Amplification of chromosomal region 20q11.21 provides a selective advantage in human embryonic stem cells: effects on growth, differentiation and further genomic instability.

Adam James Thomson Hirst

Submitted in fulfilment for the degree of doctor of philosophy

Department of Biomedical Sciences

September 2014

Abstract

Human embryonic stem (ES) cells are derived from the inner cell mass of blastocyst stage embryos. Once explanted and culture *in vitro,* human ES cells retain their pluripotent potential (i.e capacity to differentiate into all somatic cells) and acquire the ability to self-renew indefinitely. These two properties of human ES cells make them an invaluable resource for developmental biology, cell replacement therapies, drug development and toxicology screening. However, it has been widely reported that human ES cells frequently acquire karyotypic changes throughout culture *in vitro*, termed 'culture-adaptation'. These changes occur sporadically in cultures but appear to be more common in high passage cell lines. The changes also appear to be non-random, frequently involving the gain of chromosomes 1, 12, 17 and 20 suggesting that these chromosomes harbour genes that provide the cells with a selective advantage. These karyotypic abnormalities are also commonly found in human embryonal carcinoma cells and primary human tumours, a concern for the potential use of human ES cells in therapeutic applications.

A recent large-scale study by the International Stem Cell Initiative identified a small copy number variant (CNV) on chromosomal region 20q11.21, which was amplified in >20% of 125 karyotypically normal human ES cell lines. Here we show that the high prevalence of 20q11.21 amplification in human ES cells can be attributed to a strong selective advantage provided by BCL-XL. Cell lines containing the CNV show increased protection against apoptosis resulting in increased population growth rates allowing variant cells to rapidly out-compete normal diploid cells. Cell lines containing the 20q11.21 amplification also show altered differentiation patterns and increased survival of polyploidy. We also describe a method for the rapid detection of the 20q11.21 amplification in human ES cell cultures which is sensitive, cost-effective and applicable for stem cell laboratories worldwide.

List of Abbreviations

Abbreviation	Meaning
APC	Anaphase-Promoting Complex
BAD	BCL-2 Antagonist of Cell Death
BAK	BCL-2 Antagonist or Killer Protein
BAX	BCL-2-Associated X Protein
BCL-2	B-cell Lymphoma 2
bFGF	Basic Fibroblast Growth Factor
BH	BCL-2 Homology (Domain)
BMP	Bone Morphogenic Protein
BUB	Budding Uninhibited by Benzimidazole
CAG	CMV Early Enhancer/Chicken β-Actin (Promoter)
CNV	Copy Number Variant
CPC	Chromosomal Passenger Complex
CRC	Colorectal Cancer Cells
dCT	Change in CT value of the reference gene between two samples
ddCT	Change in dCT value between a reference sample and test sample
EB	Embryoid Body
EC	Embryonal Carcinoma
EpiSC	Epiblast Stem Cell
ES	Embryonic Stem (Cell)
FAM	Fluorescein Amidite
FBS	Foetal Bovine Serum
FISH	Fluorescence In Situ Hybridisation
Gab1	GRB2-Associated Binding Protein 1
GABP	GA-Binding Protein
GDF	Growth Differentiation Factor
GFP	Green Fluorescent Protein
gp130	Glycoprotein-130
ICM	Inner Cell Mass
ID	Inhibitor of Differentiation or DNA Binding
IL6	Interleukin-6
INCENP	Inner Centromere Protein
iPS	Induced Pluripotent Stem (Cell)
IRES	Internal Ribosomal Entry Site
ISCI	International Stem Cell Initiative
JAK	Janus Kinase
KLF4	Krüppel-Like Factor-4
LIF	Leukaemia Inhibitory Factor
LIFRβ	Leukaemia Inhibitory Factor Receptor-β
MAD	Mitotic Arrest Deficient
MAPK	Mitogen-Activated Protein Kinase
MCS	Multiple Cloning Site
MEF	Mouse Embryonic Fibroblast
NEBD	Nuclear Envelope Breakdown

nls	Nuclear Localisation Signal
NTERA2	Sub-line of the TERA-2 EC cell line
OMM	Outer Mitochondrial Membrane
PBS	Phosphate Buffered Saline
PI3K	Phosphatidylinositol-3 Phosphate Kinase
POU5F1	Oct3/4
qRT-PCR	Quantitative Reverse Transcription- Polymerase Chain Reaction
RFP	Red Fluorescent Protein
RhoA/ROCK	Rho-Associated Protein Kinase
SAC	Spindle-Assembly Checkpoint
SCNT	Somatic Cell Nuclear Transfer
shRNAi	Short-Hairpin RNA Interference
SSEA	Stage-Specific Embryonic Antigen
STAT3	Signal Transducers and Activators of Transcription-3
TetR	Tet Repressor
TGF-β	Transforming Growth Factor-β
TO	TetO$_2$ Sequence
UPL	Universal Probe Library

Table of Figures

3	Primary and secondary antibodies used throughout the study	32
4	Primer used to amplify cDNA for generation of over-expressing cell lines	36
5	Growth factor used to differentiate human ES cells via the formation of EBs	42
6	The sizes and descriptions of EBs formed from ESI-035 Control, CNV and BCL-XL over-expression cell lines	91
7	Summary of definitive endoderm differentiation in ESI-035 Control, CNV and BCL-XL over-expressing cell lines	99
Movies	**Description**	**Section**
1	Representative field from time-lapse imaging of HES3 Control plated at 1×10^4 cells/cm^2 and imaged for 72 hours.	4.2.1
2	Representative field from time-lapse imaging of HES3-CNV plated at 1×10^4 cells/cm^2 and imaged for 72 hours.	4.2.1
3	Representative field from time-lapse imaging of H1 Control plated at 1×10^4 cells/cm^2 and imaged for 72 hours.	4.2.1
4	Representative field from time-lapse imaging of H1-CNV plated at 1×10^4 cells/cm^2 and imaged for 72 hours.	4.2.1
5	Representative field from time-lapse imaging of ESI-035 Control plated at 1×10^4 cells/cm^2 and imaged for 72 hours.	4.2.1
6	Representative field from time-lapse imaging of ESI-035-CNV plated at 1×10^4 cells/cm^2 and imaged for 72 hours.	4.2.1
7	Representative field from time-lapse imaging of Shef5 Control plated at 1×10^4 cells/cm^2 and imaged for 72 hours.	4.2.1
8	Representative field from time-lapse imaging of Shef5-CNV plated at 1×10^4 cells/cm^2 and imaged for 72 hours.	4.2.1
9	Time-lapse imaging of the ESI-035 control cell line plated at 2×10^4 cells/cm^2 without ROCK inhibitor.	4.2.3
10	Time-lapse imaging of the ESI-035 control cell line plated at 2×10^4 cells/cm^2 in the presence ROCK inhibitor.	4.2.3
11	Time-lapse imaging of the ESI-035-CNV cell line plated at 2×10^4 cells/cm^2 without ROCK inhibitor.	4.2.3
12	Time-lapse imaging of the ESI-035-CNV cell line plated at 2×10^4 cells/cm^2 in the presence ROCK inhibitor.	4.2.3
13	Time-lapse imaging of the ESI-035-BCL-XL cell line plated at 2×10^4 cells/cm^2 without ROCK inhibitor.	4.2.3

14	Time-lapse imaging of the ESI-035-BCL-XL cell line plated at 2×10^4 cells/cm^2 in the presence ROCK inhibitor.	4.2.3
15	Time-lapse imaging of ESI-035 Control cells undergoing definitive endoderm differentiation	5.2.2
16	Time-lapse imaging of ESI-035-CNV cells undergoing definitive endoderm differentiation	5.2.2
17	Time-lapse imaging of ESI-035-BCL-XL cells undergoing definitive endoderm differentiation	5.2.2
18	An example of an ESI-035-H2B-RFP Control cell undergoing mitosis	6.2.1
19	An example of a lagging chromosome in ESI-035-H2B-RFP Control cell division	6.2.1
20	An example of a chromosomal bridge in ESI-035-H2B-RFP Control cell division	6.2.1
21	An example of a lagging chromosome in ESI-035-H2B-RFP CNV cell division	6.2.1
22	An example of a chromosomal bridge in ESI-035-H2B-RFP CNV cell division	6.2.1
23	An example of unipolar division in ESI-035-H2B-RFP CNV cell division	6.2.1

Publications

Avery, S. *et al*. BCL-XL Mediates the Strong Selective Advantage of a 20q11.21 Amplification Commonly Found in Human Embryonic Stem Cell Cultures. *Stem cell reports* 1, 379-386, doi:10.1016/j.stemcr.2013.10.005 (2013).

Contents Page

1. General Introduction

1.1. Research leading to the derivation of human embryonic stem cells

The human ES cell field has grown rapidly since their initial derivation in 1998. Human ES cells are isolated from the inner cell mass (ICM) of pre-implantation blastocyst stage embryos[1]. The ICM gives rise to the embryo proper, contributing to all three germ layers, ectoderm, mesoderm and endoderm. Once explanted and cultured *in vitro* human ES cells retain the potential to differentiate into all somatic cell types and acquire the ability to self-renew indefinitely making them a valuable resource for developmental biology, cell replacement therapies, drug development and toxicology screening[2]. It is essential to explore the research leading up to the derivation of the first human ES cell lines to fully appreciate the potential for future applications.

1.1.1. Mouse embryonal carcinoma cells

In 1954, Stevens discovered increased incidences of testicular teratocarcinoma formation in strain 129 mice, tumours that were uncommon in many other mice strains[3]. Teratocarcinomas are malignant tumours that are composed of randomly differentiated tissue of all three germ layers and are distinct from teratomas in that they also contain embryonal carcinoma (EC) cells. In 1964, Kleinsmith and Pierce[4] identified EC cells as the cancer stem cells of teratocarcinomas by injecting a single EC cell into syngeneic 129 mice, which resultedin the formation of a new teratocarcinoma. This experiment confirmed the hypothesis that the range of differentiated adult tissue present in teratocarcinomas arises from undifferentiated stem cells that have the capacity to self-renew and differentiate[5]. Since their discovery EC

- 1 -

cells were cultured *in vivo* by the serial transplant of ascites fluid intraperitoneally into mice, which resulted in a combination of solid peritoneal teratocarcinomas and ascites fluid. The ascites fluid contained a number of embryoid bodies (EBs), spherical structures resembling early mouse embryos in that they contained a core of EC cells surrounded by a layer of epithelial cells[6]. In 1967, Finch and Ephrussi[7] established the first mouse EC cell lines that could be cultured *in vitro*. EBs from ascites fluid were isolated, dissociated into single cells then plated onto a feeder layer of mitotically inactivated mouse fibroblasts (MEFs). The resulting outgrowth was homogenous in appearance and produced teratocarcinomas when injected into syngeneic 129 strain mice. In the absence of a feeder layer of MEFs the EC cells could not be maintained indefinitely suggesting an essential role for MEFs in EC cell culture. It is worth noting that shortly after, Gardner[8] demonstrated that the injection of the ICM into developing blastocysts resulted in the generation of chimeric mice, highlighting a pivotal role of the ICM cells in embryonic development. In 1970, Kahan and Ephrussi[9] generated clonal EC cell lines from teratocarcinomas derived from strain 129 mice. The isolated EC cell lines could be cultured in vitro in serum-containing medium on MEFs. The EC cells maintained an undifferentiated state with little spontaneous differentiation. Following subsequent passages in vitro, the EC cells remained in an undifferentiated state and each of the clonal EC cell lines was able to generate teratocarcinomas when injected into immunodeficient mice, demonstrating the extensive self-renewal and pluripotent potential of EC cells following culture *in vitro*.

Three years later, Karen Artzt[10] provided the first evidence of a link between EC cells and the early mouse embryo. The study showed that anti-serum derived by injecting hyper-immunized mice with the F9 EC cell line produced an anti-serum that reacted with a number of undifferentiated EC cell lines but not their differentiated derivatives or other adult mouse tissue. Strikingly, the F9 anti-serum also reacted with very early mouse embryos, suggesting

that the pluripotent potential of EC cells may be mirroring that of early embryogenesis. The full potential of EC cells was realised by work carried out by Brinster[11], who injected EC cells into developing mouse blastocysts, resulting in the generation of chimeric mice. This experiment demonstrated that EC cells and cells of the ICM were similar in their developmental potential. The similarities between EC cells and cells of the ICM led to the realisation that EC cells could be used as a model to study early embryonic development *in vitro*. EC cells could be isolated and cultured in vitro to provide large amounts of material for analysis. This provided a valuable tool, as the study of early embryonic development *in vivo* was limited due to small sample sizes and inaccessibility of embryos following implantation into the uterine wall.

In 1975, Martin and Evans[12] demonstrated the first in vitro differentiation of EC cells through the formation of EBs. EC cells were plated onto gelatine-coated dishes, which helped the attachment of EC cells in the absence of a MEF feeder-layer and promoted the growth of tightly packed spherical EBs. The EBs resembled those found in ascites fluid and when cultured appeared to undergo early embryonic development. As the EBs grew larger the outer cells differentiated into a layer of endoderm enveloping the EC cells. This was reminiscent of the blastocyst stage in mouse embryonic development, where the ICM cells are separated from the blastocoel cavity by a layer of endodermal cells. Following these results, Strickland and Mahdavl[13] successfully induced the differentiation of EC cells into endodermal cells. The serine protease plasminogen activator was found to be secreted by endodermal cells in early embryogenesis and endoderm cells derived from EC cells. In normal cultures of EC cells, plasminogen activator is present only at low levels, which increased slightly when cultures became over-confluent or in cultures with more differentiated cells. Strickland and Mahdavl induced differentiation of EC cells towards an endodermal lineage using low concentrations of retinoic acid, which was added to culture medium that would have otherwise maintained EC

cells in an undifferentiated state. EC cells rapidly differentiated in the presence of retinoic

acid, increasing the production and secretion of plasminogen activator reminiscent of

endodermal differentiation. In the same year Solter and Knowles[14] isolated a monoclonal

antibody from the F9 anti-serum, this antibody named 'stage-specific embryonic antigen 1'

(SSEA1) reacted with a range of mouse EC cell lines but not their differentiated derivatives or

other differentiated adult tissue. The SSEA1 antibody also reacted with the 8-cell stage mouse

embryo and cells within the inner cell mass, again confirming the similarities between EC cells

and the early mouse embryo.

1.1.2. Mouse embryonic stem cells

Despite earlier attempts to isolate pluripotent cells from mouse embryos, in 1981, Evans and

Kaufman [15] explanted developing mouse blastocysts and allowed them to attach to tissue-

culture treated Petri dishes. The blastocysts were grown for four days resulting in the out-

growth and differentiation of the trophoblast into giant trophectoderm cells. The ICM formed

structures similar to egg-cylinders which were picked and dissociated into single cells. The

cells were then grown on MEFs in serum-containing medium. Initial experiments showed that

the isolated ICM cells were morphologically similar to EC cells and displayed multipotent

differentiation following EB formation. As the blastocysts used in the experiments were

isolated from strain 129 mice in which teratocarcinomas are common, Evans and Kaufman

distinguished through karyotype analysis that the pluripotent cells isolated were of embryonic

origin and not contamination of EC cells. The derived cell lines contained both male (40XY)

and female (40XX) normal karyotypes, which set them apart from EC cells as the EC cells grown

within the laboratory were karyotypically abnormal and none of the lines contained a Y

chromosome. Later that year Gail Martin [16] coined the term 'embryonic stem' (ES) cell,

showing that isolated ICM cells were similar in morphology to EC cells, expressed the

embryonic antigen SSEA1 and were able to form teratocarcinomas when injected into syngeneic strain 129 mice. Following this discovery, Bradley and co-workers [17] generated chimeric mice from male mouse ES cell lines. The injection of ES cells into developing mouse embryos yielded high proportions of chimeric offspring (>50%). The group also reported that the injected cells chimerized not only somatic cells but also the gametes of male mice, demonstrating the first germ-line chimeras. This discovery revolutionised experimental mammalian genetics. ES cells grown *in vitro* could be genetically manipulated using insertional mutagenesis using retroviral vectors [18] or gene-specific targeting using homologous recombination [19, 20]. The variant ES cells could then be reintroduced into the ICM to generate germ-line chimeras therefore providing genetically distinct mice to study embryogenesis and development. Since their derivation mouse ES cells have been well characterised and remain to this day one of the most powerful tools to study mammalian genetics.

The similarities between mouse EC and ES cells led to the notion that EC cells were the malignant equivalent of ES cells. In the embryo, it is essential that ES cell proliferation and differentiation is carefully regulated in order for embryogenesis to occur without error. Without this careful regulation in the adult, EC cells resembling cells of the ICM mimic embryogenesis in adult tissue to disastrous effects.

1.1.3. <u>Human embryonal carcinoma cells</u>

Teratocarcinomas are also commonly found in humans and are one of the most common types of tumour found in young adult males [21]. Despite continuing advances in cell culture systems and rapid progress in the mouse, human EC cell lines were not established in vitro until 1975. Fogh and Trempe[22] reported the derivation of a number of human tumour cell

lines from various sources including but not limited to kidney, bladder and testes. TERA-1 and

TERA-2 were the first EC cell lines established from human testicular cancers, although it was

not clear at the time that these cells were indeed EC cells. In 1978, Jewett[23] injected TERA-1

and TERA-2 into nude mice, which resulted in the generation of teratocarcinomas showing that

the tumorigenic properties of human EC cells paralleled that of mouse EC cells. Jewett noted

that the human EC cell lines were highly aneuploid (i.e. had an abnormal number of

chromosomes) and that the teratocarcinomas generated were poorly differentiated.

Following the success of Fogh and Trempe, a number of cell lines were established from

human testicular teratocarcinomas. Andrews[24] and colleagues conducted a comparative

study of eight testicular tumour cell lines. The cellular morphology of the different EC cell lines

was variable, with five of the cell lines containing varying proportions of small round cells

typical of mouse EC cells. Of these five cell lines, two produced tumours resembling

embryonal carcinoma when injected into immunodeficient mice. Of the three cell lines that

did not contain cells with EC-like morphology, only one generated a tumour when injected into

an immunodeficient mouse. The tumour did not display the hallmark characteristics of

embryonal carcinoma suggesting that the cell line may be derived from another malignant cell

type within the testicular tumour. The study noted that human EC cells unlike their mouse

counter-parts differentiated into trophoblast cells suggesting that human EC cells may

represent a pre-blastocyst stage cell type in development. Human EC cells also were negative

for SSEA1, a marker of mouse EC and ICM cells. Instead SSEA1 appeared to mark the

differentiated derivatives of human EC cells. In 1982, Andrews and co-workers[25] found that

a number of human EC cell lines expressed the cell surface antigen SSEA3, a marker found to

be present on the cleavage stage mouse embryos but not present at the blastocyst stage or on

mouse EC cell lines. The study concluded that undifferentiated human EC cells existed in a

SSEA3+/SSEA1- state. This discovery helped to identify pluripotent human EC cells from their

differentiated derivatives. In 1984, Andrews[26, 27] showed extensive differentiation of a

clonal EC cell line derived from TERA-2 (NTERA2) following culture with retinoic acid. Evidence

of in vitro differentiation of human EC cells prior to this point was limited, despite the TERA-2

cell line displaying well differentiated tumours when injected into immunodeficient mice[28].

This provided a tool to study the differentiation of human pluripotent cells in vitro.

In subsequent years, human EC cells were characterised further, revealing differences in

surface between EC cell lines derived from mouse and human teratocarcinomas. Human EC

cells expressed a range of distinguishing cell surface markers including the globoseries

glycogen antigens SSEA3 and SSEA4, the high molecular mass proteoglycan antigens, TRA-1-60

and TRA-1-81 and protein antigens Thy1 and MHC class 1 [21]. These markers were present on

undifferentiated human EC cells but not their differentiated derivatives. These surface

antigens were not expressed by undifferentiated mouse EC/ES cells, which expressed the

surface antigen SSEA1.

1.2. <u>The derivation of human embryonic stem cells</u>

Following the derivation of mouse ES cells, many groups set out to derive ES cell lines from

different species including rabbit[29], pig[30] and sheep[31]. Although the derivation of

pluripotent cells from a number of different species constituted considerable progress in the

stem cell field, the results were somewhat controversial and translating the findings to human

was difficult due to significant differences in embryonic development. It was not until 1995

that Thomson and co-workers[32] derived ES cells from the rhesus monkey providing a more

relevant tool for studying primate development being more closely related to humans than

other mammals. The cells derived by Thomson showed remarkable resemblance to human EC

cells in terms of morphology, cell surface marker staining and differentiation *in vivo/vitro*. It

was therefore unsurprising that when Thomson derived the first human ES cell lines in 1998[1],

they also closely resembled human EC cells and were distinctly different to mouse ES/EC cells. The morphology of human ES cells was similar to rhesus monkey ES cells in that they formed flat colonies with well-defined edges, whereas mouse ES cells tended to form tightly-packed, rounded colonies with irregular borders. Human ES cells also expressed the same characteristic cell surface antigens as undifferentiated human EC cells including SSEA3, SSEA4, TRA-1-60 and TRA-1-81. Human ES cells, like human

EC cells did not stain for SSEA1, a marker of undifferentiated mouse EC/ES cells but showed increased expression upon differentiation consistent with human EC cell data[25].

Thomson and colleagues[1] also noted that human ES cells could be grown on MEFs in the absence of added leukaemia inhibitory factor (LIF) to culture medium to maintain an undifferentiated state. This highlighted possible differences between the signalling pathways governing mouse and human ES cells, as mouse ES cells could be maintained in serum free conditions with added LIF whereas human ES cells did not seem to be dependent on LIF. In the absence of feeders, human ES cells underwent rapid differentiation even if the medium was supplemented with LIF. This highlighted that the maintenance of mouse and human ES cells was governed by different pathways as LIF was essential to maintain mouse ES cells *in vitro*[33] and negated the need for co-culture with MEFs when grown on gelatin[34]. Although human ES cells displayed notable differences from their mouse equivalents, they appeared to share similar developmental potential. When injected into immunodeficient mice, human ES cells formed tumours consisting of all three primary germ layers, evidence of neural epithelium (ectoderm), bone and cartilage (mesoderm) and gut epithelium (endoderm) were present. Human ES cells also differentiated spontaneously *in vitro* either through removal of the MEF feeder-layer or growing cells to confluence. In the latter case, human chorionic gonadotropin was detected in culture medium indicating the presence of trophoblast differentiation consistent with the notion that human ES cells represent an earlier developmental stage than

mouse ES cells in which trophoblast differentiation is rare. However, more recently there has

been wide debate as to whether human ES cells are equivalent to later stages of

development[35] (see Section 1.3).

1.3. <u>Comparing human and mouse embryonic stem cells</u>

As there were marked differences between human and mouse EC cells, there were also

differences between human and mouse ES cells. During the early years of ES cell culture both

mouse and human ES cells were derived and grown in the presence of a feeder layer of MEFs,

which helped to maintain pluripotency, increase plating efficiencies and promote growth of ES

cells. The cells were cultured in serum-containing medium, with mouse ES cells requiring the

addition of LIF and bone morphogenic protein (BMP) to maintain pluripotency and human ES

cells relying on the addition of Activin A and basic fibroblast growth factor (bFGF). The

dependence on different growth factors underpins different extrinsic signaling pathways to

maintain pluripotency in human and mouse ES cells (see Section 1.5). The removal of LIF from

mouse ES cell medium induces widespread differentiation so too does the removal of

Activin/Nodal signaling in human ES cells. Mouse ES cells can be passaged as single cells

whereas human ES cells show mass cell death and differentiation upon dissociation to single

cells. Human and mouse ES cells also display different morphologies in culture, where human

ES cells form flat, round colonies with well-defined edges, mouse ES cells tend to form tightly

packed, rounded colonies. The differences observed between both the morphology and

signaling pathways governing pluripotency have led to human ES cells being more closely

compared to mouse epiblast stem cells (EpiSCs). Brons and colleagues[35] demonstrated in

2007 that isolating cells from the epiblast layer of post-implantation rodent embryos more

closely resembled human ES cells. Mouse EpiSCs displayed human ES cell-like colony

morphology and could only be derived and maintained in medium containing Activin A and

FGF2. Removal of activin/nodal signaling resulted in differentiation of EpiSCs characteristic of

human ES cells. The EpiSCs could also be differentiated *in vitro* into cells comprising the three

germ layers via the aggregation of EBs. Also injection of EpiSCs into the testes capsule of

immunodeficient mice led to the generation of teratomas consisting of a wide variety of

tissues. Interestingly the injection of EpiSCs into blastocyst resulted in poor chimera

generation suggesting that EpiSCs lack the totipotent capacity of mouse ES cells derived from

the ICM. Whether human ES cells are able to form chimeras is unknown and will remain so for

ethical and moral reasons.

1.4. <u>Induced pluripotent stem cells</u>

The derivation of induced pluripotent stem cells (iPS) built upon previous work dating back to

the 1950s. Somatic cell nuclear transfer (SCNT) was first established in the 1950s by King and

Briggs[36, 37] demonstrating that the transfer of nuclei from somatic cells into enucleated frog

oocytes resulted in the generation of cloned offspring. This work alluded to two important

points, firstly that the differentiation of pluripotent cells was a result of reversible epigenetic

changes that occur within the cell rather than irreversible genetic changes. The second

important point was that the genetic material in somatic cells still retained the potential of

totipotency. Subsequent cell fusion hybrid experiments by Miller and Ruddle[38, 39] showed

that the fusion of EC cells with somatic cells resulted in cells with EC properties showing that

the pluripotent state prevailed over the somatic cell lineage.

Another important milestone was the identification of lineage-specific transcription factors.

Transcription factors help to regulate the expression of cell type-specific genes as well as

suppressing genes from other lineages. In 1987, Davis and colleagues[40] showed that the

over-expression of the transcription factor myoD in fibroblasts induced the formation of

myofibers showing that the expression transcription factors alone could influence cell behaviour. This work was followed by Graf and co-workers[41], who showed that cells could be transdifferentiated from B and T cells into functional macrophages via the over-expression of C/EBPα, a myeloid transcription factor. More recently, insulin producing pancreatic β cells have been generated from pancreatic exocrine cells by over-expressing MafA, Pdx1 and Ngn3[42]. Transdifferentiation is not restricted to cells within the same lineage as it has been shown that fibroblasts (mesoderm origin) have been successfully converted to functional neurons (ectoderm origin) using the transcription factors Ascl1, Brn2 and Myt1l[43].

Previous work led to the realisation that pluripotency could be induced in somatic cells by introducing pluripotency associated transcription factors. In 2006 Yamanaka and Takahashi[44] derived the first iPS cell lines from mouse embryonic and adult fibroblasts. Using a method later termed 'reprogramming' they introduced key genes that were deemed to have pivotal roles in regulating the pluripotent state using retroviral vectors. They found that four genes (Oct3/4, Sox2, c-Myc and Klf4) were sufficient to revert fibroblasts back to an ES-like state. The iPS cell lines generated had ES cell like morphology, growth characteristics, expressed ES cell markers and produced teratomas with cells pertaining the three germ layers[44]. Since 2006, the iPS cell field has developed rapidly with the generation of the first human iPS cells in 2007 using the same combination of transcription factors[45] or substituting C-MYC and KLF4 with NANOG and LIN28[46]. This breakthrough had huge implications in regenerative medicine, the ability to generate patient-specific ES cell lines from somatic cells could potentially overcome limitations in donor availability and immunologic barriers.

More recently the methods to derive iPS cells have moved away from viral vectors due to obvious clinical concerns over the random integration into the genome. It has also been shown that the transfected transcription factors are not silenced once reprogrammed again

highlighting concerns for their use in therapies. The transfection of fibroblasts using modified

mRNA has been successfully used to generate iPS cell lines free from integrating viral vectors

and offers a safer and more controlled alternative to viral integration [47].

1.5. <u>The extrinsic signaling pathways regulating embryonic stem cells</u>

LIF signaling is essential in the maintenance of self-renewal and pluripotency in mouse ES cells.

Supplementing serum-containing medium with LIF negates the need for mouse ES cell co-

culture with MEFs, supporting feeder-free growth [34]. Mouse ES cells cannot be derived from

LIF-null ICM cells [48] highlighting the requirement for LIF in mouse ES cell derivation. The

culture of mouse ES cells in medium without supplemented LIF results in the down-regulation

of pluripotency related genes and cultures rapidly differentiate. Interestingly LIF is not

required in early mouse development in vivo as LIF-null mice develop normally into adulthood.

However, female mice despite being fertile are unable to undergo blastocyst implantation into

the uterus [49] a characteristic that can be rescued by transplanting the blastocyst into wild-

type female mice [50]. LIF is a member of the interleukin 6 (IL6) family of cytokines. The LIF

signaling pathway is activated by the binding of LIF to a heterodimeric receptor consisting of

the LIF receptor (LIFRβ) and glycoprotein 130 (gp130). The activation of gp130 activates a

number of downstream pathways which regulate self-renewal, survival and differentiation of

ES cells. The binding of LIF to its receptor results in the activation of receptor-associated Janus

kinases (JAKs) which mediate the phosphorylation of the tyrosine 705 residue on STAT3 (signal

transducers and activators of transcription 3). The active STAT3 molecules then form

homodimers and translocate to the nucleus where they bind specific DNA sequences to

regulate expression of target genes. The direct targets of STAT3 include GA-binding protein

(GABP), a transcription factor that positively regulates the expression of Oct3/4. Knock-down

of GABP in mouse ES cells results in Oct3/4 down-regulation and differentiation [51]. The

reprogramming factor krüppel-like factor 4 (KLF4) has also been shown to be regulated by STAT3, KLF4 is a co-factor of Oct3/4 and Sox2 [52]. Inhibition of STAT3 activation through modification of tyrosine 705 results in the loss of self-renewal and leads to the differentiation of ES cells [53]. The binding of LIF also mediates the mitogen-activated protein kinase (MAPK) pathway. The SHP-2 tyrosine phosphatase molecule binds to the active gp130 and forms a complex with Gab1 (GRB2-associated binding protein 1) which leads to the activation of ERK1/2 [54]. Perturbation of the MAPK signaling pathway through chemical inhibition of MEK, an upstream component in the activation of ERK results in enhanced growth of mouse ES cells. ERK signaling is therefore an important regulator in the balance between self-renewal and differentiation in mouse ES cells. Finally, LIF also activates the PI3K (phosphatidylinositol-3 phosphate kinase) pathway which regulates a number of cellular processes including cell cycle, proliferation and apoptosis [55]. In mouse ES cells, chemical inhibition of the PI3K pathway in the presence of LIF results in differentiation and loss of self-renewal. This response to the loss of PI3K signaling can be explained by increased ERK activity in the presence of LIF resulting in the differentiation of ES cells [56].

The pathways regulating self-renewal and pluripotency in human ES cells differ from those required for the maintenance of mouse ES cells. Human ES cells are dependent on transforming growth factor-β (TGF-β) signaling to maintain self-renewal and pluripotency. The TGF-β pathway can be divided into the TGF-β/activin/Nodal and the BMP/GDF (growth differentiation factor) pathways based on the structural similarity of ligands and the different responses that they elicit [57]. It has been shown that the two pathways regulate different aspects of stem cell culture. The TGF-β/activin/Nodal pathway is important in the regulation of self-renewal and pluripotency whereas the activation of the BMP/GDF pathway leads to differentiation. TGF-β signalling is activated through binding of TGF-β ligands to type I and II receptor serine/threonine kinases. The activation of type I receptors leads to the binding and

subsequent phosphorylation of receptor-regulated SMAD (R-SMAD) proteins 1, 2, 3, 5 and 8. The R-SMADs bind the co-mediator SMAD4 and are then transported to the nucleus where they regulate transcription of specific target genes. The TGF-β/activin/Nodal pathway signals through the binding of ligands to the type I receptors ALK-4, -5 and -7 and subsequent activation of SMAD2/3 [58]. Inhibition of TGF-β/activin/Nodal signalling leads to the loss of POU5F1 (Oct3/4) and NANOG expression in human ES cells, highlighting the importance in the maintenance of pluripotency and self-renewal [59]. The BMP/GDF pathway signals through the type I receptors ALK-1, -2, -3 and -6 and subsequent activation of SMAD1, 5 and 8. The levels of SMAD1/5 are low in undifferentiated human ES cells but upon differentiation the levels increase [59]. Inhibition of BMP signalling (and therefore SMAD1/5) through the BMP-inhibitor noggin can support self-renewal of human ES cells in MEF conditioned media in feeder-free conditions [60].

1.6. Culture-adaptation of human embryonic stem cells

It has been widely reported that upon prolonged culture in vitro human ES cells often become karyotypically abnormal, a common characteristic of human EC cells. These changes occur sporadically throughout culture but tend to become more frequent at later passage. The changes also appear to be non-random, frequently involving the duplication of chromosomes 1, 12, 17 and 20[61-66]. Although the gain of genomic material is the most common, changes can involve gains, inversions and deletions of whole or partial chromosomes. At present the mechanism(s) behind karyotypic change remain relatively unknown and it is therefore difficult to alter culture conditions to try and alleviate genetic instability. It has been suggested that errors in DNA replication and abnormal segregation of sister chromatids at mitosis may play a role in genetic instability[67]. It has been shown that human ES cells, unlike somatic cells, do not activate key S-phase checkpoint pathways in response to DNA replication stress. Human

ES cells cultured in the presence of thymidine, an inhibitor that slows DNA replication by starving cells of dCTP, results in increased apoptosis. This response to DNA replication stress may be a protection mechanism to maintain the genomic integrity of the population[68]. As mentioned above, human EC cell lines are often aneuploid, which commonly includes the amplification of 12p and 17q. These two amplifications have also been observed in a number of human cancers including breast and ovarian cancer[69, 70]. The similarity between karyotypic abnormalities between human ES cells, EC cells and primary tumours suggests a common intrinsic mechanism that results in a tumourigenic phenotype rather than a response to culture conditions. Karyotypic changes are also observed in mouse ES cells, notably duplications of chromosomes 8 and 11 are observed during culture. Interestingly the mouse chromosome 11 shows strong homology to the human chromosome 17 suggesting mouse ES cells are equally vulnerable to culture-adaptation [71].

The appearance of genetic change relies on two events, mutation and selection: firstly the cell must acquire a genetic change and as a result the change must provide the cell with a selective advantage over 'normal' cells. When looked at in their simplest form, human ES cells can follow one of three paths; self-renewal to produce two undifferentiated daughter cells, differentiate to produce specialised adult cells or undergo cell death. When looked at in this basic way it is easy to see how selection of particular variants occurs in a population of human ES cells (Figure 1). If a cell acquires a genetic change which increases its tendency to differentiate, makes it more susceptible to death or renders a cell unable to self-renew the cells would be quickly lost from a stem cell population throughout subsequent passages. However, a genetic change that inhibits differentiation to one or more lineages, increases cell survival or increases a cells capacity to self-renew would provide a selective advantage over diploid cells. This may be the reason why the genetic changes observed in human ES cells are seemingly non-random with genes on the most commonly gained chromosome contributing to

a selective advantage in one or more ways. This may also be the reason why gains of other

chromosomes such as chromosome 4 are not observed in human ES cells, making such gains

deleterious to cells. The acquisition of karyotypic abnormalities is often coupled with

decreased population doubling times and increased cloning efficiencies in human ES cells,

suggesting that abnormal cells are selected throughout culture [72]. Olariu et al modelled this

selective advantage in human ES cell culture by mixing 99% karyotypically normal cells with 1%

abnormal cells exhibiting different cytogenetic changes. The study showed that the abnormal

cells out-compete the diploid cells, taking over the culture between 20-40 passages[73].

The derivation and subsequent passage of human ES cells places them under considerable

selection pressures. It is not known if the subsequent out-growth of ES cells from the ICM

contains all ICM cells or just those cells that are capable of growth *in vitro.* The freeze/thawing

of human ES cells places the cells under considerable stress, many of the cells die upon

thawing, which may help variants with increased cell survival to out-compete normal cells.

Also, culture conditions for the maintenance of human ES cells are not optimal; it is not

uncommon for human ES cell cultures to spontaneously differentiate or undergo mass cell

death following passage. Barbaric and co-workers[74] have described certain 'bottle-necks' in

human ES cell culture that allow variant cells with increased growth rates to over-take diploid

cultures. For example, karyotypically normal human ES cells undergo mass cell death following

re-plating showing reduced motility to culture-adapted cells. This means that during

subsequent passaging of human ES cells will constantly select for variant cells with increased

growth capacities. It has also been shown that populations of human ES cells appear to be

heterogeneous in terms of gene expression levels. Single human ES cells sorted for SSEA3

show huge heterogeneity not only for lineage specific markers but also for the pluripotency

markers OCT4 and NANOG[75]. This heterogeneity has led to the theory that within a

population of stem cells, there may exist subsets of cells that are lineage-primed and upon

differentiation will favour one lineage over others. The extent to which these subsets can fluctuate between different states has yet to be determined. Developing new techniques for stem cell culture may help to alleviate the bottle-necks allowing culture-adapted cells to over-take human ES cell cultures.

Karyotypic change has also been implicated in restricting differentiation patterns of pluripotent cells. Many human EC cells appear to have restricted differentiation towards certain lineages with some EC cell lines becoming nullipotent throughout prolonged passage [25]. Culture-adapted human ES cells also appear to have altered differentiation patterns when compared to their diploid counterparts. Fazeli and co-workers[76] demonstrated this through spontaneous differentiation via the formation of EBs. The differentiation of two unrelated, diploid human ES cell lines (H7 and Shef3) showed similar differentiation patterns as determined by qRT-PCR. The diploid cell lines expressed lineage-specific genes from all three germ layers and down-regulated the pluripotency marker POU5F1. The same experiment was carried out on two culture-adapted cell lines; a sub-clone of the H7 cell line (H7.s6: amplification of 17q) and a late-passage H14 (gain of whole chromosome 17). The two karyotypically abnormal cell lines exhibited reduced differentiation capacity particularly with respect to endodermal markers. Interestingly principal component analysis of qRT-PCR data showed that although the diploid H7 and Shef3 cell lines were isolated in different labs by different techniques and that they were also of opposite genders, the expression patterns clustered closely. This was not the case for the H7 cell lines, although H7.s6 is a sub-line of the H7 cell line and share the same genetic background, isolation technique and maintenance they did not cluster together in their differentiation pattern. Instead the two culture-adapted cell lines clustered closer together potentially highlighting that karyotypic changes can effect stem cell growth through reduced differentiation capacity[76].

Figure 1. Mutations effecting human embryonic stem cell behaviour
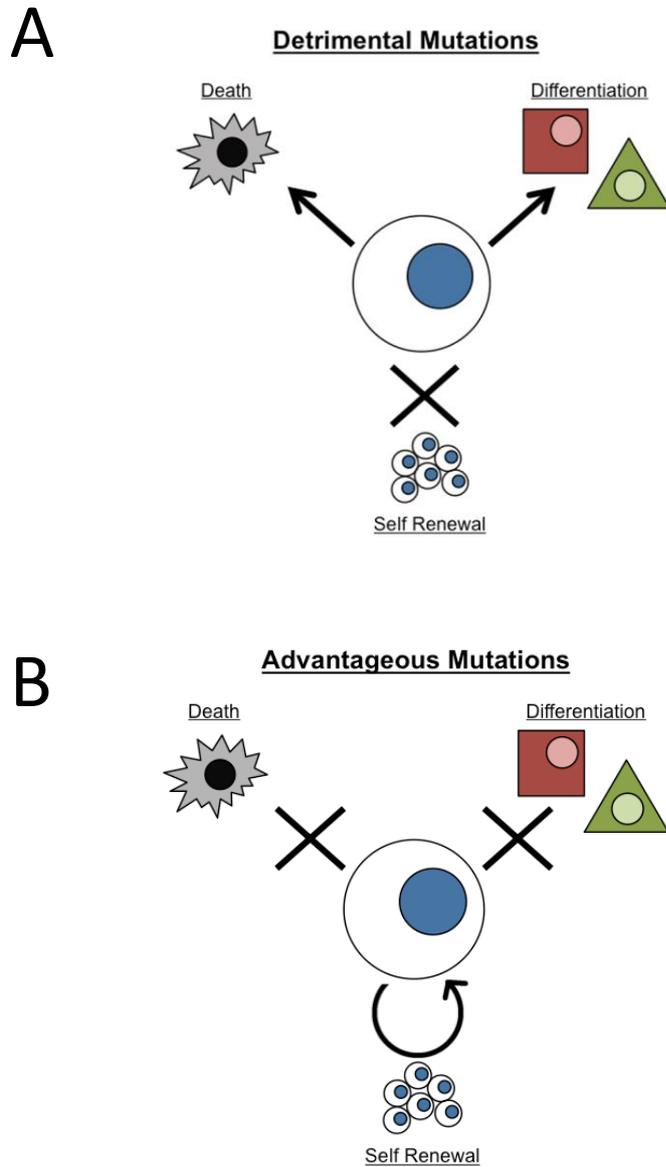
Human ES cells can ultimately follow one of three paths. They can: self-renew to create two daughter stem cells, initiate differentiation towards one of many adult cell types or undergo cell death. (A) Detrimental mutations include those that make a cell more susceptible to cell death, increase the probability that the cell will differentiate or reduce a cells capacity to self-renew. Cells that acquire these mutations would be rapidly lost from the culture. (B) Advantageous mutations include those that increase cell survival, reduce/inhibit a cells capacity to differentiate or increase proliferation. Cells acquiring these mutations would rapidly out-compete normal cells.

The International Stem Cell Initiative (ISCI) was established to characterise and compare different human ES and iPS cell lines on a global scale to better understand the potential for their eventual use in medicine. The second phase (ISCI2)[66] focussed on characterising the common genetic changes observed in human pluripotent stem cells. The study was based on 125 human ES cell lines and 11 human iPS cell lines from 38 laboratories worldwide. This sample range included human ES cell lines from a diverse ethnic and genetic background. The cell lines were analysed at two passage levels (early and late passage) to determine whether the cell line remained normal, picked up a change throughout culture or remained abnormal at late passage. Of the 125 human ES cell lines, 83 (66%) maintained a normal karyotype showing that following derivation, human ES cells are commonly diploid and maintain a normal karyotype throughout culture.

The remaining 42 cell lines were found to harbour at least one karyotypic change in early or late passage. The study showed that cells at late passage were more likely to acquire karyotypic changes than those at early passage implying that changes occur following culture and not a result of cells being abnormal upon derivation. The changes observed were consistent with those reported in the literature, notably gains of chromosomes 1, 12, 17 and 20. Unfortunately, a whole chromosome gain offers low resolution when trying to identify the gene(s) responsible for providing a selective advantage. The genomic DNA of all karyotypically normal cell lines was analysed using SNP-arrays to try and detect any small structural variants that routine karyotyping could not detect. This analysis highlighted a small copy number variant (CNV) on chromosome 20, which was common in 22 karyotypically normal cell lines. No CNVs were detected on any of the other commonly amplified chromosomes. Of the 22 cell lines identified with the 20q11.21 amplification in 17 instances the CNV was present only in the late passage cell line, the remaining 5 cell lines showed amplification both in the early and

late passages. In no instance was the CNV found in only the early passage cell line, this demonstrated that the 20q11.21 gain manifests during culture and persists throughout passage, as it is not lost once duplicated.

The 20q11.21 amplification was present in different cell lines at varying lengths (0.6Mb to 2.5Mb). However, all cell lines contained a minimal amplicon spanning roughly 0.6Mb. The minimal amplicon contains 13 coding genes, of which only three are expressed in undifferentiated human ES cells, HM13, ID1 and BCL2L1. The presence of a minimal amplicon suggests that one of the three candidate genes, once over-expressed confers a selective advantage over other diploid cells.

Here we describe the effects of 20q11.21 amplification on human ES cell culture, investigate the genes and mechanisms responsible for its high prevalence in human ES cell lines and its implications in future applications of human ES cells.

2. Methods

2.1. Matrigel coating

Tissue-culture treated vessels were coated with Matrigel (BD Biosciences: 354234) diluted 1:40 in DMEM/F12 (Sigma) for 1 hour at room temperature. Flasks/plates were then stored at 4°C until used.

2.2. Human embryonic stem cell culture

Human ES cells were grown on matrigel-coated T12.5 tissue-culture treated flasks (Falcon) supplemented with mTeSR (Stem Cell Technologies) at 37°C under a humidified atmosphere of 5% CO_2. Cultures were passaged every 4 to 6 days at a ratio of 1:3 or 1:4. Colonies were picked under the microscope (4X objective) using a fine tip pastette (Alpha Laboratories) following a 5 minute digestion at 37°C with collagenase type IV (Invitrogen). The cell suspension was then added to a new matrigel-coated flask containing mTeSR.

2.3. Embryonal carcinoma cell culture

EC cells were grown on T25 tissue-culture treated flasks in DMEM/F12 media supplemented with 10% fetal bovine serum (FBS) (Hyclone) at 37°C under a humidified atmosphere of 5% CO_2. Cultures were passaged every 4-6 days at a ratio of 1:3-1:5 following a 5 minute digestion with 0.25% trypsin-EDTA at 37°C.

2.4. Freezing of human embryonic stem cells

Human ES cells were digested with collagenase (type IV) for 5 minutes at 37°C, 5% CO_2. The collagenase was aspirated and fresh mTeSR was added to the cells. The cells were then detached from the surface of the tissue culture vessel using a fine tip pastette, taking care not to break colonies into small chunks. The suspension was then centrifuged for 3 minutes at 1000rpm and supernatant aspirated. The cell pellet was then resuspended gently in freeze medium containing 60% hESC medium, 30% KOSR and 10% DMSO. The suspension was then added to 1.5ml cryovials and placed into a Mr Frosty freezing vessel at -80°C for 24 hours. The vials were then transferred to liquid nitrogen for storage.

2.5. Thawing of human embryonic stem cells

2ml of fresh mTeSR was added to matrigel-coated T12.5 tissue culture-treated flasks and placed at 37°C, 5% CO_2 for 30 minutes. 10mls of hESC medium was added to a 15ml falcon tube and placed into a 37°C water bath for 10 minutes. Cryovials were removed from liquid nitrogen and partly submerged into a 37°C water bath until the cell pellet was partially defrosted. The pellet was then transferred to the 15ml falcon tube containing hESC medium using a 5ml strippette. The suspension was then centrifuged for 3 minutes at 1000rpm. The supernatant was aspirated and the cell pellet was tapped gently to loosen, 1ml of mTeSR from the pre-warmed flask was then used to resuspend the cells pellet, which was then added to the T12.5 flask. The cells were then placed back into the incubator at 37°C, 5% CO_2 and left for 24 hours.

2.6. Single cell dissociation

Cells were washed with sterile PBS and incubated at 37°C for 5-7 minutes in either 0.25% trypsin-EDTA (Sigma), TrypLE (GIBCO) or Accutase (Stem Cell Technologies), cells were dissociated by tapping flask sharply with palm of hand. Dissociation agent was then inactivated by mixing 1:2 with appropriate serum containing media. Cells were pelleted by centrifugation at 1200rpm for 3 minutes using a Heraeus Megafuge 1.0R centrifuge and supernatant aspirated to remove dissociation agent. Cells were then resuspended in appropriate media.

2.7. Cell counting

The live cell number counts of single cell human ES cells were determined by mixing 15uL of cell suspension at a 1:1 ratio with 0.4% Trypan Blue (Sigma). 10uL of the mix was added to an improved neubauer haemocytometer (Hawksley) and the 4 corner grids counted. The concentration of the cell suspension and total cell number were determined using the following calculations.

Concentration (*Cells/mL*) = (*Number of Cells ÷ Number of Grids Counted*) × 20,000*

Total Cell Number = Concentration (*Cells/mL*) × *Volume of Suspension* (*mL*)

*Cells were multiplied by 20,000 to account for a 2 times dilution by mixing cells 1:1 with Trypan Blue.

2.8. RNA extraction

2.8.1. Harvesting cells

2.8.1.1. Extraction of RNA from cells in monolayer

Cells were washed with sterile PBS and 600μL of Trizol (Life Technologies) was added directly into culture vessel. The Trizol was evenly dispersed by rocking flask from side to side for approximately 2 minutes until the sample became viscous. The suspension was then scraped into a corner using a sterile 25cm cell scraper (Fisher Scientific) and placed into 15ml falcon tube (STARLAB). The sample was vortexed to achieve a homogenous suspension. RNA samples were stored at -20°C if not processed immediately (frozen samples were thawed on ice before proceeding).

2.8.1.2. Extraction of RNA from cells in suspension

Cells in suspension were pelleted by centrifugation at 1000rpm for 3 minutes in a 15mL falcon tube and supernatant aspirated. The cells were resuspended thoroughly by flicking the tube and 600uL of Trizol was added. Samples were the vortexed to mix and stored as above.

2.8.2. Isolating total RNA

200μL of chloroform (Sigma) was added to the Trizol suspension and mixed thoroughly by vortexing until milky pink. Mixture was then transferred to a 1.5mL Eppendorf tube and left to stand for 10 minutes at room temperature. The sample was centrifuged at 14,000rpm for 30 minutes at 4°C to achieve phase separation. The upper aqueous layer was extracted and

placed into a sterile 1.5mL Eppendorf tube and an equal volume plus 50uL of 100% ethanol

was added and mixed by vortexing.

The sample was then processed using the RNA Clean-Up and Concentration Kit (NORGEN) to

isolate Total RNA.

2.9. RNA to cDNA reverse transcription

Total RNA was reverse transcribed to make cDNA using the High Capacity RNA-to-cDNA Kit

(Life Technologies).

2.10. Quantitative-PCR on cDNA samples

cDNA samples were quantified using a UV/Vis spectrophotometer (Geneflow:

Nanophotometer P330) and diluted to 5ng/μL using ddH$_2$O. A master mix was made on ice

consisting of 5μL Taqman Fast Universal Master Mix (Life Technologies), 0.2μL Primer Mix

(containing 5μM sense-antisense oligos), 0.1μL Roche Probe (Roche UPL) and 2.7μL ddH$_2$O per

reaction (for oligonucleotide sequences see Table 1). 8μL master mix and 2μL of gDNA

(5ng/μL) was added to each well of either a Fast 96- or 384-well PCR plate (Life Technologies)

and spun briefly. The plate was then analysed using a 7900HT Fast Real-Time PCR System

(Applied Biosystems) using the following steps, HOLD at 50°C for 5 minutes, HOLD at 95°C for

10 minutes then 45 cycles of 95°C for 15 seconds and 60°C for 1 minute. Data was then

analysed using the $\Delta\Delta$Ct method[77].

Table 1. Primer/probe pairs used for the detection of mRNA
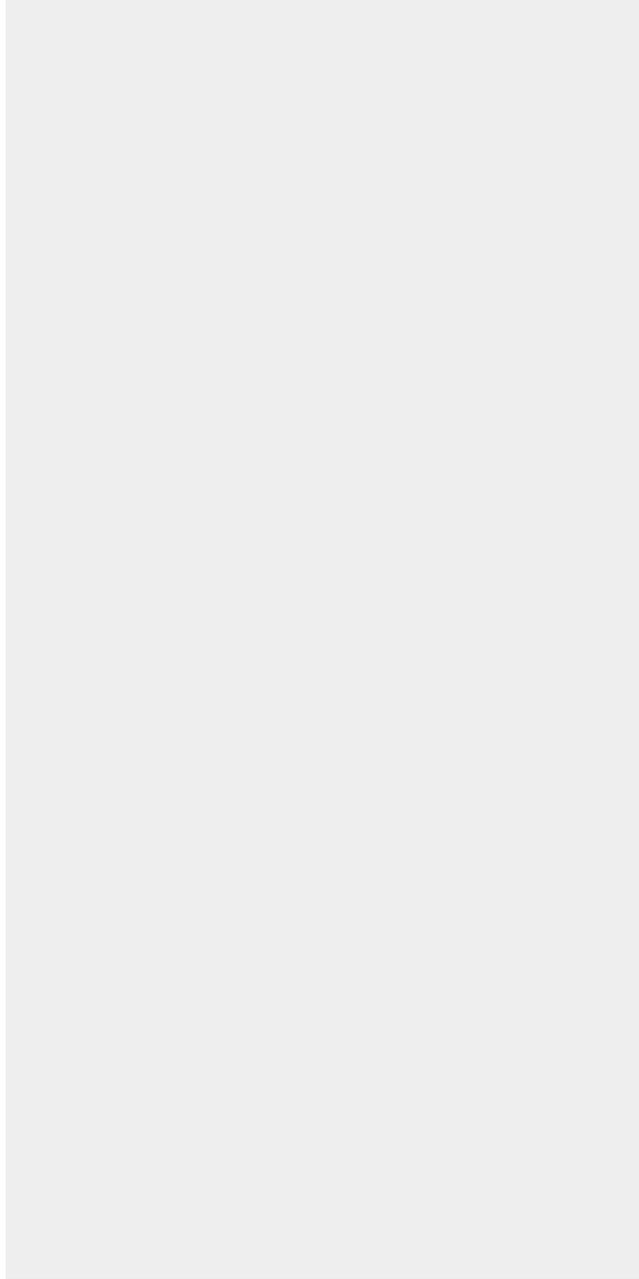
The primer/probe pairs used in the detection of mRNA using qRT-PCR are shown above. The table shows informaton on the gene name, primer sequences and probe number (UPL).

2.11. Immunodetection of protein

2.11.1. Harvesting protein

Cells were washed once with PBS and aspirated thoroughly to minimise dilution of buffer, the flask was then placed on ice. 300μL of ice cold RIPA buffer (50mM Tris-Cl pH7.7, 150mM NaCl, 1% NP40, 0.25% sodium deoxycholate, 0.1% SDS, 1x final concentration complete protease inhibitor) was added to the cells. The cells were then scraped from the surface of the flask using a 25cm cell scraper and placed into a labelled 1.5mL eppendorf tube. The sample was incubated on ice for 20 minutes before sonification (10% Amplitude, 2 seconds, 3 repeats). The sample was then spun at 14,000rpm for 20 miutes at 4°C and supernatant placed into a clean 1.5mL Eppendorf tube on ice.

2.11.2. Sample fractionation

10μg of protein sample was added to a 1.5mL eppendorf tube on ice, the volume was adjusted to 10μL using ddH$_2$O and 2.5uL of Loading Buffer was added. The sample was then incubated at 100°C for 10 minutes and then run on a 4–15% Mini-PROTEAN TGX Gel (BIO-RAD) alongside a BenchMark Pre-Stained Protein Ladder (Invitrogen). The gel was run for 15 minutes at 85V and then increased to 110V for 1 hour.

2.11.3. Visualising protein

The gel was then electroblotted onto 0.2μM polyvinylidene difluoride membrane using a Trans-blot Turbo Transfer System. The membrane was placed in 5% milk in TBST (50mM Tris, 150mM NaCl, 0.05% Tween 20, pH 7.6) for one hour. The membrane was then placed into a

50mL falcon tube containing 5% Milk in TBST and primary antibody. The tube was incubated

overnight at 4°C on a tube roller (Bibby Scientific). The membrane was washed 3 times in TBST

for 10 minutes at room temperature on a rotating plate. The membrane was then placed into

a 50mL falcon tube containing 5% Milk in TBST and secondary antibody added. The membrane

was incubated for 1 hour at 4°C on a tube roller and then washed 3 times in TBST. The HRP

was detected using SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific) and

visualised with CL-XPosure Clear Blue X-Ray Film (Thermo Scientific).

2.12.   gDNA Extraction

Cells were washed once with sterile PBS and harvested using 0.25% trypsin-EDTA. The trypsin

was inactivated by adding 10% FBS in sterile PBS at a ratio of 1:3 respectively. The cells were

centrifuged at 1000rpm for 3 minutes and supernatant aspirated. The cell pellet was then

resuspended in 200μL of PBS and added to a 1.5mL Eppendorf tube containing 20μL of

Proteinase A. The sample was then processed using the DNA Blood & Tissue Kit (QIAGEN:

69504). The optional incubation step at 56°C for 10 minutes was carried out.

2.13.   CNV Detection by Genomic DNA qPCR

Relative copy number was determined using oligonucleotide pairs designed to introns of genes

spanning the 20q11.21 region, in combination with the Universal Probe Library (Roche) (See

Table 2). Genomic DNA samples were quantified using a UV/Vis spectrophotometer

(Geneflow: Nanophotometer P330) and diluted to 5ng/μL using ddH$_2$O. A master mix was

made on ice consisting of 5μL Taqman Fast Universal Master Mix (Life Technologies: 4352042),

0.2μL Primer Mix (containing 5uM sense-antisense), 0.1μL Roche Probe and 2.7μL ddH$_2$O per

reaction. 8μL master mix and 2μL of gDNA (5ng/μL) was added to each well of either a 96- or

384-well PCR plate and spun briefly. The plate was then analysed using a 7900HT Fast Real-Time PCR System (Applied Biosystems). The reaction was held at 50°C for 5 minutes then 95°C for 10 minutes then 45 cycles of 95°C for 15 seconds then 60°C for 1 minute. Data was then analysed using the ΔΔCt method (Livak and Schmittgen, 2001).

2.14. mRNA-sequencing

43 million mapped, strad-specific 50 base reads from H1 ESCs were generated on the SOLiD3 platform (Applied Biosystems). BCL2L1 had 92.31 RPKM in the undifferentiated human ES cells.

2.15. Karyotyping

Analysis was performed by a Health Professionals Council registered Clinical Scientist in a Clinical Pathology Accredited laboratory.

2.15.1. Sample preparation

Cells were incubated at 37°C for 4-6 hours in mTeSR containing 0.06ug/mL Colcemid (GIBCO: 15212-012). The cells were then dissociated using 0.25% Trypsin-EDTA and spun at 1200rpm for 3 minutes in a 50mL falcon tube. The supernatant was then carefully decanted and cells resuspended by tapping flask sharply. 8mLs of hypotonic solution (37.5mM potassium chloride, 7.8mM trisodium citrate in dH_2O) was added dropwise whilst gently vortexing and then placed at 37°C for 20 minutes. The cells were then spun at 1200rpm for 5 minutes, supernatant was decanted and cells resuspended thoroughly by flicking. The cells were then fixed by adding 5mLs of ice cold methanol: acetic acid (3:1) dropwise whilst gently vortexing

and placed on ice for 20 minutes. The sample was then spun at 1500rpm for 10 minutes and

fixation step repeated. The sample was then stored at -20°C until processed.

2.15.2. Preparation of slides

Slides were washed in acetone using a Pasteur pipette and excess acetone removed by dipping

slides in dH$_2$O using tweezers. The slides were then placed on a 40°C heating plate until dry

and then dipped into ice cold 40% methanol. The excess was drained off by placing edge on

absorbent paper and then placed on 40°C heating plate until dry (slides could be stored in

100% ethanol until used for metaphase spread). Sample was dropped from approximately

45cm using a 1mL pipette (2-3 drops per slide) onto slides placed at a 45° angle and dried on a

40°C heating plate.

2.15.3. G-banding

The slides were treated with trypsin for 25 seconds and then stained with 4:1 Gurr's

Buffer/Leishmann's stain (Sigma) for 2 minutes. The slides were scanned, metaphases

captured and analysed using a Cytovision GSL-120 (Leica Microsystems) image analysis system.

A minimum of 5 metaphase spreads were analysed and a further 15 counted and scored.

2.16. Fluorescence in situ hybridisation (FISH)

The BCL2L1 FISH probe was a spectrum orange fluorescently labelled BAC (RP5-857M17,

almost 100kb) provided by BlueGnome (Illumina), and covers the genes *BCL2L1, COX4I2* and

the 3' end of *ID1*. For 20p telomere detection, the Vysis (Abbott Molecular) probes TelVysion

20p telomere (spectrum green), were used.

Probe and slides were denatured together at 72°C for 2 minutes in a PTC-200 DNA Engine (Peltier Thermal Cycler, MJ Research) and incubated at 37°C for 16 hours for hybridisation. Slides were washed in 0.4x sodium citrate (Abbott Molecular) with 0.3% Tween 20 (Sigma) and 2x sodium citrate with 0.1% Tween 20. Coverslips were mounted on the slides in 20uL, Vectashield Mounting Medium with DAPI (Vector Laboratories). One hundered interphase cells were analysed on an Olympus BX51 fluorescent microscope.

2.17.   Flow-cytometry analysis of surface markers

Cells were washed once with sterile PBS and harvested using accutase. Cells were resuspended in wash buffer and spun at 1000rpm for 3 minutes. The supernatant was aspirated and cells resuspended at $1x10^7$/mL in wash buffer. 100uL of sample was added to a 5mL FACS tube and primary antibody was added (see Table 3). The sample was incubated at 37°C for 30 minutes in the dark, occasionally flicking to mix. Cells were washed once with 4mLs of wash buffer and spun at 1000rpm for 3 minutes. Supernatant was decanted and cells resuspended in 100uL of wash buffer containing secondary antibody. The sample was incubated at 37°C for 30 minutes in the dark, occasionally flicking to mix. The sample was then washed twice with 4mLs of wash buffer and spinning at 1500rpm for 5 minutes. Supernatant was decanted and sample resuspended in 300uL PBS and analysed on flow cytometer (BD Biosciences: FACSCalibur).

Table 2. Primer/probe pairs used for the detection of gDNA
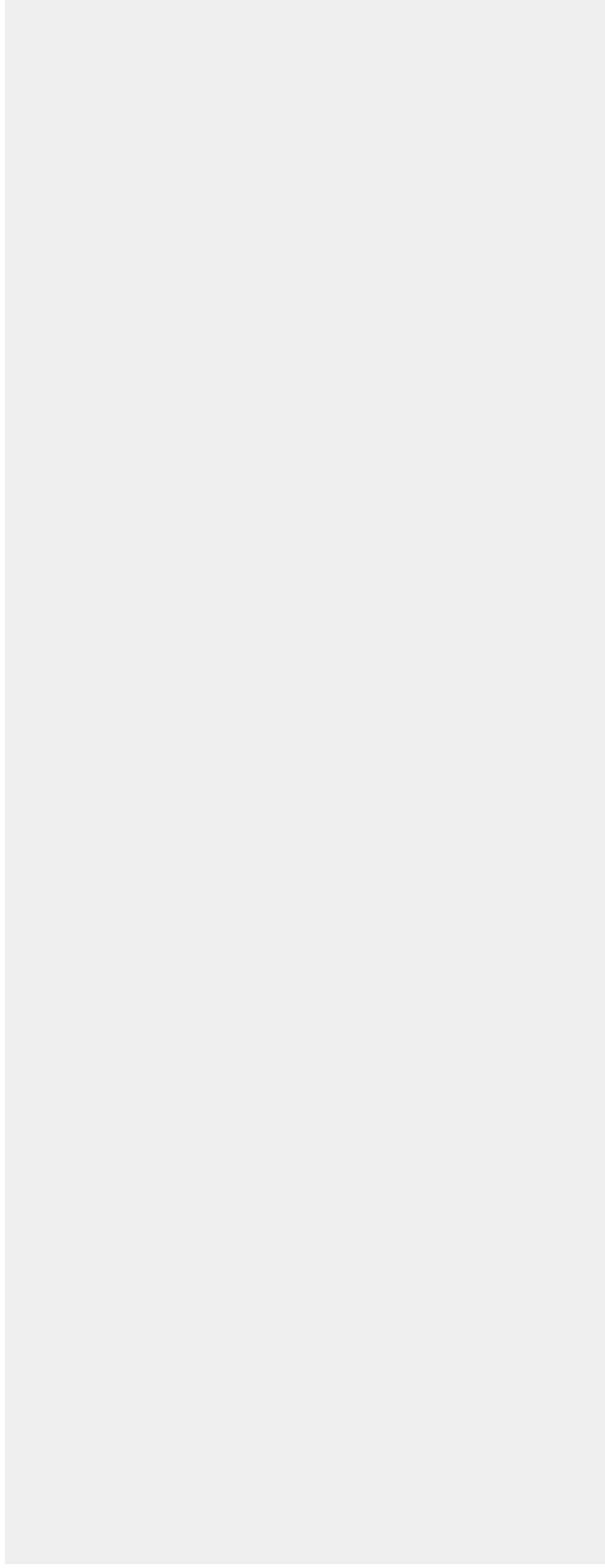
The primer/probe pairs used in the detection of gDNA using qRT-PCR are shown above. The table shows information on the gene name, primer sequences, probe number (UPL), whether they are located on an intronic region, position on the genome and genomic location relative to UCSC human genome assembly hg19 (Kent et al).

Table 3. Primary and secondary antibodies used throughout the study

Primary Antibody	Supplier	Concentration	Secondary Antibody	Concentration
P3X	Kohler and Milstein	1:10	Dylight 488-conjugated Goat anti-Mouse	1:100
TRA-1-85	Andrews	1:10	Dylight 488-conjugated Goat anti-Mouse	1:100
TRA-1-81	Andrews	1:10	Dylight 488-conjugated Goat anti-Mouse	1:100
SSEA3	Shevinsky	1:10	Dylight 488-conjugated Goat anti-Mouse	1:100
Oct3/4 (POU5F1)	Cell Signaling (C52G3)	1:100	Dylight 594-conjugated Donkey anti-Rabbit	1:100
SOX17	R&D (AF1924)	1:100	Dylight 649-conjugated Donkey anti-Goat	1:100

Table showing the concentrations of antibodies used for flow-cytometry staining and the respective secondary antibodies used.

2.18. Apoptosis assay

Cells were washed once with sterile PBS and harvested using accutase. Cells were resuspended in mTeSR and seeded into 4 matrigel-coated flasks at a density of 1×10^5/cm² (for 3, 6, 12 and 24 hour time points). Remaining sample was fixed at room temperature in the dark using 4% PFA and stored in at 4°C until processed. 3, 6, 12 and 24 hour time points were harvested following accutase treatment, fixed and stored as above. Apoptotic cells were processed using the Annexin V-FITC Apoptosis Detection Kit (abcam: ab14085) and analysed by flow cytometry (BD Biosciences: FACSCalibur).

2.19. Cell-cycle analysis

Cell cycle distribution was determined using the Click-iT EdU Alexa Fluor 647 Flow Cytometry Assay Kit (Life Technologies). Cultures ~50% confluence were incubated in mTeSR containing EdU for 45 minutes, harvested using accutase and then processed as per manufacturer's instructions. Samples were analysed by flow cytometry (BD Biosciences: BD LSRII Fllow Cytometer).

2.20. Generation of stable over-expressing cell lines

2.20.1. Constructs

BCL-XL, ID1 and HM13 expression constructs driven by the CAG promoter were prepared as follows: Full-length coding regions were amplified from hESC cDNA using oligonucliotides containing restriction endonuclease sites at the 5´ and 3´ends of the PCR product (see Table 4). BCL-XL PCR product was cloned into the vector pCAG:GFP:IRES:Puro [78, 79] following

XhoI/NotI digestion, which excised and replaced the GFP element. ID1 PCR product was cloned

into the vector pCAG:GFP:IRES:Hygro following BamHI/NotI digestion, which excised and

replaced the GFP element. HM13 PCR product was cloned in the vector pCAG:MCS:IRES:Hygro

following EcoRI/NheI digestion. pCAG:MCS:IRES:Hygro was derived by excising the GFP

fragment from pCAG:MCS:IRES:Hygro through BamHI/NotI digestion and replacing it with a

multiple cloning site. The cloning site was prepared by annealing the following

oligonucleotides:

GATCCGAATTCATCGATAAGCTTCCCGGGCTAGCTGCAGC

GGCCGCTGCAGCTAGCCCGGGAAGCTTATCGATGAATTCG

This resulted in complimentary BamHI and NotI ends. For generation of GFP expressing cells,

the pCAG:GFP:IRES:Hygro vector was used.

2.20.2. Nucleofection

Transfections were performed using the Amaxa Human Stem Cell Nucleofector Kit 1 (Lonza). A

T12.5 matrigel-coated culture flask containing mTeSR supplemented with Rock inhibitor was

incubated at 37°C at 5% CO_2 for 1 hour. Cells to be transfected were washed once with sterile

PBS and harvested using accutase. The cells were pelleted by centrifugation and resuspended

at $8x10^6$/mL in transfection master mix (82uL Solution 1, 18uL supplement). 100uL of cell

suspension was then added to 5ug of plasmid DNA and placed in cuvette. The cuvette was

then placed into a Nucleofector II device (Lonza) and pre-set program B-016 was applied. The

cell suspension was washed from the cuvette with 1mL of media from incubated T12.5 flask

and placed back into humidified incubator at 37°C, 5% CO_2.

Table 4. Primers used to amplify cDNA for generation of over-expressing cell lines

Gene	5' Oligonucleotide	3' Oligonucleotide
HM13	ACGTGAATTCATGGACTCGGCCCTCAGCGAT *EcoRI*	ACGTGCTAGCTCATTTCTCTTTCTTCTCCAGCCCCT *NheI*
ID1	ACGTGGATCCATGAAAGTCGCCAGTGGCA *BamHI*	ACGTGCGGCCGCTCAGCGACACAAGATGCGAT *NotI*
BCL-XL	ACGTCTCGAGATGTCTCAGAGCAACCGGGAG *XhoI*	ACGTGCGGCCGCTCATTTCTCTTTCTTCTCCAGCCCCT *NotI*

Note: The ACGT sequence preceding the restriction site (red) enables endonuclease docking. The sequence following the restriction site is template specific to the gene of interest.

2.21. Teratoma formation

Cells were washed once with sterile PBS and harvested following accutase treatment. The cells were counted, pelleted by centrifugation at 100rpm for 3 minutes and resuspended at 3.3×10^7/mL in 1:3 Matrigel: DMEM/F12 mix (sample kept on ice until injected). 300uL of cell suspension (~1×10^7 cells) was then injected subcutaneously at the inguinal region of NSG mice (Courtesy of Len Shultz, Jackson Laboratory). Injections were performed with approval of the Institutional Animal Care and Use Committee (Biological Resource Centre, Singapore). Mice were euthanized 4 weeks post-injection by raising CO_2 levels. The tumours were excised from surrounding tissue, weighed and fixed in 4% PFA overnight. The tumour was then paraffin-embedded, sectioned and stained with hematoxylin and eosin.

2.22. Time-lapse imaging

2.22.1. Time-lapse imaging for Sections 3.2.5 and 4.2.1

Cell division times were determined by seeding cells (following accutase treatment) at 1×10^4 cells/cm^2 on matrigel-coated 6-well plates supplemented with mTeSR and allowed to adhere for 1 hour at 37°C, 5% CO_2. The plates were then transferred to an Olympus IX600 microscope with fitted incubation chamber set to 37°C, 5% CO_2. Cells were imaged every 10 minutes for 72 hours and image stacks were assembled and analysed using FUJI (ImageJ).

2.22.2. Time-lapse imaging for Sections 6.2.1 and 6.2.2

ESI-035 and ESI-035-CNV were transfected with the pCAG:H2B-RFP:PURO plasmid using microporation. The plasmid over-expresses a histone H2B:RFP (red fluorescent protein) fusion

protein which enables the visualisation of DNA in real-time. Cells were passaged as clumps at a split ratio of 1:3 into matrigel-coated 35mm IBIDI dishes containing mTeSR. Time-lapse movies were generated by imaging cells every minute for two hours using an incubated Nikon TiE inverted fluorescence microscope fitted with an sCMOS camera using a 100X oil objective lens.

2.23.   Definitive endoderm differentiation and staining

A matrigel-coated 6-well plate was fed with mTeSR supplemented with Rock inhibitor and placed at 37°C, 5%CO_2 for 1 hour. Undifferentiated human ES cells were seeded (following accutase treatment) at a density of 6×10^5/well and placed into 37°C humidified incubator at 5% CO_2 overnight. Excess cells were pelleted by centrifugation and resuspended in 600uL of trizol to isolate RNA (day 0 time-point). The mTeSR was replaced 24 hours post-seeding with 3mL endoderm basal media (RPMI 1640, 1X B-27 Supplement, 1X Non-essential Amino Acids, 1X Glutamax (GIBCO)) supplemented with 100ng/mL activin A, 50ng/mL BMP4 and 100ng/mL bFGF (R&D Systems) and placed back at 37°C, 5% CO_2. Media was replaced at day 3 to endoderm basal media supplemented with 100ng/mL activin A and placed back at 37°C, 5% CO_2.

Following five days culture cells were washed once with sterile PBS and 4% PFA was added to cells and incubated at room temperature in the dark for 15 minutes. Cells were washed with wash buffer (PBS containing 5%FBS) and then incubated at room temperature for 1 hour in internalisation buffer (5% FBS, 0.1% Triton X-100 in PBS). The internalisation buffer was then removed and cells were then incubated at room temperature for 1 hour with POU5F1 (Cell Signaling) primary antibody at a 1:100 dilution. The cells were then washed three times with wash buffer and incubated at room temperature for 1 hour in wash buffer containing

secondary antibody. This was then repeated with SOX17 (R&D) primary and secondary antibodies (see Table 3). The cells were counter stained with Hoechst 33342.

2.24.   Cell competition assay

A HES3 normal (non-CNV) GFP reporter cell line was generated by transfecting cells with pCAG:GFP:IRES:Hygro plasmid DNA using nucleofection. The GFP expressing HES3 cells were then mixed at a 9:1 ratio with HES3-CNV, HES3-HM13, HES3-ID1 and HES3-BCL-XL over-expressing cell lines following accutase treatment. A sample of the mix was kept for analysis and 1.2×10^5 cells/cm² were seeded on matrigel-coated tissue-culture treated T12.5 flasks containing mTeSR. Cells were passaged every 4 days using accutase, a sample of the cell suspension was taken for flow cytometry analysis (BD FACSCalibur) and the remaining cells split at a ratio of 1:3.

2.25.   Stable inducible-shRNAi clones

H1 CNV cells were used to generate an inducible BCL-XL knockdown cell line following the procedures detailed previously [80]. The pSUPERIOR (Oligoengine) vector was used to construct a hairpin expression system to target BCL-XL. Annealed oligonucleotide pairs were:

GATCCGGAGATGCAGGTATTGGTGTTCAAGAGACACCAATACCTGCATCTCCCTTTTTC and

TCGAGAAAAAGGGAGATGCAGGTATTGGTGTCTCTTGAACACCAATACCTGCATCTCCG

Cells in which knockdown was desired were treated with 100 ng/ml doxycycline (Sigma) three days prior to experimental setup.

2.26. BCL-XL inhibition

Cells were harvested following accutase treatment and plated at a density of 1×10^4 cells/cm² on matrigel-coated tissue-culture treated 6-well plates in mTeSR. The cells were cultured for 5-7 days in the absence or presecnce of ABT-263 (Selleckchem) at 50, 100, 250 and 1000mM concentrations. Cell were fixed for 20 minutes in ice-cold 100% methanol and stained for 3-4 hours with Giemsa stain.

2.27. Embryoid body formation

2.27.1. Plate set-up

For each cell line four 96-well plates were prepared, one for each of the four conditions (neutral, ectoderm, mesoderm and endoderm). 3.5mLs of APEL medium was added to four 15mL falcon tubes containing differing amounts of growth factors and inhibitors (see Table 5). Once the solutions had been made, the inner 60-wells of each plate were filled with 50uL of growth factor medium (2X), the remaining outer 36-wells were filled with 100uL of PBS to avoid evaporation. Once the plates were made they were placed at 37°C at 5% CO_2 until ready to use.

2.27.2. Harvesting and plating Cells

Media was aspirated from 70-80% confluent hESC cultures and 0.5mL of TrypLE was added. The cells were then incubated at 37°C for 2 minutes and flasks were tapped sharply 3-5 times to dissociate differentiated cells. Cells were aspirated and a further 0.5mL of TrypLE was added to the flask. The cells were incubated at 37°C for 2 minutes. The flask was tapped

sharply to dissociate all other cells and 3mLs of mTeSR was used to inactivate the TrypLE. Cells were counted using the method described in section 2.7 were then spun at 1000rpm for 3 minutes and the supernatant extracted. The cells were then resuspended at 8.4×10^5 in APEL medium containing genatmycin (1:500). 50uL of cell suspension (3000 cells) was added to each well and the plates spun at 1000rpm for 3 minutes to aggregate cells at the bottom of the well. The plates were then placed in the incubator and left to grow for 10 days.

Table 5. Growth factors used to different ate human ES cells via the formation of EBs

	Neutral	Ectoderm	Mesoderm	Endoderm
Activin A	X	X	20ng/mL	100ng/mL
BMP4	X	X	20ng/mL	1ng/mL
bFGF	X	100ng/mL	X	X
DMH-1	X	1uM/mL	X	X
SB431542	X	10uM/mL	X	X

The growth factors and their concentrations added to the dif erent EB conditions, neutral, ectoder, mesoderm and endoderm.

3. Amplification of 20q11.21 Provides a Selective Advantage in Human ES Cells.

3.1. Introduction

Human ES cells often acquire extra copies of chromosomal material; these karyotypic changes most frequently involve the gains of chromosomes 1, 12, 17 and 20. These changes have been shown to provide variant cells with a selective advantage over diploid cells[73]. Interestingly, the commonly gained chromosomes often present as smaller amplifications notably 1q, 12p and 17q are more frequently observed than 1p, 12q and 17p. This suggests that smaller amplifications of genetic material may contain the gene(s) responsible for causing such an advantage. The ISCI study[66] used high-resolution molecular karyotyping to try and identify small structural variants that may be driving culture-adaptation. No copy number variants (CNVs) were observed on chromosomal regions 1q, 12p or 17q suggesting that whole or partial chromosome gains are required to provide a selective advantage. However, a small CNV on chromosome 20 (20q11.21) was identified as a possible driver of culture-adaptation. Amplification of 20q11.21 has been identified as a hotspot for copy number variation in human ES cells[81-84] and is widely reported in a number of human cancers[85-88]. The mechanism behind the selection of 20q11.21 in human ES cells is unknown and there are no reports of the growth rates of cells containing the amplification. There have been hypothesised roles for ID1, BCL2L1 and TPX2 in driving the selection of 20q11.21 *in vitro* and *in vivo*[82]. The minimal amplicon determined by SNP-array does not contain TPX2 and was therefore dismissed as a candidate driving the selection of 20q11.21 although is still within most 20q11.21 amplifications observed in human ES cells. The minimal amplicon spans approximately 60Kb and contains 13 coding genes, of which, only three are expressed in undifferentiated human ES cells:

HM13

The HM13 gene encodes a signal peptide peptidase, localised to the endoplasmic reticulum and catalyses intra-membrane proteolysis releasing signal peptides into the cytosol. The role of HM13 in human ES cells has not been investigated and is an unlikely candidate driving selection of the 20q11.21 locus. However, it has been shown to influence anchorage independent growth of SW480 cells [89].

ID1

ID (inhibitor of differentiation or DNA binding) proteins are dominant negative antagonists of basic helix-loop-helix (bHLH) transcription factors, forming heterodimers and inhibiting their DNA binding capacity. The structure of ID proteins is similar to other bHLH proteins, containing a helix-loop-helix domain but lacking the DNA binding domain of bHLH transcription factors. Increased levels of ID1 in human cancers are often correlated with aggressive tumours and poor prognosis[90]. The over-expression of ID1 in hepatocellular carcinomas has been shown to promote cell survival through activation of the NF-Kβ signalling pathway[91] and induce cell proliferation through inactivation of the p16INK4a/RB signaling pathway[92]. The increased expression of ID1 could also have an effect on the differentiation on human ES cells as it has been shown to have a high binding affinity to MYOD and NEUROD, markers of mesoderm and ectoderm respectively[90].

BCL2L1

BCL2L1 is also a strong candidate gene that could be driving selection of 20q11.21, the gene encodes two isoforms, BCL-XL (long isoform) and BCL-XS (short isoform). The two isoforms elicit anti- and pro-apoptotic functions respectively. In human ES cells the dominant isoform is BCL-XL (Figure 2). BCL-XL has been implied in a number of human cancers[93-95]. BCL2L1 amplification has been shown to increase the survival of colorectal cancer cells (CRCs) and

inhibition of BCL2L1 in CRC cells results in decreased cell viability and inhibited anchorage-independent growth [96].

The high prevalence of 20q11.21 amplification in human ES cell cultures suggests that it provides a selective growth advantage over diploid cells. To assess the growth rates of human ES cells harbouring the 20q11.21 gain, four independent human ES cell lines containing the CNV were obtained. The cell lines were also obtained at earlier passages to represent diploid cells with which to compare the effect of 20q11.21 amplification on the growth of human ES cell cultures.

Figure 2. BCL-XL is the dominant isoform expressed in human embryonic stem cells
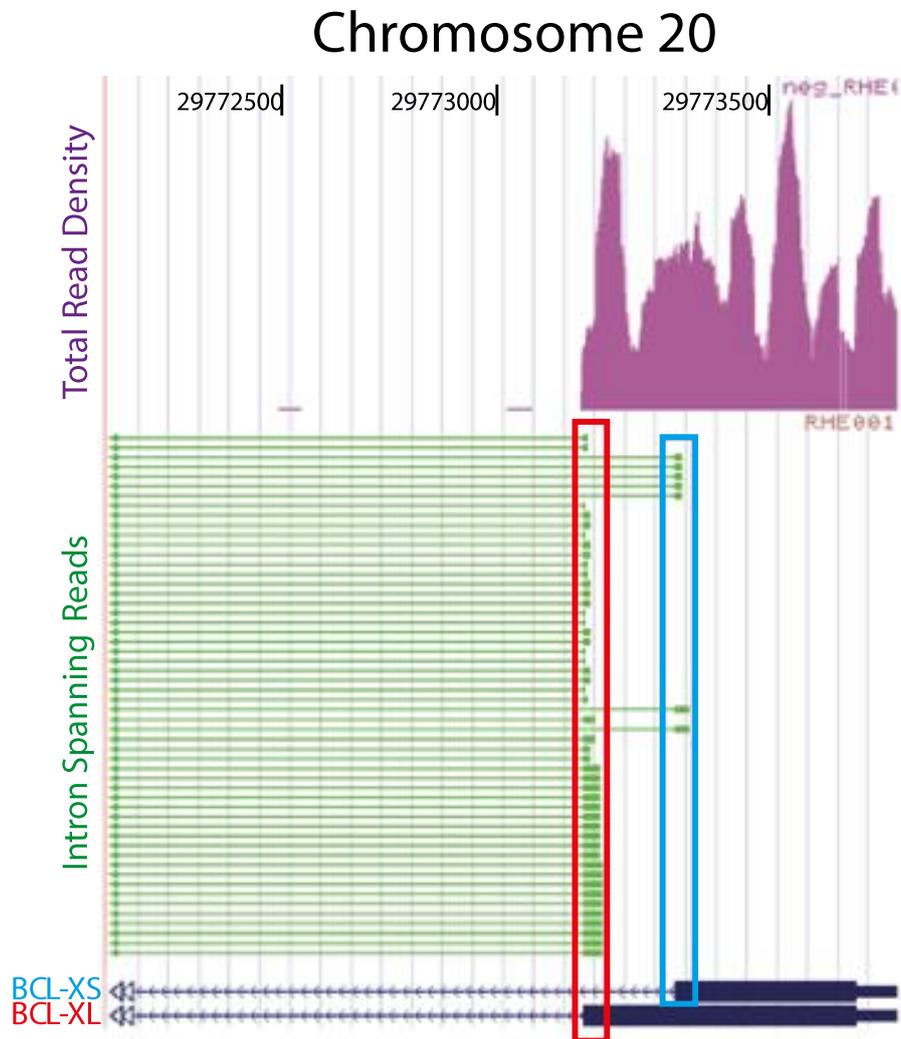
Using mRNA-Seq the intron spanning reads of BCL-XL (red bar) and BCL-XS (blue bar) were measured in the undifferentiated H1 human ES cell line. Data shows that BCL-XL is the dominant isoform expressed with 47 reads compared to BCL-XS which only displayed 7 reads. mRNA-Seq data was mapped to Ref-Seq.

3.2. Results

3.2.1. Varying lengths of 20q11.21 duplications in human ES cell lines

To study the effect that amplification of chromosomal region 20q11.21 has on the culture of human ES cells, four different cell lines were obtained, HES3[97], H1[1], ESI-035[98] and Shef5[99]. These cell lines were chosen based on SNP array data showing the presence of 20q11.21 amplification at late passage[66]. Each cell line was obtained in early and late passage, representing a 'control' and 'CNV' cell line respectively. Upon receipt, the cell lines were expanded, karyotyped and frozen to create working reference banks. The cell line banks were analysed for the absence/presence of the 20q11.21 CNV using genomic DNA based qRT-PCR and fluorescence in situ hybridization (FISH). Each of the cell lines received had a normal diploid karyotype with the exception of Shef5-CNV in which one of the X chromosomes was absent. qRT-PCR indicated the presence of the 20q11.21 amplification in all four CNV lines (Figure 3A) which was confirmed by FISH analysis for the BCL2L1 locus (Supplementary Figures 1-4). However, the control HES3 and H1 cells received were found to display a degree of mosaicism for the 20q11.21 amplification, likely highlighting the high propensity of cells to acquire the amplification during culture. The dosage of 20q11.21 in control cells was lower than that found in the CNV cell lines obtained (average 20q11.21 loci: HES3 control 2.2, HES3-CNV 3.5 and H1 control 2.5 and H1-CNV 4.2). The mosaicism observed in CNV cell lines could indicate the potential for CNV cells to acquire multiple 20q11.21 duplications throughout culture. The 20q11.21 amplification was also present at varying lengths (Figure 3B). The HES3 CNV contained the longest amplification of around 1.5Mb whereas the H1 (1.1Mb), ESI-035 (0.8Mb) and Shef5 (0.7Mb) contained the amplification at shorter lengths. All CNV cell lines contained the minimal amplicon identified in the ISCI project [66]. The closest cell line to resemble the minimal amplicon was the Shef5-CNV cell line.

Figure 3. Quantitat ve-PCR detects the presence of CNV 20q11.21 in Shef5, ESI-035, H1 and HES3 human embryonic stem cell lines
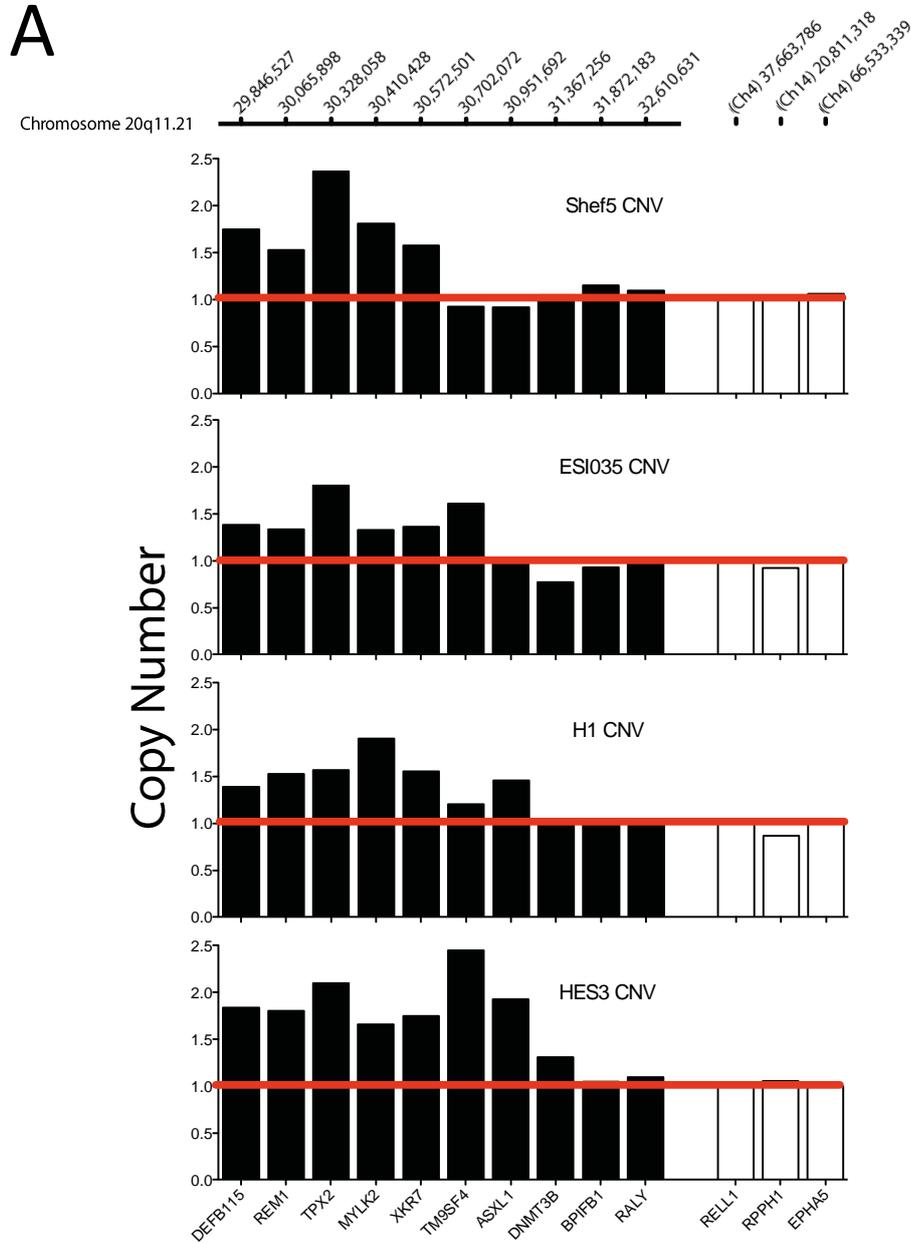
(A) Genomic qRT-PCR assay using primer/probe pairs designed to intronic regions of genes spanning the 20q11.21 locus (black bars) determines the amplicon length and copy number. Genomic positions relate to USCS human genome assembly version hg19{Kent, 2002 #236}. CT values are normalised against RELL1 (first white bar). RELL1 is located on chromosome 4 and displays a low incidence of genomic instability in hESCs. Two additional controls (white bars) confirm the suitability of the first control. All CNV cell lines were normalised against their control counter-parts.

Figure 3. Quantitat ve-PCR detects the presence of CNV 20q11.21 in Shef5, ESI-035, H1 and HES3 human embryonic stem cell lines

continued
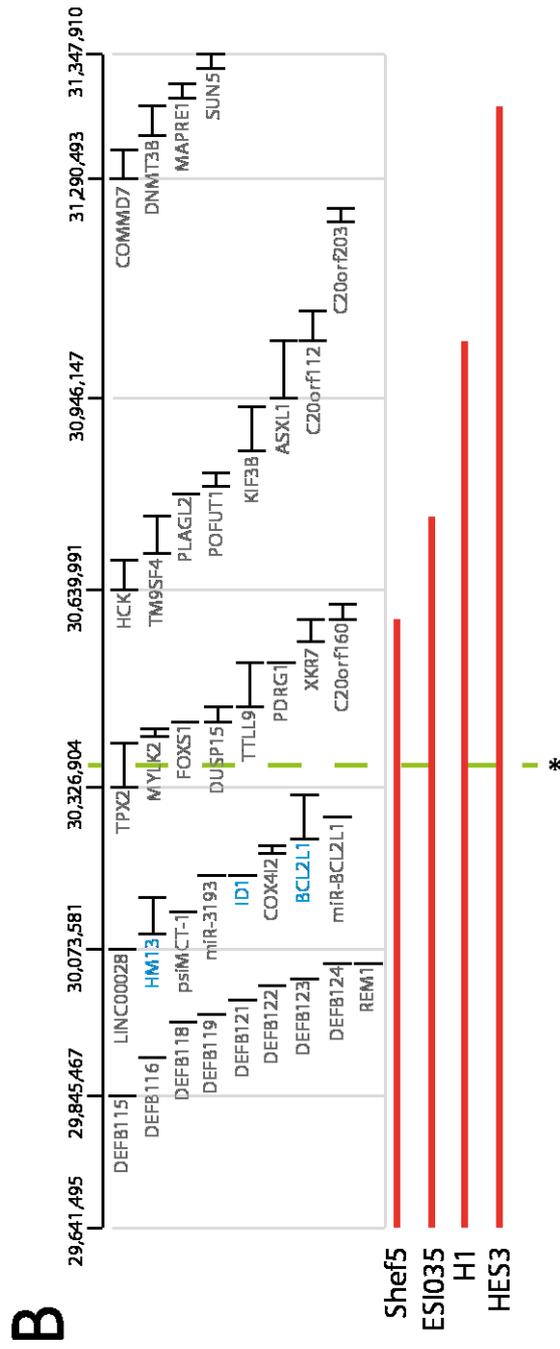
(B) Schematic representaton of ampli con lengths for the four test human ES cell lines (red lines) positioned alongside genes contained within the 20q11.21 locus. Green dotted line and asterisk represent the minimal amplicon previously described in human ES cells and genes in blue as those candidate genes located within the minimal amplicon and expressed in human ES cells.

3.2.2. Generation of Over-expression Cell Lines

To analyse the affect that increased dosages of HM13, ID1 and BCL-XL had on human ES cell culture, over-expression cell lines for each candidate gene were generated. All over-expression vectors were derived from either pCAG:GFP:IRES:PURO or pCAG:GFP:IRES:HYGRO plasmids (Figure 4A). The expression of green fluorescent protein (GFP) is driven by a CAG (CMV early enhancer/chicken β-actin) promoter, the CAG promoter simultaneously drives the expression of puromycin/hygromycin through the presence of an internal ribosomal entry site (IRES) located between the GFP and antibiotic resistance sequences. These vectors have previously shown strong, sustained expression in human ES cells [79, 80]. The pCAG:GFP:IRES:HYGRO vector was used to generate a HES3 control GFP-expressing cell line. To generate HM13, ID1 and BCL-XL over-expressing lines the GFP element was removed by restriction digest with *BamHI* and *NotI* and a multiple cloning site (MCS) was inserted to make pCAG:MCS:IRES:PURO and pCAG:MCS:IRES:HYGRO (Figure 4B). The MCS was created by annealing the sequences GATCCGAATTCATCGATAAGCTTCCCGGGCTAGCTGCAGC and GGCCGCTGCAGCTAGCCCGGGAAGCTTATCGATGAATTCG, which results in complimentary *BamHI* and *NotI* ends for correct orientation into the pCAG vector.

Each candidate gene was PCR amplified from ESI-035 control cDNA using primers with attached restriction sites (see Table 4). The products of HM13, ID1 and BCL-XL were inserted into the MCS of either the PURO or HYGRO vector depending on the suitability of restriction sites. The vectors were then transfected into HES3 and ESI-035 control lines using nucleofection (see Section 2.20). Stable transfectants were selected using the appropriate antibiotic and expanded to make working cell banks.

The ESI-035 over-expression cell lines were analysed for the characteristic human ES cell surface markers TRA-1-60 and SSEA3 using flow cytometry (Figure 5). All over-expressing cell lines showed similar expression of the stem cell markers compared with control and CNV-containing cells, demonstrating that the cells derived maintained an ES cell phenotype and that the introduction of HM13, ID1 and BCL-XL did not affect the stem cell population through increased differentiation.

Figure 4. Plasmids used in the generation of GFP-reporter cell lines and cell lines over-expressing HM13, ID1 and BCL-XL
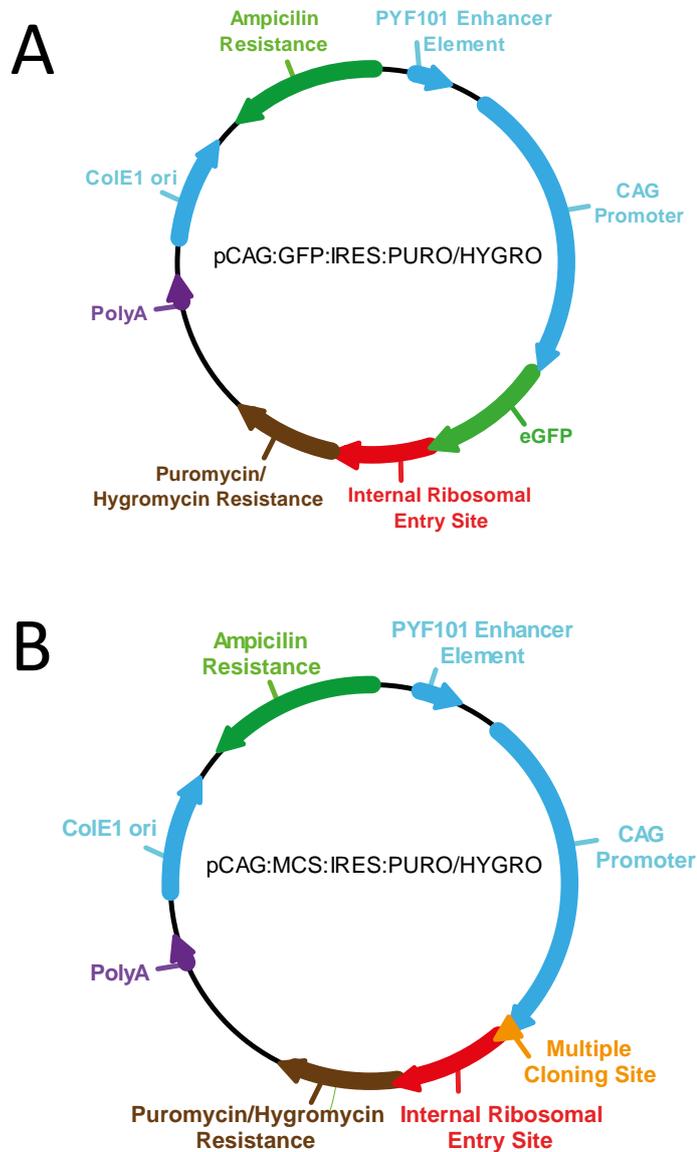
(A) The pCAG:GFP:IRES:PURO/HYGRO vector was used to generate GFP-reporter cell lines. The expression of green fluorescent protein (GFP) is driven via the CMV early enhancer/chicken β-actin promoter (pCAG). Antibiotic resistance to puromycin or hygromicin is also expressed simultaneously though the presence of an internal ribosomal entry site (IRES) to allow the selection of stable transfectants. (B) The pCAG:MCS:IRES:PURO/HYGRO was used to generate cell lines over-expressing the three candidate genes. cDNA for HM13, ID1 and BCL-XL was inserted into the multiple cloning site using compatable restriction sites.

Figure 5. Over-expression of HM13, ID1 and BCL-XL does not affect human embryonic stem cell markers
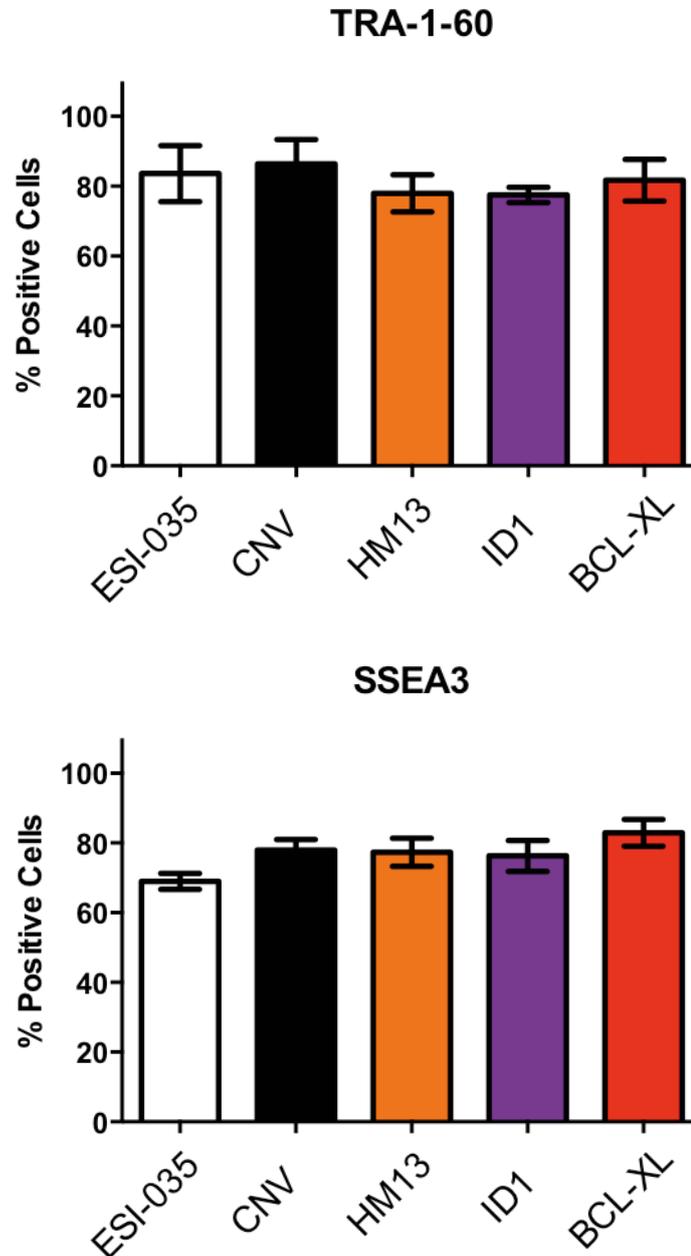
TRA-1-60

SSEA3

Over-expression of HM13 (orange bar), ID1 (purple bar) and BCL-XL (red bar) in human ES cells does not affect the expression levels of human ES cell markers (A) TRA-1-60 or (B) SSEA3. The levels are comparable with control (white bar) and CNV (black bar). Graphs show data for the over-expression cell lines generated from ESI-035 Control cells.

3.2.3.   CNV Cell Lines Exhibit Increased Population Doubling Times

To determine whether there was a difference in the growth rates between control and CNV containing cell lines, the population doubling time of each cell line was measured. Cells were plated on Matrigel in mTeSR at varying cell densities ($2x10^4$, $4x10^4$ and $8x10^4$ cells/cm²) representing clonal (sub-optimal) to routine passaging (optimal) densities. Cells were cultured for four days and the total cell counts measured. Figure 6 shows the final density of cells (cells/cm²) following four days culture. In all four cell lines, CNV containing cells displayed decreased population doubling times when compared with their control counterparts. This increase in population growth rates was observed at all three plating densities. The population doubling rate was determined by taking the total number of cells following four days culture and dividing by the initial number of cells seeded. At the lowest plating density ($2x10^4$ cells/cm²) all control cell lines were unable to surpass the original seeding density with an average population doubling rate of 0.5 over four days. This is in stark contrast to the CNV cell lines, which displayed an average population doubling rate of 4.7 at the lowest seeding density. This trend was also observed in the two other seeding densities, at $4x10^4$ cells/cm² the control cells displayed an average population doubling rate of 1.1 compared to the CNV average of 5.9. At the highest seeding density of $8x10^4$ cells/cm² the population doubling rate of control cells was higher than the initial seeding density following four days growth at an average rate of 2.0. The average of CNV cell lines was again much higher than that observed in control cell lines with an average rate of 6.1. This increase in growth rate was also observed in cells over-expressing BCL-XL mimicking that of the CNV cells with population growth rates of 6.3, 5.6 and 6.6 for the lowest to highest seeding densities respectively. Both HM13 and ID1 over-expressing cell lines also displayed higher population doubling rates than the average

control cell lines. However, the rates were similar to the rates observed in the ESI-035 control cell line from which the over-expressing cell lines were derived.

3.2.4.   Increased Growth Rates of CNV Cells Rapidly Out Compete Control Cells

The increase in population doubling rates in CNV cells indicated that amplification of 20q11.21 provides a growth advantage in human ES cells. To model the effect this has on the population dynamics in culture, cell mixing experiments were performed. The HES3 control cell line was transfected with the pCAG:GFP:IRES:PURO vector to generate a stable reporter line with which to monitor the proportions of cells within a mixed population (see section 3.2.2). The HES3-GFP line was mixed with test cell lines, HES3 control (parent line, no-GFP), HES3-CNV, HES3-HM13, HES3-ID1 and HES3-BCL-XL at a ratio of 9:1. Cells were cultured for 10 passages and the levels of GFP expressing cells measured using flow-cytometry (Figure 7). The HES3-GFP was mixed with the HES3 control cell line to determine that the random integration of the GFP vector had no effect on the growth of the cell line. The HES3-CNV and HES3-BCL-XL cell line rapidly out-competed the HES3-GFP cell line within ten passages. This was not observed in HES3 control, HES3-HM13 and HES3-ID1 cell lines, which maintained the levels of GFP-expressing cells at the initial seed ratio of 9:1.

Figure 6. CNV-containing cell lines display increased population doubling rates when compared to control cell lines
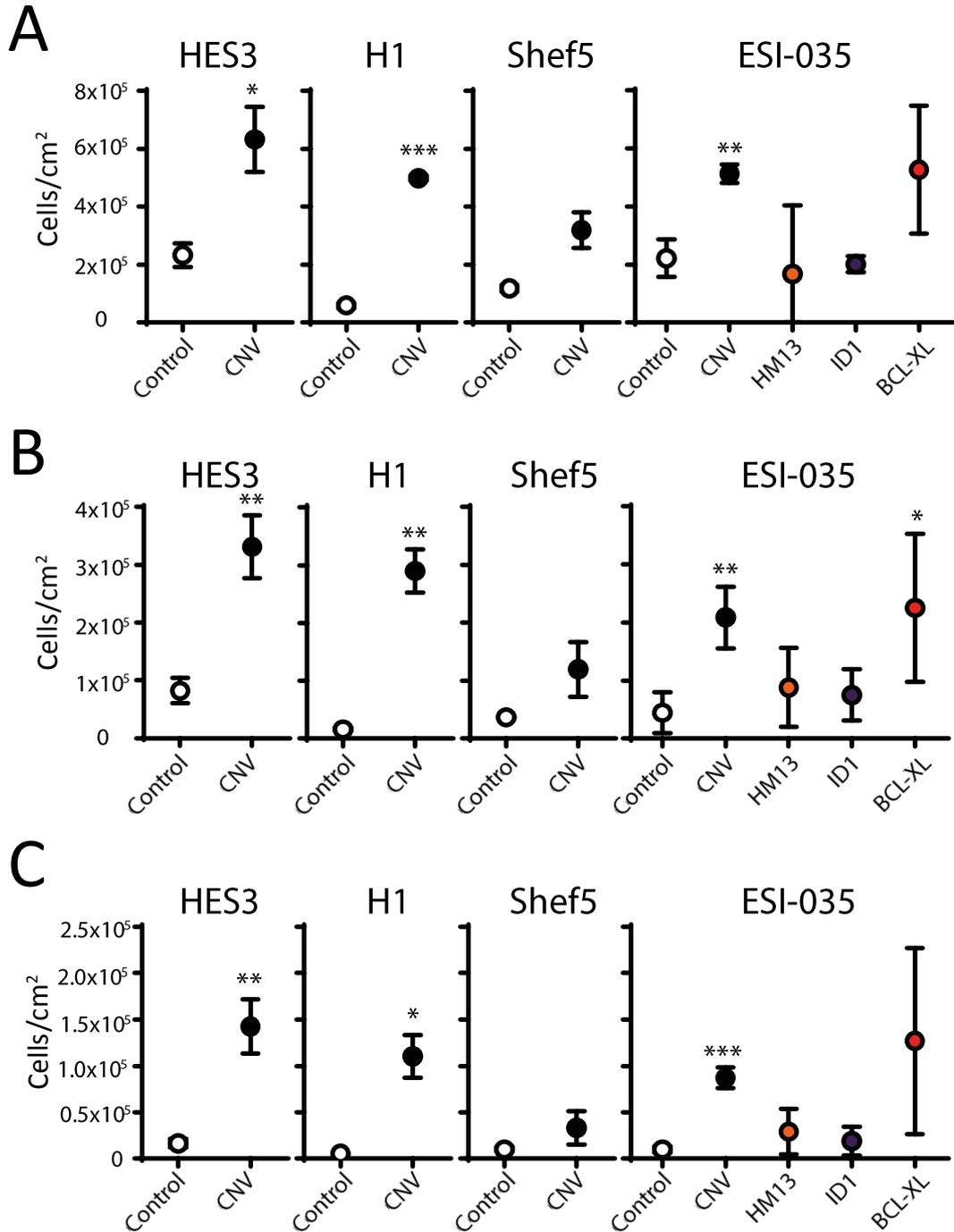
Cell densities (cells/cm²) following four days culture af er seeding at (A) 8x10⁴, (B) 4x10⁴ and (C) 2x10⁴ cells/cm². CNV cell lines (black circles) display increased population doubling times when compared with control (white circles) cell lines. This increase is mimicked by over-expression of BCL-XL (red circle) but not HM13 (orange circle) or ID1 (purple circle). Errors bars represent SEM from three biological replicates. Asterisks represent statist cal significance by two-tailed t-tests: p ≤ 0.05 (*), p ≤ 0.01 (**), p ≤ 0.001 (***).

Figure 7. Increased growth rates of CNV-containing cells rapidly out-compete diploid cells
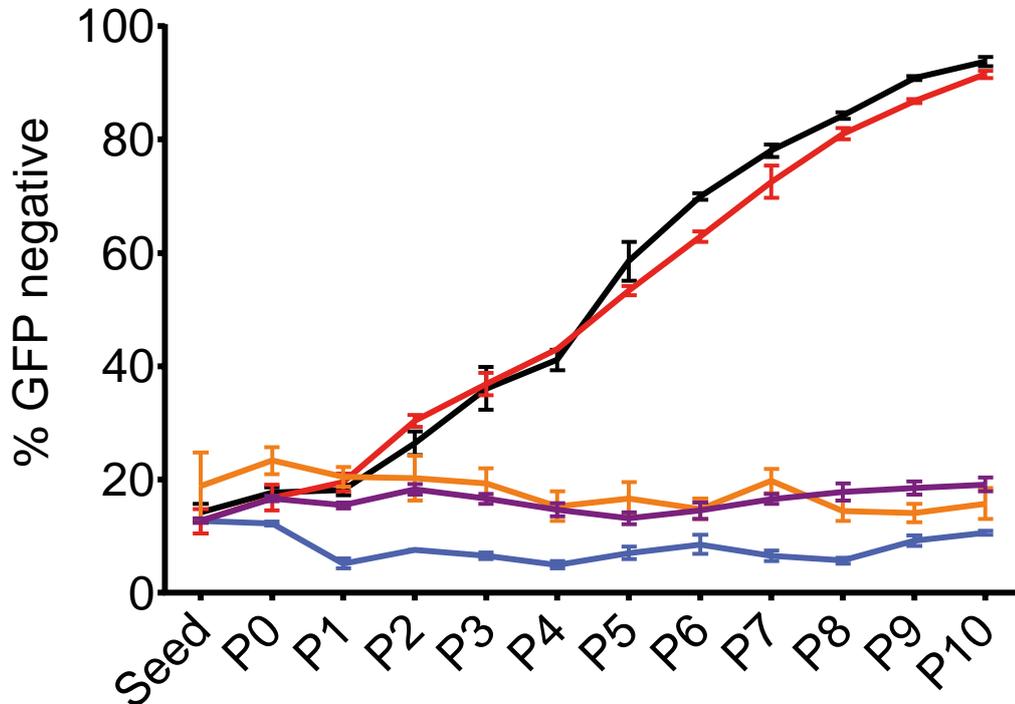
Cell competition assay using the HES3 cell line. The HES3-GFP-Control cell line was mixed at a 9:1 ratio with HES3-Control (parental line, no-GFP) (blue line), HES3-CNV (black line) and the over-expressing cell lines, HES3-HM13 (orange line), HES3-ID1 (purple line) and HES3-BCL-XL (red line). The cells were then passaged at a 1:3 split ratio for ten passages and the percentage of GFP analysed using flow-cytometry. The HES3-CNV and HES3-BCL-XL cell lines rapidly out-competed the HES3-GFP-Control cell line within ten passages. Errors bars represent SEM from three biological replicates.

3.2.5. Control and CNV Cells Display Similar Cell Cycle Distribution and Cell Cycle Times

To determine the mechanism behind the increased growth rates of CNV containing cell lines the cell cycle distribution between control and CNV cell lines were analysed (Figure 8). All cell lines showed high proportions of S-phase cells, with equal G1 and G2/M populations. The Shef5 and ESI-035 cell lines exhibited similar cell cycle distribution between control and CNV lines but the CNV lines appeared to contain slightly less cells in S-phase when compared with the control cells (ns). The HES3-CNV and H1-CNV cell lines showed significantly lower populations of S-Phase cells and increased G0/1 populations than their control counterparts. This effect was also observed in ESI-035 cells over-expressing BCL-XL, again highlighting similarities between CNV and BCL-XL over-expressing cell lines. No significant difference was observed in the G2/M populations in any of the four cell lines.

Actual cell doubling times were determined using time-lapse imaging (Figure 9), cells were plated at $1x10^4$ cells/cm^2 and imaged at 10 minute intervals. Cellular divisions were monitored manually and cell cycle times were calculated. The range of division times observed were extremely variable (7-44 hours), however, the average division times between control and CNV lines were almost identical between individual cell lines, with the exception of the H1 pair. The H1-CNV cell line displayed faster division times. However, this was not of statistical significance and can be explained by very uniform division times in CNV compared with the H1-control, which displayed a wider distribution of cell cycle times. Individual cell lines also showed a degree of variation in their division times, with Shef5 having slightly slower cell cycling times. The average division times between control and CNV-containing cell lines were 19.5 and 18.5 hours respectively. However, when these times are compared with the cell

division times calculated from population doubling experiments in Section 3.2.3 (138 and 35 hours respectively) there was a marked difference.

3.2.6.   RNA Expression in Control and CNV Cell Lines

The duplication of chromosomal region 20q11.21 in CNV cell lines increases the amount of template DNA available for gene expression. To determine the effect this increase in DNA copy number had on gene expression in undifferentiated human ES cells, quantitative RT-PCR (qRT-PCR) was used to analyse the expression levels of the three candidate genes in the 20q11.21 region (Figure 10). Total RNA was extracted from 70-80% confluent human ES cell cultures under normal growth conditions. The gene expression levels of POU5F1, TPX2 and DNMT3B were also analysed; TPX2 resides just outside the minimal amplicon but all four CNV-cell lines within the study contained the amplification of TPX2. DNMT3B is approximately 1MB distal (from the centromere) of the minimal amplicon and is only present in the HES3-CNV line. The dCT values relative to GAPDH, are displayed in Figure 10. The HES3-CNV cell line had slightly higher expression of HM13, ID1 and BCL-XL (ns) as displayed by lower dCT values. The HES3-CNV cell line also showed increased gene expression of TPX2 and DMNT3B which also reside on the 20q11.21 amplification in the HES3-CNV cell line. The levels of POU5F1 were also higher in the CNV cell line than in the HES3 control cells. Gene expression levels in the H1 cell lines showed no significant difference between control and CNV cells. Both HM13 and BCL-XL appear to be upregulated in CNV cells whereas POU5F1, HM13 and TPX2 show similar levels. DNMT3B, which is absent in the H1-CNV shows similar mRNA levels between control and CNV cell lines. The Shef5 control and CNV cell lines displayed no significant difference in the gene expression levels of HM13, ID1, BCL-XL or TPX2. However, there was a significant difference in the expression levels of POU5F1 (p=0.0031) and DNMT3B (p=0.0210). The ESI-035-CNV cell line showed increased expression of POU5F1 (ns), HM13 (ns), ID1 (p=0.0013) and TPX2

(p=0.0027). The expression levels of BCL-XL remained similar between control and CNV lines and DNMT3B showed increased expression in the control cell line (ns). The differences in mRNA expression between cell lines may represent cell line specific expression patterns. However, the differences between control and CNV cell lines within a given cell line imply that the presence of CNV20q11.21 effects the expression patterns of genes contained within the amplicon. Control cells over-expressing BCL-XL mirrored the growth rates of CNV-containing cells, although qRT-PCR data does not show significantly increased BCL-XL mRNA levels.

3.2.7. Increased BCL-XL Protein Levels in 20q11.21 CNV-containing Cell Lines

The mRNA levels of BCL-XL were not increased consistently in all CNV-containing cell lines. Western-blotting was used to analyse the levels of BCL-XL in the four human ES cell lines to determine if the increase in DNA copy number effected the level of BCL-XL protein. Protein was harvested from 70-80% confluent cultures that had been grown for three to four days on matrigel-coated culture vessels in mTeSR. BCL-XL protein levels were higher in all CNV cell lines when compared with their control counter-parts (Figure 11). There was no significant difference in the levels of BCL-XL mRNA, suggesting that BCL-XL may be stabilised by post-translational modification. Consistent with RNA-seq data, the short isoform of BCL2L1 (BCL-XS) was not detected using western-blot confirming that BCL-XL is the dominant isoform in human ES cells. There was slight variation in BCL-XL levels between the different cell lines, the H1 and HES3 control cell lines showed higher BCL-XL levels than the Shef5 and ESI-035 control lines. This can be explained by the mosaicism of the 20q11.21 amplification in the HES3 and H1 cell lines.

Figure 8. Control and CNV cell lines display similar cell cycle distribution
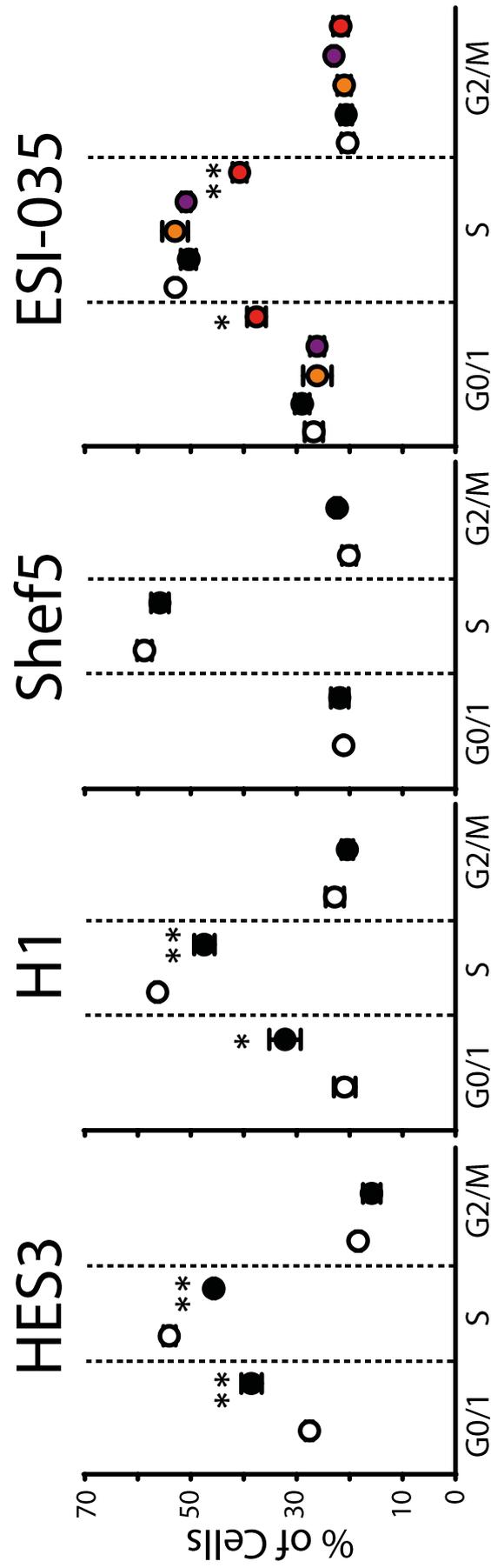
The cell cycle distribution of Control (white circles), CNV (black circles) and cell lines over-expressing HM13 (orange circle), ID1 (purple circle) and BCL-XL (red circle). The proportions of cells residing in G0/1, S and G2/M are displayed as percentages with error bars representing SEM from at least three biological replicates. Asterisks represent statistical significance by two-tailed t-tests: p ≤ 0.05 (*), p ≤ 0.01 (**).

Figure 9. Control and CNV cells display similar cell division times
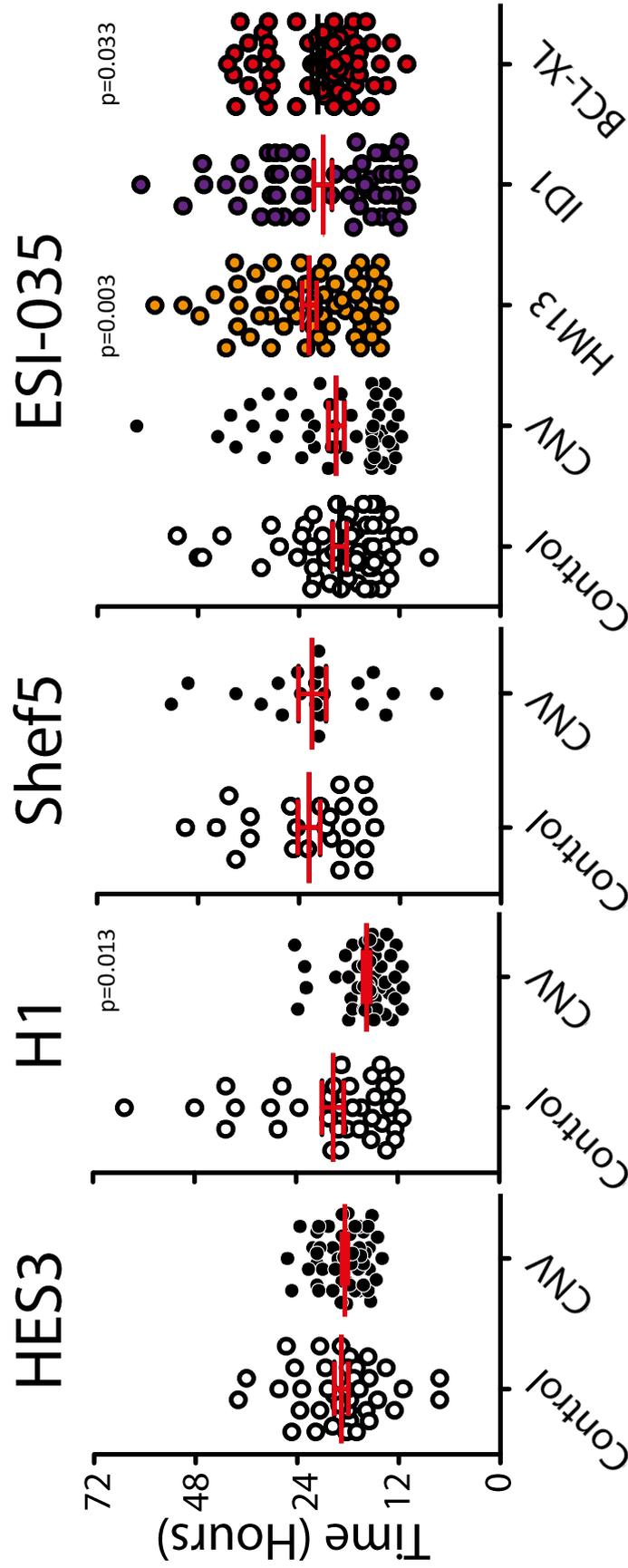
The actual cell division times were determined using live-cell time-lapse microscopy. The time between two cell divisions was measured by manually tracking individual cells; each circle represents one cell division. There was very little difference between Control (white circles) and CNV (black circles) cells within the same cell line and also between different human ES cell lines. Over-expressing cell lines were also analysed, with HM13 (orange circles) and BCL-XL (red circles) displaying slightly longer cell division times. ID1 cells (purple circles) showed no significant difference in cell division times. Error bars indicate SEM and p values are displayed above those cell lines that displayed significantly different cell division times compared with their respective control cell line.

Figure 10. Inter- and intra-cell line differences in mRNA gene-expression
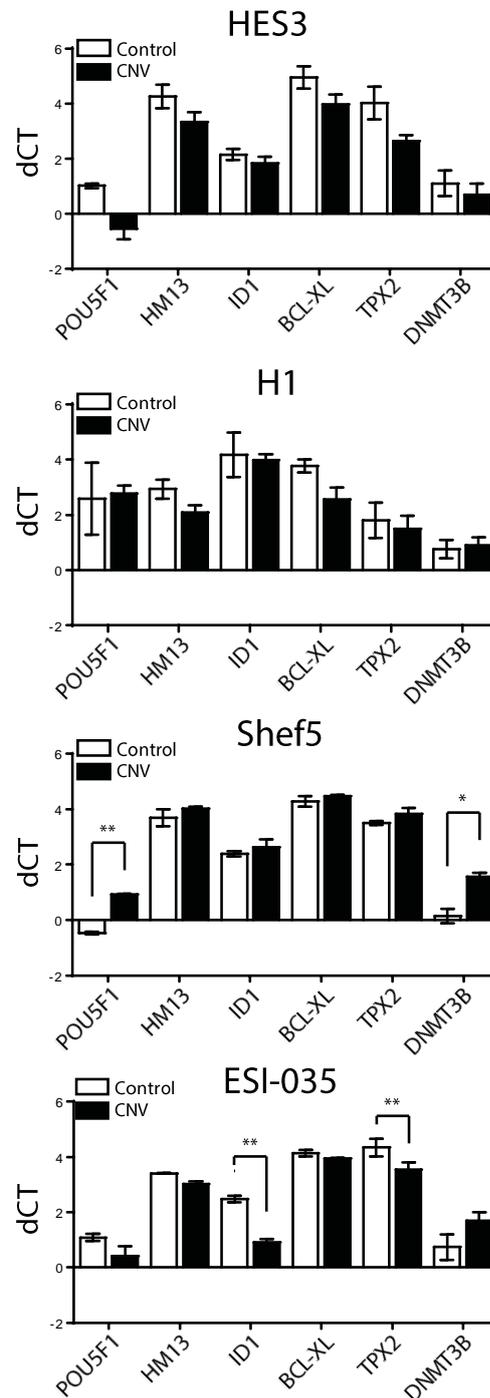
qRT-PCR comparing the gene expression levels of each control (white bars) and CNV (black bars) cell line used in the study. All genes analysed are found on chromosomal region 20q11.21 with the exception of POU5F1, which is an marker of pluripotency. dCT values are relative to the reference gene GAPDH and error bars represent SEM from three biological replicates. Asterisks represent statistical significance by two-tailed t-tests: $p \leq 0.05$ (*), $p \leq 0.01$ (**).

Figure 11. CNV-containing cell lines display increased BCL-XL protein levels
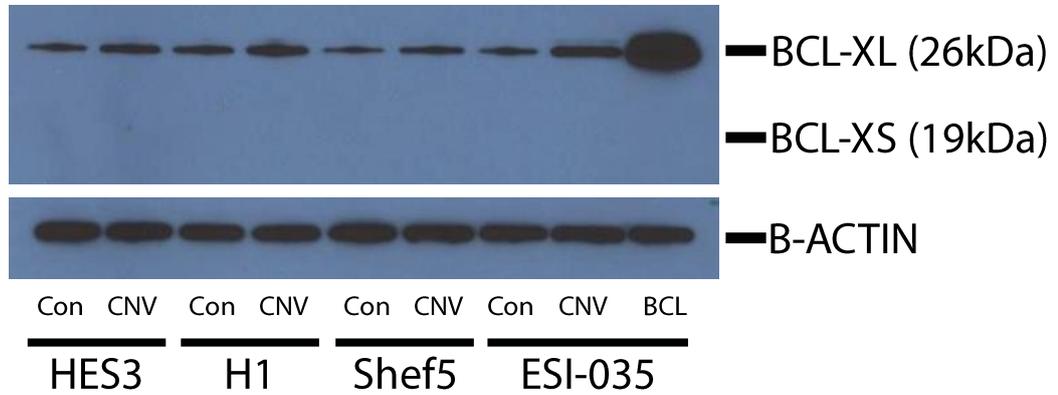
Western blot analysis on whole cell lysates show that BCL-XL protein levels are elevated in CNV and BCL-XL (BCL) over-expressing cells when compared to Control cells (Con). A universal BCL-X antibody detected the presence of BCL-XL protein but not BCL-XS confirming that the anti-apoptotic form of BCL2L1 is the dominant isoform expressed in undifferentiated human ES cells. Beta-actin was used as a loading control.

3.3. Discussion

The prevalence of 20q11.21 amplification in human ES cell lines can be attributed to the significant growth advantage that it provides. Cell lines containing the CNV consistently showed increased population doubling rates when compared with their diploid counterparts. This increase in growth rate can be mimicked by over-expression of BCL-XL in control cells, making BCL-XL a strong candidate driving the selection of 20q11.21 amplification in human ES cultures. The over-expression of HM13 and ID1 in control cells does not affect the growth rates of the cells.

Of the four cell lines obtained, two contained the CNV in a proportion of cells at early passage potentially highlighting the propensity for human ES cells to acquire the 20q11.21 amplification. HES3, which had an average of 2.2 copies per cell and H1 which had an average of 2.5. The corresponding CNV lines contained further amplification of 20q11.21 at 3.5 and 4.2 respectively. Interestingly, the HES3 and H1-CNV cell lines exhibited the highest growth rates possibly a result of increased CNV 20q11.21 gain. The cell lines obtained showed similar lengths to those found in the ISCI study and all contained the minimal amplicon. The length of CNV varies slightly between the cell lines obtained encompassing more genes with increasing length. This is interesting as CNV cell lines with larger amplicon lengths (HES3, H1) exhibited increased growth rates. This could be a result of other genes having an effect further along the CNV.

The population doubling times show that cell lines containing the 20q11.21 amplification have an increased growth capacity when compared with control lines. This characteristic is mimicked by the ESI-035-BCL-XL over-expressing cell line but not HM13 or ID1 suggesting that

BCL-XL is providing the growth advantage in CNV-containing cells. All cell lines showed variation in population doubling times both between different control cell lines and between CNV cell lines. HES3 control and ESI-035 control cell lines showed similar population doubling times in the highest seeding density, however at lower densities the HES3-control seemed less dependent on seeding density whereas the ESI-035 showed decreased population doubling times. This may be due to the HES3-control cell line having the CNV in a proportion of cells, which would explain slightly higher growth rates. Overall the CNV cell lines seem less dependent on initial seed densities with the exception of the Shef5-CNV cell line. The differences in the cell lines could be caused by both the mosaicism found within the cell lines and also the length of the CNV. At the highest seeding density which is most representative of normal passaging densities the CNV cell lines exhibit an average cell division time of 35 hours when compared to the average control cell lines of 138 hours. The selective advantage afforded by the amplification of 20q11.21 is extremely high, even at normal passaging levels and therefore would out-compete control cells rapidly in a small number of passages.

The mixing experiment was conducted to model the selective advantage of the CNV cells throughout culture *in vitro*. A GFP-expressing cell line was generated from the HES3-control cell line using the pCAG:GFP:IRES:HYGRO to create a stable line which could be detected using flow-cytometry. The HES3-GFP cell line was mixed with the parent HES3-control cell line at a ratio of 9:1, this was carried out as a control for non-random integration of the over-expression plasmid into the genome. The levels of green: non-green remained at the initial seeded levels over ten passages showing that integration of GFP into the HES3-control line was neither advantageous nor deleterious to the HES3-control cell line. The initial seed ratios remained the same when HES3-GFP was mixed with HM13 and ID1 over-expressing cells again showing that they are unlikely candidates providing the growth advantage in the 20q11.21 amplicon especially in undifferentiated conditions. The HES3-CNV and HES3-BCLXL cell lines

rapidly out-competed the HES3-GFP cell line within ten passages making BCLXL the likely candidate driving selection of the CNV.

The cell cycle distribution was analysed to see if the CNV cells contained more cells in S-phase as you would expect from a cell line that was dividing quicker. The opposite was observed in the CNV cell lines, CNV cells appeared to contain more cells in G0/1 phase than S-phase. G2/M was similar in both the control and CNV lines. There was some variation between control cell lines with HES3 and ESI-035 having slightly elevated G0/1 when compared with the other two control cell lines, this may be an effect of having a proportion of cells containing the CNV in the HES3-control. Schmitt and co-workers[100] discovered that BCL-XL has a role in regulating the G2/M checkpoint of the cell cycle, which is distinct from its role in apoptosis. CDK1 activation requires association with cyclin-B1 and the phosphorylation of the threonine-161 residue, once active; CDK1 mediates the progression from G2 to Metaphase transition. The study found that BCL-XL co-localises and binds with CDK1 following DNA damage. BCL-XL binds through a flexible loop domain located between the 41st and 60th amino acids. Deletion of this region does not impede the anti-apoptotic function of BCL-XL but hinders its ability to bind CDK1.

Becker and colleagues showed that the cell cycle times of human ES cell are relatively short, [101]. When measured using time-lapse imaging the cell division times could be determined accurately. Cells were seeded as single cells, which made it easier to track individual cells. The cells were imaged every 10 minutes and cell division times were determined by measuring the time between cells undergoing cytokinesis. The data shows that cell cycle times can vary quite dramatically from 7-44 hours, this was more apparent in the Shef5 and ESI-035 cell lines. The Shef5-control and –CNV cell lines had identical cell division times of 23 hours which is slightly longer than the average of all control cell lines which at 19.5 hours. The only cell line that displayed faster cell division times in cell containing the CNV than its corresponding control

line was the H1 cell line. The average division times were significantly lower in the late

passage (15.5 hours) than the control cell line a characteristic which can be explained by lower

variation in cell division times. The H1 cell lines also showed tighter clustering of cell division

times with the majority of divisions occurring between 12-24 hours. The ESI-035-control and -

CNV cell lines also displayed similar average cell division times which was mimicked by ESI-035

over expression cell lines, ESI-035-control cells over-expressing HM13 and BCL-XL showed

slightly slower division times which suggests the cells are not dividing faster and therefore

population doubling cannot be explained by CNV cells dividing faster.

The four cell lines obtained in this study displayed differences both between early and late

passage of the same cell line and also between different cell lines. Differences were observed

in the gene expression levels of the pluripotency gene POU5F1 as well as genes contained

within the 20q11.21 amplification, which could explain differences observed in the growth

rates of the different cell lines. The expression of RNA was highly variable both between cell

lines and between genes in the same cell line. Different cell lines expressed the three

candidate genes at differing levels, the HES3, Shef5 and ESI-035 cell lines showed high ID1

expression (dCT=1-2) and similar levels of HM13 and BCL-XL (dCT=4). The inverse was observed

in the H1 cell line with higher HM13 and BCL-XL expression (dCT=2) and lower ID1 (dCT=4).

This variability is possibly a result of taking RNA at different states of growth and with varying

levels of differentiation within the culture. That being said the levels appear to be relatively

stable with very small standard deviation between biological replicates. With the exception of

the Shef5, the CNV cell lines appear to have increased POU5F1 expression when compared

with their control counterparts, this could indicate that the CNV containing cells are more

likely to remain undifferentiated and less likely to spontaneously differentiate. Other 'culture-

adapted' cell lines also exhibit this trait [63] reiterating the notion that adaptation favours

genetic changes that positively affect a stem cells ability to retain a stem cell phenotype.

Interestingly, in an undifferentiated state the levels of BCL-XL mRNA are not significantly higher in CNV-containing cell lines when compared to their control counter-parts. Increased expression of mRNA was observed in only two out of four CNV cell lines (HES3 and H1). However, the protein levels of BCL-XL in CNV-containing cells are consistently higher in all CNV-containing cell lines. This difference in BCL-XL mRNA and protein levels was also noted in colorectal cancers. Colorectal cancers often present amplification of chromosomal region 20q [102] and increased BCL-XL levels contribute to reduced apoptosis, development of metastasis and poor response to chemotherapy [103]. Sillars-Hardebol and co-workers showed that the levels of BCL-XL mRNA were not significantly higher in colorectal cancers containing a 20q gain compared with colorectal cancers that did not have the amplification. In contrast, the BCL-XL protein levels were elevated in all colorectal cancers that contained the 20q amplification compared with those that did not [96]. This relationship between mRNA and protein levels in CNV-containing human ES cell lines and colorectal cancer suggests a common neoplastic mechanism. The increased protein levels of BCL-XL in CNV-containing cell lines suggests that BCL-XL protein may be regulated post-transcriptionally. The post-transcriptional regulation of BCL-XL is poorly understood. That being said a number of miRNAs have been predicted to regulate BCL-XL expression including miR-663, miR-296, miR-1289-1 and miR-298.

4. BCL-XL Drives 20q11.21 CNV Selection Through Increased Protection Against Stress Induced Apoptosis

4.1. Introduction

Of the three candidate genes located in the minimal amplicon of 20q11.21, BCL-XL is most likely the one responsible for the selective advantage of CNV-containing cells. Chapter 3.2.1 shows that CNV containing cells have a marked increase in growth rates which can be mimicked by over-expressing BCL-XL in control cell lines. The difference in growth rates observed is not a result of faster cell cycling times or increased cells undergoing mitosis. The four cell lines show very similar division times both between different human ES cell lines and between control and CNV cells within those cell lines. The cell cycle distribution however showed slight differences between control and CNV cell lines. CNV-containing cells show decreased populations of cells in S-phase, which is striking, as one would expect that cultures growing faster would contain more cells in S-phase. However, these subtle differences do not explain the marked difference in growth rate between control and CNV-containing cell lines. The selective advantage is therefore working on a different aspect of stem cell behaviour. The obvious place to start with BCL-XL being an anti-apoptotic gene is the survival of CNV-containing cells.

Intrinsic (or mitochondrial) apoptosis is initiated as a response to intracellular stimuli including but not limited to DNA damage, oxidative stress and increased cytoplasmic Ca^{2+} [104]. In normal circumstances lethal signals are counter-balanced by pro-survival proteins which help to sequester and translocate death signals away from the mitochondria. When the balance is disrupted and pro-apoptotic signals accumulate, the mitochondrial outer membrane becomes permeabilized resulting in dissipation of mitochondrial functions and release of cytotoxic

proteins into the cytoplasm [105]. This process is largely regulated by the B-cell lymphoma 2 (BCL-2) family of proteins, which contain members with both pro-apoptotic and pro-survival functions. The BCL2 family members interact to maintain the balance between survival and death, and can be separated into subfamilies based on the structural differences in their BCL-2 homology (BH) domains [106]. The anti-apoptotic subfamily comprises of BCL-2, the long isoform of BCL-2-like 1 (BCL-XL), BCL-2 related gene A1 (A1), myeloid cell leukemia 1 (MCL-1) and BCL-W, each containing four BH domains (BH1-4). The anti-apoptotic BCL-2 members localize mainly to the outer mitochondrial membrane (OMM) but are also present on the membrane of endoplasmic reticulum and in the cytosol. The pro-apoptotic subfamily can further be divided into two groups, the effector proteins, BCL-2-associated X protein (BAX) and BCL-2 antagonist or killer protein (BAK) and BH3-only proteins. The effector protein members contain three distinct BH domains (BH1-3) lacking the BH4 domain and are essential for inducing permeabilization of the OMM [107]. The effector proteins form homo-oligomeric pores, which allow the release of cytochrome c and other apoptogenic molecules. Anti-apoptotic members act by sequestering BAX/BAK and translocating them away from the OMM [108]. The third subfamily, the BH3-only proteins such as BCL2-like protein 11 (BIM, also known as BOD), BCL2 antagonist of cell death (BAD) and BCL2 modifying factor (BMF) are pro-apoptotic proteins that inhibit the anti-apoptotic functions of BCL2 and other apoptotic members through direct binding. Once BH3-only members bind anti-apoptotic members, the effector proteins are de-repressed and able to induce apoptosis through OMM permeabilization [106].

The BCL2L1 gene encodes two splice variants, BCL-XS and BCL-XL which exhibit pro-apoptotic and anti-apoptotic functions respectively[109]. In undifferentiated human ES cells BCL-XL is the dominant isoform (Figure 2). The main role of BCL-XL is to promote cell survival via the translocation of BAX away from mitochondrial membrane and into the cytosol[108]. BCL-XL

also promotes cell survival through direct inhibition of p53[110]. This regulatory network between anti-apoptotic BCL2 family members and their targets creates a delicate balance, which is poised to undergo rapid apoptosis should pro-apoptotic signals increase through cellular stresses. It is therefore not surprising that increased levels of BCL-XL in human ES cells following the amplification of 20q11.21 would increase the survival of CNV-containing cells.

Using time-lapse analysis to accurately measure the time between cell divisions raised two intriguing points, firstly that there were more divisions occurring in the CNV-containing cells and secondly, that many of the control cells were dying before dividing or their progeny were dying before their second cell division resulting in fewer events in control cells compared to CNV-containing cells.

4.2. Results

4.2.1. Increased Protection to Stress Induced Apoptosis of CNV Containing Cells

The difference in cell cycle times determined by population doubling and time-lapse can be explained by increased cell death in control cell lines during culture. Individual cell fates were mapped using time-lapse imaging following single cell plating (Supplementary Movies 1-8). Cells were scored on one of five outcomes: 1. Following plating the cell dies before division, 2. The cell divides and both daughter cells die within the first 24 hours of plating, 3. The cell neither dies nor divides, 4. The cell divides and one of the daughter cell survives the first 24 hours, 5. The cell divides and both daughter cells divide. Five cells were picked at random from 20 fields to generate a true representation of each cell line (100 events). Using this scoring system, there were marked differences both between different cell lines and between control and CNV within a given cell line (Figure 12). With the exception of the HES3-CNV cell line the highest proportion of cells died within the first 24 hours without dividing. The HES3 cell line showed the greatest difference between control and CNV cell lines. More than half (58%) of control cells died within the first 24 hours without making a division, of the 32% cells that survived to make a division, only 18% of divisions contained at least one surviving progeny. This highlights the stress that single human ES cells are subjected to during re-plating. In stark contrast, only 26% of cells died without dividing, less than half seen in the HES3-control cell line. Out of the 70% of cells that survived to make a cell division, 64% of events had at least one surviving daughter cell, twice as many as its control cell line. The other cell lines followed a similar trend, on average, 70% of all control cells died without making a cell division compared with 41% of CNV cells. An average of 24% of control cells survived to make a cell division with only 11% of events producing at least one surviving daughter cell. When compared to CNV cells, in which 50% of cells divided and 37% of events showed at least

one surviving progeny. These results demonstrate the high survival rate of cell lines containing

the CNV. With over three times as many successful cell divisions with surviving daughter cells

it is easy to see how the CNV-containing cells out-compete diploid cells so rapidly.

To explore this finding further, the levels of apoptosis were monitored following single cell

passage of human ES cells (Figure 13). The four different cell lines were dissociated to single

cells and seeded at 1×10^5 cells/cm^2, this density was chosen as it represented a routine

passaging density of a 1:3 split ratio. The levels of cleaved caspase3 (active form and marker

of early apoptosis) were measured using flow-cytometry over the first 24 hours following

seeding. At the zero hour time point there was little difference between the control and CNV

lines that displayed on average 5% and 3% apoptosis respectively. Increased apoptosis was

detected in control cells as early as 3 hours post-seeding. At this point 13% of cells were

undergoing apoptosis whereas only 7% of CNV cells were positive for cleaved caspase3. This

trend continued with increasing time, at 12 hours post-seeding, 52% of control cells were

positive for cleaved caspase3. In contrast, only 25% of CNV cells were apoptotic. This again

highlights the increased survival of CNV cells following cellular stress. The final time-point of

24 hours saw 62% of control cells undergoing apoptosis, more than half of the initial seeding

density. Only 33% of CNV cells were positive for cleaved caspase3, a third of the culture.

There were differences observed between the different cell lines, of the four CNV cell lines,

HES3 showed the least apoptosis after 24 hours followed by H1, Shef5 and ESI-035 showed

similar levels at 24 hours. Of the four control cell lines, HES3 showed the least apoptosis after

24 hours, followed by H1 and Shef5 cell lines, which displayed similar levels of caspase3. The

ESI-035-control cell line contained the most apoptotic cells, this was also observed in the ESI-

035-HM13 and ESI-035-ID1, which contained very similar levels of apoptotic cells after 24

hours. The ESI-035-BCL-XL cell line contained very little apoptotic cells, with only 4% of cells

positive at 24 hours.

Figure 12. CNV-containing cells show increased survival and proliferation during the first 24 hours of seeding
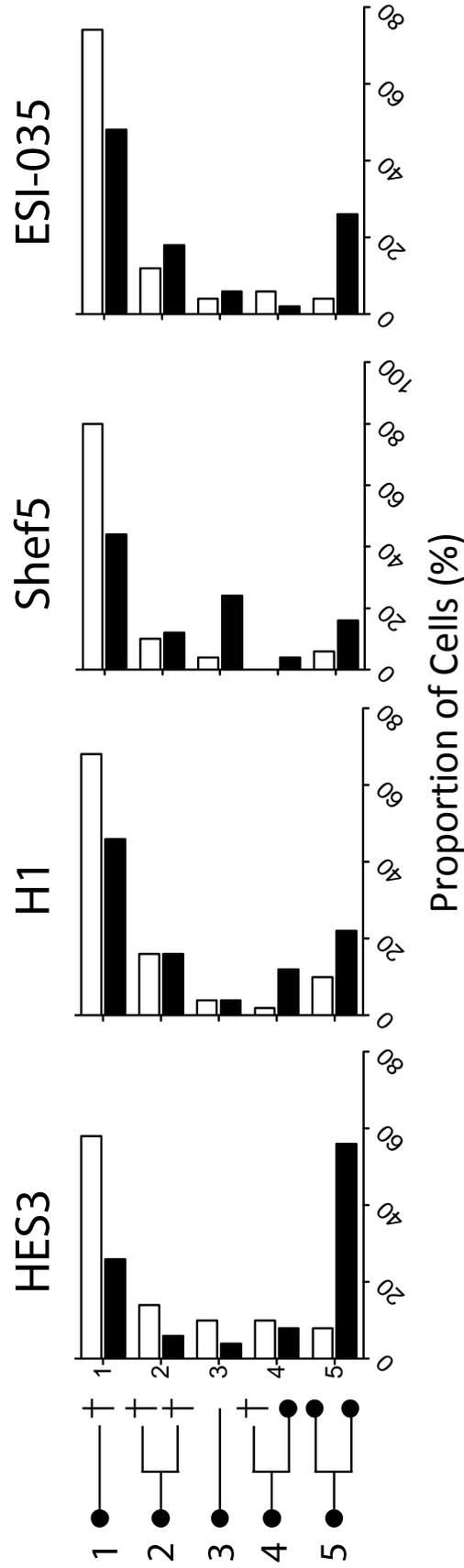
Cell fates were tracked during the first 24 hours following re-plating using live-cell time-lapse microscopy. 50 individual cells from 20 fields were tracked at random for each cell line. The cells were scored as follows: (i) the cell dies, (ii) the cell divides but both daughter cells die, (iii) the cell neither divides nor dies, (iv) the cell divides but one of the daughter cells dies, and (v) the cell divides and both daughter cells survive. CNV cells (black bars) showed reduced numbers of cells dying without making a cell division and increased numbers of cells dividing with both daughter cells surviving when compared with Control cell lines (white bars).

Figure 13. CNV-containing cell lines display increased protection against stress-induced apoptosis
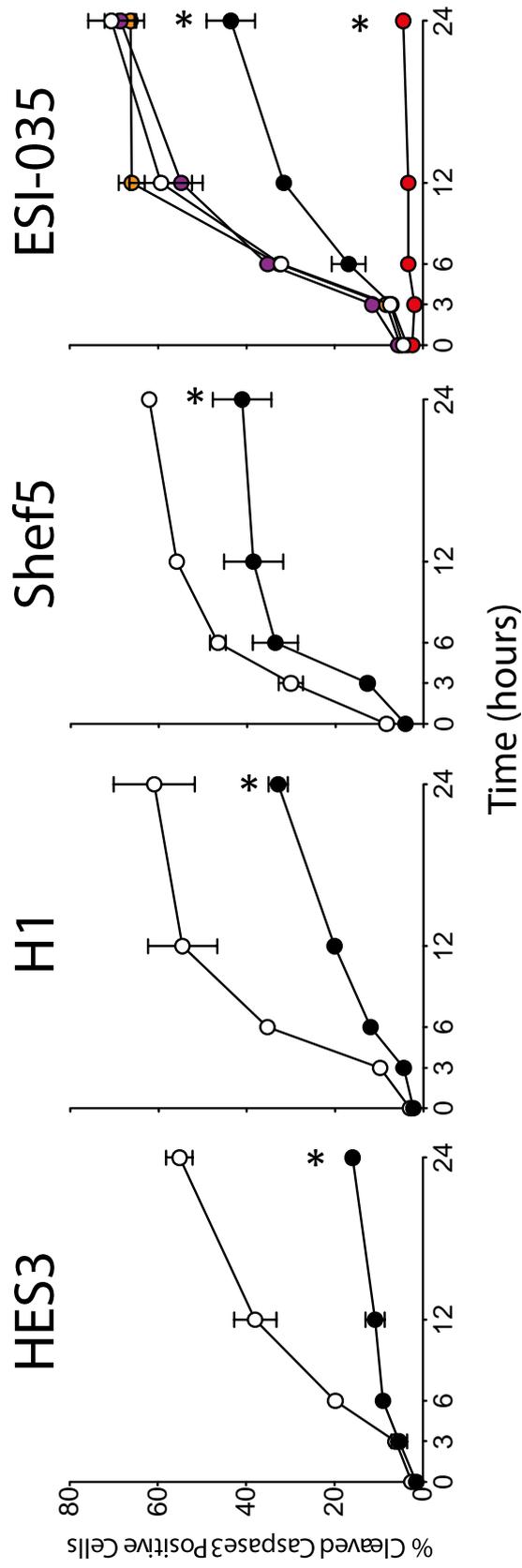
Cells were plated at 1×10^5 cells/cm^2 and the levels of apoptosis measured over the first 24 hours using flow-cytometry for cleaved caspase3. Control cells (white circles) show significantly increased levels of apoptosis just 6-12 hours following seeding when compared with CNV-containing cell lines (black circles). Error bars represent SEM of at least three biological replicates. Asterisks represent statistical significance by 2-way ANOVA: $p \leq 0.05$ (*).

4.2.2. Chemical Inhibition and Inducible Knock-Down of Bcl-xL

The increased growth capacity of CNV-containing cells can be attributed to the pro-survival properties of BCL-XL. The previous experiment showed that CNV-containing cells are more robust than their control counter-parts, particularly when seeded as single cells. Chemical inhibition and knock-down studies were conducted to further investigate the effect BCL-XL had on the survival of CNV containing cell lines. The small molecule ABT-263 is a Bad-like BH3 mimetic that inhibits the function of BCL-XL, BCL-2 and BCL-W. ABT-263 has a higher affinity for BCL-XL (<=0.5nM Ki) compared with BCL-2 and BCL-W (<=1nM Ki). Each cell line was plated at clonal densities ($1x10^4$ cells/cm²) on matrigel-coated plates in mTeSR. The mTeSR was supplemented with varying levels of ABT-263 (50nM, 100mM, 250mM and 1000mM). Control and CNV cell lines were also plated without the presence of ABT-263. The cells were cultured for 5-7 days in the presence of the inhibitor.

The difference in clonogenicity between control and CNV cell lines in the absence of ABT-263 was striking, the CNV containing cell lines were almost confluent whereas the wells containing control cells had very few colonies (Figure 14). These results are consistent with the population doubling rates observed in section 3.2.3, at the lowest seed density of $2x10^4$ cells/cm² the HES3-CNV cell line displayed the highest population doubling rates followed by the H1, ESI-035 and Shef5 CNV cell lines. The control cells also followed a similar trend to the results from section 3.1.5 with all cell lines displaying very low population growth rates. Upon treatment with ABT-263 the cloning efficiency of CNV containing cell lines could be lowered with increasing concentrations of the inhibitor. The CNV cell lines displayed similar cloning efficiencies to the control cell lines at approximately 250nM ABT-263.

To verify this result, an inducible BCL-XL knock-down cell line was generated using the Tet

Repressor (TetR) pSUPERIOR system (Figure 15A). The system utilises two plasmids, one

containing the TetR gene driven by the expression of the CAG promoter. The TetR protein

contains an N-terminus SV40 nuclear localisation signal (nls) promoting the import of TetR into

the nucleus. The second plasmid contains the short-hairpin RNA interference (shRNAi)

sequence that is driven by a modified H1 promoter. The H1 promoter contains a $TetO_2$ (TO)

sequence that is silenced by binding of TetR protein homodimers that bind with high affinity.

When the two plasmids are co-transfected into human ES cells grown in the absence of

doxycycline, the shRNAi expression is suppressed by the over-expression of TetR. When

doxycycline is added to the culture medium the molecules bind TetR protein causing a

conformational change that inhibits the binding of TetR to the TO element. This results in the

de-repression of the H1 promoter and subsequent expression of the shRNAi.

The H1-CNV cell line was used to create the inducible knock-down of BCL-XL. The

TetRnls.pCAG plasmid was transfected into the cells via nucleofection. Stable transfectants

were selected using puromycin and the clone with the highest TetR protein levels was

selected. The cell line was then transfected with the pSUPERIOR.BCL-XLshRNAi construct also

using nucleofection. Stable lines were generated using neomycin which is expressed

independently of the silenced H1 promoter. 100ng/mL of doxycycline was sufficient to

decrease the protein levels in H1-CNV cells to control levels (Figure 15B). The H1-CNV cell line

was cultured in mTeSR supplemented with 100ng/mL doxycycline for 96 hours to reduce the

levels of BCL-XL to control levels. The cells were then dissociated to single cells and plated at

$8x10^4$ cells/cm^2 in matrigel-coated 6-well plates. This density represented the highest density

used in population doubling experiments (Section 3.2.3). The cells were then cultured for 96

hours in the presence of doxycycline (100ng/mL) and final cell density was determined (Figure

15C). H1-CNV cells that had not been treated with doxycycline showed similar growth rates to

those observed in the population doubling experiment (Section 3.2.3). In the presence of doxycycline the population-doubling rate of H1-CNV cells was drastically reduced, comparable to that observed in H1 control cell line. This result confirmed BCL-XL as the main candidate driving the selection of CNV cells.

4.2.3. ROCK Inhibitor Alleviates The Selective Pressure of CNV 20q11.21

The ESI-035, ESI-035-CNV and ESI-035-BCL-XL cell lines were seeded at 2×10^4 cells/cm^2 in the presence or absence of ROCK inhibitor (Y-27632). The media was changed at 24 hours post-seeding and cultured for five days (Supplementary Movies 9-14). Total cell counts were measured and the number population doublings calculated (Figure 16). Control cells were unable to make a population doubling in the absence of ROCK inhibitor consistent with results in section 3.2.3. The CNV and BCL-XL over-expressing cell lines displayed increased growth rates with population doublings of 2.3 and 2.5 respectively. The addition of ROCK inhibitor for the first 24 hours of cell seeding increased the population growth rates of the Control cell line to that observed in CNV and BCL-XL (3.8, 4.4 and 4.3 respectively).

Figure 14. Chemical inhibition of BCL-XL in CNV-containing cells reduces cloning efficiency similar to control cell lines
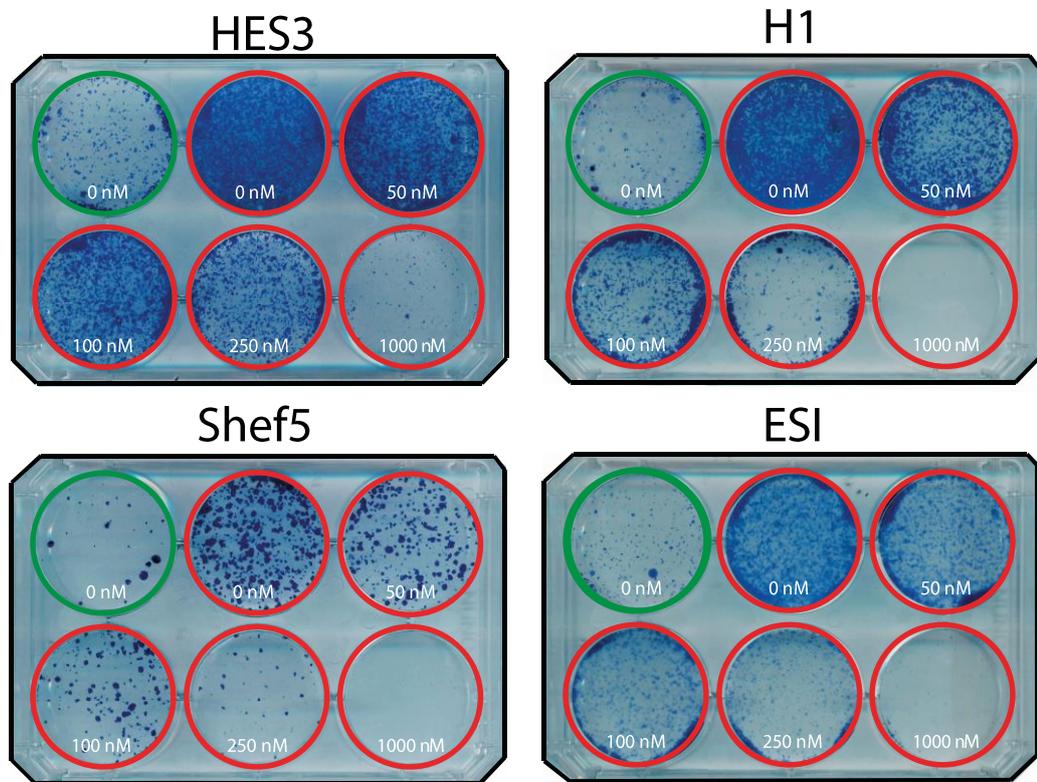
The cloning efficiency of CNV-containing cell lines could be reduced using increasing concentrations of ABT-263, a small molecule inhibitor of BCL-XL. Control (green rimmed well) and CNV (red rimmed wells) cell lines were seeded at $1x10^4$ cells/cm^2 and cultured for 5-7 days in the presence/absence of ABT-263. CNV cell lines continued to display increased cloning efficiencies compared with control cells at ABT-263 concentrations of 50nM and 100nM. Using 250nM of ABT-263 the CNV-containing cell lines exhibited cloning efficiencies similar to control cells.

Figure 15. Inducible knock-down of BCL-XL reduces growth rates of CNV-containing cells
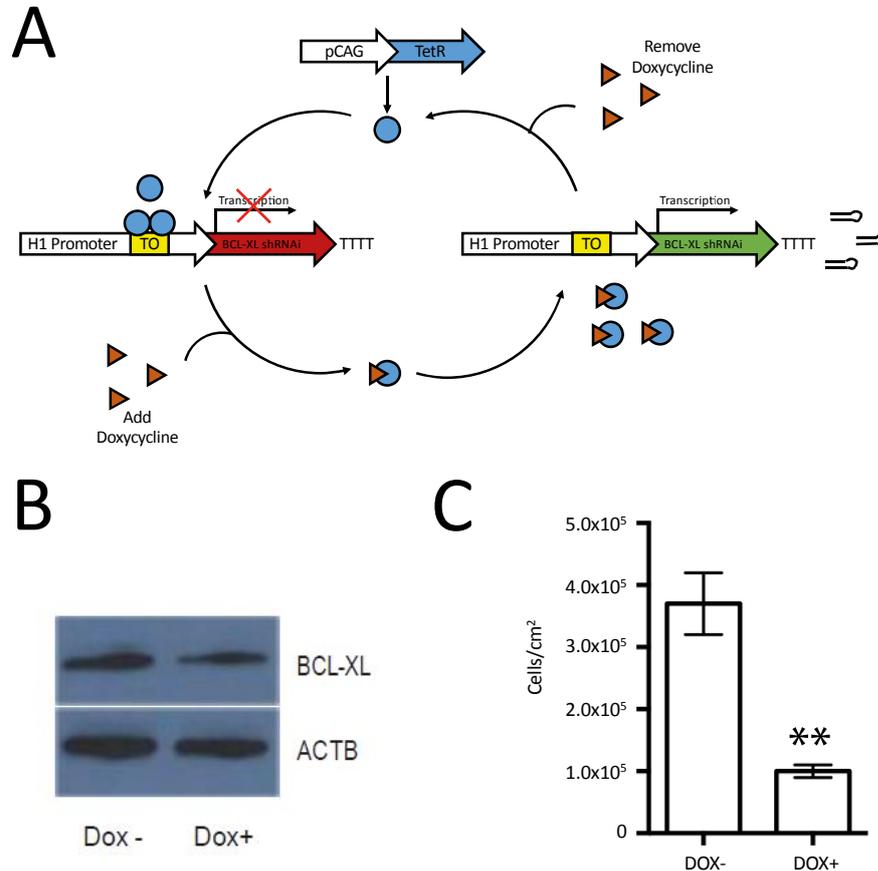
(A) Schematic showing the inducible shRNAi knock-down system used to reduce BCL-XL expression levels. Tet Repressor (TetR) – pSUPERIOR system, takes advantage of two vectors. The first drives the constitutive expression of TetR (blue circle), which binds to a TetO2 site on the promoter of a second vector, inhibiting the expression of BCL-XL-shRNAi. The addition of doxycycline (red triangles) sequesters TetR by direct binding therefore allowing the expression of BCL-XL-shRNAi. Knock-down can then be achieved by removing doxycycline {Zafarana, 2009 #107}. (B) Western blot analysis of BCL-XL using whole cell lysates prepared from the H1-CNV-BCL-XL-shRNAi cell line in the absence (Dox-) and presence (Dox+) of 100ng/mL doxycycline. (C) Cell densities (cells/cm^2) of the H1-CNV-BCL-XL-shRNAi following four days culture in the absence (DOX-) and presence (DOX+) after seeding at 8x10^4 cells/cm^2. Cells cultured in the presence of doxycycline displayed significantly reduced cell numbers compared to those not treated. Errors bars represent SEM from three biological replicates. Asterisks represent statistical significance by two-tailed t-tests: $p \leq 0.01$ (**).

Figure 16. ROCK inhibitor increases the growth rates of Control cells, alleviating the selective advantage of CNV 20q11.21
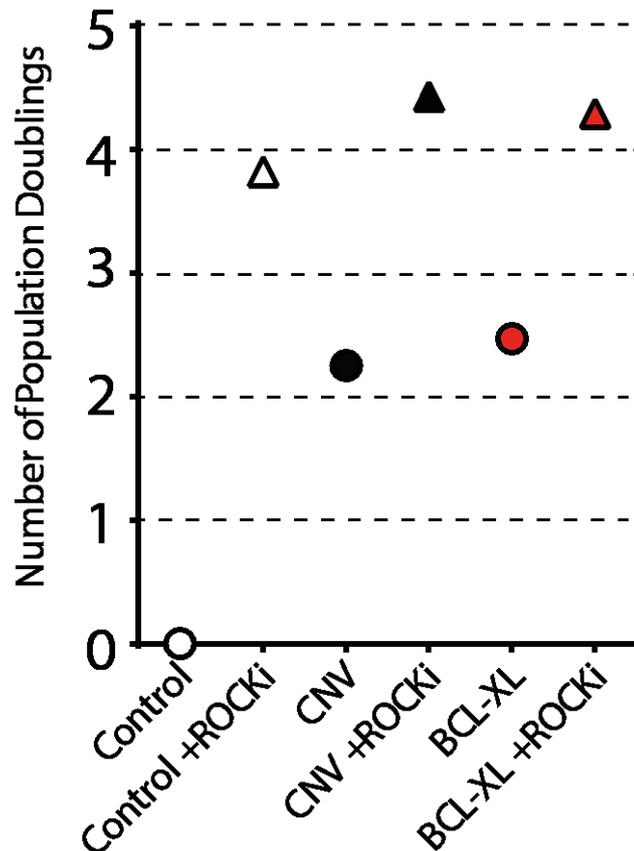
Data shows the number of population doublings following five days culture after seeding at $2x10^4$ cells/cm². In the absence of ROCK inhibitor Control cells (white circle) show decreased population doubling rates when compared with CNV (black circle) and BCL-XL over-expressing cells (red circle). This phenotype can be rescued by seeding cells in the presence of ROCK inhibitor for the first 24 hours. Control cells (white triangle) display similar population doubling rates to those observed in CNV (black triangle) and BCL-XL cell lines (red triangle).

4.3. Discussion

The high prevalence of increased 20q11.21 DNA copy number in human ES cell lines can be explained by the huge selective advantage that the CNV provides. This selective advantage can be attributed to BCL-XL, which increases the cells resistance to apoptosis following cellular stress. Increased BCL-XL activity through increased copy number or positive regulation has been implicated in a number of human cancers. Being a potent inhibitor of apoptosis, BCL-XL increases the survival of malignant cells and promotes metastasis [111-115]. Nguyen and co-workers also demonstrated the selective effect that increased BCL-XL expression has on human ES cell cultures. The group showed that BCL-XL is important for survival of human ES cells following loss of cell-cell contact [116]. Upon dissociation to single cells the RhoA/ROCK (Rho-associated protein kinase) pathway is disturbed resulting in phosphorylation of RhoA and ultimately cell death via apoptosis [117]. The anti-apoptotic properties of BCL-XL inhibit apoptosis by sequestering pro-apoptotic signals in the cytosol and translocating them away from the mitochondrial membrane. Hence increased BCL-XL levels in CNV-containing cells will decrease the sensitivity of cells to RhoA/ROCK mediated apoptosis.

Following single cell dissociation and re-plating, CNV-containing cells exhibit increased survival, which helps the culture to re-establish faster than control cell lines. When plated as single cells, the majority of control cells die within the first 24 hours without making a cell division. Of the small percentage of cells that survive to make a cell division, the majority of cases result in the death of both daughter cells. There were a very small percentage of cells that survived the first 24 hours of plating and divided with at least one daughter cell surviving. In stark contrast, the CNV cell lines showed reduced numbers of cells dying within the first 24 hours without making a division when compared with control cells. Of the cells that survived to make a cell division, there was a higher proportion of cases in which at least one daughter cell

survived. This data is consistent with Barbaric and colleagues[74], the group found that upon

single cell plating of human ES cells there existed bottle-necks in the culture of normal, diploid

cells that were overcome in karyotypically abnormal, culture-adapted cells. Firstly, upon

plating culture-adapted cells were more likely to survive than normal cells. Secondly, of the

cells that survived plating, a lower proportion of normal cells re-entered the cell cycle

compared with culture-adapted cells. Both of these bottlenecks highlight the shortfalls in

human ES cell culture conditions and can explain the speed at which abnormal cells take over

human ES cell cultures. These two bottlenecks can be alleviated via the addition of ROCK

inhibitor (γ-27632), which reduces apoptosis following single cell dissociation and increases

cloning efficiencies of human ES cells [118]. Barbaric and co-workers found that addition of

ROCK inhibitor increased the survival of normal cells but had no effect on the survival of

culture-adapted cells following single cell plating[74]. Results from section 4.2.3 show similar

findings although CNV and BCL-XL over-expressing cells also show increased growth rates in

the presence of ROCK inhibitor suggesting that other karyotypic changes (notably 12p and 17q)

may provide a selective advantage via different mechanisms.

These findings show that BCL-XL increases the survival of cells containing the 20q11.21

amplification through increased protection against caspase-mediated apoptosis. It is therefore

important to monitor human ES cells for amplification of 20q11.21 and specifically BCL2L1,

particularly if they are to be used in toxicology screens.

5. The Effect Of 20q11.21 Amplification On Differentiation

5.1. Introduction

Efficient differentiation of human pluripotent stem cells into all three primary germ layers is essential for their eventual use in cell replacement therapies. Understanding the early stages of lineage specification is fundamental to developing directed differentiation protocols into terminally differentiated, functional cells. Human ES cell populations are highly heterogeneous, even for markers of the undifferentiated phenotype. Human ES cells can express different levels of SSEA3 and POU5F1 in culture, which may suggest they are in different states of differentiation or just fluctuate between high and low expression[75, 119, 120]. The dynamics of this phenomenon are not fully understood but may be integral to understanding differentiation in human ES cells. This heterogeneity may also be a result of non-optimal culture conditions, culture methods that create homogeneous populations of human ES cells may be better for routine culture and to increase the efficiencies of current directed differentiation methodologies. There has been much debate as to the mechanisms behind directed differentiation. It seems possible that once induced to differentiate towards a specific lineage (i.e directed differentiation) all human ES cells will behave similarly and progress down a particular path regardless of their initial expression patterns. It has been suggested that inducing differentiation may not be directed at all but rather results from selection of those cells that are able to make a specific lineage. This may be the case in terms of endoderm differentiation where there is increased cell death upon induction[121, 122].

To investigate the mechanisms behind differentiation, the CNV and BCL-XL over-expressing cell lines offer interesting tools to address the question of whether directed differentiation is truly directing cell towards a particular lineage or merely selecting those cells that are able to

progress towards a particular lineage. The hypotheses are either: 1) addition of an inducer

causes cells to differentiate towards a specific lineage (directed differentiation) or 2) the cells

spontaneously differentiate along many different lineages with the addition of an inducer but

rather kills cells that have not proceeded towards a particular lineage (selected differentiation)

(Figure 17). If the first hypothesis were true, one would expect the anti-apoptotic properties

of BCL-XL would have no influence on the outcome of differentiation. If the second hypothesis

were true, then blocking apoptosis would affect the pattern of differentiation, either allowing

the differentiation of other lineages or retention of undifferentiated cells. These two

hypotheses may not be mutually exclusive and could potentially converge. The differentiation

of cells containing the 20q11.21 amplification has not yet been investigated; there may be

other genes present in the amplicon that affect the differentiation of human ES cells.

Figure 17 Model showing the possible mechanisms behind directed differentation

A <u>Directed Differentat i on of Human ES Cells</u>

| Heterogeneous stem cell population | | Despite heterogeneity all cells are capable of differentat i on | | Cell death occurs spontaneously unrelated to differentat i on condti ons |

Early stages of lineage specific dif eretiat i on

Late stages of lineage specific dif eretiat i on
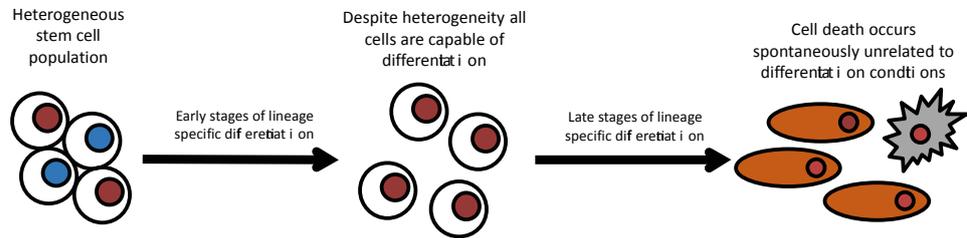
B <u>Selection of Lineage Primed Cells in Directed Dif eretiat i on</u>

| Heterogeneous stem cell population | | Cell death occurs in subsets of cells due to differentat i on condti ons | | Only cells able to make specific lineages survive differentat i on |

Early stages of lineage specific dif eretiat i on

Late stages of lineage specific dif eretiat i on
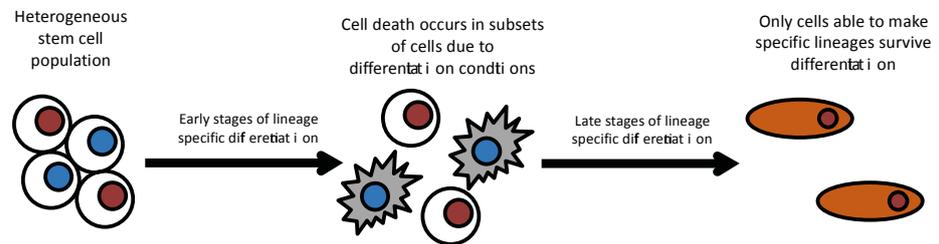
C <u>Selected Differentat i on: Ef f ect of Increased Survival</u>

| Heterogeneous stem cell population (Culture-adapted) | | Selection unable to take place due to increased survival | | Other cell lineages are able to different ate despite culture conditions |

Early stages of lineage specific dif eretiat i on

Late stages of lineage specific dif eretiat i on
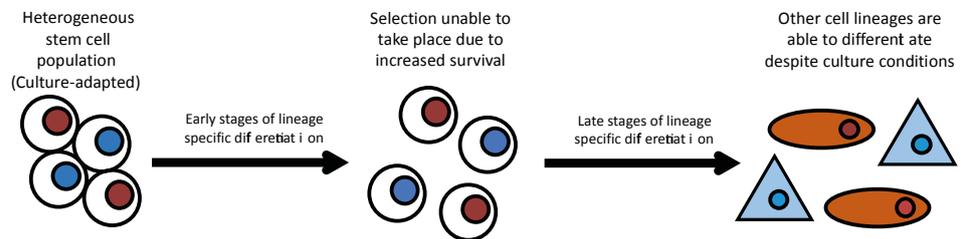
It has been widely debated whether directed differentat ion is truly di reting all cells towards a particular lineage or whether specific culture conditions select those cells that are able to different ate towards a particular lineage. (A) Directed differentat ion: the h eterogeneity of human ES cell cultures is irrelevant, during the early stages of differentat ion ES cells are able to fluctu ate between dif erent sub-states and progress to make specific lineages. Cell death is spontaneous and unrelated to culture-conditions. (B) Selecting cells capable of specific lineages: the heterogeneity of human ES cell cultures represents cells capable of different lineages, in the early stages of differentat ion cultu re conditions select for those cells capable of a specific lineage. Increased cell death is a result of selection, cells unable to make a specific lineage are removed from the culture.
(C) Increased survival in culture-adapted cells: if differentat ion is a result of selection, the increased survival of culture-adapted cells could permit the differentat ion of other linea ges despite culture conditions.

5.2.1. Spontaneous and Directed Differentiation via Formation of Embryoid Bodies

Three cell lines, ESI-035 control, ESI-035-CNV and ESI-035-BCL-XL were differentiated via the formation of embryoid bodies. The cell lines were subjected to four conditions. Neutral conditions used only APEL media with no other growth factors or inhibitors. This condition permits spontaneous differentiation into all three germ layers. The three remaining conditions contained additional growth factors and inhibitors to direct differentiation into the three primary germ layers. Ectoderm media contained of APEL supplemented with bFGF (100ng/mL), DMH-1 (1uM) a BMP inhibitor that promotes neurogenesis in human ES cells and SB431542 (10uM) a small molecule that blocks differentiation towards mesendoderm lineages by disrupting TGF-β/activin/nodal signal transduction [123] [124]. Mesoderm differentiation conditions contain APEL supplemented with Activin A (20ng/mL) and BMP4 (20ng/mL). Differentiation towards endoderm is achieved by supplementing APEL with Activin A (100ng/mL) and BMP4 (1ng/mL). Mesoderm and endoderm conditions contain identical components however at differing concentrations, higher Activin A and low BMP4 concentrations are required for endoderm specification whereas the reverse is needed for mesoderm differentiation. EBs were cultured for ten days, images were taken from four wells selected at random and RNA was extracted by pooling all EBs from the same condition.

The three cell lines formed EBs for all four differentiation conditions following aggregation in 96-well U-bottom wells (Figure 18, Table 6). Under neutral conditions the control cell line formed cystic EBs that were on average 899±87nm in size with evidence of cell death surrounding the EBs. The CNV cell line generated very uniform, non-cystic EBs that were morphologically different to the control EBs, the EBs were dense and smaller than control EBs

(648±20nm) with minimal cell death. The BCL-XL cells formed highly irregular, cystic EBs under neutral conditions, although morphologically similar to control than CNV EBs, the EBs were much larger than both control and CNV (1308±142nm). Under ectoderm differentiation conditions both control and CNV formed similar EBs, EBs were tightly packed, spherical in shape, similar in size (660±23nm and 694±20nm respectively) and minimal cell death was observed. The BCL-XL over-expressing cell line formed similar EBs, but the cells were not as tightly packed and were slightly irregular when compared to the almost perfect spherical shape observed in control and CNV cell lines. The EBs were also slightly larger than control and CNV cells with an average size of 762±109nm. The differentiation of the control cells towards a mesodermal lineage resulted in increased cell death when compared with neutral and ectoderm conditions, EBs appeared small (466±38nm) and cystic. CNV cells produced slightly larger mesoderm EBs (785±20nm) when compared with control cells although cell death was also observed. EBs also appeared slightly cystic but cells appeared more tightly packed than control cells. BCL-XL over-expressing cells produced large (1115±104nm) cystic EBs in mesoderm conditions with little cell death. EBs were very irregular in shape with regions of cystic and tightly packed cells. The differentiation towards endoderm in control cells also yielded very small, spherical EBs (398±59nm) with evidence of increased cell death. The CNV cells produced very uniform EBs under endoderm conditions resembling those seen under neutral and ectoderm conditions, EBs appeared spherical with small nodules. EBs were larger than those seen in the control cell line with an average of 671±20nm with minimal cell death observed. BCL-XL again produced the largest EBs under endoderm conditions averaging 830±109nm. EBs were irregular in shape, containing cystic and dense regions resembling those observed under neutral and mesoderm conditions.

Figure 18. Differentiation of human embryonic stem cells via the formation of embryoid bodies
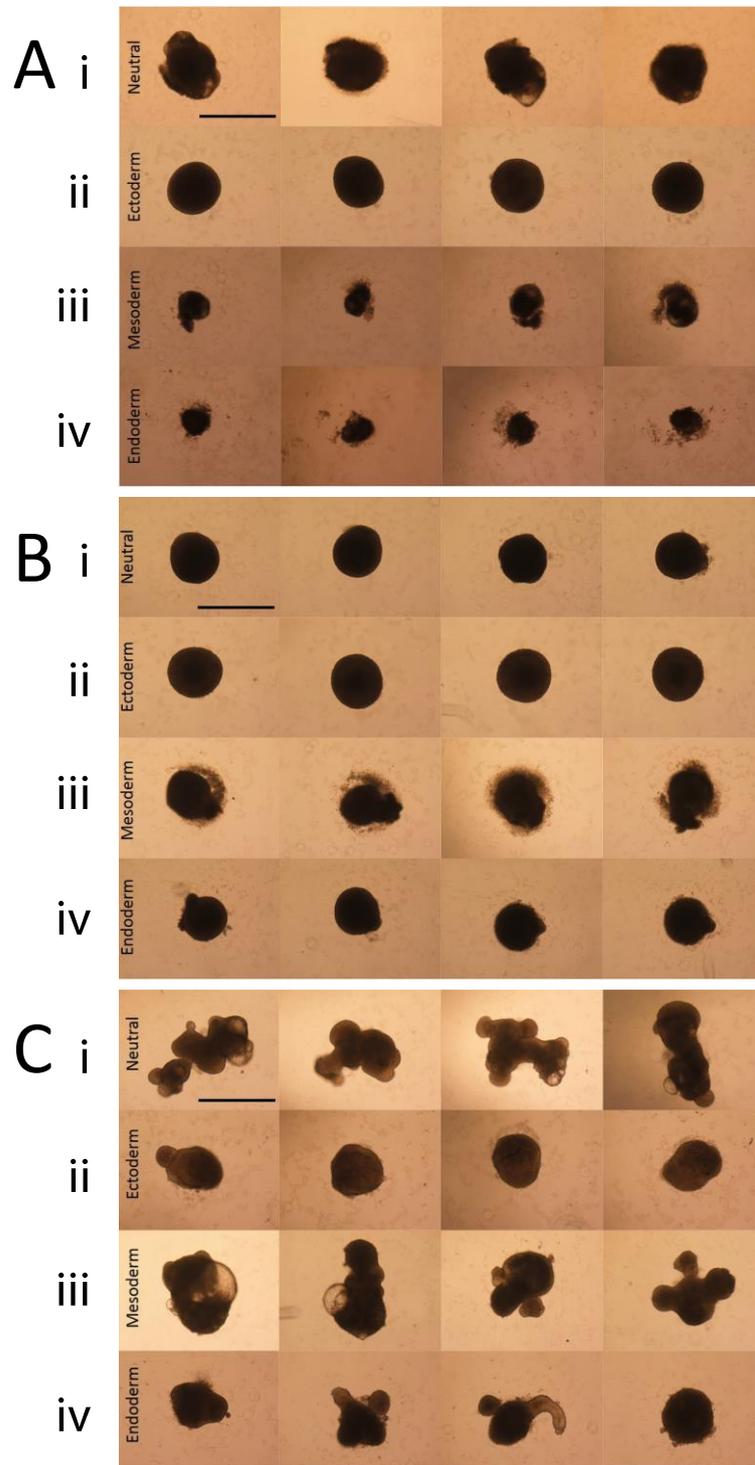
Four representative images of day 10 EBs formed from (A) ESI-035, (B) ESI-035-CNV and (C) ESI-035-BCL-XL. The EBs were subjected to four culture conditions: (i) neutral- allowing the spontaneous differentiation towards multiple lineages, (ii) ectoderm, (iii) mesoderm and (iv) endoderm. Scale bars = 1000μm.

Table 6. The sizes and descriptions of EBs formed from ESI-035 Control, CNV and BCL-XL over-expression cell lines

	Average Size (nm)	SD (nm)	Comments
ESI-035			
Neutral	899	87	Irregular, cystic and dense patches
Ectoderm	660	23	Regular spherical shape, dense
Mesoderm	466	38	Rounded, cystic, mass cell death
Endoderm	398	59	Rounded, dense, mass cell death
ESI-035-CNV			
Neutral	648	20	Regular spherical shape, dense
Ectoderm	694	20	Regular spherical shape, dense
Mesoderm	785	20	Irregular shape, dense, cell death
Endoderm	671	20	Spherical with nodules, dense
ESI-035-BCL-XL			
Neutral	1308	142	Irregular, cystic and dense patches
Ectoderm	762	109	Irregular, cystic and dense patches
Mesoderm	1115	104	Irregular, cystic and dense patches
Endoderm	830	109	Irregular, cystic and dense patches

The table shows the average sizes of EBs formed from ESI-035 Control, CNV and BCL-XL over-expressing cell lines and a description of their morphology.

This data suggests that the presence of the 20q11.21 CNV or the over-expression of BCL-XL does influence the pattern of differentiation of human ES cells. The greater size of the EBs derived from the CNV and BCL-XL over-expressing cells suggests that these differences in part might be due to the anti-apoptotic effects of BCL-XL.

Total RNA was harvested from EBs following 10 days of differentiation, the RNA was analysed using qPCR to further assess the differentiation of the three cell lines (Figure 19). The ddCT values were colour-coded based on expression level, green representing the highest expression and red representing low expression. In the undifferentiated condition as one would expect OCT4 is expressed at a high level, similar to that of GAPDH (reference gene). Under neutral conditions the control line expressed increased levels of mesoderm and endoderm markers but not ectoderm. In contrast the CNV showed increased expression of mesoderm and ectoderm markers with the BCL-XL over-expressing cell line showing slightly increased expression of all three germ layers. As the neutral condition contains only APEL medium and no added growth factors the cells are induced to differentiate spontaneously. This being the case one would expect all three lineages to be present. Under directed differentiation conditions towards ectoderm, the pluripotency gene OCT4 is down-regulated. This was observed in all cell lines and seemed most prominent in the BCL-XL and CNV cell lines indicating that all cell lines were differentiating. The endoderm markers SOX17 and GATA6 were also down-regulated under ectoderm conditions. A slight increase in the levels of mesoderm markers PECAM and CD34 was also observed in the control and CNV cell lines but not the BCL-XL over-expressing cell line. As one would expect the neural marker PAX6 was highly expressed in all cell lines following ectoderm differentiation. Surprisingly the expression levels of COL1A1, a marker of bone and skin remained unchanged or down-regulated in ectoderm conditions. Under mesoderm conditions, the control cell line showed increased expression of all three lineages, possibly highlighting the fact that there is cross-over of the

signalling pathways involved in lineage specification. This was not observed in CNV or BCL-XL cell lines, GATA6 showed increased expression but not SOX17, which remained at similar levels to undifferentiated cells. The levels of mesoderm markers PECAM and CD34 were slightly higher than undifferentiated cells but were not as high as control cells which expressed almost 100-fold and 200-fold more than undifferentiated cells respectively (CNV: 4-fold and 5-fold, BCL-XL: 9-fold and 10-fold respectively). The CNV cell line also showed increased expression of COL1A1 but not PAX6 under endoderm conditions, the reverse was observed in the BCL-XL over-expressing cell line. The control line under endoderm conditions did not down-regulate OCT4 but showed increased expression of SOX17 and GATA6, the mesoderm marker PECAM was also up-regulated along with the ectoderm markers PAX6 and COL1A1. This may highlight the need of supporting cells to efficiently differentiate towards certain lineages. In contrast the CNV and BCL-XL cell lines showed down-regulation of the endoderm markers and similar levels of mesoderm markers to undifferentiated cells. Up-regulation of PAX6 was also observed in CNV and BCL-XL cells under endoderm conditions. COL1A1 expression was slightly higher in CNV cells but not BCL-XL cells, which showed lower expression.

The larger size of EBs in CNV and BCL-XL over-expressing cells suggests that apoptosis plays a part in the ability of cells to differentiate into particular lineages, notably endoderm. For example, EBs formed from ESI-035 control cells under endoderm conditions showed increased cell death but increased expression of SOX17 and GATA6 compared with CNV and BCL-XL over-expressing cells. EBs formed from these cell lines displayed minimal cell death but did not express the endoderm markers SOX17 or GATA6. This could be a result of inner cells not receiving stimulus from the added growth factors and cell death is necessary for endoderm to develop.

Figure 19. mRNA gene expression in EBs formed from ESI-035 Control, CNV and BCL-XL over-expressing cell lines

i

Neutral

	POU5F1	PAX6	COL1A1	PECAM	CD34	SOX17	GATA6
Control	0.14	0.19	1.20	-1.47	-1.47	-3.00	-1.78
CNV	0.30	-3.24	-3.35	-2.14	-2.37	0.15	-0.33
BCL-XL	1.36	-4.07	2.92	-0.48	-0.64	-1.15	-0.52

ii

Ectoderm

	POU5F1	PAX6	COL1A1	PECAM	CD34	SOX17	GATA6
Control	4.23	-3.86	2.28	-1.30	-1.60	9.99	5.81
CNV	6.21	-8.24	-0.26	-3.01	-2.41	11.23	7.82
BCL-XL	8.78	-5.28	4.03	0.99	0.00	10.77	7.94

iii

Mesoderm

	POU5F1	PAX6	COL1A1	PECAM	CD34	SOX17	GATA6
Control	1.87	-2.61	-3.24	-6.56	-7.78	-3.41	-4.74
CNV	-0.48	0.05	-3.27	-2.24	-2.35	0.37	-1.25
BCL-XL	0.75	-3.91	0.43	-3.17	-3.40	-1.41	-2.01

iv

Endoderm

	POU5F1	PAX6	COL1A1	PECAM	CD34	SOX17	GATA6
Control	-0.29	-2.30	-0.89	-4.18	0.75	-4.64	-4.17
CNV	-0.71	-1.95	-1.42	-0.99	-0.30	1.54	1.58
BCL-XL	0.24	-3.45	2.76	0.99	1.95	1.63	1.99

Low Expression ◁_____▷ High Expression

ESI-035 Control, CNV and BCL-XL over-expressing cell lines display different lineage-specific gene expression patterns when induced to differentiate (i) spontaneously or directed towards (ii) ectoderm, (iii) mesoderm and (iv) endoderm. CT values of POU5F1 (pluripotency), PAX6 and COL1A1 (ectoderm), PECAM and CD34 (mesoderm) and SOX17 and GATA6 (endoderm) were normalised to the reference gene GAPDH. The values displayed are ddCT values, calculated by subtracting the dCT values of undifferentiated samples from differentiated samples. The values are colour coded based on the level of expression (red-green, low-high expression).

5.2.2. Directed Monolayer Differentiation Towards Definitive Endoderm

ESI-035 control, ESI-035-CNV and ESI-035-BCL-XL were also subjected to directed differentiation towards definitive endoderm. Cells were seeded at $3x10^4$ cells/well in 24-well plates coated with matrigel in mTeSR containing Y-27632 (1:1000) and left to attach for 24 hours. The mTeSR was aspirated and cells were washed once with sterile PBS (1X). The cells were then fed with endoderm inducing medium (see Section 2.24) for five days. The cells lines displayed similarities and differences in their differentiation in terms of morphological changes (Figure 20) (Supplementary Movies 15-17). Within the first 6 hours all cell lines maintained ES cell-like colony morphology, with increased evidence of cell death in the control cell line when compared with the ESI-035 CNV and ESI-BCL-XL cell lines. The cell lines continued to display compact colony morphology until 24 hours. After 24 hours the cells started to disperse into gaps between colonies with evidence of increased proliferation in the ESI-035 CNV and BCL-XL over-expressing cell lines.

The cells were stained at day five for the pluripotency marker POU5F1, the definitive endoderm marker SOX17 and counter-stained with Hoechst 33342 to stain DNA (Figure 21). Using the IN-Cell Analyzer 2000 the number of positive cells per field was generated (Table 7). Control cells displayed the lowest cell density following five days treatment with an average of 827±79 cells/field. Of the cells surviving after five days 15±8% were positive for POU5F1 and 56±10% stained positive for SOX17. In contrast, there were more surviving cells in the ESI-035 CNV cell line, with an average of 1061±46 cells/field. There were less POU5F1 (11±4%) and SOX17 (46±8%) when compared with the Control cell line. The BCL-XL over-expressing cells showed the highest final density with an average of 793±91 cells/field. The number of cells staining for POU5F1 and SOX17 were also higher than both Control and CNV cells with an average of 27±9% and 68±8% respectively. These results demonstrate that although there are

subtle differences between control and CNV cells with respect to endoderm differentiation, the CNV cells can make endoderm just as efficiently. Whether there are other lineages present in the culture would need further investigation. This result also disproves the hypothesis that directed differentiation is selecting cells that are more inclined to make a certain lineage and will undergo cell death in other conditions. This is highlighted by the BCL-XL over-expressing cells, which show increased cell numbers following five days differentiation but also contain the highest proportion of SOX17 positive cells.

Figure 20. Directed monolayer differentiation of human embryonic stem cells towards a definitive endoderm lineage
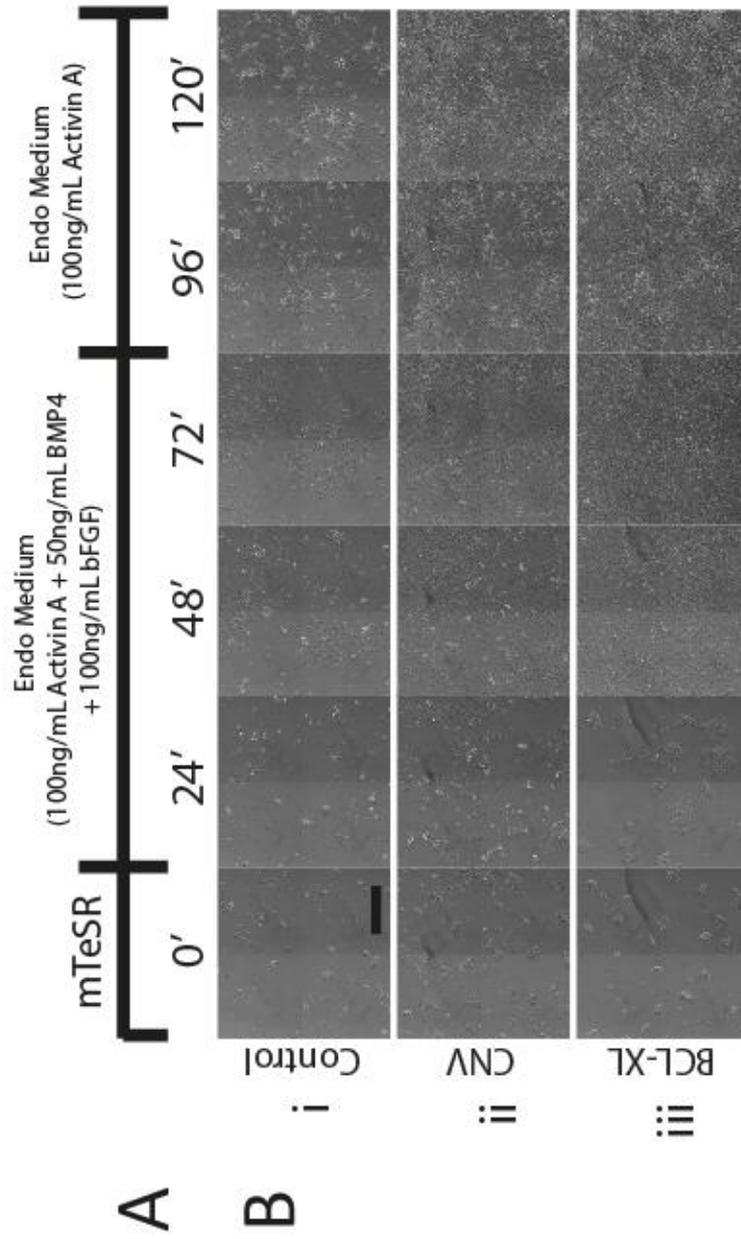
(A) schematic representation of the differentiation protocol, cells are plated as single cells in the presence of ROCK inhibitor and left to attach for 24 hours. The cells are then cultured in endoderm differentiation medium supplemented with Activin A, BMP4 and bFGF for three days. The cells are then fed with endoderm differentiation medium containing Activin A for a further two days. (B) The morphological changes of cells undergoing differentiation towards definitive endoderm. CNV (ii) and BCL-XL over-expressing cells (iii) display increased proliferation under endoderm conditions when compared to Control cells (i). Scale bar = 100µm.

Figure 21. A comparison of definitive endoderm differentiation in ESI-035 Control, CNV and BCL-XL over-expressing cell lines
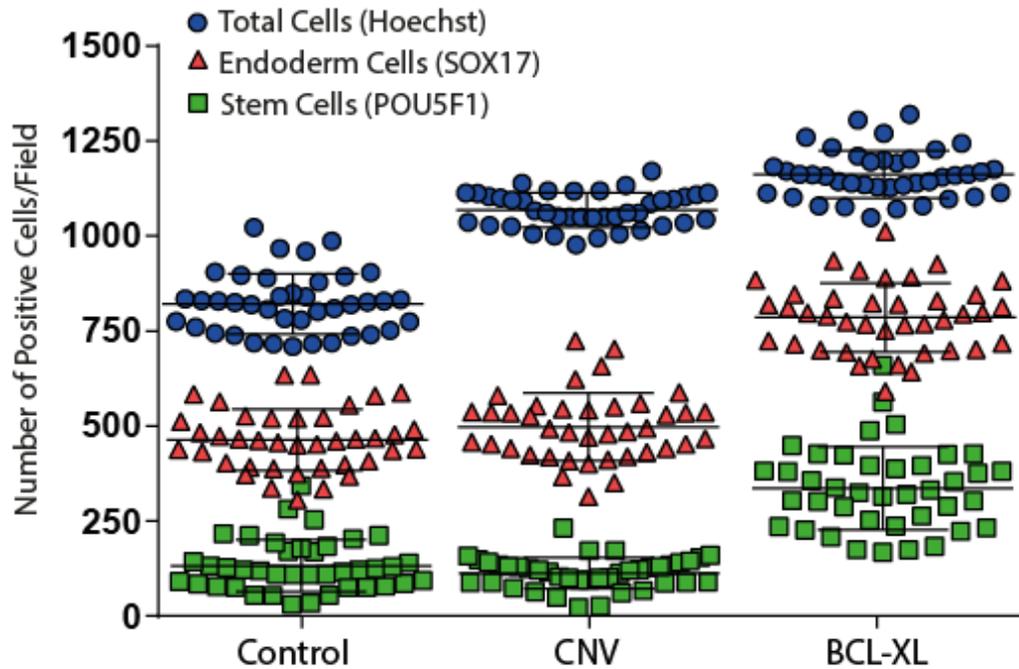
Immunostaining of day five differentiated cells under endoderm conditions, cells were stained for the pluripotency marker POU5F1 and definitive endoderm marker SOX17. DNA was counter-stained using Hoechst 33342. Each point on the graph represents the number of stained cells/field from 40 fields (of each cell line). Total number of cells (blue circles) was determined by the number of Hoechst positive cells, the number of endoderm cells was represented by cells staining positive for SOX17 (red triangles) and the number of stem cells was represented by POU5F1 positive cells (green squares). CNV and BCL-XL over-expressing cell lines display increased survival/proliferation when compared with Control cells. Control and CNV cells display similar numbers of SOX17 and POU5F1 positive cells. Theses numbers are increased in BCL-XL over-expressing cells suggesting that BCL-XL or the survival properties that it confers, do not affect the differentiation of human ES cells towards an endodermal lineage.

Table 7. Summary of definitive endoderm diferentation in ESI-035 Control, CNV and BCL-XL over-expressing cell lines

	ESI-035 Control	ESI-035 CNV	ESI-035 BCL-XL
Total Cell Number	827±79	1061±46	1158±63
Number of POU5F1+ Cells	125±68 (15±8%)	118±41 (11±4%)	323±109 (27±9%)
Number of SOX17+ Cells	466±81 (56±10%)	489±89 (46±8%)	793±91 (68±8%)
Other (%)	29%	43%	5%

The table shows the average number of cells per field (40 fields) that stained for Hoechst 33342 (DNA), POU5F1 (pluripotency marker) and SOX17 (endoderm marker). Other represent the percentage of cells that were negative for either POU5F1 or SOX17 indicatn th at they are not stem cells or definitive endoderm respectively.

5.3. Discussion

The ability to differentiate human ES cells into specific cell types is key to their eventual use in cell replacement therapies. It has not been widely investigated whether all human ES cell lines have the same capacity of differentiation and it has been shown that even in populations of cells there is heterogeneity even with respect to the pluripotency gene POU5F1[75, 119, 120]. It has also been shown that culture-adaptation leads to altered differentiation patterns. For example, Fazelli and co-workers[76] found that culture-adapted human ES cells exhibited different differentiation potential when compared with their normal counter-parts. This highlighted that cells with decreased differentiation potential could be driving selection of variant human ES cells. It is also possible that certain chromosomes that drive culture-adaptation of human ES cells through increased proliferation or survival may contain other 'hitch-hiker' genes. These genes may reduce a cells tendency to differentiate, increasing the selective advantage of the acquired chromosome or restrict the differentiation towards particular lineages. For example, cells differentiating towards endoderm have been shown to produce and release BMPs, potentially inducing differentiation of adjacent cells.

EBs have been used extensively to demonstrate the developmental potential of pluripotent cells. Using this method the ESI-035 control, CNV and BCL-XL cell lines displayed differences both when induced to differentiate spontaneously and towards individual germ layers. Under neutral conditions the control cell line formed irregular EBs that expressed mesendodermal but no ectodermal markers. The CNV produced very uniform spherical EBs, similar to those observed in the CNV under ectoderm and endoderm conditions and in the control line under ectoderm conditions. The EBs expressed high levels of ectodermal and low levels of mesodermal markers. In the BCL-XL cell line, neutral EBs looked highly irregular with both dense and cystic patches, similar to those observed in BCL-XL EBs under mesoderm and

endoderm conditions. The EBs expressed high levels of ectodermal and low levels of

mesendoderm markers. In neutral conditions the CNV and BCL-XL cells were more similar than

the control lines in expression patterns. Where the control cell line appeared to favour a

mesendodermal lineage, the CNV and BCL-XL cell lines showed high expression of PAX6, a

marker of neural differentiation.

Under ectoderm conditions, all cell lines behaved similarly, the morphology of EBs was almost

identical between control and CNV EBs. EBs formed were spherical and dense, BCL-XL EBs

were not as uniform as control or CNV lines and were not as tightly packed. The expression

patterns of all three lines were also similar with low levels of mesendoderm markers and very

high levels of PAX6 but not COL1A1.

Differentiation towards mesoderm was more variable between the cell lines, in the control,

the EBs were small in size and cystic with evidence of cell death. The EBs formed from control

cells expressed markers of all three lineages. The CNV cell line formed larger EBs than the

control (average: 785±20nm and 466±38nm respectively), cell death was also observed. The

CNV EBs expressed low levels of mesendoderm markers and increased COL1A1 expression,

PAX6 was unchanged in CNV cells. The EBs formed from BCL-XL over-expressing cells under

mesoderm conditions were larger than both the control and CNV, were very irregular in shape

and showed dense and cystic patches. The EBs showed increased expression of all markers

with the exception of COL1A1. The cell lines showed similar expression of lineage markers

with a few marked differences. The control and BCL-XL cell lines seemed to have similar

expression although BCL-XL EBs did not up-regulate COL1A1. The CNV was more distinct from

the control and BCL-XL cells in that the EBs did not express the definitive endoderm marker

SOX17 and also did not up-regulate PAX6.

Finally, differentiation towards endoderm also showed variation between the cell lines. EBs formed in control cell lines were very small with evidence of mass cell death. The control cell line up-regulated the endoderm markers SOX17 and GATA6 and showed a slight increase in PECAM, PAX6 and COL1A1 expression. The CNV produced slightly larger EBs that looked similar to those formed in the neutral and ectoderm conditions. Surprisingly the CNV EBs exhibited down-regulation of the endoderm markers unlike the control cell line, the mesoderm markers remained unchanged. There was increased expression of the ectodermal markers PAX6 and COL1A1. The BCL-XL over-expression cell line again formed irregular shaped EBs that appeared both dense and cystic. Mesendoderm markers were again down-regulated unlike the control cell line. The only marker up-regulated in the BCL-XL EBs was PAX6. These findings imply that the CNV may alter the differentiation patterns of human ES cells, some, but not all of the differences between the control and CNV lines are mimicked by over-expressing BCL-XL. This suggests that BCL-XL itself may have an effect on the differentiation of human ES cells as well as other genes contained within the CNV.

Interestingly, the changes in differentiation patterns of culture-adapted human ES cells have only been reported in cell lines containing whole chromosome gains, making potential candidates difficult to identify[76]. The 20q11.21 CNV is small, containing a smaller pool of genes, the results show that CNV cells show altered differentiation patterns following EB formation and to a lesser extent in monolayer differentiation. This effect cannot be attributed completely to the increased survival of cells through BCL-XL protection. The CNV and BCL-XL over-expressing cell lines show similarities in their differentiation, the BCL-XL cell line is more akin to the control cells suggesting another gene within the CNV is influencing differentiation.

In conclusion, directed differentiation seems to be directing cells towards a particular lineage rather than selecting for lineage-primed cells within a heterogeneous population of stem cells.

This is highlighted by the BCL-XL over-expressing cells, which display increased cell survival over control cells particularly when differentiated under endoderm conditions both in EBs and monolayer. Although BCL-XL over-expressing cell lines appear to be unable to differentiate towards endoderm in EBs they readily differentiate under monolayer conditions, suggesting the differences observed between EB and monolayer is a result of inadequate signaling in EB formation (being tightly packed with only outer cells being subject to growth factors).

The CNV-containing cell line showed similar differentiation patterns to the BCL-XL over-expressing cell line possibly due to the increased survival properties of BCL-XL. When differentiated towards endoderm in monolayer conditions the CNV-containing cell line displayed subtle differences when compared with control cells. There were increased cell numbers surviving at the end of differentiation but did not show increased SOX17 positive cells, which was observed in the BCL-XL over-expressing cell line. This could be the effect of other genes within the 20q11.21 amplicon that effect the differentiation of human ES cells.

6. Investigating the role of 20q11.21 CNV in genomic instability of human embryonic stem cells

6.1. Introduction

It has been well established that human ES cells acquire non-random genetic changes throughout culture *in vitro*. However, the mechanism(s) behind the gain of chromosomal material in human ES cells remain elusive. Broadly, genetic changes in human ES cells can be catagorised as follows: (i) numerical aneuploidies – the gain (trisomy) or loss (monosomy) of individual chromosomes, (ii) structural aneuploidies – genetically unbalanced structural rearrangements, including derivative chromosomes from translocations, (iii) point mutations and (iv) epigenetic changes. Numerical and structural aneuploidies represent the majority of reported genomic changes in human ES cells. For example, structural aneuploidies were the most common type of chromosome abnormality observed in the ISCI project[66], accounting for 29 of the 73 (39.7%) individual chromosome abnormalities identified. The second most common abberation was trisomy, which occurred in 27 out of 73 (36.9%) individual chomosome abnormalities identified.

One can draw predictions on the origin of chromosomal abnormalities based on the most commonly gained chromosomes and how they are observed. Chromosomes 1 and 17 often appear as a result of unbalanced translocations with other chromosomes, intra-chromosomal duplication as well as trisomy and isochromosome formation[65, 66]. This suggests that structural rearrangements are a product of non-homologous recombination or errors in DNA repair mechanisms. The gain of chromosome 12 is most commonly observed as a whole chromosome gain or as an isochromosome suggesting errors in sister chromatid separation

during mitosis. However, it remains unclear whether human ES cells are prone to erroneous mitoses.

A key surveillance mechanism responsible for ensuring the correct alignment and segregation of sister chromatids during mitosis is the spindle-assembly checkpoint (SAC). The SAC components are highly conserved throughout evolution and consist of the mitotic-arrest deficient (MAD) genes, MAD1, MAD2 and MAD3 (BUBR1 in humans), and the budding uninhibited by benzimidazole (BUB) genes, BUB1 and BUB3[125]. These proteins localise at the kinetochores at metaphase, and accumulate to high levels on those unattached to microtubules. Once the kinetochore is attached to the spindle, SAC proteins are displaced and degraded by proteolysis. The cell is unable to progress to anaphase until all kinetochores are attached and under equal tension to opposite poles. Li and Nicklas[126] showed by micromanipulation of chromosomes that just one unattached kinetochore was sufficient to delay the onset of anaphase. Once all chromatids are attached to the microtubules and aligned at the spindle equator the irreversible metaphase to anaphase transition is initiated. The final separation of sister chromatids is catalysed by separase, a cysteine protease that is activated by the anaphase-promoting complex (APC). The APC is activated by the binding of an APC-activator such as Cdc20 or Cdh1, creating specific complexes (APC^{Cdc20} and APC^{Cdh1}). APC^{Cdc20} is responsible for the activation of separase, which hydrolyses cohesin resulting in the separation of sister chromatids[125]. The activity of APC is controlled by sequestering activator proteins, at prophase Emi1 binds and inhibits Cdc20. In prometaphase, Cdc20 is sequestered by SAC proteins MAD2 and BUBR1[127]. SAC related proteins are expressed at high levels in cancer cells, possibly linked to increased proliferation of these cells and misregulation of SAC proteins has been linked with aneuploidy [128]. For example, it has been shown that over-expression or knock-down of MAD2 and BUB1 results in premature sister chromosome separation, chromosomal bridging and failure to undergo cytokinesis[129].

Mantel and colleagues[130] reported that although the SAC is fully functional in human ES cells, they fail to prevent polyploidy through re-replication when treated with microtubule poisons. Human somatic cells undergo caspase-mediated apoptosis following prolonged mitotic block showing a coupling of SAC and apoptosis. This SAC-apoptosis coupling seems to be absent in human ES cells that showed a high tolerance for abnormal mitotic divisions. The study concluded that the SAC is uncoupled from apoptosis in human ES cells and becomes coupled during early commitment to differentiation.

Another key player in regulating the bi-orientation of chromosomes at the spindle equator is Aurora B. The Aurora Kinase family are a group of serine/threonine protein kinases and consist of three members, A, B and C, which have different roles in cell division[131]. The three members exhibit strong homology, their catalytic domains share 67-76% similarity in their amino acid sequences. Despite this sequence similarity, Aurora A and B have very different roles in mitosis[132]. In metaphase Aurora A localises on microtubules close to the spindle poles. Using Aurora A-null mouse embryonic fibroblasts, Cowley and colleagues showed that Aurora A it is essential for bipolar spindle formation. Cells deficient of Aurora A displayed delayed mitotic entry and formed a monopolar spindle. The cells failed to undergo metaphase, anaphase or telophase and instead exited the cell cycle without undergoing cytokinesis and often initiated mitotic arrest [133]. At prometaphase Aurora B localises at the centromere where it functions to attach the centromere to the mitotic spindle. Aurora B along with INCENP (Inner Centromere Protein) and BIRC5 (also known as Survivin) form the chromosomal passenger complex (CPC). The CPC regulates key mitotic events including correction of chromosome-microtubule attachment errors, activation of the SAC and regulation of cytokinesis[134]. INCENP and Survivin are required for Aurora B localisation and in turn Aurora B is required for the correct bi-orientation of sister chromatids. Disruption of Aurora B results in defects in chromosomal alignment and failure to undergo cytokinesis[131].

Aurora C is the least studied of the family and is thought to have a similar role to Aurora B. Aurora C has been shown to regulate the correct chromosome segregation during meiosis of mammalian oocytes[135].

The increased levels of Aurora kinases in many human cancers (breast, ovarian, colon and pancreatic) have been linked with increased genomic instability and proliferation. This link has led to the development of Aurora kinase inhibitors to be used in anticancer therapy either as a monothearpy or to increase the efficacy of other treatments such as chemotherapy and ionising radiation[136]. AZD1152 is one such small molecule that is rapidly converted into the active form AZD1152-HQPA (AZD1152-hydroxyquinazoline pyrazol anilide) following administration. AZD1152 is a potent and specific inhibitor of AURORA B (Ki= 1,369nM) being 50-fold more selective than for AURORA C (Ki=17nM) and 1000-fold more selective than for AURORA A (Ki=0.36nM)[137]. AZD1152 has been shown to cause gross aneuploidy, cell cycle disregulation and increased apoptosis in pancreatic and colon cancer cells[136]. The effect of AZD1152 on human ES cells has not been studied and could provide a method of forcing abberant mitoses in human ES cells to study the effect that increased survival of CNV 20q11.21 cells has on acquire further mutations.

Dumitru and co-workers[138] reported that human ES cells are primed to undergo rapid apoptosis following DNA damage. The study found that activated BAX localises at the trans Golgi network and in response to DNA damage rapidly translocates to mitochondria to initiate apoptosis. This effect may be an in vivo fall back mechanism employed to stop the potentially devastating effects of abnormal ES cells progressing through embryogenesis. The increase of BCL-XL in CNV cells may hinder this regulation mechanism by inhibiting BAX initiating apoptosis and may allow abnormal cells to survive and proliferate. Minn and colleagues[139] also showed that BCL-XL over-expression in mouse polymphocytic cells increased the occurrence of

tetraploidy throughout culture. The study also showed that following mitotic spindle damage BCL-XL over-expressing cells were able to progress through the cell cycle whereas control cells underwent mass cell death. The gain of CNV 20q11.21 in human ES cells increases a cells protection against apoptosis, if aneuploid cells are removed from human ES cell cultures via apoptosis, increased CNV protection may result in further karyotypic changes.

The increased genomic instability observed in human ES cells is a major concern for their eventual use in cell replacement therapies. The abnormalities observed point towards errors in mitosis as a likely mechanism behind aneuploidy. The frequency of abnormal mitotic events has not been investigated in human ES cells, insight could provide clues as to the origin of the karyotypic abnormalities. The increased survival observed in cells containing the CNV 20q11.21 may also provide a platform for further genetic changes, if the cells are unable to undergo apoptosis following erroneous mitoses aneuploid cells may escape safety mechanisms protecting the genomic integrity of human ES cells.

Karyotypically normal human ES cell line (ESI-035) and its culture-adapted counterpart

harbouring chromosome 20q CNV (ESI-035-CNV) were transfected with histone H2B:RFP (red

fluorescent protein) fusion protein to enable the visualisation of mitotic chromosomes and

interphase chromatin in real-time. Metaphase to anaphase transition time was determined as

previously described by Meraldi and colleagues [140]. The time taken for mitotic cells to

undergo nuclear envelope breakdown (NEBD) to the complete segregation of sister chromatids

was measured. The first indication of NEBD was set to T=0 and three landmark mitotic events

were determined, prometaphase was determined from NEBD to correct alignment of

chromotids at the spindle equator (metaphase), the initiation of chromosome segregation

(anaphase I) to the complete separation of the two sister chromatids at the cell poles

(anaphase II). This is summarised in Figure 22 (Supplementary Movie 18) demonstrating a

normal metaphase to anaphase transition. White arrows indicate chromosomal material that

has not aligned on the spindle equator. This analysis was applied to both control and CNV cell

lines. Control cells exhibited increased variability and slightly longer metaphase to anaphase

transition times compared with CNV cells. Control cells took on average 29.0±10.3 minutes

from NEBD to complete segregation of chromatids at anaphase II compared to CNV cells, that

lasted on average 22.0±4.5 minutes. This decreased metaphase to anaphase transition times

can be explained by faster prometaphase, metaphase and anaphase times in CNV cells.

Control cells exhibited slightly longer prometaphase times, lasting on average 14.9±4.5

minutes compared to CNV times of 13.4±4.3 minutes. Control cells also displayed longer

metaphase times, lasting on average 9.8±8.1 minutes, which was slightly longer and more

variable than CNV cells (5.0±3.1 minutes). The time taken to complete anaphase was also

slightly longer and more variable in control cells (4.3±1.2 minutes) compared with CNV cells (3.6±3.7 minutes).

The cells were also analysed for abnormal chromosome segregation (Figure 23). In 21 out of 73 events (29%) the control cells showed evidence of abnormal mitoses, where lagging chromosomes or chromosomal bridges were observed (Figure 24) (Supplementary Movies 19-23). These events were decreased in CNV cells with only 10 events out of 73 (14%). However, there were an increased number of mitotic events with evidence of micronuclei in CNV cells (6 out of 73) than control cells (2 out of 73). There were also six incidences of daughter cells dying shortly after mitosis in control cells; there was no obvious abnormality in two of these events, although in the remaining cases, one died during metaphase without initiating anaphase and the reamaining three events were preceded by a noticeable abnormality in chromosome segregation. The CNV cells showed only one case of daughter cell death following mitosis, this event was not preceded by any noticeable abnormality in its division. This high proportion of erroneous mitoses in control cells could underpin the high frequency of chromosomal abnormalities in human ES cells. The fact that CNV cells display less abnormal divisions may suggest that genes within the amplified region may help to stabilise genetic instability.

Figure 22. Measuring metaphase to anaphase transition in human embryonic stem cells
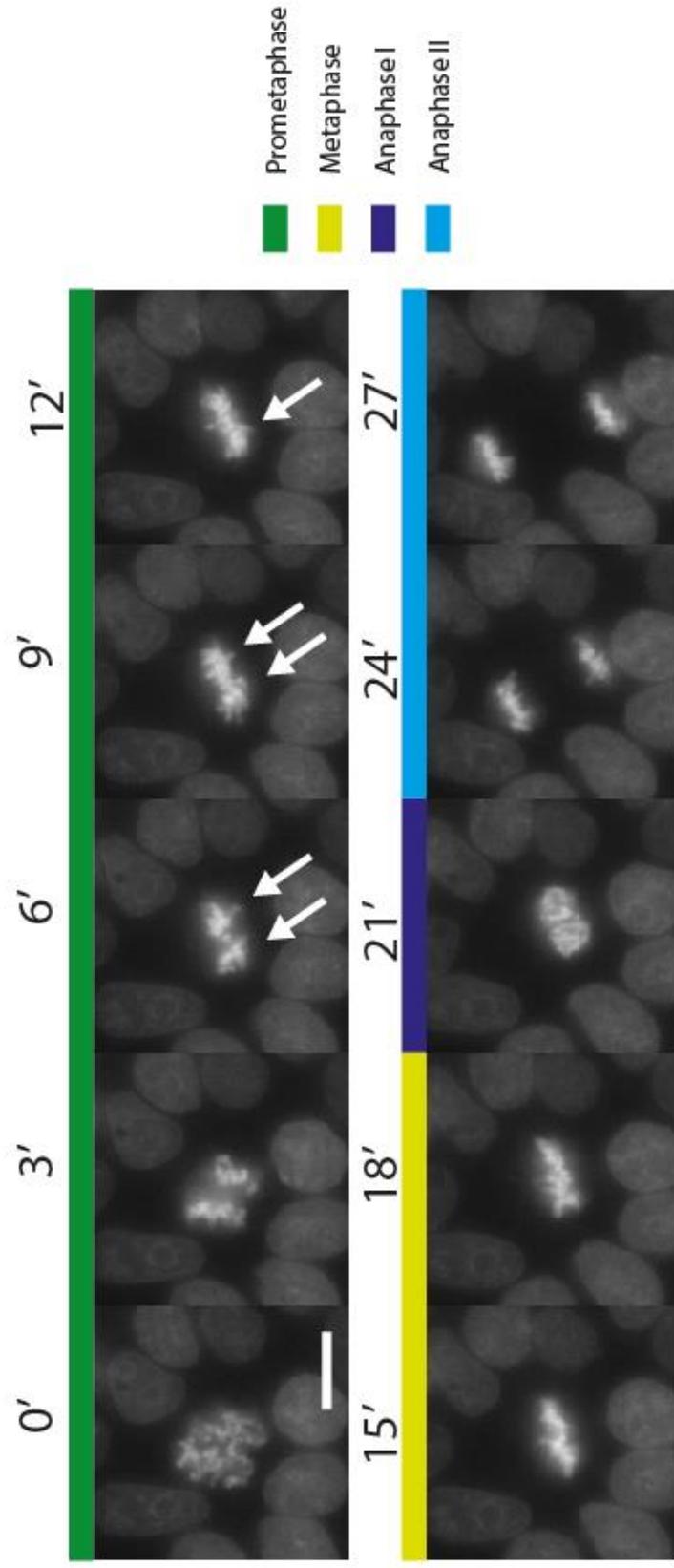
Example of the key events when measuring metaphase to anaphase transition in human ES cells line. The ESI-035-H2B-RPF cell line was imaged every minute using live-cell time-lapse microscopy. Key events in the progression from nuclear envelope breakdown (T=0) to the complete segregation of sister chromatids at anaphase are colour coded; the condensation of chromosomes and organisation of chromosomes on the spindle equator during prometaphase (green), the correct alignment at metaphase (yellow), the initiation of anaphase (purple) and complete segregation of sister chromatids at the spindle poles (blue). Unnattached chromatids that have not aligned on at the spindle equator are marked with a white arrows. The time in minutes is indicated above each frame. Scale bar = 10μm.

Figure 23. Human embryonic stem cells are prone to erroneous mitosis
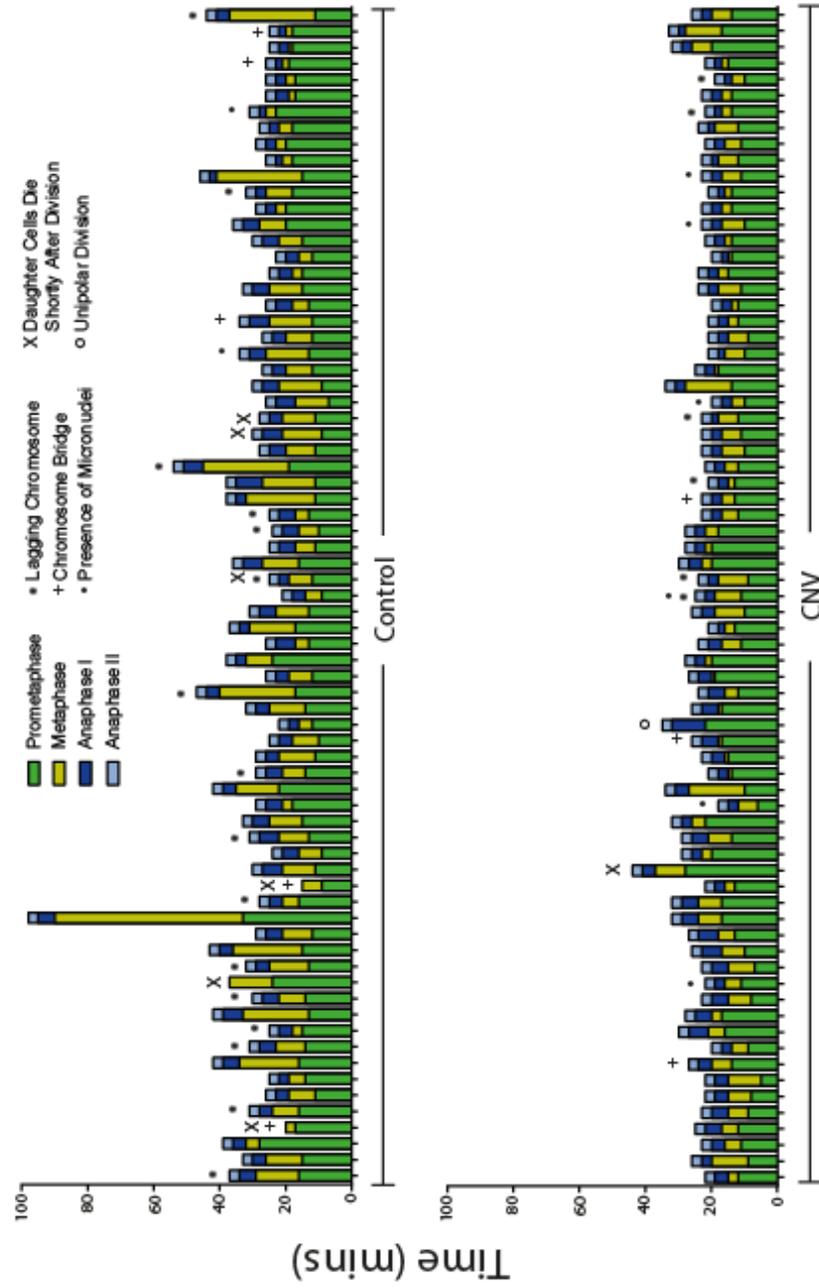
Individual mitosis events in ESI-035-H2B-RFP Control and CNV cell lines were analysed using live-cell time-lapse imaging. The time taken for individual cells to complete mitosis are represented by separate bars. Control cells displayed slightly longer and more variable mitosis times. ESI-035 Control cells also displayed increased numbers of erroneous mitosis events compared with the ESI-035 CNV cell line. Increased incidences of lagging chromosomes (*) and chromosomal bridges (+) were observed. Control cells were also more likely to die shortly after division (X) than CNV cells. Evidence of micronuclei (·) and unipolar division (o) were also observed in dividing cells.

Figure 24. Examples of abnormal chromosome segregation in human ES cells
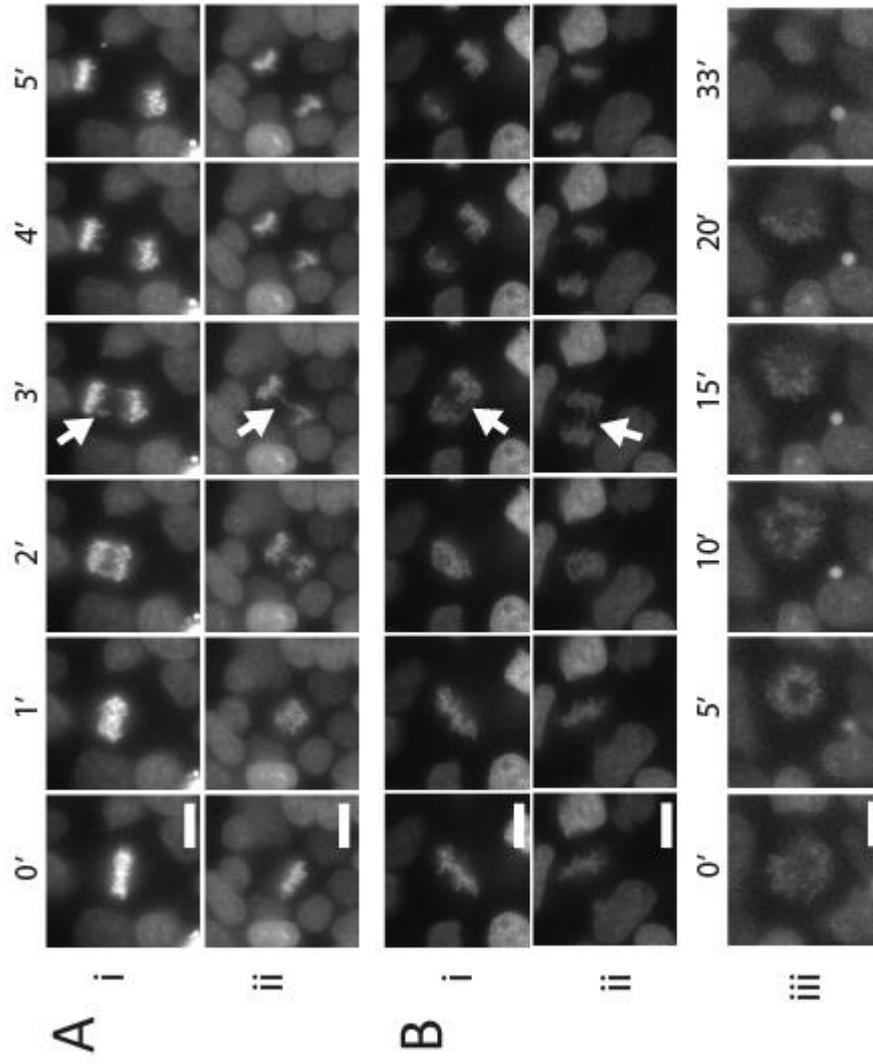
(A) Control and (B) CNV cells are prone to errors in chromosomal segregation, evidence of lagging chromosomes (i) and chromosomal bridges (ii) were observed. (iii) shows the observed unipolar division. White arrows highlight abnormal divisions. The time in minutes is indicated above each frame. Scale bar = 10µm.

6.2.2. Aurora Kinase Inhibition Causes Polyploidy in Human ES Cells

To explore the hypothesis that CNV cells would be more likely to survive abnormal chromosomal segregation, abberant mitoses were induced by culturing the cells in the presence of AZD1152. ESI-035 control, CNV and BCL-XL over-expressing cells were cultured in mTeSR containing either DMSO or 100nM AZD1152 for 24 hours. Following treatment the cells were live-stained with Hoechst 34442 and analysed using flow-cytometry (Figure 25A). Cells cultured in the presence of DMSO displayed normal cell cycle profiles with peaks representing G1 and G2/M and intermediates representing cells in S-phase. ESI-035 control and CNV cell lines displayed similar cell cycle profiles with BCL-XL over-expressing cells displaying slightly larger G2/M populations. However, upon AZD1152 treatment cells displayed grossly abnormal cell cycle profiles. Both control and CNV cell lines displayed sub G1 peaks characteristic of apoptotic cells. This population was absent in BCL-XL over-expressing cells demonstrating their increased survival. G1 populations were greatly diminished in all three cell lines, instead large G2/M populations were observed. A large peak of cells representing an 8N population was detected in all cell lines, this population was largest in the CNV cell line. ESI-035 control and BCL-XL over-expressing cells displayed similar DNA profiles following AZD1152 treatment with low 2N, tight 4N and a large 8N population. The ESI-035 CNV cell line displayed a low 2N population, smaller 4N population and larger 8N population.

To confirm this result, ESI-035 control and CNV H2B:RFP reporter cell lines were seeded as described in Section 3.2.3 and grown for two days in mTeSR. The cells were then subjected to 100nM of AZD1152 inhibitor. Cells were imaged every ten minutes for 24 hours and time-lapse movies were analysed. Both the ESI-035 control and CNV cells showed grossly abnormal cell division following culture in AZD1152. Cells appeared to be unable to organise proper chromosomal alignment, undergo proper chromosomal segregation at anaphase or initiate

cytokinesis. Instead cells appeared to exit the cell cycle at prometaphase resulting in cells becoming polyploid either 4N or 8N (Figure 25B). This progression through cell cycle checkpoints leads to increased DNA content.

6.2.3. Amplification of 20q11.21 Decreases Sensitivity To Aurora Kinase Inhibition

The ESI-035 control, CNV and BCL-XL overexpressing cell lines were seeded at 5.5×10^4 cells/cm^2 in 6-well plates coated with matrigel supplemented with Y-27632 and left to attach for 24 hours. Cells were washed once with PBS (1X) and mTeSR containing 100nM AZD1152 was added to wells. Cells were harvested at 0, 6, 12, 24 and 48 hours following inhibition. The 0 hour time point indicates the levels of apoptosis in unperturbed cells before addition of AZD1152 inhibitor. One well was washed following 24 hours inhibition, washed once and fed with mTeSR without AZD1152 inhibitor in order to 'release' cells so that they are able to undergo mitosis.

Control cells show increased apoptosis upon AZD1152 treatment over 48 hours. Cells appear to initiate apoptotic pathways between 12 and 24 hours with only 20% of healthy cells present at 48 hours. The cells can be rescued by removing the inhibitor at 24 hours and culturing in mTeSR. This increases the amount of healthy cells at 48 hours and lower cells undergoing apoptosis. In contrast the ESI-035-CNV cell line shows protection to the aurora kinase inhibitor, there are three times as many healthy cells than control cell lines. The CNV cells also show lower apoptotic cells following release from the inhibitor. Cells over-expressing BCL-XL showed no signs of undergoing apoptosis in response to AZD1152 treatment.

Figure 25. Inhibition of AURORA KINASE B causes polyploidy in human embryonic stem cells
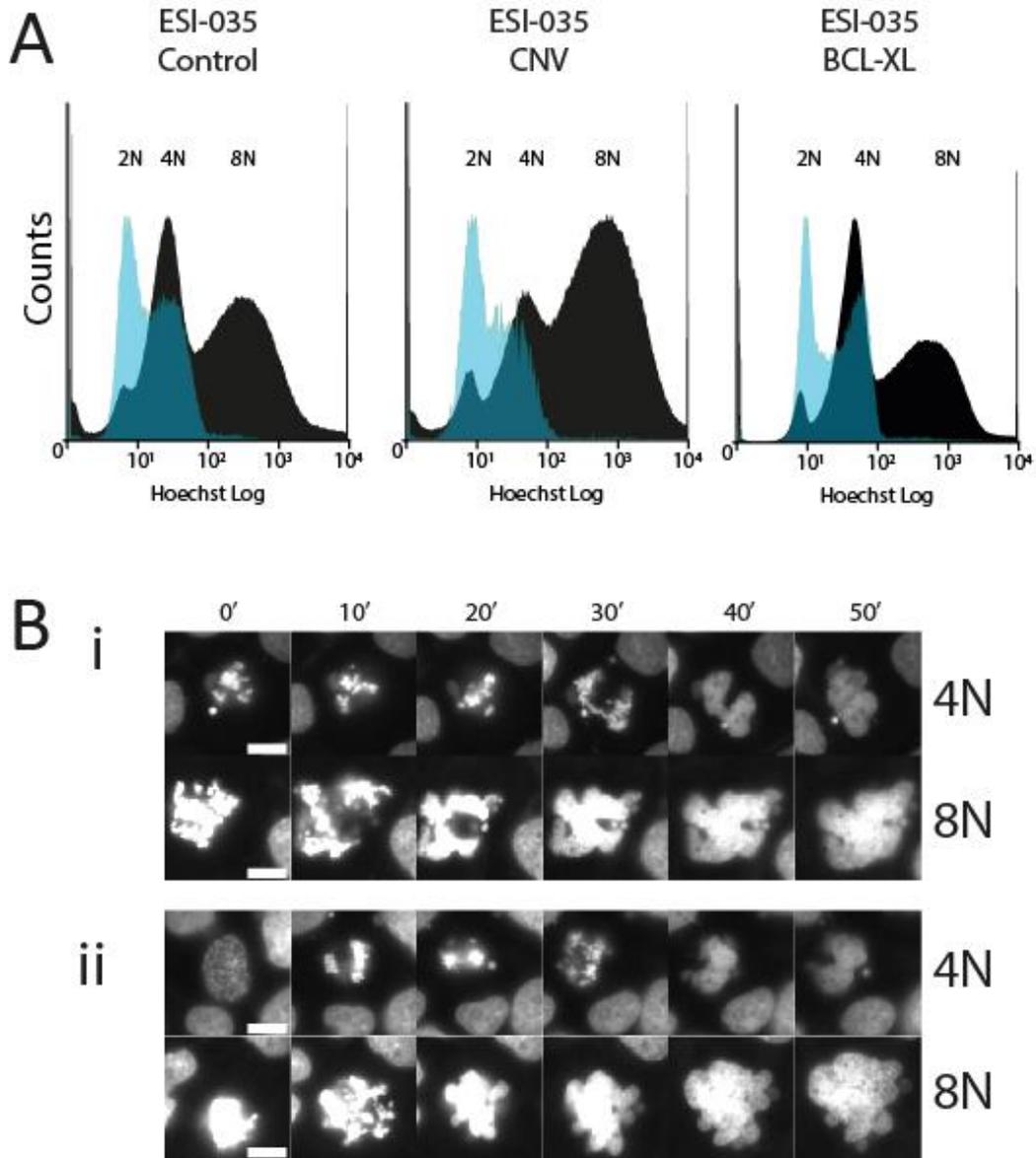
(A) ESI-035 Control, CNV and BCL-XL over-expressing cell lines were cultured either in the presence of DMSO (blue histogram) or 100nM AZD1152 (black histogram) for 24 hours. The cells were then stained using Hoechst 33342 and their DNA content analysed. Following treatment with AZD1152, all cell lines become polyploid displaying large 4N and 8N populations. (B) Both the ESI-035-H2B-RFP Control and CNV cell lines were imaged every ten minutes using live-cell time-lapse microscopy following treatment with AZD1152 (100nM). (i) Control and CNV (ii) cell lines fail to achieve metaphase or undergo cytokinesis resulting in the appearance of 4N and subsequently 8N cells. The time in minutes is indicated above each frame. Scale bar = 10µm.

Figure 26. CNV-containing cells display increased survival following AURORA KINASE B inhibition
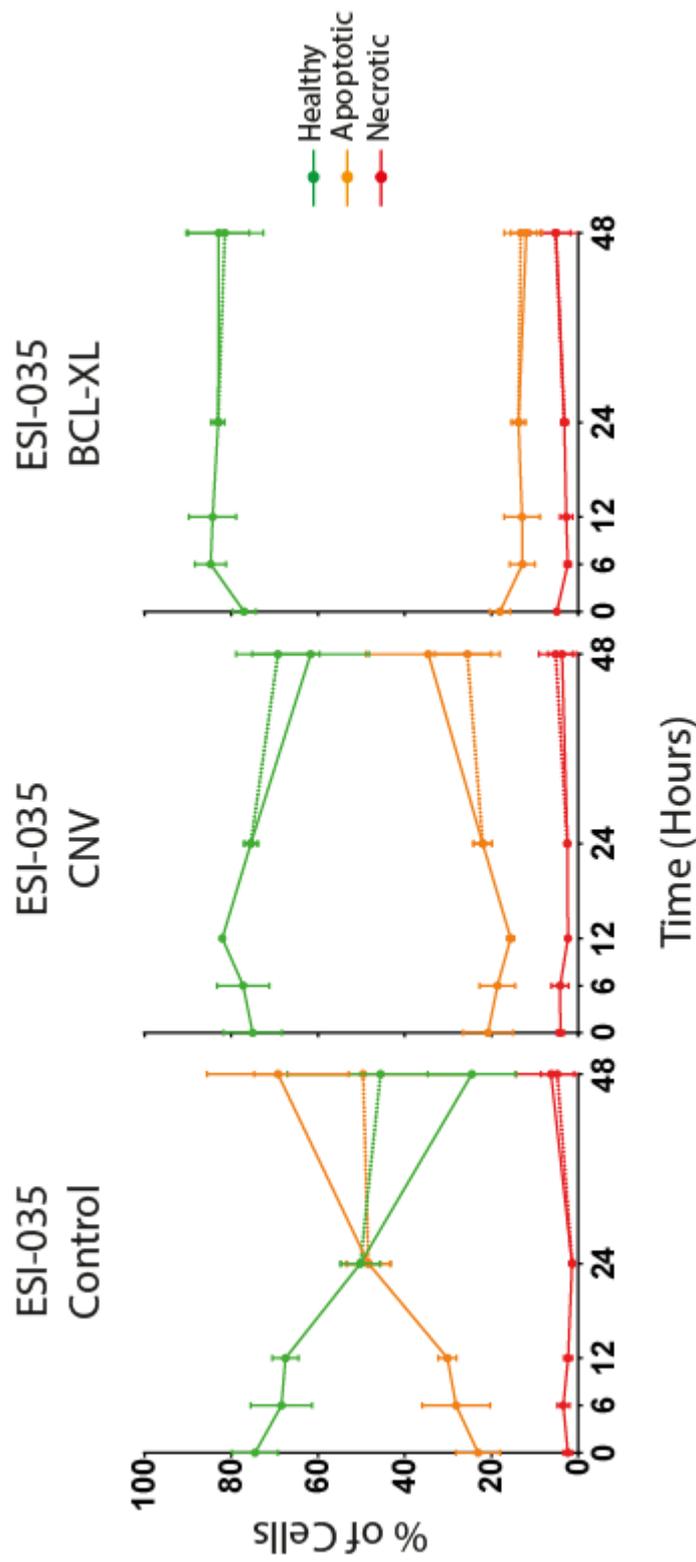
ESI-035 Control, CNV and BCL-XL over-expressing cell lines were cultured for 48 hours in the presence of 100nM AZD1152. The levels of apoptosis were measured at 6, 12, 24 and 48 hours following inhibitor exposure. Solid lines indicate: healthy cells (green), apoptotic cells (orange) and necrotic cells (red). Dashed lines indicate the removal of AZD1152 inhibitor following 24 hours exposure. Control cells display increased cell death following 24 hours exposure when compared with CNV and BCL-XL over-expression cells. Error bars represent SEM from three biological replicates.

6.3. Discussion

The cell cycle times as determined by time-lapse microscopy in Section 3.2.5 showed that the difference between the ESI-035 control and CNV cell lines were very similar (19.5 and 18.5 hours respectively). Interestingly the metaphase to anaphase transition times were significantly different (p= <0.0001) with the control cell line displaying times of 29±10.3 minutes and the CNV cell line displaying faster and more uniform times of 22±4.5 minutes. The increased variability in transition times was observed in the time taken for control cells to initiate chromosomal segregation following proper alignment of sister chromatids at the spindle equator. The control cells took on average 9.8±8.1 minutes to progress from metaphase to initiation of anaphase whereas the CNV-containing cells took 5.0±3.1 minutes. Prolonged metaphase is observed in cells that have undergone significant DNA damage, Mikhailov and colleagues[141] showed that by inducing DNA breaks in late prophase cells delayed the metaphase to anaphase transition in human cells. This delay was not a result of ATM-kinase mediated DNA damage checkpoint but due to defective kinetochore attachment via a MAD-2 dependent mechanism. The difference in the number of cells delayed in metaphase between control and CNV cells could highlight one of two possible explanations; either the control cells are more susceptible to DNA damage than CNV cells or that CNV cells are able to over-ride checkpoints and progress from metaphase to anaphase without being delayed. Prolonged metaphase transition has also been suggested to precede abnormal chromosome segregation Much of the heterogeneity in metaphase timings is a result some cells taking longer than others to correctly organise chromosomes at the spindle equator[140].

The control cells displayed increased incidents of abnormal divisions with 29% (21 out of 73) of anaphase events containing a lagging chromosome (16 events) or a chromosomal bridge (5 events). This was in stark contrast to the CNV cells in which only 13% of divisions were found

to be abnormal. There were greatly reduced numbers of lagging chromosomes (6 events) and fewer chromosomal bridges (3 events) observed. A single event of unipolar division was observed in the CNV cells, an event that was not observed in the control cell line. There was also more dividing cells that contained evidence of micronuclei in CNV cells (6 events) compared with control cells (2 events). The decreased number of abnormal divisions in CNV cells could highlight one of two possibilities, CNV cells are somehow more controlled in terms of chromosomal segregation or are more efficient at initiating anaphase.

The increased incidents of abnormal divisions in control cells may underpin the frequent gain of chromosomal material in human ES cells. It has not been addressed whether all chromosomes are equally susceptible to abnormal segregation, or that the frequent gain of specific chromosomes is a result of their tendency to lag or bridge. Just under a third of cell divisions analysed displayed a noticeable abnormality, whether these cells die later will need to be addressed. The limitation in investigating this notion is the resolution needed to identify these events, the cells were imaged every minute at 100X in order to capture these events, which limits the time that the cells can be imaged due to photo-bleaching and photo-toxicity. It is striking that CNV-containing cells showed decreased numbers of abnormal divisions suggesting that the CNV may somehow stabilise the genetic integrity of the cells so that it is less likely to acquire further changes once the cells have acquired the 20q11.21 CNV. However, there were also decreased incidents of daughter cells dying following abnormal divisions. This observation may be a result of the limited time that the cells were imaged after the events were observed, more cells may die later.

Aurora Kinase inhibitor was used to promote polyploidy in human ES cells. The control and CNV cell lines showed evidence of polyploidy through two cell divisions producing 4N and 8N cells respectively. The control and BCL-XL over-expressing cell lines displayed similar DNA

profiles, they displayed small 2N (G0/1) population, with tight 4N (G2/M) population and a large 8N population. The DNA profile of the CNV cell line following AZD1152 treatment was slightly different. The CNV also contained a small 2N population, but displayed a smaller 4N and larger 8N peak suggesting that another gene within the CNV could be affecting the cells. This DNA profile was assessed 24 hours following AZD1152 treatment, apoptosis assays show that over half of the control cells had died within the first 24 hours. Longer treatment showed further apoptosis down to 70% of cells. Removing the inhibitor after 24 hours and replacing with fresh mTeSR, the cells stabilised at 50% apoptosis. The CNV cells showed reduced sensitivity to AZD1152 treatment where only 25% of cells were undergoing apoptosis following 24 hours. This supports the notion that CNV cells are more likely to survive aneuploidy. Interestingly when cells were sorted based on a cells DNA profile using FACS, under DMSO conditions the CNV cells displayed increased cloning efficiency based on their cell number when sorted at 2N and 4N. Sorting cells after AZD1152 treatment reduced the number of cells following 4 days growth when compared with cells treated with DMSO indicating that the cells are not proliferating as fast as normal.

In conclusion, we have shown that human ES cells are prone to chromosome segregation errors. These abnormal events appear to be more abundant in the ESI-035 Control cell line than in the ESI-035 CNV cell line. This may suggest that other genes on the 20q11.21 amplification may help to control these events in CNV-containing cells. Although there were increased abnormal events in the control cell line, more of these events resulted in cell death of daughter cells shortly after division, this was not observed in CNV-containing cells. This observation was limited as time-lapse movies were relatively short (2-4 hours) and would require further investigation. Whether the rates of abnormal mitoses events are the same in all human ES cell lines is unknown but could be investigated using H2B-RFP reporter lines. The CNV-containing cell line was more robust when treated with AURORA KINASE B inhibitor,

which induced polyploidy in human ES cells. This could suggest that the CNV-containing cells

are more likely to survive abnormal mitoses events and could lead to further genomic changes.

7. Developing Detection Methods for CNV 20q11.21

7.1. Introduction

It has been widely reported that human ES cells can acquire karyotypic changes throughout culture in vitro, which raises safety concerns for their potential use in regenerative medicine[61, 65, 66]. The fact that many human cancers also acquire similar karyotypic changes exacerbates this concern[142]. It is therefore imperative that karyotypic changes are further investigated to better understand how human ES cells are affected and to develop methodologies to detect, monitor and manage genomic instability.

Another commonly used method for analysing the genetic integrity of human ES cell lines is the g-banding of metaphase spreads. Cells are arrested in metaphase by exposing them to spindle poisons, such as colcemid, and this is then followed by preparation of chromosome spreads on slides and staining to visualise chromosome banding pattern. G-banding allows examination of the entire complement of cells' chromosomes in a single assay, but it cannot readily detect duplications or deletions smaller than 5 Mb. The limited resolution of g-banding is particularly concerning given the recent observations that CNV 20q11.21 is a particularly frequent mutation in human ES cells[66, 81-84]. Due to the size of the 20q11.21 amplification (0.6Mb to 2.5Mb) it falls below the detection limit of g-banding. An alternative method that does allow detection of small duplications and deletions is Fluorescence in-situ hybridisation (FISH).

FISH can be used for a variety of different applications to visualise DNA and RNA. FISH is routinely used as a diagnostic screen by cytogenetics for the detection of DNA copy numbers to predict genetic disorders or disease prognosis[143]. Sequence-specific probes tagged with

fluorescent dyes are hybridised to targeted regions on DNA allowing the number of copies to be measured. As with g-banding there are limitations to FISH, for example, where g-banding can be used to analyse the whole genome, FISH focuses on smaller regions of interest on particular chromosomes. Both g-banding and FISH usually require sending samples to experienced cytogeneticists, particularly to identify some of the more complicated rearrangements observed in human ES cells. Such labour-intensive and expensive methods are limiting the frequency of assessing genetic stability of human ES cells during routine maintenance. Hence, there is a need for quick, cost-effective assays that could be employed in common laboratory practice for regular screening of human ES cells. Quantitative-PCR (qRT-PCR) offers an alternative method for detecting karyotypic changes[144, 145]. The use of qRT-PCR to detect changes in DNA copy number is still relatively new but offers many advantages such as detection of multiple loci in single experiments, simple assay design, fast turnaround and low cost compared to other methodologies.

The Roche Universal Probe Library (UPL) was used to develop assays for the detection of 20q11.21 amplification (Figure 27). The UPL utilises 165 short hydrolysis probes (8-9 nucleotides) which are labelled with a fluorescein amidite (FAM) dye at the 5' end and a dark quencher dye at the 3' end. The dark quencher molecule absorbs the fluorescent light emitted from FAM when in close proximity therefore fluorescence is not detected until the FAM molecule and dark quencher are separated. As the temperature rises during the polymerase chain reaction, double stranded DNA is denatured resulting in single strands. The temperature is then decreased in the annealing stage and the hydrolysis probe and primer sequences bind complimentary sequences. The DNA polymerase associates with the primers and initiates strand synthesis in a 5'-3' direction, the exonuclease activity of DNA polymerase removes the nucleotides of the hydrolysis probe thereby releasing the FAM dye where the fluorescence can be detected. The hybridisation probes take advantage of locked nucleic acids, which increase

specificity through higher melting temperatures. The probes bind sequence specific regions along the genome, the short lengths ensure maximum coverage of a number of species including but not limited to human, mouse, rat and zebrafish. The specificity comes from the pairing with designed primers, which will amplify the regions flanking the probe.

The aim was to develop a qRT-PCR method to detect the presence of CNV 20q11.21 in human ES cells that was sensitive and both time and cost-effective.

Figure 27. Quantitative Polymerase Chain Reaction: Universal Probe Library
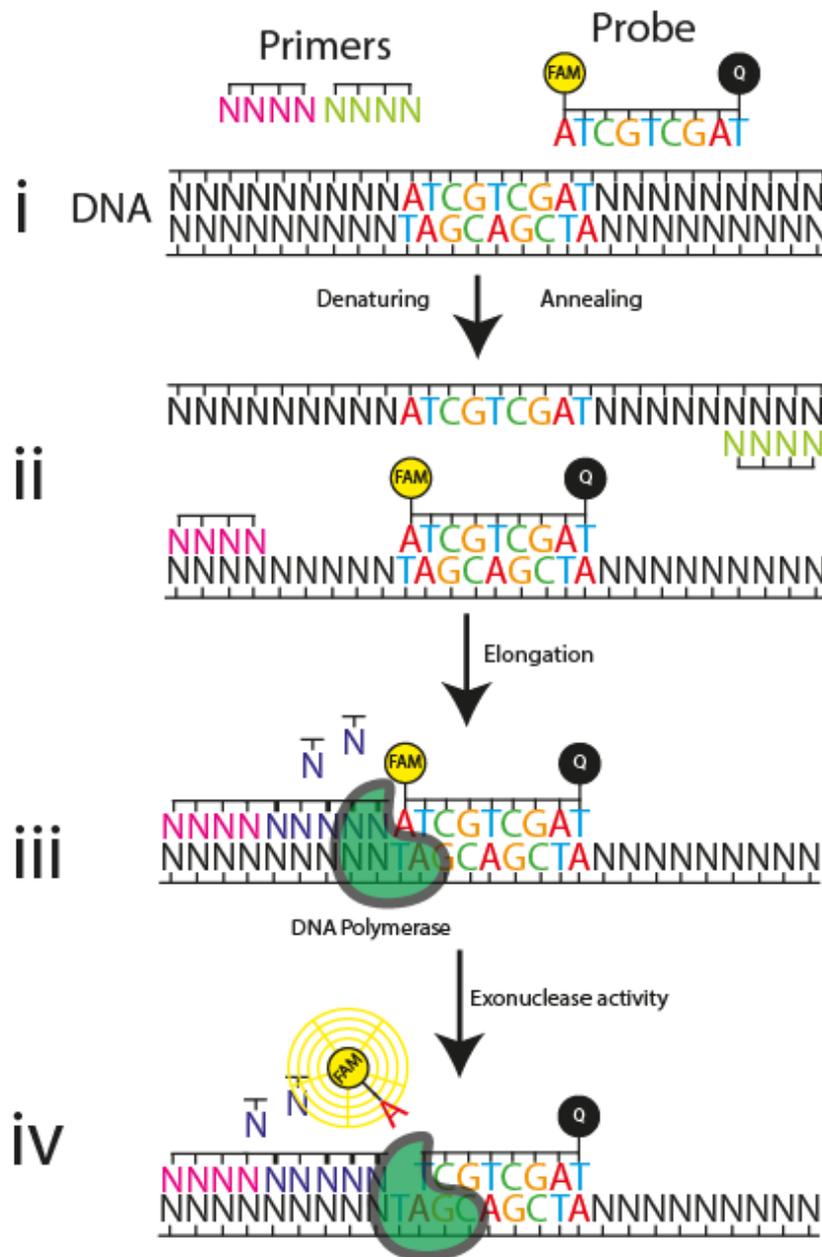
Schematic representation of the theory behind the Roche UPL: UPL probes are labelled with a fluorescein amidite (FAM) dye at the 5' end and a dark quencher dye (Q) at the 3' end. The dark quencher molecule absorbs the fluorescent light emitted from FAM when in close proximity. (i) The reaction requires the presence of DNA and specific primer/probe pairs, (ii) as the temperature rises, double stranded DNA is denatured, the temperature is the lowered allowing the binding of primers and probes to sequence specific locations on the single stranded DNA, (iii) the subsequent elongation stage allows binding of DNA polymerase which binds to primers and creates a new DNA strand in a 5' to 3' manner, (iv) the exonuclease activity of DNA polymerase removes the probe nucleotides therefore releasing the FAM dye where it can be detected.

7.2. Results

7.2.1.   Primer and Assay Design for qPCR of CNV 20q11.21

Primers were designed to intronic regions of 10 genes spanning the 20q11.21 region. This would allow the length of CNV to be determined as well as the average number of copies in a population. Figure 28 shows the location of primers and where they are situated along the CNV. To allow normalisation, primers were also designed for reference genes, which were chosen on chromosomes that are not commonly amplified in human ES cells. The chosen reference genes were two genes from chromosome 4 (RELL1- 4p14 and EPHA5-4q13.1) and one from chromosome 14 (RPPH1- 14q11.2).

The genomic sequences of target genes were obtained from the UCSC Genome Browser website (https://genome.ucsc.edu). Intronic regions were then analysed using the Roche Universal Probe Library Assay Design Center (http://lifescience.roche.com/shop/en/us/overviews/brand/universal-probe-library) for suitable primer-probe pairs. The software returns a number of primer-probe sets, which are ranked in order of their overall score, based on amplicon length and *in silico* PCR. The primers from each set were analysed using the National Center for Biotechnology Information (NCBI) Primer Blast software (http://www.ncbi.nlm.nih.gov/tools/primer-blast) to check for unspecific targets or primer homo/heterodimers. Two primer-probe sets for each gene were obtained and tested on known samples. The primer-probes used are shown in Table 2.

Figure 28. The location of primer/probe pairs designed along the 20q11.21 region
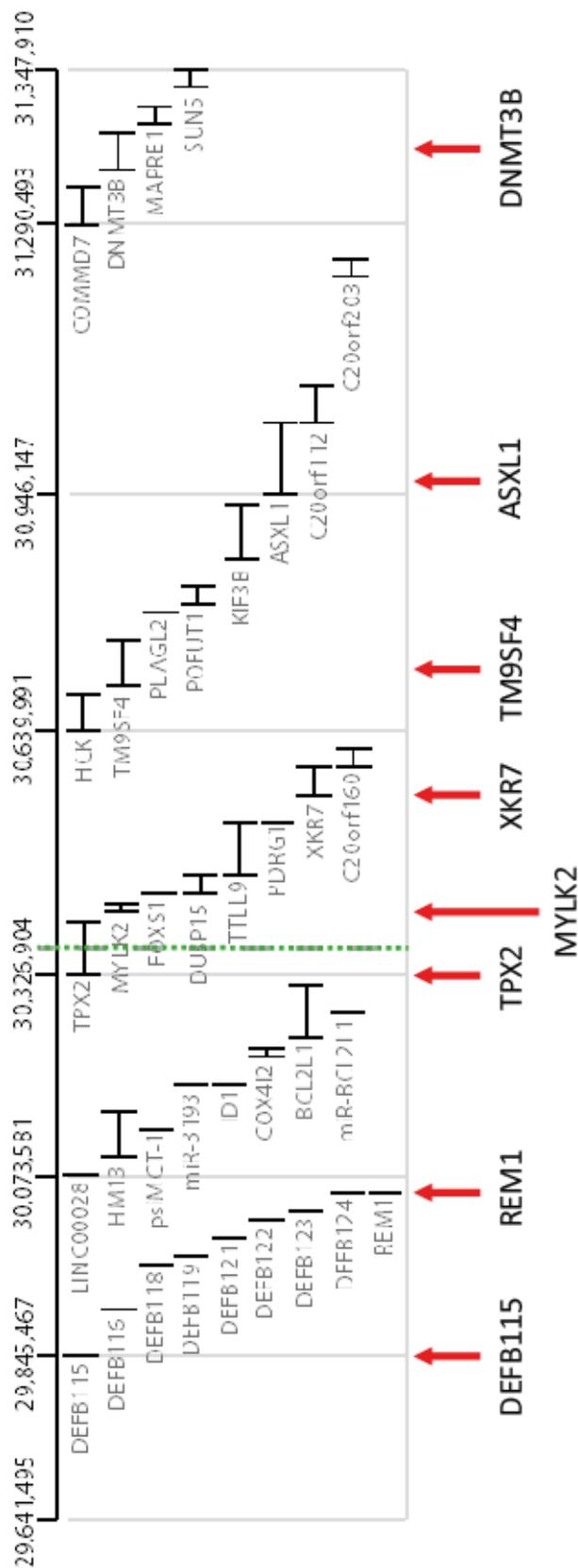
Schematic representation of the 20q11.21 region, the red arrows show the locations of primer/probe pairs designed to detect the length of amplifications in human ES cell lines. The green dashed line indicates the location of the minimal amplicon. Genomic positions relate to USCS human genome assembly version hg19{Kent, 2002 #236}.

7.2.2. Detection of CNV 20q11.21 In Human ES Cells Using qRT-PCR

To test the ability of the designed assay to detect CNV 20q11.21, genomic DNA was prepared from the four test cell line pairs (Control and CNV-containing cell lines from Shef5, ESI-035, H1 and HES3), which were previously shown to contain the CNV by FISH. These cell lines also contained the varying lengths of the CNV as determined by SNP-array[66]. The CT values of genes spanning the 20q11.21 locus were normalised to CT values of the reference gene RELL1 to obtain ΔCT (dCT) values. The dCT values of CNV-containing cell lines were then normalised to their control counter-parts to generate ΔΔCT values (ddCT). Copy number was then calculated using the 2^{-ddCT} method. Thus, the qPCR method was able to detect the CNV in all of the four cell lines. Moreover, by using the primers for ten genes spanning the CNV, this assay showed the varying lengths of CNV. For example, the Shef5 CNV cell line was found to contain an amplicon length of ~0.7Mb whereas the HES3 CNV was found to contain an amplicon length of ~1.5Mb (Figure 3A). These amplicon lengths are consistent with those found by the ISCI[66], which were found to be 0.8Mb and 1.7Mb respectively. Once, the assay was validated on the four control cell lines known to have an amplification off CNV, it was then used to test the presence of the CNV in a number of cell lines grown in the lab (Figure 29).

The Shef6 cell line was subjected to the assay and initially did not show amplification of the CNV (Figure 30A). Given the potential low sensitivity of the qPCR method in detecting mosaicism in culture, the Shef6 cell line was cloned by single cell deposition in 96 well plates to obtain clonal control cell lines that could be used as a reference for normalising test samples. Six clones were screened to check for absence of the CNV, but in the process one clone containing the CNV was found (Figure 30 B-C). Subsequently, Shef6 clones with (Shef6-CNV) and without the CNV (Shef6 Control) were grown out and kept as reference samples.

Figure 29. Genomic qRT-PCR for the 20q11.21 locus detects amplification in other human embryonic stem cell lines
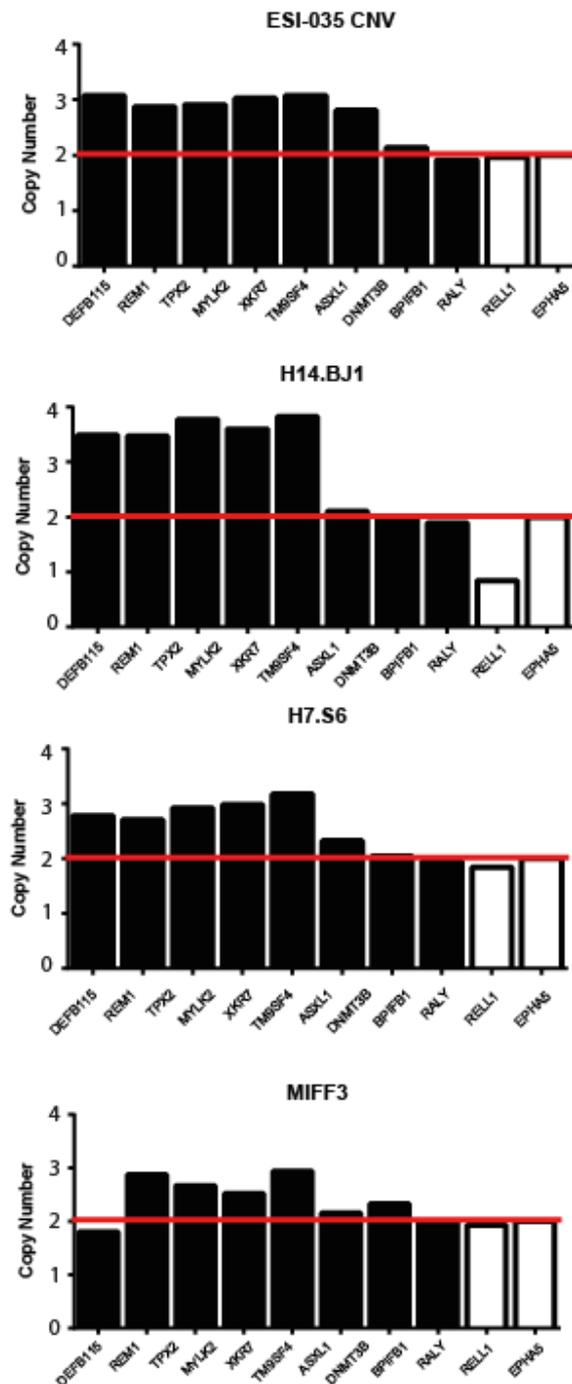
A number of human ES cell lines used within the laboratory were tested for the amplification of 20q11.21. The culture-adapted human ES cell lines H7.s6 and H14.BJ1, and the integration free iPS cell line MIFF3 showed the presence of CNV 20q11.21 amplification. ESI-035 CNV was used as a positive control for 20q11.21 amplification. CT values were normalised to EPHA5 (Chr4), RELL1 (Chr4) was used as a second control. ddCT values were obtained by normalising dCT values to ESI-035 Control samples and copy number generated using the 2^{-ddCT} method.

Figure 30. Single cell cloning reveals underlying population of cells containing the 20q11.21 amplification
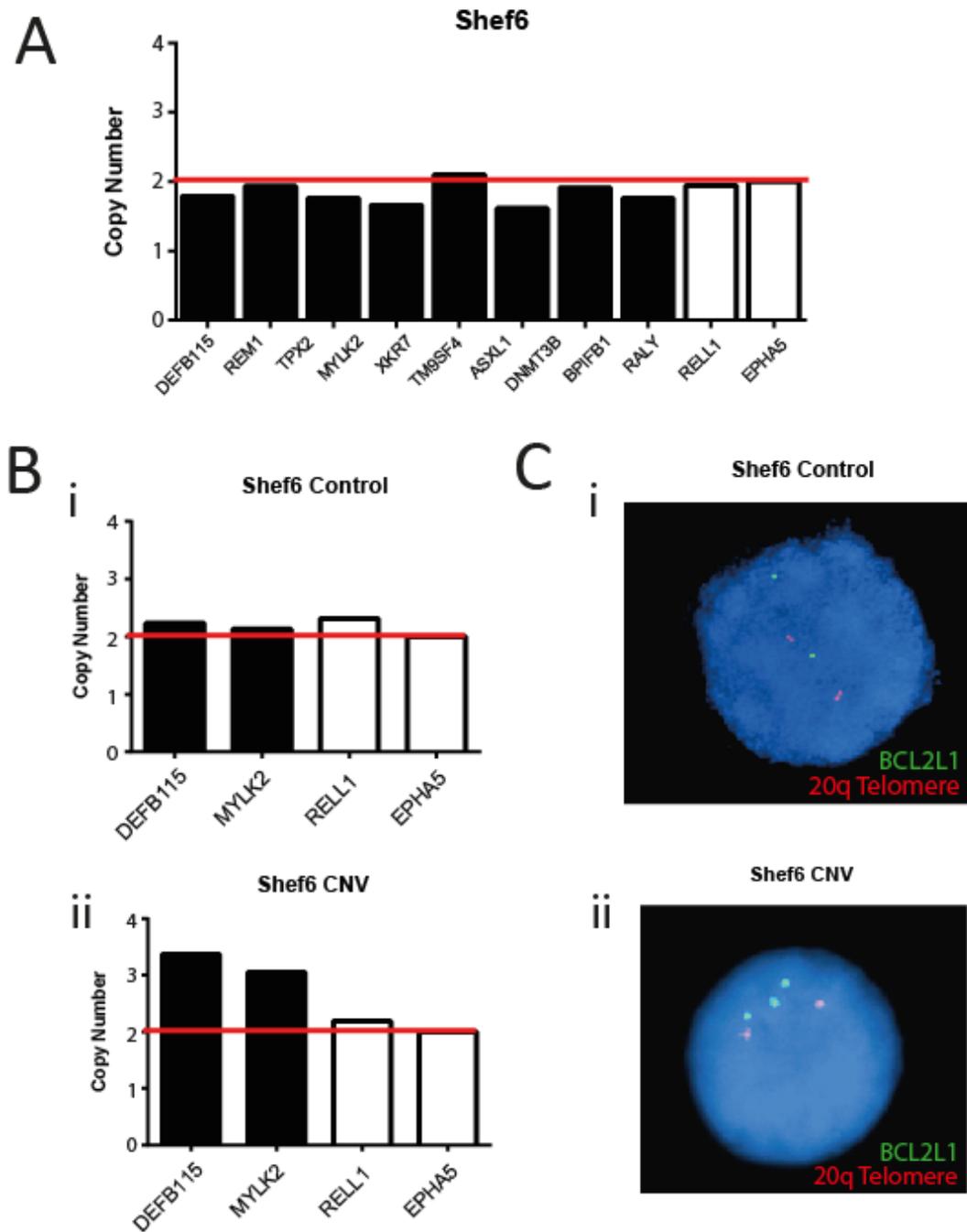
During a screen of human ES cell lines, the Shef6 cell line showed no amplification of 20q11.21 as detected by qRT-PCR. The subsequent cloning and screening of clones revealed one of six clones contained the 20q11.21 amplification as detected by qRT-PCR and FISH. (A) qRT-PCR performed on the parental Shef6 cell line before cloning. (B) Two clones showing the absence (i) and presence (ii) of 20q11.21 amplification using qRT-PCR. (C) These results were confirmed using FISH to detect BCL2L1 (green) copy number in Shef6 control (i) and Shef6 CNV (ii).

7.2.3. Assay Sensitivity

To determine the sensitivity of the assay in detecting the gain of CNV20q11.21 in mosaic human ES cell cultures cell-mixing experiments were performed. For this, the clonal Shef6 cell lines with or without the CNV gain were used (see section 7.2.2). The two cell lines were dissociated to single cells using TrypLE and total cell counts determined. The cells were then mixed at known ratios of Shef6 CNV within the Shef6 Control cells: 0.01%, 0.1%, 0.5%, 1%, 5%, 10%, 20%, 40%, 80% and 100%. Each sample was then split into two tubes, one of which was used for FISH and the other one for gDNA preparation. The results of the triplicate experiments were consistent, showing that qPCR successfully detected the presence of mosaicism in samples of where the variant cells were present at 20% or more (Figure 31Ai-Ci). In comparison, the detectable limit of FISH was much lower than qPCR, with the detection limit as low as 5-10% (Figure 31 Aii-Cii).

7.3. qRT-PCR can be used to detect other common genetic mutations

Culture-adaptation is often associated with non-random gains of chromosomal material, the gain of partial or whole chromosomes 1, 12, 17 and 20 are frequently observed in human ES cell cultures. To detect amplification of these regions, a primer-probe pair was designed to the long (q) arm and short (p) arm of each chromosome on a region that was most commonly amplified (see Table 2). This allowed the distinction between whole and partial gains of chromosomal material. Figure 32A shows the positions of each primer-probe pair along with the reported gains of each chromosome. Two genes, RELL1 (4p) and EPHA5 (4q) were chosen as reference genes due to the apparent genomic stability of chromosome 4, where there have been no reported gains or losses in the literature and would provide a more stable reference with which to normalise target sequences. The genes NANOG (12p) and MDM2 (12q) were

chosen to detect changes on chromosome 12, genes COPS3 (17p) and BIRC5 (17q) were chosen for the detection of chromosome 17 amplifications, and the gene BCL2L1 was chosen to detect the presence of 20q11.21 amplification.

7.3.1. Analysing Human ES Cell Lines for Karyotypic Changes

Six human ES cell lines were tested for common karyotypic changes using qRT-PCR (Figure 32B). The H7 normal (H7.s14) and culture-adapted (H7.s6), the H14 normal (H14.s9) and culture-adapted (H14.BJ1) and two iPS clones generated using mRNA reprogramming (MIFF1 and MIFF3). The early passage H7.s14 cell line showed no amplification of any of the regions tested but the later passage sub-line H7.s6 showed a single amplification (3 copies) of both BCL2L1 and BIRC5 situated on 20q11.21 and 17q25.3 respectively. The early passage H14.s9 cell line displayed amplification of COPS3 (4 copies) and loss of EPHA5 (1 copy) situated on 17p11.2 and 4q13.1 respectively. These abnormalities were also observed in the culture-adapted H14 sub-line (H14.BJ1) suggesting that these abnormalities were present at derivation or occurred early in culture. The H14.BJ1 also showed amplification of NANOG (3 copies), MDM2 (4 copies), BIRC5 (3 copies) and BCL2L1 (4 copies) located on 12p13.31, 12q15, 17q25.3 and 20q11.21 respectively. The iPS cell lines MIFF1 and MIFF3 also showed increased copy number of COPS3 (4 copies) possibly suggesting an abnormality in the original transfected population of fibroblasts. MIFF3 also displayed amplification of BCL2L1 (3 copies).

Figure 31. A comparison of assay sensitivity between qRT-PCR and FISH in detecting amplification of 20q11.21
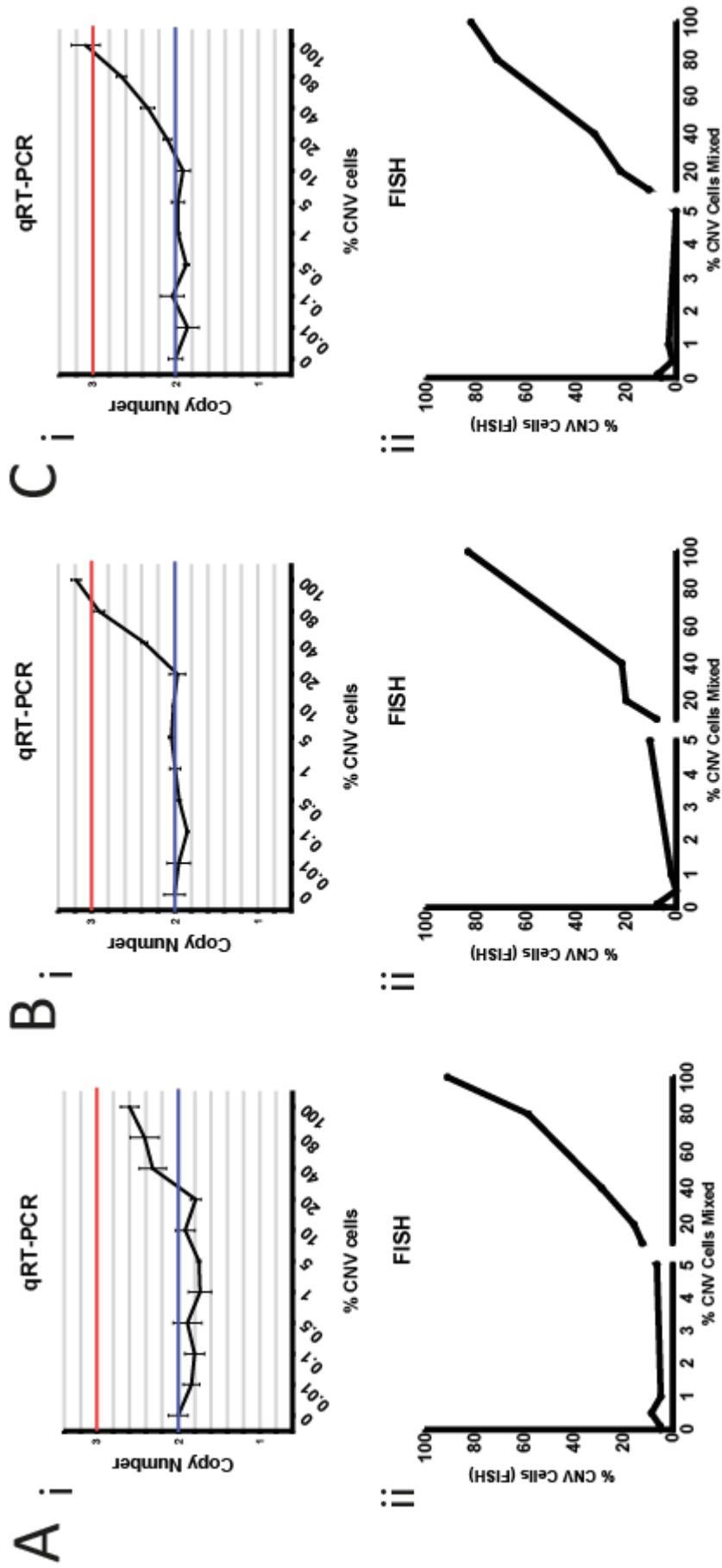
The detectable limit of qRT-PCR and FISH was determined using mixing experiments. A-C are biological replicates showing the sensitivity of (i) qRT-PCR and (ii) FISH. The level of 20q11.21 amplification was determined using primer/probe pairs designed to BCL2L1 and normalised to the reference gene EPHA5. To detect the amplification by FISH, a probe covering the BCL2L1 locus was used.

Figure 32. Screening human embryonic stem cells for common karyotypic changes.
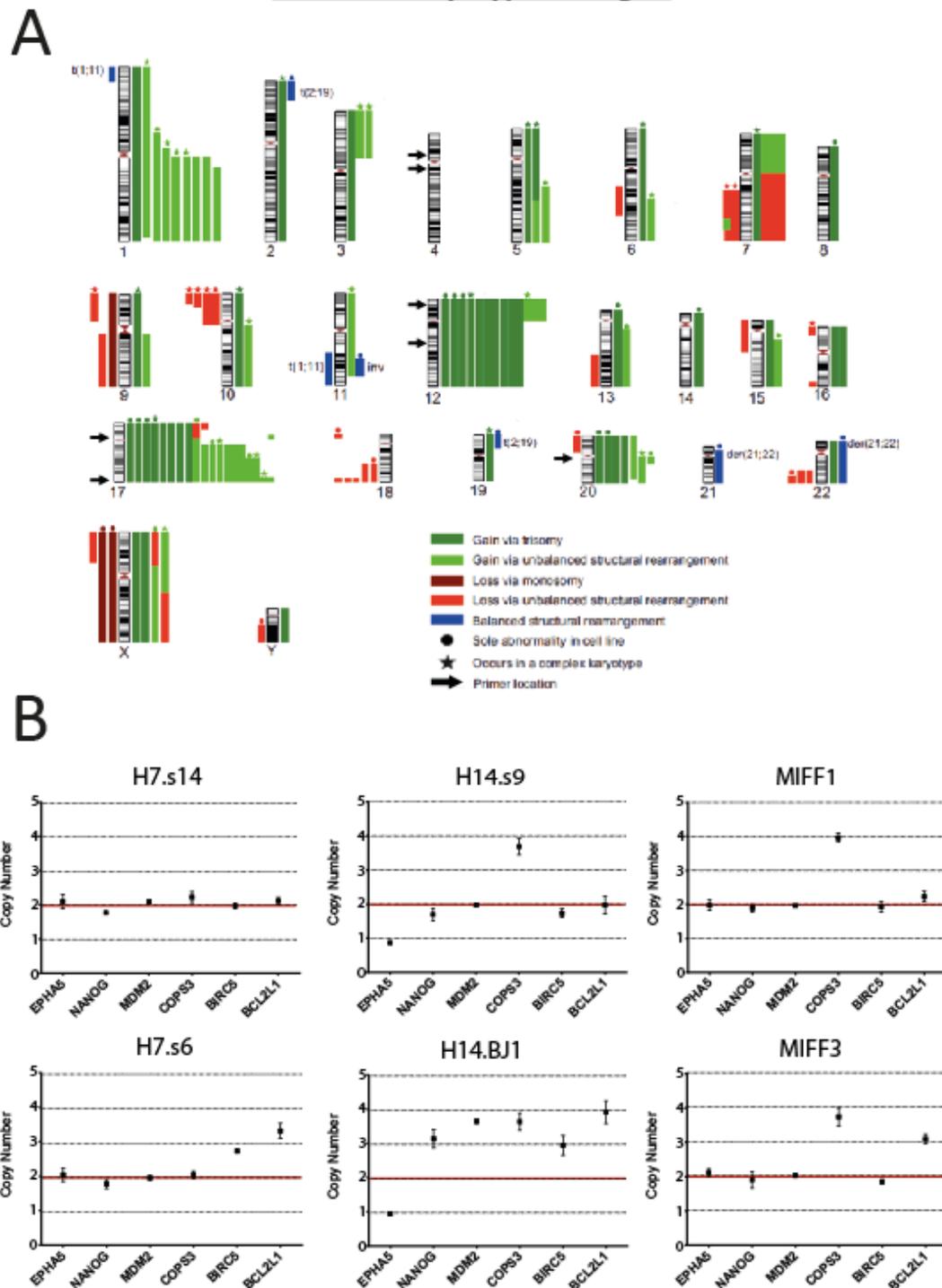
(A) An ideogram of the human chromosomes showing the reported karyotypic changes observed in human ES cells. The locations of primer/probe pairs designed to detect the common karyotypic changes are marked with black arrows.
(B) qRT-PCR karyotyping of six human ES cell lines using primer/probe pairs designed to commonly amplified regions of the human genome. The genes analysed are located on the following regions: EPHA5-4q, NANOG-12p, MDM2-12q, COPS3-17p, BIRC5-17q and BCL2L1-20q.

7.4. Discussion

The results show that qRT-PCR can be used as a method for the rapid and sensitive detection of karyotypic changes in human ES cells. There are a few advantages and disadvantages of using qRT-PCR or FISH in detecting aberrations in human ES cells. FISH is more sensitive than qPCR, with FISH being able to detect down to 5-10% in mosaic cultures and qRT-PCR down to 20%. However, using FISH to detect the 20q11.21 is not without problems, with the signal being within such close proximity they can sometimes present as a single signal making it only distinguishable by signal intensity. FISH would otherwise be more sensitive at detecting whole chromosome gains where distinct signals are separated. One advantage of using qRT-PCR is the rapid turn around of samples making the lower sensitivity not overly disadvantageous as steps can be taken to avoid compromising results/cell lines. The cell line could be thawed at an earlier passage or cloned to remove CNV cells from the culture. Cells in culture can be harvested, the genomic DNA extracted and qRT-PCR performed in the same day. FISH has a turn around of approximately two weeks, although urgent samples can be obtained within 3-5 days depending on the facilities case load and availability of relevant probes. This could potentially be disadvantageous, as we have shown that the 20q11.21 can rapidly overtake the culture. qRT-PCR is also much cheaper, the cost of running one sample for one region, for example BCL2L1, the qRT-PCR would need to run three primer-probe pairs, the reference genes EPHA5, RELL1 and the test region BCL2L1. Two reference samples would also need to be run in parallel to normalise data to control and check against a known CNV cell line sample. The cost would be ~£8 with the cost of extra samples ~£2. Using a full 384-well plate would cost ~£114 but could analyse 8 samples for up to 16 regions which is cheaper than analysing BCL2L1 copy number in just one sample using FISH, which would cost approximately £158.

The adaptability of qRT-PCR also offers advantages, as the assay can be designed to any region on the human genome. Thus, primer/probe pairs could be designed to a number of different regions to detect amplifications/deletions of multiple regions in a single assay. The caveat with both qRT-PCR and FISH is the small regions that are analysed, they do not cover the whole genome. However, due to the non-random karyotypic changes observed in human ES cells it may be possible to design a panel of primer/probe pairs that cover the majority of common changes. Specific genes could be used that are predicted to provide selective advantages, for example, NANOG is a strong candidate for the selection of chromosome 12 due its role in maintaining pluripotency in human ES cells. MDM2 is a negative regulator of the p53 pathway, it has been shown to bind p53 where it functions as a E3 ubiquitin ligase targeting p53 for degradation[146]. Increased MDM2 expression would therefore inhibit the apoptotic and cell cycle arrest functions of p53[147]. COPS3 amplification has been observed in osteosarcomas, the amplification of COPS3 coupled with mutations of p53 lead to metastasis and poor prognosis[148]. siRNA knock-down of COPS3 has been shown to reduce the proliferation and survival of lung cancer cells[149]. BIRC5 (Survivin) is a potent inhibitor of apoptosis and has been shown to be upregulated in a number of human cancers[150]. Increased BIRC5 expression is also observed in teratomas formed from human ES cells[151]. BIRC5 also has a role in mitosis, regulating the G1-S phase transition and if disrupted leads to polyploidy and caspase-mediated apoptosis[152].

The qRT-PCR offers advantages over other methods of detecting karyotypic changes. G-banding has the advantage of analysing the whole genome but lacks the resolution of other techniques not being able to detect <5Mb making the detection of small variants difficult (i.e. CNV 20q11.21). Both G-banding and FISH often require an experienced cytogeneticist as well as FISH being expensive and limiting in the number of regions that can be analysed. Another method of detecting changes has been proposed called 'e-karyotyping'[153], which offers

similar advantages as the qRT-PCR assay on genomic DNA. One caveat to this technique is that amplified regions do not always show increased levels of mRNA (Figure 10). For example, HES3 and H1 CNV cell lines displayed increased BCL-XL expression whereas in ESI-035 and Shef5 CNV cell lines the levels were similar to those observed in their respective control counter-parts.

Culture-adapted human ES cells rapidly out-compete normal cells in culture making early detection of abnormalities essential[73]. The accessibility, cost and turn around times of qRT-PCR for the detection of common chromosomal abnormalities make it an ideal method for routinely screening human ES cell lines. Multiple cell lines could be screened on a weekly/monthly basis to monitor human ES cell lines for abnormalities and can be applied to any stem cell laboratory world-wide.

8. General Discussion

8.1. The amplification of 20q11.21 provides a strong selective advantage in human ES cells

Our results show that human ES cell lines containing the 20q11.21 amplification display increased growth rates and rapidly out-compete normal diploid cells. This increased growth capacity can explain the high prevalence of 20q11.21 amplification in different human ES cell lines. The fact that this amplification is observed in a number of different human ES cell lines from various laboratories rules out the possibility that the 20q11.21 is a result of specific genetic backgrounds or culture systems. It is interesting that many human cancers also show amplification of 20q11.21 suggesting a common intrinsic mechanism behind 20q11.21 gain. This could be a result of increased proliferation and faster cell cycle times compared with somatic cells. However, human ES cells appear to be especially susceptible to 20q11.21 gain with two of the four cell lines (HES3 and H1) showing mosaic amplification of the 20q11.21 region at early and late passage (Supplementary Figure 1). For example, FISH analysis of the HES3 control cell line showed that 80% of cells were diploid but contained an underlying population (20%) of cells with 3 copies of the 20q11.21 locus. The later passage HES3 CNV cell line showed further amplification of the 20q11.21 locus with 62% of cells with 3 copies, 28% of cells with 4 copies and 10% of cells with 5 copies. This could highlight the tendency of human ES cells to acquire 20q11.21 gains or that once a cell line has acquired 20q11.21 amplification it is prone to further amplification. The four cell lines obtained in the study also displayed varying lengths of 20q11.21 amplification consistent with previous studies[66, 116]. Our results suggest that the length of 20q11.21 amplification could confer varying levels of selective advantage. For example, the HES3 CNV cell line contained the largest 20q11.21 amplicon showed the highest population growth rates whereas the Shef5 CNV contained the

shortest amplicon and showed reduced growth rates. This observation suggests that other genes along the 20q11.21 amplification could also be providing smaller selective advantages, which increase population growth or contain positive regulators of BCL2L1 increasing expression or stability of mRNA/protein.

The prevalence of 20q11.21 amplification paired with the strong selective advantage that it confers make the early detection of the amplification paramount so that appropriate steps can be taken to avoid compromising data or the transplant of potentially malignant cells. The frequent amplification of 20q in a number of cancers underlines this issue, particularly if cells are to be used for therapeutic applications or toxicology screens.

8.2. The increased growth rates of human ES cells containing the 20q11.21 amplification can be attributed to the anti-apoptotic effects of BCL-XL

The increased growth rates of human ES cell lines containing the 20q11.21 amplification can be attributed to BCL-XL. Over-expression of BCL-XL, the dominant isoform of BCL2L1 in undifferentiated human ES cells, increases the growth rates of cell lines that do not contain the 20q11.21 amplification to levels comparable with cell lines that show amplification (Figure 6). The increased growth rates were shown to be an effect of increased protection against apoptosis, rather than increased proliferation. This is consistent with BCL-XL being an anti-apoptotic gene. CNV-containing cells display reduced cell death following plating at low densities and survive to make successful cell divisions (Figure 12). Barbaric and colleagues[74] showed that certain bottle-necks exist in human ES cell culture, one of which is the survival of cells following re-plating. Our results show that the plating of control cells, even at high density is very in efficient and population growth rates are low (Figure 6). This is the bottle-neck that allows the CNV-containing cell lines to over-take the culture rapidly due to their

increased survival properties particularly within the first 24 hours of seeding (Figure 12). The

effects of the 20q11.21 CNV can be reduced either by the chemical inhibition (protein) or

knock-down of BCL-XL (mRNA), which show decreased cloning efficiencies (Figure 14) and

reduced growth rates (Figure 15) respectively. Although these results confirm that BCL-XL is

the main gene driving selection of the 20q11.21 amplification they are not practical methods

for removing CNV cells from human ES cultures. This being said, our results show that the

main bottle-neck during re-plating of human ES cells can be alleviated using ROCK inhibitor

when seeding cells (Figure 16). The presence of ROCK inhibitor for the first 24 hours of re-

plating drastically increases the growth rates of control cells similar CNV-containing cells.

In the case of 20q11.21 amplification, the gene driving selection and the mechanism behind

the growth advantage were less troublesome to investigate than other common chromosomal

changes. The relatively small amplicon provided only a handful of candidate genes driving

20q11.21 selection. The gene(s) driving the selection of larger chromosomal changes such as

chromosomes 1, 12 and 17 are more difficult to identify due to the number of possible

candidates. Identifying the mechanisms behind selection could narrow down the potential list

of candidates but selection will more than likely be provided by more than one gene, which

could reside on different regions along the genome.

8.3. Amplification of 20q11.21 alters the differentiation of human ES cells

The potential uses of human ES cells or their differentiated derivatives in therapeutic

applications rely on understanding the mechanisms behind stem cell behaviour. Two areas

that need to be addressed are, firstly, understanding the mechanisms behind lineage

specification when human ES cells differentiate, to successfully generate functional adult cell

types of different cell lineages. The second is understanding the mechanisms behind culture-

adaptation and the frequent gain of chromosomal material. It could be that certain genetic

changes are in fact irrelevant in therapeutic applications. However, the similarities in

karyotypic changes between ES cells, EC cells and primary tumours suggest that the changes

share a common neoplastic mechanism. These two points are also intertwined, as karyotypic

changes have been shown to alter differentiation patterns in human ES cells and could disrupt

the mechanisms behind differentiation[76].

Our results show that the CNV-containing cell line displays altered differentiation patterns

compared to control cells when induced to differentiate via the formation of EBs. This is

particularly evident when induced to differentiate towards mesoderm and endodermal

lineages (Figure 19). This altered pattern of differentiation was also observed in control cells

over-expressing BCL-XL suggesting that either differentiation is affected by the increased

survival of cells through BCL-XL or that BCL-XL is directly influencing differentiation. Our

results suggest that the latter hypothesis is true, as the EBs formed from CNV and BCL-XL over-

expressing cell lines under mesoderm and endoderm conditions are much larger than control

cells (Figure 18, Table 6). This is supported by results showing that BCL-XL over-expressing

cells successfully differentiate towards a definitive endoderm lineage using a monolayer

system (Figure 21). Interestingly, in this instance the CNV-containing cells do not closely

mirror the phenotype of BCL-XL over-expressing cells. The CNV-containing cells show

increased cell numbers compared with control cells but do not display increased endoderm

differentiation determined by SOX17 positive cells. This observation suggests that there could

be other genes on 20q11.21 that can affect the differentiation of human ES cells.

8.4. Human ES cells are prone to erroneous mitoses

The mechanisms behind the frequent and non-random karyotypic changes that occur in human ES cells are poorly understood. Evidence suggests that errors in chromosomal segregation at anaphase could be behind some of the abnormalities observed, particularly in cases where the gain of whole chromosomes or isochromosomes is observed[67]. Our results show that errors in chromosomal segregation at anaphase are a frequent event in human ES cells (Figures 23 and 24). Using live-cell time-lapse imaging the erroneous mitotic events could be observed in real time. The control cell line displayed a high proportion (29%) of abnormal chromosomal segregation events showing that human ES cells are particularly susceptible to erroneous mitoses. Whether this is similar in all human ES cell lines remains to be addressed but would explain the frequency of which karyotypic abnormalities are observed. It is still unknown whether certain chromosomes are particularly susceptible to lag or bridge, which could explain the non-random nature of cytogenetic changes. Interestingy, CNV-containing cells displayed reduced variability in metaphase to anaphase transition times and fewer events of abnormal chromosome segregation (14%) than control cells. This potentially highlights that gene(s) within the 20q11.21 amplification somehow regulate the later stages of mitosis. TPX2 is a candidate as it has been shown to… Although this suggests that CNV-containing cells are less prone to erroneous mitoses, they were also more likely to survive these abnormal events, with 0 out of 10 dying following abnormal division compared to 4 out of 21 in control cells. Whether this observation is limited by the length of time-lapse movies will need to be addressed, it may be that more of the control and CNV-containing cells die later in the cell cycle or that CNV-containing cells are protected against abnormal divisions. Longer time-lapse movies would have to be generated to determine this hypothesis.

AZD1152, an inhibitor of AURORA KINASE B was used to investigate the survival of control and

CNV cells following abnormal mitosis. Using live-cell time-lapse imaging, we showed that

AZD1152 induces polyploidy in human ES cells, generating cells with 4N and 8N DNA content

(Figure 25). We also monitored the levels of apoptosis following treatment of AZD1152 in

control, CNV and BCL-XL over-expressing cells. Results showed that CNV and BCL-XL were less

susceptible to apoptosis in the presence of AZD1152 suggesting that they are more likely to

survive abnormal divisions (Figure 26). To investigate this further we propose isolating

different populations of cells based on DNA content using live cell staining and sorting for

Hoechst 33342 and generating time-lapse movies to monitor the behaviour of abnormal cells.

The caveat of using AZD1152 is that it promotes polyploidy, an event that was not observed in

our study of mitotic divisions. Further investigation using the spindle poison nocodazole would

provide a more accurate representation of the events occurring in human ES cell cultures.

[154]. Nocodazole arrests cells at G2/M, occasionally cells can escape this block and initiate

anaphase and cytokinesis in an event termed 'mitotic slippage' [155]. This event can result in

abnormal chromosome segregation and promote aneuploidy more closely mirroring events

observed in Chapter 6.

8.5. <u>Detection of common karyotypic changes by qRT-PCR</u>

The amplification of 20q11.21 in human ES cells provides cells with an increased growth

advantage. We have shown that CNV-containing cells rapidly out-compete their control

counterparts within ten passages (Figure 7). This makes the early detection of 20q11.21

amplification of paramount importance. We have devised a quick, cost-effective and sensitive

method for the detection of 20q11.21 using qRT-PCR. The assay can also be adapted to detect

other commonly affected regions on other chromosomes (for instance chromosomes 1, 12 and

17). The qRT-PCR method offers advantages over other methods including, time-scale, cost, availability and adaptability. Results can be obtained on the same day as the sample is harvested making identification of abnormalities almost immediate. For the price of analysing one sample using FISH, qRT-PCR can analyse multiple samples in parallel whilst looking at a number of different chromosomal regions. The method also requires minimal equipment/reagents and in transferrable to any lab worldwide. Finally the assay is adaptable, primer/probe sets are simple to design and can be targeted to almost every region on the human genome.

9. <u>Concluding Remarks</u>

We have shown that the frequent amplification of 20q11.21 in human ES cells is a result of their increased growth rates conferred by BCL-XL. Cells containing the 20q11.21 CNV show increased protection against apoptosis and rapidly out-compete diploid cells. The CNV may also affect the differentiation patterns of human ES cells and make them more susceptible to acquiring further genomic changes. This data together with the presence of 20q11.21 amplification in EC cells and primary tumours warrant the monitoring for 20q11.21 amplification in human ES cell cultures. We have devised a simple and quick method that can be applied in laboratories worldwide to detect the 20q11.21 amplification. This method can also be used to detect other genomic changes in human ES cells.

References

1.   Thomson, J.A., et al., *Embryonic stem cell lines derived from human blastocysts.* Science, 1998. **282**(5391): p. 1145-7.
2.   Murry, C.E. and G. Keller, *Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development.* Cell, 2008. **132**(4): p. 661-80.
3.   Stevens, L.C. and C.C. Little, *Spontaneous Testicular Teratomas in an Inbred Strain of Mice.* Proc Natl Acad Sci U S A, 1954. **40**(11): p. 1080-7.
4.   Kleinsmith, L.J. and G.B. Pierce, Jr., *Multipotentiality of Single Embryonal Carcinoma Cells.* Cancer Res, 1964. **24**: p. 1544-51.
5.   Stevens, L.C., *Embryology of testicular teratomas in strain 129 mice.* J Natl Cancer Inst, 1959. **23**: p. 1249-95.
6.   Stevens, L.C., *Embryonic potency of embryoid bodies derived from a transplantable testicular teratoma of the mouse.* Dev Biol, 1960. **2**: p. 285-97.
7.   Finch, B.W. and B. Ephrussi, *RETENTION OF MULTIPLE DEVELOPMENTAL POTENTIALITIES BY CELLS OF A MOUSE TESTICULAR TERATOCARCINOMA DURING PROLONGED CULTURE in vitro AND THEIR EXTINCTION UPON HYBRIDIZATION WITH CELLS OF PERMANENT LINES.* Proc Natl Acad Sci U S A, 1967. **57**(3): p. 615-21.
8.   Gardner, R.L., *Mouse chimeras obtained by the injection of cells into the blastocyst.* Nature, 1968. **220**(5167): p. 596-7.
9.   Kahan, B.W. and B. Ephrussi, *Developmental potentialities of clonal in vitro cultures of mouse testicular teratoma.* J Natl Cancer Inst, 1970. **44**(5): p. 1015-36.
10.   Artzt, K., et al., *Surface antigens common to mouse cleavage embryos and primitive teratocarcinoma cells in culture.* Proc Natl Acad Sci U S A, 1973. **70**(10): p. 2988-92.
11.   Brinster, R.L., *The effect of cells transferred into the mouse blastocyst on subsequent development.* J Exp Med, 1974. **140**(4): p. 1049-56.
12.   Martin, G.R. and M.J. Evans, *Differentiation of clonal lines of teratocarcinoma cells: formation of embryoid bodies in vitro.* Proc Natl Acad Sci U S A, 1975. **72**(4): p. 1441-5.
13.   Strickland, S. and V. Mahdavi, *The induction of differentiation in teratocarcinoma stem cells by retinoic acid.* Cell, 1978. **15**(2): p. 393-403.
14.   Solter, D. and B.B. Knowles, *Monoclonal antibody defining a stage-specific mouse embryonic antigen (SSEA-1).* Proc Natl Acad Sci U S A, 1978. **75**(11): p. 5565-9.
15.   Evans, M.J. and M.H. Kaufman, *Establishment in culture of pluripotential cells from mouse embryos.* Nature, 1981. **292**(5819): p. 154-6.
16.   Martin, G.R., *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**(12): p. 7634-8.
17.   Bradley, A., et al., *Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines.* Nature, 1984. **309**(5965): p. 255-6.
18.   Evans, M.J., et al., *The ability of EK cells to form chimeras after selection of clones in G418 and some observations on the integration of retroviral vector proviral DNA into EK cells.* Cold Spring Harb Symp Quant Biol, 1985. **50**: p. 685-9.
19.   Doetschman, T., et al., *Targetted correction of a mutant HPRT gene in mouse embryonic stem cells.* Nature, 1987. **330**(6148): p. 576-8.
20.   Thomas, K.R. and M.R. Capecchi, *Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells.* Cell, 1987. **51**(3): p. 503-12.
21.   Andrews, P.W., et al., *Comparative analysis of cell surface antigens expressed by cell lines derived from human germ cell tumours.* Int J Cancer, 1996. **66**(6): p. 806-16.
22.   Fogh, J. and G. Trempe, *New Human Tumor Cell Lines*, in *Human Tumor Cells in Vitro*, J. Fogh, Editor. 1975, Springer US. p. 115-159.

23. Jewett, M.A., *Testis carcinoma: transplantation into nude mice.* Natl Cancer Inst Monogr, 1978(49): p. 65-6.

24. Andrews, P.W., et al., *A comparative study of eight cell lines derived from human testicular teratocarcinoma.* Int J Cancer, 1980. **26**(3): p. 269-80.

25. Andrews, P.W., et al., *Cell-surface antigens of a clonal human embryonal carcinoma cell line: morphological and antigenic differentiation in culture.* Int J Cancer, 1982. **29**(5): p. 523-31.

26. Andrews, P.W., *Retinoic acid induces neuronal differentiation of a cloned human embryonal carcinoma cell line in vitro.* Dev Biol, 1984. **103**(2): p. 285-93.

27. Andrews, P.W., et al., *Pluripotent embryonal carcinoma clones derived from the human teratocarcinoma cell line Tera-2. Differentiation in vivo and in vitro.* Lab Invest, 1984. **50**(2): p. 147-62.

28. Andrews, P., *The Characteristics of Cell Lines Derived from Human Germ Cell Tumors*, in *The Human Teratomas*, I. Damjanov, B. Knowles, and D. Solter, Editors. 1983, Humana Press. p. 285-311.

29. Graves, K.H. and R.W. Moreadith, *Derivation and characterization of putative pluripotential embryonic stem cells from preimplantation rabbit embryos.* Mol Reprod Dev, 1993. **36**(4): p. 424-33.

30. Notarianni, E., et al., *Maintenance and differentiation in culture of pluripotential embryonic cell lines from pig blastocysts.* J Reprod Fertil Suppl, 1990. **41**: p. 51-6.

31. Notarianni, E., et al., *Derivation of pluripotent, embryonic cell lines from the pig and sheep.* J Reprod Fertil Suppl, 1991. **43**: p. 255-60.

32. Thomson, J.A., et al., *Isolation of a primate embryonic stem cell line.* Proc Natl Acad Sci U S A, 1995. **92**(17): p. 7844-8.

33. Williams, R.L., et al., *Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells.* Nature, 1988. **336**(6200): p. 684-7.

34. Smith, A.G., et al., *Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides.* Nature, 1988. **336**(6200): p. 688-90.

35. Brons, I.G., et al., *Derivation of pluripotent epiblast stem cells from mammalian embryos.* Nature, 2007. **448**(7150): p. 191-5.

36. Briggs, R. and T.J. King, *Transplantation of Living Nuclei From Blastula Cells into Enucleated Frogs' Eggs.* Proc Natl Acad Sci U S A, 1952. **38**(5): p. 455-63.

37. King, T.J. and R. Briggs, *Changes in the Nuclei of Differentiating Gastrula Cells, as Demonstrated by Nuclear Transplantation.* Proc Natl Acad Sci U S A, 1955. **41**(5): p. 321-5.

38. Miller, R.A. and F.H. Ruddle, *Pluripotent teratocarcinoma-thymus somatic cell hybrids.* Cell, 1976. **9**(1): p. 45-55.

39. Miller, R.A. and F.H. Ruddle, *Teratocarcinoma X friend erythroleukemia cell hybrids resemble their pluripotent embryonal carcinoma parent.* Dev Biol, 1977. **56**(1): p. 157-73.

40. Davis, R.L., H. Weintraub, and A.B. Lassar, *Expression of a single transfected cDNA converts fibroblasts to myoblasts.* Cell, 1987. **51**(6): p. 987-1000.

41. Xie, H., et al., *Stepwise reprogramming of B cells into macrophages.* Cell, 2004. **117**(5): p. 663-76.

42. Zhou, Q., et al., *In vivo reprogramming of adult pancreatic exocrine cells to beta-cells.* Nature, 2008. **455**(7213): p. 627-32.

43. Vierbuchen, T., et al., *Direct conversion of fibroblasts to functional neurons by defined factors.* Nature, 2010. **463**(7284): p. 1035-41.

44. Takahashi, K. and S. Yamanaka, *Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors.* Cell, 2006. **126**(4): p. 663-76.

45. Takahashi, K., et al., *Induction of pluripotent stem cells from adult human fibroblasts by defined factors.* Cell, 2007. **131**(5): p. 861-72.

46. Yu, J., et al., *Induced pluripotent stem cell lines derived from human somatic cells.* Science, 2007. **318**(5858): p. 1917-20.

47. Warren, L., et al., *Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA.* Cell Stem Cell, 2010. **7**(5): p. 618-30.

48. Boiani, M. and H.R. Scholer, *Regulatory networks in embryo-derived pluripotent stem cells.* Nat Rev Mol Cell Biol, 2005. **6**(11): p. 872-84.

49. Stewart, C.L., et al., *Blastocyst implantation depends on maternal expression of leukaemia inhibitory factor.* Nature, 1992. **359**(6390): p. 76-9.

50. Hu, W., et al., *p53 regulates maternal reproduction through LIF.* Nature, 2007. **450**(7170): p. 721-4.

51. Kinoshita, K., et al., *GABPalpha regulates Oct-3/4 expression in mouse embryonic stem cells.* Biochem Biophys Res Commun, 2007. **353**(3): p. 686-91.

52. Nakatake, Y., et al., *Klf4 cooperates with Oct3/4 and Sox2 to activate the Lefty1 core promoter in embryonic stem cells.* Mol Cell Biol, 2006. **26**(20): p. 7772-82.

53. Niwa, H., et al., *Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3.* Genes Dev, 1998. **12**(13): p. 2048-60.

54. Takahashi-Tezuka, M., et al., *Gab1 acts as an adapter molecule linking the cytokine receptor gp130 to ERK mitogen-activated protein kinase.* Mol Cell Biol, 1998. **18**(7): p. 4109-17.

55. Brazil, D.P., Z.Z. Yang, and B.A. Hemmings, *Advances in protein kinase B signalling: AKTion on multiple fronts.* Trends Biochem Sci, 2004. **29**(5): p. 233-42.

56. Paling, N.R., et al., *Regulation of embryonic stem cell self-renewal by phosphoinositide 3-kinase-dependent signaling.* J Biol Chem, 2004. **279**(46): p. 48063-70.

57. Avery, S., K. Inniss, and H. Moore, *The regulation of self-renewal in human embryonic stem cells.* Stem Cells Dev, 2006. **15**(5): p. 729-40.

58. Shi, Y. and J. Massague, *Mechanisms of TGF-beta signaling from cell membrane to the nucleus.* Cell, 2003. **113**(6): p. 685-700.

59. James, D., et al., *TGFbeta/activin/nodal signaling is necessary for the maintenance of pluripotency in human embryonic stem cells.* Development, 2005. **132**(6): p. 1273-82.

60. Wang, G., et al., *Noggin and bFGF cooperate to maintain the pluripotency of human embryonic stem cells in the absence of feeder layers.* Biochem Biophys Res Commun, 2005. **330**(3): p. 934-42.

61. Draper, J.S., et al., *Culture and characterization of human embryonic stem cells.* Stem Cells Dev, 2004. **13**(4): p. 325-36.

62. Draper, J.S., et al., *Recurrent gain of chromosomes 17q and 12 in cultured human embryonic stem cells.* Nat Biotechnol, 2004. **22**(1): p. 53-4.

63. Enver, T., et al., *Cellular differentiation hierarchies in normal and culture-adapted human embryonic stem cells.* Hum Mol Genet, 2005. **14**(21): p. 3129-40.

64. Imreh, M.P., et al., *In vitro culture conditions favoring selection of chromosomal abnormalities in human ES cells.* J Cell Biochem, 2006. **99**(2): p. 508-16.

65. Baker, D.E., et al., *Adaptation to culture of human embryonic stem cells and oncogenesis in vivo.* Nat Biotechnol, 2007. **25**(2): p. 207-15.

66. International Stem Cell, I., et al., *Screening ethnically diverse human embryonic stem cells identifies a chromosome 20 minimal amplicon conferring growth advantage.* Nat Biotechnol, 2011. **29**(12): p. 1132-44.

67. Na, J., et al., *Aneuploidy in pluripotent stem cells and implications for cancerous transformation.* Protein Cell, 2014.

68. Desmarais, J.A., et al., *Human embryonic stem cells fail to activate CHK1 and commit to apoptosis in response to DNA replication stress.* Stem Cells, 2012. **30**(7): p. 1385-93.

69. Ford, D., et al., *Risks of cancer in BRCA1-mutation carriers. Breast Cancer Linkage Consortium.* Lancet, 1994. **343**(8899): p. 692-5.

70. Miki, Y., et al., *A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1.* Science, 1994. **266**(5182): p. 66-71.

71. Liu, X., et al., *Trisomy eight in ES cells is a common potential problem in gene targeting and interferes with germ line transmission.* Dev Dyn, 1997. **209**(1): p. 85-91.

72. Harrison, N.J., D. Baker, and P.W. Andrews, *Culture adaptation of embryonic stem cells echoes germ cell malignancy.* Int J Androl, 2007. **30**(4): p. 275-81; discussion 281.

73. Olariu, V., et al., *Modeling the evolution of culture-adapted human embryonic stem cells.* Stem Cell Res, 2010. **4**(1): p. 50-6.

74. Barbaric, I., et al., *Time-Lapse Analysis of Human Embryonic Stem Cells Reveals Multiple Bottlenecks Restricting Colony Formation and Their Relief upon Culture Adaptation.* Stem Cell Reports, 2014. **3**(1): p. 142-55.

75. Narsinh, K.H., et al., *Single cell transcriptional profiling reveals heterogeneity of human induced pluripotent stem cells.* J Clin Invest, 2011. **121**(3): p. 1217-21.

76. Fazeli, A., et al., *Altered patterns of differentiation in karyotypically abnormal human embryonic stem cells.* Int J Dev Biol, 2011. **55**(2): p. 175-80.

77. Livak, K.J. and T.D. Schmittgen, *Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method.* Methods, 2001. **25**(4): p. 402-8.

78. Chambers, I., et al., *Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells.* Cell, 2003. **113**(5): p. 643-55.

79. Liew, C.G., et al., *Transient and stable transgene expression in human embryonic stem cells.* Stem Cells, 2007. **25**(6): p. 1521-8.

80. Avery, S., et al., *The role of SMAD4 in human embryonic stem cell self-renewal and stem cell fate.* Stem Cells, 2010. **28**(5): p. 863-73.

81. Lefort, N., et al., *Human embryonic stem cells reveal recurrent genomic instability at 20q11.21.* Nat Biotechnol, 2008. **26**(12): p. 1364-6.

82. Spits, C., et al., *Recurrent chromosomal abnormalities in human embryonic stem cells.* Nat Biotechnol, 2008. **26**(12): p. 1361-3.

83. Wu, H., et al., *Copy number variant analysis of human embryonic stem cells.* Stem Cells, 2008. **26**(6): p. 1484-9.

84. Narva, E., et al., *High-resolution DNA analysis of human embryonic stem cell lines reveals culture-induced copy number changes and loss of heterozygosity.* Nat Biotechnol, 2010. **28**(4): p. 371-7.

85. Anzick, S.L., et al., *AIB1, a steroid receptor coactivator amplified in breast and ovarian cancer.* Science, 1997. **277**(5328): p. 965-8.

86. Tabach, Y., et al., *Amplification of the 20q chromosomal arm occurs early in tumorigenic transformation and may initiate cancer.* PLoS One, 2011. **6**(1): p. e14632.

87. Watanabe, T., et al., *Differentially regulated genes as putative targets of amplifications at 20q in ovarian cancers.* Jpn J Cancer Res, 2002. **93**(10): p. 1114-22.

88. Engler, D.A., et al., *Genome wide DNA copy number analysis of serous type ovarian carcinomas identifies genetic markers predictive of clinical outcome.* PLoS One, 2012. **7**(2): p. e30996.

89. Sillars-Hardebol, A.H., et al., *TPX2 and AURKA promote 20q amplicon-driven colorectal adenoma to carcinoma progression.* Gut, 2012. **61**(11): p. 1568-75.

90. Qian, Y. and X. Chen, *ID1, inhibitor of differentiation/DNA binding, is an effector of the p53-dependent DNA damage response pathway.* J Biol Chem, 2008. **283**(33): p. 22410-6.

91. Ling, M.T., et al., *Id-1 expression promotes cell survival through activation of NF-kappaB signalling pathway in prostate cancer cells.* Oncogene, 2003. **22**(29): p. 4498-508.

92. Lee, T.K., et al., *Over-expression of Id-1 induces cell proliferation in hepatocellular carcinoma through inactivation of p16INK4a/RB pathway.* Carcinogenesis, 2003. **24**(11): p. 1729-36.

93. Smith, L.T., et al., *20q11.1 amplification in giant-cell tumor of bone: Array CGH, FISH, and association with outcome.* Genes Chromosomes Cancer, 2006. **45**(10): p. 957-66.

94. Beroukhim, R., et al., *The landscape of somatic copy-number alteration across human cancers.* Nature, 2010. **463**(7283): p. 899-905.

95. Sillars-Hardebol, A.H., et al., *Identification of key genes for carcinogenic pathways associated with colorectal adenoma-to-carcinoma progression.* Tumour Biol, 2010. **31**(2): p. 89-96.

96. Sillars-Hardebol, A.H., et al., *BCL2L1 has a functional role in colorectal cancer and its protein expression is associated with chromosome 20q gain.* J Pathol, 2012. **226**(3): p. 442-50.

97. Reubinoff, B.E., et al., *Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro.* Nat Biotechnol, 2000. **18**(4): p. 399-404.

98. Crook, J.M., et al., *The Generation of Six Clinical-Grade Human Embryonic Stem Cell Lines.* Cell Stem Cell, 2007. **1**(5): p. 490-494.

99. Aflatoonian, B., et al., *Generation of Sheffield (Shef) human embryonic stem cell lines using a microdrop culture system.* In Vitro Cell Dev Biol Anim, 2010. **46**(3-4): p. 236-41.

100. Schmitt, E., M. Beauchemin, and R. Bertrand, *Nuclear colocalization and interaction between bcl-xL and cdk1(cdc2) during G2/M cell-cycle checkpoint.* Oncogene, 2007. **26**(40): p. 5851-65.

101. Becker, K.A., et al., *Self-renewal of human embryonic stem cells is supported by a shortened G1 cell cycle phase.* J Cell Physiol, 2006. **209**(3): p. 883-93.

102. Korn, W.M., et al., *Chromosome arm 20q gains and other genomic alterations in colorectal cancer metastatic to liver, as analyzed by comparative genomic hybridization and fluorescence in situ hybridization.* Genes Chromosomes Cancer, 1999. **25**(2): p. 82-90.

103. Fernandez, Y., et al., *Inhibition of apoptosis in human breast cancer cells: role in tumor progression to the metastatic state.* Int J Cancer, 2002. **101**(4): p. 317-26.

104. Taylor, R.C., S.P. Cullen, and S.J. Martin, *Apoptosis: controlled demolition at the cellular level.* Nat Rev Mol Cell Biol, 2008. **9**(3): p. 231-41.

105. Tait, S.W. and D.R. Green, *Mitochondria and cell death: outer membrane permeabilization and beyond.* Nat Rev Mol Cell Biol, 2010. **11**(9): p. 621-32.

106. Youle, R.J. and A. Strasser, *The BCL-2 protein family: opposing activities that mediate cell death.* Nat Rev Mol Cell Biol, 2008. **9**(1): p. 47-59.

107. Volkmann, N., et al., *The rheostat in the membrane: BCL-2 family proteins and apoptosis.* Cell Death Differ, 2014. **21**(2): p. 206-15.

108. Edlich, F., et al., *Bcl-x(L) retrotranslocates Bax from the mitochondria into the cytosol.* Cell, 2011. **145**(1): p. 104-16.

109. Minn, A.J., L.H. Boise, and C.B. Thompson, *Bcl-x(S) anatagonizes the protective effects of Bcl-x(L).* J Biol Chem, 1996. **271**(11): p. 6306-12.

110. Michels, J., et al., *Functions of BCL-X L at the Interface between Cell Death and Metabolism.* Int J Cell Biol, 2013. **2013**: p. 705294.

111. Kumar, B., et al., *EGFR, p16, HPV Titer, Bcl-xL and p53, sex, and smoking as indicators of response to therapy and survival in oropharyngeal cancer.* J Clin Oncol, 2008. **26**(19): p. 3128-37.

112. Kumar, R., et al., *Overexpression of HER2 modulates bcl-2, bcl-XL, and tamoxifen-induced apoptosis in human MCF-7 breast cancer cells.* Clin Cancer Res, 1996. **2**(7): p. 1215-9.

113. Lebedeva, I., et al., *Bcl-xL in prostate cancer cells: effects of overexpression and down-regulation on chemosensitivity.* Cancer Res, 2000. **60**(21): p. 6052-60.

114. Nagane, M., et al., *Drug resistance of human glioblastoma cells conferred by a tumor-specific mutant epidermal growth factor receptor through modulation of Bcl-XL and caspase-3-like proteases.* Proc Natl Acad Sci U S A, 1998. **95**(10): p. 5724-9.

115. Parrizas, M. and D. LeRoith, *Insulin-like growth factor-1 inhibition of apoptosis is associated with increased expression of the bcl-xL gene product.* Endocrinology, 1997. **138**(3): p. 1355-8.

116. Nguyen, H.T., et al., *Gain of 20q11.21 in human embryonic stem cells improves cell survival by increased expression of Bcl-xL.* Mol Hum Reprod, 2014. **20**(2): p. 168-77.

117. Ohgushi, M., et al., *Molecular pathway and cell state responsible for dissociation-induced apoptosis in human pluripotent stem cells.* Cell Stem Cell, 2010. **7**(2): p. 225-39.

118. Watanabe, K., et al., *A ROCK inhibitor permits survival of dissociated human embryonic stem cells.* Nat Biotechnol, 2007. **25**(6): p. 681-6.

119. Enver, T., et al., *Stem cell states, fates, and the rules of attraction.* Cell Stem Cell, 2009. **4**(5): p. 387-97.

120. Canham, M.A., et al., *Functional heterogeneity of embryonic stem cells revealed through translational amplification of an early endodermal transcript.* PLoS Biol, 2010. **8**(5): p. e1000379.

121. Mfopou, J.K., et al., *Efficient definitive endoderm induction from mouse embryonic stem cell adherent cultures: a rapid screening model for differentiation studies.* Stem Cell Res, 2014. **12**(1): p. 166-77.

122. Task, K., M. Jaramillo, and I. Banerjee, *Population based model of human embryonic stem cell (hESC) differentiation during endoderm induction.* PLoS One, 2012. **7**(3): p. e32975.

123. Neely, M.D., et al., *DMH1, a highly selective small molecule BMP inhibitor promotes neurogenesis of hiPSCs: comparison of PAX6 and SOX1 expression during neural induction.* ACS Chem Neurosci, 2012. **3**(6): p. 482-91.

124. Mahmood, A., et al., *Enhanced differentiation of human embryonic stem cells to mesenchymal progenitors by inhibition of TGF-beta/activin/nodal signaling using SB-431542.* J Bone Miner Res, 2010. **25**(6): p. 1216-33.

125. Peters, J.M., *The anaphase-promoting complex: proteolysis in mitosis and beyond.* Mol Cell, 2002. **9**(5): p. 931-43.

126. Li, X. and R.B. Nicklas, *Mitotic forces control a cell-cycle checkpoint.* Nature, 1995. **373**(6515): p. 630-2.

127. Cleveland, D.W., Y. Mao, and K.F. Sullivan, *Centromeres and kinetochores: from epigenetics to mitotic checkpoint signaling.* Cell, 2003. **112**(4): p. 407-21.

128. Holland, A.J. and D.W. Cleveland, *Boveri revisited: chromosomal instability, aneuploidy and tumorigenesis.* Nat Rev Mol Cell Biol, 2009. **10**(7): p. 478-87.

129. Duijf, P.H. and R. Benezra, *The cancer biology of whole-chromosome instability.* Oncogene, 2013. **32**(40): p. 4727-36.

130. Mantel, C., et al., *Checkpoint-apoptosis uncoupling in human and mouse embryonic stem cells: a source of karyotpic instability.* Blood, 2007. **109**(10): p. 4518-27.

131. Andrews, P.D., et al., *Mitotic mechanics: the auroras come into view.* Curr Opin Cell Biol, 2003. **15**(6): p. 672-83.

132. Carmena, M. and W.C. Earnshaw, *The cellular geography of aurora kinases.* Nat Rev Mol Cell Biol, 2003. **4**(11): p. 842-54.

133. Cowley, D.O., et al., *Aurora-A kinase is essential for bipolar spindle formation and early development.* Mol Cell Biol, 2009. **29**(4): p. 1059-71.

134. Carmena, M., et al., *The chromosomal passenger complex (CPC): from easy rider to the godfather of mitosis.* Nat Rev Mol Cell Biol, 2012. **13**(12): p. 789-803.

135. Balboula, A.Z. and K. Schindler, *Selective disruption of aurora C kinase reveals distinct functions from aurora B kinase during meiosis in mouse oocytes.* PLoS Genet, 2014. **10**(2): p. e1004194.

136. Azzariti, A., et al., *Aurora B kinase inhibitor AZD1152: determinants of action and ability to enhance chemotherapeutics effectiveness in pancreatic and colon cancer.* Br J Cancer, 2011. **104**(5): p. 769-80.

137. Yang, J., et al., *AZD1152, a novel and selective aurora B kinase inhibitor, induces growth arrest, apoptosis, and sensitization for tubulin depolymerizing agent or topoisomerase II inhibitor in human acute leukemia cells in vitro and in vivo.* Blood, 2007. **110**(6): p. 2034-40.

138. Dumitru, R., et al., *Human embryonic stem cells have constitutively active Bax at the Golgi and are primed to undergo rapid apoptosis.* Mol Cell, 2012. **46**(5): p. 573-83.

139. Minn, A.J., L.H. Boise, and C.B. Thompson, *Expression of Bcl-xL and loss of p53 can cooperate to overcome a cell cycle checkpoint induced by mitotic spindle damage.* Genes Dev, 1996. **10**(20): p. 2621-31.

140. Meraldi, P., V.M. Draviam, and P.K. Sorger, *Timing and checkpoints in the regulation of mitotic progression.* Dev Cell, 2004. **7**(1): p. 45-60.

141. Mikhailov, A., R.W. Cole, and C.L. Rieder, *DNA damage during mitosis in human cells delays the metaphase/anaphase transition via the spindle-assembly checkpoint.* Curr Biol, 2002. **12**(21): p. 1797-806.

142. Harrison, N.J., et al., *CD30 expression reveals that culture adaptation of human embryonic stem cells can occur through differing routes.* Stem Cells, 2009. **27**(5): p. 1057-65.

143. Kearney, L., *Molecular cytogenetics.* Best Pract Res Clin Haematol, 2001. **14**(3): p. 645-69.

144. D'Haene, B., J. Vandesompele, and J. Hellemans, *Accurate and objective copy number profiling using real-time quantitative PCR.* Methods, 2010. **50**(4): p. 262-70.

145. D'Hulst, C., et al., *Fast Quantitative Real-Time PCR-Based Screening for Common Chromosomal Aneuploidies in Mouse Embryonic Stem Cells.* Stem Cell Reports, 2013. **1**(4): p. 350-9.

146. Haupt, Y., et al., *Mdm2 promotes the rapid degradation of p53.* Nature, 1997. **387**(6630): p. 296-9.

147. Deb, S.P., *Cell cycle regulatory functions of the human oncoprotein MDM2.* Mol Cancer Res, 2003. **1**(14): p. 1009-16.

148. Yan, T., et al., *COPS3 amplification and clinical outcome in osteosarcoma.* Cancer, 2007. **109**(9): p. 1870-6.

149. Wang, X.M., et al., *Silencing of the COPS3 gene by siRNA reduces proliferation of lung cancer cells most likely via induction of cell cycle arrest and apoptosis.* Asian Pac J Cancer Prev, 2012. **13**(3): p. 1043-8.

150. Fukuda, S. and L.M. Pelus, *Survivin, a cancer target with an emerging role in normal adult tissues.* Mol Cancer Ther, 2006. **5**(5): p. 1087-98.

151. Blum, B., et al., *The anti-apoptotic gene survivin contributes to teratoma formation by human embryonic stem cells.* Nat Biotechnol, 2009. **27**(3): p. 281-7.

152. Li, F., et al., *Pleiotropic cell-division defects and apoptosis induced by interference with survivin function.* Nat Cell Biol, 1999. **1**(8): p. 461-6.

153. Ben-David, U., Y. Mayshar, and N. Benvenisty, *Virtual karyotyping of pluripotent stem cells on the basis of their global gene expression profiles.* Nat Protoc, 2013. **8**(5): p. 989-97.

154. Elhajouji, A., M. Cunha, and M. Kirsch-Volders, *Spindle poisons can induce polyploidy by mitotic slippage and micronucleate mononucleates in the cytokinesis-block assay.* Mutagenesis, 1998. **13**(2): p. 193-8.

155. Mena, A.L., E.W. Lam, and S. Chatterjee, *Sustained spindle-assembly checkpoint response requires de novo transcription and translation of cyclin B1.* PLoS One, 2010. **5**(9).

Supplementary Figure 1. Karyotype and FISH analysis of the HES3 cell lines

A

HES3 Control
p42+11+7+9

46,XX (70%); 46, XX dup(20)(q11.21q11.21) 30%**

20q11.21 copy number: 2 (80%), 3 (20%)

B

HES3 CNV
p73+3+3+14

46, XX (20%); 46, XX dup(20)(q11.21q11.21) 80%***

20q11.21 copy number: 3 (62%), 4 (28%), 5 (10%)
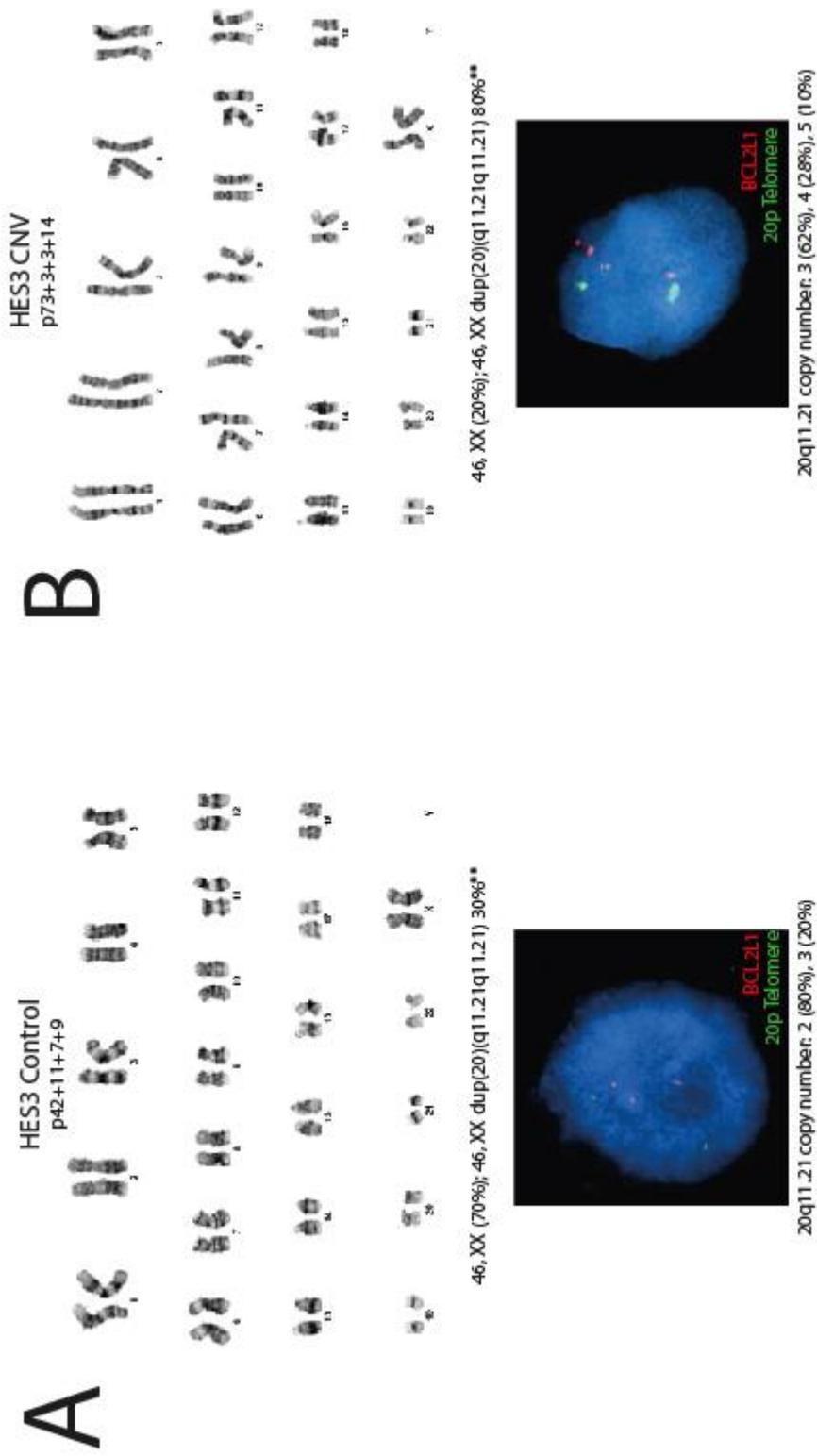
Karyotype and FISH analysis of HES3 Control (A) and HES3-CNV (B) cell lines.

Supplementary Figure 2. Karyotype and FISH analysis of the H1 cell lines

A

H1 Control
p42+10

46, XY

20q11.21 copy number: 2 (56%), 3 (44%)*

B

H1 CNV
p43+3

46, XY

20q11.21 copy number: 3 (11%), 4 (56%), 5 (33%)

BCL2L1
20p Telomere
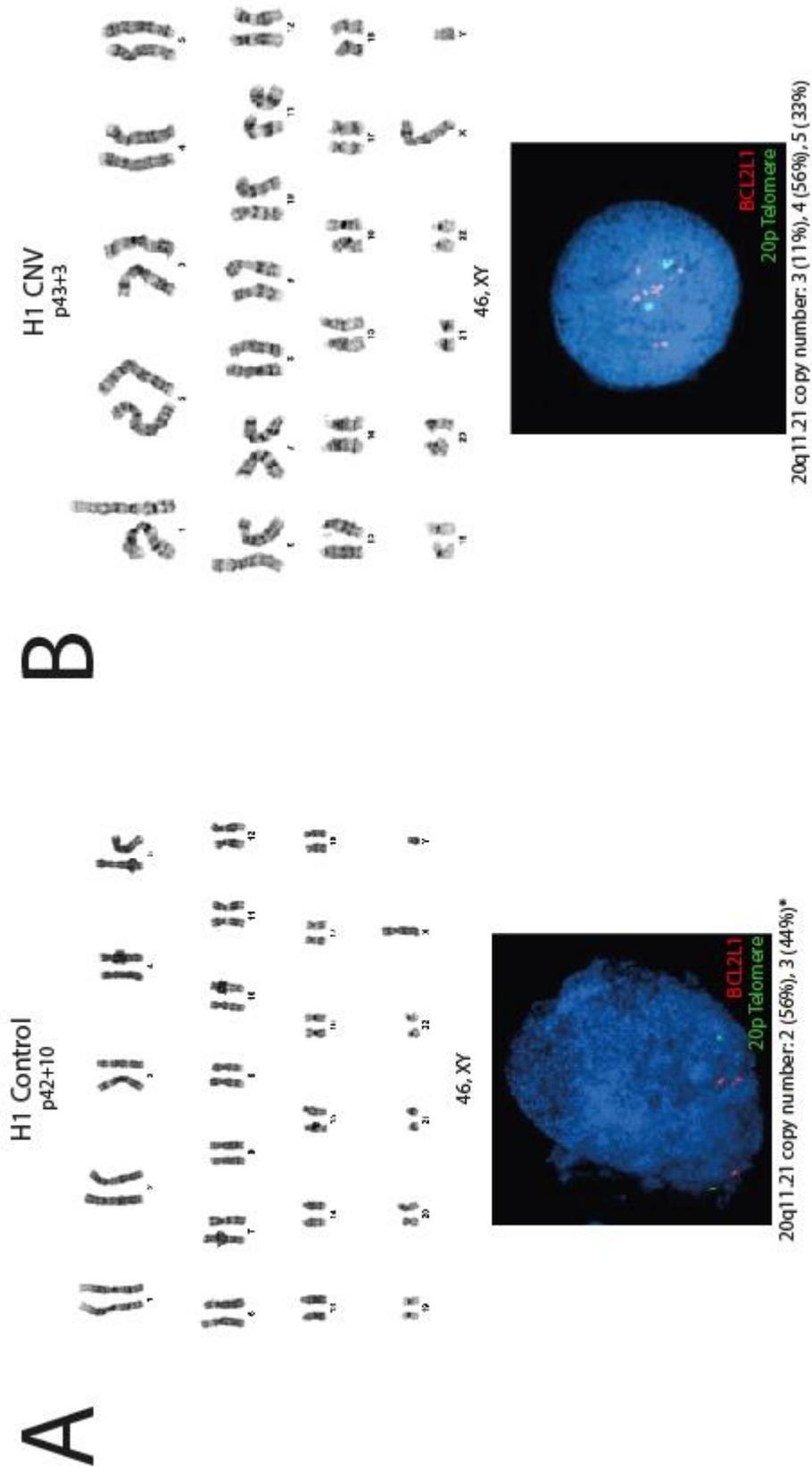
Karyotype and FISH analysis of H1 Control (A) and H1-CNV (B) cell lines.

Supplementary Figure 3. Karyotype and FISH analysis of the ESI-035 cell lines
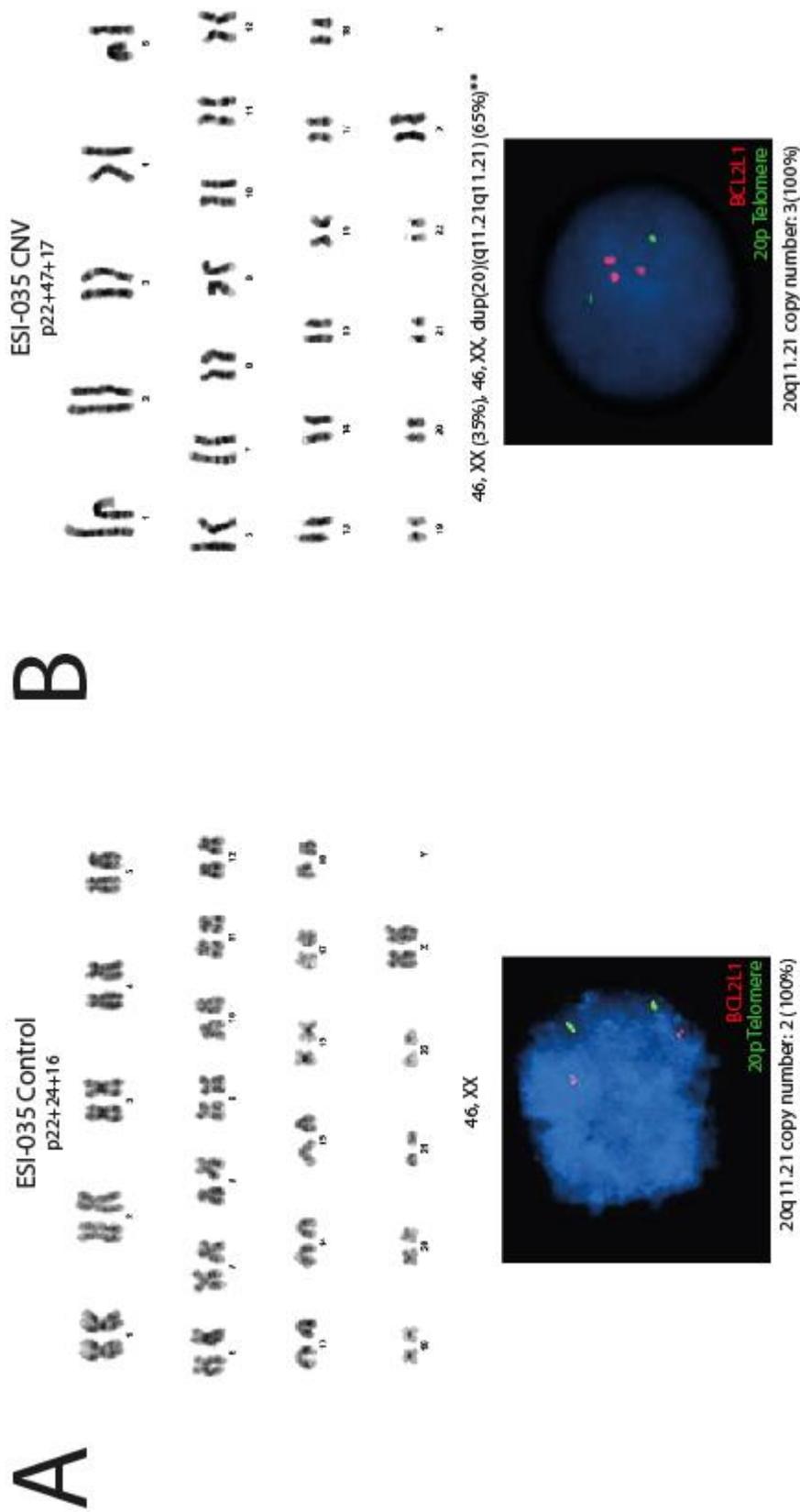
Karyotype and FISH analysis of ESI-035 Control (A) and ESI-035-CNV (B) cell lines.

Supplementary Figure 4. Karyotype and FISH analysis of the Shef5 cell lines
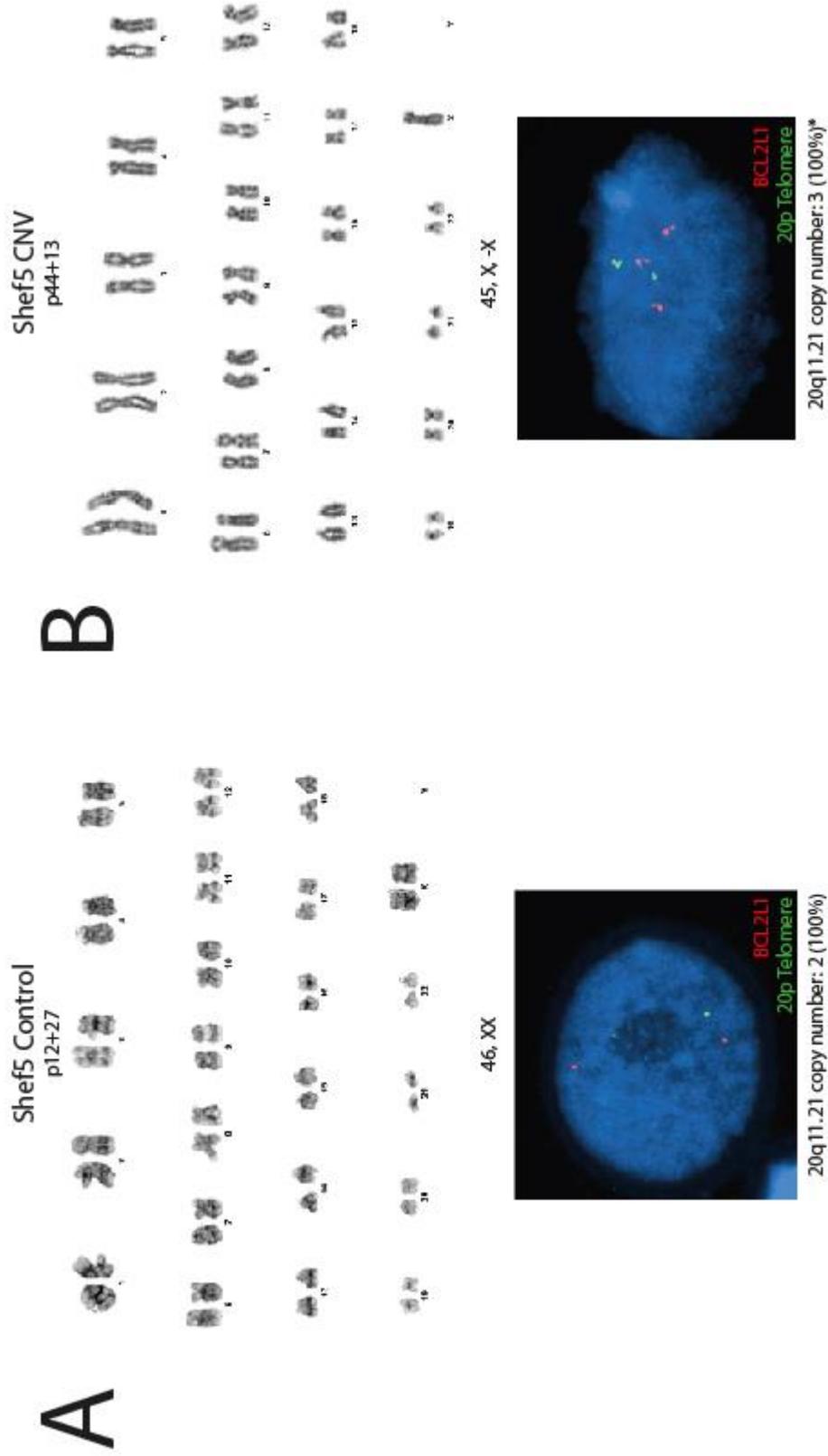
Karyotype and FISH analysis of Shef5 Control (A) and Shef5-CNV (B) cell lines.