UNIVERSITY OF LJUBLJANA
FACULTY OF PHARMACY

NUŠA KRIVEC

MASTER'S THESIS

MASTER’S STUDY PROGRAMME LABORATORY BIOMEDICINE

Ljubljana, 2018

UNIVERSITY OF LJUBLJANA
FACULTY OF PHARMACY

NUŠA KRIVEC

PROPERTIES OF HUMAN EMBRYONIC STEM CELLS WITH 18q DELETION IN CELL CULTURE AND DURING THEIR *IN VITRO* DIFFERENTIATION

LASTNOSTI ČLOVEŠKIH EMBRIONALNIH MATIČNIH CELIC Z DELECIJO 18q V CELIČNI KULTURI IN MED NJIHOVO DIFERENCIACIJO *IN VITRO*

MASTER'S STUDY PROGRAMME LABORATORY BIOMEDICINE

Ljubljana, 2018

The research was carried out at the Vrije Universiteit Brussel, Faculty of Medicine and Pharmacy in the Research Group of Reproduction and Genetics under supervision of my promoter and co-mentor Prof. Dr. Mieke Geens. My supervisor and mentor at the Faculty of Pharmacy was Assoc. Prof. Dr. Matjaž Jeras.

ACKNOWLEDGMENTS

I would like to express my gratitude to the Research Group of Reproduction and Genetics at Vrije Universiteit Brussel for giving me a chance to complete the research work for my thesis at their department and for the welcoming atmosphere I experienced during my stay. Especially, I would like to thank my promoter Prof. Dr. Mieke Geens for guidance and support in the research as well as for her time she dedicated to this thesis. I also give thanks to my mentor Assoc. Prof. Dr. Matjaž Jeras, who patiently led me through the educational process from thought-provoking lectures, to the graduation.

Furthermore, I am thankful to my fellow students without whom I cannot imagine my Erasmus journey. Thank you Alba, Ela and Jodie, for being the best company and teachers in the cell laboratory and for your encouragement and positive energy each day. A huge thank you also to Alex, Christina, Joke and Silvie, who were always there to help and gave me new ideas and perspectives.

Finally, I am thankful to my family and friends back home: Barbara, Bojan, Klemen, Metka and Tjaša, who were my safe haven and support when I met challenges of writing a thesis and living abroad.

Statement

I declare that this Master's thesis was done by me under the supervision of Assoc. Prof. Dr. Matjaž Jeras and Prof. Dr. Mieke Geens.

Nuša Krivec

Ljubljana, 10. 10. 2018

Head of the Committee: Prof. Dr. Marija Sollner Dolenc

Member of the Committee: Assist. Dr. Nejc Horvat

TABLE OF CONTENT

ABSTRACT

Human embryonic stem cells (hESCs) are pluripotent cells that possess properties of self-renewal and the capacity to differentiate into any cell type of the adult human body. Embryonic stem cells are valuable for their use in pharmaceutical research, cell therapies and transplantation procedures. However, when hESCs are kept in *in vitro* culture, they tend to acquire chromosomal abnormalities, which, in case they provide the hESCs with a selective advantage, can cause aberrant cells that can take over the culture. One of the reported abnormalities that confer a selective advantage to cells when co-cultured with mouse embryonic fibroblasts (MEFs), is a derivative chromosome 18. In this thesis we focused on progression of the 18q deletion during undifferentiated hESCs culture and upon their differentiation, when cultured on a more clinically-relevant feeder-free cell culture system based on Laminin-152.

Mosaic cultures were created by mixing different ratios of wild-type (WT) and mutant (MT) cells (10% of MT and 90% of WT, 50% of MT and 50% of WT). Real time polymerase chain reactions, with copy number assays for genes *ATP9B* and *NFATC1* and gene expression assays for *LINC00116* were used to discriminate between wild-type and mutant cells in mixes. At the protein level, antibodies against cytoplasmic protein TSPYL5 were used. They were detected with immunofluorescent microscopy, an automated cell counter and flow cytometry.

Genes *ATP9B, NFATC1* and *LINC00116* are suitable genomic markers for cells carrying the 18q deletion, but protein marker TSPYL5 failed to distinguish between the WT and MT cells. In the culture of undifferentiated hESCs, we observed fluctuations in ratios of MT vs WT cells. However, partial take-over was only observed in the culture containing 50% MT of one cell line, while in all other combinations there was no significant increase observed between passages. In contrast to spontaneous differentiation, where cells with 18q deletion seem to display a selective advantage upon directed hepatoblast differentiation, the ratios of WT and MT cells remained similar. Further research should be carried out to confirm and validate our results.

Keywords: *human embryonic stem cells, chromosome instability, 18q deletion, mosaicism*

RAZŠIRJEN POVZETEK

Uvod. Človeške embrionalne matične celice (EMC) so pluripotentne celice, ki imajo sposobnost samoobnavljanja in diferenciacije v katerikoli celico odraslega človeškega telesa. Embrionalne matične celice lahko uporabljamo na področju farmacevtskih raziskav, celičnih terapij in presaditev tkiv in organov. Najpogosteje jih pridobivamo z izolacijo iz notranjega celičnega skupka blastociste embrijev, pridobljenih s postopkom oploditve *in vitro*. Leta 2006 sta Takahashi and Yamanaka uspešno reprogramirala somatske celice v pluripotentne matične celice z uporabo genov za transkripcijske dejavnike Oct4, Sox2, Klf4 and c-Myc, ki se močno izražajo v pluripotentnih EMC.

Za gojenje EMC v *in vitro* kulturi so sprva uporabljali hranilne plasti mišjih embrionalnih fibroblastov (MEF) in gojišče z dodanimi rasnimi dejavniki in govejim serumom. Ta standardni protokol vsebuje veliko živalskih produktov, ki lahko vsebujejo imunogene in bi pri klinični uporabi predstavljali nevarnost za reakcijo z imunskim sistemom prejemnika. Razvoj gre v smeri uporabe pogojev brez hranilnih plasti in gojišč brez seruma. Večinoma se uporabljajo ekstracelularni matriksi, kot so Matrigel, kolagen IV, laminin, fibronektin in gojišča s humanim serumom ali njegovimi nadomestki.

Embrionalne in inducirane pluripotentne matične celice lahko med *in vitro* kulturo pridobijo kromosomske nepravilnosti, ki jim omogočajo selektivno prednost in s tem razrast v celični kulturi. Najpogostejše kariotipske spremembe vključujejo pridobitve na kromosomih 1, 12, 17, 20 in X. Te mutacije so lahko posledica anafaznega zaostajanja, kromosomskega nerazdvajanja in poškodb DNA. Spits in sodelavci (2008) so prvi poročali o deleciji na dolgem kraku kromosoma 18. Delecija se je v celicah pojavila skupaj z duplikacijo na kromosomu 5q ali 7q. Kot posledica nepravilne poprave je nastal derivativni kromosom 18, ki je vseboval del kromosoma 18 in del kromosoma 5 ali 7. Na lokusu 18q je prisotnih veliko tumor-supresorskih genov, kot so SMAD4, SMAD2 in DCC.

Namen. V tej magistrski nalogi smo spremljali človeške EMC z delecijo 18q v kulturi in med diferenciacijo. Najprej smo želeli poiskati označevalec, ki bi omogočil razlikovanje med celicami mutanta in divjega tipa. Naš cilj je bil raziskati, ali nediferencirane celice z delecijo 18q, ki so gojene na Lamininu-521, pridobijo selektivno prednost, kot se to zgodi v kokulturi z MEF. Zanimale so nas tudi lastnosti mozaičnih kultur med spontano in usmerjeno diferenciacijo v hepatoblast.

Metode. Mozaične kulture smo ustvarili z mešanjem različnih deležev celic divjega tipa (DT) in mutanta (MT) (10 % MT z 90 % DT in 50 % MT z 50 % DT). Za razlikovanje med celicami DT in MT smo uporabili metodo verižne reakcije s polimerazo v realnem času, s testom za določanje variabilnega števila kopij genov *ATP9B* in *NFATC1* in testom za ugotavljanje stopnje izražanja gena *LINC00116*. Na nivoju proteinov smo uporabili protitelesa proti citoplazemskemu proteinu TSPYL5, ki smo jih zaznavali z imunofluorescenčno mikroskopijo, avtomatskim celičnim števcem in pretočnim citometrom.

Rezultati. Preverili smo občutljivost testov in njihovo zmožnost zaznavanja celic mutanta v mozaičnih kulturah. Pripravili smo mešanice DNA in RNA z različnim deležem DT in MT. Rezultati so pokazali, da lahko spremembo v številu kopij in izražanju gena zanesljivo zaznamo v primeru, da je razlika v vsebnosti mutanta vsaj 20 %.

Mozaične kulture nediferenciranih celic smo vzdrževali v *in vitro* pogojih 13 pasaž, med katerimi smo izmenično določali izražanje in število kopij genov. Količina mutanta je med posameznimi pasažami nihala, vendar pa v nobenem primeru ni prišlo do razrasta mutanta v kulturi. Najbolj se je prevzetju približala kultura s 50 % začetno vsebnostjo mutanta celične linije VUB04_CF, saj je količina mutanta narasla v prvih treh pasažah in ostala precej konstantna do konca poskusa. Dokazali smo tudi, da je pritrjevanje celic na laminin enako učinkovito tako pri divjem tipu kot pri mutantu in da se razmerje med njima ne spreminja na podlagi tega dejavnika.

Po odvzemu dejavnika, ki vzdržuje pluripotentnost celic, začnejo EMC spontano diferencirati, kar vpliva na signalne poti in celični cikel. Drugačen profil izražanja genov celic z delecijo 18q lahko privede do prednosti pred divjim tipom v procesu specializacije celic. Spremembe v našem eksperimentu so se pokazale tako na fenotipu kot na znižanju števila kopij in znižanju izražanja gena ter s tem dokazale selektivno prednost mutanta. Enako smo pričakovali tudi pri diferenciaciji celic v hepatoblast, saj so izpostavljene stresnim dejavnikom med diferenciacijo v specifičen tip celic. Izkazalo se je, da celice z mutacijo na kromosomu 18 ne pridobijo prednosti v kulturi med usmerjeno diferenciacijo v hepatoblast.

Na nivoju proteinov smo želeli razliko med divjim tipom in mutantom pokazati z uporabo protitelesa proti proteinu TSPYL5. Z barvanjem jeder z barvilom Hoechst in analizo s

pretočnim citometrom smo uspeli pokazati spremembe v celičnem ciklu med diferenciacijo, medtem ko je bilo razlikovanje med divjim tipom in mutantom neuspešno, saj se je protitelo vezalo tako v citoplazmi divjega tipa kot mutanta pred in po diferenciaciji.

Diskusija. Po podatkih iz literature je meja zaznave mozaicizma z metodo qPCR med 5 % in 10 %. Z našimi testi za število kopij in izražanje gena smo razliko med MT in DT zaznali pri spremembi količine MT za 20 %, kar je bilo za naše eksperimente sprejemljivo. Mešanice enakih deležev MT in DT so ustvarile enake pogoje za preučevanje selektivne prednosti, medtem ko so mešanice z 10 % MT služile kot realistični primeri naravnega mozaicizma, pri katerih bi lahko zaznali prevzem kulture.

Po predhodnih podatkih so celice z delecijo 18q prevzele kulturo v prvih nekaj pasažah, v primeru gojenja celic na hranilni plasti MEF, medtem ko v kulturi na Lamininu-521 razrasta mutanta nismo opazili. Razlog je lahko v uporabi različnih tehnik disociacije pri pasažah. Za sproščanje celic z laminina smo uporabili encime in tako pridobili suspenzijo celotne populacije celic v gojilni posodi, vključno s poškodovanimi in senscentnimi celicami. Kolonije v kulturi na hranilni plasti pa je izbral operator, jih mehansko razrezal na majhne kose in jih prenesel v novo gojilno posodo. Pri tej tehniki obstaja večja možnost pristranskosti pri izbiri kolonij, ki lahko vpliva na količino mutanta v novi pasaži.

Pri spontani diferenciaciji preživi več celic z delecijo 18q kot celic divjega tipa. Na kromosomu 18q se nahajajo številni tumor-supresorski geni, kar lahko pojasni selektivno prednost celic z delecijo. Proces diferenciacije in manj tumor-supresorskih genov v mutantu bi lahko sprožili maligno transformacijo. Med usmerjeno diferenciacijo so celice izpostavljene večjemu selekcijskemu pritisku, saj dodani faktorji, ki sprožijo diferenciacijo, vplivajo na spremembe v signalnih poteh. Čeprav smo pričakovali, da bodo rezultati podobni kot pri spontani diferenciaciji, je razmerje med divjim tipom in mutantom ostalo enako, kar je možno pojasniti s tem, da spremembe v bioloških procesih med usmerjeno diferenciacijo v hepatoblast ne vplivajo na celice z delecijo 18q in njihov molekulski profil ni pomemben za razvoj endoderma.

Glede na rezultate analize izražanja gena *TSPYL5* smo pričakovali, da bo protein, ki ga gen kodira, ustrezen označevalec celic z delecijo 18q, vendar se je izkazal kot neuspešen. Številni dejavniki, kot so posttranskripcija, translacija in hitrost razgrajevanja proteina, vplivajo na povezavo med RNA in proteini ter lahko povzročajo razlike v izražanju.

S to raziskavo smo dobili vpogled v lastnosti EMC z delecijo 18q, vendar so za potrditev rezultatov in za boljšo statistično značilnost potrebni dodatni poskusi z več biološkimi ponovitvami.

Ključne besede: *človeške embrionalne matične celice, kromosomska nestabilnost, delecija 18q, mozaicizem*

LIST OF ABBREVIATIONS

aCGH	Array-based comparative genomic hybridization
bFGF	Basic fibroblast growth factor
BMP4	Bone morphogenetic protein 4
BSA	Bovine serum albumin
cDNA	Complementary DNA
CD45	Cluster of differentiation 45; lymphocyte common antigen
CNA	Copy number assay
CNVs	Copy number variations
DCC	Deleted in colorectal cancer gene
ddPCR	Digital droplet PCR
EBs	Embryoid bodies
ESCs	Embryonic stem cells
FACS	Fluorescence-activated cell sorting
FBS	Fetal bovine serum
FISH	Fluorescent in situ hybridization
GEA	Gene expression assay
GFP	Green fluorescent protein
hESCs	Human embryonic stem cells
HDAC inhibitor	Histone deacetylase inhibitors
ICM	Inner cell mass
iPSCs	Induced pluripotent stem cells
MEFs	Mouse embryonic fibroblasts
mESCs	Mouse embryonic stem cells
MT	Mutant
PBS	Phosphate buffered saline
PI3K-inhibitor	A phosphoinositide 3-kinase inhibitor
qPCR	Real time polymerase chain reaction
REGE	Reproduction and Genetics
RT-qPCR	Reverse transcription qPCR
SALL3	Spalt like transcription factor 3; gene
SMAD2, SMAD4	Mothers against decapentaplegic homolog 2 and 4; genes
TSPYL5	Testis specific Y-like 5 protein
VEGF	Vascular endothelial growth factor
wnt3a	Signaling protein
WT	Wild-type
XCI	X chromosome inactivation

1 INTRODUCTION

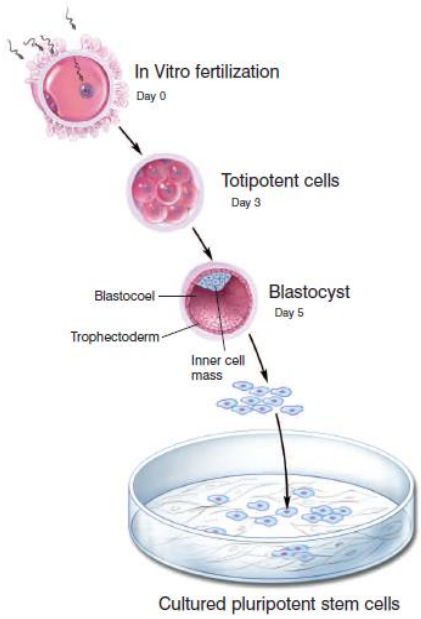
1.1 PLURIPOTENT STEM CELLS

Stem cells are unspecialized cells found in the embryo, fetus or adult human and have the ability of unlimited self-renewal and differentiation into specialized cells with specific function. They hold great promise for tissue repair and regeneration and can be used as a tool for disease modelling and drug discovery. The potential of clinical application for stem cells is wide, as they can be used for treatment of spinal cord injuries, diabetes, Parkinson's disease, Alzheimer's disease, osteoarthritis, cancer and heart disease. Most stem cell treatments are still in the phase of clinical trials, however there are a few therapies that are accepted as safe and effective. For example, the most used is hematopoietic stem cell transplantation for treating hematological malignancies, such as leukemia. [1–4]

Pluripotency is a property of individual cells that have the capacity to give rise to any of the three embryonic germ layers of the early embryo, i.e. ectoderm, endoderm and mesoderm. Therefore, they can develop into every cell type of the adult body. In mammals, pluripotency is restricted to early embryonic cells, primordial germ cells, and the stem cells of tumors derived from pluripotent cells (embryonal carcinomas). [5,6]

After fertilization and formation of the totipotent zygote, a cluster of cells called blastomeres is generated through multiple cleavage divisions. The blastomeres finally form a blastocyst that consist of inner cell mass (ICM) also called embryoblast, and outer cell mass called trophoblast. The ICM develops into the embryo while the trophoblast forms extra-embryonic tissue. [1,7]

1.1.1 EMBRIONIC STEM CELLS



Embryonic stem cells (ESCs) are pluripotent cells derived from embryos that are produced by *in vitro* fertilization for clinical purposes. They are usually isolated from the ICM of a blastocyst (Figure 1) and can maintain the same properties of self-renewal and pluripotency *in vitro* as ICM cells *in vivo*. The first mammalian ES cell lines were derived from a mouse blastocyst in 1981 (Evans and Kaufman, 1981; Martin, 1981) followed by the isolation of human embryonic stem cells (hESCs) in 1998 by Thompson *et al.* and Reubinoff *et al.* [1,6,8,9]

Figure 1: Derivation of human embryonic stem cells.[6]

1.1.2 INDUCED PLURIPOTENT STEM CELLS

In 2006 Takahashi and Yamanaka were successful at reprogramming somatic cells to acquire properties of ESCs. These cells are called induced pluripotent stem cells (iPSCs) and they are highly similar to ESCs in terms of morphology, gene expression, development potential and their epigenetic pattern. iPSCs were generated by transient overexpression of four embryonic transcription factors, i.e. Oct4, Sox2, Klf4 and c-Myc, currently known as Yamanaka factors. As iPSCs can be derived directly from somatic cells of volunteers or patients, they have several advantages over ESCs, such as: reduced risk of immune rejection when used for transplantation, potential for generating different disease models, which can be linked to a relevant disease phenotype, and minimal, if any, ethical and religious conflicts. However, there are some limitations associated with iPSCs as well. Despite improvements in reprogramming methods, the derivation yield remains relatively low. Furthermore, incomplete reprogramming of iPSCs can lead to transformation into cancer cells.[10–12]

1.2 GROWTH AND MAINTENANCE OF hESCs IN CULTURE

After derivation of an ESC line from the ICM of a blastocyst, stem cells are cultured in conditions that maintain them in an undifferentiated state with their key characteristics of

pluripotency and self-renewal. Researchers are trying to mimic the *in vivo* microenvironment called niche in which stem cells grow and it includes adhesion of cells to matrix, intercellular interactions and growth factors.[13]

The first systems implemented for hESC culture applied feeder cells to support the hESC propagation, usually mouse embryonic fibroblasts (MEFs) (Figure 2). These secrete essential growth factors, cytokines and extracellular matrices such as TGFβ, activin A, laminin-511 and vitronectin, necessary to maintain the pluripotent state. Growth factors are also present in the medium supplemented with a 20% fetal bovine serum. These standard culture protocols contain a variety of animal products and therefore hESCs cultured through these processes, cannot be used for clinical application, as they may contain immunogens that evoke immune responses and lead to rejection upon transplantation, as well as pose a risk of contamination by animal pathogens, such as viruses and prions.[6,7,13–17]

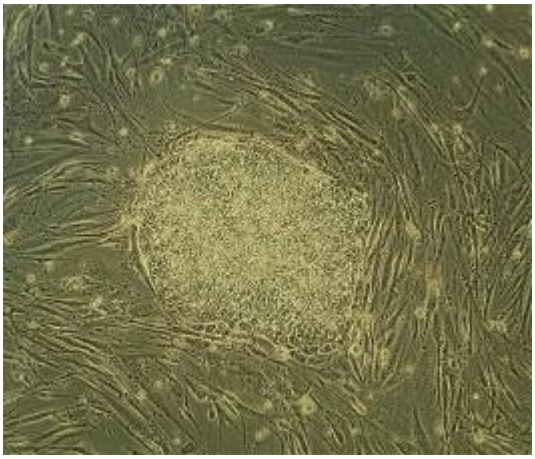


Figure 2: Stem cells cultured on MEFs.[12]

Research took a course towards growing hESCs in feeder-free conditions and in 2001 Xu et al. reported the use of Matrigel matrix with a 100% MEF-conditioned medium supplemented with a basic fibroblast growth factor (bFGF). Matrigel is an extracellular matrix isolated from a mouse Engelbreth-Holm-Swarm teratocarcinoma cells. It contains multiple components, such as collagen type IV, laminin, heparin sulfate proteoglycans, entactin and other growth factors. Even though Matrigel was one of the most commonly used substrates, it is not an ideal substitute for feeder-free culture, because of its undefined characteristics, lot-to-lot variability and risk associated with unwanted xenogeneic contaminants. [6,7,13–16]

In order to grow therapeutic grade hESC lines, chemically defined matrices and xeno-free cultures have to be established. Several components required for hESC growth have been identified and studied. Defined matrices such as collagen IV, laminin, vitronectin and fibronectin can support feeder-free hESC growth in combination with a medium that contains human serum or serum replacement. [6,7,13–16]

1.2.1 DIFFERENTIATION

During the developmental stage of an organism, undifferentiated embryonic cells undergo differentiation to produce specialized cells. Lineage-specific *in vitro* differentiation of hESCs can help us study development of many cell types and establish new therapeutic strategies. Three general approaches are used to initiate ES cell differentiation. The first one is the formation of three-dimensional aggregates known as embryoid bodies (EBs). The second method is a co-culture with stromal cells, while the third one involves differentiation of ESCs in a monolayer on extracellular matrix proteins. Differentiation protocols should enable efficient development of the aimed cell type, easy isolation of differentiated cells and mimic the developmental program in the early embryo. Undifferentiated and differentiated cells express receptors for different soluble growth factors that affect their developmental pathways. Differentiation can be partially directed by the addition of growth factors to culture media, which then alter gene expression profiles of cultured cells.[7,18,19]

Endoderm linages that have been derived are mostly pancreatic cells, hepatocytes and lung epithelium, because of their clinical potential for the treatment of type I diabetes, liver failure and pulmonary diseases. Generation of definitive endoderm can be achieved by activin A treatment, which activates the TGF-beta pathway. However, this pathway is not the only one that directs ESCs into definitive endoderm. Usually some additional growth factors are used for differentiation, for example wnt3a, PI3K-inhibitors and HDAC inhibitors. [7,20,21]

Hematopoietic, vascular and cardiac lineages are most frequently derived from mesoderm and are the easiest to generate. Differentiation to hematopoietic cells can be achieved with the co-culture with mouse bone marrow stromal cells or with the generation of EBs. After formation of EBs, cytokines interleukin 3 (IL-3) and interleukin 6 (IL-6), BMP4 and VEGF are added to the culture. The majority of the cells is converted to hematopoietic progenitor cells, as indicated by the expression of the surface CD45 marker.[7,18]

Ectodermal differentiation can give rise to non-neuronal cells, like keratinocytes and retinal pigment epithelium, as well as to neural cells, such as neurons, astrocytes and oligodendrocytes. Generating neuroectoderm and neural cells from hESCs is an important step towards treating neurodegenerative diseases. The first neural progenitor cells were isolated after *in vitro* cultivation of hESCs for four to seven weeks, at a high cell density. Some cells expressed an embryonic polysialylated neural cell adhesion molecule (PSA-

NCAM) and changed morphology. They were mechanically dissected, expanded in serum-free media and plated on coverslips coated with poly-D-lysine and laminin. Later, B27 supplement, human recombinant epidermal factor EGF and the mitogen bFGF, were added to the medium during expansion, which resulted in a successful differentiation into all three neural lineages. Neural cells were identified by the expression of neural surface markers, including the neurofilament protein and beta-tubulin, and containing glutamic acid decarboxylase and glutamate.[6,7]

1.2.2 CHARACTERIZATION

hESCs grow in relatively flat, compact colonies and maintain defined borders. They have a high nucleus to cytoplasm ratio and fast population-doubling times of about 36 hours. The self-renewal and differentiation potential depend on surface molecules and transcription factors. Undifferentiated hESCs express stage-specific embryonic antigens SSEA-3, SSEA-4, keratan sulfate antigens TRA-1-60 and TRA-1-81, alkaline phosphatase, POU5F1, NANOG and SOX2. These genes are associated with the pluripotent state and are downregulated upon differentiation. POU5F1, also known as OCT4, is a DNA binding protein, which regulates the expression of genes by binding to the octamer motif 5'-ATGCAAAT-3' and is highly expressed in undifferentiated hESCs, embryonic carcinoma and embryonic germ cells. Oct4 deficient mouse embryos cannot form ICM *in vivo* or mESC colonies *in vitro*. SOX2 is a transcription factor and forms a heterodimer with OCT4. Deletion of SOX2 in ESCs results in trophoblast differentiation.[1,22]

hESCs also express high levels of telomerase activity after prolonged *in vitro* cultivation. Telomerase is an enzyme that adds telomere repeats to chromosome ends, which is highly correlated with immortality of human cell lines. hESCs have demonstrated stable karyotypes, however, during prolonged maintenance *in vitro* they can acquire karyotypic abnormalities that are often correlated to tumorigenesis *in vivo* [7,19,20,23,24].

1.3 GENETIC AND EPIGENETIC INSTABILITIES

1.3.1 KARYOTYPIC ABNORMALITIES

It is known that hESCs and hiPSCs are prone to acquire genetic and epigenetic abnormalities during *in vitro* culture (Figure 3). These mutations can provide a selective advantage and eventually enable them to take over the culture. Genomic changes usually arise after long-

term culture. However, some cell lines become genetically aberrant in early passages and others stay intact even in very late passages. The most frequent changes involve gains of chromosomes 1, 12, 17, 20 and X (Figure 3). The trisomy of chromosome 12 is a predominant abnormality and was found in more than 40% of hESC lines and 30% of iPSC lines worldwide.[2,25–27]

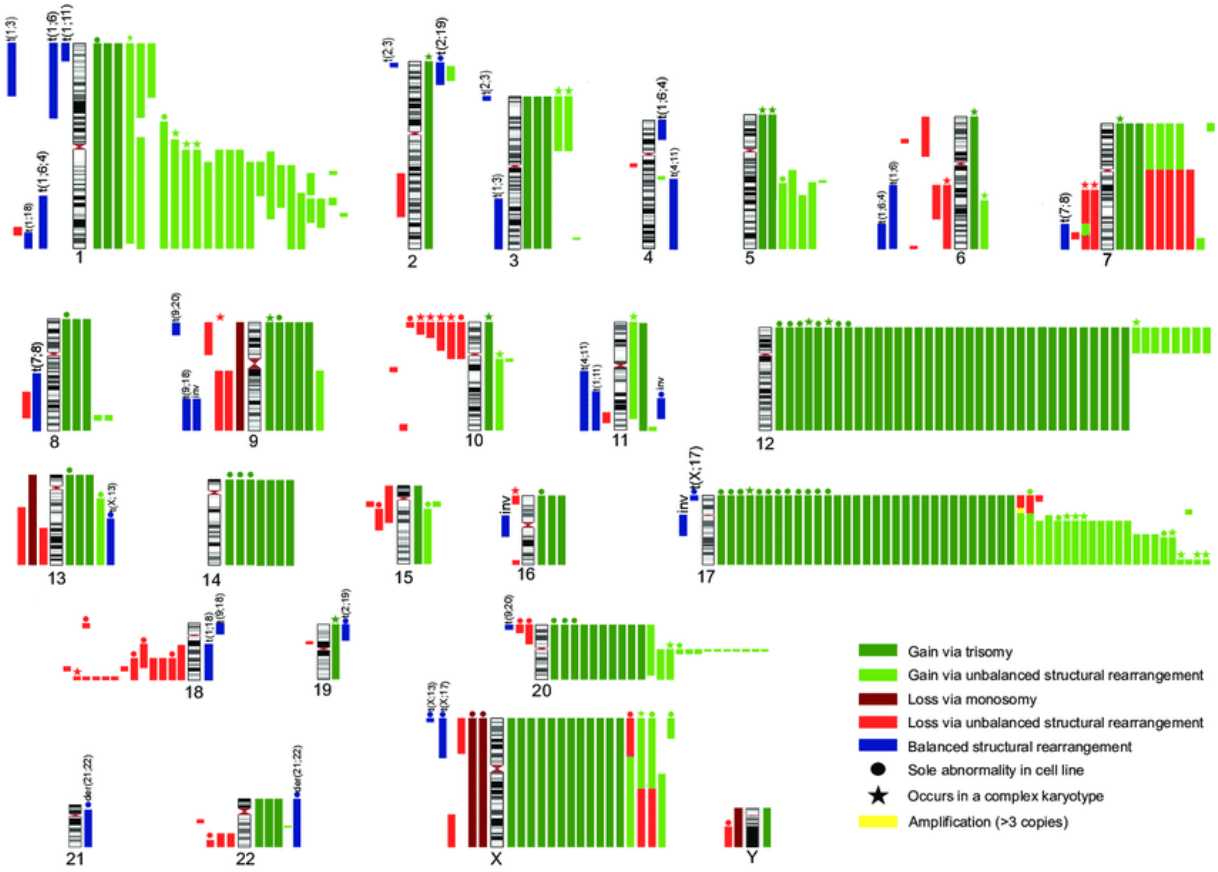


Figure 3: Ideogram of reported chromosome abnormalities.[27]

It has been suggested that mutations can be caused by anaphase lagging, chromosome non-disjunction during mitotic division and DNA damage. There are many mechanisms that help eliminate abnormal cells. They can undergo cell death, differentiate or up-regulate DNA repair pathways to repair the damage. However such repair is not always accurate.[2,25,26]

1.3.1.1 18q DELETION

Apart from full chromosome gains or losses, segmental chromosomal aberrations are also often observed in hESC and hiPSC. An example of such a recurrent aberration is the deletion of a part of the long arm of chromosome 18. Spits et al. (2008) were the first to report the

presence of a derivative chromosome 18 in hESC lines, studied by array-based comparative genomic hybridization (aCGH). Aberrations appeared in cultures of three hESC lines at early passages, ranging from passage 10 to 66. The 18q deletion appeared together with duplication at 5q or 7q. The derivative chromosome 18 was a consequence of abnormal double-strand break repair and it contained a part of chromosome 18 and part of chromosomes 5q or 7q. 18q holds many tumor-suppressor gene candidates, like SMAD4, SMAD2 and DCC, which may explain the selective advantage of cells with 18q deletion in culture. This region also includes the gene *SALL3*, which plays an important role in mammalian retinal development.[28–31]

While in the meantime this aberration has been reported to take over a mixed hESC culture, it seems that all hESC lines which obtained this abnormality were cultured in standard culture system, based on co-cultures with MEFs. Nowadays, however, this culture system has become obsolete and more defined, feeder-free culture systems are state-of-the-art. Moreover, no data is available on this aberration during differentiation of hESC.[32]

1.3.2 CULTURE CONDITIONS

A lot of research has been dedicated on how the composition of the culture medium, passaging method and oxygen tension can influence genetic and epigenetic instability of hESCs and hiPSCs. It appears that mechanical passaging may cover up the existence of aneuploid clones, which become more evident when using chemical or enzymatic passaging. Some hESC lines seem to be more prone to karyotypic abnormalities than others, regardless of culture conditions. Some groups claim that hESCs show lower propensity towards spontaneous differentiation and less chromosomal abnormalities when cultured in hypoxic conditions, which may better reflect the *in vivo* environment. Epigenetic modifications are more sensitive to stress and therefore more affected by *in vitro* manipulation. The change to high oxygen levels has been reported to induce X chromosome inactivation (XCI) and alter DNA methylation status.[25,26]

1.4 DETECTION OF GENETIC AND EPIGENETIC INSTABILITY

The conventional methods used for analysis of genetic changes in hESCs are based on cytogenetic banding techniques, e.g. Giemsa banding (G-banding) (Figure 4). Although

karyotyping provides information on the entire cell genome or large aberrations, it has low resolution and is unable to detect low-grade mosaicism. Examination is limited to a small number of dividing cells and the results must be analysed by an expert. G-banding is still used for detection of balanced rearrangements and partially abnormal mosaic cell populations.[8,27,33–35]

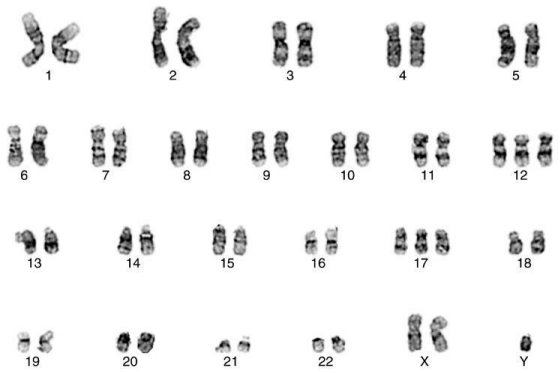


Figure 4: G-banded karyotype of hESCs. Gains of chromosomes 12 and 17 are seen.[7]

For unbalanced genomic changes array-based comparative genomic hybridization (aCGH) or shallow whole genome sequencing are used. The higher resolution of these techniques allows the detection of copy number variants (CNVs) at the kilobase level. However, aCGH is not sensitive enough to detect low-level mosaicism in hESC cultures.[8,27]

Fluorescent in situ hybridization (FISH) analyses on interphase cells is useful for detecting small amplifications and deletions. However, it is an expensive technique that requires highly experienced staff for results interpretation. Its sensitivity enables to detect 0,5% - 5% abnormal cells with specific chromosome amplification.[8,27]

Virtual karyotyping (e-karyotyping) uses gene expression microarray data from the cell line of interest and compares it with public databases, for example the National Center for Biotechnology information Gene Expression Omnibus. It uses two different bioinformatic analyses and the result shows which chromosomal aberrations are present within the cell line. The sensitivity of e-karyotyping to detect mosaicism is around 30%. [27,36]

There are also novel methods, like next-generation sequencing and single cell sequencing, that have a high resolution in detecting low-grade mosaicism but are also more expensive and time consuming, therefore they cannot be used in routine clinical practice. Baker et al. (2016) reported that commonly used methods, such as qPCR, digital droplet PCR (ddPCR) and FISH, have a limit of mosaicism detection at around 5% - 10%, which suggests that a small population of cells carrying genetic aberrations can stay undetected. Consequences and risks of genetically changed hESCs and hiPSCs still have to be determined and studied.[27,34]

2 AIM OF THE STUDY

In this study we aimed at elucidating the progression of the 18q deletion, during undifferentiated hESC culture and upon differentiation of those cells. Our first goal was to find a molecular marker to identify mutant hESCs with an 18q deletion that would be significantly different from wild-type hESC lines. Further, we wanted to check whether the derivative chromosome 18 confers a culture advantage at the undifferentiated level, when hESCs are cultured on Laminin-521 and in a defined medium, as it seems that it does so in the MEF-based cultures. Finally, we aimed to study the progression of the 18q deletion during spontaneous differentiation and directed differentiation into hepatoblasts.

3 MATERIALS AND METHODS

3.1 CELL CULTURE AND PASSAGING

hESCs were cultured in dishes coated with human recombinant Laminin-521 (Biolamina, Sweden). Laminin-521 was diluted to the final concentration of 5µg/ml in phosphate buffered saline (PBS) (Thermo Fisher, USA) with Ca^{++}/Mg^{++}. Before use, dishes were stored overnight at 4 °C. NutriStem™ (Biological Industries, Israel) medium with 10mM penicillin/streptomycin was changed daily. hESC cultures were kept at 37 °C and 5% CO_2. For passaging, cells were washed with PBS and incubated with recombinant enzyme mixture TrypLE (Thermo Fisher, USA) for 10 minutes at 37 °C for single cell dissociation. TrypLE was deactivated with NutriStem™ medium and the cells were centrifuged at 1.500 rpm for 5 min. The obtained cell pellet was resuspended in 1 ml of medium and counted with the image-based cytometer Tali™ (Invitrogen, USA) if needed, or transferred in a new dish in a passaging ratio between 1:20 to 1:30.

3.2 PELLET COLLECTION

Cell suspensions that were not used for passaging but were collected for further analyses were obtained using TrypLE as described above. In this case, 4 ml of PBS was added, and the cell suspensions were spun down at 1.500 rpm for 5 min. The pellet was resuspended in 1 ml of PBS, transferred to an 1,5 ml Eppendorf tube and spun down at 13.000 rpm for 1 min. The supernatant was aspirated, and the dry pellet was first snap-frozen in liquid nitrogen and then stored at -80 °C.

3.3 MIXING EXPERIMENTS

Cells were counted using image-based cytometer Tali™ (Invitrogen, USA) to obtain cultures with different ratios of mutant and wild-type cells. The number of cells that needed to be mixed was calculated based on the cell concentration present in suspension.

3.4 SPONTANEOUS DIFFERENTIATION

Spontaneous differentiation of hESC was performed by changing from NutriStem™ to Embryoid Body (EB) formation medium, which was changed every two days. EB formation

medium does not contain bFGF, a crucial factor for proliferation of undifferentiated hESC. After 21 days, differentiated cells were collected for analysis.

EB medium:

- ~ KnockOut DMEM (Thermo Fisher Scientific, USA)
- ~ 20% Fetal bovine serum (FBS) (Thermo Fisher, USA)
- ~ 0,1mM 2-mercaptoethanol (Sigma-Aldrich, USA)
- ~ 2mM L-glutamine (Thermo Fisher, USA)
- ~ MEM Non-Essential Amino Acids Solution (100x) (Thermo Fisher, USA)
- ~ 10mM Penicillin/Streptomycin (Invitrogen, USA)

3.5 HEPATOBLAST DIFFERENTIATION

For derivation of liver progenitors 35.000 hESCs per cm^2 were plated on a dish, coated with Laminin-521 in order to reach appropriate cell density. After 24 hours the medium was changed from NutriStem to RPMI medium. The next day, RPMI medium without glycogen synthase kinase 3 (GSK3) inhibitor CHIR99021, was used and for the next 6 days, the hESC culture was treated by Hepatoblast medium that was changed daily.

RPMI Medium:	Hepatoblast Medium:
~ RPMI 1640 (Life Technologies, USA)	~ Knockout DMEM
~ 0,5% B27 supplement (Invitrogen, USA)	~ 20% Knockout Serum Replacement
~ 100 ng/ml human activin A (R&D Systems, USA)	~ 0,5x GlutaMAX
~ 3μM/ml CHIR99021 (Stemgent, USA)	~ 1% MEM Non-Essential Amino Acids Solution (100x) (Thermo Fisher, USA)
~ 10mM Penicillin/Streptomycin (Invitrogen, USA)	~ 0,1mM 2-mercaptoethanol (Sigma-Aldrich, USA)
	~ 1% (v/v) DMSO
	~ 10mM Penicillin/Streptomycin (Invitrogen, USA)

3.6 IMMUNOSTAINING

hESCs were fixed in the exponential phase of growth. After removing media, the cells were washed 3 times with PBS and incubated for 10 min with ice cold methanol (Merck Millipore, USA) at room temperature to fix/permeabilize cells. Then the cells were washed again 3 times with PBS. To avoid unspecific binding of antibodies, blocking was performed with a 2-hour incubation with 3% bovine serum albumin (BSA) in PBS. TSPYL5 antibody (antibodies-online, Germany), conjugated with Alexa Fluor 488 was diluted in a 1:100 ratio with 1,5% BSA and incubated with cells for 1 hour in the dark at room temperature. After washing 3 times with PBS, 1:2.000 dilution of Hoechst 33342 stain (Life Technologies, USA) was added and incubated for 15 min at room temperature for nuclear staining. A drop of SlowFade Gold Antifade Reagents (Life Technologies, USA) was added and covered with a coverslip. The cells were analysed with the fluorescent microscope IX-81 (Olympus, Japan).

3.7 FLOW CYTOMETRY

All centrifugation steps were carried out at 1.800 rpm for 5 min. Cells were washed 2 times with PBS. Differentiated cells were detached with 1 ml of Trypsine-EDTA, 10x solution (Sigma-Aldrich, USA) and incubated for 5 min at 37 °C, while hESCs were detached using TrypLE under the same conditions as described for passaging above. The cell suspension was gently pipetted up and down with 2 ml of medium to obtain a single cell suspension and then spun down. Medium was aspirated, and cells were resuspended in 1 ml of PBS. The suspension was filtered through a 40 μm filter to eliminate cell clumps. The number of cells was determined by Tali™ (Invitrogen, USA). One hundred thousand (100.000) cells were used for controls (unstained cells and cells stained only with Hoechst 33342) and 300.000 cells for staining with the antibody. The cells were centrifuged and fixed/permeabilized with 1ml of ice-cold methanol, for 10 min at room temperature. After centrifugation, methanol was aspirated, and cells were washed with PBS. PBS with 1% BSA was added to the cell pellet, which was then resuspended and spun down. Cells were incubated with 100 μl TSPYL5 antibody, conjugated with Alexa Fluor 488 for 20-30 min in the dark at room temperature. The antibody was beforehand dissolved in 1% BSA in ratios 1:100 and 1:200, with addition of Hoechst 33342 (Life Technologies, USA) in ratio 1:2.000. The cells were washed with 1% BSA, spun down, resuspended in 1ml of 1% BSA and transferred into flow

cytometry tubes. The analysis was performed on the Attune flow cytometer (Applied Biosystems, USA). The specificity was increased by gating for cells of interest based on size and fluorescence of controls.

The same protocol was used when analysing differentiated and undifferentiated lines with Tali™ (Invitrogen, USA). Unstained controls were used for minimizing the effect of background.

3.8 RNA EXTRACTION

RNA was extracted from cell pellets with the RNeasy Mini Kit (Qiagen, Germany) following the producer's protocol. TRIZOL® Reagent (Life Technologies, USA), combined with RNeasy Mini Kit (Qiagen, Germany), was used to extract RNA of some cell pellets. In these cases the cell pellet was resuspended in 500 µl of Trizol. After 5 min of incubation at room temperature, 100 µl of chloroform (Merck Millipore, USA) was added and the samples were incubated for 2 min at room temperature. This mixture was spun down at 12.000 x g for 15 min. The transparent upper phase was transferred into a new tube; 70% ethanol (Honeywell, USA) was added and the whole volume was transferred onto the RNeasy Mini spin column. RNA was isolated in-line through the producer's protocol. The concentration of the obtained RNA was measured with a NanoDrop 1000 (Thermo Fisher Scientific, USA). RNA was stored at -80°C.

3.9 cDNA SYNTHESIS

A reverse transcription of RNA to cDNA was performed with the First-Strand cDNA Synthesis Kit (GE Healthcare, USA), according to producer's instructions. The cDNA was stored at -20 °C.

3.10 DNA EXTRACTION

Genomic DNA was extracted with the DNeasy Blood & Tissue Kit (Qiagen, Germany) following the producer's instructions. DNA concentration was assessed with NanoDrop 1000 (Thermo Fisher Scientific, USA). DNA was stored at +4 °C.

3.11 qPCR

Real time polymerase chain reaction was performed on the ViiA™ 7 system (Applied Biosystems, USA). The total volume of each qPCR reaction was 20µl and it contained 10µl of 2x qPCR Master Mix Low ROX (Eurogentec, Belgium), 1µl of 20x TaqMan Copy Number Assay (Life Technologies, USA), 1µl of nuclease-free water and 40 ng of DNA. TaqMan™ Copy Number Reference Assay RNase P (Applied Biosystems, USA 4403328) was added instead of nuclease-free water when performing multiplex reactions. Genomic DNA from leukocytes of a healthy donor was used as control. No template control was tested in each experiment to exclude contamination. Cycling parameters are described in Table I. Each sample was tested in triplicates. The used copy number assays were *ATP9B* (Thermo Fisher, USA Hs06498356_cn) and *NFATC1* (Thermo Fisher, USA Hs06489123_cn). The experiments were run as either comparative C_t (cycle threshold) or relative standard curve analyses. Data analysis was performed by ViiA^{TM} 7 v2.0 or Copy Caller v2.1 software (Applied Biosystems, USA).

Table I: Cycling parameters of qPCR

50 °C	2 min	
95 °C	10 min	40 cycles
95 °C	15 sec	
60 °C	1 min	

3.12 RT-qPCR

Reverse transcription qPCR was performed on the ViiA™ 7 system (Applied Biosystems, USA). The total volume of each qPCR reaction was 20µl and it contained 10µl of 2x qPCR Master Mix Low ROX (Eurogentec, Belgium), 1µl of 20x TaqMan Gene Expression Assay (Life Technologies, USA), 1µl of nuclease-free water and 40 ng of cDNA. For our own designed primers and probes, the following volumes were used: 10µl of 2x qPCR Master Mix Low ROX (Eurogentec, Belgium), 0,72µl of forward primer, 0,72µl of reverse primer, 0,5µl of probe, and 40 ng of cDNA. Final concentration of forward and reverse primer was 1,8 µM, while the concentration of the probe was 250 nM. Each sample was tested in triplicates. Reference genes used in these experiments were *GUSB* and *UBC*, while the gene of interest was *LINC00116* (Table II). Cycling parameters are described in Table I.

Experiments were run as comparative C_t (cycle threshold) analysis. Data analysis was performed by ViiA™ 7 v2.0 software.

Table II: Reference and testing genes

Gene	Primer/probe	Sequence/Reference number	Producer
UBC	Forward	5'-CGC-AGC-CGG-GAT-CAT-TTG-3'	IDT
	Reverse	5'-TCA-AGT-GAC-GAT-CAG-AGC-GA-3'	IDT
	Probe	6-FAM-TCG-CAG-TTC-TTG-TTT-GTG-MGB	IDT
GUSB		Hs 99999908_m1	Life Technologies
LINC00116		Hs 00381665_m1	Life Technologies

3.13 hESC LINES USED

For each experiment, comparison was done between two different wild-type (WT) and mutant (MT) cell lines. The first pair was WT and MT of the cell line VUB04_CF and the second one was WT of the cell line VUB03_DM1 and MT of VUB13_FXS (Table III). Karyotype of all cell lines was determined by array-based comparative genomic hybridization and confirmed by fluorescent *in situ* hybridization or G-banding.

Table III: Information about starting passages and karyotypes of cell lines used in our experiments.

Name	*Passage*	*Karyotype*
VUB04-CF	P15	46, XX
VUB04-CF	P94	46, XX, dup(5)(q14.2qter),del(18)(q21.2qter)
VUB03-DM1	P23	46, XX, dup(20)(q11.21)
VUB13_FXS	P56	46, XX, dup(5)(q21.3qter),del(18)(q12.1qter)

4 RESULTS

4.1 EVALTUATION OF POSSIBLE MOLECULAR MARKERS' SENSITIVITIES

4.1.1 DNA qPCR

Copy number assays were chosen, based on genes that are present in the common region of the 18q deletion. Four different assays for genes *SALL3*, *ATP9B* and *NFATC1* were tested to determine which one is the most suitable for detection of different ratios of mutant cells in culture. Ideally, an assay would detect two copies in wild-type (WT) and one copy in mutant (MT) cells because of the deletion and would also be able to distinguish mosaic cultures.

We first evaluated the efficiencies of primer and probe combinations in the assays, by analysing the standard curves obtained using decreasing amounts of hESC DNA (Figure 5, Table IV). While the standard curves for copy number assays *ATP9B* and *NFATC1* showed linear amplification with R^2 values, being >0,98, the amplifications for CNA *SALL3_1* and *SALL3_2* appear slightly less efficient, with R^2 values of 0,970 and 0,965.

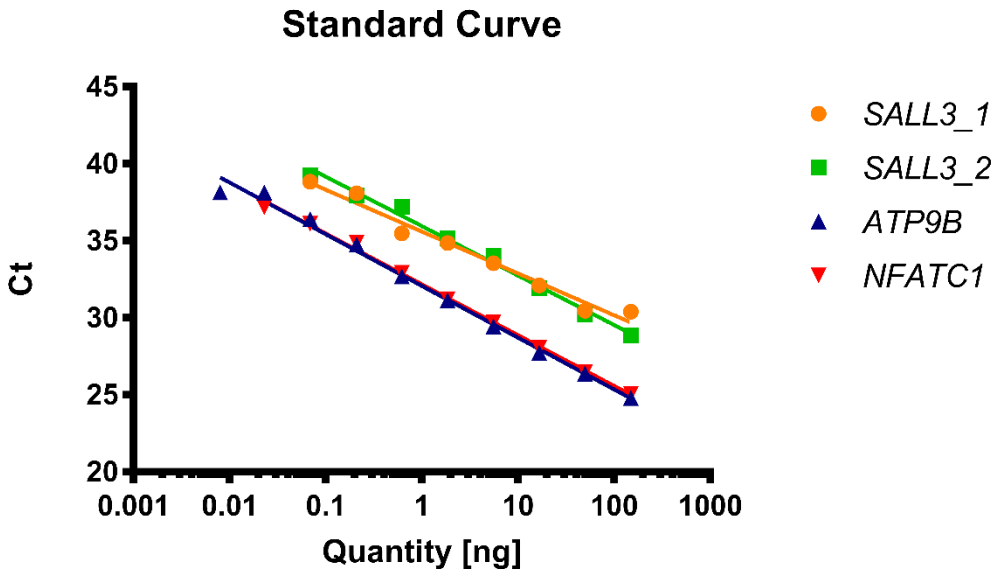


Figure 5: Sandard curves for CNA SALL3_1, SALL3_2, ATP9B and NFATC1.

Table IV: Slope, R^2 and efficiency for standard curves of CNA SALL3_1, SALL3_2, ATP9B and NFATC1.

	● *SALL3_1*	■ *SALL3_2*	▲ *ATP9B*	▼ *NFATC1*
Slope	-3,364	-3,052	-3,404	-3,295
R^2	0,970	0,965	0,987	0,991
E	98,266 %	112,665 %	96,678 %	101,132 %

More importantly, differences were observed when we applied these assays on DNA of WT and MT cells. The copy numbers obtained with CNA for *SALL3_1* and *SALL3_2* differed significantly from the expected results for an ideal CNA, while the copy numbers detected with assays for *ATP9B* and *NFATC1* were much closer (Figure 6). Due to these inconsistent results of the *SALL3* assays, we decided to further continue our experiments with CNA for *ATP9B* and *NFATC1*.

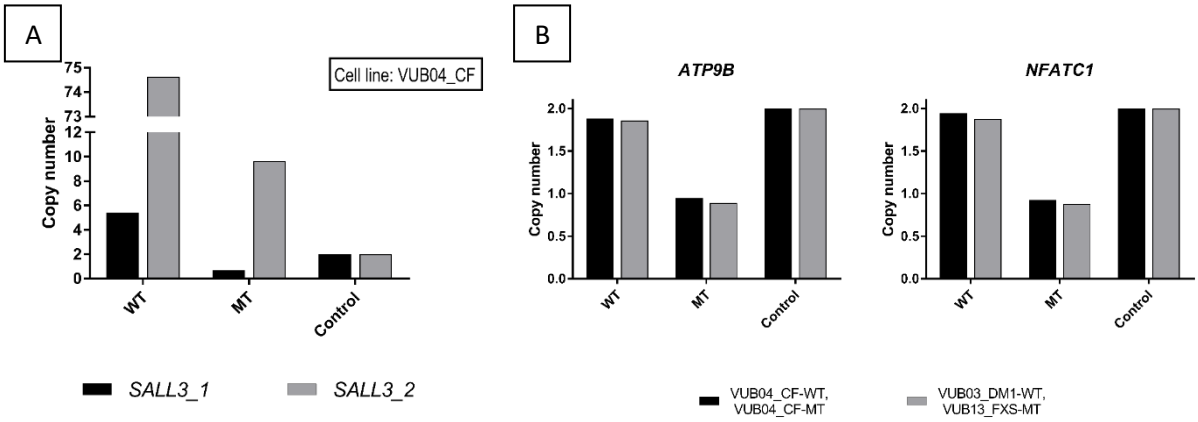


Figure 6: Comparison of copy numbers detected (A) with CNA SALL3_1 and SALL3_2 and (B) with CNA ATP9B and NFATC1. WT= wild-type, MT= mutant, Control=genomic DNA from leukocytes of a healthy donor.

Consistent results were obtained using CNA for *ATP9B* and *NFATC1*, in order to differentiate between WT and MT after spontaneous differentiation (Figure 7), which made these markers appropriate for further use.

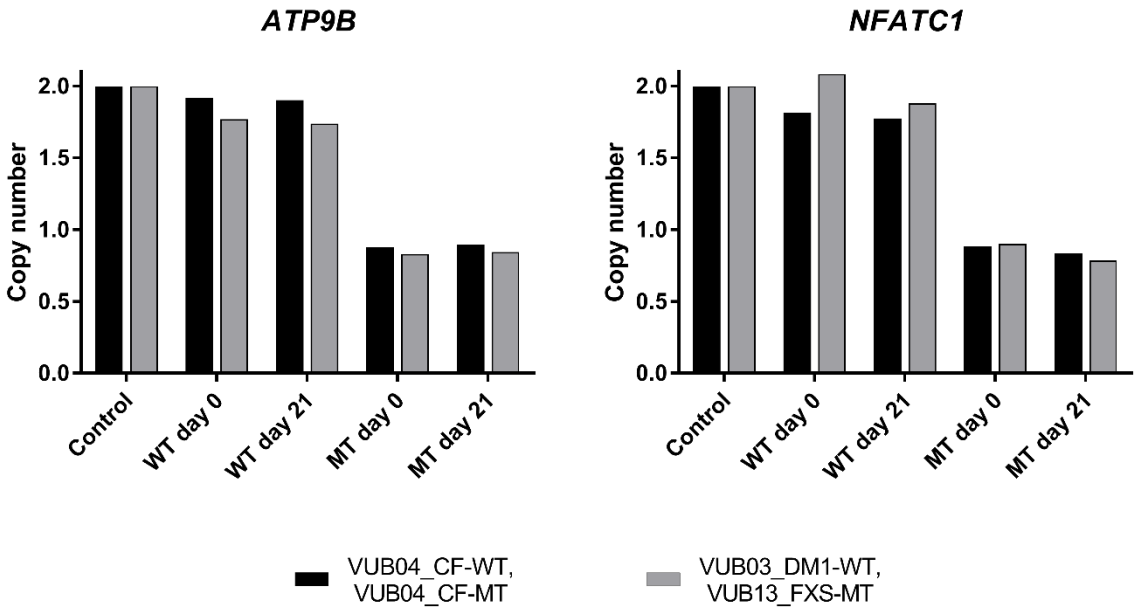


Figure 7: WT and MT cell lines before and after spontaneous differentiation for both CNA. WT= wild-type, MT= mutant, Control=genomic DNA from leukocytes of a healthy donor.

In order to analyse the sensitivity of these CNA, for determination of the mosaicism ratios, DNA mixes with different percentages of MT versus WT cell input were analysed. Two biological replicates were done for each assay. We determined that, as the amount of the MT DNA decreased, the copy numbers decreased almost linearly (Figure 8). Changes in copy numbers could be accurately detected, when the difference in amount of MT DNA was at least 20%.

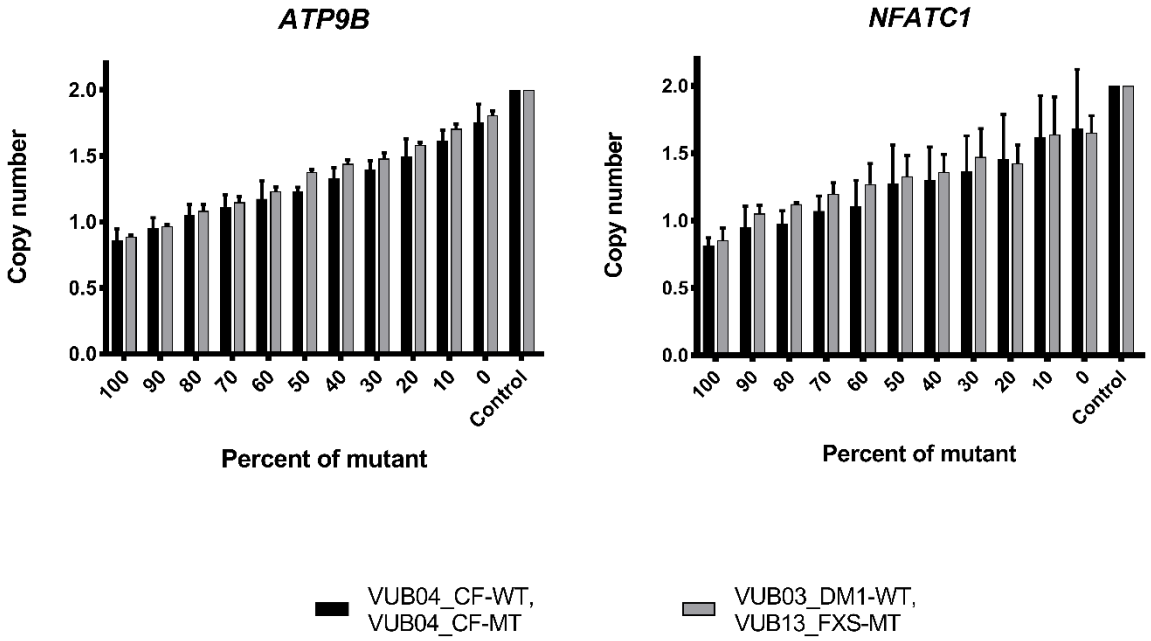


Figure 8: Detecting copy numbers in DNA mixes with different ratios of WT and MT cells. Control=genomic DNA from leukocytes of a healthy donor.

4.1.2 cDNA RT-qPCR

Gene expression assays were chosen based on the RNA sequencing results that were previously obtained in the REGE laboratory. Genes *LINC00116*, *SERPINB3* and *TSPYL5* were expected to be downregulated in MT cells, while *ARPIN/C 15orf38-AP3S2* and *PLA2G2A* should be upregulated, both in the undifferentiated state, as well as after differentiation. Expression of these 5 genes was tested in WT and MT cells, both in the undifferentiated state and after 21 days of spontaneous differentiation (Figures 9, 10). We learnt that, for the supposedly upregulated genes, only the expression of the *PLA2G2A* gene was higher in MT than in WT, and only in the undifferentiated state (Figure 9). Genes *LINC00116* and *TSPYL5*, on the other hand, showed a consistent lower expression in MT when compared to WT cells, both in the undifferentiated state, as well as in the following spontaneous differentiation (Figure 10). The *SERPINB3* gene was only expressed in undifferentiated WT cells (Figure 9).

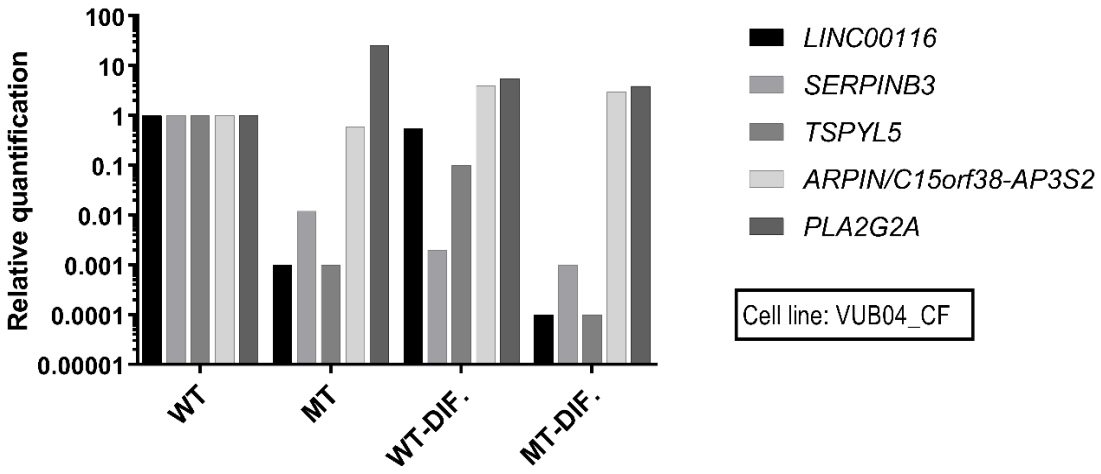


Figure 9: The results of testing different GEA on undifferentiated and spontaneously differentiated hESCs. WT= wild-type, MT= mutant

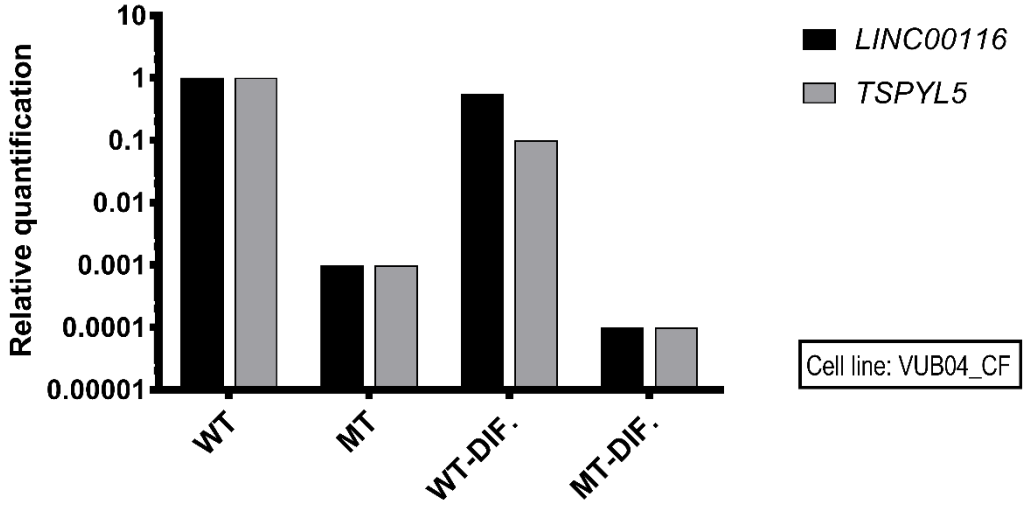


Figure 10: The expression of genes LINC00116 and TSPYL5 in WT and MT cells remained different after spontaneous differentiation. WT= wild-type, MT= mutant, DIF= differentiated cells

Because the *LINC00116* gene expression showed the biggest difference between WT and MT cells, even after 21 days of spontaneous differentiation, we further evaluated the sensitivity of GEA for *LINC00116* with different ratios of WT and MT cDNAs. RNA mixes were made from 0% to 10% MT, increasing by 10%, to evaluate the sensitivity of the assay. The assay proved sensitive enough to detect a difference in expression when a difference of at least 20% MT RNA was present (Figure 11).

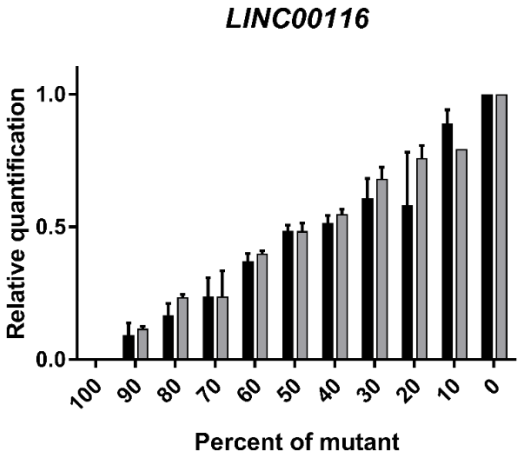


Figure 11: The expression of gene LINC00116 in cDNA mixes with different ratios of WT and MT cells.

4.2 CULTURE TAKE-OVER IN CASE OF UNDIFFERENTIATED hESCs

Two mixes with different ratios of WT and MT cells were established in cultures; the one with 90% WT and 10% MT and the other with 50% WT and 50% MT. They were kept and observed to see, whether the cells carrying the 18q deletion take over the culture at certain passage, and during spontaneous and directed differentiation.

4.2.1 COPY NUMBER

Mixed cultures (10% and 50% MT) were cultured for 13 passages. Potential take-over by aberrant cells was tested with two different copy number assays. Cell pellets were collected every other passage. Figures 12 and 13 show the results of two technical replicates for each experiment.

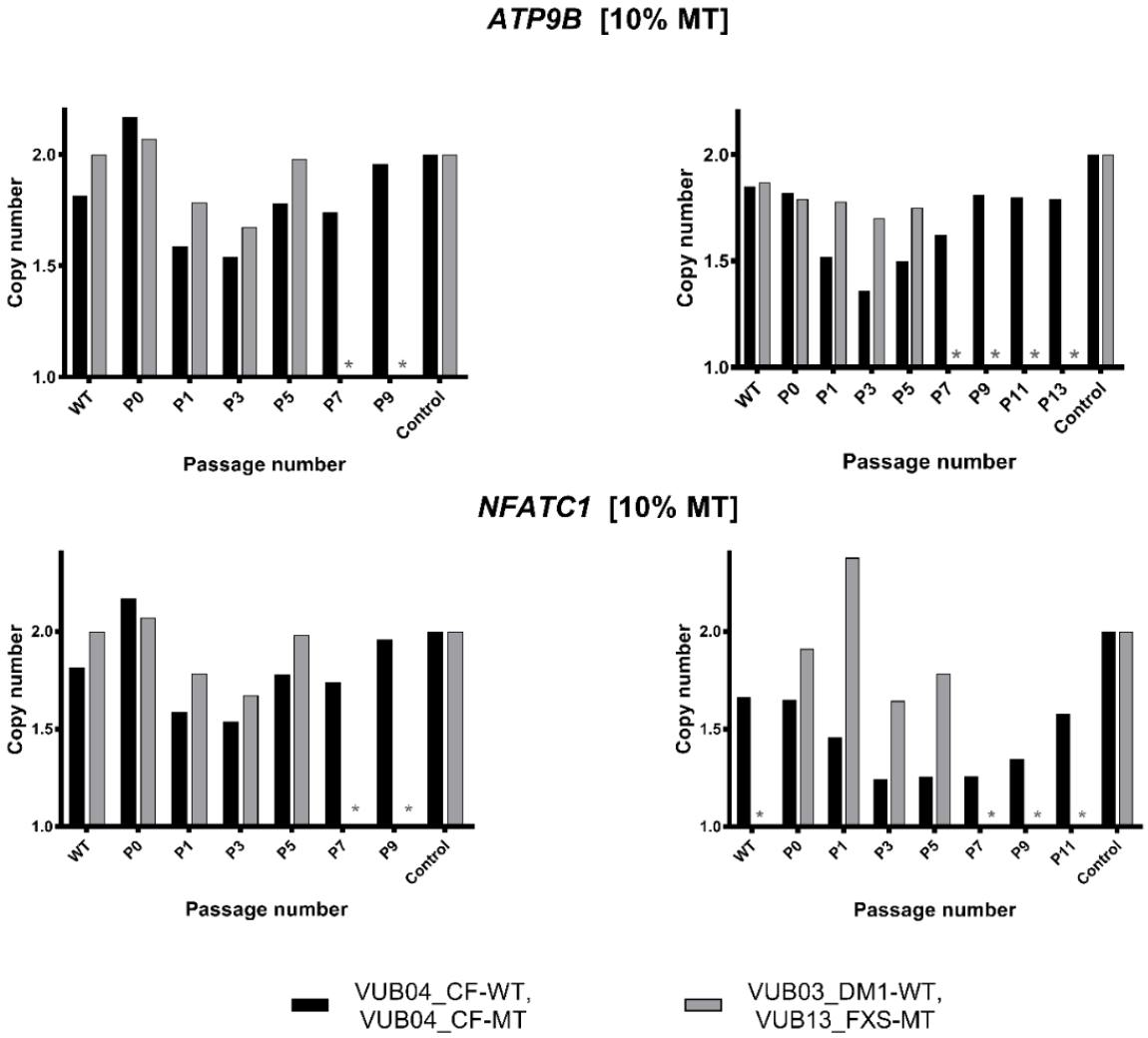


Figure 12: Progression of 18q deletion in hESC culture during 13 passages. Mixed cultures containing 10% MT cells were tested with CNA ATP9B and NFATC1. Data labeled with the asterisk (*) were missing or were not tested due to loss of pellet or poor DNA extraction yield. Control=genomic DNA from leukocytes of a healthy donor.

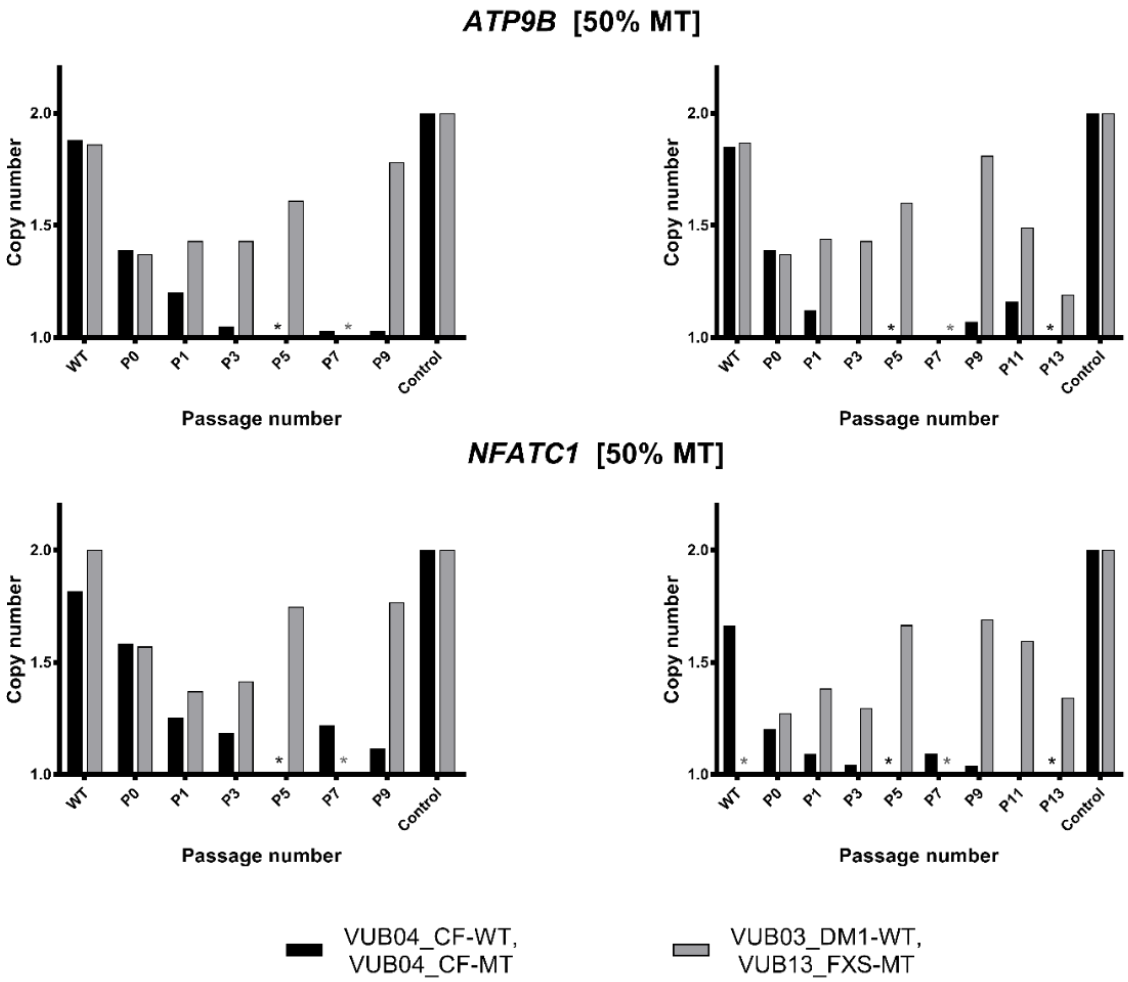


Figure 13: Progression of 18q deletion in hESC culture during 13 passages. Mixed cultures containing 50% MT cells were tested with CNA ATP9B and NFATC1. Data labeled with the asterisk (*) were missing or were not tested due to loss of pellet or poor DNA extraction yield. Control=genomic DNA from leukocytes of a healthy donor.

Graphs in Figures 12 and 13 show variations between passages with no significant decrease in copy numbers, except in the 50% MT cell mixes of cell line VUB04_CF. In this mix the copy number dropped during the first 3 passages and then changed minimally during the next 10 passages. However, the MT cell did not take over the culture entirely. The copy number in the mix containing 50% MT of the VUB13_FXS cell line was a bit lower due to higher amounts of MT cells, but it was still not consistent after 13 passages. There is also some inconsistency between two replicates, as in some experiments copy numbers appear to be higher than in others.

In some experiments, to our surprise, we observed an increase in copy numbers already at the first passage, thus we investigated whether there may be a plating advantage for the WT

cells over the MT cells, which biases the mix ratios upon starting the experiment. An experiment for plating advantage was therefore carried out to show that the observed variation in copy numbers is not a consequence of a better survival or easier attachment of WT or MT cells to Laminin-521, when the cells were passaged.

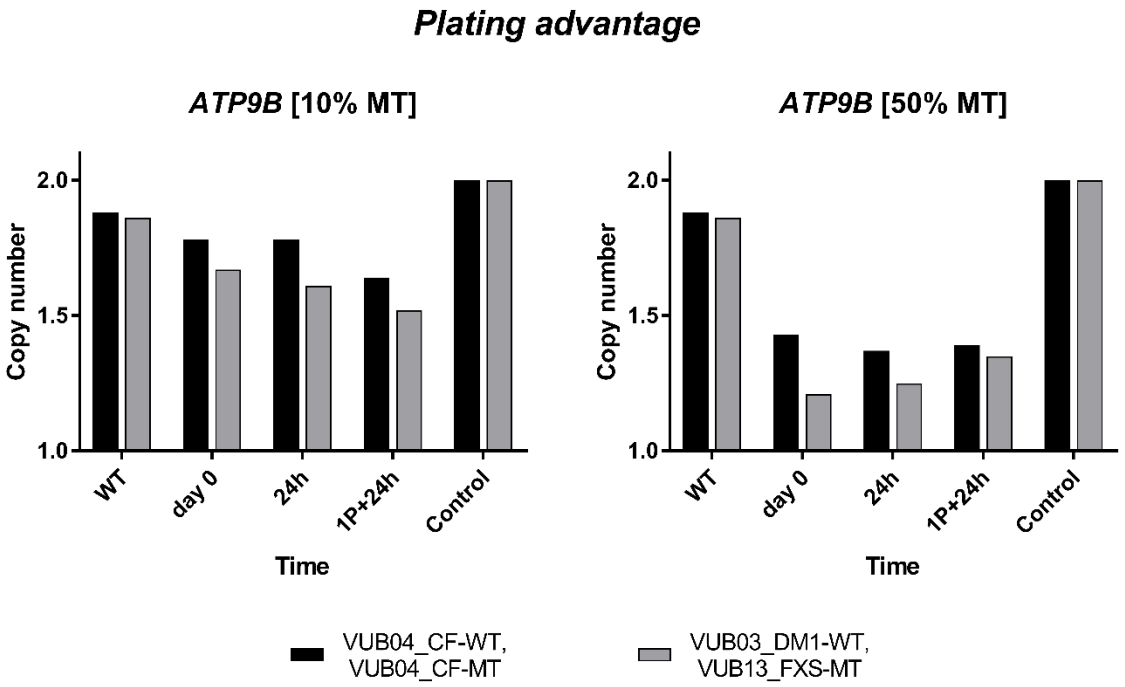


Figure 14: Plating advantage of cells in 10% MT and 50% MT mixes. Control=genomic DNA from leukocytes of a healthy donor.

In the mixes there was no plating advantage for WT or for MT, which means that cell survival and proliferation stays the same between passages. Cell pellet was collected on day 0, after 24 hours, after the first passage and then 24 hours after it. The extracted DNA was tested with CNA *ATP9B* (Figure 14).

4.2.2 GENE EXPRESSION

Mixes with 10% and 50% MT cells were kept in culture and were tested for *LINC00116* expression at every other passage. Graphs 15 and 16 show the results of two technical replicates.

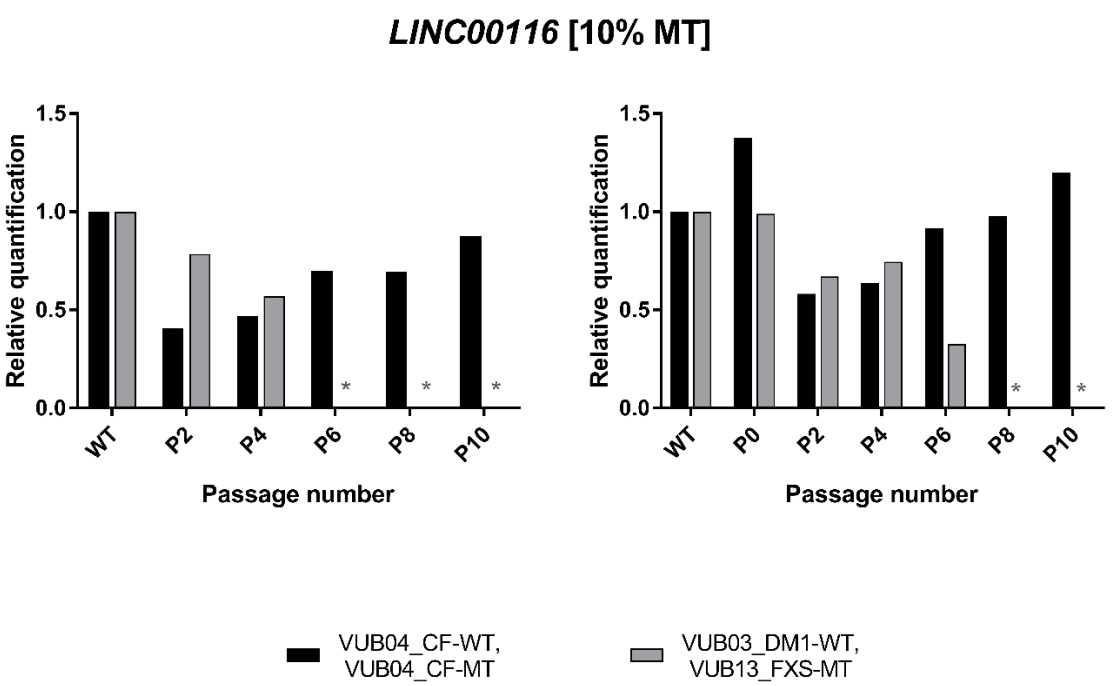


Figure 15 : Progression of 18q deletion in hESC culture during 10 passages. Mixed cultures containing 10% MT cells were tested with GEA LINC00116. Data labeled with the asterisk (*) were missing or not tested due to loss of pellet or poor RNA extraction yield.

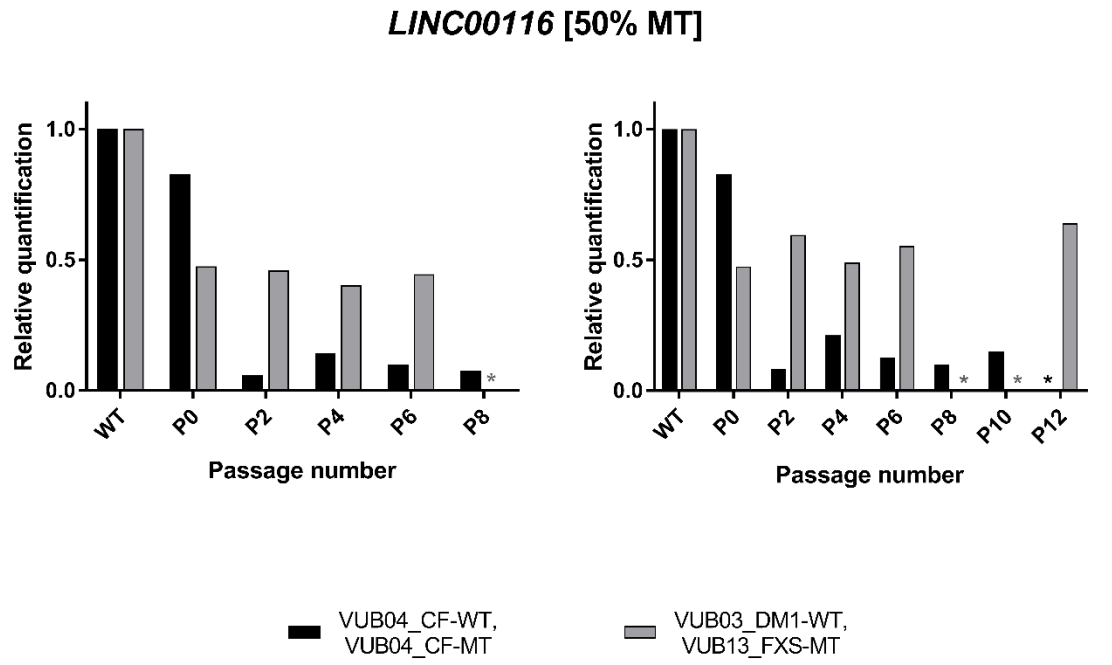


Figure 16: Progression of 18q deletion in hESC culture during 12 passages. Mixed cultures containing 50% MT cells were tested with GEA LINC00116. Data labeled with the asterisk (*) were missing or not tested due to loss of pellet or poor RNA extraction yield.

Variation of gene expression was present in mixes containing 10% MT, but MT cells did not seem to have selective advantage (Figure 15). When we compared the mixes with 50% VUB03 and 50% VUB13 cells, the gene expression stayed nearly the same after 12 passages, while in mixes containing 50% of VUB04-WT and 50% of VUB04-MT cells, we found an increase in MT cells, but they did not make a complete take-over of the culture (Figure 16).

4.3 PROGRESSION OF THE 18q DELETION DURING DIFFERENTIATION

4.3.1 SPONTANEOUS DIFFERENTIATION

Human ESCs will spontaneously differentiate once the factors for maintaining their pluripotent character are withdrawn from the culture. In this process, several signaling pathways and the cell cycle itself, are highly affected. Differential gene expression, due to chromosomal aberration may lead to a specific culture (dis)advantage of cells carrying this aberration. A possible take-over of MT cells was monitored with CNA in mixed cultures, during their major shift to more specialized cells.

In the first experiment (left graphs in Figures 17 - 20) the cell pellets were collected on days 0, 10 and 21, while in the second experiment (right graphs in Figures 17 - 20) the cell pellets of the cell line VUB04_CF were collected on days 0, 10 and 16, and those of the cell lines VUB3_DM1 and VUB13_FXS on days 0, 9 and 18.

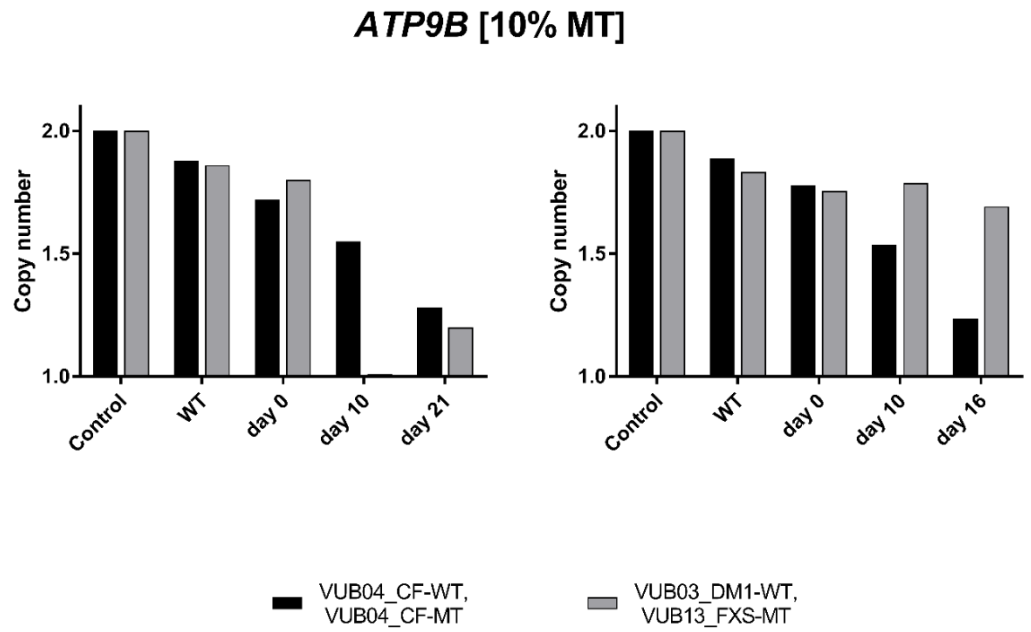


Figure 17: Change in copy numbers during spontaneous differentiation. Mixed cultures containing 10% MT cells were tested with CNA ATP9B. Control=genomic DNA from leukocytes of a healthy donor.

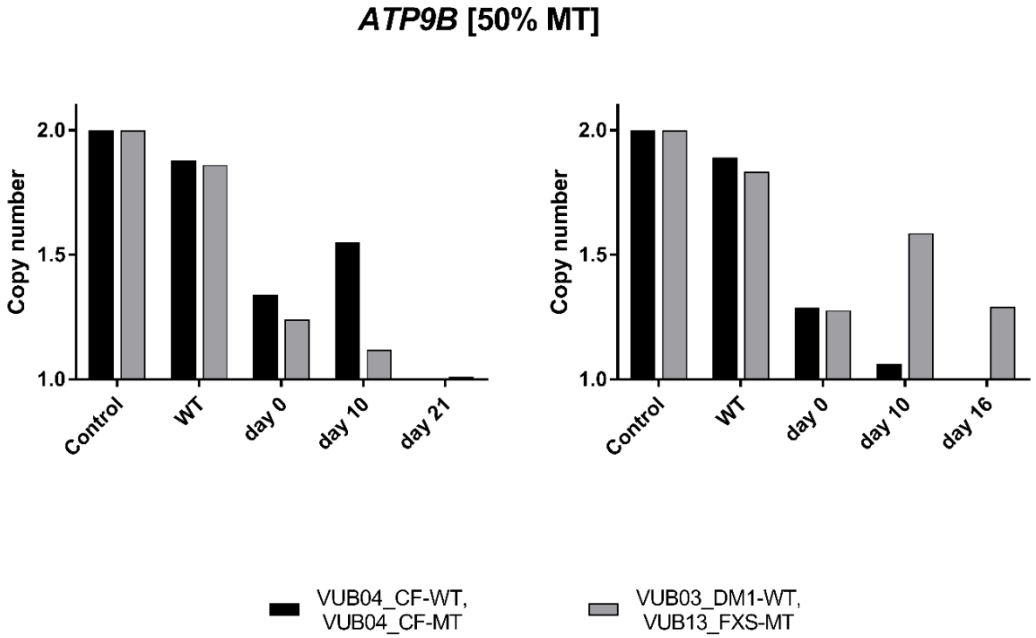


Figure 18: Change in copy numbers during spontaneous differentiation. Mixed cultures containing 50% MT cells were tested with CNA ATP9B. Control=genomic DNA from leukocytes of a healthy donor.

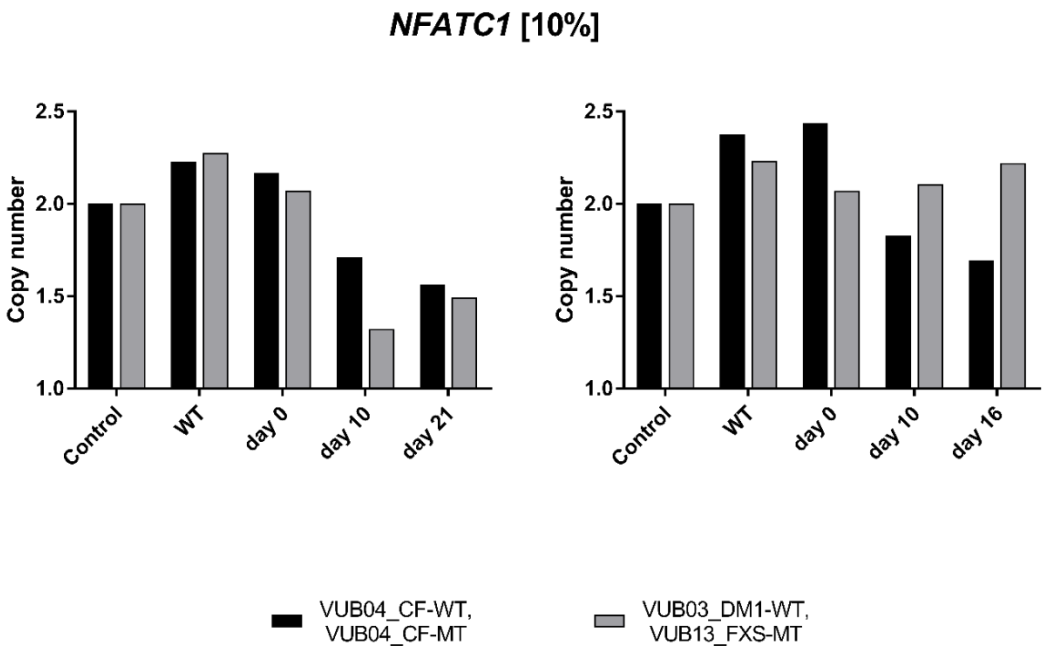


Figure 19: Change in copy numbers during spontaneous differentiation. Mixed cultures containing 10% MT cells were tested with CNA NFATC1. Control=genomic DNA from leukocytes of a healthy donor.

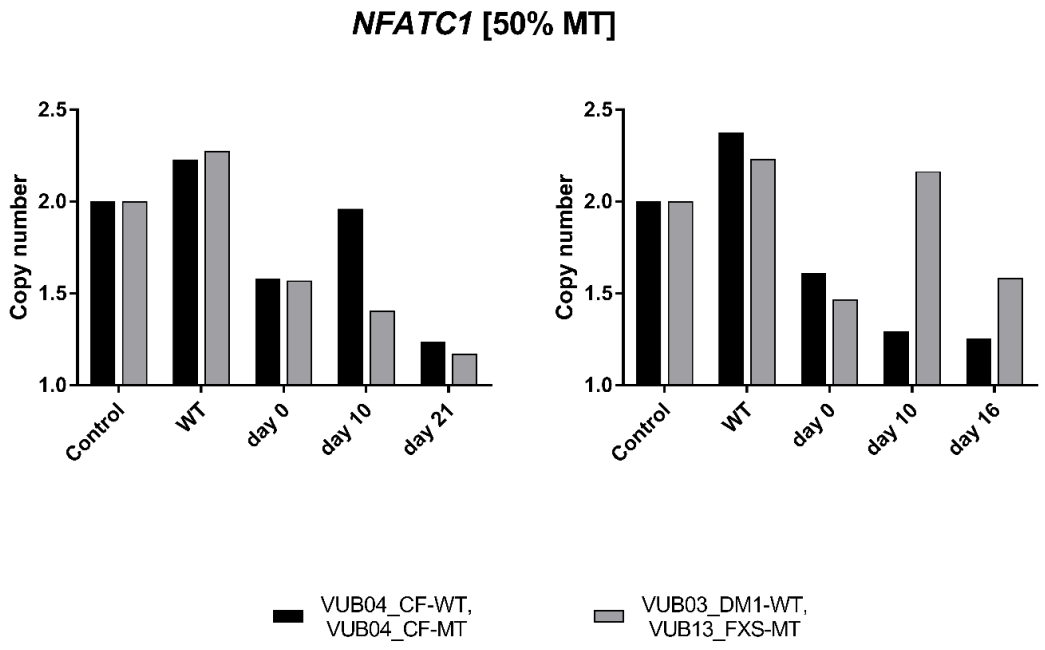


Figure 20: Change in copy numbers during spontaneous differentiation. Mixed cultures containing 50% MT cells were tested with CNA NFATC1. Control=genomic DNA from leukocytes of a healthy donor.

Figures 17 - 20 overall show a trend of decreasing copy numbers during spontaneous differentiation, implying that MT cells, at least at a certain extent, took over the culture. In some experiments with cell lines VUB03_DM and VUB13_FXS, however, no significant decrease in copy numbers, after 18 days of spontaneous differentiation, was observed. In

some cases, even higher copy numbers on day 10, as compared to day 0, were observed, which decreased at the endpoint of differentiation. In general, we did not detect a consistent downward trend of copy numbers.

Gene expression was also monitored during spontaneous differentiation. The two graphs in Figures 21 and 22 represent two biological replicates. Cell pellets were collected on the same days as in the case of copy number determination assays.

The expression of *LINC00116* was lower after 10 and 21 days of spontaneous differentiation, although it has to be taken into account that WT cells already show lower expression after differentiation, as presented in Figure 6. However, a decrease in expression of both mixes was significant (Figures 21 and 22). Differentiation was also observed at the phenotype level, as the cells became rounder and darker in the center, as seen in Figure 23.

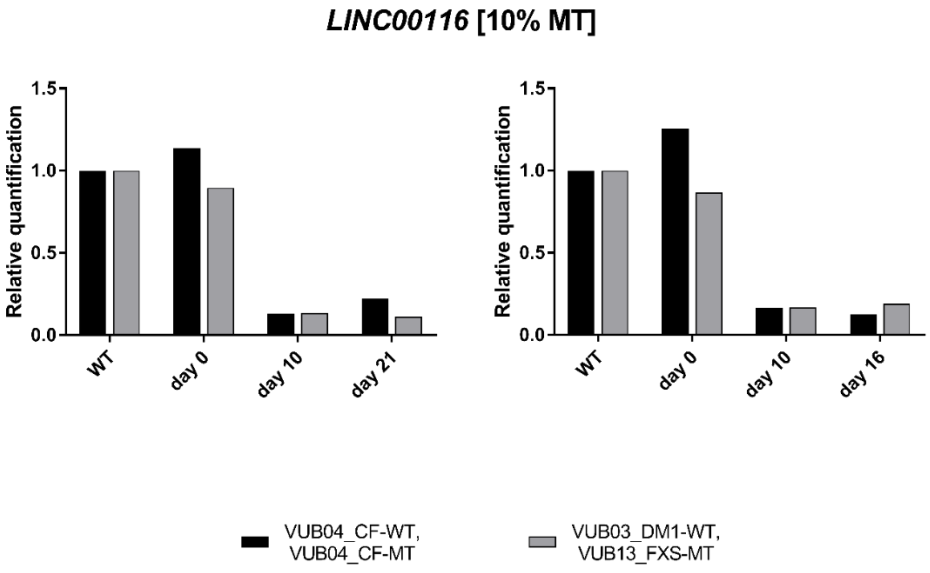


Figure 21: LINC00116 expression in mixed cultures containing 10% of MT cells during spontaneous differentiation.

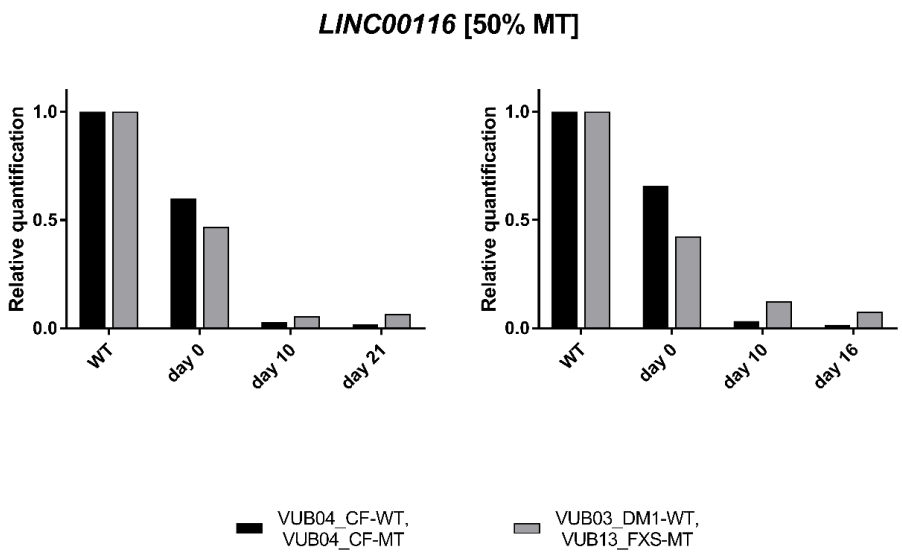


Figure 23: LINC00116 expression in mixed cultures containing 50% of MT cells during spontaneous differentiation.

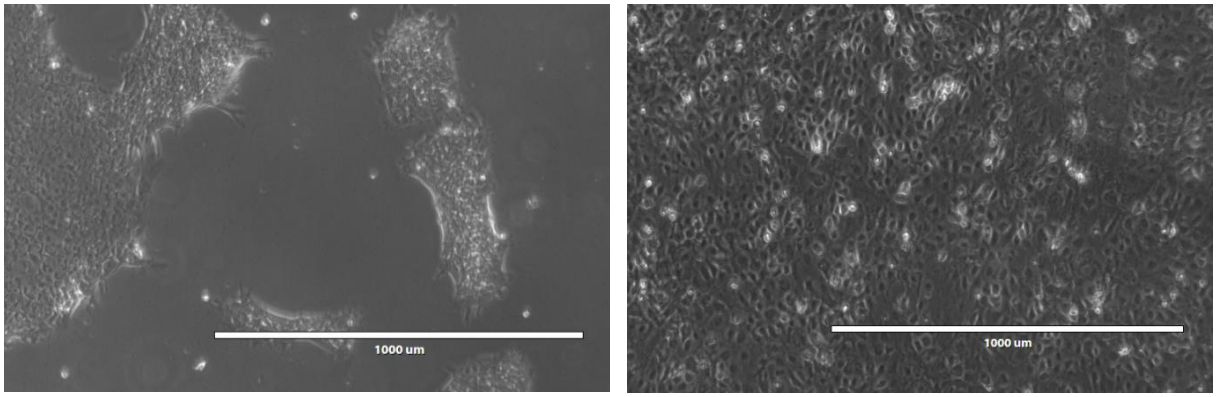


Figure 22: VUB03_DM1 cell line at passage 16, before (left) and after (right) spontaneous differentiation at passage 4.

4.3.2 HEPATOBLAST-DIRECTED DIFFERENTIATION

In contrast to spontaneous differentiation, during directed differentiation, the cells are forced to differentiate towards a specific lineage. This can expose them to different stress factors and bottlenecks with additional factors playing a role in a possible advantage during cell culture. We wanted to see if the MT cells will take over the culture during directed differentiation towards the hepatic lineage (the lineage that is studied more in detail in the lab). The differentiation to hepatocyte progenitor cells was carried out with WT and MT cells and their mixes (10% and 50% of MT cells) and was evaluated by CNA *ATP9B* on days 0 and 8 (Figures 24 and 25).

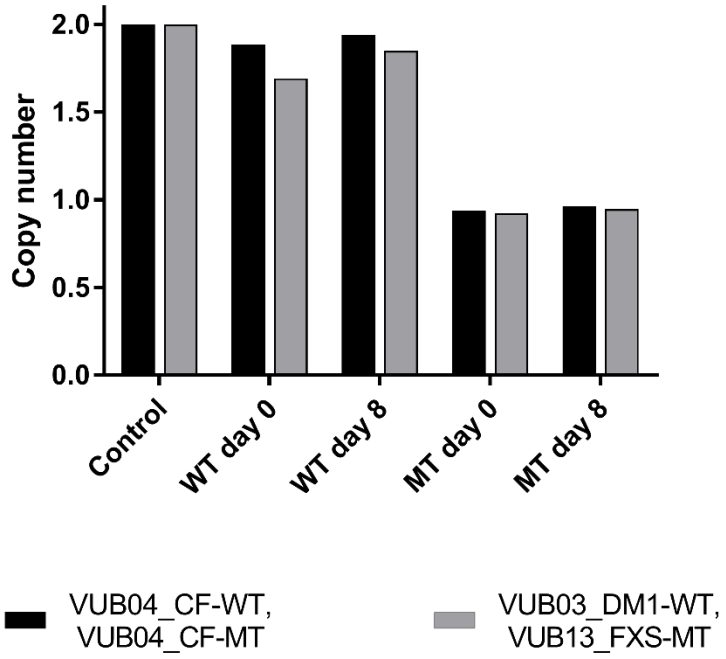


Figure 24: WT and MT cell lines before and after 8 days of directed hepatoblast differentiation. Copy numbers were detected with CNA ATP9B. Control=genomic DNA from leukocytes of a healthy donor.

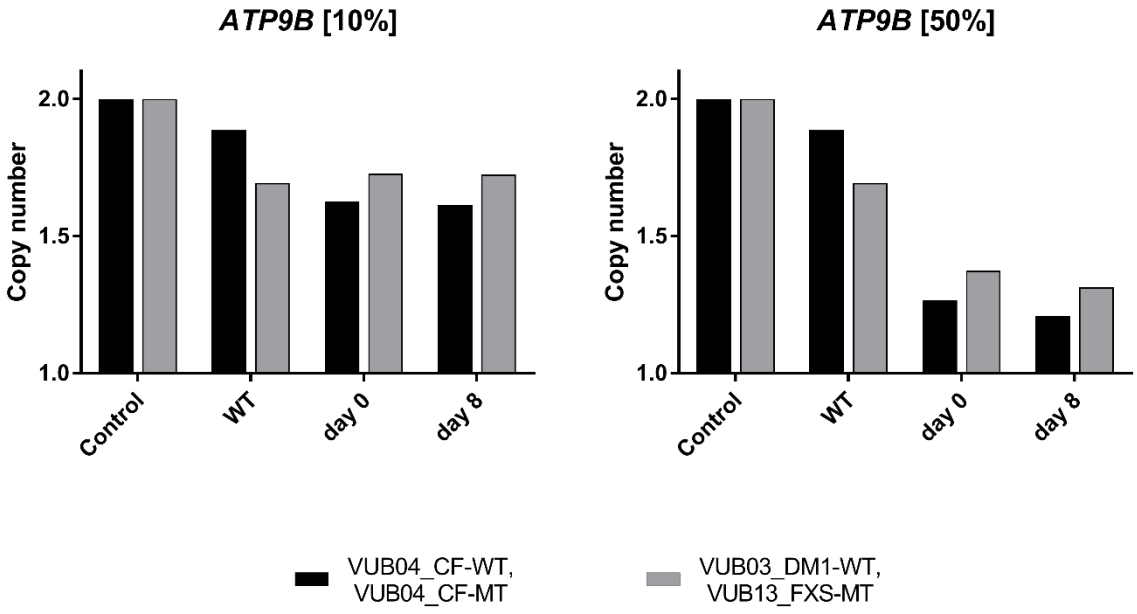


Figure 25: Mixes containing 10% and 50% of MT cells before and after directed hepatoblast differentiation. Copy numbers were detected with CNA ATP9B. Control=genomic DNA from leukocytes of a healthy donor.

Copy numbers of *ATP9B* gene in cultures of WT and MT cells remained stable during the 8 days of directed hepatoblast differentiation (Figure 24). Also, in our experiments with mixes, the copy numbers remained unaltered regardless of the ratio between WT and MT cells,

suggesting that there was no selective advantage for either the MT or the WT cells (Figure 25).

In case of *LINC00116* gene expression, we determined that it practically halved after 8 days of directed hepatoblast differentiation in WT cells. MT cells, on the other hand, showed no expression of *LINC00116*, both in the undifferentiated state, as well as after differentiation (Figure 26). However, it is important to note that this experiment was performed only once and with only one cell line, due to lack of time.

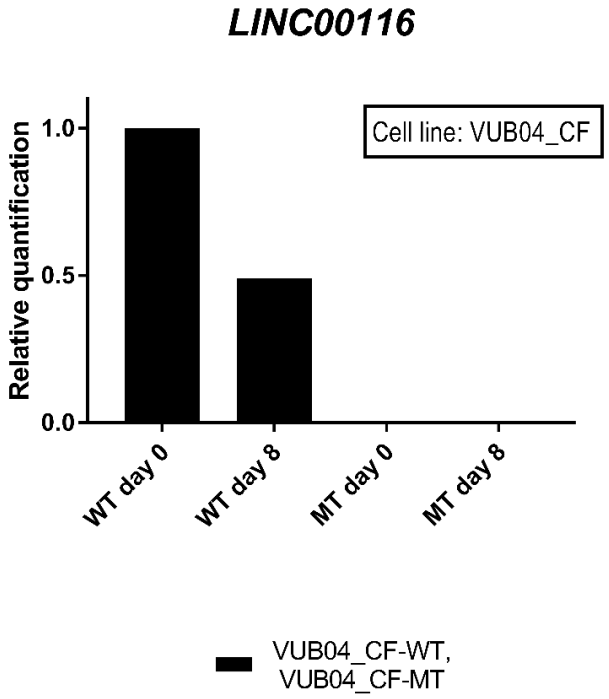


Figure 26: Expression of LINC00116 in WT and MT hESCs and hepatoblasts.

LINC00116 gene expression in mixes decreased in a similar way as observed in pure WT cell cultures, suggesting that the ratio of WT in the mixes remained similar, which means that there was no selective advantage for cells with 18q deletion (Figure 27).

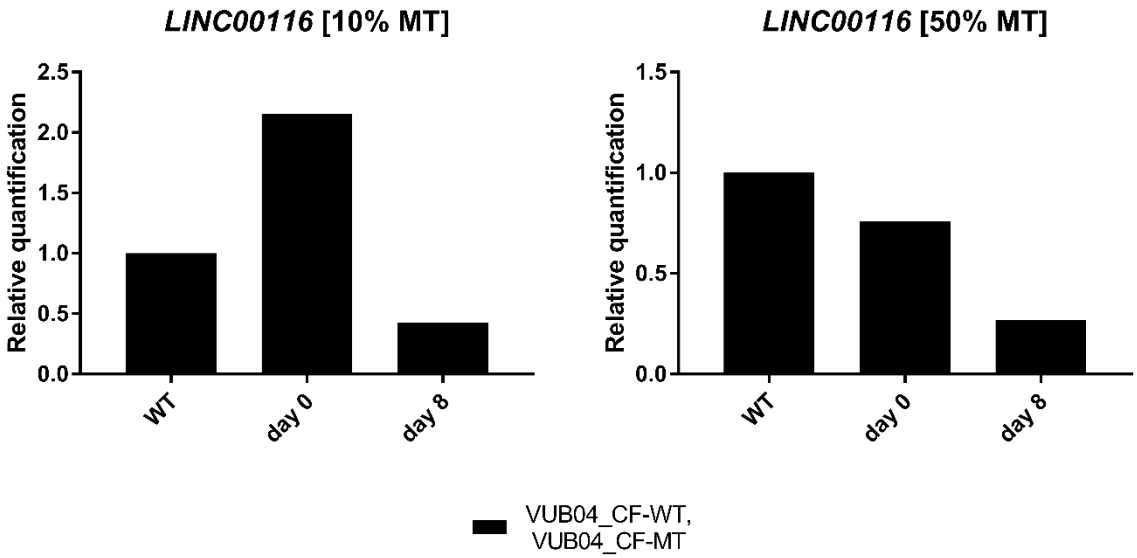


Figure 27: Expression of LINC00116 in cell culture mixes containing 10% and 50% of MT cells, before and after directed hepatoblast differentiation.

4.4 IMMUNOCYTOCHEMISTRY

Apart from monitoring gene expression at DNA and RNA levels, we were also hoping to be able to discriminate between WT and MT at the protein level, by using immunofluorescent microscopy, an automated cell counter or flow cytometry. Moreover, this would allow us to more precisely quantify WT and MT cells within a mosaic culture. Of the markers identified at the RNA level, we chose TSPYL5 as a target for our protein level studies. TSPYL5 is testis specific Y-like 5 protein, expressed in the cytoplasm.

4.4.1 FLUORESCENCE MICROSCOPY

As the cytoplasm of cells was GFP (green fluorescent protein) positive in both WT and MT cells in their undifferentiated state as well as after differentiation, TSPYL5 was unsuitable as a marker to discriminate between WT and MT cells by immunofluorescence (Figure 28).

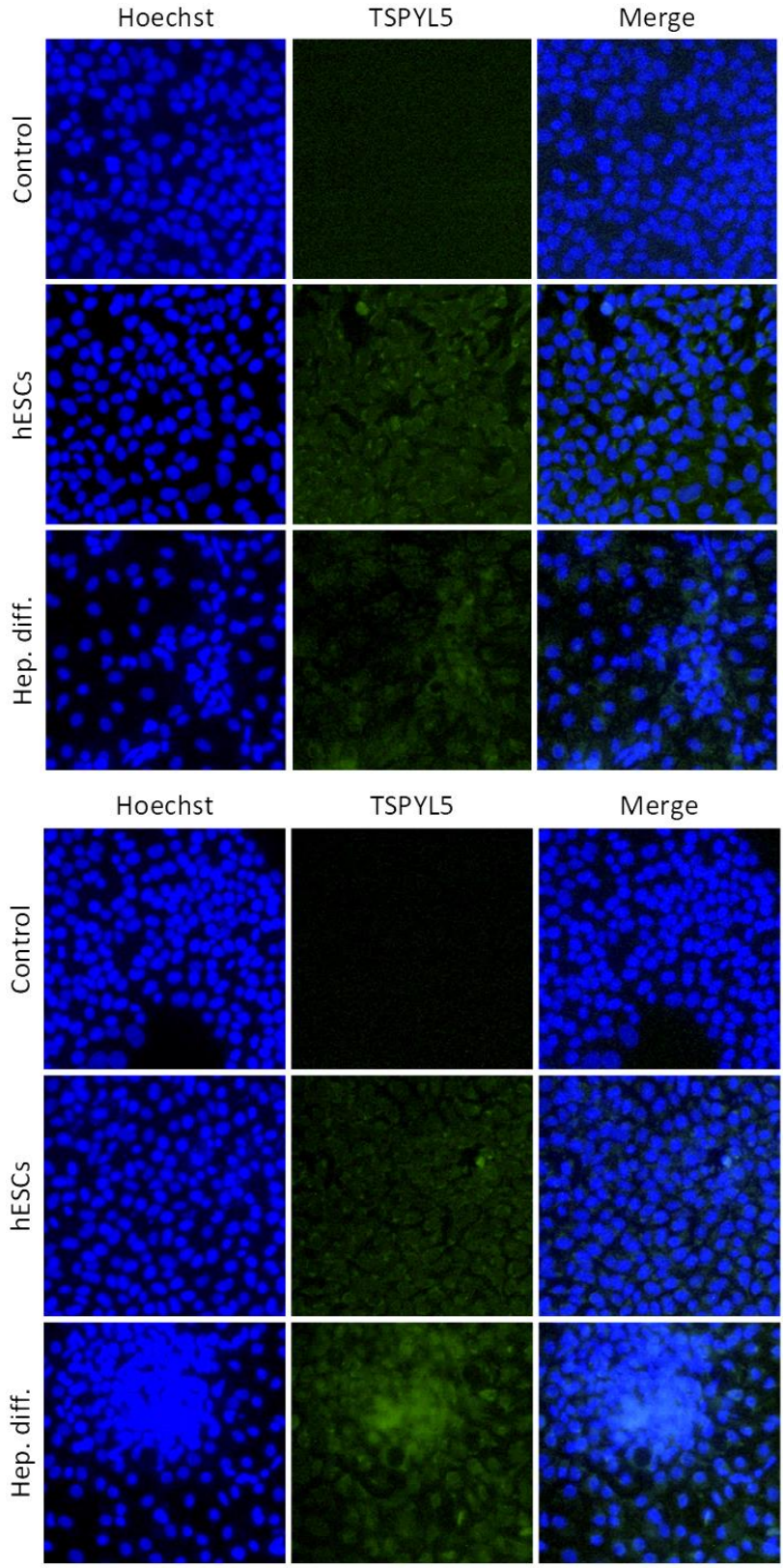
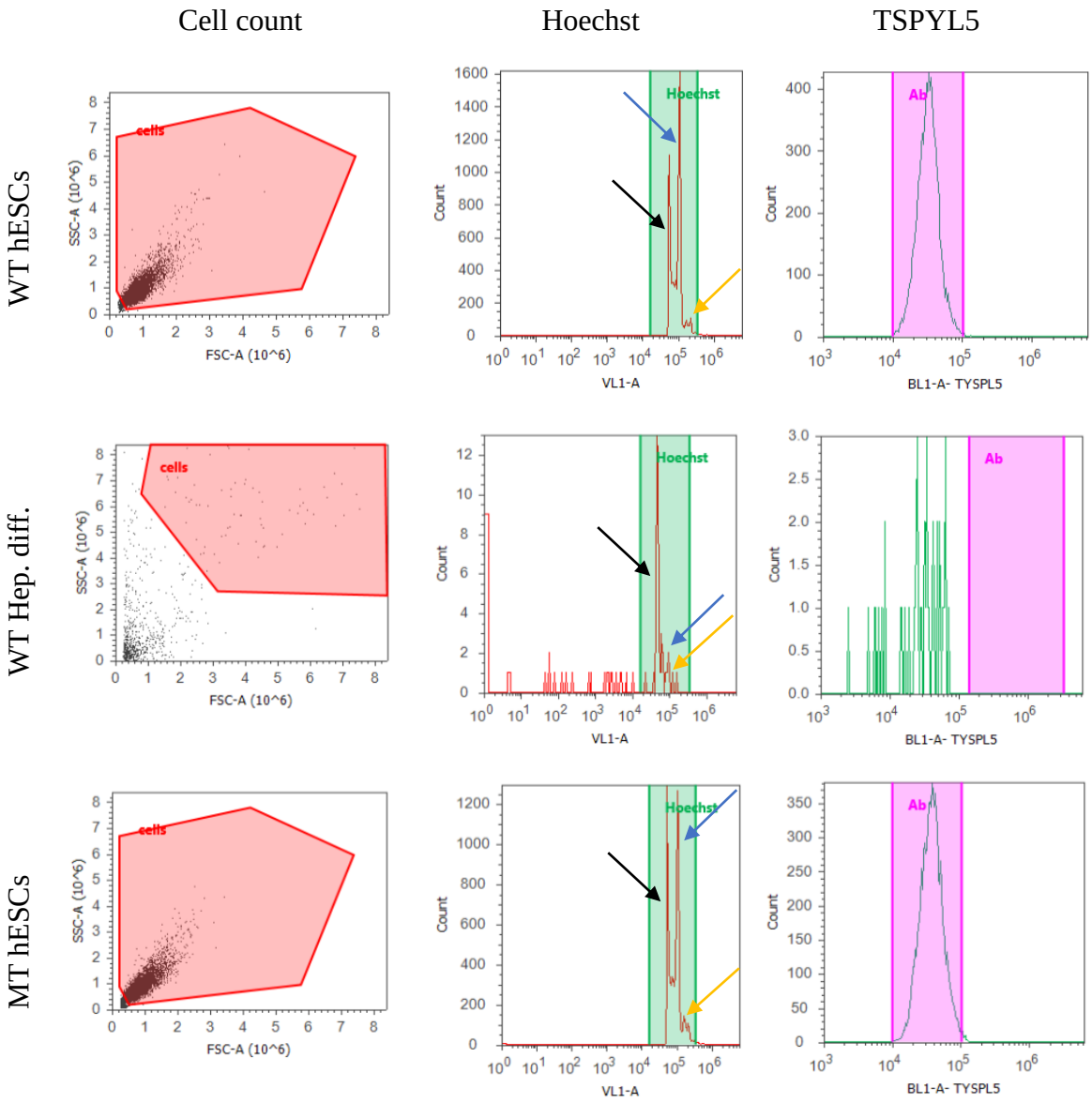


Figure 28: Immunostaining of VUB04_CF WT (top) and VUB04_CF MT (bottom). Position of cell nuclei was determined by using Hoechst staining, while the antibody for protein TSPYL5 labels the cytoplasm of cells with green fluorescence. The experiment was done on hESCs and differentiated cells of both WT and MT type. Control was stained only with Hoechst stain. Hep. diff. = hepatoblast differentiation

4.4.2 FLUORESCENCE-ACTIVATED CELL SORTING (FACS)

Flow cytometry was used to identify and differentiate between WT and MT cells in their undifferentiated and differentiated states (Figure 29). It has a higher sensitivity and specificity than fluorescence microscopy techniques.



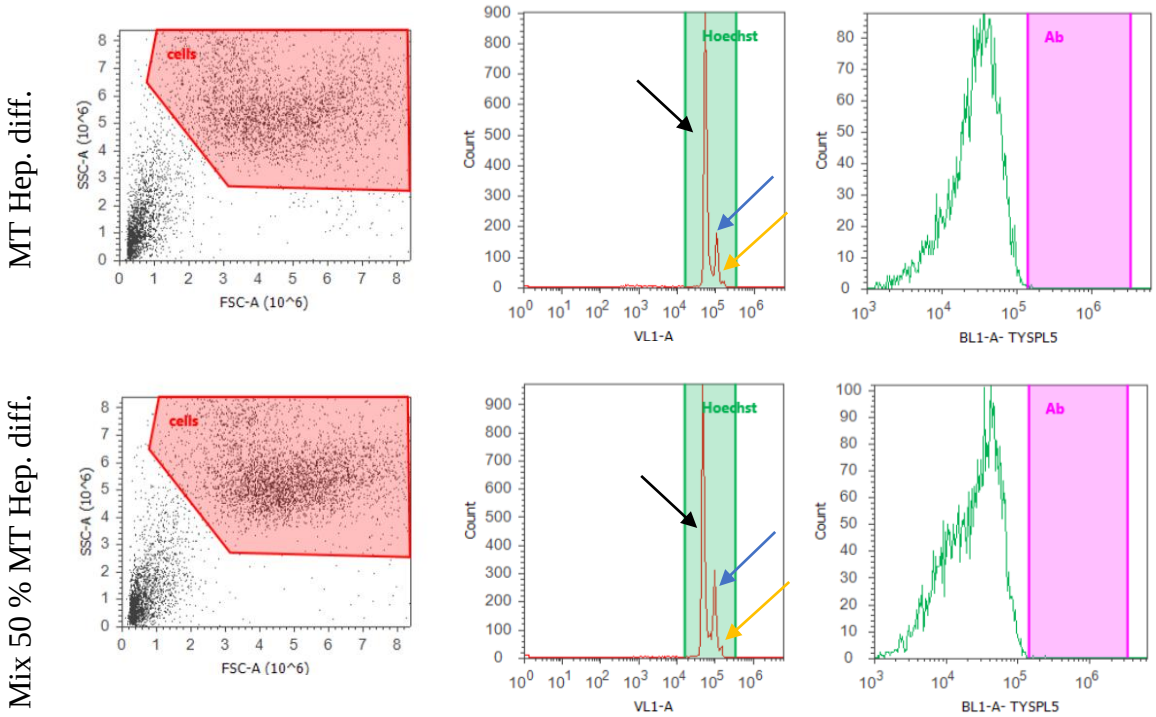


Figure 29: Cell counts assessed by flow cytometry, based on cell size (left) and fluorescence of Hoechst stain (middle) and Alexa Fluor 488 (right). Arrows are denoting cell cycle phases; G1-black, S- blue, G2-yellow. Experiment was carried out on hESCs, differentiated WT and MT cells and a mix containing 50% of MT cells.

Nuclear staining with Hoechst stain enabled the determination of proportion of cells in each phase of a cell cycle. ESCs usually exhibit short G1 and G2 phases, while most of the cells are in the S phase. This is true for WT hESCs. However, the amount of MT hESCs was the same in G1 and S phase. When ESCs begin to differentiate and become more specialized, the G1 phase gets longer (Figure 29). Fluorescence for TSPYL5 protein produced a high sharp peak in both, the WT and the MT cells. The peak got wider upon differentiation. However, it was impossible to discriminate between WT or MT cells for the expression of the TSPYL5 protein by flow cytometry in our experiments.

4.4.3 TALI IMAGE-BASED CYTOMETRIC RESULTS

Fluorescence was also measured with the image-based cytometer (TALI) to determine the amount of WT and MT cells in cell suspensions. The results are presented in Table V.

Table V: Numbers and percentages of cells detected with TALI, based on the green fluorescence emitted by the labeled antibody against TSPYL5.

Sample	GFP positive [cells/ml]	GFP positive [%]
Control	0	0
WT undiff.	$7{,}42\times10^4$	44
WT diff.	$2{,}57\times10^4$	41
MT undiff.	$4{,}72\times10^4$	10

In this experiment, we identified lower GFP positivity in MT hESCs than in WT hESCs. However, this experiment was done only once with a small number of cells and with the technique which is less specific than flow cytometry. The results should therefore be further validated before making any conclusions.

5 DISCUSSION

Human embryonic stem cells (hESCs) hold great promise in clinical and research fields, especially in regenerative medicine and drug development. However, their use in routine practice is still not within reach, since numerous studies are still needed to be done, in order to guarantee their quality and safety. Karyotypic abnormalities of these cells are one of the crucial obstacles. They have been reported to occur in long term cultures of hESCs. Deletion of a part of the long arm of chromosome 18 has recently been spotted to recurrently appear in hESC cultures. It is therefore believed that this abnormality may confer a selective advantage to abnormal cells during culture. The research presented in this thesis aimed to explore how this mutation affects hESCs in an updated feeder-free culture system and whether the mutant (MT) cells display a selective advantage in comparison to wild-type (WT) ones, upon their spontaneous and directed differentiation.

Based on the literature data, qPCR has a detection limit for mosaicism at around 5-10%, which is precise enough for our experiments.[27] Two copy number assays for genes *ATP9B* and *NFATC1* were carried out in order to detect amounts of WT and MT cells in cultures. Both genes are present in the 18q region and upon deletion only one copy can be detected. However, in contrast to the published thresholds, in our experiments a minimum of 20% difference in proportions of MT cells was needed for reliable differentiation between various mosaic cultures.[27] Similar results were also obtained in *LINC00116* gene expression assays. *LINC00116* is a long intergenic non-protein coding RNA, present on chromosome 2, and its expression is influenced by the 18q deletion. We hoped to be able to find an even more sensitive assay to score mosaicism at the protein level. For this purpose, we used a fluorescent antibody directed against cytoplasmic protein TSPYL5 for immunocytochemical staining, detected with fluorescence microscope, FACS and TALI. However, these experiments will still need further optimization.

Mixes of WT and MT cells were made to mimic mosaicism that can naturally occur in cultured hESCs. The mixes containing the same amounts of WT and MT cells created equal conditions for both cell types in order to study possible selective advantage of karyotypically abnormal cells. For the mixes containing only 10% of MT cells, it is important to note that, while the sensitivity of our PCR-based assays (20%) may not be able to reliably detect the 10% mutant cells at the start of the experiment, it would certainly have been enough to detect

a culture take-over. Moreover, a mosaic culture containing 10% or less MT cells is probably the most realistic condition. Cultures with detectable abnormalities would most probably not be used for applications, while those where the abnormalities remain under the detection limit, could be considered safe, though the abnormal cells, if taking over the culture in a further stage, could have an important influence on the applicability of the culture for clinical purposes.

5.1 hESCs IN CULTURE

It has been previously reported that hESCs with derivative chromosome 18, gain selective advantage over wild-type cells when cultured on MEFs[32], so we wanted to check if the same applies to feeder cell-free culture conditions. In our cell mixing experiments with different ratios of MT cells (10% and 50%), hESCs were monitored during 13 passages, growing on Laminin-521 in a defined, xeno- and serum-free medium. Although, cell cultures containing only 10% of MT cells at the beginning of the experiment, showed inconsistency in WT: MT ratios between passages, it did not appear that aberrant cells took over the culture. Mixes with 50% of VUB13_FXS mutant cell line and 50% of VUB03_DM1 wild-type cell line, showed similar results, as in the case of mixes containing 10% MT cells. However, in the cell culture consisting of 50% VUB04_CF mutants, we observed a fast increase in MT cell portions during early passages, though later on these remained consistent and did not result in a complete culture take-over. These differences between different cell lines may be explained with different genetic trait and background. Namely, the VUB13_FXS cell line has been derived from an embryo, affected with fragile X syndrome, the VUB03-DM1 cells originate from an embryo affected with myotonic dystrophy type I and the wild-type cells carry a duplication in chromosome 20. All these factors could affect metabolism or proliferation rates, therefore contributing to a possible culture advantage of particular cells. Another important factor can also arise from the fact that in one case, we compared WT and MT cells of different cell lines (VUB03-DM1 and VUB13_FXS), while the other cell pair (WT and MT) belonged to the same cell line, i.e. VUB04_CF. Cell lines can have different proneness to DNA damage and can be affected by previous culture conditions, which makes direct comparisons between them very difficult. We also wondered why do the hESCs carrying 18q deletion take over the culture immediately when cultured on MEFs but not on Laminin-521. The reason might be in different passaging techniques applied in these two culture conditions.

Passaging in the laminin-based culture system is performed by using enzymes. After detachment, a single cell suspension is obtained, containing the whole population of cells that was present in the culture dish, even the ones that were damaged, senescent or in early stages of differentiation. In this case, single cell survival or differential attachment capacity of WT versus MT cells could result in a possibility of passaging lower amounts of MT or WT and could be the reason for observed variation in gene copy numbers and expression. However, with our plating advantage experiment we showed that neither WT nor MT had poorer attachment to laminin after passaging.

For MEF-based cultures, on the other hand, the colonies were mechanically cut into small pieces which were then transferred to a new dish. As the colonies to be cut, were chosen by the operator, better looking colonies may have been selected during this procedure. These, however, might be correlated to cells with the 18q deletion, apparently having lower epithelial-mesenchymal-transition at the borders of the colonies. In this case, the subsequent culture take-over may be unintentionally caused by the operator, rather than selected due to a real culture advantage conferred to the cells.

5.2 SPONTANEOUS AND DIRECTED DIFFERENTIATION

Progression of the 18q deletion during spontaneous and directed differentiation is important, to understand possible consequences of the presence of this aberration, for further use of hESCs in regenerative medicine. Since no data were published regarding this topic, our aim was to get some insight of the behavior of aberrant cells during their shift to more specialized cells. Mixes of hESCs (10% or 50% MT) cultured on Laminin-521 were exposed to a differentiation medium without growth factors, necessary for maintaining their pluripotency. As we look at results, we can acknowledge that even when starting with a low ratio of MT cells in the culture, there were higher numbers of aberrant cells present at the end of 21 days of spontaneous differentiation. This was always true for the VUB04_CF cell line, while in some experiments the ratios between VUB03_DM1 and VUB13_FXS MT cells remained unchanged. Survival of cells carrying the 18q deletion seems to increase during spontaneous differentiation, which leads to overgrowth of the chromosomally abnormal cells in culture. Many tumor-suppressor genes are located within this region of chromosome 18q, which could explain the observed selective advantage of cells carrying deletion in this locus. The

process of differentiation and decrease in tumor-suppressor genes in mutant cells, might trigger their malignant transformation. In few experiments, higher percentage of WT cells was detected by copy number assays on day 10 than on day 0, although by the day 21 aberrant cells were mostly present in a culture dish. This might be a consequence of different proliferation rates of WT and MT cells. The observed morphological changes reflect many biological processes occurring during the loss of pluripotency, which may affect the MT cells differently than the WT ones.

hESCs that undergo directed differentiation *in vitro* are exposed to high selective pressure, due to supplementation of the medium with factors, which alter signaling pathways and force the cells to differentiate towards a specific lineage. We expected that the directed hepatoblast differentiation would favour mutant cells and enhance their proliferation rate, as it was observed in spontaneous differentiation. However, the ratio of WT and MT cells in our mixes remained the same before and after 8 days of hepatoblast differentiation, regardless of the amount of MT cells present in the culture at the beginning of experiment. This indicates that changes in biological processes, during this differentiation protocol, did not affect the cells with the 18q deletion. The molecular signature imposed by this deletion might not be important for development of endoderm and can therefore pass the checkpoints within the cell cycle.

Our aim was also to find a protein marker that would allow us to detect cell culture mosaicism and differentiate between undifferentiated and differentiated WT and MT cells. Based on gene expression studies, the protein TSPYL5 was the most suitable candidate, because its RNA levels were much higher in WT than in MT cells. After applying anti-TSPYL5 antibodies, followed by fluorescence microscopy and flow cytometry, it became evident that we were unable to discriminate between WT and MT cells with this marker. TSPYL5 was positive in the cytoplasm of all cells which provoke the question how good the correlation between RNA expression and protein level really is. The concentration of a given protein in cells is regulated by numerous processes like post-transcription, translation and protein degradation rates, therefore many factors can contribute to the variability between RNA and protein levels.[37,38] The correlation between RNA and protein levels can be as low as 40%, depending on the protein.[37] This can explain why TSPYL5 turned out to be an improper protein marker for the cells carrying the 18q deletion.

5.3 LIMITATIONS AND FUTURE RESEARCH

Many challenges were met due to the limited time for the thesis research. For example, experiments should be repeated more times to prove reproducibility of the study. We could use more wild-type/mutant pairs from different cell lines, which would add to the significance of results. Finding and testing candidates for most suitable protein marker for hESCs with 18q deletion would also require more time. There are some technical limitations as collecting small number of cells or low yield of DNA and RNA extractions, which prevented us from continuously tracking changes in the cells.

In order to further advance this research, it would be important to use more biological and technical replicates, to determine standard deviations of results and thereby their statistical significance. Longer differentiation into hepatocytes could also be carried out to study the influence of directed differentiation, as well as differentiation to the other two germ layers (mesoderm and ectoderm) which follow different developmental trajectories. At the protein level, a more suitable biomarker should be found, which would allow us to differentiate between WT and MT cells in both, undifferentiated and differentiated states, and to detect the amounts of each cell type in culture, as well as use other visualization techniques, such as FISH or fluorescent RNA cytochemistry. Also, studies at the single cell level would be more reliable than bulk analysis, giving us better insight into the dynamic of WT and MT cells during their long-term culture and differentiation.

6 CONCLUSION

Human ESCs carrying a deletion in chromosome 18q were successfully identified in mosaic cultures at the genomic level. Changes in copy numbers of genes *ATP9B* and *NFATC1* from two to one, and a decrease in expression of *LINC00116* were indicators of higher numbers of MT cells present in culture. By using immunocytochemistry and FACS we discovered a low correlation between RNA and protein levels of the selected TSPY5 biomarker, therefore further research is needed to identify a suitable marker for detecting mutant hESCs on the protein level.

As previously discovered, the 18q deletion confers a selective advantage to hESCs, which then take over the *in vitro* culture when cultured on MEFs. In this study the cells were growing on Laminin-521 and the take-over was only observed in cultures containing 50% of MT and 50% of WT cells of the same hESC cell line. The reason for such different outcomes, under different culture conditions, may be the use of different passaging techniques.

Upon spontaneous differentiation, aberrant cells overgrew the cultures in most experiments, which suggests that the presence of tumor-suppressor genes on 18q loci may have an important role in determining the cell fate. Directed differentiation to hepatoblasts did not have any effect on cells carrying 18q deletion, as the ratio between WT and MT cells remained unchanged. However, more replicate experiments should be done before drawing any conclusions.

7 REFERENCES

1. Huang, G., Ye, S., Zhou, X., Liu, D. & Ying, Q.-L. Molecular basis of embryonic stem cell self-renewal: from signaling pathways to pluripotency network. *Cell. Mol. Life Sci.* **72,** 1741–57 (2015).

2. Keller, A. *et al.* Genetic and epigenetic factors which modulate differentiation propensity in human pluripotent stem cells. *Hum. Reprod. Update* 1–14 (2018). doi:10.1093/humupd/dmx042

3. Mayo Clinic Staff. Stem cells: What they are and what they do. (2018). Available at: https://www.mayoclinic.org/tests-procedures/bone-marrow-transplant/in-depth/stem-cells/art-20048117. (Accessed: 23rd November 2018)

4. The International Society for Stem Cell Research. Nine Things To Know About Stem Cell Treatments – A Closer Look at Stem Cells. Available at: https://www.closerlookatstemcells.org/stem-cells-medicine/nine-things-to-know-about-stem-cell-treatments/. (Accessed: 23rd November 2018)

5. Reubinoff, B. E., Pera, M. F., Fong, C.-Y., Trounson, A. & Bongso, A. Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro. *Nat. Biotechnol.* **18,** 399–404 (2000).

6. Amit, M. & Itskovitz-Eldor, J. Derivation and spontaneous differentiation of human embryonic stem cells. *J. Anat.* **200,** 225–32 (2002).

7. Vazin, T. & Freed, W. J. Human embryonic stem cells: derivation, culture, and differentiation: a review. *Restor. Neurol. Neurosci.* **28,** 589–603 (2010).

8. Baker, D. E. C. *et al.* Adaptation to culture of human embryonic stem cells and oncogenesis in vivo. *Nat. Biotechnol.* **25,** 207–215 (2007).

9. Yu Junying & Thomson James A. Embryonic Stem Cells | stemcells.nih.gov. Available at: https://stemcells.nih.gov/info/Regenerative_Medicine/2006Chapter1.htm. (Accessed: 6th May 2018)

10. Gomes, K. M. S. *et al.* Induced pluripotent stem cells reprogramming: Epigenetics and applications in the regenerative medicine. *Rev. Assoc. Med. Bras.* **63,** 180–189

(2017).

11. Liang, G. & Zhang, Y. Embryonic stem cell and induced pluripotent stem cell: an epigenetic perspective. *Cell Res.* **23,** 49–69 (2013).

12. Singh, V. K., Kalsan, M., Kumar, N., Saini, A. & Chandra, R. Induced pluripotent stem cells: applications in regenerative medicine, disease modeling, and drug discovery. *Front. cell Dev. Biol.* **3,** 2 (2015).

13. Borowski, M. *et al. Basic pluripotent stem cell culture protocols*. *StemBook* (Harvard Stem Cell Institute, 2008).

14. Desai, N., Rambhia, P. & Gishto, A. Human embryonic stem cell cultivation: historical perspective and evolution of xeno-free culture systems. *Reprod. Biol. Endocrinol.* **13,** 9 (2015).

15. Duvillié, B. *et al.* Phenotypic alterations in insulin-deficient mutant mice. *Proc. Natl. Acad. Sci. U. S. A.* **94,** 5137–40 (1997).

16. Villa-Diaz, L. G., Ross, A. M., Lahann, J. & Krebsbach, P. H. Concise review: The evolution of human pluripotent stem cell culture: from feeder cells to synthetic coatings. *Stem Cells* **31,** 1–7 (2013).

17. Mouse embryonic fibroblast. Available at: https://en.wikipedia.org/wiki/Mouse_embryonic_fibroblast.

18. Keller, G. Embryonic stem cell differentiation: Emergence of a new era in biology and medicine. *Genes Dev.* **19,** 1129–1155 (2005).

19. Dvash, T. *et al.* Temporal gene expression during differentiation of human embryonic stem cells and embryoid bodies. *Hum. Reprod.* **19,** 2875–2883 (2004).

20. Odorico, J. S., Kaufman, D. S. & Thomson, J. A. Multilineage Differentiation from Human Embryonic Stem Cell Lines. *Stem Cells* **19,** 193–204 (2001).

21. Naujok, O., Diekmann, U. & Lenzen, S. The Generation of Definitive Endoderm from Human Embryonic Stem Cells is Initially Independent from Activin A but Requires Canonical Wnt-Signaling. *Stem Cell Rev. Reports* **10,** 480–493 (2014).

22. Zeineddine, D., Hammoud, A. A., Mortada, M. & Boeuf, H. The Oct4 protein: more

than a magic stemness marker. *Am. J. Stem Cells* **3,** 74–82 (2014).

23. Thomson, J. A. *et al.* Embryonic stem cell lines derived from human blastocysts. *Science* **282,** 1145–7 (1998).

24. Characterization of human embryonic stem cell lines by the International Stem Cell Initiative. *Nat. Biotechnol.* **25,** 803–816 (2007).

25. Nguyen, H. T., Geens, M. & Spits, C. Genetic and epigenetic instability in human pluripotent stem cells. *Hum. Reprod. Update* **19,** 187–205 (2013).

26. Lund, R. J., Närvä, E. & Lahesmaa, R. Genetic and epigenetic stability of human pluripotent stem cells. *Nat. Rev. Genet.* **13,** 732–744 (2012).

27. Baker, D. *et al.* Detecting Genetic Mosaicism in Cultures of Human Pluripotent Stem Cells. *Stem Cell Reports* **7,** 998–1012 (2016).

28. Spits, C. *et al.* Recurrent chromosomal abnormalities in human embryonic stem cells. *Nat. Biotechnol.* **26,** 1361–1363 (2008).

29. de Melo, J., Peng, G.-H., Chen, S. & Blackshaw, S. The Spalt family transcription factor Sall3 regulates the development of cone photoreceptors and retinal horizontal interneurons. *Development* **138,** 2325–2336 (2011).

30. Tanaka, T. *et al.* Chromosome 18q deletion and Smad4 protein inactivation correlate with liver metastasis: a study matched for T- and N- classification. *Br. J. Cancer* **95,** 1562–1567 (2006).

31. International Stem Cell Initiative, T. I. S. C. *et al.* Screening ethnically diverse human embryonic stem cells identifies a chromosome 20 minimal amplicon conferring growth advantage. *Nat. Biotechnol.* **29,** 1132–44 (2011).

32. Nguyen, H. T. *et al.* Human embryonic stem cells show low-grade microsatellite instability. *Mol. Hum. Reprod.* **20,** 981–989 (2014).

33. Chung, T.-L. Genetic and Epigenetic Instability of Human Pluripotent Stem Cells. *Open Stem Cell J.* **3,** 52–61 (2011).

34. Jacobs, K. *et al.* Low-grade chromosomal mosaicism in human somatic and embryonic stem cell populations. *Nat. Commun.* **5,** 1–10 (2014).

35. Lund, R. J., Närvä, E. & Lahesmaa, R. Genetic and epigenetic stability of human pluripotent stem cells. *Nat. Rev. Genet.* **13,** 732–744 (2012).

36. Ben-David, U., Mayshar, Y. & Benvenisty, N. Virtual karyotyping of pluripotent stem cells on the basis of their global gene expression profiles. *Nat. Protoc.* **8,** 989–997 (2013).

37. Cenik, C. *et al.* Integrative analysis of RNA, translation, and protein levels reveals distinct regulatory variation across humans. *Genome Res.* **25,** 1610–21 (2015).

38. Vogel, C. & Marcotte, E. M. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat. Rev. Genet.* **13,** 227–32 (2012).