tollefsbol T. Academic Press; 2011:49–61.

79. Mercer TR, Mattick JS: **Structure and function of long noncoding RNAs in epigenetic regulation.** *Nat Struct Mol Biol* 2013, **20**:300–7.

80. Morris K V, Mattick JS: **The rise of regulatory RNA.** *Nat Rev Genet* 2014, **15**:423–37.

81. Patel DJ, Wang Z: **Readout of epigenetic modifications.** *Annu Rev Biochem* 2013, **82**:81–118.

82. Lessard J, Crabtree GR: **Chromatin Regulatory Mechanisms in Pluripotency**. *Annu Rev Cell Dev Biol* 2010, **26**:503–532.

83. Arents G, Burlingame RW, Wang BC, Love WE, Moudrianakis EN: **The nucleosomal core histone octamer at 3.1 A resolution: a tripartite protein assembly and a left-handed superhelix.** *Proc Natl Acad Sci U S A* 1991, **88**:10148–52.

84. Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ: **Crystal structure of the nucleosome core particle at 2.8 A resolution.** *Nature* 1997, **389**:251–60.

85. Luger K, Richmond TJ: **DNA binding within the nucleosome core.** *Curr Opin Struct Biol* 1998, **8**:33–40.

86. Wolffe AP, Khochbin S, Dimitrov S: **What do linker histones do in chromatin?**. *Bioessays* 1997, **19**:249–55.

87. Moindrot B, Bouvet P, Mongelard F: **Chromatin structure and organization: the relation with gene expression during development and disease.** *Subcell Biochem* 2013, **61**:373–96.

88. Woodcock CL, Dimitrov S: **Higher-order structure of chromatin and chromosomes.** *Curr Opin Genet Dev* 2001, **11**:130–5.

89. Du Preez LL, Patterton H-G: **Secondary structures of the core histone N-terminal tails: their role in regulating chromatin structure.** *Subcell Biochem* 2013, **61**:37–55.

90. **Structural organization of core histones** [http://www.studyblue.com/notes/note/n/chapter-4-dna-chromosomes-and-genomes/deck/7133377]

91. Turner BM: *Chromatin and Gene Regulation: Mechanisms in Epigenetics*. Blackwell Science; 2001.

92. **The Histone octamer assembly & The Nucleosome** [http://www.nobelprize.org/educational/medicine/dna/a/transcription/histoneoctamer_nucleosome.html]

93. Wolffe AP: *Chromatin: Structure and Function*. Third edit. Academic Press; 1998.

94. Talbert PB, Henikoff S: **Histone variants--ancient wrap artists of the epigenome.** *Nat Rev Mol Cell Biol* 2010, **11**:264–75.

95. Kumari S, Swaminathan A, Chatterjee S, Senapati P, Boopathi R, Kundu TK: **Chromatin organization, epigenetics and differentiation: an evolutionary perspective.** *Subcell Biochem* 2013, **61**:3–35.

96. Allan J, Harborne N, Rau DC, Gould H: **Participation of core histone "tails" in the stabilization of the chromatin solenoid.** *J Cell Biol* 1982, **93**:285–97.

97. Ausio J, Dong F, van Holde KE: **Use of selectively trypsinized nucleosome core particles to analyze the role of the histone "tails" in the stabilization of the nucleosome.** *J Mol Biol* 1989, **206**:451–63.

98. Phillips DM: **The presence of acetyl groups in histones.** *Biochem J* 1963, **87**:258–63.

99. Kouzarides T: **Chromatin modifications and their function**. *Cell* 2007, **128**:693–705.

100. Tan M, Luo H, Lee S, Jin F, Yang JS, Montellier E, Buchou T, Cheng Z, Rousseaux S, Rajagopal N, Lu Z, Ye Z, Zhu Q, Wysocka J, Ye Y, Khochbin S, Ren B, Zhao Y: **Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification.** *Cell* 2011, **146**:1016–28.

101. Sakabe K, Wang Z, Hart GW: **Beta-N-acetylglucosamine (O-GlcNAc) is part of the histone code.** *Proc Natl Acad Sci U S A* 2010, **107**:19915–20.

102. Xie Z, Dai J, Dai L, Tan M, Cheng Z, Wu Y, Boeke JD, Zhao Y: **Lysine succinylation and lysine malonylation in histones.** *Mol Cell Proteomics* 2012, **11**:100–7.

103. Bhaumik SR, Smith E, Shilatifard A: **Covalent modifications of histones during development and disease pathogenesis.** *Nat Struct Mol Biol* 2007, **14**:1008–16.

104. Arnaudo AM, Garcia BA: **Proteomic characterization of novel histone post-translational modifications.** *Epigenetics Chromatin* 2013, **6**:24.

105. Karch KR, Denizio JE, Black BE, Garcia BA: **Identification and interrogation of combinatorial histone modifications.** *Front Genet* 2013, **4**:264.

106. Musselman CA, Lalonde M-E, Côté J, Kutateladze TG: **Perceiving the epigenetic landscape through histone readers.** *Nat Struct Mol Biol* 2012, **19**:1218–27.

107. Suganuma T, Workman JL: **Crosstalk among Histone Modifications**. *Cell* 2008, **135**:604–607.

108. Lee BM, Mahadevan LC: **Stability of histone modifications across mammalian genomes: implications for "epigenetic" marking.** *J Cell Biochem* 2009, **108**:22–34.

109. Juliandi B, Abematsu M, Nakashima K: **Epigenetics, Stem Cells and Cellular Differentiation**. In *Handb Epigenetics new Mol Med Genet*. Edited by Tollefsbol T. Academic Press; 2011:315–328.

110. Weber CM, Henikoff S: **Histone variants: dynamic punctuation in transcription.** *Genes Dev* 2014, **28**:672–82.

111. Loyola A, Almouzni G: **Marking histone H3 variants: how, when and why?**. *Trends Biochem Sci* 2007, **32**:425–33.

112. Duncan EM: **PhD thesis: Regulated Histone H3 Proteolysis During Mouse Embryonic Stem Cell Differentiation**. Rockefeller University; 2009.

113. Ahmad K, Henikoff S: **The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly**. *Mol Cell* 2002, **9**:1191–1200.

114. D'Urso A, Brickner JH: **Mechanisms of epigenetic memory.** *Trends Genet* 2014, **30**:230–236.

115. Xu M, Long C, Chen X, Huang C, Chen S, Zhu B: **Partitioning of histone H3-H4 tetramers during DNA replication-dependent chromatin assembly**. *Science (80-)* 2010, **328**:94–98.

116. Greer EL, Shi Y: **Histone methylation: a dynamic mark in health, disease and inheritance.** *Nat Rev Genet* 2012, **13**:343–57.

117. Anand P, Kunnumakkara AB, Kunnumakara AB, Sundaram C, Harikumar KB, Tharakan ST, Lai OS, Sung B, Aggarwal BB: **Cancer is a preventable disease that requires major lifestyle changes.** *Pharm Res* 2008, **25**:2097–116.

118. Waddington CH: *The Strategy of the Genes: A Discussion of Some Aspects of Theoretical Biology*. Allen & Unwin; 1957.

119. Waddington CH: **Canalization of development and the inheritance of acquired characters**. *Nature* 1942, **150**:563–565.

120. Ladewig J, Koch P, Brüstle O: **Leveling Waddington: the emergence of direct programming and the loss of cell fate hierarchies.** *Nat Rev Mol Cell Biol* 2013, **14**:225–36.

121. **Conrad Hal Waddington | Cell Landscapes on WordPress.com** [http://mhairimastersproject.wordpress.com/2011/11/08/conrad-hal-waddington/]

122. Pietersen AM, van Lohuizen M: **Stem cell regulation by polycomb repressors: postponing commitment**. *Curr Opin Cell Biol* 2008, **20**:201–207.

123. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, Jaenisch R, Wagschal A, Feil R, Schreiber SL, Lander ES: **A bivalent chromatin structure marks key developmental genes in embryonic stem cells**. *Cell* 2006, **125**:315–326.

124. Xie W, Song C, Young NL, Sperling AS, Xu F, Sridharan R, Conway AE, Garcia BA, Plath K, Clark AT, Grunstein M: **Histone h3 lysine 56 acetylation is linked to the core transcriptional network in human embryonic stem cells**. *Mol Cell* 2009, **33**:417–427.

125. Slade DJ, Subramanian V, Thompson PR: **Pluripotency: Citrullination unravels stem cells.** *Nat Chem Biol* 2014, **10**:327–8.

126. Bannister AJ, Schneider R, Kouzarides T: **Histone methylation: dynamic or static?**. *Cell* 2002, **109**:801–6.

127. Duncan EM, Muratore-Schroeder TL, Cook RG, Garcia BA, Shabanowitz J, Hunt DF, Allis CD: **Cathepsin L proteolytically processes histone H3 during mouse embryonic stem cell differentiation**. *Cell* 2008, **135**:284–294.

128. Mullis KB, Faloona FA: **Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction.** *Methods Enzymol* 1987, **155**:335–50.

129. Biogazelle: **Course on qPCR-experiment design and data analysis**. Gent, BE; 2010.

130. Arya M, Shergill IS, Williamson M, Gommersall L, Arya N, Patel HRH: **Basic principles of real-time quantitative PCR.** *Expert Rev Mol Diagn* 2005, **5**:209–19.

131. **Eurogentec qPCR-booklet** [http://www.eurogentec.com/uploads/FileBrowse/Technical Guides/Genomic Solutions/qPCR-booklet.pdf]

132. **NCBI PrimerBLAST - Primer designing tool** [http://www.ncbi.nlm.nih.gov/tools/primer-blast/]

133. Lefever S, Vandesompele J, Speleman F, Pattyn F: **RTPrimerDB: the portal for real-time PCR primers and probes**. *Nucleic Acids Res* 2009, **37**(Database issue):D942–5.

134. **Primer3 Input (version 0.4.0)** [http://bioinfo.ut.ee/primer3-0.4.0/]

135. Higuchi R, Dollinger G, Walsh PS, Griffith R: **Simultaneous amplification and detection of specific DNA sequences.** *Biotechnology (N Y)* 1992, **10**:413–7.

136. Kubista M, Andrade JM, Bengtsson M, Forootan A, Jonák J, Lind K, Sindelka R, Sjöback R, Sjögreen B, Strömbom L, Ståhlberg A, Zoric N: **The real-time polymerase chain reaction.** *Mol Aspects Med* 2006, **27**:95–125.

137. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F: **Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes**. *Genome Biol* 2002, **3**:Research0034.1–0034.11.

138. Pfaffl MW, Tichopad A, Prgomet C, Neuvians TP: **Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper--Excel-based tool using pair-wise correlations.** *Biotechnol Lett* 2004, **26**:509–15.

139. Andersen CL, Jensen JL, Ørntoft TF: **Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets.** *Cancer Res* 2004, **64**:5245–50.

140. Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J: **qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data**. *Genome Biol* 2007, **8**:R19.

141. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT: **The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments**. *Clin Chem* 2009, **55**:611–622.

142. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT: **Primer Sequence Disclosure: A Clarification of the MIQE Guidelines**. *Clin Chem* 2011, **57**:919–921.

143. Bustin SA, Benes V, Garson J, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley G, Wittwer CT, Schjerling P, Day PJ, Abreu M, Aguado B, Beaulieu J-F, Beckers A, Bogaert S, Browne JA, Carrasco-Ramiro F, Ceelen L, Ciborowski K, Cornillie P, Coulon S, Cuypers A, De Brouwer S, De Ceuninck L, De Craene J, De Naeyer H, De Spiegelaere W, et al.: **The need for transparency and good practices in the qPCR literature.** *Nat Methods* 2013, **10**:1063–7.

144. Vogelstein B, Kinzler KW: **Digital PCR.** *Proc Natl Acad Sci U S A* 1999, **96**:9236–41.

145. Huggett JF, Foy CA, Benes V, Emslie K, Garson JA, Haynes R, Hellemans J, Kubista M, Mueller RD, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT, Bustin SA: **The digital MIQE guidelines: Minimum Information for Publication of Quantitative Digital PCR Experiments.** *Clin Chem* 2013, **59**:892–902.

146. Amit M, Carpenter MK, Inokuma MS, Chiu CP, Harris CP, Waknitz MA, Itskovitz-Eldor J, Thomson JA: **Clonally derived human embryonic stem cell lines maintain pluripotency and proliferative potential for prolonged periods of culture**. *Dev Biol* 2000, **227**:271–278.

147. Bhattacharya B, Miura T, Brandenberger R, Mejido J, Luo Y, Yang AX, Joshi BH, Ginis I, Thies RS, Amit M, Lyons I, Condie BG, Itskovitz-Eldor J, Rao MS, Puri RK: **Gene expression in human embryonic stem cell lines: unique molecular signature**. *Blood* 2004, **103**:2956–2964.

148. Mountford JC: **Human embryonic stem cells: origins, characteristics and potential for regenerative therapy**. *Transfus Med* 2008, **18**:1–12.

149. Giadrossi S, Dvorkina M, Fisher AG: **Chromatin organization and differentiation in embryonic stem cell models**. *Curr Opin Genet Dev* 2007, **17**:132–138.

150. Zhang Z: **Chapter 6 - Neural Differentiation**. In *Hum Cell Cult Vol VI Embryonic Stem Cells*. Edited by Masters JR. Springer Netherlands; 2007:85–108.

151. Willems E, Leyns L, Vandesompele J: **Standardization of real-time PCR gene expression data from independent biological replicates**. *Anal Biochem* 2008, **379**:127–129.

152. Marullo M, Zuccato C, Mariotti C, Lahiri N, Tabrizi SJ, Di Donato S, Cattaneo E: **Expressed Alu repeats as a novel, reliable tool for normalization of real-time quantitative RT-qPCR data**. *Genome Biol* 2010, **11**:R9.

153. Huggett J, Dheda K, Bustin S, Zumla A: **Real-time RT-PCR normalisation; strategies and considerations**. *Genes Immun* 2005, **6**:279–284.

154. Solanas M, Moral R, Escrich E: **Unsuitability of using ribosomal RNA as loading control for Northern blot analyses related to the imbalance between messenger and ribosomal RNA content in rat mammary tumors**. *Anal Biochem* 2001, **288**:99–102.

155. Willems E, Mateizel I, Kemp C, Cauffman G, Sermon K, Leyns L: **Selection of reference genes in mouse embryos and in differentiating human and mouse ES cells**. *Int J Dev Biol* 2006, **50**:627–635.

156. Synnergren J, Giesler TL, Adak S, Tandon R, Noaksson K, Lindahl A, Nilsson P, Nelson D, Olsson B, Englund MC, Abbot S, Sartipy P: **Differentiating human embryonic stem cells express a unique housekeeping gene signature**. *Stem Cells* 2007, **25**:473–480.

157. Batzer MA, Deininger PL: **Alu repeats and human genomic diversity**. *Nat Rev Genet* 2002, **3**:370–379.

158. Rowold DJ, Herrera RJ: **Alu elements and the human genome**. *Genetica* 2000, **108**:57–72.

159. Shaikh TH, Roy AM, Kim J, Batzer MA, Deininger PL: **cDNAs derived from primary and small cytoplasmic Alu (scAlu) transcripts**. *J Mol Biol* 1997, **271**:222–234.

160. Rihani A, Van Maerken T, Pattyn F, Van Peer G, Beckers A, De Brouwer S, Kumps C, Mets E, Van der Meulen J, Rondou P, Leonelli C, Mestdagh P, Speleman F, Vandesompele J: **Effective Alu repeat based RT-Qpcr normalization in cancer cell perturbation experiments.** *PLoS One* 2013, **8**:e71776.

161. O'Leary T, Heindryckx B, Lierman S, Van der Jeught M, Menten B, Deforce D, Cornelissen R, de Sousa Lopes SC, De Sutter P: **The influence of early embryo traits on human embryonic stem cell derivation efficiency**. *Stem Cells Dev* 2011, **20**:785–793.

162. Lupberger J, Kreuzer KA, Baskaynak G, Peters UR, le Coutre P, Schmidt CA: **Quantitative analysis of beta-actin, beta-2-microglobulin and porphobilinogen deaminase mRNA and their comparison as control transcripts for RT-PCR**. *Mol Cell Probes* 2002, **16**:25–30.

163. Piehler AP, Grimholt RM, Ovstebø R, Berg JP: **Gene expression results in lipopolysaccharide-stimulated monocytes depend significantly on the choice of reference genes.** *BMC Immunol* 2010, **11**:21.

164. Valceckiene V, Kontenyte R, Jakubauskas A, Griskevicius L: **Selection of reference genes for quantitative polymerase chain reaction studies in purified B cells from B cell chronic lymphocytic leukaemia patients.** *Br J Haematol* 2010, **151**:232–8.

165. Quiroz FG, Posada OM, Gallego-Perez D, Higuita-Castro N, Sarassa C, Hansford DJ, Agudelo-Florez P, Lopez LE: **Housekeeping gene stability influences the quantification of osteogenic**

markers during stem cell differentiation to the osteogenic lineage. *Cytotechnology* 2010, **62**:109–120.

166. Jesnowski R, Backhaus C, Ringel J, Lohr M: **Ribosomal highly basic 23-kDa protein as a reliable standard for gene expression analysis**. *Pancreatology* 2002, **2**:421–424.

167. Curtis KM, Gomez L a, Rios C, Garbayo E, Raval AP, Perez-Pinzon M a, Schiller PC: **EF1alpha and RPL13a represent normalization genes suitable for RT-qPCR analysis of bone marrow derived mesenchymal stem cells.** *BMC Mol Biol* 2010, **11**:61.

168. Wang Y, Han Z, Yan S, Mao A, Wang B, Ren H, Chi Y: **Evaluation of suitable reference gene for real-time PCR in human umbilical cord mesenchymal stem cells with long-term in vitro expansion**. *Vitr Cell Dev Biol Anim* 2010, **46**:595–599.

169. Kenmochi N, Kawaguchi T, Rozen S, Davis E, Goodman N, Hudson TJ, Tanaka T, Page DC: **A map of 75 human ribosomal protein genes.** *Genome Res* 1998, **8**:509–23.

170. Vermeulen J, De Preter K, Lefever S, Nuytens J, De Vloed F, Derveaux S, Hellemans J, Speleman F, Vandesompele J: **Measurable impact of RNA quality on gene expression results from quantitative PCR**. *Nucleic Acids Res* 2011, **39**:e63.

171. Bernstein E, Hake SB: **The nucleosome: a little variation goes a long way**. *Biochem Cell Biol* 2006, **84**:505–517.

172. Skene PJ, Henikoff S: **Histone variants in pluripotency and disease.** *Development* 2013, **140**:2513–24.

173. Duncan EM, Allis CD: **Errors in erasure: links between histone lysine methylation removal and disease.** *Prog Drug Res* 2011, **67**:69–90.

174. Adams-Cioaba MA, Krupa JC, Xu C, Mort JS, Min J: **Structural basis for the recognition and cleavage of histone H3 by cathepsin L**. *Nat Commun* 2011, **2**:197.

175. Tesar M, Marquardt O: **Foot-and-mouth disease virus protease 3C inhibits cellular transcription and mediates cleavage of histone H3**. *Virology* 1990, **174**:364–374.

176. Falk MM, Grigera PR, Bergmann IE, Zibert A, Multhaup G, Beck E: **Foot-and-mouth disease virus protease 3C induces specific proteolytic cleavage of host cell histone H3**. *J Virol* 1990, **64**:748–756.

177. Gonzalo S: **Epigenetic alterations in aging**. *J Appl Physiol* 2010, **109**:586–597.

178. Mahendra G, Kanungo MS: **Age-related and steroid induced changes in the histones of the quail liver.** *Arch Gerontol Geriatr* 2000, **30**:109–14.

179. Santos-Rosa H, Kirmizis A, Nelson C, Bartke T, Saksouk N, Cote J, Kouzarides T: **Histone H3 tail clipping regulates gene expression**. *Nat Struct Mol Biol* 2009, **16**:17–22.

180. Allis CD, Bowen JK, Abraham GN, Glover C V, Gorovsky MA: **Proteolytic processing of histone H3 in chromatin: a physiologically regulated event in Tetrahymena micronuclei**. *Cell* 1980, **20**:55–64.

181. Allis CD, Wiggins JC: **Proteolytic processing of micronuclear H3 and histone phosphorylation during conjugation in Tetrahymena thermophila**. *Exp Cell Res* 1984, **153**:287–298.

182. Mandal P, Azad GK, Tomar RS: **Identification of a novel histone H3 specific protease activity in nuclei of chicken liver.** *Biochem Biophys Res Commun* 2012, **421**:261–7.

183. Mandal P, Verma N, Chauhan S, Tomar RS: **Unexpected Histone H3 Tail-clipping Activity of Glutamate Dehydrogenase.** *J Biol Chem* 2013, **288**:18743–57.

184. Zwaka TP, Thomson JA: **Homologous recombination in human embryonic stem cells.** *Nat Biotechnol* 2003, **21**:319–21.

185. Zhang C, Liu Y, Andrews PC: **Quantification of histone modifications using ^{15}N metabolic labeling.** *Methods* 2013, **61**:236–43.

186. Glibert P, Van Steendam K, Dhaenens M, Deforce D: **iTRAQ as a method for optimization: enhancing peptide recovery after gel fractionation**. *Proteomics* 2014, **14**:680–684.

187. Garcia BA, Mollah S, Ueberheide BM, Busby SA, Muratore TL, Shabanowitz J, Hunt DF: **Chemical derivatization of histones for facilitated analysis by mass spectrometry.** *Nat Protoc* 2007, **2**:933–8.

188. Johns EW: **Studies on histones. 7. Preparative methods for histone fractions from calf thymus.** *Biochem J* 1964, **92**:55–9.

189. Phillips DM, Johns EW: **A study of the proteinase content and the chromatography of thymus histones.** *Biochem J* 1959, **72**:538–44.

190. Azad GK, Tomar RS: **Proteolytic clipping of histone tails: the emerging role of histone proteases in regulation of various biological processes.** *Mol Biol Rep* 2014, **41**:2717–2730.

191. Xue Y, Vashisht AA, Tan Y, Su T, Wohlschlegel JA: **PRB1 Is Required for Clipping of the Histone H3 N Terminal Tail in Saccharomyces cerevisiae.** *PLoS One* 2014, **9**:e90496.

192. Minami J, Takada K, Aoki K, Shimada Y, Okawa Y, Usui N, Ohkawa K: **Purification and characterization of C-terminal truncated forms of histone H2A in monocytic THP-1 cells**. *Int J Biochem Cell Biol* 2007, **39**:171–180.

193. Okawa Y, Takada K, Minami J, Aoki K, Shibayama H, Ohkawa K: **Purification of N-terminally truncated histone H2A-monoubiquitin conjugates from leukemic cell nuclei: probable proteolytic products of ubiquitinated H2A**. *Int J Biochem Cell Biol* 2003, **35**:1588–1600.

194. Glibert P, Vossaert L, Van Steendam K, Lambrecht S, Van Nieuwerburgh F, Offner F, Kipps T, Dhaenens M, Deforce D: **Quantitative Proteomics to Characterize Specific Histone H2A Proteolysis in Chronic Lymphocytic Leukemia and the Myeloid THP-1 Cell Line.** *Int J Mol Sci* 2014, **15**:9407–21.

195. Dhaenens M, Glibert P, Lambrecht S, Vossaert L, Van Steendam K, Elewaut D, Deforce D: **Neutrophil Elastase in the capacity of the "H2A-specific protease."** *Int J Biochem Cell Biol* 2014.

196. Dhaenens M, Glibert P, Meert P, Vossaert L, Deforce D: **Histone proteolysis: A proposal for categorization into "clipping" and "degradation."** *Bioessays* 2014, **36**.

197. Bartley J, Chalkley R: **Further studies of a thymus nucleohistone-associated protease**. *J Biol Chem* 1970, **245**:4286–4292.

198. Chong MT, Garrard WT, Bonner J: **Purification and properties of a neutral protease from rat liver chromatin.** *Biochemistry* 1974, **13**:5128–34.

199. Dumuis-Kervabon A, Encontre I, Etienne G, Jauregui-Adell J, Méry J, Mesnier D, Parello J: **A chromatin core particle obtained by selective cleavage of histones by clostripain.** *EMBO J* 1986, **5**:1735–42.

200. Böhm L, Briand G, Sautière P, Crane-Robinson C: **Proteolytic digestion studies of chromatin core-histone structure. Identification of the limit peptides of histones H3 and H4.** *Eur J Biochem* 1981, **119**:67–74.

201. Ivanov A, Pawlikowski J, Manoharan I, van Tuyn J, Nelson DM, Rai TS, Shah PP, Hewitt G, Korolchuk VI, Passos JF, Wu H, Berger SL, Adams PD: **Lysosome-mediated processing of chromatin in senescence.** *J Cell Biol* 2013, **202**:129–43.

202. Qian M-X, Pang Y, Liu CH, Haratake K, Du B-Y, Ji D-Y, Wang G-F, Zhu Q-Q, Song W, Yu Y, Zhang X-X, Huang H-T, Miao S, Chen L-B, Zhang Z-H, Liang Y-N, Liu S, Cha H, Yang D, Zhai Y, Komatsu T, Tsuruta F, Li H, Cao C, Li W, Li G-H, Cheng Y, Chiba T, Wang L, Goldberg AL, et al.: **Acetylation-mediated proteasomal degradation of core histones during DNA repair and spermatogenesis.** *Cell* 2013, **153**:1012–24.

203. Hammoud SS, Nix DA, Zhang H, Purwar J, Carrell DT, Cairns BR: **Distinctive chromatin in human sperm packages genes for embryo development.** *Nature* 2009, **460**:473–8.

204. Puchi M, Garcia-Huidobro J, Cordova C, Aguilar R, Dufey E, Imschenetzky M, Bustos P, Morin V: **A new nuclear protease with cathepsin L properties is present in Hela and Caco-2 cells**. *J Cell Biochem* 2010.

205. Papayannopoulos V, Metzler KD, Hakkim A, Zychlinsky A: **Neutrophil elastase and myeloperoxidase regulate the formation of neutrophil extracellular traps.** *J Cell Biol* 2010, **191**:677–91.

206. Zhang D, Pasternack MS, Beresford PJ, Wagner L, Greenberg AH, Lieberman J: **Induction of rapid histone degradation by the cytotoxic T lymphocyte protease Granzyme A.** *J Biol Chem* 2001, **276**:3683–90.

207. Kayne PS, Kim UJ, Han M, Mullen JR, Yoshizaki F, Grunstein M: **Extremely conserved histone H4 N terminus is dispensable for growth but essential for repressing the silent mating loci in yeast.** *Cell* 1988, **55**:27–39.

208. Wallis JW, Rykowski M, Grunstein M: **Yeast histone H2B containing large amino terminus deletions can function in vivo.** *Cell* 1983, **35**(3 Pt 2):711–9.

209. Durrin LK, Mann RK, Kayne PS, Grunstein M: **Yeast histone H4 N-terminal sequence is required for promoter activation in vivo.** *Cell* 1991, **65**:1023–31.

210. Schuster T, Han M, Grunstein M: **Yeast histone H2A and H2B amino termini have interchangeable functions.** *Cell* 1986, **45**:445–51.

211. Karaczyn AA, Cheng RYS, Buzard GS, Hartley J, Esposito D, Kasprzak KS: **Truncation of histone H2A's C-terminal tail, as is typical for Ni(II)-assisted specific peptide bond hydrolysis, has gene expression altering effects.** *Ann Clin Lab Sci* 2009, **39**:251–62.

212. Vogler C, Huber C, Waldmann T, Ettig R, Braun L, Izzo A, Daujat S, Chassignet I, Lopez-Contreras AJ, Fernandez-Capetillo O, Dundr M, Rippe K, Längst G, Schneider R: **Histone H2A C-terminus regulates chromatin dynamics, remodeling, and histone H1 binding.** *PLoS Genet* 2010, **6**:e1001234.

213. Eickbush TH, Watson DK, Moudrianakis EN: **A chromatin-bound proteolytic activity with unique specificity for histone H2A**. *Cell* 1976, **9**(4 PT 2):785–792.

214. Watson DK, Moudrianakis EN: **Histone-dependent reconstitution and nucleosomal localization of a nonhistone chromosomal protein: the H2A-specific protease**. *Biochemistry* 1982, **21**:248–256.

215. Davie JR, Numerow L, Delcuve GP: **The nonhistone chromosomal protein, H2A-specific protease, is selectively associated with nucleosomes containing histone H1**. *J Biol Chem* 1986, **261**:10410–10416.

216. Elia MC, Moudrianakis EN: **Regulation of H2a-specific proteolysis by the histone H3:H4 tetramer**. *J Biol Chem* 1988, **263**:9958–9964.

217. Eickbush TH, Godfrey JE, Elia MC, Moudrianakis EN: **H2a-specific proteolysis as a unique probe in the analysis of the histone octamer**. *J Biol Chem* 1988, **263**:18972–18978.

218. Pantazis P, Sarin PS, Gallo RC: **Detection of the histone-2A related polypeptide in differentiated human myeloid cells (HL-60) and its distribution in human acute leukemia.** *Int J Cancer* 1981, **27**:585–92.

219. Simpkins H, Mahon K: **The histone content of chromatin preparations from leukaemic cells.** *Br J Haematol* 1977, **37**:467–73.

220. Kim HS, Yoon H, Minn I, Park CB, Lee WT, Zasloff M, Kim SC: **Pepsin-mediated processing of the cytoplasmic histone H2A to strong antimicrobial peptide buforin I.** *J Immunol* 2000, **165**:3268–74.

221. Mishra RN, Kanungo MS: **Alterations in histones of the liver and oviduct of Japanese quail during aging**. *Mol Biol Rep* 1994, **20**:15–18.

222. Haugen MH, Johansen HT, Pettersen SJ, Solberg R, Brix K, Flatmark K, Maelandsmo GM: **Nuclear legumain activity in colorectal cancer.** *PLoS One* 2013, **8**:e52980.

223. Arnandis T, Ferrer-Vicens I, Torres L, García C, Garcia-Trevijano ER, Zaragoza R, Viña JR: **Differential functions of calpain 1 during epithelial cell death and adipocyte differentiation in mammary gland involution.** *Biochem J* 2014, **459**:355–68.

224. Khalkhali-Ellis Z, Goossens W, Margaryan N V, Hendrix MJC: **Cleavage of histone 3 by cathepsin d in the involuting mammary gland.** *PLoS One* 2014, **9**:e103230.

225. Dado D, Sagi M, Levenberg S, Zemel A: **Mechanical control of stem cell differentiation.** *Regen Med* 2012, **7**:101–16.

226. Goulet B, Baruch A, Moon NS, Poirier M, Sansregret LL, Erickson A, Bogyo M, Nepveu A: **A cathepsin L isoform that is devoid of a signal peptide localizes to the nucleus in S phase and processes the CDP/Cux transcription factor**. *Mol Cell* 2004, **14**:207–219.

227. Ceru S, Konjar S, Maher K, Repnik U, Krizaj I, Bencina M, Renko M, Nepveu A, Zerovnik E, Turk B, Kopitar-Jerala N: **Stefin B interacts with histones and cathepsin L in the nucleus**. *J Biol Chem* 2010, **285**:10078–10086.

228. Tedelind S, Poliakova K, Valeta A, Hunegnaw R, Yemanaberhan EL, Heldin N-E, Kurebayashi J, Weber E, Kopitar-Jerala N, Turk B, Bogyo M, Brix K: **Nuclear cysteine cathepsin variants in thyroid carcinoma cells.** *Biol Chem* 2010, **391**:923–35.

229. Psathas JN, Zheng S, Tan S, Reese JC: **Set2-dependent K36 methylation is regulated by novel intratail interactions within H3**. *Mol Cell Biol* 2009, **29**:6413–6426.

230. Pauli A, van Bemmel JG, Oliveira RA, Itoh T, Shirahige K, van Steensel B, Nasmyth K: **A direct role for cohesin in gene regulation and ecdysone response in Drosophila salivary glands**. *Curr Biol* 2010, **20**:1787–1798.

231. Hsin JP, Sheth A, Manley JL: **RNAP II CTD phosphorylated on threonine-4 is required for histone mRNA 3′ end processing**. *Science (80-)* 2011, **334**:683–686.

232. Brookes E, Pombo A: **Modifications of RNA polymerase II are pivotal in regulating gene expression states**. *EMBO Rep* 2009, **10**:1213–1219.

233. Williams LA, Davis-Dusenbery BN, Eggan KC: **SnapShot: directed differentiation of pluripotent stem cells.** *Cell* 2012, **149**:1174–1174.e1.

234. Herceg Z, Murr R: **Mechanisms of Histone modifications**. In *Handb Epigenetics new Mol Med Genet*. Edited by Tollefsbol T. Academic Press; 2011:25–45.

235. Cuthbert GL, Daujat S, Snowden AW, Erdjument-Bromage H, Hagiwara T, Yamada M, Schneider R, Gregory PD, Tempst P, Bannister AJ, Kouzarides T: **Histone deimination antagonizes arginine methylation.** *Cell* 2004, **118**:545–53.

236. Nelson CJ, Santos-Rosa H, Kouzarides T: **Proline isomerization of histone H3 regulates lysine methylation and gene expression.** *Cell* 2006, **126**:905–16.

237. Hanover JA: **Epigenetics gets sweeter: O-GlcNAc joins the "histone code".** *Chem Biol* 2010, **17**:1272–4.

238. Chen Y, Sprung R, Tang Y, Ball H, Sangras B, Kim SC, Falck JR, Peng J, Gu W, Zhao Y: **Lysine propionylation and butyrylation are novel post-translational modifications in histones.** *Mol Cell Proteomics* 2007, **6**:812–9.

239. Liu B, Lin Y, Darwanto A, Song X, Xu G, Zhang K: **Identification and characterization of propionylation at histone H3 lysine 23 in mammalian cells.** *J Biol Chem* 2009, **284**:32288–95.

240. Filenko NA, Kolar C, West JT, Smith SA, Hassan YI, Borgstahl GEO, Zempleni J, Lyubchenko YL: **The role of histone H4 biotinylation in the structure of nucleosomes.** *PLoS One* 2011, **6**:e16299.

241. Kothapalli N, Camporeale G, Kueh A, Chew YC, Oommen AM, Griffin JB, Zempleni J: **Biological functions of biotinylated histones.** *J Nutr Biochem* 2005, **16**:446–8.

242. Wisniewski JR, Zougman A, Mann M: **Nepsilon-formylation of lysine is a widespread post-translational modification of nuclear proteins occurring at residues involved in regulation of chromatin function.** *Nucleic Acids Res* 2008, **36**:570–7.

CURRICULUM VITAE

LIESBETH VOSSAERT

Address: Parklaan 65
B-9000 Gent
Place and date of birth: Gent, 18/03/1987
Professional address: Laboratory for Pharmaceutical Biotechnology
Faculty of Pharmaceutical Sciences
Ottergemsesteenweg 460, B-9000 Gent
+32 9 264 83 56
Liesbeth.Vossaert@UGent.be

Education

2009 - 2014: PhD Student
Laboratory for Pharmaceutical Biotechnology
Ghent University

Dissertation: "Histone H3 clipping in human embryonic stem cells: in pursuit of an epigenetic outcast"

2007 - 2009: Master of Pharmaceutical Care
Ghent University
Distinction

Dissertation: "Effect of biosurfactants on adhesion of clinical *Staphylococcus epidermidis* biofilm producing strains" - Università degli Studi del Piemonte Orientale, Novara, Italy

2004 - 2007: Bachelor Pharmaceutical Sciences
Ghent University

1998 - 2004: Latin-Greek
St. Franciscusinstituut, Sint-Maria-Oudenhove

Scientific curriculum

Publications

- Histone proteolysis: A proposal for categorization into 'Clipping' and 'Degradation'.
  Dhaenens M., Glibert P., Meert P., Vossaert L. & Deforce D. *BioEssays* 2014, Epub ahead of print.
  (*BioEssays* IF 2013: 4.927 – JCR rank 11/85 (Q1) category "Biology")

- Detailed method description for non-invasive monitoring of differentiation status of human embryonic stem cells.

  Scheerlinck E., Van Steendam K., Vandewoestyne M., Lepez T., Gobin V., Meert P., Vossaert L., Van Nieuwerburgh F., Van Soom A., Peelman L., Heindryckx B., De Sutter P., Dhaenens M. & Deforce D. *Analytical Biochemistry* 2014, **461**: 60-66.

  (*Analytical Biochemistry* IF 2013: 2.305 – JCR rank 32/76 (Q2) category "Analytical Chemistry")

- Quantitative proteomics to characterize specific histone H2A proteolysis in chronic lymphocytic leukemia and the myeloid cell line.

  Glibert P., Vossaert L., Van Steendam K., Lambrecht S., Van Nieuwerburgh F., Offner F., Kipps T., Dhaenens M. & Deforce D. *International Journal of Molecular Sciences – Special issue: mass spectrometry application in biology* 2014, **15**: 9407-21.

  (*International Journal of Molecular Sciences* IF 2013: 2.339 – JCR rank 52/148 (Q2) category "Multidisciplinary Chemistry")

- Identification of histone H3 clipping activity in human embryonic stem cells.

  Vossaert L.*, Meert P.*, Scheerlinck E., Glibert P., Van Roy N., Heindryckx B., De Sutter P., Dhaenens M. & Deforce D. *Stem Cell Research* 2014, 13: 123-134.

  (*Stem Cell Research* IF 2013: 3.912 – JCR rank 31/165 (Q1) category "Biotechnology & Applied Microbiology")

- Neutrophil elastase in the capacity of the "H2A-specific protease".

  Dhaenens M., Glibert P., Lambrecht S., Vossaert L., Van Steendam K., Elewaut D. & Deforce D. *International Journal of Biochemistry and Cell Biology* 2014, 51: 39-44.

  (*International Journal of Biochemistry and Cell Biology* IF 2013: 4.240 – JCR rank 75/291 (Q2) category "Biochemistry & Molecular Biology")

- Reference loci for RT-qPCR analysis of differentiating human embryonic stem cells.

  Vossaert L., O'Leary T., Van Neste C., Heindryckx B., Vandesompele J., De Sutter P. & Deforce D. *BMC Molecular Biology* 2013, 14: 21.

  (*BMC Molecular Biology* IF 2012: 2.976 – JCR rank 147/290 (Q3) category "Biochemistry and Molecular Biology")

- Influence of Activin A supplementation during human embryonic stem cell derivation on germ cell differentiation potential.

  Duggal G., Heindryckx B., Warrier S., O'Leary T., Van der Jeught M., Lierman S., Vossaert L., Deroo T., Deforce D., Chuva de Sousa Lopes S. M. & De Sutter P. *Stem Cells and Development* 2013, 22: 3141-55.

  (*Stem Cells and Development* IF 2012: 4.670 – JCR rank 3/26 (Q1) category "Transplantation")

Poster presentations on national and international conferences

- Histone H3 clipping in hESC in relation to Oct4 expression and culture conditions

  Meert P.*, Vossaert L.*, Scheerlinck E., Glibert P., Van Roy N., Heindryckx B., De Sutter P., Dhaenens M. & Deforce D. First meeting of Belgian Society for Stem Cell Research, Gent, Belgium, September 12, 2014

- Histone H3-protease activity in human embryonic stem cells *(presenting author)*
  Vossaert L.*, Meert P.*, Dhaenens M. & Deforce D. *"Knowledge for Growth" Conference*, FlandersBio, Gent, Belgium, May 8, 2014

- Sensitive quantification of proteolytic products by use of AQUA peptides *(presenting author)*
  Glibert P., Vossaert L., Dhaenens M. & Deforce D. *"ProteoMMX3 Strictly quantitative" Conference*, Biochemical Society, Chester, UK, March 25-27, 2014

- Regulation of inositol 1,4,5-triphosphate receptor during human oocyte maturation and fertilization
  Nikiforaki D., Vossaert L., Vanden Meerschaut F., Qian C., Lu Y., Parys J., De Vos W. H., Leybaert L., Deforce D., Deroo T., Van den Abbeel E., Heindryckx B. & De Sutter P. *ESHRE 29th annual meeting*, London, UK, July 7-10, 2013

- Histone H3 clipping upon early human embryonic stem cell differentiation *(presenting author)*
  Vossaert L.*, Meert P.*, Dhaenens M. & Deforce D. *"Knowledge for Growth" Conference*, FlandersBio, Gent, Belgium, May 30, 2013

- Mass Spectrometry top-down characterization of histone H3 on the MS level
  Meert P.*, Vossaert L.*, Dhaenens M. & Deforce D. *"Knowledge for Growth" Conference*, FlandersBio, Gent, Belgium, May 30, 2013

- Reference genes for RT-qPCR analysis of differentiating human embryonic stem cells *(presenting author)*
  Vossaert L., O'Leary T., Heindryckx B., Van Neste C., Vandewoestyne M., Vandesompele J., De Sutter P. & Deforce D. *"Epigenetics and Stem Cells" Conference*, Abcam, Cambridge, UK, October 16-17, 2012

- Reference genes for RT-qPCR analysis of differentiating human embryonic stem cells *(presenting author)*
  Vossaert L., O'Leary T., Heindryckx B., Van Neste C., Vandewoestyne M., Vandesompele J., De Sutter P. & Deforce D. *Scientific afternoon FFW*, Gent, Belgium, May 16, 2012

Courses and conference attendances

- First meeting of Belgian Society for Stem Cell Research (BeSSCR), Gent, Belgium, September 12, 2014
- "Quantitative Proteomics and Data analysis" Course, Biochemical Society, Chester, UK, March 24-25, 2014
- Assistenten practicumtraining UGent, Gent, Belgium, February 18, 2014
- Protein Purification and Analysis Seminar BioRad, Grimbergen, Belgium, October 17, 2013
- "The Epigenetics Revolution: development and disease biology revisited" Conference, VIB, Leuven, Belgium, July 1-4, 2013
- Basisassistententraining UGent, Gent, Belgium, June 18, 2013
- Droplet Digital PCR Road show BioRad, Grimbergen, Belgium, June 13, 2013
- BePa Proteomics day, Gent, Belgium, November 29-30 2012
- Quantitative PCR Seminar Tour BioRad, Grimbergen, Belgium, October 26, 2011
- Stem cell seminar Life Technologies, VUB, Jette, Belgium, September 22, 2011
- Scientific afternoon FFW, Gent, Belgium, May 26, 2011

- "Academic English writing skills", Ghent University, Gent Belgium, February – May 2011
- KVCV Flanders Proteomics, Antwerpen, Belgium, December 16-17, 2010
- "qPCR experiment design and data-analysis" Course, Biogazelle, Gent, Belgium, November 8-9, 2010
- "Epigenetics and Stem Cells" Conference, Abcam, Copenhagen, Denmark, August 25-27, 2010
- Protein purification and analysis seminar GE Healthcare, Liège, Belgium, June 16, 2010
- Scientific afternoon FFW, Gent, Belgium, May 12, 2010
- "Introduction to Bioinformatics" Course, VIB BITS, Leuven, Belgium, May 4-5, 2010
- Stem Cell Workshop BD, Erembodegem, Belgium, January 14, 2010

Educational experience

- *Practical courses*
    - Practical course Phytochemistry (2nd Bachelor students)
    - Practical course Pharmaceutical Biotechnology (3rd Bachelor students)
    - Lab-instructor Pharmaceutical Bachelor Proof (3rd Bachelor students)
- *Supervising thesis dissertations*
    - Sarah Dewulf (Master of Science in Industrial Pharmacy – UGent, 2013/2014): "Onderzoek naar het enzym verantwoordelijk voor histonproteolyse in hESC"
    - Ine Coene (1st Master of Pharmaceutical care – UGent, 2012/2013): "Histonklieving H3 tijdens Humane stamceldifferentiatie"
    - Maarten Batens (Master of Science in Industrial Pharmacy – UGent, 2012/2013): "De klieving van histon H3 bij humane embryonale stamcellen"
    - Erika D'haenens (Master in Industrial Sciences: Biochemistry – HoGent, 2012/2013): "N-glycaananalyse van human serum en rhEPO met een DNA-sequencer"
    - Stefi Premereur (1st Master of Pharmaceutical care – UGent, 2011/2012): "Humane embryonale stamcellen: Histonklieving binnen stamceldifferentiatie"
    - Paulien Meert (1st Master of Pharmaceutical care – UGent, 2009/2010): "Histonklieving als posttranslationele modificatie tijdens stamceldifferentiatie"

DANKWOORD

En ten slotte, het thesis-hoofdstuk met de hoogste impact factor, het gepaste bedankingswerk!

Beste Dieter, een origineel dankwoord schrijven wordt er moeilijker en moeilijker op, gezien de hele reeks doctoraatsstudenten die je hier al hebt zien passeren – maar ik ben blij daar één van te zijn. Ik wil je bedanken voor alle kansen die ik tot nu toe op het labo gekregen heb én voor de gelegenheid om hier nog een tijdje te blijven. Je wist mij als promotor steeds bij te sturen waar nodig met je inzichten en vindingrijkheid, vaak met hoog doe-het-zelf gehalte. Maar ook de teambuildings, BBQ's, jaarlijkse personeelsdag-cadeautjes, enz. naast het werk worden erg geapprecieerd!

Vanzelfsprekend wil ik ook de andere leden van de lees- en examencommissie bedanken voor de tijd die ze genomen hebben om dit werk te lezen en te beoordelen. Een speciaal woord van dank aan mijn co-promotor prof. Petra De Sutter, om ook de deuren van haar labo open te zetten voor mijn onderzoek. Bedankt prof. Björn Heindryckx, Tom, Margot, Galbha, Sharat, Jitesh, Sabitri, Dimitra & Jasin om hierbij te helpen.

Bedankt prof. Jo Vandesompele, voor alle hulp bij het (uiteindelijk!) publiceren van de geNorm-paper. Nurten, bedankt voor de praktische bijstand voor alle PCR's die bij jullie gerund zijn.

En zo kom ik ook terecht bij mijn andere co-promotor, Maarten. Jouw (wetenschappelijk) enthousiasme blijft aanstekelijk! Bedankt om mij te introduceren en te begeleiden in de boeiende wereld van Epigenetica & Proteomics. Je passie voor wetenschap en je creativiteit kennen hun gelijke niet.

De administratieve motor achter Lab FBT: Astrid & Nadine – bedankt om de boel hier draaiende te houden! ;) Uiteraard ook dank je Inge, voor het enthousiast en vlot regelen van de laatste administratieve doctoraatsloodjes! David en de dames van het DNA labo, bedankt om mij de voorbije jaren op regelmatige basis van groene mapjes te voorzien (oh ironie ;)) maar meer nog voor de babbels & mopjes tussen het vele werk door om de sfeer erin te houden. Yens & Sofie, de logistieke krachten in het labo, ook jullie staan altijd klaar om met een glimlach een praktische oplossing aan te bieden.

Katleen, het gemak waarmee jij 100 projecten en evenveel externen, het bieden van hulp & goede raad aan collega's, het kwaliteitssysteem én je gezin lijkt te combineren, siert je. Ik hoop dat we in de toekomst nog veel morgen samenwerken (en demo's mogen 'verknoeien' ;)) Trees, je bureau bevindt zich intussen twee verdiepingen lager, maar onze lach-decibels kunnen dat gelukkig moeiteloos overbruggen ☺ Bart, bedankt voor de occasionele vrolijke noot op het 2e ☺ Ik bewonder je ambitie en

enorme motivatie voor je werk, die je zelfs aan de dag legt (of misschien nog meer) als het om statistiek gaat – Goe bezig! ;) Ellen, de award voor grootste doorzettingsvermogen gaat zeker naar jou. Ik wens jou en Koen veel geluk met de kleine baby die je binnenkort op de wereld gaat zetten. Elisabeth, één van onze laatste aanwinsten voor het histon-onderzoek en die 'aanwinst' mag je letterlijk nemen. Geloof in jezelf, je kan het. Lieselot, die andere 1,5 meter van Lab FBT ;) ik hoop dat het IWT intussen ook begrepen heeft wat voor een straffe madam je bent! De nieuwste lichting, Senne & Laura, nog maar net begonnen aan het Grote Doctoraatsavontuur; jullie gaan dat goed doen! 'De mannen van beneden' Yannick, Dieter, Christophe & Filip, bedankt voor het oplossen van de vele IT-probleempjes, voor de statistische hulp, voor de filosofische insteek in elk gesprek en voor de stoere verhalen tijdens de middagpauze.

Vervolgens, mijn buddies van het eerste uur en mede-IWT-gesneuvelden (nu we toch alle drie op het einde zitten, mag een mens er al eens mee lachen ☺): Veerle & Pieter. Veerle, bedankt voor alle mentale steun die je mij de voorbije 5 jaar als bureau-buurvrouw gegeven hebt ;) Van (niet)wetenschappelijke discussies, over uit de bol gaan op labo-uitstapjes & vrijgezellenfeestjes, tot een aperitiefje op de Coupure (de mogelijkheden zijn legio ;)), met jou valt er altijd wat te beleven! De courage met je laatste thesis-loodjes! Pieter, mijn tijdsbesef is er duidelijk op achteruit gegaan sinds ik mijn horloge niet meer gelijk kan zetten met jouw "het is 10 uur 's morgens"-Dinokoek. Ik mis onze big- en smalltalk gaande van het ventileren over onze resultaten en de laatst nieuwe histon-publicaties tot de laatst nieuwe series op TV. Ik was al overtuigd van je wetenschappelijke capaciteiten, maar ik weet zeker dat je ook in het apotheek-leven je plan zal trekken, succes ermee!

And last but not least, Paulien. Het klikte al meteen tijdens je thesisperiode in ons labo en vervolgens hebben we als collega's en uiteindelijk ook als vriendinnen samen de woelige histone clipping-watertjes doorzwommen. Op het moment dat ik er de brui aan wou geven, heb jij mij de stimulans gegeven om toch door te bijten. Je hebt mij steeds een luisterend oor geboden, zowel op als naast het werk. Bedankt hiervoor! En alvast een gelukkige verjaardag voor morgen ☺

Ook bedankt aan mijn ex-collegaatjes Kelly (voor alle hulp bij de start van mijn doctoraatsperiode), Marlies (voor jouw steeds nuchtere kijk, ongezouten mening & bijhorende lachbuien), Sandra (voor al het duimen), Mado (mijn verjaardagsbuddy ;) voor jouw praktische aanpak), Jip (voor de leuke gesprekjes in het koffiekot), Bert (voor alle loop-uitstapjes) & Shahid (good luck with your thesis!). En uiteraard ook een grote dankjewel aan de thesisstudenten die meegeholpen hebben aan dit werk: Sarah, Maarten, Ine, Erika & Stefi.

“Een gezonde geest in een gezond lichaam”. Mijn gezonde geest werd/wordt mede in stand gehouden dankzij Joke, Tina, Anke & Maarten. Jullie kennen mij als geen ander en we slagen er keer op keer in om elkaar perfect aan te vullen. Bedankt voor elke lach, alle ontspannende momenten en uitlaatklep die jullie mij al heel lang bieden. You rule!

Uiteraard ook een grote dankjewel aan de andere PortoPull’n/scoutsvriendinnetjes An, An-Sofie, Elke, Lynn, Sanne, Sien, Sandrine, … en alle mannen Brecht, Dirkske, Tom, Jonas, Niels, Frank, Lennert, Wouter, Stijn, Matthi, Gorik, … voor de leuke weekends & zotte verkleedfeestjes. Ik hoop dat we nog tot in het rusthuis regelmatig in aangepaste verkleedkledij een porto’ke mogen drinken met elkaar ;) Nog een speciaal woordje van dank voor An-Sofie voor alle tijd die je in het cover-ontwerp van dit boekje gestoken hebt. Amélietje, bedankt voor alle leuke babbels & veel succes met de aankomende examens! Bedankt reisgenootjes Steph, Karo, Veerle, Liesbeth (F), Karen, Liesbeth (C), Céline, Trien, Jonas, Stijn, … om van elke trip een dol avontuur van te maken! Vlien, we zien elkaar niet genoeg, maar onze internaat/kot-tijden lijken nog steeds alsof het gisteren was ;) En ten slotte nog de andere ‘farma-girls’ Marijke, Evelien, Lien, Daf, Magalie & Delphine (+ wederhelften & kindjes), we moeten dringend nog eens een volgend etentje vastleggen!! ☺

Tantien, bij onze eerste ontmoeting was ik nog niet zo enthousiast ;) maar dat is snel veranderd! Tegenwoordig lukt het afspreken iets minder, maar niettemin ben ik je dankbaar voor alles. Ilse, mijn ‘oudste’ kameraadje, jij hebt intussen zelf al twee schatten van kindjes, je doet dat goed ☺

Tantekes & nonkels, neefjes & nichtjes, ik heb geluk met jullie! Ik kijk al uit naar de volgende familie-feestjes... Bobo & Parrain, Oma & Opa, jullie hebben het tot stand komen van dit werk helaas niet meer mogen meemaken, maar ik koester de mooie herinneringen en draag deze thesis dan ook aan jullie op.

Paps & Christine, het lukt soms moeilijk om onze agenda’s te matchen, maar toch weet ik dat ik altijd bij jullie terecht kan/zal kunnen. Ik wens jullie ook veel succes met de carrière-switch die er stilaan zit aan te komen – dat gaat in orde zijn!

Godfried, bedankt voor alle steun, logistieke hulp en vrijheid om te doen wat ik wil, die je mij al jaren onvoorwaardelijk gegeven hebt. Mams, bedankt om mij steeds te motiveren het beste van mezelf te geven en zelfs meer in mij te geloven dan ik zelf deed. Ondanks je drukke job, weet je toch tijd te maken voor de occasionele moeder/dochter-uitstapjes ;) Ik ben een supertrotse “dochter van”!