

MOLECULAR CHAPERONE EXPRESSION AND FUNCTION IN BREAST CANCER AND BREAST CANCER STEM CELLS

A thesis submitted in the fulfillment of the requirements for the degree of

Master of Science in Biochemistry

of

Rhodes University

by

Jason Neville Sterrenberg

February 2012

Abstract

The Cancer Stem Cell (CSC) theory suggests that cancers arise from and are maintained by a subpopulation of cancer cells with stem cell properties. Molecular chaperones are key components of cellular regulation. The overexpression of chaperones has become synonymous with cancer cells with chaperones being recognized as *bona fide* anti-cancer drug targets. Although chaperone activity has been characterized in cancer cells, very little is known about the cellular functions of chaperones in cancer stem cells. We set out to compare the expression of selected molecular chaperones in non-stem cancer cell and cancer stem cell enriched populations isolated from breast cancer lines, in order to identify chaperones differentially expressed between the two populations for further biological characterization. In order to isolate breast cancer stem cells from the MCF-7 and MDA-MB-231 breast cancer cell lines, three cancer stem cell isolation and identification techniques were utilized based on (1) cell surface marker expression (CD44+/CD24- and CD44+/CD24-/EpCAM+ phenotypes), (2) aldehyde dehydrogenase enzyme activity (ALDH^{Hi}) and (3) ability to grow in anchorage-independent conditions. The MDA-MB-231 and MCF-7 breast cancer cell lines displayed CD44+/CD24- cell populations with the MCF-7 cell line additionally displaying a large CD44+/CD24-/EpCAM+ population. Although both cell lines showed similar ALDH^{Hi} populations, they differed substantially with respect to anchorage-independent growth. MCF-7 cells were able to form anchorage-independent colonies while the MDA-MB-231 cell line was not. Anchorage-independent MCF-7 cells showed enrichment in CD44+/CD24- and CD44+/CD24-/EpCAM+ cells compared to adherent MCF-7 cells, and were selected for gene expression studies. Gene expression studies identified 22 genes as being down-regulated at the mRNA level in the anchorage-independent MCF-7 cells, while only 2 genes (BAG1 and DNAJC12) were up-regulated. The down-regulation of selected chaperones in anchorage independent MCF-7 cells was confirmed at the protein level for selected chaperones, including DNAJB6, a type II DNAJ protein shown to be involved in the regulation of Wnt signaling. In order to characterize the effect of DNAJB6 expression on BCSCs we developed a pCMV mammalian expression plasmid for both DNAJB6 isoforms (DNAJB6L and DNAJB6S). We successfully constructed mutants of the conserved histidine-proline-aspartic acid (HPD) motif of the J domain of

DNAJB6S and DNAJB6L. These constructs will allow the analysis of the role of DNAJB6 in cancer stem cell function. To the best of our knowledge, this is the first report to focus on the comparative expression of molecular chaperones in normal and cancer stem cell enriched breast cancer populations.

Declaration

I declare that this thesis is my own, unaided work. It is being submitted for the degree of Master of Science of Rhodes University. It has not been submitted before for any degree or examination at any other university

.....
Mr Jason N Sterrenberg, February 2012

Table of Contents

Abstract	ii
Declaration	iv
Table of Contents	v
List of Figures	viii
List of Tables.....	x
List of Abbreviations.....	xi
List of Abbreviations.....	xi
List of Symbols	xiv
Acknowledgements	xv
Chapter 1 : Literature Review.....	1
1.1 Cancer Stem Cells: Origin, Evidence and Significance	2
1.1.1 Cancer Stem Cell Theory	2
1.1.2 Definition of a Cancer Stem Cell	5
1.1.3 Evidence of Cancer Stem Cells.....	5
1.2 Molecular Chaperones.....	9
1.2.1 Heat shock protein 90 kDa (HSP90)	10
1.2.2 Heat shock protein 70 kDa (HSPA)	11
1.2.3 . DNAJ.....	13
1.2.4 Chaperones as anti-cancer drug targets	13
1.2.5 Chaperones in stem cells and CSC.....	16
1.3 Problem statement	17
1.4 Hypothesis	18
1.5 Aims	18
Chapter 2 : Materials and Methods.....	19
2.1 Materials	20
2.2 Routine culture maintenance of adherent cancer cell lines.....	20
2.3 Anchorage-independent growth assay (mammosphere assay)	20
2.4 Cell surface marker analysis by flow cytometry	21
2.5 ALDEFLUOR assay analysis by flow cytometry	22
2.6 Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE)	23
2.7 Chemiluminescence-based western analysis	23
2.8 Confocal microscopy analysis of anchorage-independent and adherent MCF-7 breast cancer cells.....	24
2.9 Quantitative Reverse Transcriptase-Polymerase Chain Reaction (RT-PCR) for detection of molecular chaperones	24

2.10 Restriction digest identification of plasmid constructs.....	25
2.11 Sub-cloning of DNAJB6L and DNAJB6S coding regions into pCMV plasmid.....	26
2.12 Site-directed mutagenesis (SDM) of H31Q DNAJB6L and DNAJB6S mutants	26
Chapter 3 : Identification and Isolation of BCSC-Enriched Populations from the MCF-7 and MDA-MB-231 Breast Cancer Cell Lines.....	28
3.1 Introduction	29
3.2 Results	31
3.2.1 Cell surface marker-based identification of putative BCSC in MCF-7 and MDA-MB-231 breast cancer cell lines	31
3.2.2 Enzyme activity-based (ALDEFLUOR) identification of putative BCSCs in MCF-7 and MDA-MB-231 breast cancer cell lines	38
3.2.3 Anchorage-independent growth of MCF-7 and MDA-MB-231 breast cancer cell lines.....	41
3.2.4 CSC-like and stem cell-like nature of anchorage-independent MCF-7 colonies.....	42
3.3 Conclusion.....	46
Chapter 4 : Comparison of Molecular Chaperone Expression in Adherent and CD44+/CD24-enriched Anchorage-Independent Populations from the MCF-7 Breast Cancer Cell Line.....	47
4.1 Introduction	48
4.2 Results	49
4.2.1 Quality control tests and data filtration for RT ² PCR Array of anchorage-independent and adherent MCF-7 breast cancer cells.	49
4.2.2 Differential mRNA expression levels of chaperone and chaperone-associated genes in anchorage-independent and adherent MCF-7 breast cancer cells.	58
4.2.3 Differential protein expression levels of selected chaperone and chaperone-associated genes in anchorage-independent and adherent MCF-7 breast cancer cells.....	62
4.3 Conclusion.....	65
Chapter 5 : Development of a Tool to Study the Role of DNAJB6 in CSC Populations	66
5.1 Introduction	67
5.2 Results	69
5.2.1 Sequence analysis of DNAJB6 isoforms DNAJB6L and DNAJB6S	69
5.2.2 Generation of pCMV-DNAJB6L and pCMV-DNAJB6S mammalian expression vectors.	72
5.2.3 Site-directed mutagenesis of pCMV-DNAJB6L and pCMV-DNAJB6S J-domain.....	76
5.3 Conclusion.....	81
Chapter 6 : Discussion and Conclusions.....	82
6.1 Discussion.....	83
6.1.1 MCF-7 and MDA-MB-231 contained different CD44+/CD24-/EpCAM+ subpopulations.....	83
6.1.2 MCF-7 and MDA-MB-231 contained similar ALDEFLUOR profiles.....	86

6.1.3 MCF-7 breast cancer cells were capable of anchorage-independent growth.	86
6.1.4 Anchorage-independent MCF-7 cells were enriched in CD44+/CD24- and CD44+/CD24-/EpCAM+ cells.....	87
6.1.5 Down-regulated molecular chaperone gene expression profile.	88
6.1.6 Several DNAJ were down-regulated at the mRNA level in anchorage-independent MCF- 7 cells.	89
6.1.7 HSP70 regulating proteins, DNAJC12 and BAG1, were up-regulated in anchorage- independent MCF-7 cells.	91
6.1.8 DNAJB1 and DNAJB6 isoforms were down-regulated at both the protein and mRNA levels in anchorage-independent MCF-7 cells.	92
6.1.9 Development of a tool to study the role of DNAJB6 isoforms in cancer stem cell biology.	93
6.2 Conclusions and Future Perspectives	95
Chapter 7 : References	96
Chapter 8 : Appendix.....	109

List of Figures

Figure 1 Cancer Stem Cell theory of tumour organization.	4
Figure 2 Domain architecture of HSP90, HSPA and DNAJ molecular chaperones.	12
Figure 3 Fluorescence threshold gating strategy for CD44/CD24 BCSC cell surface marker analyses	33
Figure 4 Fluorescent gating strategy for identification of CD44/CD24/EpCAM cell populations.	35
Figure 5 Expression of CD44/CD24 and CD44/CD24/EpCAM on MCF-7 and MDA-MB-231 breast cancer cell lines.	37
Figure 6 Fluorescence gating strategy for ALDEFLUOR enzyme activity-based BCSC marker anlyses.	39
Figure 7 Enzyme activity-based (ALDEFLUOR) detection of putative BCSC populations in MCF-7 and MDA-MB-231 breast cancer cell lines.	40
Figure 8 Anchorage-independent growth of breast cancer cell lines.	43
Figure 9 Cancer stem cell-like nature of anchorage-independent MCF-7 breast cancer cells	45
Figure 10 Schematic diagram of the RT2 PCR Array procedure for the analysis and identification of chaperones and co-chaperones displaying differential mRNA levels in anchorage-independent and adherent MCF-7 cells.	52
Figure 11 Analysis of gene expression trend of chaperone and chaperone-associated genes between adherent and anchorage-independent MCF-7 breast cancer cells and selection of housekeeping gene for comparative mRNA expression analysis.	57
Figure 12 Fold-regulation scatter plots for individual chaperone and chaperone-associated gene families from Class ‘A’	60
Figure 13 Chaperone and chaperone-associated genes identified as showing differential expression between anchorage-independent and adherent MCF-7 breast cancer cells.	61
Figure 14 Protein expression levels of selected chaperone and chaperone-associated gene from PCR array.	64
Figure 15 Bioinformatic analysis of DNAJB6L and DNAJB5 isoforms.	70
Figure 16 Homology model of DNAJB6 J-domain.	71
Figure 17 Subcloning of DNAJB6L and DNAJB6S into pCMV.	73

Figure 18 Subcloning of DNAJB6L and DNAJB6S coding regions into pCMV generating pCMV-DNAJB6L and pCMV-DNAJB6S plasmids.	75
Figure 19 Site-directed mutagenesis for the generation of H31Q DNAJB6L and DNAJB6S mutants.	77
Figure 20 Restriction digest identification of H31Q pCMV-DNAJB6L and pCMV-DNAJB6S site-directed mutants.	79
Figure 21 Confirmation of H31Q mutation in pCMV-DNAJB6L and pCMV-DNAJB6S by sequencing.....	80

List of Tables

Table 1 DNAJ molecular chaperones function in cancer and expression in stem-like cells.....	15
Table 2 List of chaperone and chaperone-associated families and genes analyzed on the Heat Shock Protein RT2 PCR Array.	51
Table 3 Results and outcomes of quality control tests for the RT2 PCR Array of anchorage-independent and adherent MCF-7 breast cancer cell line	53
Table 4 Class categorization of chaperone and chaperone-associated genes based on differential Ct values between sample and the Ct threshold cut-off.....	55
Table 5 Differential DNAJ expression in stem-like cells.....	91
Appendix Table 1 Antibody and suppliers used in this study.	110

List of Abbreviations

ABC	ATP-binding cassette
ABCG2	ATP-binding cassette subfamily G member 2
ADP	Adenosine diphosphate
ALDH	Aldehyde dehydrogenase
ALDH1	Aldehyde dehydrogenase 1
AML	Acute myeloid leukemia
APC	Allophycocyanin
ATP	Adenosine diphosphate
ATPase	Adenosine 5'-triphosphatase
BAA	BODIPY-aminoacetate
BAAA	BODIPY-aminoacetylaldehyde
BCSC	Breast cancer stem cell
BLAST	Basic local alignment search tool
bFGF	basic Fibroblast growth factor
BSA	Bovine serum albumin
C-terminal	Carboxy terminal
CD	Cluster of differentiation
cDNA	Complementary deoxyribose nucleic acid
CSC	Cancer stem cell
Ct	Cycle threshold
DEAB	Diethylaminobenzaldehyde
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribose nucleic acid
DNAJ	E.coli HSP40-like/J-domain containing protein
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic Acid
EGF	Epidermal growth factor

EMT	Epithelial-to-mesenchymal transition
EpCAM	Epithelial cell adhesion molecule
ER	Estrogen receptor
FACS	Fluorescent-activated cell sorting
FCS	Fetal calf serum
FITC	Fluorescein isothiocyanate
FSC	Forward scatter
G/F	Glycine/phenylalanine
GDC	Genomic DNA contamination
GSK3 β	Glycogen synthase kinase 3 β
HPD	Histidine-Proline-Aspartic acid
HRP	Horseradish peroxidase
HSP	Heat shock protein
HSP90	Heat shock protein 90 kilodaltons
HSPA	Heat shock protein 70 kilodaltons
Ig	Immunoglobulin
KCl	Potassium chloride
MET	Mesenchymal-to-epithelial transition
mRNA	Messenger RNA
N-terminal	Amino terminal
NEF	Nucleotide exchange factor
NF κ B	Nuclear factor kappa B
NLS	Nuclear localization signal
NOD/SCID	Non-obese diabetic/Severe combined immune-deficiency
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PE	Phycoerythrin
PPC	Positive PCR control
PSA	Penicillin, streptomycin and amphotericin

QC	Quality control
qRT-PCR	Quantitative real-time polymerase chain reaction
RNA	Ribonucleic acid
RPF	Regulators of protein folding
RT	Reverse transcriptase
RT-PCR	Real-time polymerase chain reaction
SCID	Severe combined immune-deficiency
SDM	Site-directed mutagenesis
SDS	Sodium-Dodecyl-Sulphate
SDS-PAGE	Sodium-Dodecyl-Sulphate Polyacrylamide Gel Electrophoresis
SP	Side population
SSC	Side scatter
TBS	Tris-buffered Saline
TBST	Tris-buffered Saline with Tween-20
UV	Ultraviolet light
YT	Yeast extract and tryptone

List of Symbols

α	Alpha
β	Beta
$^{\circ}\text{C}$	Degree Celsius
kDa	kilodalton
M	Molar
mM	Millimolar
$\mu\text{ g}$	Micrograms
$\mu\text{ l}$	Microlitres
L	Litres
g	Grams
mg	Milligrams
kDa	Kilodaltons
ml	Millilitre(s)
%	Percent or g/100 ml
x g	Relative centrifugal force to gravity

Acknowledgements

I would like to thank and extend my gratitude to my supervisor Dr Adrienne Edkins, for her insight, guidance and never-ending support. Dr Edkins' supervision and knowledge throughout the research has been inspiring. Even through the toughest of times both directly and indirectly related to the research she has been compassionate and understanding. I would like to thank my co-supervisor Prof Gregory Blatch and Dr Eva Pesce and Dr Earl Prinsloo for their constant words of encouragement. Big thanks must be given to the members of BioBRU for their friendship and help. To my mother, Jenni, and father, Tony, I wish to thank you for your support (both emotional and financial) that has culminated in this Masters thesis. I would like to thank the NRF for funding my studies.

Chapter 1 : Literature Review

1.1 Cancer Stem Cells: Origin, Evidence and Significance

1.1.1 Cancer Stem Cell Theory

The Cancer Stem Cell (CSC) theory has come to the forefront of oncology research in the past decade with the identification of putative CSC populations in numerous malignant tissue types and a growing understanding of CSC biology. Cancers have and still are being therapeutically treated as homogenous entities, with therapies developed to target and destroy as many cancer cells as possible based on the assumption that all cells of the tumour possess shared biological properties (such a rapid proliferation). All cancer cells were considered to have equal potential to re-establish the cancer upon xenotransplantation (Reya *et al*, 2001). However, cell biological studies on malignant cells suggested an alternative hypothesis that had consequences for the design of future therapeutic strategies. Cancers have not only been found to be comprised of heterogeneous cell populations (Park *et al*, 1971), but additionally there was a requirement for xenotransplantation of large cell numbers in order for re-establishment of the disease *in vivo* (Bruce and van der Gaag, 1962). These observations opposed the theories of homogeneity and stochastic tumour organization for two reasons. Firstly, although heterogeneity can and does, to a certain extent, arise from frequent mutations due to genome instability and rapid proliferation (Grady, 2006), the extent of the heterogeneity observed suggested an organization similar to the hierarchical organization seen in normal non-cancerous tissues. Secondly, the requirement for the use of large cell numbers for disease re-establishment (Bruce and van der Gaag, 1962; Hamburger and Salmon, 1977) was unexpected if every cell displayed an equal capability to re-establish the disease. Both results led to the theory that within the hierarchical cellular organization of cancers, there resided a population of cancer cells that mimic the role of stem cells in normal tissue hierarchies (Reya *et al*, 2001).

The hierarchy model of cancer organization (Figure 1.A) follows the organization of normal non-cancerous tissues and ultimately led to the development of the CSC theory, a theory describing cancer development, maintenance and organization as being driven by a subset of cancer cells that displayed

stem-like properties (Reya *et al*, 2001; Dick, 2003). The definition of a CSC as described by the theory relied on the similarity of biological properties with normal tissue stem cells, cells that are responsible for the development and maintenance of normal tissues. Stem cells are required for organogenesis (organ development) and tissue maintenance through the constant replacement of dying cells (Alison and Islam, 2009). There are two types of stem cells, namely embryonic and adult/somatic stem cells. Pluripotent embryonic stem cells are capable of differentiation into all cells making up an organism (Sylvester and Longmaker, 2004). Adult/somatic stem cells arise from the partial differentiation of embryonic stem cells and are tissue-specific stem cells capable of differentiation into all cell types within the given tissue. These cells are believed to be present for the entire lifespan of the organism and are primarily involved in the maintenance and rejuvenation of tissues. Examples of adult stem cells are the stem cells lining the crypts of the intestine which are responsible for the replenishment of the intestinal lining cells (Barker *et al*, 2007). Stem cells are primitive cells capable of both self-renewal and differentiation (Reya *et al*, 2001; Lobo *et al*, 2006; Alison and Islam, 2009). Self-renewal is the ability to divide to generate cells displaying identical properties of the stem cell parent, a characteristic that suggests that only stem cells can generate stem cells and ensures that stem cells are maintained in the population. Differentiation is the process of cellular maturation whereby cells lose the ability to undergo self-renewal and develop functional and phenotypic characteristics of mature cells or differentiated cells (Figure 1.A). Differentiation results in the loss of stem cell characteristics, meaning that the ability to form the numerous cells making up a specific tissue system is lost. Stem cells are capable of both symmetric cell division, generating daughter stem cells displaying identical properties to the parent cell; and asymmetric cell division, resulting in the formation of a daughter stem cell and a daughter cell that undergoes differentiation resulting in the loss of stem cell characteristics (Lobo *et al*, 2006). The balance between self-renewal and differentiation is strictly regulated by both the intracellular and extracellular environment, often referred to as the stem cell niche. This is the microenvironment in which cell-to-cell interactions occur that regulate the fate of stem cells (Scadden, 2006; Morrison and Spradling, 2008).

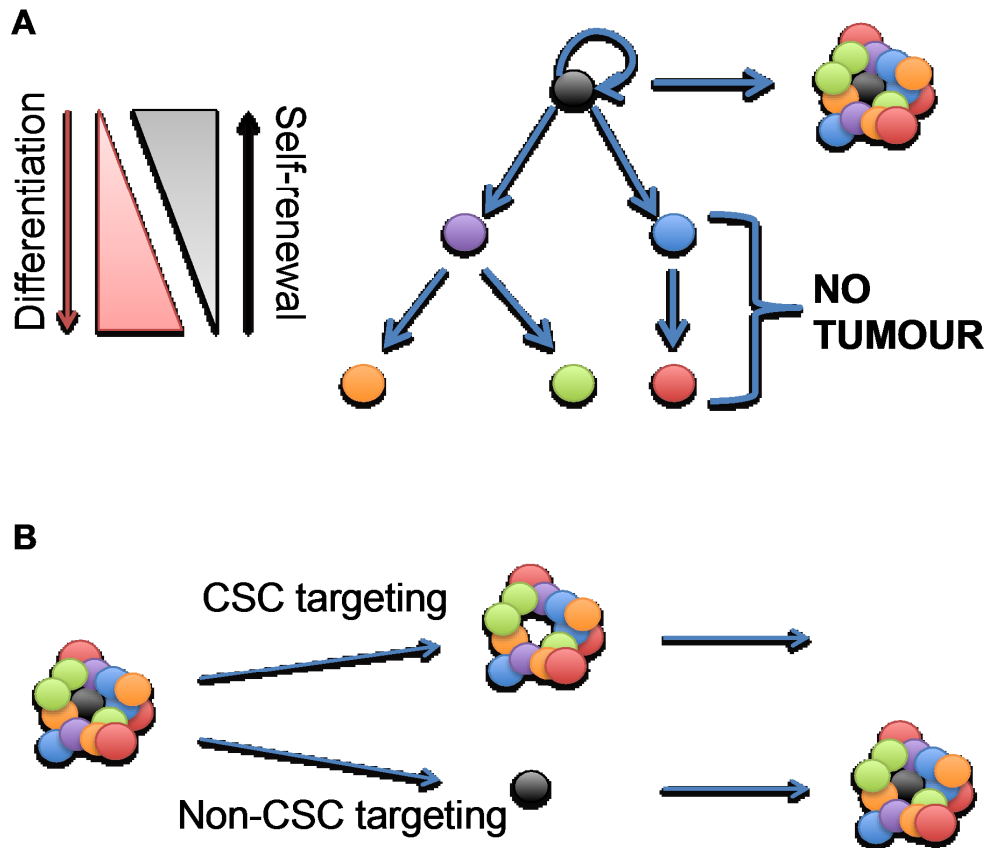


Figure 1| Cancer Stem Cell theory of tumour organization.

(A) Hierarchical structure of cancer cell organization. Cancer stem cells form the apex of tumour heterogeneity (black circles) and are capable of self-renewal (generating additional CSCs) and differentiation (generating tumour heterogeneity). Only CSCs are capable of initiating tumour formation. As cells become more differentiated they lose the ability to self-renew with self-renewal and differentiation properties seen as inversely proportional. (B) Current anti-cancer therapies (non-CSC targeting) target the majority tumour cells that form the non-stem cancer cell population. Due to natural intrinsic resistance the CSC population is able to survive the treatment and together with tumour initiating properties are capable of reforming the tumour. Cancer stem cell targeting therapies may either target the CSC population specifically or both the non-stem and CSC populations. By removing the population responsible for tumour development and maintenance the tumour will be unable to maintain its growth leading to shrinkage of the tumour and a long-term cure without possibility of recurrence.

1.1.2 Definition of a Cancer Stem Cell

Cancer stem cells are cancer cells that, like normal tissue stem cells, are capable of both self-renewal, to maintain the CSC population, and differentiation which leads to cancer cell heterogeneity. This process of cancer cell differentiation has been hypothesized as being an abnormal and uncontrolled process with further expansion of the observed heterogeneity arising from mutations occurring as a result of the rapid proliferation (Clarke *et al*, 2006). The defining characteristic of CSCs was the ability to reform a cell mass with an identical heterogeneity to the cell mass from which the CSC was isolated, suggesting the capability of these cells to undergo both self-renewal and differentiation processes (Reya *et al*, 2001; Clarke *et al*, 2006; Dick, 2008). Just as normal tissue stem cells are responsible for organogenesis and tissue maintenance, so it is thought that CSCs are responsible for tumour development (tumourigenesis) and maintenance (Dalerba *et al*, 2007). The term ‘cancer stem cell’ is used to discriminate those cancer cells that share properties with stem cells from non-stem cancer cells, which do not possess stem-like characteristics. The term ‘cancer stem cell’ does not mean that the cancer stem cell responsible for tumour initiation was necessarily derived from a somatic stem cell. There is much controversy as to which cells give rise to CSC, as it is plausible that stem cells may accumulate oncogenic mutations through the long term cell lifespan that lead to cancer initiation, or that non-stem cancer cells may dedifferentiate to gain stem-like functions leading to cancer initiation (Bjerkvig *et al*, 2005; Clarke *et al*, 2006, Takahashi *et al*, 2007; Chaffer *et al*, 2011).

1.1.3 Evidence of Cancer Stem Cells

In order to positively identify the presence of CSC populations, researchers needed to be able to isolate a distinct population of cells from malignant tissue and demonstrate the exclusive ability of this population to reform malignant tissue with an identical heterogeneity to the original malignant mass *in vivo*. Lapidot and colleagues (1994) are credited with providing the first evidence for the presence of CSCs in research performed on acute myeloid leukaemia (AML). They identified the presence of a definable subset of AML cells, identified by positive expression of CD34 and the absence or lack of

expression of CD38 (CD34+/CD38- phenotype) that displayed the exclusive ability to initiate AML formation *in vivo* in severe combined immunodeficient (SCID) mice (Lapidot *et al*, 1994). It is interesting that the AML CSC population shared an identical phenotype (CD34+/CD38-) with normal tissue stem cells from the hematopoietic system (Huang and Terstappen, 1994). Subsequent work by Bonnet and Dick (1997) led to the development of what is now seen as the gold standard CSC assay, whereby they both confirmed the presence of and determined the frequency of AML CSC using the non-obese diabetic/severe combined immunodeficiency (NOD/SCID) mouse xenotransplantation model (Bonnet and Dick, 1997). These studies identified a small yet definable population of cancer cells that were cable of self-renewal and re-establishing the disease *in vivo*, properties seen as defining characteristics of CSCs.

The first evidence of CSCs in solid tumours was provided in breast cancer, where breast cancer stem cells (BCSCs) were identified based on the expression of CD44 and absence or lack of expression of CD24 (CD44+/CD24- phenotype) (Al-Hajj *et al*, 2003). Cells displaying this phenotype were capable of *in vivo* tumour formation from as few as 200 cells (Al-Hajj *et al*, 2003). Since this important study, evidence of CSCs based on the expression of defined cell surface markers has been provided for a multitude of primary solid tumours including brain tumours (CD133+) (Singh *et al*, 2003), prostate cancer (CD133+/CD44+) (Collins *et al*, 2005), colon cancer (CD133+) (Ricci-Vitiani *et al*, 2007), head and neck squamous cell carcinoma (CD44+) (Prince *et al*, 2007), pancreatic cancer (CD44+/EpCAM+/CD24+) (Li *et al*, 2007) and melanoma (CD20+) (Fang *et al*, 2005). From these putative CSC populations, there was evidence of markers common to CSCs from different tissue tumour types. CSCs have also been isolated from cancer cell lines, most notably a number of breast cancer cell lines (Fillmore and Kuperwasser, 2007; Engelmann *et al*, 2008; Charafe-Jauffret *et al*, 2009b), prostate cancer cell lines (Pfeiffer and Schalken, 2010) and lung cancer cell lines (Ho *et al*, 2007).

Several different markers have subsequently been used to distinguish and isolate BCSCs from non-stem breast cancer cells. Studies by Wright and colleagues (2008) revealed CSC heterogeneity amongst BRCA-1 deficient mouse mammary tumours, with the isolation of two distinct BCSC

populations based on the CD44+/CD24- phenotype and the expression of CD133 (CD133+) (Wright *et al*, 2008), a marker normally associated with brain (Singh *et al*, 2003), prostate (Collins *et al*, 2005) and colon CSCs (Ricci-Vitiani *et al*, 2007). Both these CSC populations displayed high expression of aldehyde dehydrogenase 1 (ALDH1+) (Wright *et al*, 2008), a result in line with gene expression studies on CD44+/CD24- BCSCs (Shipitsin *et al*, 2007). High expression and activity of ALDH1 isoforms has been used to isolate stem/progenitor cells from both normal and malignant breast tissue (Ginestier *et al*, 2007). Comparison between ALDH1+ BCSCs and CD44+/CD24- BCSCs from primary tumours revealed a small overlapping CD44+/CD24-/ALDH1+ population that was more tumourigenic than both the ALDH1+ and CD44+/CD24- populations. Interestingly the CD44+/CD24-/ALDH1- population was unable to form tumours *in vivo*, a result that implied that the tumourigenicity of the CD44+/CD24- population was attributed to the presence of the small ALDH1+ subpopulation within the CD44+/CD24- population (Ginestier *et al*, 2007). Although these markers are commonly used for isolating CSCs, they are by no means the only markers of CSCs. In particular for breast cancer, there is much contention surrounding which markers should be used for isolating CSC populations (Lawson *et al*, 2009).

Perhaps more importantly, it should be noted that all BCSC populations identified and isolated display the same functional properties of *in vitro* and *in vivo* tumourigenicity irrespective of the markers used to isolate the population (Al-hajj *et al*, 2003; Ponti *et al*, 2005; Engelmann *et al*, 2008; Wright *et al*, 2008). Stem-like cells have been isolated and propagated based on the ability to grow anchorage-independently in serum-free non-adherent conditions for both normal neural and breast tissue (Reynolds and Rietze, 2005; Dontu *et al*, 2003), whereas non-stem cells undergo anoikis, a programmed cell death associated with loss of adhesion (Frisch and Frances, 1994; Frisch and Screaton, 2001). Ponti and colleagues adapted this technique to isolate anchorage-independent stem-cell enriched colonies derived from primary breast tumour tissues (Ponti *et al*, 2005). Anchorage-independent colonies (termed mammospheres or tumourspheres) were found to be comprised of between 95 to 98 % CD44+/CD24- expressing cells (Ponti *et al*, 2005), the most widely accepted BCSC phenotype (Al-Hajj *et al*, 2003). The anchorage independent colonies were capable of *in vivo* tumour formation at limiting cell dilutions and expressed high levels of Oct4 (Ponti *et al*, 2005), a

gene associated with stem cells (Yeom *et al*, 1996). Anchorage-independent growth properties have also allowed for the isolation of BCSCs from breast cancer cell lines (Cariati *et al*, 2008). Colonies isolated from breast cancer cell lines were enriched in *in vivo* tumour forming cells and side population cells (Cariati *et al*, 2008).

The side population (SP) refers to a small population of cells within a tissue that is capable of effluxing the fluorescent DNA-binding dyes Hoechst 33342 and Rhodamine 123 and as a result this population can be isolated based on low fluorescence levels with the aid of a flow cytometer (Dean *et al*, 2005). The reason for the efflux of these fluorescent dyes was due to the high expression levels of ATP-binding cassette (ABC) drug transporter family proteins, especially ABCG2 (Dean *et al*, 2005). Side population cells have been found to be enriched in both normal (Goodell *et al*, 1996; Alvi *et al*, 2002) and malignant (Patrawala *et al*, 2005; Cariati *et al*, 2008) stem-like cells. Side population cells isolated from breast cancers are capable of generating tumours *in vivo* (Patrawala *et al*, 2005), tumoursphere formation (Cariati *et al*, 2008) and are enriched in cells displaying the putative BCSC phenotype CD44+/CD24- (Engelmann *et al*, 2008).

The clinical importance of CSCs is related to their potential roles in cancer development as well as metastasis and cancer relapse (Dean *et al*, 2005). One aspect of the theory suggested that CSCs are naturally resistant to chemotherapy due to their slow cell cycling state, capacity for DNA repair and ABC transporter expression levels (Dean *et al*, 2005). As such this population may survive anti-cancer therapy and be capable of regenerating a new cancer mass due to the inherent tumourigenicity. Studies by Li *et al* (2008) provided evidence of this natural resistance when they found that chemotherapy treatment enriched for CD44+/CD24- cells with increased efficiencies in anchorage-independent growth and *in vivo* tumour formation (Li *et al*, 2008). A possible means of targeting CSCs is through the targeting of specific signaling pathways. Breast cancer stem cells have been found to be reliant on several signaling pathways, thereby providing a potential specific target for this therapeutically important population. Both CD44+/CD24- and tumoursphere-derived BCSC populations have been shown to be dependent on Notch receptor signaling, in particular Notch4 (Harrison *et al*, 2010). Notch signaling also plays an important role in normal mammary stem cell self-renewal and progenitor cell

proliferation (Dontu *et al*, 2004). Side population cells, expressing high levels of drug transporters, have been found to show high activity of Akt signaling with inhibition of the signaling pathway leading to reduction in side population cell number (Zhou *et al*, 2007). Current anti-cancer therapies are designed to target the bulk (differentiated) cells within a tumour, while CSCs, capable of initiating tumour growth, are largely unaffected (Figure 1.B) due to the activation of cell survival signaling pathways and overexpression of anti-apoptosis genes. Understanding the fundamental biology of CSCs and how they function may provide insight as to new therapies that can target these cells thereby removing the regeneration and maintenance abilities of cancers.

1.2 Molecular Chaperones

Molecular chaperones can be described as regulatory proteins that are involved in maintaining protein homeostasis in the cell, in order to prevent apoptosis from occurring due to the build up of non-functioning proteins which become toxic to the cell (Lindquist and Craig, 1988; Becker and Craig, 1994). Chaperones and in particular the heat shock protein (HSP) family were first identified based on their overexpression upon heat treatment (Tissieres *et al*, 1974). Although initially described as being expressed as a stress response, chaperones are now understood to play vital housekeeping functions with many HSPs being constitutively expressed. The functions of molecular chaperones include (but are not limited to) suppression of aggregation, folding of nascent and damaged proteins, translocation of proteins into cellular compartments including the nucleus, mitochondria and endoplasmic reticulum, and assistance of protein degradation (Lindquist and Craig, 1988; Becker and Craig, 1994). Of the molecular chaperone families, the Heat Shock Proteins (HSP) is the largest and best understood. The HSP family is comprised of several subfamilies including HSP90, HSPA and DNAJ. Many HSPs have essential roles under stressful conditions such as heat shock, as well as vital roles under physiological conditions (e.g. DNAJB1, HSPA8 and HSP90 β ; Vos *et al*, 2008). Chaperones usually carry out their functions as part of a multi-protein complex consisting of one or more chaperones, co-chaperones and other protein folding regulatory proteins (BAG, CCT and HSF families). Co-chaperones function in the regulation and stimulation of chaperone activity and in the

recruitment of other proteins involved in the multi-protein complex (Wall *et al*, 1994; van der Spuy *et al*, 2001; Felts and Toft, 2003).

1.2.1 Heat shock protein 90 kDa (HSP90)

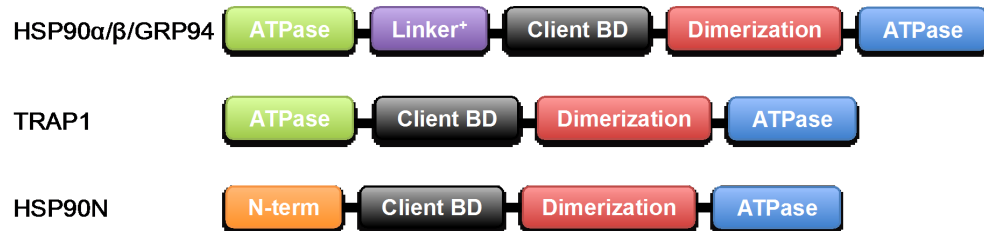
HSP90 is one of the most abundant cellular proteins, making up approximately 1 to 2 % of total protein within a cell (Mahalingam *et al*, 2009). Although HSP90 by itself is seen as a chaperone, it is commonly understood to function as part of a multi-chaperone complex consisting of other chaperones (HSPA), co-chaperones and co-factors (DNAJ) (Sreedhar *et al*, 2004) with HSP90 functioning in this complex as a phosphorylated homodimer (Minami *et al*, 1991; Minami *et al*, 1994). There are 5 HSP90 genes encoding for proteins of approximately 90 kDa in size with the exception of HSP90N. There are two cytosolic isoforms of HSP90 namely HSP90 α (inducible) and HSP90 β (constitutively expressed), as well as HSP90 isoforms located in the endoplasmic reticulum (GRP94) and the mitochondria (TRAP1). Recently an additional HSP90 member, HSP90N, has been described as a plasma membrane-associated isoform and differs from HSP90 α and HSP90 β by an N-terminal ATP binding domain truncation (Grammatikakis *et al*, 2002; Powers and Workman, 2007). HSP90 domains are highly conserved across isoforms with each member consisting of a well characterized N-terminal ATP binding domain (except HSP90N), a middle region consisting of ATP hydrolysis regulating and client binding domains, a C-terminal dimerization domain and a putative C-terminal ATP binding domain (Figure 2.A) (Podromou *et al*, 1997; Panaretou *et al*, 1998; Wandinger *et al*, 2008). Additionally HSP90 α , HSP90 β and GRP94 contain a charged linker region located between the N-terminal ATPase and middle regions. HSP90 functions include, but are not limited to, the regulation of protein conformation and stability, the transportation of proteins into cellular compartments and the stability of protein-protein interactions (Trepel *et al*, 2010). Although chaperones are generally considered to partake in active protein folding, HSP90 is unique in that the majority of its functioning is on folded or partially folded polypeptides, with the chaperone aiding in maturation of proteins rather than folding (Freeman and Morimoto, 1996; Csermley *et al*, 1998). As the chaperone has over 200 client proteins (proteins with which HSP90 interacts), many of which are involved in signaling pathways (for example steroid receptors, kinases, cytoskeletal proteins and transcription factors), HSP90 plays a vital role in cellular homeostasis (Xu *et al*, 1993; Pratt *et al*, 1998).

1.2.2 Heat shock protein 70 kDa (HSPA)

HSPA genes (HSP70, 70-kDa heat shock proteins) are commonly involved in nascent polypeptide folding and in the unfolding or degradation of misfolded or targeted proteins, amongst other functions (Kampinga and Craig, 2010). HSPA forms a critical component of the HSP90 pathway and is involved in HSP90-dependent signaling pathways. HSAs function in a cyclical manner with multiple client association and dissociation cycles controlled by ATP hydrolysis and regulated by co-chaperones, most importantly DNAJ. HSAs when associated to ATP show low client affinity, however J-domain stimulation of HSPA ATP hydrolysis to ADP increases the HSPA client affinity resulting in a stable HSPA-client complex (Greene *et al*, 1998; Michels *et al*, 1999; Fan *et al*, 2003). Nucleotide exchange factor (NEF) co-chaperones catalyze the exchange of ADP for ATP from the ATPase domain resulting in the release of the client protein (Sondermann *et al*, 2001). This process is repeated over numerous cycles in order for HSAs to actively assist in protein *de novo* protein folding (Mayer and Bukau, 2005). The HSP70 family in humans is comprised of 13 isoforms that share a relatively conserved structure with an N-terminal ATPase domain, a β -sheet structured substrate binding domain and an α -helical C-terminus (Figure 2.B) (Kampinga and Craig, 2010). Like HSP90, several HSPA members show specific subcellular localization to the endoplasmic reticulum and mitochondria.

A

HSP90



B

HSPA



C

DNAJ

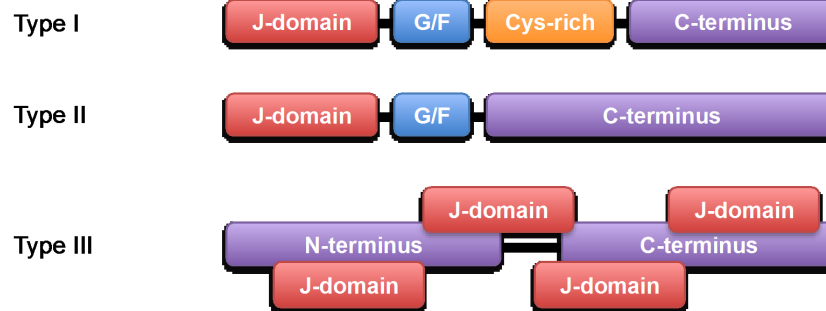


Figure 2| Domain architecture of HSP90, HSPA and DNAJ molecular chaperones.

Conserved domain structures present in HSP90 (A), HSPA (B) and DNAJ (C) molecular chaperone families. ATPase, ATPase domain. Linker⁺, charged linker region. Client BD, client binding domain. Dimerization, dimerization domain. N-term, N-terminus. α C-term, α C-terminal lid. G/F, Glycine/Phenylalanine rich region. Cys-rich, Cysteine-rich region.

1.2.3 . DNAJ

DNAJ are direct co-chaperones of HSPA where they function in the presentation of client peptides to HSPA and stimulation of the HSPA ATPase activity (Kampinga and Craig, 2010). Originally characterized according to the domain similarity to *Escherichia coli* DNAJ, human DNAJs vary both in domain architecture and molecular size, although all contain the highly conserved J-domain. The J-domain is pivotal for the stimulation of the HSPA ATPase activity (Tsai *et al*, 1996; Hennessy *et al*, 2000; Hennessy *et al*, 2005). The DNAJ family consists of 49 members and is divided into three subclasses based on structural similarities with *E.coli* DNAJ. Type I DNAJ (DNAJA isoforms) and Type II DNAJ (DNAJB isoforms) show high structural similarity to the canonical *E.coli* DNAJ, both consisting of an N-terminal J-domain, a glycine/phenylalanine rich region and a largely uncharacterized C-terminus believed to be the substrate binding domain (Lu and Cyr, 1992; Sha *et al*, 2000). The only difference between these two subclasses is that Type I DNAJ (DNAJA), like the *E.coli* DNAJ, contain a cysteine repeat region (containing the motif CXXCXGXG where X is any amino acid) following the glycine-phenylalanine rich region, while Type II DNAJ (DNAJB) lack this region (Figure 2.C) (Cheetham and Caplan, 1998; Vos *et al*, 2008; Kampinga *et al*, 2009). Type III DNAJ (DNAJC isoforms) are highly diverse, most with the only commonality with Type I and Type II DNAJ being the presence of the J-domain (Figure 2.C) (located anywhere along the peptide) (Cheetham and Caplan, 1998; Vos *et al*, 2008; Kampinga *et al*, 2009).

1.2.4 Chaperones as anti-cancer drug targets

Many HSP90 client proteins are oncogenes and thus HSP90 has been described as the cancer chaperone (Whitesell and Lindquist, 2005). In many cases, these oncogenic client proteins induce oncogenesis as a result of mutations that affect the biological functions of the proteins. These mutants have been found to display a greater requirement for HSP90 activity in order to maintain the stability of the mutant (Shimamura *et al*, 2005; Whitesell and Lindquist, 2005), thereby perpetuating the influence that HSP90 has on cancer cells. Additionally, in cancer cells, HSP90 has been found to be present in a higher order activated complex that displays a greater affinity for HSP90 inhibitors when

compared to HSP90 from non-cancerous normal cells (Kamal *et al*, 2003). This specificity, combined with the overexpression of HSP90 in cancers (Jameel *et al*, 1992; Chant *et al*, 1995; Ciocca and Calderwood, 2005), the fact that HSP90 chaperones many oncogene client proteins and the greater requirement for HSP90 to stabilize oncogene proteins (Shimamura *et al*, 2005), makes HSP90 an attractive target in cancer therapeutics. However the main concern with targeting HSP90 is the effect that HSP90 inhibition has on non-cancerous cells, which is primarily due to the essential nature of HSP90 for eukaryotic cell survival (Bagatell and Lindquist, 2004; Whitesell and Lindquist, 2005). Recently cancer cells have been shown to be able to overcome HSP90 inhibition through an alternative chaperone compensation mechanism that results in the up-regulation of numerous chaperones and co-chaperones including HSPA and HSPB members (Maloney *et al*, 2007). HSPA family of chaperones have subsequently been identified as potential anti-cancer chaperone drug targets due to the over-expression in cancers (Chant *et al*, 1995; Ralhan and Kaur, 1995) and pivotal roles that HSAs have been found to play in cell survival and proliferation (Kaur *et al*, 2000; Rohde *et al*, 2005). Additionally support for HSAs as drug target is provided by the HSP90 inhibition compensation mechanism whereby not only is HSPA overexpression a signature but HSPA function is capable of compensating for HSP90 inhibition resulting in cell survival (Bagatell *et al*, 2000; Evans *et al*, 2010). This compensatory function of HSAs is also evident when specific members of the HSPA family are inhibited resulting in the up-regulation of alternate members capable of compensating for the inhibition (Powers *et al*, 2008).

Together with HSP90 and HSAs, there is evidence suggesting that DNAJ may also serve as anti-cancer drug targets. Several DNAJs have been identified as playing significant roles in cancer cell biology through tumour property-promoting or tumour property-suppressing mechanisms (Table 1), or in the case of DNAJA3 regulating both promotion and suppression of tumour properties through the expression and activity of different isoforms (Syken *et al*, 1999; Cheng *et al*, 2005; Kim *et al*, 2004; Kim *et al*, 2005). The predominant effect of cancer-associated DNAJ functions centre around cell survival and apoptosis, in some cases promoting or associated with resistance to anti-cancer therapies (DNAJA1, DNAJA3 and DNAJB11; Table 1), or sensitizing cancer cells to anti-cancer therapies (DNAJA3, DNAJB4, DNAJC10 and DNAJC15; Table 1).

Table 1| DNAJ molecular chaperones function in cancer and expression in stem-like cells

Effect/State	DNAJ	Tissue	Expression	Reference
Tumour property-promoting	DNAJA1	Glioblastoma, Prostate		Salma <i>et al</i> , 2007; Wang <i>et al</i> , 2006
	DNAJA3	Osteosarcoma		Cheng <i>et al</i> , 2005
	DNAJB11	KSHV-associated tumours, Ovarian		Wen <i>et al</i> , 2010; Di Michele <i>et al</i> , 2010
	DNAJC9	Cervical		Lyng <i>et al</i> , 2006
Tumour property-suppressing	DNAJA3	Osteosarcoma, Melanoma, Breast, Kidney, Hematopoietic		Syken <i>et al</i> , 1999; Kim <i>et al</i> , 2004; Kim <i>et al</i> , 2005; Chen <i>et al</i> , 2005; Copeland <i>et al</i> , 2010; Dhennin-Duthille <i>et al</i> , 2010
	DNAJB4	Lung		Tsai <i>et al</i> , 2006; Lin <i>et al</i> , 2010; Chen <i>et al</i> , 2010
	DNAJB6	Breast		Mitra <i>et al</i> , 2008; Mitra <i>et al</i> , 2010
	DNAJC10	Neuroblastoma		Thoms <i>et al</i> , 2009
	DNAJC15	Ovarian, Brain		Shridhar <i>et al</i> , 2006; Lindsey <i>et al</i> , 2006; Hatle <i>et al</i> , 2007
Stem-like:				
Up-regulation	DNAJA1	CD44+ BC	RNA	Shipitsin <i>et al</i> , 2007
		AIG BCSC	RNA	Rappa and Lorico, 2010
		hESC	Protein	Van Hoof <i>et al</i> , 2006
	DNAJA2	hESC	Protein	Van Hoof <i>et al</i> , 2006
	DNAJC7	hESC	Protein	Van Hoof <i>et al</i> , 2006
	DNAJC8	hESC	Protein	Van Hoof <i>et al</i> , 2006
	DNAJC9	AIG BCSC	RNA	Rappa and Lorico, 2010
		hESC	Protein	Van Hoof <i>et al</i> , 2006
	DNAJC10	ER+ AIG BCSC	RNA	Kok <i>et al</i> , 2009
	DNAJC12	ER+ AIG BCSC	RNA	Kok <i>et al</i> , 2009
Down-regulation	DNAJB2	AIG BCSC	RNA	Rappa and Lorico, 2010
	DNAJC19	ALDH+ BCSC	RNA	Charafe-Jauffret <i>et al</i> , 2009a
Differential	DNAJB1	CD44+/CD24- BCSC	RNA	Liu <i>et al</i> , 2007 ¹
	DNAJC13	CD44+/CD24- BCSC	RNA	Liu <i>et al</i> , 2007 ¹

Adapted from Sterrenberg *et al*, 2011. ¹ Gene expression comparison made between CD44+/CD24- BCSCs and normal breast tissue AIG, Anchorage-independent growth isolated. BC, Breast cancer. BCSC, Breast cancer stem cells. ER+, Estrogen receptor positive. hESC, Human embryonic stem cells. KSHV, Kaposi's sarcoma-associated herpesvirus.

1.2.5 Chaperones in stem cells and CSC

There is a paucity of data available on the role of chaperones in stem cell and CSC biology. As chaperones are seen as regulatory proteins within cells (Becker and Craig, 1994), it is likely that chaperones and chaperone-associated genes play vital roles in the highly regulated state of stem-like cells. Several functional and gene expression studies on stem cells, CSCs and stem-like associated signaling pathways, suggested that chaperones and chaperone-associated genes may play key roles in regulating stem-like nature of these cell types (Table 1). Numerous DNAJ genes have been found to be differentially expressed in embryonic stem cells versus normal tissue and several CSC-enriched populations versus non-stem cancer cells, with DNAJA1 (Van Hoof *et al*, 2006; Shipitsin *et al*, 2007; Rappa and Lorico, 2010) and DNAJC9 (Van Hoof *et al*, 2006; Rappa and Lorico, 2010) showing up-regulation at both the mRNA and protein levels for both normal and cancerous stem-like tissues (Table 1). DNAJB6 displays tumour-suppressing properties through the regulation of the Wnt signaling pathway (Mitra *et al*, 2008; Mitra *et al*, 2010), a pathway that has found to be vital for both self-renewal and maintenance of the stem-line state of hematopoietic stem cells (Reya *et al*, 2003), embryonic stem cells (Sato *et al*, 2003), cutaneous CSCs (Malanchi *et al*, 2008) and colon CSCs (Vermeulen *et al*, 2010). Recently HSP90 has been suggested to regulate the Wnt signaling pathway through direct interactions with Glycogen synthase kinase 3 β (GSK3 β), AXIN and β -catenin (Cooper *et al*, 2011). HSP90 (in complex with HSP90 co-chaperone Cdc37) has additionally be found to stabilize and hence prevented the degradation of Fused, a hedgehog signaling molecule that promotes transcription of Hedgehog target genes through stabilization and nuclear localization of Gli transcription proteins (Murone *et al*, 2000; Kise *et al*, 2006). The association of HSP90 with Stat3 has been demonstrated in murine embryonic stem cells grown in the presence of leukaemia inhibitory factor (LIF) (Setati *et al*, 2010). Mucin-1, a tumour specific antigen, forms a complex with both HSPA and HSP90 allowing for transportation into mitochondria (Ren *et al*, 2006). As Mucin-1 has been found to be expressed by both SP BCSCs and non-SP breast cancer cells (Engelmann *et al*, 2008), it suggests that targeting of chaperones may allow for the dual targeting of both non-stem cancer cells and CSC populations. Indeed HSP90 inhibitors have shown to affect CSC population as suggested by the inhibited growth of glioma stem cells (Sauvageot *et al*, 2009) and leukemic stem

cells (Peng *et al*, 2007) after treatment with HSP90 inhibitors. Sauvageot and colleagues (2009) found that there was a greater growth inhibition effect on the CSC population compared to the non-stem cancer cell population which suggests that chaperones may offer a means of targeting both non-stem cancer cells and CSC through a single target.

1.3 Problem statement

The importance of chaperones has already been established in cancer cells with HSP90 and HSPA seen as *bona fide* drug targets. However there is increasing evidence that suggest that CSC populations display a natural resistance to anti-cancer therapies and therefore likely chaperone targeted therapies. There is additionally a lack of knowledge of the role that chaperones play in CSC biology. Indeed several of the DNAJ family members have been shown to play important role in both therapy resistance (DNAJA1, DNAJA3 and DNAJB11) and therapy sensitizing (DNAJA3, DNAJB4, DNAJC10 and DNAJC15) (Table 1). However not only does the impact of many chaperones in CSC biology remain to be determined, but also findings that alternative chaperones are capable of forming compensatory mechanisms as a result of the inhibition of specific chaperones such as the *bona fide* anti-cancer drug targets HSP90 and HSPA. Currently several gene expression studies have identified differential expression patterns of several DNAJ members between normal and stem-like cells in both malignant and normal tissues (Table 1). Although interesting, the studies only analyzed select chaperones and a complete chaperone expression pattern was not obtained. This is particularly the case with the DNAJ family as there are 49 members however the expression state of most of the family remains to determined..

1.4 Hypothesis

Breast cancer cell lines MCF7 and MDA-MB-231 contain putative CSC subpopulations that will express different levels of molecular chaperones compared to the whole cell population. Molecular chaperones that are either up-regulated or down-regulated may influence CSC properties.

1.5 Aims

- Identification and isolation of BCSC-enriched populations from the MCF-7 and MDA-MB-231 breast cancer cell lines;
- Comparison of molecular chaperone expression in whole cell and CSC-enriched populations from the MCF-7 breast cancer cell line;
- Functional analysis of DNAJB6L-H31Q mutant on CSC properties of MCF7 cell line.

Chapter 2 : Materials and Methods

2.1 Materials

MCF-7 (HTB-22) and MDA-MB-231 (HTB-26) breast cancer cell lines were a kind gift from Dr Sharon Prince, University of Cape Town, South Africa. All general reagents were purchased from Sigma-Aldrich (USA) or Saarchem (Merck, South Africa). Tissue culture reagents, 10 x Trypsin-EDTA, foetal calf serum, Dulbecco's Modified Eagle Medium with GlutaMAX™-I and Penicillin-Streptomycin–Amphotericin (PSA) were purchased from Gibco (Invitrogen, UK) and Biowhittaker (UK). Tissue culture plasticware including ultra-low attachment culture flasks were purchased from Corning Incorporated (USA). Western Blotting power pack, Hybond Support Nitrocellulose and Chemidoc™ EQ were purchased from Bio-Rad; UK. Antibodies and suppliers are listed in Appendix: Table 1. Other reagents not listed are referenced within text.

2.2 Routine culture maintenance of adherent cancer cell lines

MCF-7 and MDA-MB-231 breast cancer cell lines were routinely cultured in Dulbecco's modified Eagle's Medium (DMEM) with GlutaMAX (Gibco, Invitrogen) supplemented with 5% (v/v) FCS (Sigma) and 100 U/ml PSA at 37°C and 9% CO₂.

2.3 Anchorage-independent growth assay (mammosphere assay)

Cells were lifted using trypsin solution [0.5% (v/v) trypsin in 1% (w/v) ethylenediaminetetraacetic acid (EDTA)], washed with phosphate-buffered saline [PBS; 137mM NaCl, 27mM KCl, 4.3 mM Na₂HPO₄, 4mM KH₂PO₄, pH 7] and viable cells counted using the trypan blue exclusion method and a hemocytometer. After counting, cells were washed a further two times in PBS and collected by centrifugation at 800 x g for 2 minutes. Cells were resuspended at the desired concentrations in anchorage-independent medium consisting of DMEM with GlutaMAX (Gibco, Invitrogen), 1 x B27 supplement (Gibco, Invitrogen), 20 ng/ml basic fibroblast growth factor [bFGF] (Sigma, Germany), 20 ng/ml epidermal growth factor [EGF] (Sigma, Germany), 5 µg/ml human recombinant insulin

(Novorapid), 10 μ M Heparin (Sigma) and 100 U/ml PSA (Sigma, Germany). After resuspension, cells were passed through a 40 μ m sieve twice to ensure single cell suspension before plating in ultra-low attachment culture flasks or plates (Corning, USA). Cells were grown for 8 days with media supplementation on days 3 and 6. Media supplementation comprised of the removal of half the growth medium and replacement with 2 x concentration of fresh anchorage-independent medium. After day 8, cells were harvested and washed with PBS and used in subsequent analyses.

2.4 Cell surface marker analysis by flow cytometry

Cells were lifted using trypsin solution as described previously, washed with PBS and viable cells counted using the trypan blue exclusion method and a haemocytometer. For anchorage-independent cells, cells were collected by centrifugation, washed, incubated with trypsin solution and counted as described previously. After counting, cells were resuspended to a final concentration of 5×10^6 cells/ml in PBS. A total of 5×10^5 cells (100 μ l) were placed into FACS tubes containing between 0.5 – 1 μ g of anti-CD44, anti-CD24 and anti-EpCAM antibodies or isotype matched control antibodies (Appendix Table 1). Unstained samples containing no antibodies were included. Samples were incubated at 4°C for 60 minutes. After incubation samples were washed twice with ice cold PBS and collected by centrifugation at 800 x g for 2 minutes. For analysis, the 488 nm laser (blue) was used to excite the fluorescein isothiocyanate (FITC) and phycoerythrin (PE) fluorophores with emissions recorded in the 530/30 filter channel (FITC) and the 585/42 filter channel (PE) respectively. The 633 nm laser (red) was used to excite the allophycocyanin (APC) fluorophore with emission recorded in the 660/20 filter channel (APC). Compensation controls were used to establish compensation values for FITC and PE fluorescent overlap and this compensation was applied to all analyses. A total of 30 000 live events were recorded for all samples. Data analyses were conducted on FlowJo software (Tree Star Inc, USA). Two-colour dot plots were constructed showing FITC (CD24) fluorescence on the y-axis and APC (CD44) fluorescence on the x-axis. Isotype controls were used to establish fluorescence threshold gates with the gates set on the isotype control to exclude the major population of cells. Gates were then copied onto the respective samples in order to determine the CD44/CD24 expression profile of the sample. For EpCAM expression analyses, a histogram showing PE (EpCAM)

fluorescence was displayed for the isotype control sample. A PE fluorescence threshold gate was set such that the major isotype control population was excluded. Using a two-colour dot plot of FITC (CD24)/APC (CD44) fluorescence, the PE (EpCAM) fluorescence of the CD44+/CD24- population (Q3) was displayed as a histogram. The PE fluorescence threshold gate was copied onto the respective EpCAM histograms in order to analyze the CD44+/CD24-/EpCAM+ phenotype.

2.5 ALDEFLUOR assay analysis by flow cytometry

All reagents were prepared as per the manufacturer's instructions (StemCell Technologies) prior to first use. Cells were lifted with trypsin solution, washed with PBS and counted using the trypan blue exclusion method and a haemocytometer. For anchorage-independent cells, cells were washed and then incubated in PBS containing trypsin solution prior to counting. Cells were collected by centrifugation at 300 x g for 7 minutes and resuspended to a final concentration of 1×10^6 cells/ml in ALDEFLUOR assay buffer. Activated ALDEFLUOR reagent was added to a final concentration of 150 μ M per 1×10^6 cells (test sample). From this solution, 5×10^5 cells were removed and added to a tube containing the ALDH inhibitor, diethylaminobenzaldehyde [DEAB] at a final concentration of 150 μ M (inhibitor control). Both test and inhibitor control samples were incubated at 37°C for 45 minutes. Samples were washed twice in ice cold PBS and resuspended in ice cold ALDEFLUOR assay buffer. Samples were kept on ice until flow cytometry analysis using the FACS Aria III (BD Biosciences, Belgium). For analysis, ALDEFLUOR fluorescence of 1×10^5 cells was measured by excitation using the 488 nm blue laser and emission detection in the FITC 530/30 filter channel. ALDEFLUOR analysis was performed using FACSDiva software (BD Biosciences, Belgium) with plots constructed by displaying side scatter height (SSC-H) profiles on y-axis versus FITC height (FITC-H) fluorescence profiles on the x-axis. The ALDEFLUOR positive gates were constructed using the inhibitor control sample and were standardized such that an ALDEFLUOR positive population of 0.1% was taken as acceptable background in the inhibitor control. The gates were then copied onto the related samples lacking the inhibitor.

2.6 Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Cell lysates were prepared in 5 x SDS-PAGE sample buffer without bromophenol blue [0.05M Tris-HCL, 10% (v/v) glycerol, 5% (v/v) β -mercaptoethanol, 2% (w/v) SDS] and stored at -20°C until analysis. Prior to separation, protein concentrations were determined by absorbance at 280/360 using the Nanodrop2000 (Thermo Scientific). Lysates were diluted in 5x SDS-PAGE sample buffer containing 1% (w/v) bromophenol blue. Proteins were separated by SDS-PAGE using the protocol established by Laemmli (1970) using a 4% stacking (0.5M Tris-HCl, pH 6.8) and 12% resolving gel (1.5M Tris-HCl, pH 8.8). Gels were run at 100 V for approximately 1.5 hours in SDS-PAGE running buffer [0.25mM Tris, 192mM glycine, 1% (w/v) SDS].

2.7 Chemiluminescence-based western analysis

Cell lysates separated during SDS-PAGE were transferred to nitrocellulose membrane (Towbin *et al*, , 1979) at 100 V for 2 hours in transfer buffer [13 mM Tris-HCl, 100 mM glycine and 20 % (v/v) methanol] at 4°C. Protein transfer onto nitrocellulose was confirmed by Ponceau staining [0.5% (w/v) Ponceau S in 1% (v/v) glacial acetic acid]. Nitrocellulose membranes with transferred proteins were blocked with blotto solution [5% (w/v) fat free milk powder in Tris-buffered saline (TBS, 50 mM Tris, pH 7.5, 150 mM NaCl)] for 30 minutes and incubated with primary antibody (Appendix Table 1) overnight at 4°C, diluted to the manufacturer's recommended dilution in blotto solution. The membrane was rinsed four times in TBS containing Tween-20 [TBST; 1% (v/v) Tween-20 in TBS] for 15 minutes replacing the TBST each rinse. The membrane was incubated with species-specific secondary antibodies conjugated to horseradish peroxidase [HRP] at 1:5000 dilution in 5% (w/v) blotto solution for 45 minutes with agitation at room temperature. After secondary antibody incubation, the membrane was rinsed four times with TBST for 15 minutes each, replacing the TBST at each rinse. Detection of protein bands was performed using the Enhanced Chemiluminescence kit (GE Healthcare; UK) on the Chemidoc EQ system (Biorad; UK). Actin was used as the loading control in each case and densitometry was performed using ImageJ software.

2.8 Confocal microscopy analysis of anchorage-independent and adherent MCF-7 breast cancer cells

Anchorage-independent colonies and adherent cells were seeded into eight-well chambered coverslips overnight. Cells were fixed with a flash incubation in ice-cold methanol and left to air dry, after which the cells were washed in distilled water containing 1 $\mu\text{g/ml}$ of DNA binding dye Hoechst 33342 to stain nuclei. The coverslips were mounted onto glass slides using Dako mounting media (Sigma). Differential contrast (DIC) and fluorescent images were taken using Zeiss LSM510 meta confocal microscope (Zeiss, Germany).

2.9 Quantitative Reverse Transcriptase-Polymerase Chain Reaction (RT-PCR) for detection of molecular chaperones

Total RNA was isolated from anchorage-independent and adherent MCF-7 breast cancer cells on day 8 of growth using the NucleoSpin RNA/Protein isolation kit (Macherey-Nagel) according to manufacturer's instructions. Total RNA was quantified by measuring the absorbance at 260nm using the NanoDrop2000 spectrophotometer (Thermo Scientific). Reverse Transcriptase PCR (RT-PCR) using the RT² First Strand Kit (SA Biosciences) was used to convert 1 μg total RNA to cDNA for quantitative PCR (qPCR) analysis. After the RT-PCR procedure, cDNA was immediately added to the RT² qPCR master mix (SA Biosciences) and equal volumes loaded into each well of a 96-well PCR Array plate (SA Biosciences) as per the manufacturer's instructions. Plates were stored at -20°C overnight and centrifuged for 2 minutes at 1250 x g prior to the thermocycling. The thermocycling reaction was performed using a Bio-Rad iCycler with the following cycling parameters: Cycle 1 (1 cycle), 95°C for 10 minutes; Cycle 2 (40 repeat cycles), 95°C for 15 seconds and 60°C for 1 minute; Cycle 3 (1 cycle), 95°C for 30 seconds; Cycle 4 (80 repeat cycles), starting temperature 55°C for 10 seconds with an increase in temperature by 0.5°C increments after each repeat; Cycle 5 (1 cycle), 22°C incubation for unspecified period. Melt curves were performed to determine the presence of non-specific or contaminating products that would influence the qPCR analysis, with all wells determined as showing single peak melting points indicating the absence of contaminating or non-specific product

formation. Data analysis was performed using the manufacturer's online software (SA BioSciences; www.sabiosciences.com/pcrarraydataanalysis.php).

2.10 Restriction digest identification of plasmid constructs

The pcDNA-DNAJB6L and pcDNA-DNAJB6S constructs were a kind gift from Prof Harm Kampinga (Hageman and Kampinga, 2009). Plasmid DNA was extracted from filter paper by soaking the filter paper in 200 µl of TE buffer (10 mM Tris, 1 mM EDTA, HCl, pH 8) for 1 hour. Plasmid constructs were transformed into chemically competent JM109 *E.coli* cells. Fifty microlitres of competent cells and 1 µg of plasmid DNA were incubated on ice for 2 minutes prior to heat shock at 42°C. After heat shock, cells were incubated on ice for a further 5 minutes. Cells were mixed with 500 µl of 2 x YT broth (16 g/l Tryptone, 10 g/l Yeast extract, 5 g/l NaCl) and incubated at 37°C for 30 minutes before 150 µl of the transformation was plated on 2 x YT agar plates and incubated overnight at 37°C. Single colony transformants were used to inoculate 5 ml 2 x YT broth overnight at 37°C with shaking (200 rpm). Overnight culture of pCMV-HSJ1a was made by adding 100 µl of pCMV-HSJ1a glycerol stock to 5 ml 2 x YT broth and incubating overnight at 37°C with shaking (200 rpm). Plasmids were extracted using the QIAGEN Miniprep kit according to manufacturer's instructions (QIAGEN). Restriction digest identification and confirmation of the pcDNA-DNAJB6 constructs was performed using *Bam*HI and *Xho*I to release the DNAJB6L and DNAJB6S coding regions, and *Bam*HI to release the HSJ1a coding region from the pCMV construct. Both plasmids were incubated without restriction enzymes, with *Bam*HI or *Xho*I alone and a combination of *Bam*HI and *Xho*I for 2 hours at 37°C. Restriction digestion reactions were analysed using a 0.8% (w/v) agarose gel in Tris-Acetate-EDTA buffer (TAE; 40 mM Tris, 20 mM acetic acid, 1 mM EDTA) containing 0.5 µg/ml ethidium bromide run at 100 V for approximately 1 hour. DNA bands were visualized using UV illumination with the ChemidocEQ (BioRAD). Approximate band size (in base pairs) was estimated using *Pst*I digestion of λ DNA.

2.11 Sub-cloning of DNAJB6L and DNAJB6S coding regions into pCMV plasmid

The coding regions for DNAJB6L and DNAJB6S were released from the respective pcDNA constructs by *Bam*HI and *Xho*I digestion at 37°C for 2 hours. Additionally the pCMV-HSJ1a construct was sequentially digested with *Bam*HI and *Xho*I at 37°C for 1 hour each, in order to isolate the pCMV backbone with complimentary 5' and 3' overhangs to allow for ligation of the DNAJB6L and DNAJB6S coding regions into pCMV generating pCMV-DNAJB6L and pCMV-DNAJB6S. Digests were separated on a 0.8% (w/v) agarose TAE gel containing 0.5 µg/ml ethidium bromide at 100 V for approximately 1 hour. Bands were visualized on a UV transilluminator and bands corresponding to the sizes of DNAJB6L, DNAJB6S and pCMV backbone were excised and purified from the gel using the Zymoclean gel DNA recovery kit according to the manufacturer's instructions (Zymo Research). Isolated DNAJB6L and DNAJB6S coding regions were ligated into the pCMV backbone to generate pCMV-DNAJB6L and pCMV-DNAJB6S constructs upon incubation at 4°C with 0.5 U T4 DNA ligase. Resulting ligations were transformed into DH5α *E.coli* cells and incubated overnight on YT agar containing 50 µg/ml kanamycin. Transformants were isolated as described previously and screened for the presence of the inserts by restriction digestion with *Bam*HI and *Xho*I alone or in combination, in order to identify pCMV-DNAJB6L and pCMV-DNAJB6S transformants. The plasmids were confirmed by sequencing (Inqaba Biotec).

2.12 Site-directed mutagenesis (SDM) of H31Q DNAJB6L and DNAJB6S mutants

Complementary mutagenesis primers incorporating the H31Q mutation into the DNAJB6L and DNAJB6S coding regions were constructed using the Primer-BLAST on the NCBI website (<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>). Purified pCMV-DNAJB6 plasmids (50 ng) were added to the SDM mix containing 125 ng of each primer (forward and reverse), 10 x *Pfu* reaction buffer, distilled water and 1mM dNTPs prior to the addition of 2.5 U *Pfu* DNA polymerase. After an initial denaturation step for 30 seconds at 95°C, the reaction mixture underwent 16 cycles of denaturation (30 seconds at 95°C), annealing (1 minute at 55°C) and extension (5 minutes at 68°C).

After thermocycling, the reaction mix was placed on ice for 2 minutes prior to *DpnI* (10 U) treatment at 37°C for one hour and subsequent transformation into competent *E. coli* DH5 α . Single colony transformants were selected for plasmid isolation as previously described. Plasmids were screened for the presence of the mutation using restriction digestion with the enzyme *BbvI* and mutants confirmed by sequencing.

Chapter 3 : Identification and Isolation of BCSC-Enriched Populations from the MCF-7 and MDA-MB-231 Breast Cancer Cell Lines

3.1 Introduction

Cancer stem cells (CSC) are thought to have a role in cancer initiation and development (Reya *et al*, 2001; Visvader and Lindeman, 2008), resistance to common anti-cancer therapies (Dean *et al*, 2005; Li *et al*, 2008) and metastasis (Hermann *et al*, 2007, Pang *et al*, 2010). The CSC theory suggests that the large majority of the tumour is composed of non-stem tumour cells. Therefore, the biological characteristics of the minor CSC population (>1%) are often not observable when studying the tumour as a whole. Specific techniques are required in order to accurately identify the CSC population, and prevent the characteristics of the non-stem cancer cells, which represent the large majority of cells within a cancer, from masking the biological and functional properties of the CSCs (Clarke *et al*, 2006; Charafe-Jaffret *et al*, 2009b). It must be noted that there are as yet no incontrovertible markers of CSCs, and therefore many of the available techniques that allow for the enrichment of CSC-like populations are accepted within the field (Al-Hajj *et al*, 2003; Ponti *et al*, 2005; Ginestier *et al*, 2007).

The published methods for identification and isolation of CSCs can be separated into three groups, namely (1) cell surface marker-based techniques, relying on the expression of a single or combination of cell surface proteins; (2) functional-based techniques, which make use of stem-like biological properties to identify and isolate populations of cells; and (3) enzyme activity-based techniques that rely on both the expression and activity of a particular protein. Cell surface marker-based techniques usually rely on the expression of cell surface markers termed cluster of differentiation (CD) proteins (Clarke *et al*, 2006; Charafe-Jauffret *et al*, 2009b). Generally, but not always, a combination of markers (defined as a marker phenotype) is used to identify the desired CSC population. For BCSCs the most common and widely used marker phenotype is CD44+/CD24-, denoting the (high) expression of the CD44 cell surface protein and the lack (low) of expression of the CD24 cell surface protein (Al-Hajj *et al*, 2003; Fillmore and Kuperwasser, 2008). Functional-based techniques are based on stem-like biological properties of the cells and assess how closely the cells resemble stem-like cells in terms of biological function. The most widely used *in vitro* functional-based technique is the anchorage-independent growth assay, which tests the ability of adherent cells to survive and proliferate as spheres in an anchorage-independent serum-free environment. Somatic stem cells

isolated from normal adherent tissue are capable of surviving in a serum-free anchorage-independent environment whereas non-stem-like cells undergo a programmed cell death associated with loss of adhesion by adherent cells known as anoikis (Frisch and Francis, 1994; Frisch and Screaton, 2001; Dontu *et al*, 2003). The third category of CSC identification and isolation techniques relies on the expression of a family of proteins together with the altered or enhanced activity of these proteins for the identification of CSC enriched populations. Examples of these techniques include the side population and ALDEFLUOR assays. Both assays are fluorescence-based, with the side population assay identifying stem-like cells based on the efflux of the Hoechst 33342 dye due to overexpression of the ATP-binding cassette (ABC) family (especially the ABCG2 isoform) (Goodell *et al*, 1996). The ALDEFLUOR assay is based on the production of a fluorescent product by active aldehyde dehydrogenase (ALDH) isoforms, with the identification relying on the over-expression and activation of ALDH isoforms (especially the ALDH1 isoform), to detect a stem-like (ALDH^{Hi}) cell population (Ginestier *et al*, 2007). Both of these techniques allow for the identification and isolation of stem-like cells from both normal and malignant tissues (Ginestier *et al*, 2007).

As there is much contention and debate surrounding the choice of CSC identification and isolation techniques, as well as this being the first study at Rhodes University to establish CSC identification and isolation techniques, we undertook to evaluate three BCSC identification and isolation techniques for the two breast cancer cell lines, MCF7 and MDA-MB-231. For the marker-based technique, two common BCSC phenotypes were analyzed, that being CD44+/CD24- and CD44+/CD24-/EpCAM+ phenotypes, where '+' denotes the (high) expression of the cell surface marker while '-' denotes the absence or low expression of the cell surface marker. The ALDEFLUOR assay was used as the enzyme activity-based technique while the anchorage-independent assay was used as the functional-based CSC-enrichment technique.

3.2 Results

The proportion of putative BCSC populations in the MCF-7 and MDA-MB-231 breast cancer cell lines were established and compared using three published CSC identification and isolation techniques (as described above). The MCF-7 breast cancer cell line was selected as it is a cell line already established in BCSC research (Charafe-Jauffret *et al*, 2009b). The MDA-MB-231 breast cancer cell line, derived from metastatic and basal cell type malignant breast tissue, was also included as it has been found to express high levels of the putative BCSC cell marker CD44 (Fillmore and Kuperwasser, 2008).

3.2.1 Cell surface marker-based identification of putative BCSC in MCF-7 and MDA-MB-231 breast cancer cell lines

The expression profiles of BCSC-associated cell surface markers (CD44, CD24 and EpCAM) were determined for the MCF-7 and MDA-MB-231 breast cancer cell lines using flow cytometry. Cell surface marker-based identification of BCSC is the most widely used technique for the isolation and characterization of the population. The currently accepted BCSC phenotype is the expression of CD44 in the absence of CD24 expression (CD44+/CD24- phenotype). Additionally the combination of EpCAM expression (CD44+/CD24-/EpCAM+ phenotype) has been used to identify and isolate BCSC (Al-Hajj *et al*, 2003). In order to determine the BCSC phenotypes of the breast cancer cell lines, cells were triple stained for CD44, CD24 and EpCAM using fluorophore-conjugated antibodies against CD44 (APC), CD24 (FITC) and EpCAM (PE), allowing for the co-expression (phenotype) of the markers to be analyzed. Isotype controls, which consisted of non-specific antibodies of the same immunoglobulin class conjugated to the matched fluorophore, were used to identify non-specific antibody binding and to establish the fluorescence threshold gates (Figure 3). Figure 3 shows a diagrammatic representation of the gating strategy used to analyze the experimental data. The gating strategy began with the selection of the events (representing cells) to be analyzed (Figure 3: Panel I [Population Gate]). A population was selected based on the side scatter (SSC; Figure 3: Panel I [y-axis]) and forward scatter (FSC; Figure 3: Panel I [x-axis]) fluorescence profile. Side scatter fluorescence was inversely proportional to cell granularity, therefore the more granular the event the

lower the SSC fluorescence. Granularity was an indication of cell viability with cell debris and cells undergoing apoptosis displaying a high granularity and thus low SSC fluorescence. Forward scatter refers to the size of the events and was directly proportional to FSC fluorescence. Therefore the larger the size of an event, the greater the FSC fluorescence recorded. The population gate on the SSC versus FSC plot (Figure 3: Panel I) was used to select only a subpopulation of events, considered to be live events, for subsequent analysis.

Isotype controls were used to establish the fluorescence threshold gate for each of the antibodies in order to determine expression status of the marker. A two-colour dot plot, containing only the cells from the live gate (Figure 3; Panel I), of the isotype control sample was created displaying FITC (CD24) fluorescence (Figure 3: Panel II [y-axis]) and APC (CD44) fluorescence (Figure 3: Panel II [x-axis]). A horizontal gate representing the FITC (CD24) fluorescence threshold (Figure 3: Panel III) was established such that the major population of events on the graph were excluded. This gate recorded any event displaying a FITC (CD24) fluorescence intensity below this threshold gate as being negative or not expressing the marker (i.e. CD24-), while every event with a FITC (CD24) fluorescence intensity greater than (above) the threshold fluorescence gate was recorded as showing positive expression of the marker (i.e. CD24+). A vertical gate representing the APC (CD44) fluorescence threshold was then established in the same manner to allow for the determination of CD44+ and CD44- cells (Figure 3: Panel IV). Events displaying an APC (CD44) fluorescence intensity lower (left) than the APC (CD44) threshold fluorescence gate were recorded as not expressing the marker (i.e. CD44-), while events displaying an APC (CD44) fluorescence intensity greater than the APC (CD44) threshold fluorescence gate were recorded as being positive for the marker (i.e. CD44+). The gating strategy resulted in the identification of four distinct quadrants that corresponded to the different CD44/CD24 expression phenotypes, namely Q1: CD44-/CD24+; Q2: CD44+/CD24+; Q3: CD44+/CD24- and Q4: CD44-/CD24- (Figure 3: Panel V). The gates were set using the isotype control sample and subsequently copied onto the test samples (sample containing antibodies specific to CD44 and CD24) to identify the CD44/CD24 phenotypic profile of the cell lines (Figure 3: Panel VI).

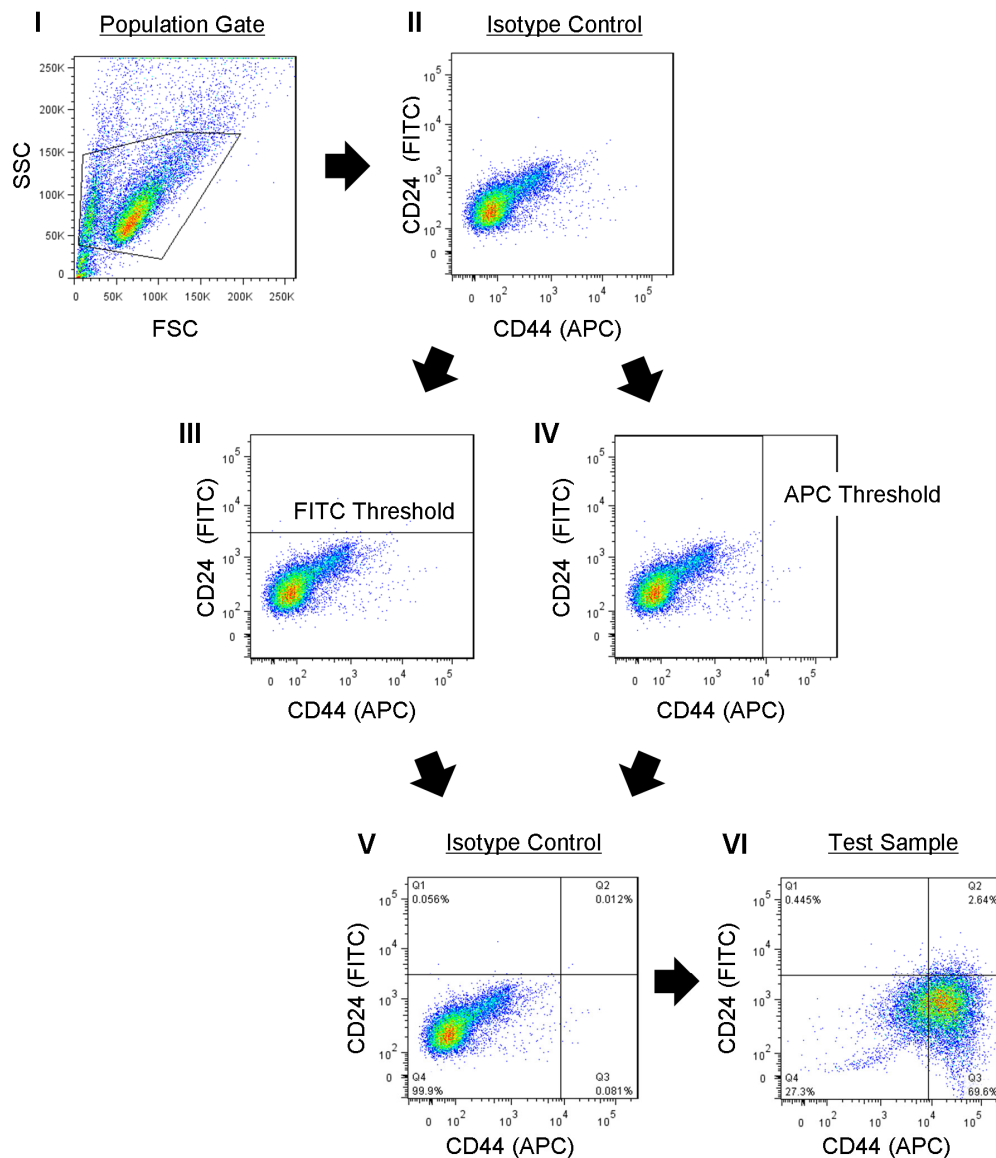


Figure 3| Fluorescence threshold gating strategy for CD44/CD24 BCSC cell surface marker analyses

The population gate was constructed using the side scatter (SSC; y-axis) and forward scatter (FSC; x-axis) two colour dot plot so as to exclude events displaying high granularity (low SSC fluorescence) and large even size (high FSC fluorescence) which signify cells undergoing or having undergone apoptosis. Only the events within the population gate were analyzed. (II) The live events (Panel I) from the isotype control sample were analyzed for FITC (CD24; y-axis) and APC (CD44; x-axis) fluorescence using a two-colour dot plot to determine non-specific binding of the respective immunoglobulin chain class to the surface of cells. (III) A FITC threshold gate was constructed such that the percentage of events in the isotype sample displaying positive FITC fluorescence was less than 0.1%. Any events with a fluorescence intensity above the threshold gate were recorded as being positive for the marker (CD24+), while events with a fluorescence intensity lower than the threshold were recorded as displaying low or no expression of the marker (CD24-). (IV) An APC threshold gate was constructed on the x-axis such that less than 0.1% of the isotype control sample displayed positive APC fluorescence. Events with a fluorescence intensity greater than the threshold (to the right) were recorded as displaying a positive expression for the marker (CD44+) while events with a fluorescence intensity lower than the threshold (to the left) were recorded as having low or no expression of the marker (CD44-). (V) The combination of both fluorescence threshold gates creates four quadrants (Q1- CD44-/CD24+; Q2- CD44+/CD24+; Q3- CD44+/CD24-; Q4- CD44-/CD24-) that can be used to characterize the expression phenotype of the two markers in question. (VI) The population and fluorescent threshold gates were then copied onto the respective test samples (sample containing antibodies specific for the markers) allowing for the identification of the proportions of cells with the different phenotypes to be identified.

As it is not possible to represent the three-colour staining on a single plot, the following gating strategy was used to identify the proportion of cells with the CD44+/CD24-/EpCAM+ phenotype associated with BCSCs. Using the isotype control sample, a histogram was constructed showing PE (EpCAM) fluorescence on the x-axis (Figure 4.A). Fluorescence threshold gates representing a negative (no or low expression) population and a positive (or high) expression population were constructed such that the negative (no expression, EpCAM-) threshold gate recorded the majority of events from the isotype sample while the positive (EpCAM+) threshold gate recorded a population corresponding to less than 0.1% of the isotype sample (Figure 4.A). Using the FITC (CD24) and APC (CD44) two-colour dot plot of the test sample (containing antibodies specific to CD24, CD44 and EpCAM), only the CD44+/CD24- population (Figure 4.B: Quadrant 3 [red rectangle]) was selected for EpCAM expression analysis. The threshold PE (EpCAM) fluorescence gates were copied from the isotype samples to the corresponding test samples to determine the proportion of CD44+/CD24- events displaying EpCAM expression (CD44+/CD24-/EpCAM+) (Figure 4.B). For visual purposes the CD44/CD24/EpCAM phenotype analyses have been illustrated as an overlay of the PE (EpCAM) isotype fluorescence (Figure 4.C: black histogram) and CD44+/CD24- population PE (EpCAM) fluorescence (Figure 4.C: blue histogram).

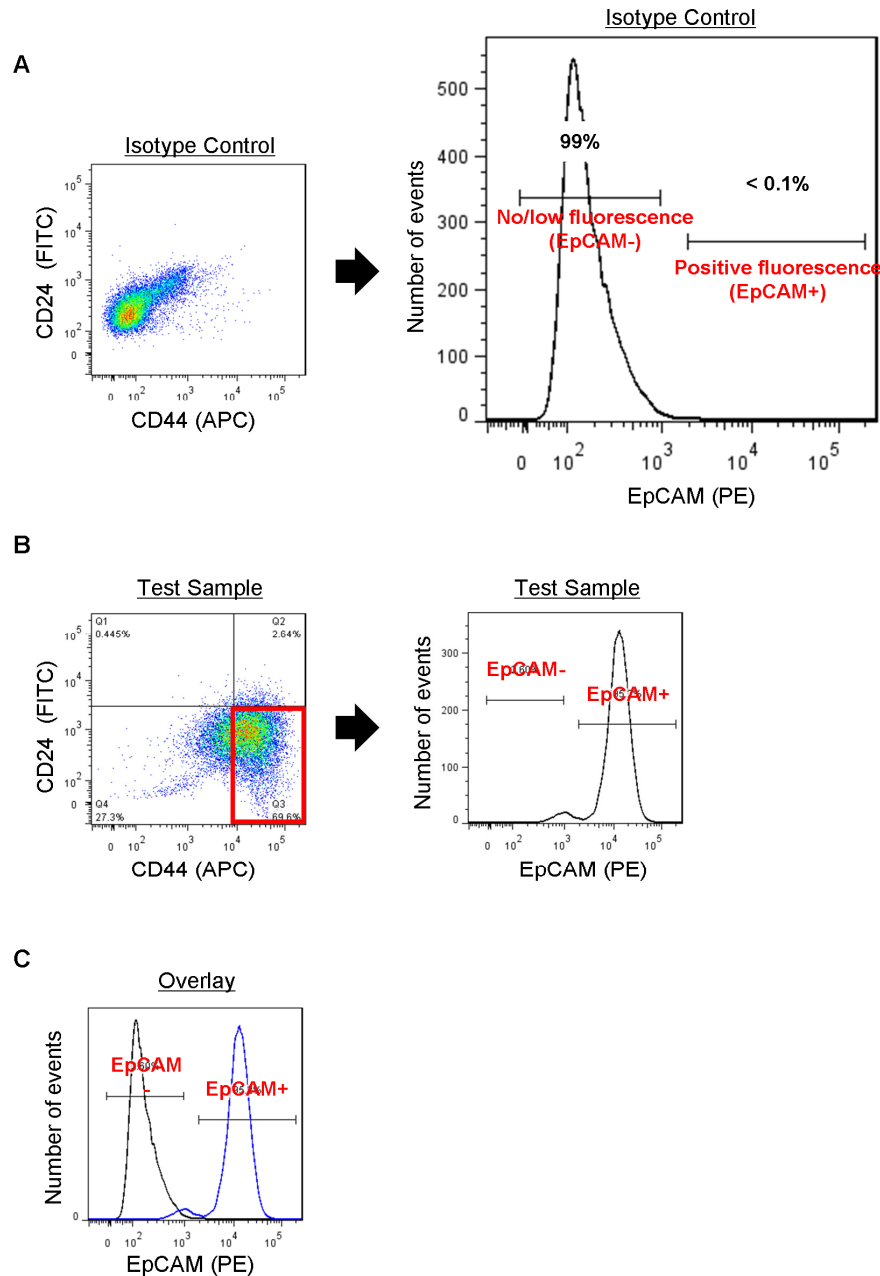


Figure 4| Fluorescent gating strategy for identification of CD44/CD24/EpCAM cell populations.

As only two fluorescence channels can be represented on a single dot plot, the analysis of the CD44/CD24/EpCAM phenotype required an additional gating strategy. (A) Using the different population gates (namely Q1 – Q4) from Panel I in Figure 3, the isotype control for EpCAM (PE) was displayed on a one-colour histogram (PE; x-axis). A negative fluorescence threshold gate was constructed to account for events displaying background fluorescence and recorded events displaying low or no expression of the marker (EpCAM-). A positive fluorescence threshold gate was constructed such that less than 0.1% of the isotype sample was recorded as displaying positive PE fluorescence. Events with PE fluorescence intensities in the range equivalent to the positive fluorescence threshold gate were recorded as displaying positive expression of the marker (EpCAM+). The PE fluorescent threshold gates were copied onto the respective test samples. (B) For EpCAM expression analysis only the CD44+/CD24- (Q3; red) population was analyzed to determine the proportion of CD44+/CD24-/EpCAM+ cells. (C) An overlay histogram of the isotype control (black) and the CD44+/CD24- population (blue) from the test sample shows that EpCAM expression was signified by a shift of the histogram to the right due to an increase in fluorescence intensity.

For each replicate, an isotype control was used to establish the respective fluorescence threshold gates (Figure 5.A: Panel I). The MCF-7 breast cancer cell line was found to be comprised of two major and two minor cell populations. The major populations consisted of CD44-/CD24- cells representing approximately half of the original cell line (55.4%; Figure 5.A Panel II, Q4), and a CD44+/CD24- cell population of approximately 32.5% of the original cell line (Figure 5.A: Panel II, Q3). Additionally, minor CD44+/CD24+ and CD44-/CD24+ populations of approximately 9% and 2% respectively were observed (Figure 5.A: Panel II, Q1 and Q2). Approximately 30% of the total population of MCF-7 breast cancer cells were determined to be CD44+/CD24-/EpCAM+, as shown by the increase in fluorescence intensity and therefore shifts of the blue histogram in Panel III (Figure 5.A) to the right of the isotype control histogram (Figure 5.A: Panel III, black histogram).

A different phenotypic profile was observed for the MDA-MB-231 breast cancer cell line, which displayed three separate populations of cells based on CD44 and CD24 expression. The major population (61.3%) of cells had the CD44+/CD24- phenotype, while the CD44+/CD24+ and CD44-/CD24- made up 33.3% and 4.88% respectively (Figure 5.B: Panel II, Q3, Q2 and Q1 respectively). The 0.49% of cells in the CD44-/CD24+ quadrant was taken as negligible (Figure 3.B: Panel II, Q1). The CD44+/CD24- cell population of the MDA-MB-231 breast cancer cell line displayed very low EpCAM expression levels (Figure 5.B: Panel III, blue histogram) and therefore most of the cells in Q3 (representing CD44+/CD24- cells; Figure 5.B: Panel II) were taken as having the CD44+/CD24-/EpCAM- phenotype. The MDA-MB-231 breast cancer cell line contained almost double (60.97%) the percentage of CD44+/CD24- cells compared to the MCF-7 breast cancer cell line (32.5%) (Figure 5.C), however the CD44+/CD24- MDA-MB-231 breast cancer cells displayed low to negligible EpCAM expression (0.35%) whereas the MCF-7 breast cancer cells was determined to have a relatively large and statistically significant ($p < 0.05$) CD44+/CD24-/EpCAM+ population (30%, Figure 5.C).

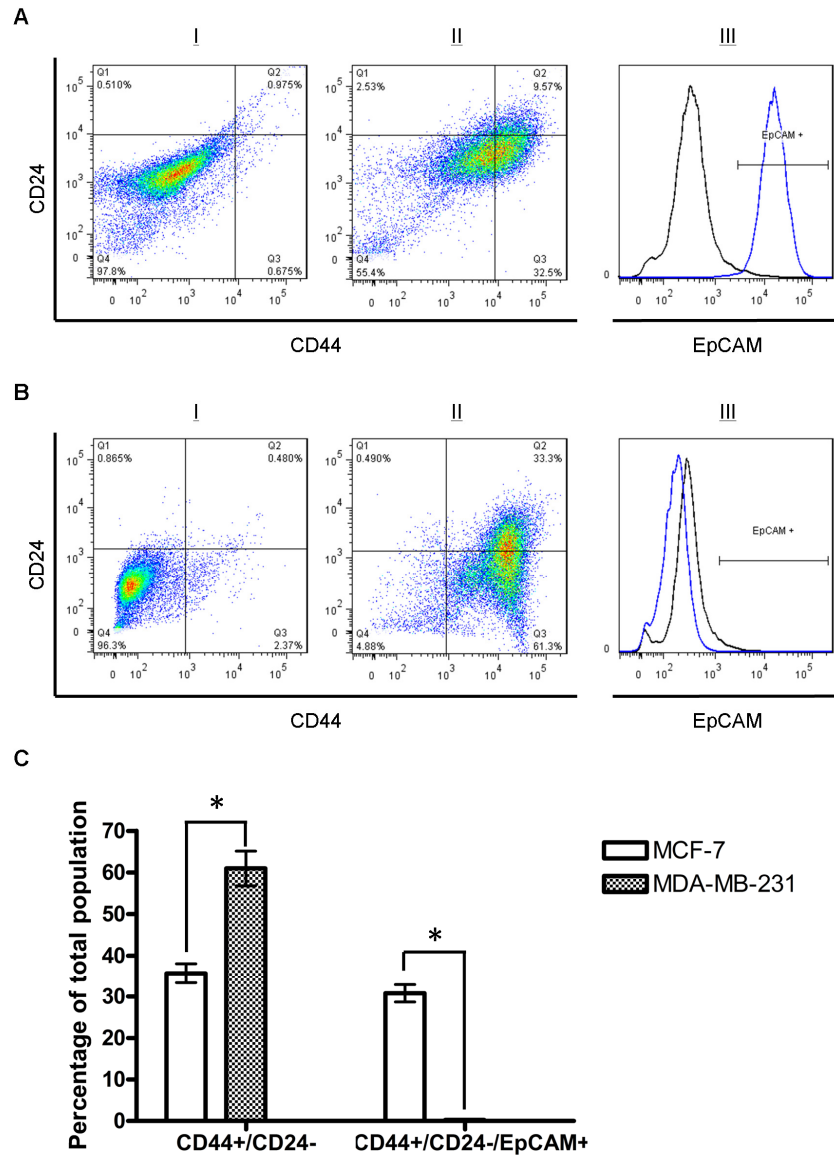


Figure 5| Expression of CD44/CD24 and CD44/CD24/EpCAM on MCF-7 and MDA-MB-231 breast cancer cell lines.

MCF-7 (A) and MDA-MB-231 (B) breast cancer cell lines were stained for the presence of the cancer stem cell markers CD44, CD24 and EpCAM and fluorescence detected by flow cytometry using the FACSaria III (BD Biosciences). Cells were incubated with fluorophore-conjugated antibodies against CD44 (APC), CD24 (FITC) and EpCAM (PE) or the corresponding fluorophore-conjugated isotype controls. In order to determine the CD44/CD24/EpCAM co-expression status of cells, the gating strategy shown in Figure 3 and 4 were used to determine the phenotypic profile of the two breast cancer cell lines. The live cell population was gated using the forward scatter (FSC) and side scatter parameters (SSC) (Figure 3: Panel I) and double colour dot plots representing the expression of CD44 (x-axis) and CD24 (y-axis) on the cell lines generated (panel I and II). For both MCF-7 (A) and MDA-MB-231 (B), panel I represents the isotype control staining profiles; while panel II represents the corresponding fluorophore-conjugated antibody staining. For the two colour dot plot, quadrant gating was applied based on the isotype control staining profile (Figure 3: Panel III and IV) and identified four distinct phenotypes (Q1: CD44⁺/CD24⁺; Q2: CD44⁺/CD24⁻; Q3: CD44⁻/CD24⁺; Q4: CD44⁻/CD24⁻) where the numbers represent the proportion of cells within each particular gate. Panel III shows a single colour histogram representing the expression of EpCAM (x axis) on the CD44⁺/CD24⁻ (Q3) subpopulation (number of cells represented on the y axis), where the black line indicates the isotype control staining and the blue line indicates the EpCAM antibody staining. For the dot plots (panel I and II) and histogram (panel III), the isotype controls were used to define the negative and positive gates as described in Figure 3 and 4. Data shown are representative of triplicate experiments. The proportion of CD44⁺/CD24⁻ and CD44⁺/CD24⁻/EpCAM⁺ cells in the MCF-7 and MDA-MB-231 cell lines was quantified using this approach in triplicate experiments and the average percentage of the different populations in the two cell lines is shown in the graph (C). * indicates statistical significance of $p < 0.05$.

3.2.2 Enzyme activity-based (ALDEFLUOR) identification of putative BCSCs in MCF-7 and MDA-MB-231 breast cancer cell lines

Putative BCSC populations in the MCF-7 and MDA-MB-231 breast cancer cell lines were investigated using the ALDEFLUOR assay (StemCell Technologies). This assay was based on the overexpression of active aldehyde dehydrogenase (ALDH), with high expression and activity of ALDH1 in particular associated with stem-like cells (Ginestier *et al*, 2007). The principle of the assay centers on the conversion of the ALDEFLUOR reagent (BODIPY-aminoacetaldehyde, BAAA) to the green fluorescent BODIPY-aminoacetate (BAA) which is retained within cells. This conversion can be measured using flow cytometry by excitation of BAA with the 488 nm blue laser and recording the emission in the FITC filter channel. An inhibitor control containing DEAB, a specific ALDH inhibitor, can be used to identify cells displaying high FITC fluorescence and hence high ALDH activity.

The gating strategy for ALDEFLUOR analyses is represented in Figure 6. A two-colour dot plot showing SSC fluorescence on the y-axis and ALDH (FITC) fluorescence on the x-axis was used for ALDEFLUOR analyses (Figure 6.A and B). The unstained sample contained no ALDEFLUOR reagent and hence displayed no ALDH (FITC) fluorescence (Figure 6.A). The inhibitor control, containing 15 μ M of the ALDH specific inhibitor DEAB, was used to establish the ALDH^{Hi} fluorescence threshold gate, which was used to identify events displaying ALDH^{Hi} activity (Figure 6.B: Panel I). The ALDH^{Hi} threshold gate was set such that an ALDH^{Hi} population of 0.1% or less was recorded in the inhibitor control sample (Figure 6.B: Panel II). With the ALDH^{Hi} gates set at 0.1% or less on the DEAB inhibitor control, individual gates were copied onto the corresponding samples without inhibitor in order to determine the proportion ALDH^{Hi} events within the sample (Figure 6.B: Panel III). Comparison of profiles generated with (Figure 7. A and B: Panel I) and without (Figure 7.A and B: Panel II) the ALDH inhibitor (DEAB) allowed for the ALDH^{Hi} population to be identified. The MCF-7 breast cancer cell line displayed a lower proportion of ALDH^{Hi} cells compared to the MDA-MB-231 breast cancer cell line (Figure 7.A and B), with the MDA-MB-231 breast cancer cell line found to display a population of 3.25%, while the MCF-7 breast cancer cell line displayed an ALDH^{Hi} population of 1.8% (Figure 7.C).

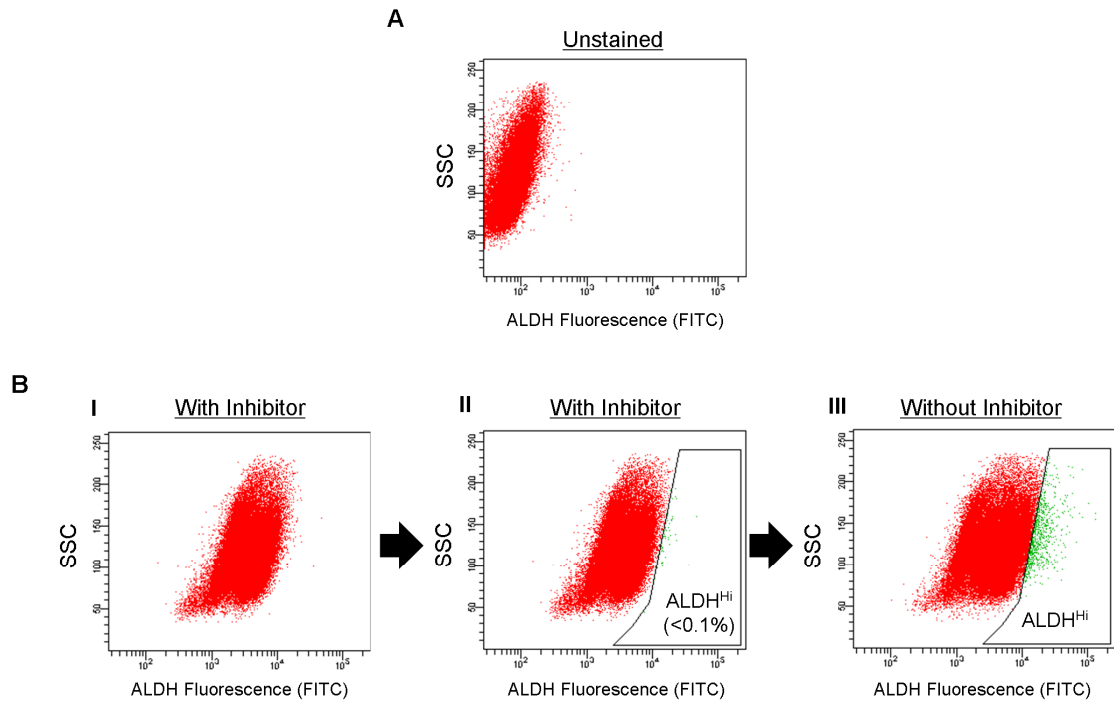


Figure 6| Fluorescence gating strategy for ALDEFLUOR enzyme activity-based BCSC marker analyses.

The ALDEFLUOR assay was visualized using a two-colour dot plot showing side scatter fluorescence (SSC; y-axis) and ALDH fluorescence (ALDH fluorescence detected in the FITC channel; x-axis). (A) Unstained samples (sample without added ALDEFLUOR reagent) showing no ALDH (FITC) fluorescence. (B) The ALDH^{Hi} fluorescence threshold gate was constructed using the DEAB (15 μ M; specific ALDH inhibitor) inhibitor control, with the fluorescence gate set such that the inhibitor control sample recorded an ALDH^{Hi} population of 0.1% or less (Panel I and II). The fluorescence gate was copied onto the corresponding sample lacking inhibitor to identify the ALDH^{Hi} population (Panel III)

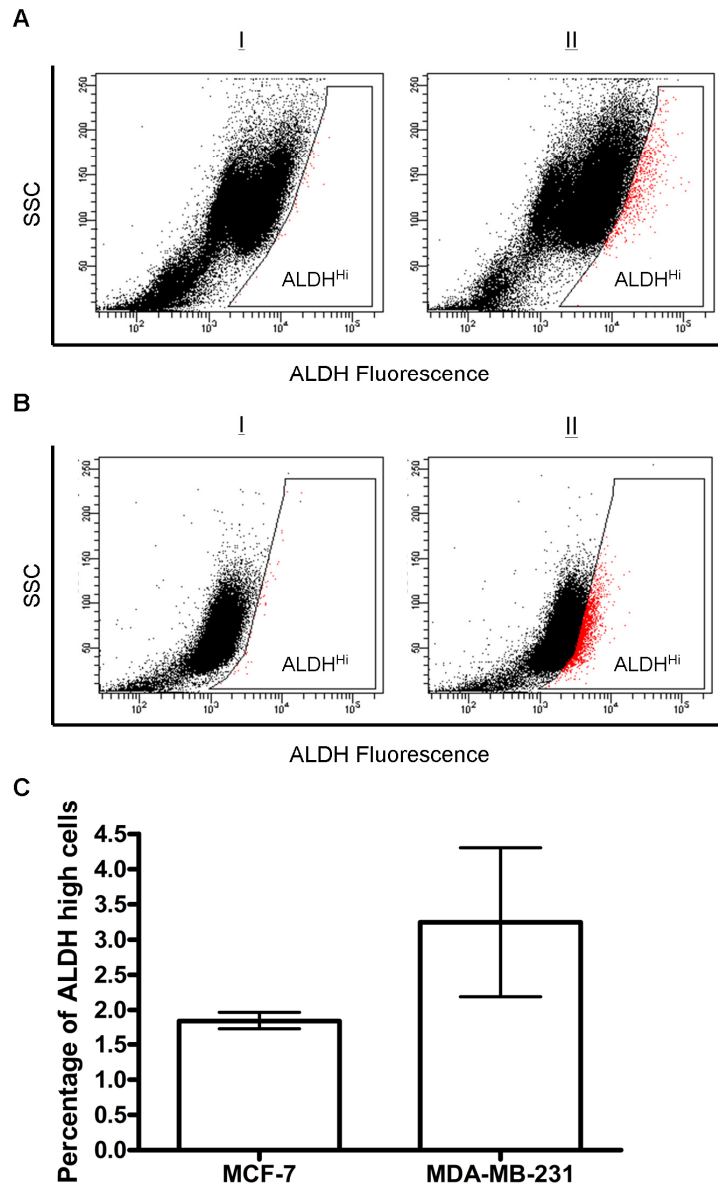


Figure 7| Enzyme activity-based (ALDEFLUOR) detection of putative BCSC populations in MCF-7 and MDA-MB-231 breast cancer cell lines.

Dot plots of side scatter (SSC; y-axis) and ALDEFLUOR fluorescence (ALDH fluorescence detected in the FITC channel; x-axis) showing the detection of aldehyde dehydrogenase (ALDH) activity in MCF-7 (A) and MDA-MB-231 (B) breast cancer cell lines using the ALDEFLUOR assay enzyme activity-based BCSC profile. Breast cancer cell lines were incubated with ALDEFLUOR reagent with (Panel I) and without (Panel II) an inhibitor of ALDH activity (DEAB; 15 μ M). The samples treated with ALDH inhibitor were used to establish the $ALDH^{hi}$ gate with a proportion of staining of 0.1% in the inhibited sample being accepted as background (Figure 6). These gates were then copied onto the corresponding ALDH without inhibitor samples to identify the proportion of cells showing $ALDH^{hi}$ activity (shown in red). The data in (A) and (B) are representative of triplicate experiments (appendix X). The proportion of ALDH high cells in the MCF-7 and MDA-MB-231 cell lines were determined in triplicate experiments and the average percentages are shown in the graph (C).

3.2.3 Anchorage-independent growth of MCF-7 and MDA-MB-231 breast cancer cell lines.

The anchorage-independent assay provided a functional-based *in vitro* technique to enrich for BCSC populations. This technique was considered independent of the expression of a single protein marker or enzyme activity. The principle of the technique was that cells displaying stem-like properties are able to survive and proliferate in serum free conditions within an anchorage-independent environment, while non stem-like cells undergo a programmed cell death associated with loss of adherence (known as anoikis) (Frisch and Francis, 1994). The anchorage-independent assay has been used for the identification and isolation of numerous stem-like cell populations from both normal (Dontu *et al*, 2003) and malignant solid tissues (Ponti *et al*, 2005; Grimshaw *et al*, 2008; Cariatì *et al*, 2008). For the assay, cells are resuspended in a serum-free anchorage-independent medium containing specialized growth factors (Materials and Methods 2.1.3). Cells were seeded into ultra-low attachment plates and grown for 8 days with the medium being replaced every 3 days.

The MCF-7 and MDA-MB-231 breast cancer cell lines were screened for anchorage-independent growth capabilities (Figure 8). Cells from both cell lines were seeded at equivalent numbers as single cells (Day 0) and grown for 8 days. For the MCF-7 breast cancer cell line, discrete colonies were observed after 4 days of culture with the colonies expanding in size and number by day 8 in culture. After 8 days of growth under anchorage independent conditions, the MCF-7 breast cancer cell line showed a greater number of anchorage-independent colonies than the MDA-MB-231 breast cancer cell line (Figure 8.A). Not only were there a greater quantity of anchorage-independent colonies observed in the MCF-7 cell line, but the size of the MCF-7 colonies were greater than that of the MDA-MB-231 colonies (Figure 8.A). A time course study of the growth of the MCF-7 breast cancer cells in anchorage-independent conditions showed the development of the colonies from single cells and continual growth over the 8 day time course (Figure 8.B).

Confocal microscopy was used to perform a morphological comparison of the anchorage-independent MCF-7 colonies and adherent MCF-7 breast cancer cells (Figure 8.C and D). Anchorage-independent

MCF-7 colonies showed a 3-dimensional spherical morphology, with an apparent concentration of cells on the circumference on the sphere, which is consistent with the morphology of mature hollow tumourspheres (Figure 8.C). An outer border was apparent in the anchorage-independent colonies (Figure 8.C), with cells appearing to line the inside of this border. The nuclei of cells making up the anchorage-independent colonies were smaller compared to the nuclei of the adherent cells (Figure 8.C and D). The adherent MCF-7 breast cancer cells displayed flattened epithelial morphologies associated with adherent cells (Figure 8.D).

3.2.4 CSC-like and stem cell-like nature of anchorage-independent MCF-7 colonies.

In order to determine whether anchorage-independent MCF-7 colonies were enriched for BCSCs, the expression of CD44, CD24 and EpCAM cell surface markers and the ALDEFLUOR activity of anchorage-independent MCF-7 colonies were compared to MCF-7 cells grown under adherent conditions (Figure 9). Both anchorage-independent and adherent cells were grown for 8 days, with media replacement at the same time. Samples were separated into single cell solutions and analyzed for the expression of BCSC cell surface markers. The gating strategies for the CD44/CD24 and CD44/CD24/EpCAM staining (Figure 9.A and B) as well as the ALDEFLUOR assay (Figure 9.C and D) were performed as described previously (Figure 3, Figure 4 and Figure 6 respectively). Isotype controls were used to establish the non-specific and positive expression populations (Figure 9.A and B: Panels I and III, black histogram) and DEAB inhibitor controls used to establish the ALDH^{Hi} populations (Figure 7.C and D: Panel I). The experiments were conducted in triplicate with the average proportions of CD44+/CD24- and CD44+/CD24-/EpCAM+ shown in Figure 9.E, and the average proportion of ALDH high cells is shown in Figure 9.F.

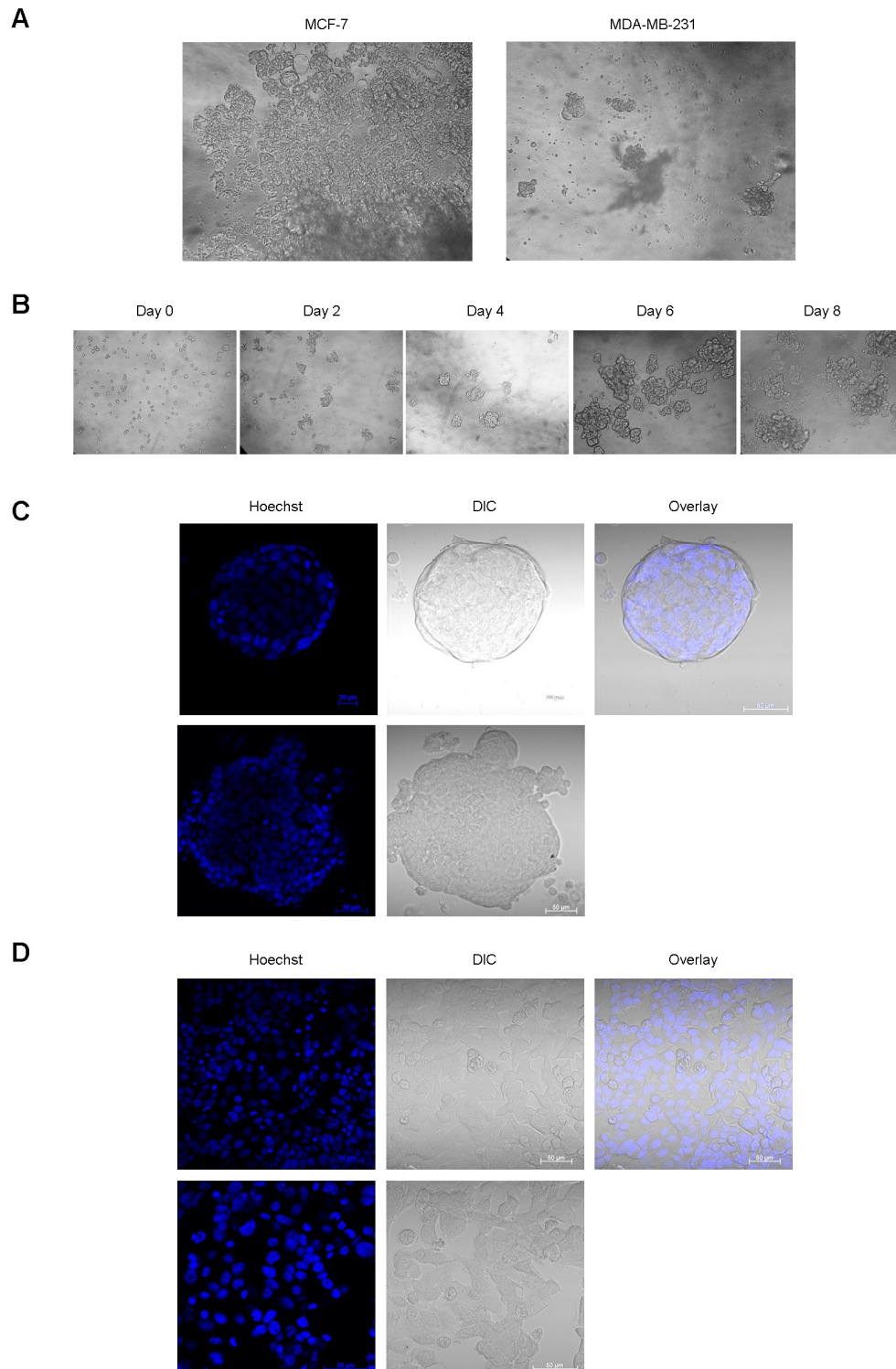


Figure 8| Anchorage-independent growth of breast cancer cell lines.

Analysis of the growth of breast cancer cell lines under anchorage-independent conditions. (A) MCF-7 and MDA-MB-231 breast cancer cells were grown in anchorage-independent conditions using modified serum-free breast tumoursphere media (as described in Materials and Methods) for 8 days after which representative pictures (images captured under 10x objective and 10x eyepiece magnification) were taken of the anchorage-independent colonies. (B) Time course study of anchorage-independent MCF-7 breast cancer cell growth in anchorage-independent conditions was conducted with representative pictures taken on days 0 (day of seeding), 2, 4, 6 and 8 days of seeding. Morphological analysis of anchorage-independent (C) and adherent (D) MCF-7 breast cancer cell cultures. Hoechst 33342 was used to stain nuclei (blue). Images were captured using either the Nikon Coolpix digital camera or the Zeiss LSM510 Meta confocal microscope. Scale bars represent 20 μ m.

Adherent MCF-7 cells displayed a similar CD44+/CD24- profile (approximately 32%) to previous analyses (Figure 5.A [Panel II] and C), although a large change was noted between the current and previous analyses with regards to the CD44+/CD24+ cell populations. In the previous analyses the MCF-7 breast cancer cell line was found to display a minor CD44+/CD24+ cell population (Figure 5.A: Panel II, Q2) while in the current analyses the CD44+/CD24+ population represented a major cell population representing more than half of the original cell line (Figure 9.A: Panel II, Q2). The proportion of EpCAM+ cells in the CD44+/CD24- population from the adherent MCF-7 cells was found to be greater than 90% (Figure 9.A: Panel III, blue histogram) a result consistent with our previous analyses (Figure 5.A: Panel III, blue histogram). The anchorage-independent MCF-7 cells displayed an increased CD44+/CD24- cell population (approximately 56%, Figure 9.B: Panel II, Q3; Figure 9.E) when compared with the adherent MCF-7 cells, and additionally displayed a larger population of CD44+/CD24-/EpCAM+ (approximately 52%, Figure 9.B: Panel III blue histogram; Figure 9.E). The CD44+/CD24- and CD44+/CD24-/EpCAM+ populations were both found to be significantly increased in the anchorage-independent MCF-7 cells (Figure 9.E, $p < 0.05$). The adherent and anchorage-independent MCF-7 cells showed similar ALDH^{Hi} profiles (Figure 9.C and D: Panel II; Figure 9.F). Although the anchorage-independent MCF-7 cells showed a greater population of ALDH^{Hi}, the difference was not determined to be statistically significant.

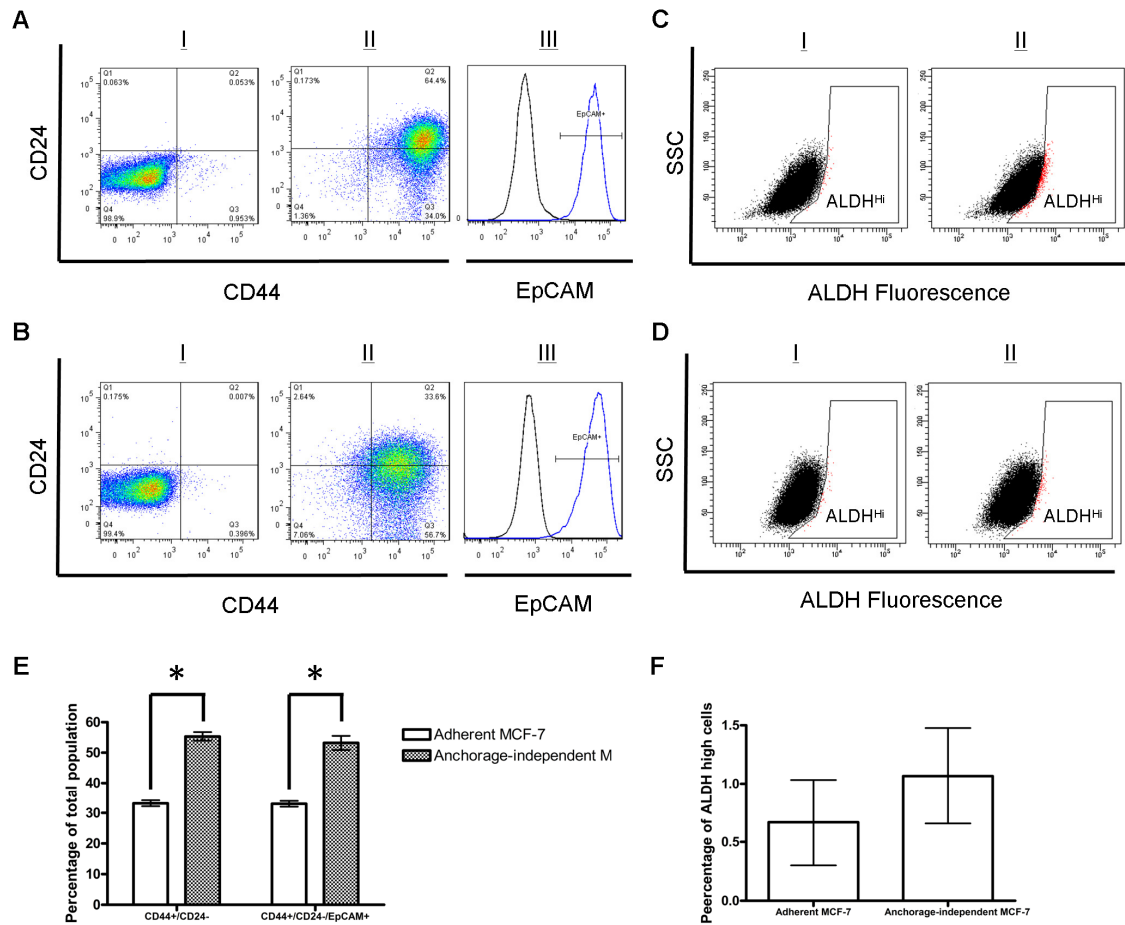


Figure 9| Cancer stem cell-like nature of anchorage-independent MCF-7 breast cancer cells

Comparison of CD44⁺/CD24⁻, CD44⁺/CD24⁻/EpCAM⁺ (A and B) and ALDH^{hi} (C and D) populations in adherent (A and C) and anchorage-independent (B and D) MCF-7 breast cancer cells. After 8 days of growth, adherent (A) and anchorage-independent (B) MCF-7 breast cancer cells were incubated with a combination of directly conjugated antibodies against CD44 (APC), CD24 (FITC) and EpCAM (PE). The y-axis in Panel I and II (A and B) represents CD24 (FITC) fluorescence while the x-axis (A and B: Panel I and II) represents CD44 (APC) fluorescence. The x-axis in Panel III (A and B) represents EpCAM (PE) fluorescence. Isotype controls (A and B: Panels I and III [black histogram]) were used to establish fluorescent threshold gates as described in Figure 3 and 4. For ALEFLUOR analyses of adherent (C) and anchorage-independent (D) MCF-7 cells, two-colour dot plots of side scatter (SSC; y-axis) and ALDH Fluorescence (FITC; x-axis) are shown in Panels I and II. Inhibitor controls (DEAB, 15 μ M; C and D: Panel I) were used to establish ALDH^{hi} threshold gate with gates set as 0.1% or less ALDH^{hi} in inhibitor control (as described in Figure 6) and then copied onto the corresponding samples without the inhibitor (C and D: Panel II). All data (A, B, C and D) are representative of three independent experiments. The proportion of CD44⁺/CD24⁻ and CD44⁺/CD24⁻/EpCAM⁺ (E) and ALDH^{hi} (F) cells in the adherent and anchorage-independent MCF-7 cells were quantified from triplicate experiments and the average percentage of the different populations in the two growth conditions is shown in the graphs. * indicates statistical significance of $p < 0.05$.

3.3 Conclusion

We demonstrated that both the MCF-7 and MDA-MB-231 breast cancer cell lines contained subpopulations displaying the ALDH^{Hi} and CD44+/CD24- BCSC phenotypes, with the MCF-7 cell line additionally showing the presence of a CD44+/CD24-/EpCAM+ BCSC phenotype. The MCF-7 cell line was able to form discrete, three-dimensional tumoursphere colonies when grown under anchorage-independent conditions, while the MDA-MB-231 cell line did not. Comparison of the expression of CSC markers on the normal and tumoursphere-derived MCF-7 cells showed that MCF-7 cells grown under anchorage-independent conditions showed an increase in the proportion of cells expressing the CD44+/CD24- and CD44+/CD24-/EpCAM+ phenotype, which is consistent with the enrichment of CSC populations reported in literature. Although not signifying a BCSC population, there was a noted large difference in CD44+/CD24+ MCF-7 populations between the analysis of the expression of BCSC phenotypes in breast cancer cell lines (Figure 5.A: Panel II, Q2) and the analysis of the cancer stem-like nature of anchorage-independent cells (Figure 9.A: Panel II, Q2). This change in phenotype expression may be associated with the prolonged growth of cell growth *in vitro*. We concluded that the tumoursphere-derived MCF-7 cells were enriched in putative CSC populations. In order to identify chaperones that may be associated with CSC characteristics, we compared the expression of chaperone transcripts in CD44+/CD24-/EpCAM+ enriched tumoursphere-derived cells with those identified in MCF-7 grown under normal adherent conditions using a commercially available quantitative RT-PCR array.

Chapter 4 : Comparison of Molecular Chaperone Expression in Adherent and CD44+/CD24- enriched Anchorage-Independent Populations from the MCF-7 Breast Cancer Cell Line

4.1 Introduction

Quantitative real time PCR (qRT-PCR) is a routinely used technique in cancer research for the identification of potential therapeutic targets based on substantial differential gene expression between non-cancerous cells and cancer cells, as well as the identification of gene expression signatures that correlate with metastatic potential and survival rates (Clarke *et al*, 2001; Han *et al*, 2002; Dooley *et al*, 2003). qRT-PCR is based on basic PCR principles however as the name suggests, and unlike PCR, qRT-PCR analyses product formation after each PCR cycle (Schmittgen and Livak, 2008). Over the PCR procedure, product formation undergoes three phases namely the exponential phase, the linear phase and the plateau phase. During the exponential product formation phase the concentration of the PCR product is increased exponentially. While during the linear and plateau phases PCR reagents become limiting factors for product formation. RT-PCR analysis allows for the determination of the exponential phases in terms of cycle number (cycle numbers through which exponential concentrations of product are produced). The raw measurement provided by the RT-PCR is the cycle at which exponential product formation ends and linear phase begins, which is usually set by a threshold baseline to ensure comparability amongst all the genes analyzed (Yuan *et al*, 2006). The baseline is set as the line of best fit before the accumulation of any detectable fluorescence above the background during the baseline cycles for all samples. The threshold cycle is then calculated by subtracting the baseline from the data generated after each cycle for each sample to ensures that any changes in background fluorescence do not influence the data.

Analysis of RT-PCR data can be achieved using one of two methods, absolute or relative comparisons. Absolute comparison refers to the comparison of the actual concentration of the gene. This technique requires the incorporation of concentration standards in the RT-PCR procedure in order to establish the concentration of the transcript (Yuan *et al*, 2006). Relative comparison analysis allows for the comparison and identification of changes in the levels of gene expression relative a particular gene, usually a housekeeping gene. The gene used for comparative purposes needs to show a constant level of expression between samples (Livak and Schmittgen, 2001; Schmittgen and Livak, 2008). This method compares the relative level of gene expression relative to the state of an unchanged gene.

The molecular chaperone gene expression profiles for CD44+/CD24-/EpCAM+ enriched anchorage-independent MCF-7 and adherent MCF-7 breast cancer cells were analyzed using the commercially available real-time reverse transcription (RT²) PCR Array for Heat Shock Proteins (SA Biosciences) to determine the relative mRNA abundances of chaperone and chaperone-associated genes. Western blot analysis of selected chaperones was performed to compare the abundance of chaperones at the protein level relative to the mRNA expression profiles. Comparisons between the anchorage-independent and adherent MCF-7 breast cancer cells were made in order to identify chaperones, particularly members of the DNAJ family of co-chaperones, whose up-regulation or down-regulation was associated with anchorage-independent growth. Anchorage independent growth was used to isolate populations that were enriched in CSC and therefore any transcripts that were differentially expressed in this population could be used to identify putative CSC-specific chaperone genes.

4.2 Results

4.2.1 Quality control tests and data filtration for RT² PCR Array of anchorage-independent and adherent MCF-7 breast cancer cells.

Relative mRNA abundances of chaperone and chaperone-associated genes were determined using the commercially available Heat Shock Protein RT² PCR Array. The list of genes included in the array is shown in Table 2. A schematic diagram showing the RT² PCR array procedure is shown in Figure 10. MCF-7 breast cancer cells were grown under anchorage-independent or adherent conditions (Chapter 3) for 8 days prior to total RNA extraction using the NucleoSpin RNA/Protein isolation kit (Macherey-Nagel). Extracted RNA (1 µg) was converted to cDNA using the RT² first strand synthesis kit (SA Biosciences). Equal amounts of cDNA were loaded in each well of the 96-well RT² PCR Array plate. Data analysis was performed using the online RT² PCR Array data analysis software provided by the manufacturer (SA Biosciences).

To ensure reliability of the analyses, several internal quality control tests were performed (results represented in Table 3). Quality control tests evaluated the array reproducibility, reverse transcriptase

(RT) efficiency and the potential of genomic DNA contamination. The array reproducibility referred to the consistency of the amount of cDNA loaded into each array plate, as well as the efficiency and reproducibility of the PCR procedure between the replicates and samples. Array data was reported as being reproducible if the positive PCR control (PPC) Ct value of each replicate is approximately 20, with a standard deviation less than 2 between arrays. The RT efficiency tested the efficiency of the RT² First Strand cDNA synthesis RT reaction, and allowed for the detection of inefficient conversion of RNA to cDNA. As the array primers in each well of the PCR array plate are designed for annealing to DNA, the inefficient conversion of RNA to cDNA during the RT² First Strand synthesis step (Figure 10) may negatively affect the mRNA expression analysis by understating the abundance of the transcripts. In order to determine the RT efficiency of the PCR Arrays, the average Ct (average of replicates of a single sample) of the PPC was subtracted from the average Ct (average of replicates of a single sample) of the reverse transcriptase control (RTC), equating to the Δ Ct, for each replicate sample. A Δ Ct value of 5 or less showed that the conversion of RNA to cDNA was efficient. In contrast, a value greater than 5 shows that there was evidence of inefficient reverse transcriptase action and will therefore be reported as failed. As the array primers may also anneal to genomic DNA, the presence of genomic DNA may overstate the abundance of transcripts. Genomic DNA contamination was evaluated using the Genomic DNA Control (GDC). The GDC Ct value was used to determine the presence of genomic DNA contamination, with a Ct of 35 or greater showing the absence of genomic DNA (reported as passing) while a Ct between 30 and 35 suggests possible but not certain genomic DNA contamination and was reported as an inquiry. Although considered acceptable, a GDC Ct ranging between 30 and 35 implied that caution was needed when making conclusions on the data. A GDC Ct below 30 was reported as failing the quality control test as genomic DNA is present at a concentration that will affect the reliability of the PCR array.

Table 2| List of chaperone and chaperone-associated families and genes analyzed on the Heat Shock Protein RT2 PCR Array.

Chaperone Family	Genes
HSP90 Family	HSP90AA1, HSP90AB1, HSP90B1
HSP70 (HSPA) Family	HSPA1A, HSPA1B, HSPS1L, HSPA2, HSPA4, HSPA4L, HSPA5, HSPA8, HSPA9
HSP60 (HSPD) Family	HSPD1
HSP40 (DNAJ) Family	DNAJA1, DNAJA2, DNAJA3, DNAJA4, DNAJB1, DNAJB2, DNAJB5, DNAJB6, DNAJB7, DNAJB8, DNAJB9, DNAJB11, DNAJB12, DNAJB13, DNAJB14, DNAJC1, DNAJC3, DNAJV4, DNAJC5, DNAJC5B, DNAJC5G, DNAJC6, DNAJC7, DNAJC8, DNAJC9, DNAJC10, DNAJC11, DNAJC12, DNAJC13, DNAJC14, DNAJC15, DNAJC16, DNAJC17, DNAJC18, DNAJC19, DNAJC21, SERPINH1
Small HSPs and HSPH	HSPB1, HSPB2, HSPB3, HSPB6, HSPB7, HSPB8, HSPE1, HSPH1
Other Regulators of Protein Folding	ATF6, BAG1, BAG2, BAG3, BAG4, BAG5, ADCK3, CCS, CCT2, CCT3, CCT4, CCT5, CCT6A, CCT6B, CCT7, CRYAA, CRYAB, HSF1, HSF2, HSF4, PDFN1, PDFN2, SIL1, TCP1, TOR1A

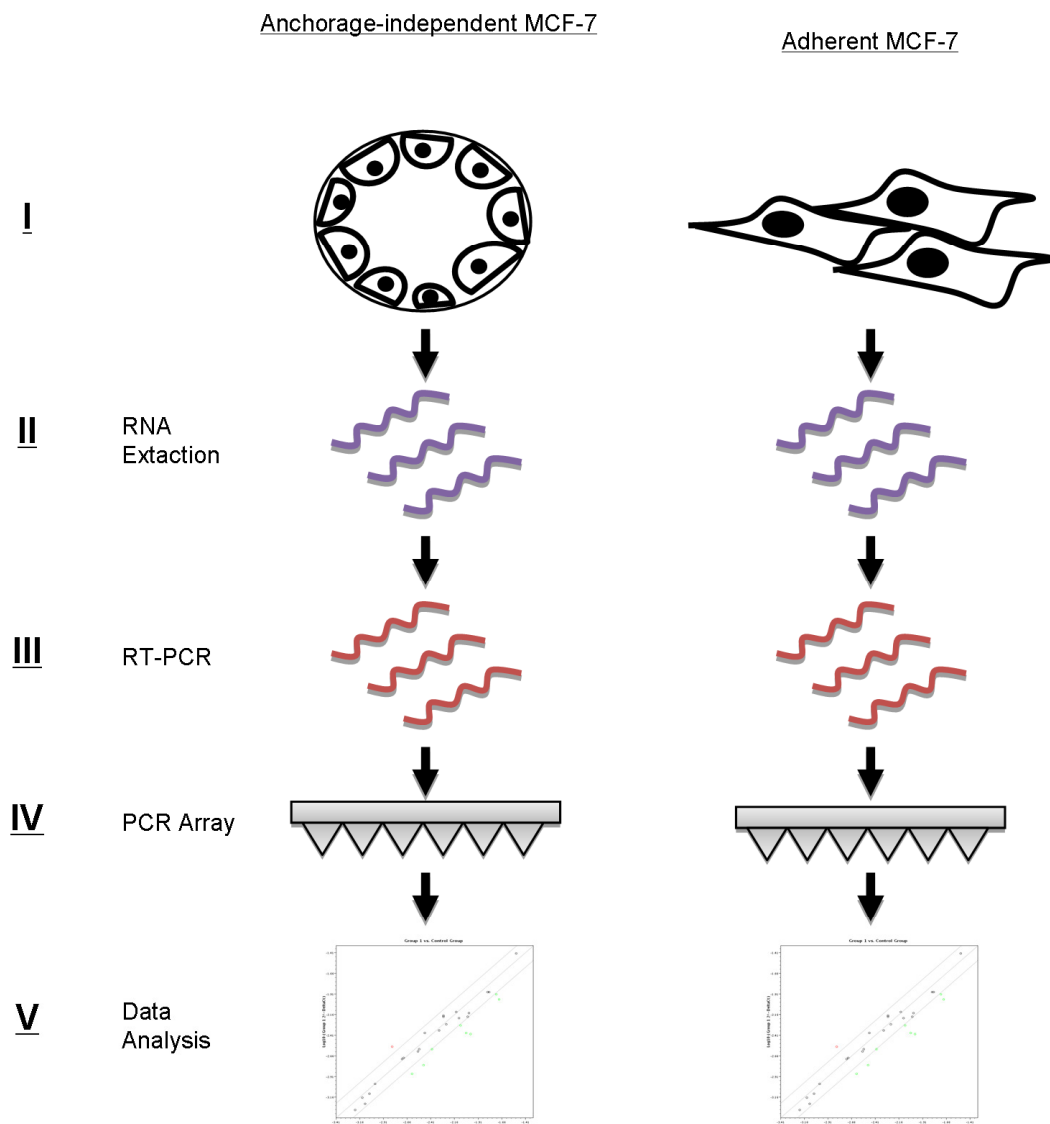


Figure 10| Schematic diagram of the RT² PCR Array procedure for the analysis and identification of chaperones and co-chaperones displaying differential mRNA levels in anchorage-independent and adherent MCF-7 cells.

I) Anchorage-independent and adherent MCF-7 breast cancer cells were grown for 8 days prior to RNA extraction. (II) Total RNA was extracted using the NucleoSpin RNA/Protein kit (Machery-Nagel) and quantified using the NanoDrop2000 (Thermo Scientific). (III) The reverse transcriptase reaction (RT-PCR) was performed using the RT² PCR First Strand synthesis kit (SA Biosciences) to convert 1 µg mRNA into cDNA for analysis. (IV) Equal cDNA was loaded into each well of the 96-well RT² PCR Array plate with thermocycling performed using the iCycler RT-PCR machine (Bio-Rad). (V) Data analyses were performed using online software provided by the manufacturer (SA Biosciences).

Table 3| Results and outcomes of quality control tests for the RT2 PCR Array of anchorage-independent and adherent MCF-7 breast cancer cell line

Quality Control	Results	Report
Reproducibility	Average Ct (PPC) – 20.42±0.16 Average Ct (RTC) – 24.87±0.37	Pass
RT Efficiency	Average Delta Ct (Avg RTC- Avg PPC) – 4.45±0.38	Pass
Genomic DNA Contamination (GDC)	Average Ct (GDC) – 31.45±0.75	Inquiry

In order to achieve statistically significant results, triplicate samples are required. Limitations with regards to number of PCR Array plates and the outsourcing of the qPCR machine from another university (Nelson Mandela Metropolitan University) meant that the analysis of only two replicates of each sample were possible. As such the data analysis has been interpreted cautiously. Nevertheless, the repeat samples were found to display reproducible results, with the average Ct of the PPC being within the 20 ± 2 value range (20.42±0.16, Table 3) therefore considered as passing the reproducibility quality control test. The RT efficiency ΔCt (average Ct^{RTC} – average Ct^{PPC}) was calculated to be 4.45 ± 0.38 (Table 3), below the threshold ΔCt of 5, suggesting the reverse transcriptase reaction was efficient. Although the reproducibility and RT efficiency were reported as passing in the quality control tests, the GDC Ct was found to be below the 35 Ct threshold (31.45±0.75, Table 3). However, the value was within the 30 to 35 Ct range, that suggested that there was a possibility of genomic DNA contamination, a result that may overstate the abundance of certain transcripts, and therefore the data were interpreted with caution. Previous optimization of the PCR Array (*data not shown*) showed that a ten-fold decrease in DNA concentration did not change the Ct value to 35 or greater.

Based on the results of the QC tests, it was determined that the data could be analysed but that caution should be taken when interpreting the data based on the possible genomic DNA contamination. This was particularly important in the case of genes that were shown to be highly expressed, where the abundance may have been overestimated by the presence of contaminating DNA. Melt-curves, analyzing the number of products generated by the PCR, revealed single peaks for each gene suggesting that no non-specific amplification had occurred during the analysis (*data not shown*). All transcripts were categorized based on their average Ct values (average between replicates within a sample) between the two different samples (anchorage-independent versus adherent). If the average Ct of the transcript was below 30 in both test sample (anchorage-independent MCF-7 breast cancer cells) and control sample (adherent MCF-7 breast cancer cells) then the gene was classified as class 'A' for analyses. In cases where the average Ct of a transcript was above 30 in either or both samples, the following classifications were attributed to the gene. A gene was categorized as class 'B' if the average Ct was above 30 in one sample and below 30 in the other sample. These genes showed relatively low expression in one sample and were relatively more highly detected in the other sample, suggesting that the actual fold-change of the gene between the two samples was likely greater than the stated value. However, more replicates would be required to conclusively determine this. Transcripts showing average Ct values greater than 30 in both control and test samples were categorized as class 'C'. These genes showed relatively low expression levels in both samples and as such the fold-change in gene expression between the samples is likely not accurate. Transcripts with Ct values reported as ND (not detected), or if the Ct was above the Ct cut-off of 35, in both samples were categorized as class 'D'. These genes were determined as being undetected therefore the fold-change was not relevant. The categorized genes are represented in Table 4.

Table 4| Class categorization of chaperone and chaperone-associated genes based on differential Ct values between sample and the Ct threshold cut-off.

Chaperone Family	Gene name
<u>Class 'A'</u>	
HSP90 Family	HSP90AB1 , HSP90AB1, HSP90B1
HSP70 (HSPA) Family	HSPA1A, HSPA1B, HSPS1L, HSPA2, HSPA4, HSPA4L, HSPA5, HSPA8, HSPA9
HSP60 (HSPD) Family	HSPD1
HSP40 (DNAJ) Family	DNAJA1, DNAJA2 DNAJA3, DNAJA4, DNAJB1 , DNAJB2, DNAJB5, DNAJB6, DNAJB9, DNAJB11, DNAJB12 , DNAJB13, DNAJB14, DNAJC1, DNAJC3, DNAJV4, DNAJC5, DNAJC5B, DNAJC7, DNAJC8, DNAJC9 , DNAJC10, DNAJC11 , DNAJC12 , DNAJC13 , DNAJC14 , DNAJC15, DNAJC16, DNAJC17, DNAJC18, DNAJC19 , DNAJC21 , SERPINH1
Small HSPs and HSPH	HSPB1 , HSPB8 , HSPE1, HSPH1
Regulators of Protein Folding	ADCK3 , ATF6, BAG1 , BAG2 , BAG3 , BAG4, BAG5 , CCS, CCT2, CCT3, CCT4 , CCT5 , CCT6A , CCT6B , CCT7, HSF1 , HSF2, PDFN2, TCP1 , TOR1A
<u>Class 'B'</u>	
HSP40 (DNAJ) Family	DNAJC6
Regulators of Protein Folding	HSF4, SIL1
<u>Class 'C'</u>	
HSP40 (DNAJ) Family	DNAJB7, DNAJB8, DNAJB13, DNAJC5G
Small HSPs and HSPH	HSPB2 , HSPB3, HSPB6 , HSPB7
Regulators of Protein Folding	CRYAA , CRYAB, PDFN1

Genes highlighted in green represent genes showing down-regulation (fold-regulation of -1.5 or less) in anchorage-independent MCF-7 breast cancer cells. Genes highlighted in red represent genes showing up-regulation (fold-regulation of 1.5 or greater) in anchorage-independent MCF-7 breast cancer cells.

The majority of genes analyzed had a Ct value between 25 – 30 in both samples. This suggested that mRNA expression levels of chaperones and chaperone-associated genes were not highly expressed in either the anchorage-independent and adherent MCF-7 breast cancer cells (Figure 11.A). However, the percentage of chaperone and chaperone-associated genes with a Ct value below 25 (high abundance) was lower in the anchorage-independent MCF-7 breast cancer cells, with 34% of genes displaying a Ct range below 25 in the anchorage-independent sample, compared to 43.23% in the adherent sample (Figure 11.A). The Ct values for the housekeeping genes displayed high reproducibility between replicate samples, with β 2-microglobulin, GAPDH and β -actin showing the highest similarities between the anchorage-independent and adherent MCF-7 breast cancer cells (Figure 11.B). The β -actin gene was selected as the housekeeping gene, and subsequent housekeeping protein, used for comparative analyses and normalization calculations due to the high reproducibility between replicates and high similarity between anchorage-independent and adherent MCF-7 samples. Although RT-PCR mRNA expression data is commonly normalized to the expression of more than one housekeeping gene, we selected a single housekeeping gene for comparative analysis at both the mRNA and protein levels therefore negating the influence of differential abundances of mRNA and corresponding protein levels between the two samples.

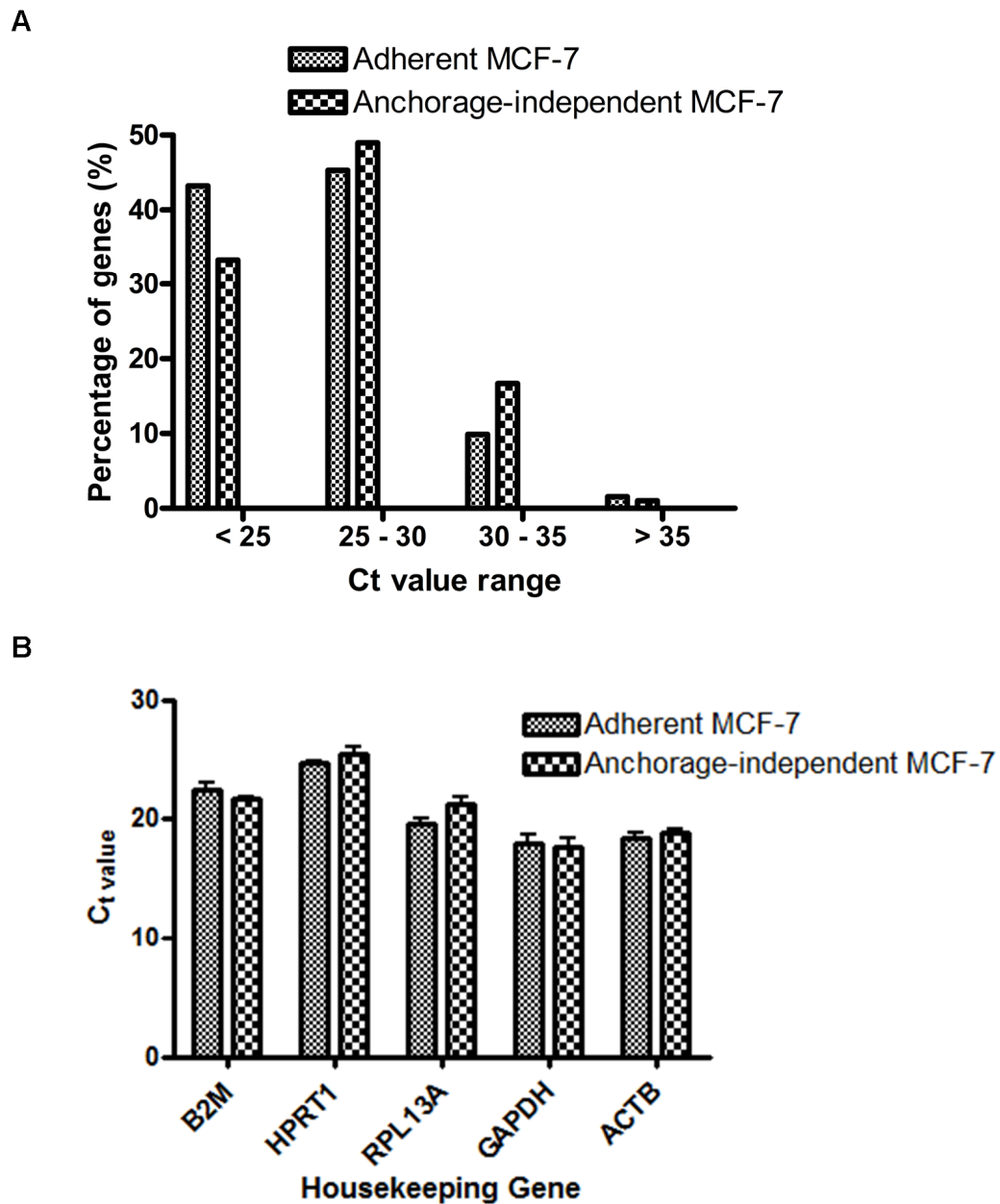


Figure 11| Analysis of gene expression trend of chaperone and chaperone-associated genes between adherent and anchorage-independent MCF-7 breast cancer cells and selection of housekeeping gene for comparative mRNA expression analysis.

(A) Proportion of chaperone and chaperone-associated genes within defined Ct value ranges in adherent and anchorage-independent MCF-7 breast cancer cells. (B) Differential housekeeping gene expression between adherent and anchorage-independent MCF-7 breast cancer cells. Graphs are representative of duplicate experiments.

4.2.2 Differential mRNA expression levels of chaperone and chaperone-associated genes in anchorage-independent and adherent MCF-7 breast cancer cells.

Only class 'A' genes (Table 4) were further analyzed for differences in the levels of mRNA expression between the two samples. For qPCR analysis, the higher the expression level of a given gene, the lower the Ct value. Additionally there is an exponential doubling of the product after each cycle. In order to determine the comparative gene expression profile of each chaperone and chaperone-associated gene, transcripts across the replicates and between samples were first normalized to the housekeeping gene, β -actin, by subtracting the average Ct for β -actin from the Ct value of each transcript, the resulting value known as the Δ Ct ($Ct^{\text{specified gene}} - Ct^{\text{average housekeeping gene}}$). To account for the inverse proportionality and exponential nature of qPCR, the normalized gene expression of a gene was represented by $2^{-\Delta Ct}$. The fold-change in gene expression between the samples were calculated using the $\Delta\Delta$ Ct method where the Δ Ct of genes from the adherent MCF-7 breast cancer cell sample were subtracted from the Δ Ct of the corresponding anchorage-independent MCF-7 breast cancer cell sample ($\Delta\Delta Ct = \Delta Ct^{\text{[gene anchorage-independent MCF-7]}} - \Delta Ct^{\text{[gene adherent MCF-7]}}$). The fold-change in the mRNA expression between anchorage-independent and adherent conditions was therefore represented as $2^{-\Delta\Delta Ct}$. For our analyses a fold-change threshold of ± 1.5 was used to identify genes either up-regulated (a $2^{-\Delta\Delta Ct}$ of 1.5 or greater) or down-regulated ($2^{-\Delta\Delta Ct}$ of -1.5 or less).

Gene expression scatter plots were constructed with the $\text{Log}_{10}(2^{-\Delta Ct})$ of the anchorage-independent MCF-7 transcripts plotted on the y-axis, while the $\text{Log}_{10}(2^{-\Delta Ct})$ of the adherent MCF-7 transcripts plotted on the x-axis. These plots were used to compare the expression profile of the different chaperone and chaperone-associated families in groups (Figure 12). The fold-regulation threshold was set at 1.5. Any genes with a fold-regulation of 1.5 or greater were determined as being up-regulated in the anchorage independent MCF-7 cells (Figure 12.A-D: Red circles) while genes with a fold-regulation of -1.5 or less were determined as being down-regulated in the anchorage independent MCF-7 cells (Figure 12.A-D: Green circles). A large proportion of DNAJ genes were found to be down-regulated in the anchorage independent MCF-7 cells, including DNAJB12, DNAJC13,

DNAJC11, DNAJC9 and DNAJC19 (Figure 12.A: Green circles, and Figure 13.B), with only DNAJC12 found to be up-regulated (Figure 12.A: Red circle, and Figure 13.A). Only HSP90AB1 was found to be differentially expressed between anchorage-independent and adherent MCF-7 breast cancer cells with no HSPA genes showing any detectable difference in expression at the mRNA level (Figure 12.B: Green circle). Interestingly there was an apparent down-regulation of the small HSP family (HSPB family) in the tumoursphere-derived MCF7 sample, with HSPB8 showing the greatest down-regulation (Figure 13.B), as well as HSPH (Figure 12.C: Green circles). The down-regulation profile of chaperone and chaperone-associated genes was further reflected in the group of genes designated “regulators of protein folding” (RPF) genes, where the majority of genes were down-regulated in the MCF-7 cells cultured under anchorage-independent conditions as opposed to adherent conditions (Figure 12.D: Green circles). Only BAG1 was observed to be up-regulated in anchorage-independent MCF-7 breast cancer cells (Figure 12.D: Red circle and Figure 13.A).

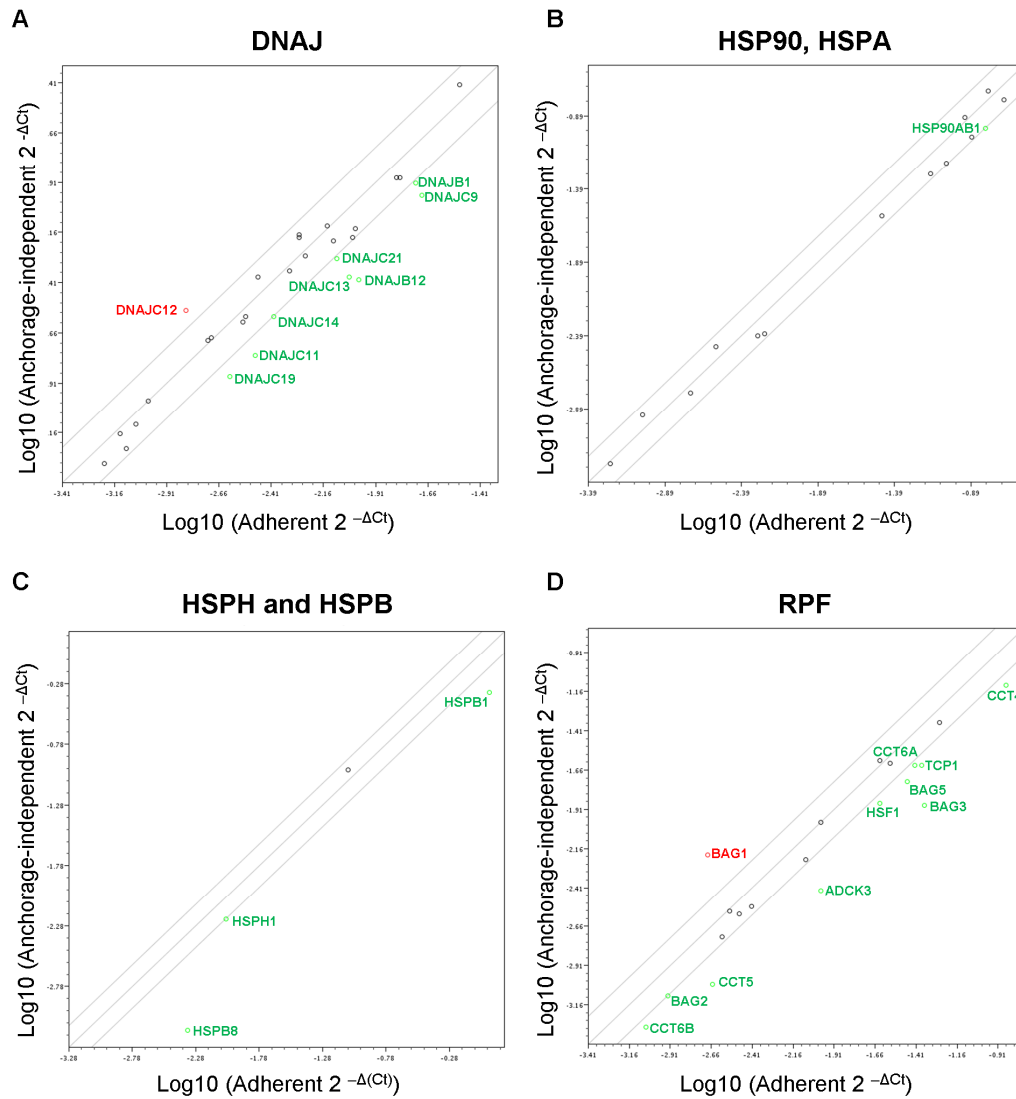


Figure 12| Fold-regulation scatter plots for individual chaperone and chaperone-associated gene families from Class 'A'.

Scatter plots of DNAJs (A), HSP90 and HSAs (B), HSPH and HSPBs (C) and RPF (Regulators of protein folding, D) showing logged β -actin-normalized gene expression levels in anchorage-independent (y-axis) and adherent (x-axis) MCF-7 breast cancer cells. Red circles represent genes showing up-regulation in anchorage-independent cells (fold-regulation of 1.5 or greater). Green circles represent genes showing down-regulation in anchorage-independent cells (fold-regulation of -1.5 or less).

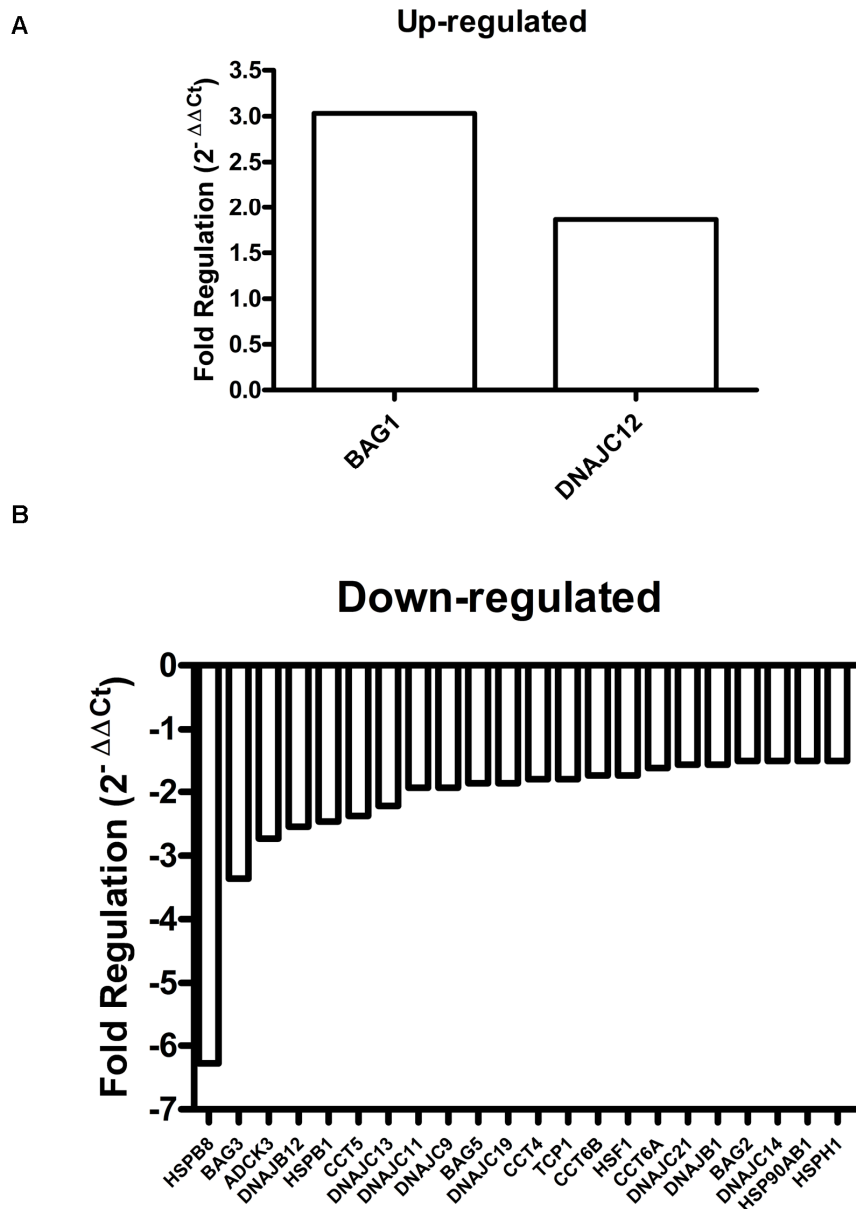


Figure 13| Chaperone and chaperone-associated genes identified as showing differential expression between anchorage-independent and adherent MCF-7 breast cancer cells.

(A) Genes determined as showing up-regulation in mRNA levels in anchorage-independent MCF-7 cells. Up-regulation was determined as a fold-regulation of 1.5 or greater. (B) Genes determined as showing down-regulation in anchorage-independent MCF-7 cells. Down-regulation was determined as a fold-regulation of -1.5 or less).

4.2.3 Differential protein expression levels of selected chaperone and chaperone-associated genes in anchorage-independent and adherent MCF-7 breast cancer cells.

To corroborate the finding of the mRNA expression profile of chaperone and chaperone-associated genes, the protein levels of selected chaperone and chaperone-associated genes were analyzed in anchorage-independent and adherent MCF-7 breast cancer cells. The genes to be analyzed were selected both on availability of commercial antibodies to specific chaperone isoforms and the difference in mRNA levels between the anchorage-independent and adherent samples. Eight chaperone and chaperone-associated genes were selected for protein analyses, including HSP90AA1/HSP90AB1, HSPA1A/HSPA8, DNAJB1, DNAJB6, DNAJC6, HSPB1 and BAG1. HSP90AB1, DNAJB1, HSPB1 and BAG1 showed differential mRNA expression levels between anchorage-independent and adherent MCF-7 breast cancer cells (down-regulation and up-regulation respectively; Figure 12.A, B and C). DNAJC6 was down-regulation in the anchorage-independent MCF-7 breast cancer cells (Table 4), although the significance of this down-regulation could not be determined. HSPA1A/HSPA8 and DNAJB6 were slightly down-regulated in anchorage-independent MCF-7 cells, although this down-regulation was not greater than the 1.5 fold-regulation threshold and therefore these genes were not highlighted in the previous analysis. Single antibodies recognizing both HSP90AA1 and HSP90AB1, and HSPA1A and HSPA8 were used to identify the HSP90 and HSPA genes therefore the protein expression levels of the individual HSP90 and HSPA family genes could not be determined.

Western blot analyses of specific chaperone and chaperone-associated proteins in anchorage dependent and independent MCF-7 cultures are represented in Figure 14. Protein analyses were performed in triplicate from independent paired anchorage-independent and adherent growth experiments. Actin was used as the protein loading control as previous analyses revealed constant Actin expression levels at both the mRNA and protein expression levels, with densitometry performed using ImageJ software to account for total protein loading and to determine the relative protein levels for comparison between growth conditions (Figure 14.A and B). A decrease in the mean

HSP90AA1/HSP90AB1, HSPA1/HSPA8, DNAJB1 and DNAJB6 protein levels was observed in anchorage-independent MCF-7 breast cancer cells, compared to adherent MCF-7 breast cancer cells. DNAJC6 and HSPB1 protein levels remained constant between the anchorage-independent and adherent MCF-7 cells (Figure 14.A and B). All cell lysate samples were quantified prior to analyses with equal concentration loaded into each well. Differential Actin protein levels were observed for the DNAJB1 and DNAJB6 however this was attributed to differential protein concentration loading rather than increased Actin expression. It was therefore necessary to determine the relative values normalized to the loading control Actin. Comparisons with the mRNA fold-regulation showed all of the genes analyzed at protein level were down-regulated at the mRNA level. HSP90AB1, DNAJB1 and HSPB1 were down-regulated by more than 1.5-fold in anchorage-independent MCF-7 breast cancer cells compared to adherent MCF-7 breast cancer cells (Figure 14.C). DNAJC6 additionally displayed down-regulation at the mRNA level, although, as this gene was categorized as class 'B', further analyses would be required to establish an accurate mRNA expression profile of the gene. Several attempts to determine the differential protein expression level of BAG1, identified as being the greatest up-regulated gene between anchorage-independent and adherent MCF-7 breast cancer cells (Figure 13.A), between anchorage-independent and adherent MCF-7 breast cancer cells were made but we were unable to visualize a band corresponding to BAG1 (*data not shown*).

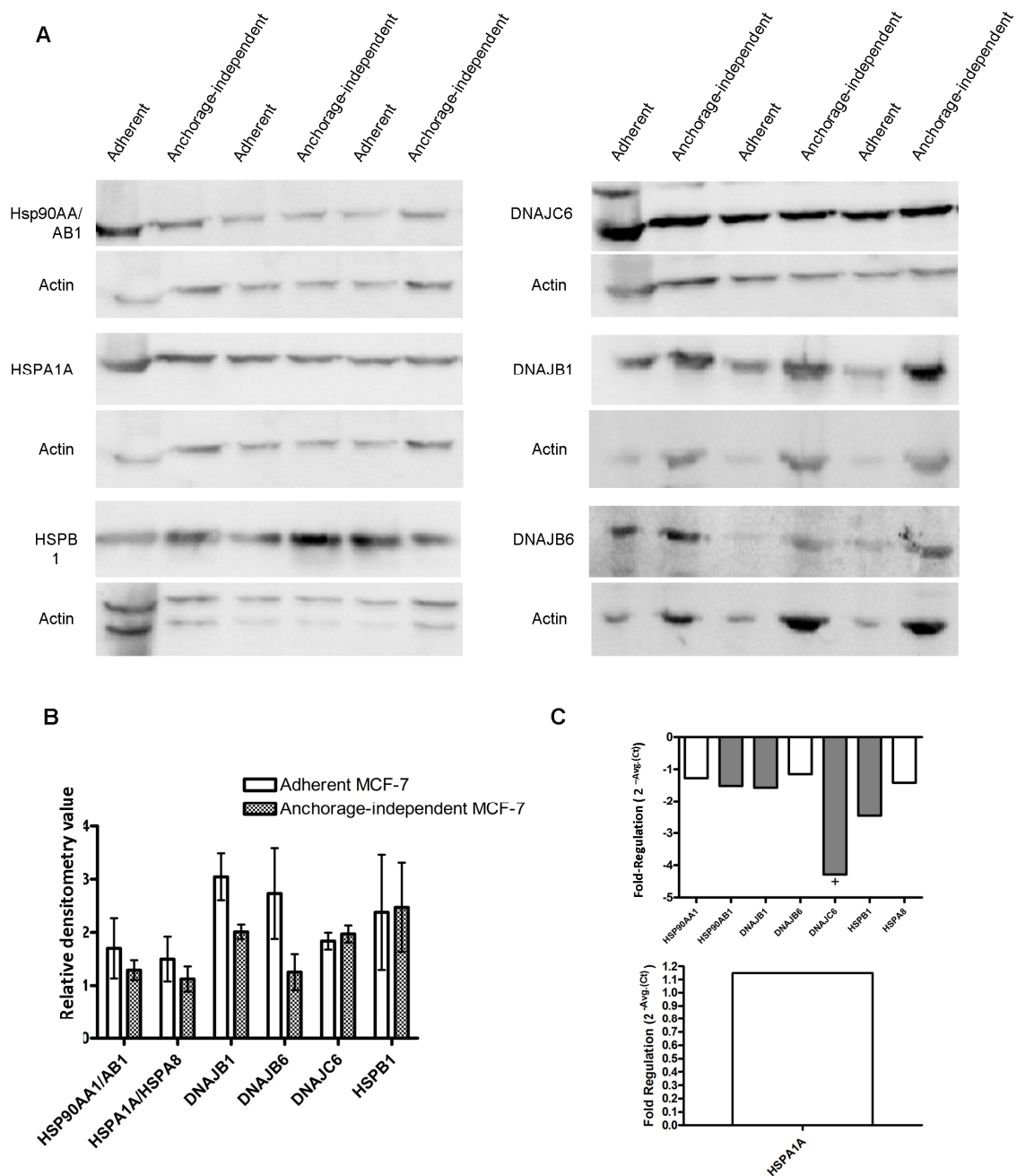


Figure 14| Protein expression levels of selected chaperone and chaperone-associated gene from PCR array.

Western blots (A) and densitometry (B) using Actin as a protein loading control were performed to determine the relative protein abundance of selected chaperone and chaperone-associated genes for validation of mRNA profiles generated by the RT2 PCR array (Figures 11 and 12). Densitometry analyses performed using ImageJ software. (C) Fold-regulation levels of selected chaperone and chaperone-associated genes (Top: selected genes downregulated at mRNA level, Bottom: selected genes up-regulated at mRNA level). Grey bars represent genes showing high degree of differential fold-regulation. '+' represent gene categorized as Class B (Table 3).

4.3 Conclusion

The general decrease in the majority of chaperone and chaperone-associated genes was unexpected in the context of cancer cells which are often thought to overexpress many chaperones (Calderwood *et al*, 2006). However not all chaperones are stress inducible with many being constitutively expressed in cells. Therefore differences in expression profiles may be due to the activity of specific chaperone pathways (encompassing specific chaperones) or activity of specific chaperones acting independent of canonical chaperone pathways. We identified DNAJB6 has having decreased expression at both the mRNA and protein levels in anchorage-independent MCF-7 cells. Previously DNAJB6 expression has been shown to negatively affect cancer properties including migration and invasion possibly through regulation of the Wnt/ β -catenin pathway. The regulatory effect of DNAJB6 on the Wnt/ β -catenin pathway (primarily due to the regulation of β -catenin levels) additionally influences the epithelial-to-mesenchymal transition (EMT), a developmental process associated with the generation of stem-like cells (Mani *et al*, 2008). We therefore selected DNAJB6 for further analysis.

Chapter 5 : Development of a Tool to Study the Role of DNAJB6 in CSC Populations

5.1 Introduction

DNAJB6 is a type II DNAJ containing a J-domain and G/F rich region. The gene is one of several DNAJ genes that have been found to encode for more than one isoform. The DNAJB6 isoforms differ by a C-terminal truncation resulting in the absence of a nuclear localization signal (NLS) in DNAJB6S (short isoform), a signaling peptide present in DNAJB6L (long isoform). As such DNAJB6S displayed a cytoplasmic localization, while DNAJB6L was located in both the cytoplasm and nucleus (Mitra *et al*, 2008; Zhang *et al*, 2008). DNAJB6S appeared to be constitutively expressed in cancer cells, while DNAJB6L was expressed only in certain cell lines; the expression of DNAJB6L being negatively correlated with aggressive properties of the breast cancer in question (Mitra *et al*, 2008). The tumour suppressing properties of DNAJB6L included the inhibition of migration, invasion, soft-agar colony formation and tumorigenesis, which has been attributed to the regulatory role played by DNAJB6L in the stem cell and CSC associated Wnt/ β -catenin signaling pathway (Reya and Clevers, 2005; Mitra *et al*, 2010; Reya *et al*, 2003; Malanchi *et al*, 2008; Vermeulen *et al*, 2010). The regulatory role of DNAJB6L, although not fully understood, is thought to arise from DNAJB6L expression being associated with DKK1 expression. DKK1 is an inhibitor of the Wnt/ β -catenin signaling pathway that functions by blocking the LRP5/6 Wnt receptor, resulting in the phosphorylation and degradation of β -catenin leading to inhibition of the Wnt/ β -catenin pathway (Reya and Clevers, 2005). The inhibition of the Wnt/ β -catenin signaling pathway, through DNAJB6L overexpression, resulted in the loss of mesenchymal cell markers and the acquisition of epithelial cell markers including Keratin 18, a known DNAJB6L client. Therefore DNAJB6L resulted in a reversion of epithelial to mesenchymal (EMT) progression associated with the generation of cells with stem-like properties (Mani *et al*, 2008) and the CD44+/CD24- BCSC phenotype (Morel *et al*, 2008).

The J-domain of DNAJ contains a highly conserved Histidine-Proline-Aspartic acid (HPD) motif that is involved with both the interaction and stimulation of the ATPase domain of HSP70 chaperones (Michels *et al*, 1999). Mutations within this highly conserved motif abrogate HSP70 interaction and stimulation therefore preventing classical DNAJ functioning. There is evidence suggesting that

DNAJs, especially type III DNAJs containing a varied arrangement and presence of non-classical DNAJ domains, have functions additional to and in some cases irrespective of HSP70 stimulation and J-domain functionality. Several DNAJ including both type III and type II have demonstrated and highlighted the potential of J-domain independent functioning, with none more so than DNAJC17. DNAJC17, absence of which is lethal to cells (Sahi *et al*, 2007), plays a role in mRNA splicing events and was shown to function independently of its J-domain with the J-domain only being necessary when interactions with partner proteins are compromised (Sahi *et al*, 2010). Similarly DNAJB8 and the smaller isoform of DNAJB6 (DNAJB6S), with both DNAJB8 and DNAJB6 genes encoding for similar domain architecture, are both capable of aggregation suppression that is independent of J-domain activity (Hageman *et al*, 2010). Recently, our hypothesis of the J-domain dependent tumour suppressor activity of DNAJB6L was confirmed by Mitra and colleagues (2012), whereby J-domain deletion mutants prevented the tumour suppressor functions of DNAJB6L. Additionally DNAJB6L was found carry out its tumour suppressing functions through interaction with HSPA8 (Mitra *et al*, 2012). However the J-domain dependent effect on CSC properties remains to be determined.

In this chapter we describe the development of mammalian expression vectors for both isoforms of DNAJB6 (DNAJB6L and DNAJB6S) which can be used to study the effect of overexpression of DNAJB6 on CSC characteristics of MCF-7 cells. In order to determine whether the effect of DNAJB6 on CSC properties of MCF-7 cells is dependent or independent of J-domain function, we developed DNAJB6 mutants (DNAJB6L[H31Q] and DNAJB6S[H31Q]) that would lack the ability to stimulate the ATPase activity of its partner HSPA. Unlike the mutants described by Mitra and colleagues (2012), the mutants produced in our study (DNAJB6L[H31Q] and DNAJB6S[H31Q]) have intact J-domain but contain a critical mutation abrogating the ability to stimulate the ATPase domain of HSPA proteins.

5.2 Results

5.2.1 Sequence analysis of DNAJB6 isoforms DNAJB6L and DNAJB6S

The DNAJB6 gene encodes two isoforms, namely DNAJB6L (large isoform) and DNAJB6S (small isoform). DNAJB6L is comprised of 326 amino acids and is approximately 85 amino acids larger than the DNAJB6S isoform (241 amino acids; Figure 15.A). The two isoforms show high sequence identity (95%), only differing in sequence in the C-terminal (Figure 15.A).

The presence of different protein domains in the two isoforms was analyzed based on sequence homology using SMART and PROSITE protein domain search engines (Schultz *et al*, 1998; Sigrist *et al*, 2010; Letunic *et al*, 2012). DNAJB6L and DNAJB6S were both found to contain a J-domain (the defining domain of DNAJ proteins) and a Glycine/Phenylalanine rich region (G/F; Figure 15.B), both domains being characteristic of Type II DNAJ proteins (Kampinga *et al*, 2009). Additionally the presence of histone deacetylase (HDAC) binding domain and an acetyltable lysine has been described (Hageman *et al*, 2010; Kampinga *et al*, 2010). A nuclear localization signal (NLS) was found in the C-terminus of DNAJB6L (Figure 15.B) but was absent from the DNAJB6S isoform due to the C-terminal truncation in the smaller isoform (Figure 15.B). As the structures of the DNAJB6 isoforms have not been resolved, homology modeling using SwissModel was used to predict the structure of the J-domain of the DNAJB6 isoforms. SwissModel identified the J-domain of DNAJB3 as displaying the highest J-domain structure similarity (83.3% identical) to the isoforms therefore this was used as the template to model the J-domain of DNAJB6L and DNAJB6S, sharing 100% sequence identity. Figure 16 depicts an overlay of the unrefined homology model of DNAJB6 J-domain (Figure 16; black) produced using SwissMol and the overlay of this model with the model of the DNAJB3 J-domain using PyMOL (Figure 16; red).

A

```

DNAJB6L      MVDYYEVLGVQRHASPEDIKKAYRKLALKWHPDKNPENKEEAERKFQVAEAYEVLSDAK 60
DNAJB6S      MVDYYEVLGVQRHASPEDIKKAYRKLALKWHPDKNPENKEEAERKFQVAEAYEVLSDAK 60
*****

DNAJB6L      KRDIYDKYGKEGLNGGGGGSHFDSPFEFGFTFRNPDDVFREFFGGRDPFSFDFFEDPFE 120
DNAJB6S      KRDIYDKYGKEGLNGGGGGSHFDSPFEFGFTFRNPDDVFREFFGGRDPFSFDFFEDPFE 120
*****

DNAJB6L      DFFGNRRGPRGSRSGTGSFFSAFSGFPSFGSGFSSFDTGFTSFGSLGHGGLTSFSSTSF 180
DNAJB6S      DFFGNRRGPRGSRSGTGSFFSAFSGFPSFGSGFSSFDTGFTSFGSLGHGGLTSFSSTSF 180
*****

DNAJB6L      GGSGMGNFKSIISTKMNVRKITTKRIVENQERVEVEEDGQLKSLTINGVADDDALAE 240
DNAJB6S      GGSGMGNFKSIISTKMNVRKITTKRIVENQERVEVEEDGQLKSLTING-----K 232
*****

DNAJB6L      ERMRRGQNALPAQPAGLRPPKPPRPASLLRHAPHCLSEEEGEQDRPRAPGPWDPLASAAG 300
DNAJB6S      EQLLRLDNK----- 241
*.: * :*

DNAJB6L      LKEGGKRKKQKQREESKKKSTKGNH 326
DNAJB6S      -----

```



Figure 15| Bioinformatic analysis of DNAJB6L and DNAJB6S isoforms.

(A) Alignment of the amino acid sequence of DNAJB6L (NP_490647.1) and DNAJB6S (NP_005485.1) using ClustalW. Stars indicate exact amino acid identity between the two isoforms, while dots indicate amino acids of similar charge or polarity. (B) Presence of predicted domain structures of DNAJB6L and DNAJB6S using SMART and PROSITE domain structure prediction search engines. G/F, Glycine/Phenylalanine rich region; NLS, Nuclear localization signal.

A

DNAJB6L	MVDYYEVLGVQRHASPEDIKKAYRKLALKWHPDKNPENKEEAERKFKQVAEAYEVLSDAK	60
DNAJB6S	MVDYYEVLGVQRHASPEDIKKAYRKLALKWHPDKNPENKEEAERKFKQVAEAYEVLSDAK	60
DNAJB3	MVDYYEVLGVQRHASPEDIKKAYRKLALKWHPDKNPENKEEAERKFKQVAEAYEVLSDAK	60
	*****.* :.*.* *****:*****	
DNAJB6L	KRDIYDKYGKEGLNGGGGGSHFDSPFEGFTFRNPDDVFREFFGGRDPPSFDFEDPFE	120
DNAJB6S	KRDIYDKYGKEGLNGGGGGSHFDSPFEGFTFRNPDDVFREFFGGRDPPSFDFEDPFE	120
DNAJB3	KRDIYDRYGEAGAEGGCTGGRPFEDPFEYVFSFRDPADVREFFGGQDPFSFDLLGNPLE	120
	*****.*: * :.*.* ** *:.*.*: *.*.* *****:*****: :.*.*	
DNAJB6L	DFFGNRRGPRGSRSRGTGSFFSAFSGFPSPFGSGFSSFDTGFTSFGSLGHGGLTSFSSTSF	180
DNAJB6S	DFFGNRRGPRGSRSRGTGSFFSAFSGFPSPFGSGFSSFDTGFTSFGSLGHGGLTSFSSTSF	180
DNAJB3	NILGGSEELLGKQKQSVCTPFLCLQ-----	145
	::*.* . *.:.*.* : * :.*.	
DNAJB6L	GGSGMGNFKSISTSTKMVNGRKITTKRIVENGQERVEVEEDGQLKSLTINGVADDDALAE	240
DNAJB6S	GGSGMGNFKSISTSTKMVNGRKITTKRIVENGQERVEVEEDGQLKSLTING-----K	232
DNAJB3	-----	
DNAJB6L	ERMRRGQNALPAQAGLRPPKPPRPASLLRHAPHCLSEEEGEQDRPRAPGPWDPLASAA	300
DNAJB6S	EQLRLDNLK-----	241
DNAJB3	-----	
DNAJB6L	LKEGGKRKKQKQREESKKKSTKGNH	326
DNAJB6S	-----	
DNAJB3	-----	

B

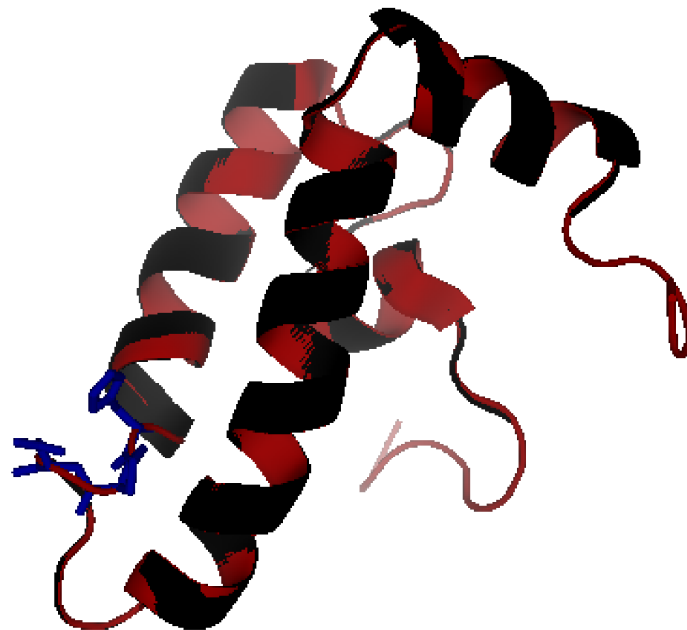


Figure 16| Homology model of DNAJB6 J-domain

(A) Protein primary amino acid sequence alignment of DNAJB6L, DNAJB6S and DNAJB3 using ClustalW. J-domain sequence highlighted in box. (B) Unrefined homology modeling of the J-domains of DNAJB6 (black) using SwissModel. DNAJB3 J-domain (2ej7, red) was used as the template for the homology modeling due to the 83.3% identity with DNAJB6L and DNAJB6S. Models annotated using PyMOL. HPD motif represented in blue. Z-score 1.32.

5.2.2 Generation of pCMV-DNAJB6L and pCMV-DNAJB6S mammalian expression vectors.

The pcDNA-DNAJB6 plasmids were used as the source of the coding regions for the DNAJB6L and DNAJB6S proteins (Hageman *et al*, 2009). Sub-cloning of the respective coding regions into the pCMV backbone was performed to generate the pCMV-DNAJB6L and pCMV-DNAJB6S vectors allowing for the expression of the isoforms in mammalian cells as fusion proteins containing an N terminal c-myc tag. Figure 17.A shows the plasmid maps of the pcDNA plasmids containing the DNAJB6 coding regions and the pCMV-HSJ1a plasmid from which the pCMV backbone was isolated. Figure 17.B shows the results of the restriction digestion analysis of the plasmids. The identity and presence of the appropriate coding regions was confirmed using *Bam*HI and *Xho*I restriction enzymes, either alone or in combination (Figure 17.B). For both the pcDNA-DNAJB6 plasmids, the DNAJB6 coding regions were flanked by unique *Bam*HI and *Xho*I restriction sites (Figure 17.A). The HSJ1a coding region in pCMV-HSJ1a was flanked by *Bam*HI restriction sites (Figure 17.A).

The uncut plasmid samples (Figure 17.B; lanes 2, 7 and 12) showed the presence of the different plasmid conformations (supercoiled, linear and circular). Digestion with either *Bam*HI (Figure 17.B; lanes 3 and 8) or *Xho*I (Figure 17.B; lanes 4 and 9) alone linearized the pcDNA-DNAJB6L and pcDNA-DNAJB6S plasmids, resulting in a single band that migrated to the approximate molecular weight of the pcDNA-DNAJB6L and pcDNA-DNAJB6S plasmids (approximately 6622 and 6169 bp respectively). *Bam*HI digestion alone linearized the pCMV vector (4300 bp) and released a smaller DNA fragment corresponding to the expected size of the HSJ1a coding region (approximately 833 bp, Figure 17.B; lane 13). The digestion of pCMV-HSJ1a with *Xho*I produced a single band corresponding to the linear size of pCMV-HSJ1a (Figure 17.B; lane 14, 5133 bp fragment). The double enzyme digestion using both *Bam*HI and *Xho*I generated two bands for all three plasmids tested, the lower molecular weight bands for each plasmid corresponded to the expected size of the coding regions, namely 1020 bp for DNAJB6L (Figure 17.B; lane 5), 725 bp for DNAJB6S (Figure 17.B; lane 10) and 833 bp for HSJ1a (Figure 17.B; lane 15).

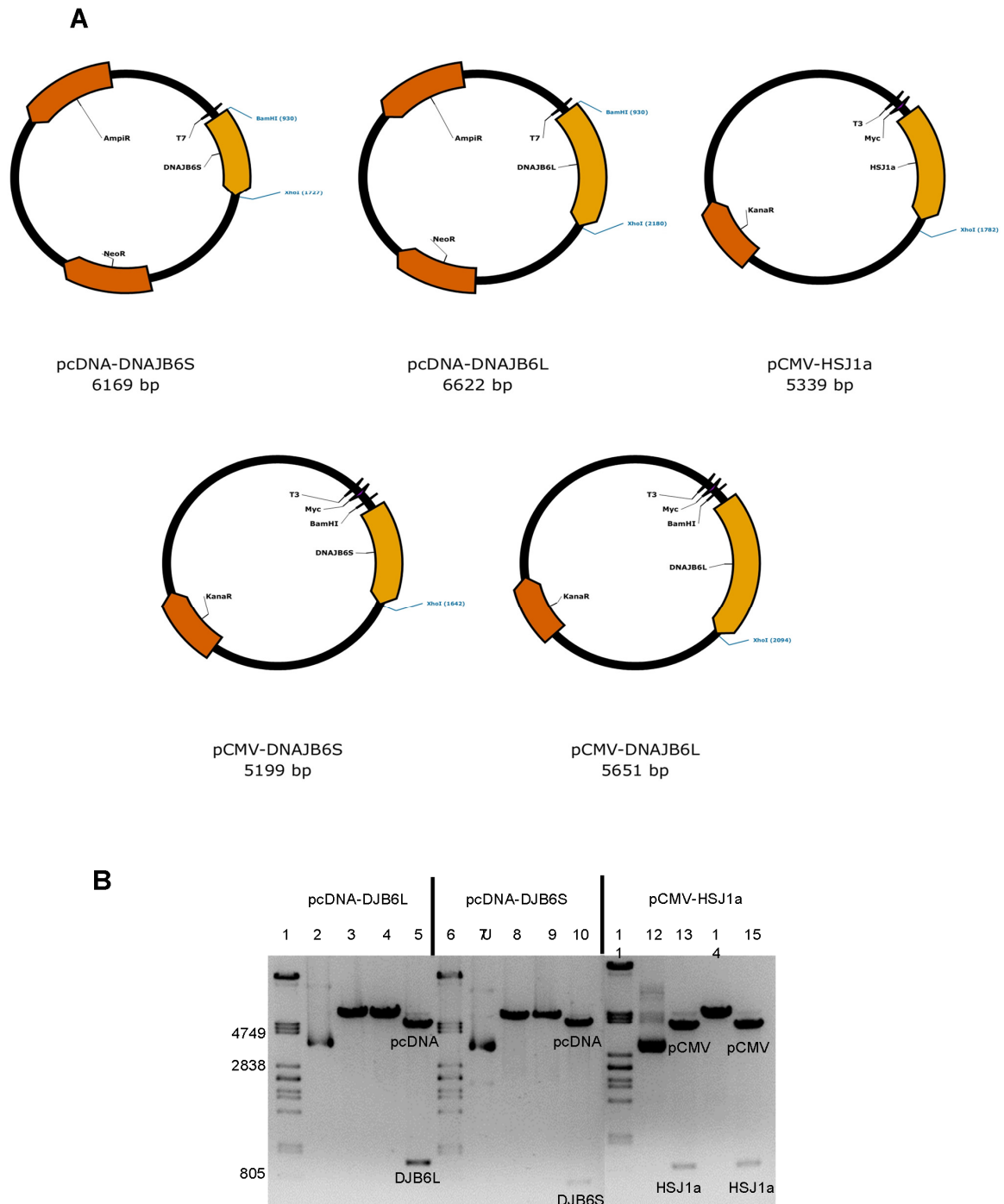


Figure 17| Subcloning of DNAJB6L and DNAJB6S into pCMV mammalian expression vector.

(A) Plasmid maps of pcDNA-DNAJB6L, pcDNA-DNAJB6S and pCMV-HSJ1a showing plasmid sizes, restriction enzyme recognition sites used for subcloning and size of DNAJB6 coding regions. (B) Restriction digestion identification of pcDNA-DNAJB6 plasmids and pCMV-HSJ1a and confirmation of the presence of DNAJB6 coding regions. λ *Pst* molecular weight marker: lanes 1, 6 and 11. Uncut DNA: lanes 2, 7, 12. *Bam*HI digestion: lanes 3, 8, 13. *Xho*I digestion: lanes 4, 9, 14. *Bam*HI and *Xho*I digestion: lanes 5, 10, 15.

The DNAJB6L and DNAJB6S coding regions and the pCMV backbone generated by *Bam*HI and *Xho*I sequential restriction enzyme digest (Figure 17.B; lanes 5, 10 and 15) were isolated and used in a ligation reaction to generate the pCMV-DNAJB6L and pCMV-DNAJB6S plasmids (Figure 17.A and 18). The DNAJB6 coding regions were ligated into the pCMV backbone with correct orientation of insertion provided by complimentary 5' and 3' overhangs created by the *Bam*HI and *Xho*I restriction digestion (Figure 17.A). Ligation reactions were transformed into DH5 α *E.coli* cells and transformants were screened for the presence of the DNAJB6 coding regions using *Bam*HI and *Xho*I single and double digestions (Figure 18). All four of the selected DNAJB6L transformants were found to contain DNA fragments corresponding to the expected size of the DNAJB6L coding region (1020 bp) and the pCMV vector backbone (4300 bp) (Figure 18.A). Both the *Bam*HI and *Xho*I single enzyme restriction digestions of the selected transformants produced single bands representing the expected linear size of the pCMV-DNAJB6L plasmid (approximately 5133 bp, Figure 18.A; *Bam*HI: lanes 3, 7, 11, 15; *Xho*I: lanes 4, 8, 12 and 16). The single band generated by the *Bam*HI and *Xho*I single enzyme digests for the pCMV-DNAJB6L transformants were larger than that of the pCMV backbone generated from the digestion of the pCMV-HSJ1a control and corresponded to the predicted molecular weight of the pCMV-DNAJB6L plasmid (approximately 5133 bp, Figure 18.A; lanes 3, 7, 11 and 15 versus lane 23 [*Bam*HI], and lanes 4, 8, 12 and 16 versus lane 24 [*Xho*I]). The digestion of the selected transformants using *Bam*HI and *Xho*I in combination produced two bands (Figure 18.A; lanes 5, 9, 13 and 17). The larger band (Figure 18.A; lanes 5, 9, 13 and 17) corresponded to the size of the pCMV backbone (4300 bp) and coincided with the pCMV band in the pCMV-HSJ1a *Bam*HI single enzyme (Figure 18.A; lane 23) and *Bam*HI and *Xho*I double enzyme digests (Figure 18.A; lane 25). The lower band from the double enzyme digests of the selected transformants (Figure 18.A; lanes 5, 9, 13 and 17) coincided with the band representing DNAJB6L present in the double enzyme digest from the pcDNA-DNAJB6L parental control plasmid (Figure 18.A; lane 21).

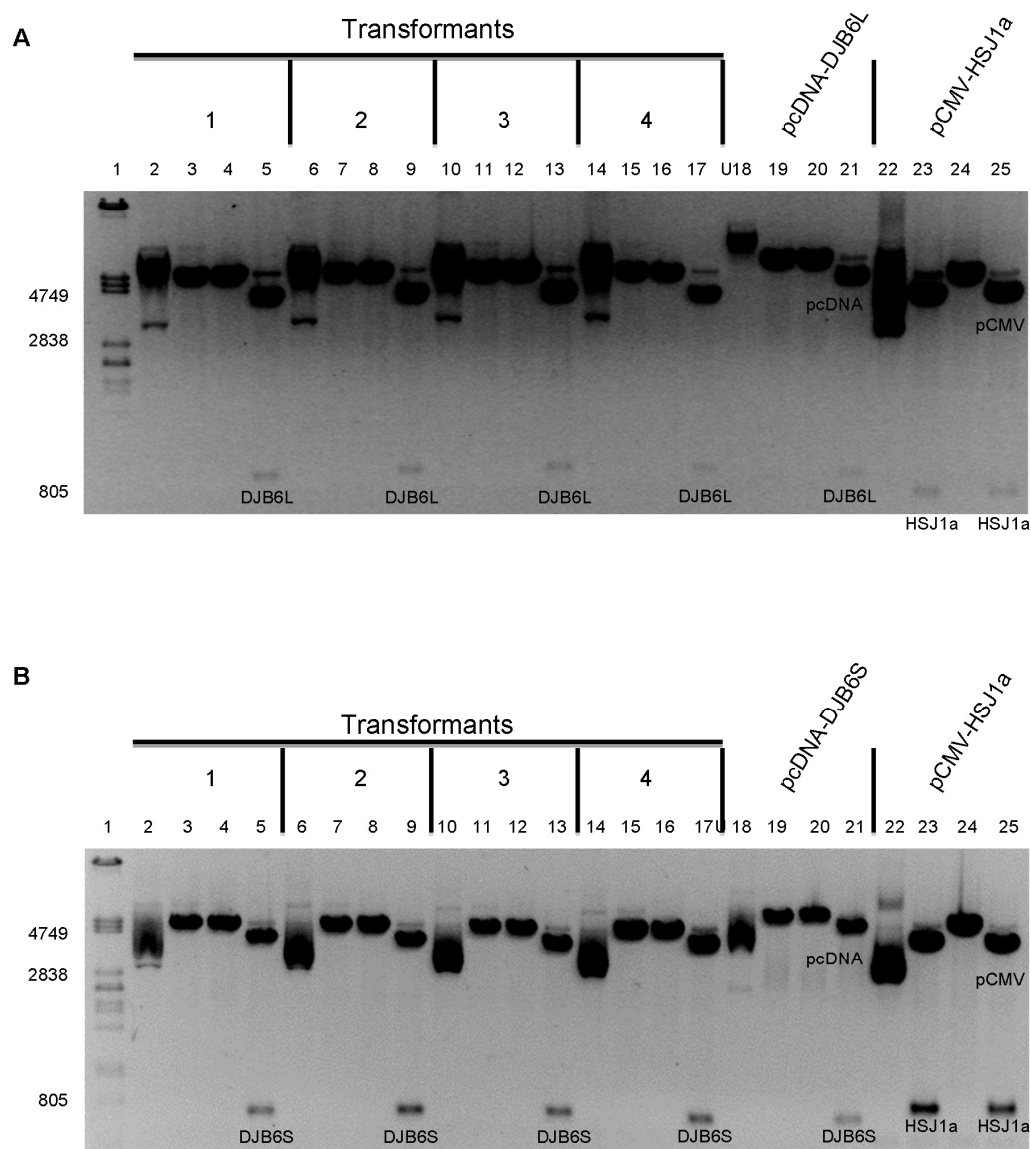


Figure 18| Subcloning of DNAJB6L and DNAJB6S coding regions into pCMV generating pCMV-DNAJB6L and pCMV-DNAJB6S plasmids.

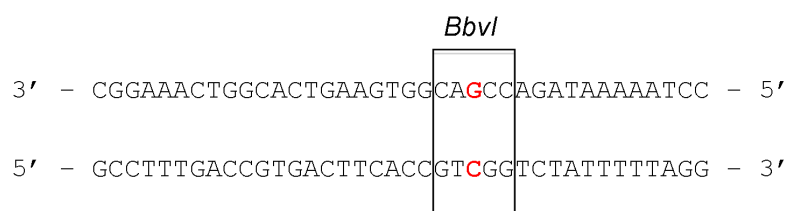
DNAJB6 coding regions and pCMV backbone generated from *Bam*HI and *Xho*I double enzyme digestions were ligated to generate pCMV-DNAJB6L and pCMV-DNAJB6S clones. Transformants of pCMV-DNAJB6L (A) and pCMV-DNAJB6S (B) were screened using restriction digestion with *Bam*HI and *Xho*I in order to identify transformants showing positive pCMV-DNAJB6L ligations. λ *Pst* molecular weight marker: lane 1. Uncut: lanes 2, 6, 10 and 14. *Bam*HI digestion: lanes 3, 7, 11 and 15. *Xho*I digestion: lanes 4, 8, 12 and 16. *Bam*HI and *Xho*I digestion: 5, 9, 13 and 17.

For the pCMV-DNAJB6S plasmid, four transformants were selected and plasmids extracted before restriction digestion to identify the presence of the DNAJB6S coding region. For the single enzyme digestions of the transformants, single bands coinciding with the pCMV band from the *Bam*HI (Figure 18.B; lane 23) and double enzyme *Bam*HI and *Xho*I digestion (Figure 18.B; lane 25) were observed (Figure 18.B; *Bam*HI: lanes 3, 7, 11, 15; *Xho*I: lanes 4, 8, 12 and 16). Additionally double enzyme digestion using *Bam*HI and *Xho*I (Figure 18.B; lanes 5, 9, 13 and 17) resulted in the generation of two bands, the larger band corresponding to the pCMV backbone (approximately 4300 bp, Figure 18.B; lane 25) while the smaller band corresponded to DNAJB6S coding region (approximately 725 bp; Figure 18.B; lane 5, 9, 13 and 17) and was confirmed by *Bam*HI and *Xho*I digestion of the pcDNA-DNAJB6S control plasmid which released a fragment of the equivalent size (Figure 18.B; lane 21).

5.2.3 Site-directed mutagenesis of pCMV-DNAJB6L and pCMV-DNAJB6S J-domain

Mutation of the histidine residue within the highly conserved HPD motif of the J-domain has been previously shown to abrogate the J-domain dependent stimulation of the ATPase domain of Hsp70 (Michels *et al*, 1999). We therefore generated H31Q mutants of DNAJB6 by converting the histidine of the HPD motif (H31) to a glutamine (Q31). In order to generate site-directed mutants, two complimentary primers were constructed integrating a single nucleotide mutation that converted the codon coding for H31 (CAT) to Q31 (CAG) (Figure 19.A; red nucleotide). The single amino acid change additionally introduced a new *Bbv*I restriction recognition site allowing for the identification of H31Q mutants by restriction digest analysis (Figure 19.A, underlined). Nucleotide sequence alignments of the H31Q site-directed mutagenesis primers showed complete identity to the DNAJB6L and DNAJB6S sequence other than the single nucleotide change (Figure 19.B).

A



B

SDM H31Q primer	-----CGGAAACTGGCACTGA	16
DNAJB6S	GCGTGCAGAGACATGCCTCACCCGAGGATATTTAAAAAGGCATATCGGAAACTGGCACTGA	240
DNAJB6L	GCGTGCAGAGACATGCCTCACCCGAGGATATTTAAAAAGGCATATCGGAAACTGGCACTGA	239

SDM H31Q primer	AGTGGCAGCCAGATAAAAAATCC-----	38
DNAJB6S	AGTGGCATTCCAGATAAAAAATCCTGAGAATAAAGAAGAAGCAGAGAGAAAATTCAGCAAG	300
DNAJB6L	AGTGGCATTCCAGATAAAAAATCCTGAGAATAAAGAAGAAGCAGAGAGAAAATTCAGCAAG	299

Figure 19| Site-directed mutagenesis for the generation of H31Q DNAJB6L and DNAJB6S mutants.

A) Sequences of complementary primers constructed for site-directed mutagenesis to introduce the H31Q single nucleotide mutation into the pCMV-DNAJB6S and pCMV-DNAJB6L constructs. Red nucleotides represent the mismatch mutation incorporating the H31Q residue change. The single nucleotide mutation incorporated an additional *BbvI* recognition site (underlined) allowing for the identification of H31Q mutants by restriction enzyme analysis. (B) Nucleotide sequence alignment of site-directed mutagenesis primers with nucleotide sequence of DNAJB6L and DNAJB6S using ClustalW. Stars represent sequences of exact identity. Red nucleotide represents the mismatched nucleotide pairing incorporating the H31Q mutation.

Site-directed mutagenesis was conducted as described in Materials and Methods. Plasmids subjected to site-directed mutagenesis were transformed into DH5 α *E.coli* cells and transformants screened for the presence of the H31Q mutation using *BbvI* restriction analysis (Figure 20). For DNAJB6L, nine transformants were screened for the presence of the H31Q mutation. Restriction digest of the parental pCMV-DNAJB6L using the *BbvI* restriction enzyme was used as the control to identify plasmids that did not possess the mutation (Figure 20.A; lane 1). Of the nine transformants screened (Figure 20.A; lanes 2 to 10), four displayed the presence of an additional *BbvI* recognition site as indicated by the band highlighted by the arrow, which was absent in the pCMV-DNAJB6L control (Figure 20.A; lanes 4, 5, 6 and 10; indicated by arrow). For DNAJB6S, five transformants were screened (Figure 20.B; lanes 2 to 6), with only a single transformant displaying the presence of the additional *BbvI* recognition site introduced by the H31Q mutation (Figure 20.B; lane 2; indicated by arrow) compared to that of the pCMV-DNAJB6S control (Figure 20.B; lane 1) detected. All plasmids including pCMV-DNAJB6L, pCMV-DNAJB6S and selected pCMV-DNAJB6L(H31Q) and pCMV-DNAJB6S(H31Q) clones (Figure 20.A; lane 6, and Figure 20.B; lane 2) were sequenced to confirm the presence of the coding regions and H31Q mutation prior to functional analyses. Sequencing results positively identified the presence of the DNAJB6L and DNAJB6S coding regions in the pCMV plasmids (Figure 21.A). The H31Q mutation in the pCMV-DNAJB6L(H31Q) and pCMV-DNAJB6S(H31Q) were also confirmed by sequencing analysis (pCMV-DNAJB6S[H31Q] J-domain shares identical sequence identity therefore the sequencing results were omitted for special requirements, Figure 21.B; boxed sequence).

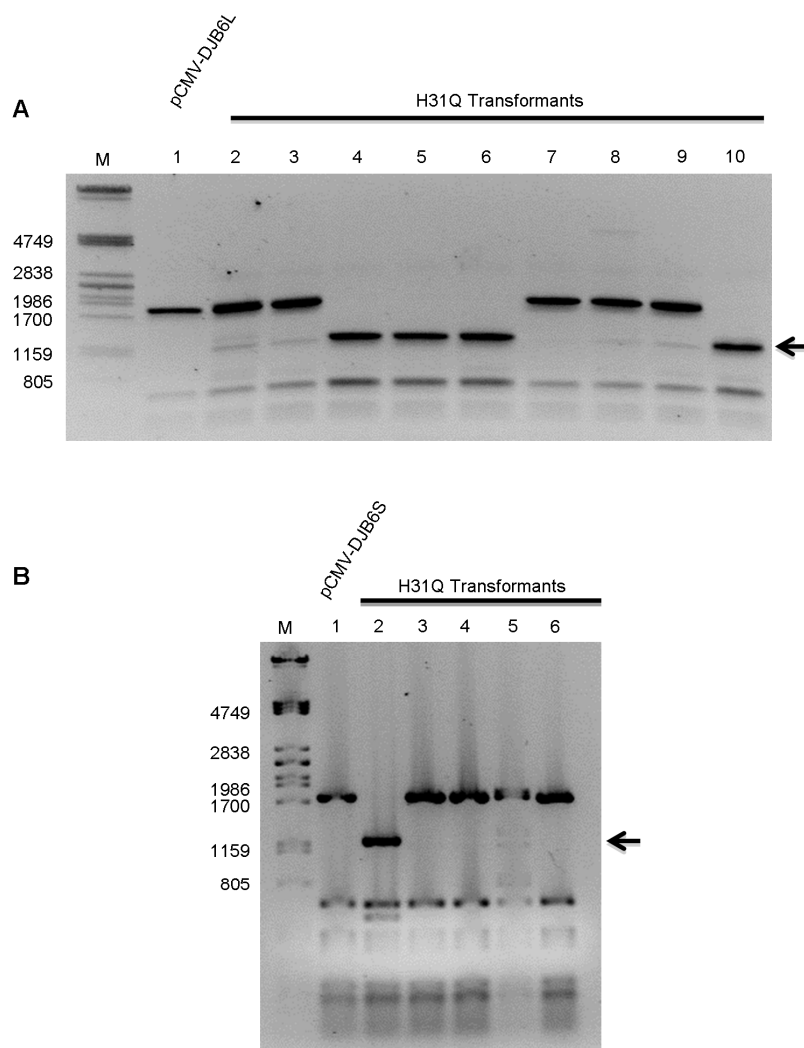


Figure 20| Restriction digest identification of H31Q pCMV-DNAJB6L and pCMV-DNAJB6S site-directed mutants.

Following site-directed mutagenesis transformants were screened by plasmids extraction and restriction digestion using *BbvI* in order to identify the presence of the H31Q mutation in the pCMV-DNAJB6L (A) and pCMV-DNAJB6S (B) plasmids. λ *Pst* molecular weight marker: lane M. pCMV-DNAJB6L (A) and pCMV-DNAJB6S (B) controls: lane 1. (A) H31Q site-directed mutagenesis clones: lanes 2 – 10. (B) H31Q site-directed mutagenesis clones: lanes 2 – 6.

A

```

pCMV_DJB6L_H31Q_T3      -----GGGGMCTAACGCGG-TGGCGGCCGCCAT----GGAGCAGAAACTC- 40
pCMV_DJB6L_T3           -----AWTRSRMYTCCSSCSGGTGGCGGCCGCCAT----GGAGCAGAAACTC- 43
DNAJB6L                  GAGAAAGGAGAGAAAGGAAAGCGCGAGG--AGCCGCCGCCACCACCAGCGCAGCAGTCT 58
                        *  **  *  *  *  *  *  *  *  *  *  *  *  *

pCMV_DJB6L_H31Q_T3      --ATCTCTGAAG-AGGATCTGGCC-----CGGGC--- 66
pCMV_DJB6L_T3           --ATCTCTGAAGTWKGATCTGGCC-----CGGGC--- 70
DNAJB6L                  GGAGCTGTGAGG-AGATTCGGGCGCTCACCCTGCCTCCCTGCGTCCCGCCACCGGCCG 117
                        *  *  *  *  *  *  *  *  *  *  *  *

pCMV_DJB6L_H31Q_T3      -----GGATCCAC-----CATGGTGGATTACTATGAAGTTCT 98
pCMV_DJB6L_T3           -----GGATCCAC-----CATGGAGGATTACTATGAAGTTCT 102
DNAJB6L                  TTCTGTCTTCGACCCATTCCAACAATCTCGTAAACATGGTGGATTACTATGAAGTTCT 177
                        ***  *  *  *  *  *  *  *  *  *  *  *

pCMV_DJB6L_H31Q_T3      AGGCGTGCAGAGACATGCCTCACCAGGATATTTAAAGGCATATCGGAACTGGCACT 158
pCMV_DJB6L_T3           AGGCGTGCAGAGACATGCCTCACCAGGATATTTAAAGGCATATCGGAACTGGCACT 162
DNAJB6L                  AGGCGTGCAGAGACATGCCTCACCAGGATATTTAAAGGCATATCGGAACTGGCACT 237
                        *****

pCMV_DJB6L_H31Q_T3      GAAGTGGCAGCCAGATAAAAAATCCTGAGAATAAAGAAGAAGCAGAGAGAAATCAAGCA 218
pCMV_DJB6L_T3           GAAGTGGCAGCCAGATAAAAAATCCTGAGAATAAAGAAGAAGCAGAGAGAAATCAAGCA 222
DNAJB6L                  GAAGTGGCAGCCAGATAAAAAATCCTGAGAATAAAGAAGAAGCAGAGAGAAATCAAGCA 297
                        *****

pCMV_DJB6L_H31Q_T3      AGTAGCGGAGGCATATGAAGTGTCTCGGATGCTAAGAAACGGGACATCTATGACAAATA 278
pCMV_DJB6L_T3           AGTAGCGGAGGCATATGAAGTGTCTCGGATGCTAAGAAACGGGACATCTATGACAAATA 282
DNAJB6L                  AGTAGCGGAGGCATATGAAGTGTCTCGGATGCTAAGAAACGGGACATCTATGACAAATA 357
                        *****

pCMV_DJB6L_H31Q_T3      TGGCAAAGAAGGATTAAATGGTGGAGGAGGAGTGAAG----- 317
pCMV_DJB6L_T3           TGGCAAAGAAGGATTAAATGGTGGAGGAGGAGTGAAG----- 318
DNAJB6L                  TGGCAAAGAAGGATTAAATGGTGGAGGAGGAGTGAAGTCAATTTGACAGTCCATTGA 417
                        *****

```

B

```

Query 75  CATGGTGGATTACTATGAAGTTCTAGGCGTGCAGAGACATGCCTCACCAGGATATTAA 134
|||
Sbjct 154 CATGGTGGATTACTATGAAGTTCTAGGCGTGCAGAGACATGCCTCACCAGGATATTAA 213

Query 135 AAAGGCATATCGGAACTGGCACTGAAGTGGCAGCCAGATAAAAAATCCTGAGAATAAAGA 194
|||
Sbjct 214 AAAGGCATATCGGAACTGGCACTGAAGTGGCAGCCAGATAAAAAATCCTGAGAATAAAGA 273

Query 195 AGAAGCAGAGAGAAAATTCAAGCAAGTAGCGGAGGCATATGAAGTGTCTCGGATGCTAA 254
|||
Sbjct 274 AGAAGCAGAGAGAAAATTCAAGCAAGTAGCGGAGGCATATGAAGTGTCTCGGATGCTAA 333

Query 255 GAAACGGGACATCTATGACAAATATGGCAAAGAAGGATTAAATGGTGGAGGAGGAGTGG 314
|||
Sbjct 334 GAAACGGGACATCTATGACAAATATGGCAAAGAAGGATTAAATGGTGGAGGAGGAGTGG 393

Query 315 AAG 317
|||
Sbjct 394 AAG 396

```

Figure 21| Confirmation of H31Q mutation in pCMV-DNAJB6L and pCMV-DNAJB6S by sequencing

(A) pCMV-DNAJB6L and pCMV-DNAJB6L(H31Q) plasmids were sent to Inqaba Biotech for sequencing using primers targeted to the pCMV and T7 promoter. Sequencing alignments (ClustalW) show presence of the single amino acid mutation (highlighted by box) incorporating the H31Q mutation in pCMV-DNAJB6(H31Q). (B) Sequence result of pCMV-DNAJB6L(H31Q) then under went BLAST analysis (NCBI) identifying DNAJB6L protein with highest identity (H31Q single amino acid mutation highlighted in box). *Data not shown: pCMV-DNAJB6S(H31Q) sequencing*

5.3 Conclusion

We were able to successfully clone DNAJB6L and DNAJB6S coding regions into mammalian expression vectors enabling the analysis of the effect of overexpression and identification of exogenous DNAJB6 through an N-terminal c-myc tag. In order to determine the J-domain dependency of DNAJB6 functions, H31Q mutants that would abrogate the ability of the DNAJB6 J domain to stimulate the ATPase activity of HSPA were successfully created using site directed mutagenesis and confirmed by sequencing analysis. Preliminary analysis can now be undertaken to determine the J-domain dependency of DNAJB6 regulation of the Wnt pathway and effect of activity on CSC properties by transfection of MCF-7 breast cancer cells. Although we showed the presence of the H31Q mutation using restriction digestion and sequencing analyses, the abrogation of HSPA ATPase stimulation by the mutation remains to be determined.

Chapter 6 : Discussion and Conclusions

6.1 Discussion

6.1.1 MCF-7 and MDA-MB-231 contained different CD44+/CD24-/EpCAM+ subpopulations.

The expression profile of a combination of cell surface markers (phenotype), rather than the expression of individual cell surface markers, has been used to identify and isolate CSC populations (Al-Hajj *et al*, 2003; Fillmore and Kuperwasser 2008). Our study revealed that the majority of MCF-7 breast cancer cells had low or undetectable expression of CD24 (CD24-). What was of particular interest was the relatively large CD44+/CD24- population, which differed to the proportion of CD44+/CD24- MCF-7 cells reported by other studies (less than 1%, Sheridan *et al*, 2006; Fillmore and Kuperwasser, 2008). The EpCAM expression profile of the MCF-7 cells was characteristic of the published data on this cell line (Fillmore and Kuperwasser, 2008), with the vast majority of the MCF-7 breast cancer cells showing positive EpCAM expression (EpCAM+). Although there have been reports that MCF-7 breast cancer cell lines from different laboratories display different biological and phenotypic properties (Osbourne *et al*, 1987), our cell line originated from the same company as that described in literature by Fillmore and Kuperwasser (2008) and as such this is unlikely to have affected our study. The difference in the phenotypic profile observed may be as a result of the environmental conditions and the effects of long-term culture of the cell line. Environmental conditions play an important regulatory role in maintaining the stem cell population in normal tissues and there is evidence to suggest a similar influence for CSC (Heddlestone *et al*, 2009; Li *et al*, 2009). Additionally there is evidence suggesting that non-stem cancer cells can be induced into displaying CSC properties, with this induction regulated by environmental and cellular signals (Roesch *et al*, 2010; Chaffer *et al*, 2011). Nevertheless, our data identified the MCF-7 breast cancer cell line to contain relatively large CD44+/CD24- and CD44+/CD24-/EpCAM+ putative BCSC populations.

The MDA-MB-231 breast cancer cell line was found to be comprised almost entirely of CD44 expressing (CD44+) cells with a relatively high CD24 expressing (CD24+) population compared to literature (Fillmore and Kuperwasser, 2008). Previous studies have shown that basal type breast cancers, including the MDA-MB-231 cell line, have a characteristically high CD44+/CD24-

phenotype (Sheridan *et al*, 2006; Fillmore and Kuperwasser, 2008; Honeth *et al*, 2008; Park *et al*, 2010; Ricardo *et al*, 2011). The fact that the MDA-MB-231 breast cancer cell line displayed a higher proportion of CD44+/CD24- cells compared to the MCF-7 cell line agreed with reports in the literature, the proportion of the population is lower than that described previously (Sheridan *et al*, 2006; Fillmore and Kuperwasser, 2008). The MDA-MB-231 cell line was found to be predominantly EpCAM-, with the proportion of CD44+/CD24-/EpCAM+ cells making up less than 1%, which is in agreement with findings in literature (Fillmore and Kuperwasser, 2008).

One of the major limitations of the use of the CD44+/CD24- phenotype for CSC identification is that there are no mechanistic data available to describe why this combination of markers should be associated with stem-like characteristics. This raises the question of which is more important – the expression of a defined cell surface marker phenotype, or the ability to exhibit the functional properties of tumorigenic stem-like cells? Interestingly, the CD44+/CD24- BCSC phenotype depicts stem-like cells as displaying no or very little expression of the cell surface marker CD24, while mammary stem cells are isolated based on high expression of the CD24 marker (Shackleton *et al*, 2006; Vaillant *et al*, 2008). However while the expression of the gene is not necessary for tumour occurrence (Gunther *et al*, 1991), there is a positive correlation between the enrichment of CD44+/CD24- cells after chemo- and radio-therapies and the resistance of putative BCSC to anti-cancer therapies. The high CD44+/CD24- phenotype in breast cancer has been linked by some authors with metastasis through the up-regulation of genes associated with invasion and a stem cell gene expression signature (Sheridan *et al*, 2006; Shipitsin *et al*, 2007). The MDA-MB-231 breast cancer cell line has been reported as having a high metastatic potential as a result of invasive and migratory properties (Price *et al*, 1990). As CD44 has been associated with metastasis (Weber *et al*, 2002; Kim *et al*, 2004; Brabletz *et al*, 2005; Balic *et al*, 2006; Sheridan *et al*, 2006) it might be expected that this cell line would show a high expression level of CD44. Although both the MCF-7 and MDA-MB-231 breast cancer cell lines contained major CD44+/CD24- populations, the EpCAM expression profile distinguished the two cell lines, with the MCF-7 breast cancer cell line being EpCAM+ while the MDA-MB-231 cell line being EpCAM-.

Some authors have suggested that the proportion of CD44+/CD24- cells within a tumour is a result of the cell type from which the cancer originates. As such the cell type origin (basal versus luminal) of breast cancers challenges the use of the CD44+/CD24- cell surface phenotype for BCSC identification and isolation. The MCF-7 and MDA-MB-231 breast cancer cell lines, although both isolated from pleural effusions, are derived from different mammary epithelial cell type origins. The MCF-7 breast cancer cell line originates from luminal type cells (expressing luminal keratins) while the MDA-MB-231 breast cancer cell line, a highly metastatic cell line, originates from basal type cells (expressing stratified epithelial keratins) (Sorlie *et al*, 2001; Sorlie *et al*, 2003). Basal and luminal breast cancers have been shown to display differential CD44/CD24 cell heterogeneity, with basal breast cancers showing characteristically high CD44+/CD24- expression. Breast cancer cell lines of luminal origin have been found to be almost entirely comprised of CD24 expressing cells (CD24+), with less than 1% of the population displaying the CD44+/CD24- phenotype associated with BCSCs (Sheridan *et al*, 2006; Fillmore and Kuperwasser, 2008; Ricardo *et al*, 2011). Additionally, Fillmore and Kuperwasser (2008) showed the CD44+/CD24- phenotype was a characteristic of the cytokeratin expression status of the breast cancer using breast cancer cell lines, where basal breast cancers showed CD44+/CD24- enrichment whereas luminal breast cancer cell lines had a low CD44+/CD24- population. It was concluded that in luminal breast cancer cell lines, where high EpCAM expression was a feature (as shown in our study using the MCF-7 cell line), BCSCs were sufficiently enriched in the CD44+/CD24- population. With basal breast cancer cell lines, the enrichment and isolation of BCSC could be obtained in the EpCAM+ population (Fillmore and Kuperwasser, 2008). Taking this into account, we found the MCF-7 breast cancer cell line to show a greater proportion of cells with the 'luminal' BCSC phenotype (CD44+/CD24-) compared to MDA-MB-231 cells displaying the 'basal' BCSC phenotype (EpCAM+). Both luminal and basal cell types originate from mammary epithelial stem cells (Villadsen *et al*, 2007). However, evidence suggests that BCSCs may not have definitive cell surface markers because BCSC populations from spontaneous murine breast tumours can be isolated using different non-overlapping cell surface marker CSC phenotypes (CD44+/CD24- and CD133+) (Wright *et al*, 2008). Although displaying different non-overlapping markers, these CSC populations displayed identical stem-like biological properties (Wright *et al*, 2008).

6.1.2 MCF-7 and MDA-MB-231 contained similar ALDEFLUOR profiles.

The MDA-MB-231 cell line showed a higher ALDH^{Hi} population compared to the MCF-7 breast cancer cell line. The ALDH^{Hi} population has been used to identify stem-like cells from both normal and malignant breast tissue (Ginestier *et al*, 2007). The proportion of ALDH^{Hi} cells in both MCF-7 (1.8%) and MDA-MB-231 (3.25%) was marginally greater than the 1% reported in previous studies (Charafe-Jauffret *et al*, 2009a). Ginestier and colleagues (2007) concluded that *in vivo* tumourigenicity, the defining CSC property (Clarke *et al*, 2006), was exclusive to ALDH^{Hi} cells. They proposed that the tumourigenic potential of the CD44+/CD24- BCSC population due to the presence of ALDH^{Hi} cells in this population (CD44+/CD24-/ALDH^{Hi} cells) meaning that tumourigenicity and CSC enrichment could be obtained irrespective of the CD44+/CD24- phenotype as there was no correlation between the two populations (Al-Hajj *et al*, 2003; Ginestier *et al*, 2007). Additionally there were a greater proportion of ALDH^{Hi} cells in basal cancers compared to luminal cancers (Charafe-Jauffret *et al*, 2009a) suggesting a greater proportion of BCSC in basal cancers. Fluctuations in absolute percentages within the ALDH^{Hi} population were observed and may have arisen from environmental changes or growth phase differences between samples. Although CSCs are maintained in the population through symmetric cell division, it remains to be determined whether the population is dynamic or maintained at a constant proportion of the tumour mass, however the recent description of the spontaneous induction of non-stem cancer cells into a CSC-like state suggests that the CSC population is dynamic (Chaffer *et al*, 2011)

6.1.3 MCF-7 breast cancer cells were capable of anchorage-independent growth.

Anchorage-independent studies revealed that the MCF-7 breast cancer cell line was capable of forming anchorage-independent colonies in serum free conditions. The anchorage-independent MCF-7 colonies showed typical three-dimensional morphologies associated with anchorage-independent growth with spherical cell clustering (Dontu *et al*, 2003; Ponti *et al*, 2005, Cariatì *et al*, 2008). Colonies additionally displayed morphological characteristics of embryoid bodies with the presence of a distinct outer membrane resembling the zona pellucida of a developing embryo (Bleil and

Wasserman, 1980). In contrast, the MDA-MB-231 cell line displayed a very low frequency of anchorage-independent colony formation after 8 days of growth. This was interesting with respect to the cell surface marker-based and enzyme activity-based identification data. As the MDA-MB-231 breast cancer cell line was found to be enriched in cells displaying the BCSC phenotype (CD44+/CD24-) as well as having a larger (although not statistically significant) ALDH^{Hi} population, it was presumed that the MDA-MB-231 cell line should have a greater proportion of cells capable of anchorage-independent growth. However, we found no correlation between the size of the CD44+/CD24- population and anchorage-independent growth properties. Additionally there was no correlation between ALDH^{Hi} proportion of cells and anchorage-independent colony formation. There was an apparent correlation between the CD44+/CD24-/EpCAM+ phenotype and anchorage-independent growth properties, suggesting that the CD44+/CD24-/EpCAM+ phenotype is more discriminating of BCSC characteristics. The observation that the MCF-7 breast cancer cell line portrayed a greater proportion of cells capable of anchorage-independent growth agreed with the suggestion that BCSCs from basal breast tumours are present in the EpCAM+ population, while BCSCs from luminal breast tumours are present in the CD44+/CD24- population (Fillmore and Kuperwasser, 2008).

6.1.4 Anchorage-independent MCF-7 cells were enriched in CD44+/CD24- and CD44+/CD24-/EpCAM+ cells.

Anchorage-independent colonies isolated from malignant breast tissues have been shown to be highly enriched, (>90% in certain cases) in CD44+/CD24- cells (Ponti *et al*, 2005). The anchorage-independent MCF-7 colonies obtained during this study were found to be enriched in CD44+/CD24- cells compared to anchorage-dependent (adherent) MCF-7 cells, with almost a 2-fold increase in the population. This increase in CD44+/CD24- cells suggested that the colonies were enrichment in BCSC-like cells, a result in agreement with previous studies (Ponti *et al*, 2005). Although there was little change observed in the proportion of ALDH^{Hi} cells between adherent and anchorage-independent MCF-7 cells, our previous data showed a lack of correlation between ALDH^{Hi} proportion of cells and anchorage-independent growth characteristics. The increase in CD44+/CD24- cells suggested the

enrichment of putative CSCs in anchorage-independent cells when compared to MCF-7 cells grown under normal conditions (adherent MCF-7 cells).

The stem-like nature of the anchorage-independent MCF-7 colonies could have been confirmed by probing for the expression of pluripotency markers, such as Oct-4, c-Myc and Sox-2. These transcription factors form part of the transcription factor cocktail capable of transforming murine and human adult fibroblast cells into induced pluripotent cells (Takahashi and Yamanaka, 2006; Takahashi *et al*, 2007). Oct-4 has been shown to be up-regulated in anchorage-independent colonies (Ponti *et al*, 2005) and c-Myc was required for the proliferation and survival of cancer stem cells in glioma (Wang *et al*, 2008), Leis and colleagues (2011) recently showed that Sox-2, and not Oct-4, overexpression was characteristic of anchorage-independent growth with knockdown of Sox-2 preventing anchorage-independent cell growth (Leis *et al*, 2011). Therefore the analysis of the levels of these pluripotency markers in the anchorage-independent colonies would support the conclusion that tumoursphere derived cells were enriched in cells with stem-like characteristics. As *in vivo* tumour formation is seen as the most important CSC characteristic (Clarke *et al*, 2006), the concluding experiment would be *in vivo* xenotransplantation assays in limiting numbers to confirm the tumorigenicity of the isolated populations where CSC-enriched population would display the exclusive *in vivo* tumour formation ability.

6.1.5 Down-regulated molecular chaperone gene expression profile.

The CD44+/CD24- phenotype remains the most widely accepted phenotype for the identification of BCSC. We showed evidence of BCSC phenotype enrichment in anchorage-independent MCF-7 cells with the CD44+/CD24- and CD44+/CD24-/EpCAM+ BCSC-associated phenotypes. We used the tumoursphere-derived MCF-7 cells as representative of a CSC-enriched sample for the analysis of chaperone gene expression. The expression of a number of molecular chaperones in adherent and anchorage independent MCF-7 colonies was compared using a commercial quantitative RT-PCR array system. The mRNA expression pattern of the chaperone and chaperone-associated genes in the MCF-7 breast cancer cell line, in general, showed that the majority of genes had low mRNA levels (Ct > 25) suggesting that not all chaperones were constitutively expressed in cancer cells. Using a fold

expression threshold of ± 1.5 fold between adherent MCF-7 and anchorage-independent cells, a number of genes were identified as being differentially regulated in the putative CSC-enriched populations. Of the total 84 genes analyzed, 22 were found to be down-regulated (< -1.5 fold) in anchorage-independent MCF-7 breast cancer cells while 2 genes were found to be up-regulated (< 1.5 fold) and 60 showed little to no change in mRNA levels between the two samples. Ideally an additional comparison with normal non-cancerous mammary cells and somatic mammary stem cells would have allowed for the identification of cancer-specific chaperones and stem-cell specific chaperones. Currently we cannot attribute the observed profile solely to the anchorage-independent state of the cells, as there may be inherent similarities between normal tissue cells and anchorage-independent cells.

6.1.6 Several DNAJ were down-regulated at the mRNA level in anchorage-independent MCF-7 cells.

The majority of the down-regulated DNAJ were of the type III (DNAJC) classification. This may be expected due to the fact that this particular DNAJ class contains the largest number of members (approximately 32) (Kampinga and Craig, 2010). However, the down-regulation of type III DNAJ may also indicate the activation or inactivation state of different cellular pathways in cancer and BCSC-like cells (Sterrenberg *et al*, 2011). Type III DNAJ are the most structurally diverse of the DNAJ classes and indeed of chaperone and chaperone-associated gene families in general. The diversity of DNAJ is evident in the sizes of the different isoforms (often greater than 40 kDa) and the domain architecture. It is thought that type III DNAJ have highly specific and stringent substrate binding domains, while type I and II show a more relaxed substrate binding domain (Kampinga and Craig, 2010). As such, the down-regulation of a specific type III DNAJ isoform may suggest a change in specific client proteins (DNAJ substrates) that are processed by particular chaperone pathways in anchorage-independent cells compared to adherent cells. Therefore chaperone-dependent processing of proteins may act as a post-translational regulatory event.

Several DNAJ have been identified as showing differential mRNA expression levels in BCSC populations (summarized in Table 5). Both DNAJB1 and DNAJC13 were found to have differential

expression profiles in CD44+/CD24- BCSC and normal breast tissue (although it was not stated whether these proteins were up-regulated or down-regulated) (Liu et al, 2007). Our study showed that both these genes were down-regulated in anchorage-independent MCF-7 cells. As we observed that anchorage-independent MCF-7 cells were enriched in CD44+/CD24- cells, our finding may corroborate that of Liu and colleagues (2007). The expression of DNAJB1 is regulated by HSF1, the heat shock protein transcription factor (Jacobs and Marnett, 2007), which was also additionally found to be down-regulated in our study. DNAJC9 was up-regulated at the RNA level in anchorage-independent BCSCs (Rappa and Lorico, 2010) and at the protein level in human embryonic cells (Van Hoof *et al*, 2006). However we found the DNAJC9 transcript to be down-regulated in our study. DNAJC19 showed decreased RNA levels in ALDH^{Hi} BCSC (Charafe-Jauffret *et al*, 2009a), a result in line with observations in our study that showed a decrease in DNAJC19 in the BCSC enriched anchorage-independent MCF-7 cells. In addition, our study identified a number of other DNAJ (DNAJB6, DNAJB12 and DNAJC14) that were down-regulated but have not been reported in the CSC gene expression literature (Table 5). The major limitations of this study along with most microarray analyses is that they do not account for the presence of isoforms or splice variants of genes. While there are 49 DNAJ genes, there is likely to be over 49 functioning DNAJ proteins due to several DNAJ genes encoding for numerous splice variants including DNAJA3, DNAJA4, DNAJB2, DNAJB5, DNAJB12, DNAJB14, DNAJC2, DNAJC7, DNAJC12 and DNAJC19. Splice variants of DNAJ genes have been shown to display opposing functions, as is the case for DNAJA3 isoforms whereby DNAJA3L (larger isoform) increased apoptosis while DNAJA3S (smaller isoform) decreased apoptosis in osteosarcoma cells treated with mitomycin C and Tumour Necrosis Factor- α (Cheng *et al*, 2005). Additionally for DNAJB6, the two isoforms display a differential subcellular localization with DNAJB6L located in both the cytoplasm and nucleus while DNAJB6S is located in the cytoplasm (Zhang *et al*, 2008).

6.1.7 HSP70 regulating proteins, DNAJC12 and BAG1, were up-regulated in anchorage-independent MCF-7 cells.

Only two genes were shown to be up-regulated (> 1.5 fold) in anchorage-independent MCF-7 cells, with both genes (DNAJC12 and BAG1) being regulators of the HSPA family of chaperones. DNAJC12 is an estrogen target gene that has been found to be up-regulated at the RNA level in ER+ compared to ER- breast cancers (De Bessa *et al*, 2006) and in ER+ anchorage-independent breast cancer cells compared ER+ parental cell lines (Kok *et al*, 2009). The MCF-7 breast cancer cell line is ER+ (Neve *et al*, 2006) and therefore the upregulation of DNAJC12 may be indicative of an increased activation of the estrogen signaling pathway in the anchorage-independent cells.

Table 5| Differential DNAJ expression in stem-like cells.

Expression State	DNAJ	Tissue	Expression	Reference
Up-regulation	DNAJA1	CD44+ BC	RNA	Shipitsin <i>et al</i> , 2007
		AIG BCSC	RNA	Rappa and Lorico, 2010
		hESC	Protein	Van Hoof <i>et al</i> , 2006
	DNAJA2	hESC	Protein	Van Hoof <i>et al</i> , 2006
	DNAJC7	hESC	Protein	Van Hoof <i>et al</i> , 2006
	DNAJC8	hESC	Protein	Van Hoof <i>et al</i> , 2006
	DNAJC9	AIG BCSC	RNA	Rappa and Lorico, 2010
		hESC	Protein	Van Hoof <i>et al</i> , 2006
	DNAJC10	ER+ AIG BCSC	RNA	Kok <i>et al</i> , 2009
	DNAJC12	ER+ AIG BCSC	RNA	Kok <i>et al</i> , 2009
		AIG BCSC	RNA	Current study ¹
Down-regulation	DNAJB1	AIG BCSC	RNA/Protein	Current study ¹
	DNAJB2	AIG BCSC	RNA	Rappa and Lorico, 2010
	DNAJB6	AIG BCSC	Protein	Current study ¹
	DNAJB12	AIG BCSC	RNA	Current study ¹
	DNAJC9	AIG BCSC	RNA	Current study ¹
	DNAJC11	AIG BCSC	RNA	Current study ¹
	DNAJC13	AIG BCSC	RNA	Current study ¹
	DNAJC14	AIG BCSC	RNA	Current study ¹
	DNAJC19	ALDH+ BCSC	RNA	Charafe-Jauffret <i>et al</i> , 2009a
		AIG BCSC	RNA	Current study ¹
	DNAJC20	AIG BCSC	RNA	Current study ¹
Differential	DNAJB1	CD44+/CD24- BCSC	RNA	Liu <i>et al</i> , 2007 ²
	DNAJC13	CD44+/CD24- BCSC	RNA	Liu <i>et al</i> , 2007 ²

Modified from Sterrenberg *et al*, 2011. ¹ AIG BCSC enriched in CD44+/CD24- phenotype. ² Gene expression comparison made between CD44+/CD24- BCSCs and normal breast tissue. AIG, Anchorage-independent growth isolated. BC, Breast cancer. BCSC, Breast cancer stem cells. ER+, Estrogen receptor positive. hESC, Human embryonic stem cells.

BAG1 showed the greatest up-regulation of all chaperone and chaperone-associated genes in the anchorage-independent MCF-7 cells. The BAG family of proteins are nucleotide-exchange factors (NEF) that are involved in the removal of ADP from the ATPase domain of HSPA proteins and replacement with ATP (Sondermann *et al*, 2001). This process of nucleotide exchange resulted in the release of the client protein from HSPA. BAG NEFs therefore function in HSPA client processing in a manner opposite to the function of DNAJ co-chaperones, and thus aid in the regulation of HSP70 function (Kampinga and Craig, 2010; Takayama *et al*, 1997). The overexpression of BAG1 has been observed in numerous cancer cells (Takayama *et al*, 1998); where members of the BAG family provided protection against apoptosis and environmental stresses including hypoxia, radiation and chemotherapeutic agents through the stabilization of Bcl-2 family proteins (Townsend *et al*, 2003). BAG1 has been shown to interact with heparin-binding EGF-like growth factors; both heparin and EGF are components of the anchorage-independent medium (Lin *et al*, 2001; Dontu *et al*, 2003). The protection against apoptosis and environmental stresses are perceived attributes of putative BCSCs, therefore supporting a possible role for BAG1. However, no correlation has been found between BAG1 mRNA and protein levels. It has been suggested that BAG1 was post-translationally regulated (Townsend *et al*, 2002) and as we were unable to determine the protein level of BAG1, whether this up-regulation at the mRNA level correlated with up-regulation at the protein level in anchorage-independent cells remains to be determined (*data not shown*). A thorough proteomics analysis (confirming the expression state of the genes at the protein level) could be performed for the majority of genes in this study. The link between transcription and protein expression is highly contentious, with evidence suggesting that there is no correlation between protein and mRNA levels (Gygi *et al*, 1999).

6.1.8 DNAJB1 and DNAJB6 isoforms were down-regulated at both the protein and mRNA levels in anchorage-independent MCF-7 cells.

In order to validate the RNA expression profile, selected genes were analyzed at the protein levels using commercially available antibodies. The protein levels were normalized to β -actin protein levels. Decreased protein levels correlated with the down-regulation at the mRNA level for all of the genes analyzed (HSP90AA1/HSP90AB1, HSPA1A/HSPA8, DNAJB1, and DNAJB6), with the exception of

DNAJC6 and HSPB1 where the protein levels remained relatively constant. Studies have shown that chaperones are highly expressed in cancer cells compared to normal tissue cells (Chant *et al*, 1995; Ciocca *et al*, 2005). It was therefore interesting to note that both the constitutively expressed cytosolic HSP90 β isoform, HSP90AB1, and the stress-inducible HSP90 α isoform, HSP90AA1, showed decreased mRNA levels (greater than -1.5 fold and less than -1.5 fold respectively) in the anchorage-independent cells. Protein analysis using an antibody that detects both isoforms identified a decrease in protein levels in anchorage-independent MCF-7 cells compared to adherent MCF-7 cells. However whether this was due to a change in a single HSP90 isoform or decrease in both isoforms remains to be determined. The decrease in the mRNA expression of DNAJB1 and DNAJB6 was confirmed at the protein level.

6.1.9 Development of a tool to study the role of DNAJB6 isoforms in cancer stem cell biology.

DNAJB6 is expressed as two isoforms, namely the DNAJB6L and DNAJB6S isoforms. For DNAJB6, the 85 amino acid truncation results in the loss of the NLS in DNAJB6S therefore the isoforms displays differential subcellular localization. The differential subcellular localization is likely to expose the DNAJB6 isoforms to different client proteins. Therefore, presuming that the C-terminal truncation does not result in the loss of important non-classical domains, the biochemical activity is likely to be the same, although the biological functions of the different DNAJB6 isoforms may vary. Hageman and colleagues (2010) found that DNAJB6S was capable of interaction with HDAC proteins. The truncation may not influence the activity of the C-terminally located HDAC binding domain present in DNAJB6 isoforms (Hageman *et al*, 2010). While there was a loss of a lysine rich portion of the protein which is likely to be acetylable in DNAJB6L, DNAJB6S does contain several lysine residues in the unique sequence located at the extreme C-terminus of this isoform.

DNAJB6 expression has been found to negatively affect breast cancer properties, leading to decreased migration and invasion. Expression of the DNAJB6L isoform has been correlated with Wnt signalling regulation (Mitra *et al*, 2008; Mitra *et al*, 2010). The fact that a decrease in DNAJB6 was observed in the anchorage-independent MCF-7 cells suggested that the activated Wnt/ β -catenin pathway may be

involved in anchorage-independent growth. The H31Q mutants of each isoform were generated in order to investigate whether the activity of DNAJB6 isoforms was dependent on the stimulation of HSPA ATPase activity. Mutation of the conserved HPD motif of the J domain is known to inhibit the J-domain dependent stimulation of HSPA ATPase activity (Michels *et al*, 1999).

We have developed a tool to determine whether the effect of DNAJB6 expression on both cancer cells and the CSC population was dependent on a functional J-domain. In order to carry out this functional analysis of DNAJB6 expression on the BCSC population, breast cancer cells will be transfected with the mammalian expression vectors (generated during the study) and analysis of the proportion of the CD44+/CD24-, CD44+/CD24-/EpCAM+, ALDH^{Hi} an anchorage-independent cells will be evaluated as described previously. As several studies have identified the possible mechanisms behind DNAJB6 tumour suppressor functions, potential insight may be provided if DNAJB6 was found to affect BCSC populations. In addition, the requirement of a functional J-domain may provide a possible drug target through the DNAJ J-domain stimulation of HSPA mechanism. These data will complement the recent findings of Mitra and colleagues, who showed that deletion of the J domain of DNAJB6L impaired the biological function (Mitra *et al*, 2012). While the authors propose that this mechanism was as a result of the loss of interaction with HSPA8, it is possible that deletion of the entire J domain (amounting to approximately one third of the protein) may have resulted in structural changes to DNAJB6 that resulted in the loss in activity. Our H31Q mutant constructs should possess a similar structure to the wild type protein and indicate whether the functional activity of DNAJB6 is dependent on the stimulation of HSPA ATPase activity.

6.2 Conclusions and Future Perspectives

We have demonstrated that the isolation of breast cancer cells from the MCF-7 breast cancer cell line based on anchorage-independent growth enriched for cells displaying the putative CD44+/CD24- and CD44+/CD24-/EpCAM+ BCSC phenotypes. In order to accurately identify this population as being a BCSC population, further analyses are needed to characterize the expression of markers of pluripotency, such as Oct4 and Sox2. Confirmation of the tumorigenic nature of the anchorage-independent MCF-7 breast cancer cells should be verified by *in vivo* tumour formation experiments. The molecular chaperone gene expression revealed several chaperone and chaperone-associated genes as being down-regulated in anchorage-independent MCF-7 breast cancer cells. As chaperone and chaperone-associated gene have been described as cellular regulators, the downregulation of several chaperones may highlight changes in signaling pathways and cellular processes that distinguish BCSC-like cells from non-stem breast cancer cells. As regulators of signaling pathways and cellular processes, chaperone and chaperone-associated genes may present potential targets to inhibit or alternate the pathways or processes thus leading to the differentiation or cell death of the therapeutically important cancer cell population. Further proteomic analysis would be required to confirm the gene expression data and to determine whether the differential mRNA levels are of biological significance. The identification of DNAJB6 as showing decreased mRNA and protein levels was interesting due to the effect that the DNAJB6 large isoform (DNAJB6L) has on the Wnt pathway, a pathway known to play an important role in stem-like nature of both normal tissue stem cells and CSCs. The affect of DNAJB6 on CSC biology has not been established. We have developed a tool to study the affect of the overexpression of DNAJB6 isoforms on cancer cells in order to determine how DNAJB6 expression would influence the CSC population. Although confirmation of the gene expression analysis at the protein level is required for many of the genes identified as displaying differential expression between non-CSC and CSC-enriched cell populations in breast cancer, this study does suggest that specific chaperone and chaperone-associated genes are required for the CSC-like biological functioning of CSCs.

Chapter 7 : References

- 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 Science* (100) 7, 3983 – 3988.
- Alison M.R. and Islam S. (2009). Attributes of adult stem cells. *Journal of Pathology* (217) 144 – 160.
- Alison M.R., Poulson R., Forbes S., and Wright N.A. (2002). An introduction to stem cells. *Journal of Pathology* (197) 419 – 423.
- Alvi A.J., Clayton H., Joshi C., Enver T., Ashworth A., Vivanco M.M., Dale T.C. and Smalley M.J. (2002). Functional and molecular characterisation of mammary side population cells. *Breast Cancer Research* 5:R1 – R5.
- Ammirante M., Rosati A., Arra C., Basile A., Falco A., Festa M., Pascale M., d'Avenia M., Marzullo L., Belisario M.A., De Marco M., Barbieri A., Giudice A., Chiappetta G., Vuttariello E., Monaco M., Bonelli P., Salvato G., Di Benedetto M., Deshmane S.L., Khalili K., Turco M.C. and Leone A. (2010). IKKgamma protein is a target of BAG3 regulatory activity in human tumour growth. *Proceedings of the National Academy of Science* 107 (16), 7497 – 7502.
- Bagatell R., Paine-Murrieta G.D., Taylor C.W., Pulcini E.J., Akinaga S., Benjamin I.G. and Whitesell L. (2000). Induction of a heat shock factor 1-dependent stress response alters the cytotoxic activity of hsp90-binding agents. *Clinical Cancer Research* 6, 3312 – 3318.
- Bagatell R. and Whitesell L. (2004). Altered Hsp90 function in cancer: A unique therapeutic opportunity. *Molecular Cancer Therapy* 3, 1021 – 1030.
- Balic M., Lin H., Young L., Hawes D., Giuliano A., McNamara G., Datar R.H. and Cole R.J. (2006). Most early disseminated cancer cells detected in bone marrow of breast cancer patients have a putative breast cancer stem cell phenotype. *Clinical Cancer Research* 12 (19), 5615 – 5621.
- Barker N., van Es J.H., Kuipers J., Kujala P., van den Born M., Cozijnsen M., Haegebarth A., Korving J., Begthel H., Peters P.J. and Clevers H. (2007). Identification of stem cells in small intestine and colon by marker gene *Lgr5*. *Nature* 449, 1003 – 1007.
- Becker J. and Craig E.A. (1994). Heat-shock proteins as molecular chaperones. *European Journal of Biochemistry* 219, 11 – 23.
- Bleil J.D. and Wasserman P.M. (1980). Structure and function of the Zona Pellucida: Identification and characterization of the proteins of the mouse oocyte's Zona Pellucida. *Developmental Biology* 76, 85 – 202.
- Bjerkvig R., Tysnes B.B., Aboody K.S., Najbauer J. and Terzis A.J.A. (2005). The origin of the cancer stem cell: current controversies and new insights. *Nature Reviews Cancer* 5, 899 – 904.
- Bonnet D. and Dick J.E. (1997). Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nature Medicine* (3) 7, 730 – 737.
- Brabletz T., Jung A., Spaderna S., Hlubek F. and Kirchner T. (2005). Migrating cancer stem cell – an integrated concept of malignant tumour progression. *Nature Reviews Cancer* 5, 744 – 749.
- Bruce W.B. and van der Gaag H. (1963). A quantitative assay for the number of murine lymphoma cells capable of proliferation *in vivo*. *Nature* (199) 79 – 80.
- Calderwood S.K., Khaleque M.A., Sawyer D.B. and Ciocca D.R. (2006). Heat shock proteins in cancer: chaperones of tumorigenesis. *Trends in Biochemical Science* 31 (3), 164 – 172.
- Cariati M., Naderi A., Brown J.P., Smalley M.J., Pinder S.E., Caldas C. and Purushotham A.D. (2008). Alpha-6 integrin is necessary for tumorigenicity of stem cell-like subpopulation within the MCF7 breast cancer cell line. *International Journal of Cancer* 122, 298 – 304.
- Chaffer C.L., Brueckmann I., Scheel C., Kaestli A.J., Wiggins P.A., Rodrigues L.O., Brooks M., Reinhardt F., Su Y., Polyak K., Arendt L.M., Kuperwasser C., Bierie B. and Weinberg R.A. (2011). Normal and neoplastic nonstem cells can spontaneously convert to a stem-like state. *Proceedings of the National Academy of Science* 106 (19) 7950 – 7955.

- Chant I.D., Rose P.E. and Morris A.G. (1995). Analysis of heat-shock protein expression in myeloid leukaemia cells by flow cytometry. *British Journal of Haematology* (90) 163 – 168.
- Charafe-Jauffret E., Ginestier C., Iovino F., Wicinski J., Cervera N., Finetti P., Hur M-H., Diebel M.E., Monville F., Dutcher J., Brown M., Viens P., Xerri L., Bertucci F., Stassi G., Dontu G., Birnbaum D. and Wicha M.S. (2009a). Breast cancer cell lines contain functional cancer stem cells with metastatic capacity and a distinct molecular signature. *Cancer Research* (4) 1302 – 1313.
- Charafe-Jauffret E., Ginestier C. and Birnbaum D. (2009b). Breast cancer stem cells: tools and models to rely on. *BMC Cancer* (9) 202.
- Cheetham M.E. and Caplan A.J. (1998). Structure, function and evolution of DnaJ: conservation and adaptation of chaperone function. *Cell Stress and Chaperones* 3, 28 - 36.
- Chen C.H., Lin H., Chuang S.M. Lin S.Y. and Chen J.J. (2010). Acidic stress facilitates tyrosine phosphorylation of HLJ1 to associate with actin cytoskeleton in lung cancer cells. *Experimental Cell Research* 316, 2910 – 2921.
- Chen S. and Smith D.F. (1998). Hop as an adaptor in the heat shock protein 70 (Hsp70) and hsp90 chaperone machinery. *Journal of Biological Chemistry* 273, 35194-35200.
- Cheng H., Cenciarelli C., Nelkin G., Tsan R., Fan D., Cheng-Mayer C. and Fidler I.J. (2005). Molecular mechanism of hTid-1, The human homolog of Drosophila tumour suppressor l(2)Tid, in the regulation of NF- κ β activity and suppression of tumour growth. *Molecular Cell Biology* (25) 44 – 59.
- Cheng H., Cenciarelli C., Tao M., Parks W.P. and Cheng-Mayer C. (2002). HTLV-1 Tax-associated hTid-1, a human DnaJ protein, is a repressor of Ikappa B kinase beta subunit. *Journal of Biological Chemistry* 277, 20605 - 20610.
- Chiosis G. (2006). Targeting chaperones in transformed systems--a focus on Hsp90 and cancer. *Expert Opinion on Therapeutic Targets* 10, 37 - 50.
- Ciocca D.R. and Calderwood S.K. (2005). Heat shock proteins in cancer: diagnostic, prognostic, predictive, and treatment implications. *Cell Stress and Chaperones* 10, 86 - 103.
- Ciocca D.R. and Calderwood S.K. (2005). Heat shock proteins in cancer: diagnostic, prognostic, predictive, and treatment implications. *Cell Stress and Chaperones* 10, 86 - 103.
- Clarke P.A., te Poele R., Wooster R. and Workman P. (2001). Gene expression microarray analysis in cancer biology, pharmacology, and drug development: progress and potential. *Biochemical Pharmacology* 62, 1311 – 1336.
- Clarke M.F., Dick J.E., Dirks P.B., Eaves C.J., Jamieson C.H.M., Jones L.D., Visvader J., Weissman I.L., and Wahl G.M. (2006). Cancer stem cells – Perspectives on current status and future direction: AACR Workshop on Cancer Stem Cells. *Cancer Research* (19) 66, 9339 – 9344.
- Collins A.T., Berry P.A., Hyde C., Stower M.J. and Maitland N.J. (2005). Prospective identification of tumorigenic prostate cancer stem cells. *Cancer Research* (23) 65, 10946 – 10951.
- Cooper L.C., Prinsloo E., Edkins A.L. and Blatch G.L. (2011). Hsp90 α / β associates with the GSK3 β /axin1/phosphor- β -catenin complex in the human MCF-7 epithelial breast cancer model. *Biochemical and Biophysical Research* 213, 550 – 554.
- Copeland E., Balgobin S., Lee C.M. and Rozakis-Adcock M. (2011). hTID-1 defines a novel regulator of c-Met Receptor signaling in renal cell carcinomas. *Oncogene*, 1 - 12.
- Csermely P., Schnaider T., Soti C., Prohaszka Z. and Nardai G. (1998). The 90-kDa molecular chaperone family: structure, function, and clinical applications. A comprehensive review. *Pharmacological Therapies* 79, 129 - 168.
- Dalerba P., Cho R.W. and Clarke M.F. (2007). Cancer stem cells: models and concepts. *Annual Review of Medicine* 58, 267 – 284.

- Dean M., Fojo T. and Bates S. (2005). Tumour stem cells and drug resistance. *Nature Reviews Cancer* (5) 275 – 284.
- Dick J.E. (2003). Breast cancer stem cells revealed. *Proceedings of the National Academy of Science* 100 (7), 3547 – 3549.
- Dick J.E. (2008). Stem cell concepts renew cancer research. *Blood* 112, 4793 – 4807.
- Di Michele M., Marcone S., Cicchillitti L., Della Corte A., Ferlini C., Scambia G., Donati .B. and Rotilio D. (2010). Glycoproteomics of paclitaxel resistance in human epithelial ovarian cancer cell lines: towards the identification of putative biomarkers. *Journal of Proteomics* 73, 879 - 898.
- De Bessa S.A., Salaorni S., Patrao D.F., Neto M.M., Brentani M.M and Nagai M.A. (2006). JDP1 (DNAJC12/Hsp40) expression in breast cancer and its association with estrogen receptor status. *International Journal of Molecular Medicine* 17, 363 - 367.
- Dhennin-Duthille I., Nyga R., Yahiaoui S., Gouilleux-Gruart V., Regnier A., Lassoued K. and Gouilleux F. (2011). The tumor suppressor hTid1 inhibits STAT5b activity via functional interaction. *Journal of Biological Chemistry* 286 (7), 5034 – 5042.
- Dontu G., Abdallah W.M., Foley J.M., Jackson K.W., Clarke M.F., Kawamura M.J. and Wicha M.S. (2003). In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells. *Genes and Development* 17, 1253 – 1270.
- Dontu G., Jackson K.W., McNicholas E., Kawamura M.J., Abdallah W.M. and Wicha M.S. (2004). Role of Notch signaling in cell-fate determination of human mammary stem/progenitor cells. *Breast Cancer Research* 6: R605 – R615.
- Dooley T.P., Reddy S.P., Wilborn T.W. and Davis R.L. (2003). Biomarkers of human cutaneous squamous cell carcinoma from tissues and cell lines identified by DNA microarrays and qRT-PCR. *Biochemical and Biophysical Research Communications* 306, 1026 – 1036.
- Engelmann K., Shen H., and Finn O.J. (2008). MCF7 Side population cells with characteristics of cancer stem/progenitor cells express the tumour antigen MUC1. *Cancer Research* (7) 68, 2419 – 2426.
- Evans C.G., Chang L. and Gestwicki J.E. (2010). Heat shock protein 70 (Hsp70) as an emerging drug target. *Journal of Medicinal Chemistry* 53 (12), 4585 – 4602.
- Fan C-Y., Lee S. and Cyr D.M. (2003). Mechanism for regulation of Hsp70 function by Hsp40. *Cell Stress and Chaperone* 8 (4), 309 – 316,
- Fang D., Nguyen T.K., Leishear K., Finko R., Kulp A.N., Hotz S., Van Belle P.A., Xu X., Elder C.E. and Herlyn M. (2005). A tumorigenic subpopulation with stem cell properties in melanomas. *Cancer Research* 65, (20) 9328 – 9337.
- Felts S.J. and Toft D.O. (2003). p23, a simple protein with complex activities. *Cell Stress and Chaperones* 8, 108 - 113.
- Fillmore C.M. and Kuperwasser C. (2008). Human breast cancer cell lines contain stem-like cells that self-renew, give rise to phenotypically diverse progeny and survive chemotherapy. *Breast Cancer Research*, 10:R25.
- Freeman B.C. and Morimoto R.I. (1996). The human cytosolic molecular chaperones Hsp90, Hsp70 (Hsc70) and hsp-1 have distinct roles in recognition of a non-native protein and protein refolding. *EMBO Journal* 15, 2969 – 2979.
- Frisch S.M. and Francis H. (1994). Disruption of epithelial cell-matrix interaction induces apoptosis. *Journal of Cell Biology* 124 (4), 619 – 626.
- Frisch S.M. and Screaton R.A. (2001). Anoikis mechanisms. *Current Opinion in Cell Biology* 13, 555 – 562.
- Ginestier C., Hur M.H., Charafe-Jauffret E., Monville F., Dutcher J., Brown M., Jacquemier J., Viens P., Kleer C., Liu S., Schott A., Hayes D., Birnbaum D., Wicha M.S. and Dontu G. (2007). ALDH1 is a marker of normal

- and malignant human mammary stem cells and a predictor of poor clinical outcome. *Cell Stem Cell* (5) 1, 555 – 567.
- Goodell M.A., Brose K., Paradis G., Connor A.S. and Mulligan C. (1996). Isolation and functional properties of murine hematopoietic stem cells that are replication *in vivo*. *Journal of Experimental Medicine* (183) 1797 – 1806.
- Grady W.M. (2006). Genomic instability and colorectal cancer. *Current Colorectal Cancer Reports* 2, (2) 66 – 71.
- Grammatikakis N., Vultur A., Ramana C.V., Sigano A., Schweinfest C.W., Watson D.K. and Raptis L. (2002). The role of HSP(N), a new member of the Hsp90 family, in signal transduction and neoplastic transformation. *Journal of Biological Chemistry* 277 (10), 8312 – 8320.
- Greene M.K., Maskos K. and Landry S.J. (1998). Role of the J-domain in the cooperation of Hsp40 with Hsp70. *Proceedings of the National Academy of Science* 95, 6108 – 6133.
- Grimshaw M.J., Cooper L., Papazisis K., Coleman J.A., Bohnenkamp H.R., Chiapero-Stanke L., Taylor-Papadimitriou J and Burchell J.M. (2008). Mammosphere culture of metastatic breast cancer cells enriched for tumorigenic breast cancer cells. *Breast Cancer Research* 10:R52.
- Gunthert U., Hofmann M., Rudy W., Reber S., Zoller M., Haubmann I., Matzku S., Wenzel A., Ponta H. and Herrlich P. (1991). A new variant of glycoprotein CD44 confers metastatic potential to rat carcinoma cells. *Cell* 85, 13 – 24.
- Gygi S.P., Rochon Y., Franz B.R. and Aebersold R. (1999). Correlation between protein and mRNA abundance in yeast. *Molecular and Cellular Biology* 19 (3), 1720 – 1730.
- Hageman J. and Kampinga H.H. (2009). Computational analysis of the human HSPH/HSPA/DNAJ family and cloning of a human HSPH/HSPA/DNAJ expression library. *Cell Stress and Chaperones* 14, 1 - 21.
- Hageman J., Rujano M.A., van Waarde M.A., Kakkar V., Dirks R.P., Govorukhina N., Oosterveld-Hut H.M., Lubsen N.H. and Kampinga H.H. (2010). A DNAJB chaperone subfamily with HDAC-dependent activities suppresses toxic protein aggregation. *Molecular Cell* 37, 355 - 369.
- Hamburger A. and Salmon S.E. (1977). Primary bioassay of human myeloma stem cells. *The Journal of Clinical Investigation* (60) 846 – 854.
- Han H., Bearss D.J., Browne L.W., Calaluce R., Nagle R.B. and Von Hoff D.D. (2002). Identification of differentially expressed genes in pancreatic cancer cells using cDNA microarray. *Cancer Research* 62, 2890 – 2896.
- Harrison H., Farnie G., Howell S.J., Rock R.E., Stylianou S., Brennan K.R., Bundred N.J. and Clarke R.B. (2010). Regulation of breast cancer stem cells activity by signalling through the Notch4 receptor. *Cancer Research* 70, (2) 709 – 718.
- Heddleston J.M., Li Z., McLendon R.E., Hjelmeland A.B. and Rich J.N. (2009). The hypoxic environment maintains glioblastoma stem cells and promotes reprogramming towards a cancer stem cell phenotype.. *Cell Cycle* 20 (8), 3274 – 3284.
- Hennessy F., Boshoff A. and Blatch G.L. (2005). Rational mutagenesis of a 40 kDa heat shock protein from *Agrobacterium tumefaciens* identifies amino acid residues critical to its *in vivo* function. *International Journal of Biochemistry and Cellular Biology* 37, 177 - 191.
- Hennessy F., Nicoll W.S., Zimmermann R., Cheetham M.E. and Blatch G.L. (2005). Not all J domains are created equal: implications for the specificity of Hsp40-Hsp70 interactions. *Protein Science* 14, 1697 - 1709.
- Hermann P.C., Huber S.L., Herrier T., Aicher A., Elwart J.W., Guba M., Bruns C.J. and Heeschen C. (2007). Distinct populations of cancer stem cells determine tumour growth and metastatic activity in human pancreatic cancer. *Cell Stem Cell* 1, 313 – 323.
- Ho M.M., Ng A.V., Lam S. and Hung J.Y. (2007). Side population in human lung cancer cell lines and tumours in enriched with stem-like cancer cells. *Cancer Research* 67, (10) 4827 – 4833.

- Hofmann M., Rudy W., Zoller M., Tolg M., Ponta H., Herrlich P. and Gunthert U. (1991). CD44 splice variants confer metastatic behavior in rats: homologous sequences are expressed in human tumour cell lines. *Cancer Research* 51, 5292 – 5297.
- Honeth G., Bendahl P-O., Ringer M., Saal L.H., Gruvberger-Saal S.K., Lovgren K., Grabau D., Ferno M., Borg A. and Hegardt C. (2008). The CD44+/CD24- phenotype is enriched in basal-like breast tumours. *Breast Cancer Research* 10:R53.
- Huang S. and Terstappen L.W.M.M. (1994). Lymphoid and myeloid differentiation of single human CD34+, HLA-DR+, CD38- Hematopoietic stem cells. *Blood* 83 (6), 1515 – 1526.
- Iwasaki M., Homma S., Hishiya A., Dolezal S.J., Reed J.C. and Takayama S. (2007). BAG3 regulates motility and adhesion of epithelial cancer cells. *Cancer Research* 67, 10252 – 10259.
- Jacobs A.T. and Marmett L.J. (2009). HSF-1 mediated BAG3 expression attenuates apoptosis in 4-hydroxynonenal-treated colon cancer cells via stabilization of anti-apoptotic Bcl-2 proteins. *Journal of Biological Chemistry* 285 (14), 9176 – 9183.
- Jameel A., Skilton R.A., Campbell T.A., Chander S.K., Coombes R.C. and Luqmani Y.A. (1992). Clinical and biological significance of HSP89 alpha in human breast cancer. *International Journal of Cancer* 50 (3), 409 – 415.
- Kamal A., Thao L., Sensintaffar J., Zhang L., Boehm M.F., Fritz L.C. and Burrows F.J. (2003). A high-affinity conformation of Hsp90 confers tumour selectivity on Hsp90 inhibitors. *Nature* 425, 407 – 410.
- Kampinga H.H. and Craig E.A. (2010). The HSP70 chaperone machinery: J proteins as drivers of functional specificity. *Nature Reviews Molecular and Cellular Biology* 11, 579 - 592.
- Kampinga H.H., Hageman J., Vos M.J., Kubota H., Tanguay R.M., Bruford E.A., Cheetham M.E., Chen B. and Hightower L.E. (2009). Guidelines for the nomenclature of the human heat shock proteins. *Cell Stress and Chaperones* 14, 105 - 111.
- Kim H-R., Wheeler M.A., Wilson C.M., Iida J., Eng D., Simpson M.A., McCarthy J.B. and Bullard K.M. (2004). Hyaluronan facilitates invasion of colon carcinoma cells *in vitro* via interaction with CD44. *Cancer Research* 64, 4569 – 4576.
- Kim S.W., Chao T.H., Xiang R., Lo J.F., Campbell M.J., Fearn C. and Lee J.D. (2004). Tid1, the human homologue of a Drosophila tumor suppressor, reduces the malignant activity of ErbB-2 in carcinoma cells. *Cancer Research* 64, 7732-7739.
- Kim S.W., Hayashi M., Lo J.F., Fearn C., Xiang R., Lazennec G., Yang Y. and Lee J.D. (2005). Tid1 negatively regulates the migratory potential of cancer cells by inhibiting the production of interleukin-8. *Cancer Research* 65, 8784 - 8791.
- Kise Y., Takenaka K., Tezuka T., Yamamoto T. and Miki H. (2006). Fused kinase is stabilized by Cdc37/Hsp90 and enhances Gli protein levels. *Biochemical and Biophysical Research Communications* 351, 78 – 84.
- Kok M., Koornstra R.H., Margarido T.C., Fles R., Armstrong N.J., Linn S.C., Van't Veer L.J and Weigelt B. (2009). Mammosphere-derived gene set predicts outcome in patients with ER-positive breast cancer. *Journal of Pathology* 218, 316 - 326.
- Laemmli U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680 – 685.
- 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 leukaemia after transplantation into SCID mice. *Nature* (367) 645 – 648.
- Lawson J.C., Blatch G.L. and Edkins A.L. (2009). Cancer stem cells in breast cancer and metastasis. *Breast Cancer Research and Treatment* DOI 10.1007/s10549-009-0524-9.

- Lee D-W., Wu X., Eisenberg E. and Greene L.E. (2006). Recruitment dynamics of GAK and auxilin to clathrin-coated pits during endocytosis. *Journal of Cell Science* 119, 3502 – 3512.
- Letunic I., Doerks T. and Bork P. (2012) SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acid Research* 40, D302 – D305.
- 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. *Cancer Research* 67, (3) 1030 – 1037.
- Li X., Lewis M.T., Huang J., Guitierrez C., Osborne C.K., Wu M-F., Hilsenbeck S.G., Pavlick A., Zhang X., Chamness G.C., Wong H., Rosen J. and Chang J.C. (2008). Intrinsic resistance of tumorigenic breast cancer cells to chemotherapy. *Journal of the National Cancer Institute* (100) 9, 672 – 679.
- Li Z., Bao S., Wu Q., Wang H., Eyler C., Sathornsumetee S., Shi Q., Cao Y., Lathia J., McLendon R.E., Hjelmeland A.B. and Rich J.N. (2009). Hypoxia-inducible factors regulate tumorigenic capacity of glioma stem cells. *Cancer Cell* 15, 501 – 513.
- Lin J., Hutchinson L., Gastoni S.M., Raab G. and Freeman M.R. (2001). BAG-1 is a novel cytoplasmic binding partner of the membrane for of heparin-binding EGF-like growth factor. *Journal of Biological Chemistry* 276 (32), 30127 – 30132.
- Lin S.Y., Hsueh C.M., Yu S.L., Su C.C., Shum W.Y., Yeh K.C., Chang G.C. and Chen J.J. (2010). HLJ1 is a novel caspase-3 substrate and its expression enhances UV-induced apoptosis in non-small cell lung carcinoma. *Nucleic Acids Research* 38, 6148 - 6158.
- Lindquist S. and Craig E.A. (1988), The heat-shock proteins. *Annual Reviews in Genetics* 22, 631 - 677.
- Lindsey J.C., Lusher M.E., Strathdee G., Brown R., Gilbertson R.J., Bailey S., Ellison D.W. and Clifford S.C. (2006). Epigenetic inactivation of MCGID1 (DNMT3L) in malignant paediatric brain tumours. *International Journal of Cancer* 118, 346 - 352.
- Liu R., Wang X., Chen G.Y., Dalerba P., Gurney A., Hoey T., Sherlock G., Lewicki J., Shedden K. and Clarke M.F. (2007). The prognostic role of a gene signature from tumorigenic breast cancer cells. *The New England Journal of Medicine* 3, (356) 217 – 226.
- Livak K.J. and Schmittgen T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method. *Methods* 25, 402 – 408.
- Lobo N.A., Shimono Y., Qian D. and Clarke M.F. (2007). The biology of cancer stem cells. *Annual Review of Cell and Developmental Biology* (23) 675 – 699.
- Lu Z. and Cyr D.M. (1998). Protein folding activity of Hsp70 is modified differentially by the hsp40 co-chaperones Sis1 and Ydj1. *Journal of Biological Chemistry* 273, 27824 - 27830.
- Lyng H., Brovig R.S., Svendsrud D.H., Holm R., Kaalhus O., Knutstad K., Oksefjell H., Sundfor K., Kristensen G.B and Stokke T. (2006). Gene expressions and copy numbers associated with metastatic phenotypes of uterine cervical cancer. *BMC Genomics* 7, 268.
- Mahalingam D., Swords R., Carew J.S., Nawrocki S.T., Bhalla K. and Giles F.J. (2009). Targeting HSP90 for cancer therapy. *British Journal of Cancer* 100, 1523 – 1529.
- Malanchi I., Peinado H., Kassen D., Hussenet T., Metzger D., Chambon P., Huber M., Hohl D., Cano A., Birchmeier W. and Huelsken J. (2008). Cutaneous cancer stem cell maintenance is dependent on beta-catenin signaling. *Nature* 452, 650 – 654.
- Maloney A., Clarke P.A., Naaby-Hansen S., Stein R., Koopman J-O., Akpan A., Yang A., Zvelebil M., Cramer R., Stimson L., Aherne W., Banerji U., Judson I., Sharp S., Powers M., deBilly E., Salmons J., Walton M., Burlingame A., Waterfield M. and Workman P. (2007). Gene and protein expression profiling of human ovarian cancer cells treated with heat shock protein 90 inhibitor 17-allylamino-17-demethoxygeldanamycin. *Cancer Research* 67, 3239 – 3253.

- Mani S.A., Guo W., Liao M.-J., Eaton E.N., Ayyanan A.,Zho A.Y., Brooks M., Reinhard F., Zhang C.C., Shipitsin M., Campbell L.L., Polyak K. Briskin C., Yang J. and Weinberg R.A. (2008). The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* 133, 704 – 715.
- Massol R.H., Boll W., Griffen A.M. and Kirchhausen T. (2006). A burst of auxilin recruitment determines the onset of clathrin-coated vesicle uncoating. *Proceedings of the National Academy of Science* 103, 10256 – 10270.
- Mayer M.P. and Bukau B. (2005). Hsp70 chaperones: Cellular functions and molecular mechanism. *Cellular and Molecular Life Sciences* 62, 670 – 684.
- Meyer H-A., Grau H., Kraft R., Kostkat S., Prehn S., Kalies K-U. and Hartmann E. (2000). Mammalian Sec61 is associated with Sec62 and Sec63. *Journal of Biological Chemistry* 275 (12), 14550 – 14557.
- Michels A.A., Kanon B., Bensaude O. and Kampinga H.H. (1999). Heat shock protein (Hsp) 40 mutants inhibit Hsp70 in mammalian cells. *Journal of biological Chemistry* 274 (51), 36757 – 36763.
- Minami Y., Kawasaki H., Miyata Y., Suzuki K. and Yahara I. (1991). Analysis of native forms and isoform compositions of the mouse 90-kDa heat shock protein, HSP90. *Journal of Biological Chemistry* 266 (16), 10099 – 10103.
- Minami Y., Kimura Y., Kawasaki H., Suzuki K. and Yahara I. (1994). The carboxy-terminal region of mammalian HSP90 is required for its dimerization and function *in vivo*. *Molecular and Cellular Biology* 14 (2), 1459 – 1464.
- Mitra A., Fillmore R.A., Metge B.J., Rajesh M., Xi Y., King J., Ju J., Pannell L., Shevde L.A. and Samant R.S. (2008). Large isoform of MRJ (DNAJB6) reduces malignant activity of breast cancer. *Breast Cancer Research* 10:R22 (doi:10.1186/bcr1874).
- Mitra A., Shevde L.A. and Samant R.S. (2009). Multi-faceted role of HSP40 in cancer. *Clinical and Experimental Metastasis* (26) 559 – 567.
- Mitra A., Menezes M.E., Shevde L.A. and Samant R.S. (2010). DNAJB6 induces degradation of beta-catenin and causes partial reversal of mesenchymal phenotype. *Journal of Biological Chemistry* 285, 24686 - 24694.
- Mitra A., Menezes M.E., Pannell L.K., Mulekar M.S., Honkanen R.E., Shevde L.A. and Samant R.S. (2012) DNAJB6 chaperones PP2A mediated dephosphorylation of GSK3 β to downregulate β -catenin transcription target, osteopontin. *Oncogene*. Online publication ahead of print 23 January 2012) doi:10.1038/onc.2011.623
- Morel A-P., Lievre M., Thomas C., Hinkal G., Ansieau S. and Puisieux A. (2008). Generation of breast cancer stem cells through the epithelial-mesenchymal transition. *PLoS one* 3 (8), e2888.
- Morrison S.J and Spradling A.C. (2008). Stem cells and niches: Mechanisms that promote stem cell maintenance throughout life. *Cell* 132, 598 – 611.
- Murone M., Luoh S-M., Stone D., Li W., Gurney A., Armanini M., Grey C., Rosenthal A. and de Sauvage F.J. (2000). Gli regulation by the opposing activities of Fused and Suppressor of Fused. *Nature Cell Biology* 2, 310 – 312.
- Neve R.M., Chin K., Fridlyand J., Yeh J., Baehner F.L., Fevr T., Clark L., Bayani N., Coppe J-P., Tong F., Speed T., Spellman P.T., DeVries S., Lupak A., Wang N.J., Kuo W-L., Stilwell J.L., Pinkel D., Albertson D.G., Waldman F.M., McCormick F., Dickson R.B., Johnson M.D., Lippman M., Ethier S., Gazdar A. and Gray J.W. (2006). A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* 10 (6), 515 – 527.
- Osbourne C.K., Hobbs K. and Trent J.M. (1987). Biological differences among MCF-7 human breast cancer cell lines from different laboratories. *Breast Cancer Research and Treatment* 9, 111 – 121.
- Panaretou B., Prodromou C., Roe S.M., O'Brien R., Ladbury J.E., Piper P.W. and Pearl L.H. (1998). ATP binding and hydrolysis are essential to the function of the Hsp90 molecular chaperone *in vivo*. *EMBO Journal* 17 (16), 4829 – 4836.

- Pang R., Law H.L., Chu A.C.Y., Poon J.T., Lam C.S.C., Chow A.K.M., Ng L., Cheung L.W.H., Lan X.R., Lan H.Y., Tan V.P.Y., Yau T.C., Poon R.T. and Wong B.C.Y. (2010). A subpopulation of CD26+ cancer stem cells with metastatic capacity in human colorectal cancer. *Cell Stem Cell* 6, 603 – 615.
- Park C.H., Bergsagel D.E. and McCulloch E.A. (1971). Mouse myeloma tumour stem cells: a primary cell culture assay. *Journal of the National Cancer Institute* 46, (2) 411 – 422.
- Patrawala L., Calhoun T., Schneider-Broussard R., Zhou J., Claypool K. and Tang D.G. (2005). Side population is enriched in tumorigenic, stem-like cancer cells, whereas ABCG2⁺ and ABCG2⁻ cancer cells are similarly tumorigenic. *Cancer Research* (14) 65, 6207 – 6219.
- Peng C., Brain J., Hu Y., Goodrich A., Kong L., Grayzel D., Pak R., Read M. and Li S. (2007). Inhibition of heat shock protein 90 prolongs survival of mice with BCR-ABL-T315I-induced leukemia and suppresses leukemic stem cells. *Blood* 110, 678 – 685.
- Pfeiffer M.J. and Schalken J.A. (2010). Stem cell characteristics in prostate cancer cell lines. *European Urology* (57) 246 - 255.
- Ponti D., Costa A., Zaffaroni N., Pratesti G., Petrangolini G., Coradini D., Pilotti S., Pierotti M.A. and Daidone M.G. (2005). Isolation and in vitro propagation of tumorigenic breast cancer cells with stem/progenitor cell properties. *Cancer Research* (13) 65, 5506 – 5511.
- Powers M.V., Clarke P.A. and Workman P. (2008). Dual targeting of HSC70 and HSP72 inhibits HSP90 function and induces tumor-specific apoptosis. *Cancer Cell* 14, 250 – 262.
- Pratt W.B. (1998). The hsp90-based chaperone system: Involvement in signal transduction from a variety of hormone and growth factor receptors. *Proceedings of the Society for Experimental Biology and Medicine* 217, 420 – 434.
- Prince M.E., Sivanandan R., Kaczorowski A., Wolff G.T., Kaplan M.J., Dalerba P., Weissman I.L., Clarke M.F., and Allies L.E. (2007). Identification of a subpopulation of cells with cancer stem cell properties in head and neck squamous cell carcinoma. *Proceeding of the National Academy of Science* (104) 3, 973 – 978.
- Prodromou C., Roe S.M., O'Brien R., Ladbury J.E., Piper P.W. and Pearl L.H. (1997). Identification and structural characterization of the ATP/ADP-binding site in the Hsp90 molecular chaperone. *Cell* 90, 65 – 75.
- Ralhan R. and Kaur J. (1995). Differential expression of Mr 70,000 heat shock protein in normal, premalignant, and malignant human uterine cervix. *Clinical Cancer Research* 1, 1217 - 1222.
- Rappa G. and Lorico A. (2010). Phenotypic characterization of mammosphere-forming cells from the human MA-11 breast carcinoma cell line. *Experimental Cell Research* 316, 1576 - 1586.
- Ren J., Bharti A., Raina D., Chen W., Ahmad R. and Kufe D. (2006). MUC1 oncoprotein is targeted to mitochondria by heregulin-induced activation of c-Src and the molecular chaperone HSP90. *Oncogene* 25, 20 – 31.
- Reya T. and Clevers H. (2005). Wnt signalling in stem cells and cancer. *Nature* 434, 843 - 850.
- Reya T., Duncan A.W., Ailles L., Domen J., Scherer D.C., Willert K., Hintz L., Nusse R. and Weissman I.L. (2003). A role for Wnt signalling in self-renewal of haematopoietic stem cells. *Nature* 423, 409 – 414.
- Reya T., Morrison S.J., Clarke M.F. and Weissman I.L. (2001). Stem cells, cancer, and cancer stem cells. *Nature* (414) 105 – 111.
- Reynolds B.A. and Rietze R.L. (2005). Neural stem cells and neurospheres – re-evaluating the relationship. *Nature Methods* (5) 2, 333 – 336.
- Ricardo S., Vieira A.F., Herhard R., Leitao D., Ointo R., Cameselle-Teijeiro J.F., Milanezi F., Schmitt F. and Paredes J. (2011). Breast cancer stem cell markers CD44, CD24 and ALDH1: expression distribution within intrinsic molecular subtype. *Journal of Clinical Pathology* 64, 937 – 946.
- Ricci-Vitiani L., Lombardi D.G., Pilozzi E., Biffoni M., Todaro M., Peschle C. and De Maria R. (2007). Identification and expression of human colon-cancer-initiating cells. *Nature* (445) 111 – 115.

- Roesch A., Fukunaga-Kalabis M., Schmidt E.C., Zabierowski S.E., Brafford P.A., Vultur A., Basu D., Gimothy P., Vogt T. and Herlyn M. (2010). A temporary distinct subpopulation of slow-cycling melanoma cells is required for continuous tumour growth. *Cell* 141, 583 – 594.
- Rohde M., Daugaard M., Jensen M.H., Helin K., Nylandsted J. and Jaattela M. (2005). Members of the heat-shock protein 70 family promote cancer cell growth by distinct mechanisms. *Genes and Development* 19, 570 – 582.
- Sahi C., Lee T., Inada M., Pleiss J.A. and Craig E.A. (2010). Cwc23, an essential J protein critical for pre-mRNA splicing with a dispensable J domain. *Molecular and Cellular Biology* 30, 33 – 42.
- Salma A., Tsiapos A. and Lazaridis I. (2007). The viral SV40 T antigen cooperates with dj2 to enhance hsc70 chaperone function. *FEBS Journal* 274, 5021 – 5027.
- Sato N., Meijer L., Skaltsounis L., Greengard P. and Brivanlou A.H. (2004). Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. *Nature Medicine* 10 (1), 55 – 63.
- Sauvageot C.M-E., Weatherbee J.L., Kesri S., Winters S.E., Barnes J., Dellagatta J., Ramakrishna N.R., Stiles C.D., Kung A.L-J., Kieran M.W. and Wen P.Y.C. (2009). Efficacy of the HSP90 inhibitor 17-AAG in human glioma cell lines and tumorigenic glioma stem cells. *Neuro-Oncology* 11, 109 – 121.
- Scadden D.T. (2006). The stem-cell niche as an entity of action. *Nature* 441, 1075 – 1079.
- Schultz J., Milpetz F., Bork P. and Ponting C.P. (1998). SMART, a simple modular architecture research tool: identification of signaling domains. *Proceedings of the National Academy of Science* 95, 5857 – 5864.
- Schmittgen T.D. and Livak K.J. (2008). Analyzing the real-time PCR data by the comparative Ct method. *Nature Protocols* 3 (6), 1101 – 1108.
- Setati M.M., Prinsloo E., Longshaw V.M., Murray P.A., Edgar D.H., and Blatch G.L. (2010) Leukaemia inhibitory factor promotes Hsp90 association with STAT3 in mouse embryonic stem cells *IUBMB Life*. 62, 61 – 6.
- Sever S., Skoch J., Newmyer S., Ramachandran R., Ko D., McKee M., Bouley R., Ausiello D., Hyman B.T. and Bacskaï B.J. (2006). Physical and functional connection between auxilin and dynamin during endocytosis. *EMBO Journal* 25, 4163 – 4174.
- Sha B., Lee S. and Cyr D.M. (2000). The crystal structure of the peptide-binding fragment from the yeast Hsp40 protein Sis1. *Structure* 8, 799 – 807.
- Shackleton M., Vaillant F., Simpson K.J., Stingl J., Smyth G.K., Asselin-Labat M-L., Wu L., Lindeman G.J. and Visvader J.E. (2006). Generation of a functional mammary gland from a single stem cell. *Nature* 439, 84 – 88.
- Sheridan C., Kishimoto H., Fuchs R.K., Mehrotra S., Bhat-Nakshatri P., Turner C.H., Goulet Jr R., Badve S. and Nakshatri H. (2006). CD44+/CD24- breast cancer cells exhibit enhanced invasive properties: an early step necessary for metastasis. *Breast Cancer Research* 8:R59
- Shimamura T., Lowell A.M., Engelman J.A. and Shapiro G.I. (2005). Epidermal growth factor receptors harboring kinase domain mutations associate with the heat shock protein 90 chaperone and are destabilized following exposure to Geldanamycins. *Cancer Research* 65, 6401 – 6408.
- Shimono Y., Zabala M., Cho R.W., Lobo N., Dalerba P., Qian D., Diehn M., Liu H., Panula S.P., Chiao E., Dirbas F.M., Somolo G., Reijo Pera R.A., Lao K. and Clarke M.F. (2009). Downregulation of miRNA-200c links breast cancer stem cells with normal stem cells. *Cell* (138) 592 – 603.
- Shipitsin M., Campbell L.L., Argani P., Weremowicz S., Bloushtain-Qimron N., Yao J., Nikolskaya T., Serebryiskaya T., Beroukhim R., Hu M., Halushka M.K., Sukumar S., Parker L.M., Anderson K.S., Harris L.N., Garber J.E., Richardson A.L., Schnitt S.S., Nikolsky Y., Gelman R.S. and Polyak K. (2007). Molecular definition of breast tumor heterogeneity. *Cancer Cell* (11) 259 – 273.

- Shridhar V., Bible K.C., Staub J., Avula R., Lee Y.K., Kalli K., Huang H., Hartmann L.C., Kaufmann S.H. and Smith D.I. (2001). Loss of expression of a new member of the DNAJ protein family confers resistance to chemotherapeutic agents used in the treatment of ovarian cancer. *Cancer Research* 61, 4258 - 4265.
- Sigrist C.J.A., Cerutti L., de Castro E., Langendijk-Genevaux P.S., Bulliard V., Bairoch A. and Hulo N. (2010). PROSITE, a protein domain database for functional characterization and annotation. *Nucleic Acid Research* 38, D161 – D166.
- Sinha D., Joshi N., Chittoor B., Samji P. and D'Silva P. (2010). Role of Magmus in protein transport and human mitochondria biogenesis. *Human Molecular Genetics* 19 (7), 1248 – 1262.
- Singh S.K., Clarke I.D., Terasaki M., Bon V.E., Hawkins C., Squire J. and Dirks P.B. (2003). Identification of a cancer stem cell in human brain tumours. *Cancer Research* (63) 5821 – 5828.
- Sittler A., Lurz R., Lueder G., Priller J., Lehrach H., Hayer-Hartl M.K., Hartl F.U. and Wanker E.E. (2001). Geldanamycin activates a heat shock response and inhibits huntingtin aggregation in a cell culture model of Huntington's disease. *Human Molecular Genetics* 10, 1307 - 1315.
- Sondermann H., Scheufler C., Schneider C., Hohfeld J., Hartl F.U. and Moarefi I. (2001). Structure of BAG-HSC70 complex: convergent functional evolution of Hsp70 nucleotide exchange factors. *Science* 291, 1553 – 1557.
- Sorlie T., Perou C.M., Tibshirani R., Aas T., Geisler S., Johnsen H., Hastie T., Eisen M.B., van de Rijn M., Jeffrey S.S., Thorsen T., Quist H., Matese J.C., Brown P.O., Botstein D., Lonning P.E. and Borresen-Dale A-L. (2001). Gene expression patterns of breast carcinomas distinguish tumour subclasses with clinical implications. *Proceeding of the National Academy of Science* 98 (19), 10869 – 10874.
- Sorlie T., Tibshirani R., Parker J., Hastie T., Marron J.S., Nobel A., Deng S., Johnsen H., Pesich R., Geisler S., Demeter J., Perou C.M., Lonning P.E., Brown P.O., Borresen-Dale A-L. and Botstein D. (2003). Repeated observation of breast tumour subtypes in independent gene expression data sets. *Proceedings of the National Academy of Science* 100 (14), 8418 – 8423.
- Sreedhar A.S., Kalmar E., Csermely P. and Shen Y-F. (1994). Hsp90 isoforms: functions, expression and clinical importance. *FEBS Letters* 562, 11 – 15.
- Sterrenberg J.N., Blatch G.L. and Edkins A.L. (2011). Human DNAJ in cancer and stem cells. *Cancer Letters* 312, 129 – 142.
- Syken J., De-Medina T. and Munger K. (1999). TID1, a human homolog of the *Drosophila* tumor suppressor *l(2)tid*, encodes two mitochondrial modulators of apoptosis with opposing functions. *Proceedings of the National Academy of Science* 96, 8499 - 8504.
- Sylvester K.G. and Longmaker M.Y. (2004). Stem Cells: Review and Update. *Archives Surgery* (139) 93 – 99.
- Takahashi K., Tanabe K., Ohnuki M., Narita M., Ichisaka T., Tomoda K. and Yamanaka S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131, 861 – 872.
- Takahashi K. and Yamanaka S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126, 663 – 676.
- Takamura Y., Yagi H., Hase K., Yoneda T., Maeda M., Akagi Y. and Sato M. (2001). JDD1, a novel member of the DnaJ family, is expressed in the germinal zone of the rat brain. *Biochemical Biophysical Research Communications* 285, 387 - 392.
- Takayama S., Krajewski S., Krajewska M., Kitada S., Zapata J.M., Kochel K., Knee D., Scudiero D., Tudor G., Miller G.J., Miyashita T., Yamada M. and Reed J.C. (1998). Expression and location of Hsp70/Hsc-binding anti-apoptotic protein BAG1 and its variants in normal tissue and tumour cell lines. *Cancer Research* 58, 3116 – 3131.
- Thomas C.G. and Spyrou G. (2009). ERdj5 sensitizes neuroblastoma cells to endoplasmic reticulum stress-induced apoptosis. *Journal of Biological Chemistry* 284, 6282 - 6290.

- Tissieres A., Mitchell H.K. and Tracy U.M. (1974). Protein synthesis in salivary glands of *Drosophila melanogaster*: Relation to chromosome puffs. *Journal of Molecular Biology* 84, 389 – 398.
- Towbin H., Staehelin T. and Gordon J. (1979). Electrophoresis transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Science* 76 (6), 4350 – 4354.
- Townsend P.A., Cutress R.I., Sharp A., Brimmel M. and Packham G. (2003). BAG-1 prevents stress-induced long-term growth inhibition of breast cell via a chaperone-dependent pathway. *Cancer Research* 63, 4150 – 4157.
- Trepel J., Mollapour M., Giaccone G. and Neckers L. (2010). Targeting the dynamic HSP90 complex in cancer. *Nature Reviews Cancer* 10, 537 – 549.
- Tsai J. and Douglas M.G. (1996). A conserved HPD sequence of the J-domain is necessary for YDJ1 stimulation of Hsp70 ATPase activity at a site distinct from substrate binding. *Journal of Biological Chemistry* 271, 9347 – 9354.
- Tsai M.F., Wang C.C., Chang G.C., Chen C.Y., Chen H.Y., Cheng C.L., Yang Y.P., Wu C.Y., Shih F.Y., Liu C.C., Lin H.P., Jou Y.S., Lin S.C., Lin C.W., Chen W.J., Chan W.K., Chen J.J. and Yang P.C. (2006). A new tumor suppressor DnaJ-like heat shock protein, HLJ1, and survival of patients with non-small-cell lung carcinoma. *Journal of the National Cancer Institute* 98, 825 – 838.
- Vaillant F., Asselin-Labat M-L., Shackleton M., Forrest N.C., Lindeman G.J. Visvader J.E. (2008). The mammary progenitor marker CD61/beta-3 integrin identifies cancer stem cells in mouse model of mammary tumorigenesis. *Cancer Research* 68, 7711 – 7717.
- van der Spuy J., Cheetham M.E., Dirr H.W. and Blatch G.L. (2001). The cochaperone murine stress-inducible protein 1: overexpression, purification, and characterization. *Protein Expression and Purification* 21, 462 – 469.
- Van Hoof D., Passier R., Ward-Van Oostwaard D., Pinkse M.W., Heck A.J., Mummery C.L. and Krijgsveld J. (2006). A quest for human and mouse embryonic stem cell-specific proteins. *Molecular and Cellular Proteomics* 5, 1261 – 1273.
- 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 – 478.
- Villadsen R., Fridriksdottir A.J., Ronnov-Jssen L., Gudjohsson T., Rank F., LaBarge M.A., Bissel M.J. and Petersen O.W. (2007). Evidence for a stem cell hierarchy in the adult human breast. *Journal of Cell Biology* 171 (1), 87 – 101.
- Visvader J.E and Lindeman G.J. (2008). Cancer stem cells in solid tumours: accumulating evidence and unresolved questions. *Nature Reviews Cancer* 8, 755 – 767.
- Vos M.J., Hageman J., Carra S. and Kampinga H.H. (2008). Structural and functional diversities between members of the human HSPB, HSPH, HSPA, and DNAJ chaperone families. *Biochemistry* 47, 7001 – 7011.
- Wall D., Zylicz M. and Georgopoulos C. (1994). The NH₂-terminal 108 amino acids of the Escherichia coli DnaJ protein stimulate the ATPase activity of DnaK and are sufficient for lambda replication. *Journal of Biological Chemistry* 269, 5446 – 5451.
- Wandinger S.K. Richter K. and Buchner J. (2008). The Hsp90 chaperone machinery. *Journal of Biological Chemistry* 283, 18473 – 18477.
- Wang C.C., Liao Y.P., Mischel P.S., Iwamoto K.S., Cacalano N.A. and McBride W.H. (2006). HDJ-2 as a target for radiosensitization of glioblastoma multiforme cells by the farnesyltransferase inhibitor R115777 and the role of the p53/p21 pathway. *Cancer Research* 66, 6756 – 6762.
- Wang J., Wang H., Li Z., Wu Q., Lathia J.D., McLendon R.E., Hjelmeland A.B and Rich J.N. (2008). C-Myc is required for maintenance of glioma cancer stem cells. *PLoS one* 3 (11), e3769

- Watson E.D., Mattar P., Schuurmans C. and Cross J.C. (2009). Neural stem cell self-renewal requires the Mrj co-chaperone. *Development Dynamics* (238) 2564 – 2574.
- Weber G.F., Bronson R.T., Ilagan J., Cantor H., Schmidts R. and Mak T.W. (2002). Absence of the CD44 gene prevents sarcoma metastasis. *Cancer Research* 62, 2281 – 2296.
- Wen K.W. and Damania B. (2010). Hsp90 and Hsp40/Erdj3 are required for the expression and anti-apoptotic function of KSHV K1. *Oncogene* 29, 3532 - 3544.
- Whitesell L. and Lindquist S.L. (2005). HSP90 and the chaperoning of cancer. *Nature Reviews Cancer* (5) 761 – 772.
- Whitesell L. and Lindquist S.L. (2009). Inhibiting the transcription factor HSF1 as an anticancer strategy. *Expert Opinion in Therapeutic Targets* 13, 469 - 478.
- Wright M.H., Calcagno A.M., Salcido C.D., Carlson M.D., Ambudkar S.V. and Varticovski L. (2008). *Brca1* breast tumours contain distinct CD44⁺/CD24⁻ and CD133⁺ cells with cancer stem cell characteristics. *Breast Cancer Research* 10:R10.
- Xu Y. and Lindquist S. (1993). Heat-shock protein hsp90 governs the activity of pp60^{v-src} kinase. *Proceedings of the National Academy of Science* 90, 7074 – 7078.
- Yeom Y., Fuhrmann G., Ovitt C.E., Brehm A., Ohbo K., Gross M., Hubner K. and Scholer H.R. (1996). Germline regulatory element of Oct 4 specific for the totipotent cycle of embryonal cells. *Development* (122) 881 – 894.
- Yuan J.S., Reed A., Chen F. and Stewart C.N. (2006). Statistical analysis of real-time PCR data. *BMC Bioinformatics* 7:85.
- Zhang Y., Yang Z., Cao Y., Zhang S., Li S., Huang Y., Ding Y-Q. and Liu X. (2008). The Hsp40 family chaperone protein DNAJB6 enhances Schlafen1 nuclear localization which is critical for promotion of cell-cycle arrest in T-cells. *Biochemistry Journal* 413, 239 – 250.
- Zhou J., Wulfeuhle J., Zhang H., Gu P., Yang Y., Deng J., Margolick J.B., Liotta L.A. Petricoin III E. and Zhang Y. (2007). Activation of the PTEN/mTOR/STAT3 pathway in breast cancer stem-like cells is required for viability and maintenance. *Proceedings of the National Academy of Science* 104, (104) 16158 – 16163.
- Zhu H., Mitsuhashi N., Klein A., Barsky L.W., Weinberg K., Barr M.L., Demetriou A. and Wu G.D. (2006). The role of the hyaluronan receptor CD44 in mesenchymal stem cell migration in the extracellular matrix. *Stem Cells* 24, 928 – 935.

Chapter 8 : Appendix

Appendix Table 1| Antibody and suppliers used in this study.

Type	Antibody	Species	Cat. No.	Supplier
CSC	CD44	Rat IgG2a k	17-0441	eBiosciences
	CD24	Mouse IgG1	11-0247	eBiosciences
	EpCAM	Mouse IgG1	12-9326	eBiosciences
Isotype Control	Rat APC	Rat IgG2a k	17-4321	eBiosciences
	Mouse FITC	Mouse IgG1	11-4714	eBiosciences
	Mouse PE	Mouse IgG1	12-4714	eBiosciences
Chaperones	HSP90AA1/AB1	Mouse IgG2a	sc-13119	Santa Cruz Biotech
	HSPA1	Mouse IgG2a	sc-24	Santa Cruz Biotech
	DNAJB1	Rabbit IgG	ab 69402	Abcam
	DNAJB6	Rabbit IgG	ab 96539	Abcam
	DNAJC6	Rabbit IgG	ab 103 321	Abcam
	BAG1	Mouse IgG2a	ab 54817	Abcam
	HSPB1	Mouse IgG2a k	smc-161B	StressMarq
Other	c-Myc	Mouse IgG1	sc-40	Santa Cruz Biotech
	Sox-2	Mouse IgG1	S1451	Sigma-Aldrich
	POU5F1 (Oct3/4)	Rabbit IgG	P0065	Sigma-Aldrich
	Actin	Rabbit IgG	A2103	Sigma-Aldrich
	Histone 3	Rabbit IgG	9715	Cell Signaling
	β-catenin	Rabbit IgG	sc-7199	Santa Cruz Biotech
	Donkey α-Mouse 488	Mouse IgG	A21202	Invitrogen
HRP	Goat α-Rat	Rat IgG	sc-2032	Santa Cruz Biotech
	Donkey α-Mouse	Mouse IgG	sc-2314	Santa Cruz Biotech
	Donkey α-Rabbit	Rabbit IgG	ab 6802	Abcam