Heterogeneity of Ewing’s sarcoma cancer stem-like cells

Mariona Chicón Bosch

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

University of Leeds

Leeds Institute of Cancer and Pathology

School of Medicine and Health

December, 2018

The candidate confirms that the work submitted is her own and that appropriate credit has been given where reference has been made to the work of others.

Acknowledgements

I am grateful for all the colleagues at the Leeds Institute of Cancer and Pathology (LICAP) that have helped me over these last 3 years. I would like to thank Dr. Adam Davison (Flow Cytometry and Imaging Facility, University of Leeds) for his advice and assistance on confocal microscopy, and for optimising with Ms. Charlotte Haunch-Smith (Children's Cancer Research Group, University of Leeds) the settings to evaluate exosomes by flow cytometry. I am grateful to Ms. Michelle Hutchinson (Clinical and Biomedical Proteomics Group, University of Leeds) for her assistance developing the survival program in R. I would like to thank Dr. Sally Fairweather (Next Generation Sequencing Facility, University of Leeds) for performing the library preparation and the RNA-sequencing of the samples used in this study. I am also grateful to Mr. Martin Fuller (Astbury Biostructure Laboratory, University of Leeds) for kindly preparing the exosomes for electron microscopy evaluation. A special thank goes for the past and present members of the Children's Cancer Research Group for their continued support and guidance over these past 3 years.

I am extremely grateful to my supervisors for their continued support. To Dr. Alastair Droop for his training in programming, bioinformatics analysis and introduction to LaTeXwriting. To Professor Susan Burchill for her guidance and support in the lab, and her continued encouraging that has helped me get this far. And to Dr. Elizabeth Roundhill, for her unceasing support in the lab and throughout the different aspects of this PhD.

Finally, to Jaume, for always being there, especially over these past 4 months. And to my family and friends, who despite the 1,400 km of distance, have been with me when I've needed them.

Abstract

The outcome for patients with Ewing's sarcoma (ES) is poor, especially for those that relapse, with overall survival <25%. Emerging evidence suggests that ES cancer stem-like cells (ES-CSCs) that are capable of self-renewal and are resistant to current treatments contribute to disease progression and relapse. I have therefore hypothesised that eradication of ES-CSCs will improve outcomes for some patients. Moreover, since exosomes play an important role in cell-cell communication I have examined whether the transfer of exosome cargo from ES cells might induce an ES-phenotype in non-ES cells and whether this cargo might include candidate biomarkers of risk.

I have successfully generated patient-derived cell cultures from primary ES (parental ES cells) and characterised these cells to better understand their heterogeneity. Differences in migration, self-renewing ability and resistance to chemotherapeutics were identified. High levels of drug resistance, measured using an in vitro assay, predicted poor patient outcome.

ES-CSCs isolated and propagated based on their self-renewing ability were more resistant to classic chemotherapies than the parental ES cells from which they were derived. Using RNA-sequencing I have identified 6 genes that were enriched in ES-CSCs compared to ES parental cells. These genes might be exploited as biomarkers of risk and candidates for the development of novel therapeutics.

Exosomes were successfully characterised from parental ES cells, ES-CSC, plasma samples and established cell lines. ES-derived exosome cargo included the ES markers CD99 and EWSR1-ETS, plus NRXN1 which was the top candidate driver of ES-CSCs. Transfer of ES-derived exosomes

from high self-renewing ES cells to low self-renewing non-ES cells induced an ES phenotype and increased the self-renewing ability of target cells.

In summary, using a functional assay I have isolated ES-CSCs and identified candidate driver genes that might be exploited as circulating biomarkers of risk and targets for the development of treatment to eradicate ES-CSCs.

Table of Contents

List of Figures

List of Tables

Abbreviations

ABC	ATP-binding casette transporter
Ago-2	Argonaute 2
ALDH	Aldehyde dehydrogenase
ANOVA	Analysis of variance
ATTC	American Type Tissue Collection
β_2M	β_2-microglobulin
bp	Base pair
BSA	Bovine Serum Albumin
c	Concordance index
CAF	Cancer-associated fibroblast
CCRG	Children's Cancer Research Group
CD99	Cluster of differentiation 99 (or MIC-2)
cDNA	Complementary DNA
CI	Confidence interval
CLL	B-cell chronic lymphocytic leukaemia
COG	Children's Oncology Group
coxph	Proportional hazards regression model
CSC	Cancer stem-like cell
CT	CellTrace™
C_t	Cycle threshold
DAPI	4', 6-diamindino-2-phentlindole
DAVID	Database fpr Annotation, Visualisation and Integrated Discovery
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
DTT	Dithiothreitol
dUC	Differential ultracentrifugation
E1AF	Early region 1A enhancer binding protein

EC50	Half maximal effective concentration
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
EE2012	Euro Ewing 2012
EEA1	Early endosome antigen 1
EFS	Event-free survival
ELFN2	Extracellular leucine rich repeat and fibronectin type III domain 2
ENSG	Ensembl Gene ID
ER	Endoplasmic reticulum
ERG	ETS related gene
ES	Ewing’s sarcoma
ESCRT	Endosomal sorting complexes required for transport
ETS	E26 transformation-specific
ETV1	ETS Variant 1
EV	Extracellular vesicle
EWSR1	Ewing Sarcoma Breakpoint Region 1
FACS	Fluorescence-activated cell sorting
FCS	Foetal calf serum
FEV	Fifth Ewing variant
FISH	Fluorescence in situ hybridisation
FLI1	Friend leukemia integration 1 transcription factor
g	Gravitational force
Grp75	Glucose-regulated protein 75
GTPase	Guanosine triphosphateases
h	Hour(s)
HR	Hazard ratio
ICC	Immunocytochemistry
IE	Ifosfamide and etoposide
IF	Immunofluorescence
IGF1	Insulin-like growth factor 1
IGF1R	IGF receptor 1
kb	Kilobase
kDα	Kilodalton

LAM	Leeds antibiotic media
LFC	Log_2 fold change
LICAP	Leeds Institute of Cancer and Pathology
M	Molar
MAPK	Mitogen-activated protein kinase
MBA	Membrane-based affinity
MEM	Minimum Essential Medium
MgCl2	Magnesium chloride
min	Minute(s)
miRNA	Micro RNA
mo.	Month old
mRNA	Messenger RNA
MSC	Mesenchymal stem cell
MT	MitoTracker™
MV	Microvesicle
MVB	Multivesicular bodies
MW	Molecular weight
MWCO	Molecular weight cut-off
n-SMase2	Neural sphingomyelinase 2
NBL	Neuroblastoma
ND	Not detected
NGS	Next Generation Sequencing
NRXN1	Neurexin 1
NSCLC	Non-small cell lung cancer
NTA	Nanoparticle Tracking Analysis
OCT	Optimal Cutting Temperature compound
OSR	Overall survival rate
OST	Osteosarcoma
PAR	Pseudoautosomal region
PARP-1	Poly ADP ribose polymerase 1
PBS	Phosphate buffered saline
PBS-T	PBS containing 0.1% Tween-20
PCR	Polymerase chain reaction

PDGF	Platelet derived growth factor
PFA	Paraformaldehyde
PI3K	Phosphoinositide 3-kinase
PME	Polymer mediated enrichment
PMSF	Phenylmethylsulfonyl fluoride
poly(A)	Polyadenylated tail
PPIA	Peptidylprolyl isomerise A
pPNET	Peripheral Primitive Neurectodermal tumours
R	Coding language R
RHA	RNA helicase A
RIN	RNA integrity number
RIPA	Radio Immuno Precipitation Assay
RNA	Ribonucleic acid
RNA-seq	RNA sequencing
rRNA	Ribosomal RNA
RT^A	Reverse transcriptase (amplification)
RT	Room temperature
RT^A-PCR	Reverse transcriptase (amplification) PCR
RT^A-qPCR	Reverse transcriptase (amplification) quantitative PCR
s	Second(s)
SC	Stem cell
SD	Standard deviation
SDS	Sodium dodecyl sulphate
SEM	Standard deviation of the mean
SFM	Serum free media
SQL	Structure query language
SR	Self-renewing
sRNA	Small RNA
SRY	Sex-determining region Y
T25	Tissue culture flask 25cm^2
T75	Tissue culture flask 75cm^2
TBE	Tris-Borate-EDTA buffer
TEM	Transmission Electron Microscopy

tRNA	Transport RNA
TSG-101	Tumor susceptibility gene 101
u	Unit
UF	Ultrafiltration
UC	Ultracentrifugation
UV	Ultraviolet
v/v	Volume/volume
VAC	Vincristine, actinomycin D and cyclophosphamide
VAI	Vincristine, actinomycin D and ifosfamide
VDC	Vincristine, doxorubicine and cyclophosphamide
VEGF	Vascular endothelial growth factor
VIDE	Vincristine, ifosfamide, doxorubicin and etoposide
w/v	Weight/volume
WB	Western blot
XIST	X-inactive specific transcript
yr.	Year old

1. Introduction

1.1. Ewing's sarcoma (ES)

1.1.1 Diagnosis and prognosis of ES

Ewing's sarcoma (ES) comprises a group of rare solid tumours including Ewing's sarcoma of bone and soft tissue, peripheral primitive neuroectodermal tumours (pPNET), Askin's tumour and neuroepithelioma. ES is the second most frequent primary bone tumour in children and young adults with a peak incidence in those of 10 to 24 years of age (Burchill, 2003). It is slightly more common in males than females (1.4:1 ratio) and has an incidence of 1.8 cases per million per year in the UK (McNally et al., 2012). The majority of ES arise in the bone (71%), being the most common sites the long bones of the leg (38 %), the pelvis (20 %) and the chest wall (16 %) (Bernstein et al., 2006). However, a proportion of cases arise in soft tissue (29%) (Stiller et al., 2013).

At a cellular level, ES is characterised by poorly differentiated small round cells, with a thin clear cytoplasm and large nuclei (Pinto et al., 2011). In addition, the cluster of differentiation 99 (CD99), also known as MIC-2, is a cell-surface glycoprotein expressed at high levels in 95–100 % of ES cells (Bernstein et al., 2006). This is a useful immunohistochemical biomarker for the differential diagnosis of ES patients from other small round cell tumours (Pinto et al., 2011), although high expression is detected in some normal tissues and other malignancies too (Grünewald et al., 2018). Based on this, other diagnosis markers have been proposed, such as the Friend leukaemia integration 1 (FLI1) transcription factor (Hornick, 2014), Caveolin-1 (Llombart-Bosch et al., 2009), NK2 homeobox 2 (NKX2-2) (Machado et al., 2018) or Leucine rich repeat and Immunoglobin-like domain-containing protein 1 (LIGNO1) (Jain et al., 2017; Town et al., 2016). However, their expression is not restricted to ES cells, and therefore, should be used in combination to discriminate false positives.

Another distinct characteristic of ES is the presence of a chromosomal translocation between the 5′-end of the Ewing Sarcoma Breakpoint Region 1 (*EWSR1*) gene (chromosome 22q12) and 3′-end of different members of the *ETS*

gene family of transcription factors (Figure 1.1). In 85 % of cases, *EWSR1* partners with *FLI1* (11q24) generating the chimeric fusion product EWSR1-FLI1 (Delattre et al., 1992). In the remaining 15 % of cases, different *EWSR1* partner genes have been identified, including *ETS* family members *ERG* (21q22) or in rarer cases *ETV1* (7p22), *E1AF* (17q12) or *FEV* (2q33) (Peter et al., 1997). Moreover, both *EWSR1* and *FLI1* chromosome breakpoint sites are dispersed over several kilobases (Kb) resulting in different fusion variants. The most common are between exon 7 of *EWSR1* and exon 6 of *FLI1* (*EWSR1-FLI1* type 1; 60 % of cases) or with exon 5 of *FLI1* (*EWSR1-FLI1* type 2; 20 % of cases) (Van Doorninck et al., 2010).

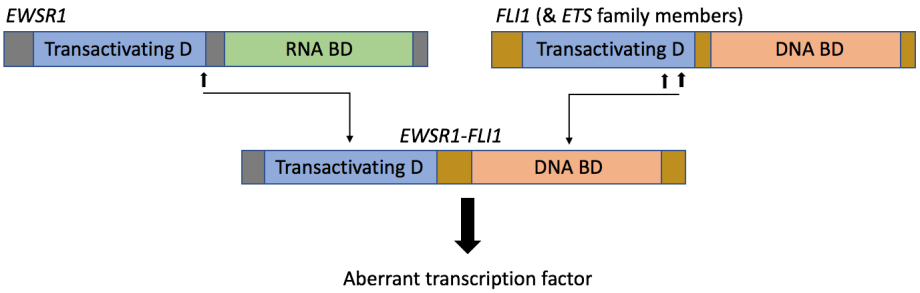


Figure 1.1: Characteristic *EWSR1-ETS* fusion in ES. The 5′ of *EWSR1* gene (22q12), involving the transactivating domain, partners with 3′ of a member of the ETS transcription factor family, which includes the DNA binding domain. This fusion generates an aberrant transcription factor. Abbreviations: D = domain, BD = Binding domain.

Many studies have investigated the role of the EWSR1-ETS fusion protein in ES. More than 2 decades ago, this chimeric fusion was reported to be an aberrant transcription activator (May et al., 1993), where the EWSR1 5′ acts as a regulator domain and the 3′ of the ETS acts as an transcription activator domain (Figure 1.1) (Bailly et al., 1994; Ohno et al., 1993). Both domains are important in the tumorigenic role of the EWSR1-FLI1 fusion (May et al., 1993). Although the DNA binding activity of FLI1 is important for the fusion's oncogenic activity, mutation studies have shown that part of the transforming activity is independent of this as fusions with an inactive DNA binding site could still transform cells (Jaishankar et al., 1999).

Several studies have investigated the post-transcriptional activity of EWSR1-FLI1, identifying several mechanisms involved in the ES phenotype with both upregulated and downregulated genes (Lessnick and Ladanyi, 2012). Amongst

these we can find many genes involved in the regulation of different pathways important for survival-proliferation balance and apoptosis, such as insulin-like growth factor (IGF) or Wnt signalling (Prieur et al., 2004); increased expression of genes associated to cell proliferation and survival, including platelet-derived growth factor C (*PDGFC*) (Zwerner and May, 2001), cyclin D1 (*CCDN1*) (Fagone et al., 2015) and *NKX2-2* (Smith et al., 2006); decreased expression of genes involved in growth inhibition and increased expression of apoptosis escape genes, including cyclin-dependent kinase inhibitor 1 (*CDKN1A*) (Nakatani et al., 2003), transforming growth factor-β (*TGF-*β) members (Im et al., 2000) or IGF-binding protein 3 (*IGFBP3*) (Prieur et al., 2004). Altogether, these results confirm the importance of the chimeric fusion in the ES tumorigenesis through different cellular mechanisms, evidence of different pathways which could be investigated for targeted therapies.

Additional structural and numerical chromosomal aberrations have been implicated in ES tumorigenesis and clinical outcome. These include gains in chromosomes 8 and 12 (Armengol et al., 1997), chromosome 20 (Roberts et al., 2008), the *CDKN2A* locus on chromosome 9 (Tirode et al., 2014) and presence of an unbalanced translocation between chromosomes 1 and 16 (der(16)t(1;16)) resulting in a gain of 1q and partial loss 16 (Mugneret et al., 1988). The presence of complex karyotypes has been associated with worse outcomes (Roberts et al., 2008), with a particular link between gain of 1q, 12 and loss of 16q to worse outcomes (Hattinger et al., 2002).

Different studies have evaluated the presence of somatic mutations in ES, finding a shortage of secondary genetic abnormalities (Brohl et al., 2014; Crompton et al., 2014; Tirode et al., 2014). Amongst the few recurrent mutations linked to ES, *STAG2* and *CDKN2A* have been reported in 17% and 12% of cases respectively, and have been found to be mutually exclusive (Tirode et al., 2014). Although less frequent, mutations in *TP53* have been reported in 7% of cases (Tirode et al., 2014). Both *CDKN2A* and *TP53* mutations have been associated to poor outcomes (Huang et al., 2005). Other less common alterations include overexpression of the epigenetic modulator *EZH2*, a known mediator of the EWSR1-ETS fusion (Richter

et al., 2009); and loss of expression of the tumour suppressor gene *TGFβRII* (Im et al., 2000; Tirode et al., 2014).

As well as the various genetic alterations that inform on the prognosis of ES patients, there are several clinical factors that have been associated to outcome. Presence of metastatic disease at diagnosis, larger tumour volume and older age are associated with poor outcome, whereas higher presence of necrotic tissue after treatment, local relapse and younger ages correlate with better outcomes (Gaspar et al., 2015; Grünewald et al., 2018).

1.1.2 Current treatment and outcome of ES

Although significant improvement has been made in the diagnosis of ES and treatment for patients with localised disease over the last decades, ES is still frequently associated with poor prognosis and low overall survival rates (OSR) (Gaspar et al., 2015). ES patients with localised disease have benefited greatly from the introduction of chemotherapy, with OSR increasing from 10 before introduction of chemothepray (40 year ago) to 75% (Balamuth and Womer, 2010). In contrast, patients diagnosed with metastatic disease (25% of ES cases) have not seen an improvement in survival, where the 5-year OSR remains under 30 % (Gaspar et al., 2015). Nevertheless, patients with lung metastasis have better survival rates (5-year OSR 50 %) compared to those with metastasis at other sites (OSR <20%) (Gaspar et al., 2015). In addition, a considerable proportion of ES patients are refractory to conventional treatment (25% with localised disease, 75% with metastasis at diagnosis) and another significant proportion (26%) may initially benefit from cytotoxic therapy but the disease will recur (5-year OSR for local relapse 25%; for distant relapse 22%) (Rodriguez-Galindo et al., 2003). Moreover, patients that relapse within 2 years of the diagnosis have worse prognosis (5-year OSR 7%) than those that relapsed later (OSR 29%) (Stahl et al., 2011). Therefore, there is a need for better treatment strategies, specially for the patients with metastatic disease at diagnosis and those who relapse, as these have the lowest survival rates.

Over the last decades, cooperative research in clinical trials has led to advances in ES treatment resulting in a multidisciplinary strategy that combines localised therapy (surgery and/or radiotherapy) to remove the main tumour and systemic

therapy (chemotherapy) to reduce the tumour size and remove micrometastatic disease (Grünewald et al., 2018). Current clinical trials are evaluating different chemotherapeutic combinations before (induction) and after (consolidation) surgery/radiotherapy in order to define the best treatment modality. Interestingly, the established induction chemotherapy differs between Europe and North America. The chemotherapeutic regime in Europe is the combination of vincristine, ifosfamide, doxorubicin and etoposide (VIDE) (Juergens et al., 2006), whereas the compressed regime (given in 2-weeks instead of 3-weeks) of vincristine, doxorubicin and cyclophosphamide alternated with ifosfamide and etoposide (VDC-IE) is the protocol adopted in North America (Womer et al., 2012). All these are cytotoxic chemotherapies, with a mechanism of action based on targeting important aspects of cell division or DNA replication. Although these are processes important for the fast-dividing cancer cells, cells of the normal tissue will be affected too, thus leading to toxicity and side-effects due to the treatment (Remesh, 2012). Moreover, any cancer cell with slower cell division will be less affected.

In Europe, the EURO-EWING 99 clinical trial consisted of giving VIDE as induction chemotherapy and then randomizing patients according to risk group for the consolidation chemotherapy. Patients with low risk received either vincristine, actinomycin D and cyclophosphamide (VAC) or vincristine, actinomycin D and ifosfamide (VAI); and patients with high risk received either VAI or high-dose chemotherapy (busulphan and melphalan; Bu-Mel) followed by stem-cell transplant (Juergens et al., 2006). Results from this indicated that cyclophosphamide could replace ifosfamide in low risk patients as consolidation therapy (Le Deley et al., 2014), as previous studies questioned the efficacy of ifosfomide and related toxicities (Suarez et al., 1991). In North America, the Children's Oncology Group (COG) trial compared the administration of VDC-IE induction chemotherapy every 2 weeks (compressed) or every 3 weeks (standard), showing a better 5-year event-free survival (EFS) for patients receiving the compressed modality (Womer et al., 2012).

The ongoing Euro Ewing 2012 (EE2012) clinical trial is assessing both induction modalities (Europe and North America) in localised and lung metastatic patients in

order to define the most effective induction chemotherapy strategy for ES patients at the same time as evaluating therapy-related toxicities (Odri et al., 2010) (Figure 1.2). In addition, the EE2012 clinical trial is also evaluating the effect of zolendronic acid in consolidation therapy, a bone resorption inhibitor that has an effect on osteolysis and tumour invasion and proliferation (Odri et al., 2010) (Figure 1.2). This is covered in more detail in Section 1.1.4.

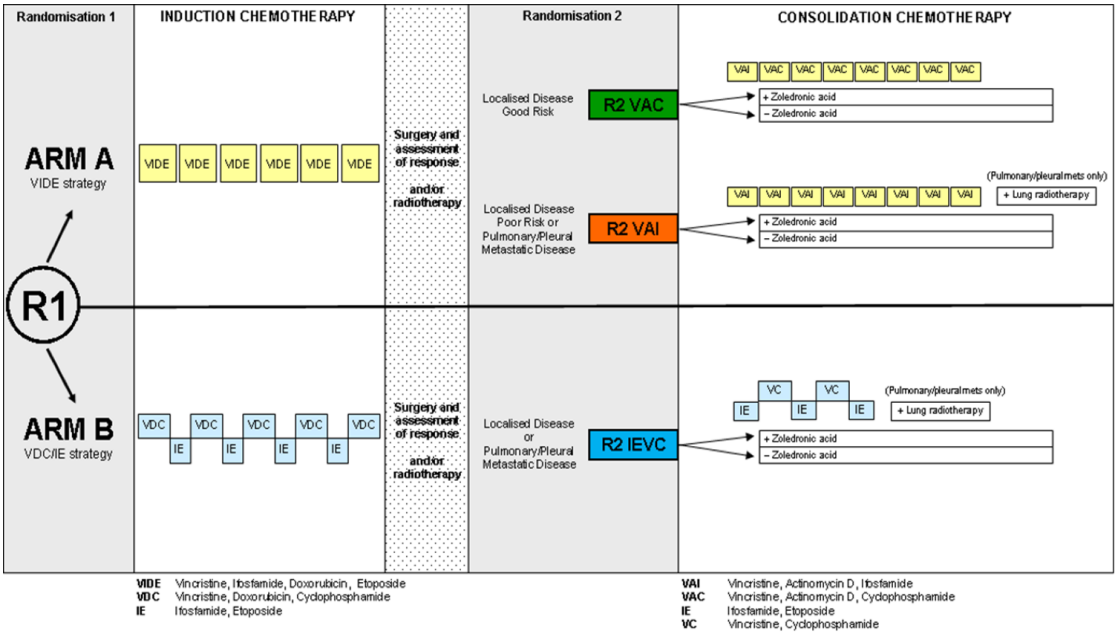


Figure 1.2: Euro Ewing 2012 randomized clinical trial for treatment of ES patients with localised and lung metastatic disease. Arm A consist of the European protocol (VIDE for induction, and VAC or VAI for consolidation chemotherapy for good risk or poor risk patients respectively). Arm B consists on the US protocol (VDC-IE compressed induction chemotherapy and IE-VC for consolidation). Euro Ewing 2012 is also comparing the effect of zoledronic acid alongside consolidation chemotherapy. Abbreviations: Vincristine, Ifosfamide, Doxorubicin, Etoposide, Actinomycin D, Cyclophosphamide. (Source: Reproduced from the Euro Ewing 2012 clinical trial protocol with permission from Dr. Bernadette Brennan and Professor Keith Wheatley).

As mentioned before, patients with refractory disease or relapse are associated with unacceptably low OS. Moreover, there is no standard treatment for these patients. The Euro Ewing Consortium (EEC) study for recurrent ES (rEECur) is currently evaluating the most effective chemotherapy regimen out of 4 different combinations with the lowest toxicity in refractory and relapse ES patients (ISRCTN 36453794). In addition, due to the treatment strategies currently used for this malignancy, ES patients are at risk of developing treatment-induced toxicities

(Grünewald et al., 2018). The cytotoxic drugs currently used for the treatment of ES have been associated to therapy-related toxicities, as mentioned before. A 9% of ES patients develop secondary malignancies and long-term ES survivors (>25 years) have an increase in mortality and morbidity compared to the general population (Ginsberg et al., 2010). In addition, a recent study showed that 30% of ES patients with 5-year EFS died within 35 years of diagnosis, being the most common causes of death late relapses (15%) and treatment-related toxicities (11%) (Marina et al., 2018).

Therefore, there is a need for better treatment strategies in ES, with the aim to improve outcomes for ES patients, sepecially those with metastatic, relapse and refractory disease. The implementation of better treatment approaches might result in a reduction of the side-effect toxicities associated to the current treatment modalities. Moreover, the identification of better predictive and prognostic biomarkers is essential, as this will result in better treatment stratification for ES patients and ultimately, better survival rates.

1.1.3 Novel treatment strategies in ES

In light of low 5-year OSR for ES patients, the high incidence of relapses and the toxicities associated to current treatment strategies, many groups in the ES community are investigating the implementation of targeted therapies against ES biological markers and signalling pathways.

The involvement of a characteristic translocation in ES pathogenesis suggests that EWSR1-FLI1 is the obvious target to focus on, as has been done in other malignancies (Capdeville et al., 2002). However, this has proven to be problematic, as the chimeric protein has absence of intrinsic enzymatic activity and lacks a binding pocket (Pishas and Lessnick, 2016). Although the fusion was identified more than 25 years ago (Delattre et al., 1992), it has not been until recently that a targeted therapy preventing the EWSR1-FLI1 oncoprotein from interacting with RNA helicase A (RHA) has reached the clinical setting. In 2009, a compound that could disassociate EWSR1-FLI1 from the RNA helicase A (RHA) was reported (Erkizan et al., 2009). RHA had previously been reported to enhance the transcription function of the chimeric fusion (Toretsky et al., 2006). This small molecule, YK-4-279, could induce apoptosis in ES cells and reduce tumour growth

in orthotopic ES mouse xenografts, in comparison to non-ES models (Erkizan et al., 2015). Since the identification of this compound, different groups have been assessing its efficacy as a novel targeted therapy for ES (Pishas and Lessnick, 2016). A recent study showed YK-4-279 to have synergistic effects with vincristine (Zöllner et al., 2017). Mice with ES xenograft treated with both compounds showed a decrease in the tumour burden and better survival rates. An ongoing phase I clinical trial is evaluating the initial safety of YK-4-279 (under the TK216 name) in relapse and refractory ES patients (NCT02657005)

With a high proportion of ES cells expressing the CD99 cell surface marker, targeting this protein has been an alternative strategy to treat ES. Moreover, high expression of CD99 has been associated with migration capacity (Kreppel et al., 2006) and inhibition of neural differentiation (Rocchi and Manara, 2010) in ES cells, indicating a role in ES pathogenesis. The use of anti-CD99 in ES cells induces apoptosis, with synergistic effects when used alongside conventional chemotherapy (Scotlandi et al., 2006). This combination therapy showed effective results in mice ES xenograft bearing localised and metastatic disease tumours. Later studies showed that CD99 engagement in ES cells induced methuosis *in vitro*, a form of non-apoptotic cell death via catastrophic vacuolation (Manara et al., 2016). This process was reported to be related to the IGF-receptor 1 (IGF1R) expression, which upon inhibition of CD99 leads to an increased level of vacuoles and cell death. As mentioned in Section 1.1.1, expression of CD99 is not limited to ES cells, and therefore, this treatment strategy could lead to off-target toxicities. However, a recent study showed promising results when evaluating anti-CD99 therapy in combination with doxorubicin (Guerzoni et al., 2015). Both *in vitro* evaluations and ES xenografts data provided evidence of cell death, mdm2 degradation and p53 reactivation in ES cells, in comparison to mesenchymal stem cells (MSC), known for their high levels of CD99 expression. Therefore, this data suggests that an altered genetic profile is needed for the efficacy of this treatment strategy, thus sparing normal tissue cells with high CD99 expression. Altogether, these studies suggest anti-CD99 targeted therapy in combination with conventional chemotherapy regimens as a suitable candidate for better treatment strategies.

Another strategy in ES is acting on the EWSR1-FLI1 downstream signalling pathways. As mentioned in Section 1.1.1, IGF signalling pathway is amongst the genes regulated by the EWSR1-FLI1 fusion protein. This pathway is a complex system which plays a role in different malignancies, including sarcomas (Larsson et al., 2005; Mancarella et al., 2018). IGF1R is activated by ligand binding (i.e. IGF1), which leads to the activation of intracellular signalling pathways initiated by an intracellular tyrosine kinase domain (Balamuth and Womer, 2010). This, in turn, activates the phosphatidylinositol 3 kinase (PI3K) and mitogen-activated protein kinase (MAPK) signalling pathways, which ultimately results in cell proliferation, invasion, metastasis and inhibition of apoptosis (Balamuth and Womer, 2010; Mancarella et al., 2018). To regulate the activation of IGF1R via the binding of its ligands (IGF1), different IGF-binding proteins (i.e. IGFBP3) can bind to circulating IGF1 and prevent the interaction to the receptor (Van Maldegem et al., 2016). In ES, an autocrine loop activates the IGF1R system in both ES cell lines and tumours (Scotlandi et al., 1996), where the EWSR1-FLI1 fusion plays an active role in IGF signalling regulation by binding to the IGFBP3 promoter and repressing its activity (Prieur et al., 2004). Supporting the role of IGF in ES, *in vitro* inhibition of IGF1R resulted in proliferation inhibition, survival reduction and inhibition of IGF pathway as well as reduction of tumour growth *in vivo* (Benini et al., 2001; Guenther et al., 2018; Manara et al., 2007, 2016; Martins et al., 2006; Scotlandi et al., 1998). Moreover, addition of exogenous IGFBP3 in ES mouse xenograft led to growth inhibition and reduction of metastasis (Benini et al., 2006).

All this evidence on the role of IGF signalling pathway in ES tumorigenesis has resulted in great interest in introducing IGF targeted strategies for the treatment of ES patients. Both monoclonal antibodies aganist IGF1R and IGF1 as well as tyrosine kinase inhibitors against IGF1R have been evaluated in the clinic, with remarkable results in a subset of ES patients (Juergens et al., 2011; Olmos et al., 2011; Tap et al., 2012). However, phase II clinical trial results were not significant enough to continue following this strategy, and most studies were not continued (Grünewald et al., 2018; Mancarella et al., 2018). Although there is a clear involvement of IGF pathway in ES tumorigenesis and that the levels of expression of IGF1 and IGF1R are predictors of prognosis (Scotlandi et al., 2011), this targeted strategy is not

resulting as effective as expected when identified in ES. Future studies would need to identity the subset of patients that can benefit mostly of this treatment and evaluate the efficacy on using IGF inhibitor as monotherapy or in combination with other treatments.

The use of trabectedin, a DNA binding drug approved for the use of soft-tissue sarcomas (Gordon et al., 2016), is showing promising results in ES as these cells are more sensitive to this drug compared to other sarcomas, being able to reverse the expression levels of altered signalling pathways (Grohar et al., 2011). This was corroborated by Amaral et al. (2015), where combination of trabectedin with anti-IGF1R showed synergistic cytotoxic effects in ES cell lines and xenografts. A phase II clinical trial by COG showed that from the 16 ES patients treated with trabectedin, only 1 achieved stable disease when used as single agent (Baruchel et al., 2012). From this data, the authors concluded that there was not enough evidence to continue evaluating this drug as single agent. Based on this, it might be that ES patients could benefit better if trabectedin was to be used in combination with other chemotherapeutic agents.

Another targeted therapy strategy that has become popular in the ES field is against the poly ADP ribose polymerase 1 (PARP-1), important in the DNA single-strand base excision repair machinery. Therefore, inhibition of PARP-1 leads to catastrophic double-stranded DNA breaks if DNA cannot be repaired before replication. In ES, PARP-1 is highly expressed compared to other sarcomas (Pishas and Lessnick, 2016) and the use of the PARP-1 inhibitor olaparib alongside standard chemotherapy showed complete response in ES xenografts as well as a direct positive feedback between EWSR1-FLI1 expression and higher levels of PARP-1 (Brenner et al., 2012; Stewart et al., 2015). Based on the different mechanisms of action of these drugs and the possibility to combine them to increase efficacy, a recent study investigated the use of the PARP-1 inhibitor olaparib alongside trabectedin (Ordóñez et al., 2015). A reduction of cell proliferation, apoptosis and cell cycle arrest was seen *in vitro*, as well as complete remission in patient-derived xenograft (PDX) models.

Other targeted therapies have exploited proteins with high expression in ES. Inhibition of lysine-specific demethylase 1 (LSD1), an epigenetic regulator

important in ES tumorigenesis, showed induction of apoptosis and reversion of the EWSR1-ETS transcription profile *in vitro* as well as tumour growth reduction as a single agent in xenograft ES models (Sankar et al., 2014). Other studies have evaluated the inhibition of the tyrosine kinase c-kit, highly expressed in ES in 30% of patients (Scotlandi et al., 2003), although the doses needed to induce an effect on ES cells were higher than for those cancers where this drug is being used in the clinic. This is consistent with data from a phase II clinical trial in patients with relapse and refractory ES, where c-kit inhibition did not show efficacy as a single agent (Bond et al., 2008).

Therefore, many different strategies are being evaluated for ES patients, some with promising results in the clinical setting. Some of them use novel agents designed for this cancer (i.e. anti-CD99 and YK-4-279) and others use drugs approved for other cancers (i.e. olaparib and trabectedin). Based on the different studies mentioned and the clinical data available, the combination of strategies targeting various mechanisms in cells suggests better results. Therefore, investigations on novel therapeutic approaches should evaluate the different compounds in combination, in order to increase the effect on ES cells, which could lead to better results. However, this has to be evaluated carefully, as this could lead to off-target toxicities.

1.1.4 Microenvironment in ES development and its therapeutic opportunities

It is increasingly clear that the behaviour of cancer cells is highly dependent on cross-talk and interaction of cancer cells with its host environment. Moreover, cancer cells have shown adaptive responses to the cellular context they arise in, with the tumour microenvironment modulating cancer progression and metastasis.

In the case of ES, as a high proportion of tumours locate in the bone, the role the bone microenvironment could play on this disease has to be taken into consideration. This "bone niche" would help the cancer phenotype, by releasing different signals such as growth factors and cytokines (Redini and Heymann, 2015). The bone destruction linked to ES could originate from an activation of osteoclasts, which would lead to bone resorption (destruction of bone tissue) (Lau et al., 2007). For this, tumour cells release osteoclast activation factors (i.e. IL-6 or

TNF-α) at the same time that bone resorption leads to the release of factors contained in bone tissue (i.e. IGF1, TGF-β and PDGF) (Guise et al., 2005). This is especially important, as many of these factors have been associated with ES tumorigenesis (Section 1.1.1). Therefore, inhibiting the activation of osteoclasts, either directly on these cells or indirectly by inhibiting the signals around them, becomes an interesting treatment strategy for ES, as it might have an impact on important upregulated ES pathways.

The most straight forward strategy would be by inhibiting bone resorption by directly targeting osteoclasts. The use of bisphophonates, a synthetic substitute of endogenous pyrophosphate, leads to the inhibition of bone destruction by inducing osteoclast apoptosis (Roelofs et al., 2006). Zolendronic acid is one of these molecules, and has gained interest in the past few years. Using xenograft ES intraosseous models, Odri et al. (2010) reported tumour growth inhibition when used as single agent and less relapses when used in combination with ifosfamide. Confirming the mechanism of action in bone, this effect was not seen in xenograft ES soft-tissue models (Odri et al., 2010). In addition, zolendronic acid showed inhibition of invasion on ES cells *in vitro* and reduction on the number of lung metastasis in xenograft ES models (Odri et al., 2014), suggesting an additional anti-tumour effect independent of osteoclast targeting. Based on these findings, this drug has been included in the EE2012 clinical trial (Section 1.1.2), to evaluate its efficacy in ES patients.

Another important factor in the tumour microenvironment is hypoxia (oxygen deprivation), which correlates with more aggressive cancers and worse prognosis in many malignancies (Muz et al., 2015). Hypoxia induces transcription of genes related to vascularization, tumour growth, remodeling of extracellular matrix component, invasion or formation of the pre-metastatic niche (Redini and Heymann, 2015). In ES, expression of hypoxia-inducible factor (HIF)-1α has been linked to up-regulation of EWSR1-FLI1, inducing invasion and colony formation under hypoxic conditions *in vitro* (Aryee et al., 2011). Moreover, hypoxic conditions increase expression of neuropeptide Y (NPY), a transcription target of EWSR1-ETS, leading to bone invasion and tumour growth in xenograft ES models (Hong et al., 2015). Inhibition of NPY induced a reduction of bone destruction (Hong et al.,

2015). Therefore, targeting of hypoxic factors in ES might provide useful in targeting tumour dissemination.

Another factor that has been related to ES progression through the microenvironment is the implication of the immune system. Berghuis et al. (2011) evaluated the relationship between tumour cells and the microenvironment immune responses in ES, showing an increase in the amount of $CD8^+$ T-cell infiltration in ES compared to the stroma cells, suggesting that this difference could come from the inflammatory status in tumour areas. Angiogenesis is an important factor in many malignancies. In ES, the blockade of angiogenesis using bevacizumab antibody, which inhibits the vascular endothelial growth factor (VEGF) interactions with VEGF receptors 1 and 2, has been exploited as a targeted therapy showing efficacy in mouse models (Gaspar et al., 2015; Lee et al., 2006), and has reached clinical trials with promising results for ES patients (Vo et al., 2016).

Altogether, these data shows the involvement of the microenvironment in modulating ES progression. Therefore, targeting these mechanisms could become another interesting option for novel therapeutic targets in ES. In addition, this importance in the microenvironment suggests an extra layer of complexity when investigating ES tumorigenesis and candidate drugs. This modulation from the microenvironment described above might have implications in the efficacy of drugs, and therefore, needs to be taken into consideration when testing new compounds in a preclinical setting.

1.1.5 Pre-clinical models to study ES

Better understanding of ES biology, tumorigenesis and identification of novel therapeutic strategies relies greatly on the preclinical models used. Contrary to other malignancies, to date ES lacks a genetic engineered mouse model, as expression of EWSR1-FLI1 results in embryonic lethality (Ent et al., 2018; Minas et al., 2017). Therefore, scientists in the ES field have been using and developing other model systems to study ES (Table 1.1).

Established ES cell lines are widely used in the pre-clinical setting, providing an accessible and easy-to-use tool (Table 1.1). However, their continued culture

results in changes in the cell cycle regulation and abnormal metabolic characteristics (Stacey, 2005; Teicher et al., 2011). Consistent with this, ES established cell lines showed a different gene expression profile when compared to patient tumours (Crompton et al., 2014). For this reason, many groups are using patient-derived cell cultures (Table 1.1), originated from the culturing of tumour tissue obtained from ES patients. As these are obtained from fresh tissue, they have not been afected by the extensive culturing established cell lines have been exposed to, leading to better models more resemblant to the patient's biology (Miserocchi et al., 2017). Moreover, the ability to match the *in vitro* profile obtained through evaluation of different cancer phenotypes to the clinical information of the patient provides a powerful mechanism for the identification of prognostic biomarkers and response to treatment. Similarly, another advantage of using patient-derived cell cultures is the ability to have sequential samples (different time-points) or access to other biological materials from the same patient (i.e. plasma or urine), which might become useful for validation studies. However, obtaining patient-derived cell cultures is limited by the low incidence of this cancer.

Although patient-derived cell cultures provide a better model than established cell lines, both are 2D cultures, thus lacking the cell-to-cell interaction present in the body as well as the microenvironment effects (Ent et al., 2018). For this, the use of animal engraftment models provides a powerful solution (Table 1.1). Although the first xenograft mouse models used subcutaneous and intravenous injection of ES cells (Scotlandi et al., 1998; Vormoor et al., 2001), more recent models have recreated better the location of ES tumours in the bone (Goldstein et al., 2015; Vormoor et al., 2014; Wang, 2009). These models have become useful tools to test novel drug treatments, as the 3D tumour structure and the interaction with the host provide better systems to study ES compared to 2D cultures. However, most of these xenografts use established cell lines, thus relying again on models that have differed from the patient cellular profile, and therefore, could be leading to biased results.

For this reason, recent efforts have been put in developing patient-derived xenografts (PDX; Table 1.1), for which a piece of the patient tumour is placed in the mouse and grown in these conditions (Ent et al., 2018; Geier et al., 2015). These

provide a model more closely resemblant to the biology of ES, showing promising results (Geier et al., 2015; Grünewald et al., 2018; Guenther et al., 2018). Of interest, the paediatric preclinical testing program (PPTP) initiative has the aim to develop ES PDX, amongst other paediatric malignancies, in order to do preclinical therapy testing (Grünewald et al., 2018; Houghton et al., 2007). Therefore, PDX might provide the most useful system to investigate ES in a preclinical setting, and should be used instead of established cell line xenografts.

It is important to mention, however, that the maintenance of mouse models requires special infrastructure and personnel, which could be a limitations for many groups. In order to overcome this, several groups have developed scaffolds to recreate the 3D interactions present in the body without the need of animal models (Ent et al., 2018) (Table 1.1). ES cells incubated in a polymer scaffold used *in vitro* showed a similar profile of gene expression than xenografts, compared to 2D cultures (Fong et al., 2013). In addition, other studies have tried to recreate the scenario ES cells are in the body, such as shear stress induced by blood flow or compression, obtaining similar gene expression profiles than those seen in *in vivo* models (Marturano-Kruik et al., 2018; Santoro et al., 2015). Therefore, these systems provide a good model for the study of ES tumorigenesis without the use of animal systems.

Further understanding of ES biology and the cell of origin might ultimately lead to the development of a genetic engineered mouse model. Before this is achieved, the systems described in Table 1.1 are the best tools to investigate ES tumorigenesis and provide preclinical systems to screen for novel therapeutics. However, it is important to bear in mind that even PDX, which are considered to be the best model, are grown in mice. The interaction of the host with the human tumour could lead to biased results, such as a different interaction between the host microenvironment or the immune system with the tumour cells, which could result in false positives and negatives when investigating novel treatment strategies. Therefore, results obtained should be evaluated carefully and if possible, in different models to account for heterogeneity.

Pre-clinical model	Description	Advantages	Disadvantages
Established cell line	Cell culture (2D) generated from tumour, extensive culture.	Widely used; accessible; easy to use.	Changes in cell cycle regulation; abnormal metabolism.
Patient-derived cell cultures	Cell culture (2D) generated from fresh tumour sample, not extensively cultured.	Less time maintained in culture (less cellular adaptations); heterogeneity of patients and tumours; match other samples and time points.	Limited access to tumour samples; hard to generate cell cultures.
Xenograft mouse model	Mouse engraftment with established or patient-derived cell culture (3D).	Recreate tumour/host interactions; study disease progression, metastasis, etc; different tumour locations.	Use cell cultures (changes in cell cycle); need of specialised personnel and facilities; expensive.
Patient-derived xenograft (PDX)	Mouse engraftment using patient tumour mass, no previous culture (3D).	Recreate tumour/host interactions; study disease progression, metastasis, etc; different tumour locations; use fresh tumour (no culture, no adaptation).	Limited access to tumour samples; need of specialised personnel and facilities; expensive.
Scaffold	Artificial structure to host cell culture and recreate cellular interactions (3D).	No use of animals (no need of especial infrastructure/personnel); design conditions tailored to research need.	Design of scaffold; complex design of cellular interactions and factors that could impact on tumour growth.

Table 1.1: Summary table of pre-clinical models in ES.

1.2. ES cellular context

1.2.1 ES cell of origin

After almost a century since ES was first described, the cell of origin is still under debate. When J. Ewing described ES in 1921, he suggested an endothelial origin (Ewing, 1921). Later studies, however, showed expression of cell surface markers related to the neuroectodermal lineage, suggesting a link between ES and the neuroectoderm (Lipinski et al., 1986).

The chimeric EWSR1-ETS fusion acts as a strong transcription factor, becoming the main driving event in ES (covered in Section 1.1.1). Because of this, many groups have investigated how expression of the fusion in different cellular contexts transforms non-ES cells towards an ES gene expression profile, with the aim to identify which might be the cell of origin for ES.

It was not until the 1990s, after the chimeric fusion was fully characterised, that the first studies transfecting the *EWSR1-FLI1* fusion took place. Experiments using murine fibroblast NIH3T3 cells showed a transformation into an ES-like phenotype upon expression of the fusion, with downstream target genes of the fusion being expressed as well (Braun et al., 1995; May et al., 1993). However, this was not observed in all fibroblasts investigated, thus suggesting a permissive cellular contex upon which ES-like transformation is possible. Nevertheless, the ability to transform non-ES cells via expression of EWSR1-FLI1 started the path towards the use of such model systems. With the introduction of more powerful gene expression technologies, it was found that the ES-like profile obtained in murine fibroblasts does not mimick the ES tumour profile in patients, and only transformed human fibroblasts correlated with patient tumours. Lessnick et al. (2002) showed that EWSR1-FLI1 expression on human fibroblasts, instead of inducing an oncogenic transformation, led to cell cycle arrest. In neuroblastoma (NBL) cells (Rorie et al., 2004) as well as in rhabdomyosarcoma (Hu-Lieskovan et al., 2005), expression of the ES oncoprotein lead to an ES-like phenotype with neural crest characteristics. Based on these findings, neural crest cells have been postulated as the cell of origin, as they tolerate the presence of the EWSR1-FLI1 fusion, initiating an ES-like profile with expression of important ES-related genes (Staege et al.,

2004; von Levetzow et al., 2011). Therefore, the neural crest profile in ES cells that was suggested to indicate a neuroectoderm origin in earlier studies could just be induced by expression of the fusion. Nevertheless, these data suggested that a specific permissive cellular environment was important to tolerate the fusion and induce an oncogenic transformation similar to that of ES.

A more recent hypothesis suggests ES to arise from a pluripotent precursor cell, where the EWSR1-FLI1 fusion would induce the tumorigenic profile and inhibit normal differentiation. Increasing evidence suggests MSCs as the candidate cell of origin for ES. These are multipotent cells that can differentiate into multiple cell types of the mesoderm lineage, including osteoblasts (bone), chondocytes (cartilage), myocytes (muscle), adipocytes (fat) and bone marrow. This is consistent with the multiple tissues in which ES can arise. Consistent with this hypothesis, Tirode et al. (2007) showed that silencing *EWSR1-FLI1* in ES cells induced a cell phenotype and gene expression profile similar to that of MSCs. Furthermore, these cells were able to differentiate into other lineages under appropriate conditions. Moreover, Riggi et al. (2008) demonstrated that MSCs tolerated the expression of the EWSR1-FLI1 oncoprotein. Following expression of EWSR1-FLI1 in MSCs, the cells acquired a gene expression profile resembling that of ES, but not that of other bone and soft tissue sarcomas. Later studies identified paediatric MSCs, rather than adult MSCs, to be more readily transformed by the expression of EWSR1-FLI1 in ES, consistent with the high incidence of ES in young people (Riggi et al., 2010). Although expression of the fusion in MSCs induces expression of important genes involved in ES tumorigenesis, such as EZH2, NKX2.2 or IGF1 (Riggi et al., 2008), this is not enough for the formation of ES *in vivo*. It has been suggested that the fusion alone might not be enough for the transformation of MSCs into ES, and that other cellular changes are needed, such as mutations in *p53* (Castillero-Trejo et al., 2005) or epigenetic regulation (Richter et al., 2009).

Although there remains uncertainty as to whether ES arise in cells of the neural crest or MSCs, these two possibilities are not mutually exclusive since MSCs could derive from an intermediate neural crest state (Takashima et al., 2007). This harmonises the different studies on ES cell of origin and is consistent with the multiple organs in which these tumours can arise. Based on an increasing

understanding of ES tumorigenesis, the role of the microenvironment and new emerging ES models, we might be close to identifing the cell of origin of this malignancy. This could ultimately lead to the generation of a genetic engineered mouse model and a deeper understanding of how ES arises and develops.

1.2.2 Cancer stem-like cells

Traditionally, cancer has been explained using the stochastic or clonal evolution model, which suggests that cancer arises and progresses by acquisition of genetic mutations and epigenetic alterations, resulting in a selective advantage and subsequent heterogeneous cancer population (Nowell, 1976) (Figure 1.3 top). However, more recent research has lead to the development of an alternative model for cancer evolution, known as the cancer stem-like cell model. This is based on a hierarchical organization in which cellular growth and progression leading to tumorigenesis arises in a minor fraction of tumour cells, called cancer stem-like cells (CSC) or tumour-initiating cells (Visvader and Lindeman, 2012) (Figure 1.3 bottom). CSCs are characterised by self-renewal ability (can create unlimited copies of themselves), differentiation (can produce mature non-stem cancer cells), increased drug resistance and initiate tumours by their enhanced tumorigenic growth and proliferation (Schatton et al., 2009) (Figure 1.4). Although CSC share some common features with physiological stem cells (SC), it should be noted that a direct relationship between both entities is not required.

The presence of these CSCs in tumours has become a plausible explanation for treatment resistance and relapse, as few CSC would be capable of surviving therapy and re-growing the tumour mass. The first evidence of such a population of cells in cancer was in acute myeloid leukaemia (Bonnet and Dick, 1997), but since then, CSC have been identified in many malignancies, including breast (Al-Hajj et al., 2003), brain (Singh et al., 2004), pancreas (Li et al., 2007), colon (Dalerba et al., 2007), skin (Schatton et al., 2008) or ovaries (Zhang et al., 2008) amongst others.

In some cancers, the CSC population has been directly linked to tissue SC that have acquired a malignant phenotype (Bonnet and Dick, 1997). However, in other cancers, the progenitor or differentiated cells are the ones that have acquired the CSC capacity (Krivtsov et al., 2006). In recent years, a new level of complexity has been identified around CSC: the role of the microenvironment. Increasing evidence

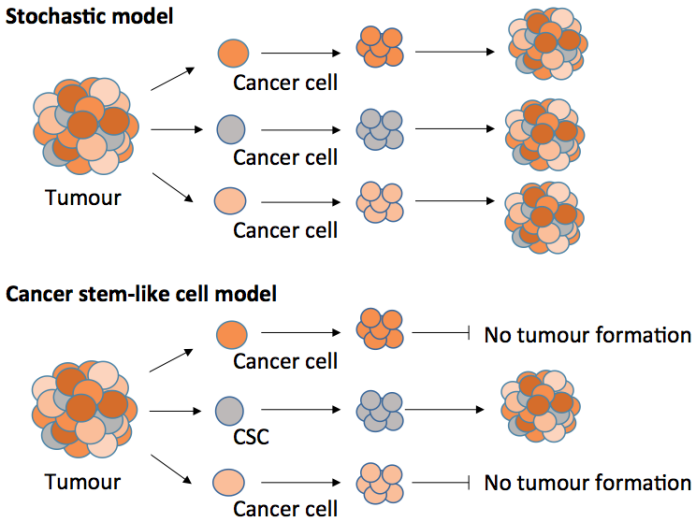


Figure 1.3: Stochastic *vs.* cancer stem-like cell (CSC) model. The stochastic or clonal evolution model (top) explains tumour growth through the acquisition of mutations leading to a selection of cancer cells with selective advantage to grow. Contrary, the CSC model (bottom) suggests that ony a fraction of cancer cells, the CSC (gray) can lead to tumour growth resulting on an heterogeneous population of cancer cells.

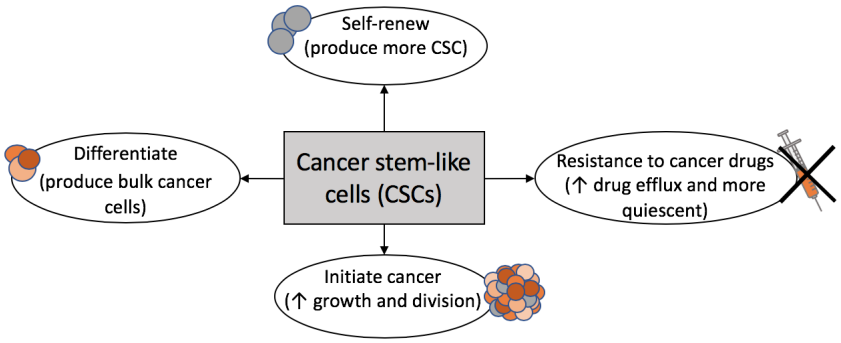


Figure 1.4: Characteristics of CSCs. CSCs are defined by their self-renewing ability, increased resistance to treatment, ability to initiate tumour growth and differentiate into different cell types.

suggests a bidirectional relationship between CSCs and the tumour microenvironment, indicating that both are dependent entities in the tumorigenic process. For instance, glioblastoma CSC induce the secretion of VEGF to promote vasculature in the tumour surroundings (Gilbertson and Rich, 2007), whereas in cutaneous squamous cell carcinoma it is VEGF from the niche that acts on the CSC capacity both through inducing self-renewing ability and modulating the microenvironment (Beck et al., 2011). Moreover, factors secreted by myofibroblasts from the niche can induce a CSC profile to differentiated tumour

cells (Vermeulen et al., 2010). This further emphasizes the importance of the microenvironment in regulating the CSC behaviour and the need to fully understand its implications in cancer progression, as mentioned in Section 1.1.4.

The role that CSC could have in disease progression and patient outcome has led to extensive research in order to elucidate the mechanisms involved in the CSC phenotype. This would not only help in understanding the biology of this cell population, but pave the way towards targeted therapies against CSCs. Although there is variability between CSC from different cancers, some shared characteristics have been defined. Many deregulated signalling pathways have been identified in CSC, including embryonic pathways such as Sex determining region Y-box 2 (SOX2), Octamer-binding transcription factor 4 (OCT4) and NANOG (Herreros-Villanueva et al., 2014) and oncogenic signalling pathways such as PTEN (Ciuffreda et al., 2014). Another important feature of CSCs is their resistance to therapy. A slower cell cycle in CSC has been described, resulting in less efficacy of cytotoxic drugs as these act on fast-dividing cells (Moore et al., 2012); and an increased expression of ATP-binding cassette (ABC) drug efflux transporters (Hirschmann-Jax et al., 2004) or higher aldehyde dehydrogenase (ALDH) activity (Liu and Zheng, 2013), which results in higher efficacy of effluxing drugs and metabolising them respectively. Other CSC features reported are an increase in anti-apoptotic proteins or increased DNA repair mechanisms (Pützer et al., 2017). To all of these, we have to include the intrinsic characteristics of CSCs, which are their increased ability to form tumours in serial transplantation (Baiocchi et al., 2010), their self-renewing potential which can be assessed by their ability to form colonies from a single cell (Pastrana et al., 2011) or their ability to differentiate into other cell types (Sell, 2004). All these characteristics described in CSCs can be exploited to identify these cellular population in tumours with the aim to better understand their phenotype.

Contrary to using intrinsic characteristics of CSCs, some groups have identified cell surface markers that are enriched in CSCs, such as CD133, CD44 or EpCAM (Visvader and Lindeman, 2012). The first isolation of putative CSC based on the expression profile of these markers was in leukaemia, where a population of CD34+ CD38- cells could induce tumorigenesis in leukaemia, in contrast to other

subsets of cells (Lapidot et al., 1994). Based on these findings, studies across different malignancies have tried to identify a cellular marker expression profile that would identify CSC from the rest of cells. However, these resulted in a more complex strategy than anticipated, as the CSCs markers identified could not be used across different cancers and were not exclusively expressed by these cells (Schatton et al., 2009). It is important to mention that the ability of these markers to identify the CSC population varies across cancers, and therefore, their use for identification purposes has to be evaluated on a case-by-case basis. Although these markers might be useful to identify a subset of cells with CSC characteristics, functional assays investigating the intrinsic mechanisms that define CSCs, such as tumorigenesis, ability to self-renew or higher drug resistance, should be employed. Using cellular markers could lead to partial loss of CSCs, as various populations with different cellular profiles might be present in a tumour, and therefore, not detected.

Classic chemotherapeutic strategies (systemic approach) have been designed to eliminate the highly-proliferative cancer cells (bulk population), as they were seen as the tumorigenic core. Based on current knowledge of the CSC model, it has been suggested that this could be the reason behind treatment failure, residual disease and relapse, as CSC would be able to survive treatment and drive tumour regrowth, progression and metastasis (Figure 1.5). Increasing evidence has shown plasticity around the CSC population, where microenvironmental factors and cross-talk between cells can lead to conversion from non-CSC to CSC phenotypes (Cabrera, 2015; Doherty et al., 2016). Therefore, the aim would be to target the CSC population with targeted therapies, and the bulk of cancer cells with conventional approaches to prevent these cells from shifting to CSC (Figure 1.5).

Different strategies to target these cells are being evaluated, based on the different characteristics identified in CSCs. For instance, Gupta et al. (2009) identified a chemical compound (salinomycin) that showed selective toxicity for breast CSC and loss of expression of genes associated to these cells both *in vitro* and *in vivo*, although the exact mechanism of action by which this acts is still unknown (Dewangan et al., 2017). Another approach would be by targeting cell surface markers identified in CSC, such as CD44 in some leukaemia (Jin et al.,

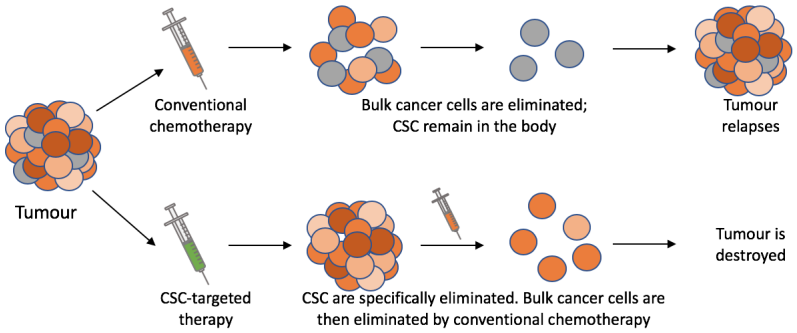


Figure 1.5: Treatment strategy for tumours containing a CSC population. Conventional chemotherapy targets fast dividing cells (bulk of the tumour; orange cells), thus leaving CSCs still in the body (gray cells) which will be able to re-populate the tumour mass and lead to relapses. However, using targeted therapies against CSCs, these slow-growing resistant cells would be eliminated. Then, using conventional chemotherapies, the bulk of the tumour can be eliminated as well, leading to effective removal of the tumour mass.

2006) or CD133 in head and neck cancer (Waldron et al., 2011), although as mentioned above, these could result in a partial-effect if the CSC markers exploited are not expressed by all CSC in the tumour. Similarly, other groups are exploiting the altered signalling pathways, such as by inhibiting the Wnt/β-catenin signalling pathway in hepatic CSC, which had an effect on the self-renewing ability and tumorigenesis both *in vitro* and *in vivo* (Kim et al., 2016). Altogether, we can see that efforts are being put in identifying these cells and findings mechanisms to target them specifically, which will depend highly on the correct isolation method used.

1.2.3 CSC in ES

The low OS and high relapse incidence in ES have been suggested to be due to a population of CSC within ES tumours. ES-CSCs were first identified by isolation of $CD133^+$ cells, which showed higher tumour initiation and growth in xenograft mouse compared to $CD133^-$ cells (Suva et al., 2009). $CD133^+$ cells were highly plastic, as they could generate both a $CD133^+$ and $CD133^-$ cell population, consistent with the CSC intrinsic properties (give rise to different cell types). Later studies from the same group identified a higher expression of stem cell genes in ES-CSC (Riggi et al., 2010) as well as a relationship with disruption of the TARBP2-dependent micro RNA (miRNA) maturation, associated with miR-143 and miR-145 expression (De Vito et al., 2012). TARBP2 is a protein involved in the

stabilization machinery of miRNA, and therefore, its disruption leads to changes in the miRNA levels of important factors for ES phenotype (De Vito et al., 2012). Restoration of TARBP2 activity as well as introduction of miR-143 or miR-145 expression (downstream TARBP2 targets) lead to a reduction of the self-renewing capacity *in vitro* and tumorigenic potential of ES-CSC *in vivo*. In a more recent publication (Cornaz-Buros et al., 2014), a combination of therapies reduced both ES-CSC and the bulk of the tumour in mouse models. For this, doxorubicin was used to target the bulk of the tumour, while enoxacin, an antibacterial agent that has been shown to enhance TARBP2 levels (Melo et al., 2011), was used to specifically target the ES-CSC population as this was found to express TARBP2 (De Vito et al., 2012). Therefore, this combination of therapies to target both the CSC and the bulk of the tumour, as described in Figure 1.5, is what would fit with the suggested strategy to improve outcomes.

However, not all ES literature agrees that $CD133^{+}$ is a good marker for ES-CSCs. For instance, Jiang et al. (2010) showed heterogeneity on the $CD133^{+}$ population across different ES cell lines. The profile associated to CSCs (i.e. sphere formation and increased resistance) was only observed in $CD133^{+}$ cells from one of the cell lines investigated. In the other ES cell lines, even when a $CD133^{+}$ cell population could be isolated, these did not show CSC properties. Therefore, these results indicated that, even though the CD133 cell surface marker is able to identify a subset of cells with CSC characteristics, it might not be a good ES-CSC marker.

Since then, other ES-CSC markers have been identified with the aim to better define and isolate the ES-CSC population. For instance, Awad et al. (2010) identified a population of ES-CSC based on their higher expression of ALDH. These cells were enriched for stem cell genes, had higher colony and sphere formation ability and were more resistant to chemotherapy *in vitro*, as well as higher tumorigenesis *in vivo*, when compared to $ALDH^{low}$ cells. Another group identified ES-CSC based on the CD57 neural crest marker expression (Wahl et al., 2010), showing that $CD57^{high}$ cells have enhanced sphere-forming ability, can invade the extracellular matrix (ECM) more efficiently and have an increased tumorigenicity *in vivo* compared to $CD57^{low}$ cells. In addition, they showed that there was no correlation between CD57 and CD133 expression, further evidence of the

heterogeneity within ES-CSC. However, recent data from Leuchte et al. (2014) suggested that selecting by the sphere formation (as was done initially by Wahl et al.) does not enrich for ES-CSC phenotype, as when comparing spheres with monolayers, similar expression of CD57 and CD133 was detected, as well as no enhancement dye efflux or tumour formation *in vivo*. Finally, the ability to identify ES-CSC based on a higher dye efflux activity (side population) has been shown in ES cell lines. This cell population, when compared to the rest of the cells, showed increased clonogenic capacity and invasion *in vitro* as well as higher resistance to chemotherapies, which was explained by their increased expression of ABC transporters (Komuro et al., 2007; Yang et al., 2010).

These studies are consistent with the hypothesis that ES have a population of cells with higher tumorigenesis, self-renewing ability and increased resistance to conventional chemotherapies compared to the majority of cancer cells (tumour bulk), suggesting that they could be behind relapse and poor outcome for ES patients. However, a universal marker to identify ES-CSC is still missing. Although the different strategies mentioned above lead to the isolation of a cell population with CSC characteristics, the data suggests that we could be observing different subsets of CSCs. Therefore, targeting only one of these subpopulations would not result in the total elimination of ES-CSCs. To specifically target these ES-CSCs with the aim of improving treatment and outcomes, better ES-CSC markers and understanding of the drivers of these cells must be identified, so that new therapies to eradicate these cells can be designed. For this, the use of functional assays, those exploiting intrinsic characteristics of a CSC cell (i.e. self-renewing ability) might provide a better method to identify the overall CSC population, rather than exploiting cell surface markers that might not be shared across all ES-CSCs in a tumour. In addition, the identification of a gene expression profile associated with ES-CSCs might result in a prognostic marker for ES patients, which might result in more efficient patient stratification, personalised medicine and subsequent improved outcomes.

1.3. Exosomes

Extracellular vesicles (EV) comprise an heterogeneous group of vesicles which originate from the plasma membrane or endosome and are released by cells in physiological and pathogenic conditions. Amongst these we can find exosomes, first described by Pan and Johnstone (1983) and further characterised in Johnstone et al. (1987). Initially, they were associated with the mechanisms for waste disposal in cells, but later it was found that they played an important role in intercellular communication (Abels and Breakefield, 2016). Exosomes are characterised by a lipid bilayer membrane (Figure 1.6), have a size of 30 to 150 nm and are cup-shaped in appearance visualised by electron microscopy (Théry et al., 2009). The exosome composition is heterogeneous, depending on the biogenesis process, cell of origin or physiologic conditions, although some common features have been described and help with the identification of exosomes from other EV types.

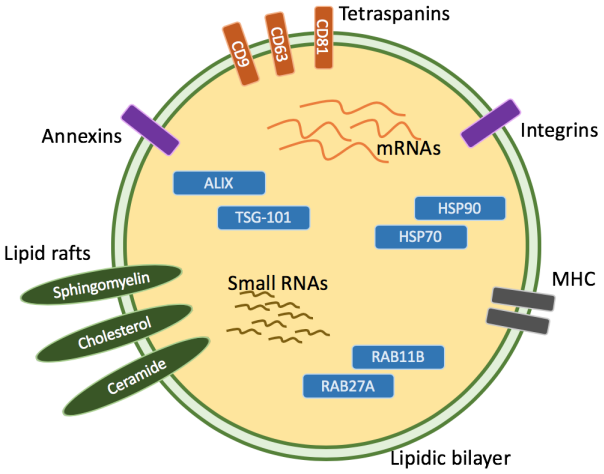


Figure 1.6: Summary of exosomes surface markers and cargo. Exosomes are formed by a lipidic bilayer (green); different proteins, mainly associated to the machinery of exosome biogenesis (rectangular); different nucleic acids such as small RNAs (lines); and different lipids (oval). Abbreviations: MHC = Major Histocompatibility Complex, HSP = Heat Shock Protein.

Proteins involved in the biogenesis of exosomes (Section 1.3.1) are frequently expressed on the exosome surface. These include components of the endosomal sorting complexes required for transport (ESCRT) such as TSG-101 or ALIX (Henne

et al., 2012); tetraspanins such as CD63, CD9 or CD81 (Pols and Klumperman, 2009); or members of the Rab GTPase family, which are involved in the release of exosomes, such as RAB11B or RAB27A (Ostrowski et al., 2010) (Figure 1.6). Reflecting on the process for the generation of exosomes, proteins associated with the golgi, nucleus or endoplasmic reticulum are not generally found in exosomes (Thery et al., 2001).

A diverse variety of genetic material has been identified as part of the exosome cargo, including DNA and RNA. Exosomes are mainly enriched for small RNAs (sRNAs), although functional mRNAs have been described (Valadi et al., 2007; Zhang et al., 2015). Other types of RNAs have been detected such as rRNAs, long and short non-coding RNAs or tRNA fragments (Abels and Breakefield, 2016) (Figure 1.6). The DNA cargo of exosomes is less well studied, but has been detected (Thakur et al., 2014). Exosomes have some lipids that reflect the cell from which they have been derived (Mathivanan et al., 2010), but may contain lipids that have been found to be enriched in exosomes, such as sphingomyelin, cholesterol or ceramides (Llorente et al., 2013) as well as the presence of lipid rafts containing certain proteins (Staubach et al., 2009).

Therefore, characteristics associated with exosomes, such as a specific size range, the presence of proteins associated to exosome biogenesis or an enrichment for sRNA can be exploited to identify exosomes from other EVs.

1.3.1 Biogenesis and function of exosomes

Exosomes are generated through the endosomal pathway (György et al., 2011). The inward vesiculation of the endosome membrane results in the formation of small vesicles, forming multivesicular bodies (MVB; Figure 1.7A). At this point, different proteins, lipids and nucleic acids are sorted into these small vesicles (Figure 1.7B) (Kharaziha et al., 2012), thus acquiring the characteristics of what defines an exosome. Different mechanisms have been identified to play a role in this process, such as tetraspanins (Pols and Klumperman, 2009) and the components of ESCRT (Henne et al., 2012), although recent studies have identified ESCRT-independent pathways for the biogenesis of exosomes, such as the sydencan-syntetin-ALIX pathway (Baietti et al., 2012) or through the raft-based

microdomains highly enriched for the sphingolipid cermaides, that lead to the segregation of the cargo into endosomes (Trajkovic et al., 2008).

Once exosomes have been accumulated into MVBs, these can fuse with the lysosome for degradation or can release their cargo (exosomes) into the extracellular space by fusion with the plasma membrane (Figure 1.7C) (Kharaziha et al., 2012). The family of Rab GTAases has been associated with this process of exosome release, where the involvement of different Rab proteins leads to exosomes displaying a different protein profile (Laulagnier et al., 2004; Ostrowski et al., 2010).

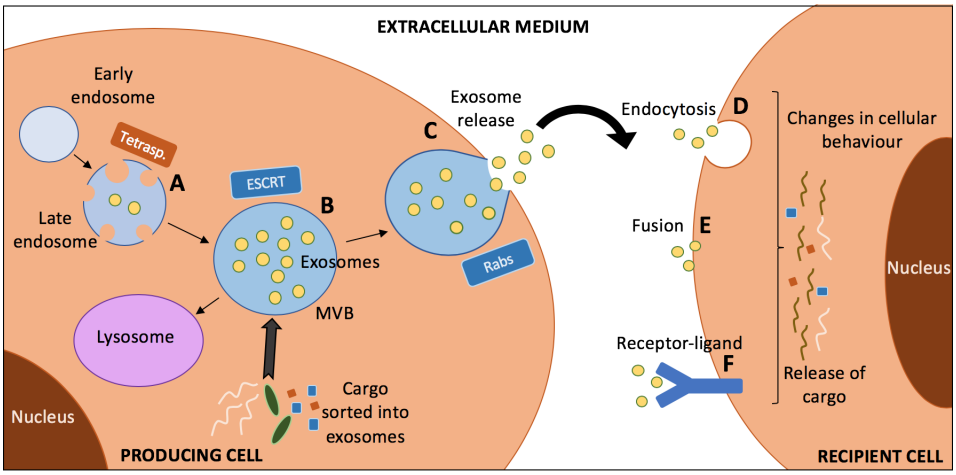


Figure 1.7: xosomes biogenesis, release and upkate by cells. Exosomes are produced through the inward budding of the late endosome membrane (**A**), at which point it becomes a multivesicular bodies (MVB) filled exosomes (**B**). At this stage, different cellular components (proteins, RNAs and lipids) are sorted into the exosomes, defining their profile. Then, MVBs can either go to recycling (lysosme) or can fuse with the plasma mebrane releasing exosomes to the extracellular medium (**C**). Once exosome are in circulation, they can be taken up by recipient cells either by endocytosis (**D**), fusion with plasma membrane (**E**) or through a receptor-ligand interaction (**F**). All these mechanisms lead to release of their cargo or activation of signalling pathways, thus resulting in changes in recipient cells.

Once exosomes are released into the extracellular space and into circulation, other cells will need to display mechanisms for the internalisation of these vesicles. Exosomes can be internalized into recipient cells by endocytosis (Figure 1.7D), for which different mechanisms have been described, including caveolin- (Svensson et al., 2013) and clathrin- (Tian et al., 2014) mediated endocytosis. Another proposed mechanism is through the fusion of exosomes with the plasma membrane (Figure 1.7E), resulting in the direct release of the cargo into cells

(Mulcahy et al., 2014), where a low pH in these has been described to facilitate this process (Parolini et al., 2009). A third mechanism proposed is the presence of specific proteins on the surface of exosomes that can interact with receptors on recipient cells (Figure 1.7F), thus inducing a response (Munich et al., 2012).

Through these processes, exosomes exchange information between the cells from which they are produced into recipient cells, thus playing an important role in cell-to-cell communication. Exosomes have important roles in the maintenance of normal physiological processes. They play a role in neurotransmission signalling in neurons, where these cells can increase the release of exosomes filled with neurotransmitters (Lachenal et al., 2011). They have been involved in the immune response, where they have been linked to activation or suppression of adaptive immune responses depending on the cellular status (Théry et al., 2009). Similarly, exosomes have been associated with tissue repair (Gatti et al., 2011) or stem cell maintenance (Ratajczak et al., 2006). Exosomes also play important roles in many different pathological processes, such as in HIV (Dias et al., 2018), neurodegenerative diseases (Jan et al., 2017) or cancer (covered in Section 1.3.2).

Therefore, cells are exploiting this ability of cross-talk through the transfer of exosomes to modulate the microenvironment and lead to favourable conditions for different pathological states. This suggests the importance of fully understanding the cargo associated to exosomes and the implications it might have when taken up by other cells.

1.3.2 Exosomes and cancer

Based on the intrinsic characteristics of exosomes, these vesicles have become a unique entity not only for the study of cancer biology, but also for their clinical and targeted therapy applications. A deeper understanding of how cancer cells could be exploiting exosomes to promote cancer progression or modulation of the tumour microenviroment might provide answers on cancer biological processes. Moreover, the stability of exosomes in circulation due to the lipid bilayer and the uptake by recipient cells make them appealing candidates for biomaker studies and targeted therapy strategies.

Different studies have reported the presence of oncoproteins in the cargo of

exosomes, suggesting that the delivery of these to non-cancer cells could lead to a transformation towards an oncogenic state. For instance, HER2 has been detected in ovarian and breast cancer-derived exosomes and MART1 in melanoma-derived exosomes (Andre et al., 2002), both oncoproteins important in the development of these diseases. Moreover, mutant KRAS was transferred from colon cancer cells to non-cancer cells, increasing the growth of the latter (Demory Beckler et al., 2013). Similar to the transfer of oncoproteins, onco-RNAs (i.e. miRNAs and mRNAs) have been identified in the transfer process between exosomes and recipient cells. The first identification of RNA transfer via exosomes was described by Valadi et al. (2007), where mRNAs contained in exosomes from human and mouse mast cell lines could be transferred to recipient cells and translated into protein, thus indicating that mRNAs contained in exosomes are functional. Since then, different groups have characterised the RNA content of exosomes in cancer cells and investigated if this results in changes in the profile of recipient cells. Exosomes from cancer associated fibroblasts (CAF) in pancreatic adenocarcinoma could release miR-146a into epithelial cells, thus inducing an increase in proliferation and survival upon treatment with gemcitabine, the chemotherapy used to treat this cancer (Richards et al., 2017). Exosomes from gastric cancer cells with high expression of ZFAS1, a long non-coding RNA associated to tumour growth and metastasis, could be internalised by gastric cells with low levels of ZFAS1 and promote increased cell growth and migration capacity (Pan et al., 2017). Finally, some groups have investigated the ability of drug resistant cells to induce an increase in resistance towards chemotherapeutic drugs via the transfer of exosomes. Exosomes from doxorubicin-resistant osteosarcoma (OST) cells could induce a resistant profile in sensitive OST cells, which was accompanied by an increase in the expression of drug efflux proteins in recipient cells by the transfer of their mRNAs via exosomes (Torreggiani et al., 2016). Therefore, cancer cells can induce changes in neighbour cells by the delivery of specific cancer-related proteins and RNAs, thus suggesting an important mechanism of cross-talk and tumour progression.

The tumour microenvironment plays an important role in cancer progression and metastasis (covered for ES in Section 1.1.4). Due to the ability of exosomes to

modulate the cellular behaviour of recipient cells, different groups have investigated if exosomes from cancer cells can modulate the tumour microenvironment, thus creating an extra layer of complexity in this process. Exosomes from glioblastoma cells under hypoxic conditions could induce angiogenesis and tumour growth when injected in xenograft mouse models compared to exosomes from cells under non-hypoxic conditions (Kucharzewska et al., 2013). Moreover, exosomes from prostate cancers induced a differentiation of MSCs towards myofibroblasts, which then led to release of pro-angiogenic factors and induction of tumour proliferation and invasion (Chowdhury et al., 2015), which was found to be modulated by the presence of TGF-β in exosomes. The implications of exosomes in the modulation of immune responses have been widely studied, as this could inform both on new mechanisms for tumour progression, but ultimately could be exploited for novel treatment strategies (Théry et al., 2009). The cargo of cancer-derived exosomes have been found to help in the suppression of the anti-tumour immune response, either by the direct contact with immune cells, or by the modification of non-immune cells in the tumour microenvironment that would ultimately lead to a change in the immune response (Whiteside, 2016). Exosomes from melanoma cells alter the differentiation of myeloid cells toward an immunosuppressive state, compared to exosomes from normal cells (Valenti et al., 2006).

Increasing evidence demonstrates that tumour-derived exosomes have an important role in the metastatic process, especially in the formation of the pre-metastatic niche. For tumour cells to be able to invade distant sites of the body, a series of events are needed which include invasion, evasion of the immune response, entering the circulation and exiting at distant organs to settle there. Exosomes have been found to play a role in this process, facilitating the creation of a pre-metastatic niche for tumour cells (Tickner et al., 2014). For instance, melanoma-derived exosomes injected in mice are seen to be captured by tissue in different sites, such as lung, liver or kidney, common metastatic sites for melanoma (Takahashi et al., 2013). Another study demonstrated that exosomes derived from CSC in renal cancer, when injected in mice, could induce a pro-metastatic niche compared to exosomes from the overall renal cancer cell

populations, which was explained by an increase level of RNAs associated to metastasis in the exosomal cargo (Grange et al., 2011).

Since exosomes play an important role in cancer progression through the transfer of their cargo, this has become an interesting area for the development of novel tools such as companion diagnostics and therapeutic strategies.

1.3.3 Therapeutic applications of exosomes in cancer

As discussed above (Section 1.3.2), cancer-derived exosomes contain cargo that relates to the cancer they are derived from, as different oncoproteins and onco-RNAs have been found as part of the cargo. As exosomes are secreated in circulation, where the cargo is protected from exogenous factors by the lipidic biolayer, the ability to identify cancer-specific markers in them has become an interesting opportunity for non-invasive biomarker strategies. A panel of 8 miRNAs that have been shown to have diagnostic potential in ovarian cancer were found in exosomes from the plasma of ovarian cancer patients at similar expression levels than in ovarian tumours, and could discriminate between benign and malignant tumours (Taylor and Gercel-Taylor, 2008). A similar result was obtained in B-cell chronic lymphocytic leukaemia (CLL), where exosomes expressed a panel of proteins identified in the cells from which they were derived, being able to discriminate between patients with CLL and healthy donors (Belov et al., 2016). Therefore, these findings suggests the ability to use exosomes as non-invasive circulating biomarkers, as these show a profile associated to the cancer cells and can discriminate from non-cancer or benign tumours. Moreover, they could be exploited for disease follow-up in order to assess the tumour burden throughout treatment. However, due to the heterogeneity of profiles often associated to cancer, it is important to use well defined cancer-related markers that can identify all patients, or use more conventional methodologies to confirm the results.

Since exosomes play an important role in cancer progression, they provide a potential therapeutic inhibitory target. Exosomes might be exploited therapeutically by inhibiting the production of exosomes by cancer cells. This might be achieved through the inhibition of the neural sphingomyelinase 2 (n-SMase2), a component of the sphingomyelin pathway that has been linked to the production of exosomes (Trajkovic et al., 2008). Using the n-SMase2 inhibitor

GW4869 in mice, Fabbri et al. (2012) saw a reduction of exosome production, leading to decrease in the lung metastatic sites. However, using this approach leads to two major caveats. On the one hand, there is still uncertainty on the different mechanisms behind exosome production and release, where differences across cells have been identified. In prostate cancer, GW4869 does not have an effect on exosome production (Phuyal et al., 2014), and therefore, other mechanisms would need to be identified. On the other hand, exosomes are important for physiological cells, and the different components that have been associated to exosome biogenesis are expressed in other cell types (El Andaloussi et al., 2013). Therefore, inhibiting such processes might lead to important off-target toxicities.

Alternatively, exosomes might be exploited to treat cancer cells via targeting specific exosome cargo associated to the cancer phenotype. For instance, melanoma-derived exosomes can induce a pro-metastatic status in bone marrow cells by the transfer of the oncoprotein MET. When the levels of MET were reduced in exosomes, a decrease in the ability to induce metastatic disease in mice was seen (Peinado et al., 2013). This approach might provide a more accurate strategy, as the reduction of oncoproteins in cancer cells might have an effect on the targeted cancer cells and the cells from the microenvironment to which exosomes would induce an effect. However, for these targeted strategies, the oncoproteins need first to be identified and targeted or inhibited in order to induce the change in the exosome cargo.

Exosomes can modulate the immune system, being able to activate or suppress immune responses (Théry et al., 2009). Therefore, they are being evaluated as cancer vaccines for their ability to present antigens through their membrane, thus triggering immune responses (Sun et al., 2018). For instance, dendritic cell-derived exosomes can induce an immunogenic state in breast cancer recipient cells, showing an increase in the activation of T-cells (Romagnoli et al., 2015), suggesting the possibility to use exosomes for immunotherapies. Following this strategy, several groups are evaluating the effect of exosomes in a clinical setting (Wang et al., 2017). As an example, dendritic cell-derived exosomes have been used alongside standard chemotherapy as a maintenance immunotherapy in non-small

cell lung cancer (NSCLC), where data from a phase II clinical trial has shown efficacy in inducing antitumour effects through activation of the natural killer cell process (Besse et al., 2016). The ability to use exosomes alongside standard chemotherapies, exploiting their ability to modulate cancer and microenvironment cells, suggests an exciting opportunity for treatment strategies. However, given the uncertainty around the biogenesis and transfer of exosomes to other cells, it is important to fully understand the implications this approach might have on cancer progression.

Exosomes might also be exploited in a clinical setting as vectors for drug delivery. Exosomes could be used to deliver chemotherapeutic drugs as part of the exosome cargo to recipient cells. In this line, exosomes have been shown to deliver doxorubicin and reduce the tumour size more efficiently than using free drug (Jang et al., 2013). Moreover, using doxorubicin-loaded exosomes, higher doses of the doxorubicin can be delivered to the tumour, which resulted in lower toxicities and a more potent anti-tumour effect in mice (Hadla et al., 2016).

Alternatively, specific miRNAs targeting drug-resistance pathways could be delivered to recipient cells, being able to induce sensitivity towards conventional chemotherapies (Li et al., 2015). Another promising role of exosomes for therapeutic strategies is via the direct modification of their surface and cargo (exosome engineering), thus being able to specifically express those markers that might induce an effect on recipient cells (Gilligan and Dwyer, 2017). For instance, TNF-related apoptosis-inducing ligand (TRAIL)-positive exosomes obtained from transduced leukemic cells could induce a necrotic response in *in vivo* myeloma and melanoma mouse models compared to mock exosomes (Rivoltini et al., 2016). However, despite clear necrosis and anti-angiogenesis signs were seen after $TRAIL^{+}$ exosome treatment, only partial responses were seen. As only a fraction of exosomes reached the tumour site, many going to other organs, could explain these absence of response (Rivoltini et al., 2016). Therefore, although the technique of using exosomes for targeted therapies looks promising, it is important to be sure that exosomes will induce an anti-tumour effect and that they will reach the tumour site, as off-target effects might lead to important toxicities.

All this evidence postulates exosomes as an exciting approach for novel therapeutic strategies. Further understanding of their role in cellular cross-talk and modulation of the microenvironment will help define strategies to target specific cells or induce changes that could help in the treatment of cancer and other diseases.

1.3.4 Exosomes in ES

Four different studies have profiled and evaluated the role of ES-derived exosomes as candidate biomarkers and their implications in tumour progression. Miller et al. (2013) were the first to evaluate ES-derived exosomes, showing that exosomes released from ES cell lines express 12 mRNAs associated with ES, including EWSR1-FLI1 fusion and previously-defined ES markers EZH2, NKX2.2 or STEAP1. From these, a panel of 5 mRNAs showed sensitivity and specificity in detecting ES-derived exosomes spiked in plasma of healthy donors, suggesting their application as biomarker for minimal residual disease. The detection of EWSR1-FLI1 fusion in ES-derived exosomes was soon confirmed (Tsugita et al., 2013), as well as the ability to detect ES-derived exosomes bearing the fusion in blood from ES mouse xenografts. Confirming previous findings (Miller et al., 2013), ES-derived exosomes from a tissue engineered 3D model contained the EZH2 mRNA (Villasante et al., 2016). Moreover, CD99 protein has been identified in exosomes derived from ES cells (Ventura et al., 2016). These findings suggest the ability to use ES-derived exosomes for biomarker studies and evaluation of residual disease, as specific ES-associated RNAs can be detected in circulation. However, these studies need to be validated in patient samples, evaluating the specificity of these profiles and the yield of ES-exosomes in circulation.

As well as exploiting ES-derived exosomes for biomarker studies, some of these groups have evaluated the role these might play in ES progression. The EWSR1-FLI1 mRNA was found to be transferred to other ES cells, inducing higher expression levels of the fusion in these cells; however, this uptake was not seen by OST cells (Tsugita et al., 2013). Similarly, the EZH2 mRNA was taken up by MSCs co-cultured with ES cells, inducing an increased level of expression for this protein (Villasante et al., 2016). In contrast, upon incubation with ES exosomes, osteoclasts showed a decrease in the EZH2 mRNA expression, whereas osteoblasts did not show an effect

(Villasante et al., 2016). Therefore, exosomes from ES cells can induce changes in other cell types, although difference in the uptake efficacy are seen across cells.

Further confirming this role of ES exosomes in tumour progression, Ventura et al. (2016) have evaluated the role that ES-derived exosomes may have in the cellular behaviour of ES recipient cells. By silencing CD99 expression in ES cells, the released exosomes had an increased expression of miR-34a, which has been shown to induce a neural differentiation in ES cells. When exosomes from CD99-silenced cells were taken up by normal ES cells, this led to the inhibition of the nuclear factor-κB (NF-κB) signalling through miR-34a repression of the Notch pathway, which was enough to induce neural differentiation in recipient cells. Therefore, the transfer of exosomal cargo from CD99-silenced to normal ES cells was enough to induce the same profile as silencing CD99 or over-expressing miR-34a in cells (Ventura et al., 2016).

Therefore, the data from these 4 publications suggests that ES-derived exosomes could be exploited for biomarker studies due to the presence of CD99 and EWSR1 fusion in their cargo, as well as ES-associated mRNAs expressed in cells. In addition, ES-derived exosomes can be taken up by non-ES cells and induce changes in their gene expression profile or behaviour. Therefore, further understanding of their role in ES tumorigenesis and the implications of their cargo might result in novel prognostic and diagnostic biomarker strategies for ES, as well as a better understanding on disease progression and relapse.

1.4. Hypothesis and aims

The hypothesis of my thesis is that identification and characterisation of ES-CSCs is essential for the development of novel targeted therapies to eradicate these cells and ultimately lead to improved outcomes for ES patients. For this, using a functional assay to isolate ES-CSCs from patient-derived ES samples will inform better on the heterogeneity of these cells across patients and their intrinsic profile. Moreover, since exosomes have been associated with cancer progression, the characterisation of ES-derived exosomes might provide further understanding on ES tumorigenesis.

The aims of my research have therefore been to:

1. Study the *in vitro* phenotype of patient-derived ES cell cultures and their correlation to the clinical data (Chapter 2).
2. Isolate ES-CSCs from patient-derived cell cultures using a functional assay (Chapter 2).
3. Examine the profile of ES-CSCs derived from patient samples using *in vitro* assays and RNA-seq gene expression profile (Chapter 3).
4. Develop a pipeline for the analysis of RNA-seq data from ES patient-derived cell cultures (Chapter 3).
5. Optimise a method to isolate and characterise ES-derived exosomes from cell culture supernatant and plasma (Chapter 4).
6. Investigate the impact on non-ES cells after uptake of ES-derived exosomes (Chapter 4).

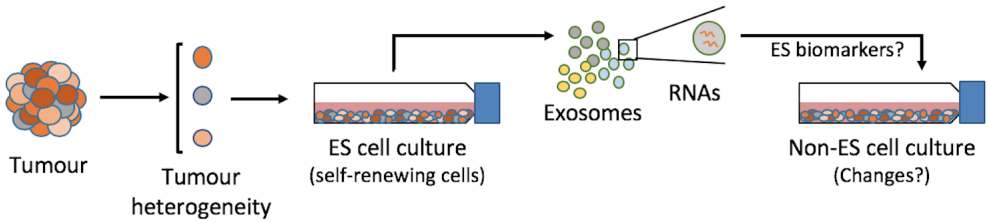


Figure 1.8: Summary of thesis hypothesis. The evaluation of patient-derived ES cell cultures will allow the study of tumour heterogeneity and profiling ES-CSCs. The isolation of ES-derived exosomes and their impact when taken up by non-ES cells might provide new insights into ES tumorigenesis.

2. Profiling of ES patient-derived cell cultures

2.1. Introduction

Despite improvements in diagnosis and treatment over the last decades, patients diagnosed with metastatic ES (25 % of cases) have a 5-year event free survival (EFS) rate which still remains unacceptably low (50 % for patients with isolated lung metastasis; < 20 % for patients with other metastatic sites) (Esiashvili et al., 2008). In addition, a considerable proportion of ES patients are refractory to conventional treatments (25 % with localised disease; 75 % with metastasis) and another significant proportion may initially benefit from cytotoxic therapy but will eventually recur (EFS at 5 years of 10 %) (Rodriguez-Galindo et al., 2003) (Section 1.1.1).

In order to better understand ES tumorigenesis process and how best to target the ES-CSCs, it is important to carry out the investigations in biological material that best informs on the patient's biology. Much research is currently based on the use of established cell lines, which have arisen from human tumours and are useful preclinical tools due to their easy handling and the ability to obtain high quantities of cells. However, their continued growth in culture has resulted in adaptation and frequently loss of normal cell cycle regulation, as well as an accumulation of abnormal metabolic characteristics (Stacey, 2005). Improvement in tissue culture techniques have allowed the maintenance of patient-derived cells in culture, conserving the tumour cell phenotype *in vitro*. Therefore, the ability to obtain primary cultures from fresh tissue has allowed more informative research into the biology of human cancers, as these cell cultures will display a phenotype and genotype more similar to that of the patient they are derived from. Although the variability of cells forming a tumour (intra-tumour heterogeneity) and the different profiles across samples with the same diagnosis (inter-tumour heterogeneity) was initially considered a disadvantage for research, it is now considered a unique resource of information for the study of tumour heterogeneity and patient stratification.

Different biological traits have been associated with the cancer phenotype (Hanahan and Weinberg, 2000, 2011). These include the evasion of apoptosis, sustainability of cell proliferation, evasion of growth suppression, invasion and metastasis, replicative immortality, sustained angiogenesis, as well as, genome instability, inflammation, energy metabolism reprogramming and evasion of immune destruction. Therefore, cancer research is often based around the profiling of these characteristics in cancer cells, in order to better understand their role in tumorigenesis with an aim to define better treatment strategies or exploit these markers for prognostic purposes.

CSCs are a small population of cells within the tumour mass responsible for tumour growth, drug resistance and relapse (Schatton et al., 2009). Emerging evidence suggests that ES-CSCs are the cells behind poor outcome in ES patients, as these cells would be able to survive treatment, leading to tumour growth and spread, eventually leading to relapse (Awad et al., 2010; Cornaz-Buros et al., 2014; Suva et al., 2009; Wahl et al., 2010) (Section 1.2). One of the intrinsic characteristics of CSCs is their ability to self-renew, and therefore, evaluation of this trait in the patient-derived ES cultures would inform on the CSC profile of these samples.

A high proportion of ES patients (25 %) have metastatic disease at diagnosis, being an important factor for the patients prognosis. Therefore, it is important to evaluate the capacity of ES cells to migrate and invade into other areas, as these characteristics have been associated to the ability of tumour cells to metastasise (Paul et al., 2017; Yamaguchi et al., 2005). This can be partially assessed *in vitro* by evaluating the ability of these cells to migrate (Kramer et al., 2013). Moreover, the expression of CD99 has been associated with the migration capacity of ES cells (Rocchi and Manara, 2010) and in other cancers (Seol et al., 2012). Therefore, it is important to evaluate if the migration index of ES cells correlates with CD99 expression and if these are predictors of outcome.

Another important factor associated with patient prognosis is resistance to chemotherapies, thus resulting in failure of treatment and relapse. To overcome this, efforts are being put into identifying signalling pathways and molecular factors that could be important for this cancer. In ES, this is especially important, as a high proportion of patients are refractory to treatment or will eventually relapse.

Therefore, new insights into the mechanisms used by ES cells to progress could ultimately lead to the identification of new targeted therapies for the treatment of ES. Amongst the different mechanisms evaluated in ES, we can find the tyrosine kinase c-kit, implicated in different malignancies (Ashman and Griffith, 2012; Conic et al., 2015; Zhao et al., 2014). This was shown to have high expression in ES tumours (Scotlandi et al., 2003), although *in vitro* evaluation (Scotlandi et al., 2003) and in a phase II clinical trial (Bond et al., 2008) have shown that inhibiting this protein has not proven advantageous for ES patients when used as single agent. However, it could be that this enhanced expression of c-kit in ES cells could be exploited for combined treatment strategies or used for diagnosis purposes.

Therefore, using ES patient-derived cell cultures and evaluating their phenotype regarding different hallmarks of cancer would help in acquiring a deeper knowledge on the phenotype and genotype of ES biology. In addition, correlating the results with the clinical data of the patients might lead to the identification of a signature associated to treatment failure and relapse, which could result in novel prognostic markers.

2.1.1 Aims of this chapter

The aims of this chapter are to:

1. Propagate and maintain in culture patient-derived ES cells from ES patient tissues.
2. Identify the ES diagnostic biomarkers (CD99 and EWSR1 fusion) in ES patient-derived cultures.
3. Evaluate the phenotypic profile of ES cell cultures based on different cancer-related *in vitro* assays (self-renewing ability, migration and chemoresistance) as well as the expression of c-kit.
4. Perform survival analyses to investigate any correlation between phenotypic characteristics and clinical data to identify prognostic indicators.

2.2. Materials and Methods

2.2.1 Clinical samples and ethics

ES samples (portion of tissue and/or bone) were received by Royal Mail at room temperature (RT) in 10 ml of Leeds Antibiotic Media (LAM; RPMI-1640 media (Sigma-Aldrich Company Ltd, Dorset, UK) containing 10 % foetal calf serum (FCS; Labtech International Ltd, Ringmer, UK), 2 mM L-glutamine (Sigma-Aldrich), 0.1 mg/ml streptomycin and 100 units penicillin (Sigma-Aldrich)) from the UK centres involved in the GenoEwing project. Samples were received at diagnosis, resection and relapse time points. All samples had clinical consent at the time of surgery (IRAS 167880). Samples diagnosed as non-ES were kept but clinical data was not obtained. All samples were pseudo anonymised on receipt and recorded in the Achiever tissue tracking software in compliance with the Human Tissue Act. Historical ES samples were obtained through CCLG biological study number 010 BS 03.

2.2.2 Sample processing

Samples were processed on day of reception. Following centrifugation at 402 g for 5 min (Eppendorf centrifuge 5804; Eppendorf UK Ltd, Stevenage, UK) media was aspirated. All tissue was placed into a Petri dish. The bone (when present) and the soft tissue were separated with a scalpel and most of the soft tissue was macerated and placed in 2.5 ml of LAM media into a 25 cm^2 tissue culture flasks (T25, Croning Primaria™, Falcon, BD Biosciences, Oxford, UK). Cells were incubated at 37 °C in 5 % CO_2 in air (Sanyo CO2 Incubator, MCO-2OAIC; Sanyo Gallenkamp Plc., Loughborough, UK). A small proportion of the tissue was kept apart (not macerated) and frozen at time of collection. For this, tissue was placed on top of a thin layer of Optimal Cutting Temperature compound (OCT; VWR International Ltd, Lutterworth, UK) on a metal block that had been previously cooled on dry ice. Once the OCT was almost solidified, the tissue was completely covered by adding more OCT until completely solidified. Then, the frozen tissue was detached from the metallic block and placed in a cassette properly labelled to be stored at −80 °C. Bone tissue received was digested by incubation in 2 ml of 1:1 1 × Trypsin (10% Trypsin (Sigma-Aldrich) v/v in 1 × phosphate buffered saline (PBS; Sigma-Aldrich)) and ethylenediaminetetraacetic acid (EDTA, 0.1 % w/v in PBS (Fisher Scientific;

Thermo Fisher Scientific, Warrington, UK)) for 30 min at 37 °C with vigorous shaking every 10 min. This was then centrifuged to pellet any cells and bone tissue, the media aspirated and the pellet was resuspended in 5 ml of fresh media and placed in a T25 tissue culture flask and incubated (Figure 2.1).

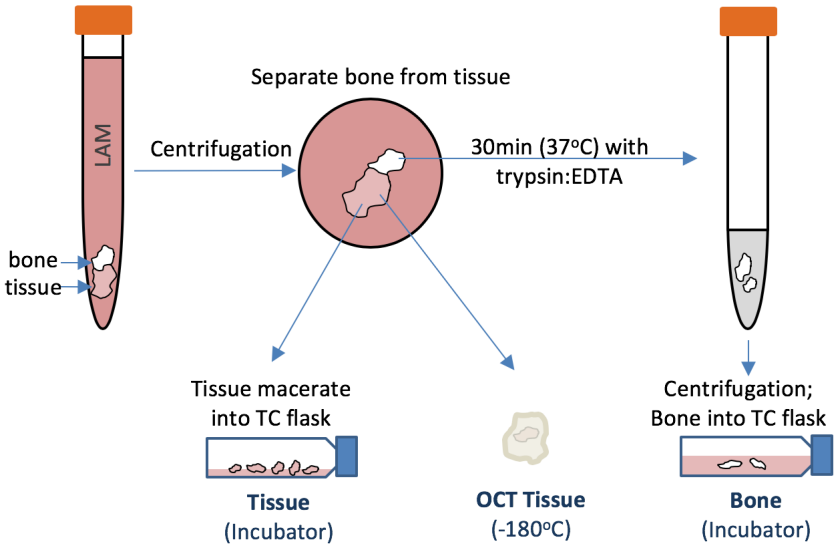


Figure 2.1: GenoEwing sample processing protocol. Sample (tissue and/or bone) received in media and centrifuged to pellet all cells and tissue. Tissue is macerated and placed in fresh media and/or frozen in OCT and stored at −180 °C. Bone is digested in trypsin and EDTA (1:1) before incubation in fresh media.

2.2.3 Tissue culture

All ES patient-derived samples and established cell lines (Table 2.1) were cultured on Primaria™ plastic 75 cm tissue culture flasks (T75; Croning Primaria™, Falcon, BD Biosciences). Media was aspirated and cells were washed with 5 ml of PBS, followed by a 2 min incubation with 5 ml EDTA at RT. After aspiration of EDTA, flasks were incubated with 5 ml of trypsin for 2 min, or until cells had completely detached from the plastic. Trypsin was neutralised with an equal volume of media, centrifuged at 402 g for 5 min (Eppendorf centrifuge 5804; Eppendorf), the supernatant aspirated and the resulting pellet resuspended in the appropriate media (Table 2.1). Cells were reseeded in T75 tissue culture flasks containing a total volume of 15 ml of media and maintained in culture at 37 °C in 5 % CO_2 in air. All cultured samples were screened for mycoplasma contamination using the EZ-PCR mycoplasma test kit (Geneflow Ltd, Staffordshire, UK).

Cell line	Fusion status	Cellular origin (source)	Media
A673	EWSR1-FLI1 Type I	ES. 15yr F (ATCC)	DMEM + 10 % FCS + 2 mM Glut.
RD-ES	EWSR1-FLI1 Type II	ES. 19yr M (ATCC)	RPMI + 10 % FCS + 2 mM Glut.
SK-ES-1	EWSR1-FLI1 Type II	ES. 18yr M (ATCC)	McCoy's 5 + 15 % FCS + 2 mM Glut.
SK-N-MC	EWSR1-FLI1 Type I	ES. 14yr F (ATCC)	DMEM F12 + 10 % FCS
SK-N-MC EC50[1]	EWSR1-FLI1 Type I	– (produced in house)	DMEM F12 + 10 % FCS + 28 nM doxorubicin
TC-32	EWSR1-FLI1 Type I	ES. 17yr F (Gift: Dr. Toretsky)	RPMI + 10 % FCS + 2 mM Glut.
TTC-466	EWSR1-ERG	ES. Metast. (Gift: Dr. Sorenson)	RPMI + 10 % FCS + 10 % condit. media + 2 mM Glut.
HOS	—	OST. 13yr F (ATCC)	RPMI + 10 % FCS + 2 mM Glut.
BE(2)C	—	NBL. 2yr M (ATCC)	DMEM + 10 % FCS + 2 mM Glut.
NB69	—	NBL. 16mo M (Gift: Prof. Haber, CCIA)	RPMI + 10 % FCS + 2 mM Glut.
SK-N-SH	—	NBL. 4yr F (ATCC)	DMEM:MEM (1:1) + 10% FCS + 2 mM Glut.
IPOO6B10[2]	—	MSC. (Gift: Dr. Jones)	StemMACS MSC Expansion Media + StemMACS supplement

Table 2.1: Characteristics and conditions of established cell lines used. Details of sample type, fusion status (only for ES samples), origin and source of the cell line and media required for culture are provided. All cell lines are of human origin. [1] Generated from normal SK-N-MC cells, with increasing doses of doxorubicin until reaching the EC50 value (28 nM concentration in the growth medium). [2] IPOO6B10 MSC cells will be referred as MSC throughout the thesis No information regarding patient or site available. Abbreviations: ES = Ewing's sarcoma, OST = osteosarcoma, NBL = neuroblastoma, yr = year old, mo = month old, M = male, F = female, Metast. = metastatic sample, ATCC = American Type Culture Collection, CCIA = Children's Cancer Institute (Sydney, Australia), DMEM = Dulbecco's Modified Eagle Medium (Sigma-Aldrich), MEM = Minimum Essential Medium (Sigma-Aldrich), Glut. = glutamine. StemMACS media from MACS (Miltenyi Biotec Ltd, Watford, UK).

2.2.4 Pelleting of cells

After harvesting (Section 2.2.3), cells were centrifuged at 402 g for 5 min. Media was aspirated and cell pellet resuspended in 1 ml PBS. Cell suspension was placed in a 1.5 ml centrifugation tube and centrifuged at 160 g for 5 min (Micro Centaur MSE, Sanyo Gallenkamp). PBS was then aspirated and pellet cells were stored at −80 °C.

2.2.5 Determination of viable cell number

Viable cell number was determined using the trypan blue exclusion assay. After harvesting (Section 2.2.3), an aliquot of the cell suspension (10 μl) was mixed with an equal volume of trypan blue dye (0.4 %, Sigma-Aldrich) and using a Neubauer haemocytometer the viable cell number was assessed. Due to the diffusion of trypan blue across permeable cell membranes, non-viable cells become blue. The number of viable cells (with intact membranes) was counted in 4 areas of the grid, the mean calculated and result multiplied by $\times 10^4$ cells/ml.

When needed, the automated Vi-cell (Beckman Coulter Ltd, High Wycombe, UK) viable cell counter was used. Following the same principle as described above to measure viable cells, this machine allows for an automated process. For this, 0.5 ml of cell suspension is taken by the machine, mixed with an equal volume of trypan blue and the viable cell number ($\times 10^6$ cells/ml) is calculated.

2.2.6 Reverse Transcription$^{\text{Amplification}}$ - Polymerase Chain Reaction (RT$^{\text{A}}$-PCR)

2.2.6.1 RNA extraction

Total RNA was extracted using the miRNeasy Micro Kit (Qiagen Ltd, Crawley, UK) following manufacturers instructions. Briefly, Qiazol (700 μL) was added to the cell pellet and disrupted by pipetting up and down and incubated for 5 min in order to lysate cells. Chloroform (Merck Millipore, Watford, UK) was added (140 μL), shaken vigorously for 15 s and incubated for 3 min to allow separation of protein (bottom), DNA (interphase) and RNA (aqueous phase).The mix was then centrifuged at 12,000 g for 15 min at 4 °C (Sigma 1-14; Sigma-Aldrich) and the upper aqueous phase (RNA) was kept and mixed with a 1.5 × volume of 100 % ethanol (approximately 700 μL) and mixed thoroughly by pipetting. The sample was then

passed through the supplied spin column and centrifuged (8,000 g, 1 min), the flow through discarded and a first wash with 350 μL RWT buffer done (8,000 g 1 min). At this stage, a DNaseI digestion was performed in order to digest any possible DNA contamination. For this, 80 μL of DNaseI diluted in RDD buffer (1:8 DNase in RDD buffer; RNase-free DNase, Qiagen) and incubated for 15 min at RT. Then, the column was further washed with RWT buffer (350 μL, 8,000 g, 1 min), RPE buffer (500 μL, 8,000 g 1 min) and 80 % ethanol (Merck-Millipore; 500 μL, 8,000 g 2 min). Finally, the column was centrifuged at 13,000 g for 1 min to remove residual flow through, moved to a new collection tube and 12 μL RNase-free H_2O was directly added to the membrane and centrifuged at 13,000 g for 1 min to eluate the RNA. RNA was stored at −80 °C.

2.2.6.2 Spectophotometric measurment of nucleic acid quantity and quality

The RNA quality and quantity was assessed by Nanodrop (Labtech International Ltd.). The Nanodrop was blanked using RNase free water and 1 μL of RNA was loaded on the anvil. The absorbance of nucleic acids was measured at 260 nm and 280 nm. The RNA concentration was determined assuming that a 40 μg/ml RNA dilution has an absorbance of 1 at 260 nm. In addition, an absorbance ratio of 1.8–2 at 260/280 nm indicated that the sample was free from contaminants from the extraction process (i.e. proteins or chemicals).

2.2.6.3 Electrophoretic characterisation of RNA

The quality of total RNA was also assessed by the Pico 6000 Agilent Bioanalyzer 2100 (Agilent Technologies Ltd, Cheadle, UK) and was expressed by the RNA Integrity Number (RIN) based on the 18S and 28S rRNA ratio. Total RNA samples were diluted in RNase free H_2O to a concentration of 5 ng/μL and denatured at 70 °C for 2 min. Gel-dye mix (9 μL; 1 μL RNA dye : 65 μL prefiltered gel) was added to the chip and spread by maximum plunger's pressure for 30 s and released. RNA conditioning was added (1 well) and 5 μL of RNA marker was pipetted to all wells, followed by 1 μL of sample in each well and 1 μL of ladder to the specified well. Chip was vortexed (Agilent Bioanalyzer vortex) for 1 min and run on the Agilent 2100 Bioanalyser.

2.2.6.4 Preparation of cDNA

To generate complementary DNA (cDNA), 400 ng RNA in RNase-free H_2O was prepared to a final volume of 10 μl, denatured by heating for 5 min at 95 °C and placed immediately on ice to prevent RNA hybridisation. RNA was then mixed with 30 μl of reverse transcription (RT) mix containing a final concentration of 1× First-Strand Buffer (Invitrogen; Thermo Fisher Scientific), 1 mM deoxyribonucleotide triphosphates (dNTPs; Invitrogen), 0.3 μg random primers (Invitrogen), 0.1 M dithiothreitol (DTT; Invitrogen), 16 units RNasin®Plus RNase Inhibitor (Promega, Southampton, UK), 8 mM magnesium chloride ($MgCl_2$; Sigma-Aldrich) and 5 units SuperScript™III Reverse Transcriptase (Invitrogen). For all samples, an RT negative (RT-) was prepared alongside to control for the presence of any contaminating genomic DNA that could invalidate the results. A tube containing H_2O instead of RNA was prepared alongside as well to control for contaminations. As positive controls, ES cell lines TC-32 (EWSR1-FLI1 type 1), RD-ES (EWSR1-FLI1 type 2) and TTC-466 (EWSR1-ERG) were used. Samples were then placed on the Techne TC-5000 Thermal Cycler (Bibby Scientific Ltd, Staffordshire, UK) at 25 °C for 5 min, 50 °C for 1 h for DNA polymerisation, and a final step of 70 °C for 15 min for RT deactivation, followed by cooling down at 4 °C and stored at −20 °C.

2.2.6.5 Polymerase Chain Reaction (PCR)

From the resulting cDNA of Section 2.2.6.4, 10 μl were mixed with 40 μl of PCR mix containing sequence specific primers (Table 2.2; Sigma-Aldrich) (Roberts et al., 2008), 1.25 units AmpliTaq Gold (Applied Biosystems, Thermo Fisher Scientific) in 1× PCR Gold Buffer (Applied Biosystems) and 2.5 % dimethyl sulfoxide (DMSO; Sigma-Aldrich); except for β_2-microglobulin (β_2M), where DMSO was not added. Samples were then heated on the Techne TC-5000 Thermal Cycler (Bibby Scientific) at 96 °C for 10 min to denature cDNA and activate the AmpliTaq Gold, followed by 50 cycles (*EWSR1-FLI1/ERG*) or 33 cycles (β_2M) at 96 °C for 30 s, 1 min at specific temperatures (Table 2.2) (Roberts et al., 2008) for each target for the annealing and a 75 °C step for 1 min. These cycles were followed by a final extension step at 75 °C for 7 min. Samples were then cooled at 4 °C and stored at −20 °C. β_2M was used as a housekeeping gene to normalise the levels of expression across the different samples.

Primer	Sequence (5'-3')	Concent. (nM)	Annealing temperature (°C)
EWS F primer	TCCTACAGCCAAGCTCCAAGTC	200	-
FLI1 R primer	GAATTGCCACAGCTGGATCTGC	200	65
ERG R primer	TGAGGGGTACTTGTACAGA	200	55
β_2M F primer	GAGTATGCCTGCCGTGTG	200	-
β_2M R primer	AATCCAAATGCGGCATCT	200	60

Table 2.2: Primers and PCR conditions used to amplify for *EWSR1-FLI1*, *EWSR1-ERG* and β_2M. F = forward, R = reverse, Concent. = concentration.

2.2.6.6 Agarose gel electrophoresis

Agarose (1.5 % w/v, Invitrogen) was dissolved in 150 ml 1× Tris-Borate-EDTA (TBE) buffer (90 mM Tris (MP Biomedicals, Thermo Fisher Scientific), 90 mM Orthoboric acid (Sigma-Aldrich), 2 mM EDTA pH 8 in H_2O) by heating in microwave. Once the solution had cooled to approximately 60 °C, Nancy-520 (1 μg/ul; Sigma-Aldrich) was added and mixed by gently rotation of the flask. Nancy-520 is fluorescent stain specific for double-stranded DNA, which allows the visualisation of the separated bands. The sides of a Horizon20-25™agarose gel tray (Invitrogen) were sealed, agarose was poured and combs were placed and gel was left to settle at RT for 40 min. Then, the combs and tape were removed carefully and the tray placed into a Horizon 20-25™agarose gel electrophoresis tank (Invitrogen). Before loading, 30 μl of PCR product were mixed with 6 μl loading buffer (0.4 mM Ficoll (Type400, Sigma-Aldrich), 17.3 mM sodium dodecyl sulphate (SDS; Fisher Scientific), 0.4 mM Bromophenol blue (Sigma-Aldrich), 5.6 mM Orange G (Sigma-Aldrich)) to be able to visualize the progress of the electrophoresis. In addition, Bromophenol blue runs at a size of 150bp DNA and Orange G at a size of 50bp DNA, helping with the visualization of DNA size separation. 1× TBE was poured on top of gel and samples (36 μl) were loaded into the wells, alongside 50bp ladder (Bio-Labs) to be able to size the products. The gel was run at 80 V for 90 min or until the nucleic acids had separated sufficiently (determined by the separation of the blue and orange dyes present in the loading buffer). Gel was imaged on the ChemiDoc MP (Bio-Rad Laboratories Ltd, Hempstead, UK) by UV light to visualise the DNA.

2.2.7 Cytospins

Pelleted cells (Section 2.2.4) were resuspended in 3 ml of PBS. Six Superfrost Plus microscope slides (Thermo Fisher Scientific) and filters (Hettich zentrifugen; Hettich

Lab, supplied by SLS, Hessle, UK) were assembled into the centrifuge sample chambers, and 0.5 ml of cell suspension were added to each. Chambers were centrifuged at 1,000 g for 3 min (Rotix 32A Hettich zentrifugen, Hettick Lab) to cytospin cell to the slide, supernatant removed, and centrifuged again at 3,000 g for 1 min. Slides were left to air dry for 30 min and stored at −20 °C wrapped in foil.

2.2.8 Fluorescence *in situ* hybridisation (FISH)

All FISH characterisation of ES patient-derived cultures was done by Dr. Elizabeth Roundhill. Cytospins (Section 2.2.7) were defrosted at RT and aged at 60 °C for 1 h (Hybaid Shake'n'Stack oven, Thermo Fisher Scientific). Slides were then placed into Coplin jar (Sigma-Aldrich) at 72 °C containing denaturing buffer (70 % formamide (Sigma-Aldrich), 2× SSC solution (20× SSC: 3 M sodium chloride (VWR), 0.3 M sodium citrate (BDH Laboratories, supplied by VWR) in dH_2O, pH 5.3) in dH_2O, pH 7–8) and incubated for 5 min. Slides were then transferred to another Coplin jar containing 70 % ethanol and immersed for 1 min, followed by 1 min in 85 % and another 1 min immersion in 100 % ethanol, to be then left to air dry. The EWSR1 break apart probe (Vysis, Abbott Molecular, Berkshire, UK) was prepared by mixing 7 μl WCP hybridisation buffer (Vysis), 1 μl EWSR1 break apart probe (Vysis) and 2 μl purified water (total volume = 4 μl per slide) and vortexed. Prepared hydridisation mix was denatured at 73 °C for 5 min on a heating block and then reduced to 45 °C. Once slides were dry, 4 μl of probe mixture was added on the cell area of the slides, placing immediately a coverslip on top and sealed with rubber cement (FixoGum, Marabu). Slides were then placed in a humidified box (airtight container containing moistened towels) at 37 °C incubator overnight to allow hybridisation of probe to target cells. Rubber cement was removed from around the coverslips and slides were placed in Coplin jar containing post-hybridisation wash buffer (PHB; 2× SSC buffer with 0.3 % nonidet NP-40, Fluka Analytical, supplied by Sigma-Aldrich) 5 min at RT to allow coverslips to float off, followed by an immersion in PHB for 2 min at 72 °C and moved back to PHB at RT for 1 min. Slides were then air dried in the dark, mounted in faramount (Aquous medium ready-to-use; Dako, Agilent Technologies) containing 0.2 μg/ml 4', 6-diamindino-2-phentlindole (DAPI, Sigma-Aldrich) and allowed to dry in the dark. Slides were imaged using the

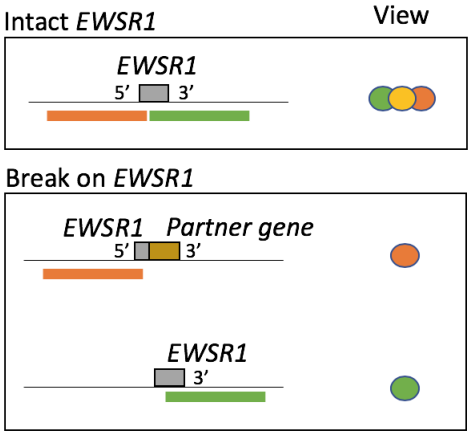


Figure 2.2: FISH *EWSR1* break apart probe for the detection of *EWSR1* fusions in ES patient-derived cultures. The Vysis probe used consists of an orange probe (497kb) flanking 5' of *EWSR1* into exon 4 and a green probe (1100kb) that flanks 3' of *EWSR1*. When the *EWSR1* gene is intact (top), both fluorophores are together so that they will be seen as yellow under the microscope. However, when there is a break apart on *EWSR1* (bottom), the 2 sections of the gene will be separate and therefore, they can be seen as independent fluorophores under the microscope. Because the known *EWSR1* break apart in ES are on introns 7 to 10, this FISH probe would detect ES characteristic fusions.

widefield fluorescent inverted microscope Nikon Eclipse Ti-E (Nikon, Kingston Upon Thames, UK) on DAPI, FITC and Texas Red filters.

When the *EWSR1* gene was intact, the probe was seen in yellow as both green and orange fluorophors were together (Figure 2.2). In contrast, when *EWSR1* gene had a break apart, both probes (3' and 5' *EWSR1*) were separate in the nucleus, thus resulting in both fluorophores emitting separately (red and green) (Figure 2.2). SK-N-MC ES cell line was used as positive control and HOS OST cell line as negative control.

2.2.9 Immunocytochemistry (ICC)

Cytospins were defrosted at RT and the Dako EnVision+ System (DAB) (Dako) protocol was followed, as summarised in Figure 2.3. Briefly, slides were fixed in fresh paraformaldehyde 4 % (w/v) in PBS (PFA; Sigma-Aldrich) for 20 min and washed in PBS. For c-kit, slides were placed in 0.1 % Triton-X 100 (in PBS; Amersham Biosciences, Little Chalfont, UK) for 5 min to permeabilise cells and then moved back to PBS. Slides were then placed into Shandon Sequenza (Thermo Scientific) and 100 μl of peroxidase block (Dako) was added to each slide. After 5 min, slides were washed 3 times with TBS-T (137 mM sodium chloride, 5 mM

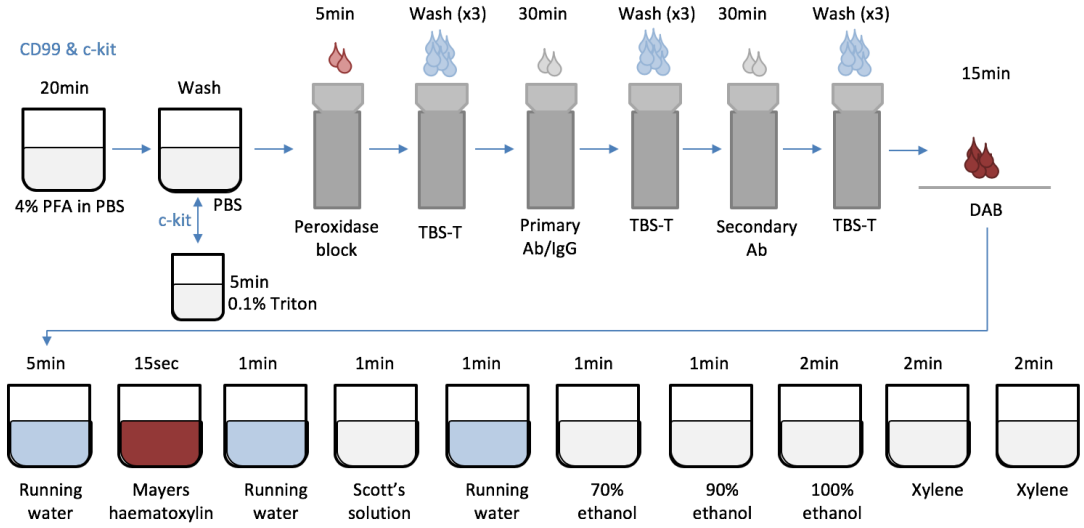


Figure 2.3: ICC protocol for CD99 and c-kit staining on ES patient-derived cell cultures. ES cytospins were stained for CD99 and c-kit primary antibodies or IgG control using the Dako EnVision+ System (Dako, Agilent Technologies) and counterstained with haematoxylin to stain the nucleus. Abbreviations: PFA = paraformaldehyde, TBS-T = washing solution (sodium chloride with TrisHCl Ultra in dH_2O), Ab = antibody, IgG = isotype control, DAB = DAB Substrate Chromogen System.

TrisHCl Ultra (MP Biomedicals) in dH_2O). Primary antibody or IgG isotype to control for unspecific binding (100 μl; Table 2.3) were added and left for 30 min. Slides were again washed 3 times with TBS-T, and secondary antibody (Peroxidase labelled polymer, 100 μl; Dako) was added and left for 30 min. After 3 more TBS-T washes, slides were removed from the Sequenza and 0.5 ml of DAB Substrate Chromogen System (1 drop of DAB Chromogen per 1 ml of Substrate buffer; Dako) were added on top of each slide and incubated for 15 min at RT, until DAB oxidation with peroxidase (brown colouring) was visible on positive control. Slides were then counterstained with haematoxylin for visualisation of the nucleus. For this, slides were placed for 5 min in running water, placed for 15 s in Mayers Haematoxylin (0.3 % w/v haematoxylin (Sigma-Aldrich), 2 % v/v ethanol, 0.03 % w/v sodium iodate (Sigma-Aldrich), 0.1 % w/v citric acid (MP Biomedicals), 5 % w/v chloral hydrate (Sigma-Aldrich), 5 % w/v aluminium potassium sulphate (Acros Organics, Fisher Scientific), 12 % v/v glycerol (Sigma-Aldrich) in dH_2O), 1 min in running water to wash excess, placed 1 min in Scott's solution in order to help achieving a good staining contrast for the nucleus (bluing reagent) (2 % w/v magnesium sulphate (Sigma-Aldrich), 0.35 % w/v sodium bicarbonate

Antibody	Species	Concentration (ng/μl)	ID	Supplier
CD99	Mouse	4.36	M3601	Dako
c-kit	Rabbit	20	ab5506	Abcam
IgG isotype	Mouse	4.36	IC0041A	R&D systems
IgG isotype	Rabbit	0.5	08-6199	Invitrogen

Table 2.3: Antibody dilutions and specifications for CD99 and c-kit ICC on ES patient-derived cell cultures. Supplier information: Abcam Plc., Cambridge, UK; R&D systems, Abingdon, UK.

(Sigma-Aldrich) in dH_2O), another minute in running water, and then 1 min in 70 % ethanol, 1 min in 90 % ethanol, 2 min in 100 % ethanol to dehydrate cells. Finally, slides are immersed for 2 min in xylene (99% Xylene; Acros Organics) and again 2 min in xylene to remove any remaining water. Finally, following dehydration, slides were mounted in DPX mountant (Sigma-Aldrich) and covered with a glass coverslip. ICC was assessed by light microscopy (Olympus CKX41; Olympus Microscopy, Southend-on-Sea, UK). TC-32 ES cell line was used as positive control for both CD99 and c-kit expression. Scoring for both markers was assessed as negative (—), low expression (+) and high (++) expression intensity by 2 independent lab members, blinded and scores merged.

2.2.10 Flow cytometry for CD99 expression on cells

The CD99 expression was evaluated by flow cytometry to determine the percentage of cells expressing CD99 marker. ES patient-derived cells were seeded into Primaria^{TM} 6-well plates at a concentration of 2×10^5 cells/well. After 24 h, cells were harvested (Section 2.2.3) and pelleted (Section 2.2.4). In order to wash any remaining tissue culture media, cells were resuspended in 1 ml of FACS buffer (0.5 % Bovine Serum Albumin (BSA; Sigma-Aldrich), 2 mM EDTA in PBS), centrifuged and supernatant aspirated. Cell pellets were then resuspended in 50 μl of FACS buffer containing the appropriate concentration of primary antibody (CD99 10 mg/ml; Santa Cruz Biotechnology, Heidelberg, Germany) or corresponding IgG isotype control (IgG_{2B} Control Isotype, 10 mg/ml; R & D systems), and incubated for 30 min at 4 °C. Cells were washed by resuspending in 1 ml of FACS buffer, centrifuged and supernatant aspirated. Cell pellets were then resuspended in 50 μl of FACS buffer containing the appropriate concentration of fluorescent secondary antibody (goat anti-mouse-RPE, 1 ng/μl; Southern Biotech, supplied by Cambridge Biosciences, Cambridge, UK) and incubated for 30 min at 4 °C. Finally, cells were

washed as described above and resuspended in 1 ml of FACS buffer and analysed on the Attune (Attune Digital 2 Laser 6 Colour Acoustic Focusing Analyser; Thermo Fisher Scientific) using an excitation laser of 480nm. A total of 10,000 events for the specified population were analysed and all experiments were performed in triplicate, with an IgG isotype control and a non-labelled control to define the CD99+ gate. The data analysis was done using the Attune Cytometric Software V2.1 (Thermo Fisher Scientific). The percentage of CD99 positivity was calculated as the mean of percentage of positive cells on the 3 CD99 repeats minus the percentage of positive cells for the IgG control ± standard deviation (SD). TC-32 ES cell line was used as positive control.

2.2.11 Self-renewing assay

For the evaluation of the self-renewing ability of the different patient-derived ES cell cultures, a self-renewing functional assay was used as summarised in Figure 2.4. The earliest possible passage (normally passage 4–5) was used for this assay. Cells were counted (Section 2.2.5) and a single cell suspension was obtained of 1 cell per 200 μL of media. Using a multichannel pipette (Gilson, Dunstable, UK), ten 96-well plates (Corning Primaria™) were filled with 200 μL of cell suspension per well, so that 1 cell fell into each well. Following a 3-week incubation at 37 °C, the self-renewing ability was measured by the presence of more than 5 cells per well (Equation 2.1).

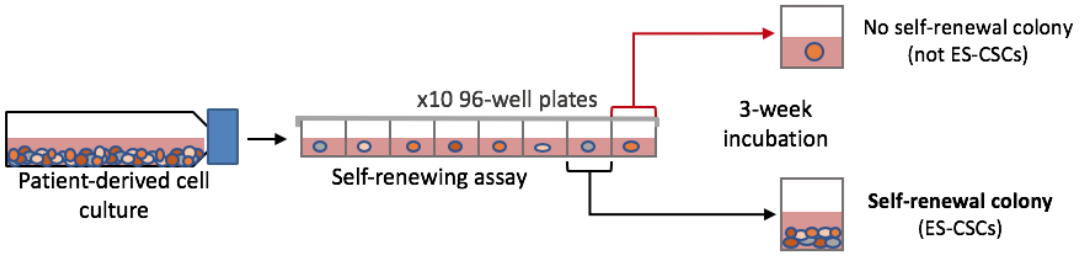


Figure 2.4: Self-renewal assay from single cell suspension. The self-renewal ability was assed by the capacity of single cells to form clones after 3 weeks of incubation. Based on the ability of cancer stem-like cells (CSC) to be able to self-renew, the wells that had a cell population after 3 weeks of incubation suggested that a CSC fell into that well; whereas if a well was empty or only 1-2 cells were present, it suggested that the cell that was seeded in that well was a non-CSC cell.

$$\text{self-renewing capacity} = \frac{\text{wells with more than 5 cells}}{\text{total of seeded wells}} \times 100 \qquad (2.1)$$

2.2.11.1 Cell per well distribution of self-renewing assay

To determine the accuracy of the single-cell seeding on Section 2.2.11, the number of cells that were seeded per well was evaluated. After 24 h of seeding the single-cell suspension in 96-well plates (Section 2.2.11), media was carefully aspirated from wells and a wash in PBS (100 μl) was performed to remove any debris. Cells were fixed in 4 % PFA (100 μl per well) for 15 min at RT, PFA discarded and wells were washed again in PBS to remove any leftover fixative. Cells were incubated in 0.2 μg/ml DAPI in 50 μl PBS per well for 30 min in the dark. Solution was removed and wells were filled with 100 μl PBS. To quantify the number of cells per well, plates were scored using the fluorescence microscopy (Olympus CKX41).

The evaluation of the seeding technique for the self-renewing assay (Section 2.2.11.1) was performed using the Chi-Square goodness of fit test. First, the observed cells per well from 2 independent SK-N-MC seeded plates were compared to the expected assuming Poisson distribution of $\lambda = 1$ (one cell per well). In addition, when data from 8 different ES patient-derived cell cultures was obtained, the analysis was repeated but this time investigating which λ fitted best for each plate. Therefore, $\lambda < 1$ would indicate that more wells than expected were empty and the probability of having multiple cells per well would be lower. Contrary, having $\lambda > 1$ would indicate that less wells are empty, but the probability of having multiple cells per well increased. In addition, the probability of having 2 or more cells per well was calculated.

2.2.12 Migration assay from a 3D spheroid

Assay previously developed in the CCRG by Dr. Lucy Shaw. For the evaluation of the migration capacity of ES patient-derived cells, a migration assay from a 3D spheroid was used. Cells (1×10^3 in 200 μl of media; Section 2.2.5) were seeded in ultra-low adherent plate wells (SLS) and incubated for 4 days to obtain cellular spheroids. The SK-N-MC cell line was used as a positive control. On day 5, wells of a 24-well plate (Corning Primaria™, Falcon, BD Biosciences) were coated with 600 μL gelatin (0.1 % in H_2O; Fluka Analytical), incubated for 1 h at 37 °C and replaced with 600 μL of media. Media (100 μL) was removed from 4 complete spheroids (selected by light microscopy, Olympus CKX41) and the remaining 100 μL (with the spheroid) were carefully moved using a P1000 pipette (Gilson) to

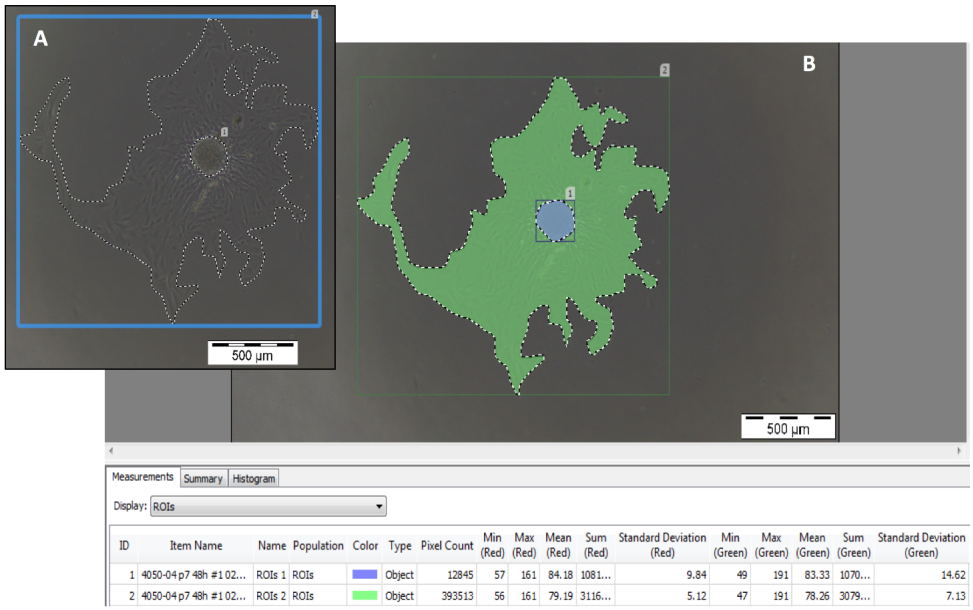


Figure 2.5: Scoring of migration assay using Volocity®6.3.0 software. **A:** The edge of the spheroid (inner circle) and the edge of the migrating area of cells were specified using the Volocity®6.3.0 software. **B:** The measurments on the Volocity®6.3.0 software provide a pixel count for each one of the defined areas as shown on the bottom of the image (spheroid in blue, migration area in green) . This pixel count was used to calculate the migration index.

the 24-well plate. Plates containing spheroids were incubated for 1 h to allow cells to adhere to bottom. Spheroids were imaged at 0 h, 24 h, 48 h and 72 h by light microscopy and Cell B software (Olympus). Images were then analysed using Volocity®6.3.0 software (PerkinElmer, Cambridge, UK) by manually outlying the core and outer migrated boundary to obtaining a pixel count (Figure 2.5). Migration index was calculated using the pixel count from Volocity®6.3.0 software relative to the core area (Equation 2.2). SK-N-MC ES cell line used as positive control.

$$\text{migration index} = \frac{\text{outer migration area - core area}}{\text{core area}} \tag{2.2}$$

2.2.13 Sensitivity to doxorubicin and vincristine

Assay based on protocol from Murray et al. (2017). Cells (1×10^3; Section 2.2.5) were seeded in triplicate into 6-well plates (Corning Primaria™, Falcon, BD Biosciences) in 2 ml of LAM and after 24 h incubation, doxorubicin and vincristine (in ddH_2O; Sigma-Aldrich) dilutions were added obtaining final concentrations of

1 nM, 5 nM, 10 nM, 25 nM, 100 nM and 200 nM for both drugs, and 3 wells were treated with vehicle control (ddH_2O). After a 72 h incubation, media containing doxorubicin or vincristine was replaced with fresh media without drug, and incubated for 7 further days (Figure 2.6A). Cells were then stained with crystal violet (0.25 % crystal violet (Fluka Analytical) in 50:50 ddH_2O:methanol, Merck-Millipore) for 20 min at room temperature and rinsed with water. TC-32 cell line was used as positive control (6×10^3 cells per well). The formation of colonies on the wells was imaged on ChemiDoc MP (Bio-Rad) and analysed by Quantity One®software (Bio-Rad) using the colony analysis tool with sensitivity of 7 (Figure 2.6).

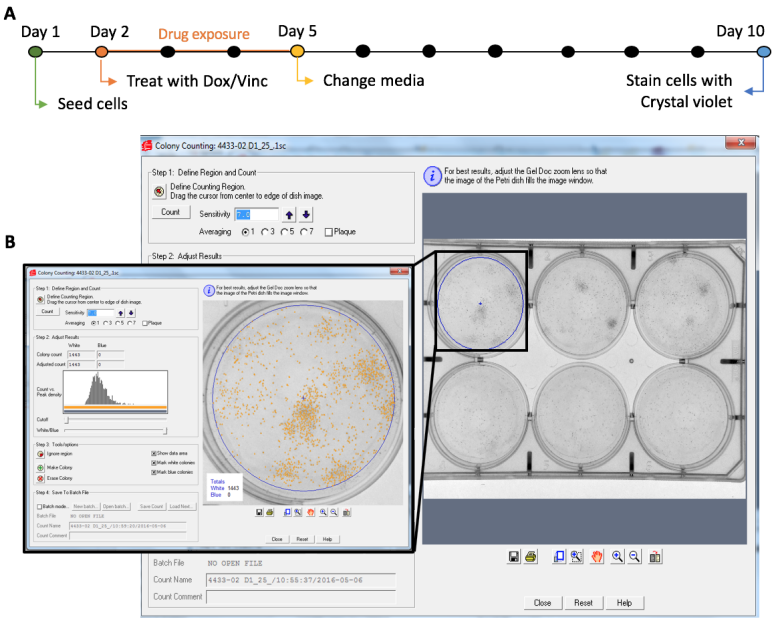


Figure 2.6: Sensitivity to doxorubicin and vincristine evaluated by colony formation analysis using Quantity One®software. **A:** The colony number was evaluated by the Quantity One®software using the colony analysis tool on sensitivity of 7. **B:** The colony count for each well was identified and scored for drug sensitivity analyses.

Percentage of colony formation was obtained relative to the colony number of untreated wells (Equation 2.3). EC50 (half maximal effective concentration) calculations were performed using a nonlinear regression on dose-response inhibition for a variable slope (GraphPad Prism 6).

$$\text{Colony formation} = \frac{\text{colony formation at drug concentration}}{\text{colony formation at untreated wells}} \times 100 \qquad (2.3)$$

2.2.14 Survival analysis

The survival analyses were performed in R (R, 2017) using the proportional hazards regression model (coxph) from package Survival (Therneau and Lumley, 2017) and Kaplan-Meier plots exported using the Survminer package (Kassambara et al., 2018). The coxph package can distinguish between continuous variables (i.e. migration capacity) and factor variables (i.e. presence of metastasis or not), and for those cases that more than 2 factors are present (i.e. metastatic disease location) provides results for each variable. In order to evaluate the data, the log rank p-value, the hazard ratio and the 95% confidence interval were assessed, only considering significant differences when $p > 0.05$.

2.2.15 Statistics

Data from phenotypic characterisation was analysed by descriptive statistics. Unpaired t-test (2 datasets) or ordinary one-way analysis of variance (ANOVA) for multiple comparisons (>2 datasets) were used, considering significant differences when $p<0.05$. All analyses were undertaken using GraphPad Prism 6 software, except the descriptive statistics that were performed on Microsoft Excel version 15.24.

2.3. Results

2.3.1 Processing and culture of ES cells from patient-derived samples

In order to study the biology of ES cells, patient-derived samples were used instead of established cell lines. Through the GenoEwing study, a total of 43 samples have been received from different UK orthopaedic surgical centres (Table A1). The successful culture rate of these samples was 79 % (34/43), but when separated according to sample type, 94 % of diagnosis biopsies (31/33) have been successfully maintained in culture, in comparison to 30 % for the post-treatment resections (3/10). From all the 43 received samples (Figure 2.7A), 58 % (25/43) were ES: 60 % (15/25) diagnosis and 40 % (10/25) post-treatment resections. Similar to the overall successful rate, ES diagnostic biopsies were successfully cultured in 87 % of cases (13/15), in comparison to only 30 % for post-treatment

resections (3/10). The rest of the samples collected (42 %, 18/43) were non-ES, comprising different cancer and non-cancer types (Figure 2.7A, Table A1).

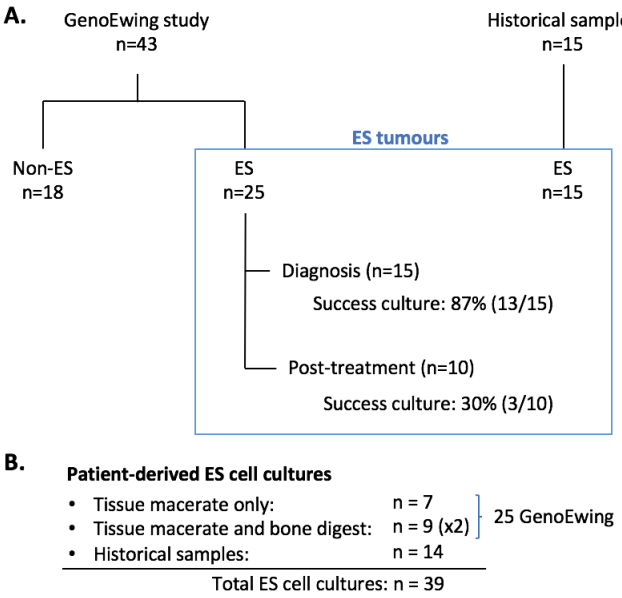


Figure 2.7: Patient-derived ES samples collected and processed. **A.** Both GenoEwing samples and historical samples are shown, separated based on ES or other cancers (non-ES). For the GenoEwing samples, the type of tissue received (diagnosis biopsy or post-treatment resection) is shown, as success in culture varied between these. **B.** Patient-derived cell cultures successfully obtained. As GenoEwing samples were processed as tissue macerate and/or bone digest depending on the tissue received, a total count of the obtained ES patient-derived cell cultures is shown.

From the 16 successfully cultured ES samples (Figure 2.7B), 37 % (6/16) had a soft-tissue fraction only (no bone) and therefore did only generate a tissue macerate derived culture (identified as Geno#T through the study). The other 63 % (10/16) had both a soft-tissue and bone fraction and were processed to generate both cell cultures (Geno#T and Geno#B respectively). From these, 90% generated both cell cultures, and only one generated tissue macerate culture only (Geno12). Altogether, this lead to a total of 25 ES patient-derived cell cultures from the GenoEwing study, originated from 16 different patients (Figure 2.7B). In addition, 15 historical ES samples (success culture rate 93%, 14/15) were used for this study (named as Hist#), generating a total of 39 ES patient-derived cell cultures from 30 different patients (Figure 2.7B). The location of the primary tumour was bone in 63% of tumours (19/30; 15 GenoEwing and 4 historicals) and soft-tissue

tumours in 33% of cases (10/30; 1 GenoEwing and 9 historicals), with 1 patient having unknown tumour location (Hist12).

ES cultured cells (39 samples) grew as substrate adherent monolayers (Figure 2.8) except for a case of suspension culture that was not possible to maintain in culture (Geno4; Table A1). Cell cultures were a mix of cells with spindle-shape morphology (black arrow) and some cells were reminiscent of MSC, with wider bodies (Figure 2.8; white arrow). All experiments were performed on the lowest possible passage, normally between passage 4-8 since generation of cell culture.

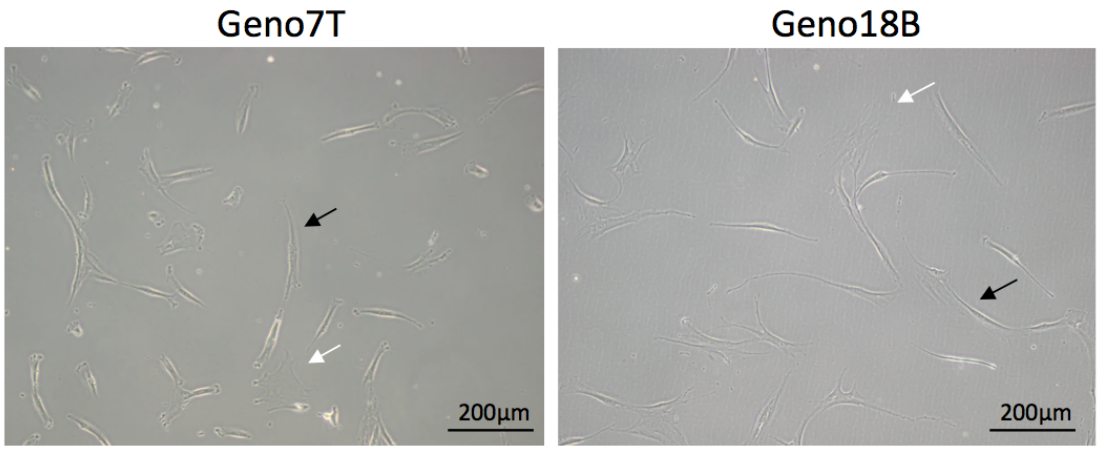


Figure 2.8: Morphology of patient-derived ES cell cultures growing as substrate adherent monolayers. As an example, Geno7T and Geno18B cell cultures are shown. A mixed population of spindle-shaped cells (black arrows) and cells reminiscent of MSC (white arrows) was observed and is representative of the patient-derived ES cultured samples.

Throughout this study, the difference between bone-derived tumours and soft-tissue tumours (tumour location) were investigated to determine if tumour location was associated to the cellular phenotype. In addition, cell cultures from tissue macerate and bone digestion (processing method) were compared, to evaluate if the processing method could result in different *in vitro* behaviours or the effect of the bone microenvironment (cells from the bone) could be maintained in culture. Finally, the GenoEwing and historical samples (project) were evaluated, to determine if sample collection and processing method could generated differences in the phenotype, or the profiles were reproducible. Statistical differences are only reported in each section when significantly different.

2.3.2 Detection of ES characteristic markers in ES patient-derived samples

Successfully cultured ES patient-derived samples were tested for the expression of CD99 assessed by ICC (39/39) and flow cytometry (25/39); and for the presence of a EWSR1 translocation by FISH with *EWSR1* break apart probes (24/39) or by RT^A-PCR against the EWSR1-FLI1 or EWSR1-ERG fusions (36/39) (Table 2.4).

2.3.2.1 CD99 profile on ES patient-derived samples

A total of 97 % of ES cell cultures (38/39) were positive for CD99 expression in either ICC or flow cytometry data, as shown in Table 2.4. The CD99 expression by ICC was evaluated in all samples (39), which showed that 82 % (32/39) samples expressed CD99, at low (14/39) or high (18/39) levels and 18 % had negative expression (7/39) (Figure 2.9A). The percentage of CD99 cells by flow cytometry was also assessed in 64 % (25/39) cultures with a median expression of 72 % (range 8–96 %) (Figure 2.9B). It is important to mention that of those 13 samples with expression under 72 %, only 1 had less than a 20 % of positive cells (Hist12).

When comparing both techniques, the samples with high expression by ICC were not always the samples with higher percentage of $CD99^+$ cells (Table 2.4). From the 7 samples with negative results by ICC, 3 showed CD99 expression by flow cytometry. The other 4 samples could not be evaluated by ICC, thus not being able to compare the results. This differences in the CD99 expression results between ICC and flow cytometry could be explained by the antibody clone type used, which could lead to different affinity for the epitope.

2.3.2.2 Presence of an *EWSR1* fusion

The presence of a EWSR1 translocation was evaluated in 37 patient-derived samples by either FISH (24/37) and/or RT^A-PCR (36/37) (Table 2.4), and a fusion involving EWSR1 was detected in 78 % of ES-derived cell cultures (29/37).

Evaluation of a break apart assay on the *EWSR1* by FISH (24/37 samples) showed that in 13 % of cases (3/24) all imaged cells had a break apart on the *EWSR1* gene, visualised as both 3' (red) and 5' (green) *EWSR1* probes being separate in the nucleus (Figure 2.10A). In addition, 87 % (21/24) had a mixture of cells with and without break apart (Figure 2.10A, Table 2.4). No samples evaluated by FISH

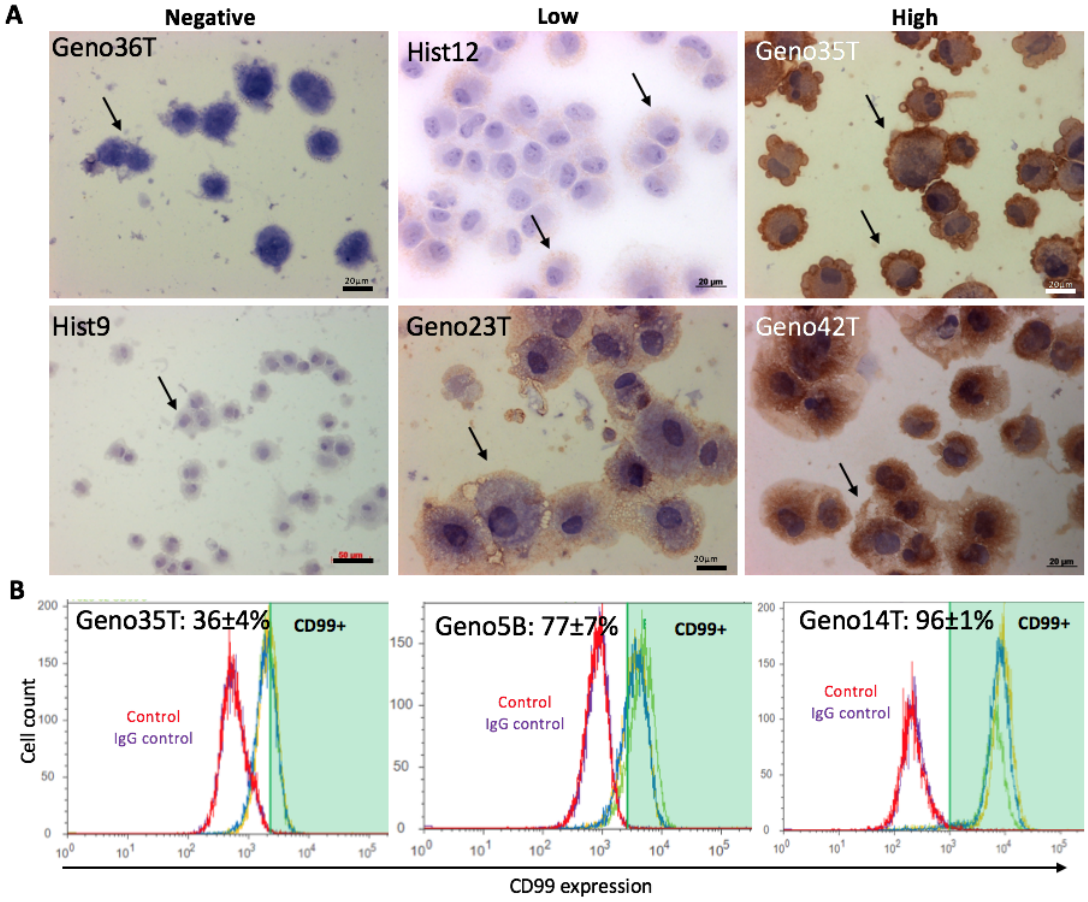


Figure 2.9: Evaluation of CD99 expression in patient-derived ES cultures. **A.** Expression of CD99 by ICC scored as negative (left), low expression (middle) and high expression (right). Arrows indicative of expression on cell surface. Scale bar = 20 μm. **B.** CD99 expression was evaluated by flow cytometry. Expression of CD99 (n= 3, green, blue and yellow) compared to the IgG (purple) and unlabelled control (red) in 3 different cell cultures, with lower to higher CD99 expression.

showed negative results (no break apart on the *EWSR1* gene), compared to the negative control used (OST HOS cell line) that always showed absence of break apart (Figure 2.10A).

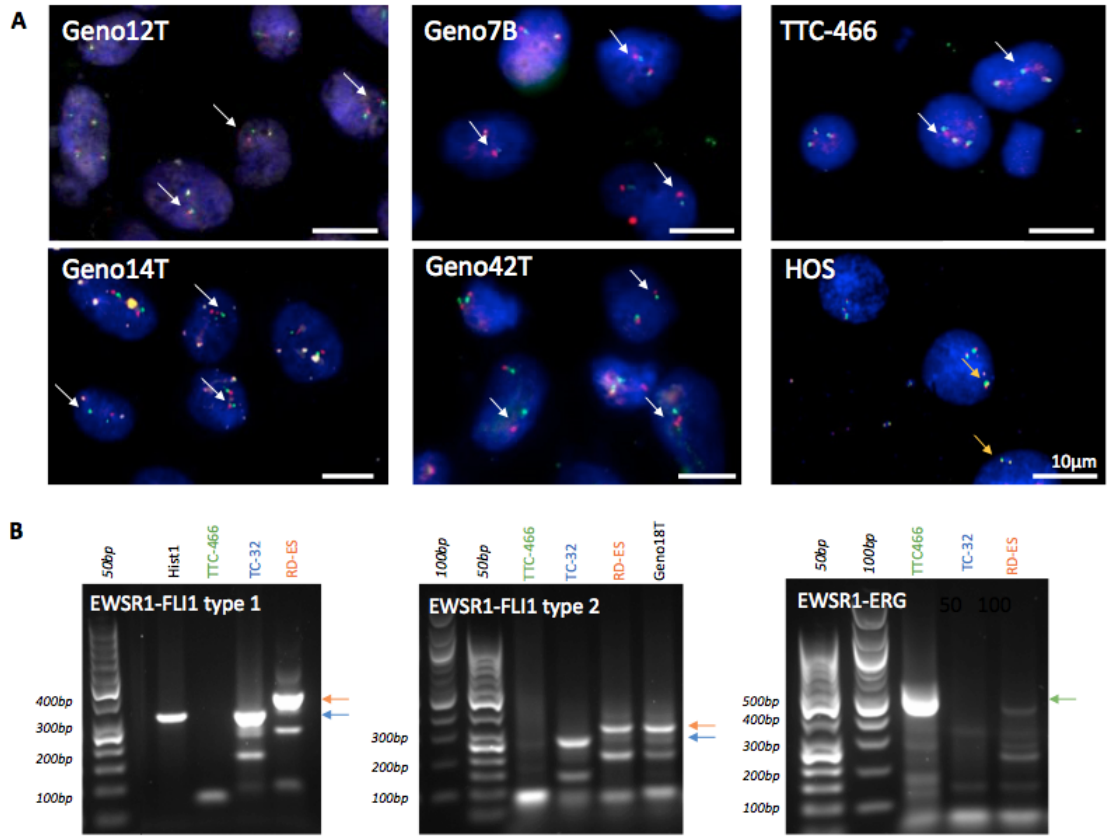


Figure 2.10: Evaluation of EWSR1 fusion on ES patient-derived cultures. **A.**The presence of an *EWSR1* fusion was evaluated by FISH using a break-apart probe against *EWSR1*. 4 patient-derived ES cultures are shown as an example, as well as ES cell line TTC-466 and OST cell line HOS (negative control). White arrows indicate break apart (red and green probes separate), yellow arrows indicate intact *EWSR1* gene (red and green together, visually yellow). **B.** The presence of an EWSR1 fusion (EWSR1-FLI1 and EWSR1-ERG) was also evaluated by RT^A-PCR. An example of EWSR1-FLI1 type 1 (left; blue arrow), EWSR1-FLI1 type 2 (middle; orange arrow) are shown on patient-derived ES cultures; EWSR1-ERG was evaluated as well but no samples expressed this fusion, and therefore TTC-466 is shown as example (right; green arrow). The positive controls are shown TC-32 for EWSR1-FLI1 type 1; RD-ES for EWSR1-FLI1 type 2; and TTC-466 as EWSR1-ERG. Size ladder included in gels (bp = base pairs).

Contrary to the FISH technique, which detects any break apart on the *EWSR1* gene after exon 4, the RT^A-PCR detects specific fusions for which the primers are designed against. Therefore, this is a more informative technique to know the specific fusion type, but any gene partner not covered by the designed primers will not be detected. RT^A-PCR data indicated that 7/36 samples had a EWSR1-FLI1 type

1 fusion, 4/24 had a EWSR1-FLI1 type 2 fusion and none showed expression of the EWSR1-ERG fusion. In the remaining 25/36 samples evaluated, no fusion was detected (Figure 2.10B, Table 2.4). This could suggest that ES patient-derived cell cultures have low yields of the fusion mRNA, which results in lack of detection by PCR in comparison to cell lines, where the fusion can be detected easily using the same amount of RNA.

When putting data from both techniques together, all negative samples by RT^A-PCR that had FISH data showed a break apart by the latter, indicating that an EWSR1 fusion was present. A possible explanation for this difference of detection could be the yield of the fusion. As FISH (DNA level) can detect the fusion in all samples, it could be suggested that at the RNA level, the fusion is less expressed. Another explanation is that fusions without the FLI1 and ERG partners, or with other break aparts that the PCR primers would not recognise, are involved in these samples, and therefore, would not be detected. The other 8 samples with negative results by RT^A-PCR could not be investigated by FISH, and therefore, we cannot conclude if the fusion is present or not. Future studies should aim to fully characterise these samples by both methods in order to compare the results.

When comparing CD99 expression and presence of an EWSR1 fusion, 69% of samples (27/39) showed expression of both ES diagnostic markers. The rest of the ES patient-derived cultures only showed expression of CD99 but no fusion (15%, 6/39), only fusion but no CD99 (5%, 2/39) or could only be evaluated for CD99 expression (5%, 2/39; Hist7 and Hist9). In addition, 2 cell cultures (5%; Hist2 and Hist11) showed negative expression for both characteristic ES markers. Therefore, 95% of patient-derived ES cultures expressed either CD99 or had an EWSR1 fusion, indicative of their ES diagnosis. The 2 samples (5%) that were negative for both markers were only evaluated by ICC (CD99) and RT^A-PCR (fusion), and therefore, evaluation using flow cytometry (CD99) and FISH (fusion) as well would be needed to further confirm their negative profile, as in other samples, these two techniques were able to confirm a positive ES profile when ICC or RT^A-PCR could not. Therefore, further characterisation of the samples that have not been analysed by FISH or flow cytometry would be needed in order to confirm their expression profile regarding the 2 ES diagnostic biomarkers.

ID	CD99 expression		EWSR1 fusion	
	ICC	Flow (%±SD)	FISH	RT^A-PCR
Hist1	neg	26 ± 8	—	type 1
Hist2	neg	—	—	ND
Hist3	+	34 ± 28	—	ND
Hist4	+	85 ± 6	—	ND
Hist5	+	80 ± 10	—	ND
Hist6	+	42 ± 52	—	type 1
Hist7	+	69 ± 5	—	—
Hist8	neg	86 ± 12	mix	ND
Hist9	+	—	—	—
Hist10	+	—	—	ND
Hist11	neg	—	—	ND
Hist12	+	8 ± 2	mix	type 1
Hist14	+	—	—	ND
Hist15	+	—	—	ND
Geno1T	neg	87 ± 5	mix	type 1
Geno2T	+	88 ± 2	mix	type 2
Geno2B	neg	—	—	type 2
Geno5T	++	95 ± 4	mix	type 1
Geno5B	+	77 ± 7	mix	type 1
Geno7T	++	92 ± 2	yes	ND
Geno7B	++	78 ± 2	mix	ND
Geno10T	++	95 ± 1	mix	ND
Geno10B	++	95 ± 2	mix	ND
Geno12T	++	—	—	ND
Geno14T	+	96 ± 1	mix	type 1
Geno18T	++	59 ± 13	—	type 2
Geno18B	++	24 ± 15	mix	ND
Geno23T	+	—	yes	ND
Geno25B	++	72 ± 3	mix	ND
Geno25B	++	68 ± 1	mix	ND
Geno34T	++	—	mix	ND
Geno34B	++	—	mix	ND
Geno35T	++	36 ± 4	mix	ND
Geno36T	neg	—	yes	ND
Geno39T	++	30 ± 8	mix	ND
Geno39B	++	35 ± 6	mix	type 2
Geno42T	++	—	mix	ND
Geno42B	++	49 ± 8	mix	ND
Geno43T	++	—	mix	—

Table 2.4: Assessment of ES markers in patient-derived ES cell cultures (n = 39). Expression of CD99 by ICC (neg = negative expression, + = low expression, ++ = high expression; blinded scoring) and flow cytometry (as percentage of cells expressing CD99 ± SD, n= 3). Presence of an *EWSR1* translocation was assessed by FISH (yes = cells have break on *EWSR1* gene; mix = mixture of cells with and without break apart) and RT-PCR (type 1 = EWSR1-FLI1 type 1, type 2 = EWSR1-FLI1 type 2, ERG = EWSR1-ERG, ND = not detected). — = assay not done.

2.3.3 Expression of c-kit in ES cultures

As discussed in Section 2.1, c-kit has been found to have high expression in ES samples (Scotlandi et al., 2003), although this did not provide an improved response to c-kit inhibitors when tested clinically for ES patients (Bond et al., 2008). In this study, the expression of this tyrosine kinase receptor was assessed in 30 of the 39 ES patient-derived cell cultures by ICC (Table 5.9) to evaluate if the expression was indeed high in them, and therefore, further c-kit inhibitors could be evaluated for the treatment of ES.

A total of 93 % of ES patient-derived cultures (28/30) showed expression of c-kit: 17 % (5/30) had low levels of expression, compared to 77 % (23/30) with high levels (Figure 2.11). Only 2 samples (Hist1 and Hist5) showed negative expression for c-kit. In addition, 43 % of the samples (13/30) presented with a mixture of cells with different levels of positive expression (Figure 2.11). Therefore, this increased expression of c-kit could be further investigated preclinically by evaluating different c-kit inhibitors used in the clinic for other malignancies and identify if it could be used as part of the treatment modalities for these patients.

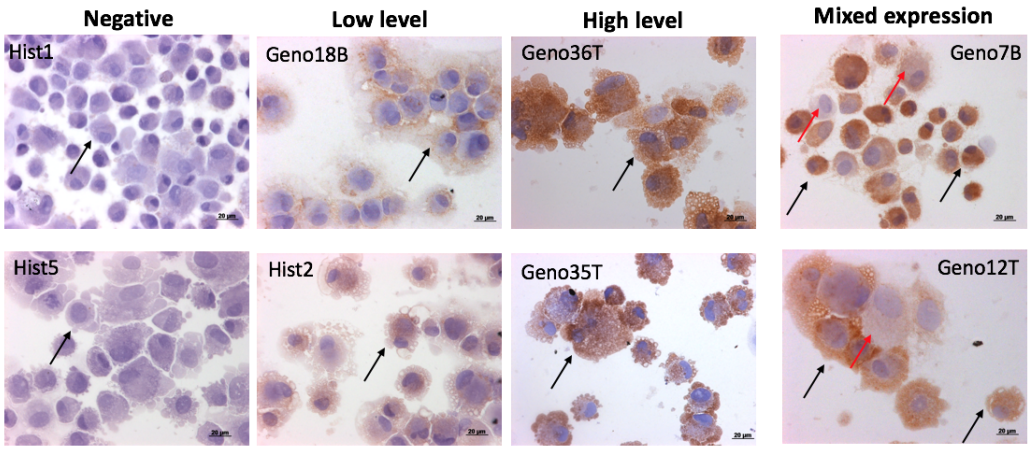


Figure 2.11: C-kit expression profile by ICC in ES patient-derived cells. Expression of c-kit was classified as (from left to right): negative expression, low expression and high expression. In addition, some samples showed a mixed expression profile, with low and high expression of c-kit (bottom). Black arrow = cells expressing c-kit; red arrow = mixed population (different level of intensity). Scale bar = 20 μm.

2.3.4 Evaluation of the single-cell seeding technique

To evaluate the ability to self-renew from a single cell, it was important to ensure that the methodology used resulted in the seeding of a single cell per well and

not more. Seeding more than one cell per well could have a great impact on the downstream analysis and interpretation of the data. Therefore, the number of cells per well from 2 independent 96-well plates of SK-N-MC ES cell line cell suspension were evaluated. The observed number of cells per well was 1.3 ± 1.3 ($n = 192$, range 0–6) and the median was 1. This cell distribution (observed) was compared with the expected number of cells per well assuming Poisson distribution of $\lambda = 0$, 1, 2 and 3 (number of cells that we expect to have into each well) to see which model fitted best with the observed data. As shown in Figure 2.12, the best fit between observed and expected data was assuming $\lambda = 1$ (probability of observed and expected on same distribution is 0.99). Therefore, the methodology used to seed 1 cell per well in order to evaluate the self-renewing ability is successfully seeding single-cell suspensions into the wells.

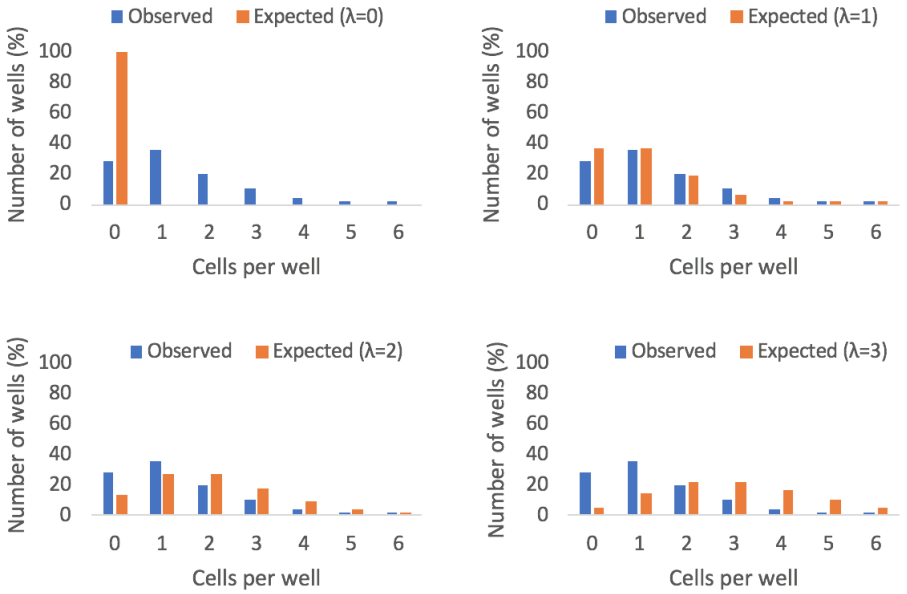


Figure 2.12: Comparison between observed and expected number of cells per well on the self-renewing assay assuming different λ. The observed number of cells per well (data from 2 independent 96-well plates of SK-N-MC cell line) was compared to the expected following Poisson distribution assuming $\lambda = 0, 1, 2$ and 3.

Following this preliminary evaluation of the technique, data was collected from 8 different ES patient-derived cultures alongside the self-renewing assay, to evaluate the cell per well distribution in these cases (Table 2.5). In all plates, the mean of cells per well was below 1, whereas the median was 0 ($n = 4$) or 1 ($n = 4$), indicating that

λ was either 1 or 0, which would lead to most wells having 0–1 cells. The probability of having more than 2 cells was evaluated (Table 2.5), confirming that the probability of having multiple cells per well was low (0.06 ± 0.06, mean ± SD). Altogether, this data indicates that the single-cell seeding technique fits a Poisson distribution of 1 cell per well, meaning that any cell colonies formed after 3 weeks of incubation will be derived from single cells with the ability to self-renew, rather than an excess of cells at time of seeding.

ID	Cells per well						Mean	Median	prob > 2 cells
	0	1	2	3	4	+5			
Geno23T	29	47	14	6	0	0	0.97	1	0.06
Geno25B	39	33	18	6	0	0	0.91	1	0.06
Geno25B	53	36	6	1	0	0	0.53	0	0.01
Geno34T	46	33	15	1	0	0	0.69	1	0.01
Geno34B	53	32	9	1	0	0	0.56	0	0.01
Geno35T	58	24	11	1	1	0	0.56	0	0.02
Geno39T	47	26	15	4	3	0	0.84	1	0.07
Geno39B	48	33	11	3	0	0	0.67	0	0.03

Table 2.5: Cell distribution for self-renewing seeding technique on ES patient-derived cell cultures evaluated after 24 h of seeding single-cell suspension. Table shows number of wells with 0–5 cells per well in a 96-well plate. Right columns of the table show the mean cell per well for each sample, the median and the probability (prob) of having more than 2 cells per well.

2.3.5 Self-renewing ability of ES patient-derived cells

The ability to self-renew from a single cell was investigated in 38 ES patient-derived cultures (Figure 2.13, Table 5.9). The percentage of self-renewing cells per cell culture was 9±8 % (mean ± SD; range 0.1–32 % n= 38), indicating that ES patient-derived cell cultures display different levels of self-renewing ability.

When comparing the self-renewing ability of bone tumours (18/28 patients) with soft-tissue tumours (10/28 patients), no statistical differences were seen between these two sample groups ($p = 0.4$), indicating that the tumour site does not have an impact on the self-renewing ability. After 3 week incubation, wells with a self-renewing colony were seen as shown in Figure 2.14.

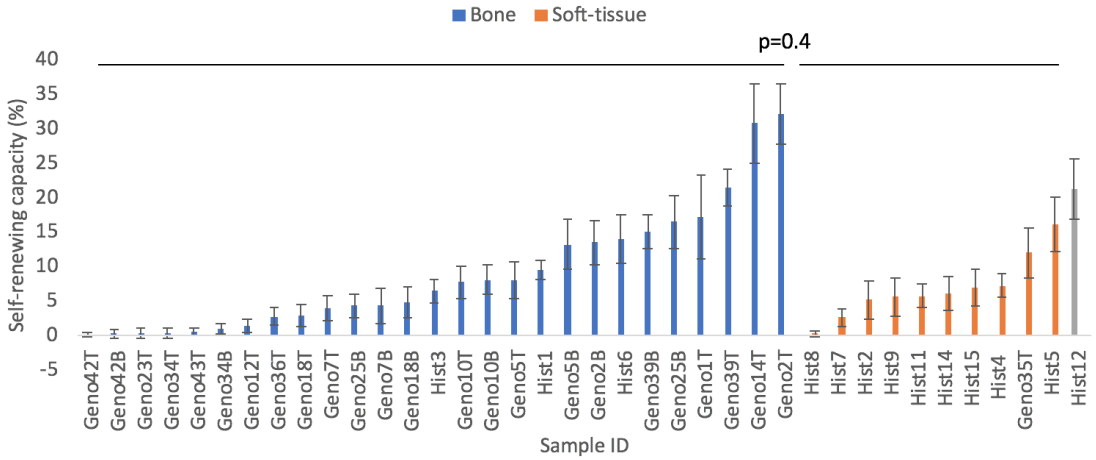


Figure 2.13: Self-renewal capacity of patient-derived ES cultures from a single cell after a 3-week incubation. Results are shown as the mean ± SD of the percentage of wells with self-renewing clones after 3 weeks of incubation from 10 different 96-well plates. Bone tumours (blue) and soft-tissue tumours (orange) shown from lowest to highest self-renewing ability, and no differences were seen across these two sample groups ($p = 0.4$). In gray (His12) sample from unknown site.

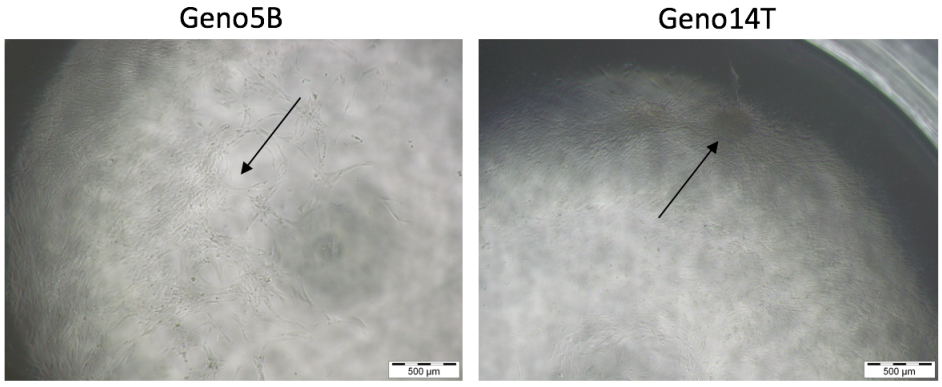


Figure 2.14: Self-renewing colonies after 3 week incubation. Single cells that can self-renew are able to produce a cell populations after 3 week of incubation. 2 patient-derived ES cell cultures are shown as example. Arrows indicate self-renewing cells. Scale bar = 500 μm.

2.3.6 ES patient-derived cells show heterogeneity on the migration index

Another characteristic of cancer cells is their ability to migrate and metastasise. Therefore, the migration index of patient-derived ES cultures was assessed on 39 samples using the 3D spheroid migration assay over a time course of 72 h (Figure 2.16, Table 5.9). The overall migration capacity at 24 h was 10 ± 5 (mean ± SD, range 4–23, n = 39), 21 ± 16 (6–84, n = 39) at 48 h and 37 ± 32 (8–171, n = 39) at 72 h. An example of how the cells migrated out of the spheroid during time is shown in Figure 2.15.

When comparing the bone tumours (19/29 patients) with the soft-tissue tumours (10/29 patients), statistical differences were seen for the migration index at 72 h ($p = 0.02$), where soft-tissue tumours displayed a higher migration index (59 ± 45, 21–170, $n = 10$) than bone tumours (29 ± 21, 8–97, $n = 19$ (Figure 2.16). In addition, when comparing GenoEwing samples (23 ± 13, 8–54, $n = 25$) with historical samples (62 ± 40, 21–171, $n = 14$), historical samples had a higher migration index at 72 h ($p < 0.0001$). Both these differences are linked, as most soft-tissue tumours (9/10) are historical samples (9/14), and therefore, the differences seen are driven by the same cell cultures. Therefore, this data could suggest that tumours from the soft-tissue have an enhanced ability to migrate, suggesting that they might be more prone to metastasise.

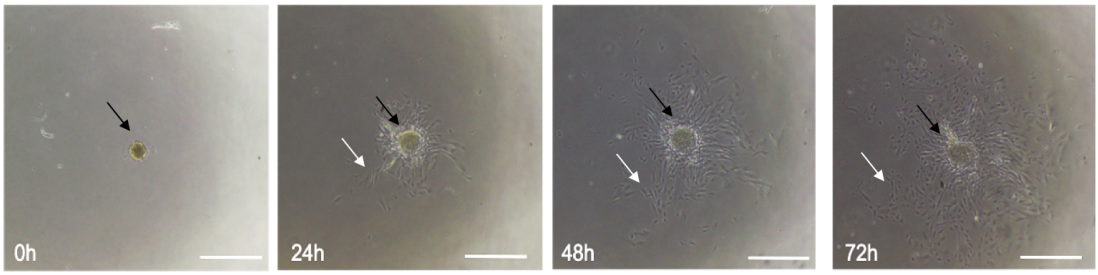


Figure 2.15: Migration index over a 72 h time course. Geno5B is shown as an example. Black arrow = spheroid, white arrows = migratory cells. Scale bar = 200 μm.

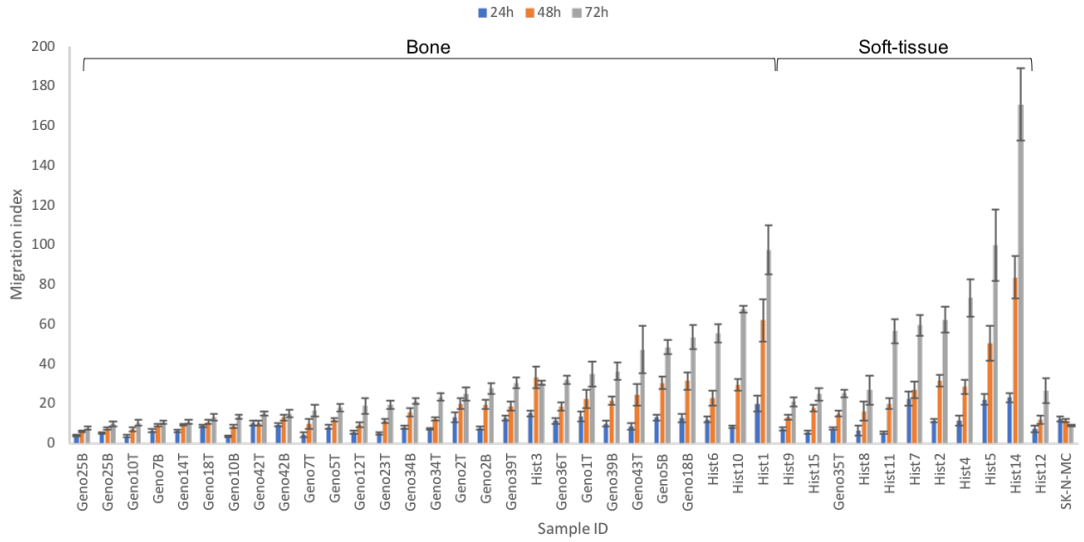


Figure 2.16: Migration index over time on patient-derived ES cultures was analysed using a 3D spheroid migration assay. Migration index of the successfully cultured ES samples ($n = 39$) is shown over a time course of 72 h. Results are shown as mean ± SEM of 3 independent repeats. Bone and soft-tissue tumours shown from lowest to highest migratory index. Hist12 from unknown site; and SK-N-MC ploted as positive control.

2.3.7 Different degrees of sensitivity to cytotoxic drugs are seen in ES cell cultures

The patient-derived ES cell cultures were evaluated for their resistance to doxorubicin and vincristine cytotoxic drugs, standard chemotherapies used for the treatment of this malignancy in the clinic. This was assessed as the percentage of colonies resistant to these cytotoxic drugs compared to untreated cells, over a range of 1, 5, 10, 25, 100 and 200 nM.

Initially, the drug concentrations only included 1, 25, 100 and 200 nM, but as data showed a significant decrease in the colony number from 1 to 25 nM for both drugs ($p < 0.0001$), 2 extra concentrations were introduced in-between: 5 and 10 nM.

The sensitivity to doxorubicin and vincristine was evaluated in 35 ES patient-derived cell cultures (Table 2.3.7, Table 5.9) and showed that on average, ES cell cultures had a 45±26 % (mean ± SD, range 10–134 %, $n = 35$) of colonies resistant to 200 nM doxorubicin and 50±34 % (8–189 %, $n = 35$) to 200 nM vincristine. Therefore, a high heterogeneity on the sensitivity to these cytotoxic drugs was seen across ES patient-derived cell cultures, with some being highly sensitive (i.e. Geno5B, Table 2.3.7) and

others being less sensitive (i.e. Hist1, Table 2.3.7). An example of how the colony number decreased with an increase in the concentration of cytotoxic drug is shown in Figure 2.17.

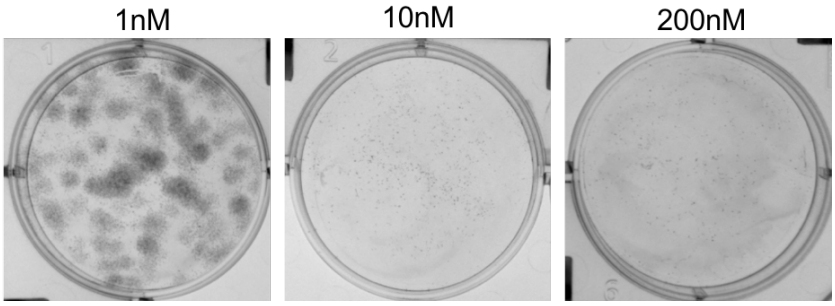


Figure 2.17: Number of resistant colonies to increasing concentration of doxorubicin. As an example, the colonies (stained with crystal violet) at 1, 10 and 200 nM of doxorubicin for Geno5B.

From these results, the EC50 values (concentration of drug which would induce a decrease of 50% on the number of cell colonies) was calculated when possible, as it was an indication of how resistant the samples are to the evaluated cytotoxic drugs. From the total of 35 samples evaluated for doxorubicin, the EC50 could be obtained in 71 % of cases (25/35), which was 10±13 nM (range 2–70 nM, Table 2.3.7). Out of the 10 samples that did not provide a clear EC50, 80 % (8/10) were samples that had only been evaluated on 1, 25, 100 and 200 nM, further indicating that the incorporation of the 2 extra drug concentrations (5 and 10 nM) were important to successfully analyse the sensitivity. A similar profile was seen for vincristine. The EC50 could be obtained for 49 % (17/35) of samples, and was 13±30 nM (range 2–130 nM, Table 2.3.7). Again, a high percentage of the samples that could not provide an EC50 value (72 %, 13/18) had only been evaluated on 1, 25, 100 and 200 nM of vincristine.

The wide range of results indicate that the ES cell cultures derived from patient tissue show an heterogeneity on their response to cytotoxic drugs. Interestingly, a high correlation was seen between the number of resistant colonies at 200 nM of doxorubicin and vincristine ($R = 0.92$) as well as the EC50 for both drugs ($R = 0.81$). This correlation between both drugs is consistent with their mechanism of action, as both target processes important in cell division although acting on different pathways (Section 1.1.2).

ID	Doxorubicin (nM)						EC50	Vincristine (nM)						EC50
	1	5	10	25	100	200	nM	1	5	10	25	100	200	nM
Geno1T	104	—	—	44	41	34	11	102	—	—	33	32	34	—
Geno2B	84	—	—	26	22	19	3	116	—	—	19	16	15	—
Geno2T	85	—	—	20	19	15	3	97	—	—	12	12	11	4
Geno5B	93	—	—	16	13	10	5	90	—	—	11	9	8	4
Geno5T	100	—	—	41	39	33	8	104	—	—	35	36	34	—
Geno7B	135	—	—	45	43	27	13	127	—	—	35	31	26	—
Geno7T	96	—	—	46	45	36	6	106	—	—	35	40	37	—
Geno10B	176	—	—	67	71	59	—	137	—	—	57	59	50	—
Geno10T	121	—	—	58	61	51	—	122	—	—	60	58	53	—
Geno14T	96	—	—	35	34	29	5	99	—	—	26	28	26	—
Geno18B	136	—	—	77	80	72	—	131	—	—	75	76	68	—
Geno18T	103	—	—	71	79	46	—	122	—	—	75	89	71	8
Geno23T	250	218	288	265	210	134	—	285	218	237	219	194	189	—
Geno25B	113	52	45	34	32	16	4	120	32	31	23	20	20	4
Geno25B	205	92	97	77	78	55	15	191	72	70	56	62	67	4
Geno34B	128	93	93	71	77	57	10	133	86	79	62	66	58	5
Geno34T	105	69	73	67	67	46	4	94	65	68	59	80	68	—
Geno35T	107	59	61	44	40	21	6	125	64	43	28	26	25	5
Geno36T	87	61	123	90	89	62	—	99	77	87	70	86	72	—
Geno39B	130	72	59	35	37	24	6	130	60	44	32	37	30	4
Geno39T	176	113	99	52	52	34	11	185	94	57	36	32	29	6
Geno42B	96	41	39	32	33	19	3	96	31	31	28	30	27	2
Geno42T	77	31	29	26	23	19	2	69	28	36	33	33	28	—
Geno43T	88	56	52	47	48	30	3	101	57	59	52	58	56	3
Hist1	104	—	—	74	64	47	70	71	—	—	52	59	53	—
Hist2	105	—	—	52	56	45	10	101	—	—	50	50	50	—
Hist3	125	—	—	66	68	49	13	92	—	—	61	70	66	—
Hist4	102	—	—	73	77	49	—	104	—	—	76	74	66	12
Hist5	86	—	—	36	34	27	4	75	—	—	27	26	26	2
Hist6	94	—	—	103	108	94	—	108	—	—	102	84	66	130
Hist7	85	—	—	79	94	88	—	126	—	—	106	138	114	—
Hist8	99	—	—	33	28	26	8	99	—	—	36	35	27	7
Hist9	181	111	124	95	101	65	19	226	130	112	94	86	82	—
Hist12	107	98	71	70	78	66	—	137	100	89	74	65	61	10
Hist15	129	—	—	81	67	57	20	123	—	—	54	48	44	15
TC-32	87	17	14	8	0	0	2	111	25	16	17	0	0	4

Table 2.6: Response to doxorubicin and vincristine evaluated by number of resistant colonies compared to vehicle control. Table summarises the number of colonies resistant to different concentrations of drug (1–200 nM) compared to the vehicle control (in %). Last column for each drug shows the EC50 values calculated from these drug responses (in nM). — = data not available for that concentration or EC50 not possible to calculate. Cell cultures investigated $n = 35$.

When comparing the percentage of resistant colonies at 200 nM on bone and soft-tissue tumours, no differences were seen in terms of sensitivity to either doxorubicin (46±29 % vs 49±22 %, $p = 0.76$) nor vincristine (54±40 % vs 55±30 %, $p = 0.93$). Similar results were seen when comparing the EC50 values for doxorubicin (10±15 % vs 11±7 %, $p = 0.88$) and for vincristine (16±38 % vs 8±5 %, $p = 0.65$). Therefore, this suggests that the location of the primary tumour does not have a link with the resistance profile of the ES cells.

2.3.8 Correlations between the different *in vitro* evaluations

After obtaining the data from the different *in vitro* assays used to profile the patient-derived ES cell cultures, the relationship between these variables was investigated. The self-renewing ability showed an anti-correlation with the resistance to vincristine ($R = -0.44$) and doxorubicin ($R = -0.31$), suggesting that samples with lower self-renewing ability have higher resistance to the cytotoxic drugs. This agrees with the mechanism of action of these drugs, as they act upon fast-dividing cells, and therefore, cells that can divide and proliferate to form colonies from a single cell would be more sensitive to these. The migration index was not associated to either of the other variables (self-renewing $R = 0.01$, doxorubicin $R = 0.09$ and vincristine $R = 0.05$), suggesting that it is an independent characteristic of the cells, not related to the ability to self-renew or the sensitivity to cytotoxic drugs.

Based on the processing method to obtain ES cell cultures, some patient samples originated both a tissue macerate-derived cell culture and a bone digest-derived cell culture (ES cell cultures from the bone and the soft-tissue section of the collected sample). Throughout the different evaluation of the patient-derived ES cell cultures, no differences were seen between processing method. This indicates that even being derived from different sections of the sample, where the microenvironment might play a different role, these cells behave in the same manner *in vitro* or have the same genotype. Similarly, cell cultures obtained through the GenoEwing project and the historical samples did not show differences of profile in most evaluations, suggesting that the origin and processing of the samples does not have an effect on the *in vitro* profile.

Finally, ES patient-derived cultures from bone ES tumours and soft tissue ES tumours were compared on the different investigations, as it could be that the primary site of the tumour has a correlation with the cell phenotype *in vitro*. However, no differences were seen between these in all evaluations except migration, suggesting that the tumour site does not have an effect on the behaviour of the cells, or at least this is lost *in vitro*. The only assay that showed differences between bone and soft-tissue tumours (which was also the only assay to display differences between GenoEwing and historical samples) was the migration index (bone vs tissue $p = 0.02$, GenoEwing vs historical $p < 0.0001$).

Because most soft-tissue tumours are from the historical cell cultures, both these results are linked. Based on the statistical differences, it seems that the historical samples are driving the migration index differences, and not the tumour location.

2.3.9 Survival analyses

2.3.9.1 Evaluation of clinical data and disease progression for ES patients

In order to evaluate if the *in vitro* characteristics of patient-derived ES cell cultures correlated with the patients clinical outcome, complete minimal clinical data was obtained from 85 % of ES patients (34/40; last clinical information collected June 2018). However, for 8 of these 34 patients, it was not possible to generate a patient-derived cell culture, and therefore, no lab data (phenotype) could be correlated to the clinical profile (Table A3). Therefore, only 26 samples were used for the survival analysis, as these had the relevant clinical data and phenotypic characterisation data (clinical data summarised in Table A3).

The median age at diagnosis of ES patients was 12 years (range 0.5–24 year, $n = 26$) and the ratio male:female was 1 : 1.3 (9 male:12 female, $n = 21$). Metastatic disease at diagnosis was present in 46 % of patients (12/26), of which 42 % located in the lungs only and 58% in other areas (33 % involved lungs and other sites and 25 % had no lung involvement). When evaluating disease progression, 69 % of ES patients had a relapse or the disease progressed (classified as event; 18/26) with a median time to event of 430 days (430 ± 335 days, range 86–1,477, $n = 18$). The overall survival status (patients that are still alive) indicated that 35 % of ES patients did not survive (9/26), being the median time to event (death) of 1158 days (1158 ± 1297 days, range 86–3,882 days, $n = 9$). In comparison, the time last seen for ES patients that were alive was 1935 days (1935 ± 2581 days, range 174–7,397 days, $n = 17$).

Following this, the different clinical factors were evaluated to see how they correlated with disease progression for this dataset ($n = 26$). The age at diagnosis was not associated with survival either for EFS (p-value = 0.06, Hazard-ratio (HR) 0.9, 95% confidence interval (CI) 0.8–1, $n = 26$) or OSR (p-value = 0.08, HR 0.9, CI 0.8–1.0, $n = 26$). Similarly, the patient's gender was not indicative of clinical response (EFS: p-value = 0.39, HR 0.6, CI 0.2–1.8; OSR: p-value = 0.96, HR 1.0, CI 0.2–3.9, $n = 21$). The tumour's primary site has been linked with outcome

(Section 1.1.2), but when evaluated in this dataset, no association was seen when comparing bone tumours to soft-tissue arisen tumours (EFS: p-value = 0.84, HR 0.9, CI 0.3–2.4, $n = 26$; OSR: p-value = 0.67, HR 1.4, CI 0.3–5.5, $n = 26$).

Another factor that has been associated with poor outcome in ES patients is the presence of metastasis at time of diagnosis (Section 1.1.2). When this was evaluated in this small cohort, no associated was seen at EFS (p-value = 0.2, HR 1.9, CI 0.7–4.9, $n = 26$) nor OSR (p-value = 0.46, HR 1.7, CI 0.4–6.3, $n = 26$). Still in relationship to metastatic disease, it has been found that patients with lung metastasis do better than patients with other metastatic sites at time of diagnosis (Section 1.1.2). When localised disease, lung metastasis and other metastatic sites were compared, a significant association was seen between other metastatic sites than lung and worse survival both at EFS (p-value = 0.03, HR 1.3, CI 1.1–10.9, $n = 26$, Figure 2.18A) and OSR (p-value = 0.04, HR 1.7, CI 1.1–28.1, $n = 26$, Figure 2.18B), thus confirming what has been described in the literature.

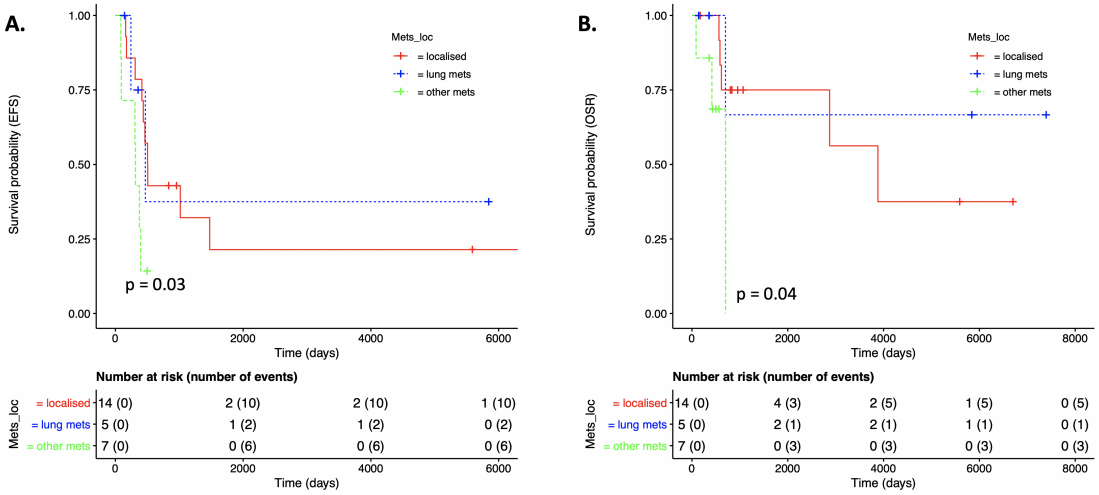


Figure 2.18: Presence of non-lung metastasis at diagnosis is indicative of worse outcome.Patients with non-lung metastasis (green) do worse ($p = 0.03$) than patients with lung metastasis (blue) or localised disease (red) when evaluating EFS (**A**) and OSR (**B**). Results obtained by cox model for factor variables.

2.3.9.2 Assessment of *in vitro* profile of ES samples to predict outcome

Patient-derived ES cell cultures were evaluated on different cancer-related phenotypes: ability to self-renew from a single cell, to migrate from a spheroid and their resistance to conventional cytotoxic drugs used in a clinical setting

(Section 2.1). Therefore, the differences seen in the characterisation of the patient-derived samples were evaluated alongside the clinical data for the ES patients, in order to identify if any of the assays performed could inform on the disease progression for ES patients.

Firstly, the ability to self-renew from a single cell was evaluated in comparison to the clinical data. No association to the patient's outcome was seen for EFS (p-value = 0.95, HR 1.0, CI 1.0–1.1, $n = 25$) nor OSR (p-value = 0.39, HR 0.9, CI 0.9–1.1, $n = 25$). To further investigate this phenotype of cancers cells, the ability to propagate ES-CSCs from these patient-derived cell cultures was investigated, in case this was an indicator of outcome for these patients (compared to the degree of self-renewal ability they had). However, no association was seen for this phenotype to ES patient's outcome (EFS: p-value = 0.95, HR 1.0, CI 0.3–2.7, $n = 25$; OSR: p-value = 0.65, HR 0.7, CI 0.2–3.0, $n = 25$).

The migration index evaluated at 72 h was not indicative of outcome either for EFS (p-value = 0.42, HR 1.0, CI 0.98–1.0, $n = 26$) nor OSR (p-value = 0.42, HR 1.0, CI 0.97–1.0, $n = 26$). Moreover, the *in vitro* resistance to different cytotoxic drugs has been linked to the resistance of tumour cells in patients, and therefore, can be an indication of the patient's response to treatment and subsequent outcome (Section 1.1.2). However, for this ES cohort of 26 patients, the number of resistant colonies to 200 nM doxorubicin compared to the untreated control was not an indicator to the patient outcome (EFS: p-value = 0.84, HR 1.0, CI 0.98–1.0, $n = 23$; OSR: p-value = 0.87, HR 1.0, CI 0.97–1.0, $n = 23$). A similar result was observed when evaluating the resistance to 200 nM vincristine (EFS: p-value = 0.42, HR 1.0, CI 0.99–1.03, $n = 23$; OSR: p-value = 0.36, HR 1.0, CI 0.99–1.03, $n = 23$).

When investigating the expression of CD99 by ICC (negative, low or high expression of CD99), no association to the patient's outcome was seen for EFS (p-value = 0.34, HR −0.7, CI 0.1–2.1, $n = 26$) nor OSR (p-value = 0.36, HR −1.0, CI 0.04–3.3, $n = 26$). The data from flow cytometry evaluation of CD99 expression was investigated too, and this confirmed the lack of associated between the expression of CD99 and the patient outcome for this ES cohort (EFS: p-value = 0.22, HR 1.0, CI 0.97–1.01, $n = 17$; OSR: p-value = 0.07, HR 0.97, CI 0.9–1.0, $n = 17$).

2.4. Discussion

In this study I have shown the ability to generate patient-derived ES cell cultures from ES samples and how these can be profiled *in vitro* to better understand the ES tumorigenesis. In addition, I have been able to study how the different phenotypic data correlates with the clinical data of the patient's behind the cell cultures generated. Although most of the variables were not indicative of the patient's outcome, it proved to be a good system to investigate the role of the different tumorigenic characteristics of cells in the cancer process. I believe that with a bigger sample number (only 26 samples were used in this dataset) with complete clinical and phenotypic data, it would have proven to be more informative of any possible associations.

CSC are characterised by their ability to self-renew, differentiate and an enhanced tumorigenic growth and proliferation, giving rise to tumour progression and relapse (Schatton et al., 2009). For this reason, the self-renewing ability of the patient-derived ES cell cultures was evaluated, thus exploiting an intrinsic characteristic of CSC profile. ES cell cultures showed a proportion of cells that had the ability to self-renew from a single cell, with variability across patients, suggesting heterogeneity on the self-renewing phenotype. When this was correlated with clinical data, it was not informative of the patients outcome. Moreover, when the ability to propagate ES-CSCs from the different samples studied on the self-renewing assay was investigated, no association with the clinical data was seen either. Different groups (Awad et al., 2010; Cornaz-Buros et al., 2014; De Vito et al., 2012; Riggi et al., 2010; Suva et al., 2009; Wahl et al., 2010) are investigating the role and phenotype of ES-CSCs in order to fully understand their implication in ES progression. Therefore, it is important to increase the sample size in this study and evaluate if the ability to self-renew or the ability to generate ES-CSCs can be informative of the clinical outcome. Moreover, it would be fundamental to fully characterise these ES-CSCs in order to increase our knowledge around these cells and see if a different phenotype is associated to these cells. This last aim would help in the development and identification of novel targeted therapies to target these cells, as suggested and investigated by other groups in the ES field (Cornaz-Buros et al., 2014; De Vito et al., 2012).

The sensitivity to two cytotoxic chemotherapies used for the treatment of ES (doxorubicin and vincristine) was also assessed *in vitro*. Due to the nature of the assay used initially (only using 4 different concentrations instead of the 6 introduced later on), it was not possible to obtain an accurate EC50 value for all samples. This made difficult the evaluation of the resistance profile of the different patient-derived cell cultures, and the percentage of the number of resistant colonies at the highest concentration (200 nM) compared to the untreated control was used. Interestingly, a high correlation was seen between the sensitivity to these two drugs *in vitro*. This can be explained by the fact that both are targeting important pathways on cell division, thus affecting fast-dividing cells. It would be interesting to evaluate if using both drugs simultaneously on these cultures, the EC50s can be reduced, thus indicating that these cytotoxic drugs act synergistically. However, when investigating if the sensitivity to these cytotoxic drugs (colonies resistant at 200 nM) was correlated with the patient's response, no effect was seen for either drug. It was not possible to include the EC50 values in this investigation due to the low number of data for this variable. Future studies should aim to increase the sample size and data available for the EC50 values in order to further investigate if any correlations on the patient's response in the clinic are seen. Such studies would provide a mean to identify those patients that would response better or worse in front of a treatment strategy, and tailor each patient to the suitable treatment modality. For this reason, groups in the ES field are moving into preclinical models more informative of the patient's profile, such as PDX, in order to fully understand the mechanisms behind relapse and what treatment strategies would be more effective for ES treatment. For instance, Dowless et al. (2018) recently published a preclinical evaluations of a CDK4/6 inhibitor, abemaciclib, to show good response in ES cells and xenograft ES, but to further confirm their results, used PDX mouse models showing that this inhibitor was effective in reducing tumour growth too. Similarly to the treatment response, an association was seen between higher amounts of necrotic tissue at resection (data from the clinic) and better survival. This data agrees with previous studies were good response to treatment (i.e. higher necrosis) was found to be a good predictor for positive outcome (van Maldegem et al., 2012). Therefore, these patient-derived cell cultures are informative of patients outcome. Further

investigation of the molecular profile between both datasets (high and low sensitivity) might provide novel predictive biomarkers for ES patients stratification and personalised medicine.

Another phenotype that was evaluated in ES patient-derived cell cultures was the migration index, which is an indication of the capacity of cancer cells to move away from the tumour bulk, in order to invade and metastasise to other areas (Hanahan and Weinberg, 2011). In this study, all ES cell cultures showed migration ability, with variation across samples. When this data was correlated to the clinical data of the patients, it was not informative of the patients outcome. Other studies in ES have shown a relationship between migration capacity and aggressive phenotypes. Krook et al. (2014) investigated the migratory capacity of ES cells using transwell chambers, which assesses the ability of cells to invade when exposed to different conditions, and saw that cells exposed to stress had higher levels of migration. In another study, Garofalo et al. (2011) used an *in vivo* mouse model to investigate the ability to metastasise of ES cells, showing that cells with more aggressive phenotype induced higher levels of metastasis. Therefore, these methodologies differ from the one used in this study, a 3D migration assay, where the ability of ES cells to migration from a core spheroid is evaluated, where no effects of the microenvironment are taken into consideration (contrary to the other studies). Future studies should therefore try to mimic those conditions, either by performing the 3D spheroid assay in models more resemblant of the tumour site (i.e. scaffolds) or by investigating the ability to migrate in *in vivo* models.

The expression of the CD99 cell surface marker, characteristic of ES phenotype, was assessed by ICC and flow cytometry. Interestingly, an heterogeneous profile was seen across the different patient-derived cell cultures, with different levels of expression. The CD99 oncogenic involvement in ES has been widely studied (Manara et al., 2018), with its function in ES pathogenesis being linked to migration, cell proliferation and growth (Kreppel et al., 2006; Rocchi and Manara, 2010). However, when the levels of expression was investigated in relation to the clinical profile of the patient's, no association to outcome was seen either for ICC nor flow cytometry data. Moreover, as described by Kreppel et al. (2006), CD99 expression is linked to the migration capacity. However, when these two

phenotypes were correlated in the present study, no correlation was seen. This discrepancy could be generated from the cells used for the study (Kreppel et al. used established cell lines) or the method to evaluate the migration index (Kreppel et al. transwell chambers). Importantly, Kreppel et al. only assessed the ability to migrate for 24 h, and therefore, it could be that the longer periods of investigation (here I have used 72 h) result in a loss of this relationship. It is important to point out that differences in the expression of CD99 were seen between ICC and flow cytometry. These could be explained by the use of different antibodies between both techniques: ICC uses the clones 12E7 whereas flow cytometry is based on the TÜ12 clone. Moreover, CD99 has been linked to different cellular processes, including cell cycle, that could be fluctuating while on culture (Manara et al., 2018). It would be interesting to evaluate the different clones used in both techniques (if available) and see if these differences are indeed due to the antibody used or due to biological cell cycle fluctuations.

When the survival of the ES patients was compared between localised disease, lung metastasis or other metastasis at diagnosis, an association to worse outcome was seen for those patients with metastatic disease other than in the lungs (EFS $p = 0.03$, OSR $p = 0.04$), whereas lung metastasis had a similar survival rate than localised disease patients. This observation in this dataset agrees with the literature (Gaspar et al., 2015), as lung metastasis is associated with better outcomes. The other clinical factors investigated in this cohort did not show any correlation with the clinical outcome, such as the age at diagnosis, the gender of the patient, the presence of metastasis or the primary site of the tumour. Amongst these, the age of diagnosis for ES patients has been frequently associated to outcome, where younger ages are associated to worse outcomes (Cotterill et al., 2000; Gupta et al., 2010; Karski et al., 2013; Lee et al., 2010). However, due to the small dataset of this study, not many patients with older ages were enrolled, and this study was difficult to perform. It the collection of samples would involve other surgical centres, maybe the patient's enrolled in this study would embrace a wider population, making these analyses easier.

In the present study, a high proportion of ES patient-derived cell cultures express high levels of c-kit, in agreement with other ES studies (González et al., 2004;

Scotlandi et al., 2003). This tyrosine kinase receptor has been extensively linked to different fundamental biological process that are the basis of tumorigenesis (Miettinen and Lasota, 2005), leading to the development of c-kit inhibitors to treat cancers with abnormal expression of this receptor such as gastrointestinal stromal tumours (GIST) (Kitamura and Hirota, 2004). Due to not having c-kit data for all patients investigated in the survival analyses, this was not possible to be investigated as done for the other variables. Other studies have shown that c-kit expression is not informative of prognosis in a cohort of 100 ES patients (Scotlandi et al., 2003). However, the high levels of expression, although not informative of prognosis, could be exploited for targeted therapies using the already-approved c-kit inhibitors. In this line, different studies have evaluated the use of tyrosine kinase inhibitors of ES cells, showing growth inhibition and synergy with cytotoxic drugs (Lissat et al., 2012). However, when evaluated in a clinical setting, patients did not show the same response seen *in vitro* (Bond et al., 2008). Novel tyrosine kinase inhibitors are emerging, which are showing promising results *in vitro* and in mouse models (Wang, 2009). Therefore, it could be that the expression of c-kit in ES patients might be beneficial once the appropriate targeted therapy has been found.

The presence of an EWSR1 fusion was investigated by FISH and RT^A-PCR. All cell cultures evaluated for FISH showed presence of break aparts in the *EWSR1* gene, indicating that a fusion was present. However, evaluation by RT^A-PCR failed to identify a fusion in 69% of the cell cultures. The first explanation for these differences is that FISH evaluates the break apart on the DNA level whereas the PCR is based on the RNA level, and therefore, could be that there is a decrease in expression at the transcriptional level. Another difference between both techniques is that FISH can evaluate the presence of a break apart cell by cell, whereas PCR is pooling together a population of cells. Therefore, if the fusion levels are too low on the patient-derived samples, PCR might no be able to detect the fusion. This could be re-evaluated by increasing the concentration of RNA used for the RT^A-PCR. One last factor that could influence the results is the presence of fusions other than with FLI1 (type 1 and 2) or ERG, as therefore, the PCR would not be able to detect them, whereas FISH would still detect a break apart. Different

studies have shown the presence of rarer forms of the fusion, where EWSR1 partners with other members of the ETS family or even other TET family members being part of the translocation instead of EWSR1 (Pinto et al., 2011). Therefore, in order to identify if the fusions present in the samples that did not show results by PCR are indeed novel fusion variants, other techniques that allow for unknown variants, such as RNA-seq, could be used.

Due to the processing method used to obtain cell cultures from the ES samples, some tissues generated both a soft-tissue derived cell culture and a cell culture from the bone fraction of the tumour. This was originally designed to evaluate if the cells that grew inside the bone compartment had a different behaviour than those of the mass residing outside the bone. As shown throughout the results section, no differences were seen between both cell cultures. This would suggest that the role that the bone microenvironment could play on the ES cell is lost *in vitro*, outside the bone influence. Different groups are trying to mimic the bone niche to better understand the biology of cells, to replicate the microenvironment the tumour would be inside a person. For instance, Villasante et al. (2016) developed a bioengineered 3D tumour model that would mimic the cell-to-cell interaction. They demonstrated that these 3D models are more resemblant to the ES patient's gene expression profile compared to monolayer cultures. Therefore, seeing how important the microenvironment has become in the tumorigenesis (covered in Section 1.1.4), it is fundamental to move to these bioengineered scaffolds to better understand the behaviour of cancer cells. In addition, most likely reflecting the necrotic profile of the chemotherapy treated cells, only a small proportion of the post-treatment samples produced successful cell cultures in comparison to the success rates for the diagnostic biopsies. This is a caveat that comes from using patient-derived samples from resection, as the tissue will be often damaged by the treatment the patient had received (necrotic tissue), and will be harder to obtain viable cell cultures. To better overcome this obstacle it would be necessary to start with larger tissue masses in order to increase the chance of retaining some viable cells.

It is important to mention that due to the small sample size, the results have to be evaluated carefully. Some links have been possible to draw between the phenotype

and the clinical data, informing on the subset of patients that would do worse in the clinical setting. However, larger patient cohorts are needed in order to accurately define these associations and identify the subsets of ES patients that would do worse in the clinic or which cellular characteristics are behind tumorigenesis. Such studies might help define the ES patients that have worse prognosis, and further investigations could try to identify a panel of markers that can distinguish between these and other ES patients with better outcome. Such findings would need then to be confirmed in a validation cohort. This is specially important in ES, as it is a heterogeneous disease, as has been observed by the phenotypic evaluations in this study and by the results on other groups, where differences across patients are linked to different outcomes (Scotlandi et al., 2009; Tirode et al., 2014). Therefore, the ability to separate patients into different risk groups according to their tumour phenotype might result in patient stratification and subsequent personalised treatment, which could ultimately lead to improved outcomes.

3. Characterisation of ES-CSCs

3.1. Introduction

Emerging evidence from ES studies suggests that CSCs (ES-CSCs) could be behind the poor outcome for ES patients due to their increased drug resistance, ability to self-renew and re-grow the tumour mass, thus resulting in relapse and tumour progression (Trucco and Loeb, 2012; Visvader and Lindeman, 2012). Therefore, it is fundamental to identify and isolate these ES-CSCs from patient samples, in order to better understand their intrinsic gene expression profile, their role in ES progression and implications on the outcome of patients. This might ultimately lead to the development of targeted therapies against these cells, with the aim to improve outcomes for ES patients. Similar to other malignancies (Al-Hajj et al., 2003; Galli et al., 2004; Hermann et al., 2007; Lapidot et al., 1994; Schatton et al., 2008; Singh et al., 2004), different groups in the ES field have been trying to identify and characterise the ES-CSCs in order to better understand their importance in ES malignancy and how best to target them (Awad et al., 2010; Komuro et al., 2007; Suva et al., 2009; Wahl et al., 2010; Yang et al., 2010).

There is an agreement in the importance of profiling the CSCs within tumours in order to better understand the tumorigenesis. However, differences across CSCs have resulted in an ongoing debate around the best approach to identify this population (Schatton et al., 2009), which has reached the ES field (Trucco and Loeb, 2012). The approaches used to distinguish the CSC from the overall cancer population include descriptive assays, those that assess for qualities expected in CSCs; and functional assays, which evaluate the ability to behave as CSCs (covered in Section 1.2). Previous studies in the ES field have exploited the expression of markers (i.e. CD133) (Cornaz-Buros et al., 2014; De Vito et al., 2012; Riggi et al., 2010; Suva et al., 2009; Wahl et al., 2010), an enhanced ALDH activity (Awad et al., 2010) or higher drug efflux (Komuro et al., 2007; Yang et al., 2010), all previously associated to the CSCs. However, as these are not intrinsic characteristics that define a CSC, such as their ability to self-renew or differentiate, these methods could be isolating only a fraction of the ES-CSCs present within tumours. Therefore,

in this study I have used a functional assay (Section 2.1) exploiting the self-renewing ability of these cells, in order to identify and isolate ES-CSCs from ES patient-derived samples. Using this approach produces cells that reflect the heterogeneity of ES and provides a unique panel of cells for future studies. Although phenotypic assays can inform on the differences between these ES-CSCs and the ES cell population, more stringent analyses are needed in order to identify candidate genes for targeted therapies.

Over the last decades, the methods to study gene expression have evolved dramatically, going from low throughput techniques such as quantitative PCR to more high throughput methodologies that enable evaluation of genome-wide profiles (Kukurba and Montgomery, 2016). Although microarrays are still widely used in biomedical research, a set of limitations are associated with their use (Byron et al., 2016). As microarrays are based on the hybridisation to already-known targets, a prior knowledge of the genomic profile is needed. In addition, there is a limitation in quantifying the expression of rare expressed genes or highly abundant genes, limiting the output data from this technique (Kukurba and Montgomery, 2016). Moreover, the identification of novel splice variants, insertions and deletions or fusion events by microarrays would only be possible if specific probes are included for the purpose (Byron et al., 2016). Altogether, Next-Generation Sequencing (NGS) provides a more powerful approach to study the differences across samples, especially when there is no prior knowledge of the expression profile this should have. Due to the low mutation burden described in ES tumours (Tirode et al., 2014), genome-based analyses would not be ideal in this tumour type. Moreover, using exome sequencing, thus only analysing exons (and not introns), we would lose all the information related to splice variants. This is an important factor, as splice variants in ES have been shown to be informative of outcome (Roundhill and Burchill, 2013). Therefore, the most informative method to investigate the expression profile of ES-CSCs and the cells from which they were derived is transcriptomic analysis (RNA-seq). With this approach, the diversity of expression of the different genes can be obtained, including splice variants and previously unknown events, being able to identify heterogeneity of gene

expression across different samples, ultimately being able to link certain expression profiles to distinct phenotypes (Byron et al., 2016; Wang et al., 2009).

The need of complex bioinformatic analyses to accurately manipulate, normalise and compare the generated data are important factors when using this technology. However, the increase in popularity of these technologies and the rise of studies that want to better understand the heterogeneity across samples has been accompanied by an improvement in the bioinformatic tools required (Costa-silva et al., 2017). The importance of choosing a software approach over another often relies on preferences of the user or the biological needs for each study. The first important step on the analysis of the data is the alignment process, where the obtained reads are aligned to a reference genome in order to map each read to a specific gene or region (Kukurba and Montgomery, 2016). Although different packages have been developed to perform this alignment step (i.e. BowTie2, STAR, Kallisto), each uses a different approach to perform this process (Costa-silva et al., 2017). For instance, BowTie2 cannot map splice junctions, and therefore, cannot handle RNA data (Kukurba and Montgomery, 2016). Kallisto is a pseudo-aligner, thus being based on the probability of a read aligning to a region of the transcriptome, instead of actually aligning it accurately (Bray et al., 2016). This results in fast quantification of the reads that match the different areas of the transcriptome, but by not using an aligning method *per se*, it is less accurate. Moreover, as it uses transcriptomes as reference, it does not allow for intron variability and can only recognise transcripts that have already been described and are part in the transcriptome reference used (Bray et al., 2016). STAR (Splice Transcripts Alignment to Reference), contrary to Kallisto, aligns the reads to the specific locations in the genome, and can identify splice variants and novel junctions, contrary to BowTie2 (Dobin et al., 2013). Therefore, for this study, the reads obtained from RNA-seq of ES-CSCs and ES cells were aligned using STAR. Moreover, STAR allows for the identification of fusions and events across non-consecutive regions of the genome using the add-on package STAR-Fusion (Haas et al., 2017).

As well as choosing the alignment method based on the requirements of each experiment, the downstream data normalisation and differential expression

analyses are important (Engström et al., 2014). A recent study comparing different programs for data normalisation indicated that differences across the programs to normalise RNA-seq data could have important implications in the results obtained (Costa-silva et al., 2017). This is a fundamental step, as it involves the normalisation of the read number across the genome for the different samples, to be able to compare across them and ensure that the differences seen are indeed biological and not technical or analysis artefacts (Zhang et al., 2014). For this reason, the bioinformatics field has and is still trying to define the best methodology to do this, comparing the different packages that are emerging and validating the data generated (Costa-silva et al., 2017). Amongst the most common tools, we can find DESeq2. This is a differential gene expression analysis method to determine the differences in gene expression from high throughput sequencing data between samples, based on a negative binomial distribution method (Love et al., 2014). When this method was compared to similar software packages used in the literature, it came amongst the best in terms of accuracy, sensitivity and precision (Costa-silva et al., 2017; Seyednasrollah et al., 2013), and was therefore selected for the present study.

3.1.1 Aims of this chapter

The aims of this chapter are to:

1. Propagate ES-CSCs in culture
2. Study their phenotype based on expression of ES markers, migration index, self-renewing ability and sensitivity to chemotherapy drugs, in order to compare to the ES cells from which they are derived.
3. Develop a bioinformatics pipeline for the analysis of RNA-seq data from ES patient-derived cell cultures.
4. Identify candidate genes differentially expressed in ES-CSCs that could be exploited for targeted therapies.

3.2. Materials and Methods

3.2.1 Phenotypic characterisation of ES-CSCs

3.2.1.1 Propagation of single cell-derived populations

Single cell-derived self-renewing colonies were isolated as part of Chapter 2 (Section 2.2.11). Once cell colonies reached a confluency of 70% in the 96-well plate where the assay was performed, they were harvested (Section 2.2.3) and moved to a 24-well plate with a final volume of 1.5 ml of media per well. After 72 h of incubation, wells were checked to determine cell growth (Equation 3.1) and when reached 70 % confluency moved to a T25 and cultured (Section 2.2.3). At this stage, cell populations were either pelleted (Section 2.2.3) or continued in culture on T75 culture flasks. The propagated single cell-derived self-renewing colonies are referred as ES-CSCs throughout this study.

$$\text{self-renewing capacity} = \frac{\text{wells with cells}}{\text{total of seeded wells}} \times 100 \tag{3.1}$$

3.2.1.2 CD99 expression on ES-CSCs

The expression of CD99 was assessed by ICC (Section 2.2.9) and when possible, by flow cytometry (Section 2.2.10).

3.2.1.3 Presence of a *EWSR1* fusion on ES-CSCs

The presence of the characteristic *EWSR1-ETS* fusion was investigated by RT^A-PCR (Section 2.2.6) and FISH (Section 2.2.8) when possible.

3.2.1.4 Self-renewing ability of ES-CSCs

The ability to self-renew from a single cell was assessed on the ES-CSCs populations using the self-renewing functional assay (Section 2.2.11).

3.2.1.5 Migration assay on ES-CSCs

The migration capacity on the ES-CSCs populations was assessed using a 3D migration assay (Section 2.2.12).

3.2.1.6 Sensitivity to doxorubicin and vincristine of ES-CSCs

The sensitivity to the cytotoxic drugs doxorubicin and vincristine was assessed using a colony assay (Section 2.2.13).

3.2.2 Total RNA sequencing of ES patient-derived cell cultures, ES-CSCs and established cell lines

3.2.2.1 RNA extraction and quality assessment for RNA-seq

Total RNA was extracted from cell pellets (Section 2.2.4) using the miRNeasy Micro kit (Section 2.2.6.1). The concentration of RNA was determined by Nanodrop (Section 2.2.6.2) and the quality of the RNA was assessed using the Agilent Bioanalyzer Pico 6000 kit (Section 2.2.6.3). For RNA-seq, only samples with RIN > 9 were used, as this indicated the RNA extracted was of good quality.

3.2.2.2 Library preparation

Total RNA libraries were prepared initially by the NGS facility (Dr. Sally Fairweather) and later in the CCRG lab by Dr. Elizabeth Roundhill. A total of 1 μg of RNA from ES patient-derived cell cultures or established cell lines was prepared using the TruSeq Stranded Total Library Preparation kit (Illumina®, Little Chesterford, UK) following manufacturers instructions (summarised in Figure 3.1). Importantly, indexing adaptors specific for each sample were used using the recommended combination from Illumina®, in order to pool samples and later tell them apart (Figure 3.1). The quality of the obtained libraries was confirmed using the D1000 ScreenTape with D1000 Reagents kit (Agilent Technologies) by identifying if the expected size fragments were present (260bp). In addition, the 2200 TapeStation System (Agilent Technologies) confirmed if the prepared samples were free from contaminants from the preparation protocol. Libraries were quantified using the Quant-iT™dsDNA Assay Kit (Thermo Fisher Scientific) according to manufacturer's instructions, in order to verify that the same amount of sample was used for RNA-seq. A total of 100 ng per sample for every 4 samples were pooled for loading (4 samples per lane).

3.2.2.3 Total RNA sequencing (RNA-seq)

Total RNA-seq was carried out at the on-site Next Generation Sequencing (NGS) Facility (Leeds Institute of Biomedical and Clinical Sciences, University of Leeds). It was performed on the Illumina HiSeq3000 (Illumina®) with a coverage of 4 samples per lane and using paired end sequencing. In order to amplify for all RNAs, random primers were used on the amplification step. A total of 151 cycles were performed. Some data used for this study (all cell lines with the exception of

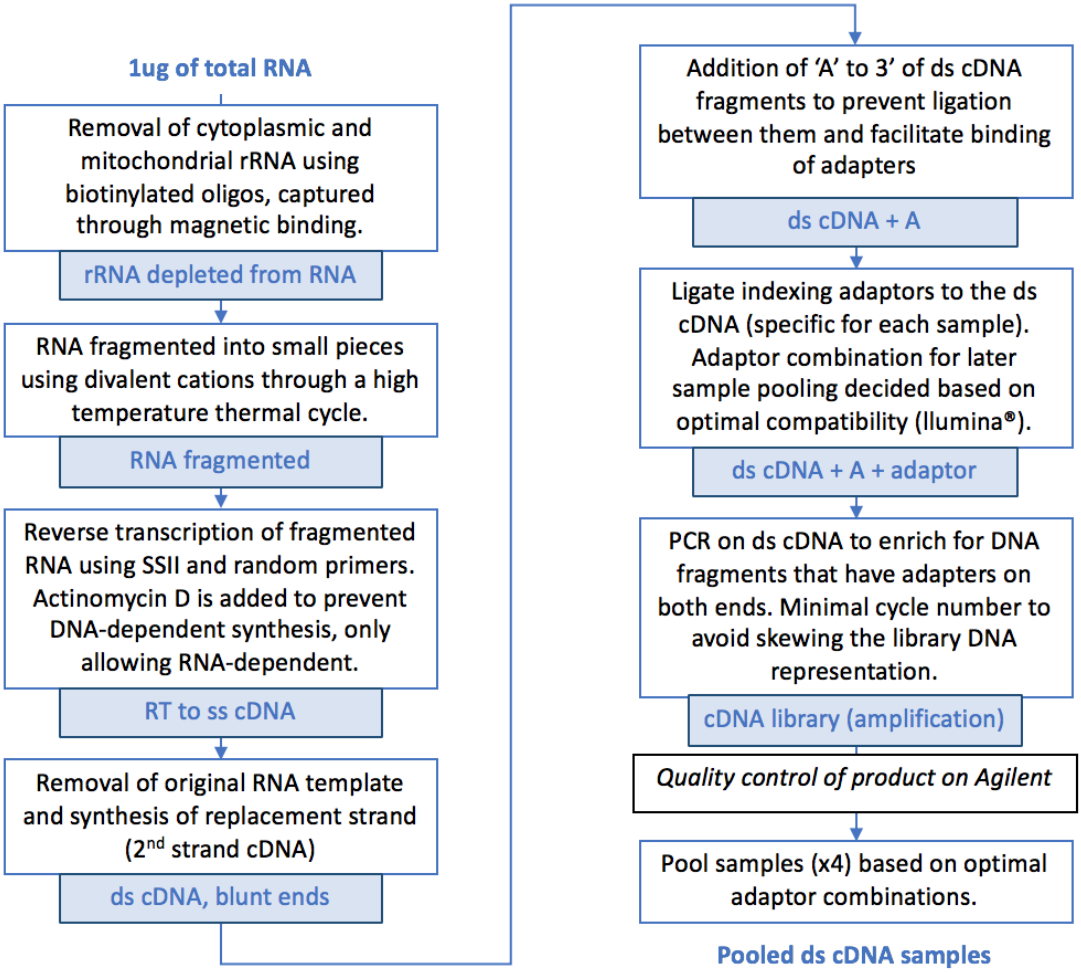


Figure 3.1: Library preparation process using the TruSeq Stranded Total Library Preparation kit (Illumina®). The quality control step was performed on the D1000 ScreenTape (Agilent) and 2200 TapeStation System (Agilent) (Section 3.2.2.2).

TTC-466 and also ES-CSC population Hist1.s) had previously been obtained by RNA-seq using the Illumina®HiSeq2500 (Illumina®) with a coverage of 11 samples per lane and using paired end sequencing and a total of 101 cycles. Raw data was obtained in FASTQ file format.

3.2.3 Bioinformatic analyses of RNA-seq data

The bioinformatic analyses of RNA-seq data were performed using R language (R, 2017). This bioinformatics approach to analyse the data generated is summarised in Figure 3.2 and each section described in detail below.

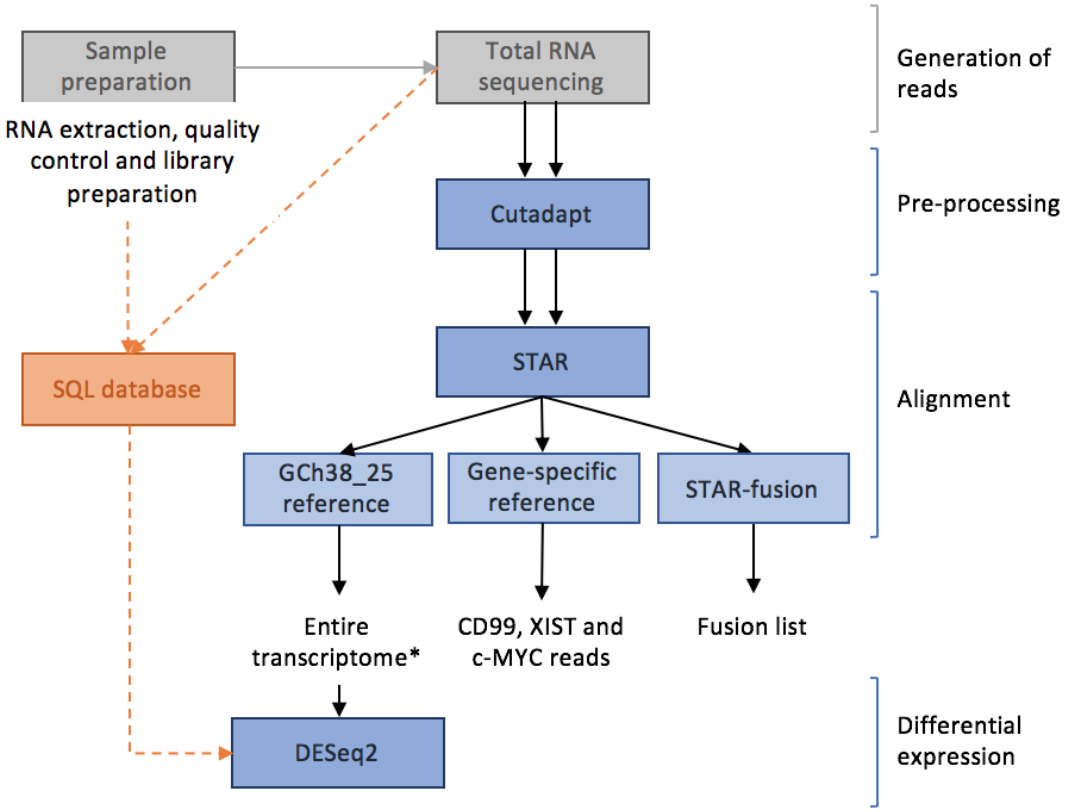


Figure 3.2: RNA-seq bioinformatics pipeline for ES samples. Sample preparation (Section 3.2.2.1, Section 3.2.2.2) and RNA-seq (Section 3.2.2.3) in gray. The different bioinformatics sections to process the reads (Section 3.2.3.2), alignment of reads using the standard reference genome (Section 3.2.3.3), the gene-specific strategy (Section 3.2.3.5) or STAR-fusion to identify fusions (Section 3.2.3.4) or the differential expression profile (Section 3.2.3.6) is shown in blue. Alongside the data analysis, metadata was stored in a database to facilitate analyses throughout the pipeline (orange). Double arrows are indicative of steps where both reads are processed (if only one read, it is discarded), until the alignment, when both reads are merged.

3.2.3.1 Database storage

The Structural Query Language (SQL; (Chamberlin and Boyce, 1976)) was used to create a database that stored the location of the FASTQ files (Section 3.2.2.3), the information related to each RNA-seq run (i.e. machine used, number of cycles, etc) and the metadata for each sample sequenced (i.e. patient-derived culture or cell line, processing method used, sample ID, etc). This facilitated the management of the different information for each sample in order to analyse the RNA-seq data (Figure 3.2).

3.2.3.2 Pre-processing of the raw data

Paired raw data (Section 3.2.2.3) was processed in order to remove any reads with low quality, which could have an effect on the downstream analyses. If a read was removed, the paired read was removed as well, always having either both retained or both discarded. For this, the Cutadapt package (Martin, 2011) was used (Figure 3.2). Low quality reads (defined with a Phred score (Ewing et al., 1998) lower than 20 indicating a base call accuracy lower than 99 %) were removed. Reads shorter than 20 nt were removed as they would provide a less reliable alignment later on. Finally, the adapters used for the library preparation (Section 3.2.2.2) and the polyadenylated (poly(A)) tails were removed. The default setting for Cutadapt to look for such contaminating sequences is to perform 1 iteration, which would mean that any reads that still have adapter or poly(A) contamination would be later on discarded. In this study, the iteration number was increased to 10, allowing for the removal of multiple contaminating sequences. The Fastqc package (Andrews, 2010) was used to check the quality of reads before and after pre-processing (pre-trimmed and post-trimmed respectively). To facilitate this process, the fqtools package (Droop, 2016) was used for easier fastqc file manipulation.

3.2.3.3 Alignment of reads

Post-trimmed reads were aligned to the GENCODE (The ENCODE Project Consortium, 2011) human reference 38 release 25 (GCh38_25) using the aligner STAR (Dobin et al., 2013) (Figure 3.2). This aligner program uses a 2-pass alignment process, where all reads are aligned against the reference genome (1-pass alignment). For any read aligning to multiple regions, only the best alignment was retained in order

to avoid ambiguous alignments. Therefore, reads that would align to more than one region of the genome, with the same alignment accuracy (same score) would not be able to provide a single unique alignment, and would be discarded from the downstream analyses. Moreover, any novel junctions (not described in the reference genome used) identified in the 1-pass alignment were then included alongside the canonical regions and a new splice junction database built to which all reads were re-aligned to (2-pass alignment). This method allows for the discovery of any novel splice junctions that might by highly frequent in one sample (enough to be detected) but rare in other samples (not enough coverage to be identified). A matrix of all samples with read coverage per gene was obtained.

3.2.3.4 Identification of fusions

Using standard alignment packages such as STAR (Section 3.2.3.3), events that involve different areas of the genome (i.e. fusions) would be discarded, as the alignment would include 2 areas of the reference that should not be together. To be able to identify fusions, the STAR-fusion package was used (Haas et al., 2017) (Figure 3.2). With this, all reads that have been discarded from STAR (i.e. ambiguous alignments or unknown reference) are analysed against a reference of all possible combinations of genes and exons within the genome. Therefore, any reads that contain a splice junction for a fusion would now be aligned accurately. The output generated is a list of all identified translocations with the number of reads aligning and the details of the alignment, such as bases on each gene or exact coordinates of the involved genes.

3.2.3.5 Gene specific reference

As discussed in Section 3.2.3.3, any read that can be aligned to multiple regions of the genome will be discarded, as it cannot provide an unambiguous alignment. This is the case of CD99, which is present on the pseudo-autosomal region (PAR) on chromosomes X and Y (Figure 3.3A), and therefore, has 2 possible alignments with the exact same sequence (Figure 3.3B). To overcome this, a gene specific reference was built: CD99 sequence ± 1000 bases at each side to account for any possible flanking read pairs (Figure 3.2). With this, all post-trimmed reads were aligned to this reference instead of the GCh38_25 reference (Section 3.2.3.3) using the STAR aligner, and as only one possible CD99 sequence was available (no longer both X

and Y chromosomes), CD99 reads would be retained (Figure 3.3C). To investigate the accuracy of this strategy, the gene specific reference was also built for XIST and c-MYC, two genes that could be aligned using the standard GCh38_25 reference (Section 3.2.3.3).

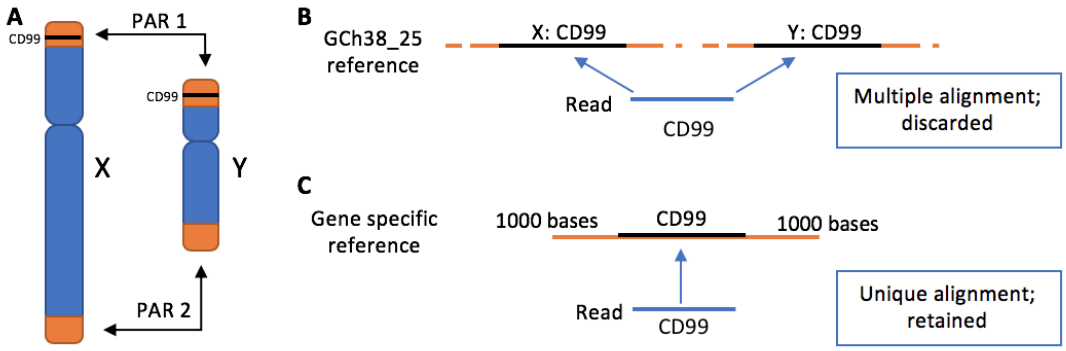


Figure 3.3: Gene specific reference for CD99 alignment. **A.** Schematic representation of the PAR on chromosomes X and Y, and the location of gene CD99 on PAR1. **B.** When using the standard human genome reference (GCh38_25), both chromosomes X and Y are present resulting in two identical copies of CD99. This results in the reads covering CD99 being discarded. **C.** When using the gene specific reference, which consists of the CD99 sequence ± flanking regions, reads covering CD99 can be properly aligned as there is only one possible match.

3.2.3.6 Differential gene expression profile

To identify differences across the sample groups, the differential gene expression was analysed using DESeq2 (Love et al., 2014). In order to group the samples for the different comparisons, an R script was generated to extract relevant samples from the SQL database for analysis in DESeq2 (Figure 3.2). A JavaScript Object Notation (JSON) script was used to define the specific analyses required and the parameters to be used, thus allowing multiple analyses to be performed consecutively. Significant differences were defined as adjusted p-value ≤ 0.001 and absolute $\log_2$ fold change (LFC) ≥ 2. Identified genes were ranked based on their adjusted p-value score (lowest to highest) and LFC (highest to lowest). Scores from both ranks were combined and the genes with the lowest scores (low adjusted p-value and high LFC) were selected as candidate genes. Differentially expressed genes were visualised using the pheatmap package (Kolde, 2019) on the full list of differentially expressed genes obtained from DESeq2. Data was scaled by row using an Eucledian distance matrix and hierarchical clustering.

3.2.3.7 Functional annotation of differentially expressed genes

In order to investigate the relationship between the differentially expressed genes enriched in each sample group, the Database for Annotation, Visualization and Integrated Discovery (DAVID) database was used (Huang et al., 2009*a*,*b*). The full list of enriched genes in each sample group (from the DESeq2 output) based on the LFC was analysed using the Functional Annotation Chart. A background correction was applied by using the full list of genes annotated by RNA-seq prior to differential expression analysis. Analysis was performed using False Discovery Rate correction and p-value threshold as $p < 0.001$.

3.2.4 Reverse Transcription^{A} - quantitative PCR (RT^{A}-qPCR)

The RT^{A} quantitative PCR (RT^{A}-qPCR) validation of the RNA-seq data was carried out by Dr. Elizabeth Roundhill. cDNA was prepared from 10 ng of RNA diluted in RNase-free H_2O as described in Section 2.2.6.4. Once the cDNA was obtained (40 μl total), 20 μl of cDNA was mixed with 80 μl of PCR mix containing sequence specific primers and probes (Table 3.1) and 1× TaqMan Universal PCR Master Mix (Applied Biosystems) in RNase-free H_2O. Samples were then transferred in triplicate (25 μl/well) to a MicroAmp™Optical 96-well Reaction Plate and covered with MicroAmp™Optical Adhesive Film (Applied Biosystems). Plate was then placed on the 7900HT Fast Real Time System (Applied Biosystems) using the SDS 2.4 platform (Thermo Fisher Scientific) and incubated for 2 min at 50 °C, 15 min at 95 °C in order to activate the DNA polymerase. Then, 50 cycles of 15 s at 95 °C to denature the cDNA, 1 min at 66 °C and 90 s at 72 °C for the annealing and extension phase respectively. Data was exported and analysed using the RQ Manager 1.2.1 (Thermo Fisher Scientific) and cycle threshold (C_t) calculated using a threshold of 0.05. In addition, RT- products (samples in the absence of reverse transcriptase enzyme) and H_2O were used as negative controls to control for contaminants. The peptidylprolyl isomerise A (PPIA) was used as the housekeeping gene to be able to normalise and compare expression levels across the different samples. In order to compare RT^{A}-qPCR data with RNA-seq data, value of CD99 expression for both techniques were normalised to PPIA gene expression obtained from RT^{A}-qPCR and RNA-seq respectively (performed by Dr. Elizabeth Roundhill).

Primer or probe	Sequence (5'-3')	Concentration (nM)
CD99 (hs00365982_m1)	ChrX (2,691,133–2,741,309 on GRCh38)	-
PPIA F primer	GGACCCAACACAAATGGTTCC	200
PPIA R primer	CTTTCACTTTGCCAAACACCA	200
PPIA probe	ATGCTTGCCATCCAACCACTCAGTCTTG	100

Table 3.1: Primers and probes used to amplify for CD99 and PPIA. * CD99 was an assay on demand developed by Thermo Fisher, TaqMan™Gene Expression Assays. Sequence is company propriety, present in chromosome X on the positions described in the table. F = forward, R = reverse.

3.2.5 Statistics

Comparisons between ES-CSCs and the parental cell cultures from which they were derived was performed using unpaired t-test (2 datasets) or one-way ANOVA for multiple comparisons (>2 datasets), with significant difference when $p>0.05$. These analyses were performed in GraphPad Prism 6 software, except the descriptive statistics that were performed in Microsoft Excel version 15.24.

The differential expression profile was performed using DESeq2 as described in Section 3.2.3.6. Heatmaps were obtained using the R package Pheatmap (Kolde, 2015) using Pearson correlation and scaling by rows. Correlations calculated throughout the study are based on a Pearson correlation coefficient.

3.3. Results

3.3.1 Characterisation of single-cell derived clones (ES-CSC)

3.3.1.1 Successful culture of putative ES-CSCs

A functional assay was used in Chapter 2 (Section 2.3.5) to assess the ability to self-renew from a single cell in ES patient-derived cell cultures. From the 38 ES patient-derived cell cultures evaluated (Section 2.3.5), 58 % (22/38) produced self-renewing populations (identified as ES-CSCs from now on) that could be moved up from a 96-well plate to a 24-well plate (Table 3.2). The number of ES-CSC populations obtained per sample was 26 ± 21 (range 1–72, total number = 577). Therefore, a high heterogeneity in the ability to form self-renewing colonies was observed. Nevertheless, a direct correlation was seen between the self-renewing ability of the patient-derived cell cultures (Section 2.3.5) and the number of ES-CSCs colonies moved to a 24-well plate ($R^2 = 0.55$; Table 3.2). 82 % of the ES-CSCs colonies moved

to a 24-well plate (473/577) grew after this first cell harvesting; but only 9 % of the total (52/577) were successfully cultured up to a T25 tissue culture flask, and this further decreased in reaching T75 (only 13 ES-CSCs, from 5 different cell cultures from 4 patients).

ID	SR ability (%)	SR population moved	SR population grown	SR population cultured
Hist1	10	11	36% (4)	9% (1)
Hist3	7	30	57% (17)	0
Hist4	7	28	57% (16)	0
Hist5	16	24	71% (17)	4% (1)
Hist6	14	8	38% (3)	0
Hist12	21	72	78% (56)	7% (5)
Hist14	6	6	0	0
Geno1T	17	48	92% (44)	4% (2)
Geno2T	32	48	88% (42)	4% (2)
Geno2B	14	46	74% (34)	2% (1)
Geno5T	8	19	79% (15)	11% (2)
Geno5B	13	69	86% (59)	13% (9)
Geno7T	4	1	100% (1)	100% (1)
Geno7B	4	4	100% (4)	100% (4)
Geno10T	8	7	100% (7)	14% (1)
Geno10B	8	11	100% (11)	0
Geno14T	31	47	100% (47)	21% (10)
Geno25B	17	9	78% (7)	0
Geno25B	4	24	100% (24)	8% (2)
Geno35T	12	17	100% (17)	12% (2)
Geno39T	21	24	100% (24)	8% (2)
Geno39B	15	24	100% (24)	29% (7)
Total	13±8%	577	473	52

Table 3.2: Success in culturing self-renewing (SR) cell populations from the self-renewing assay. The SR ability (%) column describes the self-renewing ability evaluated in Section 2.3.5. The following columns indicate the number of SR cell populations moved to a 24-well plate; the number of these that successfully grew (percentage from the total moved and number between brackets); and the number of SR cell populations that were successfully cultured up to a T25 (percentage from the total moved and number between brackets). SR = self-renewing.

Therefore, not all the ES-CSCs colonies obtained in the functional assay could be successfully propagated in culture. Whereas some samples showed a high ability to form ES-CSCs (Geno5B; Table 3.2), others did not form any (Hist14; Table 3.2) or formed initially but they did not reach a T25 culture (Hist3; Table 3.2). The differences observed in the self-renewing ability of the patient-derived cell cultures could be explained by a different profile of the ES tumours across patients, having a higher percentage of ES-CSC within. However, to explain the differences in propagating the isolated ES-CSCs in culture, it could be that a subset of ES-CSCs

does not have the ability to grow *in vitro*, as they might need factors from the microenvironment and other cells within the tumour to grow. Altogether, culturing the self-renewing clones from a 96-well plate to a T75 procured putative ES-CSC patient-derived cell population that could be studied in comparison to the profile of the cells they were derived from (identified as parental populations from now on).

Similarly to the patient-derived cultures these ES-CSCs were derived from (Section 2.3.1), the 13 ES-CSCs populations grew as adherent monolayers and resembled the parental cultures in terms of morphology (Figure 3.4).

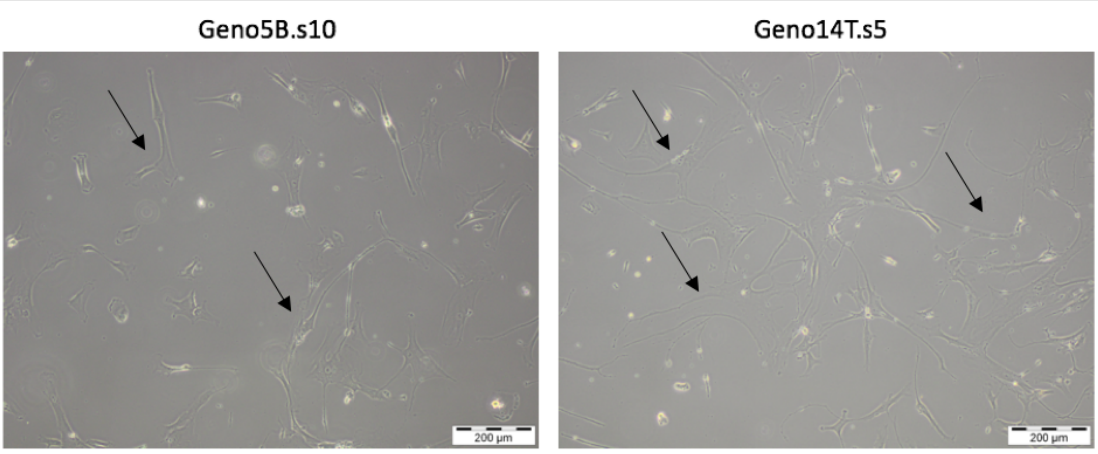


Figure 3.4: Morphology of single cell-derived ES cultures growing as substrate adherent monolayers. As an example, Geno5B.s10 and Geno14T.s5 are shown. Arrow indicate characteristic ES-CSC cell morphology.

3.3.1.2 ES-CSCs express CD99 and have a *EWSR1* fusion

As these ES-CSC were derived from a single cell, it was interesting to explore their expression of CD99 and the presence of a fusion compared to the parental cultures they were derived from, to see if they conserved the same expression or differences were seen across ES-CSCs and the parental cultures.

The expression of CD99 was assessed in the 13 successfully cultured ES-CSCs (Table 3.3, Figure 3.5). In all cases, ES-CSCs expressed CD99: 54 % at high level of expression across all cells (++) and 46 % at low level of expression across all cells (+). Interestingly, heterogeneity of expression could be seen amongst ES-CSCs derived from the same parental populations. For example, Geno14T (parental) had a low level of CD99 expression, whereas the different ES-CSCs derived from it had

ID	CD99 expression		EWSR1 fusion		c-kit expression	
	ICC	Flow (%)	FISH	RT-PCR	c-kit	mixed exp.
Hist12	+	8.5	mix	type 1	+	not
Hist12.s3	++	—	—	type 1	—	—
Hist12.s5	++	—	—	ND	—	—
Hist12.s9	+	—	—	type 1	—	—
Geno5B	+	77.4	mix	type 1	+	not
Geno5B.s3	+	—	mix	ND	+	not
Geno5B.s4	++	—	mix	ND	++	mix
Geno5B.s8	++	—	mix	ND	++	not
Geno5B.s10	++	4.3	—	ND	+	not
Geno5B.s11	++	—	—	—	+	not
Geno14T	+	95.6	mix	type 1	++	mix
Geno14T.s3	++	1.0	yes	type 1	++	not
Geno14T.s5	+	0.9	—	ND	++	not
Geno14T.s7	+	—	mix	ND	++	not
Geno39T	++	29.5	mix	ND	++	mix
Geno39T.s1	+	—	mix	—	++	not
Geno39B	++	34.5	mix	type 2	++	mix
Geno39B.s11	+	—	yes	—	++	not

Table 3.3: Assessment of CD99 expression, *EWSR1* fusion and expression of c-kit in ES-CSC compared to the parental cultures they are derived from. In blue, the profile of the parental cultures is shown above the derived ES-CSCs. Expression of CD99 by ICC (0 = negative expression, + = low expression, ++ = high expression) and flow cytometry (as percentage of cells expressing CD99). Presence of an *EWSR1* translocation was assessed by FISH (yes = cells have break on *EWSR1* gene; mix = mixture of cells with and without break apart) and RT^{A}-PCR (type 1 = EWSR1-FLI1 type 1, type 2 = EWSR1-FLI1 type 2, ERG = EWSR1-ERG, ND = not detected). Expression of c-kit by ICC using the same scale as CD99. The presence of a mixture of low and high expression is indicated as mix or not when expression was homogeneous.

either low expression ($n = 2$) or high expression ($n = 1$). Therefore, not all derived ES-CSC behaved in the same manner as the cultures they are derived from.

Moreover, the evaluation of the CD99 expression profile by flow cytometry could be performed in 3 ES-CSC only, derived from 2 different parental cultures (Table 3.3). Data indicated that the expression of CD99 using this methodology was much lower than the parental cultures, contrary to what ICC data indicated (Table 3.3). For instance, Geno5B (parental) had a 77 % of cells expressing CD99, whereas Geno5.s10 (ES-CSC) had only 4 %. A similar result was obtained for Geno14, where the parental had a 96 % of CD99 expression, and the Geno14.s3 (ES-CSC) only 1 %. Therefore, in some ES-CSCs (i.e. Geno14T.s3) ICC indicates high expression of CD99 (++), whereas flow cytometry only showed a 1% of positivity. This difference between the flow cytometry and ICC characterisation agrees with data from Chapter 2 (Section 2.3.2.1), where in some cases, no correlation of expression was seen between these techniques.

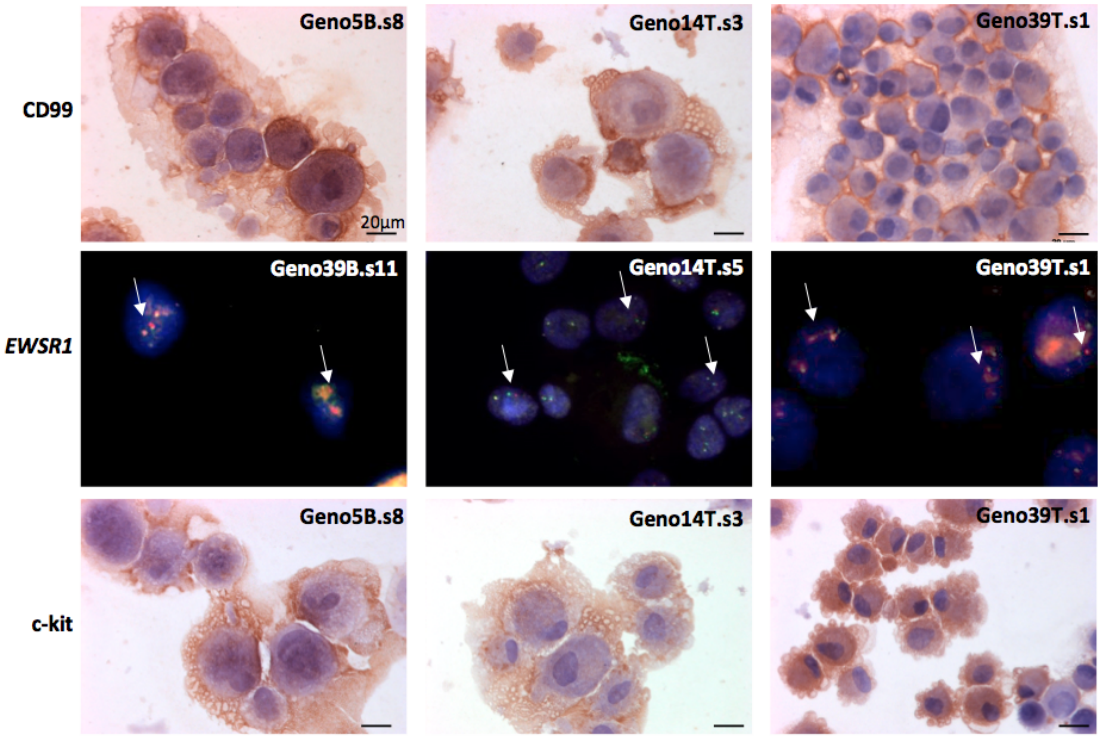


Figure 3.5: Expression of CD99, *EWSR1* and c-kit in ES-CSCs. Expression of CD99 is shown by ICC. Presence of an *EWSR1* break apart by FISH. Yellow probe = *EWSR1* intact probe, green probe = 5' of *EWSR1* probe, red probe = 3' *EWSR1* (involved in the fusion). Therefore, red and green probe separated indicative of *EWSR1* fusion (arrows). Expression of c-kit by ICC. 3 different ES-CSC are shown as an example for each evaluation. Scale bar in ICC = 20 μm.

Regarding the characterisation of the presence of the EWSR1-ETS fusion in ES-CSCs, this was performed in 10/13 samples. RT^A-PCR data identified a fusion in 3 ES-CSCs (Hist12.s3, Hist12.s9, Geno14T.s3) all agreeing with the fusion type identified in the parental sample (Table 3.3). For the other 8 ES-CSCs, no fusion was detected by this technique. Moreover, the presence of the fusion could be investigated by FISH in 7/13 samples, in all of which break apart on the EWSR1 probe (indicative of a EWSR1 break where the fusion is described) was seen (Table 3.3, Figure 3.5). When putting both datasets together, in samples that the fusion was not detected by RT^A-PCR, FISH could detect a break apart, confirming the presence of a fusion. The low detection of fusion by RT^A-PCR compared with FISH is in agreement with data from the previous chapter, where in samples that no fusion could be detected by RT^A-PCR, FISH was able to detect it.

3.3.1.3 ES-CSCs express c-kit

Similarly to the patient-derived cultures, the expression of the tyrosine kinase receptor c-kit was investigated in ES-CSCs, as this has been proposed as a candidate for targeted therapies in ES (Section 1.1.3). Data from 10/13 ES-CSCs, derived from 4 different cultures, indicated that all evaluated ES-CSC express c-kit (30 % at low levels and 70 % and high levels; Table 3.3, Figure 3.5). Similarly to CD99 expression by ICC (Section 3.3.1.2), some ES-CSC derived from the same parental culture (Geno5B) showed heterogeneity of expression, where 3 ES-CSC had low expression and 2 had high expression. Therefore, heterogeneity in the expresion of c-kit can be seen across ES-CSCs from the same sample, and across different ES-CSCs.

There was no correlation between CD99 and c-kit expression, as the 3 ES-CSCs that had low levels of c-kit expression (Geno5B.s3, Geno5B.s10, Geno5B.s11), only 1 had low levels of CD99 as well (Geno5B.s3); similarly, the samples that had low levels of CD99 (Geno14T.s3, Geno14T.s5, Geno39T.s1, Geno39B.s11), 3 had high levels of c-kit expression.

3.3.1.4 Self-renewing ability is decreased in ES-CSCs

The ability to self-renew from a single cell was assessed on the ES-CSCs, to evaluate if these cells, after being isolated and propagated in culture, still retained the ability to self-renew from a single cell.

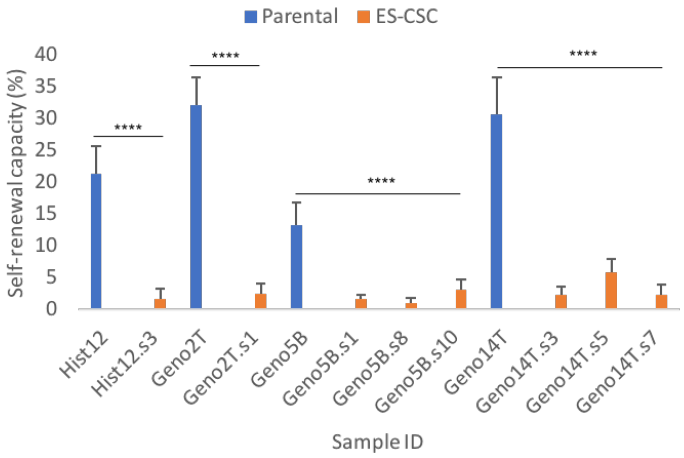


Figure 3.6: Self-renewal capacity of ES-CSC compared to the parental cultures after a 3-week incubation. Results are shown as the mean ± SD of the percentage of wells with self-renewing colonies after 3 weeks of incubation of 10 96 well-plates evaluated. Blue = parental, orange = ES-CSC. Statistical differences (paired t-test) between parental cultures and ES-CSC evaluated (****: $p < 0.0001$).

This was evaluated in 62 % of ES-CSCs (8/13), derived from 4 different parental populations, showing a self-renewing ability of 2.5±1.5 % (range 0.9–5.8 %; Figure 3.6). When compared to the self-renewing ability of the parental populations (24±9 %, range 13–32 %, $n = 4$), a decrease was seen on the ES-CSCs capacity to grow from a single cell ($p < 0.0001$, Figure 3.6). In addition, the few single-cell colonies that were obtained could not be maintained in culture, as they did not grow enough to be moved up in culture.

3.3.1.5 ES-CSCs show heterogeneity in the migration index

The ability to migrate from a 3D spheroid was assessed in 92 % of ES-CSCs (12/13) derived from 4 different parental populations, as had been done for the parental culture (Section 2.3.6). The migration index of ES-CSCs at 72 h was 13±7 (range 3–31; Figure 3.7).

When compared to the migration index of the 4 parental populations (28±16, range 11–49), the ES-CSCs had a lower ability to migrate ($p = 0.017$). However, as shown in Figure 3.7, differences in the migration index were seen across ES-CSCs from the different samples and across ES-CSCs obtained from the same culture. In most cases (9/13) ES-CSCs have a different migration profile compared to the parental cultures they are derived from. For example, sample Geno5B, all derived clones

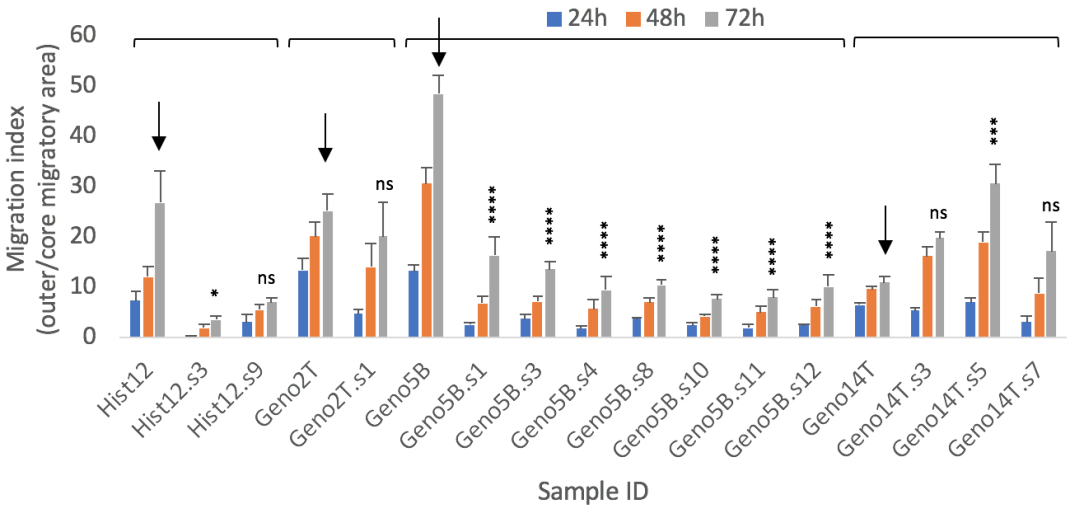


Figure 3.7: Migration index of ES-CSC compared to parental cultures. Migration index of both parental (arrow) and derived ES-CSC (following parental) is shown over a time course of 72 h. Results are shown as mean ± SEM of 2 independent repeats (Hist12.s9 1 repeat; Geno5B.s8 3 repeats). Statistical differences between the ES-CSC and the paired parental are shown as: ns = not significant, * $p < 0.01$, *** $p < 0.0005$, **** $p < 0.0001$.

($n = 7$) had a decreased migratory index compared to the parental population ($p < 0.0001$; Figure 3.7). The same was seen for ES-CSCs derived from Hist12, where the migration index was reduced compared to the parental culture for ES-CSC Hist12.s3 ($p = 0.01$), but not for Hist12.s9 ($p = 0.09$). In contrast, ES-CSCs derived from sample Geno14T showed an increased migratory index compared to the parental culture, only significant for Geno14T.s5 ($p = 0.0005$). Finally, ES-CSC Geno2T.s1 did not show a different migration index to its parental culture (Figure 3.7). Therefore, we can conclude that different migration capacities of ES-CSCs exist within these samples, with some having higher migratory index than the cells they are derived from; and others lower migratory index.

3.3.1.6 ES-CSCs have an increased resistance to cytotoxic drugs

One of the intrinsic characteristics of CSCs is an increased resistance to cytotoxic drugs (Section 5.2). For this reason, the sensitivity to doxorubicin and vincristine was assessed in 46 % (6/13) of ES-CSCs, originated from 2 different parental populations (Geno5B and Geno14T).

When evaluating all ES-CSCs together, the percentage of resistant colonies

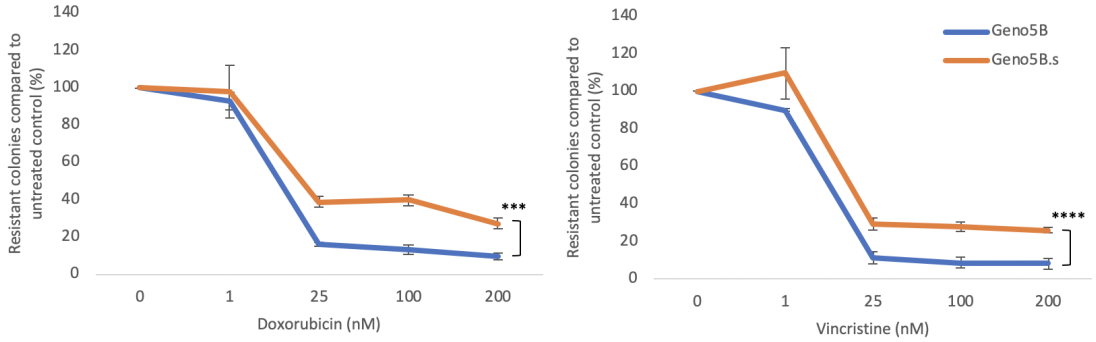


Figure 3.8: Differences in the sensitivity to cytotoxic drugs between ES-CSCs and parental cultures evaluated by number of colonies resistant to increasing concentrations of doxorubicin (right) and vincristine (left) compared to untreated cells. As an example, the parental culture Geno5B (blue) is plotted in comparison to the mean of 3 ES-CSCs derived from this (Geno5B.s4, s8 and s10; as Geno5B.s). Data as mean ± SEM of 2 independent repeats for Geno5B and 1 repeat (3 replicate wells) for ES-CSCs (3 different samples). Statistical differences between all Geno5B ES-CSC and parental culture Geno5B as *** $p = 0.0002$, **** $p < 0.0001$.

compared to the vehicle control was 37±16 % (range 24–65 %) for doxorubicin and 36±10 % (range 24–49 %) for vincristine. To investigate if ES-CSCs had an increased resistance compared to the patient-derived cell cultures they were isolated from, all ES-CSC from the same sample were grouped and compared to parental culture they were derived from. ES-CSCs from Geno5B (Geno5B.s) showed an increased resistance to both doxorubicin ($p = 0.0002$; Figure 3.8) and vincristine ($p < 0.0001$; Figure 3.8), compared to the parental culture (Geno5B). The same was seen for the ES-CSCs originated from Geno14T, which were more resistant to doxorubicin ($p = 0.006$) and vincristine ($p = 0.01$) than the parental culture.

As heterogeneity was seen for the previous evaluations, it was important to understand if ES-CSCs from the same culture showed differences in the resistance profile. Indeed, differences where only seen for Geno5B.s10 for doxorubicin ($p = 0.002$), whereas all Geno5B.s ES-CSC had increased resistance to vincristine ($p < 0.001$). Similarly, for ES-CSC derived from Geno14T, an increase in the doxorubicin resistance was only seen for Geno14T.s3 and Geno14T.s7 ($p = 0.002$ and $p = 0.0004$ respectively); whereas vincristine did not show differences compared to the effect it had on the parental culture.

3.3.2 RNA-seq pipeline validation

3.3.2.1 Quality control of generated reads

The total RNA profile of 46 samples was investigated by RNA-seq. For the analysis, 7 established cell lines (6 ES and 1 NBL), 26 patient-derived cell cultures (2 non-ES) and 13 ES-CSCs derived from patient-derived cultures were included. Due to the sample processing method used to derive patient-derived cultures (Section 2.2.3), the 24 ES patient-derived cell cultures were originated from 17 different patients, as 7 of the samples had both soft-tissue macerate and bone cell cultures. In addition, the 13 different ES-CSC cell populations were obtained from 5 different samples. The variety of samples included in the analysis allowed for comparisons between cell lines and patient-derived cultures; differences between obtaining cell cultures through soft-tissue maceration or bone digestion from the same sample; and differences between the patient-derived cultures and the self-renewing cell populations isolated from them.

Regarding the quality of the data generated, the raw read count per sample was $5.9 \times 10^7 \pm 2.7 \times 10^7$ (mean ± SD, range 1.3×10^6–1.2×10^8). Once these reads were processed to remove any low quality reads and adaptor contamination, a 99±1 % of reads was retained per sample (range 92–100 %). The annotation success (number of reads that could be successfully aligned to the GCh38_25 reference) of these post-trimmed reads was 62±9 % (range 37–82 %). This suggests that a percentage of the reads generated through RNA-seq could not be aligned properly, which could be explained by multiple alignment (thus discarded), unknown reference or not possible to map accurately (aligning to different areas of the reference). Moreover, the cell lines had a higher alignment score compared to the patient-derived cultures (76±10 % for cell lines, 60±6 % patient-derived, $p < 0.0001$).

3.3.2.2 Correlation of gene expression results with clinical information

To test that the output data correlated with the previously-known data from the samples, the expression of 2 characteristic gender-associated genes was investigated and compared to the gender of the patients. For this, the X-inactive specific transcript (*XIST*) and the sex-determining region Y (*SRY*) expression were investigated from the read counts generated. The first gene, *XIST* is present on

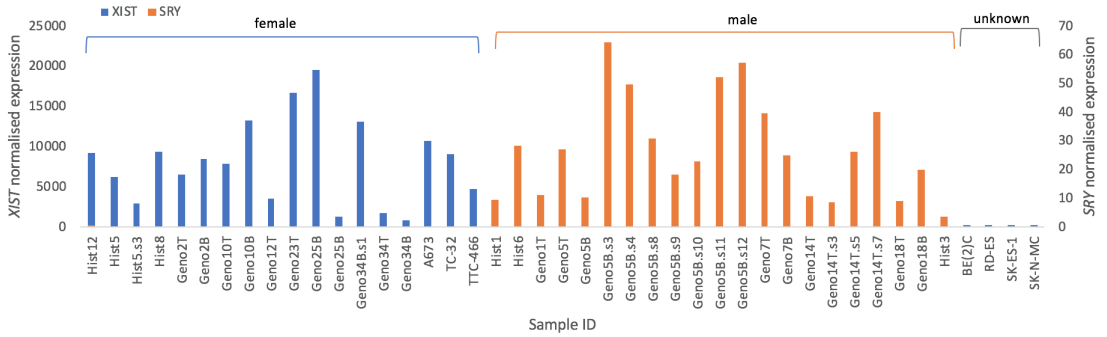


Figure 3.9: Expression of sex-determining genes by RNA-seq to validate output data. The expression of XIST (blue; expressed in females) and SRY (orange; expressed in males) was evaluated on the RNA-seq output data and correlated with the gender of the patients (indicated on top of the graph). The samples for which no expression was seen for any of these genes are plotted in the end of the graph.

chromosome X and plays a major role in the inactivation of the X chromosome in females, and therefore, should be expressed only in females . The second, *SRY*, is a DNA-binding protein that initiates the male sex determination in humans, and therefore, should mainly be expressed in males.

As shown in Figure 3.9, all samples originated from a female expressed *XIST* (blue bars) and all samples from males expressed *SRY* (orange bars), with none of the samples showing ambiguous results (expression of both). Finally, most of the established cell lines (4/6) fall under the category of unknown gender. Although the gender of the patients behind these cell cultures is known, the RNA-seq data suggests absence of expression for both *XIST* and *SRY*. These results agree with previous cytogenetic data from the CCRG, as these cell lines were shown to have a complex karyotype, indicating a loss of chromosomes where these genes would be present. Altogether, these results indicate that the RNA-seq data obtained correlate with the previously known information of the samples, suggesting that both the RNA-seq and the downstream analyses (aligning and data normalisation) provided accurate results.

3.3.2.3 CD99 expression using the gene specific reference

Because expression of CD99 is a characteristic of ES cells (Section 1.1.1), the RNA-seq data provided another method to further evaluate the CD99 expression of the ES patient-derived cultures previously assessed in Section 2.3.2.1.

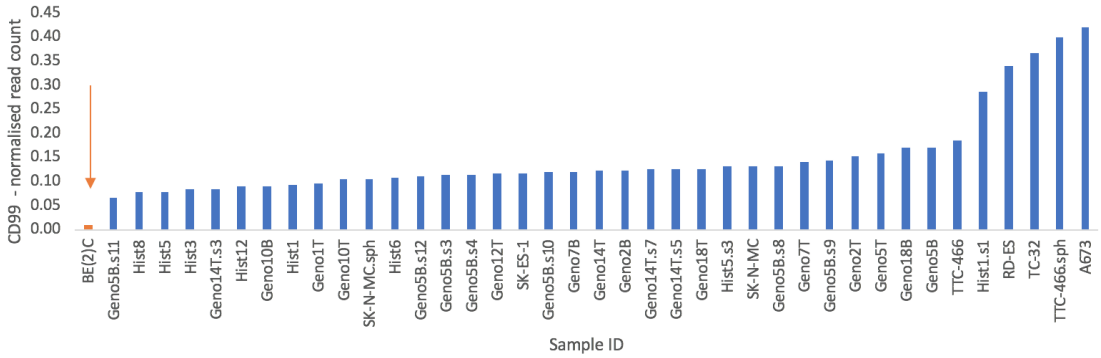


Figure 3.10: CD99 expression by RNA-seq using the gene specific references. Data is shown as normalised read number (reads against CD99 compared to the total number of reads). Samples are shown as ES (blue) and NBL (orange; arrow) in ascending expression of CD99.

The CD99 gene is located in PAR, present both in chromosome X and Y, which results in two identical copies of the gene in the reference genome. Therefore, the aligner cannot map the reads to CD99 unambiguously, thus discarding them. To solve this problem, the gene specific reference strategy described in the methods (Section 3.2.3.5) was used. Using a reference for only the CD99 sequence, it was possible to align all reads unambiguously, as only 1 copy of the gene was present.

Using this strategy, all ES samples showed expression of CD99 (Figure 3.10, blue), whereas the NBL cell line BE(2)C had the lowest levels of expression (Figure 3.10, orange arrow). Interestingly, when correlating the ICC data (Section 2.2.9) with the normalised read count for the 32 samples that had information for both assays, no correlation was seen ($R^2 = 0.13$). This lack of correlation in expression could be due to the different nature of CD99: ICC evaluates protein expression, whereas RNA-seq evaluates RNA levels of expression. Moreover, RNA-seq is measuring the overall CD99 expression in a population, and therefore, if a sample has a mixture of cells with high expression and cells with low expression, this variability will be lost in the read count.

It was important to confirm that the gene specific reference strategy was accurate in the alignment process and the data generated could be trusted. Therefore, 2 genes that could be aligned using the standard GENCODE reference as only one copy was present in the genome were selected. One of these was *XIST*, the gene

that was used for the gender interrogation, as it would give absence or presence of expression. The other gene, *c-MYC*, showed different levels of gene expression, thus being able to investigate if these were consistent across both alignment methods (gene specific reference and GENCODE reference). The read output generated through both strategies was compared and this indicated that the results were the same (XIST: $R^2 = 0.9985$; c-MYC: $R^2 = 0.9997$). Therefore, the gene specific reference strategy was generating accurate results compared to the standard reference, meaning that the CD99 data obtained was accurate.

To further validate if the data generated through the RNA-seq pipeline was validating the biological profile of the samples, the expression of CD99 by RT^A-qPCR was assessed in some cell cultures (performed by Dr. Elizabeth Roundhill). A total of 33 cell cultures were included in the analysis (18 ES patient-derived cell cultures, 6 ES-CSC cultures, 6 ES established cell lines and 3 non-ES cell cultures), as data was available both for RNA-seq and RT^A-qPCR. Normalising both datasets with PPIA expression, so values could be compared, a high correlation was observed ($R^2 = 0.89$), further confirming that the RNA-seq results could be trusted.

3.3.2.4 Identification of EWSR1-ETS fusion including novel variants

One of the advantages of NGS over other methods such as PCR is the ability to identify the expression of transcripts that had previously not been described, such as unknown fusions or isoforms. This is specially important in ES, where the *EWSR1* fusion is a characteristic of this cancer, and different variants and partner genes have been described (Section 1.1.1). Using the STAR-fusion package (Section 3.2.3.4), all the reads that had been discarded from the general alignment (using GENCODE reference) were investigated to determine if any reads could map to 2 different areas of the genome, thus indicating that they were part of a fusion, deletion or insertion. Therefore, all the reads aligned to the fusion could be detected via this method.

The *EWSR1-FLI1* fusion was detected in 6 different samples, 5 of which were established ES cell lines and 1 a patient-derived cell culture (Geno18T). In addition, the *EWSR1-ERG* fusion was identified in the TTC-466 cell line (Table 3.4). Therefore, the characteristic *EWSR1-ETS* fusion was detected in all established ES cell lines

using RNA-seq, confirming the RT^{A}-PCR data. Interestingly, the coverage for these fusions varied from 185 reads falling into the junction for *EWSR1-ERG* on TTC-466 to 31 reads for *EWSR1-FLI1* on RD-ES or only 6 for SKES-1. Therefore, it can be seen that although the RNA-seq can detect these already-defined fusions, the coverage of putting 4 samples per lane on the sequencer can lead to only 6 reads falling into the fusion junction. For example, the patient-derived sample that had an *EWSR1-FLI1* (Geno18T) had only 2 reads falling into the junction, whereas PCR data identified a EWSR1-FLI1 fusion. This difference in the detection levels could be explained by how the fusions are detected in RNA-seq. In order to identify a read as being part of a fusion, the break between both genes (i.e. EWSR1 and FLI1) has to fall in the read, allowing STAR-fusion to identify that event as part of the fusion. However, all the reads that align either to EWSR1 or FLI1, not falling into the junction, will be seen as endogenous genes. For this reason, PCR methods, amplifying for the fusion sequence of interest, will be more sensitive.

Samples	Gene 1 (5'-3')	Chr. 1	Gene 2 (5'-3')	Chr. 2
6	*EWSR1*	22	*FLI1*	11
1	*EWSR1*	22	*ERG*	21
1	*EWSR1*	22	*KATNB1*	16
1	*EWSR1*	22	*CHD4*	12
1	*EWSR1*	22	*PTK7*	6
1	*TEAD1*	11	*EWSR1*	22
1	*SF3B3*	16	*EWSR1*	22
1	*SLC6A6*	3	*EWSR1*	22
1	*CAMSAP2*	1	*EWSR1*	22
1	*ETV5*	3	*EWSR1*	22

Table 3.4: Known and novel EWSR1 fusions identified by RNA-seq. The number of sequenced samples containing each of the fusions is described, as well as the partners genes of *EWSR1* and the chromosome location.

As shown in Table 3.4, 8 other *EWSR1* fusions could be detected in a total of 7 different samples (5 patient-derived cultures, 1 ES-CSC culture and the ES cell line TTC-466). It is important to mention that of these, 5 have *EWSR1* in the 3' region of the fusion, contrary to what is seen for the canonical *EWSR1-ETS* fusions. All these novel *EWSR1* fusions have a coverage of 2–4 reads falling into the fusion junction. To confirm the presence of these fusions in the sample, these novel fusion should be validated by RT^{A}-PCR. Moreover, it would be important to determine if these are

main driver events for these samples, or just secondary events not driving the ES phenotype.

The fusions identified in the patient-derived ES cell cultures detected by RNA-seq were correlated with the RT^A-PCR data from Section 2.3.2.2. From the 6 novel fusions identified in patient-derived cultures, 3 (*CAMSAP2-EWSR1*, *ETV5-EWSR1* and *EWSR1-CHD4*) were detected in Geno7T-B samples, in which no fusion was detected by RT^A-PCR. However, a break apart could be seen when evaluating by FISH. Therefore, it could be that the break apart identified by FISH was for the fusions detected by RNA-seq, suggesting why the EWSR1-FLI1 or EWSR1-ERG fusion were not detected by RT^A-PCR. The other 3 fusions (*SLC6A6-EWSR1*, *TEAD1-EWSR1* and *SF3B3-EWSR1*) were detected in samples where a EWSR1-FLI1 fusion was detected by RT-PCR (Geno14T, Geno18T and Geno2B respectively; Section 2.3.2.2). Therefore, it could be that these are by-stander fusions, present alongside the canonical EWSR1-FLI1.

3.3.3 Differences between patient-derived cell cultures and established cell lines

Before comparing parental cell cultures and ES-CSCs, the different sample groups were analysed to see how different or similar they were. For this, DESeq2 was used to compare all ES patient-derived cultures to the established cell lines. This would inform if these 2 cell culture types share a common profile due to the same diagnosis; or because of the extensive culturing *in vitro* of established cell lines, differences are seen, as has been described in the literature (Section 1.1.5).

As shown in Figure 3.11, these 2 sample groups have a different expression profile, resulting in a independent clustering between both groups. In total, 14,060 differentially expressed genes were identified (adjusted p-value ≤ 0.001): 0 being increased in the cell lines and 0 decreased compared to the patient-derived cultures (Figure 3.11). Therefore, even originated from the same cancer type, the sequencing data indicates that there are important differences between both datasets. When investigating the expression profile of the NBL established cell line in comparison to the ES cell cultures, this clustered with the ES established cell lines (indicated by a black arrow in Figure 3.11). Therefore, the expression profile of ES established cell lines was more similar to the NBL cell line than to the ES

patient-derived cell cultures. This would suggest that the continued culturing of the established cell lines has resulted in a drift from the profile observed in ES patients.

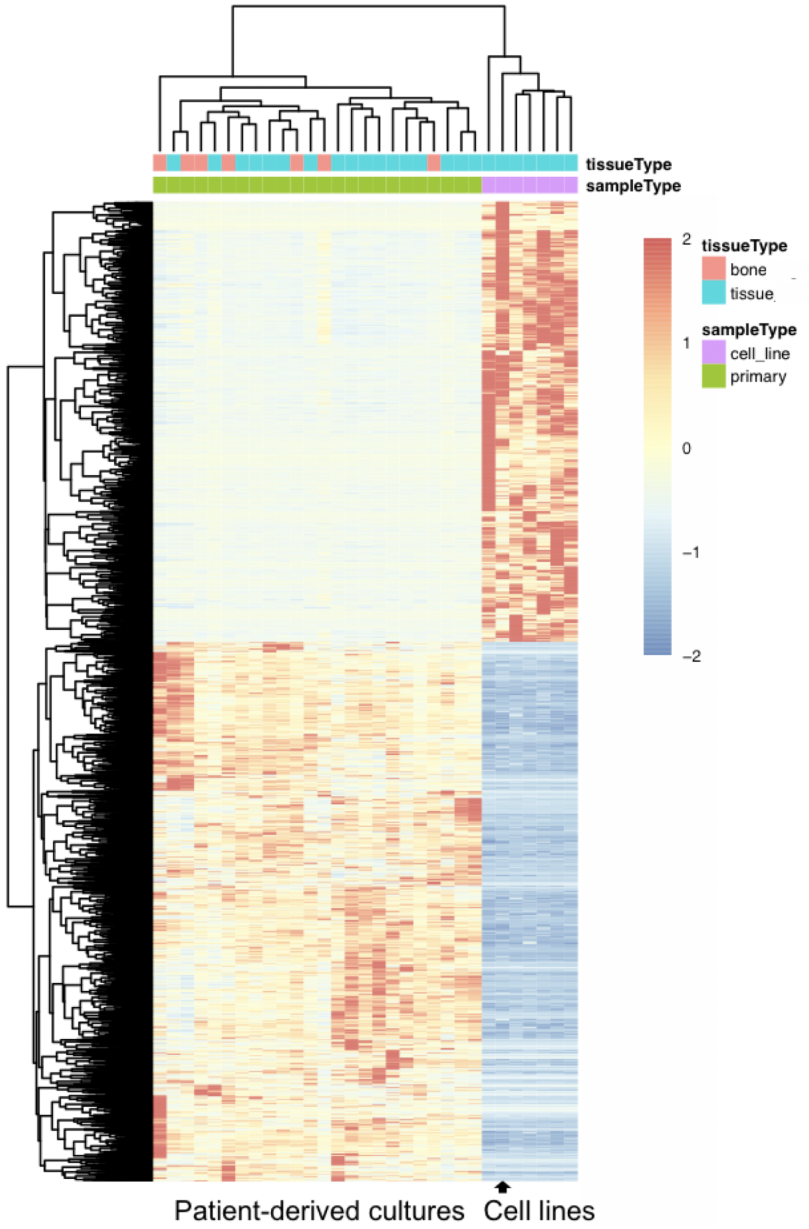


Figure 3.11: Differences in the gene expression profile between established cell lines and patient-derived cultures evaluated by DESeq2. Heatmap shows all differentially expressed genes ($n = 14060$) with enriched expression (red) or decreased expression (blue). Black arrow in cell lines indicates where NBL cell line clusters. Genes used from DESeq2 output using Euclidean distance matrix and scaled by row.

To further investigate what was driving the differences between established cell lines and patient-derived cell cultures, the differentially expressed genes were analysed using a pathway analysis tool in order to identify which molecular functions were enriched in each cell group. Amongst the genes enriched in established cell lines (3764 genes mapped by DAVID), they were classified in 182 different functional groups (Table A4). As can be seen in Table 3.5 (top), where the top 20 most significantly enriched pathways are shown (from the total 182), all functions are related to cell cycle processes, cell division and DNA replication. This would suggest that established cell lines, in comparison to patient-derived cell cultures, have higher expression of genes involved in such processes. In contrast, when evaluating the enriched functions amongst the genes enriched in patient-derived cell cultures (4041 genes mapped by DAVID), a total of 381 different functional families were identified (Table A5). Amongst the top 20 when ranking by the most significant pathways, functions associated to cellular adhesion and signalling cascades can be seen (Table 3.5, bottom). Therefore, there is a clear differences in the genetic functions enriched in each cell type, confirming that the differences observed in the differentially expressed genes by RNA-seq are linked to a different cellular behaviour.

In those cases that both a piece of bone and some soft tissue were collected as part of the ES tumour (Section 2.3.1), a comparison was done to investigate the differences between processing methods. As seen when evaluating the phenotypic data on Section 2.3, where no differences between both sample types were seen, RNA-seq data indicated that no genes came up as being differentially expressed. Therefore, this analysis further confirms that even processed through a different methodology (tissue maceration or bone digestion), both cell cultures have the same expression profile (seen here) and the same phenotype (seen in Section 2.3).

Term	Enriched in	Genes	P-Value	Fold enrichment
Cell cycle	Cell lines	223	1.10×10^{-28}	2,1
Nucleoplasm	Cell lines	661	3.00×10^{-26}	1,4
DNA replication	Cell lines	82	5.80×10^{-24}	3,1
Nucleus	Cell lines	1095	8.80×10^{-24}	1,3
DNA replication	Cell lines	55	1.80×10^{-19}	3,5
Mitosis	Cell lines	105	2.00×10^{-19}	2,4
Cell division	Cell lines	137	3.40×10^{-19}	2,1
Cell division	Cell lines	128	9.60×10^{-19}	2,1
Chromatin binding	Cell lines	132	2.40×10^{-16}	2
Phosphoprotein	Cell lines	1562	3.80×10^{-14}	1,1
Centromere	Cell lines	60	1.30×10^{-13}	2,7
Mitotic nuclear division	Cell lines	91	1.40×10^{-13}	2,2
DNA replication	Cell lines	27	1.70×10^{-13}	4,5
Chromosome	Cell lines	125	1.90×10^{-13}	1,9
DNA repair	Cell lines	86	4.30×10^{-13}	2,2
Sister chromatid cohesion	Cell lines	50	4.40×10^{-13}	2,9
G1/S transition of mitotic cell cycle	Cell lines	49	1.20×10^{-12}	2,8
DNA damage	Cell lines	111	3.40×10^{-12}	1,9
DNA repair	Cell lines	96	8.50×10^{-12}	2
Kinetochore	Cell lines	46	1.60×10^{-11}	2,8
Glycoprotein	Patient-derived	1027	5.60×10^{-35}	1,4
Focal adhesion	Patient-derived	167	6.50×10^{-33}	2,5
Glycosylation site:N-linked	Patient-derived	971	7.80×10^{-33}	1,4
Signal peptide	Patient-derived	797	6.80×10^{-32}	1,4
Signal	Patient-derived	891	1.10×10^{-30}	1,4
Extracellular exosome	Patient-derived	683	4.40×10^{-28}	1,4
Actin-binding	Patient-derived	114	2.50×10^{-22}	2,5
Membrane	Patient-derived	1467	4.40×10^{-22}	1,2
Extracellular matrix organization	Patient-derived	91	2.00×10^{-21}	2,7
Focal adhesion	Patient-derived	94	1.80×10^{-20}	2,6
Cell adhesion	Patient-derived	160	4.40×10^{-20}	2
Disulfide bond	Patient-derived	730	1.20×10^{-19}	1,3
Extracellular matrix	Patient-derived	104	1.40×10^{-19}	2,4
Secreted	Patient-derived	454	1.20×10^{-18}	1,4
Calcium	Patient-derived	248	2.50×10^{-18}	1,7
Extracellular matrix	Patient-derived	113	4.80×10^{-18}	2,2
Cell adhesion	Patient-derived	156	5.00×10^{-18}	2
Extracellular space	Patient-derived	343	1.70×10^{-16}	1,5
Actin binding	Patient-derived	106	2.30×10^{-16}	2,2
Proteinaceous extracellular matrix	Patient-derived	100	1.20×10^{-14}	2,1

Table 3.5: Pathway analysis results of established cell lines and patient-derived cell cultures. Functional families are shown for the top 20 (based on lowest p-value) of established cell lines (top) and patient-derived cell cultures (bottom) when investigating by DAVID using p-value threshold of 0.001. Data as number of genes involved in that function (genes), p-value and fold-change enrichment compared to expected. Functional annotation compared to all genes identified in RNA-seq analysis (background correction).

3.3.4 Correlation of the RNA-seq profile with clinical data

Having matched phenotypic and clinical data for the sequenced samples (Section 2.3), the next step was to evaluate if a different profile could be linked to any of the previously described variables in Chapter 2. Therefore, the bioinformatics pipeline was used to group samples according to the different phenotypes or clinical data groups evaluated before (Section 2.3) and see if the gene profile correlated with that variable. This would inform if predictive prognostic biomarkers could be identified.

A differential profile was observed between male and female-derived samples, with a total of 94 genes differentially expressed. When the top 20 by adjusted p-value and LFC were evaluated, the results correlated with the patient's gender. Specifically, the 17/20 genes had increased expression in the male samples, and all were located in chromosome Y. In contrast, the differentially expressed gene in females was TSIX (antisense XIST transcript), located in chromosome X. The other 2 genes were unannotated, and therefore, no data could be obtained from them.

When the ES patient-derived cell cultures were investigated according to the patients OS, a total of 41 genes were differentially expressed between alive and deceased patients. From the candidate genes (lower adjusted p-value and higher LFC), a total of 11 genes were increased in deceased patients, compared to 9 for the alive. Interestingly, amongst the genes differentially increased in patients with good OS, 2 genes involved in the major histocompatibility complex were found (HLA-DMB and HLA-DRA). The major histocompatibility complex comprises a series of cell-surface proteins involved in the regulation of the immune system (Trowsdale and Knight, 2013). Therefore, it is interesting to see that a higher expression of these genes is linked to patients that are doing better in the clinical setting. Therefore, these genes could be further investigated in a larger set of samples, to see if they could predict outcome, thus becoming putative prognostic biomarkers for ES.

3.3.5 A differential expression profile is identified between ES-CSCs and parental cell cultures

As mentioned previously, the propagated ES-CSCs provide a unique resource to study the phenotype of these cellular population, and compare their profile to the cells they are derived from. Phenotypic data *in vitro* showed that ES-CSCs have a different profile than the parental cultures when evaluating the self-renewing ability (Section 3.3.1.4), the migration index (Section 3.3.1.5) and the sensitivity to cytotoxic drugs (Section 3.3.1.6). The next step was to compare the expression profile of these two sample groups to further evaluate the differences across them. More importantly, the evaluation of ES-CSCs through RNA-seq provides a unique opportunity to identify candidate genes enriched in the ES-CSCs, which could be behind their different phenotype and could be ultimately exploited for targeted therapies.

Firstly, all ES-CSC (n= 13) were compared to the parental cell cultures they were originated from (n= 5). This identified a total of 347 genes that were differentially expressed in these 2 datasets, with 175 being increased in ES-CSCs and 172 decreased in ES-CSCs compared to the parental cultures. These results indicate that a different expression profile exists between both cell cultures, as can be seen in Figure 3.12 where they cluster separately. There is the exception of Geno39T.s1 (indicated by arrow in Figure 3.12) which clusters together with patient-derived cell cultures (parentals). Moreover, we can see that amongst the ES-CSCs, they cluster with the other ES-CSCs derived from the same parental, as indicated with colour legends on the bottom of Figure 3.12. When investigating the genes present in the top 50 of both LFC and adjusted p-value, only 9 candidate genes remained (Table 3.6), 6 with an increased expression in ES-CSC and 3 with a decreased expression, when compared to the parental cultures.

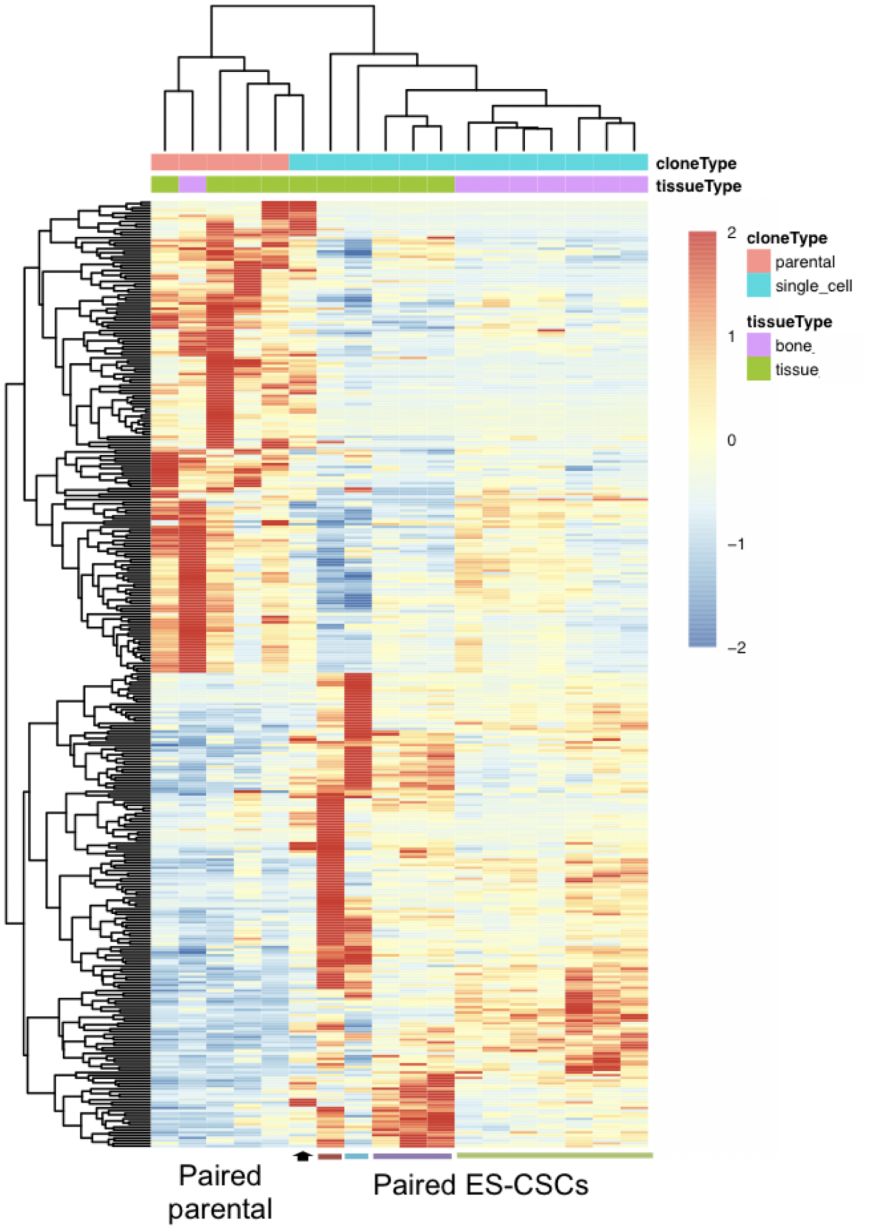


Figure 3.12: Differential expression profile between ES-CSCs and patient-derived cell cultures (parental cultures) they were derived from evaluated by DESeq2. Heatmap shows all differentially expressed genes ($n = 347$) with enriched expression (red) or decreased expression (blue). Colour lines at bottom of heatmap for ES-CSC indicate ES-CSCs originated from the same patient-derived cell culture. Black arrow indicates ES-CSC that clusters closely to patient-derived cultures. Genes used from DESeq2 output using Euclidean distance matrix and scaled by row.

Gene	Increased in	p-adj	p-adj rank	LFC	LFC rank
ELFN2	ES-CSC	1.6×10^{-4}	—	3.2	1
SLC38A11	ES-CSC	1.4×10^{-12}	5	1.8	9
NRXN1	ES-CSC	4.1×10^{-10}	11	1.6	12
TLR4	ES-CSC	1.1×10^{-13}	3	1.2	30
PIEZO2	ES-CSC	2.4×10^{-8}	29	1.3	29
CCDC190	ES-CSC	6.2×10^{-9}	22	1.1	46
SLC1A3	parent	8.0×10^{-10}	13	-1.3	27
LRRC32	parent	2.2×10^{-9}	18	-1.3	28
PIK3IP1	parent	6.6×10^{-10}	12	-1.2	37

Table 3.6: Candidate genes differentially expressed between ES-CSCs and paired parental cultures selected from the top 50 on adjusted p-value (p-adj) and $\log_2$ fold change (LFC). ELFN2 was not present on the top 50 for adjusted p-value but due to its high LFC, it is included.

It is important to identify an expression profile informative of the ES-CSCs. Therefore, instead of comparing ES-CSCs to only the patient-derived cell cultures they were derived from (n= 5), the next analysis included all ES patient-derived cell cultures (n= 24). Therefore, differences in the expression profile identified using this strategy would be more robust, as the differences seen would be due to the different nature of ES-CSCs, and not due to heterogeneity across cell cultures. Through this, a total of 973 differentially expressed genes were identified, with 454 increased and 519 decreased in ES-CSCs compared to the patient-derived cell cultures (Figure 3.13). After ranking by LFC and adjusted p-value and only keeping those genes that were present in both ranks, a total of 18 candidate genes were left, with only 2 increased in ES-CSCs and 16 decreased compared to the patient-derived cultures (Table 3.7). Interestingly, NRXN1 (increased in ES-CSCs; Table 3.7) had been identified in the previous analysis performed with only parental cultures; and ELFN2 (increased in ES-CSCs; Table 3.7) showed the highest LFC in the previous analysis, although its adjusted p-value was not on the top 50. Therefore, these two genes, present when assessing all ES patient-derived cell cultures, suggest that they could be important for the ES-CSC profile. These results open a door towards further validation and the possibility to be exploited for targeted therapies, as their high expression is linked to ES-CSC.

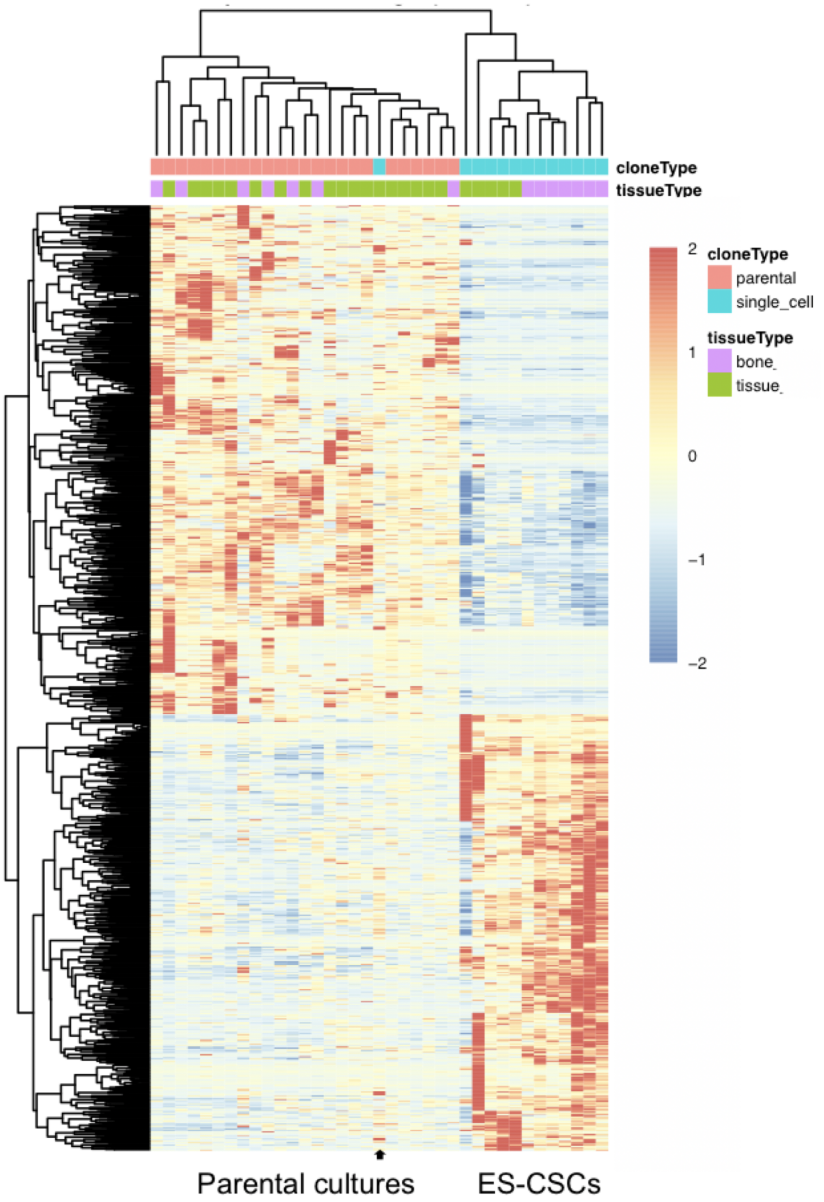


Figure 3.13: Differential expression profile between ES-CSCs and all patient-derived cell evaluated by DESeq2. Heatmap shows all differentially expressed genes ($n = 973$) with enriched expression (red) or decreased expression (blue). Black arrow indicates ES-CSC that clusters more closely to patient-derived cell cultures. Genes used from DESeq2 output using Euclidean distance matrix and scaled by row.

To further investigate the role of the differentially expressed genes in patient-derived cell cultures and ES-CSCs, a pathway analysis was performed, similarly to what was done for established cell lines and patient-derived cell cultures (Section 3.3.3). From the genes enriched in ES-CSCs (185 genes mapped by DAVID), they could be classified in 5 different functional groups, as shown in Table 3.8 (top), showing pathways involved in signalling and glycosylation. When investigating the genes enriched in patient-derived cell cultures compared to ES-CSC (285 genes mapped by DAVID), these could be mapped in 26 different functional families (Table 3.8 bottom). As can be seen, similar pathways are enriched in this dataset (glycosylation, signalling) as well as some involved in extracellular components. Therefore, we could conclude that in terms of functional groups, not many differences are seen between ES-CSCs and patient-derived cell cultures.

Gene	Increased in	p-adj	p-adj rank	LFC	LFC rank	Score
NRXN1	ES-CSC	1.5×10^{-15}	26	3.2	1	27
ELFN2	ES-CSC	5.7×10^{-14}	45	3.2	3	48
CECR7	parent	5.9×10^{-24}	7	-4.4	1	8
TDRD9	parent	5.2×10^{-24}	6	-4.0	2	8
HAS1	parent	1.2×10^{-25}	5	-3.6	7	12
GAS7	parent	2.1×10^{-23}	8	-3.7	4	12
DOK6	parent	7.1×10^{-23}	9	-3.6	6	15
CD40	parent	7.4×10^{-22}	11	-3.4	9	20
EYA1	parent	3.5×10^{-17}	18	-3.3	10	28
A2M	parent	1.9×10^{-19}	14	-3.1	16	30
FRAS1	parent	2.4×10^{-16}	23	-2.9	28	51
COL15A1	parent	2.0×10^{-15}	27	-2.9	29	56
LRRN1	parent	1.0×10^{-16}	20	-2.8	43	63
MXRA5	parent	3.0×10^{-16}	24	-2.8	44	68
CYP7B1	parent	1.1×10^{-13}	48	-3.0	24	72
MARVELD2	parent	2.8×10^{-14}	42	-2.9	32	74
KCND3	parent	5.0×10^{-15}	30	-2.8	46	76
IGFBPL1	parent	1.1×10^{-13}	47	-2.9	33	80

Table 3.7: Candidate genes differentially expressed between ES-CSCs and all patient-derived cell cultures. Table shows candidate genes identified from the top 50 of both adjusted p-value (p-adj) and log2 fold change (LFC), based on their ranking. Score = rank p-adj + rank LFC.

Term	Enriched in	Genes	P-Value	Fold enrichment
Glycoprotein	ES-CSCs	57	1.90×10^{-5}	1,7
Signal	ES-CSCs	49	1.20×10^{-4}	1,7
Signal peptide	ES-CSCs	45	1.00×10^{-4}	1,8
Glycosylation site:N-linked	ES-CSCs	53	1.20×10^{-4}	1,6
Calcium	ES-CSCs	18	3.10×10^{-4}	2,7
Glycoprotein	Parental	113	1.60×10^{-12}	1,9
Glycosylation site:N-linked	Parental	108	2.40×10^{-12}	1,9
Developmental protein	Parental	37	2.50×10^{-8}	2,9
Signal peptide	Parental	81	6.90×10^{-8}	1,8
Signal	Parental	89	1.40×10^{-7}	1,7
Proteinaceous extracellular matrix	Parental	17	1.50×10^{-6}	4,5
Secreted	Parental	49	1.20×10^{-5}	1,9
Protein heterodimerization activity	Parental	21	1.40×10^{-5}	3,1
Extracellular matrix	Parental	16	1.90×10^{-5}	3,9
Dendrite	Parental	17	2.20×10^{-5}	3,6
Extracellular region	Parental	42	3.80×10^{-5}	2,0
Disulfide bond	Parental	71	4.90×10^{-5}	1,6
Cytoplasmic domain	Parental	73	6.20×10^{-5}	1,6
Angiogenesis	Parental	13	9.60×10^{-5}	4,0
Heparin binding	Parental	11	1.30×10^{-4}	4,7
Extracellular matrix organization	Parental	12	1.10×10^{-4}	4,3
Extracellular space	Parental	36	2.10×10^{-4}	1,9
Ectodermal dysplasia	Parental	6	2.70×10^{-4}	10,3
Growth factor	Parental	9	4.60×10^{-4}	5,0
Inflammatory response	Parental	16	4.10×10^{-4}	2,9
Head region	Parental	7	4.50×10^{-4}	7,1
Tail region	Parental	7	5.20×10^{-4}	6,9
Disulfide bond	Parental	60	5.30×10^{-4}	1,5
Differentiation	Parental	21	7.60×10^{-4}	2,3
Extracellular domain	Parental	58	7.00×10^{-4}	1,5
Polymorphism	Parental	188	9.70×10^{-4}	1,2

Table 3.8: Pathway analysis results of ES-CSCs and patient-derived cell cultures. Functional families are shown for genes enriched in ES-CSC ($n = 5$) and enriched in patient-derived cell cultures ($n = 26$) when investigating by DAVID using p-value threshold of 0.001. Data as number of genes involved in that function (genes), p-value and fold-change enrichment compared to expected. Functional annotation compared to all genes identified in RNA-seq analysis (background correction).

3.4. Discussion

In this study I have profiled ES-CSCs derived from a self-renewing single cell, which have a different expression profile and are more resistant to conventional chemotherapies than the cells they are derived from. To my knowledge, this is the first time CSCs isolated from ES cells have been evaluated by RNA-seq and compared to the parental cell population. Previous studies have assessed gene expression on ES-CSC by PCR (Awad et al., 2010; Cornaz-Buros et al., 2014; De Vito et al., 2012; Riggi et al., 2010) or array (De Vito et al., 2012; Riggi et al., 2010). ES-CSCs are suggested to be the reason behind poor outcome and refractory disease in ES patients, as they survive treatment, leading to tumour growth and relapse (Trucco and Loeb, 2012). Therefore, the correct isolation and profiling of these ES-CSCs is fundamental to better understand their biology and identify novel targeted strategies against these cells (Section 1.2.3). Using a functional self-renewing assay, a population of ES cells with self-renewing abilities could be identified and propagated in culture. These ES-CSC cultures are a unique material to determine the specific profile of these cells, with the aim of identifying novel candidate genes for targeted therapies and biomarkers that would help with patient stratification and personalised medicine.

Total RNA-seq of ES-CSCs and the patient-derived ES cell cultures they were derived from identified a different expression profile between both sample groups. This difference was present both when only assessing the paired parental cultures (347 differentially expressed genes) or when using all the patient-derived samples (973 differentially expressed genes). These analyses involved ES-CSCs derived from different samples, suggesting that the differences identified are not due to the underlying biology of the patient, but due to a different profile of the ES-CSCs. In order to elucidate if the patient-derived cell cultures and ES-CSCs had an underlying biology that differed, the differentially expressed genes were evaluated using pathway analyses. Interestingly, the functional groups that were enriched in ES-CSCs and patient-derived cell cultures separately involved the same biological functions (i.e. glycosilation or signalling, amongst others). Therefore, the differences in phenotype are not due to massive changes in the molecular

machinery and functions, but due to differences in the expression of particular genes.

This study was focused on identifying ES-CSCs markers, which would have a higher expression in ES-CSC samples than the patient-derived cell cultures. Amongst the different candidate genes enriched in ES-CSCs, neurexin 1 (*NRXN1*) is the only candidate gene that comes up in both comparisons (paired samples: adjusted p-value 4×10^{-10}, LFC 1.6; all samples: adjusted p-value 2×10^{-15}, LFC 3.2), suggesting that its higher expression is linked to an ES-CSC profile. It is a neural differentiation related gene which has previously been found to have high expression on ES cells and xenograft ES mouse tumours (Tanaka et al., 2014). Another candidate gene of interest enriched in the ES-CSCs is *ELFN2* (extracellular leucine rich repeat and fibronectin type III domain containing 2), that alongside *NRXN1* comes up as being enriched in ES-CSCs compared to the overall ES cultures (paired samples: adjusted p-vale 2×10^{-4}, LFC 3.2; all samples: adjusted p-value 6×10^{-14}, LFC 3.2). However, ELFN2 expression indicates that it is a weaker marker, as when comparing to the 5 cell cultures the ES-CSC were derived from, it did not come up on the adjusted p-value rank (although it had the highest LFC). ELFN2 has previously been linked to a more aggressive gastric cancer profile (Zhang, 2017). Interestingly, both these genes are members of the synaptic adhesion molecules (SAMs), involved in the synapse mechanisms in mammalian brain (Rudenko, 2017). As well as ELFN2 and NRXN1, this family of proteins includes LINGO1, a protein that has been shown to be enriched in ES cells (Town et al., 2016). Therefore, the relationship between the 2 ES-CSC markers identified in this study and LINGO1 should be further investigated, as a link between these molecules could be important for ES biology. In order to elucidate the role these markers could have in the ES-CSC profile, validation studies should investigate how silencing or overexpressing these genes changes the self-renewing ability or phenotype of ES-CSCs. This would help understand if these markers are important for maintaining the CSC profile in these cells, and if they could be linked to prognosis. A limitation of these markers is that they are involved in synapse processes in the brain, and therefore, I expect that higher levels of expression are found in this tissue. If targeted therapies were to be designed against NRXN1 and

ELFN2 to eliminate the ES-CSCs, it is fundamental to evaluate the expression in other tissues and the possible toxicities that such treatments could generate.

Previous studies on the ES field have used expression of markers (Cornaz-Buros et al., 2014; De Vito et al., 2012; Riggi et al., 2010; Suva et al., 2009; Wahl et al., 2010), higher drug efflux (Komuro et al., 2007; Yang et al., 2010) or increased activity of ALDH (Awad et al., 2010) to identify and investigate the CSC population within ES samples. Within these, expression of cell surface markers is a surrogate method for the identification of CSC, as their expression is not an intrinsic characteristic of CSC populations. Therefore, it is possible that these assays are not identifying the overall CSC populations within ES tumours, but rather a subset of ES-CSCs. In my dataset, CD133 (common CSC marker used in the literature) was not differentially expressed in ES-CSCs compared to patient-derived cell cultures. In fact, when evaluating the expression levels between these 2 sample groups, higher expression was seen in patient-derived cell cultures (3 ± 5, $n = 26$) than in ES-CSCs (0.2 ± 0.6, $n = 13$; fold-change 14). This would explain why, when researchers have tried to compare the ES-CSC populations identified through these different methods, no correlation was found (Jiang et al., 2010; Leuchte et al., 2014; Wahl et al., 2010). Therefore, using a functional assay that exploits an intrinsic characteristic of CSCs, such as the ability to self-renew, we should be able to isolate the different subsets of ES-CSCs independent of their expression profile. This is consistent with the phenotypic characterisation of ES-CSCs, which showed heterogeneity across the different cell cultures obtained (i.e. differences in the migration index or CD99 expression). Therefore, I would say that the results obtained when investigating the RNA profile of the ES-CSCs in this study provide a more robust picture of the ES-CSC profile than only investigating a subset of ES-CSCs. However, a limitation of isolating ES-CSCs based on their ability to self-renew and propagate in culture is that an important fraction of these cells cannot be successfully cultured up to T75 (from 577 isolated, only 13 reached T75 confluency). Therefore, it could be that the successfully cultured ES-CSCs investigated in this study have a different profile than the ones that could not grow in *in vitro* conditions. Future studies could aim to adapt the self-renewing assay *in vitro* to other models, such as the use of scaffolds which recreate better the *in vivo*

profile. This might result with higher success rates in growing these ES-CSCs, thus being able to study the overall ES-CSC populations.

Propagated ES-CSCs were profiled *in vitro* using the same phenotypic assays used for the parental cultures in Chapter 2 (Section 2.3). A more resistant profile to paired parental cultures was seen for both doxorubicin and vincristine. This agrees with data previously obtained in the CCRG where more ABC transporters could be detected in ES-CSCs than in the parental cultures they were derived from (30 expressed in ES-CSC vs. 8 in parental cells; Dr. Elizabeth Roundhill, personal communication). An increased expression of ABC transporters is linked to CSCs (Schatton et al., 2009), as cells would be able to efflux drug more efficiently, thus resulting in less sensitivity to chemotherapies. This increased expression of ABC transporters is in agreement with data from other ES studies in CSC populations (Awad et al., 2010; Komuro et al., 2007; Yang et al., 2010), where an increase on dye efflux was seen, which could be blocked using verapamil, an ABC transporter inhibitor. However, only one ABC transporter (ABCB1) had increased expression in ES-CSCs compared to patient derived cell cultures (LFC 2.6, adj p-value 8×10^{-9}). The other ABC transporter evaluated, although expressed in ES-CSCs, they had similar levels of expression than patient-derived cell cultures.

When the ability to self-renew from a single cell was evaluated, this showed that these ES-CSCs population had lost this ability compared to the parental cultures. This observation could be explained by the fact that ES-CSCs are slower growing than their parental populations, and therefore, might need longer periods of time to cross the threshold of 5 cells (used for the self-renewing assay evaluation). Another possible explanation is that the ability to self-renew from a single cell *in vitro* is lost after the first selection. The microenvironment has been shown to be important in tumour modulation and progression (Hu et al., 2015; Xu et al., 2018). Therefore, this loss in the self-renewing ability from a single cell might be due to the absence of factors from other ES cells (bulk of the tumour) or from other cells in the tumour niche. Serial transplantation in mice is the most common method to evaluate the ability to self-renew in ES-CSC, where most ES-CSC publications report the ability of ES-CSC to form tumours *in vivo* (Awad et al., 2010; Cornaz-Buros et al., 2014; De Vito et al., 2012; Riggi et al., 2010; Suva et al., 2009; Wahl et al., 2010). Therefore,

future studies should try to evaluate the self-renewing ability of ES-CSC in better models, such as scaffolds recreating the microenvironment and factors present in the body, or placing them in mouse xenografts and PDX to see if they lead to higher levels of tumorigenesis than when placing ES cells (overall populations).

Moreover, ES-CSC showed heterogeneity on their migration index compared to the parental cultures. In most cases, the migration index decreased compared to the parental culture; with the exception of one culture where the ES-CSCs had a higher migration index. Similarly, changes in the expression of CD99 and c-kit compared to the parental culture were seen. Not only the level of expression differed in some cases from the parental culture, but differences across ES-CSCs derived from the same sample were seen. These results suggest that ES-CSCs have a different phenotype to the ES cell cultures they are derived from, and at the same time, not all ES-CSCs derived from the same sample share the same phenotype. This is an important finding, as it suggests that there are different subpopulations of ES-CSCs behaving differently. Therefore, profiling of all these cells to investigate the shared profile and the differences that are driving these changes is important to further understand the ES-CSC biology. Due to the small sample size (13 ES-CSCs on RNA-seq), these differences in the various phenotypes could not be investigated. Future studies should analyse a bigger panel of ES-CSCs, from different samples and using multiple colonies from the same sample, in order to increase the knowledge on these differences and similarities across ES-CSCs.

The transcriptomic analysis performed in this study has involved both patient-derived ES cell cultures (Section 2.3) and established ES cell lines. This permitted the comparison of both datasets, in order to evaluate if the ES gene expression profile was retained in the established cell lines, resembling thus the recently obtained patient-derived cultures; or extensive culturing had led to a drift in the expression profile. A total of 14,060 differentially expressed genes were identified between both sample groups, resulting in samples from both groups clustering separately. Moreover, ES cell lines clustered more closely to the NBL cell line (BE(2)C) than to the patient-derived ES cultures, indicating that the profile driving the clustering was not ES-related, as that would have resulted in ES established cell lines clustering more closely to ES patient-derived cell cultures,

and separate from the NBL cell line. This is in agreement with data from another transcriptomic analysis in ES (Crompton et al., 2014), where a different gene expression profile was seen between established cell lines and patient tumours. When they performed principal components analysis, the largest amount of variance separated both datasets into different clusters, indicating that these two sample groups did not share the same profile. When performing functional analyses on the differentially expressed genes on established cell lines and patient-derived cell cultures separately, differences in the enriched functional pathways were seen. Agreeing with previous studies (Brohl et al., 2014; Crompton et al., 2014; Stacey, 2005), established cell lines have have increased expression of genes involved in cell cycle regulation and DNA replication, processes important for cellular division. In contrast, patient-derived cell cultures, as seen in the ES-CSCs evaluation, are enriched for genes involved in glycosilation and cellular adhesion, amongst others. Therefore, as well as seeing that these two sample groups have many differentially expressed genes, it can be seen that the molecular mechanisms that control their cellular biology differ. This is an important observation, as many studies use established cell lines to identify novel therapeutic targets or genes of interest to better understand the biology of the cancer. Therefore, if these cell lines have acquired a different profile due to the prolonged culturing *in vitro*, the results reported in terms of drug efficacy or pathways altered could be biased, not accurately representing the pathological conditions. Altogether, these observations further exemplify the need to use patient-derived cell cultures or tumours when studying the biology of samples, as established cell lines have acquired a different profile that could have important implications in the analyses.

One of the advantages of using RNA-seq is the ability to detect unknown variants in our dataset. This is particularly important in ES, where the characteristic EWSR1 fusion has been reported to partner with less common ETS genes (Section 1.1.1). Using the STAR-Fusion package (Haas et al., 2017), novel fusion variants have been identified in the patient-derived ES cultures, which would have otherwise been discarded due to mapping in different areas of the genome. These include *EWSR1-KATNB1*, *EWSR1-CHD4* and *EWSR1-PTK7* fusion variants, as well as the

reciprocal translocations CAMSAP2-EWSR1, *ETV5-EWSR1*, *SLC6A6-EWSR1*, *TEAD1-EWSR1* and *SF3B3-EWSR1*. Some of these novel variants where detected in samples that failed to show positive results by RT^A-PCR when evaluating the known EWSR1-FLI1 or -ERG fusions, but did show a break apart on *EWSR1* by FISH. Therefore, we could be observing novel variants of the EWSR1 fusion, further expanding the known list of EWSR1 fusion described in the literature (Romeo and Tos, 2010). Of special interest, both *TEAD1* and *ETV5* are transcription factors, which could play a similar role to FLI1. Moreover, *ETV5* is a member of the ETS family of transcription factors, from which are members other *EWSR1* fusion partners in ES, such as *ETV4* (or *E1AF*), *ETV1* or *FEV* (Peter et al., 1997). However, it is important to validate these results in order to investigate if these fusions are the drivers of the malignant profile in the samples they have been identified; or they are just by-stander effects thus not having a major role in ES biology.

In addition to these novel findings, known EWSR1 fusions were detected in all established ES cell lines, agreeing with PCR data from the CCRG (some described in Section 2.3.2.2). However, RNA-seq detected an EWSR1 fusion in only 1 patient-derived cell culture, failing to identify fusions in samples where an EWSR1 fusion had already been reported by RT^A-PCR (Section 2.3.2.2). This lack of detection could be explained by the differences in sensitivity between RT^A-PCR and RNA-seq. In the first technique, the primers used are directed against the fusion only, therefore amplifying for only the target of interest. However, RNA-seq, although it can detect novel variants, it amplifies the full transcriptome, and to identify them, the junction between both genes has to fall within the read. Therefore, all the reads that fall in either of the genes involved, not overlapping the junction, will be seen as endogenous expression. For instance, an established cell line with high levels of expression seen by RT^A-PCR only had 31 reads falling into the fusion junction. Therefore, if the aim of the study was to identify novel fusions, the coverage of the sequenced samples would need to be increased from 4 samples per lane, in order to increase the chance of having reads fall into the junction, or use targeted RNA-seq to evaluate already-known fusions.

In order to evaluate the expression of CD99 on this ES sample cohort, a specific strategy had to be developed. Due to the CD99 gene being present in the PAR of

chromosome X and Y (Mangs and Morris, 2007), STAR aligner used in this study (Dobin et al., 2013) would only retain reads that aligned unambiguously to one regions of the genome. This would result in all CD99 reads being discarded, not allowing for the evaluation of CD99 expression in ES samples. Therefore, building the gene-specific reference for only the CD99 sequence permitted the evaluation of CD99 expression in ES samples. The strategy proposed here could be used for other genes or sequences that are present in multiple areas of the genome, such as other genes in PAR areas.

Another important output from this study was the generation of a bioinformatics program that permitted the pre-processing, alignment, normalisation and comparison of the different samples used. In order to make this process easier for the user, a metadata database (in this case SQL) was employed. This allowed for the collection of all data relevant from the RNA-seq runs and the metadata of the samples analysed (i.e. cancer or cell culture type). Therefore, when performing the different comparisons between the diverse sample groups (established cell lines vs. patient-derived cell cultures), an easy selection of the samples could be done through different filters based on this metadata. With minimal interaction from the user, the pipeline was more protected from human error, which could result in important implications in downstream analyses. Moreover, due to the constant increase in the metadata behind each sample (i.e.' migration index, OSR status or c-kit expression), another program was developed to account for this extra information. Not having to manually edit the SQL database every time, a script was generated to easily select and classify the samples needed for the evaluation based on their ID. However, if this tool is to be used constantly, it should be automated in order to reduce the human error when inputting the sample IDs, similarly to what was done for the main pipeline. Nevertheless, these scripts have become very useful in the group, allowing powerful comparisons across samples with minimal user interaction.

The downstream analyses of RNA-seq are complex, from the manipulation of the raw reads to the comparison of different datasets obtained. The wide use of the technique has allowed for more user-friendly and straight-forward analysis tools that help scientists with lack of bioinformatics expertise to be able to perform

these analyses. These include easy to follow guides and optimised default settings, being able to perform the desired analyses with minimal optimisation. Nevertheless, the data manipulation and interpretation needs to be accurately planned in order to get the most of the data generated (Costa-silva et al., 2017). This includes aligner characteristics, such as time or computational power needed; the correct programs for the desired analyses (i.e. ability to identify fusions); how to interrogate the metadata; or what statistical program should be used for the normalisation and evaluation of the data generated (Hrdlickova et al., 2017). In this study, STAR with its sub-package STAR-fusion has allowed for the identification of EWSR1 fusions in this ES dataset. A recent evaluation of 14 different aligners used in the literature was performed by Baruzzo et al. (2017), showing that STAR was amongst the most reliable methods. Moreover, in a recent study by Costa-silva et al. (2017), some of the most common normalisation programs for RNA-seq data were compared, using the same starting raw reads and evaluating the output data generated. They identified that upon correct selection of the alignment method, the normalisation and comparison step had the most impact on the results generated and subsequent downstream analyses. In my study, the DESeq2 package (Love et al., 2014) was used. This decision was based on this method being amongst the best normalisation and comparison programs for gene expression data (Costa-silva et al., 2017; Love et al., 2014; Zhang et al., 2014), showing a good balance between accuracy, precision and sensitivity, and performing well with low sequencing depth. It would be interesting to use the raw reads obtained in this study to analyse them through another of the programs described in the literature, in order to evaluate in this ES dataset if DESeq2 is indeed the best program, or other methods are starting to supersede it.

4. The exosome profile from ES samples

4.1. Introduction

Exosomes are a type of extracellular vesicle that mediate cellular communication by the active transport and exchange of their cargo inducing changes in cellular behaviour (Zhang et al., 2015). To date, 4 different studies have suggested ES-derived exosomes as candidate circulating biomarkers (Miller et al., 2013; Tsugita et al., 2013; Villasante et al., 2016) and their putative role in ES cancer progression (Ventura et al., 2016) (Section 1.3).

ES-derived exosomes have been isolated using the gold-standard differential ultracentrifugation (dUC) method (Miller et al., 2013; Ventura et al., 2016); using commercial methods that are based on the precipitation of exosomes (Ventura et al., 2016; Villasante et al., 2016) or using a filtration-based approach (Tsugita et al., 2013). ES-derived exosomes share common features with exosomes isolated from other cancers, including the size, presence of characteristic surface markers and enrichment of sRNAs (Miller et al., 2013; Tsugita et al., 2013; Ventura et al., 2016; Villasante et al., 2016). The EWSR1-ETS fusion has been identified as part of the ES-derived exosome cargo (Miller et al., 2013; Tsugita et al., 2013) as well as the expression of CD99 cell surface marker (Ventura et al., 2016). For this reason, ES-derived exosomes could be exploited as circulating biomarkers, as the 2 diagnostic ES markers are present in these. In addition, the RNA profiling of ES-derived exosomes to date has been performed on mRNAs of interest by PCR (Miller et al., 2013; Tsugita et al., 2013; Ventura et al., 2016; Villasante et al., 2016) or by micro arrays (Miller et al., 2013). To my knowledge, no one has investigated the overall RNA profile of ES exosome cargo using RNA-seq.

Exosomes have been implicated in different pathogenic processes through the transfer of their cargo to recipient cells (El Andaloussi et al., 2013) (Section 1.3.2). For instance, breast cancer-derived exosomes with a stemness and metastatic RNA profile could induce a transformation towards stemness expression profile in recipient cells, as well as an increase in invasion and proliferation (Rodríguez et al., 2015). Moreover, the injection of these exosomes in orthotopic breast cancer mice

induced an increase in tumour growth and metastasis. Similarly, exosomes from CAF induced an increase in clonogenicity and tumour growth in CSCs (Hu et al., 2015). In the ES field, only one study has evaluated the implications of ES-derived exosomes on recipient cells (Ventura et al., 2016). The transfer of exosomes from CD99-silenced cells to native ES cells induced the same pro-differentiated profile as directly silencing CD99. This is consistent with the transfer of miRNA-34a, highly expressed in CD99-silenced cells, as this can induce a reduction of Notch expression and decrease in the NF-κB activity, involved in the differentiation of ES. Therefore, ES-derived exosomes can induce changes to recipient cells based on the release of specific cargo. ES exosomes have also been shown to transfer the EWSR1-ETS fusion (Tsugita et al., 2013) and EZH2 mRNA (Villasante et al., 2016) to recipient cells, although the cellular effects of this transfer have not been evaluated.

The mechanism of exosome uptake by cells is poorly understood. However, the uptake of exosomes by cells can be monitored by labelling the isolated exosomes with generic dyes, so when they are taken up there is an increase in fluorescence in recipient cells which can be measured. Several dyes have been used for this purpose including long-chain dialkylcarbocyanines (Horibe et al., 2018; Tian et al., 2010) or CFSE (Escrevente et al., 2011; Zhang et al., 2011). The uptake of exosomes by recipient cells is quantified by flow cytometry, investigating the increase in fluorescence of the cells and by immunofluorescence (IF), to evaluate if labelled exosomes can be seen in the recipient cells (Franzen et al., 2014; Zhang et al., 2011). Cellular uptake of exosomes is an active regulated process demonstrated by a decrease in the uptake by recipient cells when using inhibitors against the pathways involved in exosome uptake (Atai et al., 2013; Franzen et al., 2014; Horibe et al., 2018; Tian et al., 2014) or incubation of recipient cells at non-physiological conditions (4 °C) (Horibe et al., 2018; Tian et al., 2014). Exosome uptake is a dose and time-dependent effect and appears to take place most quickly during the first hours of incubation (2–4 h) (Franzen et al., 2014; Horibe et al., 2018; Zhang et al., 2011), with a report even observing uptake after 30 min of incubation (Escrevente et al., 2011). To date, the uptake mechanisms of ES-derived exosomes on recipient cells have not been investigated.

Other groups, instead of studying directly the uptake mechanisms of exosomes by recipient cells, they use the changes in the recipient cells to confirm this exosomes internalization. For instance, an increased expression of oncoproteins was detected in colorectal cancer cells after being incubated with exosomes from a KRAS-mutated cell line (Demory Beckler et al., 2013). In ES, similar studies have been performed, in which the successful uptake of ES-derived exosomes by recipient cells was evaluated by the transfer of the EWSR1-ETS fusion to other cells (Tsugita et al., 2013), changes in the CD99 expression (Ventura et al., 2016) or an increased expression of EZH2 (Villasante et al., 2016), compared to cells not exposed to exosomes. Therefore, although the actual process of exosomes internalization is not investigated, the use of surrogate markers (i.e. change in the expression profile) can be used to evaluate the successful uptake of exosomes.

Due to this increased interest in the role these vesicles play in different biological processes and pathologies, there have been important developments in the isolation and characterisation methods of exosomes. In Table 4.1, the most common methods used for the isolation of exosomes are described. These commonly-used methodologies embrace diverse principles for the successful isolation of exosomes, each with different advantages and disadvantages that must be considered according to the characteristics of each experiment. As a result of this heterogeneity of techniques and the importance of the quality of the isolated exosome for downstream applications, several investigations have compared them (Lobb et al., 2015; Martins et al., 2018; Nordin et al., 2015; Paolini et al., 2016; Tauro et al., 2012). The overall conclusion from the different reports is that although all methods isolate exosomes with the expected size, expression of markers and enrichment for sRNAs, there are differences in the yield or purity of the eluate. As the quality of the isolated exosomes have great implications in the downstream analyses, the decision on what method to use has to be made on a case-by-case basis. For this reason, the International Society of Extracellular Vesicles (ISEV) has published a series of recommendation and guidelines for the correct isolation of exosomes (Lötvall et al., 2014; Witwer et al., 2013).

Method	Description	Advantages	Disadvantages	Commercial kits	References
Differential UC	Different centrifugation steps to remove contaminants and pellet exosomes.	Most commonly used; low cost.	Requires of special equipment; low purity; less efficient for viscous samples (i.e. plasma).	–	Théry et al. (2006)
Density gradient centrifugation	Gradient of sucrose or iodoxinol to separate based on density.	High purity; separation of aggregates and protein contaminants.	Lengthy; requires special equipment; sensitive to centrifugation time.	OptiPrep™(supplied by different companies)	Kalra et al. (2013); Van Deun et al. (2014)
Precipitation	Using synthetic polymers, alter the solubility of exosomes for easier separation.	Common technique; relatively fast and simple; low centrifugation speeds (less effect on exosomes).	Low purity; presence of polymer on precipitates (downstream implications).	ExoQuick (System Biosciences); Total Exosome Isolation (Thermo Fisher); PME (Analytik-Jena)	Martins et al. (2018); Schageman et al. (2013)
Immunoaffinity capture	Capture of exosomes based on the presence of specific markers on their surface, using magnetic beads.	High specificity; high purity.	Prior knowledge of exosome profile; expensive when using large volumes.	–	Pedersen et al. (2013); Rupp et al. (2011); Tauro et al. (2012)
Size exclusion chromatography	Separation of exosomes based on size, using column-based chromatography.	High purity; relatively easy; low centrifugation force; collection of different vesicle populations.	Loss of exosomes in column; lengthy; less efficient for viscous material (i.e. plasma).	qEV (Izon); Exo-spin™(Cell Guidance Systems)	Hong et al. (2016); Taylor et al. (2011)
Ultrafiltration	Exosomes separated using different size filters.	High purity, relatively fast and simple, direct RNA extraction.	Loss of exosomes in membrane; possible clogging of membrane.	VivaSpin filters (Sartorius)	Cheruvanky et al. (2008)

Table 4.1: Summary table of the most commonly used exosome isolation methods.

In these, the different factors that can impact the isolation process, recommendations for each biological material from which to isolate exosomes or the minimal experiments required to prove that the isolated vesicles are indeed exosomes are described. With this, ISEV has the aim to help researchers in this growing field, as future understanding of the exosome biological processes will depend greatly from the correct isolation of exosomes.

Besides the commonly-used methods for the isolation of exosomes (Table 4.1), new approaches are still being developed. Amongst these we can find a membrane-based affinity (MBA) method by Qiagen; the ExoEasy Maxi kit (Enderle et al., 2015). The principle is to use the bioproperties of the exosome membrane to capture these vesicles into a membrane, while all entities with different properties (i.e. protein aggregates) are discarded. Similar to the precipitation methods, it does not need special equipment; the RNA extraction can be performed directly on the membrane, as described for ultrafiltration; and can be used in a wide range of starting volumes and materials without an impact on result nor cost.

Based on the different methods described in Table 4.1 and this new approach, the MBA method from Qiagen appeared to provide the best advantages for this study since it allows the isolation of exosomes both from large cell culture media volumes or small plasma aliquots; it is cost-effective; it does not require special equipment; the RNA extraction can be done directly onto the membrane, using the miRNeasy extraction kit routinely used in the laboratory; and the eluted exosomes can be easily concentrated by ultracentrifugation. Different groups have used this kit to study exosomes from cell culture (Kim et al., 2017; Liang et al., 2018; Stolzenburg and Harris, 2018), serum and plasma (Enderle et al., 2015; Liu et al., 2018; Stranska et al., 2018; Xu et al., 2018) and urine (Delić et al., 2016), showing the expected characteristics for exosomes.

4.1.1 Aims of this chapter

The aims of this chapter were to:

1. Optimise the protocol to isolate exosomes from ES cell culture media.
2. Characterise exosomes from ES patient-derived cell cultures and plasma.
3. Investigate the RNA cargo of ES exosomes.

4. Study the transfer of exosomes from ES cells to non-ES cells.
5. Investigate the impact on cellular behaviour of uptake of ES-derived exosomes on non-ES cells.

4.2. Materials and Methods

4.2.1 Collection and processing of plasma samples

Plasma samples were obtained from patients enrolled in the EE2012 clinical trial (Research Ethics Committee 12/NW/0827) at diagnosis or from healthy volunteers. Whole blood (5 ml) was collected into an EDTA PAXgene Blood RNA Tubes (PreAnalytiX, BD Biosciences) and within 2 h, centrifuged at 1,200 g for 10 min at RT in order to separate the plasma and cellular fraction. The plasma was carefully transferred into cryovials (Nunc, Thermo Fisher Scientific) in 0.5 ml aliquots and stored at −80 °C. The cellular fraction was discarded.

4.2.2 Media collection for exosome isolation

Cells were harvested (Section 2.2.3) and seeded into T75 flasks with 15 ml of appropriate media with FCS to reach an 80–90 % confluency at 72 h (time of media collection; cell number adapted for each cell type). After seeding, cells were incubated for 24 h in 5 % CO_2 at 37 °C to allow them to adhere to the plastic of the flask. Then, media was aspirated and cells were washed 3 times with 5 ml of pre-warmed PBS. Serum-free media (SFM; 15 ml) was added to each flask and incubated for 48 h, at which point flasks were 80–90 % confluent. The change to SFM was important to remove any exosomal contamination from the serum. Media was collected by pipetting and the contents of two flasks combined. Cellular debris and larger vesicles were removed from collected media by centrifugation at 2,000 g for 30 min (Eppendorf centrifuge 5804) and filtration through a 0.8 μm filter (Sartorius Minisart; Sartorius, Surrey, UK).

4.2.3 Optimisation of media collection protocol

To identify the optimal method for the media collection, different modifications of the Section 4.2.2 protocol were evaluated as summarised in Figure 4.1. In order to get the optimal exosome yield without effecting cell viability, which would impact

on the quality of the isolated exosomes, the effect of incubation in SFM over time (24–120 h) was assessed (Figure 4.1A). The pH of the cell culture media was measured using a pH meter (Hanna Instruments pH 210 Microprocessor pH Meter; Hanna Instruments Ltd, Bedfordshire, UK) as well as cell viability by trypan blue exclusion assay (Section 2.2.5). Another evaluation was done on the method used to remove cellular debris and larger vesicles. As summarised in Figure 4.1B, different centrifugation and filtration options were compared to identify which led to a higher exosome yield.

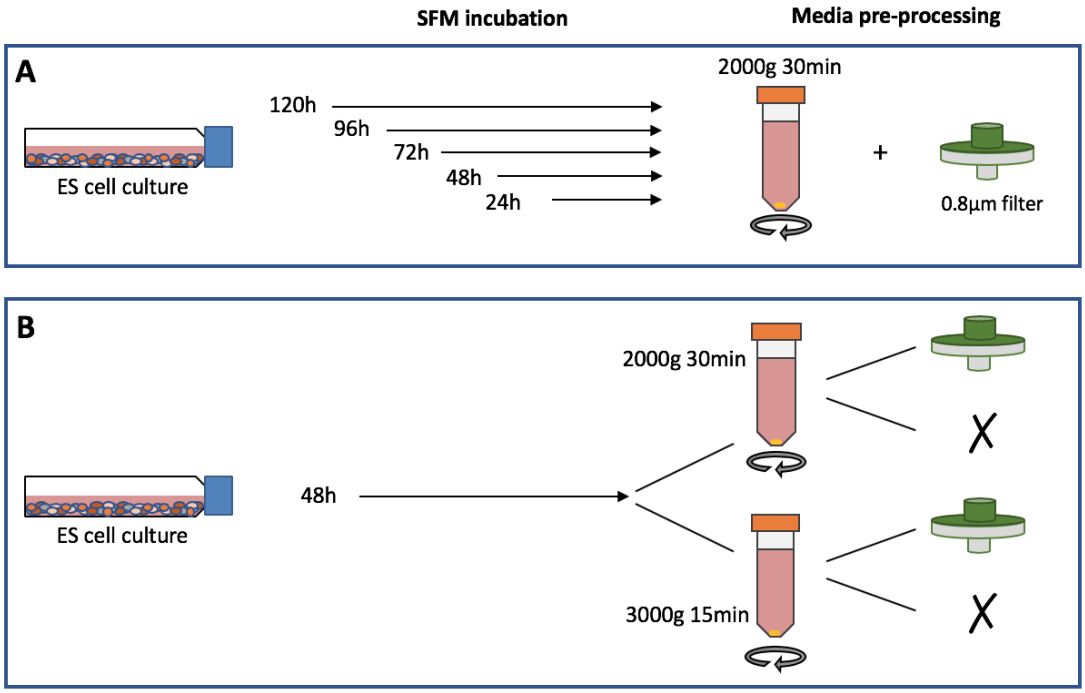


Figure 4.1: Optimisation of media collection conditions for exosome isolation. **A.** The impact of different incubation times in SFM was evaluated, using the protocol described in Section 4.2.2 to remove cellular contaminants. pH of media at collection point was assessed. **B.** Comparison of different protocols to remove cellular debris and larger vesicles, after an incubation of cells in SFM for 48 h. Cell viability at collection time was evaluated to ensure the exosomes originated from viable cells. All experiments were performed in triplicate.

4.2.4 Collection of exosome producing cells

At the time of media collection, cells were harvested (Section 2.2.3) and viable cell number counted (Section 2.2.5). This was to ensure the viability of cells producing exosomes as well as to obtain the number of cells from which exosomes had been isolated to allow normalisation of exosome number in downstream analyses. Cells

were collected by centrifugation (Section 2.2.3) and utilised as a control for downstream applications to be compared to isolated exosomes.

4.2.5 Isolation of exosomes using membrane-based affinity (MBA) method

For the isolation of exosomes from cell culture supernatant, the MBA method exoEasy Maxi kit (Qiagen) was used as summarised in Figure 4.2 (details of buffers remain company proprietary). In brief, an equal volume of conditioned culture media was combined with XBP buffer (binding buffer), mixed gently by inversion, and then passed through the spin column by centrifugation at 500 g for 1 min in order to bind exosomes to the membrane. This was followed by centrifugation at 5,000 g for 5 min to dry the membrane. The membrane was washed with 10 ml of XWP buffer (washing buffer) by loading this on the membrane and then centrifugation of the column at 500 g for 5 min. The spin column was transferred to a new collection tube and 400 μl of XE buffer (elution buffer) was applied to the membrane, incubated for 1 min at RT and the column centrifuged at 500 g for 5 min. The elute was re-applied to the membrane, incubated for 1 min and centrifuged at 5,000 g for 5 min. Exosomes were pelleted from the eluate by ultracentrifugation (UC) at 100,000 g for 70 min at 4 °C (Optima TL UC 100,000rpm with rotor TLA-55; Beckman Coulter). The supernatant was carefully removed and discarded (Figure 4.2).

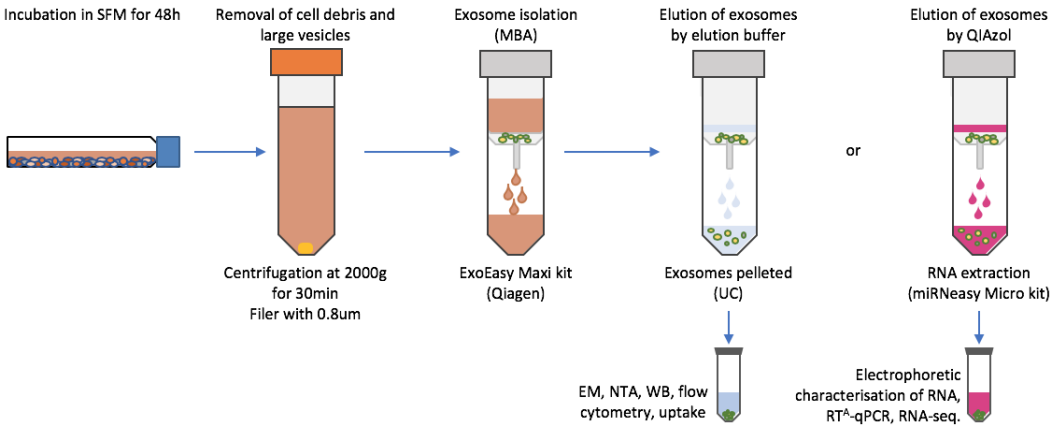


Figure 4.2: Summary of method to isolate and characterise exosomes using the membrane-based affinity (MBA) method exoEasy Maxi kit (Qiagen). Abbreviations: EM = electron microscopy, NTA = nanoparticle tracking analysis, WB = western blot.

The extraction of RNA from exosomes was performed on exosomes bound to the membrane instead of eluting exosomes in the XE buffer. Therefore, 700 µl of Qiazol (Qiagen) was applied to the membrane instead of 400 µl XE. Protocol for RNA extraction continues in Section 4.2.9.7.

4.2.6 Isolation of exosomes by differential ultracentrifugation (dUC)

The differential ultracentrifugation method (dUC) (Théry et al., 2006) was used with minor modifications (Figure 4.3). Briefly, cell culture supernatant was collected and centrifuged at 2,000 g for 30 min as described in Section 4.2.2. Then, supernatant was placed in polycarbonate bottles (Beckman Coulter) and centrifuged again for 30 min at 10,000 g (Beckman L-80 Ultracentrifuge with rotor Type 45Ti) in order to remove any remaining cell debris and large vesicles. The supernatant was carefully transferred to a new UC bottle and the sample centrifuged at 100,000 g for 70 min to pellet the exosomes. The media was aspirated and discarded, and the pelleted exosomes were collected from the bottom of the bottle by gentle pipetting to disaggregate the exosomes into 1 ml filtered PBS (0.1 µm filter, Sartorius Minisart) and were transferred to a UC tube and pelleted again 100,000 g for 70 min as described in Section 4.2.5. If the volume of media was less than 50 ml, filtered PBS was added up to the required volume for UC. The comparison between enrichment of exosomes using dUC and the MBA method (Section 4.2.2, Section 4.2.5) is summarised in Figure 4.3.

4.2.7 Isolation of exosomes by polymer mediated enrichment (PME)

For the isolation of exosomes from cell culture supernatant, the polymer-mediated enrichment (PME; Analytik-Jena AG, Jena, Germany) was evaluated. In this method, exosomes are captured by a polymer, which can be dissolved in buffer to produce a suspension of exosomes for use in downstream applications. Contents of buffers are company proprietary. Briefly, 1 ml of cell culture supernatant, processed as described in Section 4.2.2, was mixed with 30 µl of VCR 1 reagent and vortexed briefly. Then, 150 µl of VCR 2 reagent was added to the sample which was mixed by a second vortex. The sample was incubated for 1 min at RT before centrifugation at 13,000 g for 30 min. At this stage, a viscous substance was

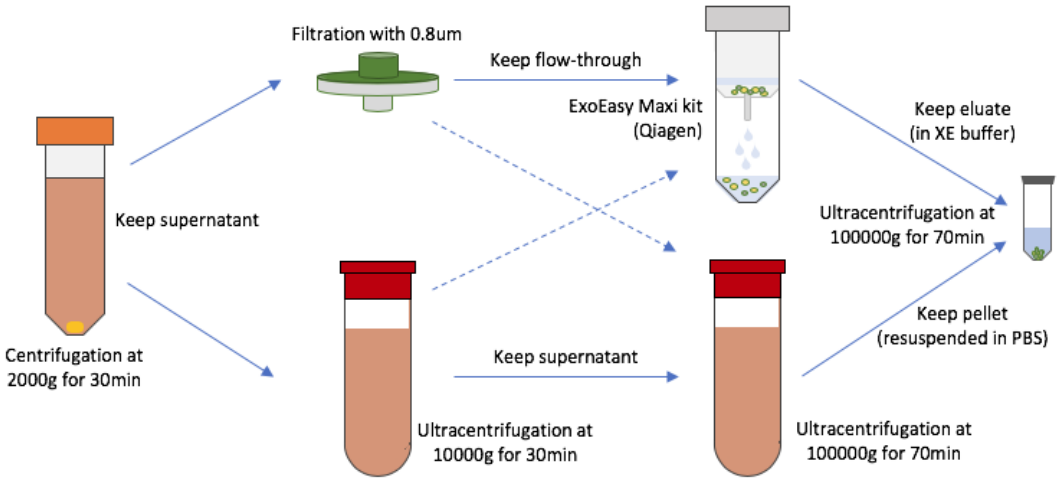


Figure 4.3: Comparison between MBA and dUC for the isolation of exosomes. As well as comparing both protocols (MBA top, dUC bottom; full lines), I evaluated the impact of removing cellular debris and larger vesicles by filtration or ultracentrifugation (dotted lines) on each isolation protocol.

visible on the sides and bottom of tube, which was the polymer capturing the exosomes. The supernatant was then carefully removed and the tube centrifuged for 10 s and any residual supernatant discarded. Protocol for RNA extraction continues in Section 4.2.9.7.

4.2.8 Plasma preparation and isolation of exosomes

For the isolation of exosomes from human plasma (0.5 ml), this was mixed with PBS to obtain 4 ml of final volume. This mixture was then passed through the 0.8 μm filter to remove cells and larger vesicles. For the isolation of exosomes, the MBA method described in Section 4.2.5 as used, with washes performed in a smaller volume (3.5 ml XWP buffer) to minimise loss of exosomes from plasma.

4.2.9 Characterisation of Exosomes

4.2.9.1 Nanoparticle Tracking Analysis

The size distribution and concentration of exosomes released from cells was performed by Nanoparticle Tracking Analysis (NTA) using the ZetaView® (Particle Metrix GmbH; Inning am Ammersee, Germany). NTA is based on the Brownian motion of particles, as the velocity of the random movement of particles suspended in liquid can be related to the size of the particle. Exosomes pelleted by UC (Section 4.2.5) were resuspended in 200 μl of filtered PBS and analysed by Cell

Guidance Systems (Babraham Research Campus, Cambridge, UK). Filtered PBS was analysed to assess the level of background noise and as a negative control. The reliability of NTA was evaluated by analysing polystyrene beads of known size (100 and 200 nm; 3000 Series Nanosphere™ Size Standards; Fisher Scientific). All analyses were performed using minimal brightness of 20 and analysing 11 different field positions.

4.2.9.2 Transmission Electron Microscopy

The protocol for the examination of isolated exosomes on transmission electron microscopy (TEM) was adapted from that of Théry et al. (2006). Pelleted exosomes (Section 4.2.5) were resuspended and fixed in 5 μl 2 % PFA in PBS for 30 min at RT. From this point, samples were processed by Mr. Martin Fuller (Electron Microscopy technician, Astbury Biostructure Laboratory, University of Leeds). Briefly, fixed exosomes were pipetted onto Parafilm M (Bemis NA, Sigma-Aldrich) and a 3.05 mm diameter formvar and carbon coated copper grid (AGAR Scientific, Stansted, UK) was inverted on top of the sample and incubated for 20 min at RT. Grids were then washed 6 times for 3 min each on ddH_2O. For visualisation by TEM, a 1 % uranyl acetate solution was applied to samples for 10 s to stain preparations, and left to air dry. Samples were imaged using a JEM 1400 TEM (JEOL, Zaventem, Belgium) at 100 kV, and images captured on AMT 1k CCD running on AMT v620 software.

4.2.9.3 Protein extraction

Cell pellets (used as controls, Section 2.2.4) were resuspended in 50 μL of Radio Immuno Precipitation Assay (RIPA) lysis buffer (1× PBS, 1 % Nonidet P-40, 0.5 % sodium deoxycholate (BDH), 0.1 % SDS) containing a cocktail of protease inhibitors (100 μl/ml phenylmethylsulfonyl fluoride (PMSF, Sigma-Aldrich), 10 μl/ml sodium orthovanadate (Sigma-Aldrich), 1 μl/ml leupeptin (Sigma-Aldrich), 30 μl/ml aproptinin (Sigma-Aldrich)) and incubated for 30 min on ice; samples were agitated by vortexing twice during this incubation. Protein lysate was then centrifuged at 12,470 g for 10 min at 4 °C (Microfuge 1-14, Sigma-Aldrich) and the supernatant retained. Then, 5 μL of the supernatant was mixed with 45 μL of ddH_2O and put aside for the determination of the protein concentration (Section 4.2.9.4). The remaining protein lysate was divided into 2 aliquots mixed with an equal volume of 2× SDS loading buffer (100 mM Tris-HCl pH 8.3, 4 % SDS, 0.2 % bromophenol blue,

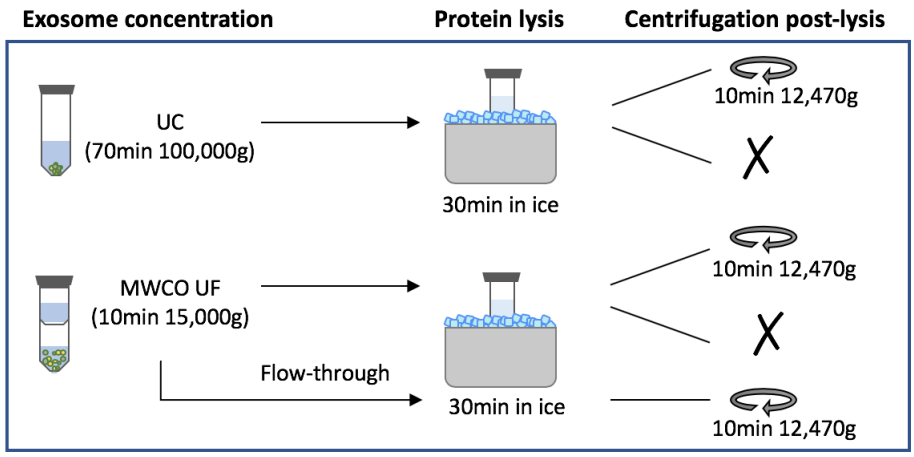


Figure 4.4: Optimisation of protein extraction protocol from exosomes. To test which protocol led to a higher protein yield from the isolated exosomes, ultracentrifugation (UC) and ultrafiltration (UF) were compared. In addition, performing or not the centrifugation step after the protein lysis was evaluated to see if this could have an impact in the final yield. The flow-through from the UF process was retained to test if exosomes were lost by the filter.

20 % glycerol (Sigma-Aldrich)): one with 200 mM DTT and the other without; and stored at −20 °C until required. DTT was used as reducing agent, and therefore, when proteins did not require reduction, DTT was not added.

Exosome pellets (Section 4.2.5) were resuspended in 25 μL of lysis buffer, incubated on ice for 30 min; samples were agitated by vortexing twice during this incubation. The centrifugation step was not performed, so the total lysate volume was retained. For the protein concentration assay (Section 4.2.9.4), 2 μl was mixed with 3 μl of ddH_2O. The remaining protein lysate was divided into 2 aliquots of 2× SDS loading buffer with or without DTT.

For the lysis of exosomal protein, different methods were compared as summarised in Figure 4.4. The UC step performed to concentrate the exosome eluate from the XE buffer (Section 4.2.5) was compared to ultrafiltration (UF) using a molecular weight centrifugal concentrator (MWCO; Vivaspin® 500, 100 kDα, Sartorius Minisart) in case exosomes were lost as part of the supernatant of UC. For this, exosomes eluted in XE buffer (400 μl) were added to the UF column and centrifuged at 15,000 g for 10 min at 4 °C, and the eluate (100 μl approximately) was collected. Protein lysis of the concentrated exosomes from both UC and UF was performed as described above, as well as from the flow-through of the UF in

case exosomes were not collected by the filter. In addition, performing or not the 10 min 12,470 g centrifugation after the 30 min lysis step on ice was evaluated (Figure 4.4), as both options are described in the literature (Enderle et al., 2015; Felicetti et al., 2016; Li et al., 2016; Paolini et al., 2016; Valadi et al., 2007).

4.2.9.4 Assessment of protein concentration

Protein concentration was determined using the BioRad DC Protein Assay (Bio-Rad) based on the Lowry Assay (Lowry et al., 1951). Briefly, BSA was used to prepare a standard dilution curve ranging from 0 to 2 mg/ml in 10 % lysis buffer from the protein lysis (Section 4.2.9.3). The average absorbance for each standard was calculated and a standard curve of protein concentration versus absorbance at 690 nm was generated using a line of best fit.

All standard and cellular protein lysates (5 μl of 1:5 dilution in ddH_2O) were added in triplicate into a 96 well plate (F96 Maxisorp Nunc-Immuno plate; Thermo Fisher Scientific); the protein content of exosome lysates (5 μl of 2:3 dilution in ddH_2O) was measured only once to minimise the amount of sample used. Reagent S (20 μl) was added to 1 ml of reagent A, and 25 μl of resulting mix was added to each well. Then, 200 μl of Reagent B was applied to all wells, the plate was carefully shaken manually for 10 s, and incubated for 15 min at RT. The absorbance of all wells at 690 nm was read using a multiscan microplate reader (Original Multiskan EX; Thermo Electron Corporation, Thermo Fisher Scientific). The protein concentration of each sample was determined by reading the absorbance of each sample from the BSA standard curve.

4.2.9.5 Western Blot

The BioRad Mini-Protean Tetra System (Bio-Rad) was used to perform western blots (WB). Protein lysate from exosomes (15 μL or normalised based on cell number) and cells (5–20 μg) were loaded into a Mini Protean® TGX™ 4–15 % gel (Bio-Rad), alongside 5 μL Geneflow Protein Ladder (BLUeye Prestained Protein Ladder; Geneflow) and transferred to an electrophoresis tank containing 2× SDS running buffer (25 mM Tris-HCl pH 8.3, 250 mM glycine (Sigma-Aldrich), 0.1 % SDS (w/v)). The gel was electrophoresed at a constant voltage of 100 V until the bromophenol blue dye had reached the bottom of the gel.

Antibody	Species	Concent. (µg/ml)	Id	Supplier	Target
EEA1	Rb	0.1	ab2900	Abcam	E. endosome
Grp75	Ms	0.11	ab2799	Abcam	Mitochondria
Pan cadherin	Ms	1	ab22744	Abcam	Pl. membrane
TATA TBP	Ms	0.7	ab818	Abcam	Nucleus
Ago-2	Ms	0.45	ab32381	Abcam	RNA-assoc. prot.
CD99	Ms	0.1	sc-53148	Santa Cruz	ES*
TSG-101	Ms	2.5	ab83	Abcam	Exosome
CD63	Ms	1	ab59479	Abcam	Exosome
CD81	Ms	1	sc-166029	Santa Cruz	Exosome
CD9	Ms	97	ab92726	Abcam	Exosome
Anti-mouse	Gt	1:5000	170-6516	Bio-Rad	*Secondary-HRP*
Anti-rabbit	Gt	1:5000	4010-05	South. Biotech	*Secondary-HRP*

Table 4.2: Optimised antibody concentrations or dilutions for detection of proteins by western blot analysis. All primary antibodies were incubated overnight at 4 °C, and secondary antibodies for 2 h at RT. Abbreviations: Ms = mouse, Rb = rabbit, Gt = goat, E. = Early, Pl. = plasma, RNA-assoc. prot. = RNA-associated protein. ES* = Ewing's sarcoma, binds cell surface glycoprotein. HRP = Horseradish peroxidase tagged.

Proteins were then transferred to nitrocellulose membrane 0.45 µm (Bio-Rad) by placing the gel and the membrane in-between 3 mm filter paper (Whatmann, supplied by GE Healthcare, Staffordshire, UK) and foams, held together in a transfer tank filled with transfer buffer (25 mM Tris pH 8.3, 192 mM glycine, 20 % methanol, 0.01 % SDS). Buffer was stirred and an ice block was placed in the tank to disperse heat generated during the transfer process. Protein transfer was performed for 2 h at 300 mA. After transfer, the membrane was removed from the Mini-Protean system and blocked by incubating in Odyssey Blocking Buffer (LI-COR; supplied by VWR) for 30 min on a rotating platform. The membrane was then incubated with the optimised antibody concentration (Table 4.2) in incubation solution (1:1 Odyssey Blocking Buffer and PBS, with 0.1 % Tween-20; Sigma-Aldrich) overnight at 4 °C on a rotating platform. Following incubation, membranes were washed 3 times for 5 min in PBS containing 0.1 % Tween-20 (v/v; PBS-T) in order to remove excess antibody. Membranes were then incubated for 2 h at RT on a rotating platform with secondary antibody in incubation solution (Table 4.2). Membranes were washed again in PBS-T and a final 5 min wash in PBS was performed to remove Tween-20 from the membrane. For visualisation, the membrane was incubated for 1 min with HRP substrate (Luminata™ forte HRP substrate; Merck-Millipore) and membranes imaged on ChemiDoc MP (Bio-Rad).

4.2.9.6 Flow cytometry for exosomes

Pelleted exosomes (Section 4.2.5) were resuspended in filtered PBS at a concentration based on viable cell number at time of collection (20 μl per 2.5×10^5 cells). For plasma samples, the exosome pellet was resuspended in 70 μl of filtered PBS. Then, 5 μl of the exosome dilution (exosomes from 6.25×10^4 cells) were incubated with the appropriate antibody dilution (Table 4.3; total volume 20 μl) in PBS for 30 min at 4 °C. Then, exosome dilution was resuspended in an equal volume of 1 μM CellTrace™Far Red (CT, Invitrogen; final concentration 0.5 μM in PBS; Table 4.3) and incubated for 20 min at 37 °C. Finally, labelled exosomes were diluted in PBS to a final volume of 250 μL and analysed immediately on the Cytoflex S flow cytometer (Beckman Coulter).

For the successful visualisation of exosomes on the flow cytometer, the gain settings for all lasers were set to the maximum (3000) in order to maximise the sensitivity and a specific threshold was defined for each laser used (7000 for APC and PB450, 8000 for PE and 5000 for FITC); these parameters were optimised by Dr. Adam Davison and Miss Charlotte Haunch-Smith. Samples were analysed for the same duration of time (1 min at 30 μl/ min) or until 10000 events had been evaluated. In order to set the size gates to define the population of interest, a mixture of bead sizes (100, 300, 500 and 900 nm beads; Megamix-Plus FSC; BioCytex SARL, Marseille, France) was used.

Dye or Antibody	Species	Fluorophore	Concentration	Id	Supplier
CT	-	APC	0.5 μM	C34564	Invitrogen
CD81	Mouse	FITC	0.125 μg/ml	551108	BD
CD63	Mouse	PB450	0.25 μg/ml	561984	BD
CD99	Mouse	PE	3.75 μg/μl	555689	BD

Table 4.3: Antibody and dye concentrations for flow cytometry on exosomes. CellTrace™incubation was performed for 20 min at 37 °C; antibody incubations were performed for 30 min at 4 °C. All dilutions were performed in filtered PBS. CT = CellTrace, generic dye.

Flow cytometry for ES exosomes was optimised by assessing the expression of CD99 ES marker using the anti-CD99 antibody (Table 4.3). A dilution curve of exosomes from a CD99-positive sample (TC-32) and a negative sample (plasma from healthy donor) was used. Exosomes were isolated and prepared from both as described

in Section 4.2.5 and Section 4.2.9.6. Then, a dilution curve of both samples (TC-32:plasma) was done for the ratios 100:0, 90:10, 75:25, 50:50, 25:75, 10:90 and 0:100 and each labelled with the same concentration of anti-CD63 and anti-CD99 (dual labelled). The exosome expression profile was evaluated on the Cytoflex S flow cytometer as described above.

4.2.9.7 RNA extraction from exosomes

For exosomal total RNA extraction, the miRNeasy Micro kit (Qiagen), previously described in Section 2.2.6.1, was adapted. As described in Section 4.2.5, 700 µl of Qiazol (Qiagen) was applied to the membrane of the MBA (exoEasy Maxi kit). Then, the column was centrifuged for 5,000 g for 5 min to collect the exosomes, which was then transferred to a 1.5 ml collection tube. After brief agitation by vortex, the RNA extraction protocol continues as described in Section 2.2.6.1.

When exosomes were isolated using the PME method (Section 4.2.7), the PME exosome miRNA extraction kit (version 12/15; Analytik-Jena) was used to extract RNA. To each pellet obtained (Section 4.2.7), 400 µl of Lysis Solution CBVE was added to the tube. The contents were mixed by gentle pipetting after which 20 µl of Proteinase K (Analytic-Jena) was added, mixed by vortexing for 10 s and incubated at 70 °C for 20 min under continuous agitation to increase lysis efficiency. Then, the tube was centrifuged at 13,000 g for 2 min, the supernatant was transferred to a new collection tube and 800 µl of ethanol was added and the tube contents mixed by pipetting. Then, a spin filter was placed into a new tube, and 650 µl of sample was transferred to the filter and centrifuged at 11,000 g for 1 min. This process was repeated until all the sample had been loaded onto the filter. Then, the tube containing the flow-through was discarded and the filter transferred to a new collection tube. At this stage, a DNA digestion step was performed in order to eliminate the cell free DNA that could be isolated alongside exosomes. For this, the DNase I (Qiagen) was used as described in Section 2.2.6.1. Following this, 500 µl of Washing Solution HS was added to the filter, the tube was centrifuged at 11,000 g for 1 min and the spin filter then moved to a new collection tube. Then, 650 µl Washing Solution LS was added to the filter, centrifuged again at 11,000 g for 1 min, and the spin filter moved again to a new collection tube, which was centrifuged at 13,000 g for 3 min to remove all remaining ethanol.

Finally, the spin filter was placed in a new collection tube, 25 µl of RNase free H_2O was applied to the filter which was incubated for 2 min at RT and centrifuged at 11,000 g for 1 min. This step was repeated with 15 µl of RNase free H_2O in order to increase the miRNA yield. The miRNA in a total volume of 40 µl was stored at −20 °C.

4.2.9.8 Electrophoretic characterisation of RNA for exosomes

The RNA quality and quantity was assessed using the Nanodrop (Section 2.2.6.2) and Pico 6000 Agilent Bioanalyser 2100 (Section 2.2.6.3). In addition, the Small RNA Agilent Bioanalyser 2100 was employed to evaluate the miRNA and small RNA concentration. Sample preparation was as for the Pico 6000 Agilent chip (Section 2.2.6.3) with some minor differences. For chip loading, 2 µl of dye (previously vortexed and centrifuged) was mixed with 40 µl of prefiltered gel and centrifuged at 13,000 g for 10 min at RT. The gel-dye mix (9 µl) was loaded onto the chip using minimal pressure for 60 s.

4.2.9.9 Total RNA-seq of exosomes

A total of 10 µl of exosomal RNA from each samples was used in the Total RNA-seq. This volume was equivalent to 83 % of the RNA extracted from exosomes (10/12 µl) collected from two T75 flasks of cells. Volume of sample was used to standardise the amount of exosomal RNA analysed using RNA-seq, as RNA concentration of exosomes cannot be quantified accurately. The RNA was prepared (Section 3.2.2.2) and sequenced (Section 3.2.2.3) as previously described for cells. The bioinformatic downstream analysis of the obtained reads was performed (Section 3.2.3) and the differential expression profile between producing cells and exosomes was evaluated by DESeq2 (Section 3.2.3.6). The RNA-seq data of the cells from which exosomes had been collected is described in Chapter 3 (Section 3.3).

4.2.9.10 RT^A-qPCR on exosomes

For the detection of the EWSR1-ETS fusion gene and the ES-CSC marker NRXN1 mRNA levels as part of the exosomal cargo, the maximum volume of exosomal RNA was analysed (10 µl; exosomes from 2 T75 flasks); it was not possible to quantify the concentration accurately. Cells from which exosomes were collected were used as a control, running a dilution curve from 0.4–4,000 ng in 10 µl in RNase-free H_2O. cDNA was prepared (Section 2.2.6.4) and qPCR performed (Section 3.2.4) using sequence

Primer or probe	Sequence (5'-3')	Concentration (nM)
EWSR1 F primer	GTCAACCTCAATCTAGCACAGGG	200
FLI1 R primer	CTGTCGGAGAGCAGCTCCAG	200
ERG R primer	CTGTCCGACAGGAGCTCCAG	200
EWSR1 probe	CTCCTACCAGCTATTCCTCTACACAGCCGACT	200
NRXN1*	Chr2 (4,991,850–51,032,536) on GRCh38	-
PPIA F primer	GGACCCAACACAAATGGTTCC	200
PPIA R primer	CTTTCACTTTGCCAAACACCA	200
PPIA probe	ATGCTTGCCATCCAACCACTCAGTCTTG	100

Table 4.4: Primers and probes used to amplify for EWSR1-FLI1, EWSR1-ERG, NRXN1 and PPIA. * NRXN1 was an assay on demand (hs00611317) developed by ThermoFisher, TaqMan™Gene Expression Assays, with unknown sequence but present on chromosome 2 on between the positions described above. F = forward, R = reverse.

specific primers and probes (Table 4.4; EWSR1-ETS protocol based on Peter et al. (2001)). For evaluation of the fusions, 50 cycles of 15 s at 95 °C to denature the cDNA, 1 min at 66 °C for the annealing of primers and 90 s at 72 °C for the extension phase were used. For NRXN1 evaluation, the PCR conditions were as described (Section 3.2.4) using an assay on demand (Table 4.4).

Data was exported and analysed using the RQ Manager 1.2.1 (Thermo Fisher Scientific) and C_t calculated using a threshold of 0.05. Expression of PPIA was used as housekeeping gene for cells and as a PCR positive control for exosomes: this was validated by RNA-seq data as reads for PPIA were detected in exosomes. In addition, RT^A- products (samples in the absence of reverse transcriptase enzyme) and H_2O were used as negative controls, to control for contaminants.

4.2.10 Exosome preparation for uptake experiments

For the evaluation of exosome uptake by cells, exosomes were labelled with CT (Section 4.2.9.6) at a ratio of 2.5×10^5 producing cells in 0.5 μM CT or using DMSO (vehicle) as control. After incubation (Section 4.2.9.6), 1 ml PBS was added to dilute any excess dye, mixed thoroughly and centrifuged at 100,000 g for 70 min at 4 °C to pellet the labelled exosomes. The pellet was then resuspended in media to obtain a dilution of exosomes from 10× recipient cells in 10 μl (i.e. when 5×10^4 recipient cells, exosomes from 50×10^4 producing cells were resuspended in 10 μl). Exosomes were then added to target cells and incubated at 37 °C to investigate their uptake by cells (Section 4.2.11) and impact on cellular behaviour (Section 4.2.12), as summarised in Figure 4.5.

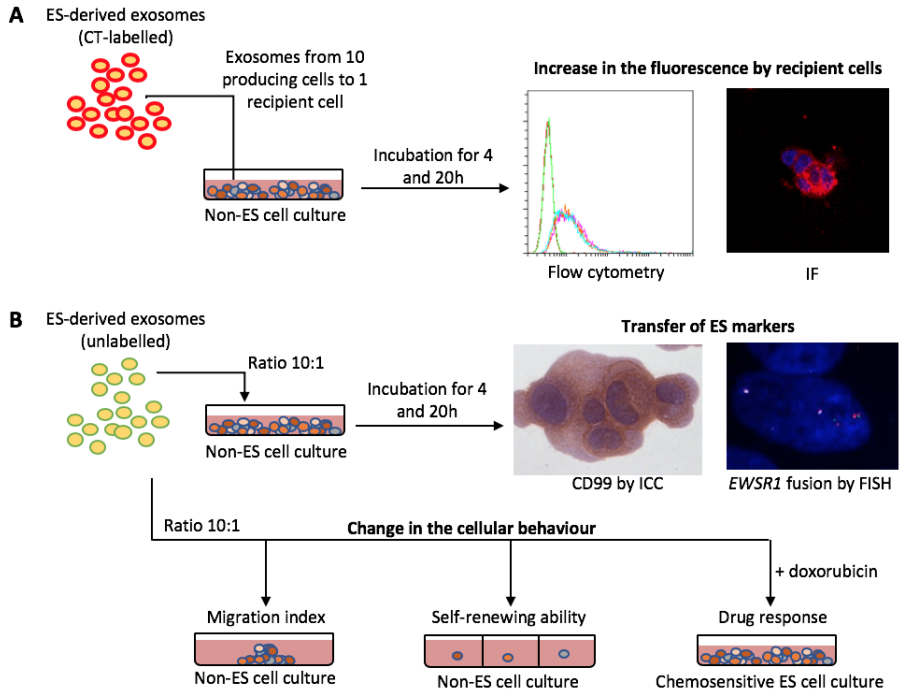


Figure 4.5: Summary of the ES exosome uptake experiments on non-ES cells. **A.** Evaluation of the uptake of ES-derived exosomes by non-ES cells based on the increase in fluorescence (Section 4.2.11). **B.** Evaluation of the transfer of ES markers CD99 and *EWSR1* fusion to non-ES cells (top) and changes in the migration, self-renewing and drug sensitivity (bottom) of recipient cells upon uptake of ES-derived exosomes (Section 4.2.12).

4.2.11 Evaluation of exosome uptake by cells

4.2.11.1 Assessment of exosome uptake by flow cytometry

Cells (5×10^4) were seeded in 24-well plates in appropriate media containing FCS and incubated for 24 h to let cells adhere to bottom of the plate. Then, labelled exosomes or DMSO control were added to wells (Section 4.2.10) and cells incubated at 37 °C (native conditions) or 4 °C to test if uptake by cells was an active process, and therefore, could be inhibited at lower temperatures. After 4 and 20 h incubation, cells were harvested (Section 2.2.3) and resuspended in FACS buffer (Section 2.2.10) and analysed immediately on the Cytoflex S flow cytometer (Beckman Coulter) (parameters used: forward and side scatter 32; APC (CT^+) 300). The uptake of SK-N-MC exosomes by cells was first evaluated in SK-N-MC cells, and then assessed in TTC-466 (ES), NB69 (NBL) and SK-N-SH (NBL) using the same conditions.

To identify the optimal exosome incubation time, different exosome exposure lengths (1, 2, 4, 6, 10, 20 h) were tested, both at 37 °C and 4 °C. To identify the optimal amount of exosomes to add to recipient cells, different ratios (1:1, 5:1, 10:1, 25:1 and 50:1 producing cell : receiving cell) were tested.

4.2.11.2 Assessment of exosome uptake by immunofluorescence

For direct visualisation of exosome uptake into cells, immunofluorescence (IF) on whole cells was used. For this, glass coverslips (22 mm × 22 mm; VWR) were sterilised by submerging them in 100 % ethanol, allowing them to air dry and then placing one per well of a 6-well plate. Cells were harvested (Section 2.2.3) and seeded at a density of 1×10^4 cells in 200 µl of media (Table 2.1) on the coverslips. Once cells had adhered to the coverslip (16 h), wells were gently flooded with 1 ml of media and incubated for 2 more days to allow cells to recover from harvest. Then, labelled exosomes or DMSO labelled control (Section 4.2.10) were added and cells incubated at 37 °C or 4 °C.

After incubation for 4 h or 20 h, media was aspirated and cells washed with 1 ml of PBS. When a live dye was used to label cells (MT or CT; Table 4.5) this was added 30 min prior to aspirating the media so the dye could be taken up prior to fixing the cells. The protocol used was as recommended by Sigma-Aldrich for Prestige antibodies (Table 4.5). Briefly, cells were fixed with ice cold 4 % PFA in 10 % FCS for 15 min at RT, washed in 1 ml PBS and permeabilised in 1 ml of 0.1 % Triton-X in PBS 3 times for 5 min. Cells were washed again with 1 ml PBS to remove any residual Triton-X and then incubated overnight at 4 °C with 100 µl of primary antibody diluted in 4 % FCS in PBS (Table 4.5). Cells were then washed 4 times in 1 ml PBS for 10 min to remove any unbound antibody and cells incubated at RT in the dark for 1.5 h with 100 µl of secondary antibody diluted in 4 % FCS in PBS (Table 4.5), after which cells were washed in PBS to remove any residual antibody. To stain the nucleus, 0.2 µg/ml of DAPI diluted in 100 µl PBS was added to cells and incubated at RT in the dark for 20 min. Finally, cells were washed 4 times in 1 ml PBS for 10 min and mounted using Faramount. Slides were analysed using the Nikon A1R confocal microscope (Nikon) with the NIS-Elements C software (Nikon).

Dye or antibody	Species	Fluorophore	Concent.	ID (supplier)	Target
CT (D)	–	Cy5	0.5 μM	C34564 (Invitrogen)	–
MT (D)	–	TexasRed	0.1 μM	M7512 (Invitrogen)	Mitoch.
EZR (A)	Ms	–	2 μg/ml	AMAb90967 (Prestige)	Pl. memb.
PDIA3 (A)	Ms	–	2 μg/ml	AMAb90988 (Prestige)	ER
GOLPH4 (A)	Rb	–	0.5 μg/ml	HPA001677 (Atlas Ab.)	Golgi
CD63 (A)	Ms	-	2 μg/ml	ab8219 (Abcam)	Exosomes
IgG (A)	Ms	–	2 μg/ml	X0931 (Dako)	–
IgG (A)	Rb	–	0.5 μg/ml	08-6199 (Invitrogen)	–
Anti-Ms (A)	Gt	FITC	4 μg/ml	A11001 (Invitrogen)	–
Anti-Rb (A)	Gt	TexasRed	4 μg/ml	A11036 (Invitrogen)	–

Table 4.5: Dye (top) and antibody (bottom) concentrations for IF. Abbreviations: CT = CellTrace™FarRed, MT = MitoTracker CMXRos, D = dye, A = antibody, Ms = mouse, Rb = rabbit, Gt = goat, ID = identifier, Mitoch. = mitochondria, Pl. memb. = plasma membrane, ER = endoplasmic reticulum. Suppliers information: Prestige Antibodies is part of Sigma-Aldrich; Atlas Antibodies is part of Cambridge Biosciences; Invitrogen is part of Thermo Fisher Scientific.

4.2.12 Evaluation of changes on recipient cells by exosome uptake

4.2.12.1 CD99 expression after exosome uptake by ICC

To evaluate the change in CD99 expression on non-ES cells (NB69 and SK-N-SH NBL cells; MSC as non-cancer cells; Table 2.1) induced by ES exosomes, cytospins of cells (Section 2.2.7) after incubation with exosomes for 4 and 20 h (Section 4.2.10) were done and CD99 expression evaluated by ICC (Section 2.2.9). As a control, non-ES cells were not exposed to ES exosomes to assess the basal levels of CD99 expression on these cells. SK-N-MC cells were used as positive control.

4.2.12.2 FISH after exosome uptake

To evaluate the transfer of the *EWSR1* fusion gene via ES exosomes to non-ES cells (NB69 and SK-N-SH NBL cells; MSC as non-cancer cells; Table 2.1), cytospins of cells (Section 2.2.7) after incubation with exosomes for 4 and 20 h (Section 4.2.10) were evaluated by FISH (Section 2.2.8). As described in Figure 2.2, intact probe would be seen as a yellow fluorophore (both 3' and 5' of *EWSR1* probe together), which is expected in non-ES cells as these do not have the characteristic *EWSR1* fusion. However, if the fusion is transferred via ES exosomes to non-ES cells, an increase in red probe should be seen, as this is the 5' region of *EWSR1*, present on the *EWSR1* gene fusion. As a negative control, recipient cells not exposed to exosomes were

used, as these should not contain the *EWSR1* gene fusion. SK-N-MC cells were used as positive control.

4.2.12.3 Self-renewing ability of non-ES cells after incubation with exosomes from high self-renewing cells

In order to evaluate if exosomes could induce changes in the self-renewing ability of recipient cells, exosomes from SK-N-MC, a high self-renewing ES cell line, were incubated with non-ES cells (NB69 and SK-N-SH as NBL cell lines; MSC cells as non-cancer cells; Table 2.1. For this, the self-renewing assay (Section 2.2.11) was used adding exosomes (Section 4.2.10) from 1×10^4 SK-N-MC cells (ratio 10 : 1 producing cells to receiving cells) to the single-cell suspension prior to seeding in the ten 96-well plates. To compare the effect of exosomes on recipient cells, 10 control plates containing cells that had not been exposed to exosomes were included, confirming the self-renewing ability of these cells in the absence of exosomes.

4.2.12.4 Migration index of non-ES cells after incubation with ES exosomes

To test if the transfer of SK-N-MC exosomes to non-ES cells (SK-N-SH cell line; Table 2.1) could have effects on the migration capacity of recipient cells, the migration from a 3D spheroid was evaluated (Section 2.2.12). Exosomes (ratio 10 : 1 producing cells to receiving cells; Section 4.2.10) were added to the wells containing the already-formed spheroids at 0 h (when spheroids were transferred from the low-adherent plates to the gelatin-coated wells). As a control, cells not incubated with exosomes were evaluated alongside to confirm the migratory index without the effect of exosomes. SK-N-MC cells were used as positive control for the migration assay.

4.2.12.5 Drug sensitivity of cells after uptake of exosomes from drug resistant cells

To assess if the uptake of exosomes from resistant cell lines conferred a survival advantage towards cytotoxic drugs, native SK-N-MC cells (1×10^4) were seeded in triplicate in 24-well plates. After 24 h, exosomes from doxorubicin-resistant SK-N-MC cells (SK-N-MC cultured in 28 nM doxorubicin until growing as stable cell line, performed by Dr. Elizabeth Roundhill) were isolated and added to wells containing native SK-N-MC cells at a ratio of 10 producing cells to 1 receiving cell (Section 4.2.10) and incubated for 24 and 48 h. These 2 incubation times in

exosomes were used to test if time exposure to exosomes had an effect on the change of phenotype on recipient cells. Cells were then treated with different concentrations of doxorubicin (Section 2.2.13; 1, 5, 10, 25, 50, 100 and 200 nM) or vehicle control (ddH_2O) in triplicate. After a 48 h incubation in the presence of the drug, cells were harvested keeping the supernatant as described in Section 2.2.3, pelleted and resuspended in 0.5 ml of media and the viability of cells was assessed using trypan blue exclusion assay (Section 2.2.5). As a control, the same experiment was performed in SK-N-MC cells that did not receive exosomes from doxorubicin-resistant SK-N-MC cells in order to compare the sensitivity to doxorubicin with or without exposure to exosomes.

4.2.13 Statistical analysis

Data from phenotypic characterisation was analysed by descriptive statistics. Unpaired t-test (2 datasets) or ordinary one-way ANOVA for multiple comparisons (>2 datasets) were used, differences were considered significant when $p<0.05$. EC50 evaluation was performed using a nonlinear regression on dose-response inhibition for a variable slope. All analyses were undertaken using GraphPad Prism 6 software, except the descriptive statistics that were performed on Microsoft Excel version 15.24. For the multidimensional scaling (MDS), the distance between samples was calculated using Spearman measurment and ploted using R. For the differential expression analysis of the total RNA-seq data between cells and exosomes, the DESeq2 package was used (Section 3.2.3.6).

4.3. Results

4.3.1 Optimisation of cell culture conditions for the isolation of exosomes

In order to obtain an optimal exosome yield from cell cultures, different incubation times in SFM were tested to identify the exposure to SFM that led to an increased exosome product without affecting cell viability or exosome quality.

Cells exposed to SFM for longer periods of time (96–120 h) were less viable and had lower cell number compared to cells cultured in normal serum conditions (96 h $p = 0.0004$, 120 h $p < 0.0001$). In contrast, cells exposed to SFM for 24–72 h were

not affected (Figure 4.7A,B), with the exception of 24 h, that the cell number was higher than the control ($p = 0.01$). The pH of cell culture media at time of collection was inversely correlated with cell viability ($R = -0.49$). Whereas cells incubated in normal serum conditions (control) had a pH of 6.7, all cell culture conditions in SFM had higher pH, which increased the longer the cells were maintained in SFM (Figure 4.7B).

To further decide which incubation time provided a higher exosome yield, the RNA concentration was evaluated from cells and isolated exosomes. An inverse correlation was seen between cells (24–120 h: range 2,452–7 ng/μl) and exosomes (24–120 h: range 217–558 ng/μl) ($R = -0.92$), where the higher the cellular RNA concentration, the lower the exosomal RNA concentration (Figure 4.6A). Similarly, miRNA concentration showed an inverse correlation between cells (24–120 h: range 19–761 pg/μl) and exosomes (24–120 h: range 5,535–1,440 pg/μl) ($R = -0.89$), indicating that the higher the miRNA concentration in exosomes, the lower it was for cells (Figure 4.6B). When comparing the sRNA and miRNA concentrations of exosomes across the different incubation times (Figure 4.7C), the RNA concentration was reduced with prolonged incubation times (although not significant across time), consistent with a lower cell viability of the producing cells.

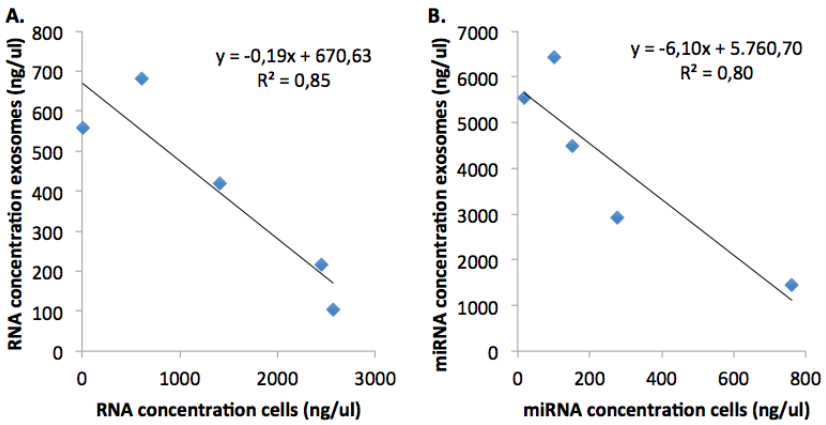


Figure 4.6: Correlation of RNA concentration between cells and exosomes. Inverse correlation between total RNA concentration **A** or miRNA concentration **B** of cells and exosomes isolated from these across time (24–120 h).

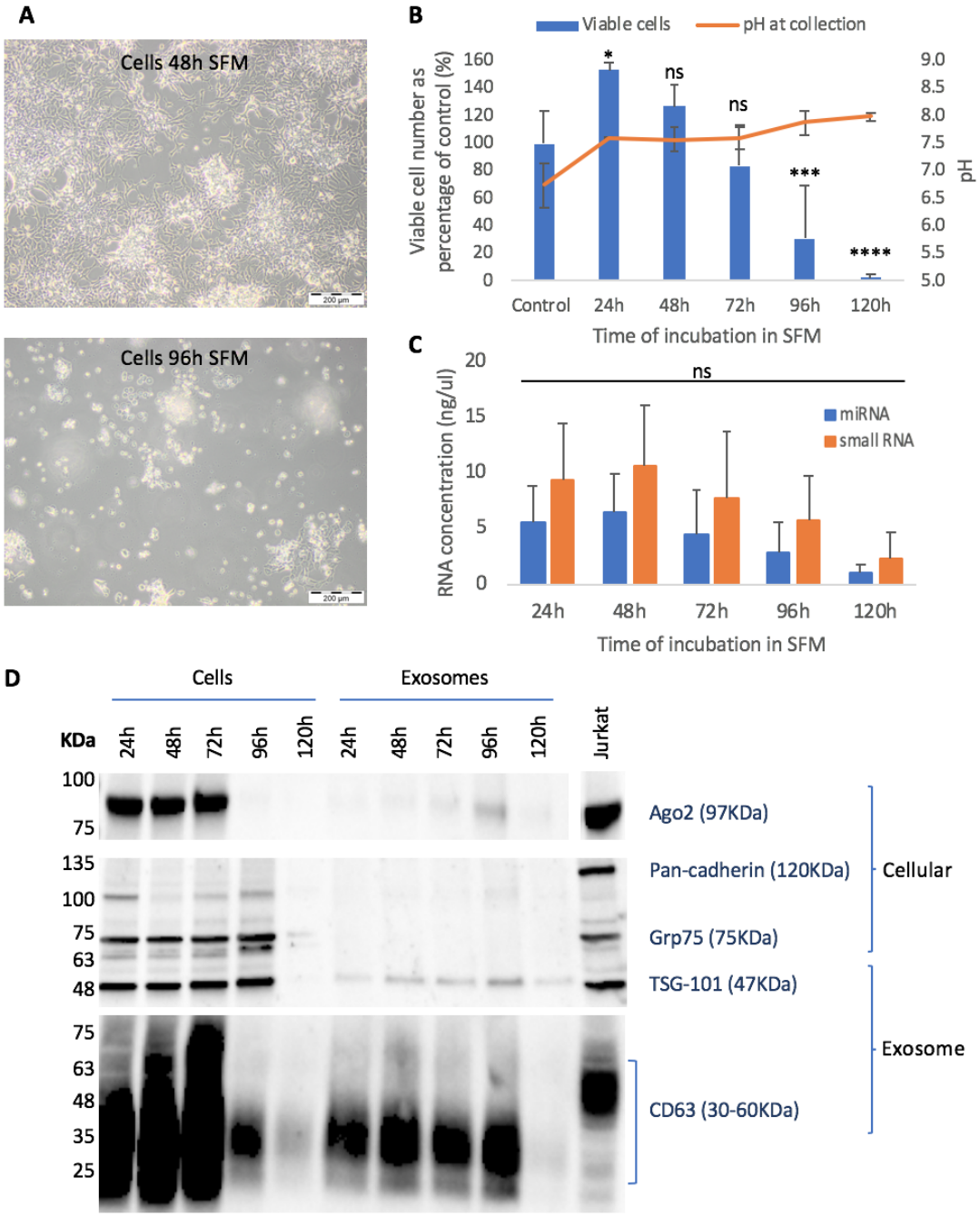


Figure 4.7: Optimisation of the incubation time in SFM for the isolation of exosomes from viable SK-N-MC cells. **A.** Representative photographs of cells after 48 h or 96 h in SFM. **B.** Viable cell number expressed as a percentage of control (cells incubated in normal serum conditions) and pH of media over the different incubation times. **C.** sRNA and miRNA concentration (ng/μl) of exosomes after different incubation times in SFM. In B and C data are shown as mean ± SD of 3 independent repeats. **D.** Expression of exosomal markers and cellular markers in isolated exosomes and producing cells following incubation in SFM for 24–120 h. Cells (20 μg or in the case of cells after 120 h incubation maximum volume was loaded, 15 μl) and exosomes (15 μl) were electrophoresed. Figure representative of 3 independent repeats. KDa = Molecular weight marker.

Finally, to examine the yield and purity of isolated exosomes, protein expression of exosomal markers and cellular contaminants was assessed (Figure 4.7D). WB revealed expression of different exosomal markers (TSG-101 and CD63) in exosomes isolated from 24–96 h, with hardly any expression at 120 h, reflecting the low numbers of viable cells at this time point. In addition, expression of cellular contaminants (Pan-cadherin, Grp75 and Ago2) was absent in exosomes from 24–72 h, suggesting the isolated exosomes were free from contaminants. In contrast, exosomes from 96 h incubation expressed Ago2, reflecting the presence of RNA-protein complexes that could impact on downstream RNA analysis.

Combining all the evidence, 48 h was chosen as the optimal incubation time in SFM for the isolation of exosomes, as this provided the best balance between viability of producing cells and the obtained exosome quality and yield.

4.3.2 Optimisation of the media collection method prior to exosomes isolation

To remove all cellular debris and larger vesicles, 4 pre-processing centrifugation steps were evaluated; 2,000 g for 30 min or 3,000 g for 15 min, with or without filtration (Figure 4.8).

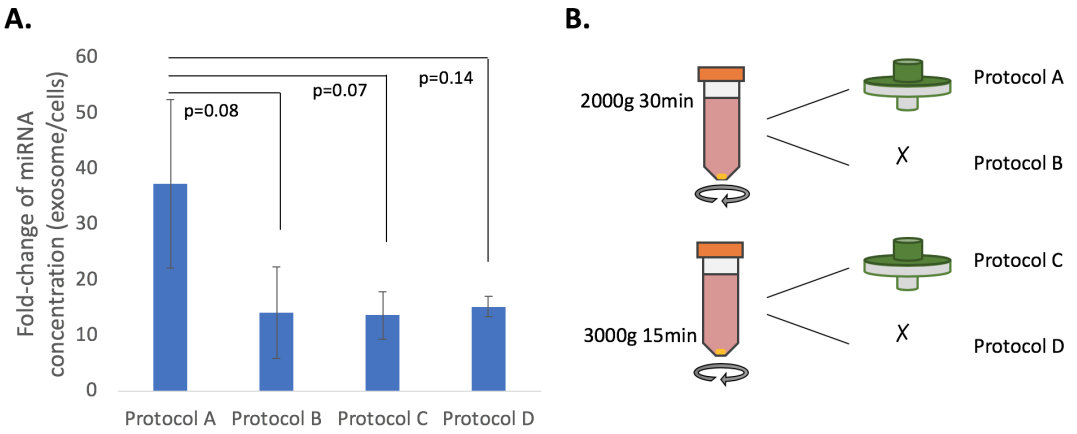


Figure 4.8: Differences in miRNA concentration of exosomes isolated from different media pre-processing protocols. **A.** Fold-change increase in miRNA concentration of exosomes compared to producing cells for the 4 different pre-processing methods. Data as mean ± SD from 3 independent repeats (protocol D 2 repeats). ANOVA was used to determine statistical differences across protocols. **B.** Description of each pre-processing protocol.

Exosome product from the 4 different protocols was evaluated by its miRNA concentration compared to the producing cells. Protocol A (centrifugation at 2,000 g for 30 min and filtration with a 0.8 μm filter) (Figure 4.8B) had the highest enrichment of miRNAs in exosomes compared to the other pre-processing protocols (Figure 4.8A). However, due to the variability of miRNA yield through the repeats (high SD), the differences were not statistically significant (A vs B $p = 0.08$, A vs C $p = 0.07$, A vs D $p = 0.14$). Taken together, protocol A was chosen for the pre-processing of media prior to the isolation of exosomes.

4.3.3 Adaptation of the protein extraction protocol for exosomes

Different methods were compared to identify the best protocol for the lysis of exosomes and recovery of exosomal protein, as summarised in Figure 4.9A.

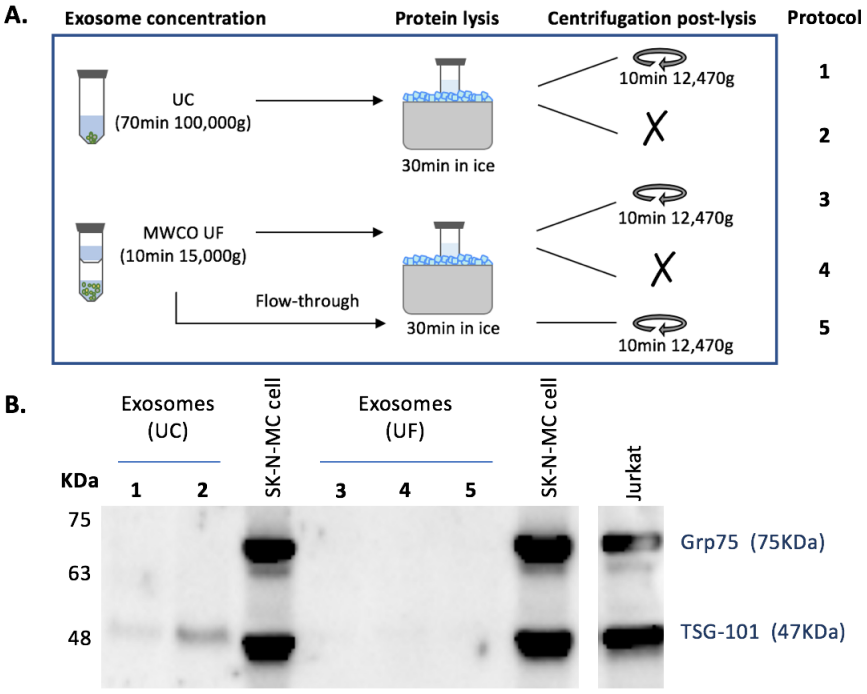


Figure 4.9: Optimisation of protein extraction from exosomes evaluated by WB. **A.** Different protocols compared as described in Section 4.2.9.3. **B.** Exosome yield evaluated by expression of TSG-101 exosomal marker across the different methods. Cellular marker Grp75 as negative control; SK-N-MC and Jurkat cells (50 μg) as positive controls; exosomes (15 μl) from same starting volume across methods. WB is representative of 2 independent repeats. IDs from diagram (1-5) match samples loaded in WB.

Data from two independent repeats revealed that using UC to concentrate the eluted exosomes and not performing the lysis centrifugation step (protocol 2) led to the highest protein yield as shown in Figure 4.9B. Therefore, this method was taken forward for the extraction of protein from eluted exosomes.

4.3.4 Exosome protein expression profile can be evaluated by flow cytometry

Most flow cytometers have been developed for the study of cells, and therefore, are not prepared to study particles as small as exosomes (Pospichalova et al., 2015). To overcome loss of exosomes in the so called 'noise' of flow cytometry, all the laser gains on the CytoFlex S were set to maximum (Section 4.2.9.6). A mixture of different fluorescence size beads (100–900 nm) was used to define the position of different populations by size on the scatter plot (Figure 4.10A) and filtered PBS was run to stablish the background noise (Figure 4.10B). As unlabelled exosomes were still too small to be successfully distinguished based on the size scatter (Figure 4.10C), they were labelled with fluorescent dyes and antibodies so they could be detected by their increase in fluorescence. First, events were selected based on the size gates generated in Figure 4.10A, as to only investigate events from 100–900 nm. Then, each dye or antibody was run alone (matching concentrations used to label exosomes) to set the negative gates (Figure 4.10D), so when the labelled exosomes were run they could be distinguished from the free dye or antibody.

Firstly, exosomes were labelled with CT (generic dye) in order to verify that the isolation method had worked: CT labels any vesicle independent of its expression profile. Exosomes were also labelled with antibodies against the exosome markers CD63 and CD81 and the ES marker CD99 (Figure 4.10D). As shown in Figure 4.10D, a population can be distinguished from the free dye or antibody, which are the labelled exosomes. Knowing that the use of these dye and antibodies allow the detection of exosomes by flow cytometry, the actual strategy to evaluate the expression profile of the isolated exosomes was performed. For this, exosomes were labelled with both CT (generic labelling) and an antibody against one of the markers of interest (i.e. CT + CD63), which allowed quantification of the number of exosomes (CT+) expressing the specific marker (CD63+).

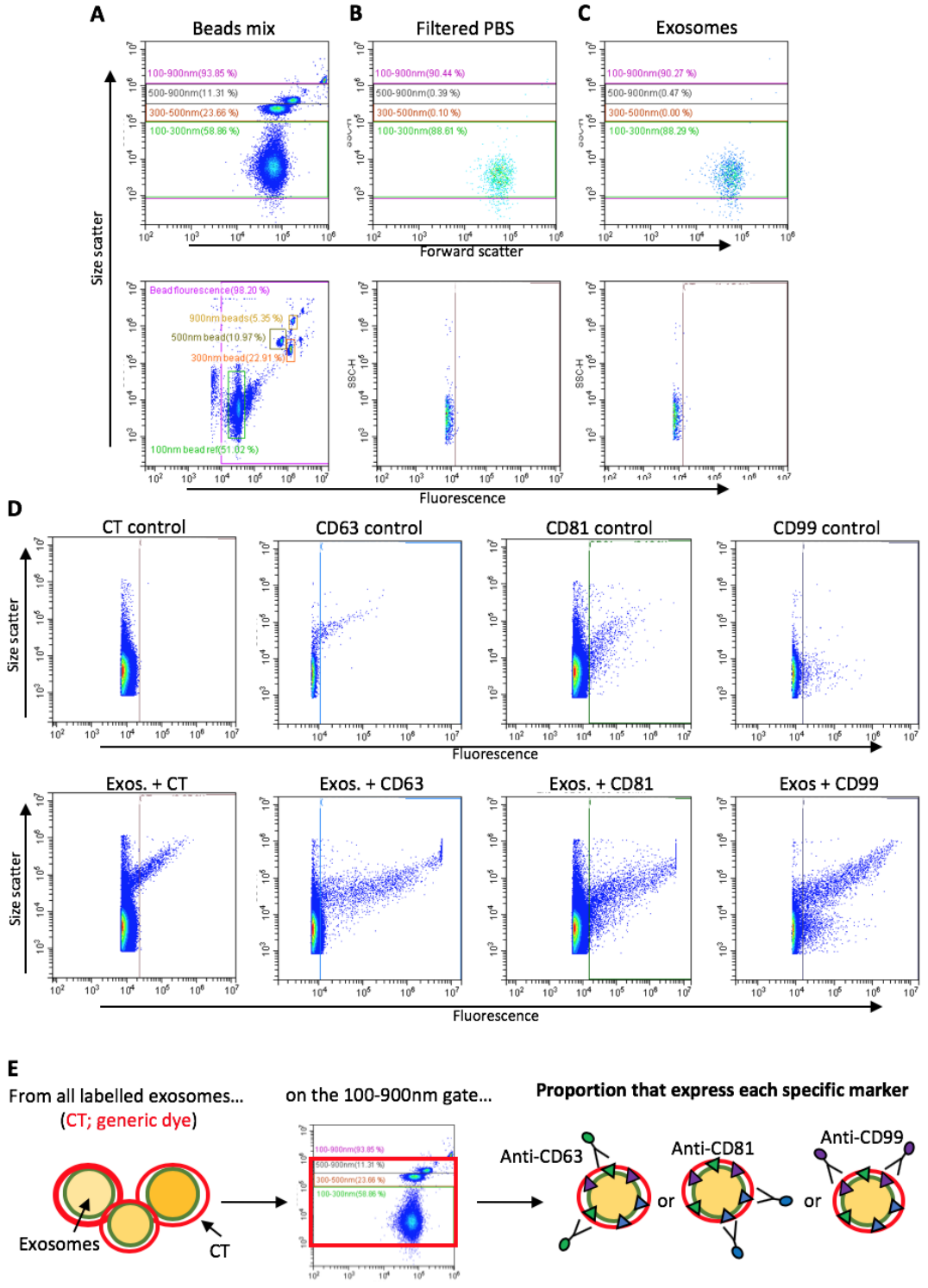


Figure 4.10: Flow cytometry strategy to study exosomes. Different gating and negative control strategies were used for the successful separation of exosomes from the background noise. **A.** Beads of mixed sizes were used to establish the different size ranges. **B.** Filtered PBS and **C.** unlabelled exosomes were analysed to confirm they did not produce fluorescence. **D.** Profiling of dye/antibody alone (control) to determine the negative gate (top) and labelled exosomes to determine their expression (bottom) based on the intensity of fluorescence. **E.** Strategy for the dual labelling of exosomes.

This strategy allowed the gating of all exosomes between 100–900 nm labelled by CT, and from these, investigate what percentage expressed each of the antibodies investigated (Figure 4.10E). This same strategy was used for dual antibody labelling (i.e. CD63-CD99), to quantify how many CD63+ exosomes expressed CD99.

To evaluate the specific binding of CD99, the expression of CD99 was evaluated in a dilution curve from CD99-positive and negative exosomes (TC-32 and control plasma respectively). As shown in Figure 4.11A, with increasing number of exosomes from TC-32 cells there was higher expression of CD99, with a shift from CD63 only to CD63-CD99 dual expression. When comparing the percentage of exosomes that express CD63 only or express both CD63 and CD99, a gradual change is observed across the different ratios investigated (Figure 4.11B). It is important to mention that due to the different origin of the exosomes (cell culture vs. plasma), the yield could not be normalised in the same manner, leading to higher yield for the plasma samples. Altogether, it can be concluded that the CD99 labelling on exosomes is specific.

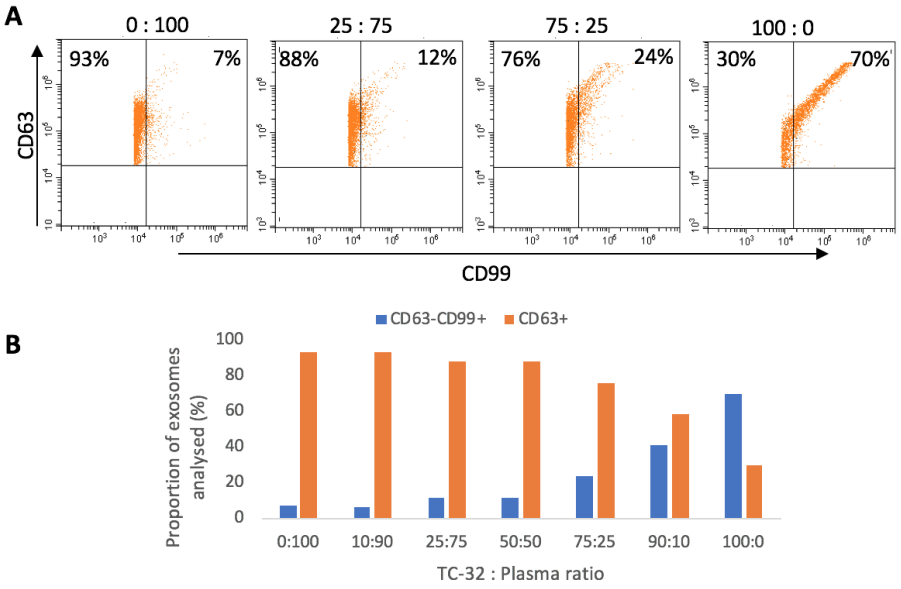


Figure 4.11: Evaluation of CD99-specific labelling on exosomes. Dilution curve of CD99-positive exosomes (TC-32) and CD99-negative exosomes (healthy donor plasma) labelled with CD63 and CD99. **A.** Dual labelling of exosomes on ratios 0:100, 25:75, 75:25 and 100:0 (TC32 : healthy plasma). The change in the exosome profile is shown as the shift of exosomes from CD63 only (top-left quadrant) to CD63-CD99 labelling (top-right quadrant). **B.** Change on the detection of CD63+ only exosomes (orange) to CD63-CD99+ exosomes (blue) across the different ratios. Only events with size 100–900 nm and expressing the primary marker (CD63) are gated. Data as percentage of exosomes with each profile.

4.3.5 Evaluation of different exosome isolation techniques

The PME methodology is developed to isolate exosomes from 1 ml of starting volume (in this case cell culture supernatant). Therefore, in order to compare this to the MBA protocol (Section 4.2.5), the starting volume was reduced to 1 ml of pre-processed cell culture supernatant and the volume increased with PBS to the recommended volume for extraction. To establish which protocol was more efficient in the isolation of exosomes, the miRNA concentration of the isolated exosomes was compared.

Data from 3 independent repeats showed that a higher miRNA concentration was obtained through the PME protocol (1755±957 pg/μl, range 567–3,104 pg/μl, $n = 9$) compared to the MBA (659±397 pg/μl, range 258–1,358 pg/μl, $n = 9$) ($p = 0.006$, Figure 4.12). Therefore, for small volumes of media (≤1 ml), the PME methodology was more useful for isolating exosomes than the MBA method when evaluating the yield based on miRNA concentration.

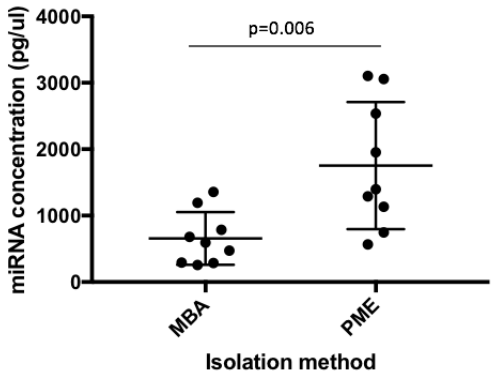


Figure 4.12: Comparison of MBA and PME exosome isolation methods assessed by the miRNA concentration (pg/μl) obtained from exosomes isolated from 1 ml of cell culture supernatant. Data expressed as mean ± SD from 3 independent repeats.

The optimised MBA isolation protocol (Section 4.2.5) was compared to the gold standard dUC method (Théry et al., 2006). A third strategy combining the media pre-processing step (UC vs. filtration) with the isolation step (UC vs. MBA) was performed to evaluate if differences in the exosome yield were due the pre-processing of media rather than the isolation method (Section 4.2.6).

Particle number data (NTA) of 3 independent repeats showed a higher number of exosomes when using the MBA protocol ($3 \times 10^{10} \pm 6 \times 10^{9}$ exosomes, range

2×10^{10}–4×10^{10} exosomes, $n = 6$) compared to the UC protocol ($6 \times 10^{9} \pm 7 \times 10^{9}$ exosomes, range 6×10^{7}–2×10^{10} exosomes, $n = 5$) ($p = 0.0005$; Figure 4.13A). When the differences in pre-processing protocol were taken into consideration, data showed no statistical differences between using filtration or UC (Figure 4.13A).

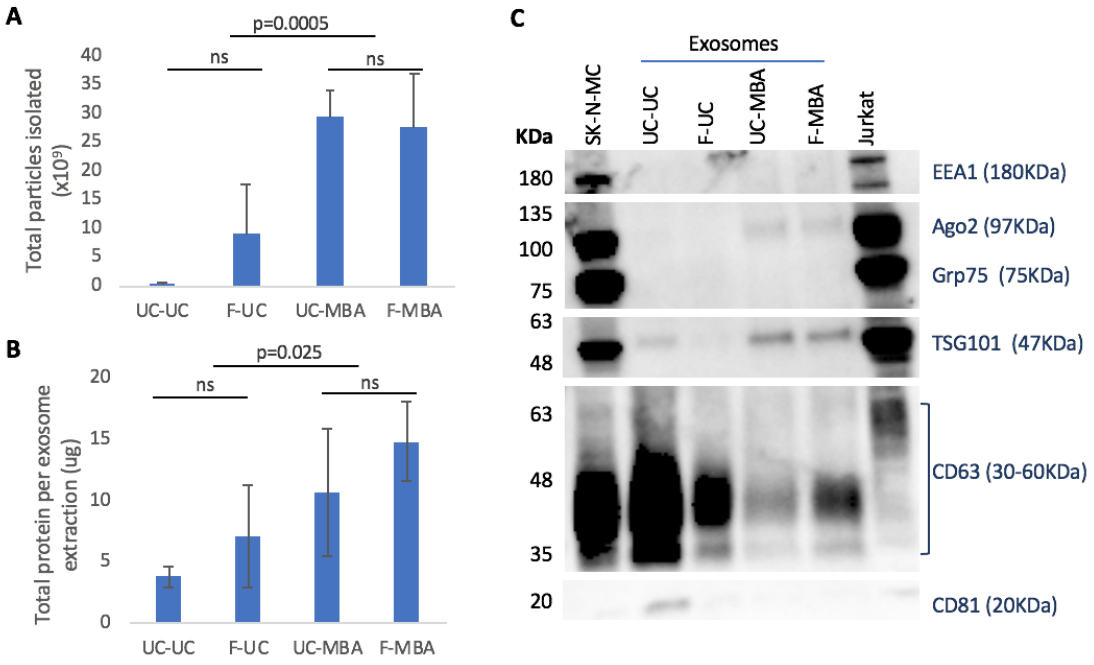


Figure 4.13: Comparison of MBA and dUC exosome isolation methods. Exosomes from the same cell culture media were pre-processed using filtration (F) or ultracentrifugation (UC), and exosomes from this media were isolated using either the MBA protocol (MBA) or ultracentrifugation (UC). **A.** Comparison of exosome yield using NTA. **B.** Comparison of exosome yield by protein concentration. Data from A and B expressed as mean ± SD from 3 independent repeats (except NTA UC-UC 2 repeats). **C.** Expression of exosomal markers (TSG-101, CD63 and CD81) and cellular contaminants (EEA1, Ago2 and Grp75) across the 4 different protocols. Cells (10 μg) and exosomes (10 μl protein lysate, equivalent of 1/2 tissue culture flask) were loaded. SK-N-MC and Jurkat cells were included as positive control. WB representative of 3 independent repeats.

Similarly than the NTA, data from the exosome protein concentration showed a higher protein amount when using the MBA method (13±4 pg, range 6–18 pg, $n = 6$) compared to the UC method (5±3 pg, range 3–11 pg, $n = 6$) ($p = 0.025$; Figure 4.13B). Again, no differences were seen between the different pre-processing protocols for either of the isolation methods (Figure 4.13B).

In order to assess that the yield obtained was actually from exosomes, and not other vesicles or contaminants of similar size, the expression of different markers

was evaluated. Contrary to the results from NTA and protein concentration, WB showed higher expression of the exosomal marker CD63 when using UC isolation in contrast to MBA, that had lower yield (Figure 4.13C). However, the exosomal marker TSG-101 was more highly expressed on the exosomes obtained by MBA than the UC. In addition, expression of CD81 was only visible in 3/6 UC elute, and none of the MBA eluates. Regarding cellular contaminants, all protocols showed absence of these, indicating that the eluates were clean from other vesicle types.

Therefore, NTA, protein concentration and expression of TSG-101 all demonstrate that MBA protocol leads to higher exosome yields compared to the UC, with these being free from contaminating vesicles. However, expression profiles of CD63 and CD81 contradict the above, as these were higher on UC isolated exosomes than MBA, even when the number of exosomes were the lowest. These data would indicate that dUC enriches for a small fraction of exosomes that are highly enriched for tetraspanins (CD63 and CD81), but other exosomes (TSG-101) are not captured, thus resulting in lower particles numbers and protein concentration.

To further clarify the above data, this was confirmed by flow cytometry, as this allowed quantification of particles (CT labelled) and expression of the different markers (CD63 and CD81). As shown in Table 4.6, a lower yield of exosomes was seen when isolating by UC independently to the pre-processing method used, compared to MBA, that showed higher exosome yields both when using CT (all exosomes labelled; $p = 0.001$) or specific exosomal markers (CD63: $p < 0.0001$, CD81: $p < 0.0001$). Again, comparison across the different pre-processing protocols did not show differences in the yield obtained. Therefore, flow cytometry confirms the results seen before, where MBA provides higher exosome yields with increased expression in TSG-101, whereas dUC leads to lower yields (both in particle number and protein concentration) but an enrichment on tetraspanins. Due to the low particle number, this enrichment cannot be appreciated by flow cytometry. Combining all these evidence, the MBA protocol provided the optimal results for downstream evaluation of exosomes.

Isolation protocol	CT+ events	CD63+ events	CD81+ events
UC-UC	8 ± 9	3 ± 0.5	2 ± 2
F-UC	17 ± 120	4 ± 0.4	16 ± 4
UC-MBA	727 ± 752	6608 ± 349	1815 ± 90
F-MBA	727 ± 618	5507 ± 649	1644 ± 142
UC vs MBA	$p = 0.001$	$p < 0.0001$	$p < 0.0001$

Table 4.6: Evaluation of exosome yield by flow cytometry using dUC and MBA methods to isolate exosomes. The number of events labelled by CT (generic dye), CD63 or CD81 (exosome markers) was evaluated in all 4 protocols. Differences between UC and MBA isolation were seen (bottom table) for all markers. Data as mean ± SD from triplicates.

4.3.6 Profiling of ES cell line-derived exosomes

Exosomes from the SK-N-MC cell line were evaluated using the optimised MBA protocol (Section 4.2.5), before starting the characterisation of exosomes from patient-derived ES cell cultures.

SK-N-MC exosomes isolated had a diameter of around 100 nm by TEM (Figure 4.14A), consistent with the data from NTA (median size ± SD: 89 ± 3, $n = 4$; Figure 4.14B). WB revealed expression of both TSG-101 and CD63 exosomal markers in SK-N-MC derived exosomes, but interestingly, expression of the exosomal marker CD9 was not detected (Figure 4.14C). The isolated exosomes did not express the cellular markers (EEA1, Ago2, Pan-cadherin or Grp75), indicating that these were free from contaminants. Moreover, the expression of CD99, highly expressed in ES cells, was investigated, to see if ES-derived exosomes express this marker. As shown in Figure 4.14C, this was not detected in the eluted exosomes.

In addition, the RNA profile of exosomes and the SK-N-MC cells from which they were derived revealed an enrichment of miRNAs in exosomes (Figure 4.14D), both when evaluating the percentage of miRNA from the total sRNA (exosomes 60 ± 6%, cells 11 ± 5%, $p < 0.0001$, n=7) and the miRNA concentration (exosomes 5052±2,985 pg/μl, cells 113±115 pg/μl, $p = 0.004$, n=7).

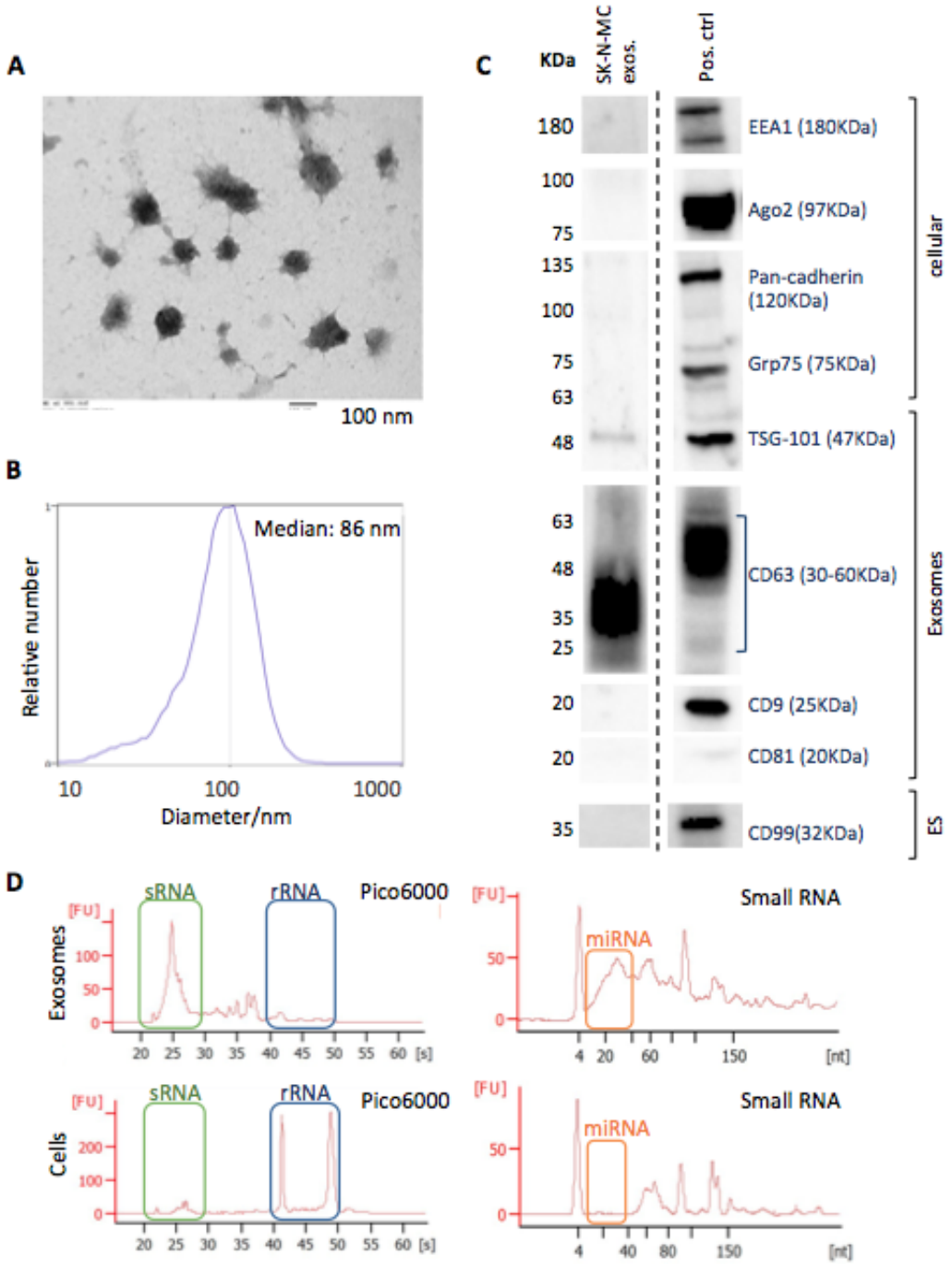


Figure 4.14: Profile of SK-N-MC-derived exosomes. **A.** TEM of SK-N-MC exosomes. Scale bar = 100 nm. **B:** Size distribution of SK-N-MC exosomes using NTA. Graph is representative of 4 independent repeats. **C.** Expression of exosomal markers (TSG-101, CD63, CD9 and CD81) and cellular contaminants (EEA1, Pan-cadherin, Grp75 and Ago2) in exosomes from SK-N-MC cells compared to positive control (Jurkat cells). Expression of the ES marker CD99 was also evaluated (SK-N-MC cells as positive control). Cells (20 μg of protein) and exosomes (from 10×10^6 cells) were loaded. **D.** Electrophoretic characterisation of RNA by Pico6000 Agilent (left panel) and Small RNA (right panel) on exosomes (top panel) and cells (bottom panel). sRNA (green), rRNA (blue) and miRNA (orange) RNAs are highlighted. RNA data representative of 7 independent repeats.

Evaluation by flow cytometry confirmed that SK-N-MC derived exosomes express the exosomal markers CD81 and CD63, as well as the ES marker CD99 (Figure 4.15A). This is in agreement with CD63 expression on WB (Figure 4.14C), at the same time that complements WB by showing expression of CD81 and CD99, not detected on the protein lysates (Figure 4.14C). When evaluating the percentage of exosomes (CT+) that expressed each marker (Figure 4.15B), CD63 expression was identified in a high proportion of exosomes (97 ± 0.2%), confirming the enrichment of this exosomal marker in the isolated exosomes, as seen by WB (Figure 4.14C). Moreover, the ES characteristic marker CD99 was detected in 72 ± 0.4% of isolated exosomes (CT+) and in 64 ± 1% when only evaluating the CD63+ exosome population (Figure 4.15B).

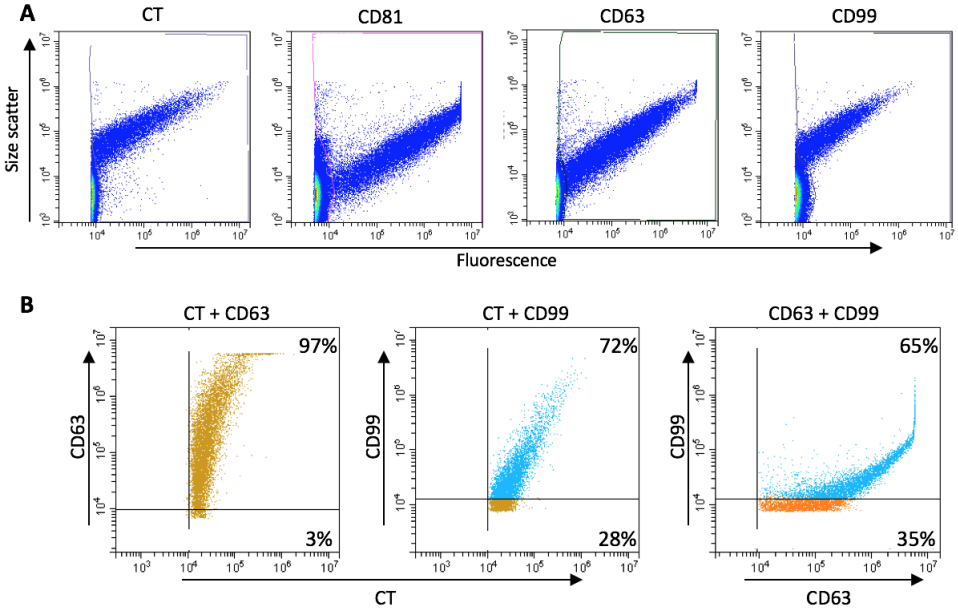


Figure 4.15: Characterisation of SK-N-MC exosomes by flow cytometry. **A.** SK-N-MC exosomes single-labelled (left to right) with CT, CD81 and CD63 exosomal markers and CD99 ES marker to evaluate the expression of these markers in these exosomes. **B.** SK-N-MC exosomes dually labelled with (left to right) CT + CD63, CT + CD99 and CD63 + CD99. Presence of exosome population on the top-right quadrant indicative of dually-labelled exosomes. Only events with size 100–900 nm and expressing the primary marker (CT or CD63) are gated. Plots representative of 3 independent repeats.

Therefore, the characteristic ES marker CD99 can be detected in a proportion of ES-derived exosomes, confirming their cellular origin, whereas most exosomes express the CD63+ exosomal marker, confirming WB data. This indicates that there is an heterogeneity in the exosomes released by SK-N-MC cells, as not all

exosomes (CD63+) express the ES marker CD99.

4.3.7 Profiling of exosomes from ES patient-derived cell cultures

Exosomes from ES patient-derived cell cultures (Chapter 2) were investigated using the MBA method (Section 4.2.5). Exosomes were isolated from samples of 12 different patients, resulting in a total of 19 different samples (11 tissue macerate, 7 bone digested, from which 6 were paired; and 1 ES-CSC derived culture; Table 4.7). Exosomes were isolated from these samples at passages lower than 15 in all cases. A summary of the samples investigated and the results obtained is shown in Table 4.7.

4.3.7.1 Patient-derived ES exosomes have the expected size

Exosomes from a selection of ES patient-derived cell cultures ($n = 4$) and ES-CSC culture ($n = 1$) were characterised by NTA (Figure 4.16A, Table 4.7). The median diameter size was 97 ± 12 nm (84–111 nm), which agreed with the previous data obtained by SK-N-MC exosomes (89 ± 3 nm, $n = 4$, Section 4.3.6). In addition, NTA also provides data on the total yield of exosomes isolated, which was $2 \times 10^5 \pm 3 \times 10^5$ exosomes per cell (range 1×10^4–7×10^5 exosomes per cell).

4.3.7.2 Patient-derived exosomes express exosomal markers and CD99

Exosomes from 15 ES patient-derived cell cultures were evaluated for their expression of exosome markers TSG-101, CD63 and CD81, the presence of cellular contaminants EEA1, Ago2 and Grp75, and the expression of CD99 ES marker by WB (Table 4.7; full blots in Figure A1, Figure A2, Figure A3). Both TSG-101 and CD63 exosomal markers were expressed in the majority of cases (13/13 for TSG-101, 14/15 for CD63), with a broad range of expression levels (Figure 4.16B). The exosomal marker CD81 was only detected in 3/15 samples (low expression levels), agreeing with SK-N-MC data where CD81 was not detected. Regarding the expression of cellular markers, Grp75 was detected in only 1/12, but both EEA1 and Ago2 were detected in 7/15 and 7/12 samples respectively but at low expression levels. The expression of the ES marker CD99 was evaluated in exosomes from patient-derived samples, and as seen for SK-N-MC, CD99 was not detected in any of the exosomal lysates (0/15), compared to the cells from which they were derived, that showed expression of CD99 (Figure 4.16B)

ID	NTA		WB							Flow cytometry			Agilent				Other		
	Size (nm)	Exos./cell	CD63	CD81	TSG-101	Ago2	Grp75	EEA1	CD99	% CD63+	% CD81+	% CD99+	Exos. (%)	Cells (%)	Exos. (ng/ul)	Cells (ng/ul)	Plasma	Fusion	RNA-seq
SK-N-MC	89	1×10^3	1	0	1	0	0	0	0	97%	91%	73%	49	2	2840	9	–	✓	✓
Geno1T	92	1×10^5	1	1	1	–	–	1	0	92%	99%	90%	77	14	6067	220	✓	–	✓
Geno2T	–	–	1	0	–	–	–	1	0	–	–	–	80	20	2151	103	–	–	–
Geno2B	–	–	1	0	1	1	0	1	0	–	–	–	69	7	1218	28	–	–	–
Geno5T	–	–	0	0	1	0	0	0	0	–	–	–	–	–	–	–	–	–	–
Geno5B	89	1×10^4	1	0	–	–	–	1	0	82%	99%	89%	69	12	3096	48	–	✗	–
Geno7T	–	–	1	1	1	1	0	1	0	–	–	–	74	28	8573	396	✓	–	✓
Geno10T	–	–	–	–	–	–	–	–	–	–	–	–	77	–	1487	–	–	–	–
Geno10B	–	–	–	–	–	–	–	–	–	–	–	–	72	8	1182	49	–	–	–
Geno14T	84	5×10^4	1	0	1	1	0	1	0	94%	99%	93%	64	9	4201	128	✓	✗	✓
Geno18T	–	–	1	0	1	1	0	0	0	–	–	–	64	–	894	–	–	–	–
Geno18B	–	–	1	0	1	1	0	0	0	–	–	–	68	9	672	43	–	–	–
Geno23T	–	–	1	0	1	0	1	0	0	–	–	–	66	13	854	53	–	–	–
Geno25B	–	–	1	0	1	1	0	0	0	–	–	–	66	11	1310	73	–	–	–
Geno25B	–	–	–	–	–	–	–	–	–	–	–	–	59	26	621	283	–	–	–
Geno34B	–	–	1	0	1	0	0	0	0	–	–	–	63	6	989	41	–	–	–
Geno35T	107	2×10^5	1	0	1	0	0	0	0	83%	99%	81%	69	5	1863	16	–	–	–
Geno39T	–	–	1	0	1	0	0	0	0	–	–	–	69	4	2491	79	–	–	–
Geno39B	–	–	1	1	1	1	0	1	0	–	–	–	75	12	3252	19	–	–	–
Geno14T.s3	111	7×10^5	–	–	–	–	–	–	–	93%	99%	92%	66	24	445	110	–	–	–
Total: 19	5	5	15	15	13	12	12	15	15	5	5	5	18	16	18	16	3	2	3

Table 4.7: Summary table of exosome characterisation from patient-derived ES samples. NTA (Section 4.3.7.1) as size and concentration normalised to cell number. WB (Section 4.3.7.2) as expression of the different exosomal (CD63, CD81, TSG-101), cellular (Ago2, Grp75, EEA1) and ES (CD99) markers as 0 = not expressed, 1 = expressed. Flow cytometry (Section 4.3.7.2) evaluation as percentage of exosomes (CT) expressing each marker (CD63, CD81 and CD99). Agilent RNA characterisation of exosomes and cells (Section 4.3.7.3) as % miRNA from the total sRNA; and miRNA concentration (pg/μl). Other characterisation indicating if plasma was evaluated by flow cytometry, fusion was detected by RT-qPCR and RNA-seq was performed on exosomes (✓ = yes, ✗ = no). – = data not available. Bottom row (Total) indicative of the number of samples (patient-derived) evaluated by each assay/marker. Top row displaying information on SK-N-MC cell line for easier comparison.

Moreover, exosomes isolated from 4 patient-derived cell cultures (3 replicates each) and 1 ES-CSC (1 replicate) were further evaluated for their expression of exosomal markers CD63 and CD81, and the expression of the ES marker CD99 by flow cytometry (Figure 4.16C, Table 4.7). Similarly to what was seen for SK-N-MC exosomes (Section 4.3.6), exosomes derived from patient samples and ES-CSC express exosomal markers CD81 and CD63, as well as ES marker CD99. When the percentage of exosomes (CT+) that expressed each one of the markers was evaluated, a similar profile as for SK-N-MC was seen across all the patient-derived exosome samples (Figure 4.16C). A high proportion of exosomes expressed CD63 (88 ± 6%, n= 13) and CD81 (99 ± 0.1%, n= 13), further confirming that the isolated vesicles were indeed exosomes. Interestingly, when evaluating the expression of the ES marker CD99 on the patient-derived exosomes, a higher percentage was seen than for SK-N-MC (98 ± 1%, n= 13). Therefore, patient-derived cell cultures produce exosomes with different profiles, as the amount of exosomes expressing CD99 is not as high as the exosomes expressing CD63 or CD81 markers. In addition, it is important to mention that different to what was seen for SK-N-MC exosomes (Section 4.3.6), where an homogeneous population was seen on the scatter plot, exosomes derived from patient cell cultures had an heterogeneous profile, with subpopulation of different sizes and levels of expression (indicated by arrows; in Figure 4.16C).

4.3.7.3 Patient-derived exosomes have an enrichment for sRNA

The RNA profile of exosomes from 18 cell cultures and producing cells (17 patient-derived ES cultures, 1 ES-CSC culture; Table 4.7) was compared to establish whether the exosome cargo reflects the RNA profile of the producing cells, or whether there is a specific sorting of cargo into these vesicles.

When evaluating the total RNA profile, higher levels of rRNA were found in cells compared to the exosomes they generated, which had low rRNA abundance (Figure 4.17A). The isolated exosomes showed a significant enrichment of miRNAs from the total population of sRNAs (69 ± 6%; range 59–80 %, $n = 18$) compared to cells, which had low levels of miRNAs (13 ± 8%; range 4–28 %, $n = 16$) ($p < 0.0001$, Table 4.7), resulting in a fold-change increase of 7 ± 4. The concentration of miRNAs from both sample types was again statistically significant,

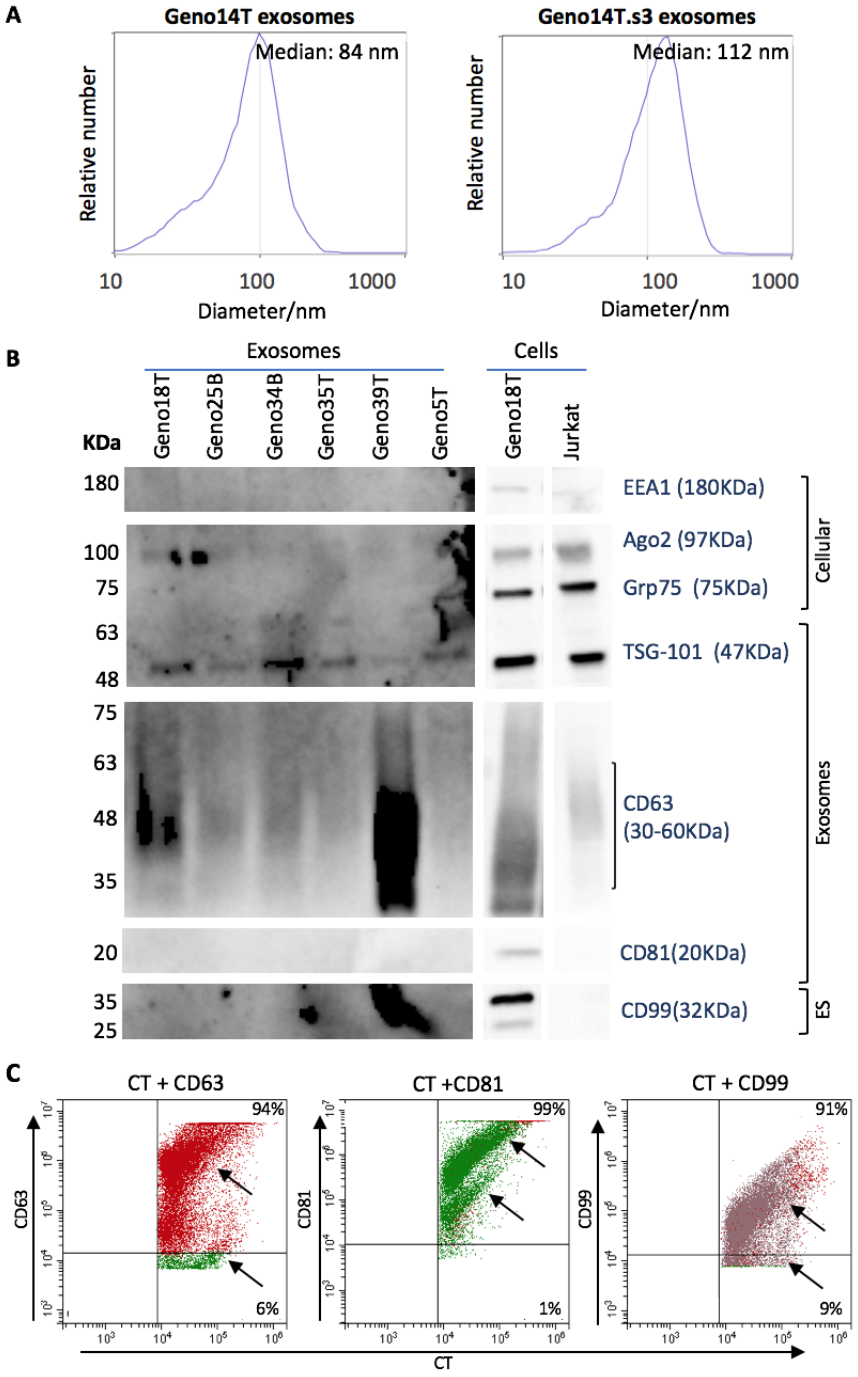


Figure 4.16: Characterisation of ES patient-derived exosomes. **A.** Size distribution of Geno14T (left) and Geno14T.s3 (right) exosomes using NTA. **B.** Evaluation of expression of exosomal (TSG-101, CD63, CD81), cellular (EEA1, Ago2, Grp75) and ES marker (CD99) on exosomes (Geno18T, Geno25B, Geno34B, Geno35T, Geno39T, Geno5T) compared to cells (Geno18T and positive control Jurkat). Exosomes (protein from 5×10^5 cells) and cells ($20\,\mu g$ protein) were loaded. **C.** Expression profile of exosomes from Geno14T cells by flow cytometry, as percentage of exosomes (CT+) expressing CD81, CD63 or CD99. Arrows indicating different exosome populations based on size and expression. Only events with size 100–900 nm and labelled by CT are gated. Plots representative of 3 independent repeats.

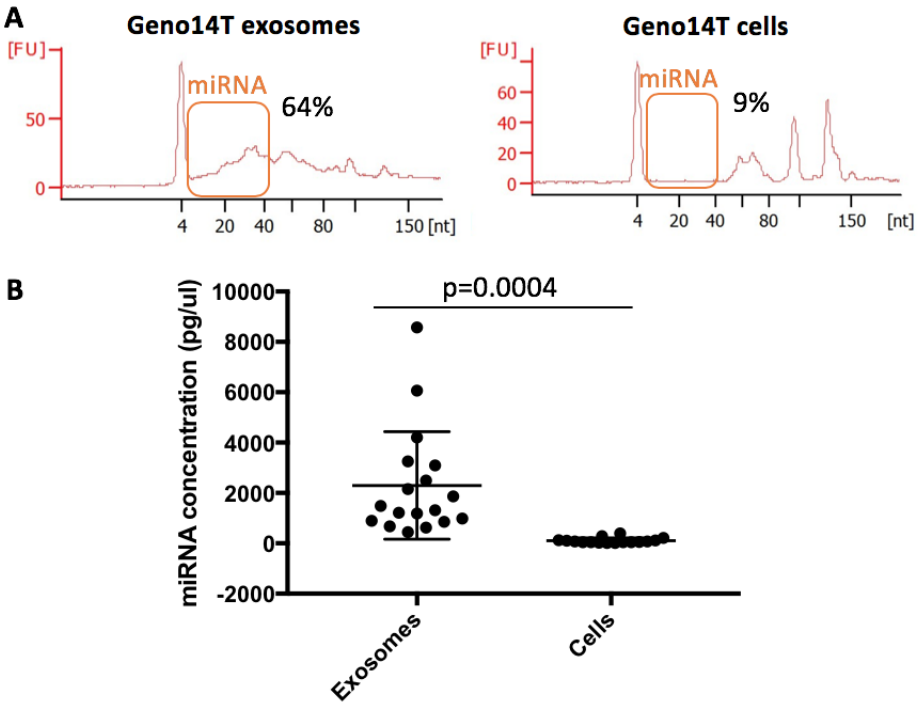


Figure 4.17: Comparison of the RNA levels between patient-derived cells and exosomes. **A.** Electrophoretic characterisation of RNA comparing Geno14T exosomes (left) and cells (right), with the miRNA highlighted in orange. **B.** Comparison of the miRNA concentration between exosomes and cells they are originated from of 18 different patient-derived cell cultures.

with higher concentration on the exosomes eluates (2298±2,138 pg/μl, range 445±8,573 pg/μl, n = 18) than in cells (106±107 pg/μl, range 16±396 pg/μl, = 16) (p = 0.0004; Figure 4.17, Table 4.7). The fold-change increase in this case between paired exosomes and cells was 39 ± 44. Therefore, agreeing with SK-N-MC data (Section 4.3.6), exosomes have an enrichment of miRNAs compared to the cells they are derived from, suggesting that there is a specific sorting of the cargo into these vesicles.

4.3.8 Exosomes from plasma of ES patients express CD99

Since human-derived exosomes may be a powerful source of circulating biomarkers, I wanted to investigate whether ES-derived exosomes could be isolated from plasma samples from ES patients. To optimise the method of isolating exosomes from plasma, I used plasma from healthy donors as this was relatively easy to obtain compared to the small amounts of plasma from patients. With this, I

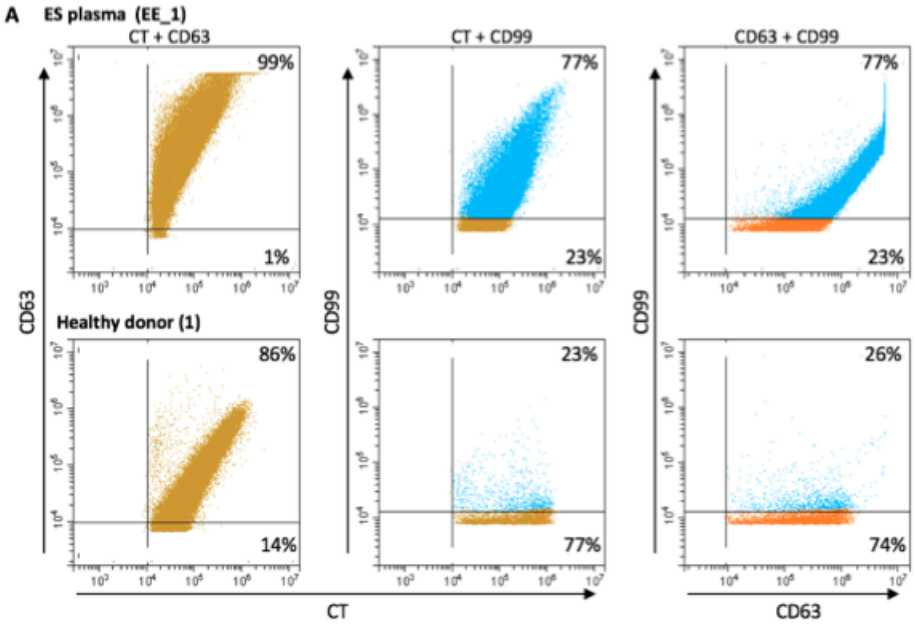
investigated the optimal starting volume of plasma for which exosome could be evaluated by flow cytometry, in order to minimise the plasma loss. As shown in Table 4.8, exosomes could be successfully isolated and analysed from as little as 5 μl of starting volume of plasma, which after dividing into the different tubes for analysis, was equivalent of 0.04 μl of plasma analysed (9 ± 2 CT+ events, $n = 12$). As expected, the exosome yield increased with increasing the starting volume of plasma (Table 4.8), as can be seen in the "CT+ events" column. To evaluate the loss of exosomes across the different volumes, the number of exosomes identified (CT events) was related to the μl of starting volume. If the starting volume did not have an impact on the yield, a similar number of exosomes per μl would be identified. However, as can be seen in the last column of Table 4.8, the smaller the plasma volume used for the extraction, the less exosomes per μl were detected. Taking this data into account, and having in mind the small amounts of plasma from patients, it was decided to use 100 μl of plasma as starting volume for the characterisation of exosomes.

Extracted plasma (ul)	Plasma (ul) per tube	Plasma analysed (ul)	CT+ events	CT+ events per ul plasma
5	0.4	0.04	9±2	219
50	3.6	0.43	108±20	253
100	7.1	0.86	269±51	314
200	14.3	1.71	569±123	332
500	35.7	4.29	1470±312	343
2000	142.9	17.14	6849±2106	400

Table 4.8: Plasma dilution curve for optimal exosomes isolation yield by flow cytometry. Decreasing volumes of plasma (2,000–5 μl) were used for exosomes isolation and the yield was evaluated by flow cytometry to identify the minimal yield of plasma from which exosome characterisation could be performed.

Exosomes from plasma of 4 ES patients (3 of which had matching patient-derived cell cultures; Table 4.7) and 3 healthy donors were evaluated by flow cytometry (Figure 4.18). As seen for the patient-derived ES samples, exosomes from all plasma samples expressed the CD63 exosome marker (93 ± 2%, range 75–100 %, n= 7), confirming that CD63 is an exosomal marker rather than linked to ES exosomes. Similar to exosomes from SK-N-MC (Section 4.3.6) and patient-derived cells (Section 4.3.7.2), exosomes (CT+) from all ES plasma samples showed expression of CD99 ES marker (71 ± 3%, 58–85 %, n= 4). A similar profile was seen

when evaluating CD99 expression on CD63+ exosomes only (49 ± 3%, 39–69 %, n= 4; Figure 4.18 top row of plots).



B **Percentage of events expressing both markers (% ± SD; 3 replicates per sample)**

Sample ID	CD63+ from CT+	CD99+ from CT+	CD99+ from CD63+
EE_1* plasma	100 ± 0 %	58 ± 0.2%	45 ± 1 %
Geno1T plasma	100 ± 0 %	85 ± 1 %	39 ± 0.1 %
Geno7T plasma	100 ± 0 %	79 ± 1 %	68 ± 1 %
Geno14T plasma	100 ± 0 %	61 ± 1 %	43 ± 1 %
Healthy1 plasma	85 ± 0.3 %	19 ± 2 %	21 ± 0.3 %
Healthy2 plasma	75 ± 0.3 %	11 ± 1 %	15 ± 1 %
Healthy3 plasma	90 ± 1 %	8 ± 1 %	12 ± 1 %

Figure 4.18: Profiling of exosomes from ES and healthy donor plasma by flow cytometry. **A.** Dual labelling of exosomes using (left to right) CT + CD63, CT + CD99 and CD63 + CD99 on a ES plasma sample (top) and a healthy donor plasma (bottom). Exosome population on the top-right quadrant indicative of dually-labelled exosomes. Only events with size 100–900 nm and expressing the primary marker (CT or CD63) are gated. Plots representative of 3 independent repeats. **B.** Summary table of exosomes (%) from ES and healthy donor plasma samples expressing both markers ± SD from 3 repeats. Plasma sample EE_1 does not have paired cell culture.

In contrast, exosomes from healthy donors did not express the CD99 ES marker, as can be seen in Figure 4.18A (bottom row of plots). The successful isolation of exosomes was confirmed, as CD63 expression is high amongst the CT exosomes (84 ± 2%, range 75–91 %, n= 3). However, the defined CD99+ exosomes population observed for ES plasma samples is lost here, and all CD99 events are due to the

scatter of CT or CD63-positive exosomes (12 ± 2% and 16 ± 1% respectively, n= 3; Figure 4.18A). Combining both datasets together (ES plasma and healthy donors), statistical differences were seen in the CD99 expression between ES patients and healthy donors when gating for CT+ exosomes ($p < 0.0001$) and CD63+ exosomes ($p < 0.0001$).

Therefore, patient-derived exosomes in plasma express CD99, agreeing with data on exosomes isolated from cell cultures originated from these same patients (Section 4.3.7.2). In contrast, CD99 exosomes are not detected in healthy donor plasma, suggesting that exosomes in blood circulation could be exploited for circulating biomarker strategies in ES.

4.3.9 Characterisation of the RNA cargo from ES-derived exosomes

As described in Chapter 1, exosome cargo has a role in tumour progression and may be exploited to identify circulating biomarkers for diagnosis and disease progression. Therefore, exosomes from a selected panel of ES samples (based on availability of clinical data and matched plasma samples) were further characterise by the analysis of the RNA cargo by RT^A-qPCR and RNA-seq.

4.3.9.1 ES exosomes contain the EWSR1 fusion and ES-CSC marker NRXN1

The main driver of the ES phenotype is the EWSR1-ETS fusion (Chapter 1). I therefore investigated if this was carried in exosomes derived from ES. For this, exosomes from the ES cell lines SK-N-MC, TC-32 (both EWSR1-FLI1 type 1) and TTC-466 (EWSR1-ERG), as well as the ES patient-derived cell cultures Geno5B and Geno14T (both EWSR1-FLI1 type 1; Table 4.7) were evaluated for the presence of the characteristic fusion by RT^A-qPCR. All samples ($n = 5$) were analysed by RT^A-qPCR for PPIA to confirm the quality of RNA extracted from exosomes; RNA-seq confirmed PPIA was expressed in exosomes.

Exosomes from all 3 ES cell lines contain the EWSR1-ETS fusion gene as part of the exosome cargo (Table 4.9). Unexpectedly, the EWSR1-FLI1 fusion gene was not detected in the exosomes from the ES patient-derived cell cultures (Table 4.9). Nevertheless, confirming the quality of the exosomal RNA, PPIA was detected in all

ID	Target	C_t target	C_t PPIA	ΔC_t
SK-N-MC	EWSR1-FLI1	35.2	23.2	12.0
TC-32	EWSR1-FLI1	34.4	22.9	11.5
TTC-466	EWSR1-ERG	34.6	24.7	9.9
Geno5B	EWSR1-FLI1	–	22.9	–
Geno14T	EWSR1-FLI1	–	25.0	–
SK-N-MC	NRXN1	35.0	19.0	16.0

Table 4.9: Detection of EWSR1-ETS fusion and ES-CSC marker NRXN1 in exosome cargo by RT^AqPCR. Data as C_t (cycle threshold) or ΔC_t (difference between target and housekeeping gene cycle threshold). The EWSR1-FLI1 fusion could not be detected in exosomes from patient-derived cells (– = not detected).

cases at similar levels (C_t 23–25), even for the patient-derived exosomes in which the fusion was not identified (Table 4.9). This difference in the detection of the fusion between cell lines and patient-derived cell cultures agrees with previous RT^A-PCR data (Section 2.3.2.2), where using the same amount of RNA, the fusion gene was easily detected in the cell lines in contrast to the yield observed for patient-derived cultures.

In addition to assessing the presence of the characteristic EWSR1-ETS fusion gene in the cargo of ES-derived exosomes, the NRXN1 ES-CSC marker identified in Section 3.3.5 was assessed in SK-N-MC exosomes. As seen in Table 4.9, this ES-CSC candidate marker is carried as part of the exosome cargo, although at lower levels than those seen for EWSR1-FLI1 on the same sample (ΔC_t 12 for EWSR1-FLI1; ΔC_t 16 for NRXN1).

Altogether, these data indicates that ES-derived exosomes have the characteristic EWSR1-ETS fusion and the ES-CSC NRXN1 marker as part of their cargo. This further confirms the ability to use ES-derived exosomes as circulating biomarkers for ES diagnosis. At the same time, the presence of the EWSR1-ETS fusion gene and the ES-CSC marker NRXN1 could have implications in cancer progression if ES exosomes can transfer these RNAs to recipient cells.

4.3.9.2 A different RNA expression profile was detected between exosomes and producing cells using RNA-seq

The RNA cargo from exosomes of SK-N-MC and 3 patient-derived ES cell cultures (Geno1T, Geno7T and Geno14T; Table 4.7) was investigated by Total RNA-seq. From the total number of reads ($9 \times 10^7 \pm 7 \times 10^7$), 97 % of reads were conserved after processing ($9 \times 10^7 \pm 7 \times 10^7$), similar to the rates obtained for cells (Section 3.3.2.1), indicating that the RNA-seq from exosomes generated good quality data. From these post-processed reads, the annotation success (number of reads that can be annotated to the reference genome) was 31 % (19–57 %; 1×10^7–1×10^8, Table 4.10). Three of the the exosome samples (SK-N-MC, Geno1T and Geno7T) had an annotation rate of 22 %±3 %, whereas the exosomes from Geno14T had a higher annotation rate (57 %), much similar to the one obtained for cells (60% for matched cells; Section 3.3.2.1). In addition, exosomes from Geno14T had 4 fold greater number of raw reads compared to the other exosome samples, which resulted in the higher annotation rate (Table 4.10). This data suggests that the cargo of Geno14T exosomes has a profile different from the rest of exosomes, with higher RNA content and a variety of RNA species that can be more easily aligned.

ID	Total reads	Total RNAs	Read coverage per RNA	CD99 reads	EWSR1 fusion
SK-N-MC	1.01×10^7	23150	437.8	2522	—
Geno1T	1.33×10^7	20433	651.7	4571	—
Geno7T	1.36×10^7	20372	662.6	4431	—
Geno14T	1.13×10^8	29598	3823.5	47886	—
Total	1.50×10^8	34096			

Table 4.10: Read coverage for RNA-seq data on exosomes. The number of aligned reads (total reads) and the number of unique RNA (total RNAs) is shown for each sample, as well as the mean coverage per RNA. Reads aligning to both CD99 and EWSR1-ETS fusion are shown. — indicative of no detection. Bottom line (total) shows the total number of reads and the unique RNAs identified across the 4 samples.

Regarding the variability of RNAs expressed, a total of 34,096 unique Ensembl gene (ENSG) ID were identified across the 4 samples, which related to 23,388±4,337 unique genes per sample (Table 4.10). The average read coverage per gene was 584±127, with the exception again of Geno14T, which was 6× higher (3,823 reads per gene; Table 4.10).

The expression of CD99 was evaluated using the gene specific reference strategy

described in Section 3.3.2.3 for the characterisation of cells by RNA-seq. Corroborating with the cell from which exosomes were derived, all samples showed expression of CD99 (2,522–47,886 reads, again Geno14T being 12× higher in coverage; Table 4.10). This data agrees with the expression of CD99 on isolated exosomes identified by flow cytometry (Section 4.3.6, Section 4.3.7.2). However, the EWSR1-ETS fusions were not identified in exosomes by RNA-seq, even for the samples which the fusion was detected by RT^A-qPCR (Section 4.3.9.1). These results agree with data from cells (Chapter 3), where RNA-seq showed less sensitivity in detecting the EWSR1-ETS fusion than RT^A-PCR (Section 3.3.2.4).

The exosome RNA profile of the 3 patient-derived ES samples was compared to see how much of the gene expression profile was shared or private. As shown in Figure 4.19A, 53 % of the identified RNAs are shared amongst the 3 exosome samples whereas 28 % are uniquely expressed; consistent with earlier results, exosomes from Geno14T had ten-fold greater expression of private RNAs. As the RNA-seq profile from the 4 matching cells was obtained previously (Chapter 3), the differences in RNA profile were investigated (Figure 4.19B). Exosomes and cells shared 78 % of unique RNAs, whereas 3 % and 18 % of RNAs were only identified in exosomes or cells respectively.

To further investigate the differences in RNA profile between the exosomes and the cells, multidimensional scaling was used to see the relative distance between the samples (Figure 4.19C). The 3 ES patient-derived samples (Geno1T, Geno7T and Geno14T; orange) cluster together and separate from the SK-N-MC cell line, as seen before (Section 3.3.3). As predicted, the exosomes derived from these samples (blue) cluster separate from the cells they are derived from. Exosomes from Geno14T clustered more closely to the Geno14T cells than the other patient-derived exosomes, further suggesting that exosomes from this sample have a different profile.

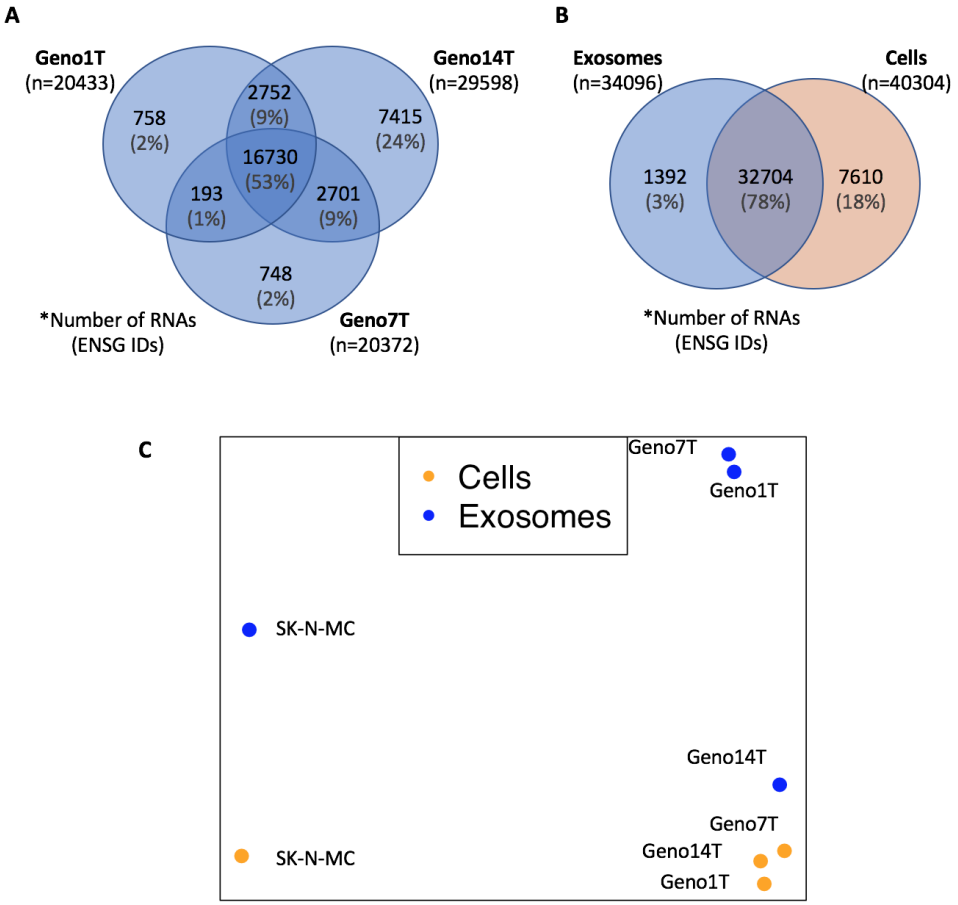


Figure 4.19: Distinctive RNA profile of exosomes and producing cells by RNA-seq. **A.** Comparison of RNA profile of exosomes from 3 ES patient-derived cell cultures. **B.** Comparison between the RNA profile of exosomes ($n = 4$) and matching cells ($n = 4$). A and B shown as total number of unique ENSG IDs and percentage from the overall RNAs identified. **C.** Multidimensional scaling graph representing the distance between the different exosomes (blue) and matching cells (orange), indicative of the similarities between RNA profiles.

Finally, each exosome sample and matched cell was evaluated separately to identify the degree of overlap and how many RNAs were uniquely identified in each sample. A high degree of overlap between exosomes and cells from the same sample was detected (65±10 %; Table 4.11), agreeing with what was seen when comparing all samples together (Figure 4.19B). The number of RNAs only detected in cells (30±13 %) and exosomes (5±4 %) resembled the analysis performed in all samples together (18 % and 3 % respectively, Figure 4.19B). Therefore, RNA-seq data suggests that there is high overlap in the RNA profile between cells and

exosomes derived from these. It is important to mention that the few RNA species only detected in exosomes (5%) could be due to low yield of exosomes material sequenced, as we would expect all genetic material detected in the cargo of exosomes to be present in the cells they are derived from. In the same line, the greater amount of RNA species only detected in cells could be due to higher RNA sequenced from these. In order to fully validate which RNA species are not present in the cargo of exosomes compared to the producing cells, a method to normalise the RNA content sequenced should be used, although this is not yet described.

ID	Shared	Only in cells	Only in exosomes
SK-N-MC	67%	22%	10%
Geno1T	56%	43%	1%
Geno7T	58%	40%	2%
Geno14T	79%	16%	5%
Mean:	65 ± 10%	30 ± 13%	5 ± 4%

Table 4.11: Differences in the RNA profile between exosomes and matching cells investigated by RNA-seq. Data as number of unique RNAs from each sample pair (cell and exosomes) identified in both (shared), only in cells or only in exosomes.

Using DESeq2 (Section 3.2.3.6), 652 RNA species were found to be significantly differentially expressed between exosomes and cells; 392 RNAs were enriched in exosomes and 260 decreased in exosomes (Figure 4.20). Amongst the 260 RNAs decreased in exosomes (higher in cells), 94 % were protein coding RNAs (244/260), compared to only 3 % of sRNAs (7/260). The remaining 3 % included other RNA species such as pseudogenes or antisense RNAs. In comparison, the 392 RNAs enriched in exosomes contained 41 % of protein coding RNAs (162/392) and 30 % of sRNAs (119/392). This enrichment of sRNAs in exosomes was consistent with previous data (Section 4.3.7.3). The remaining 29 % included antisense RNAs (14 %), pseudogenes (7 %) and other species (8 %).

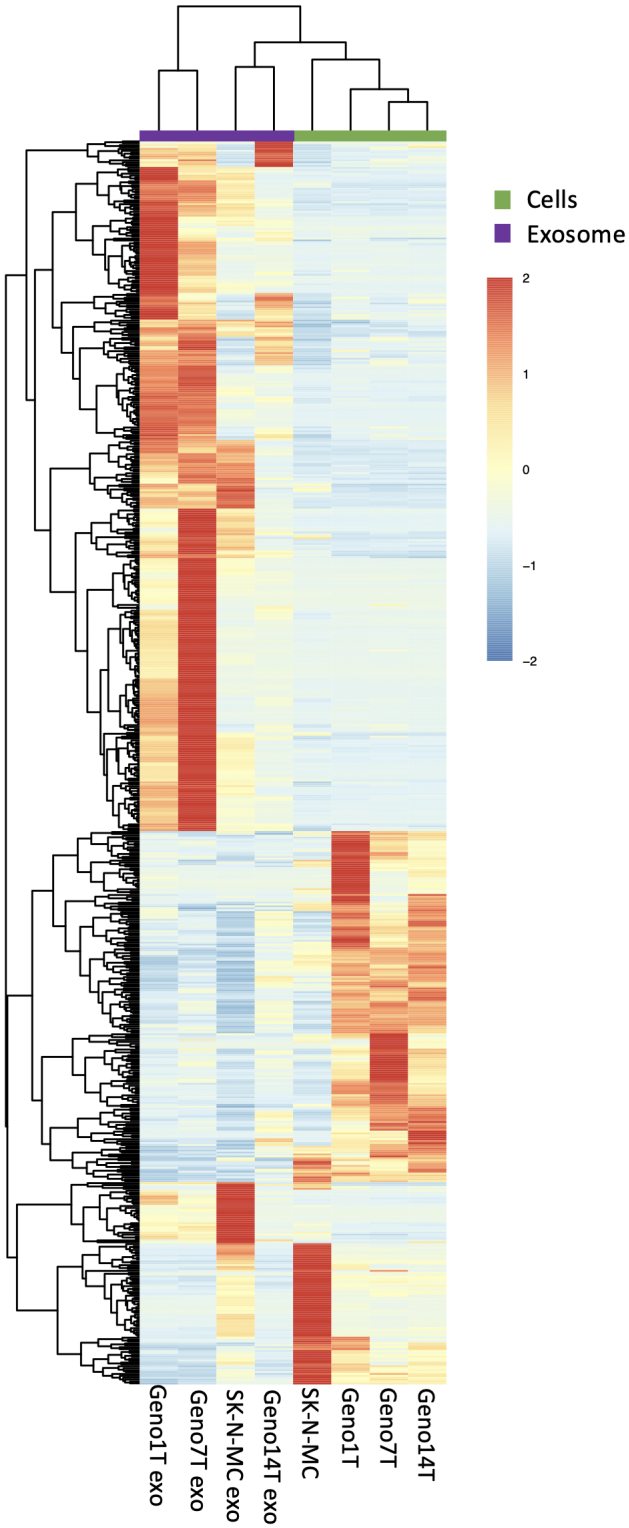


Figure 4.20: Differential expression profile between ES cell cultures and exosomes derived from these evaluated by DESeq2. Heatmap shows all differentially expressed genes (n = 652) with enriched expression (red) or decreased expression (blue).

Differentially expressed genes between both datasets were ranked based on LFC and adjusted p-value (Section 3.2.3.6), and a total of 18 RNAs was found on the top 50 for both ranks (Table 4.12). From these, only 2/18 were decreased in exosomes compared to producing cells and 16/18 were enriched in exosomes. Of these, 62.5 % (10/16) were non-coding RNAs and only 37.5 % (6/16) were protein-coding genes. This profiling confirms the characterisation of RNAs so far, as it indicates that indeed exosomes are enriched for non-coding RNAs compared to cells, that mainly express protein-coding genes. Moreover, these differences in the profile between cells and exosomes further confirm that there is an active sorting of the cargo into exosomes.

Gene	Increased in	p-adj	p-adj rank	LFC	LFC rank	Score (rank)	RNA type
MIR320A	Exos.	9.28×10^{-28}	2	8.68	4	1	non-coding
AC020911.2	Exos.	4.96×10^{-13}	6	9.61	1	2	non-coding
RN7SL278P	Exos.	1.48×10^{-11}	13	9.46	3	3	non-coding
AC087521.1	Exos.	7.51×10^{-13}	7	7.95	10	4	non-coding
ADH1B	Exos.	4.47×10^{-17}	3	7.54	16	5	protein coding
RNU1-55P	Exos.	7.80×10^{-10}	26	8.05	8	6	non-coding
FUT6	Exos.	3.92×10^{-10}	23	7.72	12	7	protein coding
U3	Exos.	1.10×10^{-11}	11	7.27	27	8	non-coding
APOF	Exos.	6.45×10^{-10}	25	7.37	19	9	protein coding
FAM166A	Exos.	7.01×10^{-16}	4	6.74	42	10	protein coding
HIST1H1T	Exos.	1.24×10^{-9}	32	7.28	26	11	protein coding
MIR3193	Exos.	1.70×10^{-9}	35	7.28	25	12	non-coding
MIR425	Exos.	3.54×10^{-9}	39	7.32	22	13	non-coding
AC005593.1	Exos.	1.56×10^{-9}	33	7.08	30	14	non-coding
SCARA5	Exos.	3.51×10^{-10}	20	6.71	43	15	protein coding
AC124784.1	Exos.	2.25×10^{-10}	16	6.52	48	16	non-coding
NCAM1	Cells	9.32×10^{-13}	8	-7.34	20	1	protein coding
FRAS1	Cells	3.45×10^{-9}	38	-6.84	37	2	protein coding

Table 4.12: Candidate RNAs differentially expressed between exosomes and cells identified by DESeq2. The RNAs present in both top 50 of p-adj and LFC were identified. Score as ranking position of p-adj value and LFC value.

4.3.10 Optimisation of the exosome uptake conditions in SK-N-MC cells

The profile of ES-derived exosomes has been evaluated throughout this study, assessing both the properties (Section 4.3.7) and the cargo (Section 4.3.9) of these vesicles. The next steps were to investigate the transfer of exosomes between different cell types and if this uptake led to changes in the cellular behaviour of recipient cells.

First, the conditions to investigate the uptake of exosomes had to be optimised. In order to identify if exosomes were taken up by recipient cells, I first needed to know how stable the labelled exosomes were with time. CT labelled exosomes from SK-N-MC, TC-32 and TTC-466 ES cell lines were monitored for fluorescence over time up to 24 h. When comparing to the number of CT events (exosomes) at 0 h, a decrease in the fluorescence was seen for the 3 cell lines over time (Figure 4.21). This was more significant in TC-32 exosomes, for which a decrease in the number of labelled exosomes was already visible at 1 h (31 % reduction; $p = 0.04$) and reached a decrease by half at 4 h (47 % reduction; $p = 0.004$; Figure 4.21). A similar response was seen in SK-N-MC exosomes, which had a reduction of 46 % after 1 h incubation ($p = 0.03$), and continued decreasing until reaching a 80% reduction at 24 h ($p = 0.001$). Finally, TTC-466 labelled exosomes showed a decrease in the number compared to 0 h, but this was not significant until reaching the 24 h of incubation (81 % reduction; $p = 0.02$; Figure 4.21). This rapid loss of fluorescence suggests that the exosome transfer experiments would need to be evaluated in a time window shorter than 24 h. However, these exosomes were labelled and analysed in a tube *in vitro*, whereas for the exosome transfer experiments, exosomes would be in contact with cells (more physiological state), which could lead to differences in the *in vitro* stability.

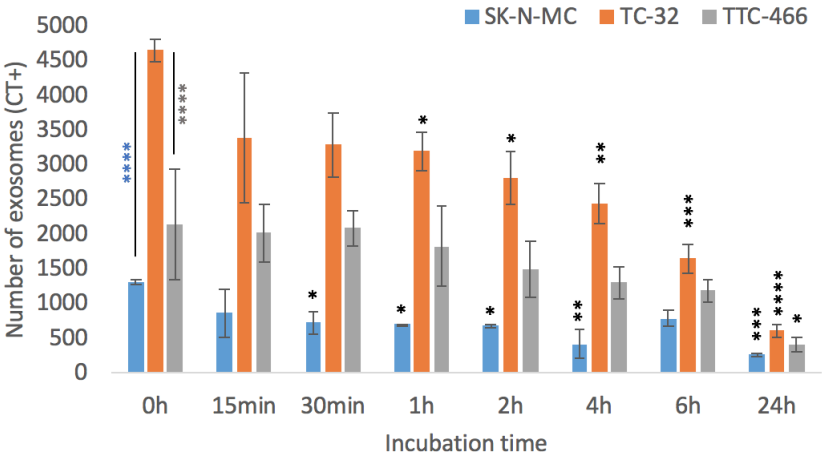


Figure 4.21: *In vitro* stability of labelled exosomes over 24 h evaluated by flow cytometry. Exosomes from SK-N-MC (blue), TC-32 (orange) and TTC-466 (gray) were labelled and incubated at 37 °C to evaluate the decrease in fluorescence over time (2 replicates per time point). Y axis represents number of CT positive events (exosomes). Statistical differences between number of exosomes at 0 h and later time points indicated by * (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.00001$). Difference on the exosome yield across cell lines indicated at 0 h (blue = SK-N-MC vs. TC-32, gray = TTC-466 vs. TC-32).

Moreover, from this experiment it was observed that the exosome production was heterogeneous across the different cell lines (Figure 4.21). The number of exosomes analysed was normalised to the number of producing cells, indicating that TC-32 cells produced more exosomes (4656 ± 160, $n = 2$) than SK-N-MC cells (1306 ± 39, $n = 2$; $p < 0.0001$) and TTC-466 cells (2146 ± 792, $n = 2$; $p < 0.0001$); and SK-N-MC and TTC-466 produce similar amounts ($p = 0.07$) (Figure 4.21).

Based on this data, and considering that exosomes had to have enough time to be taken up by cells in culture, SK-N-MC labelled exosomes were incubated with SK-N-MC cells over a time course of 20 h (1 h, 2 h, 4 h, 6 h, 10 h and 20 h). To investigate if exosomes had been taken up by cells, the increase in fluorescence in cells (number of cells and intensity of fluorescence) was investigated over time. To evaluate if the uptake was an active process, cells were incubated both at 37 °C and 4 °C, as the latter would affect uptake in case this was energy-dependent. As shown in Figure 4.22A, the increase in fluorescence is higher when cells are incubated at 37 °C, indicating that the exosome uptake is indeed an active process and can be partially blocked by reducing the metabolic activity of cells by keeping them at 4 °C.

In addition, data over the different time points indicated that the exosome uptake increased over time (Figure 4.22B).

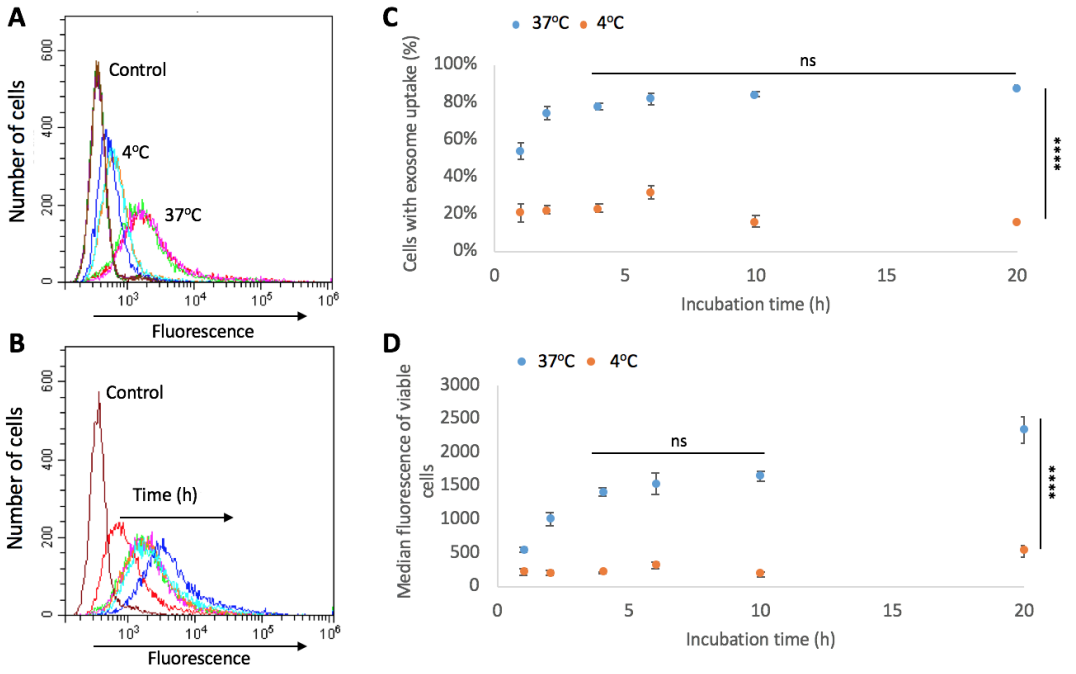


Figure 4.22: Evaluation of optimal incubation time for exosome uptake using flow cytometry. **A.** Exosome uptake is inhibited by incubation at 4 °C compared to 37 °C. 4 h time point shown as example. **B.** Increase in the CT fluorescence of recipient cells over time, compared to cells not exposed to exosomes. A and B graphs representative of 3 independent repeats. **C.** Number of cells with exosome uptake over time at 37 °C (blue) and 4 °C (orange). **D.** Median fluorescence of recipient cells over time at 37 °C (blue) and 4 °C (orange). C and D graphs data as mean ± SEM of 3 independent repeats. Significant differences (****$p < 0.0001$) between 37 °C and 4 °C for both evaluations. Plateau indicated by $p = ns$.

First, the number of cells that had taken up exosomes at 37 °C increased over time (1 h: 54 ± 14% cells, 4 h: 78 ± 4% cells; $p < 0.0001$), reaching a plateau at 4 h (78–88% from 4 h onwards, no statistical differences; Figure 4.22C). In comparison, the uptake at 4 °C did not change over time (0–20 h: 21–16 % cells with uptake; Figure 4.22C). When both temperatures were compared, a difference in uptake was seen at all time points ($p < 0.0001$) (Figure 4.22C). Then, the increase in fluorescence in cells was evaluated, as this was an indicator of the amount of exosomes that were taken up by cells. As can be seen in Figure 4.22D, there is a gradual increase during the first time points at 37 °C (546–11,401 median fluorescence; $p = 0.01$), reaching a plateau from 4 h to 10 h (1,401–1,640 median

fluorescence) and a final increase in fluorescence at 20 h (2337 ± 600 median fluorescence, $p < 0.0001$). Therefore, although the number of cells that uptake exosomes reaches a plateau over time, an increase in the amount of exosomes captured per cell is still visible at 20 h suggesting that exosomes are still being taken up by cells. When comparing both temperatures, there was no increase in the fluorescence of recipient cells at 4 °C, in comparison to 37 °C cells ($p < 0.0001$, except 1 h, $p = 0.06$), consistent with the conclusion that uptake of exosomes is an active process.

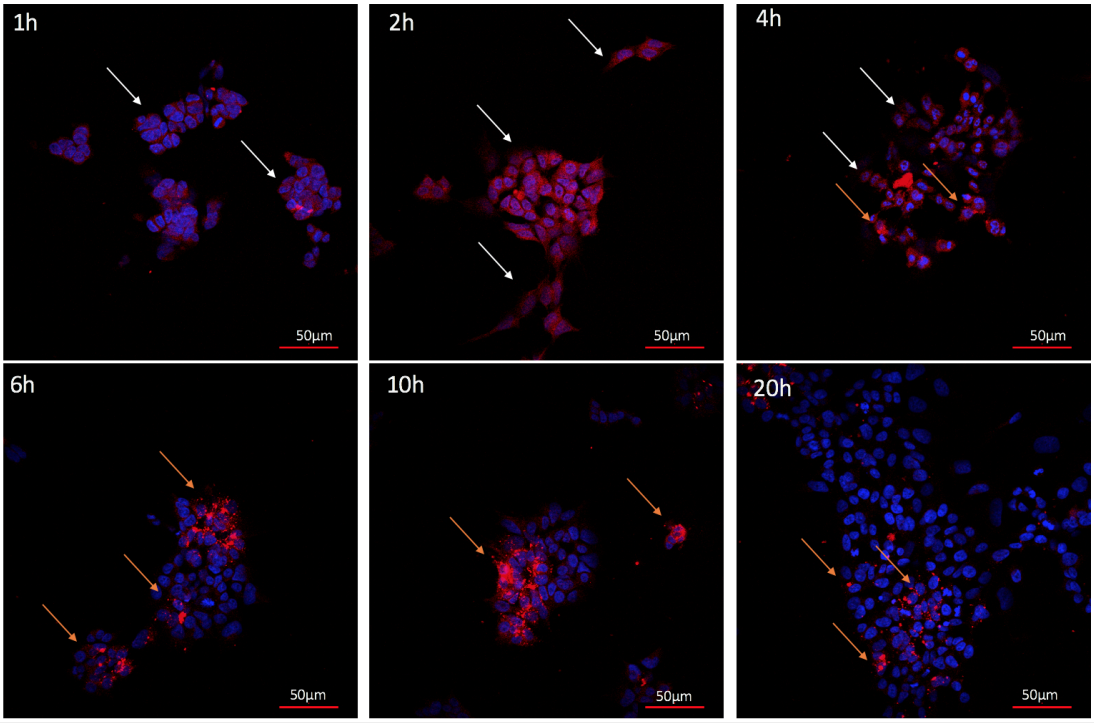


Figure 4.23: Exosome uptake over time evaluated by IF. Representative images of exosome uptake by cells at 37 °C over 20 h time course. Red = CT (exosomes); blue = DAPI (nucleus). White arrows = exosomes on cell surface; orange arrows = exosomes as intracellular clusters. Scale bar = 50 μm.

Uptake of exosomes was also analysed by IF of cells, to see if the exosome uptake could be visualised by microscopy (Figure 4.23). Data from the different time points suggests that exosomes initially cover the surface of the cells within 1 h of incubation. After 4–6 h, intracellular clusters of exosomes start to appear and the overall surface expression of exosomes was no longer visible. Therefore, an active uptake and internal trafficking of the exosomes taken up by cells can be seen by IF. Data suggests that after 4 h of incubation, exosomes leave the surface of the cells and are re-distributed in different areas of the cell, forming clusters of fluorescence (Figure 4.23), which is visible up to 20 h.

Based on both flow cytometry and IF data, the time points 4 h and 20 h were chosen to investigate the uptake of exosomes, as differences in the uptake were seen between both time points, and exosomes had already been distributed into specific areas of the cell.

To further test if the uptake of exosomes is concentration dependent, different concentrations of exosomes were added to cells and evaluated at 4 and 20 h, based on the ratio between exosomes from producing cells to recipient cells (i.e. exosomes from 10 cells into 1 recipient cell = 10:1 ratio). Flow cytometry data indicated that at 4 h, the number of cells taking up exosomes increased across the different ratios (from 1:1 to 25:1 ratio increase from 7±3 % to 86±5 %, differences $p < 0.0001$) reaching a plateau at 25:1 (25:1 86±5 %, 50:1 90±2 % cells with exosome uptake; $p = 0.9$; Figure 4.24A). In contrast, the increase in cells with exosome uptake at 20 h reached the plateau at the ratio 10:1 (from 1:1 to 10:1 increase from 10±5 % to 85±6 %, $p < 0.0001$; from 10:1 to 50:1 ratios plateau in uptake at 85–93 %, $p > 0.17$). Therefore, only ratios 1:1, 5:1 and 10:1 showed differences in the uptake between 4 h and 20 h (Figure 4.24A).

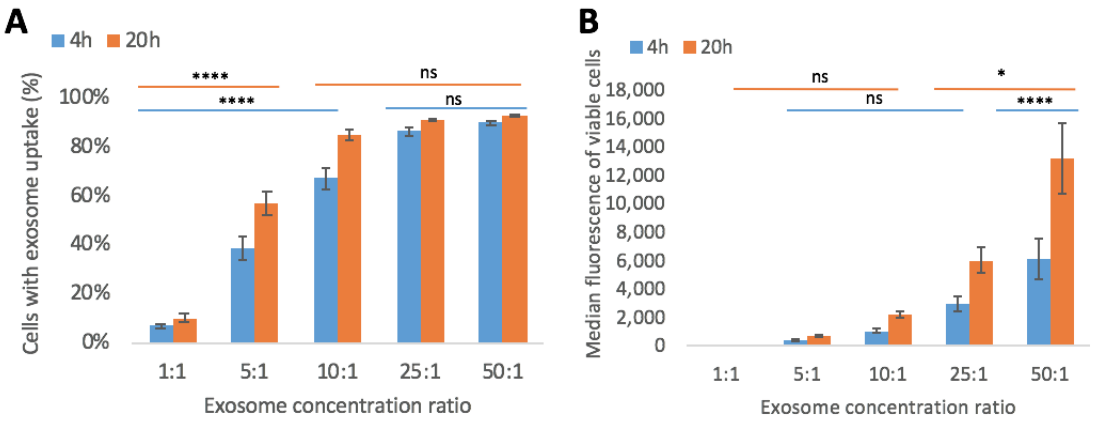


Figure 4.24: Uptake of exosomes by cells using different exosome:cell ratios evaluated by flow cytometry. Different amounts of exosomes from producing cell (×1, 5, 10, 25 and 50) to recipient cells (×1) were evaluated to identify the best exosome to cell ratio for uptake experiments. **A.** Number of cells with exosome uptake at 4 and 20 h across the different comparisons. An increase in uptake is seen for the lowest ratios, whereas it reaches a plateau at higher ratios at both time points. **B.** Median fluorescence of cells as indication of amount of exosomes taken up at 4 and 20 h using the different comparisons. A significant increase in the uptake was seen from 25:1 or 10:1 ratio at 4 h and 20 h respectively. No differences were seen across lower comparisons. Data as mean ± SEM of 3 independent repeats using SK-N-MC exosomes into SK-N-MC cells. ns = not significant differences, * $p < 0.05$, **** $p < 0.0001$.

When the median fluorescence of recipient cells was evaluated, this showed again that the fluorescence increased with higher exosome concentration for both time points. At 4 h, ratios 5:1 to 25:1 showed similar exosome uptake (391–2,981 median fluorescence, $p > 0.07$), and ratio 50:1 was the only one different from the rest (6076 median fluorescence, $p < 0.0001$; Figure 4.24B). However, at 20 h, differences were starting to be significant at ratio 25:1 (5,999–13,179 median fluorescence; differences to lower ratios $p < 0.02$; Figure 4.24B). Based on this data, the ratio 10:1 was chosen as an increase on the cells taking up exosomes could be seen between both time points, and the median fluorescence (number of exosomes going into the cells) was still increasing.

4.3.11 Evaluation of the exosome biodistribution in recipient cells

Having demonstrated that SK-N-MC cells take up SK-N-MC exosomes, I investigated the subcellular localisation of exosomes and markers of different cellular compartments using IF by confocal microscopy. Although the visualisation of single exosomes is not possible due to their small size, clusters of these vesicles can be

seen on the cells, as shown previously (Figure 4.23). First, exosomes labelled with CT were incubated with cells alongside exosomes labelled with anti-CD63, as previous data indicated that ES-derived exosomes express this exosomal marker (Section 4.3.6).

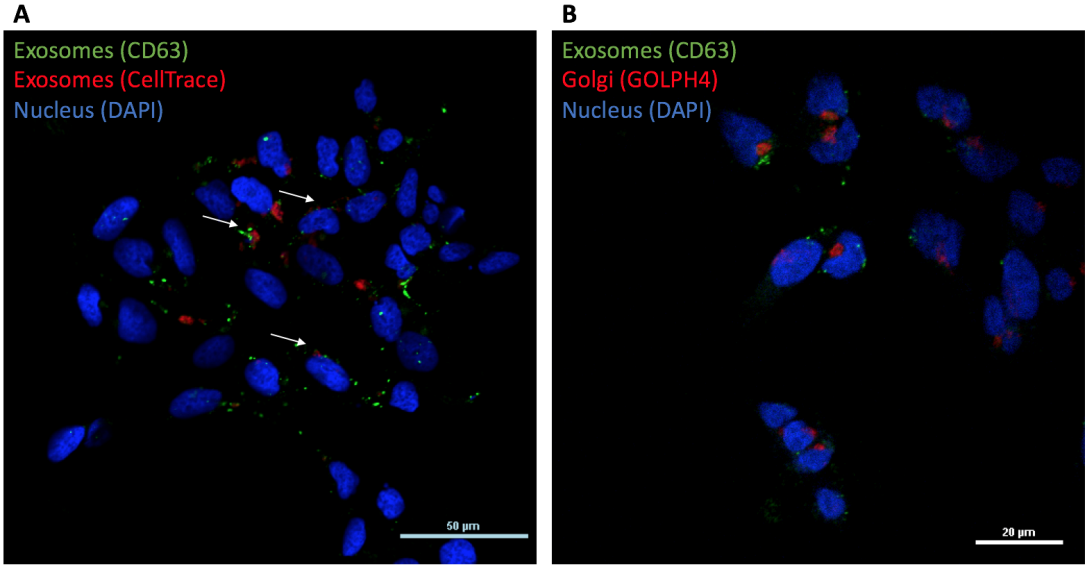


Figure 4.25: Biodistribution of CD63-labelled exosomes. **A.** SK-N-MC exosomes labelled with CT (red) or CD63 (green) after 4 h incubation with SK-N-MC cells. Arrows indicate areas where both exosome populations cluster nearby. Scale bar = 50 μm. **B.** CD63-labelled exosomes (green) after 4 h incubation with SK-N-MC cells, labelled with Golgi marker (GOLPH4, red). Scale bar = 20 μm.

As shown in Figure 4.25A, both exosome populations (anti-CD63-labelled and CT-labelled) localise closely in the cells, showing the same cluster morphology. Evaluation of CD63 exosomes on SK-N-MC cells showed that these vesicles do not co-localise with the Golgi (anti-GOLPH4), clustering in different areas of the cytoplasm (Figure 4.25B). Finally, SK-N-MC cells that had taken up CT-labelled exosomes showed that these vesicles clustered near the endoplasmic reticulum (green) but separate from the Golgi (Figure 4.26 top), as seen before for CD63-labelled exosomes. In addition, the evaluation of CT-labelled exosomes indicated that these vesicles cluster in specific areas of the cytoplasm (within the limits of the plasma membrane) and separate from the mitochondria (Figure 4.26 bottom). Therefore, SK-N-MC exosomes (CT labelled) cluster in specific areas of the cytoplasm but these do not co-localise with ER, mitochondria nor Golgi.

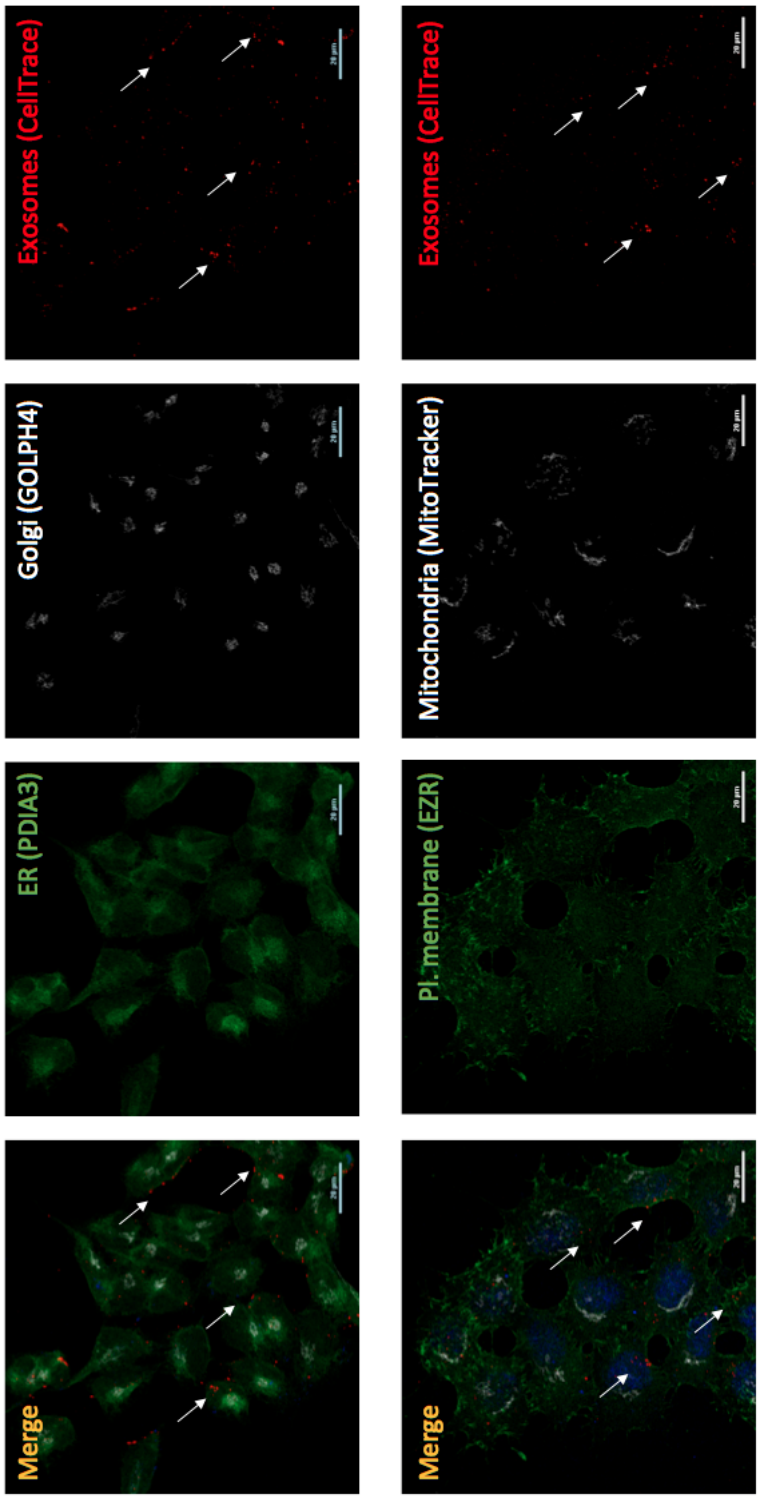


Figure 4.26: Biodistribution of exosomes inside recipient cells evaluated by IF. The localisation of labelled SK-N-MC exosomes (CT; red) after 4 h incubation with SK-N-MC cells was assessed. (Top) SK-N-MC cells labelled for endoplasmatic reticulum (ER; anti-PDIA3, green), Golgi apparatus (anti-GOLPH4, white); (Bottom) or plasma membrane (anti-EZR, green) and mitochondria (MitoTracker, white). Arrows indicate exosomes clusters in merged image and red field. Nucleus (DAPI, blue). Scale bar = 20 μm.

4.3.12 Non-ES cells can uptake ES exosomes

Having established that exosomes can be transferred to recipient SK-N-MC cells (ES exosomes to ES cells), the next step was to investigate if other cells could uptake SK-N-MC exosomes. These studies were undertaken to confirm that exosomes could be transferred between different cell types in my model system, so that it could be used to evaluate the role of exosomes in cross-talk between cancer and normal cells. For this, SK-N-MC exosomes (ES) were incubated with cells from another cancer type, NBL (SK-N-SH) over 4 and 20 h. ES cells TTC-466, which have a different fusion type than SK-N-MC, were evaluated as well to see if a different ES profile could have an implication on the uptake capacity.

As shown in Figure 4.27A, the 3 cell types could uptake exosomes at 4 h. This process was reduced when cells were incubated at 4 °C, consistent with previous studies (Section 4.3.10). Exosomes uptake increased over time (4 h to 20 h), with differences in the amount of cells that had taken up exosomes (NB69: 59% to 79% cells, $p = 0.0006$; SK-N-SH: 66% to 81% cells, $p = 0.0002$; TTC-466: 64% to 73% cells, $p = 0.01$; Figure 4.27B). This is consistent with previous data on SK-N-MC cells, which indicated that differences on the uptake were seen between these 2 time points (Section 4.3.10), seen again in here for SK-N-MC (used as positive control; 40% to 66% cells, $p = 0.0007$). In addition, an inhibition of the uptake could be seen at 4 °C (assessed at 4 h) across all samples, indicating that the uptake was indeed an active process in all cases (NB69: 40% vs 29% cells, $p < 0.0001$; SK-N-SH: 59% vs 34% cells, $p < 0.0001$; TTC-466: 64% vs 27% cells, $p = 0.0003$; Figure 4.27C).

The increase in fluorescence in cells was investigated across time, to evaluate the amount of exosomes taken up by the cells, rather than the cells that could uptake exosomes. Similar to previous data (Section 4.3.10), an increase could be seen between 4 h and 20 h, indicating that cells could still uptake exosomes after 4 h, not reaching a plateau in uptake (Figure 4.27C).

Differences across cell types could be seen regarding the exosome uptake efficacy. SK-N-MC cells had a lower uptake rate than NB69 (4 h: 40% vs. 59%, $p = 0.004$; 20 h: 66% vs. 79%, $p = 0.02$), SK-N-SH (4 h: 40% vs. 66%, $p = 0.0003$; 20 h: 66% vs. 81%,

$p = 0.02$) and TTC-466 (only at 4 h: 40% vs. 64%, $p = 0.0009$); but differences were not seen across the 3 cell lines. This indicates that SK-N-MC cells are less efficient on taking up exosomes than the rest of cell lines evaluated, suggesting that different pathways could be involved in the internalisation of exosomes depending on the recipient cells. When considering the amount of exosomes taken up per cell (median fluorescence), no differences were seen across the different cell lines.

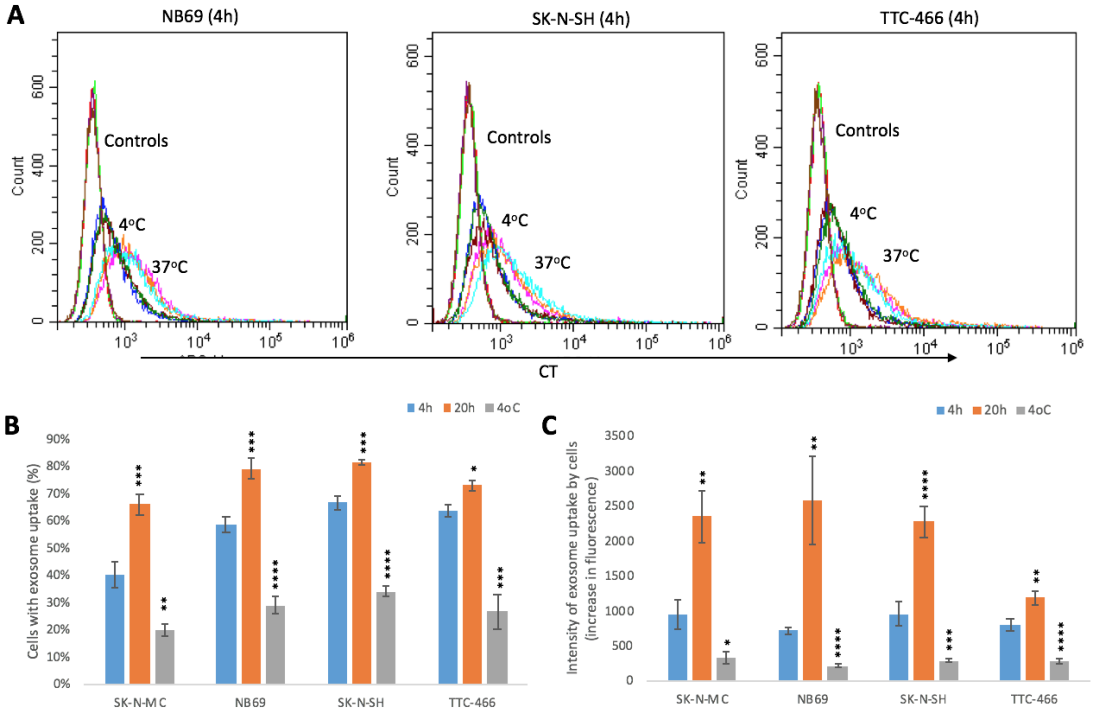


Figure 4.27: Evaluation of SK-N-MC exosome uptake by NB69, SK-N-SH and TTC-466 cell lines evaluated by flow cytometry. **A.** SK-N-MC exosomes taken up by (left to right) NB69 cells (NBL), SK-N-SH cells (NBL) and TTC-466 cells (ES) at 4 h at 37 °C and 4 °C. Graphs representative of 3 independent repeats. **B.** Percentage of cells with exosome uptake and **C.** intensity of exosome uptake (as median fluorescence on cells) at both 4 (blue) and 20 h (orange) at 37 °C. Uptake at 4 h at 4 °C (gray) to control the active process. Significant differences were seen on uptake between 37 and 4 °C (indicated on top of 4 °C bar). Significant differences were seen on the uptake between 4 and 20 h, indicating that cells did not reach a plateau (indicated on top of 20 h bar). Data as mean ± SEM of 3 independent repeats. P-values as: * 0.05–0.01, ** 0.01–0.001, *** 0.001–0.000,1, **** < 0.0001.

Similar to the evaluation performed on SK-N-MC cells (Section 4.3.11), the uptake of ES-labelled exosomes was evaluated by IF. SK-N-MC exosomes, once taken up by NB69 and SK-N-SH NBL cells, cluster in specific areas of the cytoplasm (Figure 4.28, red arrows), suggesting an active trafficking in different parts of the cell. This agrees with previous data on uptake by SK-N-MC cells, suggesting that this distribution of

exosomes is not particular of ES cells, but a general process of exosome uptake. In addition, uptake by ES cell line TTC-466 could be seen, indicating that the presence of a different fusion type did not have implications on the uptake capacity by cells (Figure 4.28).

Agreeing with previous data (Section 4.3.10), the uptake of exosomes by non-ES cells had already taken place at 4 h of incubation. The presence of labelled exosome clusters was still visible 20 h after adding the ES exosomes in culture (Figure 4.28), agreeing with the data for SK-N-MC exosomes taken up by SK-N-MC cells used for the optimisation studies.

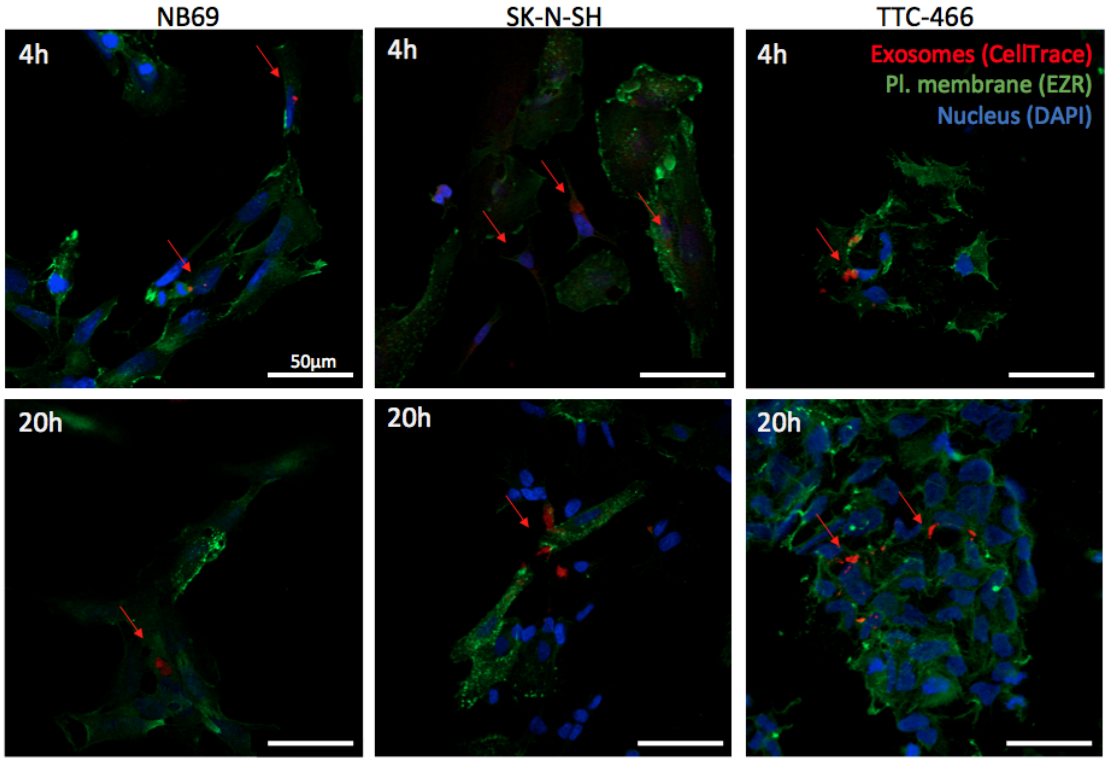


Figure 4.28: Evaluation of SK-N-MC exosome uptake by NB69, SK-N-SH and TTC-466 cell lines evaluated by IF. SK-N-MC exosomes (CellTrace, red arrows) taken up by (left to right) NB69 NBL cells, SK-N-SH NBL cells and TTC-466 ES cells at 4 h and 20 h. Red = CT (exosomes); Green = Plasma membrane (anti-EZR); blue = DAPI (nucleus).

4.3.13 MSC cells can uptake ES-derived exosomes

Several studies have suggested MSC to be the cell of origin in ES, as expression of the EWSR1-FLI1 fusion in these cells induces a transformation towards an ES

phenotype (Section 1.2.1). Therefore, I investigated if MSC could uptake ES-derived exosomes by IF, similar to what I have shown for NBL cells (Section 4.3.12). The confirmation that these cells could uptake ES-derived exosomes could suggest a new mechanism by which ES cells could transform MSC cells towards ES cells, thus playing a role in tumour progression and cellular cross-talk in this disease.

As shown in Figure 4.29, MSC cells can uptake ES-derived exosomes (CellTrace labelled) after 4 h of incubation. Instead of appearing as a cluster, as seen before at 4 h for NBL cells, exosomes in MSC are distributed throughout the cell cytoplasm (red arrow, Figure 4.29), and only in some areas they start to form clusters (yellow arrows, Figure 4.29). This could suggest a different intracellular trafficking in these cells, or a slower mechanism compared to NBL and ES cells, thus needing more time to appear in cluster formation.

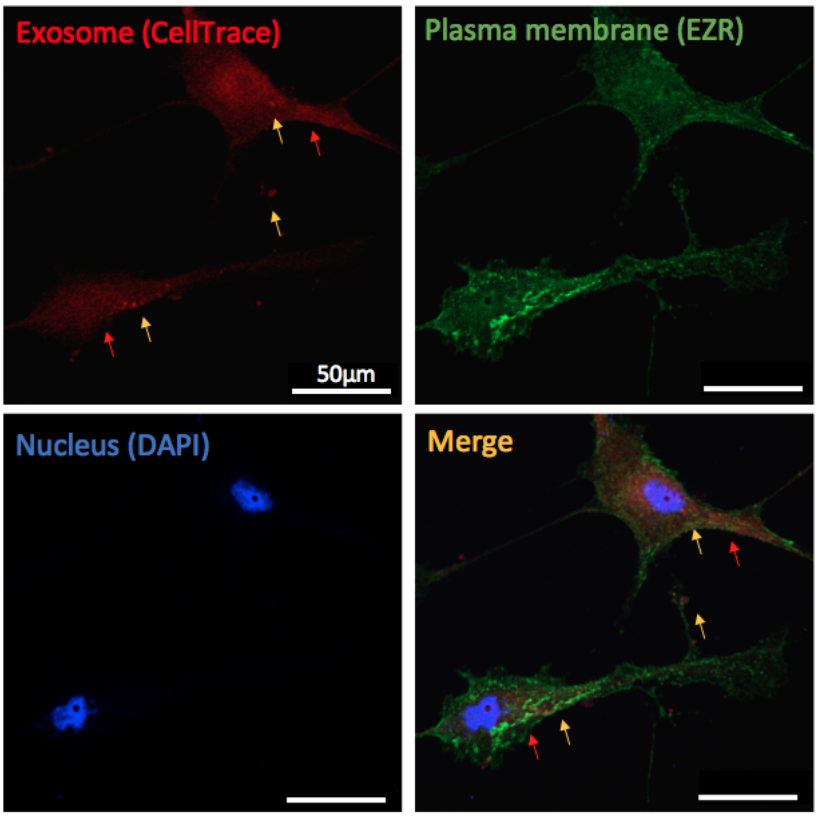


Figure 4.29: Visualisation of SK-N-MC exosome uptake by MSC cells by IF. MSC cells incubated with SK-N-MC exosomes (CellTrace, red) for 4 h. Red arrows indicate exosomes populations in MSC cytoplasm; yellow arrows indicate formation of clusters of exosomes.

Therefore, the ability of MSC to uptake ES exosomes opens a door to new evaluations of ES cross-talk and cancer progression, as ES exosomes could be transferring important ES cargo to MSC, which could lead to cellular transformation.

4.3.14 CD99 can be transferred from ES cells to non-ES cells via exosomes

Once it was confirmed that non-ES cells could uptake ES exosomes (Section 4.3.12), it was important to investigate if CD99, an important factor in ES linked to migration and cellular transformation (Section 1.1.3), could be transferred to non-ES cells through exosome transfer. Previous data on exosome profiling had shown that ES-derived exosomes contain the CD99 in both protein (Section 4.3.6) and RNA (Section 4.3.9.2) form.

For this, NBL cells (NB69 and SK-N-SH) were incubated with SK-N-MC-derived exosomes for 4 h and 20 h to evaluate if an increase in the CD99 expression was detected after exposure to ES-derived exosomes. This was compared to cells that had not been exposed to ES exosomes.

As shown in Figure 4.30A, cells that were not exposed to exosomes had a negative to low expression of CD99, but in some cases, some degree of expression of CD99 was seen in cells not exposed to ES exosomes (Figure 4.30B). This background CD99 expression on non-ES cells could not be explained by technical or incubation errors, as cell preparations from the same experiment showed differences in expression, as well as across different experiments. CD99 expression has been linked to cell cycle, and therefore, this could be a possible explanation on differences in expression across cell preparations. Nevertheless, in all cases, the CD99 expression of non-ES cells incubated with exosomes was higher than the control, even when these showed expression of CD99 (Figure 4.30C-D).

When evaluating the data from NB69 and SK-N-SH NBL cells incubated with ES-derived exosomes, an increase in the CD99 expression was seen (Figure 4.31), visible already at 4 h. This increase in CD99 expression was accompanied by a change in morphology, inducing larger cytoplasms (black arrows) and appearance of apoptotic cells (red arrows). This change was more visible for SK-N-SH cells,

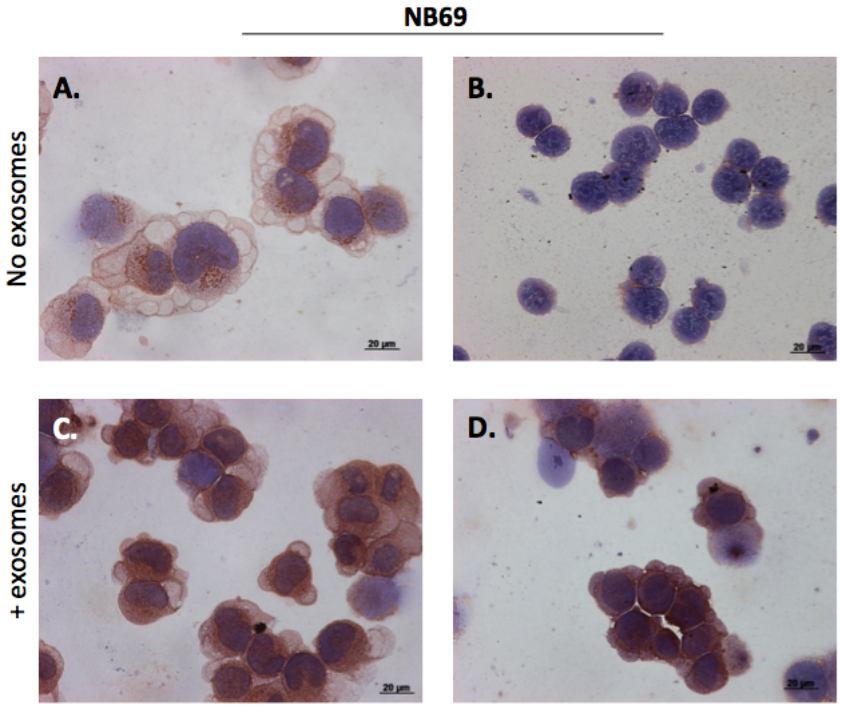


Figure 4.30: Variation in CD99 expression in NBL cells (NB69). CD99 expression on NBL cells (no exosomes added) varied across experiments. Cells not exposed to exosomes (A-B) show different levels of CD99 expression. However, NB69 cells exposed to ES exosomes had higher levels of CD99 expression (C-D) in all cases. NB69 cells showed as example.

especially at 20 h, when cellular blebbing and vacuolar formation (red arrows), indicative of apoptotic cells, was seen. (Figure 4.31).

Furthermore, the subcellular localisation of CD99 after exposure to ES exosomes was different than when these vesicles were tracked by IF (Section 4.3.12). Here, CD99 expression is visible around the cell surface (blue arrow), whereas IF evaluation of exosome trafficking in NBL cells showed clusters of exosomes in specific areas of the cytoplasm. Therefore, this could suggest that CD99 transferred from exosomes goes to specific areas of the cells, instead of following the exosomes-trafficking pathway. Further evaluation of the routes exosomes follow once internalised and how CD99 is sorted inside recipient cell would further clarify this. Therefore, ES exosomes can transfer CD99 from ES cells to non-ES, thus inducing a change in the CD99 cellular expression.

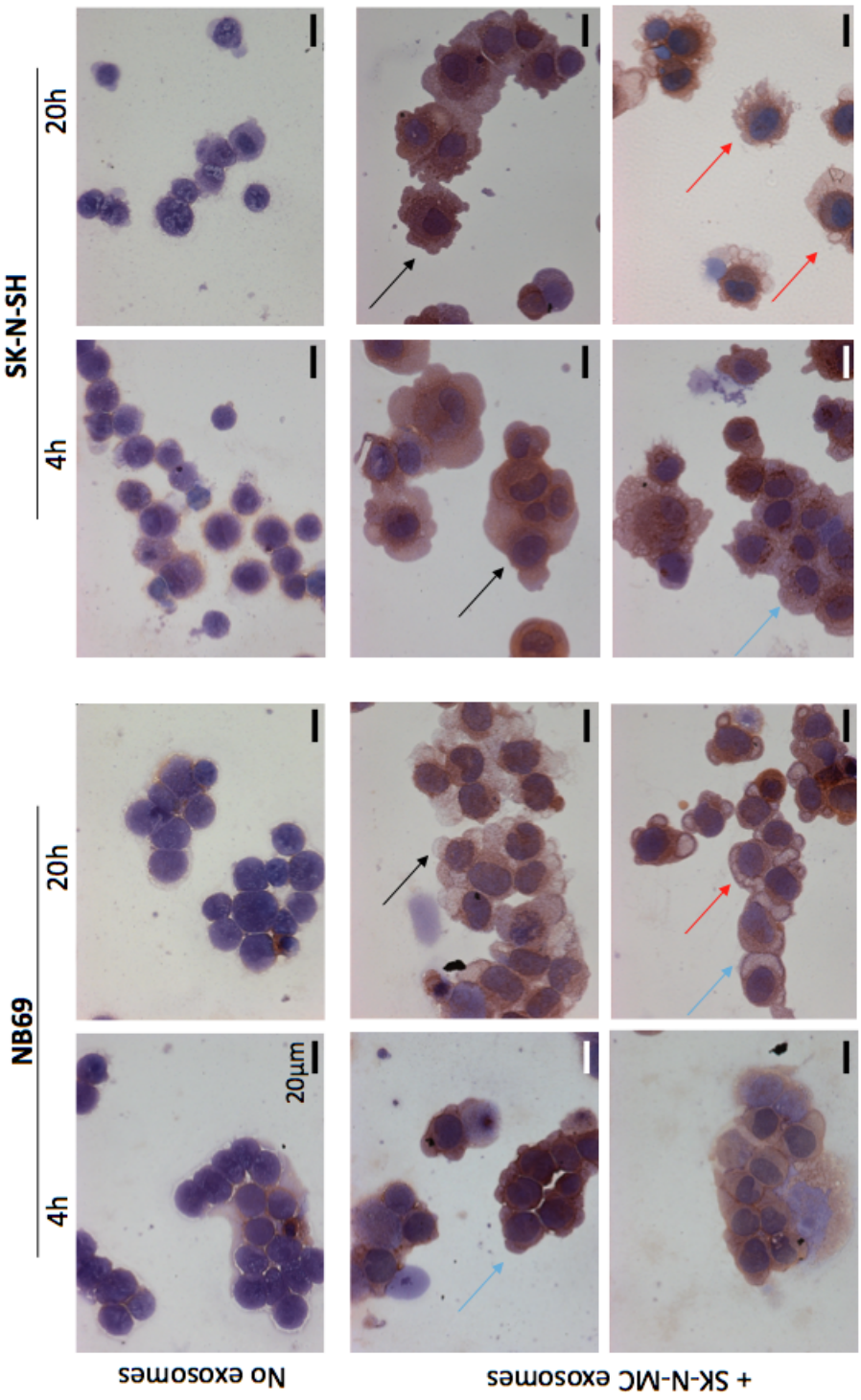


Figure 4.31: Increase in CD99 expression after ES exosomes are taken up by NB69 and SK-N-SH NBL cells evaluated by ICC. CD99 expression of NB69 (left) and SK-N-SH (right) after 4 h and 20 h of incubation with SK-N-MC exosomes. Cells without exosome incubation used at negative control (top). Data representative of 3 (NB69) and 2 (SK-N-SH) independent repeats. Black arrow = increased cytoplasm, red arrow = apoptotic cell appearance, blue arrow = CD99 expression on cell surface.

As previous data (Section 4.3.13 indicated that MSC cells can uptake ES-derived exosomes, it was important to assess if CD99 expression could be induced in MSC through the uptake of ES exosomes, as shown for NBL cells. The induction of expression or transfer of CD99 from ES cells to MSC is important as the latter cell type has been proposed as the cell of origin in ES (Section 1.2.1), and therefore, this could be a mechanism by which ES cells can lead to tumour progression and cellular transformation.

Similar to what was seen on NBL cells (Section 4.3.14, MSC have an increase in the CD99 expression after being incubated for 4 h with SK-N-MC exosomes. However, the change is less obvious due to the basal expression levels of CD99 in MSC cells, as can be seen in Figure 4.32.

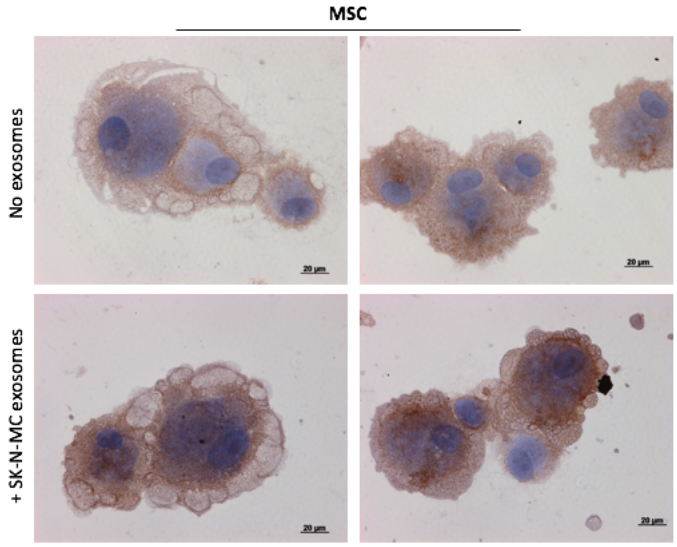


Figure 4.32: Increase in the CD99 expression by MSC after SK-N-MC exosomes uptake evaluated by ICC. MSC incubated for 4 h without (top) or with (bottom) SK-N-MC exosomes.

4.3.15 The *EWSR1* fusion can be transferred to non-ES cells via exosomes

As well as investigating the transfer of CD99 to non-ES cells, the uptake of the *EWSR1* fusion gene (DNA) from ES-derived exosomes was evaluated in NBL cell lines NB69 and SK-N-SH, and the normal MSC cells, as these cells do not have the fusion. Data from previous exosome characterisation had shown that SK-N-MC exosomes contain the gene fusion, investigated by RT^A-qPCR data (Section 4.3.9.1).

As discussed in Section 2.2.8, the *EWSR1* DNA probe used identifies when a break apart is present after exon 4. Therefore, if the probe is intact (no break apart), both 3' and 5' regions of the gene will be together, detecting yellow fluorescence. However, if only the 5' region (involved in the characteristic *EWSR1* ES fusion) is transferred to non-ES cells, an increase in the red probe will be detected, as the 3' region of the gene (green probe) will not be transferred.

Intact *EWSR1* probe was mainly detected in NB69 cells not exposed to ES exosomes (84 ± 22 % of detected probes, number of cells analysed $n = 41$), and only a small fraction involved the 5' probe (5 ± 11 % of detected probes, $n = 41$), which could be explained by instability in the karyotype of cancer cells, a common phenomenon in cancer. However, after incubation with ES-derived exosomes for 4 h, an increase in the amount of 5' *EWSR1* probe was detected (11 ± 20 % detected probe, $n = 167$, $p = 0.06$), whereas the amount of intact probe did not change compared to the control (80 ± 30 % detected probe, $n = 167$, $p = 0.35$; Figure 4.33A, Figure 4.34), consistent with transfer of the fusion to these cells. This increase was more significant after 20 h of incubation with ES-exosomes. Whereas the amount of intact *EWSR1* did not change (control 87 ± 24 % detected probe, $n = 50$; exposed to exosomes 82 ± 29 % detected probe, $n = 251$, $p = 0.25$), the number of 5' probe present in cells significantly increased after exposure to ES exosomes (control 2 ± 7 % detected probe, $n = 50$; exposed to exosomes 13 ± 22 % detected probe,$n = 251$, $p = 0.0007$; Figure 4.33A). This data suggests that even still having the full *EWSR1* gene in NB69 cells, there is an increase in the 5' probe, suggesting the transfer of the characteristic *EWSR1* fusion from ES cells to NB69 cells. This increase is more significant over time, thus confirming that ES exosomes are still being taken up after 4 h of incubation.

Evaluation of the uptake of the *EWSR1* fusion into SK-N-SH NBL cells showed that a high proportion of *EWSR1* probe is intact in SK-N-SH cells not exposed to ES exosomes (79 ± 33 % detected probe, $n = 34$), although a reduction was seen when exposed to these vesicles (63 ± 35 % detected probe, $n = 101$, $p = 0.03$; Figure 4.33B). However, the amount of 5' *EWSR1* probe detected increased from cells not exposed to exosomes (9 ± 18 % detected probe, $n = 34$) to cells exposed

to ES exosomes (18 ± 21 % detected probe, $n = 101$, $p = 0.05$) after 4 h incubation (Figure 4.33B, Figure 4.34).

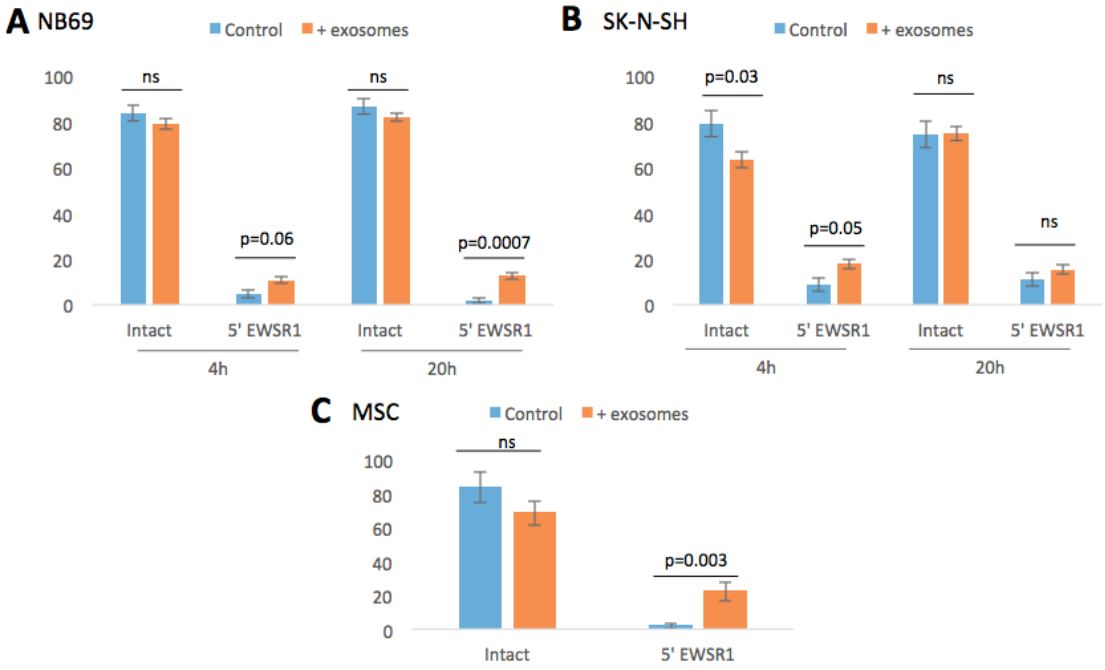


Figure 4.33: Transfer of the *EWSR1* fusion gene to non-ES cells via ES exosome uptake evaluated by FISH. The number of intact *EWSR1* and 5' *EWSR1* probe (involved in *EWSR1* fusion) in each cell was quantified in cells exposed for 4 h and 20 h with ES exosomes or not (control). Data as mean ± SEM of all scored cells at each condition on **A.** NB69, **B.** SK-N-SH and **C.** MSC cells (only 4 h data). Differences between control and cells exposed to exosomes are described.

This data, in agreement with NB69 characterisation, suggests that ES exosomes can transfer the *EWSR1* fusion to SK-N-SH cells, as an increase in the 5' *EWSR1* probe can be seen after exposure to exosomes. However, this increase is lost after 20 h, when no differences are seen in the amount of 5' probe between cells exposed to ES exosomes (15 ± 23% detected probe, $n = 126$) and cells not exposed (11 ± 19% detected probe, $n = 44$; $p = 0.24$; Figure 4.33B). Therefore, in contrast to what is seen in NB69 cells, SK-N-SH NBL cells can uptake the *EWSR1* fusion through the transfer of ES exosomes, but this is not maintained in time, as a decrease in the 5' probe is seen from 4 to 20 h incubation in exosomes. This could suggest that the SK-N-SH cells do not tolerate the fusion inside them. This could explain why these cells had a more apoptotic appearance by ICC when evaluating the expression of CD99 after ES exosome uptake (Section 4.3.14).

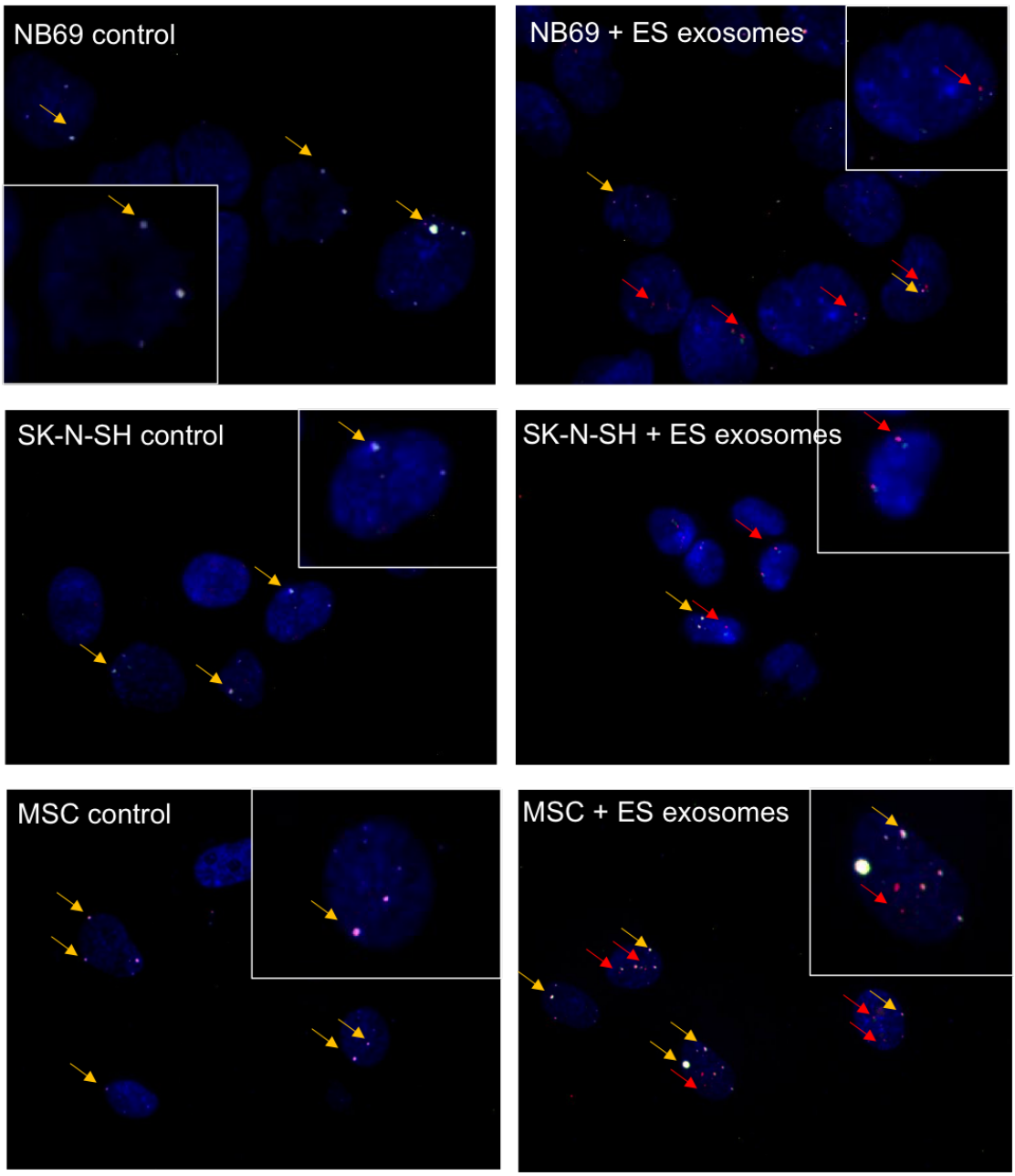


Figure 4.34: *EWSR1* 5' region transferred to non-ES cells by exosome uptake. FISH data on NB69, SK-N-SH and MSC cells showed presence of intact *EWSR1* gene (white-yellow probe, yellow arrows) on control (no incubation with exosomes). However, after 4 h incubation with SK-N-MC exosomes, an increase in the number of 5' *EWSR1* probe (red probe, red arrows) was detected. Data representative of 2 independent repeats (MSC only 1).

Finally, the profile of MSC cells after being exposed for 4 h to ES-derived exosomes was evaluated. Similar amounts of intact *EWSR1* probe were detected in MSC not exposed to exosomes (84 ± 32 % probe detected, $n = 12$) and those exposed for 4 h to ES exosomes (68 ± 29 % probe detected, $n = 16$, $p = 0.17$; Figure 4.33C). However, when the proportion of 5' *EWSR1* probe was evaluated, a significant difference was seen between control cells (2 ± 6 % probe detected, $n = 12$) and cells incubated with exosomes (23 ± 22 % probe detected, $n = 16$, $p = 0.003$; Figure 4.33C, Figure 4.34), confirming that ES exosomes can transfer the *EWSR1* fusion to MSC cells, suggested by the increase in 5' *EWSR1* DNA probe.

Altogether, this data indicates that ES exosomes can transfer the characteristic *EWSR1* fusion to non-ES cells by releasing their cargo, but this does not inform on the ability of the recipient cells to support the *EWSR1* fusion or the induction of changes in the phenotype.

4.3.16 The uptake of ES exosomes by non-ES cells induces changes in the self-renewing capacity

After evaluating that specific ES biomarkers can be transferred from ES cells to non-ES cells through exosomes, it was important to investigate if this uptake could induce changes in recipient cells. As the self-renewing ability is an important trait of the CSC profile and has been widely associated to cancer progression (Section 5.2), the ability of ES exosomes to induce changes in this cancer phenotype was studied.

SK-N-MC cells have a high self-renewing ability (26 ± 5 %; Figure 4.35), and therefore, we might hypothesise that uptake of SK-N-MC exosomes by cells with less self-renewing ability would induce an increase in this. Therefore, using the same strategy as described in Section 2.3.5, the self-renewing assay was used for NB69, SK-N-SH and MSC cells incubated or not to ES-derived exosomes throughout the assay. Previous data indicated that ES exosomes can be taken up by these cells, and this is visible still after 20 h. However, due to the duration of the self-renewing assay (3 weeks), this experiment will inform on the ability of ES exosomes to induce a change in the phenotype, and if this is retained after 3 weeks of incubation.

NB69 cells showed an increase in the self-renewing ability when exposed to ES

exosomes (5.8 ± 4.3 %) compared to cells not exposed to these (1.9 ± 1.7 %, $p < 0.0001$; Figure 4.35). A similar change was seen in MSC, as exposure to ES exosomes induced an increase in the self-renewing ability (2.3 ± 1.7 %) when compared to MSC not exposed to exosomes (1 ± 1 %; $p = 0.002$; Figure 4.35).

Contrary to NB69, which have a low self-renewing ability, SK-N-SH cells showed a high intrinsic self-renewing ability (not exposed to exosomes; 24 ± 4 %). When exposed to ES exosomes, no differences in the increase on the self-renewing ability were seen (26 ± 2 %; Figure 4.35). Although this observation is from a single repeat, this difference could suggest that as SK-N-SH have already a high ability to self-renew, ES exosomes cannot induce a change in profile as significant as in the cells with low self-renewing ability.

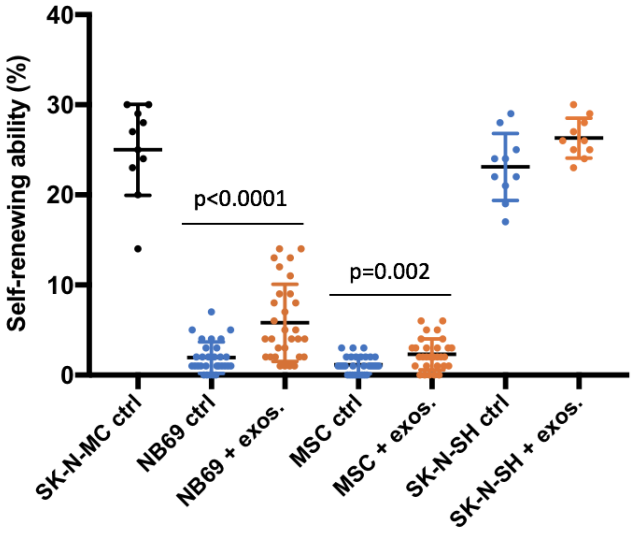


Figure 4.35: Changes in the self-renewing ability of non-ES cells after incubation with SK-N-MC exosomes. SK-N-MC cells (exosome origin; black dots, $n = 10$) have a high ability to self-renew. NB69, MSC and SK-N-SH cells were exposed (orange) or not (blue) to SK-N-MC exosomes for the duration of the experiment. Differences between control and exosome-treated cells are shown for each cell line. Data as mean ± SD of 3 independent repeats, 10 plates per repeat (each dot represents a plate, $n = 30$); SK-N-SH only 1 repeat ($n = 10$). Ctrl. = control, Exos. = exposed to ES exosomes.

Therefore, ES exosomes not only can be taken up by non-ES cells, but these can induce a change in the self-renewing ability of recipient cells, confirming the role of exosomes in inducing changes in cellular behaviour and their importance in cancer progression. It is important to consider, however, that based on these experiments

we do not know if the changes induced are ES-related (dependent on the cargo of exosomes from SK-N-MC cells) or exosome-related. To test this, we would need to include an additional control of adding exosomes from other cell types to these recipient cells (NBL and MSC), and see if the changes induced are maintained or not.

4.3.17 The migration index does not changes after uptake of ES exosomes by non-ES cells

The migration ability of cells is another important phenotype of cancer progression, as it is correlated with the ability of cells to migrate to other areas (Section 2.3.6). Initially, NB69 (NBL cell line) had to be used to investigate the impact of ES exosomes to the migration index of non-ES cancer cells. However, NB69 cells cannot form spheroids, and therefore, the migration assay could not be performed.Therefore, the migration index was investigated in SK-N-SH cells (NBL cell line) over 72 h after treatment with ES-derived exosomes, as described in Section 2.3.6. To study if SK-N-MC exosomes could have an effect on the migration capacity of recipient cells, spheroids were incubated with and without exosomes.

No differences in the migration index were observed between SK-N-SH cells exposed or not to ES exosomes (72 h; control 5 ± 1 %; + exosomes 5 ± 1 %; $p = 0.90$; Figure 4.36A). Alongside the evaluation of SK-N-SH NBL cells, SK-N-MC cells were evaluated as well, to see if an increased exposure to exosomes from these cells (as using ratio 10:1) could induce a change in SK-N-MC migration index. Again, no differences were seen between control and exposure to exosomes (72 h; control 6 ± 1 %; + exosomes 6 ± 1 %; $p = 0.51$; Figure 4.36B). Therefore, SK-N-MC exosomes cannot induce a change in the migration index on SK-N-SH cells, or at least, with the conditions used for this study. It is important to mention that using this methodology to test migration changes, it is not possible to test if exosomes can be taken up by cells inside the sphere. This could lead to only cells on the outside being exposed to exosomes, and not having enough capacity to change the behaviour of the overall spheroid population.

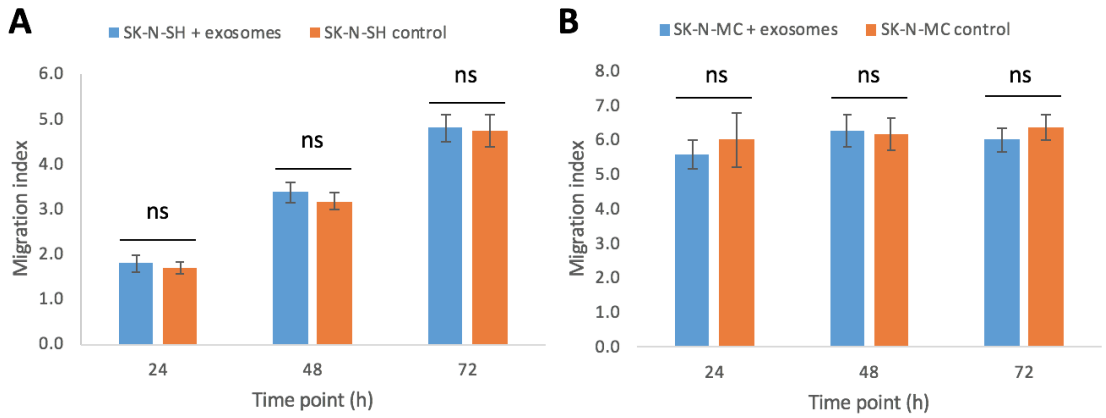


Figure 4.36: Evaluation of migration index on SK-N-SH cells after exposure to ES-derived exosomes. **A.** SK-N-SH migration assay performed with (blue) and without (orange) SK-N-MC exosomes over 72 h incubation time. **B.** SK-N-MC the migration index evaluated with (blue) and without (orange) SK-N-MC exosomes. Data as mean ± SEM of 2 independent repeats. ns = not significant differences.

4.3.18 Exosomes from ES chemoresistant cells do not transfer a resistant profile to chemosensitive cells

In addition to studying the impact of ES exosomes on inducing changes in the self-renewing ability and migration index on recipient cells, the ability to change the sensitivity to chemotherapeutic drugs was evaluated. Exosomes have been linked to the transfer of specific cargo that alters the phenotype of recipient cells, inducing more resistant profiles (Section 1.3.2). Therefore, I investigated if exosomes from a resistant ES cell line (SK-N-MC EC50 cells; cultured in 28 nM doxorubicin) could induce a change in the sensitivity to doxorubicin when taken up by native SK-N-MC, in comparison to SK-N-MC cells not exposed to exosomes.

The EC50 of the resistant and non-resistant SK-N-MC cells was calculated to demonstrate the difference in sensitivity to doxorubicin. As shown in Figure 4.37A, resistant SK-N-MC had a higher EC50 (82 ± 7 nM) compared to native SK-N-MC (21 ± 9 nM, $p = 0.02$).

When native SK-N-MC cells were incubated with exosomes from the resistant SK-N-MC cells for 24 h prior to the treatment with doxorubicin, no difference in sensitivity to doxorubicin were seen between exposure to exosomes (EC50: 24 ± 7 nM) and control cells (EC50: 21 ± 9 nM, $p = 0.74$; Figure 4.37A). A similar result was obtained when cells were incubated with resistant exosomes for 48 h prior to doxorubicin

treatment, as control had a similar sensitivity (EC50: 45 ± 17 nM) to cells exposed to exosomes (EC50: 38 ± 18 nM, $p = 0.88$; Figure 4.37B).

Therefore, in the conditions used, exosomes derived from a resistant cell line cannot induce changes in the resistance profile of recipient cells to doxorubicin. Although cells were incubated up to 48 h with exosomes prior to treatment with doxorubicin, it could be that this is not enough time to induce a change in the recipient cell profile. In Section 4.3.16, a change in the self-renewing profile was seen after 3 weeks of incubation with exosomes. Therefore, longer exposure to exosomes could be evaluated, as 48 h may not be enough to induce a selective advantage for doxorubicin resistance.

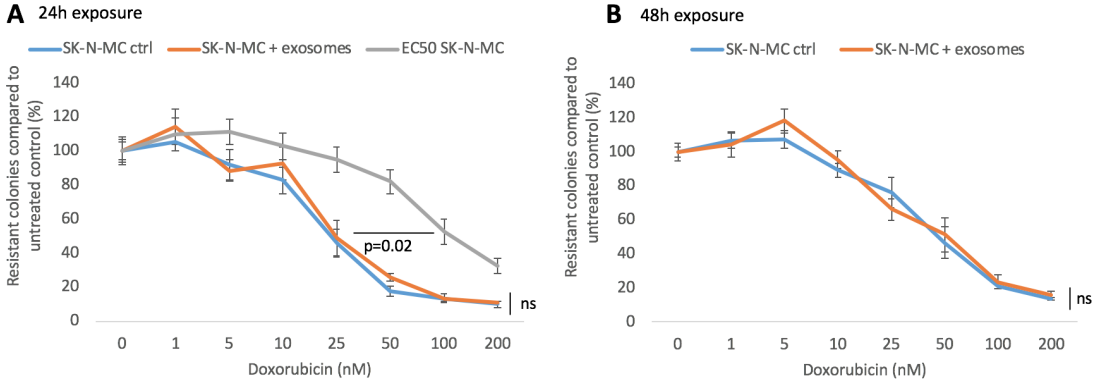


Figure 4.37: Evaluation of drug response on chemosensitive cells exposed to exosomes from resistant cells. SK-N-MC sensitive cells incubated for **A.** 24 h and **B.** 48 h with exosomes derived from resistant cell line (SK-N-MC EC50). Control cells (no exosomes added; blue line), cells + resistant exosomes (orange line), SK-N-MC resistant cells (where exosomes are derived from; gray line). Data is represented as mean ± SEM of 2 independent repeats. Differences between sensitive and resistant SK-N-MC cells indicated in A ($p < 0.0001$).

4.4. Discussion

In the current study I have shown that ES-derived exosomes can induce changes in the self-renewing ability of non-ES cells. In addition, the characteristic ES markers CD99 and *EWSR1* fusion can be transferred to non-ES cell via exosome uptake. Moreover, to my knowledge, this has been the most comprehensive profiling of ES exosomes from patient-derived samples, including RNA-seq characterisation of the cargo and evaluation of different exosome-related profiles. To date, the cargo of ES-derived exosomes had only been evaluated by RT^A-qPCR studying specific targets

(Miller et al., 2013; Tsugita et al., 2013; Ventura et al., 2016; Villasante et al., 2016) or micro arrays (Miller et al., 2013).

ES-derived exosomes (SK-N-MC) can induce an increase on the self-renewing ability on non-ES cells, including MSC. This was shown after the uptake of ES exosomes by NBL cells and MSC cells was confirmed, suggesting that exosomes could be involved in the cross-talk between different cell types. These findings agree with data from other groups, that have reported cancer-derived exosomes to induce a change in the phenotype of recipient cells, towards a more oncogenic state (Hu et al., 2015; Rodríguez et al., 2015; Sun et al., 2018). Similar to what has been observed in other cancers (El Andaloussi et al., 2013), ES exosomes could be involved in the tumour progression and microenvironment modelling via the release of their cargo to non-ES cells, inducing changes in their phenotype. This agrees with data from another study in ES exosomes (Ventura et al., 2016), in which the transfer of cargo from CD99-silenced ES cells to wild-type ES cells was enough to induce a change in phenotype mimicking the CD99 silencing. Contrary to the investigation by Ventura et al. (2016), in which the exact pathway behind the change in profile was investigated, the results presented here have only evaluated the change in the self-renewing ability. Therefore, future work should investigate the differences in the gene expression profile of non-ES cells before and after being incubated with ES exosomes in order to identify which pathways are involved in this process.

Unlike the change in self-renewing ability, ES-derived exosomes failed to induce changes in the migration index nor the resistance profile to chemotherapeutic drugs across different cell types. Although this was the first time these changes in recipient cells were evaluated using ES exosomes, other groups have shown that exosomes from other cancers can induce changes in the migration (Zheng et al., 2018) or an increase in the resistance when the transfer goes from resistant cells to sensitive cells (Lv et al., 2014). Contrary to the results presented here, Zheng et al. (2018) described the ability of exosomes from tumour-associated macrophages to induce a change in the gastric cancer cell migration and invasion capacity upon co-culture, playing a role both *in vitro* and *in vivo* mouse models. A possible explanation for the negative results in my studies is that I have

investigated the migration capacity of cells forming an spheroid, and therefore, I have not taken into consideration the possibility that exosomes might not reach the inner core. If that was the case, only cells on the outer layers of the spheroid would be exposed to the ES-derived exosomes. In contrast to this method, Zheng et al. (2018) used a transwell assay where they placed gastric cancer cells previously co-cultured (or not) with macrophage exosomes, and evaluated the ability to migrate or invade (through matrigel) towards to lower chamber, where macrophages with FBS were placed. Therefore, using this method, gastric cells were exposed to exosomes in a monolayer format. To overcome this, future studies should evaluate the changes in migration using an assay such as transwell or scratch assay, where all cells would be exposed to exosomes. Another explanation for the absence of change, which could be applied to the lack of induction of resistance too, is the nature of the cells used. It could be that the cells used in these assays (SK-N-SH for migration and sensitive SK-N-MC for drug assay) lack plasticity towards a change induced by the exosomes they are exposed to. In order to fully evaluate if this is the case, other cell types could be investigated as recipients, as it could be that ES-derived exosomes induce different degrees of change in depending on the recipient cell type. In addition, when comparing to the studies by Lv et al. (2014), a longer incubation with exosomes was performed (72 h) and the exposure to the chemotherapeutic evaluated (docetaxel) was for 18 h only, instead of 48 h. Therefore, next studies in ES exosomes should evaluate longer exposure times with exosomes to investigate if cells need more time to change their resistance profile.

The presence of the characteristic EWSR1 fusion was confirmed on exosomes from established ES cell lines, both on EWSR1-FLI1 and EWSR1-ERG fusion types. However, when the presence of the fusion was investigated in exosomes from patient-derived cell cultures, this was not detected. As the PPIA levels (used as positive control) are comparable between exosomes from cell lines and patient-derived cell cultures, this differences in the detection could be due to the expression levels of the fusion rather than a difference in the exosome yield evaluated. This difference in expression between these two sample types agrees with previous data (Section 2.3.2.2, Section 3.3.2.4), where the characterisation of

RNA levels of EWSR1 in cells showed that the fusion was easily detected in established cell lines compared to patient-derived cell cultures (RT^A-PCR, Section 2.3.2.2; RNA-seq, Section 3.3.2.4). In addition, the fusion could not be detected in exosomes by RNA-seq, even in the SK-N-MC-derived exosomes where the fusion was identified by RT^A-qPCR. This lower sensitivity by RNA-seq is in agreement with data from (Section 3.3.2.4, in which RNA-seq failed to identify the fusion in samples where RT^A-PCR did.

When non-ES cells, including MSCs, were incubated with ES exosomes, an increase in the amount of 5' *EWSR1* DNA probe (contained as part of the characteristic gene fusion, instead of the 3' or the full *EWSR1* gene) was detected compared to non-ES cells not exposed to ES exosomes. For this, a FISH DNA protocol was used instead of the RT^A-qPCR, as this would provide information on the presence of the EWSR1-ETS gene fusion per cell, being more sensitive as shown in Section 2.3.2.2. This transfer of the gene fusion agrees with previous data on the ES field (Tsugita et al., 2013; Villasante et al., 2016). These findings are particularly important, as the EWSR1 gene fusion is the main driving event in ES (May et al., 1993), its expression in MSCs (Riggi et al., 2008) or NBL (Rorie et al., 2004) cells being sufficient to induce a change in these cells towards an ES-like phenotype. To see that MSC can uptake the gene fusion via ES exosomes suggests a new possibility for ES tumorigenesis, as MSCs have been proposed as the ES cell of origin (Riggi et al., 2008; Tirode et al., 2007). However, the transfer of the EWSR1 fusion was only evaluated at the DNA level (as this FISH probe is for DNA), whereas the presence of the fusion in ES-derived exosomes was done at the RNA level (RT^A-qPCR). Future studies should therefore investigate this transfer at RNA level or protein level, first to validate if the transfer of the RNA cargo is effective via exosomes and an increase in the presence of EWSR1 fusion is found on non-ES cells, and second to evaluate in the transferred genetic material (either DNA or RNA) can be transcribed and translated into EWSR1-fusion protein, confirming that the fusion is functional. Once this is confirmed, future studies should investigate if the fusion, once internalised into MSC and NBL cells, can lead to changes in the expression profile towards ES biology by evaluating the gene expression profile of MSC exposed or not to ES-derived exosomes. Altogether, this would inform on the

ability of ES exosomes to modulate tumour progression by transforming cells from the microenvironment into ES-like cells or initiating transformation towards a metastatic niche.

Exosomes from patient-derived cell cultures, ES-CSCs, established cell lines and plasma from ES patients have shown expression of CD99 on the exosome surface (Section 4.3.6, Section 4.3.7.2, Section 4.3.8) and as part of the RNA cargo in established cell lines and patient-derived exosomes (Section 4.3.9.2). The presence of CD99 on ES-derived exosomes agrees with previous published data (Ventura et al., 2016), although this was investigated by WB whereas in my study WB analyses failed to identify CD99 protein expression on ES exosomes (Section 4.3.7.2). When evaluating the uptake of ES exosomes by non-ES cells, an increase on the CD99 protein expression was seen in both NBL and MSC cells, suggesting that ES-derived exsosomes can transfer CD99 to non-ES cells, changing their protein expression profile. This is an important finding, as CD99 has been implicated in ES phenotype (Manara et al., 2016) and targeting of this molecule leads to cell death in ES *in vitro* and *in vivo* mouse models (Guerzoni et al., 2015; Scotlandi et al., 2006). Moreover, as CD99 mRNA was identified in the cargo of exosomes, it would be crucial to evaluate if these mRNAs are transferred to recipient cells and can be translated into functional protein. This would help understand if the increase in CD99 protein expression on non-ES recipient cells is due to the translation of mRNA or due to the transfer of protein CD99. In this study, the expression of CD99 protein was only evaluated up to 20 h after exposure to ES exosomes, already seeing differences in the CD99 expression and a change in morphology suggestive of an apoptotic response. Similar to the gene fusion uptake experiments, the next step would be to investigate the changes in non-ES cells upon internalization of CD99 via exosomes, to confirm if this uptake is indeed triggering an apoptotic response and what are the pathways involved in this. Previous studies on ES exosomes reported the effect of exosomes from CD99-silenced cells into native ES cells, similar to the one induced by silencing the CD99 expression (Ventura et al., 2016), although this study was performed on ES cells only. Therefore, it would be interesting to see how CD99 could affect the

phenotype of recipient cells by comparing the data generated in this study with those from exosomes of CD99 knock down ES cells.

The presence of the 2 primary characteristic ES biomarkers (EWSR1-ETS gene fusion and CD99) in ES-derived exosomes might be exploited as circulating biomarker studies for the diagnosis and disease progression on ES patients. Profiling of exosomes from plasma of ES patients showed expression of CD99, in contrast to exosomes from healthy donor plasma, which did not express CD99. To my knowledge, this is the first study that evaluated the expression profile of CD99 on exosomes from ES plasma samples. A previous study on ES exosomes (Miller et al., 2013) showed that a panel of 5 mRNAs associated to ES, including the EWSR1-FLI1 fusion, were present in ES-derived exosomes and could be exploited for biomarker strategies, as they provided a sensitive and specific signal when spiked in blood of healthy donors. This is in agreement with another ES study (Tsugita et al., 2013) in which exosomes with the characteristic EWSR1-ETS fusion could be detected in circulation of ES xenograft mouse models. Therefore. the presence of these 2 ES biomarkers as part of the exosome cargo could be further evaluated in a clinical setting, in order to investigate their utility as non-invasive circulating biomarkers. This could be applied not only for diagnosis purposes, but for disease monitoring and response to treatment as well. To further investigate this possibility, exosomes isolated from plasma of ES patients could be further profiled, as only CD99 expression has been investigated in this study. This should include detection of the EWSR1-ETS fusion and to determine if a specific gene signature is present in ES-derived exosomes, as previously done for ES-derived exosomes from established cell lines (Miller et al., 2013). The identification of such profile in ES-derived exosomes might allow its application in distinguishing ES patients from different risk groups, as has been shown for ovarian cancer (Taylor and Gercel-Taylor, 2008). This could have important implications in patient stratification and personalised treatment strategies.

It has been more than a decade since functional mRNAs and miRNAs were found to be contained in exosomes and involved in the cell-to-cell communication through their transfer to other cells (Valadi et al., 2007). However, due to the enrichment of sRNAs these vesicles contain and the advantages to use them as circulating

biomarkers (Section 1.3.2), many groups are focusing on screening for sRNAs in order to fully elucidate the function they may have in tumour progression (Bortoluzzi et al., 2017; Wang et al., 2018). Several groups have described specific mRNAs in ES-derived exosomes (Miller et al., 2013; Tsugita et al., 2013; Ventura et al., 2016; Villasante et al., 2016) with only one group using a profiling approach consisting of a micro array covering 36,079 different RNAs of the human genome (Miller et al., 2013).

Therefore, this is the first time the exosome cargo from ES samples has been investigated by comprehensive transcriptomic analysis (RNA-seq). In agreement with results from other cancers (Jenjaroenpun et al., 2013; Schageman et al., 2013), I have found that exosomes have a different expression profiles to those of the producing cells from which they were generated. Interestingly, for the first time I have shown an enrichment of non-coding RNAs. Amongst the candidate genes identified to be enriched in ES exosomes compared to the producing cells, miR-320a and ADH1B are found amongst the top (positions 1 and 5 respectively). miR-320a has previously been reported to be a suppressor of invasion, cell proliferation and migration in nasopharyngeal carcinoma (Qi et al., 2014), in multiple myeloma (Lu et al., 2016) and in renal cancer (Zhao et al., 2018). In addition, exosomes from cancer-associated stromal fibroblasts were able to transfer miR-320a to hepatocellular carcinoma cells, inducing antitumour effects leading to a suppression of cell proliferation, migration and metastatic capacity both *in vitro* and *in vivo* (Zhang et al., 2017). In contrast, Wang et al. (2016) reported miR-320a to be up-regulated in pancreatic cells resistant to 5-Fluorouracil. Moreover, a recent publication on the profiling of different subsets of exosomes based on the expression of different surface markers showed enriched expression of miR-320a only in one subset of exosomes, suggesting a relationship between the exosomes subtype and the RNA cargo (Kaur et al., 2018).

The other candidate gene of interest, ADH1B, encodes for the ADH1B enzyme involved in the metabolism of ethanol and has been associated with tumorigenesis of alcohol-related cancers. Consistent with a potential tumour-suppressor role, defective ADH1B leads to increased cancer risks (Lee et al., 2008; Okada et al., 2018). Due to the young age of diagnosis of ES, the role of ADH1B does not

suggest a link to alcohol in this disease. However, a defective ADH1B passed from the parents could be linked to this profile. Although this connection between ADH1B and ES has not been reported in the literature, an association between ADH1B and ALDH, a known markers for ES-CSC cells, has been described in malignancies (Jelski and Szmitkowski, 2008), as both are implicated in the degradation of acetaldehydes. A recent transcriptomic investigation on several cancers identified ADH1B to be consistently downregulated, further evidence of its tumour-suppressive role (Li et al., 2017). Interestingly, higher expression of ADH1B is associated with a subset of soft tissue sarcomas that have lower probability to metastasise (Skubitz et al., 2014). Before validation studies on the role these genes could have in tumorigenesis, transcriptomic analyses should be done in a larger cohort of ES exosomes, as to obtain a more robust exosome-associated profile. It would also be necessary to include non-ES exosome samples to investigate if the RNA profile enriched in exosomes is ES-specific or rather indicative of a cancer-related phenotype. Once this is accomplished, future studies should investigate the role of these RNAs in recipient cells. Modification of the expression on producing cells (i.e. silencing miR-320a or ADH1B) should lead to a change in the exosomal cargo from these cells, and therefore, the differences in the profile of recipient cells from altered exosomes or wild type exosomes could be investigated.

The involvement of CSC in cancer progression has been associated to the transfer of exosomes in the tumour microenviroment (Rodríguez et al., 2015; Sun et al., 2018). Therefore, the presence of the NRXN1 ES-CSC marker identified in this thesis (Section 3.4) was investigated in ES-derived exosomes. Data from RT^A-qPCR and RNA-seq showed expression of this RNA in the exosome cargo of ES established cell lines and patient-derived cell cultures. Therefore, ES cells might be specifically packaging this RNA in order to be delivered to other cells and induce a CSC profile. This data could be linked to the increased self-renewing ability induced by the uptake of ES exosomes to non-ES cells. This is an important finding, as it agrees with the initial hypothesis of ES-derived exosomes being important for tumour progression. Therefore, further evaluation of NRXN1 RNA and the role it might have in changing the cellular behaviour of recipient cells should be

investigated. As suggested in Chapter 3, experiments in which this ES-CSC gene is silenced could inform on the role of NRXN1 RNA in ES cells and ES tumorigenesis.

In this study I have been able to monitor the uptake of ES-labelled exosomes by recipient cells both by IF and flow cytometry, as has been shown in other cancers (Franzen et al., 2014; Zhang et al., 2011). In addition, I have demonstrated that the uptake of ES-derived exosomes by cells is energy-dependent, as incubation of cells at 4 °C leads to a significant decrease in the uptake efficiency, as seen by others (Horibe et al., 2018). In agreement with published data (Escrevente et al., 2011; Horibe et al., 2018; Zhang et al., 2011), I have shown that the ES-exosome uptake process occurs in the first hours of incubation, with internalisation already observed after 1 h. Interestingly, an internal trafficking of exosomes in recipient cells was observed, as after 4–6 h incubation, labelled-exosomes left the surface of cells and formed defined clusters inside. Evaluation of the location of exosomes once internalised by cells showed no co-localisation with ER, Golgi nor mitochondria. Future studies should use a wider panel of cellular markers, specially of those postulated to be involved in the exosome uptake mechanism (i.e. endosome and lysosme) (El Andaloussi et al., 2013) to see if exosomes are sorted into these organelles and what mechanisms are implicated. To help with this aim, a new range of probes by Thermo Fisher Scientific, the LysoSensor probes, might provide useful information on this trafficking process. These probes are activated by low pH (a characteristic of organelles such as endosome and lysosome) and therefore, would become a powerful tool to investigate the internalisation of exosomes into these organelles upon uptake by cells. Moreover, inhibitors against the components associated to exosome biogenesis, such as Rab GTPases or ESCRT components (El Andaloussi et al., 2013; Essandoh et al., 2015; Kosaka et al., 2010) could be used, and the effects of their blockage investigated. This is especially important as different groups are considering the inhibition of exosome-associated mechanisms (production, release and uptake) in order to be exploited as therapeutic strategy (El Andaloussi et al., 2013; Fabbri et al., 2012; Horibe et al., 2018; Tian et al., 2014), and all these information in ES exosome biogenesis and uptake would help in understanding the different mechanisms involved.

Flow cytometry is a technique that allows for the identification of the size distribution and the expression of multiple markers, allowing for in-depth expression marker characterisation. However, as discussed in Section 4.3.4, due to small size of exosomes, this analysis is not as straightforward as would be when analysing cells. For this reason, different strategies are being used in order to analyse exosomes. One of the strategies that is used is the labelling of exosomes with fluorescent dyes or fluorescent-tagged antibodies so they are detectable based on the increase of intensity (Hoen et al., 2012; Pospichalova et al., 2015; Stoner et al., 2016). In this study I show for the first time that this technique allows the characterisation of ES-derived exosomes, not only on their protein expression profile of different markers (exosomal and ES), but on the quantification of the percentage of exosomes expressing each marker as well, based on the total exosome population that has been labelled with a generic dye (Section 4.3.4). Previous data on ES exosomes has investigated the protein expression profile by flow cytometry using latex beads, in order to increase the size of exosomes so they can be measured by conventional flow cytometers (Miller et al., 2013). Therefore, the method used in my study allows the characterisation of all exosomes, independent of their protein expression profile or ability to bind to the latex beads. This is an advantage compared to using a previously-described exosomal marker that could be only detecting a proportion of exosomes. However, this is assuming an homogeneous labelling of all exosomes with the generic dye, and therefore, does not take into consideration that different subpopulations of exosomes could be more or less efficiently labelled, and therefore, could be lost or underrepresented in the gated exosome population. Future developments in flow cytometry will be important to the exosome field, as exosomes are still near the 'noise' population. Therefore, more sensitive techniques that allow evaluation of smaller particles will provide the ability to gate exosomes by size only (as done with cells) rather than fluorescence, thus capturing all exosomes independent to their efficiency in taking up dyes.

For this study, the MBA method developed by Qiagen (exoEasy Maxi kit) was used for the isolation of exosomes, as it allowed isolation from both large volumes of cell culture media and small aliquots of plasma; the RNA extraction could be performed

directly on the membrane in which exosomes were captured; and did not require special equipment, providing an easy and fast methodology. Profiling of ES-derived exosomes using this method agreed with data from other exosome studies using the MBA technology (Enderle et al., 2015; Kim et al., 2017; Liang et al., 2018; Stolzenburg and Harris, 2018), consistent with the exosome characteristics described in the literature. To investigate the differences in isolation methods, the MBA technology was compared to two other isolation methods: a precipitation method and the gold-standard dUC (summarised in Table 4.1).

For the precipitation method, the PME kit from Analytik Jena was assessed. This technology looked promising when using small volumes of starting material, compared to the MBA technology, as a higher miRNA yield was recovered from equal starting volumes of cell culture supernatant. However, due to the nature of this study (large volumes of cell culture supernatant), the PME method was not convenient. If Analytik Jena were to develop an adaptation of the method for bigger starting volumes, I would assess again the output material, evaluating if they isolate the same exosomes populations, both by expression of surface markers and the RNA profile contained in the cargo. This is important as differences in profile have been described across different isolation methods used (Tauro et al., 2012; Van Deun et al., 2014). The other method to which the MBA technology was compared to was the widely-used dUC (Théry et al., 2006). Exosomes isolated through this method had the expected characteristics described in the literature. Interestingly, the difference in protein expression and yield between dUC and MBA were in agreement with previous studies in which higher tetraspanin yield was observed in dUC compared to other isolation methods, as well as a lower overall protein yield (Van Deun et al., 2014). A possible explanation for these differences in the protein yield could be that the dUC is enriching for tetraspanin-positive exosomes (i.e. $CD63^{+}$),and would explain why the overall protein yield is lower. However, MBA might be isolating these and other exosome populations, thus showing expression of different exosome-associated markers and higher protein yield. This would be corroborated by the NTA data obtained in my study, where the number of exosomes isolated by dUC is lower than for MBA (NTA data), as it might be that other exosome populations are not being pelleted by the

dUC method, and thus lost in the process. This explanation is in agreement with a previous study where a comparison between MBA and dUC showed that the later had a lower recovery rate indicted by lower exosome numbers (Enderle et al., 2015). Therefore, as mentioned in Section 4.1, it is important to decide the method of isolation for the specific study, as differences in the exosome profile and yield are seen across isolation methodologies, and these would have an important impact in downstream analyses. It is fundamental to characterise the isolated exosomes, independently on the method used, by different methodologies covering expression of exosomal markers, size and RNA cargo, in order to confirm the successful isolation of these vesicles.

5. Final discussion

The main findings of my thesis are:

1. Patient-derived cell cultures to identify drivers of disease and poor outcome
2. Identification of novel candidate markers of ES-CSCs
3. Implications of the heterogeneity across ES patients
4. Use of a functional assay to isolate ES-CSCs
5. Preclinical models to study ES biology
6. ES-derived exosome cargo as a circulating biomarker for ES
7. Uptake of ES-derived exosome cargo induces an ES-like phenotype in MSC and NBL cells
8. The importance of larger sample numbers

5.1. Patient-derived cell cultures to identify drivers of disease and poor outcome

In this study I have been able to identify drivers of disease and outcome using patient-derived ES cell cultures. This is important, and suggests that the *in vitro* profile obtained from studying cell cultures from a diverse range of ES patients can successfully identify markers that are associated to the clinical prognosis.

The top ES-CSC candidate marker identified in this study, NRXN1, is enriched in ES-CSCs compared to the cells from which they are derived. Interestingly, increased levels of protein expression of this marker in patient-derived cell cultures are associated with worse outcome when investigated in 36 ES tumours ($p = 0.008$, data kindly provided by Dr. Elizabeth Roundhill). Moreover, when this was validated in a public available ES dataset (GSE17618) containing 44 ES tumours (Savola et al., 2011), the same results were obtained ($p = 0.009$, data kindly provided by Dr. Elizabeth Roundhill). Therefore, higher levels of expression

of NRXN1, which suggests higher amounts of ES-CSCs cells within these tumours, are indicative of poor outcome.

The other phenotypic evaluations performed *in vitro* (self-renewing ability, migration index, sensitivity to chemotherapeutic drugs and CD99 expression) were not indicative of patient prognosis nor outcome (Section 2.3.9.2). As discussed already in Section 2.4, this lack of association of phenotypes associated to tumorigenesis to patient's outcome could be explained by the sample size used in my study ($n = 26$). Nevertheless, the ability to profile these patient-derived cell cultures *in vitro* and correlate the results with the clinical data of the patients allows for the survival evaluations with the aim in identifying predictive and prognosis signatures to stratify ES patients.

Moreover, using RNA-seq profiling of the patient-derived cell cultures to compare between ES patients who were alive or deceased at the time of this study (June 2018) allowed the identification of a panel of genes to be differentially expressed between both patient groups. Amongst them, higher expression of major histocompatibility complex genes was associated with better outcome, being consistent with their role in the immune system (Trowsdale and Knight, 2013). Therefore, using transcriptomic approaches to investigate the different risk groups could result in a panel of prognostic biomarkers. To date, the most informative prognostic indicators are associated to clinical information, such as tumour site, presence of metastasis at diagnosis; or copy number alterations in the karyotype (Gaspar et al., 2015; Grünewald et al., 2018). Therefore, identification of molecular markers associated to prognosis could help categorise ES patients into different risk groups, giving different treatment accordingly.

Altogether, this data indicates that the patient-derived cell cultures are a good model to investigate drivers of ES disease, as validation studies using ES tumours have shown the same results as seen in these cell cultures. Therefore, studies like this will help identify potential prognostic and predictive biomarkers that could provide future patient stratification strategies with the aim to implement personalised medicine. This is especially important in ES, where a high proportion of patients are refractory to treatment or eventually relapse (Gaspar et al., 2015).

5.2. Identification of novel candidate markers of ES-CSCs

To my knowledge, this study was the first that profiled CSCs isolated from ES samples using RNA-seq. One of the aims of my work was to investigate the profile of the isolated ES-CSCs in comparison to ES cells from which they were derived, in order to identify differentially expressed genes that could be used for targeted strategies against ES-CSCs. Previous studies in ES-CSCs have investigated expression of stemness genes (Awad et al., 2010; Cornaz-Buros et al., 2014; De Vito et al., 2012; Riggi et al., 2010) or ES-associated genes and previously reported CSC markers (Riggi et al., 2010) by PCR, as well as array expression of the full genome (Riggi et al., 2010) or all the miRNAs known (De Vito et al., 2012).

When the 13 ES-CSCs were compared to the 5 ES patient-derived cell cultures from which they were derived, a total of 347 genes were differentially expressed, with 6 candidate genes enriched in ES-CSCs (Section 3.3.5). None of these genes have been reported to be markers of ES-CSCs in previous ES studies, confirming that they are novel ES-CSC markers. To further validate these results, the expression profile of the 13 ES-CSCs was compared to the 26 patient-derived cell cultures (irrespective of producing or not ES-CSCs), in order to take into consideration the heterogeneity across all ES samples and identify if these markers could define an ES-CSC from the rest of the ES cell population. A total of 973 differentially expressed genes were identified in this comparison, further evidencing that differences exist between the isolated ES-CSCs and the patient-derived cell cultures. When pathway analysis was performed on the enriched genes on ES-CSC and patient-derived cell cultures respectively, the same functional groups enriched in ES-CSC (n = 5, associated to glycosilation and signalling) were enriched in patient-derived cell cultures as well (n = 26) amongst other functions such as extracellular-associated machinery. Interestingly, 2 of the novel ES-CSC markers (NRXN1 and ELFN2) were present in the top ranks for both evaluation (ES-CSC vs. paired patient-derived cell cultures or all patient-derived cell cultures). Therefore, this strategy allows the identification of markers of ES-CSCs populations across the heterogeneity of the samples investigated, suggesting that they are strong candidates.

An increased expression of NRXN1, a member of the neurexin family involved in synapsis processes, has previously been identified in ES cells (Tanaka et al., 2014). The other candidate gene, ELFN2, although not previously linked to ES biology, has been associated to an aggressive form of gastric cancer (Zhang, 2017). Interestingly, both these genes and LINGO1, a candidate diagnostic biomarker in the ES field (Jain et al., 2017; Town et al., 2016), are members of the synaptic adhesion molecules (SAMs) important in synapsis of mammalian brain, which have been shown to be linked to neurophysiological and neurodevelopmental diseases (Rudenko, 2017). However, expression of LINGO1 was not differentially expressed in ES-CSCs but could be detected in the same levels than ES patient-derived cells, suggesting that LINGO1 expression would not identify the ES-CSC population. Moreover, 6 other SAM where identified across the differentially expressed genes between ES-CSC and ES ($p < 2 \times 10^{-5}$). Therefore, this link between the 2 candidate ES-CSC genes (and another 6 in the list) and LINGO1, which has been shown to be enriched in ES cells (Town et al., 2016), suggests a relationship that could be important for ES biology. Interestingly, SAMs are implicated in cellular adhesion and interactions with components of the extracellular matrix and cellular surface (Rudenko, 2017). Therefore, this could suggest that ES-CSCs have an enrichment for adhesion-associated proteins which might regulate their adaptation and survival in the tumour mass and interact with neighbouring cells. In order to fully validate these markers as candidates for targeted therapies and their role in ES-CSC biology, the expression in other tissues should be investigated, to evaluate possible off-target toxicities and if the enrichment is particular of these cells. As SAMs are important proteins for the brain synapses (Rudenko, 2017), I expect that high levels of expression would be present in brain tissue. Therefore, novel targeted strategies should take this into consideration and develop treatments that do not cross the blood-brain barrier, and thus prevent toxicities.

Future studies should validate the importance of these candidate genes in the ES-CSC profile, in order to elucidate if their expression is a driver of a self-renewing cell population. To investigate this, these genes could be overexpressed in patient-derived cell cultures to see if this leads to enhanced self-renewing ability (characteristic of CSCs), and therefore, would suggest their involvement in the

ES-CSC profile, and not just be a surrogate markers. Moreover, similar to the data from NRXN1 predicting outcome (Section 5.1), expression of these other 5 ES-CSC candidate markers should be validated using other approaches (i.e. protein expression) and investigate if they are predictors of outcome.

It is suggested that CSCs are behind treatment failure and relapse in many malignancies (Visvader and Lindeman, 2012), including ES (Trucco and Loeb, 2012). Therefore, the ability to identify candidate markers enriched in ES-CSCs isolated from patient-derived cell cultures is an important step towards the development of targeted therapies against this cell population, which might ultimately lead to better outcomes in ES patients.

5.3. Implications of the heterogeneity across ES patients

ES is an heterogeneous malignancy. Patients can present with metastatic disease at diagnosis, be refractory to treatment or initially respond but eventually relapse (Gaspar et al., 2015). In addition, differences across the molecular profile have been identified, including various fusion types, presence of cancer-associated mutations and increased levels of important cellular pathways (Grünewald et al., 2018). This heterogeneity across patients makes it challenging for the development of new treatment strategies.

Different DNA mutations and alterations in the expression levels of different genes have been associated to ES biology and prognosis (*TP53, STAG2, CDK4* and *CDK6*) (Huang et al., 2005; Tirode et al., 2014). When the expression levels of these genes was investigated in ES patient-derived cell cultures and ES-CSCs (data kindly provided by Dr. Elizabeth Roundhill), no difference in the RNA expression level were observed (fold change < 2), suggesting that these events are not associated to the ES-CSC profile. Super-enhancers, regions of open chromatin that can induce increased transcription factor of genes (Hnisz et al., 2013), have become popular in the ES field, as they have been associated to the ES signature regulating important signalling pathways in this malignancy (Kennedy et al., 2015; Lin et al., 2019; Tomazou et al., 2015). However, the expression of 8 previously-reported

super-enhancers was not different between ES patient-derived cell cultures and ES-CSCs (fold change < 1.7, data from Dr. Elizabeth Roundhill), thus suggesting that these effectors are not behind the ES-CSC profile. Therefore, the differences in the expression profile between ES-CSC and the cell cultures from which they are derived are not driven by previously-known genes associated to ES, thus suggesting an extra layer of heterogeneity and complexity in this disease.

Using RNA-seq, 8 novel EWSR1 fusions were identified in ES patient-derived cell cultures, comprising both 3' EWSR1 translocations and reciprocal translocations (Section 3.3.2.4). To my knowledge, none of these fusions have been previously identified in ES, and therefore, could be novel fusions increasing the list of rare forms of EWSR1 fusion variants (Romeo and Tos, 2010). Future studies should validate the RNA-seq findings using a more sensitive technique, such as PCR, in order to confirm the presence of these fusions in ES samples. Moreover, it is important to understand if these represent drivers of the ES phenotype or are by-standers, thus not having an important role in the biology of the tumour. These results will inform if these fusions represent an extra layer of heterogeneity in ES samples, as they could be associated to different phenotypic profiles.

Moreover, *in vitro* characterisation of 39 ES patient-derived cell cultures identified differences in the migration index, self-renewing ability, sensitivity to chemotherapies and expression of CD99 and c-kit (Section 2.3), where no correlation was seen between the different profiles. A similar result was obtained when investigating the phenotype of ES-CSCs *in vitro*, as heterogeneity in the migration index, CD99 and c-kit expression were shown across ES-CSCs derived from the same cell cultures and between different cell cultures (Section 3.3.1). Therefore, it is important to investigate if any of the phenotypes observed across patients could be exploited for patient stratification and personalised medicine.

Due to the poor outcome associated to ES and the lack of targeted therapies to treat ES patients, many groups are evaluating new compounds for the treatment of ES (Amaral et al., 2015; Brenner et al., 2012; Sankar et al., 2014), some of which have reached clinical trials (Baruchel et al., 2012; Bond et al., 2008). Based on the data from my study, it is important to investigate the differences across patients, as a specific ES subgroup associated to a specific profile could benefit from one

treatment strategy over another. Therefore, future studies should evaluate the differences across patients with the aim to identify novel treatment approaches designed for each patient profile, leading to personalised medicine strategies. This might ultimately lead to better treatment response and less therapy-related toxicities, as patients would be given the most adequate treatment protocol for their tumour profile.

5.4. Use of a functional assay to isolate ES-CSCs

I have hypothesised that ES-CSCs play an important role in ES progression and relapse, consistent with their increased resistance to conventional chemotherapy and self-renewing ability, leading to tumour growth (Schatton et al., 2009). For this reason, ES-CSCs have been isolated from ES patient-derived cell cultures. A better understanding of their profile and identification of CSC-associated drivers could result in novel targeted therapies against these cells.

In this study I have used a functional assay that exploits an intrinsic characteristic of CSCs, the ability to self-renew. Using this strategy, I have been able to identify a subset of cells in ES samples that can self-renew and have higher resistance to chemotherapies (Section 3.3.1). Using these ES-CSCs, I have been able to identify candidate genes differentially expressed in ES-CSCs compared to the cells from which they are derived (Section 5.2), providing a good method for the profiling of such a cell population within tumours.

Previous studies in the ES field have used the expression of cell surface makers associated to a CSC profile to isolated putative ES-CSCs from ES samples (Cornaz-Buros et al., 2014; De Vito et al., 2012; Riggi et al., 2010; Suva et al., 2009; Wahl et al., 2010). Other ES-CSC studies have exploited mechanisms linked to the CSC profile, such as increased activity of ALDH (Awad et al., 2010) or higher drug efflux (Komuro et al., 2007; Yang et al., 2010), which are surrogate markers of their increased resistance to drugs. Although all these studies successfully identified a subset of ES cells with CSC characteristics, such as increased tumorigenesis or higher drug efflux, a lack of correlation across the different proposed ES-CSCs markers has been described (Jiang et al., 2010; Leuchte et al.,

2014; Wahl et al., 2010). This would suggest that these strategies are not isolating the overall ES-CSC population, but rather a subset of ES-CSCs.

One of the most widely debated markers of CSCs is CD133, which has been used for the isolation of CSCs across different cancers (Irollo and Pirozzi, 2013). Importantly, this was the marker used to isolate ES-CSCs for the first time (Suva et al., 2009), and has continued to be used in other ES-CSC studies (Cornaz-Buros et al., 2014; De Vito et al., 2012; Riggi et al., 2010; Wahl et al., 2010). However, data from other malignancies has identified $CD133^-$ to have a similar profile than $CD133^+$, which are supposed to be the CSCs populations. Wang et al. (2012) showed that $CD133^-$ colorectal cancer cells had an increased invasive potential, suggesting a profile associated to CSCs, contrary to the hypothesis of $CD133^+$ cells to be CSCs. Similarly, $CD133^-$ has been shown not to be a good marker for the identification of CSCs in melanoma, as successful isolation of cells with CSC phenotype was obtained through spheroid formation ability rather than CD133 expression, seeing differences between sphere-derived and $CD133^+$ cells (Fomeshi et al., 2015). Therefore, these observation in other cancers could explain the discrepancies in ES-CSCs studies, where all groups have isolated a population with CSC characteristics, as different isolation methods do not correlate between them. This suggests that the markers exploited in ES studies are isolating different subsets of ES-CSCs, and therefore, the results obtained might not be shared across all ES-CSCs within a tumour. When the levels of CD133 were investigated in my RNA-seq dataset, higher levels expression were seen in patient-derived cell cultures than in ES-CSC (fold-change 14), suggesting that there is no association between the isolated ES-CSCs and expression of this previously-postulated ES-CSC marker. Therefore, using a functional assay exploiting an intrinsic CSC characteristic provides a better model to identify and profile ES-CSCs rather than markers that could be only isolating a subset of these cells.

When evaluating if ES-CSCs conserved the ability to self-renew from a single cell after being isolated and maintained *in vitro*, a loss of this phenotype was observed (Section 3.3.1.4). A possible explanation for this change of phenotype is that ES-CSCs are slower growing than the patient-derived cell cultures, and therefore, longer periods of time (> 3 weeks) might be needed to detect the colonies formed from

a single cell. Another possibility is that after continued culture *in vitro*, ES-CSCs lose the capacity to grow from a single-cell. The microenvironment has been shown to play an important role in cancer progression and modulation (Xu, 2018), and therefore, it could be that the absence of factors from either cells present in the body or from other ES cells leads to a change in their profile. To overcome this, future studies should investigate the self-renewing ability in conditions more similar to the body, such as scaffolds recreating the tumour microenvironment or investigating if they have an enhanced ability to form tumours in xenograft mouse models, as has been done in other ES-CSC studies (Awad et al., 2010; Cornaz-Buros et al., 2014; De Vito et al., 2012; Riggi et al., 2010; Suva et al., 2009; Wahl et al., 2010).

Although the functional assay used in this study provides a powerful method to isolate the overall ES-CSC population, some limitations are associated to this approach. By isolating ES-CSCs based on their ability to self-renew from a single cell *in vitro*, we could be losing a subset of these cells that lack the ability to propagate in culture, as they might need extrinsic factors to grow. As suggested before, this assay could be adapted to scaffolds in order to increase the success in isolating ES-CSCs by better mimicking the *in vivo* conditions (Ent et al., 2018; Marturano-Kruik et al., 2018; Santoro et al., 2015). Moreover, by exploiting the self-renewing ability, only cells that have been able to grow from a single cell are obtained. Therefore, what would be the negative populations (cells that cannot self-renew) is lost and cannot be used for comparative evaluations. This leads in ES-CSCs being compared to the overall ES cell populations, which will contain ES-CSCs within, thus biasing slightly the results.

In order to overcome both limitations, the results obtained in this study could be employed to generate more informative profiles. To do this, the strategy used by Town et al. (2016) to investigate the surfaceome of ES cells could be used. In order to do that, the surfaceome of successfully isolated ES-CSCs should be investigated in order to define a panel of markers that can selectively identify these cells. With that, cell sorting analysis on the ES patient-derived cell cultures would result in the separation of the ES-CSC^{+} fraction versus the ES-CSC^{-} fraction, thus resulting in more informative expression profiles when comparing both datasets. This might result in a deeper understanding of the ES-CSC profile compared to the overall ES

cell populations, helping in the development of novel treatment strategies and potential prognostic biomarkers.

5.5. Preclinical models to study ES biology

There is an urgent need for better treatment strategies for ES patients, as chemotherapeutic approaches have not changed much in the last decades, thus not improving patient outcomes (Balamuth and Womer, 2010). For this reason, preclinical models should be the most similar to the tumour's patient biology in order to provide reliable results and move forwards in terms of treatment success.

In this study I have used patient-derived cell cultures originated from ES tumours, which were obtained from patients with different clinical profiles. Using these cell cultures, I have been able to study the aggressiveness of these cells *in vitro* (Section 2.3), evaluate putative prognostic markers (Section 2.3.9.1) and isolate ES-CSCs (Section 3.3.1), suggesting that they are a good model to investigate ES biology.

RNA-seq profiling of patient-derived ES cell cultures and ES established cell lines indicated that a different gene expression profile was seen across both sample groups (Section 3.3.3). Moreover, ES established cell lines had a more similar profile to a NBL cell line than the patient-derived ES cell cultures. This data suggests that what is driving the clustering between established cell lines and patient-derived cell cultures is independent of ES biology, as if that was the case, we would expect ES established cell lines to cluster with the ES patient-derived cell cultures rather than the NBL cell line. Moreover, pathway analysis confirmed this hypothesis, as established cell lines were enriched for genes associated with cell cycle and DNA replication, whereas patient-derived cell cultures had mainly enrichment of genes associated to glycosilation and signalling. Therefore, based on the data in this study, we can further confirm that established cell lines are not a good model to study ES, as extensive culturing has resulted in a drift in their expression profile from the original ES profile present in patient samples.

These results are consistent with previous data from the ES field (Crompton et al., 2014), where a difference in the gene expression profile was identified between ES

tumours and established cell lines. Moreover, principal components analysis showed that the two principal components driving the majority of the differences across the investigated samples resulted in the independent clustering of established cell lines versus tumours (Crompton et al., 2014), similar to what I have seen in my study. Moreover, Brohl et al. (2014) investigated the expression levels of ES-associated genes, such as CDKN2A or TP53, and saw higher expression of these genes in established cell lines than in tumour samples. A similar observation was seen in this dataset, where ES established cell lines had higher levels of expression of ES-related genes such as NKX2-2, CDKN2A or TP53. This is an important observation, as many studies are still using established cell lines, some of them isolated from patients 30 years ago (Teicher et al., 2011) to identify novel targeted therapies for ES patients. Therefore, these differences in the expression profile could lead to false positive results, that when moved from preclinic evaluations to clinical trials, could result in no benefit for ES patients.

Therefore, future investigations with the aim to identify novel targeted therapies and better understand the ES biology should be performed in informative preclinical models, such as patient-derived cell cultures. One of the future studies that could continue from my research is the identification of novel treatment approaches. Using the RNA-seq data obtained on the patient-derived cell cultures, a search for targets of available inhibitors could be performed, with the aim to identify those targets that are enriched in ES cells and therefore, could be exploited for targeted therapies. From this identification, inhibitors against these proteins could be evaluated in the patient-derived cell cultures to establish if they are sensitive to these strategies. Results from these analyses might then provide a better understanding of the response that the patients might have in the clinic, rather than using established cell lines.

It is important to remember, however, that although ES patient-derived cell cultures are informative of the patient profile, they are still 2D monolayers, and therefore, do not grow in the same conditions they would do in the patient's body. Recent evidence has shown that scaffolds recreating the microenvironment and cell-to-cell interactions provide a better model for cancer biology. These systems, using biomaterials and 3D structures, allow for interactions between cells

(compared to 2D monolayer culture), the co-culture with other cell types to recreate the tumour microenvironment and induce release and uptake of factor between cell types, and even the addition of microfludics and shear stress to mimic the blood pressure in the human body (Ent et al., 2018; Marturano-Kruik et al., 2018; Santoro et al., 2015). A recent study by Villasante et al. (2016) showed that using tissue-engineered scaffolds resulted in a phenotype and expression profile more similar to the *in vivo* ES models than 2D monolayers. Therefore, a future strategy might be to place these patient-derived cell cultures in scaffolds in order to study drug response and aggressiveness. Another option would be to obtain PDX from these cells, as these have been shown to be good preclinical models of ES, resulting in the ability to test the treatment response of ES cells (Dowless et al., 2018; Guenther et al., 2018).

5.6. ES-derived exosome cargo as circulating biomarker for ES

In this study I have optimised a protocol for the isolation and characterisation of exosomes from cell culture media and plasma samples using a membrane-based affinity method. This has allowed the study of ES-derived exosomes from patient-derived ES cell cultures, ES-CSCs, plasma from ES patients and healthy donors and established ES cell lines. To my knowledge, this is the widest exosome profiling from patient-derived ES samples to date.

ES exosomes from patient-derived cell cultures, ES-CSCs and established cell lines express CD99 on their surface. This is in agreement with data from Ventura et al. (2016), where CD99 expression was reported in ES-derived exosomes from an ES established cell line. The expression of CD99 is used for the diagnosis of ES, as it is highly expressed in this malignancy (Bernstein et al., 2006). To test if ES-derived exosomes could be exploited for circulating biomarker strategies in ES, I investigated the expression of CD99 on exosomes from plasma of ES patients and healthy donors, and found that exosomes in blood circulation of ES patients express CD99, in contrast to exosomes from healthy donors. Moreover, CD99 RNA was detected in the cargo of exosomes by RNA-seq. Other studies in the exosome field have shown that functional mRNAs are packed in exosomes, and once internalized

into recipient cells, they can be translated (Valadi et al., 2007). As CD99 expression has been associated to migration capacity and proliferation in ES (Manara et al., 2018), future studies should investigate if CD99 mRNA contained in exosomes is functional and therefore, could induce a change in phenotype when internalised by recipient cells. If confirmed, this would indicate a mechanism by which ES cells could be modulating other cells in the tumour microenvironment.

To further confirm if ES-derived exosomes could be exploited for circulating biomarker strategies, the presence of the EWSR1-ETS fusion gene, the other ES diagnostic biomarker, was investigated. RT^A-qPCR showed that EWSR1-FLI1 fusion mRNA was present in the cargo of SK-N-MC and TC-32 ES established cell lines, which are characterised for having this fusion. This is in agreement with data from previous studies, where EWSR1-FLI1 fusion RNA has been identified in exosomes from ES cell lines (Miller et al., 2013; Tsugita et al., 2013). For the first time, I have identified the EWSR1-ERG fusion RNA in the cargo of ES-derived exosomes from an EWSR1-ERG ES established cell line, confirming that the presence of the fusion in the cargo is independent to the fusion type. To validate these results, the presence of the EWSR1-ETS fusion RNA should be investigated in exosomes from ES patient plasma, as has been done for CD99 in this study, and compare their profile with healthy donor plasma. Other groups in the ES field have suggested similar strategies, as they have been able to detect ES-derived exosomes (containing EWSR1-ETS fusion RNA) in circulation of ES mouse xenograft (Tsugita et al., 2013) or spiked in plasma of healthy donors (Miller et al., 2013). Therefore, the identification of both ES diagnostic biomarkers in ES-derived exosomes suggests the potential role these vesicles could have in defining novel non-invasive circulating biomarker strategies for ES diagnosis and disease monitoring. Similar strategies that do not involve exosomes are being investigated, such as the detection of circulating tumour cells expressing CD99 (Benini et al., 2018) or the EWSR1 fusion in circulation (Schleiermacher et al., 2003). Therefore, we could be near the implementation of novel circulating biomarker strategies for ES, which could result in more effective treatment response monitoring and disease follow up.

Another important observation of this study is that exosomes and the cells from which they are derived have a high RNA expression overlap (78% of identified RNAs are present in cells and exosomes), thus suggesting that exosomes have a profile that inform on the producing cell. This is consistent with the detection of CD99 and EWSR1-ETS fusion RNA in the exosome cargo. Moreover, in this study I have shown that the novel candidate of ES-CSCs NRXN1 RNA (Section 5.2) is present in the cargo of ES-derived exosomes. Validation of the role of NRXN1 proposed in Section 5.2 could help determine if this expression could be exploited for prognostic biomarker strategies, as expression of NRXN1 in tumours is associated with outcome (Section 5.1). Therefore, this high degree of expression overlap between cells and exosomes and the identification of a specific ES-associated signature could be exploited for prognostic and diagnostic biomarker strategies. In a previous study on ES-derived exosomes, Miller et al. (2013) identified a panel of 12 mRNAs associated to ES biology that could be detected in ES-derived exosomes from established cell lines. Interestingly, 5 of these mRNAs (NRB01, NKX2-2, STEAP1, LIPI and the EWSR1-FLI1 fusion) could be used for a sensitive and specific detection of ES exosomes spiked in plasma from healthy donors. When this panel of markers was investigated in the RNA-seq profile obtained in my study, NRB01, NKX2-2 and LIPI were mainly expressed in SK-N-MC-derived exosomes and not detected in patient-derived exosomes. Only STEAP1 was identified across the 4 exosome samples sequenced. Therefore, although such a panel becomes an exciting strategy for ES, it is important to evaluate the expression levels of these markers in exosomes from patients, rather than using established cell lines.

Future studies should evaluate the RNA-seq profile obtained in this study to identify a subset of RNAs that could be exploited for diagnostic biomarker strategies. Such a panel should inform on the tumour expression profile and discriminate from other malignancies, and therefore, would need to be validated in exosomes from ES plasma samples and other malignancies or healthy donors. Moreover, using prognostic indicators identified in this study (Section 5.1) or in other studies, it might be possible to define a panel of RNAs contained in exosomes that can predict the treatment response and the prognosis of the patient. This has already been

described in exosomes from OST (Xu et al., 2017), where a panel of miRNAs and mRNAs in the cargo of OST-derived exosomes from plasmsa can be used as predictive biomarker of response to treatment. Therefore, the exosome cargo becomes a candidate non-invasive circulating biomarker, which might inform on diagnostic, predictive and prognostic factors. This could result in better patient stratification, personalised medicine and more effective disease monitoring.

5.7. Uptake of ES-derived exosome cargo induce an ES-like phenotype in MSC and NBL cells

In this study I have shown that ES exosomes can transfer CD99 protein and EWSR1 fusion DNA to MSC and NBL cells, and increase the self-renewing ability of these upon internalisation of the exosome cargo. Exosomes have been shown to be important in cellular cross-talk and microenvironment modulation (El Andaloussi et al., 2013). Moreover, in other cancers, exosomes can induce tumour growth (Hu et al., 2015) or anti-tumour immunity (Zhang et al., 2011) depending on the cell of origin; or regulation of target genes in normal cells (Tian et al., 2014). In a previous study in ES, exosomes from a CD99-silenced ES cell could induce a change of phenotype similar to when directly silencing CD99 in recipient cells, thus suggesting the ability of exosome cargo to modulate ES cellular behaviour (Ventura et al., 2016).

Exosomes derived from a high self-renewing ES cell line, upon incubation with MSC and NBL cells, induced an increase in the self-renewing ability of these cells. The self-renewing ability is associated with the CSC profile (Schatton et al., 2009), and ES-CSCs have been postulated to be the reason behind tumour progression and relapse in ES (Trucco and Loeb, 2012). This is the first time that the effects of ES exosomes on non-ES cells have been investigated, as previous studies have only evaluated the uptake of specific markers by investigating the change in the expression profile (Tsugita et al., 2013; Villasante et al., 2016). My results are consistent with data from other cancers, where exosomes from CSCs can induce a transformation towards a CSC profile in recipient cells (Hu et al., 2015; Rodríguez et al., 2015). Therefore, this data suggests that ES exosomes could induce a change in the cellular phenotype, from less self-renewing to more self-renewing

cells, postulating a putative mechanism of cancer cells to induce changes in the neighbouring cells and lead to more aggressive phenotypes.

As discussed before (Section 5.6), the ES-CSC marker NRXN1 was present in the cargo of exosomes. In addition, RNA-seq evaluation of the ES exosome cargo identified 392 RNAs that were differentially increased in exosomes compared to the cells from which they were derived (Section 4.3.9.2), with an enrichment for sRNAs, in agreement with other RNA profiling studies of exosomes (Jenjaroenpun et al., 2013). Amongst these RNAs, genes previously linked to cancer were found, such as the miRNA miR-320a or ADH1B. Both have been associated to malignant phenotypes (Lu et al., 2016; Okada et al., 2018; Qi et al., 2014; Skubitz et al., 2014; Zhao et al., 2018). Therefore, future studies should evaluate the role of these RNAs contained in exosomes and if the transfer to recipient cells induces a change in the phenotype, such as the increased self-renewing ability observed in MSC and NBL cells.

For the first time, I have shown that ES exosomes can inducing an increase in the expression of CD99 and the transfer of EWSR1 fusion gene to recipient MSC and NBL cells. Although the presence of both markers has already been described in exosomes (Miller et al., 2013; Tsugita et al., 2013; Ventura et al., 2016), this is the first time the transfer to non-ES cells has been identified. Tsugita et al. (2013) evaluated if labelled-EWSR1 RNA from ES cells could be internalised by other ES or OST cells. Their results indicated that the fusion could only be transferred to ES cells, as no detection of labelled-EWSR1 fusion was seen on OST cells. Moreover, my results are consistent with work by Villasante et al. (2016), in which ES exosomes could be taken up by MSCs, evaluated by an increase in the fluorescence due to the transfer of EZH2-labelled mRNA via ES exosomes. Therefore, my study increases the current knowledge on ES cross-talk to non-ES cells, suggesting that ES cells might be using the transfer of exosomal cargo to induce a more aggressive profile in neighbouring cells.

In my study, however, only the transfer of these 2 biomarkers has been evaluated. Characterisation of exosomes (discussed in Section 5.6) indicated that CD99 is present in protein and RNA form. Therefore, future investigations should assess if the increase in CD99 expression after uptake is due the translation of CD99 RNA in

recipient cells, thus confirming the mRNA is functional; or is due to the transfer of the CD99 protein only. Moreover, expression of EWSR1-FLI1 has been shown to induce an ES-like profile when expressed in MSC (Riggi et al., 2008) and NBL cells (Rorie et al., 2004). In my studies I have only shown that EWSR1-fusion (DNA) is transferred to non-ES cells. It would be important to investigate if the mRNA (present in the exosome cargo) is transferred to recipient cells as well, and if so, this leads to the translation into EWSR1-FLI1 protein, confirming it's functionality. Moreover, future studies should evaluate if the fusion's uptake via exosomes by MSC and NBL cells is enough to induce a transformation on the cellular molecular machinery, mimicking the transformation induced by direct expression of the fusion (Riggi et al., 2010, 2008; Rorie et al., 2004). To evaluate this, non-ES cells should be incubated with ES exosomes, and evaluate over time the expression of EWSR1-FLI1 RNA and protein and the changes in downstream signalling pathways. This same strategy might identify the processes driving the increase in self-renewing ability upon ES-derived exosome incubation. These studies might help understand the mechanisms of ES tumour progression and the ability of ES cells to modulate the microenvironment, which could ultimately result in the identification of putative targeted therapies against these processes.

The ability of ES exosomes to induce changes in the migration capacity of non-ES cells was investigated, and contrary to the changes seen in the self-renewing ability, ES exosomes could not modify the migration capacity upon uptake by NBL cells. A change in the migration profile has been reported in other cancers upon incubation with exosomes (Chowdhury et al., 2015; Pan et al., 2017; Zheng et al., 2018). However, these other studies have used the transwell migration assay (Pan et al., 2017; Zheng et al., 2018) or the scratch assay (Chowdhury et al., 2015) in order to evaluate the effect of exosomes upon recipient cells. Therefore, cells being in a 2D monolayer, they assure that most if not all cells will be exposed to exosomes. However, in my study I have evaluated the change in migration in recipient cells by incubating spheroids formed by the cell type under study with exosomes. This raises an important caveat, as only cells in the outer layers of the spheroid will be in direct contact with ES-derived exosomes, and I did not evaluate if exosomes could internalise and send their cargo into the inner layers of the

spheroid. Another consideration is that the other studies have used exosomes from cells that have already been found to induce an increase in the migration and invasive profile, as are prostate cancer-derived exosomes into MSCs (Chowdhury et al., 2015), exosomes from gastric cancer cells containing an lncRNA involved in metastasis (Pan et al., 2017) or tumour-associated macrophages into gastric cancer (Zheng et al., 2018). Therefore, it could be that exosomes from SK-N-MC cells (used for the migration assay) lack the ability to transform NBL cells used in my study, or that the latter are resistant to the effect induced from ES exosomes. Future studies should therefore use exosomes from a variety of ES cells, such as patient-derived cell cultures with high migration index, or evaluate the change in profile in a wider set of recipient cells. Another possibility would be to add ES-derived exosomes into models that mimic better the tumour microenvironment, such as scaffolds with ES cells or PDX, in order to see if the presence of ES exosomes induces an increase in the migration and metastasis, as has been identified in other malignancies (Kucharzewska et al., 2013; Tickner et al., 2014; Zheng et al., 2018).

Another important factor in which exosomes are involved is in drug resistance (Jiang et al., 2018), as they can modulate the recipient cells towards a more chemoresistant profile. In the literature, different groups have described the ability of exosomes derived from resistant cells to be able to induce a change in the sensitivity to chemotherapeutic drugs when taken up by sensitive cells (Lu et al., 2013; Lv et al., 2014; Sousa et al., 2015; Torreggiani et al., 2016). However, in this study, when exosomes from a chemoresistant ES cell line where incubated with sensitive ES cells, no change in the response to doxorubicin was seen (Section 4.3.18). A possible explanation for this discrepancy of results is that only 2 diferent incubation times with exosomes were evaluated: 24 and 48 h. Therefore, it could be that cells do not have enough time to transform their phenotype towards a more resistant profile, thus not seeing an effect upon doxorubicin treatment. For instance, the study of Lv et al. (2014) incubated exosomes with cells for 72 h prior to treatment with chemotherapeutic drug, seeing an increase in the resistance to docetaxel, the chemotherapy evaluated. Therefore, future studies should aim to investigate longer exosome exposure times prior to treatment with doxorubicin, in

order to see if this increase in resistance is dependent on time. Similarly than for the changes on the migration index, a broader panel of ES cells-producing exosomes and sensitive recipient cells should be evaluated, as it could be that the signals send by the cells used in this study are not strong enough to lead to changes in the chemoresistance of recipient cells.

It is important to mention that this study was based on the changes ES-derived exosomes could induce in non-ES cells or more chemosensitive cells, in order to investigate their putative role in tumorigenesis. However, exosomes from normal tissue cells have been shown to modulate cancer cells towards a less oncogenic state, thus being exploited for therapeutic approaches (Romagnoli et al., 2015; Théry et al., 2009; Zhang et al., 2011). Such an strategy has resulted in exosomes being included in a phase II clinical trial for NSCLC patients (Besse et al., 2016), where dendritic cell-derived exosomes are administered alongside the standard chemotherapy, seeing an improvement in response via the activation of an anti-tumour response. Therefore, it could be interesting to evaluate if exosomes from normal cells (i.e. immune system) can reverse the oncogenic profile of ES cells upon incubation. This might result in the identification of novel companion immunotherapies to be used alongside chemotherapy, with the aim to improve treatment efficacy in ES patients.

A limitation of my studies in exosomes is that the mechanisms of production and uptake of ES exosomes have not been investigated. Although there is still some uncertainty towards the specific pathways involved, different mechanisms have been identified in cells for exosome production (Baietti et al., 2012; Henne et al., 2012; Trajkovic et al., 2008) and uptake (Mulcahy et al., 2014; Munich et al., 2012; Svensson et al., 2013; Tian et al., 2014). In this study I have shown that ES, NBL and MSC cells can uptake ES-derived exosomes, thus suggesting that the mechanism behind the uptake is not ES-specific. To date, the production of ES exosomes has not been described in the literature, and therefore, future studies should aim to elucidate the specific mechanisms behind this process. This is important as different groups are exploiting the inhibition and modification of exosome biogenesis, release and uptake processes as therapeutic strategies (El Andaloussi et al., 2013; Fabbri et al., 2012). Therefore, understanding the

pathways used by ES cells to produce, release and uptake ES exosomes might help in identifying putative treatment strategies.

Therefore, future studies should aim to identify the exact mechanism of ES-exosome uptake, as this might lead to a better understanding of the tumorigenesis process in ES, which might ultimately result in targeted strategies against these mechanisms. These studies should be performed in a variety of patient-derived cell cultures, as these samples will provide a better model to study heterogeneity across ES patients and identify if the mechanisms behind ES exosome production and uptake are ES specific or change across patients. Altogether, further understanding of these processes might help investigate what happens to CD99 and EWSR1-ETS fusion once these are internalised into recipient cells. Such information, although specific for ES biology, could help the general cancer field in understanding how exosomes can induce changes in recipient cells.

In summary, the findings from this study reveal the importance exosomes play in modulating the cellular behaviour of recipient cells and how they can induce oncogenic changes to these cells. Recent studies in ES have emphasised the role of the microenvironment in tumour progression (Jiang et al., 2018; Redini and Heymann, 2015). Moreover, exosomes have been involved in the modulation of the pre-metastatic niche, thus suggesting that the transfer of specific cargo to non-cancer cells play an important role in tumour progression (Peinado et al., 2013). Future work should further investigate the role of ES exosomes in modulating cancer progression and relapse, as they could play an important role in the poor outcome of ES patients. Therefore, further understanding of how exosomes modulate neighbouring cells could help design novel treatment strategies to help with disease progression and relapse.

5.8. The importance of larger sample numbers

In this study I have been able to obtain interesting results associated to the patient prognostic profile (Section 5.1), the identification of candidate ES-CSC markers (Section 5.2), evaluation of the heterogeneity across ES samples and identification of putative novel EWSR1 fusions (Section 5.3), postulate exosomes for circulating biomarker strategies (Section 5.6) or their putative role in the ES oncogenic

transformation (Section 5.7). However, most of these results were obtained from a small sample size, and therefore, should be further validated in a bigger sample cohort.

Both clinical and phenotypic data was only available for 26 ES patient-derived cell cultures, thus resulting in a small sample group to perform the prognostic evaluation of the phenotypic *in vitro* profiling (Section 5.1). Although an association between metastatic location and patient outcome could be observed, future studies should include a bigger sample size in order to have more robust results. Similarly, increasing the number of samples in the RNA-seq evaluations would allow for a multivariate analysis combining clinical data, phenotypic data and RNA expression profile, which could result in the ability to identify prognostic and predictive markers associated to a specific patient group. This would help in implementing patient stratification strategies in ES, leading to more specific treatment approaches for each risk group and ultimately, better outcomes.

Similarly, the RNA-seq profiling of exosomes was only performed in 4 samples (1 established cell line and 3 patient-derived cell cultures) (Section 5.6). Therefore, in order to identify a panel of markers that could be exploited for diagnostic and prognostic strategies, more exosome samples should be interrogated, in order to increase the robustness of the results and correlate the profile with the ES producing cells and the clinical response of the patients. Moreover, in order to identify such panel of markers, exosomes from other malignancies and derived from normal tissue should be included, to be able to select those RNAs specifically contained in ES-derived exosomes.

5.9. Final conclusions

In this study I have identified candidate drivers of the ES-CSC profile using a functional self-renewing assay, which might be exploited to develop targeted therapies against these cells. Therefore, futures studies should validate the expression of these ES-CSC markers in a wider panel of samples to confirm if they can be exploited for targeted therapies against these cell populations, thus ultimately leading to improved outcomes for ES patients.

Moreover, I have shown that ES-derived exosomes can induce an ES-like phenotype when taken up by non-ES cells, suggesting mechanisms by which ES cells could be modulating other cells in tumours. Further investigation of the pathways involved in this oncogenic modulation, specially upon uptake of the EWSR1-ETS fusion gene, might result in understanding novel mechanisms by which ES cells induce tumour progression and metastasis.

Commercial suppliers

Abbott Molecular	Abbott Molecular Abbott House, Vanwall Business Park Berkshire, SL6 4XF, UK **www.molecular.abbott**
Abcam	Abcam Plc. Cambridge Science Park Cambridge, CB4 0FL, UK **www.abcam.com**
Agar Scientific	Agar Scientific Ltd M11 Business Link, Stansted Essex, CM24 8GF, UK **www.agarscientific.com**
Agilent Technologies	Agilent Technologies LDA UK Limited Cheadle Royal Business Park, Cheadle Cheshire SK8 3GR, UK **www.agilent.com**
Amersham Biosciences	Amersham Biosciences UK Ltd. Amersham Place, Little Chalfont, Buckinghamshire, HP7 9NA, UK **www.amershambiosciences.com**
Analytik-Jena	Konrad-Zuse-Strasse 1 07745 Jena, Germany **www.analytik-jena.de**
Applied Biosystems	Part of Thermo Fisher Scientific
BD Biosciences	BD Biosciences Edmund Halley Road, Oxford Science Park Oxford, OX4 4DQ, UK **www.bdbiosciences.com**
BDH Laboratory Supplies	Supplied by VWR
Beckman Coulter	Beckman Coulter (UK) Ltd Oakley Court, Kingsmead Business Park High Wycombe, HP11 1JU, UK **www.beckmancoulter.com**
Bibby Scientific	Bibby Scientific UK Beacon Road, Stone Staffordshire, ST15 OS, UK

	www.bibby-scientific.com
Bio-Rad	Bio-Rad Laboratories Ltd. Bio-Rad House, Hemel Hempstead, Hertfordshire, HP2 7DX, UK **www.bio-rad.com**
BioCytex	BioCytex SARL Chemin de l'Armée d'Afrique 13010 Marseille, France **www.biocytex.com**
Cambridge Bioscience	Cambridge Bioscience Ltd Munro House, Bar Hill Cambridge, CB23 8SQ, UK **www.bioscience.co.uk**
Cell Guidance Systems	Cell Guidance Systems Ltd Maia Building, Babraham Bioscience Campus Cambridge, CB22 3AT, UK **www.cellgs.com**
Dako	Part of Agilent Technologies
Eppendorf	Eppendorf UK Ltd Arlington Business Park Stevenage SG1 2FP, UK **www.eppendorf.com**
Fisher Scientific	Part of Thermo Fisher Scientific
Fluka Analytical	Supplied by Sigma-Aldrich
GE Healthcare	GE Healthcare UK Pollards Wood, Chalfont St Giles Buckinghamshire, HP8 4SP, UK **www.gehealthcare.com/uken**
Geneflow	Geneflow Ltd Fradley Business Centre, Fradley Staffordshire, WS13 8NF, UK **www.geneflow.co.uk**
Gilson	Gilson Scientific Ltd. Woodside Estate, Dunstable Bedfordshire, LU5 4TP, UK **www.gilsonuk.com**
Hanna instruments	Hanna Instruments Ltd Pages Industrial Park, Leighton Buzzard Bedfordshire, LU7 4AD, UK **www.hannainstruments.co.uk**

Hettich Lab	Supplied by SLS
Illumina	Illumina UK Chesterford Research Park, Little Chesterford Essex, CB10 1XL, UK **www.emea.illumina.com**
Invitrogen	Part of Thermo Fisher Scientific
Izon	Izon Science Magdalen Centre, Oxford Science Park Oxford, OX4 4GA, UK **www.izon.com**
JEOL	JEOL (Europe) BV Leuvensesteenweg 542 1930 Zaventem, Belgium **www.jeolbenelux.com**
Labtech	Labtech International Ltd. Maple House, Ringmer East Sussex, BN8 5NN, UK **www.labtech.co.uk**
LI-COR Biosciences	Supplied by VWR
Life Technologies	Part of Thermo Fisher Scientific
Merck Millipore	Merck Millipore Croxley Green Business Park, Watford Hertfordshire, WD18 8YH, UK **www.merckmillipore.com**
Miltenyi Biotec	Miltenyi Biotec Ltd. Almac House, Church Lane Bisley Surrey, GU24 9DR, UK **www.miltenyibiotec.com**
MP Biomedicals	Part of Thermo Fisher Scientific
Nikon	Nikon UK Limited Richmond Road, Kingston Upon Thames Surrey, KT2 5PR, UK **www.nikoninstruments.com**
Olympus	Olympus Microscopy KeyMed House, Southend-on-Sea Essex, SS2 5QH, UK **www.olympus.co.uk**
Particle Metrix	Particle Metrix GmbH

	Wildmoos 4 82266 Inning am Ammersee, Germany **www.particle-metrix.de**
Perkin Elmer	PerkinElmer Saxon Way, Cambridge Cambridgeshire, CB23 8SL, UK **www.perkinelmer.co.uk**
Promega	Promega UK Delta House, Southampton Science Park Southampton, SO16 7NS, UK **www.promega.com**
Qiagen	QIAGEN Ltd. QIAGEN House, Fleming Way, Crawley West Sussex, RH10 9NQ, UK **www.qiagen.com**
R&D Systems	R&D Systems Europe, Ltd. Abingdon Science Park Abingdon, OX14 3NB, UK **www.rndsystems.com**
Santa Cruz	Santa Cruz Biotechnology, Inc. Bergheimer Str. 89-2 69115 Heidelberg, Germany **www.scbt.com**
Sanyo Gallenkamp	Sanyo Gallenkamp Plc. The Office Village, Loughborough Leicestershire, LE11 1QJ, UK **www.sanyo-biomedical.co.uk**
Sartorius	Sartorius Stedim UK Ltd. Longmead Business Centre, Epsom Surrey, KT19 9QQ, UK **www.sartorius.co.uk**
Sigma-Aldrich	Sigma-Aldrich Company Ltd. The Old Brickyard, Gillingham Dorset, SP8 4XT, UK **www.sigmaaldrich.com**
SLS	Scientific Laboratory Supplies (SLS) Ltd. Orchard House, Hessle East Riding of Yorkshire, HU13 0AE, UK **www.scientificlabs.co.uk**
Southern Biotech	Supplied by Cambridge Bioscience
Systems Biosciences	Supplied by Cambridge Bioscience

Thermo Fisher Scientific	Thermo Fisher Scientific Ltd. Lingley House, Birchwood Warrington, WA3 7QH, UK **www.thermofisher.com**
VWR	VWR International Ltd. Hunter Boulevard, Magna Park Lutterworth Leicestershire, LE17 4XN, UK **www.uk.vwr.com**
Whatman	Supplied by GE Healthcare

Appendices

ID	Time point	Diag.	Site	Processing method			Culture
				Tissue	Bone	OCT	
Hist1	Diagnosis	ES	bone	✓	—	—	adherent
Hist2	Diagnosis	ES	soft-tissue	✓	—	—	adherent
Hist3	Diagnosis	ES	bone	✓	—	—	adherent
Hist4	Diagnosis	ES	soft-tissue	✓	—	—	adherent
Hist5	Relapse	ES	soft-tissue	✓	—	—	adherent
Hist6	Diagnosis	ES	bone	✓	—	—	adherent
Hist7	Diagnosis	ES	soft-tissue	✓	—	—	adherent
Hist8	Relapse	ES	soft-tissue	✓	—	—	adherent
Hist9	Diagnosis	ES	soft-tissue	✓	—	—	adherent
Hist10	Diagnosis	ES	bone	✓	—	—	adherent
Hist11	PT-Resect.	ES	soft-tissue	✓	—	—	adherent
Hist12	Diagnosis	ES	?	✓	—	—	adherent
Hist13	?	ES	bone	✗	—	—	—
Hist14	Diagnosis	ES	soft-tissue	✓	—	—	adherent
Hist15	Diagnosis	ES	soft-tissue	✓	—	—	adherent
Geno1	Diagnosis	ES	bone	✓	—	—	adherent
Geno2	Diagnosis	ES	bone	✓	✓	—	adherent
Geno3	Diagnosis	OST	—	✓	✓	—	adherent
Geno4	Diagnosis	ES	bone	✗	✗	—	suspension
Geno5	Diagnosis	ES	bone	✓	✓	—	adherent
Geno6	Diagnosis	OM	—	✓	✓	—	adherent
Geno7	Diagnosis	ES	bone	✓	✓	—	adherent
Geno8	Relapse	ES	bone	✗	—	—	—
Geno9	Diagnosis	OST	—	✓	✓	—	adherent
Geno10	Diagnosis	ES	bone	✓	✓	—	adherent
Geno11	Diagnosis	NHL	—	✓	✓	—	adherent
Geno12	Diagnosis	ES	bone	✓	✓	OCT	adherent
Geno13	Diagnosis	L	—	✓	✓	OCT	adherent
Geno14	Diagnosis	ES	bone	✓	—	OCT	adherent
Geno15	PT-Resect.	ES	bone	✗	✗	OCT	—
Geno16	PT-Resect.	ES	bone	✗	✗	OCT	—
Geno17	Diagnosis	NBL	—	✓	—	—	adherent
Geno18	Diagnosis	ES	bone	✓	✓	OCT	adherent
Geno19	Diagnosis	OST	—	✓	✓	OCT	adherent
Geno20	Diagnosis	OM	—	✓	—	—	adherent
Geno21	Diagnosis	N	—	✓	✓	OCT	adherent
Geno22	Diagnosis	N	—	✓	✓	OCT	adherent
Geno23	Diagnosis	ES	bone	✓	—	OCT	adherent
Geno24	PT-Resect.	ES	bone	✗	—	OCT	—
Geno25	Diagnosis	ES	bone	✓	✓	—	adherent
Geno26	Diagnosis	ES	bone	—	✗	—	—
Geno27	Diagnosis	OST	—	✓	✓	OCT	adherent
Geno28	PT-Resect.	ES	bone	✗	—	OCT	—
Geno29	Diagnosis	N	—	✓	✓	OCT	adherent
Geno30	Diagnosis	OST	—	✓	✓	OCT	adherent

ID	Time point	Diag.	Site	Processing method			Culture
				Tissue	Bone	OCT	
Geno31	Diagnosis	N	—	✓	—	OCT	adherent
Geno32	Diagnosis	N	—	✓	—	OCT	adherent
Geno33	Diagnosis	OST	—	✓	✓	OCT	adherent
Geno34	Diagnosis	ES	bone	✓	✓	OCT	adherent
Geno35	PT-Resect.	ES	soft-tissue	✓	—	OCT	adherent
Geno36	PT-Resect.	ES	bone	✓	—	OCT	adherent
Geno37	Diagnosis	N	—	—	✓	—	adherent
Geno38	Relapse	ES	soft-tissue	—	—	—	slides
Geno39	Diagnosis	ES	bone	✓	✓	OCT	adherent
Geno40	Diagnosis	N	—	✓	✓	OCT	adherent
Geno41	PT-Resect.	ES	bone	✗	✗	OCT	—
Geno42	Diagnosis	ES	bone	✓	✓	OCT	adherent
Geno43	PT-Resect.	ES	bone	✓	—	OCT	adherent

Table A1: Patient-derived samples collected and processed in the lab. ID for historical (Hist) and GenoEwing (Geno) samples. Time point of samples as diagonosis, post-treatment resection (PT-Resect.) or relapse. Tumour site for ES samples as bone or soft-tissue. Diagnosis as of Ewing's sarcoma (ES),osteosarcoma (OST), osteomyelitis (OM), Non-Hodgkin's lymphoma (NHL), lymphoma (L), neuroblastoma (NBL) or diagnosis negative for Ewing's sarcoma or osteosarcoma and not known (N). Processing method (Section 2.2.2) as soft-tissue macerate, bone digest or tissue frozen in OCT, and culture success (✓ = yes; ✗ = no, — = not applicable). Obtained cell cultures grown as adherent culture, suspension, or — when unsuccessful culture).

ID	c-kit		Self-renewing		Migration	Doxorubicin		Vincristine	
	c-kit	mixed exp.	%	Clones	72 h	200 nM	EC50	200 nM	EC50
Hist1	-	mix	10	1	98	47	70	53	—
Hist2	+	not	5	0	62	45	10	50	—
Hist3	—	—	7	0	31	49	13	66	—
Hist4	—	—	7	0	73	49	—	66	12
Hist5	-	not	16	1	100	27	4	26	2
Hist6	—	—	14	0	56	94	—	66	130
Hist7	—	—	3	0	60	88	—	114	—
Hist8	++	not	0	0	27	26	8	27	7
Hist9	++	not	6	0	21	65	19	82	—
Hist10	—	—	—	—	68	—	—	—	—
Hist11	—	—	6	0	57	—	—	—	—
Hist12	+	not	21	5	27	66	—	61	10
Hist14	—	—	6	0	171	—	—	—	—
Hist15	—	—	7	0	25	57	20	44	14
Geno1T	++	not	17	2	35	34	11	34	—
Geno2T	++	not	32	2	25	15	3	11	4
Geno2B	—	—	14	1	28	19	3	15	—
Geno5T	++	not	8	2	18	33	8	34	—
Geno5B	+	not	13	9	49	10	5	8	3
Geno7T	+	not	4	1	17	36	6	37	—
Geno7B	++	mix	4	4	11	27	13	26	—
Geno10T	++	not	8	1	10	51	—	53	—
Geno10B	++	mix	8	0	13	59	—	50	—
Geno12T	++	mix	2	0	19	—	—	—	—
Geno14T	++	mix	31	10	11	29	5	26	—
Geno18T	++	not	3	0	13	46	—	71	8
Geno18B	+	not	5	0	54	72	—	68	—
Geno23T	++	not	0	0	20	134	—	189	—
Geno25B	++	mix	17	2	10	16	4	20	3
Geno25B	++	mix	4	0	8	55	15	67	4
Geno34T	++	not	0	0	23	46	4	68	—
Geno34B	++	mix	1	0	21	57	10	58	5
Geno35T	++	mix	12	2	25	21	6	25	5
Geno36T	++	not	3	0	32	62	—	72	—
Geno39T	++	mix	21	2	31	34	11	29	6
Geno39B	++	mix	15	7	36	24	6	30	4
Geno42T	++	mix	0	0	15	19	2	28	—
Geno42B	++	mix	0	0	15	19	3	27	2
Geno43T	++	—	1	0	47	30	3	56	3

Table A2: Phenotypic characterisation of ES cell cultures. c-kit and CD99 expression (ICC) as 0 = negative, + = low or ++ = high expression; and if it contained mixed levels of expression (mix) or not (not). Self-renewing ability described as percentage of wells with colonies after 3-week incubation and the number of colonies propagated in cultures. Migration index at 72 h. Response to doxorubicin and vincristine as number of resistant colonies compared to vehicle control at 200 nM and the EC50 value when possible to obtain. — = data not available.

ID	Gender	Age	Mets	Mets loc.	EFS status	EFS time	OSR status	OSR time	CD99 ICC	CD99 flow	Self-renew.	Migrat. 72 h	Doxorubicin 200 nM	Doxorubicin EC50	Vincristine 200 nM	Vincristine EC50
Hist1	M	14	✓	1	1	469	1	700	0	26	10	98	47	70	53	–
Hist2	M	5	✗	–	1	1016	1	3882	0	–	5	62	45	10	50	–
Hist3	–	6	✗	–	1	310	1	585	1	34	7	31	49	13	66	–
Hist4	F	15	✗	–	0	5587	0	5587	1	85	7	73	49	–	66	12
Hist5	F	14	✓	1	1	243	0	7397	1	80	16	100	27	4	26	2
Hist6	M	24	✗	–	0	6699	0	6699	1	42	14	56	94	–	66	130
Hist7	F	2	✓	2	1	86	1	86	1	69	3	60	88	–	114	–
Hist8	F	9	✗	–	1	1477	1	2874	0	86	0	27	26	8	27	7
Hist9	F	17	✗	–	1	437	1	615	1	–	6	21	65	19	82	–
Hist10	–	23	✗	–	1	174	0	174	1	–	–	68	–	–	–	–
Hist11	M	16	✗	–	1	415	1	564	0	–	6	57	66	–	61	–
Hist14	M	22	✓	1	0	5844	0	5844	1	–	6	171	–	–	–	–
Hist15	–	0	✗	–	1	161	0	360	1	–	7	25	57	20	44	15
Geno1	M	12	✗	–	1	463	0	1066	0	87	17	35	34	11	34	–
Geno2	F	9	✓	2	1	314	1	702	1	44	23	27	17	3	13	4
Geno5	M	15	✓	2	1	93	1	417	2	40	11	33	21	8	21	–
Geno7	M	16	✗	–	0	957	0	957	2	85	4	14	31	6	31	–
Geno10	F	9	✗	–	0	833	0	833	2	95	8	12	55	–	52	–
Geno12	F	12	✗	–	1	507	0	830	2	–	2	19	–	–	–	–
Geno14	M	12	✗	–	1	504	0	803	1	96	31	11	29	5	26	–
Geno23	F	11	✓	2	1	395	0	562	1	2	0	20	134	–	189	–
Geno25	F	10	✓	2	0	498	0	498	2	70	10	9	36	15	43	4
Geno34	F	7	✓	2	1	376	0	430	2	–	1	22	52	4	63	–
Geno36	F	16	✓	2	1	306	0	361	0	–	3	32	62	–	72	–
Geno39	–	17	✓	1	0	355	0	355	2	34	18	34	29	11	29	6
Geno42	–	13	✓	1	0	142	0	142	2	25	0	15	19	2	28	–

Table A3: Clinical data from the ES patients used in the survival analyses (patients with complete clinical data and available phenotypic data). Gender as M = male, F = female. Age at diagnosis as years. Metastasis (Mets) at diagnosis as ✓ = yes, ✗ = no, and location (mets loc.) as 1 = lungs, 2 = other location. EFS (event-free survival) and OSR (overall survival rate) as status (0 = free of relapse or alive respectively, 1 = relapsed or deceased respectively); when no relapse (EFS = 0) or not deceased (OSR = 0), time last seen is shown. CD99 data as ICC (0 = negative, 1 = low and 2 = high expression) and flow cytometry (percentage of CD99 positive cells). Self-renewing ability (self-renew.) as percentage of wells with CSCs grown. Migration index at 72 h. Sensitivity to doxorubicin and vincristine as percentage of colonies resistant at 200 nM and EC50 (nM) for both, respectively. — = data not available.

Pathway analysis results for established cell lines enriched genes

Term	Genes	P-Value	Fold enrichment
Cell cycle	223	1,10E-28	2,1
Nucleoplasm	661	3,00E-26	1,4
DNA replication	82	5,80E-24	3,1
Nucleus	1095	8,80E-24	1,3
DNA replication	55	1,80E-19	3,5
Mitosis	105	2,00E-19	2,4
Cell division	137	3,40E-19	2,1
Cell division	128	9,60E-19	2,1
Chromatin binding	132	2,40E-16	2
Phosphoprotein	1562	3,80E-14	1,1
Centromere	60	1,30E-13	2,7
Mitotic nuclear division	91	1,40E-13	2,2
DNA replication	27	1,70E-13	4,5
Chromosome	125	1,90E-13	1,9
DNA repair	86	4,30E-13	2,2
Sister chromatid cohesion	50	4,40E-13	2,9
G1/S transition of mitotic cell cycle	49	1,20E-12	2,8
DNA damage	111	3,40E-12	1,9
DNA repair	96	8,50E-12	2
Kinetochore	46	1,60E-11	2,8
Cell cycle	51	7,80E-11	2,5
Poly(A) RNA binding	271	9,80E-11	1,4
Condensed chromosome kinetochore	41	1,50E-10	2,8
Nervous system development	92	4,60E-10	1,9
Nucleus	1039	6,80E-10	1,2
Alternative splicing	1916	1,10E-09	1,1
Telomere maintenance via recombination	22	1,60E-09	4
DNA replication initiation	21	6,30E-09	3,9
Base excision repair	21	1,30E-08	3,9
Chromosome, centromeric region	29	1,60E-08	3
DNA recombination	33	4,70E-08	2,7
DNA-binding	422	5,40E-08	1,3
DNA recombination	33	1,00E-07	2,6
Splice variant	1410	1,10E-07	1,1
Base-excision repair	21	1,20E-07	3,5
Pro-rich region	224	1,40E-07	1,4
repeat:PXXP 1	13	1,70E-07	5,1
repeat:PXXP 3	13	1,70E-07	5,1
repeat:PXXP 2	13	1,70E-07	5,1
repeat:PXXP 5	13	1,70E-07	5,1
repeat:PXXP 4	13	1,70E-07	5,1
DNA strand elongation involved in DNA replication	13	1,90E-07	5
RNA-binding	158	2,00E-07	1,5
Chromosome segregation	31	2,20E-07	2,6
Chromatin	36	2,80E-07	2,4
Regulation of transcription in G1/S transition	15	7,00E-07	4,1
Spindle pole	41	7,40E-07	2,2
Mitotic sister chromatid segregation	16	9,50E-07	3,9

Term	Genes	P-Value	Fold enrichment
Leber hereditary optic neuropathy	10	9,90E-07	5,9
Synapse	96	1,10E-06	1,6
mRNA splicing, via spliceosome	66	1,30E-06	1,8
Spliceosomal complex	36	1,40E-06	2,3
mRNA splicing	73	1,50E-06	1,7
Mismatch repair	15	1,90E-06	3,9
Damaged DNA binding	27	2,10E-06	2,6
Kinesin, motor domain	22	2,20E-06	3
KISc	22	2,20E-06	2,9
Nuclear chromatin	60	2,30E-06	1,8
Kinesin-motor domain	22	2,60E-06	2,9
Nucleotide phosphate-binding region:ATP	218	3,70E-06	1,3
mRNA processing	87	3,90E-06	1,6
Kinetochore	31	5,80E-06	2,3
Single-stranded DNA binding	33	5,80E-06	2,2
Microtubule	84	6,80E-06	1,6
nucleotide-excision repair, DNA gap filling	15	7,00E-06	3,6
regulation of signal transduction by p53	42	7,00E-06	2
Microtubule	77	7,00E-06	1,6
Spliceosome	43	7,10E-06	2
Kinesin, motor region, conserved site	20	1,10E-05	2,9
nucleotide binding	90	1,10E-05	1,6
Methylation	216	1,20E-05	1,3
nucleolus	189	1,40E-05	1,3
chromosome	36	1,50E-05	2,1
methylosome	10	1,70E-05	4,9
ATP binding	309	2,10E-05	1,2
RNA splicing	50	2,20E-05	1,8
single-stranded DNA-dependent ATPase activity	9	2,50E-05	5,2
double-strand break repair via homologous recombination	28	2,50E-05	2,3
Homologous recombination	15	4,20E-05	3,2
Spliceosome	40	4,40E-05	1,9
protein sumoylation	38	4,50E-05	1,9
RRM	58	4,50E-05	1,7
G2/M transition of mitotic cell cycle	43	4,60E-05	1,9
Fanconi anemia pathway	22	4,70E-05	2,5
cell junction	111	5,00E-05	1,4
Mitochondrion	237	5,00E-05	1,3
Cadherin 5 domain	35	5,10E-05	2
mitotic nuclear envelope disassembly	20	5,20E-05	2,6
RNA binding	125	5,30E-05	1,4
Pyrimidine metabolism	33	5,80E-05	2
kinesin complex	22	6,20E-05	2,5
RNA polymerase II core promoter region	90	6,40E-05	1,5
ATP-binding	283	6,60E-05	1,2
spindle organization	11	6,70E-05	4
cholesterol biosynthetic process	18	7,40E-05	2,7
Ubl conjugation	338	7,40E-05	1,2
Homoxhilic cell adhesion via plasma membrane	47	9,20E-05	1,8
Neurogenesis	66	9,20E-05	1,6

Term	Genes	P-Value	Fold enrichment
Steroid biosynthesis	12	9,30E-05	3,6
DNA binding	331	9,40E-05	1,2
Homeobox, conserved site	53	9,60E-05	1,7
Chromatin regulator	74	9,60E-05	1,5
regulation of cell cycle	39	9,90E-05	1,9
Homeodomain-like	79	9,90E-05	1,5
histone binding	38	1,00E-04	1,9
MCM complex	8	1,10E-04	5,2
region of interest:5 X 4 AA repeats of P-X-X-P	8	1,10E-04	5,2
protein binding	1574	1,10E-04	1,1
Cadherin	37	1,30E-04	1,9
Parkinson’s disease	41	1,30E-04	1,8
postsynaptic density	51	1,30E-04	1,7
Purine metabolism	49	1,30E-04	1,7
GPCR, family 2-like	23	1,40E-04	2,3
mRNA transport	33	1,40E-04	2
replication fork protection complex	7	1,50E-04	5,8
compositionally biased region:Poly-His	21	1,50E-04	2,4
RNA recognition motif domain	59	1,60E-04	1,6
domain:Cadherin 2	36	1,70E-04	1,9
domain:Cadherin 1	36	1,70E-04	1,9
Viral nucleoprotein	14	1,80E-04	3,1
mismatch repair	15	1,80E-04	2,9
microtubule binding	57	1,80E-04	1,6
viral nucleocapsid	14	1,90E-04	3
Cadherin-like	37	1,90E-04	1,9
CA	36	2,00E-04	1,9
microtubule motor activity	28	2,10E-04	2,1
transcription regulatory region DNA binding	58	2,10E-04	1,6
compositionally biased region:Poly-Gly	73	2,10E-04	1,5
GPCR, family 2, secretin-like	20	2,20E-04	2,4
domain:Cadherin 4	35	2,20E-04	1,9
domain:Cadherin 3	35	2,20E-04	1,9
Metabolic pathways	242	2,20E-04	1,2
domain:Cadherin 6	28	2,30E-04	2
DNA-binding region:Homeobox	53	2,30E-04	1,6
DNA synthesis involved in DNA repair	16	2,40E-04	2,7
telomerase RNA binding	10	2,50E-04	3,9
microtubule-based movement	28	2,70E-04	2
interstrand cross-link repair	20	2,90E-04	2,4
CDK Regulation of DNA Replication	11	3,10E-04	3,4
nuclear matrix	31	3,20E-04	1,9
Nucleotidyltransferase	25	3,40E-04	2,1
brain development	52	3,50E-04	1,6
Oxidative phosphorylation	38	3,90E-04	1,8
centriole	35	3,90E-04	1,8
histone methyltransferase complex	12	4,00E-04	3,2
compositionally biased region:Poly-Lys	56	4,00E-04	1,6
Cholesterol biosynthesis	11	4,10E-04	3,4
Fanconi anemia	11	4,10E-04	3,4

Term	Genes	P-Value	Fold enrichment
Nucleic acid-binding, OB-fold	26	4,10E-04	2
cell cycle	58	4,20E-04	1,6
termination of RNA polymerase II transcription	23	4,30E-04	2,1
four-way junction DNA binding	9	4,40E-04	4
CENP-A containing nucleosome assembly	18	4,60E-04	2,4
positive regulation of excitatory postsynaptic potential	11	4,70E-04	3,4
neuron cell-cell adhesion	10	4,80E-04	3,6
mitochondrial inner membrane	103	4,80E-04	1,4
NAD	48	5,00E-04	1,6
Ctf18 RFC-like complex	7	5,10E-04	5,1
repeat:PXXP 6	7	5,10E-04	5,1
domain:MCM	7	5,10E-04	5,1
region of interest:6 X 4 AA repeats of P-X-X-P	7	5,10E-04	5,1
reciprocal meiotic recombination	14	5,10E-04	2,8
Virion	15	5,20E-04	2,7
nuclear chromosome, telomeric region	38	5,30E-04	1,7
compositionally biased region:Gln-rich	42	5,50E-04	1,7
Cadherin conserved site	34	5,80E-04	1,8
compositionally biased region:Asp/Glu-rich (acidic)	34	5,90E-04	1,8
HOX	59	6,10E-04	1,5
nucleosomal DNA binding	18	6,20E-04	2,4
Developmental protein	197	6,20E-04	1,2
Repressor	129	6,50E-04	1,3
mitochondrion	268	6,70E-04	1,2
Sm	11	6,80E-04	3,2
Ribonucleoprotein LSM domain	11	6,80E-04	3,2
mRNA export from nucleus	31	6,80E-04	1,8
Mini-chromosome maintenance	6	7,20E-04	5,9
telomerase holoenzyme complex	11	7,40E-04	3,2
Meiosis	27	7,80E-04	1,9
chromatin remodeling	28	7,80E-04	1,9
lateral element	9	8,20E-04	3,8
1.RBphosphoE2F	8	8,70E-04	4,2
postsynaptic membrane	54	1,00E-03	1,5

Table A4: Pathway analysis results enriched in established cell lines. Functional families (n = 182, ranked by p-value) for genes enriched in established cell lines compared to patient-derived cell cultures. Results from DAVID using p-value threshold of 0.001. Data as number of genes involved in that function (genes), p-value and fold-change enrichment compared to expected. Functional annotation compared to all genes identified in RNA-seq analysis (background correction).

Pathway analysis results for patient-derived cell cultures

Term	Genes	P-Value	Fold enrichment
Glycoprotein	1027	5,60E-35	1,4
focal adhesion	167	6,50E-33	2,5
glycosylation site:N-linked (GlcNAc...)	971	7,80E-33	1,4
signal peptide	797	6,80E-32	1,4
Signal	891	1,10E-30	1,4
extracellular exosome	683	4,40E-28	1,4
Actin-binding	114	2,50E-22	2,5
Membrane	1467	4,40E-22	1,2
extracellular matrix organization	91	2,00E-21	2,7
Focal adhesion	94	1,80E-20	2,6
cell adhesion	160	4,40E-20	2
Disulfide bond	730	1,20E-19	1,3
Extracellular matrix	104	1,40E-19	2,4
Secreted	454	1,20E-18	1,4
Calcium	248	2,50E-18	1,7
extracellular matrix	113	4,80E-18	2,2
Cell adhesion	156	5,00E-18	2
extracellular space	343	1,70E-16	1,5
actin binding	106	2,30E-16	2,2
proteinaceous extracellular matrix	100	1,20E-14	2,1
Golgi apparatus	221	2,50E-14	1,6
plasma membrane	850	4,10E-14	1,2
integrin binding	52	4,30E-14	2,9
Muscle protein	38	5,60E-14	3,6
ECM-receptor interaction	46	1,50E-13	3
Polymorphism	2188	3,20E-13	1,1
calcium ion binding	201	2,90E-13	1,6
endoplasmic reticulum lumen	74	1,70E-12	2,3
actin cytoskeleton	81	3,10E-12	2,2
disulfide bond	613	3,40E-12	1,3
angiogenesis	82	6,20E-12	2,1
muscle contraction	50	7,70E-12	2,7
Golgi apparatus	226	1,50E-11	1,5
Golgi membrane	167	1,50E-11	1,6
cell surface	155	1,70E-11	1,7
actin filament binding	56	1,80E-11	2,5
stress fiber	32	2,90E-11	3,4
Epidermal growth factor-like domain	82	2,60E-11	2,1
Dilated cardiomyopathy	42	5,70E-11	2,8
Hypertrophic cardiomyopathy (HCM)	40	6,10E-11	2,9
short sequence motif:Cell attachment site	43	5,50E-11	2,8
perinuclear region of cytoplasm	170	1,70E-10	1,6
sequence variant	2237	2,00E-10	1,1
collagen fibril organization	25	2,80E-10	3,8
sarcolemma	41	3,70E-10	2,7
extracellular matrix structural constituent	35	3,70E-10	3
cytoskeleton	113	4,40E-10	1,8
Arrhythmogenic right ventricular cardiomyopathy (ARVC)	35	5,40E-10	2,9

Term	Genes	P-Value	Fold enrichment
positive regulation of cell migration	68	3,70E-10	2,1
EGF-like, conserved site	71	4,10E-10	2,1
Cytoskeleton	270	5,40E-10	1,4
Insulin-like growth factor binding protein, N-terminal	56	4,50E-10	2,3
EGF-like domain	79	7,70E-10	2
positive regulation of angiogenesis	49	7,10E-10	2,4
cell-cell adherens junction	100	1,10E-09	1,8
collagen binding	32	1,20E-09	3,1
positive regulation of endothelial cell migration	27	2,00E-09	3,4
Angiogenesis	50	3,40E-09	2,3
Immunoglobulin-like fold	175	3,50E-09	1,5
topological domain:Cytoplasmic	697	4,30E-09	1,2
Alternative splicing	1918	7,90E-09	1,1
EF-hand-like domain	88	6,50E-09	1,8
Cell membrane	621	9,20E-09	1,2
integrin-mediated signaling pathway	42	7,00E-09	2,5
skeletal muscle tissue development	27	1,30E-08	3,2
lysosomal lumen	38	2,30E-08	2,5
extracellular region	343	2,40E-08	1,3
basement membrane	35	2,50E-08	2,6
Myosin	27	3,00E-08	3,1
heparin binding	58	2,80E-08	2,1
Insulin-like growth factor-binding protein, IGFBP	16	2,70E-08	4,6
domain:IGFBP N-terminal	16	2,60E-08	4,6
Proteoglycans in cancer	68	4,30E-08	1,9
muscle filament sliding	23	3,30E-08	3,4
EGF-like calcium-binding, conserved site	42	3,90E-08	2,4
Regulation of actin cytoskeleton	71	5,60E-08	1,9
Cell junction	167	5,70E-08	1,5
Transmembrane helix	1049	5,80E-08	1,1
Transmembrane	1051	7,10E-08	1,1
EGF-like calcium-binding	48	6,90E-08	2,2
EGF	67	9,60E-08	1,9
membrane raft	68	9,20E-08	1,9
fibronectin binding	18	9,00E-08	3,9
Cytoplasm	922	1,10E-07	1,1
Endoplasmic reticulum	242	1,20E-07	1,4
IB	15	1,40E-07	4,3
Lysosome	78	1,80E-07	1,8
Glycosaminoglycan biosynthesis	15	2,10E-07	4,3
chondroitin sulfate biosynthetic process	17	1,60E-07	4
CH	32	2,20E-07	2,5
lamellipodium	55	2,40E-07	2
Disease mutation	519	2,60E-07	1,2
Pathways in cancer	110	3,60E-07	1,6
topological domain:Lumenal	120	2,40E-07	1,6
cell-matrix adhesion	36	2,50E-07	2,4
Pleckstrin homology-like domain	114	2,80E-07	1,6
Heparin-binding	36	3,80E-07	2,4
Calponin homology domain	33	3,10E-07	2,5

Term	Genes	P-Value	Fold enrichment
EGF-type aspartate/asparagine hydroxylation site	41	3,20E-07	2,3
Signal-anchor	118	4,70E-07	1,6
Z disc	44	4,70E-07	2,2
endoplasmic reticulum	197	5,40E-07	1,4
cadherin binding involved in cell-cell adhesion	84	5,30E-07	1,7
transmembrane region	958	4,60E-07	1,1
actin filament	29	7,00E-07	2,6
integrin complex	17	7,30E-07	3,7
domain:EF-hand 1	60	6,00E-07	1,9
response to hypoxia	57	7,40E-07	1,9
cell migration	57	7,40E-07	1,9
Aortic aneurysm	10	1,00E-06	5,9
Phosphoprotein	1498	1,40E-06	1,1
EGF CA	48	1,50E-06	2
cell adhesion mediated by integrin	12	1,20E-06	4,8
wound healing	33	1,20E-06	2,4
region of interest:Actin-binding	17	1,30E-06	3,6
domain:EF-hand 2	59	1,40E-06	1,9
structural constituent of muscle	22	1,60E-06	3
extracellular matrix disassembly	32	1,50E-06	2,4
EF-hand domain	69	1,80E-06	1,8
signal transduction	261	1,70E-06	1,3
positive regulation of endothelial cell proliferation	30	2,10E-06	2,4
EF-Hand 1, calcium-binding site	57	2,50E-06	1,8
brush border	27	3,10E-06	2,6
muscle organ development	34	2,60E-06	2,3
Ehlers-Danlos syndrome	11	3,60E-06	5
inflammatory response	102	3,10E-06	1,5
cell-cell junction	55	4,00E-06	1,8
substrate adhesion-dependent cell spreading	20	3,40E-06	3
positive regulation of osteoblast differentiation	27	3,50E-06	2,5
lysosome	67	5,00E-06	1,7
sarcomere	20	5,10E-06	3
Phagosome	50	6,10E-06	1,9
cellular response to mechanical stimulus	30	4,20E-06	2,4
cellular response to hypoxia	36	4,20E-06	2,2
SMAD binding	20	5,10E-06	3
extracellular matrix binding	16	5,10E-06	3,5
domain:Myosin head-like	19	4,60E-06	3,1
insulin-like growth factor binding	13	5,70E-06	4,1
CUB domain	25	5,60E-06	2,6
Integrin	21	7,20E-06	2,9
regulation of cell shape	47	5,60E-06	1,9
Pleckstrin homology domain	75	6,40E-06	1,7
negative regulation of cell proliferation	103	6,20E-06	1,5
Endosome	118	8,30E-06	1,5
positive regulation of Iκ B kinase/NF$-\kappa$ B signaling	51	6,60E-06	1,9
cell junction assembly	10	6,70E-06	5,1
Lipoprotein	190	9,10E-06	1,3
protease binding	36	8,20E-06	2,1

Term	Genes	P-Value	Fold enrichment
muscle myosin complex	12	9,10E-06	4,3
positive regulation of ERK1 and ERK2 cascade	55	7,30E-06	1,8
transforming growth factor beta binding	12	9,10E-06	4,3
domain:LIM zinc-binding 3	15	7,90E-06	3,6
glycosaminoglycan biosynthetic process	21	8,30E-06	2,8
ruffle	33	1,20E-05	2,2
PI3K-Akt signaling pathway	94	1,40E-05	1,5
domain:EGF-like 6	21	9,80E-06	2,8
platelet degranulation	37	1,00E-05	2,1
Myosin head, motor domain	20	1,20E-05	2,9
Metal-binding	691	1,70E-05	1,1
Basement membrane	18	1,70E-05	3
Fibronectin, type III	61	1,50E-05	1,7
laminin binding	15	1,70E-05	3,4
repeat:Spectrin 2	15	1,50E-05	3,4
repeat:Spectrin 1	15	1,50E-05	3,4
positive regulation of protein kinase B signaling	32	1,70E-05	2,2
integral component of membrane	964	2,20E-05	1,1
Myosin-like IQ motif-containing domain	12	2,00E-05	4
membrane	449	2,50E-05	1,2
domain:VWFC	12	1,90E-05	4
response to mechanical stimulus	25	2,10E-05	2,4
microfilament motor activity	13	2,70E-05	3,7
repeat:Spectrin 4	13	2,40E-05	3,7
RNA polymerase II core promoter proximal region	14	3,00E-05	3,5
Limb-girdle muscular dystrophy	16	3,50E-05	3,1
receptor-mediated endocytosis	43	2,60E-05	1,9
domain:CH 1	15	2,70E-05	3,3
domain:CH 2	15	2,70E-05	3,3
Lysosome	41	4,10E-05	1,9
endoplasmic reticulum membrane	194	3,60E-05	1,3
Cell projection	163	3,80E-05	1,3
MAPK signaling pathway	72	4,50E-05	1,6
positive regulation of gene expression	73	3,20E-05	1,6
sarcomere organization	16	3,30E-05	3,1
platelet-derived growth factor receptor signaling pathway	16	3,30E-05	3,1
antigen processing and presentation of peptide antigen	16	3,30E-05	3,1
Adherens junction	28	5,10E-05	2,2
MYSc	20	5,30E-05	2,6
growth factor activity	50	4,60E-05	1,8
domain:Fibronectin type-III 4	25	4,30E-05	2,4
splice variant	1401	4,60E-05	1,1
LIM domain	27	6,20E-05	2,2
growth factor binding	15	5,40E-05	3,2
repeat:Spectrin 3	13	4,70E-05	3,6
cell-cell adhesion	74	5,20E-05	1,6
Amoebiasis	36	7,70E-05	1,9
lysosomal membrane	74	7,00E-05	1,6
Cytokine-cytokine receptor interaction	68	8,60E-05	1,6
endoplasmic reticulum-Golgi intermediate compartment	26	7,50E-05	2,2

Term	Genes	P-Value	Fold enrichment
regulation of actin cytoskeleton organization	21	6,40E-05	2,5
Growth factor	41	8,60E-05	1,8
immune response	91	6,80E-05	1,5
CUB	22	9,80E-05	2,4
positive regulation of GTPase activity	134	7,40E-05	1,4
domain:Fibronectin type-III 3	30	7,30E-05	2,1
response to cytokine	22	7,50E-05	2,4
Complement and coagulation cascades	27	1,10E-04	2,1
Spectrin/alpha-actinin	15	8,40E-05	3,1
collagen catabolic process	25	7,80E-05	2,3
Coated pit	21	1,00E-04	2,5
Insulin-like growth factor binding protein	9	8,70E-05	4,7
phagocytic vesicle membrane	23	1,10E-04	2,3
Actinin-type, actin-binding, conserved site	13	9,20E-05	3,4
domain:Actin-binding	12	8,40E-05	3,6
Zinc finger, LIM-type	27	9,40E-05	2,2
blood coagulation	54	8,80E-05	1,7
JAK-STAT cascade	16	8,90E-05	2,9
regulation of cell adhesion	19	9,00E-05	2,6
Immunoglobulin E-set	35	9,90E-05	1,9
Thick filament	11	1,30E-04	3,8
TNF signaling pathway	36	1,50E-04	1,9
Antiviral defense	36	1,40E-04	1,9
endodermal cell differentiation	14	1,00E-04	3,2
type I interferon signaling pathway	25	1,10E-04	2,2
positive regulation of cell-matrix adhesion	13	1,10E-04	3,3
Bacterial invasion of epithelial cells	29	1,60E-04	2
SH3 domain	59	1,50E-04	1,6
topological domain:Extracellular	535	1,20E-04	1,1
domain:EGF-like 2	31	1,20E-04	2
regulation of cell growth	29	1,30E-04	2,1
actin cytoskeleton organization	41	1,40E-04	1,8
filopodium	26	1,70E-04	2,1
Thrombospondin, type 1 repeat	25	1,50E-04	2,2
Annexin	9	1,90E-04	4,4
negative regulation of TGFbR signaling pathway	25	1,40E-04	2,2
blood vessel remodeling	16	1,40E-04	2,8
integral component of plasma membrane	294	1,80E-04	1,2
metal ion-binding site:Zinc; catalytic	40	1,40E-04	1,8
Thyroglobulin type-1	11	1,60E-04	3,7
endosome	62	1,90E-04	1,6
domain:EGF-like 3; calcium-binding	17	1,50E-04	2,7
VEGFR signaling pathway	27	1,50E-04	2,1
domain:Fibronectin type-III 8	13	1,50E-04	3,2
domain:Fibronectin type-III 7	13	1,50E-04	3,2
Cell division and chromosome partitioning	34	3,60E-04	1,8
domain:EGF-like 4	23	1,70E-04	2,3
propeptide:Removed in mature form	63	1,70E-04	1,6
protein binding	1602	2,00E-04	1,1
negative regulation of chondrocyte differentiation	11	1,80E-04	3,6

Term	Genes	P-Value	Fold enrichment
PH	73	2,40E-04	1,5
Growth factor binding	11	2,40E-04	3,6
domain:EGF-like 1	38	1,90E-04	1,8
SPEC	15	2,60E-04	2,8
positive regulation of fibroblast proliferation	22	2,00E-04	2,3
Annexin repeat	9	2,20E-04	4,3
Annexin repeat, conserved site	9	2,20E-04	4,3
Annexin	9	2,20E-04	4,3
Calmodulin-binding	44	2,80E-04	1,7
domain:Fibronectin type-III 5	19	2,10E-04	2,5
positive regulation of MAP kinase activity	23	2,20E-04	2,2
repeat:Annexin 4	9	2,10E-04	4,3
repeat:Annexin 1	9	2,10E-04	4,3
repeat:Annexin 2	9	2,10E-04	4,3
repeat:Annexin 3	9	2,10E-04	4,3
small GTPase mediated signal transduction	66	2,20E-04	1,5
transport vesicle	32	2,80E-04	1,9
FN3	47	3,10E-04	1,7
Prenylation	47	3,00E-04	1,7
MicroRNAs in cancer	69	3,50E-04	1,5
negative regulation of pathway-restricted SMAD	9	2,40E-04	4,2
Hydroxylation	30	3,30E-04	2
Spectrin repeat	13	2,70E-04	3,1
platelet-derived growth factor receptor binding	10	2,90E-04	3,8
Myosin, N-terminal, SH3-like	10	2,80E-04	3,8
Sulfation	20	3,40E-04	2,3
negative regulation of cell migration	32	2,60E-04	1,9
domain:EGF-like 5	18	2,60E-04	2,5
Integrin alpha chain	11	3,00E-04	3,5
Integrin alpha-2	11	3,00E-04	3,5
TY	11	3,80E-04	3,3
filamentous actin	15	3,50E-04	2,7
calcium-binding region:2	37	2,80E-04	1,8
ER to Golgi transport vesicle membrane	20	3,80E-04	2,3
myosin complex	20	3,80E-04	2,3
ER-Golgi transport	30	4,00E-04	1,9
osteoblast differentiation	34	3,10E-04	1,9
positive regulation of muscle cell differentiation	13	3,10E-04	3,1
interferon-gamma-mediated signaling pathway	25	3,10E-04	2,1
epithelial to mesenchymal transition	16	3,20E-04	2,7
positive regulation of focal adhesion assembly	12	3,30E-04	3,2
artery morphogenesis	12	3,30E-04	3,2
ossification	28	3,30E-04	2
platelet-derived growth factor binding	8	3,80E-04	4,5
ANX	9	4,60E-04	3,9
intracellular signal transduction	97	3,40E-04	1,4
Complement Clr-like EGF domain	14	3,80E-04	2,9
Sushi/SCR/CCP	22	3,90E-04	2,2
Immunoglobulin I-set	42	3,90E-04	1,7
endosome membrane	51	4,60E-04	1,6

Term	Genes	P-Value	Fold enrichment
domain:CH	18	3,70E-04	2,5
LIM	27	5,10E-04	2
domain:PDZ	35	3,80E-04	1,8
Cleavage on pair of basic residues	71	5,20E-04	1,5
adherens junction	20	5,10E-04	2,3
regulation of cell migration	26	4,10E-04	2
platelet alpha granule lumen	21	5,40E-04	2,2
costamere	11	5,60E-04	3,3
dystrophin-associated glycoprotein complex	10	5,60E-04	3,5
myosin filament	10	5,60E-04	3,5
domain:PH	62	4,40E-04	1,5
receptor complex	38	5,70E-04	1,7
external side of plasma membrane	52	5,70E-04	1,6
Sushi	21	6,20E-04	2,2
cell cortex	37	6,00E-04	1,7
domain:Thyroglobulin type-1	9	4,70E-04	4
FG-GAP repeat	11	5,30E-04	3,3
Integrin alpha beta-propellor	11	5,30E-04	3,3
acetylglucosaminyltransferase activity	10	5,60E-04	3,5
cell adhesion molecule binding	23	5,80E-04	2,1
domain:Ig-like C2-type 1	52	5,00E-04	1,6
kidney development	29	5,10E-04	1,9
repeat:FG-GAP 6	11	5,10E-04	3,3
repeat:FG-GAP 7	11	5,10E-04	3,3
repeat:FG-GAP 5	11	5,10E-04	3,3
repeat:FG-GAP 3	11	5,10E-04	3,3
repeat:FG-GAP 2	11	5,10E-04	3,3
repeat:FG-GAP 4	11	5,10E-04	3,3
repeat:FG-GAP 1	11	5,10E-04	3,3
Coiled coil	563	6,90E-04	1,1
actin filament organization	25	5,20E-04	2
positive regulation of NF-kappaB transcription factor activity	39	5,20E-04	1,7
Platelet activation	39	7,70E-04	1,7
domain:Fibronectin type-III 9	10	5,10E-04	3,6
TSP1	25	7,20E-04	2
Zymogen	56	7,10E-04	1,5
positive regulation of autophagy	17	5,60E-04	2,5
positive regulation of substrate adhesion-dependent cell spreading	15	5,70E-04	2,6
repeat:Spectrin 19	7	5,70E-04	5
domain:Fibronectin type-III 14	7	5,70E-04	5
domain:Fibronectin type-III 15	7	5,70E-04	5
repeat:Spectrin 20	7	5,70E-04	5
negative regulation of endopeptidase activity	37	5,70E-04	1,8
Rho GTPase activation protein	29	6,20E-04	1,9
domain:Ig-like C2-type 2	52	5,70E-04	1,6
Proteoglycan	19	7,80E-04	2,3
domain:EGF-like 2; calcium-binding	21	6,20E-04	2,2
Protease inhibitor	36	8,30E-04	1,7
positive regulation of cell adhesion	18	6,30E-04	2,4
Immunoglobulin domain	104	8,60E-04	1,4

Term	Genes	P-Value	Fold enrichment
defense response to virus	46	6,50E-04	1,6
Protease inhibitor I4, serpin, conserved site	15	7,10E-04	2,6
podosome	12	8,40E-04	3
antigen processing and presentation	19	6,90E-04	2,3
cytokine activity	49	8,00E-04	1,6
ruffle membrane	27	9,00E-04	1,9
signal transducer activity	54	8,30E-04	1,5
receptor binding	86	8,30E-04	1,4
Coagulation factor 5/8 C-terminal type domain	12	7,90E-04	3
phosphatidylserine binding	16	8,90E-04	2,5
activin binding	8	8,90E-04	4,1
Laminin G domain	21	8,50E-04	2,1
cortical actin cytoskeleton	17	9,90E-04	2,4
DHR-1 domain	8	8,50E-04	4,2
Dedicator of cytokinesis	8	8,50E-04	4,2
Dedicator of cytokinesis C-terminal	8	8,50E-04	4,2
DHR-2 domain	8	8,50E-04	4,2
GTPase activator activity	69	9,40E-04	1,5
regulation of inflammatory response	23	8,20E-04	2,1
repeat:Spectrin 10	8	8,30E-04	4,2
repeat:Spectrin 12	8	8,30E-04	4,2
repeat:Spectrin 16	8	8,30E-04	4,2
repeat:Spectrin 15	8	8,30E-04	4,2
repeat:Spectrin 17	8	8,30E-04	4,2
repeat:Spectrin 14	8	8,30E-04	4,2
domain:DHR-1	8	8,30E-04	4,2
repeat:Spectrin 11	8	8,30E-04	4,2
repeat:Spectrin 13	8	8,30E-04	4,2
domain:DHR-2	8	8,30E-04	4,2
positive regulation of actin filament polymerization	18	8,60E-04	2,3
region of interest:Triple-helical region	11	8,60E-04	3,2
regulation of tumor necrosis factor production	6	9,00E-04	5,6
positive regulation of immunoglobulin secretion	8	9,30E-04	4,1

Table A5: Pathway analysis results enriched in patient-derived cell cultures. Functional families (n = 381, ranked by p-value) for genes enriched in patient-derived cell cultures compared to established cell lines. Results from DAVID using p-value threshold of 0.001. Data as number of genes involved in that function (genes), p-value and fold-change enrichment compared to expected. Functional annotation compared to all genes identified in RNA-seq analysis (background correction).

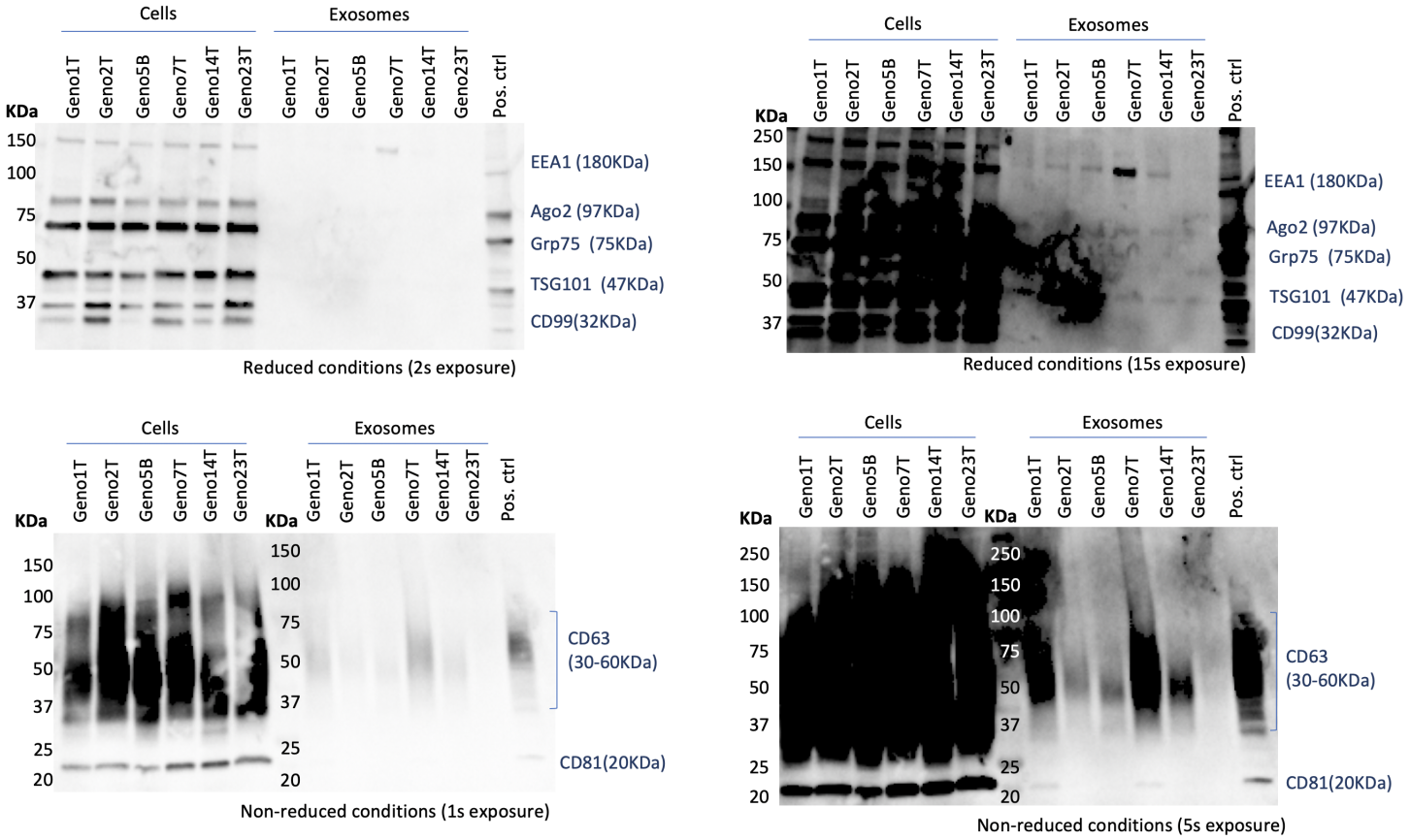


Figure A1: Exosome and patient-derived ES cell cultures characterised by WB - figure 1. Protein extracted from cells (left) and exosomes (right) characterised by WB. Reduced conditions for EEA1, Ago2, Grp75, TSG101 and CD99; non-reduced conditions for CD63 and CD81. Low exposure for clear visualisation of cell extractions; long exposures to visualise exosome markers. KDa = size ladder.

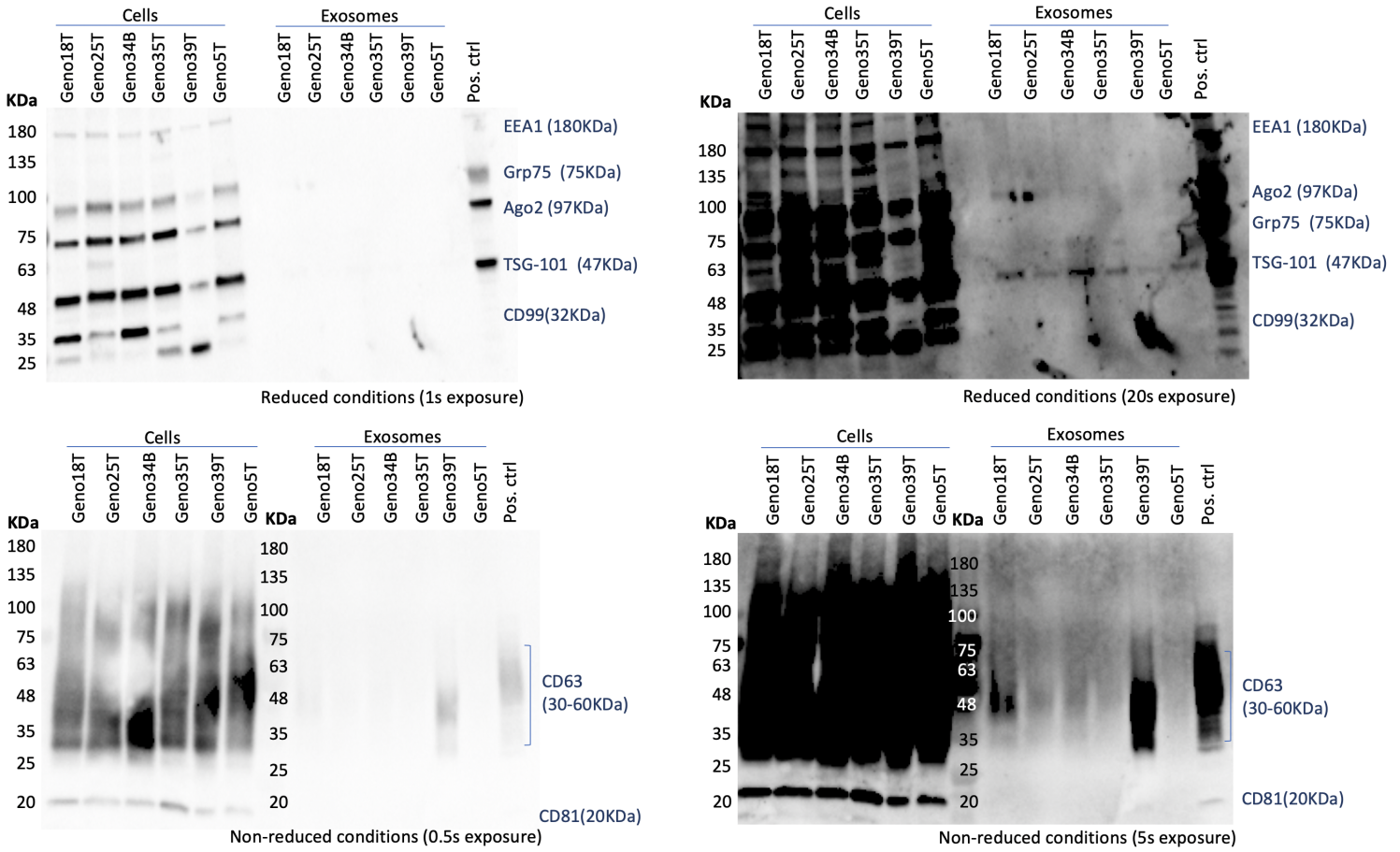


Figure A2: Exosome and patient-derived ES cell cultures characterised by WB - figure 2. Protein extracted from cells (left) and exosomes (right) characterised by WB. Reduced conditions for EEA1, Ago2, Grp75, TSG101 and CD99; non-reduced conditions for CD63 and CD81. Low exposure for clear visualisation of cell extractions; long exposures to visualise exosome markers. KDa = size ladder.

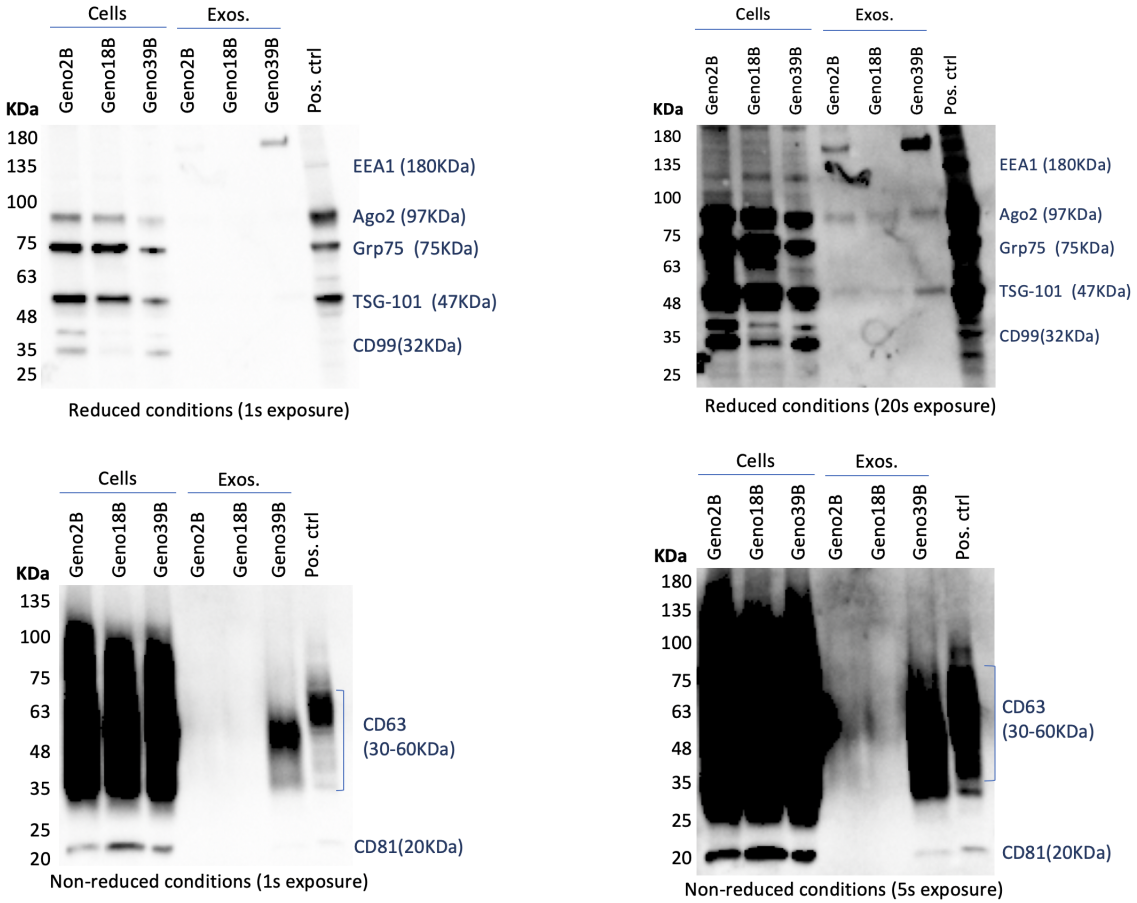


Figure A3: Exosome and patient-derived ES cell cultures characterised by WB - figure 3. Protein extracted from cells (left) and exosomes (right) characterised by WB. Reduced conditions for EEA1, Ago2, Grp75, TSG101 and CD99; non-reduced conditions for CD63 and CD81. Low exposure for clear visualisation of cell extractions; long exposures to visualise exosome markers. KDa = size ladder.

References

Abels, E. R. and Breakefield, X. O. (2016), 'Introduction to Extracellular Vesicles: Biogenesis, RNA Cargo Selection, Content, Release, and Uptake', *Cellular and Molecular Neurobiology* **36**(3), 301–312. 26, 27

Al-Hajj, M., Wicha, M. S., Benito-Hernandez, A., Morrison, S. J. and Clarke, M. F. (2003), 'Prospective identification of tumorigenic breast cancer cells', *Proceedings of the National Academy of Sciences* **100**(7), 3983–3988. 19, 83

Amaral, A. T., Garofalo, C., Frapolli, R., Manara, M. C., Mancarella, C., Uboldi, S., Di Giandomenico, S., Ordóñez, J. L., Sevillano, V., Malaguarnera, R., Picci, P., Hassan, A. B., De Alava, E., D'Incalci, M. and Scotlandi, K. (2015), 'Trabectedin efficacy in Ewing sarcoma is greatly increased by combination with anti-IGF signaling agents', *Clinical Cancer Research* **21**(6), 1373–1382. 10, 220

Andre, F., Schartz, N. E. C., Movassagh, M., Flament, C., Pautier, P., Morice, P., Pomel, C., Lhomme, C., Escudier, B., Chevalier, T. L., Tursz, T., Amigorena, S., Raposo, G., Angevin, E. and Zitvogel, L. (2002), 'Malignant effusions and immunogenic tumour-derived exosomes', *Lancet* **360**, 295–305. 30

Andrews, S. (2010), 'FastQC: a quality control tool for high throughput sequence data'. 91

Armengol, G., Tarkkanen, M., Virolainen, M., Forus, A., Valle, J., Böhling, T., Asko-Seljavaara, S., Blomqvist, C., Elomaa, I., Karaharju, E., Kivioja, A. H., Siimes, M. A., Tukiainen, E., Caballín, M. R., Myklebost, O. and Knuutila, S. (1997), 'Recurrent gains of 1q, 8 and 12 in the Ewing family of tumours by comparative genomic hybridization', *British Journal of Cancer* **75**(10), 1403–1409. 3

Aryee, D. N. T., Niedan, S., Kauer, M., Schwentner, R., Idriss, M., Ban, J., Muehlbacher, K., Kreppel, M., Walker, R. L., Poremba, C., Kofler, R. and Kovar, H. (2011), 'Hypoxia modulates EWS-FLI1 transcriptional signature and enhances malignant properties of Ewing's sarcoma cells in-vitro', *Cancer Research* **70**(10), 4015–4023. 12

Ashman, L. K. and Griffith, R. (2012), 'Therapeutic targeting of c-KIT in cancer', *Expert Opinion on Investigational Drugs* **22**(1), 103–115. 40

Atai, N. A., Balaj, L., Van Veen, H., Breakefield, X. O., Jarzyna, P. A., Van Noorden, C. J., Skog, J. and Maguire, C. A. (2013), 'Heparin blocks transfer of extracellular vesicles between donor and recipient cells', *Journal of Neuro-Oncology* **115**(3), 343–351. 130

Awad, O., Yustein, J. T., Shah, P., Gul, N., Katuri, V., Neill, A. O., Kong, Y., Brown, M. L., Toretsky, J. A. and Loeb, D. M. (2010), 'High ALDH Activity Identifies Chemotherapy-Resistant Ewing's Sarcoma Stem Cells That Retain Sensitivity to EWS-FLI1 Inhibition', *PloS one* **5**(11). 24, 39, 76, 83, 120, 122, 123, 217, 221, 223

Baietti, M. F., Zhang, Z., Mortier, E., Melchior, A., Degeest, G., Geeraerts, A., Ivarsson, Y., Depoortere, F., Coomans, C., Vermeiren, E., Zimmermann, P. and David, G. (2012), 'Syndecan-syntenin-ALIX regulates the biogenesis of exosomes', *Nature Cell Biology* **14**(7), 677–685. 27, 233

Bailly, R.-A., Bosselut, R., Zucman, J., Cormier, F., Delattre, O., Roussel, M., Thomas, G. and Ghysdaell, J. (1994), 'DNA-Binding and Transcriptional Activation Properties of the EWS-FLI-1 Fusion Protein Resulting from the t(11;22) Translocation in Ewing Sarcoma', *Molecular and Cellular Biology* **14**(5), 3230–3241. 2

Baiocchi, M., Biffoni, M., Ricci-Vitiani, L., Pilozzi, E. and De Maria, R. (2010), 'New models for cancer research: Human cancer stem cell xenografts', *Current Opinion in Pharmacology* **10**(4), 380–384. 21

Balamuth, N. J. and Womer, R. B. (2010), 'Ewing's sarcoma', *The Lancet Oncology* **11**(2), 184–192. 4, 9, 224

Baruchel, S., Pappo, A., Krailo, M., Baker, K. S., Wu, B., Villaluna, D., Lee-Scott, M., Adamson, P. C. and Blaney, S. M. (2012), 'A phase 2 trial of trabectedin in children with recurrent rhabdomyosarcoma, Ewing sarcoma and non-rhabdomyosarcoma soft tissue sarcomas: A report from the Children's Oncology Group', *European Journal of Cancer* **48**(4), 579–585. 10, 220

Baruzzo, G., Hayer, K. E., Kim, E. J., Camillo, B. D., FitzGerald, G. A. and Grantiacomo, G. R. (2017), 'Simulation-based comprehensive benchmarking of RNA-seq aligners', *Nature Methods* **14**(2), 135–139. 128

Beck, B., Driessens, G., Goossens, S., Youssef, K. K., Kuchnio, A., Caauwe, A., Sotiropoulou, P. A., Loges, S., Lapouge, G., Candi, A., Mascre, G., Drogat, B., Dekoninck, S., Haigh, J. J., Carmeliet, P. and Blanpain, C. (2011), 'A vascular niche and a VEGF-Nrp1 loop regulate the initiation and stemness of skin tumours', *Nature* **478**(7369), 399–403. 20

Belov, L., Matic, K. J., Hallal, S., Best, O. G., Mulligan, S. P. and Christopherson, R. I. (2016), 'Extensive surface protein profiles of extracellular vesicles from cancer cells may provide diagnostic signatures from blood samples', *Journal of Extracellular Vesicles* **5**(1), 1–12. 32

Benini, S., Gamberi, G., Cocchi, S., Garbetta, J., Alberti, L., Righi, A., Gambarotti, M., Picci, P. and Ferrari, S. (2018), 'Detection of circulating tumor cells in liquid biopsy from Ewing sarcoma patients', *Cancer Management and Research* **10**, 49–60. 227

Benini, S., Manara, M. C., Baldini, N., Cerisano, V., Serra, M., Mercuri, M., Lollini, P.-l., Nanni, P., Picci, P. and Scotlandi, K. (2001), 'Inhibition of Insulin-like Growth Factor I Receptor Increases the Antitumor Activity of Doxorubicin and Vincristine against Ewing's Sarcoma Cells', *Clinical Cancer Research* **7**(June), 1790–1797. 9

Benini, S., Zuntini, M., Manara, M. C., Cohen, P., Nicoletti, G., Nanni, P., Oh, Y., Picci, P. and Scotlandi, K. (2006), 'Insulin-like growth factor binding protein 3 as an anticancer molecule in Ewing's sarcoma', *International Journal of Cancer* **119**(5), 1039–1046. 9

Berghuis, D., Santos, S. J., Baelde, H. J., Taminiau, A. H., Maarten Egeler, R., Schilham, M. W., Hogendoorn, P. C. and Lankester, A. C. (2011), 'Pro-inflammatory chemokine-chemokine receptor interactions within the Ewing sarcoma microenvironment determine CD8+ T-lymphocyte infiltration and affect tumour progression', *Journal of Pathology* **223**(3), 347–357. 13

Bernstein, M., Kovar, H., Paulussen, M., Randall, R. L., Schuck, A., Teot, L. A. and Juergens, H. (2006), 'Ewing's sarcoma family of tumors: Current management', *The Oncologist* **11**(5), 503–519. 1, 226

Besse, B., Charrier, M., Lapierre, V., Dansin, E., Lantz, O., Planchard, D., Le Chevalier, T., Livartoski, A., Barlesi, F., Laplanche, A., Ploix, S., Vimond, N., Peguillet, I., Théry, C., Lacroix, L., Zoernig, I., Dhodapkar, K., Dhodapkar, M., Viaud, S., Soria, J. C., Reiners, K. S., Pogge von Strandmann, E., Vély, F., Rusakiewicz, S., Eggermont, A., Pitt, J. M., Zitvogel, L. and Chaput, N. (2016), 'Dendritic cell-derived exosomes as maintenance immunotherapy after first line chemotherapy in NSCLC', *OncoImmunology* **5**(4), 1–13. 34, 233

Bond, M., Bernstein, M. L., Pappo, A., Schultz, K. R., Krailo, M., Blaney, S. M. and Adamson, P. C. (2008), 'A Phase II Study of Imatinib Mesylate in Children With Refractory or Relapsed Solid Tumors: A Children's Oncology Group Study', *Pediatr Blood Cancer* **50**, 254–258. 11, 40, 64, 80, 220

Bonnet, D. and Dick, J. E. (1997), 'Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell', *Nature* **3**(7), 730–7737. 19

Bortoluzzi, S., Lovisa, F., Gaffo, E. and Mussolin, L. (2017), 'Small RNAs in Circulating Exosomes of Cancer Patients: A Minireview', *High-Throughput* **6**(4), 13. 209

Braun, B. S., Frieden, R., Lessnick, S. L., May, W. A. and Denny, C. T. (1995), 'Identification of target genes for the Ewing's sarcoma EWS/FLI fusion protein by representational difference analysis.', *Molecular and cellular biology* **15**(8), 4623–30. 17

Bray, N. L., Pimentel, H., Melsted, P. and Pachter, L. (2016), 'Near-optimal probabilistic RNA-seq quantification', *Nature Biotechnology* **34**(5), 525–527. 85

Brenner, J. C., Feng, F. Y., Han, S., Patel, S., Goyal, S. V., Bou-Maroun, L. M., Liu, M., Lonigro, R., Prensner, J. R., Tomlins, S. A. and Chinnaiyan, A. M. (2012), 'PARP-1 inhibition as a targeted strategy to treat Ewing's sarcoma', *Cancer Research* **72**(7), 1608–1613. 10, 220

Brohl, A. S., Solomon, D. A., Chang, W., Wang, J., Song, Y., Sindiri, S., Patidar, R., Hurd, L., Chen, L., Shern, J. F., Liao, H., Wen, X., Gerard, J., Kim, J. S., Lopez Guerrero, J. A., Machado, I., Wai, D. H., Picci, P., Triche, T., Horvai, A. E., Miettinen, M., Wei, J. S., Catchpool, D., Llombart-Bosch, A., Waldman, T. and Khan, J. (2014), 'The Genomic Landscape of the Ewing Sarcoma Family of Tumors Reveals Recurrent STAG2 Mutation', *PLoS Genetics* **10**(7). 3, 125, 225

Burchill, S. A. (2003), 'Ewing's sarcoma: diagnostic, prognostic, and therapeutic implications of molecular abnormalities', *Journal of clinical pathology* **56**(2), 96–102. 1

Byron, S. A., Van Keuren-Jensen, K. R., Engelthaler, D. M., Carpten, J. D. and Craig, D. W. (2016), 'Translating RNA sequencing into clinical diagnostics: Opportunities and challenges', *Nature Reviews Genetics* **17**(5), 257–271. 84, 85

Cabrera, M. C. (2015), 'Cancer stem cell plasticity and tumor hierarchy', *World Journal of Stem Cells* **7**(1), 27. 22

Capdeville, R., Buchdunger, E., Zimmermann, J. and Matter, A. (2002), 'Glivec (STI571, imatinib), a rationally developed, targeted anticancer drug', *Nature reviews. Drug discovery* **1**(July), 493–502. 7

Castillero-Trejo, Y., Eliazer, S., Xiang, L., Richardson, J. A. and Ilaria, R. L. (2005), 'Expression of the EWS/FLI-1 oncogene in murine primary bone-derived cells results in EWS/FLI-1-dependent, Ewing sarcoma-like tumors', *Cancer Research* **65**(19), 8698–8705. 18

Chamberlin, D. D. and Boyce, R. F. (1976), 'Sequel', *Proceedings of the 1976 ACM SIGFIDET (now SIGMOD) workshop on Data description, access and control - FIDET '76* pp. 249–264. 91

Cheruvanky, A., Zhou, H., Pisitkun, T., Kopp, J. B., Knepper, M. A., Yuen, P. S. T. and Star, R. A. (2008), 'Rapid isolation of urinary exosomal biomarkers using a nanomembrane ultrafiltration concentrator', *Am J physiol Renal Physiol* **292**(5), 1–12. 132

Chowdhury, R., Webber, J. P., Gurney, M., Mason, M. D., Tabi, Z. and Clayton, A. (2015), 'Cancer exosomes trigger mesenchymal stem cell differentiation into pro-angiogenic and pro-invasive myofibroblasts', *Oncotarget* **6**(2), 715–731. 31, 231, 232

Ciuffreda, L., Falcone, I., Cesta Incani, U., Del Curatolo, A., Conciatori, F., Matteoni, S., Vari, S., Vaccaro, V., Cognetti, F. and Milella, M. (2014), 'PTEN expression and function in adult cancer stem cells and prospects for therapeutic targeting', *Advances in Biological Regulation* **56**, 66–80. 21

Conic, I., Stanojevic, Z., Jankovic Velickovic, L., Stojnev, S., Ristic Petrovic, A., Krstic, M., Stanojevic, M., Bogdanovic, D. and Stefanovic, V. (2015), 'Epithelial ovarian cancer with CD117 phenotype is highly aggressive and resistant to chemotherapy', *Journal of Obstetrics and Gynaecology Research* **41**(10), 1630–1637. 40

Cornaz-Buros, S., Riggi, N., DeVito, C., Sarre, A., Letovanec, I., Provero, P. and Stamenkovic, I. (2014), 'Targeting cancer stem-like cells as an approach to defeating cellular heterogeneity in Ewing sarcoma.', *Cancer research* **74**(22), 6610–22. 24, 39, 76, 83, 120, 122, 123, 217, 221, 222, 223

Costa-silva, J., Domingues, D. and Lopes, F. M. (2017), 'RNA-Seq differential expression analysis : An extended review and a software tool', *PLoS ONE* **12**(12), 1–18. 85, 86, 128

Cotterill, B. S. J., Ahrens, S., Paulussen, M. and Ju, H. F. (2000), 'Prognostic Factors in Ewing's Tumor of Bone: Analysis of 975 Patients From the European Intergroup Cooperative Ewing's Sarcoma Study Group', *Journal of Clinical Oncology* **18**(17), 3108–3114. 79

Crompton, B. D., Stewart, C., Taylor-Weiner, A., Alexe, G., Kurek, K. C., Calicchio, M. L., Kiezun, A., Carter, S. L., Shukla, S. A., Mehta, S. S., Thorner, A. R., de Torres, C., Lavarino, C., Suñol, M., McKenna, A., Sivachenko, A., Cibulskis, K., Lawrence, M. S., Stojanov, P., Rosenberg, M., Ambrogio, L., Auclair, D., Seepo, S., Blumenstiel, B., DeFelice, M., Imaz-Rosshandler, I., Celis, A. S. C., Rivera, M. N., Rodriguez-Galindo, C., Fleming, M. D., Golub, T. R., Getz, G., Mora, J. and Stegmaier, K. (2014), 'The genomic landscape of pediatric Ewing sarcoma', *Cancer Discovery* **4**(11), 1326–1341. 3, 14, 125, 224, 225

Dalerba, P., Dylla, S. J., Park, I.-K., Liu, R., Wang, X., Cho, R. W., Hoey, T., Gurney, A., Huang, E. h., Simone, D. M., Shelton, A. A., Parmiani, G., Castelli, C. and Clarke, M. F. (2007), 'Phenotypic characterization of human colorectal cancer', *PNAS* **104**(24), 10158–10163. 19

De Vito, C., Riggi, N., Cornaz, S., Suva, M., Baumer, K., Provero, P. and Stamenkovic, I. (2012), 'A TARBP2-Dependent miRNA Expression Profile Underlies Cancer Stem Cell Properties and Provides Candidate Therapeutic Reagents in Ewing Sarcoma', *Cancer Cell* **21**(6), 807–821. 23, 24, 76, 83, 120, 122, 123, 217, 221, 222, 223

Delattre, O., Zucman, J., Plougastel, B., Desmaze, C., Melot, T., Peter, M., Kovar, H., Joubert, I., de Jong, P. and Rouleau, G. (1992), 'Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours', *Nature* **359**(6391), 162–165. 2, 7

Delić, D., Eisele, C., Schmid, R., Baum, P., Wiech, F., Gerl, M., Zimdahl, H., Pullen, S. S. and Urquhart, R. (2016), 'Urinary exosomal miRNA signature in type II diabetic nephropathy patients', *PLoS ONE* **11**(3), 1–16. 133

Demory Beckler, M., Higginbotham, J. N., Franklin, J. L., Ham, A.-J., Halvey, P. J., Imasuen, I. E., Whitwell, C., Li, M., Liebler, D. C. and Coffey, R. J. (2013), 'Proteomic Analysis of Exosomes from Mutant KRAS Colon Cancer Cells Identifies Intercellular Transfer of Mutant KRAS', *Molecular & Cellular Proteomics* **12**(2), 343–355. 30, 131

Dewangan, J., Srivastava, S. and Rath, S. K. (2017), 'Salinomycin: A new paradigm in cancer therapy', *Tumor Biology* **39**(3). 22

Dias, M. V. S., Costa, C. S. and DaSilva, L. L. P. (2018), 'The Ambiguous Roles of Extracellular Vesicles in HIV Replication and Pathogenesis', *Frontiers in Microbiology* **9**(October), 1–13. 29

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T. R. (2013), 'STAR: ultrafast universal RNA-seq aligner', *Bioinformatics* **29**(1), 15–21. 85, 91, 127

Doherty, M. R., Smigiel, J. M., Junk, D. J. and Jackson, M. W. (2016), 'Cancer stem cell plasticity drives therapeutic resistance', *Cancers* **8**(1), 1–13. 22

Dowless, M. S., Lowery, C. D., Shackleford, T. J., Renschler, M., Stephens, J. R., Flack, R., Blosser, W., Gupta, S., Stewart, J., Webster, Y., Dempsey, J., VanWye, A. B., Ebert, P. J., Iversen, P., Olsen, J. B., Gong, X., Buchanan, S. G., Houghton, P. J. and Stancato, L. F. (2018), 'Abemaciclib is Active in Preclinical

Models of Ewing's Sarcoma via Multi-pronged Regulation of Cell Cycle, DNA Methylation, and Interferon Pathway Signaling', *Clinical Cancer Research* **24**(15), clincanres.1256.2018. 77, 226

Droop, A. P. (2016), 'Qsubsec: A lightweight template system for defining sun grid engine workflows', *Bioinformatics* **32**(8), 1267–1268. 91

El Andaloussi, S., Mäger, I., Breakefield, X. O. and Wood, M. J. (2013), 'Extracellular vesicles: Biology and emerging therapeutic opportunities', *Nature Reviews Drug Discovery* **12**(5), 347–357. 33, 129, 204, 211, 229, 233

Enderle, D., Spiel, A., Coticchia, C. M., Berghoff, E., Mueller, R., Schlumpberger, M., Sprenger-Haussels, M., Shaffer, J. M., Lader, E., Skog, J. and Noerholm, M. (2015), 'Characterization of RNA from exosomes and other extracellular vesicles isolated by a novel spin column-based method', *PLoS ONE* **10**(8), 1–19. 133, 141, 213, 214

Engström, P. G., Steijger, T., Sipos, B., Grant, G. R. and Kahles, A. (2014), 'Europe PMC Funders Group Systematic evaluation of spliced alignment programs for RNA- seq data', *Nature Methods* **10**(12), 1185–1191. 86

Ent, W. V. D., Sand, L. G. L. and Hogendoorn, P. C. W. (2018), 'Molecular genetics of Ewing sarcoma , model systems and finding novel (immuno-) therapeutic targets', *Transl Genet Genom* **2**(10), 1–12. 13, 14, 15, 223, 226

Erkizan, H. V., Kong, Y., Merchant, M., Schlottmann, S., Barber-Rotenberg, J. S., Abaan, O. D., Chou, T.-h., Dakshanamurthy, S., Brown, M. L., Üren, A. and Toretsky, J. A. (2009), 'Small molecule selected to disrupt oncogenic protein EWS-FLI1 interaction with RNA Helicase A inhibits Ewing's Sarcoma Hayriye', *Nature medicine* **15**(7), 750–756. 7

Erkizan, H. V., Schneider, J. A., Sajwan, K., Graham, G. T., Griffin, B., Chasovskikh, S., Youbi, S. E., Kallarakal, A., Chruszcz, M., Padmanabhan, R., Casey, J. L., Uren, A. and Toretsky, J. A. (2015), 'RNA helicase A activity is inhibited by oncogenic transcription factor EWS-FLI1', *Nucleic Acids Research* **43**(2), 1069–1080. 8

Escrevente, C., Keller, S., Altevogt, P. and Costa, J. (2011), 'Interaction and uptake of exosomes by ovarian cancer cells', *BMC Cancer* **11**. 130, 211

Esiashvili, N., Goodman, M. and Marcus, R. B. (2008), 'Changes in incidence and survival of Ewing sarcoma patients over the past 3 decades: Surveillance Epidemiology and End Results data', *Journal of pediatric hematology/oncology* **30**(6), 425–30. 38

Essandoh, K., Yang, L., Wang, X., Huang, W., Qin, D., Hao, J., Wang, Y., Zingarelli, B., Peng, T. and Fan, G. C. (2015), 'Blockade of exosome generation with GW4869 dampens the sepsis-induced inflammation and cardiac dysfunction', *Biochimica et Biophysica Acta - Molecular Basis of Disease* **1852**(11), 2362–2371. 211

Ewing, B., Ewing, B., Hillier, L., Hillier, L., Wendl, M. C., Wendl, M. C., Green, P. and Green, P. (1998), 'Base-Calling of Automated Sequencer Traces Using', *Genome Research* **8**, 175–185. 91

Ewing, J. (1921), 'Diffuse Endothelioma of Bone', *Proceedings of the New York Pathological Society* **21**(1), 17–24. 17

Fabbri, M., Paone, A., Calore, F., Galli, R., Gaudio, E., Santhanam, R., Lovat, F., Fadda, P., Mao, C., Nuovo, G. J., Zanesi, N., Crawford, M., Ozer, G. H., Wernicke, D., Alder, H., Caligiuri, M. A., Nana-Sinkam, P., Perrotti, D. and Croce, C. M. (2012), 'MicroRNAs bind to Toll-like receptors to induce prometastatic inflammatory response', *PNAS* **109**(31), 2110–2116. 33, 211, 233

Fagone, P., Nicoletti, F., Salvatorelli, L., Musumeci, G. and Magro, G. (2015), 'Cyclin D1 and Ewing's sarcoma/PNET: A microarray analysis', *Acta Histochemica* **117**(8), 824–828. 3

Felicetti, F., De Feo, A., Coscia, C., Puglisi, R., Pedini, F., Pasquini, L., Bellenghi, M., Errico, M. C., Pagani, E. and Carè, A. (2016), 'Exosome-mediated transfer of miR-222 is sufficient to increase tumor malignancy in melanoma', *Journal of Translational Medicine* **14**(1), 56. 141

Fomeshi, M. R., Ebrahimi, M., Mowla, S. J., Khosravani, P., Firouzi, J. and Khayatzadeh, H. (2015), 'Evaluation of the expressions pattern of miR-10b, 21, 200c, 373 and 520c to find the correlation between epithelial-to-mesenchymal transition and melanoma stem cell potential in isolated cancer stem cells', *Cellular and Molecular Biology Letters* **20**(3). 222

Fong, E. L. S., Lamhamedi-Cherradi, S.-E., Burdett, E., Ramamoorthy, V., Lazar, A. J., Kasper, F. K., Farach-Carson, M. C., Vishwamitra, D., Demicco, E. G., Menegaz, B. A., Amin, H. M., Mikos, A. G. and Ludwig, J. A. (2013), 'Modeling Ewing sarcoma tumors in vitro with 3D scaffolds', *Proceedings of the National Academy of Sciences* **110**(16), 6500–6505. 15

Franzen, C. A., Simms, P. E., Van Huis, A. F., Foreman, K. E., Kuo, P. C. and Gupta, G. N. (2014), 'Characterization of uptake and internalization of exosomes by bladder cancer cells', *BioMed Research International* **2014**. 130, 211

Galli, R., Binda, E., Orfanelli, U., Cipelletti, B., Gritti, A., Vitis, S. D., Fiocco, R., Foroni, C., Dimeco, F. and Vescovi, A. (2004), 'Isolation and Characterization of Tumorigenic, Stem-like Neural Precursors from Human Glioblastoma', *Cancer research* **64**, 7011–7021. 83

Garofalo, C., Manara, M. C., Nicoletti, G., Marino, M. T., Lollini, P.-L., Astolfi, a., Pandini, G., López-Guerrero, J. a., Schaefer, K.-L., Belfiore, a., Picci, P. and Scotlandi, K. (2011), 'Efficacy of and resistance to anti-IGF-1R therapies in Ewing's sarcoma is dependent on insulin receptor signaling.', *Oncogene* **30**(24), 2730–2740. 78

Gaspar, N., Hawkins, D. S., Dirksen, U., Lewis, I. J., Ferrari, S., Le Deley, M. C., Kovar, H., Grimer, R., Whelan, J., Claude, L., Delattre, O., Paulussen, M., Picci, P., Hall, K. S., Van Den Berg, H., Ladenstein, R., Michon, J., Hjorth, L., Judson, I., Luksch, R., Bernstein, M. L., Marec-Bérard, P., Brennan, B., Craft, A. W., Womer, R. B., Juergens, H. and Oberlin, O. (2015), 'Ewing sarcoma: Current management and future approaches through collaboration', *Journal of Clinical Oncology* **33**(27), 3036–3046. 4, 13, 79, 216, 219

Gatti, S., Bruno, S., Deregibus, M. C., Sordi, A., Cantaluppi, V., Tetta, C. and Camussi, G. (2011), 'Microvesicles derived from human adult mesenchymal stem cells protect against ischaemia-reperfusion-induced acute and chronic kidney injury.', *Nephrology Dialysis Transplantation* **26**(5), 1474–83. 29

Geier, B., Kurmashev, D., Kurmasheva, R. T. and Houghton, P. J. (2015), 'Preclinical Childhood Sarcoma Models: Drug Efficacy Biomarker Identification and Validation', *Frontiers in Oncology* **5**(August), 1–15. 14, 15

Gilbertson, R. J. and Rich, J. N. (2007), 'Making a tumour's bed: Glioblastoma stem cells and the vascular niche', *Nature Reviews Cancer* **7**(10), 733–736. 20

Gilligan, K. E. and Dwyer, R. M. (2017), 'Engineering exosomes for cancer therapy', *International Journal of Molecular Sciences* **18**(6). 34

Ginsberg, J. P., Goodman, P., Leisenring, W., Ness, K. K., Meyers, P. A., Wolden, S. L., Smith, S. M., Stovall, M., Hammond, S., Robison, L. L. and Oeffinger, K. C. (2010), 'Long-term survivors of childhood ewing sarcoma: Report from the childhood cancer survivor study', *Journal of the National Cancer Institute* **102**(16), 1272–1283. 7

Goldstein, S. D., Hayashi, M., Albert, C. M., Jackson, K. W. and Loeb, D. M. (2015), 'An orthotopic xenograft model with survival hindlimb amputation allows investigation of the effect of tumor microenvironment on sarcoma metastasis', *Clinical and Experimental Metastasis* **32**(7), 703–715. 14

González, I., Andreu, E. J., Panizo, A., Inoge, S., Fontalba, A., Ferna, L., Gaboli, M., Sierrasesu, L., Martı, S., Pardo, J. and Pro, F. (2004), 'Imatinib Inhibits Proliferation of Ewing Tumor Cells Mediated by the Stem Cell Factor / KIT Receptor Pathway, and Sensitizes Cells to Vincristine and Doxorubicin-Induced Apoptosis', *Clinical Cancer Research* **10**, 751–761. 79

Gordon, E. M., Sankhala, K. K., Chawla, N. and Chawla, S. P. (2016), 'Trabectedin for Soft Tissue Sarcoma: Current Status and Future Perspectives', *Advances in Therapy* **33**(7), 1055–1071. 10

Grange, C., Tapparo, M., Collino, F., Vitillo, L., Damasco, C., Deregibus, M. C., Tetta, C., Bussolati, B. and Camussi, G. (2011), 'Microvesicles released from human renal cancer stem cells stimulate angiogenesis and formation of lung premetastatic niche', *Cancer Research* **71**(15), 5346–5356. 32

Grohar, P. J., Griffin, L. B., Yeung, C., Chen, Q.-R., Pommier, Y., Khanna, C., Khan, J. and Helman, L. J. (2011), 'Ecteinascidin 743 Interferes with the Activity of EWS-FLI1 in Ewing Sarcoma Cells', *Neoplasia* **13**(2), 145–IN10. 10

Grünewald, T. G. P., Cidre-Aranaz, F., Surdez, D., Tomazou, E. M., de Álava, E., Kovar, H., Sorensen, P. H., Delattre, O. and Dirksen, U. (2018), 'Ewing sarcoma', *Nature Reviews Disease Primers* **4**(1), 5. 1, 4, 5, 7, 9, 15, 216, 219

Guenther, L. M., Dharia, N. V., Ross, L., Saur Conway, A., Robichaud, A. L., Catlett, J. L., Wechsler, C., Frank, E. S., Goodale, A. B., Church, A. J., Tseng, Y.-Y., Guha, R., McKnight, C., Janeway, K. A., Boehm, J. S., Mora, J., Davis, M. I., Alexe, G., Piccioni, F. and Stegmaier, K. (2018), 'A combination CDK4/6 and IGF1R inhibitor strategy for Ewing sarcoma', *Clinical Cancer Research* p. clincanres.0372.2018. 9, 15, 226

Guerzoni, C., Fiori, V., Terracciano, M., Manara, M. C., Moricoli, D., Pasello, M., Sciandra, M., Nicoletti, G., Gellini, M., Dominici, S., Chiodoni, C., Fornasari, P. M., Lollini, P. L., Colombo, M. P., Picci, P., Cianfriglia, M., Magnani, M. and Scotlandi, K. (2015), 'CD99 triggering in Ewing sarcoma delivers a lethal signal through p53 pathway reactivation and cooperates with doxorubicin', *Clinical Cancer Research* **21**(1), 146–156. 8, 207

Guise, T. A., Kozlow, W. M., Heras-Herzig, A., Padalecki, S. S., Yin, J. J. and Chirgwin, J. M. (2005), 'Molecular mechanisms of breast cancer metastases to bone', *Clinical Breast Cancer* **5**(2), 46–53. 12

Gupta, A. A., Pappo, A., Saunders, N., Hopyan, S., Ferguson, P., Wunder, J., O'Sullivan, B., Catton, C., Greenberg, M. and Blackstein, M. (2010), 'Clinical outcome of children and adults with localized ewing sarcoma: Impact of chemotherapy dose and timing of local therapy', *Cancer* **116**(13), 3189–3194. 79

Gupta, P. B., Onder, T. T., Jiang, G., Tao, K., Kuperwasser, C., Weinberg, R. A. and Lander, E. S. (2009), 'Identification of Selective Inhibitors of Cancer Stem Cells by High-Throughput Screening', *Cell* **138**(4), 645–659. 22

György, B., Szabó, T. G., Pásztói, M., Pál, Z., Misják, P., Aradi, B., László, V., Pállinger, É., Pap, E., Kittel, Á., Nagy, G., Falus, A. and Buzás, E. I. (2011), 'Membrane vesicles, current state-of-the-art: Emerging role of extracellular vesicles', *Cellular and Molecular Life Sciences* **68**(16), 2667–2688. 27

Haas, B. J., Dobin, A., Stransky, N., Li, B., Yang, X., Tickle, T., Bankapur, A., Ganote, C., Doak, T. G., Pochet, N., Sun, J., Wu, C. J. and Thomas, R. (2017), 'STAR-Fusion : Fast and Accurate Fusion Transcript Detection from RNA-Seq', *bioRxiv* . 85, 92, 125

Hadla, M., Palazzolo, S., Corona, G., Caligiuri, I., Canzonieri, V., Toffoli, G. and Rizzolio, F. (2016), 'Exosomes increase the therapeutic index of doxorubicin in breast and ovarian cancer mouse models', *Nanomedicine* **11**(18), 2431–2441. 34

Hanahan, D. and Weinberg, R. A. (2000), 'The Hallmarks of Cancer', *Cell* **100**, 57–70. 39

Hanahan, D. and Weinberg, R. A. (2011), 'Hallmarks of cancer: The next generation', *Cell* **144**(5), 646–674. 39, 78

Hattinger, C. M., Pötschger, U., Tarkkanen, M., Squire, J., Zielenska, M., Kiuru-Kuhlefelt, S., Kager, L., Thorner, P., Knuutila, S., Niggli, F. K., Ambros, P. F., Gadner, H. and Betts, D. R. (2002), 'Prognostic impact of chromosomal aberrations in Ewing tumours', *British journal of cancer* **86**(11), 1763–9. 3

Henne, W. M., Buchkovich, N. J., Zhao, Y. and Emr, S. D. (2012), 'The endosomal sorting complex ESCRT-II mediates the assembly and architecture of ESCRT-III helices', *Cell* **151**(2), 356–371. 26, 27, 233

Hermann, P. C., Huber, S. L., Herrler, T., Aicher, A., Ellwart, J. W., Guba, M., Bruns, C. J. and Heeschen, C. (2007), 'Distinct Populations of Cancer Stem Cells Determine Tumor Growth and Metastatic Activity in Human Pancreatic Cancer', *Cell Stem Cell* **1**(3), 313–323. 83

Herreros-Villanueva, M., Bujanda, L., Billadeau, D. D. and Zhang, J. S. (2014), 'Embryonic stem cell factors and pancreatic cancer', *World Journal of Gastroenterology* **20**(9), 2247–2254. 21

Hirschmann-Jax, C., Foster, A. E., Wulf, G. G., Nuchtern, J. G., Jax, T. W., Gobel, U., Goodell, M. A. and Brenner, M. K. (2004), 'A distinct "side population" of cells with high drug efflux capacity in human tumor cells.', *Proceedings of the National Academy of Sciences of the United States of America* **101**(39), 14228–33. 21

Hnisz, D., Abraham, B. J., Lee, T. I., Lau, A., Saint-André, V., Sigova, A. A., Hoke, H. A. and Young, R. A. (2013), 'XSuper-enhancers in the control of cell identity and disease', *Cell* **155**(4), 934. 219

Hoen, E. N. t., van der Vlist, E. J., Aalberts, M., Mertens, H. C., Bosch, B. J., Bartelink, W., Mastrobattista, E., van Gaal, E. V., Stoorvogel, W., Arkesteijn, G. J. and Wauben, M. H. (2012), 'Quantitative and qualitative flow cytometric analysis of nanosized cell-derived membrane vesicles', *Nanomedicine: Nanotechnology, Biology, and Medicine* **8**(5), 712–720. 212

Hong, C. S., Funk, S., Muller, L., Boyiadzis, M. and Whiteside, T. L. (2016), 'Isolation of biologically active and morphologically intact exosomes from plasma of patients with cancer', *Journal of Extracellular Vesicles* **5**(1), 1–11. 132

Hong, S.-H., Tilan, J. U., Galli, S., Izycka-Swieszewska, E., Polk, T., Horton, M., Mahajan, A., Christian, D., Jenkins, S., Acree, R., Connors, K., Ledo, P., Lu, C., Lee, Y.-C., Rodriguez, O., Toretsky, J. A., Albanese, C. and Kitlinska, J. (2015), 'High neuropeptide Y release associates with Ewing sarcoma bone dissemination - in vivo model of site-specific metastases', *Oncotarget* **6**(9), 7151–7165. 12

Horibe, S., Tanahashi, T., Kawauchi, S., Murakami, Y. and Rikitake, Y. (2018), 'Mechanism of recipient cell-dependent differences in exosome uptake', *BMC Cancer* **18**(1), 1–9. 130, 211

Hornick, J. L. (2014), 'Novel uses of immunohistochemistry in the diagnosis and classification of soft tissue tumors', *Modern Pathology* **27**(SUPPL. 1), S47–S63. 1

Houghton, P. J., Morton, C. L., Tucker, C., Payne, D., Favours, E., Cole, C., Gorlick, R., Kolb, E. A., Zhang, W., Lock, R., Carol, H., Tajbakhsh, M., Reynolds, C. P., Maris, J. M., Courtright, J., Keir, S. T., Friedman, H. S., Stopford, C., Zeidner, J., Wu, J., Liu, T., Billups, C. A., Khan, J., Ansher, S., Zhang, J. and Smith, M. A. (2007), 'The Pediatric Preclinical Testing Program: Description of Models and Early Testing Results', *Pediatr Blood Cancer* **49**, 928–940. 15

Hrdlickova, R., Toloue, M., Tian, B., Genetics, M. and New, R. (2017), 'RNA-Seq methods for transcriptome analysis', *Wiley Interdiscip Rev RNA* **8**(1), 1–24. 128

Hu-Lieskovan, S., Zhang, J., Wu, L., Shimada, H., Schofield, D. E. and Triche, T. J. (2005), 'EWS-FLI1 fusion protein up-regulates critical genes in neural crest development and is responsible for the observed phenotype of Ewing's family of tumors', *Cancer Research* **65**(11), 4633–4644. 17

Hu, Y., Yan, C., Mu, L., Huang, K., Li, X., Tao, D., Wu, Y. and Qin, J. (2015), 'Fibroblast-derived exosomes contribute to chemoresistance through priming cancer stem cells in colorectal cancer', *PLoS ONE* **10**(5), 1–17. 123, 130, 204, 229

Huang, D. W., Sherman, B. T. and Lempicki, R. A. (2009*a*), 'Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists', *Nucleic Acids Research* **37**(1), 1–13. 94

Huang, D. W., Sherman, B. T. and Lempicki, R. A. (2009*b*), 'Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources.', *Nature protocols* **4**(1), 44–57. 94

Huang, H. Y., Illei, P. B., Zhao, Z., Mazumdar, M., Huvos, A. G., Healey, J. H., Wexler, L. H., Gorlick, R., Meyers, P. and Ladanyi, M. (2005), 'Ewing sarcomas with p53 mutation or p16/p14ARF homozygous deletion: A highly lethal subset associated with poor chemoresponse', *Journal of Clinical Oncology* **23**(3), 548–558. 3, 219

Im, Y.-h., Kim, H. T., Lee, C., Poulin, D., Welford, S., Sorensen, P. H. B., Denny, C. T. and Kim, S.-j. (2000), 'EWS-FLI1, EWS-ERG, and EWS-ETV1 oncoproteins of Ewing tumor family all suppress transcription of Transforming Growth Factor β Type II Receptor Gene', *Cancer Research* **60**(301), 1536–1540. 3, 4

Irollo, E. and Pirozzi, G. (2013), 'CD133: To be or not to be, is this the real question?', *American Journal of Translational Research* **5**(6), 563–581. 222

Jain, A., Zhang, J. and Rabbitts, T. (2017), 'Clinics in Oncology LINGO-1 is a New Therapy Target and Biomarker for Ewing Sarcoma', *Clinics in Oncology* **2**, 1–4. 1, 218

Jaishankar, S., Zhang, J., Roussel, M. F. and Baker, S. J. (1999), 'Transforming activity of EWS/FLI is not strictly dependent upon DNA-binding activity', *Oncogene* **18**, 5592–5597. 2

Jan, A. T., Malik, M. A., Rahman, S., Yeo, H. R., Lee, E. J., Abdullah, T. S. and Choi, I. (2017), 'Perspective insights of exosomes in neurodegenerative diseases: A critical appraisal', *Frontiers in Aging Neuroscience* **9**(SEP), 1–8. 29

Jang, S. C., Kim, O. Y., Yoon, C. M., Choi, D. S., Roh, T. Y., Park, J., Nilsson, J., Lötvall, J., Kim, Y. K. and Gho, Y. S. (2013), 'Bioinspired exosome-mimetic nanovesicles for targeted delivery of chemotherapeutics to malignant tumors', *ACS Nano* **7**(9), 7698–7710. 34

Jelski, W. and Szmitkowski, M. (2008), 'Alcohol dehydrogenase (ADH) and aldehyde dehydrogenase (ALDH) in the cancer diseases', *Clinica Chimica Acta* **395**(1-2), 1–5. 210

Jenjaroenpun, P., Kremenska, Y., Nair, V. M., Kremenskoy, M., Joseph, B. and Kurochkin, I. V. (2013), 'Characterization of RNA in exosomes secreted by human breast cancer cell lines using next-generation sequencing.', *PeerJ* **1**(November), e201. 209, 230

Jiang, K., Dong, C., Yin, Z., Li, R., Wang, Q. and Wang, L. (2018), 'The critical role of exosomes in tumor biology', *Journal of Cellular Biochemistry* **1**(September), 1–13. 232, 234

Jiang, X., Gwye, Y., Russell, D., Cao, C., Douglas, D., Hung, L., Kovar, H., Triche, T. J. and Lawlor, E. R. (2010), 'CD133 expression in chemo-resistant Ewing sarcoma cells.', *BMC cancer* **10**, 116. 24, 122, 221

Jin, L., Hope, K. J., Zhai, Q., Smadja-Joffe, F. and Dick, J. E. (2006), 'Targeting of CD44 eradicates human acute myeloid leukemic stem cells', *Nature Medicine* **12**(10), 1167–1174. 22

Johnstone, R. M., Adam, M., Hammond, J. R., Orr, L. and Turbide, C. (1987), 'Vesicle formation during reticulocyte maturation. Association of plasma membrane activities with released vesicles (exosomes).', *Journal of Biological Chemistry* **262**(19), 9412–9420. 26

Juergens, C., Weston, C., Lewis, I., Whelan, J., Paulussen, M., Oberlin, O., Michon, J., Zoubek, A., Juergens, H. and Craft, A. (2006), 'Safety Assessment of Intensive Induction With Vincristine, Ifosfamide, Doxorubicin, and Etoposide (VIDE) in the Treatment of Ewing Tumors in the EURO-E.W.I.N.G. 99 Clinical Trial', *Pediatr Blood Cancer* **47**, 22–29. 5

Juergens, H., Daw, N. C., Geoerger, B., Ferrari, S., Villarroel, M., Aerts, I., Whelan, J., Dirksen, U., Hixon, M. L., Yin, D., Wang, T., Green, S., Paccagnella, L. and Gualberto, A. (2011), 'Preliminary efficacy of the anti-insulin-like growth factor type 1 receptor antibody figitumumab in patients with refractory Ewing sarcoma', *Journal of Clinical Oncology* **29**(34), 4534–4540. 9

Kalra, H., Adda, C. G., Liem, M., Ang, C. S., Mechler, A., Simpson, R. J., Hulett, M. D. and Mathivanan, S. (2013), 'Comparative proteomics evaluation of plasma exosome isolation techniques and assessment of the stability of exosomes in normal human blood plasma', *Proteomics* **13**(22), 3354–3364. 132

Karski, E. E., Matthay, K. K., Neuhaus, J. M., Goldsby, R. E. and DuBois, S. G. (2013), 'Characteristics and outcomes of patients with Ewing sarcoma over 40 years of age at diagnosis', *Cancer Epidemiology* **37**(1), 29–33. 79

Kassambara, A., Kosinski, M., Biecek, P. and Fabian, S. (2018), 'Drawing Survival Curves using 'ggplot2'', *CRAN* **0.4.2**. 56

Kaur, S., Elkahloun, A. G., Arakelyan, A., Young, L., Myers, T. G., Otaizo-Carrasquero, F., Wu, W., Margolis, L. and Roberts, D. D. (2018), 'CD63, MHC class 1, and CD47 identify subsets of extracellular vesicles containing distinct populations of noncoding RNAs', *Scientific Reports* **8**(1), 1–17. 209

Kennedy, A. L., Vallurupalli, M., Chen, L., Crompton, B., Cowley, G., Vazquez, F., Weir, B. A., Tsherniak, A., Parasuraman, S., Kim, S., Alexe, G. and Stegmaier, K. (2015), 'Functional, chemical genomic, and super-enhancer screening identify sensitivity to cyclin D1/CDK4 pathway inhibition in Ewing sarcoma', *Oncotarget* **6**(30). 219

Kharaziha, P., Ceder, S., Li, Q. and Panaretakis, T. (2012), 'Tumor cell-derived exosomes: A message in a bottle', *Biochimica et Biophysica Acta - Reviews on Cancer* **1826**(1), 103–111. 27, 28

Kim, J., Lee, H., Park, K., Choi, Y., Nam, J. and Hong, I. (2016), 'CWP232228 targets liver cancer stem cells through Wnt/β-catenin signaling: a novel therapeutic approach for liver cancer treatment', *Oncotarget* **7**(15), 20395–20409. 23

Kim, Y. J., mi Yoo, S., Park, H. H., Lim, H. J., Kim, Y. L., Lee, S., Seo, K. W. and Kang, K. S. (2017), 'Exosomes derived from human umbilical cord blood mesenchymal stem cells stimulates rejuvenation of human skin', *Biochemical and Biophysical Research Communications* **493**(2), 1102–1108. 133, 213

Kitamura, Y. and Hirota, S. (2004), 'Kit as a human oncogenic tyrosine kinase', *Cellular and Molecular Life Sciences* **61**(23), 2924–2931. 80

Kolde, R. (2015), 'pheatmap: Pretty Heatmaps', *CRAN* . 95

Kolde, R. (2019), 'Package 'pheatmap''. 93

Komuro, H., Saihara, R., Shinya, M., Takita, J., Kaneko, S., Kaneko, M. and Hayashi, Y. (2007), 'Identification of side population cells (stem-like cell population) in pediatric solid tumor cell lines', *Journal of Pediatric Surgery* **42**(12), 2040–2045. 25, 83, 122, 123, 221

Kosaka, N., Iguchi, H., Yoshioka, Y., Takeshita, F., Matsuki, Y. and Ochiya, T. (2010), 'Secretory mechanisms and intercellular transfer of microRNAs in living cells', *Journal of Biological Chemistry* **285**(23), 17442–17452. 211

Kramer, N., Walzl, A., Unger, C., Rosner, M., Krupitza, G., Hengstschläger, M. and Dolznig, H. (2013), 'In vitro cell migration and invasion assays', *Mutation Research - Reviews in Mutation Research* **752**(1), 10–24. 39

Kreppel, M., Aryee, D. N. T., Schaefern, K. L., Amann, G., Kofler, R., Poremba, C. and Kovar, H. (2006), 'Suppression of KCMF1 by constitutive high CD99 expression is involved in the migratory ability of Ewing's sarcoma cells', *Oncogene* **25**(19), 2795–2800. 8, 78

Krivtsov, A. V., Twomey, D., Feng, Z., Stubbs, M. C., Wang, Y., Faber, J., Levine, J. E., Wang, J., Hahn, W. C., Gilliland, D. G., Golub, T. R. and Armstrong, S. A. (2006), 'Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9', *Nature* **442**(7104), 818–822. 19

Krook, M. A., Nicholls, L. A., Scannell, C. A., Chugh, R., Thomas, D. G. and Lawlor, E. R. (2014), 'Stress-induced CXCR4 promotes migration and invasion of Ewing Sarcoma', *Molecular cancer research* **12**(6), 953–64. 78

Kucharzewska, P., Christianson, H. C., Welch, J. E., Svensson, K. J., Fredlund, E., Ringner, M., Morgelin, M., Bourseau-Guilmain, E., Bengzon, J. and Belting, M. (2013), 'Exosomes reflect the hypoxic status of glioma cells and mediate hypoxia-dependent activation of vascular cells during tumor development', *Proceedings of the National Academy of Sciences* **110**(18), 7312–7317. 31, 232

Kukurba, K. R. and Montgomery, S. B. (2016), 'RNA Sequencing and Analysis', *Cold Spring Harb Protoc.* **2015**(11), 951–969. 84, 85

Lachenal, G., Pernet-Gallay, K., Chivet, M., Hemming, F. J., Belly, A., Bodon, G., Blot, B., Haase, G., Goldberg, Y. and Sadoul, R. (2011), 'Release of exosomes from differentiated neurons and its regulation by synaptic glutamatergic activity', *Molecular and Cellular Neuroscience* **46**(2), 409–418. 29

Lapidot, T., Sirard, C., Vormoor, J., Murdoch, B., Hoang, T., Caceres-Cortes, J., Minden, M., Paterson, B., Caligiuri, M. A. and Dick, J. E. (1994), 'A cell initiating human acute myeloid leukaemia after transplantation into SCID mice', *Nature* **367**(6464), 645–648. 22, 83

Larsson, O., Girnita, A. and Girnita, L. (2005), 'Role of insulin-like growth factor I receptor signalling in cancer', *British Journal of Cancer* **92**(12), 2097–2101. 9

Lau, Y. S., Adamopoulos, I. E., Sabokbar, A., Giele, H., Gibbons, C. L. and Athanasou, N. A. (2007), 'Cellular and humoral mechanisms of osteoclast formation in Ewing's sarcoma', *British Journal of Cancer* **96**(11), 1716–1722. 11

Laulagnier, K., Motta, C., Hamdi, S., Roy, S., Fauvelle, F., Pageaux, J.-F., Kobayashi, T., Salles, J.-P., Perret, B., Bonnerot, C. and Record, M. (2004), 'Mast cell- and dendritic cell-derived exosomes display a specific lipid composition and an unusual membrane organization', *Biochemical Journal* **380**(1), 161–171. 28

Le Deley, M. C., Paulussen, M., Lewis, I., Brennan, B., Ranft, A., Whelan, J., Le Teuff, G., Michon, J., Ladenstein, R., Marec-Bérard, P., Van Den Berg, H., Hjorth, L., Wheatley, K., Judson, I., Juergens, H., Craft, A., Oberlin, O. and Dirksen, U. (2014), 'Cyclophosphamide compared with ifosfamide in consolidation treatment of standard-risk Ewing sarcoma: Results of the randomized noninferiority Euro-EWING99-R1 trial', *Journal of Clinical Oncology* **32**(23), 2440–2448. 5

Lee, C. H., Lee, J. M., Wu, D. C., Goan, Y. G., Chou, S. H., Wu, I. C., Kao, E. L., Chan, T. F., Huang, M. C., Chen, P. S., Lee, C. Y., Huang, C. T., Huang, H. L., Hu, C. Y., Hung, Y. H. and Wu, M. T. (2008), 'Carcinogenetic impact of ADH1B and ALDH2 genes on squamous cell carcinoma risk of the esophagus with regard to the consumption of alcohol, tobacco and betel quid', *International Journal of Cancer* **122**(6), 1347–1356. 209

Lee, J., Hoang, B. H., Ziogas, A. and Zell, J. A. (2010), 'Analysis of prognostic factors in Ewing sarcoma using a population-based cancer registry', *Cancer* **116**(8), 1964–1973. 79

Lee, T. H., Bolontrade, M. F., Worth, L. L., Guan, H., Ellis, L. M. and Kleinerman, E. S. (2006), 'Production of VEGF165 by Ewing's sarcoma cells induces vasculogenesis and the incorporation of CD34+ stem cells into the expanding tumor vasculature', *International Journal of Cancer* **119**(4), 839–846. 13

Lessnick, S. L., Dacwag, C. S. and Golub, T. R. (2002), 'The Ewing's sarcoma oncoprotein EWS/FLI induces a p53-dependent growth arrest in primary human fibroblasts', *Cancer Cell* **1**(4), 393–401. 17

Lessnick, S. L. and Ladanyi, M. (2012), 'Molecular pathogenesis of Ewing Sarcoma: New Therapeutic and Transcriptional Targets', *Annu Rev Pathol* **7**, 145–159. 2

Leuchte, K., Altvater, B., Hoffschlag, S., Potratz, J., Meltzer, J., Clemens, D., Luecke, A., Hardes, J., Dirksen, U., Juergens, H., Kailayangiri, S. and Rossig, C. (2014), 'Anchorage-independent growth of ewing sarcoma cells under serum-free conditions is not associated with stem-cell like phenotype and function', *Oncology Reports* **32**(2), 845–852. 25, 122, 221

Li, C., Heidt, D. G., Dalerba, P., Burant, C. F., Zhang, L., Adsay, V., Wicha, M., Clarke, M. F. and Simeone, D. M. (2007), 'Identification of Pancreatic Cancer Stem Cells', *American Association of Cancer Research Journals* **67**(3), 1030–1038. 19

Li, H., Zhang, P., Sun, X., Sun, Y., Shi, C., Liu, H. and Liu, X. (2015), 'MicroRNA-181a regulates epithelial-mesenchymal transition by targeting PTEN in drug-resistant lung adenocarcinoma cells', *International Journal of Oncology* **47**(4), 1379–1392. 34

Li, M., Sun, Q. and Wang, X. (2017), 'Transcriptional landscape of human cancers', *Oncotarget* **8**(21), 34534–34551. 210

Li, X.-Q., Liu, J.-T., Fan, L.-L., Liu, Y., Cheng, L., Wang, F., Yu, H.-Q., Gao, J., Wei, W., Wang, H. and Sun, G.-P. (2016), 'Exosomes derived from gefitinib-treated EGFR-mutant lung cancer cells alter cisplatin sensitivity via up-regulating autophagy', *Oncotarget* **7**(17), 24585–24595. 141

Liang, G., Kan, S., Zhu, Y., Feng, S., Feng, W. and Gao, S. (2018), 'Engineered exosome-mediated delivery of functionally active miR-26a and its enhanced suppression effect in HepG2 cells', *International Journal of Nanomedicine* **13**, 585–599. 133, 213

Lin, L., Huang, M., Shi, X., Mayakonda, A., Hu, K., Jiang, Y.-Y., Guo, X., Chen, L., Pang, B., Doan, N., Said, J. W., Xie, J., Gery, S., Cheng, X., Lin, Z., Li, J., Berman, B. P., Yin, D., Lin, D.-C. and Koeffler, H. (2019), 'Super-enhancer-associated MEIS1 promotes transcriptional dysregulation in Ewing sarcoma in co-operation with EWS-FLI1', *Nucleic Acids Research* **47**(3), 1255–1267. 219

Lipinski, M., Braham, K., Philip, I., Wiels, J., Philip, T., Dellagi, K., Goridis, C., Lenoir, G. M. and Tursz, T. (1986), 'Phenotypic characterization of Ewing sarcoma cell lines with monoclonal antibodies', *Journal of Cellular Biochemistry* **31**(4), 289–296. 17

Lissat, A., Chao, M. M. and Kontny, U. (2012), 'Targeted Therapy in Ewing Sarcoma', *ISRN Oncology* **2012**, 1–9. 80

Liu, S.-Y. and Zheng, P.-S. (2013), 'High aldehyde dehydrogenase activity identifies cancer stem cells in human cervical cancer.', *Oncotarget* **4**(12), 2462–75. 21

Liu, X., Yuan, W., Yang, L., Li, J. and Cai, J. (2018), 'miRNA Profiling of Exosomes from Spontaneous Hypertensive Rats Using Next-Generation Sequencing', *Journal of Cardiovascular Translational Research* pp. 1–9. 133

Llombart-Bosch, A., Machado, I., Navarro, S., Bertoni, F., Bacchini, P., Alberghini, M., Karzeladze, A., Savelov, N., Petrov, S., Alvarado-Cabrero, I., Mihaila, D., Terrier, P., Lopez-Guerrero, J. A. and Picci, P. (2009), 'Histological heterogeneity of Ewing's sarcoma/PNET: An immunohistochemical analysis of 415 genetically confirmed cases with clinical support', *Virchows Archiv* **455**(5), 397–411. 1

Llorente, A., Skotland, T., Sylvänne, T., Kauhanen, D., Róg, T., Orłowski, A., Vattulainen, I., Ekroos, K. and Sandvig, K. (2013), 'Molecular lipidomics of exosomes released by PC-3 prostate cancer cells', *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids* **1831**(7), 1302–1309. 27

Lobb, R. J., Becker, M., Wen, S. W., Wong, C. S. F., Wiegmans, A. P., Leimgruber, A. and Möller, A. (2015), 'Optimized exosome isolation protocol for cell culture supernatant and human plasma', *Journal of Extracellular Vesicles* **1**(27031), 1–11. 131

Lötvall, J., Hill, A. F., Hochberg, F., Buzás, E. I., Di Vizio, D., Gardiner, C., Gho, Y. S., Kurochkin, I. V., Mathivanan, S., Quesenberry, P., Sahoo, S., Tahara, H., Wauben, M. H., Witwer, K. W. and Théry, C. (2014), 'Minimal experimental requirements for definition of extracellular vesicles and their functions: a position statement from the International Society for Extracellular Vesicles.', *Journal of extracellular vesicles* **3**, 26913. 131

Love, M. I., Huber, W. and Anders, S. (2014), 'Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.', *Genome biology* **15**(12), 550. 86, 93, 128

Lowry, O. H., Rosebrough, N. J., Farr, A. L. and Randall, R. J. (1951), 'Protein measurement with the folin phenol reagent', *Readings* **193**(1), 265–275. 141

Lu, J. F., Luk, F., Gong, J., Jaiswal, R., Grau, G. E. and Bebawy, M. (2013), 'Microparticles mediate MRP1 intercellular transfer and the re-templating of intrinsic resistance pathways', *Pharmacological Research* **76**, 77–83. 232

Lu, Y., Wu, D., Wang, J., Li, Y., Chai, X. and Kang, Q. (2016), 'MiR-320a regulates cell proliferation and apoptosis in multiple myeloma by targeting pre-B-cell leukemia transcription factor 3', *Biochemical and Biophysical Research Communications* **473**(4), 1315–1320. 209, 230

Lv, M.-m., Zhu, X.-y., Chen, W.-x., Zhong, S.-l., Hu, Q., Ma, T.-f., Zhang, J., Chen, L., Tang, J.-h. and Zhao, J.-h. (2014), 'Exosomes mediate drug resistance transfer in MCF-7 breast cancer cells and a probable mechanism is delivery of P-glycoprotein', *Tumor Biology* **35**(11), 10773–10779. 204, 205, 232

Machado, I., Yoshida, A., Morales, M. G. N., Abrahão-Machado, L. F., Navarro, S., Cruz, J., Lavernia, J., Parafioriti, A., Picci, P. and Llombart-Bosch, A. (2018), 'Review with novel markers facilitates precise categorization of 41 cases of diagnostically challenging, "undifferentiated small round cell tumors". A clinicopathologic, immunophenotypic and molecular analysis', *Annals of Diagnostic Pathology* **34**, 1–12. 1

Manara, M. C., Landuzzi, L., Nanni, P., Nicoletti, G., Zambelli, D., Lollini, P. L., Nanni, C., Hofmann, F., García-Echeverría, C., Picci, P. and Scotlandi, K. (2007), 'Preclinical in vivo study of new insulin-like growth factor-I receptor-specific inhibitor in Ewing's sarcoma', *Clinical Cancer Research* **13**(4), 1322–1330. 9

Manara, M. C., Pasello, M. and Scotlandi, K. (2018), 'CD99: A Cell Surface Protein with an Oncojanus Role in Tumors', *Genes* **9**(159), 1–17. 78, 79, 227

Manara, M. C., Terracciano, M., Mancarella, C., Sciandra, M., Guerzoni, C., Pasello, M., Grilli, A., Zini, N., Picci, P., Colombo, M. P., Morrione, A., Scotlandi, K., Manara, M. C., Terracciano, M., Mancarella, C., Sciandra, M., Guerzoni, C., Pasello, M., Grilli, A., Zini, N., Picci, P., Colombo, M. P., Morrione, A. and Scotlandi, K. (2016), 'CD99 triggering induces methuosis of Ewing sarcoma cells through IGF-1R/RAS/Rac1 signaling', *Oncotarget* **7**(48), 79925–79942. 8, 9, 207

Mancarella, C., Pasello, M., Ventura, S., Grilli, A., Calzolari, L., Toracchio, L., Lollini, P.-L., Donati, D. M., Picci, P., Ferrari, S. and Scotlandi, K. (2018), 'Insulin-Like Growth Factor 2 mRNA-Binding Protein 3 is a Novel Post-Transcriptional Regulator of Ewing Sarcoma Malignancy', *Clinical Cancer Research* **24**(15), 3704–3716. 9

Mangs, A. H. and Morris, B. J. (2007), 'The Human Pseudoautosomal Region (PAR): Origin, Function and Future', *Current Genomics* **8**, 129–136. 127

Marina, N. M., Liu, Q., Donaldson, S. S., Sklar, C. A., Armstrong, G. T., Oeffinger, K. C., Leisenring, W. M., Ginsberg, J. P., Henderson, T. O., Neglia, J. P., Stovall, M. A., Yasui, Y., Randall, R. L., Geller, D. S., Robison, L. L. and Ness, K. K. (2018), 'Longitudinal Follow-Up of Adult Survivors of Ewing Sarcoma: A Report from the Childhood Cancer Survivor Study (CCSS)', *Cancer* **123**(13), 2551–2560. 7

Martin, M. (2011), 'Cutadapt removes adapter sequences from high-throughput sequencing reads', *EMBnet.journal* **17**(1), 10. 91

Martins, A. S., Mackintosh, C., Herrero Martín, D., Campos, M., Hernández, T., Ordóñez, J. L. and De Alava, E. (2006), 'Insulin-like growth factor I receptor pathway inhibition by ADW742, alone or in combination with imatinib, doxorubicin, or vincristine, is a novel therapeutic approach in Ewing tumor', *Clinical Cancer Research* **12**(11 I), 3532–3540. 9

Martins, S., Rosa, I. M., Odete, A. B. and Henriques, A. G. (2018), 'Exosome isolation from distinct biofluids using precipitation and column-based approaches', *PLoS ONE* pp. 1–16. 131, 132

Marturano-Kruik, A., Villasante, A., Yaeger, K., Ambati, S. R., Chramiec, A., Raimondi, M. T. and Vunjak-Novakovic, G. (2018), 'Biomechanical regulation of drug sensitivity in an engineered model of human tumor', *Biomaterials* **150**, 150–161. 15, 223, 226

Mathivanan, S., Lim, J. W. E., Tauro, B. J., Ji, H., Moritz, R. L. and Simpson, R. J. (2010), 'Proteomics Analysis of A33 Immunoaffinity-purified Exosomes Released from the Human Colon Tumor Cell Line LIM1215 Reveals a Tissue-specific Protein Signature', *Molecular & Cellular Proteomics* **9**(2), 197–208. 27

May, W. A., Gishizky, M. L., Lessnick, S. L., Lunsford, L. B., Lewis, B. C., Delattre, O., Zucman, J., Thomas, G. and Denny, C. T. (1993), 'Ewing Sarcoma 11;22 translocation produces a chimeric transcription factor that requires the DNA-binding domain encoded by FLI1 for transformation', *ChemPlusChem* **90**(12), 5752–5756. 2, 17, 206

McNally, R. J., Blakey, K., Parslow, R. C., James, P. W., Pozo, B. G., Stiller, C., Vincent, T. J., Norman, P., McKinney, P. A., Murphy, M. F., Craft, A. W. and Feltbower, R. G. (2012), 'Small-area analyses of bone cancer diagnosed in Great Britain provide clues to aetiology', *BMC Cancer* **12**(1), 1. 1

Melo, S., Villanueva, A., Moutinho, C., Davalos, V., Spizzo, R., Ivan, C., Rossi, S., Setien, F., Casanovas, O., Simo-Riudalbas, L., Carmona, J., Carrere, J., Vidal, A., Aytes, A., Puertas, S., Ropero, S., Kalluri, R., Croce, C. M., Calin, G. A. and Esteller, M. (2011), 'Small molecule enoxacin is a cancer-specific growth inhibitor that acts by enhancing TAR RNA-binding protein 2-mediated microRNA processing', *Proceedings of the National Academy of Sciences* **108**(11), 4394–4399. 24

Miettinen, M. and Lasota, J. (2005), 'KIT (CD117): A review on expression in normal and neoplastic tissues, and mutations and their clinicopathologic correlation', *Applied Immunohistochemistry and Molecular Morphology* **13**(3), 205–220. 80

Miller, I. V., Raposo, G., Welsch, U., Prazeres da Costa, O., Thiel, U., Lebar, M., Maurer, M., Bender, H. U., von Luettichau, I., Richter, G. H. S., Burdach, S. and Grunewald, T. G. P. (2013), 'First identification of Ewing's sarcoma-derived extracellular vesicles and exploration of their biological and potential diagnostic implications', *Biology of the Cell* **105**(7), 289–303. 35, 129, 204, 208, 209, 212, 227, 228, 230

Minas, T. Z., Surdez, D., Javaheri, T., Tanaka, M., Howarth, M., Kang, H.-J., Han, J., Han, Z.-Y., Sax, B., Kream, B. E., Hong, S.-H., Çelik, H., Tirode, F., Tuckermann, J., Toretsky, J. A., Kenner, L., Kovar, H., Lee, S., Sweet-Cordero, E. A., Nakamura, T., Moriggl, R., Delattre, O. and Üren, A. (2017), 'Combined experience of six independent laboratories attempting to create an Ewing sarcoma mouse model', *Oncotarget* **8**(21), 34141–34163. 13

Miserocchi, G., Mercatali, L., Liverani, C., De Vita, A., Spadazzi, C., Pieri, F., Bongiovanni, A., Recine, F., Amadori, D. and Ibrahim, T. (2017), 'Management and potentialities of primary cancer cultures in preclinical and translational studies', *Journal of Translational Medicine* **15**(1), 1–16. 14

Moore, N., Houghton, J. and Lyle, S. (2012), 'Slow-Cycling Therapy-Resistant Cancer Cells', *Stem Cells and Development* **21**(10), 1822–1830. 21

Mugneret, F., Lizard, S., Aurias, A. and Turc-Carel, C. (1988), 'Chromosomes in Ewing's Sarcoma. II. Nonrandom Additional Changes, Trisomy 8 and der(16)t(1;16)', *Cancer Genet Cytogenet* **32**, 239–245. 3

Mulcahy, L. A., Pink, R. C., Raul, D. and Carter, F. (2014), 'Routes and mechanisms of extracellular vesicle uptake', *Journal of Extracellular Vesicles* **3**, 1–14. 29, 233

Munich, S., Sobo-Vujanovic, A., Buchser, W. J., Beer-Stolz, D. and Vujanovic, N. L. (2012), 'Dendritic cell exosomes directly kill tumor cells and activate natural killer cells via TNF superfamily ligands', *OncoImmunology* **1**(7), 1074–1083. 29, 233

Murray, J., Valli, E., Yu, D. M. T., Truong, A. M., Gifford, A. K., Eden, G. L., Gamble, L. D., Hanssen, K. M., Flemming, C. L., Tan, A., Tivnan, A., Allan, S., Saletta, F., Cheung, L., Ruhle, M., Schuetz, J. D., Henderson, M. J., Byrne, J. A., Norris, M. D., Haber, M. and Fletcher, J. I. (2017), 'Suppression of the ATP-binding cassette transporter ABCC4 impairs neuroblastoma tumor growth and sensitizes to irinotecan in vivo.', *European Journal of Cancer* **83**, 132–141. 54

Muz, B., de la Puente, P., Azab, F. and Azab, A. K. (2015), 'The role of hypoxia in cancer progression, angiogenesis, metastasis, and resistance to therapy', *Hypoxia* **3**, 83–92. 12

Nakatani, F., Tanaka, K., Sakimura, R., Matsumoto, Y., Matsunobu, T., Li, X., Hanada, M., Okada, T. and Iwamoto, Y. (2003), 'Identification of p21WAF1/CIP1 as a direct target of EWS-Fli1 oncogenic fusion protein', *Journal of Biological Chemistry* **278**(17), 15105–15115. 3

Nordin, J. Z., Lee, Y., Vader, P., Mäger, I., Johansson, H. J., Heusermann, W., Wiklander, O. P., Hällbrink, M., Seow, Y., Bultema, J. J., Gilthorpe, J., Davies, T., Fairchild, P. J., Gabrielsson, S., Meisner-Kober, N. C., Lehtiö, J., Smith, C. I., Wood, M. J. and Andaloussi, S. E. (2015), 'Ultrafiltration with size-exclusion liquid chromatography for high yield isolation of extracellular vesicles preserving intact biophysical and functional properties', *Nanomedicine: Nanotechnology, Biology, and Medicine* **11**(4), 879–883. 131

Nowell, P. C. (1976), 'The Clonal Evolution of Tumor Cell Populations', *Science* **194**(4260), 23–28. 19

Odri, G. A., Dumoucel, S., Picarda, G., Battaglia, S., Lamoureux, F., Corradini, N., Rousseau, J., Tirode, F., Laud, K., Delattre, O., Gouin, F., Heymann, D. and Redini, F. (2010), 'Zoledronic acid as a new adjuvant therapeutic strategy for Ewing's sarcoma patients', *Cancer Research* **70**(19), 7610–7619. 6, 12

Odri, G., Kim, P. P., Lamoureux, F., Charrier, C., Battaglia, S., Amiaud, J., Heymann, D., Gouin, F. and Redini, F. (2014), 'Zoledronic acid inhibits pulmonary metastasis dissemination in a preclinical model of Ewing's sarcoma via inhibition of cell migration', *BMC Cancer* **14**(1). 12

Ohno, T., Rao, V. N. and Reddy, E. S. P. (1993), 'EWS/Fli-1 Chimeric Protein Is a Transcriptional Activator', *Cancer Research* **53**(24), 5859–5863. 2

Okada, Y., Momozawa, Y., Sakaue, S., Kanai, M., Ishigaki, K., Akiyama, M., Kishikawa, T., Arai, Y., Sasaki, T., Kosaki, K., Suematsu, M., Matsuda, K., Yamamoto, K., Kubo, M., Hirose, N. and Kamatani, Y. (2018), 'Deep whole-genome sequencing reveals recent selection signatures linked to evolution and disease risk of Japanese', *Nature Communications* **9**(1), 1–10. 209, 230

Olmos, D., Martins, A. S., Jones, R. L., Alam, S., Scurr, M. and Judson, I. R. (2011), 'Targeting the Insulin-Like Growth Factor 1 Receptor in Ewing's Sarcoma: Reality and Expectations', *Sarcoma* **2011**, 1–13. 9

Ordóñez, J. L., Amaral, A. T., Carcaboso, A. M., Herrero-Martín, D., García-Macías, M. d. C., Sevillano, V., Alonso, D., Pascual-Pasto, G., San-Segundo, L., Vila-Ubach, M., Rodrigues, T., Fraile, S., Teodosio, C., Mayo-Iscar, A., Aracil, M., Galmarini, C. M., Tirado, O. M., Mora, J. and de Álava, E. (2015), 'The PARP inhibitor olaparib enhances the sensitivity of Ewing sarcoma to trabectedin', *Oncotarget* **6**(22), 18875–18890. 10

Ostrowski, M., Carmo, N. B., Krumeich, S., Fanget, I., Raposo, G., Savina, A., Moita, C. F., Schauer, K., Hume, A. N., Freitas, R. P., Goud, B., Benaroch, P., Hacohen, N., Fukuda, M., Desnos, C., Seabra, M. C. and Darchen, F. (2010), 'Rab27a and Rab27b control different steps of the exosome secretion pathway', *Nature Cell Biology* **12**(1), 19–30. 27, 28

Pan, B. T. and Johnstone, R. M. (1983), 'Fate of the transferrin receptor during maturation of sheep reticulocytes in vitro: Selective externalization of the receptor', *Cell* **33**(3), 967–978. 26

Pan, L., Liang, W., Fu, M., hua Huang, Z., Li, X., Zhang, W., Zhang, P., Qian, H., cheng Jiang, P., rong Xu, W. and Zhang, X. (2017), 'Exosomes-mediated transfer of long noncoding RNA ZFAS1 promotes gastric cancer progression', *Journal of Cancer Research and Clinical Oncology* **143**(6), 991–1004. 30, 231, 232

Paolini, L., Zendrini, A., Noto, G. D., Busatto, S., Lottini, E., Radeghieri, A., Dossi, A., Caneschi, A., Ricotta, D. and Bergese, P. (2016), 'Residual matrix from different separation techniques impacts exosome biological activity', *Scientific Reports* **6**, 23550. 131, 141

Parolini, I., Federici, C., Raggi, C., Lugini, L., Palleschi, S., De Milito, A., Coscia, C., Iessi, E., Logozzi, M., Molinari, A., Colone, M., Tatti, M., Sargiacomo, M. and Fais, S. (2009), 'Microenvironmental pH is a key factor for exosome traffic in tumor cells', *Journal of Biological Chemistry* **284**(49), 34211–34222. 29

Pastrana, E., Silva-Vargas, V. and Doetsch, F. (2011), 'Eyes Wide Open: A Critical Review of Sphere-Formation as an Assay For Stem Cells', *Cell Stem Cell* **8**(5), 486–498. 21

Paul, C. D., Mistriotis, P. and Konstantopoulos, K. (2017), 'Cancer cell motility: Lessons from migration in confined spaces', *Nature Reviews Cancer* **17**(2), 131–140. 39

Pedersen, K. W., Kierulf, B., Oksvold, M. P., Li, M., Vlassov, A. V., Roos, N., Kullmann, A. and Neurauter, A. (2013), 'Isolation and Characterization of Exosomes using Magnetic Beads', *BioProbes* **71**, 10–13. 132

Peinado, H., Aleckovic, M., Lavotshkin, S., Matei, I., Costa-Silva, B., Moreno-Bueno, G., Hergueta-Redondo, M., Williams, C., García-Santos, G., Nitadori-hoshino, A., Hoffman, C., Badal, K., Garcia, B. a., Callahan, M. K., Yuan, J., Martins, V. R., Skog, J., Rosandra, N., Brady, M. S., Wolchok, J. D., Chapman, P. B. and Kang, Y. (2013), 'Melanoma exosomes educate bone marrow progenitor cells toward a pro-metastatic phenotype through MET', *Nature Medicine* **18**(6), 883–891. 33, 234

Peter, M., Couturier, J., Pacquement, H., Michon, J., Thomas, G., Magdelenat, H. and Delattre, O. (1997), 'A new member of the ETS family fused to EWS in Ewing tumors', *Oncogene* **14**(10), 1159–1164. 2, 126

Peter, M., Gilbert, E. and Delattre, O. (2001), 'A multiplex Real-Time PCR assay for the detection of gene fusions observed in solid tumors.', *Laboratory Investigation* **81**(6), 905–12. 146

Phuyal, S., Hessvik, N. P., Skotland, T., Sandvig, K. and Llorente, A. (2014), 'Regulation of exosome release by glycosphingolipids and flotillins', *FEBS Journal* **281**, 2214–2227. 33

Pinto, A., Dickman, P. and Parham, D. (2011), 'Pathobiologic markers of the Ewing sarcoma family of tumors: State of the art and prediction of behaviour', *Sarcoma* **2011**, 1–15. 1, 81

Pishas, K. I. and Lessnick, S. L. (2016), 'Recent advances in targeted therapy for Ewing sarcoma', *F1000Research* **5**(0), 2077. 7, 8, 10

Pols, M. S. and Klumperman, J. (2009), 'Trafficking and function of the tetraspanin CD63', *Experimental Cell Research* **315**(9), 1584–1592. 27

Pospichalova, V., Svoboda, J., Dave, Z., Kotrbova, A., Kaiser, K., Klemova, D., Ilkovics, L., Hampl, A., Crha, I., Jandakova, E., Minar, L., Weinberger, V. and Bryja, V. (2015), 'Simplified protocol for flow cytometry analysis of fluorescently labeled exosomes and microvesicles using dedicated flow cytometer.', *Journal of extracellular vesicles* **4**, 25530. 156, 212

Prieur, A., Tirode, F., Cohen, P. and Delattre, O. (2004), 'EWS/FLI-1 Silencing and Gene Profiling of Ewing Cells Reveal Downstream Oncogenic Pathways and a Crucial Role for Repression of Insulin-Like Growth Factor Binding Protein 3', *Molecular and Cellular Biology* **24**(16), 7275–7283. 3, 9

Pützer, B. M., Solanki, M. and Herchenröder, O. (2017), 'Advances in cancer stem cell targeting: How to strike the evil at its root', *Advanced Drug Delivery Reviews* **120**, 89–107. 21

Qi, X., Li, J., Zhou, C., Lv, C. and Tian, M. (2014), 'MicroRNA-320a inhibits cell proliferation, migration and invasion by targeting BMI-1 in nasopharyngeal carcinoma', *FEBS Letters* **588**(20), 3732–3738. 209, 230

R (2017), 'R: A Language and Environment for Statistical Computing'.
URL: *https://www.r-project.org/* 56, 90

Ratajczak, J., Miekus, K., Kucia, M., Zhang, J., Reca, R., Dvorak, P. and Ratajczak, M. Z. (2006), 'Embryonic stem cell-derived microvesicles reprogram hematopoietic progenitors: Evidence for horizontal transfer of mRNA and protein delivery', *Leukemia* **20**(5), 847–856. 29

Redini, F. and Heymann, D. (2015), 'Bone Tumor Environment as a Potential Therapeutic Target in Ewing Sarcoma', *Frontiers in Oncology* **5**(December), 1–11. 11, 12, 234

Remesh, A. (2012), 'Toxicities of anticancer drugs and its management', *International Journal of Basic & Clinical Pharmacology* **1**(1), 2. 5

Richards, K. E., Zeleniak, A. E., Fishel, M. L., Wu, J., Littlepage, L. E. and Hill, R. (2017), 'Cancer-associated fibroblast exosomes regulate survival adn proliferation of pancreatic cancer cells', *Oncogene* **36**(13), 1770–1778. 30

Richter, G. H. S., Plehm, S., Fasan, A., Rossler, S., Unland, R., Bennani-Baiti, I. M., Hotfilder, M., Lowel, D., von Luettichau, I., Mossbrugger, I., Quintanilla-Martinez, L., Kovar, H., Staege, M. S., Muller-Tidow, C. and Burdach, S. (2009), 'EZH2 is a mediator of EWS/FLI1 driven tumor growth and metastasis blocking endothelial and neuro-ectodermal differentiation', *Proceedings of the National Academy of Sciences* **106**(13), 5324–5329. 3, 18

Riggi, N., Suva, M., De Vito, C., Provero, P., Stehle, J., Baumer, K., Cironi, L., Janiszewska, M., Petricevic, T., Suva, D., Joseph, J.-m., Guillou, L. and Stamenkovic, I. (2010), 'EWS-FLI-1 modulates miRNA145 and SOX2 expression to initiate mesenchymal stem cell reprogramming toward Ewing sarcoma cancer stem cells', *Genes & development* **24**, 916–932. 18, 23, 76, 83, 120, 122, 123, 217, 221, 222, 223, 231

Riggi, N., Suvà, M. L., Suvà, D., Cironi, L., Provero, P., Tercier, S., Joseph, J. M., Stehle, J. C., Baumer, K., Kindler, V. and Stamenkovic, I. (2008), 'EWS-FLI-1 expression triggers a Ewing's Sarcoma initiation program in primary human mesenchymal stem cells', *Cancer Research* **68**(7), 2176–2185. 18, 206, 231

Rivoltini, L., Chiodoni, C., Squarcina, P., Tortoreto, M., Villa, A., Vergani, B., Bürdek, M., Botti, L., Arioli, I., Cova, A., Mauri, G., Vergani, E., Bianchi, B., Mina, P. D., Cantone, L., Bollati, V., Zaffaroni, N., Gianni, A. M., Colombo, M. P. and Huber, V. (2016), 'TNF-related apoptosis-inducing ligand (trail)-armed exosomes deliver proapoptotic signals to tumor site', *Clinical Cancer Research* **22**(14), 3499–3512. 34

Roberts, P., Burchill, S. A., Brownhill, S., Cullinane, C. J., Johnston, C., Griffiths, M. J., McMullan, D. J., Brown, N. P., Morris, S. P. and Lewis, I. J. (2008), 'Ploidy and Karyotype Complexity Are Powerful Prognostic Indicators in the Ewing's Sarcoma Family of Tumors: A Study by the United Kingdom Cancer Cytogenetics and the Children's Cancer and Leukaemia Group', *Genes, Chromosomes & Cancer* **47**, 207–220. 3, 46

Rocchi, A. and Manara, M. (2010), 'CD99 inhibits neural differentiation of human Ewing sarcoma cells and thereby contributes to oncogenesis', *The Journal of Clinical Investigations* **120**(3), 668–680. 8, 39, 78

Rodriguez-Galindo, C., Spunt, S. L. and Pappo, A. S. (2003), 'Treatment of Ewing sarcoma family of tumors: Current status and outlook for the future', *Medical and Pediatric Oncology* **40**(5), 276–287. 4, 38

Rodríguez, M., Silva, J., Herrera, A., Herrera, M., Peña, C., Martín, P., Gil-Calderón, B., Larriba, M. J., Coronado, M. J., Soldevilla, B., Turrión, V. S., Provencio, M., Sánchez, A., Bonilla, F. and García-Barberán, V. (2015), 'Exosomes enriched in stemness/metastatic-related mRNAS promote oncogenic potential in breast cancer', *Oncotarget* **6**(38). 129, 204, 210, 229

Roelofs, A. J., Thompson, K., Gordon, S., Rogers, M. J., Boyce, Roodman, Suva, Coleman, Weilbaecher, Bruland, Vessella and Lipton (2006), 'Molecular mechanisms of action of bisphosphonates: Current status', *Clinical Cancer Research* **12**(20 PART 2), 6222–6231. 12

Romagnoli, G. G., Zelante, B. B., Toniolo, P. A., Migliori, I. K. and Barbuto, J. A. M. (2015), 'Dendritic cell-derived exosomes may be a tool for cancer immunotherapy by converting tumor cells into immunogenic targets', *Frontiers in Immunology* **5**(JAN), 1–5. 33, 233

Romeo, S. and Tos, A. P. D. (2010), 'Soft tissue tumors associated with EWSR1 translocation', *Virchows Arch* **456**, 219–234. 126, 220

Rorie, C. J., Thomas, V. D., Chen, P., Pierce, H. H., O'Bryan, J. P. and Weissman, B. E. (2004), 'The Ews/Fli-1 Fusion Gene Switches the Differentiation Program of Neuroblastomas to Ewing Sarcoma/Peripheral Primitive Neuroectodermal Tumors', *Cancer Research* **64**(4), 1266–1277. 17, 206, 231

Roundhill, E. and Burchill, S. (2013), 'Membrane expression of MRP-1, but not MRP-1 splicing or Pgp expression, predicts survival in patients with ESFT', *British Journal of Cancer* **109**(1), 195–206. 84

Rudenko, G. (2017), 'Dynamic control of synaptic adhesion and organizing molecules in synaptic plasticity', *Neural Plasticity* **2017**, 1–14. 121, 218

Rupp, A. K., Rupp, C., Keller, S., Brase, J. C., Ehehalt, R., Fogel, M., Moldenhauer, G., Marmé, F., Sültmann, H. and Altevogt, P. (2011), 'Loss of EpCAM expression in breast cancer derived serum exosomes: Role of proteolytic cleavage', *Gynecologic Oncology* **122**(2), 437–446. 132

Sankar, S., Theisen, E. R., Bearss, J., Mulvihill, T., Hoffman, L. M., Sorna, V., Beckerle, M. C., Sharma, S. and Lessnick, S. L. (2014), 'Reversible LSD1 inhibition interferes with global EWS/ETS transcriptional activity and impedes Ewing sarcoma tumor growth', *Clinical Cancer Research* **20**(17), 4584–4597. 11, 220

Santoro, M., Lamhamedi-Cherradi, S.-E., Menegaz, B. A., Ludwig, J. A. and Mikos, A. G. (2015), 'Flow perfusion effects on three-dimensional culture and drug sensitivity of Ewing sarcoma', *Proceedings of the National Academy of Sciences* **112**(33), 10304–10309. 15, 223, 226

Savola, S., Klami, A., Myllykangas, S., Manara, C., Scotlandi, K., Picci, P., Knuutila, S. and Vakkila, J. (2011), 'High Expression of Complement Component 5 (C5) at Tumor Site Associates with Superior Survival in Ewing's Sarcoma Family of Tumour Patients', *ISRN Oncology* **2011**, 1–10. 215

Schageman, J., Zeringer, E., Li, M., Barta, T., Lea, K., Gu, J., Magdaleno, S., Setterquist, R., Vlassov, A. V., Schageman, J., Zeringer, E., Li, M., Barta, T., Lea, K., Gu, J., Magdaleno, S., Setterquist, R. and Vlassov, A. V. (2013), 'The Complete Exosome Workflow Solution: From Isolation to Characterization of RNA Cargo', *BioMed Research International* **2013**, 1–15. 132, 209

Schatton, T., Frank, N. Y. and Frank, M. H. (2009), 'Identification and targeting of cancer stem cells', *BioEssays* **31**(10), 1038–1049. 19, 22, 39, 76, 83, 123, 221, 229

Schatton, T., Murphy, G. F., Frank, N. Y., Yamura, K., Waaga-Gasser, A. M., Gasser, M., Zhan, Q., Jordan, S., Duncan, L. M., Weishaupt, C., Fuhlbrigge, R. C., Kupper, T. S., Sayegh, M. H. and Frank, M. H. (2008), 'Identification of cells initiating human melanomas', *Nature* **451**(7176), 345–349. 19, 83

Schleiermacher, G., Peter, M., Oberlin, O., Philip, T., Rubie, H., Mechinaud, F., Sommelet-Olive, D., Landman-Parker, J., Bours, D., Michon, J. and Delattre, O. (2003), 'Increased risk of systemic relapses associated with bone marrow micrometastasis and circulating tumor cells in localized ewing tumor', *Journal of Clinical Oncology* **21**(1), 85–91. 227

Scotlandi, K., Benini, S., Nanni, P., Lollini, P. L., Nicoletti, G., Landuzzi, L., Serra, M., Manara, M. C., Picci, P. and Baldini, N. (1998), 'Blockage of insulin-like growth factor-I receptor inhibits the growth of Ewing's sarcoma in athymic mice', *Cancer Research* **58**(18), 4127–4131. 9, 14

Scotlandi, K., Benini, S., Sarti, M., Serra, M., Loffini, P.-l., Maurici, D., Picci, N., Manara, M. C. and Baidini, N. (1996), 'Insulin-like Growth Factor I Receptor-mediated Circuit in Ewing's Sarcoma/Peripheral Neuroectodermal Tumor: A Possible Therapeutic Target', *Cancer Research* **56**, 4570–4574. 9

Scotlandi, K., Manara, M. C., Serra, M., Marino, M. T., Ventura, S., Garofalo, C., Alberghini, M., Magagnoli, G., Ferrari, S., Lopez-Guerrero, J. A., Llombard-Bosch, A. and Picci, P. (2011), 'Expression of insulin-like growth factor system components in Ewing's sarcoma and their association with survival', *European Journal of Cancer* **47**(8), 1258–1266. 9

Scotlandi, K., Manara, M. C., Strammiello, R., Landuzzi, L., Benini, S., Perdichizzi, S., Serra, M., Astolfi, A., Nicoletti, G., Lollini, P. L., Bertoni, F., Nanni, P. and Picci, P. (2003), 'c-kit Receptor expression in Ewing's sarcoma: Lack of prognostic value but therapeutic targeting opportunities in appropriate conditions', *Journal of Clinical Oncology* **21**(10), 1952–1960. 11, 40, 64, 80

Scotlandi, K., Perdichizzi, S., Bernard, G., Nicoletti, G., Nanni, P., Lollini, P. L., Curti, A., Manara, M. C., Benini, S., Bernard, A. and Picci, P. (2006), 'Targeting CD99 in association with doxorubicin: An effective combined treatment for Ewing's sarcoma', *European Journal of Cancer* **42**(1), 91–96. 8, 207

Scotlandi, K., Remondini, D., Castellani, G., Manara, M. C., Nardi, F., Cantiani, L., Francesconi, M., Mercuri, M., Caccuri, A. M., Serra, M., Knuutila, S. and Picci, P. (2009), 'Overcoming resistance to conventional drugs in Ewing sarcoma and identification of molecular predictors of outcome', *Journal of Clinical Oncology* **27**(13), 2209–2216. 82

Sell, S. (2004), 'Stem cell origin of cancer and differentiation therapy', *Critical Reviews in Oncology/Hematology* **51**(1), 1–28. 21

Seol, H. J., Chang, J. H., Yamamoto, J., Romagnuolo, R., Suh, Y., Weeks, A., Agnihotri, S., Smith, C. A. and Rutka, J. T. (2012), 'Overexpression of CD99 Increases the Migration and Invasiveness of Human Malignant Glioma Cells', *Genes and Cancer* **3**(9-10), 535–549. 39

Seyednasrollah, F., Laiho, A. and Elo, L. L. (2013), 'Comparison of software packages for detecting differential expression in RNA-seq studies', *Briefings in Bioinformatics* **16**(1), 59–70. 86

Singh, S. K., Hawkins, C., Clarke, I. D., Squire, J. A., Bayani, J., Hide, T., Henkelman, R. M., Cusimano, M. D. and Dirks, P. B. (2004), 'Identification of human brain tumour initiating cells', *Nature* **432**(November), 396–401. 19, 83

Skubitz, K. M., Skubitz, A. P. N., Xu, W. W., Luo, X., Lagarde, P., Coindre, J. M. and Chibon, F. (2014), 'Gene expression identifies heterogeneity of metastatic behavior among high-grade non-translocation associated soft tissue sarcomas', *Journal of Translational Medicine* **12**(1), 1–10. 210, 230

Smith, R., Owen, L. A., Trem, D. J., Wong, J. S., Whangbo, J. S., Golub, T. R. and Lessnick, S. L. (2006), 'Expression profiling of EWS/FLI identifies NKX2.2 as a critical target gene in Ewing's sarcoma', *Cancer Cell* **9**(5), 405–416. 3

Sousa, D., Lima, R. T. and Vasconcelos, M. H. (2015), 'Intercellular Transfer of Cancer Drug Resistance Traits by Extracellular Vesicles', *Trends in Molecular Medicine* **21**(10), 595–608. 232

Stacey, G. (2005), 'Primary Cell Cultures and Immortal Cell Lines', *Encyclopedia of Life Sciences* pp. 1–6. 14, 38, 125

Staege, M. S., Hutter, C., Neumann, I., Foja, S., Hattenhorst, U. E., Hansen, G., Afar, D. and Burdach, S. E. (2004), 'DNA microarrays reveal relationship of Ewing family tumors to both endothelial and fetal neural crest-derived cells and define novel targets', *Cancer Research* **64**(22), 8213–8221. 17

Stahl, M., Ranft, A., Paulussen, M., Bölling, T., Vieth, V., Bielack, S., Gortitz, I., Braun-Munzinger, G., Hardes, J., Jurgens, H. and Dirksen, U. (2011), 'Risk of Recurrence and Survival After Relapse in Patients With Ewing Sarcoma', *Pediatr Blood Cancer* **57**, 549–553. 4

Staubach, S., Razawi, H. and Hanisch, F. G. (2009), 'Proteomics of MUC1-containing lipid rafts from plasma membranes and exosomes of human breast carcinoma cells MCF-7', *Proteomics* **9**(10), 2820–2835. 27

Stewart, E., Goshorn, R., Bradley, C., Griffiths, L. M., Benavente, C., Twarog, N. R., Miller, G. M., Caufield, W., Iii, B. B. F., Bahrami, A., Pappo, A., Wu, J., Loh, A., Karlström, Å., Gordon, B., Tsurkan, L., Hatfield, M. J., Potter, P. M., Snyder, S., Thiagarajan, S., Shirinifard, A., Sablauer, A., Anang, A. and Dyer, M. A. (2015), 'Targeting the DNA Repair Pathway in Ewing Sarcoma', *Cell Rep* **9**(3), 829–841. 10

Stiller, C. A., Trama, A., Serraino, D. and et al (2013), 'Descriptive epidemiology of sarcomas in Europe: Report from the RARECARE project', *European Journal of Cancer* **49**(3), 684–695. 1

Stolzenburg, L. R. and Harris, A. (2018), 'Microvesicle-mediated delivery of miR-1343: impact on markers of fibrosis', *Cell and Tissue Research* **371**(2), 325–338. 133, 213

Stoner, S. A., Duggan, E., Condello, D., Guerrero, A., Turk, J. R., Narayanan, P. K. and Nolan, J. P. (2016), 'High sensitivity flow cytometry of membrane vesicles', *Cytometry Part A* **89**(2), 196–206. 212

Stranska, R., Gysbrechts, L., Wouters, J., Vermeersch, P., Bloch, K., Dierickx, D., Andrei, G. and Snoeck, R. (2018), 'Comparison of membrane affinity-based method with size-exclusion chromatography for isolation of exosome-like vesicles from human plasma', *Journal of Translational Medicine* **16**(1), 1–9. 133

Suarez, A., McDowell, H., Niaudet, P., Comoy, E. and Flamant, F. (1991), 'Long-term follow-up of ifosfamide renal toxicity in children treated for malignant mesenchymal tumors: An International Society of Pediatric Oncology Report', *Journal of Clinical Oncology* **9**(12), 2177–2182. 5

Sun, W., Luo, J. D., Jiang, H. and Duan, D. D. (2018), 'Tumor exosomes: A double-edged sword in cancer therapy', *Acta Pharmacologica Sinica* **39**(4), 534–541. 33, 204, 210

Suva, M., Riggi, N., Stehle, J. C., Baumer, K., Tercier, S., Joseph, J. M., Suva, D., Clement, V., Provero, P., Cironi, L., Osterheld, M. C., Guillou, L. and Stamenkovic, I. (2009), 'Identification of cancer stem cells in Ewing's sarcoma', *Cancer Research* **69**(5), 1776–1781. 23, 39, 76, 83, 122, 123, 221, 222, 223

Svensson, K. J., Christianson, H. C., Wittrup, A., Bourseau-Guilmain, E., Lindqvist, E., Svensson, L. M., Mörgelin, M. and Belting, M. (2013), 'Exosome uptake depends on ERK1/2-heat shock protein 27 signaling and lipid raft-mediated endocytosis negatively regulated by caveolin-1', *Journal of Biological Chemistry* **288**(24), 17713–17724. 28, 233

Takahashi, Y., Nishikawa, M., Shinotsuka, H., Matsui, Y., Ohara, S., Imai, T. and Takakura, Y. (2013), 'Visualization and in vivo tracking of the exosomes of murine melanoma B16-BL6 cells in mice after intravenous injection', *Journal of Biotechnology* **165**(2), 77–84. 31

Takashima, Y., Era, T., Nakao, K., Kondo, S., Kasuga, M., Smith, A. G. and Nishikawa, S. I. (2007), 'Neuroepithelial Cells Supply an Initial Transient Wave of MSC Differentiation', *Cell* **129**(7), 1377–1388. 18

Tanaka, M., Yamazaki, Y., Kanno, Y. and Igarashi, K. (2014), 'Ewing's sarcoma precursors are highly enriched in embryonic osteochondrogenic progenitors', *The Journal of Clinical Investigations* **124**(7), 3061–3074. 121, 218

Tap, W. D., Demetri, G., Barnette, P., Desai, J., Kavan, P., Tozer, R., Benedetto, P. W., Friberg, G., Deng, H., McCaffery, I., Leitch, I., Badola, S., Chang, S., Zhu, M. and Tolcher, A. (2012), 'Phase II study of ganitumab, a fully human anti-type-1 insulin-like growth factor receptor antibody, in patients with metastatic Ewing family tumors or desmoplastic small round cell tumors', *Journal of Clinical Oncology* **30**(15), 1849–1856. 9

Tauro, B. J., Greening, D. W., Mathias, R. A., Ji, H., Mathivanan, S., Scott, A. M. and Simpson, R. J. (2012), 'Comparison of ultracentrifugation, density gradient separation, and immunoaffinity capture methods for isolating human colon cancer cell line LIM1863-derived exosomes', *Methods* **56**(2), 293–304. 131, 132, 213

Taylor, D. D. and Gercel-Taylor, C. (2008), 'MicroRNA signatures of tumor-derived exosomes as diagnostic biomarkers of ovarian cancer', *Gynecologic Oncology* **110**(1), 13–21. 32, 208

Taylor, D. D., Zacharias, W. and Gercel-Taylor, C. (2011), Exosome Isolation for Proteomic Analyses and RNA profiling, *in* R. J. Greening, S. W. and David, eds, 'Serum/Plasma Proteomics: Methods and Protocols', Vol. 728, Springer Science+Business Media, chapter Chapter 15, pp. 235–246.
URL: *http://www.springerlink.com/index/10.1007/978-1-61779-068-3* 132

Teicher, B. A., Bagley, R. G., Rouleau, C., Kruger, A., Ren, Y. and Kurtzberg, L. (2011), 'Characteristics of human Ewing/PNET sarcoma models', *Ann Saudi Med* **31**(2), 174–182. 14, 225

Thakur, B. K., Zhang, H., Becker, A., Matei, I., Huang, Y., Costa-Silva, B., Zheng, Y., Hoshino, A., Brazier, H., Xiang, J., Williams, C., Rodriguez-Barrueco, R., Silva, J. M., Zhang, W., Hearn, S., Elemento, O., Paknejad, N., Manova-Todorova, K., Welte, K., Bromberg, J., Peinado, H. and Lyden, D. (2014), 'Double-stranded DNA in exosomes: a novel biomarker in cancer detection', *Cell Research* **24**(6), 766–769. 27

The ENCODE Project Consortium (2011), 'A User's Guide to the Encyclopedia of DNA Elements (ENCODE)', *PLOS Biology* **9**(4), 1–21. 91

Therneau, T. M. and Lumley, T. (2017), 'Survival Analysis'. 56

Théry, C., Amigorena, S., Raposo, G. and Clayton, A. (2006), 'Isolation and characterization of exosomes from cell culture supernatants and biological fluids.', *Current Protocols in Cell Biology* **Chapter 3**, Unit 3.22. 132, 137, 139, 159, 213

Thery, C., Boussac, M., Veron, P., Ricciardi-Castagnoli, P., Raposo, G., Garin, J. and Amigorena, S. (2001), 'Proteomic Analysis of Dendritic Cell-Derived Exosomes: A Secreted Subcellular Compartment Distinct from Apoptotic Vesicles', *The Journal of Immunology* **166**(12), 7309–7318. 27

Théry, C., Ostrowski, M. and Segura, E. (2009), 'Membrane vesicles as conveyors of immune responses.', *Nature reviews. Immunology* **9**(8), 581–93. 26, 29, 31, 33, 233

Tian, T., Wang, Y., Wang, H., Zhu, Z. and Xiao, Z. (2010), 'Visualizing of the cellular uptake and intracellular trafficking of exosomes by live-cell microscopy', *Journal of Cellular Biochemistry* **111**(2), 488–496. 130

Tian, T., Zhu, Y. L., Zhou, Y. Y., Liang, G. F., Wang, Y. Y., Hu, F. H. and Xiao, Z. D. (2014), 'Exosome uptake through clathrin-mediated endocytosis and macropinocytosis and mediating miR-21 delivery', *Journal of Biological Chemistry* **289**(32), 22258–22267. 28, 130, 211, 229, 233

Tickner, J. A., Urquhart, A. J., Stephenson, S.-A., Richard, D. J. and O'Byrne, K. J. (2014), 'Functions and therapeutic roles of exosomes in cancer.', *Frontiers in oncology* **4**(May), 127. 31, 232

Tirode, F., Laud-Duval, K., Prieur, A., Delorme, B., Charbord, P. and Delattre, O. (2007), 'Mesenchymal Stem Cell Features of Ewing Tumors', *Cancer Cell* **11**(5), 421–429. 18, 206

Tirode, F., Surdez, D., Ma, X., Parker, M., Cécile, M., Deley, L., Bahrami, A., Zhang, Z., Lapouble, E., Grossetête-lalami, S., Reynaud, S., Rio-frio, T., Hedlund, E., Wu, G., Chen, X., Pierron, G., Oberlin, O., Zaidi, S., Lemmon, G. and Gupta, P. (2014), 'Genomic landscape of Ewing Sarcoma defines an aggressive subtype with co-association of STAG2 and TP53 mutations', *Cancer discovery* **4**(11), 1342–1353. 3, 4, 82, 84, 219

Tomazou, E. M., Sheffield, N. C., Schmidl, C., Schuster, M., Schönegger, A., Datlinger, P., Kubicek, S., Bock, C. and Kovar, H. (2015), 'Epigenome Mapping Reveals Distinct Modes of Gene Regulation and Widespread Enhancer Reprogramming by the Oncogenic Fusion Protein EWS-FLI1', *Cell Reports* **10**(7), 1082–1095. 219

Toretsky, J. A., Erkizan, V., Levenson, A., Abaan, O. D., Parvin, J. D., Cripe, T. P., Rice, A. M., Lee, S. B. and Üren, A. (2006), 'Oncoprotein EWS-FLI1 activity is enhanced by RNA helicase A', *Cancer Research* **66**(11), 5574–5581. 7

Torreggiani, E., Roncuzzi, L., Perut, F., Zini, N. and Baldini, N. (2016), 'Multimodal transfer of MDR by exosomes in human osteosarcoma', *International Journal of Oncology* **49**(1), 189–196. 30, 232

Town, J., Pais, H., Harrison, S., Stead, L. F., Bataille, C., Bunjobpol, W. and Zhang, J. (2016), 'Exploring the surfaceome of Ewing sarcoma identifies a new and unique therapeutic target', *PNAS* **113**(13), 3603–3608. 1, 121, 218, 223

Trajkovic, K., Hsu, C., Chiantia, S., Rajendran, L., Wenzel, D., Wieland, F., Schwille, P., Brügger, B. and Simons, M. (2008), 'Ceramide Triggers Budding of Exosome Vesicles into Multivesicular Endosomes', *Science* **319**, 1244–1247. 28, 32, 233

Trowsdale, J. and Knight, J. C. (2013), 'Major Histocompatibility Complex Genomics and Human Disease', *Annual Review of Genomics and Human Genetics* **14**(1), 301–323. 113, 216

Trucco, M. and Loeb, D. (2012), 'Sarcoma stem cells: Do we know what we are looking for?', *Sarcoma* **2012**. 83, 120, 219, 229

Tsugita, M., Yamada, N., Noguchi, S., Yamada, K., Moritake, H., Shimizu, K., Akao, Y. and Ohno, T. (2013), 'Ewing sarcoma cells secrete EWS/Fli-1 fusion mRNA via microvesicles', *PloS one* **8**(10), e77416. 35, 129, 130, 131, 204, 206, 208, 209, 227, 229, 230

Valadi, H., Ekström, K., Bossios, A., Sjöstrand, M., Lee, J. J. and Lötvall, J. O. (2007), 'Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells', *Nature Cell Biology* **9**(6), 654–659. 27, 30, 141, 208, 227

Valenti, R., Huber, V., Filipazzi, P., Pilla, L., Sovena, G., Villa, A., Corbelli, A., Fais, S., Parmiani, G. and Rivoltini, L. (2006), 'Human tumor-released microvesicles promote the differentiation of myeloid cells with transforming growth factor-β-mediated suppressive activity on T lymphocytes', *Cancer Research* **66**(18), 9290–9298. 31

Van Deun, J., Mestdagh, P., Sormunen, R., Cocquyt, V., Vermaelen, K., Vandesompele, J., Bracke, M., De Wever, O. and Hendrix, A. (2014), 'The impact of disparate isolation methods for extracellular vesicles on downstream RNA profiling', *Journal of Extracellular Vesicles* **3**(1), 1–14. 132, 213

Van Doorninck, J. A., Ji, L., Schaub, B., Shimada, H., Wing, M. R., Krailo, M. D., Lessnick, S. L., Marina, N., Triche, T. J., Sposto, R., Womer, R. B. and Lawlor, E. R. (2010), 'Current treatment protocols have eliminated the prognostic advantage of type 1 fusions in Ewing Sarcoma: A report from the children's oncology group', *Journal of Clinical Oncology* **28**(12), 1989–1994. 2

Van Maldegem, A. M., Bovée, J. V., Peterse, E. F., Hogendoorn, P. C. and Gelderblom, H. (2016), 'Ewing sarcoma: The clinical relevance of the insulin-like growth factor 1 and the poly-ADP-ribose-polymerase pathway', *European Journal of Cancer* **53**, 171–180. 9

van Maldegem, A. M., Hogendoorn, P. C. and Hassan, A. B. (2012), 'The clinical use of biomarkers as prognostic factors in Ewing sarcoma', *Clinical Sarcoma Research* **2**(1), 7. 77

Ventura, S., Aryee, D. N., Felicetti, F., De Feo, A., Mancarella, C., Manara, M. C., Picci, P., Colombo, M. P., Kovar, H., Care, A. and Scotlandi, K. (2016), 'CD99 regulates neural differentiation of Ewing sarcoma cells through miR-34a-Notch-mediated control of NF-κB signaling', *Oncogene* **35**(30), 3944–3954. 35, 36, 129, 130, 131, 204, 207, 209, 226, 229, 230

Vermeulen, L., De Sousa E Melo, F., Van Der Heijden, M., Cameron, K., De Jong, J. H., Borovski, T., Tuynman, J. B., Todaro, M., Merz, C., Rodermond, H., Sprick, M. R., Kemper, K., Richel, D. J., Stassi, G. and Medema, J. P. (2010), 'Wnt activity defines colon cancer stem cells and is regulated by the microenvironment', *Nature Cell Biology* **12**(5), 468–476. 21

Villasante, A., Marturano-Kruik, A., Ambati, S. R., Liu, Z., Godier-Furnemont, A., Parsa, H., Lee, B. W., Moore, M. A. and Vunjak-Novakovic, G. (2016), 'Recapitulating the size and cargo of tumor exosomes in a tissue-engineered model', *Theranostics* **6**(8), 1119–1130. 35, 36, 81, 129, 130, 131, 204, 206, 209, 226, 229, 230

Visvader, J. E. and Lindeman, G. J. (2012), 'Cancer stem cells: Current status and evolving complexities', *Cell Stem Cell* **10**(6), 717–728. 19, 21, 83, 219

Vo, K. T., Matthay, K. K. and DuBois, S. G. (2016), 'Targeted antiangiogenic agents in combination with cytotoxic chemotherapy in preclinical and clinical studies in sarcoma', *Clinical Sarcoma Research* **6**(1), 9. 13

von Levetzow, C., Jiang, X., Gwye, Y., von Levetzow, G., Hung, L., Cooper, A., Hsu, J. H. R. and Lawlor, E. R. (2011), 'Modeling initiation of Ewing sarcoma in human neural crest cells', *PLoS ONE* **6**(4), 1–10. 18

Vormoor, B., Knizia, H. K., Batey, M. A., Almeida, G. S., Wilson, I., Dildey, P., Sharma, A., Blair, H., Hide, I. G., Heidenreich, O., Vormoor, J., Maxwell, R. J. and Bacon, C. M. (2014), 'Development of a preclinical orthotopic xenograft model of ewing sarcoma and other human malignant bone disease using advanced in vivo imaging', *PLoS ONE* **9**(1). 14

Vormoor, J., Baersch, G., Decker, S., Hotfilder, M., Schäfer, K. L., Pelken, L., Rübe, C., Van Valen, F., Jürgens, H. and Dockhorn-Dworniczak, B. (2001), 'Establishment of an in vivo model for pediatric ewing tumors by transplantation into NOD/scid mice', *Pediatric Research* **49**(3), 332–341. 14

Wahl, J., Bogatyreva, L., Boukamp, P., Rojewski, M., Van Valen, F., Fiedler, J., Hipp, N., Debatin, K. M. and Beltinger, C. (2010), 'Ewing's sarcoma cells with CD57-associated increase of tumorigenicity and with neural crest-like differentiation capacity', *International Journal of Cancer* **127**(6), 1295–1307. 24, 39, 76, 83, 122, 123, 221, 222, 223

Waldron, N. N., Kaufman, D. S., Oh, S., Inde, Z., Hexum, M. K., Ohlfest, J. R. and Vallera, D. A. (2011), 'Targeting Tumor-Initiating Cancer Cells with dCD133KDEL Shows Impressive Tumor Reductions in a Xenotransplant Model of Human Head and Neck Cancer', *Molecular Cancer Therapeutics* **10**(10), 1829–1839. 23

Wang, J.-P., Tang, Y.-Y., Fan, C.-M., Guo, C., Zhou, Y.-H., Li, Z., Li, X.-L., Li, Y., Li, G.-Y., Xiong, W., Zeng, Z.-Y. and Xiong, F. (2018), 'The role of exosomal non-coding RNAs in cancer metastasis', *Oncotarget* **9**(15), 12487–12502. 209

Wang, J., Zheng, Y. and Zhao, M. (2017), 'Exosome-based cancer therapy: Implication for targeting cancer stem cells', *Frontiers in Pharmacology* **7**(JAN), 1–11. 33

Wang, K., Xu, J., Zhang, J. and Huang, J. (2012), 'Prognostic role of CD133 expression in colorectal cancer: A meta-analysis', *BMC Cancer* **12**(1), 1. 222

Wang, W., Zhao, L., Wei, X., Wang, L., Liu, S., Yang, Y., Wang, F., Sun, G., Zhang, J., Ma, Y., Zhao, Y. and Yu, J. (2016), 'MicroRNA-320a promotes 5-FU resistance in human pancreatic cancer cells', *Scientific Reports* **6**, 1–11. 209

Wang, Y. (2009), 'Inhibiting platelet-derived growth factor β reduces Ewing's sarcoma growth and metastasis in a novel orthotopic human xenograft model', *In Vivo* **23**(6), 903–909. 14, 80

Wang, Z., Gerstein, M. and Snyder, M. (2009), 'RNA-Seq: a revolutionary tool for transcriptomics.', *Nature reviews. Genetics* **10**(1), 57–63. 85

Whiteside, T. L. (2016), 'Exosomes and tumor-mediated immune suppression', *Journal of Clinical Investigation* **126**(4), 1216–1223. 31

Witwer, K. W., Buzás, E. I., Bemis, L. T., Bora, A., Lässer, C., Lötvall, J., Nolte-'t Hoen, E. N., Piper, M. G., Sivaraman, S., Skog, J., Théry, C., Wauben, M. H. and Hochberg, F. (2013), 'Standardization of sample collection, isolation and analysis methods in extracellular vesicle research.', *Journal of extracellular vesicles* **2**, 1–25. 131

Womer, R. B., West, D. C., Krailo, M. D., Dickman, P. S., Pawel, B. R., Grier, H. E., Marcus, K., Sailer, S., Healey, J. H., Dormans, J. P. and Weiss, A. R. (2012), 'Randomized controlled trial of interval-compressed chemotherapy for the treatment of localized Ewing Sarcoma: A report from the Children's Oncology Group', *Journal of Clinical Oncology* **30**(33), 4148–4154. 5

Xu, J. (2018), 'Exosomes Regulate the Transformation of Cancer Cells in Cancer Stem Cell Homeostasis', *Stem Cell International* **2018**. 223

Xu, J.-F., Wang, Y.-P., Zhang, S.-J., Chen, Y., Gu, H.-F., Dou, X.-F., Xia, B., Bi, Q. and Fan, S.-W. (2017), 'Exosomes containing differential expression of microRNA and mRNA in osteosarcoma that can predict response to chemotherapy', *Oncotarget* **8**(44), 75968–75978. 229

Xu, Q., Zhao, Y., Zhou, X., Luan, J., Cui, Y. and Han, J. (2018), 'Comparison of the extraction and determination of serum exosome and miRNA in serum and the detection of miR-27a-3p in serum exosome of ALS patients', *Intractable and Rare Diseases Research* **7**(1), 13–18. 123, 133

Yamaguchi, H., Wyckoff, J. and Condeelis, J. (2005), 'Cell migration in tumors', *Current Opinion in Cell Biology* **17**(5 SPEC. ISS.), 559–564. 39

Yang, M., Zhang, R., Yan, M., Ye, Z., Liang, W. and Luo, Z. (2010), 'Detection and characterization of side population in Ewing's sarcoma SK-ES-1 cells in vitro', *Biochemical and Biophysical Research Communications* **391**(1), 1062–1066. 25, 83, 122, 123, 221

Zhang, H., Xie, Y., Li, W., Chibbar, R., Xiong, S. and Xiang, J. (2011), 'CD4 T cell-released exosomes inhibit CD8 cytotoxic T-lymphocyte responses and antitumor immunity', *Cellular and Molecular Immunology* **8**(1), 23–30. 130, 211, 229, 233

Zhang, J., Li, S., Li, L., Li, M., Guo, C., Yao, J. and Mi, S. (2015), 'Exosome and exosomal microRNA: Trafficking, sorting, and function', *Genomics, Proteomics and Bioinformatics* **13**(1), 17–24. 27, 129

Zhang, S., Balch, C., Chan, M. W., Lai, H.-C., Matei, D., Schilder, J. M., Yan, P. S., Huang, T. H.-M. and Nephew, K. P. (2008), 'Identification and Characterization of Ovarian Cancer-Initiating Cells from Primary Human Tumors', *Cancer Research* **68**(11), 4311–4320. 19

Zhang, Y. (2017), 'Overexpression of CST4 promotes gastric cancer aggressiveness by activating the ELFN2 signalling pathway', *Am J Cancer Res* **7**(11), 2290–2304. 121, 218

Zhang, Z. H., Jhaveri, D. J., Marshall, V. M., Bauer, D. C., Edson, J., Narayanan, R. K., Robinson, G. J., Lundberg, A. E., Bartlett, P. F., Wray, N. R. and Zhao, Q.-y. (2014), 'A Comparative Study of Techniques for Differential Expression Analysis on RNA-Seq Data', *PLoS ONE* **9**(8). 86, 128

Zhang, Z., Li, X., Sun, W., Yue, S., Yang, J., Li, J., Ma, B., Wang, J., Yang, X., Pu, M., Ruan, B., Zhao, G., Huang, Q., Wang, L., Tao, K. and Dou, K. (2017), 'Loss of exosomal miR-320a from cancer-associated fibroblasts contributes to HCC proliferation and metastasis', *Cancer Letters* **397**, 33–42. 209

Zhao, F., Chen, Y., Wu, Q., Wang, Z. and Lu, J. (2014), 'Prognostic value of CD117 in cancer: A meta-analysis', *International Journal of Clinical and Experimental Pathology* **7**(3), 1012–1021. 40

Zhao, S., Wang, Y., Lou, Y., Wang, Y., Sun, J., Luo, M., Li, W. and Miao, L. (2018), 'MicroRNA-320a suppresses tumour cell proliferation and invasion of renal cancer cells by targeting FoxM1', *Oncology Reports* **40**, 1917–1926. 209, 230

Zheng, P., Luo, Q., Wang, W., Li, J., Wang, T., Wang, P., Chen, L., Zhang, P., Chen, H., Liu, Y., Dong, P., Xie, G., Ma, Y., Jiang, L., Yuan, X. and Shen, L. (2018), 'Tumor-associated macrophages-derived exosomes promote the migration of gastric cancer cells by transfer of functional Apolipoprotein E', *Cell Death and Disease* **9**(434). 204, 205, 231, 232

Zöllner, S. K., Selvanathan, S. P., Graham, G. T., Commins, R. M., Hong, S. H., Moseley, E., Parks, S., Haladyna, J. N., Erkizan, H. V., Dirksen, U., Hogarty, M. D., Üren, A. and Toretsky, J. A. (2017), 'Inhibition of the oncogenic fusion protein EWS-FLI1 causes G2-M cell cycle arrest and enhanced vincristine sensitivity in Ewing's sarcoma', *Science Signaling* **10**(499). 8

Zwerner, J. P. and May, W. A. (2001), 'PDGF-C is an EWS/FLI induced transforming growth factor in Ewing Family Tumors', *Oncogene* **5**(20), 1–27. 3