ABERRATIONS OF *DPPA3* (*STELLA*), *EDR1* (*PHC1*), *GDF3*, AND *NANOG*, PUTATIVE STEM CELL GENES ON CHROMOSOME 12, IN BREAST CARCINOMA

Thesis submitted for the degree of

Doctor of Philosophy

at the University of Leicester

by

Chetana Ruangpratheep

Department of Cancer Studies and Molecular Medicine

University of Leicester

December 2011

Chetana Ruangpratheep

Aberrations of *DPPA3* (*STELLA*), *EDR1* (*PHC1*), *GDF3*, and *NANOG*, putative stem cell-associated genes on chromosome 12, in breast carcinoma

ABSTRACT

INTRODUCTION: Chromosome 12p13 has been reported to show gain/amplification in some breast cancers. This region contains putative stem cell-associated genes: *DPPA3*, *EDR1*, *GDF3*, and *NANOG*, but these genes have not been investigated in breast previously. Hence, the aim of this thesis was to study their role in breast carcinomas.

MATERIALS AND METHODS: The mRNA expression was evaluated in normal and malignant breast tissues using TaqMan® gene expression assays. Western blotting and immunohistochemistry were used for determination of protein expression. Copy number variations (CNVs) were assessed by TaqMan® copy number assays (CNAs) and Affymetrix® genome-wide human single nucleotide polymorphism (SNP) array 6.0.

RESULTS: Expression of *DPPA3*, *EDR1*, and *NANOG* was undetectable in normal breast tissue, but there was variable expression in breast carcinomas (BC) where expression of these genes tended to be higher in surrounding normal breast tissue. *GDF3* was not expressed in BC. At the 95% confidence interval, higher expression of *DPPA3* in BC was related to axillary lymph node metastasis; lower expression of *DPPA3*, *EDR1*, and *NANOG* correlated with high grade; and lower expression of *NANOG* was found in tumours of size > 2.0 cm. Both TaqMan® CNAs targeting each gene individually and SNP 6.0 genome wide array revealed complicated patterns of CNVs for these genes. The majority of BC had gain of *DPPA3* but loss (deletion) of *EDR1* and *NANOG*. However, there was no significant correlation between CNVs and either mRNA expression or protein expression.

CONCLUSION: Variable aberrations in copy number and expression of *DPPA3*, *EDR1*, and *NANOG* genes in the chromosome 12p13 region are associated with aggressive characteristics of breast carcinomas.

KEYWORDS: Breast carcinoma; Chromosome 12; Putative stem cell-associated genes

ACKNOWLEDGEMENTS

This PhD thesis would not have been possible unless I was supported by the following:

- ♥ The Ministry of Science and Technology of the Royal Thai Government for a full scholarship to study abroad in Master's and PhD degrees;
- ♥ Department of Pathology, Phramongkutklao College of Medicine, Thailand for an offer of this scholarship to me;
- ♥ Dr. Jacqueline Shaw and Professor Rosemary Walker, my great enthusiastic supervisors, at Department of Cancer Studies and Molecular Medicine, the University of Leicester for my wonderful and excellent experience of PhD study based on their expertise and motivational recommendations;
- ♥ Dr. James Howard Pringle who is the first person opened the gate of my fantastic study at University of Leicester since I did Master of Science (MSc) in Molecular Pathology and Toxicology at this University in 2006;
- ♥ All friendly laboratory technicians/scientists at this Department, including Ms. Karen Kulbicki prepared good breast tissue sections for all analyses in this thesis; and Ms. Linda Potter, Ms. Angie Gillies, and Ms. Lindsay Primrose assisted me in learning and doing all laboratory procedures which are used in this thesis;
- ♥ Dr. Karen Page for inspirational discussions with me about the real-time polymerase chain reaction (PCR);

- ♥ Mr. Kevin Blighe, one of my friends who is doing his PhD in the same research group, for kindly and helpful bioinformatic analysis of Affymetrix® Genome-Wide Human SNP Array 6.0 data on my 10 breast carcinoma samples;
- ♥ My parents for their great encouragement that I could finally achieve my PhD study;
- ♥ All my Thai and non-Thai friends who I met in Leicester, including Mrs. Duangduan Siriwittayawan, Dr. Danaya Holloway PhD, Miss Naowarat Keawseangtham, Miss Naparat Prueksuralai, Miss Teonchit Nuamchit, Mr. Ali Alharth, and Mr. Eftychios Papadogeorgakis for plenty of impressive and memorable academic and non-academic issues whilst I was doing PhD in this University.

Lastly, I owe my deepest gratitude to the Ministry of Science and Technology of the Royal Thai Government, the University of Leicester, and the United Kingdom for all the worth experience of my first overseas post-graduate study. Thank you.

Chetana Ruangpratheep

December 2011

TABLE OF CONTENTS

1. Page, K., Hava, N., Ward, B., Brown, J., Guttery, D.S., **Ruangpratheep, C.**, Blighe, K., Sharma, A., Walker, R.A., Coombes, R.C. & Shaw, J.A. 2011, "Detection of HER2 amplification in circulating free DNA in patients with breast cancer", *British journal of cancer,* vol. 104, no. 8, pp. 1342-1348.

2. Shaw, J.A., Page, K., Blighe, K., Hava, N., Guttery, D., Ward, B., Brown, J., **Ruangpratheep, C.**, Stebbing, J., Payne, R., Palmieri, C., Cleator, S., Walker, R.A. & Coombes, R.C. 2011, "Genomic analysis of circulating cell free DNA infers breast cancer dormancy", *Genome research.*

LIST OF FIGURES

CHAPTER 1 INTRODUCTION

CHAPTER 2 MATERIALS AND METHODS

CHAPTER 3 EXPRESSION OF DPPA3 (STELLA), EDR1 (PHC1), GDF3, AND NANOG mRNA AND PROTEINS IN BREAST CARCINOMA

CHAPTER 4 COPY NUMBER VARIATIONS (CNVs) OF *DPPA3* (*STELLA*), *EDR1* (*PHC1*), AND *NANOG* GENES IN BREAST CARCINOMA

LIST OF TABLES

CHAPTER 1 INTRODUCTION

CHAPTER 2 MATERIALS AND METHODS

CHAPTER 3 EXPRESSION OF DPPA3 (STELLA), EDR1 (PHC1), GDF3, AND NANOG mRNA AND PROTEINS IN BREAST CARCINOMA

CHAPTER 4 COPY NUMBER VARIATIONS (CNVs) OF *DPPA3* (*STELLA*), *EDR1* (*PHC1*), AND *NANOG* GENES IN BREAST CARCINOMA

ABBREVIATIONS

A	Adenine
ACC	Adenoid cystic carcinoma
aCGH	Array comparative genomic hybridisation
ADH	Atypical ductal hyperplasia
AIs	Aromatase inhibitors
AJ	Alex Jeffreys buffer
AJCC	American Joint Committee on Cancer
ALDH1	Aldehyde dehydrogenase 1
ALH	Atypical lobular hyperplasia
AMV	Avian Myeloblastoma Virus
ANOVA	Analysis of variance
APoC	Apocrine carcinoma
APS	Ammonium persulfate
AR	Antigen retrieval
array CGH	Microarray-based comparative genomic hybridization
AS	Alternative splicing of pre-mRNA
b	The y-intercept of standard curve
BC	Breast carcinoma
BCA	Bicinchoninic acid
BCA CL	Breast cancer cell lines
BMI	Body mass index
BSA	Bovine serum albumin

C	Cytosine
CA-125	Cancer antigen-125
CCL	Columnar cell lesion
CDH1	E-cadherin
CEA	Carcinoembryonic antigen
cDNA	Complementary deoxyribonucleic acid
CGH	Comparative genomic hybridization
CI	Confidence interval
CIS	Carcinoma in situ component
CK	Cytokeratin
CLDN	Claudin
CN	Copy number
CNAs	Copy Number Assays
CNVs	Copy number variations
CSC	Cancer stem cell
C_t	Threshold cycle
Cyt	Cytoplasmic immunostaining
Cyt (& Nuc)	Cytoplasmic and occasionally nuclear immunostainings
Da	Dalton
DAB	3,3'-diaminobenzidine
DCIS	Ductal carcinoma in situ
DCIS & IDC-NST	Combined ductal carcinoma in situ and invasive (infiltrating) ductal carcinoma of no special type
DEPC	Diethylpyrocarbonate
df	Degrees of freedom

DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
DPPA3	*Developmental Pluripotency-associated Gene 3* gene
ΔC_t	Delta threshold cycle
$\Delta\Delta C_t$	Delta delta threshold cycle
DMEM	Dulbecco's modified Eagle's medium
DMSO	Dimethyl sulfoxide
dNTPs	Deoxyribonucleotide triphosphates
DPBS	Dulbecco's phosphate buffered saline
E	The efficiency of real-time polymerase chain reaction assay
EDR1	*Early Development Regulator 1* gene
EDTA	Ethylenediamine tetraacetic acid
EGFR	Epidermal growth factor receptor
ER	Estrogen receptor
ERBB2	See HER2
ERD	Estrogen receptor downregulator
ESA	Epithelial-specific antigen
FBS	Foetal bovine serum
FEA	Flat epithelial atypia
FFPE	Formalin-fixed, paraffin-embedded
FISH	Fluorescence in situ hybridisation
G	Guanine
GA	Gestational age
GAPDH	*Glyceraldehyde-3-Phosphate Dehydrogenase* gene
GDF3	*Growth Differentiation Factor 3* gene

gDNA	Genomic deoxyribonucleic acid
GnRH	Gonadotropin-releasing hormone
H_2O_2	Hydrogen peroxide
H & E	Hematoxylin and Eosin staining
HER2 (ERBB2)	Human epidermal growth factor receptor 2
HGDNA	Normal human genomic deoxyribonucleic acid
HGNC	HUGO Gene Nomenclature Committee
HIER	Heat-induced epitope retrieval
HMECs	Human mammary epithelial cells
HPF	High-power field
HPRT1	*Hypoxanthine Phosphoribosyltransferase 1* (*Lesch-Nyhan syndrome*) gene
HRP	Horseradish peroxidase
HRT	Hormone replacement therapy
H-score	Histochemical scoring
HUT	Hyperplasia of usual type
IAA	Isoamyl alcohol
IC	Invasive (infiltrating) carcinoma
ICa	Invasive (infiltrating) carcinoma component
IDC-NST	Invasive (infiltrating) ductal carcinoma of no special type
IDC-OCGC	Invasive (infiltrating) ductal carcinoma with osteoclast-like giant cells
Ig	Immunoglobulin
IHC	Immunohistochemistry
ILC	Invasive (infiltrating) lobular carcinoma

IMS	Industrial Methylated Spirit
K	Keratin
LCIS	Lobular carcinoma in situ
LH	Luteinising hormone
LHRH	Luteinising hormone-releasing hormone
LN	Lobular neoplasia
LOH	Loss of heterozygosity
m	The slope of standard curve
MaSc	Mammary stem cell
MC	Mucinous carcinoma
MDC	Medullary carcinoma
MGA	Microglandular adenosis
MGCT CL	Mixed germ cell tumour cell line
MiPC	Micropapillary carcinoma
MPC	Metaplastic carcinoma
mRNA	Messenger ribonucleic acid
MW	Microwave
N_2	Nitrogen
NA	No available tissue
NB	Normal female breast tissue
NCBI	National Centre for Biotechnology Information
NCHG	Nottingham Combined Histologic Grade
NEC	Neuroendocrine carcinoma
NI	No information
NME	Normal mammary epithelial cell component

NOD	Non-obese diabetic
NOS	Not otherwise specified
NPA	No primary antibody
ns	No statistical significance
NST	No special type
NTC	No template (target) control
Nuc	Nuclear immunostaining
Nuc (& Cyt)	Nuclear and occasionally cytoplasmic immunostainings
Nuc & Cyt	Nuclear and cytoplasmic immunostainings
OCGC	Osteoclast-like giant cell
PAGE	Polyacrylamide gel electrophoresis
PC	Pressure cooker
PCR	Polymerase chain reaction
PHC1	*Polyhomeotic Homolog 1* (*Drosophila*) gene
PLCIS	Pleomorphic lobular carcinoma in situ
PR	Progesterone receptor
PROM 1	Prominin 1
PSA	Prostate-specific antigen
PTMs	Post-translational modifications of protein
PVDF	Polyvinylidene fluoride
QPCR	Quantitative polymerase chain reaction
QRT-PCR	Quantitative reverse transcriptase polymerase chain reaction
r	Correlation coefficient
r^2	Coefficient of determination
RANK	Receptor activator of nuclear factor kappa-B

REC	Research Ethics Committee
RefSeq	Reference Sequences
RNA	Ribonucleic acid
[RNA]	Concentration of RNA
RNase P	*Ribonuclease P* gene
rpm	Revolutions per minute
RPMI	Roswell Park Memorial Institute
RQ	Relative quantification
RS	Radial scar
RT	Reverse transcriptase
RT-PCR	Reverse transcriptase polymerase chain reaction
SA	Sclerosing adenosis
SCID	Severe combined immunodeficient
SD	Standard deviation
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate-Polyacrylamide gel electrophoresis
Sec	Second
SERMs	Selective estrogen receptor modulators
SNB	Surrounding normal breast tissue paired with breast carcinoma
SNP	Single Nucleotide Polymorphism
SP	Setting program
STAT1	Signal transducer and activator of transcription 1
T	Thymine

TAE	Tris-Acetate-Ethylenediamine tetraacetic acid
TBS	Tris buffered saline
TBS-T	Tris buffered saline in 0.1% Tween® 20
TC	Tubular carcinoma
TDLU	Terminal duct lobular unit
TEMED	N,N,N´,N´-tetramethylethylenediamine
TERT	*Telomerase reverse transcriptase* gene
TFRC	*Transferring Receptor* gene
TICs	Tumor initiating cells
TNM	Tumour, Node, and Metastasis
TOPO2A	Topoisomerase II alpha
TPMT	Thiopurine methyl-transferase gene
UK	The United Kingdom
USA	The United States of America
UV	Ultraviolet
V	Volt
VEGF	Vascular endothelial growth factor
VI	Vascular invasion component
WB	Western blotting
WGA	Whole genome amplification
WR	Working reagent

CHAPTER 1

INTRODUCTION

1.1. *THE FEMALE BREAST*

The adult female breast is a cone-shaped formation and is composed of skin, subcutaneous tissue, and mammary tissue. It lies on the pectoralis major muscle of the anterior thoracic wall. The mammary tissue is divided into 15 – 25 ill-defined lobes which are separated and supported by the suspensory ligaments of Cooper. Each lobe is divided further into 20 – 40 lobules (alveoli) which subdivide into many functional units called acini producing milk. The lobes, lobules, and acini are all connected by thin tubes called ducts. Hence, the terminal duct lobular unit (TDLU) of the breast lobules is formed from a grapelike cluster of small acini together with the terminal ducts. The units drain into a series of ducts and subsequently into the major lactiferous (single interlobular collecting) ducts. Each of the lactiferous ducts form a subareolar dilatation called the lactiferous sinus reaching the nipple. The lumens of both ducts and lobular units are lined by a continuous single layer of cuboidal to low columnar epithelial cells, which are surrounded by a low, flattened discontinuous layer of myoepithelial cells. These lie on the basement membrane and are contractile cells containing myofilaments. The contraction of myoepithelium leads to milk ejection during lactation (Lester 2005, Pandya, Moore 2011, Robinson, Huether 2002, Sharkey, Allred & Valente 1996). The myoepithelial cells also have the other functions such as regulators for integration of multiple signalling from adjacent cells and parenchyma; the maintenance of luminal epithelial differentiation and polarisation; and natural tumour suppressor by inhibition of tumour cell growth and angiogenesis (Adriance et al. 2005, Deugnier et al. 2002,

Gudjonsson et al. 2005, Gudjonsson et al. 2002). The breast stroma is dense fibrous connective tissue mixed with interlobular adipose tissue (Lester 2005, Pandya, Moore 2011, Robinson, Huether 2002, Sharkey, Allred & Valente 1996). The upper outer quadrant of the breast usually contains the largest proportion of the TDLU. In addition, the epithelial cells of the outer quadrant have been reported as showing a higher level of genetic instability, which may relate to the development of breast cancer commonly found in this area (Ellsworth et al. 2004, Hutson, Cowen & Bird 1985, Lee 2005).

During reproductive life, the mammary epithelium and stroma display a physiologic response to the levels of estrogen and progesterone associated with the menstrual cycle (Lester 2005, Longacre, Bartow 1986, Robinson, Huether 2002, Sharkey, Allred & Valente 1996). A small fraction of phenotypically and behaviourally progenitor (committed adult stem) cells in the TDLU are presumed to regenerate adult breast epithelium as part of the cyclic changes (Bocker et al. 2002, Boecker, Buerger 2003, Bombonati, Sgroi 2011, Dontu et al. 2003a, Petersen, Polyak 2010, Stingl et al. 2005, Villadsen et al. 2007). With the approach of the menopause, the breast starts to change with smaller ducts, decreased size and number of the lobules, and increased interlobular adipose tissue (Lester 2005).

1.2. *EPIDEMIOLOGY OF BREAST CANCER*

Breast cancer is the most common non-skin malignant neoplasm in females (Lester 2005) worldwide, including both economically developed and developing countries. Moreover, globally it is the leading cause of cancer deaths among females (Jemal et al. 2011). The most common clinical presentation of breast cancer, including non-invasive and invasive types, is a painless palpable lump (Bhattacharya, Adhikary 2006, Mathis et al. 2010), particularly in the upper outer quadrant of the breast (Bhattacharya, Adhikary 2006, Lee 2005, MacLean 2004, Saber 2000, Sohn et al. 2008). Several personal and environmental factors are related to the development of breast cancer. These factors can be categorised according to strength of evidence as convincing (well established) (Table 1-1), possible (less certain) (Table 1-2), and indecisive (Table 1-3) (http://info.cancerresearchuk.org/cancerstats/types/breast/riskfactors/#source1, Amir et al. 2010, Evans, Howell 2007, Iwasaki, Tsugane 2011, Lee, Park & Park 2008, Parsa, Parsa 2009, Trichopoulos et al. 2008, Warren, Devine 2004).

Table 1-1 Convincing (well established) risk factors for breast cancer (Adapted from http://info.cancerresearchuk.org/cancerstats/types/breast/riskfactors/#source1, Amir et al. 2010, Evans, Howell 2007, Iwasaki, Tsugane 2011, Lee, Park & Park 2008, Parsa, Parsa 2009, Trichopoulos et al. 2008, Warren, Devine 2004)

Convincing (well established) risk factor	High risk	Low risk
1. Age	> 50 years old	< 50 years old
2. Place of residence	North America, Northern Europe	Africa, Asia
3. Reproductive factors		
3.1. Menstrual factors		
3.1.1. Age at menarche	< 14 years old	≥ 16 years old
3.1.2. Age at menopause	≥ 54 years old	< 48 years old
3.1.3. Artificial menopause by bilateral oophorectomy	No removal	Pre-menopausal removal
3.2. Pregnancy factors		
3.2.1. Age at first full-term birth	≥ 30 years old	< 22 years old
3.2.2. Parity	Nulliparity	To have at least one child
3.3. Hormonal factors		
3.3.1. Endogenous estrogen	High levels	
3.3.2. Exogenous hormones		
3.3.2.1. Oral contraceptive pills (OCPs)	Use	No use
3.3.2.2. Hormone replacement therapy (HRT)	Long-term Use (> 5 years)	No use
3.4. Breast feeding	Never	Long term
4. Mammary epithelial cells from breast biopsy	Atypical epithelial hyperplasia	No epithelial hyperplasia
5. Family history of breast cancer	At least one first-degree relatives having breast cancer	No first-degree relatives having breast cancer
6. *BRCA1* and *BRCA2* genes	Mutation	No mutation
7. Non-reproductive lifestyle factors		
7.1. Height	≥ 160 cm	≤ 148 cm
7.2. Body mass index (BMI)		
7.2.1. Pre-menopausal women	22.5 - 24.9 kg/m^2 (Lean)	≥ 30 kg/m^2 (Obese)
7.2.2. Post-menopausal women	≥ 30 kg/m^2 (Obese)	< 19 kg/m^2 (Underweight)
7.3. Alcohol consumption	10 g increment of Ethanol/day (continuous) or >150 g of Ethanol/week	No drinking
7.4. Radiation exposure to the chest	High exposure	No exposure

Note: Menarche is the age of beginning regular menstruation.; Menopause is the age of ending menstruation.; Bilateral oophorectomy is the surgical removal of both ovaries.; Parity is the number of times having given birth.; Nulliparity is the state of never having given birth.; First-degree relatives are mother, father, daughters, sons, sisters, and brothers.

Table 1-2 Possible (less certain) risk factors for breast cancer (Adapted from http://info.cancerresearchuk.org/cancerstats/types/breast/riskfactors/#source1, Amir et al. 2010, Evans, Howell 2007, Iwasaki, Tsugane 2011, Lee, Park & Park 2008, Parsa, Parsa 2009, Trichopoulos et al. 2008, Warren, Devine 2004)

Possible (less certain) risk factor	High risk	Low risk
1. Breast density on mammograms	≥ 75% of Total breast area	≤ 10% of Total breast area
2. Non-reproductive lifestyle factors		
2.1. Socio-economic status	High status	Low status
2.2. Exercise	< 3 days/month	≥ 3 days/week
2.3. Diet		
2.3.1. Consumption of whole grain products	Small amounts	Large amounts
2.3.2. The level of folic acid in the diet of women who drink alcohol regularly	Low level	High level

Note: Mammographic density is the radiographic appearances of white and black areas associated with the connective and epithelial components of the breast on mammogram. Mammographic dense tissue is correlated with proliferation of either epithelium or stroma (Boyd et al. 2000).

Table 1-3 Indecisive risk factors for breast cancer (Adapted from http://info.cancerresearchuk.org/cancerstats/types/breast/riskfactors/#source1, Amir et al. 2010, Evans, Howell 2007, Iwasaki, Tsugane 2011, Lee, Park & Park 2008, Parsa, Parsa 2009, Trichopoulos et al. 2008, Warren, Devine 2004)

Indecisive risk factor	High risk	Low risk
1 Reproductive history		
1.1. Being twin	Non-identical twin	Identical twin
1.2. Mother had pre-eclampsia during woman's pregnancy	No	Yes
2. Non-reproductive lifestyle factors		
2.1. Diet		
2.1.1. Consumption of soy products containing Isoflavones	Low intake in Western	High intake in Asian
2.1.2. Consumption of dairy products	High intake	Low intake
2.1.3. Consumption of red meats or cured meats	High intake	Low intake
2.1.4. Consumption of well-done meat	High intake	Low intake
2.1.5. Consumption of fruits and vegetables	Low intake	High intake
2.1.6. Fibre in the diet	Small amounts	Large amounts
2.2. Working on the night shift	Yes	No
2.3. Smoking and passive smoking		
2.3.1. Cigarette smoking	Heavy smoking	No smoking
2.3.2. Exposure to tobacco smoke	High exposure	No exposure
2.4. Exposure to electromagnetic fields	High exposure	No exposure

Note: Red meats are beef, pork, and lamb.**;** Cured meats include hot dogs, sausage, salami, bacon, luncheon meat, and some types of ham.**;** Pre-eclampsia is a complication of pregnancy characterised by hypertension, edema, and proteinuria.

A proportion of breast cancers are inherited due to germline alterations of breast cancer susceptibility genes (Kenemans, Verstraeten & Verheijen 2004, van der Groep et al. 2006). Hence, breast cancer is generally characterised as "sporadic breast cancer" and "familial/hereditary breast cancer" (Berliner, Fay & Practice Issues Subcommittee of the National Society of Genetic Counselors' Familial Cancer Risk Counseling Special Interest Group 2007, Lynch et al. 1994, Lynch, Lynch 1996). It has been reported that approximately 75% of the development of breast cancers are correlated with environmental and lifestyle factors, whereas hereditary factors are involved in around 25% of breast cancers (Key, Verkasalo & Banks 2001, Lichtenstein et al. 2000). Most women who have an affected mother, sister, and/or daughter are unlikely to develop breast cancer (Collaborative Group on Hormonal Factors in Breast Cancer 2001, Key, Verkasalo & Banks 2001). Inheritance of mutations in *BRCA1* and *BRCA2* genes accounts for 2% - 6% of breast cancer overall (Esteves et al. 2009, Key, Verkasalo & Banks 2001, Papelard et al. 2000, Peto 2001, Schwartz et al. 2008, Wooster, Weber 2003). Nearly 90% of breast cancer patients have no family history of breast cancer in the first-degree relatives (Collaborative Group on Hormonal Factors in Breast Cancer 2001, Key, Verkasalo & Banks 2001). Hence, the majority of breast cancers are classed as sporadic and arise due to sporadic somatic mutations without germline mutations (Didraga et al. 2011, Joosse et al. 2009, Olopade et al. 2008).

1.3. *BREAST CARCINOGENESIS BASED ON HISTOPATHOLOGIC FEATURES*

There is evidence from retrospective studies of benign biopsies excised from women who subsequently developed breast cancer, and from the histological evaluation of changes present within breasts that have developed a cancer, that certain changes, particularly atypical hyperplasia, are associated with an increased risk of developing

breast cancer (Bombonati, Sgroi 2011, Briegel 2006, Kumar, Abbas & Fausto 2005a, Lester 2005, O'Shaughnessy et al. 2002, Recareanu et al. 2010, Tsubura et al. 2007, Weinberg 2007a) (Table 1-4). These epithelial changes are accompanied by genetic and epigenetic alterations; atypical lobular hyperplasia exhibits similar genetic alterations to lobular carcinoma in situ and invasive lobular carcinoma (Allred, Mohsin & Fuqua 2001, O'Malley 2010). Carcinoma in situ (CIS) (a pre-cancerous non-invasive breast lesion) forms when the mammary epithelium is totally replaced by atypical dysplastic cells (called severe dysplastic change), but these cells do not break through the basement membrane and invade underlying stromal tissues. Once the pre-cancerous cells move beyond the basement membrane, this lesion is considered to be invasive carcinoma (malignant tumour). The development of most breast cancers is generally a multi-step process and proceeds over long periods of time (Bombonati, Sgroi 2011, Briegel 2006, Kumar, Abbas & Fausto 2005b, Lester 2005, O'Shaughnessy et al. 2002, Recareanu et al. 2010, Tsubura et al. 2007, Weinberg 2007b).

The most common histopathological type of invasive carcinomas is invasive (infiltrating) ductal carcinoma (IDC) of no special type (NST) (Albrektsen, Heuch & Thoresen 2010, Chaiwun et al. 2010, Lester 2005, Roy, Othieno 2011, Weigelt et al. 2008). The other special types of invasive breast carcinomas are shown in Table 1-5. Invasive ductal and invasive lobular carcinomas are named by their characteristics of in situ component. Nevertheless, these descriptive terms do not relate to the site or cell type of origin because all carcinomas arise from the TDLUs of the mammary gland (Bombonati, Sgroi 2011, Lester 2005, Lopez-Garcia et al. 2010, Weigelt, Geyer & Reis-Filho 2010).

Table 1-4 Proliferative and carcinoma in situ (CIS) lesions of the breast and the relative risk of developing invasive carcinoma (Adapted from Lopez-Garcia et al. 2010)

Pathologic lesion		Relative risk
Proliferative	Apocrine metaplasia	Not available
	Columnar cell lesion	1.5 – 2.0
	Hyperplasia of usual type	1.2 – 2.0
	Sclerosing adenosis	1.5 – 2.0
	Radial scar	1.5 – 2.2
	Flat epithelial atypia	1.5 – 2.0
	Atypical ductal hyperplasia	3.0 – 5.0
	Atypical lobular hyperplasia	4.0 – 5.0
	Microglandular adenosis	Not available
CIS	Low grade ductal CIS	8.0 – 10.0
	High grade ductal CIS	8.0 – 10.0
	Classical lobular CIS	8.0 – 10.0
	Pleomorphic lobular CIS	Not available

Table 1-5 Frequency of histopathological types of invasive breast carcinoma (Adapted from Weigelt et al. 2008)

Histopathological type	Frequency
Invasive ductal carcinoma of no special type	50% – 80%
Invasive lobular carcinoma	5% – 15%
Medullary carcinoma	1% – 7%
Tubular carcinoma	1% – 6%
Mucinous carcinoma	< 5%
Metaplastic carcinoma	< 5%
Apocrine carcinoma	0.3% – 4%

1.4. *CHARACTERISTICS OF BREAST CANCER*

1.4.1. *Histological grade*

The differentiation of invasive breast carcinoma is assessed to provide information about prognosis using the Nottingham Combined Histologic Grade (NCHG) system (Elston-Ellis modification of Scarff-Bloom-Richardson grading system). The NCHG is based on the evaluation of 3 histopathological features: degree of tubule (gland) formation, nuclear pleomorphism, and mitotic count (Elston, Ellis 1991, Filho, Ignatiadis & Sotiriou 2011, Fitzgibbons et al. 2000, Lester 2005, McGuire et al. 1990, Meyer et al. 2005, Parham, Hagen & Brown 1992, Rakha et al. 2010, Robbins et al. 1995, Simpson et al. 2000; Weigelt, Geyer & Reis-Filho 2010) (Table 1-6). Patients with high-grade (grade III) breast cancer tend to have early recurrence and poor prognosis (Bundred 2001, Rakha et al. 2010, Schnitt 2001). However, the major problem of this histologic grading is morphological heterogeneity of the tumour leading to a lack of reproducibility and inter-observer variability, especially the degree of mitotic count and nuclear atypia (Filho, Ignatiadis & Sotiriou 2011, Komaki, Sano & Tangoku 2006).

Table 1-6 The Nottingham Combined Histologic Grade (NCHG) system (Elston-Ellis modification of Scarff-Bloom-Richardson grading system) of invasive breast carcinoma (Elston, Ellis 1991, McGuire et al. 1990, Meyer et al. 2005, Parham, Hagen & Brown 1992, Robbins et al. 1995, Simpson et al. 2000)

Tubule formation	
Score 1	> 75%
Score 2	10% – 75%
Score 3	< 10%
Nuclear pleomorphism	
Score 1	< 10 µm in greatest diameter, regular outlines, and uniform chromatin.
Score 2	10 – 13 µm in greatest diameter, variation in nuclear shape and size
Score 3	Marked nuclear variation
Mitotic counts/10 High power fields (HPFs)*	
Score 1	0 – 11 mitoses
Score 2	12 – 22 mitoses
Score 3	> 23 mitoses
Total score	**Grade/Differentiation[§]**
3 – 5	I/Well
6 – 7	II/Moderate
8 - 9	III/Poor

Note: * It is essential to calibrate the field diameter of the microscope before assessment of mitotic counts by using 40x objective (magnification x400) (NHS Cancer Screening Programmes jointly with The Royal College of Pathologists 2005).; § Grade I and II invasive breast cancers are generally categorised as low grade cancer (Morishita et al. 1997, Sato et al. 2011, Shet et al. 2007).

1.4.2. *Tumour size*

It has been previously reported that patients with increased size of breast cancer are related to low survival rate (Vorgias et al. 2001), particularly the risk of death in the first 5 years (Arriagada et al. 2006).

1.4.3. *The status of axillary lymph nodes*

In breast cancer, the status of axillary lymph node provides information about regional metastasis, guidance on adjuvant treatment, survival rate, and the risk of tumour recurrence (Hammer, Fanning & Crowe 2008, Nieweg et al. 2002). The presence of axillary lymph node metastasis is associated with large tumour size (Wasuthit et al. 2011) and is more likely to advance to distant metastasis (Comen, Norton & Massague 2011).

1.4.4. *Breast cancer staging*

After a definite diagnosis of breast cancer, the anatomical extent of cancer is evaluated within the breast or the other parts of the body. This process is called staging which is applied for planning the most appropriate treatment. The most common system used to describe the stages of breast cancer is the American Joint Committee on Cancer (AJCC) TNM system according to primary tumour size (T), lymph node status (N), and metastasis (M) (Wittekind, Greene, Hutter, Klimpfinger & Sobin 2008). The stage grouping is subsequently determined by a combination of the T, N, and M categories (http://www.cancer.gov/cancertopics/pdq/treatment/breast/healthprofessional/page3, Hammer, Fanning & Crowe 2008) (Table 1-7).

Table 1-7 TNM breast cancer staging (Adapted from Hammer, Fanning & Crowe 2008)

Stage		Primary tumour (T)	Regional lymph node metastasis (N)	Distant metastasis (M)	5-year survival rate
0		CIS	No	No	99%*
I		Microinvasion of ≤ 1 mm in greastest dimension	No	No	92%*
		Size ≤ 20 mm in greatest dimension	No	No	
II	**IIA**	No evidence of primary tumour	1 – 3 Nodes	No	82%*
		Size ≤ 20 mm in greatest dimension	1 – 3 Nodes	No	
		Size > 20 mm but ≤ 50 mm in greatest dimension	No	No	
	IIB	Size > 20 mm but ≤ 50 mm in greatest dimension	1 – 3 Nodes	No	65%*
		Size > 50 mm in greatest dimension	No	No	
III	**IIIA**	No evidence of primary tumour	4 – 10 Nodes	No	47%*
		Microinvasion of ≤ 1 mm in greastest dimension	4 – 10 Nodes	No	
		Size ≤ 20 mm in greatest dimension	4 – 10 Nodes	No	
		Size > 20 mm but ≤ 50 mm in greatest dimension	4 – 10 Nodes	No	
		Size > 50 mm in greatest dimension	1 – 3 Nodes	No	
		Size > 50 mm in greatest dimension	4 – 10 Nodes	No	
	IIIB	Any size with direct extension to chest wall or skin	No	No	44%*
		Any size with direct extension to chest wall or skin	1 – 3 Nodes	No	
		Any size with direct extension to chest wall or skin	4 – 10 Nodes	No	
	IIIC	Any tumour	> 10 Nodes	No	41.3%§
IV		Any tumour	Any lymph node	Yes	14%*

Note: * From Taneja et al. 2010 and § From Kuru 2011.

1.4.5. *Standard treatment for breast cancer*

The selection of the optimal treatment for breast cancer is usually considered from (I) the clinicopathological characteristics of the primary tumour, i.e. histopathological types, histological grade, tumour size, presence or absence of detectable tumour metastasis in axillary lymph nodes and/or the other sites, staging, expression of hormonal receptors [estrogen receptor (ER) and progesterone receptor (PR)] and human epidermal growth factor receptor 2 [HER2 (ERBB2)] protein in the cancer cells (See Section 1.5.3), and multi-gene testing; and (II) the patient's conditions, i.e. age, menopausal status, general health, and personal preference [http://cancerhelp.cancerresearchuk.org/type/breast-cancer/treatment/which-treatment for-breast-cancer, Breakthrough Breast Cancer 2009, National Comprehensive Cancer Network (NCCN) 2011 and 2012)]. Standard breast cancer treatment is generally classified into 2 main types: (I) local treatment, i.e. surgery and radiation therapy (radiotherapy); and (II) systemic treatment, i.e. chemotherapy, hormone (endocrine) therapy, and biological (targeted) therapy (http://www.breastcancer.org/treatment/ hormonal/, http://www.breastcancer.org/treatment/targeted_therapies/, http://www.

cancer.org/Cancer/BreastCancer/DetailedGuide/breast-cancer-treating-general-info, http://cancerhelp.cancerresearchuk.org/type/breast-cancer/treatment/which-treatment-for-breast-cancer, http://www.macmillan.org.uk/Cancerinformation/Cancertypes/Breast/

Treatingbreastcancer/Hormonaltherapies/Ovarianablation.aspx, http://www.nhs.uk/

Conditions/Cancer-of-the-breast-female/Pages/Treatment.aspx, American Cancer Society 2011, Breakthrough Breast Cancer 2009, Hammer, Fanning & Crowe 2008, National Comprehensive Cancer Network (NCCN) 2011 and 2012, Susan G. Komen for the Cure 2010 and 2011a and b).

1.4.5.1. Local treatment

Local treatment is aimed to treat the cancer at the breast and axillary (armpit) area without involvement in the other sites of the body (Table 1-8).

1.4.5.1 (1). Surgery

Surgery is mostly the first treatment for patients with breast cancer. The primary goals of breast cancer surgery are to remove the entire cancerous breast tissue or as much of the cancer as possible and to assess the accurate stage of breast cancer. In addition, the microscopic examination of at least one surgically removed regional lymph node from the axilla (armpit) will determine the presence or absence of axillary lymph node metastasis, the need for subsequent systemic therapy, the survival rate, and the risk of cancer recurrence. The surgical removal of breast cancer consists of 2 basic procedures, including lumpectomy (breast conserving surgery or wide local excision) and mastectomy.

1.4.5.1 (2). Radiation therapy (radiotherapy)

Radiation therapy (radiotherapy) could reduce the risk of local breast cancer recurrence.

Table 1-8 Local treatment for breast cancer

Local treatment	Description
1. Surgery	
1.1. Lumpectomy (breast conserving surgery or wide local excision)	The removal of only the breast tumour (lump) and a small amount of surrounding normal breast tissue
1.2. Mastectomy	The removal of either a large part of or the whole breast tissue, sometimes along with other nearby tissues
2. Radiation therapy (radiotherapy)	Treatment with high-energy rays or particles to destroy any remaining cancer cells in the targeted tissue after surgery

1.4.5.2. Systemic treatment

Systemic treatment is aimed to use oral or intravenous medications for eliminating the primary cancer and any cancer cells which may have migrated from the original site and subsequently located in the other parts of the body (Table 1-9). This treatment could be categorised according to the use as (I) neoadjuvant therapy is pre-operative systemic treatment for shrinking the tumour size, especially larger than 5 cm size (National Comprehensive Cancer Network (NCCN) 2011), and a subsequent allowance of the less extensive surgical removal of the cancerous tissue.; and (II) adjuvant therapy is post-operative systemic treatment based on the clinicopathological findings of individual cases for killing any remaining and/or migrating cancer cells.

Table 1-9 Systemic treatment for breast cancer

Systemic treatment	Description
1. Chemotherapy	Oral or intravenous medications using one or more anticancer (cytotoxic) drugs to destroy any cancer cells at the primary and/or metastatic sites.
2. Hormone (endocrine) therapy	Systemic treatment for estrogen receptor (ER)-positive breast cancer which is diagnosed by post-operative immunohistochemical study
2.1. Anti-estrogen drugs	For inhibition of the estrogen effect on breast cancer cell growth
2.1.1. Selective estrogen receptor modulators (SERMs)	For temporary attachment to the ER inside the breast cancer cells and prevent estrogen binding to its receptor ➢ Raloxifene (Evista®), Tamoxifen (Nolvadex, Istubal, and Valodex), and Toremifene (Fareston®)
2.2.1. Estrogen receptor downregulator (ERD)	• Working in a similar mechanism to SERMs • For decreasing the number of ERs and changing the shape of ER within the breast cancer cells leading to inefficient function of their receptors ➢ Fulvestrant (Faslodex®)

Table 1-9 (Continued) Systemic treatment for breast cancer

Systemic treatment	Description
2. Hormone (endocrine) therapy (Continued)	
2.2. Aromatase inhibitors (AIs)	Blocking the action of aromatase on the estrogen synthesis in post-menopausal women ➢ Competitive non-steroidal AI [Letrozole (Femara®)], Irreversible steroidal aromatase inactivator [Exemestane (Aromasin®)], and Selective non-steroidal AI [Anastrozole (Arimidex®)]
2.3. Ovarian shutdown	Treatment for breast cancer in pre-menopausal women by temporary or permanent shutdown of releasing estrogen from the ovaries
2.3.1. Ovarian suppression (temporary ovarian shutdown)	• The suppression of producing estrogen from the ovaries in pre-menopausal women with breast cancer by using drugs called gonadotropin-releasing hormone (GnRH) or luteinising hormone-releasing hormone (LHRH) analogues. These drugs block the anterior pituitary gland releasing the luteinising hormone (LH) which stimulates the ovaries producing estrogen. • The ovaries could start working again after no continuous medication. ➢ Goserelin (Zoladex®) and Leuprolide (Lupron®)/Leuprorelin (Prostap®)
2.3.2. Ovarian ablation (permanent ovarian shutdown)	The complete reduction in estrogen circulating in the body of pre-menopausal women with breast cancer by the surgical removal of the ovaries (oophorectomy) or by giving radiotherapy to the ovaries

Table 1-9 (Continued) Systemic treatment for breast cancer

Systemic treatment	Description
3. Biological (targeted) therapy	Treatment for stopping the growth and spread of the breast cancer cells by using monoclonal antibodies or enzyme inhibitor to the specific proteins in the cancer cells with minimal damage to normal cells and tissues
3.1. Monoclonal antibodies	Monoclonal antibodies for breast cancer treatment are synthesised in the laboratory.
3.1.1. Bevacizuma (Avastin®)	• Binding to the vascular endothelial growth factor (VEGF) protein in the cancer cells. This protein basically stimulates the formation of new blood vessels from pre-existing vessels, which is called angiogenesis, into the tumour mass. • Interfering with tumour angiogenesis that is the transportation of oxygen and nutrients to the growing and dividing cancer cells.
3.1.2. Trastuzumab (Herceptin®)	• Attachment to the growth-promoting protein called human epidermal growth factor receptor 2 (HER2) on the surface of the breast cancer cells. By blocking this protein, the cancer cells cannot grow. • Treatment for HER2-positive breast cancer which is diagnosed by post-operative immunohistochemistry and/or fluorescence in situ hybridisation (FISH)
3.2. Enzyme inhibitor [Lapatinib (Tykerb®)]	• Inhibition of the tyrosine kinase activity for epidermal growth factor receptor (EGFR) and HER2 proteins on the surface of the growing and dividing breast cancer cells. • Treatment for metastatic HER2-positive breast cancer that has stopped responding to Trastuzumab (Herceptin®).

1.4.5.2 (1). Chemotherapy

Basically, chemotherapeutic drugs enter blood circulation in the whole body and damage any cells having active growth and division. Therefore, the drugs affect normal body cells as well as cancer cells. This action of chemotherapy on normal cells leads to the development of short- and long-term side effects.

1.4.5.2 (2). Hormone (endocrine) therapy

In pre-menopausal women, estrogen hormone is mainly produced by the ovaries. In post-menopausal women, the ovaries stop producing estrogen, but androgen hormone produced by the adrenal glands is converted into small amounts of estrogen by aromatase enzyme in fat tissue. For breast cancer patients, estrogen in the body attaches to estrogen receptor (ER) within the cancer cells and the signals from the binding estrogen promote the growth and multiplication of breast cancer cells. The purposes of hormone therapy are to reduce the production of estrogen in the body and to prevent the attachment of estrogen to its receptor within any breast cancer cells remaining after surgical removal. Therefore, hormone therapy can reduce the risks of developing recurrent and metastatic tumours in early-stage ER-positive breast cancer and can also decrease the growth or the size of advanced-stage ER-positive breast cancer

1.4.5.2 (3). Biological (targeted) therapy

Biological (targeted) therapy is treatment for stopping the growth and spread of the breast cancer cells with minimal damage to normal cells and tissues by using monoclonal antibodies or enzyme inhibitor.

1.5. *BIOMARKERS (BIOLOGICAL MARKERS) IN CANCER*

1.5.1. *Characteristics of cancer biomarker*

A biomarker (biological marker) is defined as a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention (Biomarkers Definitions Working Group. 2001). Usually, biomarkers are detectable by various methods, including biochemical analysis of blood and/or tissue samples and biomedical imaging (Nass, Moses 2007). Since the development of cancers is associated with genetic and/or epigenetic changes resulting in alterations of protein expression in the affected cells, a cancer biomarker is either a molecule produced by the cancer cell or a specific response of the body to the presence of cancer. Thus, cancer biomarkers are present in the cancerous tissues and/or body fluids, including deoxyribonucleic acid (DNA), messenger ribonucleic acid (mRNA), transcription factors, cell surface receptors, secreted proteins, and small metabolites (Bensalah, Montorsi & Shariat 2007, Bhatt et al. 2010, Falasca 2012, Maruvada et al. 2005, Mishra, Verma 2010).

Based on clinical applications (Nass, Moses 2007, Bensalah, Montorsi & Shariat 2007, Falasca 2012, Mayeux 2004, Research Advocacy Network 2010), there are 2 main types of cancer biomarkers, including (I) biomarker of exposure (antecedent biomarker) – predisposition (risk assessment) biomarker; and (II) biomarkers of disease, i.e. screening, diagnostic, prognostic, predictive, pharmacological (including pharmacokenetics and pharmacodynamics), surrogate endpoint (efficacy-response or efficacy-surrogate), and recurrence biomarkers (Table 1-10).

Table 1-10 Types of cancer biomarkers based on clinical applications (Nass, Moses 2007, Bensalah, Montorsi & Shariat 2007, Falasca 2012, Mayeux 2004, Research Advocacy Network 2010)

Type of biomarker	Description
1. Biomarker of exposure (antecedent biomarker)	
❖ Predisposition (risk assessment)	For an evaluation of the likelihood of individually developing cancer
2. Biomarkers of disease	
2.1. Pre-diagnosis	
❖ Screening biomarker	For the detection of early-stage cancer in asymptomatic patients ➢ Serum prostate-specific antigen (PSA) in prostate cancer
2.2. Diagnosis	
❖ Diagnostic biomarker	For the definitive establishment of the presence of cancer ➢ Serum cancer antigen-125 (CA-125) in ovarian cancer
2.3. Post-diagnosis with treatment	
2.3.1. Prognostic biomarker	For the identification of the probable clinical outcome of cancer regardless of therapy and for the selection of appropriate treatment ➢ Serum carcinoembryonic antigen (CEA) in colorectal cancer
2.3.2. Predictive biomarker	For an evaluation of the most likely effectiveness of individual treatment ➢ Estrogen (ER), progesterone (PR), and HER2 receptors in breast cancer tissue
2.3.3. Pharmacological biomarkers	➢ Mutation in thiopurine methyl-transferase (*TPMT*) gene for mercaptopurine (a leukemic drug) metabolism
❖ Pharmacokinetic biomarker	For the determination of drug metabolism
❖ Pharmacodynamic biomarker	For the optimisation of dose response
2.3.4. Surrogate endpoint (efficacy-response or efficacy-surrogate) biomarker	For the early determination of the ultimate clinical outcome related to drug development ➢ Positron emission tomography (PET) scanning with the tracer fluorine-18 fluorodeoxyglucose (^{18}F-FDG) called ^{18}F-FDG PET for assessment of therapeutic response in non-small cell lung cancer (Frank, Hargreaves 2003, Nabi, Zubeldia 2002, van Loon et al. 2011, Zhao, Schwartz & Larson 2009)
2.4. Post-treatment	
❖ Recurrence biomarker	For an evaluation of the likelihood of or monitoring recurrent cancer after treatment ➢ Onco*type* DX® for stage I or II ER-positive breast cancer with no axillary lymph node metastasis)

1.5.2. *The development of cancer biomarkers*

In cancer, the development of any novel biomarkers is comprised of the following (Bensalah, Montorsi & Shariat 2007, Rifai, Gillette & Carr 2006, Verderio et al. 2010) as: Discovery phase – Preclinical testing is to identify candidate biomarkers in vitro or in animal models by a combination of current molecular technologies, i.e. genomics for DNA analysis, transcriptomics for analysis of mRNA expression in a cell or tissue at a given time, proteomics for measurement of protein expression in a cell or tissue at a given time, and metabolomics for measurement of entire metabolite profile in a cell or tissue and metabonomics for measurement of metabolite profile in a cell or tissue under given conditions (Nass, Moses 2007, Maruvada et al. 2005, Research Advocacy Network 2010, Ilyin, Belkowski & Plata-Salaman 2004, Ludwig, Weinstein 2005).; Phase 0 – Qualification and verification is (I) to confirm expression of the identified candidate biomarkers on healthy and patient samples using methods differed from Discovery assay (qualification) and (II) to preliminarily determine sensitivity (the ability to detect a cancer case having a positive biomarker result) and specificity (the ability to detect a healthy case having a negative biomarker result) of candidate biomarkers (verification).; Phase 1 – Optimisation is to precisely determine candidate biomarkers belonging to the cancer cases and to establish the assay cut-off points.; Phase 2 – Analytical validation is to evaluate the performance characteristics of the candidate biomarker assay, including reproducibility (precision) (the closeness of agreement among results of measurements performed under different conditions) and accuracy (the closeness of agreement between the value of a measurement and the true concentration of the quantity intended to be measured in that sample.; Phase 3 – Clinical validation is to assess the sensitivity and the specificity of candidate biomarkers on clinical practice in the large patient populations other than the samples in the Discovery

phase.; and Phase 4 – Commercialisation is post-approval reporting and testing of the validated cancer biomarkers for the industrialised clinical utility. However, in comparison with the currently used cancer biomarkers, the clinical implementation of the newly established biomarkers need to address the 4 following concepts: easier, better, faster, and cheaper (Bensalah, Montorsi & Shariat 2007).

1.5.3. *Predictive biomarkers for breast cancer*

Predictive biomarkers basically provide information as to whether a patient is likely to respond to a specific therapeutic intervention and are associated with sensitivity or resistance of tumour to that therapy. These biomarkers may be the target of a specific therapy itself (Weigel, Dowsett 2010). The current predictive biomarkers for breast cancer are generally determined by immunohistochemistry, although standardisation of tissue fixation, methodology of immunostaining, and interpretation of immunoreactivity is required (Walker 2008). The immunohistochemical biomarkers are categorised as: (I) Established markers for routine clinical use (see Section 1.5.1 – 1.5.2), i.e. estrogen receptor (ER), progesterone receptor (PR), and HER2 (ERBB2) receptor; (II) Potential markers for clinical use which need refinement of antibodies and scoring systems, i.e. Epidermal growth factor receptor (EGFR) for chemotherapeutic response and basal marker, Ki67 (MIB-1) for tumour cell proliferation, and Topoisomerase II alpha (TOPO2A) for response to Anthracycline chemotherapeutic drug; and (III) Research markers of interest which are less likely to be used clinically such as p53 protein for mutation analysis of *TP53* tumour suppressor gene; Bcl-2, bcl-x, bax, and survivin for apoptotic markers corresponding to chemotherapeutic response; and Cyclin D1, cyclin E, p21, and p27 for tumour cell proliferation markers (Bundred 2001, Duffy 2005, Morabito et al. 2003, Schnitt 2001, Taneja et al. 2010, Walker 2008,

Weigel, Dowsett 2010). In breast cancer, the ER is the most reliable biomarkers and the best example of a predictive markers for therapeutic response (Walker 2008).

1.5.3.1. The status of estrogen receptor (ER) and progesterone receptor (PR)

Expression of ER and PR in breast cancer tissue is immunohistochemically evaluated to aid prediction of response to hormone therapy (Duffy 2005, Walker 2008). The accurate interpretation of these hormonal receptors is related to optimal staining, quality assurance, and standard scoring (Walker 2008). Dunnwald and colleagues (2007) reported that 155,175 breast cancer patients with known the status of ER and PR showed 63% of ER-positive/PR-positive tumours, 21% of ER-negative/PR-negative tumours, 13% of ER-positive/PR-negative tumours, and 3% of ER-negative/PR-positive tumours. These tumours had various races/ethnics and clinical manifestations. Compared to patients with ER-positive/PR-positive tumours, higher mortality risk was observed in ER-positive/PR-negative, ER-negative/PR-positive, and ER-negative/PR-negative tumours (Dunnwald, Rossing & Li 2007).

1.5.3.2. The status of *HER2* (*ERBB2*) gene

The status of *HER2* gene in breast cancer tissue is currently determined by the analysis of protein expression by immunohistochemistry (IHC) and DNA copy number change by fluorescence in situ hybridization (FISH) (Taneja et al. 2010, Walker 2008, Wolff et al. 2007). As the immunostaining of ER and PR, HER2 immunoreactivity is affected by variations in laboratory techniques and intra- and inter-observer interpretation. The main problems in immunohistochemical assessment of HER2 expression are at the 1+ (negative)/2+ (equivocal) borderline and the 2+ (equivocal)/3+ (positive) borderline (Walker 2008). Approximately 15% - 30% of breast cancers have overexpression and/or amplification of this gene and are correlated with poor prognosis

(Daniele, Sapino 2009, Schmitt 2009). Patients with HER2-positive breast cancer receive Trastuzumab (Herceptin®), a humanised monoclonal antibody against HER2 protein, which can improve survival and time to progression (Burstein 2005, Daniele, Sapino 2009, Ferretti et al. 2010, Schmitt 2009).

1.6. *BREAST CARCINOGENESIS BASED ON MOLECULAR CHARACTERISTICS*

Molecular approaches using microarray-based gene expression profiling were introduced over 10 years ago in order to identify diagnostic, prognostic, and predictive information about breast cancer, over that which can be obtained from pathology and receptor analyses (Perou et al. 2000, Sorlie et al. 2001, van de Vijver et al. 2002). These approaches are now being used to examine cancers from different ethnic groups, breast cancer subgroups, and for the selection of therapy. Breast cancers have been subdivided into 5 main groups: Luminal A, Luminal B, HER2-enriched, Basal-like, and Normal breast-like cancers (Table 1-8). In addition, Molecular apocrine, Claudin-low, and Interferon-related groups are also recognised. Generally, the luminal A breast cancer is the most common molecular subtype (~ 45%) but normal breast-like tumour is not frequently observed. These breast cancer subtypes showed expression of ER, PR, and HER2 as: (I) Luminal A cancers are ER-positive and PR-positive but HER2-negative; (II) Luminal B cancers are ER-positive but PR-negative and HER2-negative; (III) HER2-enriched cancers are ER-negative and PR-negative and HER2-positive; (IV) Basal-like cancers lack expression of all ER, PR, and HER2; and (V) Normal breast-like cancers show similar expression of markers to the Luminal B subtype (Correa Geyer, Reis-Filho 2009, Reis-Filho, Pusztai 2011, Weigelt, Baehner & Reis-Filho 2010, Weigelt, Geyer & Reis-Filho 2010).

Table 1-11 The characteristics of molecular subgroups of breast cancer (Adapted from Correa Geyer, Reis-Filho 2009, Reis-Filho, Pusztai 2011, Weigelt, Baehner & Reis-Filho 2010, Weigelt, Geyer & Reis-Filho 2010)

			Molecular subgroups				
			Luminal A	**Luminal B**	**HER2-enriched**	**Basal-like**	**Normal breast-like**
Frequency	Calza et al. 2006 (N = 369)		33.06%	14.63%	11.65%	15.99%	24.66%
	Sihto et al. 2008 (N = 1,236)		68.30%	9.50%	9.70%	7.90%	
	Lin et al. 2009 (N = 1,028)		62.00%	9.00%	12.00%	13.00%	
	Al Tamimi et al. 2010 (N = 231)		3.90%	16.00%	17.30%	10.00%	
	Ben Abdelkrim et al. 2010 (N = 194)		51.50%	16.00%	14.50%	18.00%	
	Su et al. 2011 (N = 2,791)		48.60%	16.70%	13.70%	12.90%	
	Mean		**44.56%**	**13.64%**	**13.14%**	**12.97%**	
IHC	**ER**	**Positive**	91% - 100%	91% - 100%	29% - 59%	0% - 19%	44% - 100%
		Negative	0% - 9%	0% - 9%	41% - 71%	81% - 100%	0% - 56%
		Summary*	**Positive**	**Positive**	**Negative**	**Negative**	**Positive**
	PR	**Positive**	70% - 74%	41% - 53%	25% - 30%	6% - 13%	22% - 63%
		Negative	26% - 30%	47% - 59%	70% - 75%	87% - 94%	37% - 78%
		Summary*	**Positive**	**Negative**	**Negative**	**Negative**	**Negative**
	HER2	**Positive**	8% - 11%	15% - 24%	66% - 71%	9% - 13%	0% - 13%
		Negative	89% - 92%	76% - 85%	29% - 34%	87% - 91%	87% - 100%
		Summary*	**Negative**	**Negative**	**Positive**	**Negative**	**Negative**
	Ki67 (Cell proliferation marker)		Low	High	High	High	Low or Intermediate
	Basal markers [Cytokeratin (CK) 5/6 and EGFR]		Negative	Negative	Predominantly Negative	Positive	Predominantly Negative

Note: IHC = Immunohistochemistry; * The presumptive immunoexpression is summarily based on the frequency of negative and positive staining.

Table 1-11 (Continued) The characteristics of molecular subgroups of breast cancer (Adapted from Correa Geyer, Reis-Filho 2009, Reis-Filho, Pusztai 2011, Weigelt, Baehner & Reis-Filho 2010, Weigelt, Geyer & Reis-Filho 2010)

		Molecular subgroup				
		Luminal A	**Luminal B**	**HER2-enriched**	**Basal-like**	**Normal breast-like**
***TP53* mutation**		Low	Intermediate	High	High	Low
Other markers		*FOXA1*: High	*FGFR1*: Amplification *ZIC3*: Amplification	*GRB7*: High	*RB1*: Low or Negative *CDKN2A*: High *BRCA1*: Low or Negative *FGFR2*: Amplification	
Histological grade	**I and II (Low)**	70% - 87%	38% - 59%	11% - 45%	7% - 12%	37% - 80%
	III (High)	13% - 30%	41% - 62%	55% - 89%	88% - 93%	20% - 63%
	Summary♥	**I and II (Low)**	**III (High)**	**III (High)**	**III (High)**	**I and II (Low)**
Most suitable histopathological type		Cribiform carcinoma Classic ILC Tubular carcinoma	Micropapillary carcinoma	Apocrine carcinoma Classic ILC Micropapillary carcinoma Pleomorphic ILC	Acinic cell carcinoma Adenoid cystic carcinoma Medullary carcinoma Metaplastic carcinoma Pleomorphic ILC Secretory carcinoma	Medullary carcinoma Metaplastic carcinoma
Outcome		Good	Intermediate or Poor	Poor	Poor	Intermediate
Benefit from chemotherapy		Low (0% - 5% pCR)	Intermediate (10% - 20% pCR)	Intermediate (25% - 40% pCR)	High (≥ 40% pCR)	Low (0% - 5% pCR)

Note: pCR = Pathological complete response after neo-adjuvant chemotherapy; ♥ The presumptive histological grade is summarily based on the frequency of grade I and II (low grade) and grade III (high grade).

Based on estrogen receptor (ER) status and molecular characteristics, a more complex mechanism for the development of breast cancer (Lopez-Garcia et al. 2010) (Figure 1-1) has been proposed. The luminal breast cancer develops from the ER-positive precursor lesions, including columnar cell lesion (CCL), columnar cell hyperplasia (CCH), flat epithelial atypia (FEA), atypical ductal hyperplasia (ADH), lobular neoplasia [both atypical lobular hyperplasia (ALH) and lobular carcinoma in situ (LCIS)], and pleomorphic LCIS (PLCIS). Atypical apocrine hyperplasia (APH) and microglandular adenosis (MGA) are considered as the ER-negative non-obligate precursor lesions. These lesions can lead to the development of high-grade ductal carcinoma in situ (DCIS) and subsequent grade III (high-grade) invasive ductal carcinoma (IDC). Although genetic aberrations are more common in the high-grade cancer than in the low-grade cancer, gain of chromosome 16p and deletion of chromosome 16q are rare in high-grade cancer. Hence, the majority of high-grade DCIS is supposed to arise either de novo or from a precursor other than ADH/low-grade DCIS (Lopez-Garcia et al. 2010).

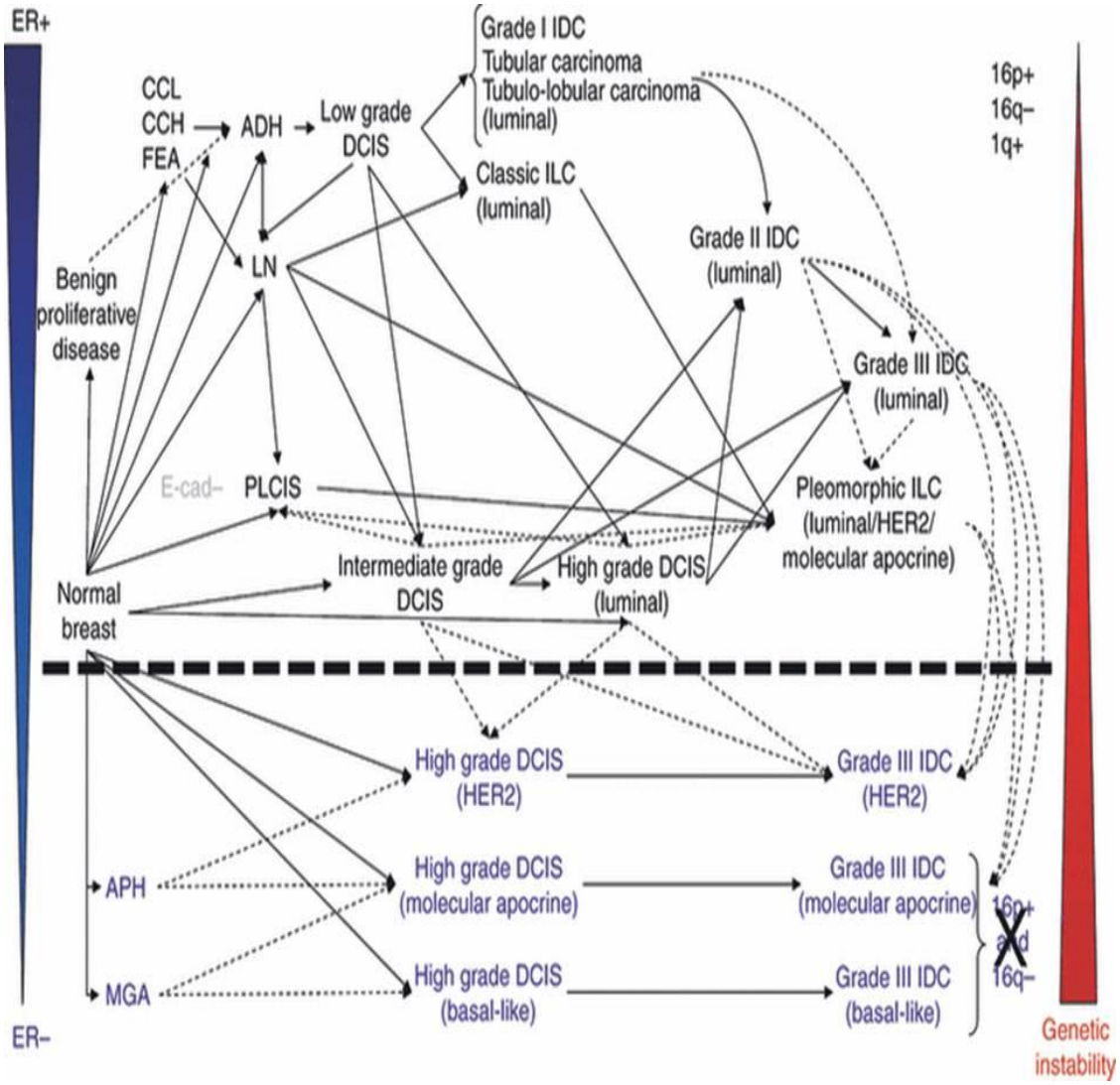


Figure 1-1 A molecular model of multi-step breast carcinogenesis (From Lopez-Garcia et al. 2010)**.** The putative mechanisms of breast cancer development from proliferative lesions to invasive carcinoma is based on expression of estrogen receptor (ER) and molecular subtypes. The levels of ER expression and genetic instability are shown in the left (Blue) and the right (Red) side of the figure, respectively. CCL = Columnar cell lesion; CCH = Columnar cell hyperplasia; FEA = Flat epithelial atypia; ADH = Atypical ductal hyperplasia; LN = Lobular neoplasia [both atypical lobular hyperplasia (ALH) and lobular carcinoma in situ (LCIS)]; PLCIS = Pleomorphic LCIS; APH = Atypical apocrine hyperplasia; MGA = Microglandular adenosis; DCIS = Ductal carcinoma in situ; IDC = Invasive ductal carcinoma; and E-cad = E-cadherin

1.7. *BREAST STEM CELLS AND CARCINOGENESIS*

Stem cells are cells with the ability to self-renew and to generate progeny that can differentiate into various cell types comprising each adult or mature tissue (Hart et al. 2004, Kuroda et al. 2005, Smalley, Ashworth 2003, Wicha, Liu & Dontu 2006). A small number of mammary stem cells (MaSC) are usually located in the suprabasal position of the terminal duct lobular units and are surrounded by proliferating progenitor cells. The MaSC differentiates into common (bipotent) progenitors which gives rise to committed luminal and myoepthelial progenitors. These committed progenitors subsequently differentiate into luminal epithelial and myoepithelial cells, respectively (Bombonati, Sgroi 2011, Lim et al. 2009, Polyak 2007, Prat, Perou 2009) (Figure 1-2). According to normal breast development, the adult mammary gland requires stem cells or a stem cell-like activity to proliferate mammary epithelial cells during pregnancy and lactation and replace lost cells due to routine cell turnover (Dontu et al. 2003b, Dontu, El-Ashry & Wicha 2004, Smalley, Ashworth 2003, Woodward et al. 2005, Tan et al. 2006). The number and activity of MaSc and occasional progenitor cells are controlled via paracrine signalling from the receptor activator of nuclear factor kappa-B (RANK) ligand released by adjacent progesterone receptor (PR)-positive luminal epithelial cells (Bombonati, Sgroi 2011, Gonzalez-Suarez et al. 2010) (Figure 1-3). The mammary stem cells can be separated and isolated from breast tissue by using in vitro cell culture and in vivo animal model.

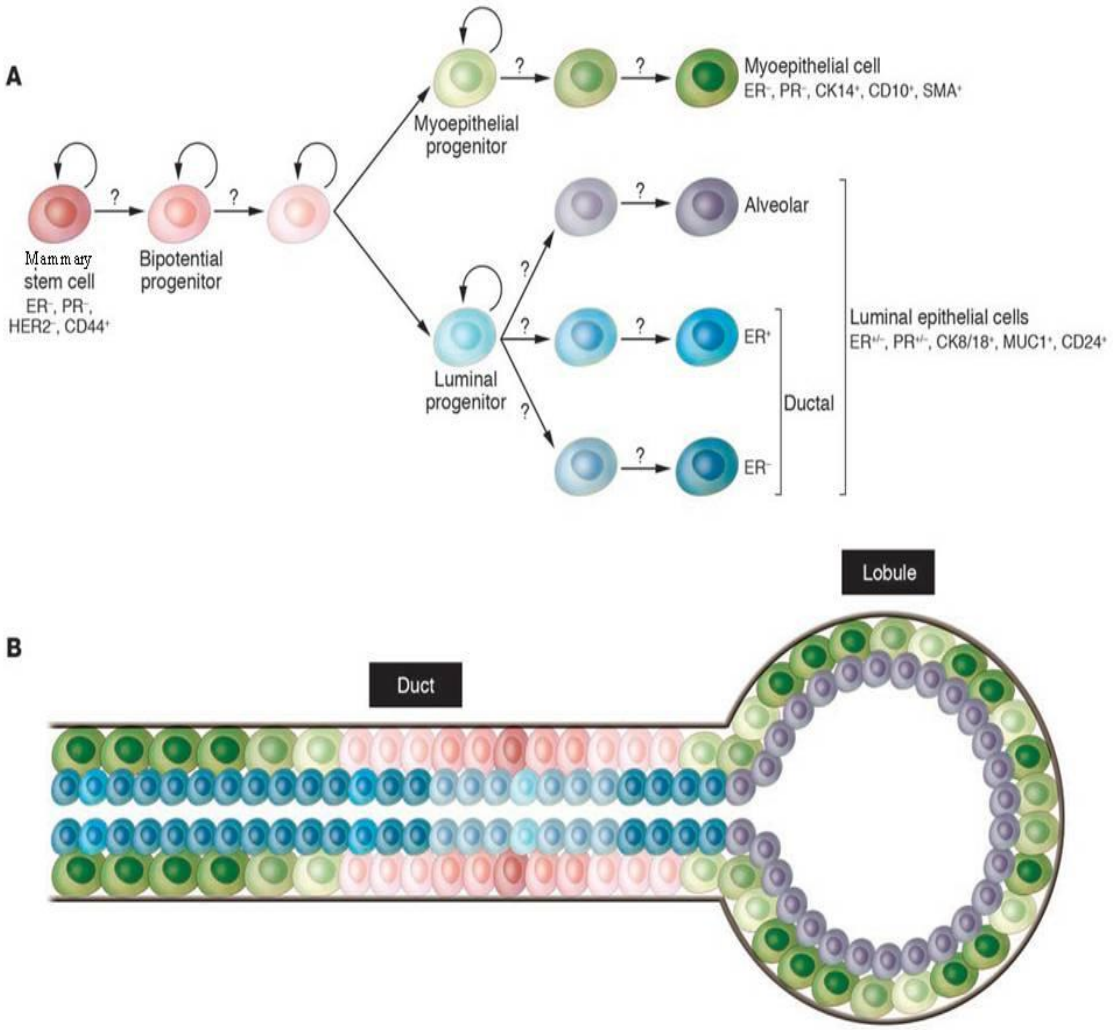


Figure 1-2 Hypothetical model of human mammary stem cell and its differentiation (From Polyak 2007)**.** (A) Mammary stem cell with intrinsic self-renewal potential (Curved black arrow) differentiates into the luminal epithelial and myoepithelial cells. However, the intermediary steps and their regulation are still unknown (Question mark). (B) Mammary stem cell and variously differentiated cells are located in the terminal duct lobular unit. These cells are lined by the basement membrane. The colour of cell types correlated with the cells in (A).

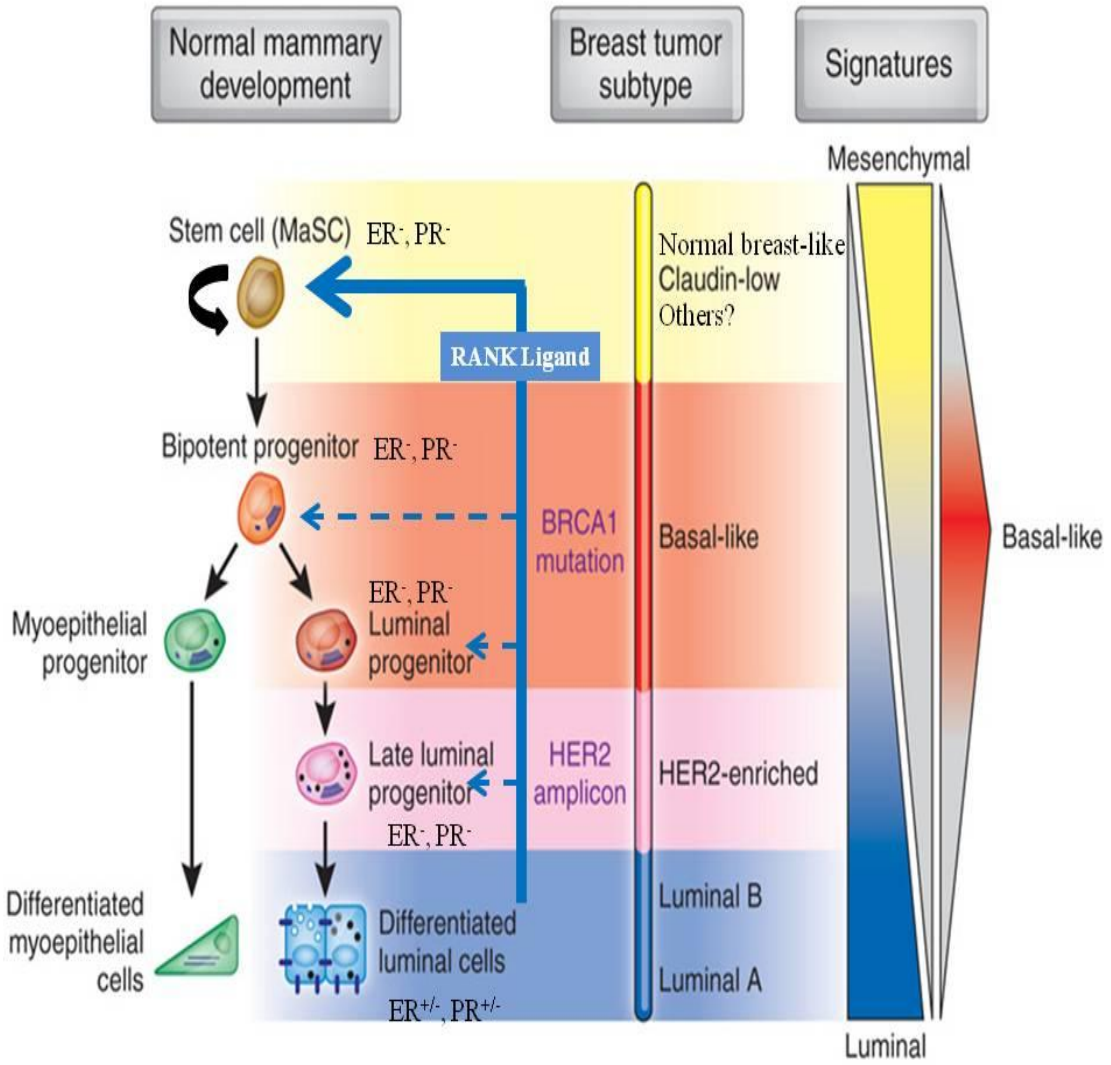


Figure 1-3 A model of human mammary epithelial hierarchy correlated with molecular subtypes of breast cancer (Adapted from Bombonati, Sgroi 2011, Lim et al. 2009, Prat, Perou 2009)**.** Each of the molecular phenotypes of invasive breast carcinoma is related to the development of normal breast epithelium from mammary stem cell (MaSC) to differentiated luminal epithelial cell. The self-renewal characteristic of MaSC is displayed by curved black arrow. During the reproductive cycle and pregnancy, the paracrine signalling of RANK ligand (Blue line) is secreted by neighbouring PR-positive mammary epithelial cells for proliferation of MaSC and occasional progenitors. Therefore, this hormonal control may play an additional role in breast carcinogenesis. The rightmost bars represent the level of gene expression patterns of luminal (Blue), mesenchymal or claudin-low (Yellow), and basal-like (Red) cells.

For in vitro cell culture, Dontu and colleagues have propagated the colonies of undifferentiated human mammary epithelial cells (HMECs) as non-adherent mammospheres, obtained by mechanical and enzymatic dissociation from reduction mammoplasties, based on their ability to proliferate in suspension. These mammospheres highly contain cells with self-renewal potential and are able to differentiate into all cell lineages of adult mammary gland, including the myoepithelial cell lineage and the luminal epithelial lineage (consists of ductal and secretory alveolar subtypes). Additionally, the mammosphere-derived cells are able to clonally generate the tree-like ductal-acinar structure of mammary gland. Transcriptional profiling of cells isolated from non-adherent mammospheres reveals significant overlapping genetic programs with other stem cells. Therefore, the isolation and characterisation of human mammary stem cells using this in vitro cultivation system could demonstrate a full explanation of the molecular pathways that govern normal development and carcinogenesis of the breast (Dontu et al. 2003).

For in vivo animal model, Al-Hajj and colleagues have reported the initial evidence of putative breast cancer stem cells. They were able to distinguish the tumorigenic (tumor-initiating) cancer cells from non-tumorigenic cancer cells by using a xenograft nonobese diabetic (NOD)/severe combined immunodeficient (SCID) mouse model for human breast cancer. The tumorigenic cells from primary tumor express CD^{44+} $CD^{24-/low}$ on their cell surface and have the ability to generate new tumor containing additional tumorigenic CD^{44+} $CD^{24-/low}$ cells with phenotypically diverse non-tumorigenic cells. Thus, the tumorigenic cancer cells seem to exhibit properties of stem cells, including self-renewal and differentiation (Al-Hajj et al. 2003).

In breast cancer, besides the cancer cells with stem cell phenotype having CD^{44+} $CD^{24-/low}$ expression (Charafe-Jauffret, Ginestier & Birnbaum 2009, Chen, Chen 2010,

Klonisch et al. 2008), they could be identified by the other stem cell markers, including aldehyde dehydrogenase 1 (ALDH1) (Ginestier et al. 2007), CD 133 [Prominin 1 (PROM1)] (Wright et al. 2008), epithelial-specific antigen (ESA) (Hwang-Verslues et al. 2009), p63 (Du et al. 2010, Hanker et al. 2010, McKeon 2004), PROCR (Hwang-Verslues et al. 2009), and SOX2 (Leis et al. 2012).

Since stem cells characteristically have a long life and a large potential for replication, they can accumulate multiple non-lethal DNA damage or mutations in normal regulatory genes correlating with the first step of carcinogenesis (Ashkenazi, Gentry & Jackson 2008, Dontu & Wicha 2005, Giuliano et al. 2005, Hart et al. 2004, Killeen 2004, Kumar, Abbas & Fausto 2005a and b, Liu, Dontu & Wicha 2005, Smalley, Ashworth 2003, Tai et al. 2005, Trosko et al. 2005, Weinberg RA 2007, Wicha, Liu & Dontu 2006,). Therefore, both stem and cancer cells have been shown to share common properties (Dontu et al. 2003b, Dontu, El-Ashry & Wicha 2004, Wicha, Liu & Dontu 2006), including: (a) self-renewal capacity; (b) differentiation capacity (Ashkenazi, Gentry & Jackson 2008, Charafe-Jauffret et al. 2008, Cho, Clarke 2008, Lynch, Cariati & Purushotham 2006, Woodward et al. 2005, Wu 2008); (c) immortality due to active telomerase expression and activation of anti-apoptotic pathways; (d) resistance to damaging agents; and (e) the capacity for migration and metastasis with independent survival (Charafe-Jauffret et al. 2008).

The different molecular subtypes of breast cancer may relate to the features of breast stem cells (Bombonati, Sgroi 2011, Lim et al. 2009, Melchor, Benitez 2008, Polyak 2007, Prat, Perou 2009). Genomic aberrations of mammary stem cells (MaSC) or each progenitor may result in the development of different molecular phenotypes of breast carcinoma (Figure 1-3). Normal breast-like and claudin-low subtypes are

considered to be derived from the estrogen receptor (ER)-negative MaSCs. Basal-like breast cancer could develop from either the ER-negative common (bipotent) or luminal progenitor cells with *BRCA1* mutation. The HER2-enriched subtype can arise from the ER-negative late luminal progenitor having HER2 amplification. The luminal A and B breast cancers originate from the differentiated luminal epithelial cells. For luminal A subtype, the tumour develops from the ER-positive epithelial cells, whereas luminal B cancer can arise from the ER-negative or –positive luminal epithelial cells.

1.8. *CHROMOSOME 12 ABERRATIONS AND BREAST CANCER*

Chromosome 12 aberrations have been demonstrated in a number of human cancers, including glioblastoma (Mischel et al. 2003), breast cancer (Bui et al. 1997, Engel et al. 1998, Farabegoli et al. 2001, Letessier et al. 2006, Mark et al. 1999, Schondorf et al. 1997, Yao et al. 2006), esophageal adenocarcinoma (Miller et al. 2003), lung cancer (Miller et al. 2003, Wikman et al. 2005), gastric cancer (Gorringe et al. 2005), pancreatic carcinoma (Heidenblad et al. 2002), testicular germ cell tumors (Bourdon et al. 2002, Goddard et al. 2007, Rodriguez et al. 2003, Zafarana et al. 2003), ovarian carcinoma (Bourdon et al. 2002), bone and soft tissue tumors (Chibon et al. 2004, Heidenblad et al. 2006, Man et al. 2004, Rozeman et al. 2006, Weng et al. 2004,), and diffuse large B-cell lymphoma (Chen et al. 2006). In breast cancer, amplification of chromosome 12p13 correlates with high-grade (poorly differentiated) tumours (Letessier et al. 2006). Many regions on chromosome 12 have been reported to show gene amplification in breast carcinomas by using microarray-based comparative genomic hybridization (array CGH), for example amplification of *H2AFJ* at 12p12 and *EPS8* at 12p13 (Yao et al. 2006), and these have been considered to be novel candidate

breast cancer oncogenes (Table 1-9). Data from the Wellcome Trust Sanger Institute website (http://www.sanger.ac.uk/cgi-bin/genetics/CGP/cghviewer/Cgh Viewer.cgi) show both gain and loss of chromosome 12 in breast cancer cell lines (Table 1-10). This is reported as an intensity value which is the DNA copy number of chromosome 12p and 12q from a microarray analysis of each cell line. A value of 1.0 is equivalent to a diploid complement (2 copies) in a normal cell. Gene deletion is indicated by the copy number dropping to 0 (Stratton, Futreal 2007). These gene copy number data might be of value when considering different chromosomal regions for validation and evaluation of therapeutic targets and prognostic factors in breast cancer cases.

Table 1-12 Frequency of gene amplification on chromosome 12 in breast carcinoma cases by using microarray-based comparative genomic hybridization (array CGH)

Location		Amplified gene	Frequency (%)	Number of breast carcinoma cases	Reference
12p	12p12	*H2AFJ*	8.51	4/47	Yao et al. 2006
	12p12.1	*KRAS2*	37.04	10/27	Daigo et al. 2001
			8.51	4/47	Yao et al. 2006
	12p13	*EPS8*	8.51	4/47	Yao et al. 2006
		ETV6 (*TEL*)	10.64	5/47	Yao et al. 2006
	12p13.2	*CD69*, *EDR1* (*PHC1*), *NANOG*, *PRR4*, and *STELLA* (*DPPA3*)	9.09	3/33 (Medullary breast carcinoma)	Vincent-Salomon et al. 2007
		BCL2L14 and *ETV6* (*TEL*)	9.09	3/33 (Medullary breast carcinoma)	Vincent-Salomon et al. 2007
	12p13.2 to p13.31	*CD9*, *LTBR*, *NTF3*, and *TNFRSF1A*	9.09	3/33 (Medullary breast carcinoma)	Vincent-Salomon et al. 2007
	12p13.32 to p13.33	*CCND2*	6.06	2/33 (Medullary breast carcinoma)	Vincent-Salomon et al. 2007

Table 1-12 (Continued) Frequency of gene amplification on chromosome 12 in breast carcinoma cases by microarray-based comparative genomic hybridization (array CGH)

Location		Amplified gene	Frequency (%)	Number of breast carcinoma cases	Reference
12q	12q13.12 to q14.1	*CDK2*, *CDK4*, *ERBB3*, *GLI1*, *GPD1*, *ITGA7*, *MMP19*, and *SAS*	92.31	12/13 [Invasive (infiltrating) lobular carcinoma]	Reis-Filho et al. 2006
	12q13.2 to q13.3	*GLI1*	21.43	6/28	Daigo et al. 2001
			29.03	9/31	Nessling et al. 2005
	12q13.3	*SAS* (*TSPAN31*)	29.03	9/31	Nessling et al. 2005
	12q14	*CDK4*	29.03	9/31	Nessling et al. 2005
			-	1/1 (JIMT-1 breast cancer cell line)	Rennstam et al. 2007
		TIP120A (*CAND1*)	29.03	9/31	Nessling et al. 2005
	12q14.3 to q15	*MDM2*	22.22	6/27	Daigo et al. 2001
			29.03	9/31	Nessling et al. 2005

Table 1-13 The intensity value for the DNA copy number changes on short (p) and long (q) arms of chromosome 12 in breast cancer cell lines (From http://www.sanger.ac.uk/cgi-bin/genetics/CGP/cghviewer/Cgh Viewer.cgi)

Breast cancer cell lines	Intensity value		Breast cancer cell lines	Intensity value	
	12p region	12q region		12p region	12q region
AU565	0.65 – 1.10	0.5 – 1.25	Hs-578-T	0.60 – 1.00	0.50 – 1.00
BT-20	0.50 – 1.60	0.20 – 1.75	KPL-1	0.70 – 1.50	0.80 – 1.50
BT-549	0.50 – 1.00	0.45 – 0.95	MCF7	0.50 – 1.50	0.80 – 1.75
CAL-51	0.90 – 1.25	0.80 – 1.30	MDA-MB-134-VI	1.00 – 2.00	1.00 – 2.00
CAL-85-1	0.50 – 1.90	0.50 – 2.00	MDA-MB-157	0.70 – 0.95	0.55 – 0.90
CAL-120	1.40 – 1.60	0.50 – 5.00	MDA-MB-175-VII	0.55 – 1.20	0.70 – 1.20
CAL-148	0.90 – 1.10	0.80 – 1.10	MDA-MB-231	0.50 – 0.80	0.50 – 1.20
CAMA-1	0.30 – 1.45	0.40 – 1.50	MDA-MB-361	1.25 – 1.70	1.00 – 2.35
COLO-824	0.90 – 1.80	0.25 – 2.10	MDA-MB-415	0.30 – 1.25	0.80 – 1.25
EVSA-T	0.60 – 1.60	0.75 – 1.60	MDA-MB-435	0.70 – 1.30	0.60 – 1.00
HCC38	0.80 – 1.50	0.45 – 1.20	MDA-MB-453	0.70 – 1.25	0.70 – 1.50
HCC1143	0.75 – 1.00	0.75 – 3.30	MDA-MB-468	0.60 – 0.90	0.50 – 0.80
HCC1187	0.50 – 1.50	0.50 – 1.00	MFM-223	0.70 – 1.70	0.75 – 1.70
HCC1395	0.80 – 1.70	0.40 – 1.00	MRK-nu-1	0.80 – 1.40	0.70 – 1.60
HCC1419	0.75 – 1.50	0.50 – 1.90	MT-3	0.80 – 1.25	0.70 – 1.40
HCC1569	0.75 – 1.05	0.70 – 1.20	NCI-ADR-RES	0.70 – 1.30	0.70 – 1.60
HCC1599	0.35 – 1.00	0.80 – 1.30	OCUB-M	1.00 – 1.80	0.90 – 1.30
HCC1806	0.90 – 1.75	0.70 – 1.50	T47D	0.30 – 1.35	1.00 – 1.50
HCC1937	0.90 – 1.50	0.70 – 1.20	UACC-812	0.50 – 1.20	0.75 – 1.70
HCC1954	0.30 – 1.40	0.60 – 2.60	UACC-893	0.90 – 1.80	0.80 – 1.80
HCC2157	0.60 – 0.80	0.30 – 0.90	YMB-1	1.25 – 2.00	0.75 – 2.40
HCC2218	0.90 – 1.20	0.90 – 1.30	ZR-75-1	1.00 – 2.60	0.50 – 2.20

1.9. *PUTATIVE STEM CELL-ASSOCIATED GENES ON CHROMOSOME 12 AND BREAST CANCER*

Early Development Regulator 1 (*EDR1*) / *Polyhomeotic Homolog 1* (*PHC1*) (*Drosophila*) (Vincent-Salomon et al. 2007), *Growth Differentiation Factor 3* (*GDF3*) (Ezeh et al. 2005), *Homeobox Transcription Factor Nanog* (*NANOG*), and *STELLA* (*Developmental Pluripotency-associated Gene 3* or *DPPA3*) (Ezeh et al. 2005, Vincent-Salomon et al. 2007) are genes encoding putative transcription factors located in a cluster on chromosome 12p13 which specifically maintain the pluripotency (undifferentiated state) of embryonic stem cells and germ lineage cells (Boyer et al. 2005, Giuliano et al. 2005, Hart et al. 2004, Korkola et al. 2006, Kuroda et al. 2005, Pain et al. 2005). According to the concept of breast stem cells and carcinogenesis, the putative stem cell-associated genes on this chromosome could be activated in the early stage of the development of breast cancer in the adult. These genes may therefore play a role in tumor progression due to promoting the self-renewal (immortal) of tumor cells (Ezeh et al. 2005, Hart et al. 2004, Korkola et al. 2006, Letessier et al. 2006, Tai et al. 2005, Trosko et al. 2005). Additionally, the putative stem-cell associated genes may relate to overexpression and/or amplification of the other genes on chromosome 12 in breast cancer.

Thus far there are only limited data concerning these putative stem cell-associated genes in breast cancer. *DPPA3*, *EDR1*, and *NANOG* genes have been reported previously to be amplified in 3 out of 33 cases (9.09%) of medullary breast carcinoma by using microarray-based comparative genomic hybridization (array CGH) (Vincent-Salomon et al. 2007). Expression of *DPPA3*, *GDF3*, and *NANOG* genes was identified in MCF7 cells and a single case of invasive (infiltrating) ductal carcinoma (IDC) stage 3

by polymerase chain reaction (PCR) (Ezeh et al. 2005); and expression of the *NANOG* gene was correlated with high-grade (grade III) breast cancers and poor clinical outcome (Ben-Porath et al. 2008).

1.10. *CHROMOSOME 12 ABERRATIONS AND CLINICOPATHOLOGICAL IMPLICATIONS*

Generally, chemotherapeutic drugs are either cytotoxic or cytostatic and act on cells with are dividing (both normal and cancer cells). They spare the undifferentiated, very slowly dividing, compartment which contains cancer stem cells. Subsequently, recurrent tumour and/or metastasis may occur. Therapeutic agents which selectively eliminate cancer stem cells so resulting in potential cancer cure are therefore desired (Cariati, Purushotham 2008, Cho, Clarke 2008, Dontu et al. 2003b, Dontu, El-Ashry & Wicha 2004, Gudjonsson, Magnusson 2005, Liu, Dontu & Wicha 2005, Smalley, Ashworth 2003, Tan et al. 2006, Wicha, Liu & Dontu 2006). Hence, an evaluation of genomic instability, especially amplification of putative stem cell-associated genes, of chromosome 12 in breast cancer can potentially provide: (I) better understanding of mammary carcinogenesis (Bertucci et al. 2006); (II) potential diagnostic biomarkers (Dontu et al. 2003b, Dontu, El-Ashry & Wicha 2004, Gudjonsson, Magnusson 2005, Letessier et al. 2006, Tan et al. 2006, Wicha, Liu & Dontu 2006); (III) prognostic information (Letessier et al. 2006, Pain et al. 2005); and (IV) in the long term of more effective therapeutic drugs (Letessier et al. 2006, Pain et al. 2005, Smalley, Ashworth 2003).

1.11. *HYPOTHESIS*

The hypothesis to be tested in this thesis is that expression of putative stem cell-associated genes on chromosome 12 may be important in the development and progression of breast cancers and may provide new information to stratify breast cancers for appropriate treatment.

1.12. *AIMS AND OBJECTIVES*

The aim of this thesis was to investigate the role of putative stem cell-associated genes on chromosome 12 in breast carcinomas.

The objectives were to:

(I) Evaluate expression of putative stem cell-associated genes in normal female breast tissues, breast carcinoma tissues, surrounding normal breast tissues corresponding to the cancers, and breast cancer cell lines using Western blotting, immunohistochemistry (IHC) and TaqMan® quantitative reverse transcriptase polymerase chain reaction (QRT-PCR);

(II) Evaluate copy number changes of putative stem cell-associated genes using TaqMan® copy number assays (CNAs) and relate to Affymetrix® Genome-Wide Human Single Nucleotide Polymorphism (SNP) Array 6.0 data;

(III) Correlate copy number and expression of putative stem cell-associated genes with clinicopathological information.

CHAPTER 2

MATERIALS AND METHODS

2.1. *MATERIALS*

2.1.1. *Tissue samples*

The tissue samples in this thesis consisted of 121 breast tissues, including non-neoplastic and malignant tissues. They were obtained from the Department of Histopathology, University Hospitals of Leicester. The 121 samples comprised: 28 normal female breast tissues obtained from reduction mammoplasties; 3 non-invasive breast cancer tissues [ductal carcinoma in situ (DCIS)]; 52 invasive breast carcinoma tissues; 38 surrounding normal breast tissues taken at least 4 cm distant from the corresponding breast cancer. The characteristics of the 52 invasive breast cancers are shown in Table 2-1. FFPE tissue was available for all tissue samples and frozen tissue samples for the majority. There was ethical approval for the use of all tissues [Leicestershire, Nottinghamshire, and Rutland Research Ethics Committee (REC) (06/Q2502/70)].

For frozen samples, the fresh tissue was sampled approximately 0.8 x 0.4 x 0.3 cm size, then frozen rapidly in liquid nitrogen, then stored in liquid nitrogen (N_2) tank until use. For FFPE samples, the fresh tissue was sampled approximately 2.0 x 1.0 x 0.3 cm size and fixed in 10% Formol saline solution (Appendix 1) for 18-24 hours. The formalin fixed tissues were processed overnight using the Leica ASP300 S Automated Vacuum Tissue Processor (Leica Microsystems, UK) and subsequently embedded in paraffin wax. The FFPE tissue blocks were stored at room temperature until required.

Table 2-1 The characteristics of 52 invasive breast cancers

Characteristics	N (%)
Histopathology	
DCIS & IDC-NST	42/52 (80.77%)
IDC-NST	7/52 (13.46%)
ILC	1/52 (1.92%)
MC	2/52 (3.85%)
Age (Years old)	
≤ 50	32/52 (61.54%)
> 50	20/52 (38.46%)
Tumour size (cm)	
≤ 2.0	17/52 (32.69%)
> 2.0	35/52 (67.31%)
Tumour grade (differentiation)	
Low grade [Grade I (Well) & II (Moderate)] ♦ *Grade I (Well)* ♦ *Grade II (Moderate)*	22/52 (42.31%) *2/52 (3.85%)* *20/52 (38.46%)*
High grade [Grade 3 (Poor)]	30/52 (57.69%)
Axillary lymph node	
No metastasis	26/52 (50.00%)
Metastasis	26/52 (50.00%)
ER	
Negative	10/51 (19.61%)
Positive ♦ *Low level* ♦ *High level*	41/51 (80.39%) *2/51 (3.92%)* *39/51 (76.47%)*
No information	1/52 (1.92%)
PR	
Negative	10/51 (19.61%)
Positive ♦ *Low level* ♦ *High level*	41/51 (80.39%) *7/51 (13.72%)* *34/51 (66.67%)*
No information	1/52 (1.92%)
HER2 (ERBB2)	
Negative	11/16 (68.75%)
Positive	5/16 (31.26%)
No information	36/52 (69.23%)

Note: DCIS & IDC-NST = Combined ductal carcinoma in situ (DCIS) and invasive (infiltrating) ductal carcinoma of no special type (IDC-NST); ILC = Invasive (infiltrating) lobular carcinoma; and MC = Mucinous carcinoma

2.1.2. *Cell culture materials*

2.1.2.1. Cell lines

Eight cell lines comprised one mixed germ cell tumour cell line (NCCIT); one immortalised normal breast cell line (HBL-100); and 6 breast cancer cell lines (MCF7, MDA-MB-231, MDA-MB-436, MDA-MB-468, T47D, and ZR-75-1) (Table 2-2).

2.1.2.2. Reagents

- 200 mM (100x) L-Glutamine (29.2 mg/mL) (100 mL) (25030-024; Gibco® - Invitrogen, UK)
- Dimethyl sulfoxide (DMSO) (D5879; Sigma-Aldrich®, UK)
- Dulbecco's modified Eagle's medium (DMEM) without L-Glutamine (D5921; Sigma-Aldrich®, UK)
- Dulbecco's phosphate buffered saline (DPBS) without calcium and magnesium (14190-094; Gibco® - Invitrogen, UK)
- Foetal bovine serum (FBS) (10270-098; Gibco® - Invitrogen, UK)
- Roswell Park Memorial Institute (RPMI)-1640 with L-Glutamine (R8758; Sigma-Aldrich®, UK)
- Trypsin-EDTA (1:250) UV Inactivated (1x) (L11-660; PAA The Cell Culture Company, UK)

Table 2-2 A brief description of cell lines

Cell line	Catalogue number	Cell type	Gender	Ethnicity	Age (Years old)	Source	Supplier
NCCIT	CRL-2073™	Malignant germ cell	Male	Japanese	Unknown	Mediastinal pleuripotent embryonal carcinoma; Teratocarcinoma	ATCC®, UK
HBL-100	HTB-124™	Immortalised epithelial cell	Female	Caucasian	27	Normal mammary epithelial cells from an early lactation	ATCC®, UK
MCF7	HTB-22™	Malignant epithelial cell	Female	Caucasian	69	Metastatic breast carcinoma	ATCC®, UK
MDA-MB-231	HTB-26™	Malignant epithelial cell	Female	Caucasian	51	Metastatic breast carcinoma	ATCC®, UK
MDA-MB-436	HTB-130™	Malignant epithelial cell	Female	Caucasian	43	Metastatic breast carcinoma	ATCC®, UK
MDA-MB-468	HTB-132™	Malignant epithelial cell	Female	Black	51	Metastatic breast carcinoma	ATCC®, UK
T47D	HTB-133™	Malignant epithelial cell	Female	Unknown	54	Metastatic breast carcinoma	ATCC®, UK
ZR-75-1	CRL-1500™	Malignant epithelial cell	Female	Caucasian	63	Metastatic breast carcinoma	ATCC®, UK

2.1.3. *Commonly used chemicals and reagents*

- 1x Tris-Acetate-Ethylenediaminetetraacetic acid (EDTA) (TAE)
- Bovine serum albumin (BSA) (Albumin Standard Ampoules, 2 mg/mL)
- Glycine (G8898; Sigma-Aldrich®, UK)
- SeaKem® LE Agarose gel (Cambrex, USA)
- Sodium dodecyl sulfate (SDS) (Lauryl sulfate) (L-3771; Sigma-Aldrich®, UK)
- TRIGENE ADVANCE Disinfectant Concentrate (TR104; Medichem International, UK)
- Tris base [Tris (hydroxymethyl) aminomethane] (BPE152-1; Fisher Scientific, UK)

2.1.4. *Equipment and reagents for real-time polymerase chain reaction (PCR)*

2.1.4.1. Equipment

- Nanodrop® ND-1000 Ultraviolet (UV)-Visible Spectrophotometer (Labtech International, UK) for measurement of nucleic acid concentration
- StepOnePlus™ Real-Time PCR System (Applied Biosystems, UK) for analysis of gene expression and CNVs

2.1.4.2. Reagents for quantitative PCR (QPCR)

- 2x TaqMan® Genotyping Master Mix (Applied Biosystems, UK)
- 20x TaqMan® Copy Number Gene Assays and 20x TaqMan® Copy Number Reference Assays (Applied Biosystems, UK) (Table 2-3)

Table 2-3 TaqMan® Copy Number Assays (CNAs) for the analysis of copy number variations (CNVs)

Gene		Assay identity	Supplier
Target	*DPPA3*	Hs03835309_cn	Applied Biosystems, UK
	EDR1	Hs04413714_cn	Applied Biosystems, UK
	NANOG	Hs03820140_cn	Applied Biosystems, UK
Endogenous reference	*RNase P*		Applied Biosystems, UK
	TERT		Applied Biosystems, UK

2.1.4.3. Reagents for quantitative reverse transcriptase PCR (QRT-PCR)

- TaqMan® Fast Universal PCR Master Mix and TaqMan® Gene Expression Assays (Applied Biosystems, UK) (Table 2-4)

Table 2-4 TaqMan® Gene Expression Assays for quantitative reverse transcriptase polymerase chain reaction (QRT-PCR)

Gene		Assay identity	Amplicon length (bp)	Supplier
Target	*DPPA3*	Hs01931905_g1	131	Applied Biosystems, UK
	EDR1	Hs01863307_s1	148	Applied Biosystems, UK
	GDF3	Hs00220998_m1	65	Applied Biosystems, UK
	NANOG	Hs02387400_g1	109	Applied Biosystems, UK
Endogenous reference	*GAPDH*	Hs02758991_g1	93	Applied Biosystems, UK
	HPRT1	Hs99999909_m1	100	Applied Biosystems, UK
	TFRC	Hs00174609_m1	79	Applied Biosystems, UK

2.1.5. *Antibodies for analysis of protein expression*

- Western blotting antibodies (Table 2-5)
- Immunohistochemical antibodies (Table 2-6)

Table 2-5 Antibodies for Western blotting

Antibody			Source	Type	Working dilution	Immunogen		Catalogue number	Supplier
						Amino acids	Predicted band size (kDa)		
Primary	Target	DPPA3 (STELLA)	Rabbit	Polyclonal	1:1,000	1 – 159 (Full length)	20	Sc-67250	Santa Cruz Biotechnology, Germany
		EDR1 (PHC1)	Mouse	Monoclonal	1:1,000	751 - 851	43	ab54954	abcam®, UK
		NANOG	Rabbit	Polyclonal	1:1,000	N-terminus	42	3580	Cell Signaling Technology, UK
	Positive control	Vinculin	Mouse	Monoclonal	1:5,000	Vinculin from human uterus	116	V9131	Sigma®, UK
Secondary	Anti-mouse Ig/HRP		Goat	Polyclonal	1:2,500			P0447	DakoCytomation, UK
	Anti-rabbit Ig/HRP		Goat	Polyclonal	1:2,000			P0448	DakoCytomation, UK

Table 2-6 Immunohistochemical antibodies

Antibody		Source	Type	Species reactivity	Immunogen	Recommended dilution	Cellular localisation	Catalogue number	Supplier
Primary	DPPA3 (STELLA)	Rabbit	Polyclonal	Mouse*	Residues 100 to the C-terminus	1:50 – 1:250	Cytoplasmic and nuclear	ab19878	abcam®, UK
	EDR1 (PHC1)	Rabbit	Polyclonal, mono-specific	Human	Protein Epitope Signature Tag (PrEST)	1:50 – 1:100	Cytoplasmic (and occasionally membranous) and mainly nuclear$	HPA006973	Sigma-Aldrich®, UK
	NANOG	Rabbit	Polyclonal	Human	The internal region	1:10 – 1:50	Cytoplasmic (and occasionally membranous) and nuclear¥	PAB1887	AMS Biotechnology, UK
Secondary	NovoLink™ Polymer (Anti-mouse/rabbit IgG-Poly-HRP reagent) from NovoLink™ Polymer Detection System								Novocastra Laboratories, UK

Note: * No test in other species; $ http://www.proteinatlas.org/ENSG00000111752/cancer/breast+cancer; and

¥ http://www.proteinatlas.org/ENSG00000111704/cancer/breast+cancer

2.1.6. *Assay kits*

- Illustra™ GenomiPhi™ V2 DNA Amplification Kit (GE Healthcare, UK)
- RNeasy Mini Kit (74106; Qiagen, UK) for extraction of ribonucleic acid (RNA)
- Pierce® Bicinchoninic Acid (BCA) Protein Assay Kit (Thermo Fisher Scientific, UK) for determination of protein concentration used for Western blotting
- NovoLink™ Polymer Detection System (Novocastra™ Laboratories, UK) for immunohistochemistry (IHC), including:
  - DAB (3,3'-Diaminobenzidine) Chromogen
  - DAB Substrate Buffer (Polymer)
  - Peroxidase Block [3% Hydrogen peroxide (H_2O_2)]
  - Polymer (Anti-mouse/rabbit IgG-Poly-HRP reagent)
  - Post Primary Block
  - Protein Block

2.2. *METHODS*

2.2.1. *General aseptic techniques*

Sterile techniques were applied throughout my research for the prevention of contamination with (I) dust, particulates, microbes in cell culture (Cell Culture Basics Companion Handbook, Fundamental Techniques in Cell Culture, Masters, Stacey 2007, Ryan, Mariano); (II) any complementary (copy) deoxyribonucleic acid (cDNA) in QRT-PCR; and (III) foreign genomic DNA (gDNA) in analysis of CNVs (Bustin, Nolan , Bustin, Nolan , Sproul 2006). The procedures used included: (1) wearing a clean laboratory coat and gloves at all times; (2) using separate coats for cell culture and other labs; (3) sanitising gloved hands with 70% ethanol before commencing laboratory work; (4) frequent changing of gloves; (5) cleaning the working laboratory bench and equipment with 70% ethanol or TRIGENE ADVANCE Disinfectant at a dilution of 1:10 before, during, and after each use; and (6) using fresh reagents for each real-time PCR assay.

2.2.2. *Cell culture*

The frozen cell lines (Section 2.1.2.1 and Table 2-2) were placed in 37 °C water bath (Grant Instruments, UK) for approximately 4 minutes or until the cell lines were thawed. The cell culture media, 10% FBS, L-Glutamine, and DPBS solution were also warmed in 37 °C water bath for at least 30 minutes before use. The pre-warmed DMEM medium was supplemented with 10% FBS and 2mM L-Glutamine, whereas the pre-warmed RPMI-1640 medium was supplemented with 10% FBS only. The thawed cell lines were washed and buffered with 8 mL of DPBS in 15 mL Corning® plastic centrifuge tube [430053; Sigma-Aldrich®, UK] then spun at 1,000 revolutions per minute (rpm) in Heraeus® Megafuge® 1.0 centrifuge (75003490; DJB Labcare, UK) for

5 minutes at room temperature. After supernatant fluid was discarded, the pellets of each cell line were re-suspended and sub-cultivated in individual Corning® 75 cm^2 rectangular canted neck cell culture flask (430641; Sigma-Aldrich®, UK) containing their appropriate mixture of pre-warmed media at a recommended ratio as demonstrated in Table 2-7. The morphology of the cultured cells in flasks was observed under the Olympus CK2 Inverted Microscope (Olympus Optical, UK) and subsequently incubated at 37 °C with humidified atmosphere of 5% carbon dioxide (CO_2).

Table 2-7 Culture media and a sub-cultivation ratio for cell lines

Culture media	Cell line	Recommended sub-cultivation ratio
DMEM	MCF7	1:3 to 1:6
	MDA-MB-231	1:2 to 1:4
	MDA-MB-436	1:2
	MDA-MB-468	1:2 to 1:4
RPMI-1640	NCCIT	1:4 to 1:8
	HBL-100	1:2
	T47D	1:3 to 1:5
	ZR-75-1	1:4 to 1:6

Note: DMEM = Dulbecco's modified Eagle's medium; RPMI = Roswell Park Memorial Institute

After sub-cultivation for 24 hours, the incubated cells were daily and microscopically examined for cell growth, degree of confluency, bacterial and/or fungal contaminants. The growth media in each incubated flask were renewed 2 – 3 times per week. The cultured cells were harvested at approximately 80% confluency. The previous media mixtures, FBS, DPBS and Trypsin-EDTA were warmed in 37 °C water bath for at least 30 minutes before harvest. After the spent media were discarded, the cells were washed and buffered with 5 mL of pre-warmed DPBS then detached from the cell culture substrate by adding 4 mL of pre-warmed Trypsin-EDTA. The flasks were returned to a 37 °C humid CO_2 incubator and left for 5 minutes then the cells were assessed their dissociation under the inverted microscope. The dissociated cells were washed and buffered with 5 mL of pre-warmed DPBS by slowly pipetting up and down 2 – 3 times then transferred to individual 15 mL Corning® plastic centrifuge tubes. All tubes were spun at 1,000 rpm in a centrifuge for 5 minutes at room temperature and the supernatant was discarded. The cells were subsequently prepared for cryopreservation and pellets.

For cryopreservation, each of trypsinised cells was individually re-suspended in freezing medium (a combination of 8 mL of pre-warmed DMEM or RPMI mixture, 1 mL of pre-warmed FBS, and 1 mL of pre-warmed DMSO). One mL of each cell aliquot was transferred to a sterile cryogenic storage vial and stored at -80 °C overnight then transferred to liquid N_2 tank for permanent storage. For cell pellets, each sample of trypsinised cells was washed and buffered with 1 mL of pre-warmed DPBS by slowly pipetting up and down 2 – 3 times then transferred to individual 1.5 mL Eppendorf® microtubes. All microtubes were spun at 1,000 rpm in a microcentrifuge for 1 minute at room temperature and the supernatant was discarded. The cell pellets were stored at -20 °C until use in extraction of deoxynucleic acid (DNA) and ribonucleic acid (RNA).

2.2.3. *Hematoxylin and Eosin (H&E) staining*

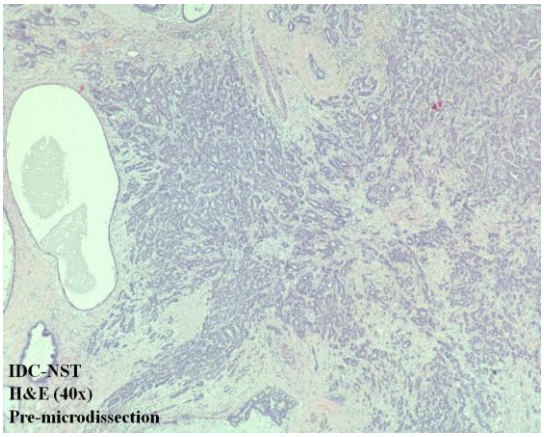
The tissue sections of 4 µm frozen and FFPE samples (Section 2.1.1) were deparaffinized in 2 changes of xylene (Genta Medical, UK) for 3 minutes each. Rehydration of the deparaffinized sections was performed in 99%, 99%, and 95% Industrial Methylated Spirit (IMS)® (Genta Medical, UK) for one minute each. They were subsequently washed in running tap water for one minute. Staining in Mayer's Hematoxylin solution was carried out for 5 minutes followed by the removal of excessive blue stain in running tap water for 5 minutes. Eosin solution was counterstained for 3 minutes and then the slides were placed in running tap water for 30 seconds. All stained slides were dehydrated through 95%, 99%, and 99% IMS® for one minute each. Afterwards they were cleared in 2 changes of xylene for 3 minutes each. Finally, the stained sections were applied to Resinous (DPX) mountant for microscopy® (VWR International, UK). The histopathological appearances were observed under Leitz Dialux® light microscope (Leica Microsystems, UK).

2.2.4. *Analysis of nucleic acid expression*

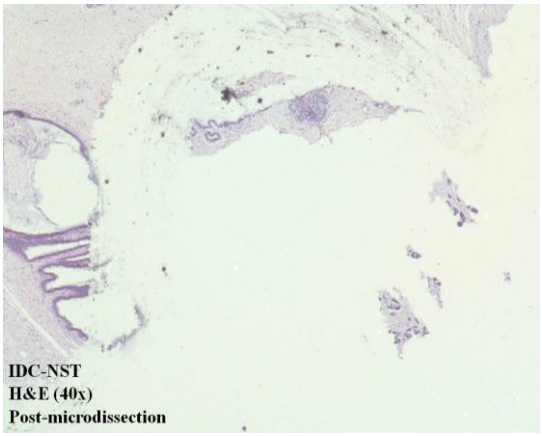
2.2.4.1. Microdissection

The microdissection procedure was modified from the method of Zhuang and colleagues (Zhuang et al. 1995). Under the light microscope, fields of invasive carcinoma in each of the H&E-stained slides were selected and carefully rimmed by hand with a permanent marker. The unstained tumor tissue sections were deparaffinized twice in xylene for 3 minutes each. Then they were rehydrated in 99%, 99%, and 95% IMS® for one minute each and air dried. The tumor areas on those unstained sections were matched to the marked fields on the respective H&E-stained slides and subsequently scraped with a sterile pipette tip. The microdissected tumor cells were

immediately digested in 500 µL of 0.05 M Tris of pH 8.0/0.1% SDS and transferred to each individual 1.5 mL Eppendorf® microtubes for extraction of DNA. The microdissected slides were stained with H&E for a demonstration of correctly selected tumor areas (Figure 2).



(A)



(B)

Figure 2 Manual microdissection. The H&E staining of the invasive breast carcinoma tissue samples shows an area of cancer cells before (A) and after (B) manual microdissection.

2.2.4.2. Extraction of genomic DNA (gDNA)

Each of the pellets of cell lines (Section 2.2.2) was resuspended in 1,000 µL of 0.05 M Tris of pH 8.0/0.1% SDS. For inactivation of deoxyribonuclease (DNase) and degradation of protein, 50 µL of 10 mg/mL of proteinase K was added. The microtubes containing microdissected tumour cells (Section 2.2.4.1) were repeatedly reacted with 25 µL of 10 mg/mL of proteinase K for 3 consecutive days. The digested cell lines were incubated at 37 °C for 1 hour, whereas the microdissected tumours were incubated at 55 °C overnight for each addition of proteinase K.

All incubated gDNA lysates were separated from other cellular components twice by applying 500 µL of phenol : chloroform : IAA (25 : 24 : 1) saturated with 10 mM Tris of pH 8.0, 1 mM EDTA (P3803; Sigma-Aldrich®, UK) then spun at 13,000 rpm in a microcentrifuge for 3 minutes. The supernatant was extracted with 500 µL of chloroform : IAA (24 : 1) (C0549; Sigma-Aldrich®, UK) then spun at 13,000 rpm in a microcentrifuge for 3 minutes. The gDNA in the last supernatant was precipitated in 120 µL of 1 M sodium chloride and 900 µL of –20 °C absolute ethanol (E/0650DF/17; Fisher Scientific, UK). For gDNA extracted from tissue sections, one µl of 20 mg/mL of glycogen (Invitrogen, UK) was additionally used as a carrier to aid nucleic acid precipitation in ethanol. The isolated gDNA solution from cell lines and tissue sections were incubated at –20 °C for 30 minutes and 24 hours, respectively. They were spun at 13,000 rpm in a microcentrifuge at 4 °C for 10 minutes, fluid discarded, rinsed with 200 µL of 70 % ethanol, fluid discarded, and spun at 13,000 rpm in a microcentrifuge for 5 minutes. The gDNA extracted from tissue sections and cell lines were resuspended in sterile ultrapure water in 30 and 200 µL of final volumes, respectively. The concentration of isolated gDNA was measured from the absorbance at 260 nm on the

Nanodrop® ND-1000 Ultraviolet (UV) - Visible Spectrophotometer. The gDNA was stored at 4 °C until use.

2.2.4.3. Whole genome amplification (WGA) of DNA samples

The gDNA from the 10 selected BC tissue samples in Section 2.2.4.2 was prepared for Affymetrix® Genome-Wide Human Single Nucleotide Polymorphism (SNP) Array 6.0 using Illustra™ GenomiPhi™ V2 DNA Amplification Kit. A 5 µL aliquot of 10 ng of gDNA was mixed with 20 µL of GenomiPhi Sample Buffer, then incubated at 95 °C for 3 minutes and subsequently cooled on ice for 5 minutes. The cooled gDNA sample was amplified with 22.5 µL of GenomiPhi V2 Reaction Buffer and 2.5 µL of GenomiPhi V2 Enzyme Mix (Phi 29 polymerase). The mixture was incubated overnight at 30 °C and amplification reaction inactivated afterwards at 65 °C for 10 minutes. Whole genome amplified DNA aliquot was dissolved in 150 µL of 1x Tris-Low EDTA buffer solution for long-term storage at –20 °C. Then it was quantified by the absorbance at 260 nm on the Nanodrop® ND-1000 Ultraviolet (UV) - Visible Spectrophotometer. Whole genome amplified DNA samples were determined for their integrity and size on 1% SeaKem® LE Agarose gel electrophoresis in 1x Tris-Acetate-EDTA (TAE) with 5 µL of HyperLadder™ I. The DNA was subsequently visualized with ethidium bromide under UV light [UV Transilluminator (TM-40; UVP, UK)].

2.2.4.4. Extraction of ribonucleic acid (RNA)

A 100 mg of 4 µm frozen tissue sections of normal breast, surrounding normal breast, and breast carcinoma tissues (Section 2.1.1) and the pellets of cell lines (Section 2.2.2) were added to 1,000 µL of TRI Reagent® (T9424; Sigma-Aldrich®, UK) and 200 µL of Chloroform (C2432; Sigma-Aldrich®, UK) then spun at 13,000 rpm in a microcentrifuge at 4 °C for 15 minutes. The upper clear fluid was transferred to a fresh

tube and added to 1.25x the volume of absolute ethanol. RNA was isolated using RNeasy Mini Spin Column from the RNeasy Mini Kit which was spun twice at 13,000 rpm in a microcentrifuge at room temperature for 15 seconds. The membrane-bound RNA on the Column was washed with 700 µL of Buffer RW1 and spun at 13,000 rpm in a microcentrifuge at room temperature for 15 seconds. The Column was subsequently washed twice with 500 µL of Buffer RPE and spun at 13,000 rpm in a microcentrifuge at room temperature for 15 seconds and 2 minutes, respectively. Afterwards the Column was washed with 100 µL of RNase-Free Water and spun at 13,000 rpm in a microcentrifuge at room temperature for one minute. The isolated RNA was measured for concentration using absorbance at 260 nm on the Nanodrop® ND-1000 Ultraviolet (UV) - Visible Spectrophotometer. The RNA aliquots were stored at –20 °C until use.

2.2.4.5. Synthesis of complementary DNA (cDNA)

The cDNA was synthesised in a 25 µL of reverse transcriptase (RT) reaction containing 10 µL of RNA aliquots (Section 2.2.4.4); 5 µL of Avian Myeloblastoma Virus (AMV) RT 5x Reaction Buffer (M515A; Promega, UK); 2.5 µL of 10 mM Deoxyribonucleotide triphosphates (dNTPs)/Diethylpyrocarbonate (DEPC); 0.62 µL (25U) of RNasin® Ribonuclease Inhibitor (40 U/µL) (N2115; Promega, UK) for protection against RNA degradation; 0.5 µL (5U) of AMV RT (10U/µL) (M510A; Promega, UK); and 6.38 µL of sterile ultrapure water. Negative controls were obtained by performing a reaction without RT. Both reactions were incubated at 42 ºC for 1 hour by using a Perkin Elmer DNA Thermal Cycler 480. The cDNA aliquots were stored at 4 ºC until use.

2.2.4.6. Real-time polymerase chain reaction (PCR)

2.2.4.6 (1). General principle

The real-time PCR was performed using triplicate reaction in MicroAmp® Fast Optical 96-Well Reaction Plate with Barcode (0.1 mL) (Applied Biosystems, UK) on the StepOnePlus™ Real-Time PCR System. The threshold cycle (C_t) value during the PCR was performed on a manual threshold (the level of signal in the region associated with an exponential phase of PCR) with an automatic baseline (a little change in reporter fluorescent signal during the PCR cycles 3 to 15). This C_t value is the cycle number at which the fluorescence generated within a reaction crosses the threshold associated with the first detected amplicon. Basically, the results of DNA and mRNA expression were determined by the difference in the C_t value (ΔC_t) between Ct value from target gene and Ct value from endogenous reference genes. The levels of both DNA and mRNA expression were inversely related to the ΔC_t value.

2.2.4.6 (2). Selection of no template (target) control (NTC) wells

The different positions of triplicate NTC wells in the real-time PCR plate set-up, i.e. the upper row, the middle row, and the lower row, were assessed for the risk of any DNA contamination. The test assay was performed on 50 amplification cycles with manual threshold at 0.3 and automatic baseline.

2.2.4.6 (3). Standard curves of the real-time PCR assay

The quality of the real-time assay was initially determined by constructing a standard curve in Microsoft Office Excel program based on 7 dilution series (1:5; 1:25; 1:125; 1:625; 1:3,125; 1:15,625; and 1:78,125) of gDNA or cDNA as template. The plot was related to the straight line with the equation "y = mx + b", where "y" is the C_t

value; "m" is slope; "x" is $\log_{10}$ template quantity; and "b" is y-intercept. The slope of the line is related to the PCR efficiency of the assay. The y-intercept indicates the assay sensitivity and the quantitative accuracy of the template. The coefficient of determination (r^2) of the standard curve represents the accuracy (validity) of the dilutions and precision (reproducibility) of pipetting. A good assay should have 95% - 105% of amplification efficiency; a y-intercept value between 33 and 37 cycles, and an r^2 value close to 1.00 (Adams 2006, Durham, Chinnery 2006).

2.2.4.6 (4). Determination of the optimal endogenous reference genes

The optimal endogenous reference gene for each target gene was determined from the mean ΔC_t value close to 0 and the low standard deviation (SD) of ΔC_t calculated using a commercial normal human genomic DNA (HGDNA) (0.2 µg/µL) (11691112001, Roche Applied Science, UK) as a template.

2.2.4.6 (5). Normalisation and precision (reproducibility) of the real-time PCR assay

The real-time PCR assay was normalised to normal or non-neoplastic cell and also evaluated their precision (reproducibility) by a repeat assay.

2.2.4.6 (6). Analysis of copy number variations (CNVs) by quantitative PCR (QPCR)

TaqMan® Copy Number Assays (CNAs) (Table 2-3) was applied for analysis of alterations in a DNA copy number (CN) of *DPPA3*, *EDR1*, and *NANOG* genes in HGDNA, normal breast tissues, surrounding normal breast tissues, breast carcinoma tissues, and cell lines. The calibrator (reference) sample was HGDNA. The cell lines were used for a parallel analysis. Sterile ultrapure water was used as a NTC sample. The

endogenous reference genes were either Ribonuclease P (*RNase P*) gene or Telomerase reverse transcriptase (*TERT*) gene. According to triplicate reaction mentioned in Section 2.2.4.6 (1), a total volume of 10 µL of each mixture consisted of 5 µL of 2x TaqMan® Genotyping Master Mix; 0.5 µL of 20x TaqMan® Copy Number Gene Assay; 0.5 µL of 20x TaqMan® Copy Number Reference Assay; 2 µL of sterile ultrapure water; and 2 µL of 5 ng/µL of gDNA sample (Section 2.2.3.2) diluted with sterile ultrapure water.

The StepOnePlus™ Real-Time PCR System was established using the parameters as shown in Table 2-8. The 60 amplification cycles were run and initiated by denaturation of gDNA at 95 °C for 10 minutes followed by 95 °C for 15 seconds and 60 °C for one minute. The C_t value during the PCR was performed on a manual threshold at 0.2 with an automatic baseline. Expression of *DPPA3*, *EDR1*, and *NANOG* genes was determined on the relative quantification (RQ) which is a fold change in target gene expression in a sample relative to the same gene expression in the calibrator sample (HGDNA). This value was calculated by the following equation: "$RQ = 2^{-\Delta\Delta Ct}$", where "$\Delta C_t = C_{t\ \text{Target gene}} - C_{t\ \text{Endogenous reference gene}}$" and "$\Delta\Delta C_t$ = The mean $\Delta C_{t\ \text{Target gene in gDNA sample}}$ – The mean $\Delta C_{t\ \text{Target gene in calibrator sample (HGDNA)}}$" (O'Leary et al. 2003, Pfaffl 2004, Pfaffl 2006, Schmittgen 2006).

According to the manufacturer's instruction, the CN value for each sample was calculated from the above-mentioned RQ value as "$CN_{Sample} = RQ_{Sample}$ x $CN_{Calibrator\ (HGDNA)}$", where CN of HGDNA = 2. The predicted CN value was obtained from rounding the calculated CN and defined as: CN of 0 = Homozygous deletion; CN of 1 = Heterozygous deletion (Loss of heterozygosity); CN of 2 = Normal diploid; CN of 3 = Single copy gain; CN of 4 = Two-copy gain; and CN of ≥ 5 = Amplification (CN ≥ 8 = High-level amplification).

Table 2-8 The parameters of StepOnePlus™ Real-Time PCR System for TaqMan® Copy Number Assays (CNAs)

Parameters		
Run		Standard
Experiment		Advanced Setup
		Quantitation-Standard Curve
Reporter	***TaqMan® Copy Number Gene Assay***	FAM
	TaqMan® Copy Number Reference Assay	VIC
Quencher	***TaqMan® Copy Number Gene Assay***	NFQ-MGB
	TaqMan® Copy Number Reference Assay	TAMRA

2.2.4.6 (7). Analysis of single nucleotide polymorphisms (SNPs)

The whole genome amplified DNA samples (Section 2.2.4.3) were diluted in sterile ultrapure water to a final concentration of 100 ng/µL in total volume of 15 µL. These DNA samples were subsequently applied for analysis of SNPs in breast carcinomas using Affymetrix® Genome-Wide Human SNP Array 6.0. This analysis was performed by Almac Diagnostics, UK.

2.2.4.6 (8). Analysis of messenger RNA (mRNA) expression by quantitative reverse transcriptase PCR (QRT-PCR)

TaqMan® Gene Expression Assays (Table 2-4) were applied for an evaluation of expression of DPPA3, EDR1, and NANOG mRNA in normal breast tissues, surrounding normal breast tissues, breast carcinoma tissues, and cell lines. Sterile ultrapure water was used as a NTC sample. The endogenous reference genes were

Glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*), Hypoxanthine Phosphoribosyltransferase 1 (Lesch-Nyhan syndrome) (*HPRT1*), and Transferrin Receptor (*TFRC*).

According to triplicate reaction mentioned in Section 2.2.4.6 (1), a total volume of 10 µL of each mixture consisted of 5 µL of TaqMan® Fast Universal PCR Master Mix; 0.5 µL of TaqMan® Gene Expression Assays; and 4.5 µL of cDNA sample from Section 2.4.2.2 diluted with sterile ultrapure water at a 1:5 ratio. Based on the StepOnePlus™ Real-Time PCR System, 40 amplification cycles were run and initiated by denaturation of cDNA at 95 °C for 20 seconds followed by 95 °C for 3 seconds and 60 °C for 30 seconds. The C_t value during the PCR was performed on a manual threshold at 0.3 with an automatic baseline.

2.2.5. *Analysis of protein expression*

2.2.5.1. Extraction of protein

The 500 µL of Gold lysis buffer working solution (Appendix 2) was added to each of 8 cell line pellets (5×10^6 cells/pellet) and 3 frozen breast cancer tissues. The mixture was left on ice for 10 minutes. Afterwards it underwent more lysis by passage 5 times through a 25 G sterile needle with a 1 mL syringe. The lysate was spun at 13,000 rpm in a microcentrifuge for 3 minutes. The protein supernatant was subsequently transferred to a new 1.5 mL Eppendorf® microtube and stored at -20 °C until use.

2.2.5.2. Determination of protein concentration

Pierce® Bicinchoninic Acid (BCA) Protein Assay Kit was used for protein quantification. A standard curve was constructed from a serial dilution of bovine serum albumin (BSA) (Albumin Standard Ampoules, 2 mg/mL) in the working range of 0.020

to 2 µg/µL (Table 2-9). The BCA Working Reagent (WR) was prepared from a mixture of 50 parts of BCA Reagent A and one part of BCA Reagent B (a 50:1 dilution of A to B). Then the protein lysate from 2.2.4.1 was added to the WR at a dilution of 1:20 and subsequently incubated at 37 °C for 30 minutes. Afterwards the absorbance of the standard and sample mixtures was measured twice at 562 nm using disposable 1 mL (10 × 4 × 45 mm) cuvettes (67.742; Sarstedt, Germany) with a UV-visible spectrophotometer (Thermo Spectronic, Genesys^{TM} 10 Series).

The mean absorbance measurement of the blank standard vial was subtracted from the mean absorbance measurement of all other individual standard vials and sample mixtures for the corrected absorbance measurement. The corrected absorbance of each BSA standard and its concentration in µg/µL was used for the construction of a standard curve in Microsoft Office Excel program. The plot was determined theoretical amounts of protein related to the straight line with the equation "y = mx + c", where "y" is the absorbance at 562 nm and "x" is protein amount. The concentration of the protein samples was calculated on the basis of the corrected absorbance and the mentioned equation for the final protein amount of 25 µg.

Table 2-9 Preparation of diluted bovine serum albumin (BSA) standards

Vial	Sterile ultrapure water (µL)	BSA (µL)	Final BSA concentration (µg/µL)
A	0	300	2.000
B	125	375	1.500
C	325	325	1.000
D	175	175 of vial B dilution	0.750
E	325	325 of vial C dilution	0.500
F	325	325 of vial E dilution	0.250
G	325	325 of vial F dilution	0.125
H	400	100 of vial G dilution	0.025
I (Blank)	400	0	0.000

2.2.5.3. Western blotting

2.2.5.3 (1). <u>Electrophoretic separation of protein</u>

Tris-glycine SDS-polyacrylamide gel electrophoresis (SDS-PAGE) was applied for a protein separation from the cell line samples using the 15% resolving and 5% stacking gels (Appendix 2). The gels were polymerised in a gel caster consisted of a short glass plate (Bio-Rad, UK), a 1.5 mm spacer plate (Bio-Rad, UK), and a 1.5 mm loading comb (Bio-Rad, UK) combining the Mini-PROTEAN® 3 Cell Assembly (Bio-Rad, UK). The polymerised gel was subsequently immersed in 1x running buffer. Each well in the submerged stocking gel was flushed out with the running buffer for the removal of gel debris and air bubbles. The protein lysate from Section 2.2.4.1 was mixed with the Western loading buffer at a 10:1 dilution. The mixture was incubated at 95 ºC for 5

minutes by using a Hybaid OmniGene Thermal Cycler. Each sample was individually loaded into the wells. Twenty µL of Precision Plus Protein™ Dual Color Standards (161-0374; Bio-Rad, UK) was used for the determination of the protein size as the following molecular weights: 250, 150, 100, 75, 50, 37, 25, 20, 15, and 10 kDa. The gel electrophoresis was run with constant voltage of 200 volts (V) for 45 minutes.

2.2.5.3 (2). Immunoblotting transfer

For the removal of residual SDS and salts and equilibration to methanol-containing buffer, the electrophoresed gel was rinsed with 1x transfer buffer (Appendix 2) for 10 minutes by using the automatic rocker (Luckham 4RT Rocking Table). The proteins on the gel were transferred to Polyvinylidene fluoride (PVDF) membrane [Immobilon™ Transfer Membrane (a pore size of 0.45 µm) (Millipore, UK)] by using the protein blotting cassette which was assembled as follows: the clear side of the cassette, the wet fibre pad, the wet filter paper (Gel-Blotting-Papers) (D37582; Schleicher & Schuell BioScience GmbH, Germany), the wet PVDF membrane, the electrophoresed gel, the wet filter paper, the wet fibre pad, and the black side of the cassette. The cassette was placed into the protein blotting module filled with the 1x transfer buffer. The clip side of the cassette was uppermost in the module and the black side of the cassette faced the black side of the module. The protein-transferring process was run overnight with constant voltage of 20 V at room temperature.

2.2.5.3 (3). Immunoblotting detection

The blotted membrane was placed into a plastic tray containing 1x TBS/0.1% Tween® 20 (TBS-T) (Appendix 2) on the automatic rocker for 5 minutes at room temperature for the prevention of non-specific background binding of the primary and/or secondary antibodies. The membrane was incubated overnight with primary

antibody diluted in 10 mL of the blocking solution [5% (w/v) Marvel skimmed milk powder in TBS-T] at 4 ºC on the automatic rocker. The primary antibodies and their working dilutions were shown in Table 2-5.

The membrane was subsequently rinsed with TBS-T three times at room temperature for 5 minutes per wash on the automatic rocker for the removal of residual primary antibody. Afterwards it was incubated with secondary antibody (Table 2-5) in 10 mL of the blocking solution at room temperature for 45 minutes on the automatic rocker. Polyclonal Goat Anti-Mouse Immunoglobulins (Ig)/Horseradish Peroxidase (HRP) was applied for EDR1 and Vinculin antibodies. Polyclonal Goat Anti-Rabbit Ig/HRP was applied for DPPA3 and NANOG antibodies.

Then the membrane was rinsed with TBS-T three times at room temperature for 5 minutes per wash on the automatic rocker for the removal of residual secondary antibody. It was subsequently incubated with a chemiluminescent substrate [SuperSignal® West Dura Extended Duration Substrate (Thermo Fisher Scientific, UK)] for detecting HRP on immunoblots. The working detection reagent was the mixture of SuperSignal West Dura Luminal/Enhancer Solution and SuperSignal West Dura Stable Peroxide Solution at a 1:1 ratio. The incubation time was 1 minute for the detection of DPPA3, EDR1, and NANOG proteins and was 3 seconds for the detection of Vinculin protein. The WB result was detected with the X-ray films (GE Healthcare, UK) at 1, 3, 5, 10, and 15 minutes of exposure time.

2.2.5.4. Immunohistochemistry (IHC)

2.2.5.4 (1). General principle

The unstained sections of 4 µm FFPE tissue were mounted onto glass slides coated with VECTABOND® Reagent (Vector Laboratories, UK) for the enhancement of tissue adhesion according to the manufacturer's protocol. After melting paraffin at 65 °C in an air incubator for 10 minutes, all sections were applied for deparaffinization in 2 changes of xylene and rehydration in 99% and 95% IMS® for 3 minutes each. They were subsequently washed in running tap water for 5 minutes. The IHC was performed by heat-induced epitope retrieval (HIER) and NovoLink™ Polymer Detection System.

For elimination of endogenous peroxidase activity, the sections were incubated with 100 µL of Novocastra™ Peroxidase Block [3% hydrogen peroxide (H_2O_2)] at room temperature for 5 minutes. Then they were incubated with 100 µL of Novocastra™ Protein Block to reduce non-specific binding of primary antibody and NovoLink™ Polymer at room temperature for 5 minutes. The sections were subsequently incubated overnight with 100 µL of optimally diluted primary antibody (Table 2-6) in a humid chamber at 4 °C. The 50 mM TBS of pH 7.6 solution was used as a diluent of primary antibody. The negative control sections were treated with 100 µL of TBS with no primary antibody (NPA).

The overnight sections were incubated with 100 µL of Novocastra™ Post Primary Block to enhance penetration of the subsequent NovoLink™ DAB (3,3'-diaminobenzidine) Substrate Buffer (Polymer) at room temperature for 30 minutes. The tissue-bound primary antibody in the sections was detected by 100 µL of NovoLink™ Polymer [Anti-mouse/rabbit IgG-Poly-HRP reagent (secondary antibody)] at room temperature for 30 minutes. The sections were further incubated with 100 µl of the

substrate/chromogen (DAB) working solution for the development of final brown product at room temperature for 5 minutes. This working solution was prepared from the mixture of Novocastra™ DAB Chromogen and NovoLink™ DAB Substrate Buffer (Polymer) at a 1:20 ratio. The buffer solution was also comprised of 0.05% H_2O_2 to inhibit endogenous peroxidase activity in tissue. Afterwards counterstaining was performed with Mayer's Hematoxylin for one minute. The immunostained sections were subsequently dehydrated through 95%, 99%, and 99% IMS® then 2 changes of xylene for 3 minutes each. They were finally covered by Resinous (DPX) mountant for microscopy®.

2.2.5.4 (2). Immunohistochemical optimisation

The immunohistochemical study was initially evaluated the optimal staining conditions, including antigen retrieval (AR) method, retrieval time, and primary antibody concentration. The most appropriate staining condition was determined from the histochemical scoring (H-score) [see Section 2.2.5.4 (3)] and the staining quality.

2.2.5.4 (2.1). *Determination of the optimal AR method*

The immunostaining was performed by using 2 heat-mediated AR methods, i.e. pressure cooker [Pascal® Microprocessor Controlled Pressure Chamber (S2800; DakoCytomation, UK)] and microwave (Tecnolec® Superwave 750). The deparaffinised sections were submerged by 1x 10 mM citrate buffer of pH 6.0 for both methods. The setting program (SP) of pressure cooker was SP1 at 120 °C for 30 seconds; SP2 (cooling down) at 90 °C for 10 seconds; and SP limit at 10 °C. For microwave AR, antigens in the deparaffinized sections were improved by the pre-treatment with the highest heat of a 750-watt microwave oven (Tecnolec® Superwave 750) in the buffer for 20 minutes. The starting dilution of primary antibody was

randomly selected from their recommended concentrations (Table 2-6) as a 1:150 dilution of DPPA3, a 1:50 dilution of EDR1, and a 1:30 dilution of NANOG. After the optimal AR method had been established, the retrieval time and primary antibody concentration were adjusted further.

2.2.5.4 (2.2). ***Determination of the optimal heating time for AR***

The immunoreactivity was performed on 3 different heating times of the established AR method from Section 2.2.5.4 (2.1), for example pressure cooker retrieval time of 30, 45, and 60 seconds and microwave retrieval time of 20, 30, and 40 minutes. The primary antibody was used the same dilution as Section 2.2.5.4 (2.1).

2.2.5.4 (2.3). ***Determination of the optimal dilution of primary antibody***

After the appropriate AR method and time had been encountered, the immunoreactivity was performed on 3 different dilution of each primary antibody: (I) 1:150, 1:100, and 1:50 for DPPA3; (II) 1:50, 1:30, and 1:15 for EDR1; and (III) 1:30, 1:20, and 1:10 for NANOG.

2.2.5.4 (3). Interpretation of immunoreactivity

An evaluation of immunohistochemical staining was performed on Leitz Dialux® light microscope first using 10x objective (magnification of 100x) in order to scan and locate the histopathological appearances. Then the 40x objective (magnification of 400x) was subsequently applied for more detailed information on the staining (Keller, Goldman 2006, Nicholson et al. 1991). T47D cell line was used as a control staining. The staining results were assessed by myself (a Thai board-certified pathologist) on 2 separate occasions for 4 weeks which were undertaken without knowledge of the first results.

The immunoreactivity of DPPA3, EDR1, and NANOG proteins was evaluated for 4 histopathological components of each breast cancer tissue sections, i.e. normal mammary epithelial cells, DCIS, invasive carcinoma, and vascular invasion. Since Allred (Quick) scoring system for breast cancer considers only overall appearance of nuclear immunostaining (Collins, Botero & Schnitt 2005, Detre, Saclani Jotti & Dowsett 1995, Harvey et al. 1999, Leake et al. 2000, NHS Cancer Screening Programmes jointly with The Royal College of Pathologists 2005, Qureshi, Pervez 2010) (Appendix 7), but DPPA3, EDR1, and NANOG immunoexpression could show various patterns in the cancer tissues (see Section 3.4 Discussion in Chapter 3) and should be firstly determined the appropriate cut-off points for their positive staining. Therefore, an interpretation was based on the histochemical scoring (H-score) assessment incorporating both the staining intensity (*i*) and a percentage of stained cells at each intensity level (*Pi*). The *i* values were indicated as 0 (no evidence of staining), 1 (weak staining), 2 (moderate staining), and 3 (strong staining). The *Pi* values varied from 0% to 100%. The final H-score was derived from the sum of *i* multiplied by *Pi* as the equation shown below. This score, therefore, was in the range of 0 to 300 (McCarty et al. 1985, McCarty et al. 1986, Tadrous 2007).

$$\text{H-score} = (0 \times P_0) + (1 \times P_1) + (2 \times P_2) + (3 \times P_3)$$

In each sample, the first and second H-scores of the invasive cancer component were calculated for their mean values. Additionally, the mean H-score was used to determine the optimal cut-off H-score values for negative and positive immunoexpression of DPPA3, EDR1, and NANOG.

2.2.6. *Statistical analysis*

The GraphPad Prism® 5 programme was used for analysis of the results of mRNA expression from QRT-PCR, protein expression from IHC, DNA copy number changes, and SNPs for *DPPA3*, *EDR1*, and *NANOG* genes at the 95% confidence interval (CI). In addition, the breast cancer cases were evaluated for any correlation between expression of mRNA, protein and CN and their clinicopathological features [the age (≤ 50 years old and > 50 years old); tumour size (≤ 2.0 cm and > 2.0 cm); tumour grade/differentiation; axillary lymph node metastasis; and the status of ER, PR, and HER2 (ERBB2)].

2.2.6.1. Statistical parameters and tests for mRNA expression

(1). Precision (reproducibility) of TaqMan® QRT-PCR assay

- Ratio paired *t*-test

(2). Comparison of the ΔCt value for mRNA expression to cell lines, normal breast tissue (NB), surrounding normal breast tissue (SNB), and invasive breast cancer tissue (IC)

- Unpaired *t*-test

(3). Comparison of the ΔCt value for mRNA expression between IC and its corresponding SNB

- Ratio paired *t*-test

(4). Co-expression of DPPA3, EDR1, and NANOG mRNA in cell lines, NB, SNB, and IC

- Pearson correlation coefficient (r)

(5). The effect of a percentage of invasive carcinoma component (≤ 50% and > 50%) in breast cancer tissue section on the QRT-PCR results

- Unpaired *t*-test

(6). The correlation between the ΔCt value for mRNA expression in IC and clinicopathological features

- Unpaired *t*-test

2.2.6.2. Statistical parameters and tests for protein expression

(1). Intra-observer variation of immunostaining interpretation

- Ratio paired *t*-test

(2). Determination of the optimal cut-off H-score for positive immunoexpression in breast cancer tissue

- One-way analysis of variance (ANOVA) with Tukey's multiple comparison test
- Unpaired *t*-test for confirmation of the one-way ANOVA results

(3). Co-immunoexpression of DPPA3, EDR1, and NANOG proteins in breast cancer tissue

- Pearson correlation coefficient

(4). Comparison of positive immunoreactivity between invasive carcinoma component and its corresponding normal mammary epithelial component

- Ratio paired *t*-test

(5). The correlation between immunoexpression in breast cancer and clinicopathological features

- Fisher exact test

(6). The correlation between positive immunoexpression in breast cancer and the clinicopathological features

- Unpaired *t*-test

2.2.6.3. Statistical parameters and tests for concordance between mRNA expression and immunoexpression

(1). The correlation between the ΔCt value for mRNA expression and immunostaining H-score in breast cancer

- Pearson correlation coefficient

(2). Comparison of the level of mRNA expression in breast cancer between negative and positive immunoreactivity

- Ratio paired *t*-test

2.2.6.4. Statistical parameters and tests for copy number variations (CNVs)

(1). Precision (reproducibility) of TaqMan® CNAs

- Paired *t*-test

(2). Comparison of the ΔCt value for CNVs to each cell line

- One-way ANOVA

(3). Comparison of the ΔCt value for CNVs to NB, SNB, and IC

- Unpaired *t*-test

(4). The effect of a percentage of invasive carcinoma component (≤ 50% and > 50%) in breast cancer tissue section on the results of TaqMan® CNAs

- Unpaired *t*-test

(5). The correlation between the ΔCt value for CNVs in IC and clinicopathological features

- Unpaired *t*-test

2.2.6.5. Statistical parameters and tests for concordance between mRNA expression, immunoexpression, and CNVs

(1). The correlation between the ΔCt value for mRNA expression and the ΔCt value for CNVs in NB, SNB, and IC

- Pearson correlation coefficient

(2). The correlation between the immunostaining H-score of invasive carcinoma component and the ΔCt value for CNVs in breast cancer

- Pearson correlation coefficient

2.2.6.6. Statistical test for an agreement of genome wide copy number analysis between TaqMan® CNAs and Affymetrix® Genome-Wide Human SNP Array 6.0

- Ratio paired *t*-test

CHAPTER 3

EXPRESSION OF DPPA3 (STELLA), EDR1 (PHC1), GDF3, AND NANOG mRNA AND PROTEINS IN BREAST CARCINOMA

3.1. *INTRODUCTION*

Breast cancer can be considered a genetic disease because genomic alterations, including genetic and chromosomal aberrations, are involved in the development and progression of breast cancer, i.e. changes in the chromosomal number, abnormalities of chromosomal structure, point mutations, and gene amplifications (Lerebours, Lidereau 2002). In relation to the concept of cancer stem cells in carcinogenesis (Polyak, Hahn 2006, Reya et al. 2001, Yu, Bian 2009), it has been proposed that breast carcinomas develop from aberrant differentiation of a small number of mammary stem cell (MaSc) or progenitor cells in the terminal duct lobular units (TDLUs) of the mammary gland. Since both cells remain in the TDLUs for a long time, they are prone to accumulate genetic and/or epigenetic modifications (Bombonati, Sgroi 2011, Dontu et al. 2003b, Lim et al. 2009, Lindeman, Visvader 2010, Luo et al. 2010, Petersen, Polyak 2010, Prat, Perou 2009).

Stem cell-associated genes on chromosome 12, i.e. *DPPA3* (*STELLA* or *PGC7*), *EDR1* (*PHC1* , *HPH1*, or *RAE28*), *GDF3*, and *NANOG*, have been correlated previously with the pathogenesis of both non-hematologic (Tanaka et al. 2009) and hematologic (Tokimasa et al. 2001) malignancies. In addition, suppression of these genes results in loss of tumourigenic potential of embryonal carcinomas (Giuliano et al. 2005, Sperger et al. 2003). Ezeh and colleagues (2005) used a PCR technique to identify *DPPA3* and *GDF3* in the MCF7 breast carcinoma cell line and *NANOG* in both MCF7

and one case of IDC stage 3 (Ezeh et al. 2005). Approximately 10% of invasive breast cancers show low expression of *EDR1* gene (Sanchez-Beato et al. 2006).

3.2. *AIMS*

The aim of this chapter is to determine the mRNA and protein expression of stem cell-associated genes on chromosome 12 (*DPPA3*, *EDR1*, *GDF3* and *NANOG*) in breast cancer cell lines, normal breast, and breast carcinomas and their associated surrounding normal breast and to relate the findings for the cancers to their clinicopathological features.

3.3. *RESULTS*

3.3.1. *Analysis of mRNA expression*

3.3.1.1. Quantitative reverse transcriptase polymerase chain reaction (QRT-PCR) plate preparation: Selection of no template (target) control (NTC) wells

The QRT-PCR plate set-up was assessed to determine if the position of the triplicate NTC wells had any influence on contamination with cDNA samples. The test assay was performed on 50 amplification cycles with manual threshold at 0.3 and automatic baseline. The result showed no contamination in the upper, the middle, and the lower positions of the NTC wells (Table 3-1). This indicated that the positions of the NTC wells on the QRT-PCR set-up plate were unlikely to be correlated with any issue of cDNA contamination.

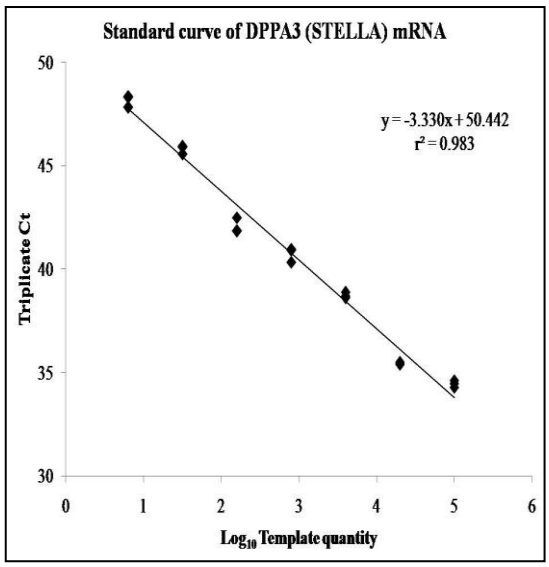
Table 3-1 The quantitative reverse transcriptase polymerase chain reaction (QRT-PCR) set-up plate for the evaluation of decontamination of the triplicate NTC wells. The QRT-PCR set-up plate was assessed the most appropriate position of the triplicate NTC wells for eliminating or minimizing contamination with cDNA samples. The assay was performed on 50 amplification cycles with manual threshold at 0.3 and automatic baseline. The numbers in each well represented the threshold cycle (Ct) value for TaqMan® Gene Expression Assays of target genes [*DPPA3* (*STELLA*), *EDR1* (*PHC1*), *GDF3*, and *NANOG*] and endogenous reference genes (*GAPDH*, *HPRT1*, and *TFRC*) in a mixed germ cell tumour cell line (NCCIT).

Target gene	**Sample (1:5 Dilution)**		**Triplicate Ct**											
			Target gene			**Endogenous (reference) gene**								
						GAPDH			***HPRT1***			***TFRC***		
			1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
DPPA3	**Water (NTC)**	**A**	U	U	U	U	U	U	U	U	U	U	U	U
	NCCIT	**B**	33.094	33.074	33.204	28.177	28.042	28.007	34.428	35.197	34.936	35.409	35.639	35.325
EDR1	**NCCIT**	**C**	32.149	32.263	32.176	27.521	27.849	27.836	34.458	34.367	34.730	35.692	35.683	36.025
	Water (NTC)	**D**	U	U	U	U	U	U	U	U	U	U	U	U
GDF3	**Water (NTC)**	**E**	U	U	U	U	U	U	U	U	U	U	U	U
	NCCIT	**F**	34.603	35.052	34.450	27.505	27.859	27.774	35.276	34.634	34.529	36.083	36.047	35.924
NANOG	**NCCIT**	**G**	32.427	32.340	33.332	27.990	27.940	27.758	35.133	35.051	34.436	35.372	35.963	35.788
	Water (NTC)	**H**	U	U	U	U	U	U	U	U	U	U	U	U

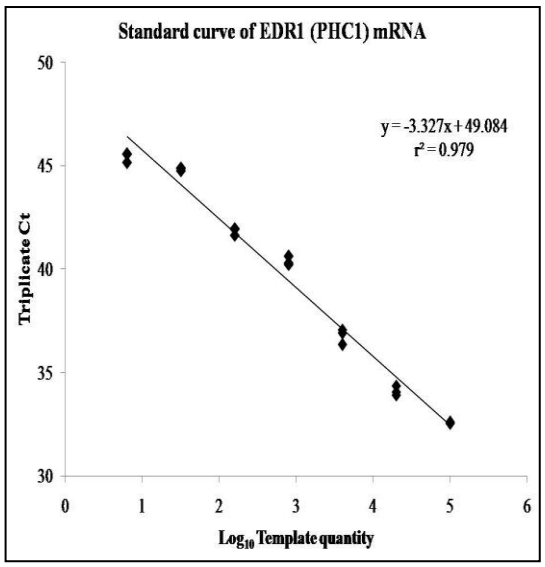
Note: U = Undetermined result

3.3.1.2. Standard curves of TaqMan® QRT-PCR assay

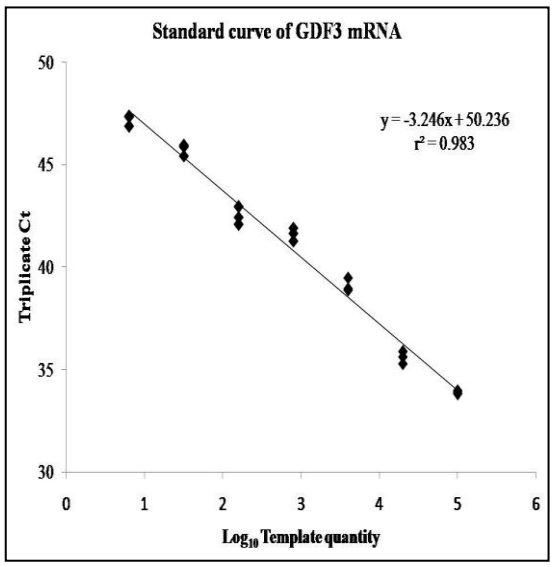
The standard curves of TaqMan® QRT-PCR assay for DPPA3, EDR1, GDF3, and NANOG mRNA were generated by triplicate threshold cycle (Ct) values for 7 serial dilutions of cDNA isolated from a mixed germ cell tumour cell line (NCCIT) as a template (Figure 3-1 and Appendix 3). The calculation of the QRT-PCR efficiency (E) from the slope (m) of each standard curve yielded specific amplification of the TaqMan® Gene Expression Assays with an acceptable efficiency of 95% – 105% (99.66% of DPPA3 mRNA; 99.79% of EDR1; 103.27% of GDF3 mRNA; and 99.17% of NANOG mRNA). The coefficient of determination (r^2) of each reaction was close to 1.0, indicating the QRT-PCR assay had the accuracy (validity) of the dilutions and precision (reproducibility) of pipetting. Since a y-intercept value between 33 and 37 cycles is correlated with 95% - 100% of amplification efficiency [see Section 2.2.4.6 (3) Standard curves of the real-time PCR assay in Chapter 2], but the y-intercept of each reaction in this study of mRNA expression was considerably higher than 37 cycles (50.442 cycles of DPPA3 mRNA; 49.084 cycles of EDR1; 50.236 cycles of GDF3 mRNA; and 48.550 cycles of NANOG mRNA), indicating that the determination of the amount of cDNA template would be inaccurate at this level (Table 3-2).



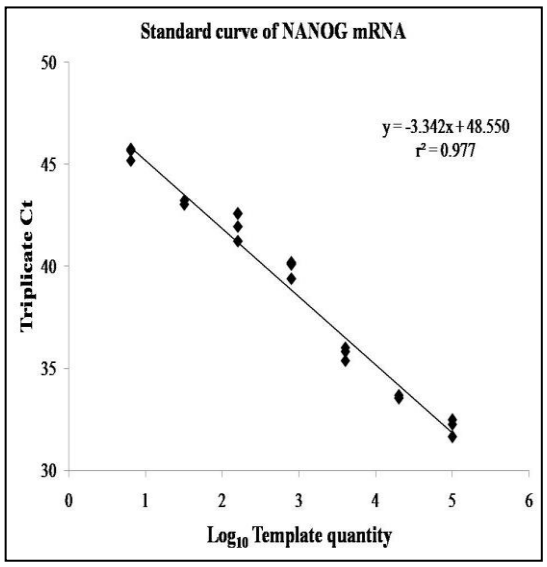
(A)



(B)



(C)



(D)

Figure 3-1 Standard curves of TaqMan® quantitative reverse transcriptase polymerase chain reaction (QRT-PCR) assay. The standard curves of TaqMan® QRT-PCR assay for DPPA3, EDR1, GDF3, and NANOG mRNA were constructed by triplicate Ct values for the 7 serial dilutions of the isolated cDNA isolated from a mixed germ cell tumour cell line (NCCIT) as a template. The assay was performed on 50 amplification cycles with manual threshold at 0.3 and automatic baseline. (A) DPPA3 mRNA; (B) EDR1 mRNA; (C) GDF3 mRNA; and (D) NANOG mRNA

Table 3-2 Standard curve parameters. The calculation of the linear regression on Microsoft Excel® programme was applied to the evaluation of the slope (m), the QRT-PCR efficiency (E), the y-intercept (b), and the coefficient of determination (r^2) for the standard curves in Figure 3-1.

mRNA	Slope (m)	QRT-PCR efficiency (E)* (%)	y-intercept (b) (Cycle)	Coefficient of determination (r^2)
DPPA3 (STELLA)	-3.330	99.66	50.442	0.983
EDR1 (PHC1)	-3.327	99.79	49.084	0.979
GDF3	-3.246	103.27	50.236	0.983
NANOG	-3.342	99.17	48.550	0.977

Note: * QRT-PCR efficiency (E) = $\{[10^{(-1/\text{Slope})}] - 1\} \times 100\%$

3.3.1.3. Normalisation, the optimal endogenous reference genes, and precision (reproducibility) of TaqMan® QRT-PCR assay

The results of TaqMan® QRT-PCR assays were normalised to HBL-100 as a non-neoplastic control; this cell line was derived from immortalised normal mammary epithelial cells in an early lactation. Referring to the previous reports on more accurate quantitative mRNA expression measurements (Thellin et al. 1999, Tricarico et al. 2002, Vandesompele et al. 2002, Warrington et al. 2000), the mean of each Ct value derived from the 3 endogenous reference genes, including *GAPDH*, *HPRT1*, and *TFRC*, was applied for calculation of the ΔCt value. Expressions of DPPA3, EDR1, GDF3, and NANOG mRNA in NCCIT, HBL-100, and MCF7 were determined twice to assess the precision (reproducibility) of the QRT-PCR assay. The paired assays showed no significant difference in the mean ΔCt values for mRNA expression in NCCIT, HBL-100, and MCF7 [ratio paired *t*-tests at the 95% confidence interval (CI), $p > 0.050$] (Table 3-3), showing good reproducibility.

Table 3-3 Precision (reproducibility) of TaqMan® QRT-PCR assay. Based on expression of DPPA3 (STELLA), EDR1 (PHC1), and NANOG mRNA in NCCIT, HBL-100, and MCF7, the ratio paired *t*-test at the 95% confidence interval (CI) was used for determination of precision (reproducibility) of TaqMan® QRT-PCR assay.

Expression of mRNA	**ΔCt**				**Ratio paired *t*-test at 95% CI**		
	1st assay		**2nd assay**				
	X	**SD**	**X**	**SD**	***t***	***df***	***p*-value**
DPPA3 (N = 3/3)	0.550	0.283	0.553	0.280	0.210	8	0.8392^{ns}
EDR1 (N = 3/3)	0.177	0.425	0.129	0.472	0.650	5	0.5446^{ns}
GDF3 (N = 1/3)	-0.191	0.182	-0.215	0.311	0.208	2	0.8547^{ns}
NANOG (N = 3/3)	0.541	0.015	0.609	0.068	2.035	2	0.1788^{ns}

Note: ns = No statistical significance

3.3.1.4. Expression of DPPA3 (STELLA), EDR1 (PHC1), GDF3, and NANOG mRNA in cell lines

A mixed germ cell tumour cell line (NCCIT) derived from teratocarcinoma was used as a positive control because this cell line usually express various stem cell markers, including *DPPA3* (*STELLA*) and *NANOG* genes (Pascal et al. 2009, Taranger et al. 2005). No GDF3 mRNA expressed in the following cell lines: HBL-100, MCF7, MDA-MB-231, MDA-MB-436, MDA-MB-468, T47D, and ZR-75-1. The unpaired *t*-test at the 95% CI was applied to compare the mean ΔCt value from all 8 cell lines. Four cell lines (NCCIT, MCF7, MDA-MB-436, and ZR-75-1) had a statistically significant mean ΔCt value for expression of DPPA3, EDR1, and NANOG mRNA that was less than HBL-100, indicating that mRNA expression was higher in these cell lines ($p \leq 0.050$). MDA-MB-231 and MDA-MB-468 showed no significant difference in ΔCt value for EDR1 mRNA from HBL-100 ($p > 0.050$). MDA-MB-231 showed no evidence of expression of DPPA3 and NANOG mRNA. Of note, a similar pattern of expression of DPPA3, EDR1, and NANOG mRNA was present in each cell line. T47D showed remarkably lower expression of these mRNA in comparison to NCCIT, HBL-100, MCF7, MDA-MB-436, MDA-MB-468, and particularly ZR-75-1 (Figure 3-2).

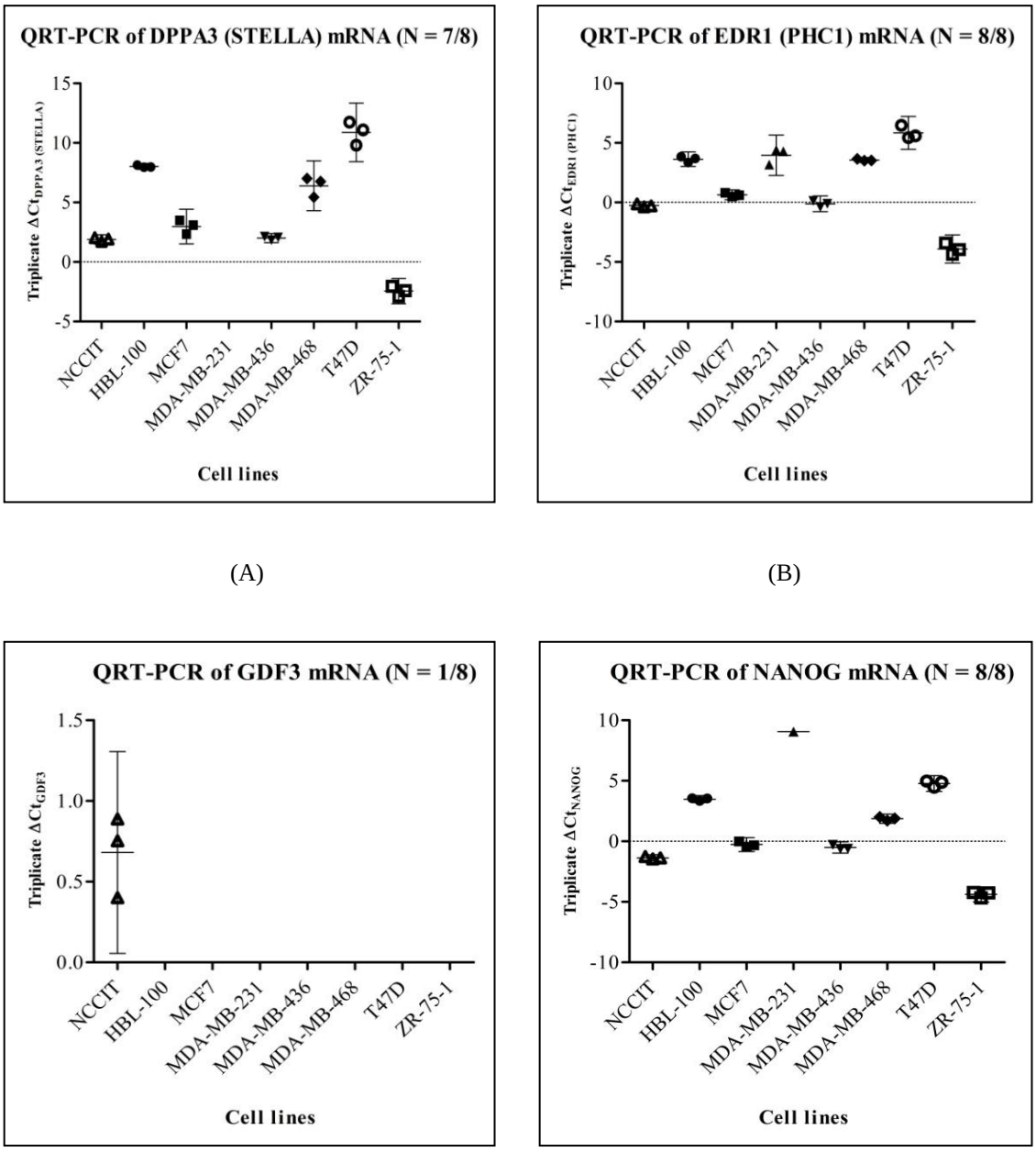


Figure 3-2 Expression of DPPA3 (STELLA), EDR1 (PHC1), GDF3, and NANOG mRNA in cell lines. The triplicate ΔCt values from 8 cell lines, including NCCIT, HBL-100, MCF7, MDA-MB-231, MDA-MB-436, MDA-MB-468, T47D, and ZR-75-1, were plotted with a horizontal line at the mean and error bars at the 95% CI. (A) DPPA3 mRNA; (B) EDR1 mRNA; (C) GDF3 mRNA; and (D) NANOG mRNA

3.3.1.5. Co-expression of DPPA3 (STELLA), EDR1 (PHC1), and NANOG mRNA in cell lines

Based on Pearson correlation coefficient (r) at the 95% CI, 7 cell lines (HBL-100, MCF7, MDA-MB-231, MDA-MB-436, MDA-MB-468, T47D, and ZR-75-1) showed a significant linear correlation of ΔCt values as follows: DPPA3 mRNA versus EDR1 mRNA ($p < 0.0001$); DPPA3 mRNA versus NANOG mRNA ($p < 0.0001$); and EDR1 mRNA versus NANOG mRNA ($p < 0.0001$) (Figure 3-3), indicating that DPPA3, EDR1, and NANOG mRNAs were co-expressed in breast cancer cell lines. Hence, investigation of co-expression of these putative stem cell genes in human breast cancers was indicated as worthwhile based on the cell line data.

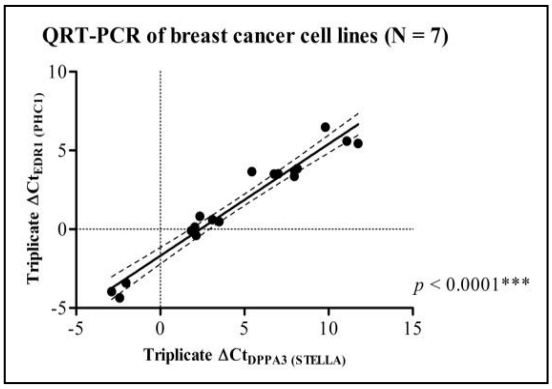
3.3.1.6. Expression of DPPA3 (STELLA), EDR1 (PHC1), and NANOG mRNA in normal breast (NB), surrounding normal breast (SNB), ductal carcinoma in situ (DCIS), and invasive breast carcinoma (IC) tissues

Expression of DPPA3, EDR1, and NANOG mRNA was evaluated in 28 NB, 38 SNB, 3 DCIS, and 52 IC tissues. The unpaired *t*-test at the 95% CI was applied to compare the mean ΔCt value from each of tissues (Figure 3-4). Only 2 NB tissues showed expression of DPPA3 mRNA. The SNB had higher expression of DPPA3 mRNA compared to the IC ($p < 0.0001$), but this result was based on only 9 SNB samples. There was no expression in any of 3 DCIS tissues. For expression of EDR1 mRNA, NB had the lowest expression, then higher in SNB ($p < 0.0001$) and IC ($p = 0.0108$). In addition, SNB had a high expression of EDR1 mRNA in comparison to IC ($p < 0.0001$). All 3 DCIS samples had expression of EDR1 mRNA at a similar level to IC. Expression of NANOG mRNA was infrequent in NB but there were more SNB cases showing expression. In addition, the level of this mRNA expression was

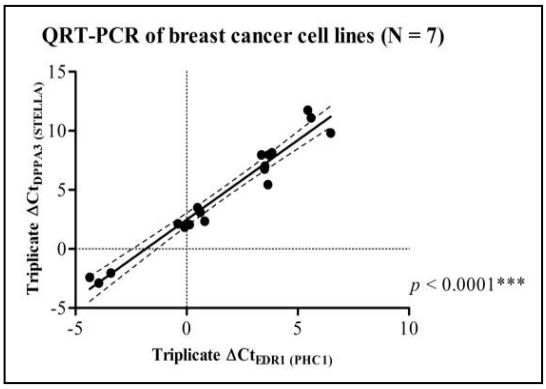
significantly higher in SNB than NB ($p < 0.001$). NANOG mRNA was expressed in the majority of IC samples and expression was at a similar level to NB ($p = 0.2330$). However, IC had a statistically significant lower expression of NANOG mRNA compared to SNB ($p < 0.0001$). Expression of NANOG mRNA was also detected in 2 DCIS samples. When each SNB was compared with its corresponding IC, the ratio paired *t*-test showed that mRNA expression of all 3 putative stem cell genes was significantly lower in IC ($p < 0.050$). Interestingly, there was a similar pattern of expression of DPPA3, EDR1, and NANOG mRNA in each tissue sample; the level of these mRNA expression was lowest in NB and higher in SNB but reduced again in IC.

3.3.1.7. Co-expression of DPPA3 (STELLA), EDR1 (PHC1), and NANOG mRNA in NB, SNB, DCIS, and IC

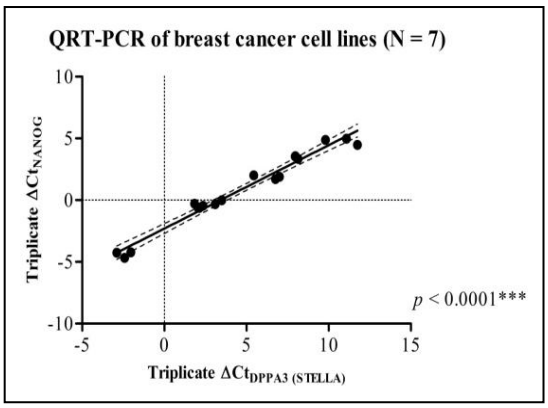
Expression of DPPA3 mRNA in NB, SNB, and IC was always detected in combination with both EDR1 and NANOG mRNA. Co-expression of EDR1 and NANOG mRNA was seen in NB, SNB, and IC. Only EDR1 mRNA had independent expression in 8 NB, 12 SNB, 1 DCIS, and 2 IC samples (Figure 3-5). There was a significant linear correlation of expression of EDR1 mRNA with NANOG mRNA in NB ($p = 0.0013$), SNB ($p < 0.0001$), and IC ($p < 0.0001$) (Figure 3-6).



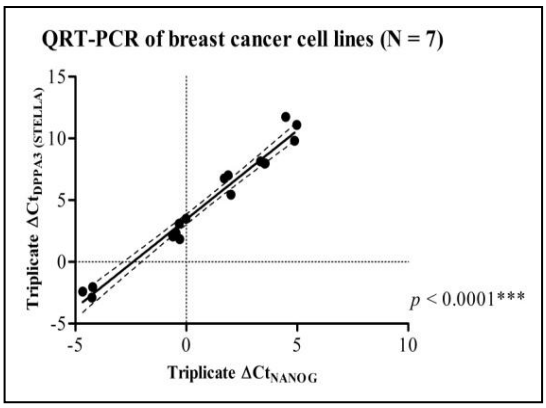
(A)



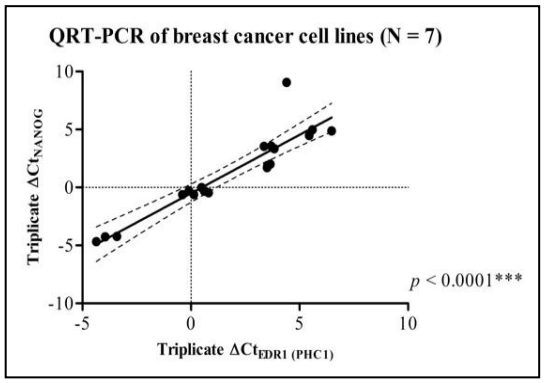
(B)



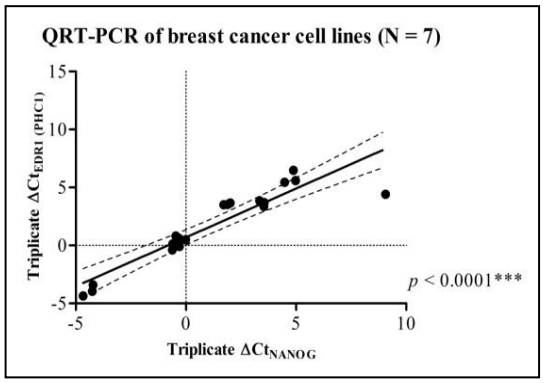
(C)



(D)

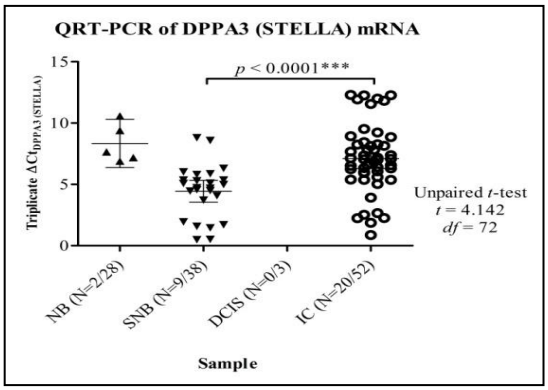


(E)

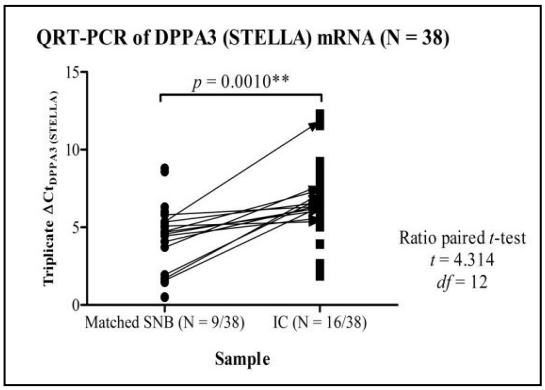


(F)

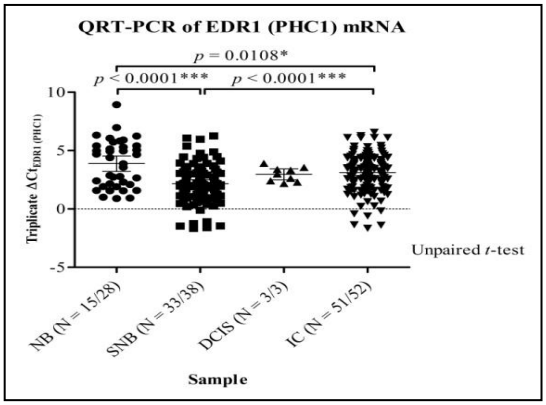
Figure 3-3 The correlation between expression of DPPA3 (STELLA), EDR1 (PHC1), and NANOG mRNA in cell lines. The Pearson correlation coefficient (r) at the 95% CI was used for analysis of the correlation of ΔCt values between DPPA3, EDR1, and NANOG mRNA in 7 cell lines (HBL-100, MCF7, MDA-MB-231, MDA-MB-436, MDA-MB-468, T47D, and ZR-75-1). (A) and (B) DPPA3 mRNA versus EDR1 mRNA; (C) and (D) DPPA3 mRNA versus NANOG mRNA; (E) and (F) EDR1 mRNA versus NANOG mRNA



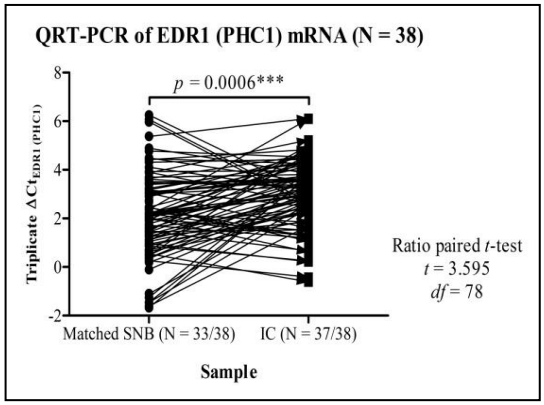
(A)



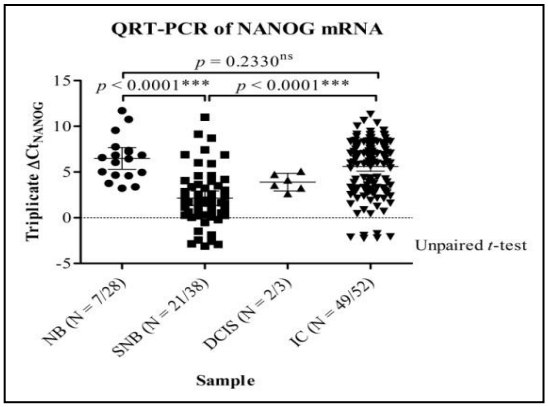
(B)



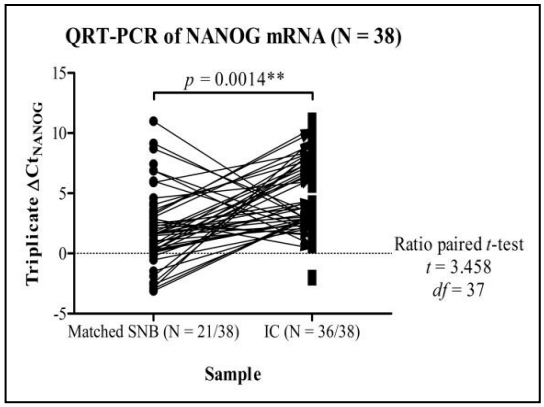
(C)



(D)



(E)



(F)

Figure 3-4 QRT-PCR analysis of DPPA3 (STELLA), EDR1 (PHC1), and NANOG mRNA in normal breast (NB), surrounding normal breast (SNB), ductal carcinoma in situ (DCIS), and invasive breast carcinoma (IC). Scatter diagrams were plotted from the triplicate ΔCt values for DPPA3, EDR1, and NANOG mRNA in 28 NB, 38 SNB, 3 DCIS, and 52 IC tissues. Results show a horizontal line at the mean and error bars at the 95% CI. The unpaired *t*-test was used for comparison with ΔCt values. In addition, SNB matched with IC were compared by using the ratio paired *t*-test. (A) and (B) DPPA3 mRNA; (C) and (D) EDR1 mRNA; (E) and (F) NANOG mRNA

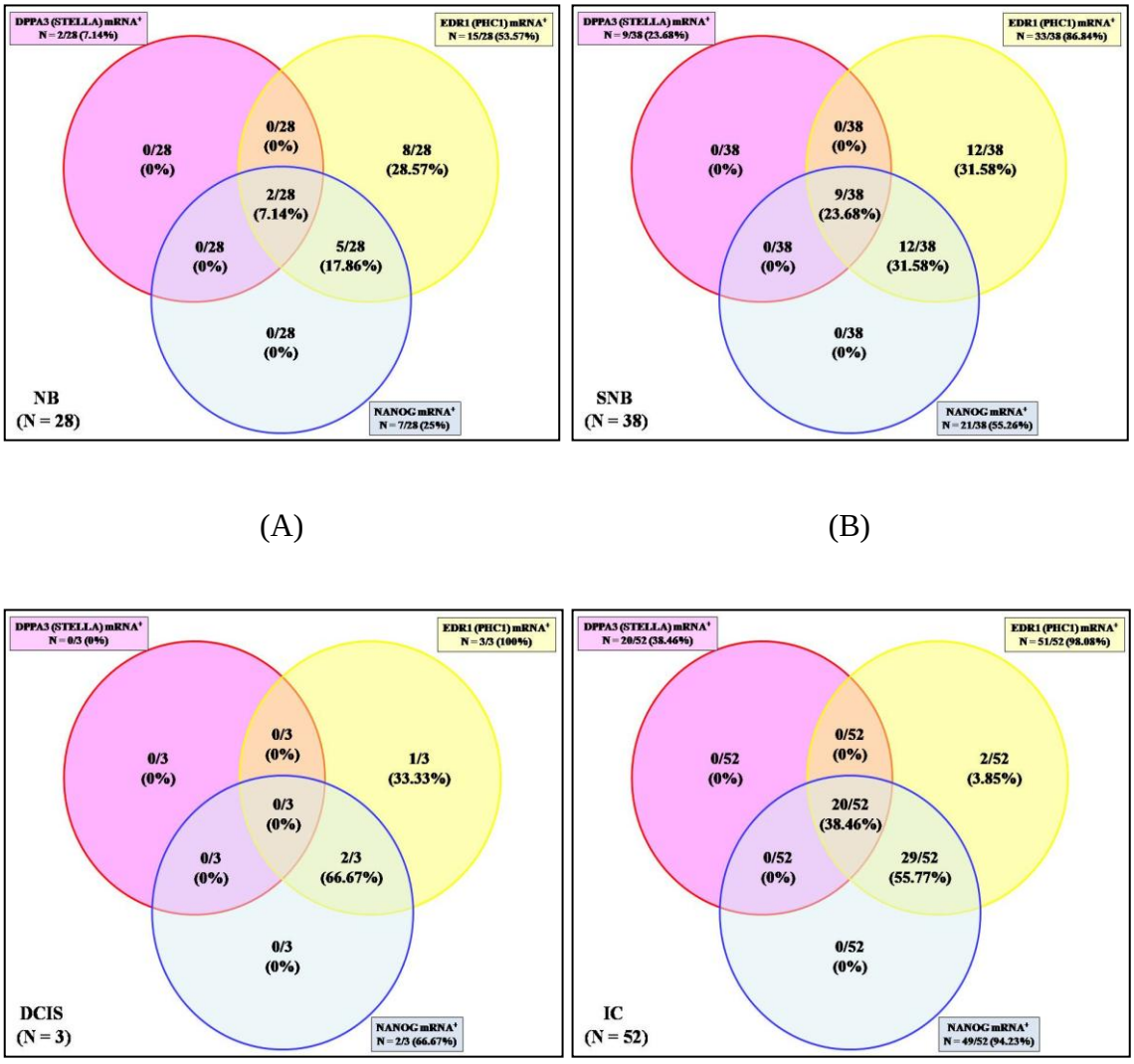
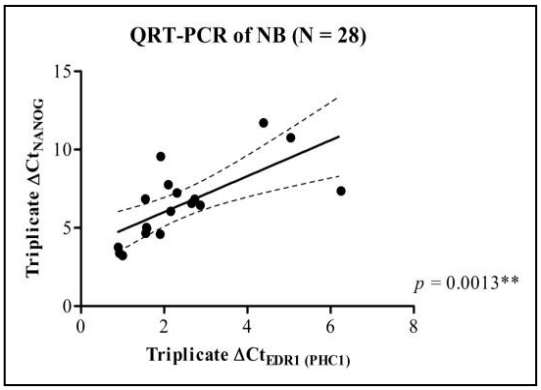
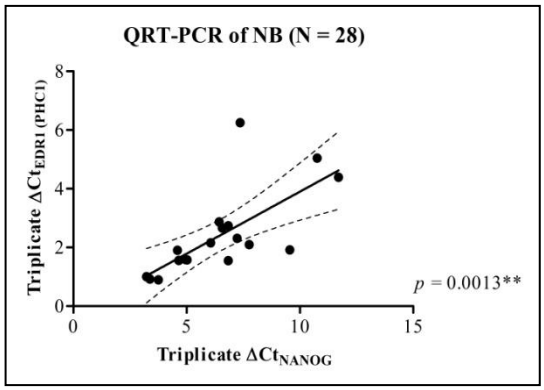


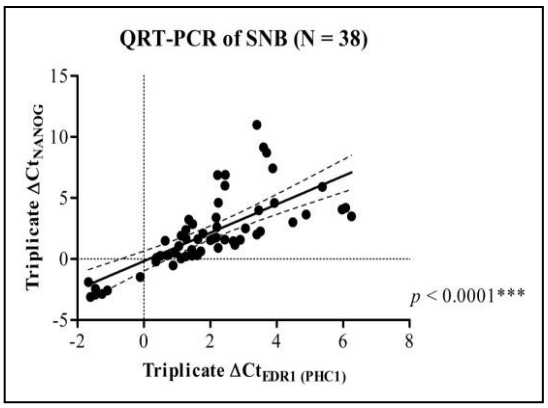
Figure 3-5 Co-expression of DPPA3 (STELLA), EDR1 (PHC1), and NANOG mRNA in NB, SNB, DCIS, and IC. Venn diagrams show the number of tissue samples (NB, SNB, DCIS, and IC) with positive expression and co-expression of DPPA3, EDR1, and NANOG mRNA. (A) NB; (B) SNB; (C) DCIS; and (D) IC



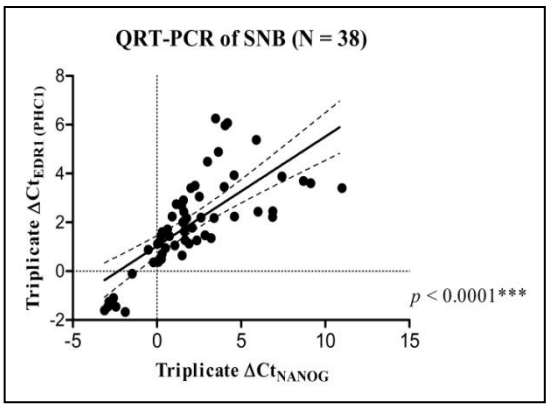
(A)



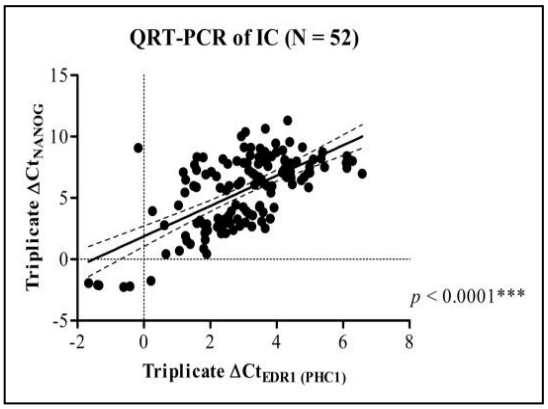
(B)



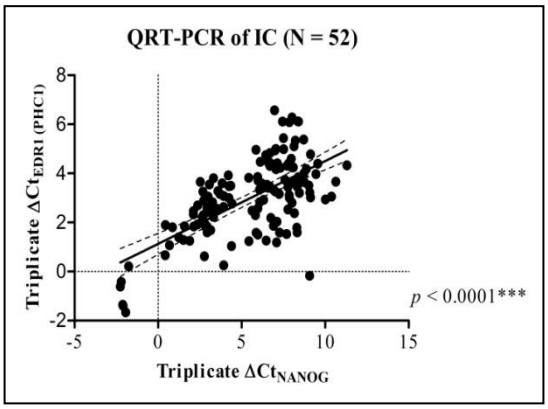
(C)



(D)



(E)



(F)

Figure 3-6 The correlation between EDR1 (PHC1) and NANOG mRNA in NB, SNB, and IC. The Pearson correlation coefficient (r) at the 95% CI was used for analysis of the correlation of ΔCt values between EDR1 and NANOG mRNA in 28 NB, 38 SNB, and 52 IC tissues at the 95% CI. (A) and (B) NB; (C) and (D) SNB; (E) and (F) IC

3.3.1.8. The effect of a percentage of invasive carcinoma component in breast cancer tissue section on the QRT-PCR results

The percentage of invasive carcinoma presented within tissue sections (≤ 50% and > 50%) was related to expression of DPPA3, EDR1, and NANOG mRNA for 39 of 52 (75%) breast cancer cases, using the unpaired *t*-test at the 95% CI (Table 3-4). There were 9 and 30 cases showing the invasive carcinoma component of ≤ 50% and > 50%, respectively. This assessment of the percentage was performed by myself. Breast cancers that expressed DPPA3 mRNA were not analysed since only 1 of the 9 case with the invasive carcinoma component of ≤ 50% showed expression of this mRNA. The statistical analysis showed that the ΔCt values for expression of EDR1 and NANOG mRNA showed no relationship with percentage of invasive carcinoma component using a 50% cut-off ($p > 0.050$).

Table 3-4 The effect of a percentage of invasive carcinoma component in breast cancer tissue section on the QRT-PCR results. The unpaired *t*-test at the 95% CI was applied to the analysis of the relationship between a percentage of invasive carcinoma component and the ΔCt values for expression of DPPA3 (STELLA), EDR1 (PHC1), and NANOG mRNA in 39 breast tissue samples.

Expression of mRNA [N (%)]	**Percentage of invasive carcinoma component (N = 39)**						***p*-value[#]**
	≤ 50% [N = 9/39 (23.08%)]			**> 50% [N = 30/29 (76.92%)]**			
	N (%)	**$X_{\Delta Ct}$**	**$SD_{\Delta Ct}$**	**N (%)**	**$X_{\Delta Ct}$**	**$SD_{\Delta Ct}$**	
DPPA3 [N = 9/39 (23.08%)]	1/39 (2.56%)	8.255	0.018	8/39 (20.51%)	7.757	4.691	NA
EDR1 [N = 38/39 (97.44%)]	9/39 (23.08%)	2.752	1.354	29/39 (74.36%)	3.398	1.725	0.0823^{ns}
NANOG [N = 36/39 (92.31%)]	9/39 (23.08%)	5.683	2.878	27/39 (69.23%)	6.782	2.968	0.1161^{ns}

Note: # Unpaired *t*-test at the 95% CI; NA = No statistical analysis; and ns = No statistical significance

3.3.1.9. The correlation of expression of DPPA3 (STELLA), EDR1 (PHC1), and NANOG mRNA in invasive breast carcinoma (IC) with clinicopathological features

The ΔCt values for expression of DPPA3, EDR1, and NANOG mRNA in 52 IC tissue samples were analysed for any correlation with clinicopathological features [the age (≤ 50 years old and > 50 years old); tumour size (≤ 2.0 cm and > 2 cm); tumour grade/differentiation; axillary lymph node metastasis; and the status of ER, PR, and HER2 (ERBB2)], using the unpaired *t*-test at the 95% CI (Table 3-5 to 3-7). However, no information was available for the status of ER and PR in one case and for the status of HER2 in 36 (69.23%) cases. Hence, the final number of IC used for evaluation were as follows: (I) 52 cases for the relationship to the age group, tumour size, tumour grade/differentiation, and axillary lymph node metastasis; (II) 51 cases for the relationship to the status of ER and PR; and (III) 16 cases for the relationship to HER2 status.

For expression of DPPA3 mRNA (Table 3-4), the presence of axillary lymph node metastasis was significantly correlated with a higher level of this mRNA expression ($p = 0.0029$). There was no correlation between expression of DPPA3 mRNA and the following features: the age group ($p = 0.9658$), tumour size ($p = 0.6103$), tumour grade/differentiation ($p = 0.7944$), and the status of ER ($p = 0.0875$) and PR ($p = 0.0875$). Nevertheless, these results were based on only 4 – 16 IC samples with detectable expression of DPPA3 mRNA in each individual group of the clinicopathological features. No statistical analysis was applied for the relationship to HER2 status because 4 of the 16 cases with this mRNA expression showed a lack of HER2 expression and none were HER2-positive tumour.

The majority of cancers [N = 51/52 (98.08%)] showed that expression of EDR1 mRNA did not correlate with tumour size (p = 0.4179). Approximately 60% of high-graded (grade III)/poorly differentiated carcinoma had a lower level of expression of EDR1 mRNA compared to cancers with low-grade (grade I and II)/well and moderately differentiation (Table 3-5). In addition, low level of expression of NANOG mRNA related to size greater than 2.0 cm (p = 0.0270) and high-grade tumour (p = 0.0002) (Table 3-6).

There was no statistically significant correlation between mRNA expression for both *EDR1* and *NANOG* genes and the other clinicopathological features such as the age group, axillary lymph node metastasis, and the status of ER and PR ($p > 0.050$). The 16 cancers with known HER2 status showed that expression of EDR1 mRNA was not related (p = 0.0533), but expression of NANOG mRNA was lower in the small number of cases with positive HER2 expression (p = 0.0117).

Table 3-5 The correlation of the ΔCt value for expression of DPPA3 (STELLA) mRNA in the IC tissue with clinicopathological features. The unpaired *t*-test at the 95% CI was applied to the analysis of the correlation between the ΔCt value for expression of DPPA3 mRNA in 52 IC tissue samples and their clinicopathological features.

Clinicopathological features	All ICs (N = 52)					
	Triplicate $\Delta Ct_{DPPA3\ (STELLA)\ mRNA}$			Unpaired *t*-test (95% CI)		
	N (%)	X	SD	*t*	*df*	*p*-value
Age (Years old)						
≤ 50 [N = 32/52 (61.54%)]	10/52 (19.23%)	7.124	3.062	0.043	47	0.9658^{ns}
> 50 [N = 20/52 (38.46%)]	10/52 (19.23%)	7.089	2.626			
Tumour size (cm)						
≤ 2.0 [N = 17/52 (32.69%)]	9/52 (17.31%)	7.328	2.294	0.513	47	0.6103^{ns}
> 2.0 [N = 35/52 (67.31%)]	11/52 (21.15%)	6.911	3.245			
Tumour grade/differentiation						
Low (I & II)/Well & Moderate [N =22/52 (42.31%)]	11/52 (21.15%)	7.014	1.198	0.262	47	0.7944^{ns}
High (III)/Poor [N = 30/52 (57.69%)]	9/52 (17.31%)	7.229	4.135			
Axillary lymph node						
No metastasis [N = 26/52 (50.00%)]	13/52 (25.00%)	7.954	2.297	3.147	47	**0.0029****
Metastasis [N = 26/52 (50.00%)]	7/52 (13.46%)	5.510	3.075			
ER						
Negative [N = 10/51 (19.61%)]	4/51 (7.84%)	5.747	4.865	1.745	47	0.0875^{ns}
Positive [N = 41/51 (80.39%)]	16/51 (31.37%)	7.455	1.953			
No information [N = 1/52 (1.92%)]						
PR						
Negative [N = 10/51 (19.61%)]	4/51 (7.84%)	5.747	4.865	1.745	47	0.0875^{ns}
Positive [N = 41/51 (80.39%)]	16/51 (31.37%)	7.455	1.953			
No information [N = 1/52 (1.92%)]						
HER2 (ERBB2)						
Negative [N = 11/16 (68.75%)]	4/16 (25.00%)	4.567	3.543	No statistical analysis		
Positive [N = 5/16 (31.26%)]	0/16 (0.00%)	-	-			
No information [N = 36/52 (69.23%)]						

Note: ns = No statistical significance

Table 3-6 The correlation of the ΔCt value for expression of EDR1 (PHC1) mRNA in the IC tissue with clinicopathological features. The unpaired *t*-test at the 95% CI was applied to the analysis of the correlation between the ΔCt value for expression of EDR1 mRNA in 52 IC tissue samples and their clinicopathological features.

Clinicopathological features	All ICs (N = 52)					
	Triplicate $\Delta Ct_{EDR1\ (PHC1)\ mRNA}$			Unpaired *t*-test (95% CI)		
	N (%)	X	SD	*t*	*df*	*p*-value
Age (Years old)						
≤ 50 [N = 32/52 (61.54%)]	31/52 (59.62%)	3.248	1.556	1.350	150	0.1790^{ns}
> 50 [N = 20/52 (38.46%)]	20/52 (38.46%)	2.907	1.459			
Tumour size (cm)						
≤ 2.0 [N = 17/52 (32.69%)]	16/52 (30.77%)	2.968	1.241	0.812	150	0.4179^{ns}
> 2.0 [N = 35/52 (67.31%)]	35/52 (67.31%)	3.184	1.639			
Tumour grade/differentiation						
Low (I & II)/Well & Moderate [N =22/52 (42.31%)]	21/52 (40.38%)	2.778	1.178	2.298	150	**0.0229***
High (III)/Poor [N = 30/52 (57.69%)]	30/52 (57.69%)	3.348	1.689			
Axillary lymph node						
No metastasis [N = 26/52 (50.00%)]	25/52 (48.08%)	3.270	1.338	1.240	150	0.2169^{ns}
Metastasis [N = 26/52 (50.00%)]	26/52 (50.00%)	2.965	1.680			
ER						
Negative [N = 10/51 (19.61%)]	10/51 (19.61%)	3.273	2.391	0.755	147	0.4517^{ns}
Positive [N = 41/51 (80.39%)]	40/51 (78.43%)	3.038	1.223			
No information [N = 1/52 (1.92%)]						
PR						
Negative [N = 10/51 (19.61%)]	10/51 (19.61%)	3.460	2.458	1.512	147	0.1327^{ns}
Positive [N = 41/51 (80.39%)]	40/51 (78.43%)	2.991	1.176			
No information [N = 1/52 (1.92%)]						
HER2 (ERBB2)						
Negative [N = 11/16 (68.75%)]	11/16 (68.75%)	2.789	2.198	1.983	46	0.0533^{ns}
Positive [N = 5/16 (31.25%)]	5/16 (31.25%)	3.971	0.995			
No information [N = 36/52 (69.23%)]						

Note: ns = No statistical significance

Table 3-7 The correlation of the ΔCt value for expression of NANOG mRNA in the IC tissue with clinicopathological features. The unpaired *t*-test at the 95% CI was applied to the analysis of the correlation between the ΔCt value for expression of NANOG mRNA in 52 IC tissue samples and their clinicopathological features.

Clinicopathological features	All ICs (N = 52)					
	Triplicate $\Delta Ct_{NANOG\ mRNA}$			Unpaired *t*-test (95% CI)		
	N (%)	X	SD	*t*	*df*	*p*-value
Age (Years old)						
≤ 50 [N = 32/52 (61.54%)]	29/52 (55.77%)	5.966	2.750	1.700	134	0.0914^{ns}
> 50 [N = 20/52 (38.46%)]	20/52 (38.46%)	5.085	3.265			
Tumour size (cm)						
≤ 2.0 [N = 17/52 (32.69%)]	15/52 (28.85%)	4.757	2.514	2.236	134	**0.0270***
> 2.0 [N = 35/52 (67.31%)]	34/52 (65.38%)	5.981	3.122			
Tumour grade/differentiation						
Low (I & II)/Well & Moderate [N =22/52 (42.31%)]	20/52 (38.46%)	4.481	2.285	3.778	134	**0.0002*****
High (III)/Poor [N = 30/52 (57.69%)]	29/52 (55.77%)	6.366	3.184			
Axillary lymph node						
No metastasis [N = 26/52 (50.00%)]	23/52 (44.23%)	5.347	2.746	0.942	134	0.3480^{ns}
Metastasis [N = 26/52 (50.00%)]	26/52 (50.00%)	5.831	3.198			
ER						
Negative [N = 10/51 (19.61%)]	10/51 (19.61%)	5.371	4.276	0.365	131	0.7157^{ns}
Positive [N = 41/51 (80.39%)]	38/51 (74.51%)	5.613	2.645			
No information [N = 1/52 (1.92%)]						
PR						
Negative [N = 10/51 (19.61%)]	10/51 (19.61%)	5.349	4.155	0.416	131	0.6779^{ns}
Positive [N = 41/51 (80.39%)]	38/51 (74.51%)	5.621	2.671			
No information [N = 1/52 (1.92%)]						
HER2/neu (c-erb B2)						
Negative [N = 11/16 (68.75%)]	11/16 (68.75%)	5.148	3.891	2.631	44	**0.0117***
Positive [N = 5/16 (31.25%)]	5/16 (31.25%)	7.977	1.407			
No information [N = 36/52 (69.23%)]						

Note: ns = No statistical significance

3.3.1.10. Summary of findings for DPPA3 (STELLA), EDR1 (PHC1), and NANOG mRNAs expression in breast tissue

Expression of DPPA3, EDR1, and NANOG mRNA was lowest in NB, higher in SNB, and lower in IC. DCIS had no expression of DPPA3 but showed expression of EDR1 and NANOG mRNA. EDR1 mRNA was independently expressed in 28% of NB, 32% of SNB, and 4% of IC. Co-expression of EDR1 and NANOG mRNAs in the breast was detected in descending order as follows: IC (56%), SNB (32%), and NB (18%). Concordant expression of all 3 putative stem cell genes was infrequently present in NB (7%) but identified in SNB (24%) and IC (38%), respectively.

The invasive breast carcinoma with axillary lymph node metastasis (13%) was associated with a high level of DPPA3 mRNA expression. Almost 60% of the high-graded (grade III)/poorly differentiated carcinomas had a lower expression of EDR1 and NANOG mRNA. Approximately 65% of IC had a correlation between tumour size greater than 2.0 cm and a lower level of NANOG mRNA expression. The small number of HER2-positive tumours also had lower expression of NANOG mRNA. The other clinicopathological features such as the patient age group and expression of ER and PR had no significant relationship to expression of DPPA3, EDR1, and NANOG mRNA.

3.3.2. *Analysis of protein expression*

3.3.2.1. Western blotting

3.3.2.1 (1). Protein concentration

The concentration of protein isolated from 8 cell lines and 3 breast carcinoma tissues was calculated on the basis of a standard curve of the corrected UV absorbance at 562 nm using the Pierce® Bicinchoninic Acid (BCA) Protein Assay Kit (Appendix 4 – 6). The mean protein concentration for 7 BCA CLs and 3 frozen BC tissues was 1.555 and 0.874 μg/μL, respectively.

3.3.2.1 (2). Preliminary study of expression of DPPA3 (STELLA), EDR1 (PHC1), and NANOG proteins in cell lines and breast carcinoma tissues

The preliminary results of Western blotting for 8 cell lines (NCCIT used as a positive control) and 3 breast carcinoma tissues are shown in Figure 3-7, Vinculin was used as the loading control showing a band at 116 kDa. For DPPA3, the expected 20 kDa band was not seen in all breast cancer cell lines (BCA CLs) and breast carcinoma tissues (BC). However, 6 BCA CLs (HBL-100, MCF7, MDA-MB-436, MDA-MB-468, T47D, and ZR-75-1) and all 3 BC tissues revealed additional bands at molecular weights of 28 – 56 kDa, but there was no additional band seen for MDA-MB-231. NCCIT (lane 1) had a very faint unexpected band at 25 kDa. EDR1 was identified at the expected molecular weight of 43 kDa in all cell lines and tissue with an additional band at 56 kDa. NANOG was found in NCCIT lane (lane 1) at the expected 42 kDa weight with an additional band at 33 kDa. Five BCA CLs (HBL-100, MCF7, MDA-MB-436, MDA-MB-468, and ZR-75-1) and a sample of DCIS with IDC-NST showed differing levels of NANOG at molecular weight 31 – 33 kDa. There was no expression of

NANOG in MDA-MB-231 and T47D. The sample of DCIS with IDC-NST showed high expression of both EDR1 and NANOG. DCIS and IDC-NST tissues had different protein sizes of EDR1 at 56 kDa but they had no expression of NANOG.

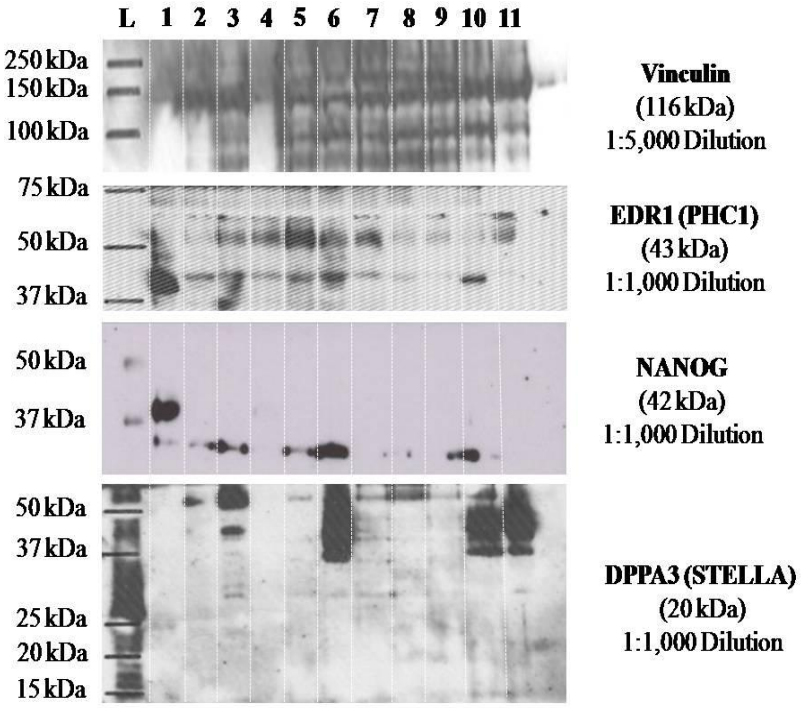


Figure 3-7 Western blotting for expression of DPPA3 (STELLA), EDR1 (PHC1), and NANOG proteins in 8 cell lines and 3 breast carcinoma tissues. Expression of DPPA3, EDR1, and NANOG proteins in one mixed germ cell tumour cell line (NCCIT), 7 breast cancer cell lines (HBL-100, MCF7, MDA-MB-231, MDA-MB-436, MDA-MB-468, T47D, and ZR-75-1) and 3 different samples of breast carcinoma tissues (DCIS, DCIS & IDC-NST, and IDC-NST) was detected by Western blotting technique. Vinculin protein was used as a loading equivalent control. NCCIT was a positive sample control for protein expression of these putative stem cell genes. Twenty-five μg of protein sample was loaded in each lane. The protein bands were visualised after 15 minutes exposure to an X-ray film. L = Protein ladder; Lane 1 = NCCIT; Lane 2 = HBL-100; Lane 3 = MCF7; Lane 4 = MDA-MB-231; Lane 5 = MDA-MB-436; Lane 6 = MDA-MB-468; Lane 7 = T47D; Lane 8 = ZR-75-1; Lane 9 = DCIS; Lane 10 = DCIS & IDC-NST; and Lane 11 = IDC-NST

From this preliminary Western blotting results, there were 2 mainly important problems, including high background and the presence of unexpected protein bands. Hence, the optimisation of Western blotting should to be rigorously applied prior to evaluating protein expression. The unsatisfying results were likely caused by non-specific binding associated with incomplete blocking, insufficient washing, and/or inappropriate concentrations of primary and/or secondary antibodies. The possible solutions were higher concentrations of blocking and/or washing buffers, the prolonged blocking and/or washing steps, the use of another blocking buffer such as bovine serum antigen (BSA), the assessment of suitable dilutions of primary and/or secondary antibodies, and the use of antibodies from different suppliers (Thermo Fisher Scientific 2010).

Theoretically, an ideal protein used as the internal reference (loading) control is correlated with the normalisation of the detecting signal from the proteins of interest because of equality of the loaded protein amount in each lane and the assay efficiency (http://www.labome.com/method/Loading-Controls-for-Western-Blots.html). Besides Vinculin (116 kDa of predicted band size), the other proteins are commonly used as the loading control in Western blotting of breast cancer research: Actin [particularly Beta (β)-actin] (approximately 42 kDa band size) (DeGraffenried et al. 2004, Liang, Brekken & Hyder 2006, Malkas et al. 2006), Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (approximately 40 kDa band size) (Lin et al. 2011, Gomes et al. 2012), and Tubulin [both Alpha (α)- and Beta (β)-tubulin] (approximately 50 kDa band size) (Drabsch, Robert & Gonda 2010, Ellison-Zelski, Alarid 2010, Tomlinson, Knowles & Speirs 2012). Usually, different sample preparations and experimental conditions require different loading controls. In addition, the predicted band size of the loading control should be different from the detecting bands of the protein samples (

http://www.labome.com/method/Loading-Controls-for-Western-Blots.html, Yu et al. 2011). Thus, Vinculin was able to use as the loading control in this study because the expected band sizes of DPPA3, EDR1, and NANOG were 20, 43, and 42 kDas, respectively. Nevertheless, this Western blotting evaluated an expression of proteins, that were isolated from both cell lines and tissues, on the same time. Therefore, expression of DPPA3, EDR1, and NANOG proteins in different sample types should be assessed on the separate immunoblotting membranes. Additionally, the appropriate loading controls might be optimised and selected for each of protein samples and individual sample preparations, for example based on the predicted band sizes, Actin, GAPDH, and Tubulin could be used as the loading control for DPPA3; and only Tubulin could be used as the loading control for EDR1 and NANOG.

Although expression of DPPA3, EDR1, and NANOG proteins in BCA CLs and some BC tissues did not correlate with mRNA expression from the QRT-PCR results, variations in the level of these protein expressions should be investigated further in more invasive breast carcinoma tissues.

3.3.2.2. Immunohistochemistry (IHC)

3.3.2.2 (1). Optimisation

The primary antibodies against DPPA3 (STELLA), EDR1 (PHC1), and NANOG proteins were tested to find the most appropriate immunostaining conditions: antigen retrieval (AR) method, the AR time, and the concentration of primary antibody.

3.3.2.2 (1.1). *Antigen retrieval (AR) methods*

Heat-induced epitope retrieval (HIER) method was used with a comparison of microwave and pressure cooker for DPPA3, EDR1, and NANOG proteins (Table 3-7). One randomly selected FFPE tissue sample of IDC-NST was used. The preliminary immunostaining was performed with a 1:150 dilution of DPPA3, a 1:50 dilution of EDR1, and a 1:30 dilution of NANOG. These antibody dilutions were randomly selected from the manufacturer's instructions.

Immunoexpression of DPPA3 and NANOG was mainly in the cytoplasm of cancer cells, but EDR1 immunostaining was predominantly present in the nuclei of cancer cells (Figure 3-8). Both microwave and pressure cooker AR methods revealed mostly weak intensity of DPPA3 and their H-scores were 90 and 95, respectively. EDR1 and NANOG had a higher H-score and stronger intensity with pressure cooker AR, particularly expression of EDR1. Therefore, on the basis of higher H-score values, pressure cooker was the AR method of choice for immunoexpression of DPPA3, EDR1, and NANOG proteins.

Table 3-8 A comparison of heat-induced epitope retrieval (HIER) methods between microwave and pressure cooker. The optimal AR technique was determined between microwave and pressure cooker for immunohistochemical study of DPPA3 (STELLA), EDR1 (PHC1), and NANOG proteins in invasive breast carcinoma tissue.

Parameter			**HIER method**	
			Microwave	**Pressure cooker**
Setting	**Equipment**		Tecnolec® Superwave 750	Pascal®
	AR buffer		1x 10 mM citrate buffer of pH 6.0	
	Temperature		Power level 7	SP1: 120 °C; SP2: 90 °C; SP limit: 10 °C
	Heating time		20 minutes	SP1: 30 seconds; SP2: 10 seconds
Sample			FFPE tissue of IDC-NST	
Immunostaining	**Kit**		NovoLink™ Polymer Detection System	
	Dilution	DPPA3	1:150	
		EDR1	1:50	
		NANOG	1:30	
	Location	DPPA3	Cytoplasm	Cytoplasm
		EDR1	Nucleus	Nucleus
		NANOG	Cytoplasm	Cytoplasm
	Intensity (%)	DPPA3	NoS (10%); Weak (90%)	NoS (5%); Weak (95%)
		EDR1	NoS (5%); Weak (5%); Moderate (95%)	Moderate (5%); Strong (95%)
		NANOG	NoS (5%); Weak (95%)	Weak (5%); Moderate (95%)
	H-score	DPPA3	90	95
		EDR1	195	295
		NANOG	95	195

Note: # Pascal® Microprocessor Controlled Pressure Chamber; NoS = No staining; and SP = Setting program

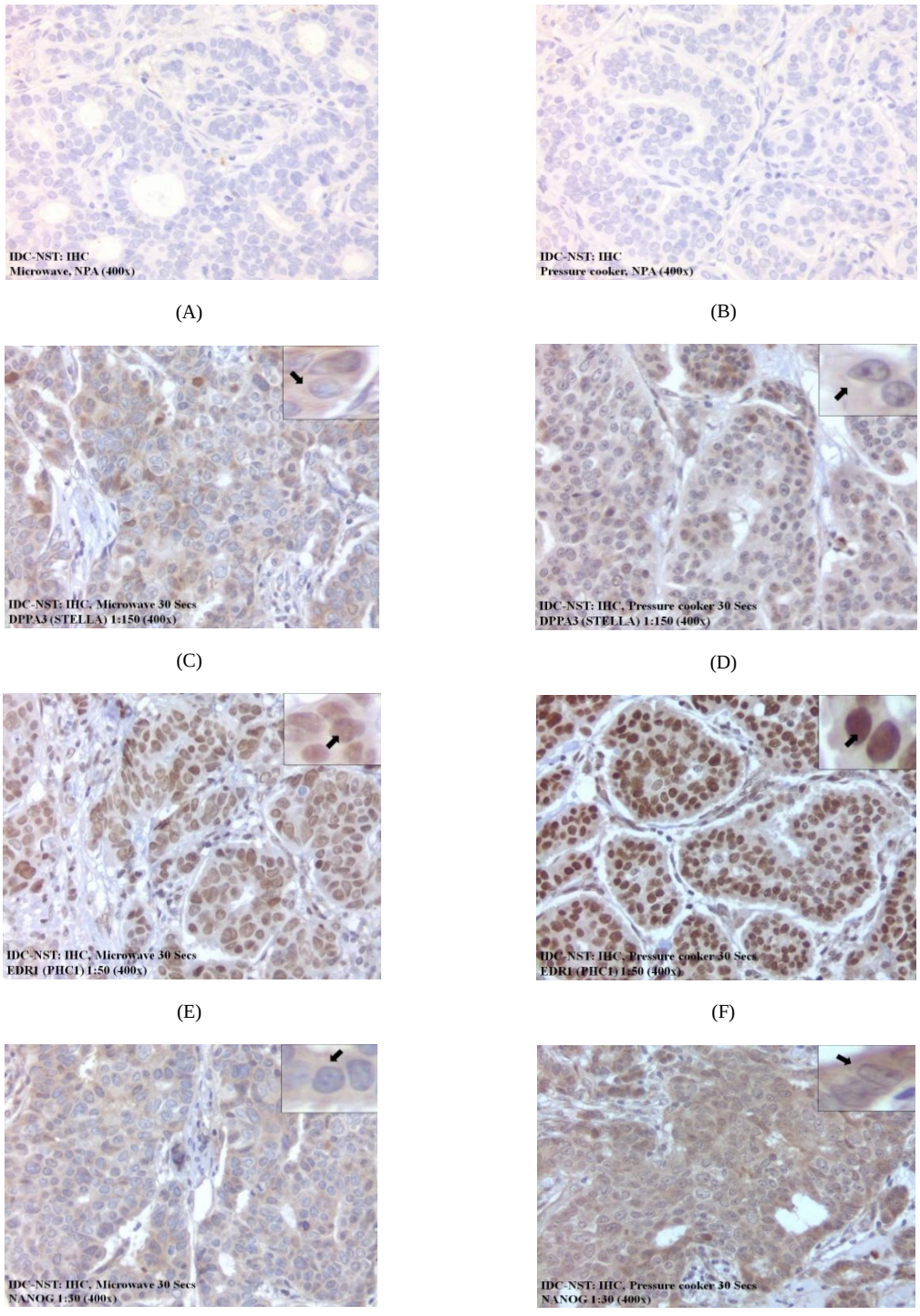


Figure 3-8 Immunohistochemistry (IHC) of invasive breast carcinoma (IC) tissue based on microwave and pressure cooker antigen retrieval (AR) methods. The immunoreactivity of DPPA3, EDR1, and NANOG proteins were assessed in the IDC-NST tissue using microwave (MW) (Left) and pressure cooker (PC) (Right) antigen retrieval methods. The main location of immunostaining (Arrow) is illustrated with the inset. (A) MW with no primary antibody (NPA); (B) PC with NPA; (C) MW with DPPA3; (D) PC with DPPA3; (E) MW with EDR1; (F) PC with EDR1; (G) MW with NANOG; and (H) PC with NANOG

3.3.2.2 (1.2). ***The heating time for pressure cooker antigen retrieval***

The optimal heating time for pressure cooker AR was evaluated for immunoreactivity of DPPA3, EDR1, and NANOG proteins in the same FFPE tissue sample as described in Section 3.3.2.2 (1.1). The AR times of 30, 45, and 60 seconds were tested with the same primary antibody concentration as described in the aforementioned Section (Table 3-9 and Figure 3-9). For 30, 45, and 60 seconds of AR, the cancer cells showed the same H-score of 95 for weak cytoplasmic staining of DPPA3. The EDR1-immunostained cancer cells yielded mainly strong nuclear staining and the same H-score of 295 for the AR time of 30 and 45 seconds. On the contrary, the AR time of 60 seconds for EDR1 gave mostly moderate nuclear staining and lower H-score of 210. For 30 seconds AR, NANOG showed weak cytoplasmic staining and the H-score of 95. The immunoreactivity of NANOG demonstrated slightly stronger cytoplasmic staining and the same H-score of 105 for the heating times of 45 and 60 seconds. In addition, there was occasionally weak nuclear staining for 60 seconds AR.

Hence, the appropriate heating time for pressure cooker AR was considered from the higher H-score value and the immunostaining appearances: (I) 30 seconds pressure cooker AR was selected for DPPA3 and EDR1, which would reduce the risk of section loss that can occur with longer AR; (II) 45 seconds AR for NANOG was chosen, which yielded more stronger and clearly cytoplasmic staining.

Table 3-9 A comparison of the heating time for pressure cooker antigen retrieval (AR) method. The optimal heating time for pressure cooker AR method was determined for immunohistochemical study of DPPA3 (STELLA), EDR1 (PHC1), and NANOG proteins in invasive breast carcinoma tissue. The primary antibody concentration was the same as described in Table 3-8.

Immunostaining		**Heating time for pressure cooker AR (Seconds)**		
		30	**45**	**60**
Sample		FFPE tissue of IDC-NST		
Dilution	DPPA3	1:150		
	EDR1	1:50		
	NANOG	1:30		
Location	DPPA3	Cytoplasm	Cytoplasm	Cytoplasm
	EDR1	Nucleus	Nucleus	Nucleus
	NANOG	Cytoplasm	Cytoplasm	Cytoplasm & occasional nucleus
Intensity (%)	DPPA3	NoS (5%); Weak (95%)	NoS (5%); Weak (95%)	NoS (5%); Weak (95%)
	EDR1	Moderate (5%); Strong (95%)	Moderate (5%); Strong (95%)	Moderate (90%); Strong (10%)
	NANOG	NoS (5%); Weak (95%)	Weak (95%); Moderate (5%)	Weak (95%) in cytoplasm (and nucleus); Moderate (5%) in cytoplasm
H-score	DPPA3	95	95	95
	EDR1	295	295	210
	NANOG	95	105	105

Note: NoS = No staining

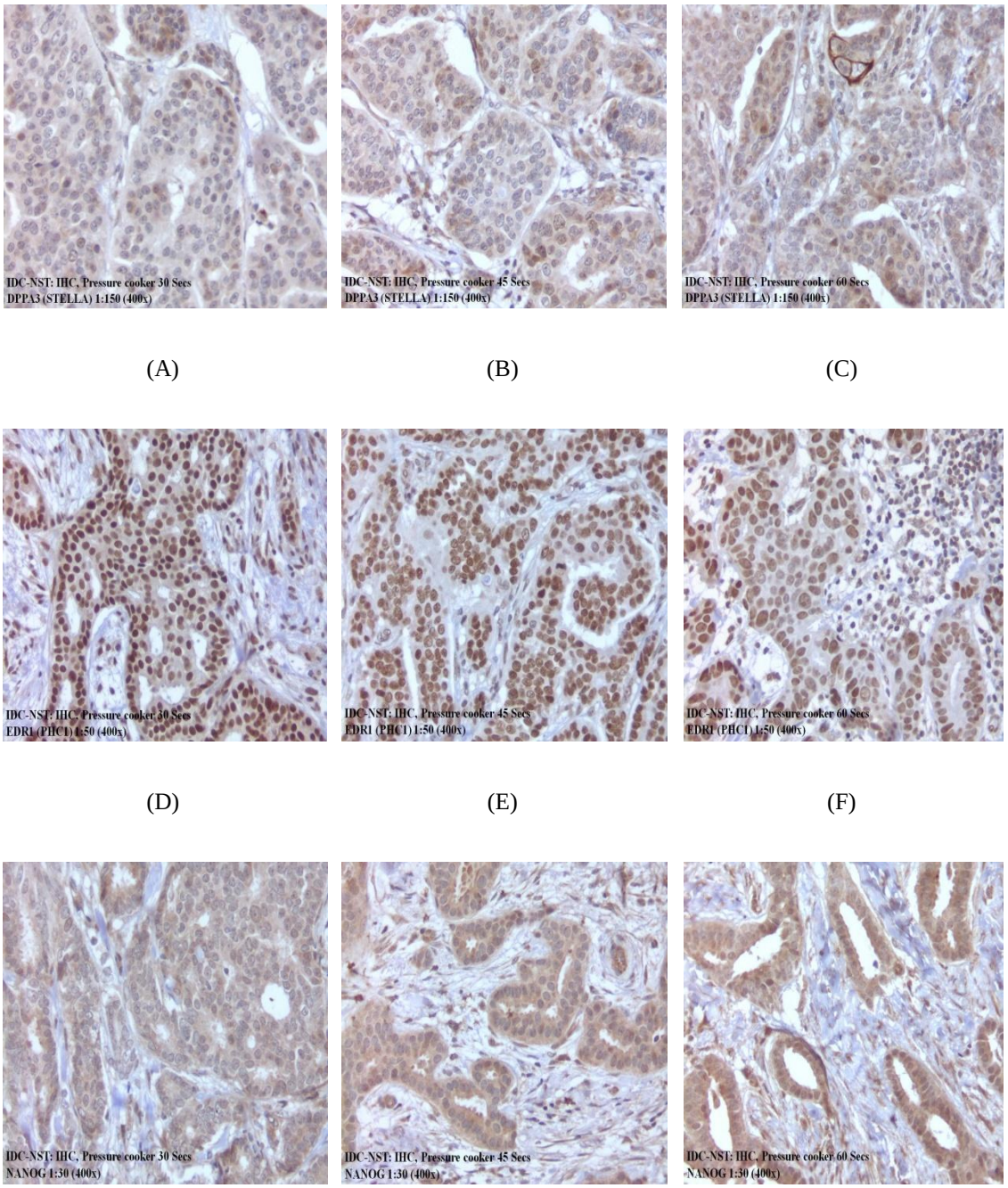


Figure 3-9 IHC of the IC tissue based on 30, 45, and 60 seconds pressure cooker AR method. The immunoreactivity of DPPA3 (Upper), EDR1 (Middle), and NANOG (Lower) proteins were performed on the IDC-NST tissue using pressure cooker AR method for 30 (A, D, and G), 45 (B, E, and H), and 60 (C, F, and I) seconds.

3.3.2.2 (1.3). ***Primary antibody concentration***

The dilutions of primary antibodies against DPPA3, EDR1, and NANOG proteins were optimised using the same tissue sample of IDC-NST as described in Section 3.3.2.2 (1.1). IHC was performed using the AR time as summarised in Section 3.4.2.2 (1.2) with 3 different dilutions of each primary antibody: (I) 1:150, 1:100, and 1:50 for DPPA3; (II) 1:50, 1:30, and 1:15 for EDR1; and (III) 1:30, 1:20, and 1:10 for NANOG (Table 3-10 and Figure 3-10).

All 3 dilutions for DPPA3 yielded weak cytoplasmic staining and the same H-score of 95, but a dilution of 1:150 gave more clearly defined cytoplasmic staining. The immunostaining of EDR1 showed mainly moderate staining in the nuclei of the cancer cells for dilutions of 1:50 (H-score of 205). There was stronger nuclear staining and weak to moderate cytoplasmic staining with the higher concentration of EDR1 antibody (H-scores of 299 for the dilutions of 1:30 and 1:15). The majority of NANOG-immunostained cancer cells revealed weak cytoplasmic staining for the dilutions of 1:30 (H-score of 95), 1:20 (H-score of 95), and 1:10 (H-score of 110). Additionally, the higher concentration of NANOG antibody yield more nuclear staining.

Therefore, regardless of the H-score values, the optimal immunoreactivity of DPPA3, EDR1, and NANOG proteins in the BC tissue was obtained by using dilutions of 1:150, 1:50, and 1:30, respectively, which gave more specific cytoplasmic staining for DPPA3 and NANOG and nuclear staining for EDR1.

Table 3-10 A comparison of primary antibody concentration. The optimal concentrations of primary antibodies were determined for immunohistochemical study of DPPA3 (STELLA), EDR1 (PHC1), and NANOG proteins in the same IDC-NST tissue as described in Section 3.3.2.2 (1.1). The Heating time for pressure cooker AR was the same as summarised in Section 3.4.2.2 (1.2).

Antibodies	Dilution	Sample	Heating time (Seconds)	Immunostaining location	Immunostaining intensity (%)	H-score
DPPA3	1:150	FFPE tissue of IDC-NST	30	Cytoplasm	NoS (5%); Weak (95%)	95
	1:100		30	Cytoplasm & occasional nucleus	NoS (5%); Weak (95%) in cytoplasm (and nucleus)	95
	1:50		30	Cytoplasm & occasional nucleus	NoS (5%); Weak (95%) in cytoplasm (and nucleus)	95
EDR1	1:50		30	Nucleus	Moderate (95%); Strong (5%)	205
	1:30		30	Nucleus & occasional cytoplasm	Moderate (1%) and strong (99%) in nucleus (Weak to moderate in cytoplasm)	299
	1:15		30	Nucleus & occasional cytoplasm	Moderate (1%) and strong (99%) in nucleus (Weak to moderate in cytoplasm)	299
NANOG	1:30		45	Cytoplasm	NoS (5%); Weak (95%)	95
	1:20		45	Cytoplasm & occasional nucleus	NoS (5%); Weak (95%) in cytoplasm (and nucleus)	95
	1:10		45	Cytoplasm & occasional nucleus	Weak (90%) in cytoplasm (and nucleus); Moderate (10%) in cytoplasm	110

Note: NoS = No staining

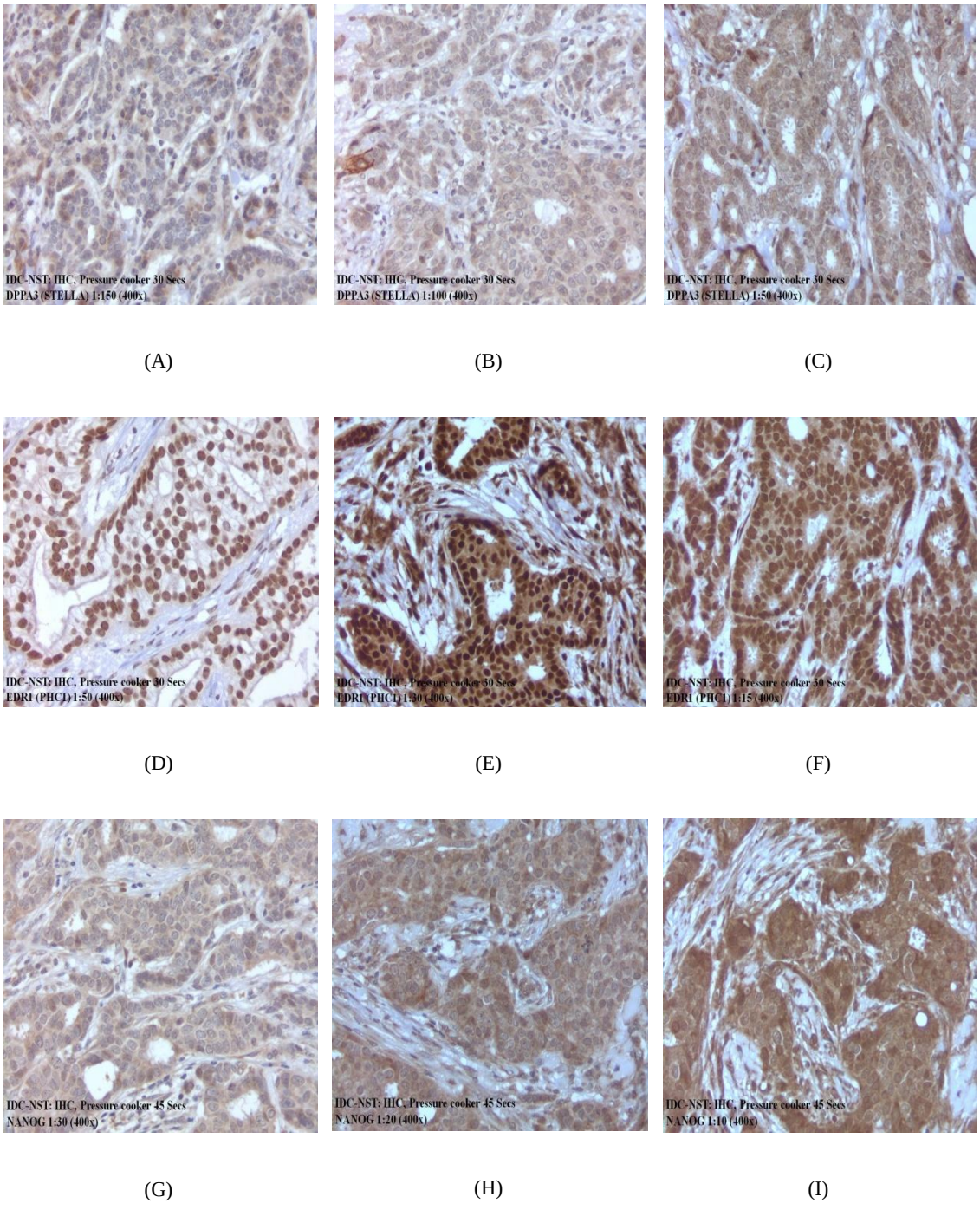


Figure 3-10 IHC of the IC tissue based on pressure cooker AR method with different primary antibody concentrations. The immunoreactivity of DPPA3 (Upper), EDR1 (Middle), and NANOG (Lower) proteins were performed on the IDC-NST tissue using pressure cooker AR method with 3 different concentrations of primary antibody. (A) 1:150 dilution of DPPA3; (B) 1:100 dilution of DPPA3; (C) 1:50 dilution of DPPA3; (D) 1:50 dilution of EDR1; (E) 1:30 dilution of EDR1; (F) 1:15 dilution of EDR1; (G) 1:30 dilution of NANOG; (H) 1:20 dilution of NANOG; and (I) 1:10 dilution of NANOG

3.3.2.2 (1.4). ***Summary of optimised immunostaining techniques***

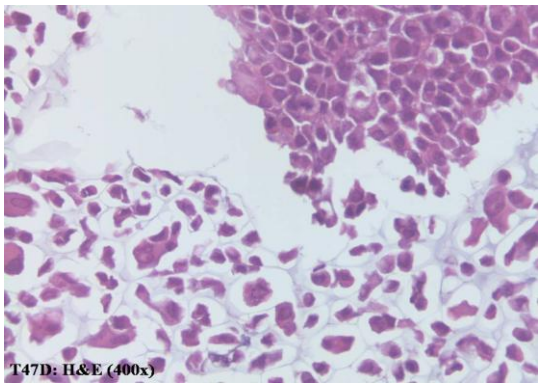
The immunohistochemical studies of DPPA3, EDR1, and NANOG proteins in invasive breast carcinoma tissue were performed using pressure cooker AR at the following times and primary antibody dilutions: (I) the retrieval time of 30 seconds with 1:150 dilution of DPPA3 and 1:50 dilution of EDR1 and (II) the retrieval time of 45 seconds with 1:30 dilution of NANOG.

3.3.2.2 (2). General immunohistochemical appearances of DPPA3 (STELLA), EDR1 (PHC1), and NANOG in invasive breast carcinoma tissue

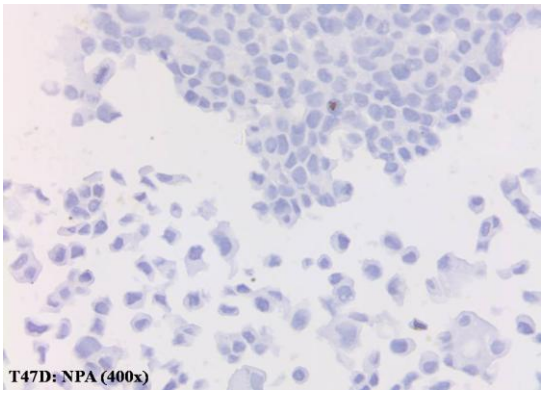
Based on 52 frozen invasive breast carcinoma tissue samples from Section 3.3.1 for analysis of mRNA expression, 44 (84.62%) cases had FFPE tissue samples available. The immunoreactivity of DPPA3, EDR1, and NANOG proteins in invasive breast carcinoma tissue was assessed in 4 histopathological components: normal mammary epithelial cells (NME); carcinoma in situ (CIS); invasive carcinoma (ICa); and vascular invasion (VI) when present. One tissue section of IDC-NST contained only NME and surrounding fibrofatty tissue, and there was only sufficient for DPPA3 staining. Therefore, the numbers of FFPE tissue of the invasive breast carcinoma sample were 44, 43, and 43 cases of DPPA3, EDR1, and NANOG immunostainings, respectively

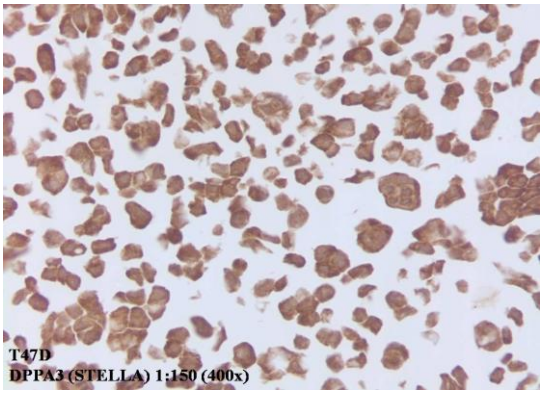
According to the manufacturers' information (Table 2-4 in Chapter 2 Materials and Methods), DPPA3, EDR1, and NANOG proteins could be expressed in either cytoplasm or nucleus. The immunohistochemical study was performed as Section 3.3.2.2 (1.4). T47D breast cancer cell line was used as a positive control and showed predominantly cytoplasmic staining of all 3 proteins (Figure 3-11). The NME, CIS, ICa,

and VI components had various staining appearances, being either negative or having weak, moderate, or strong staining (Figure 3-12 to 3-14). DPPA3 and NANOG were always detected in the CIS and VI components, but strong staining for NANOG protein was not identified in these components. For all 4 components, DPPA3 and NANOG proteins were expressed mainly in the cytoplasm, whereas EDR1 immunoreactivity was predominantly in the nucleus.

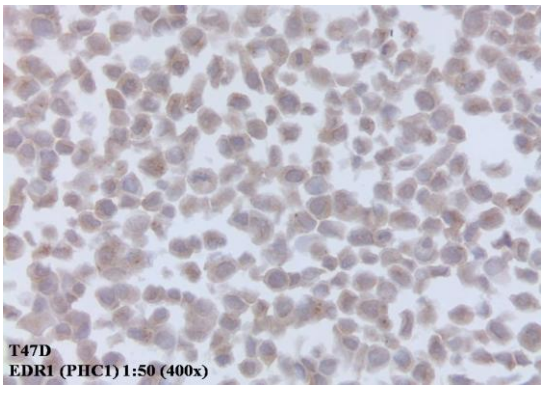


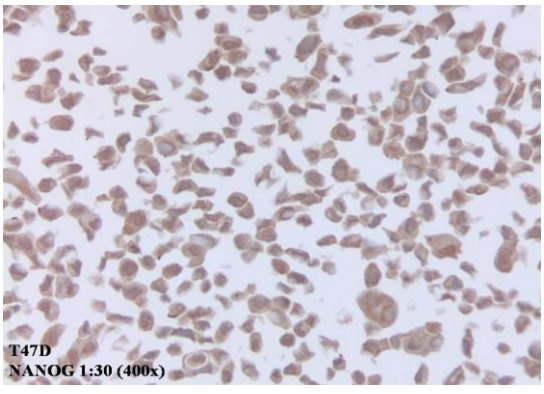
(A)





(B) (C)





(D) (E)

Figure 3-11 The immunoexpression of DPPA3 (STELLA), EDR1 (PHC1), and NANOG proteins in T47D breast cancer cell line. T47D breast cancer cell line was used as a positive control for DPPA3, EDR1, and NANOG stainings. The immunoreactivity of all 3 proteins was predominantly in the cytoplasm. (A) H&E staining; (B) No primary antibody (NPA); (C) 1:150 dilution of DPPA3; (D) 1:50 dilution of EDR1; and (E) 1:30 dilution of NANOG

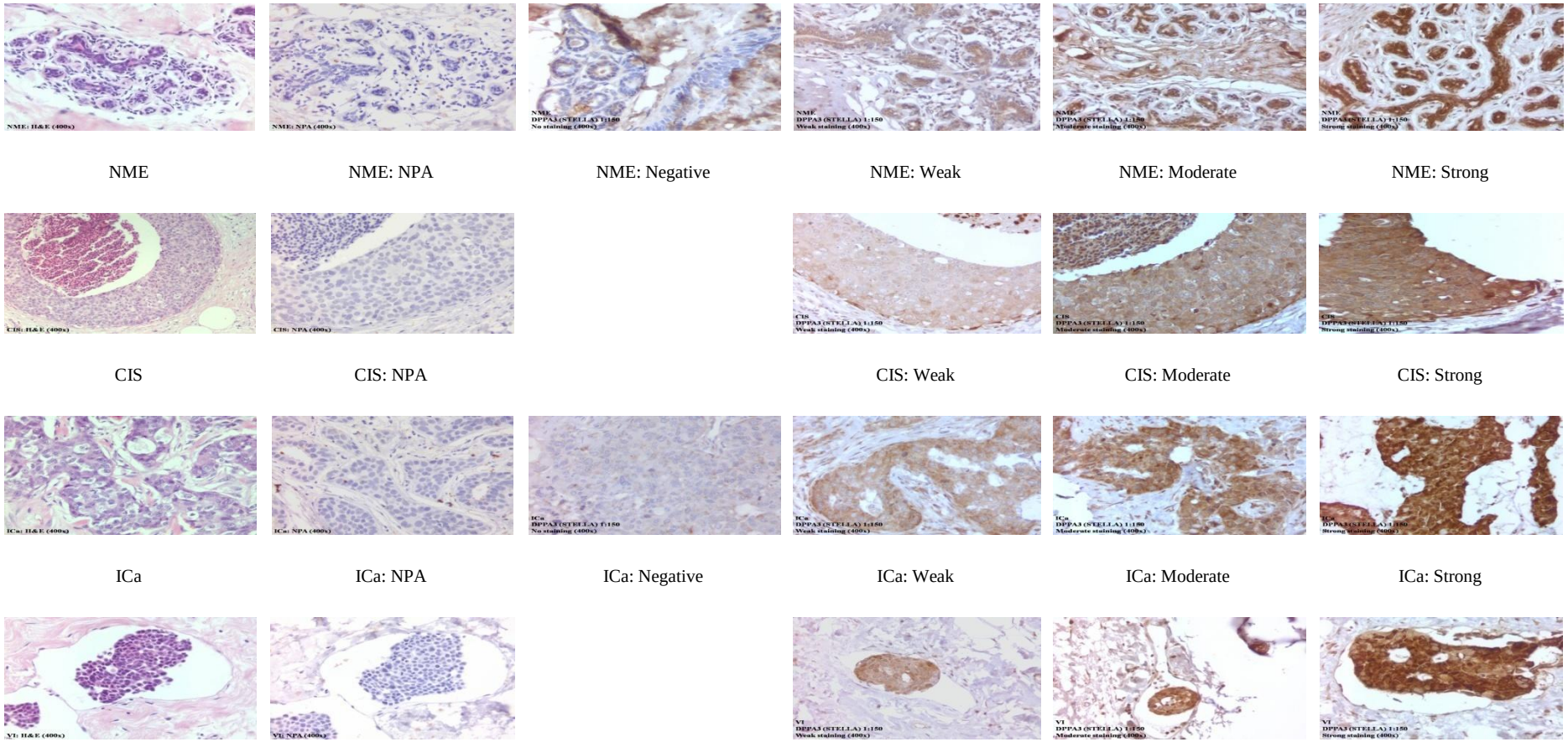
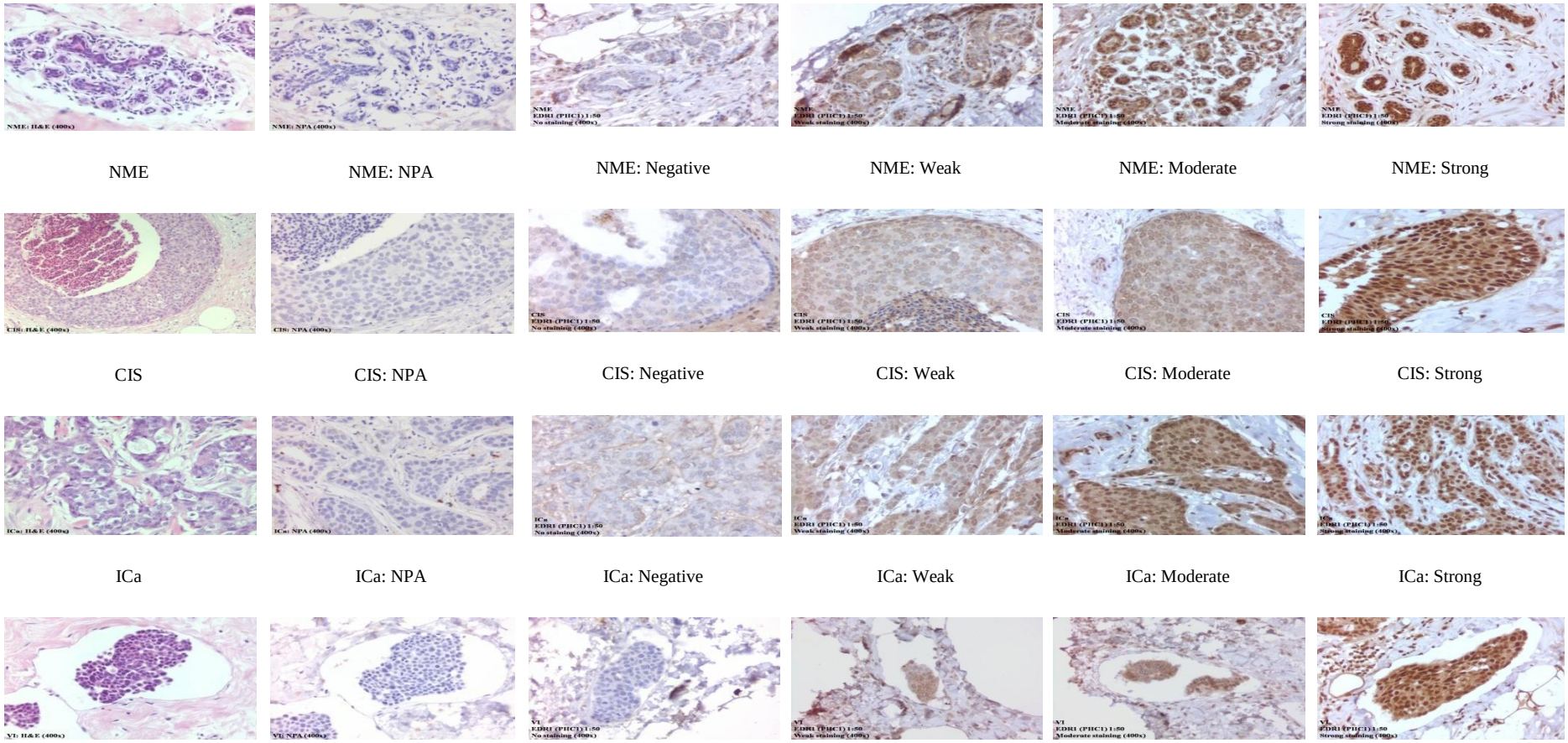


Figure 3-12 The immunostaining appearances of DPPA3 (STELLA) protein in invasive breast carcinoma tissue. The NME, CIS, ICa, and VI components of 44 breast carcinoma tissue samples showed various appearances of DPPA3 immunoexpression, including negative, weak, moderate, and strong stainings. Absence of protein was not found in the CIS and VI components. NPA = No primary antibody



NME | NME: NPA | NME: Negative | NME: Weak | NME: Moderate | NME: Strong

CIS | CIS: NPA | CIS: Negative | CIS: Weak | CIS: Moderate | CIS: Strong

ICa | ICa: NPA | ICa: Negative | ICa: Weak | ICa: Moderate | ICa: Strong

VI | VI: NPA | VI: Negative | VI: Weak | VI: Moderate | VI: Strong

Figure 3-13 The immunostaining appearances of EDR1 (PHC1) protein in invasive breast carcinoma tissue. The NME, CIS, ICa, and VI components of 43 breast carcinoma tissue samples showed various appearances of EDR1 immunoexpression, including negative, weak, moderate, and strong staining. NPA = No primary antibody

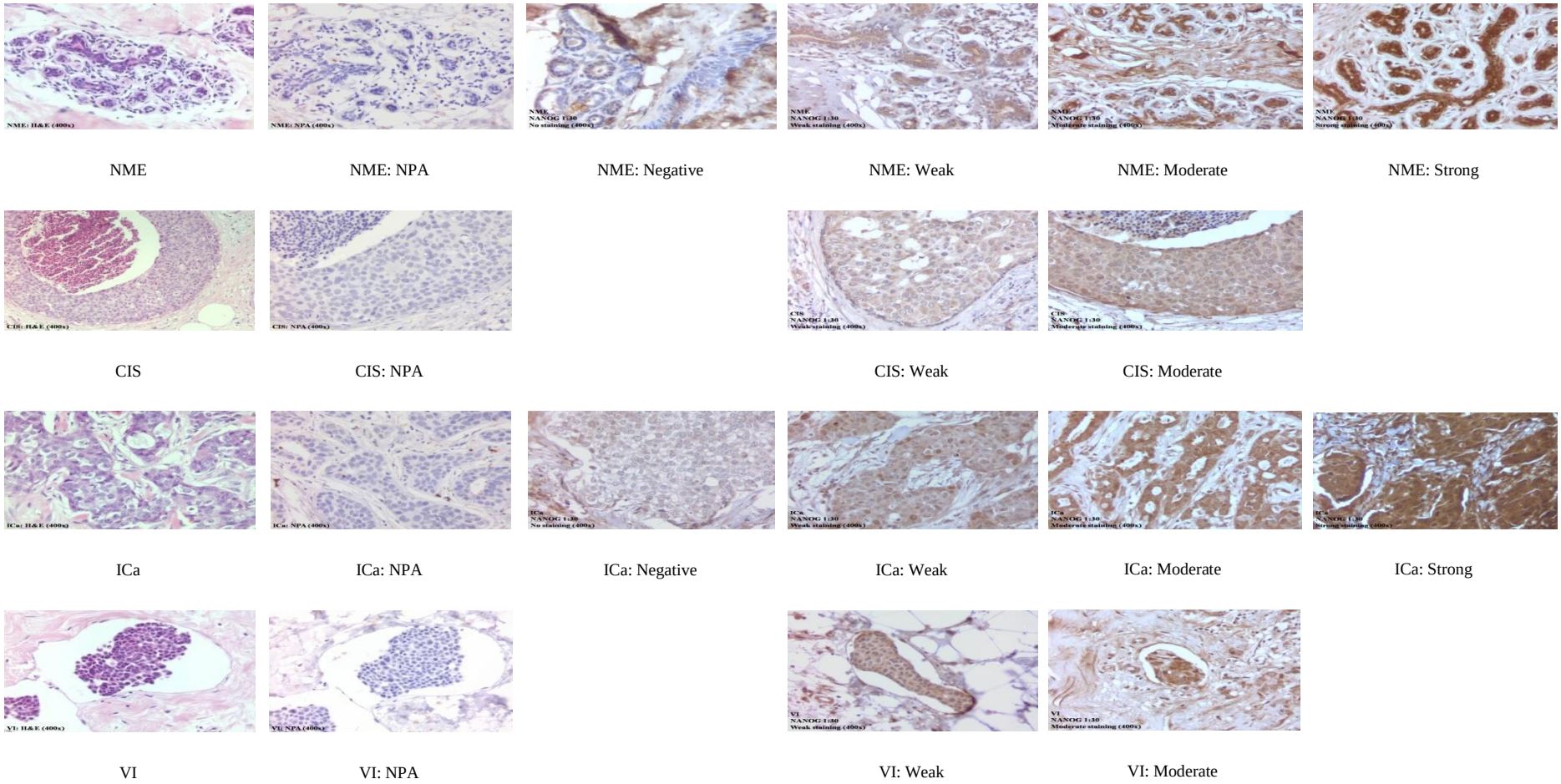


Figure 3-14 The immunostaining appearances of NANOG protein in invasive breast carcinoma tissue. The NME, CIS, ICa, and VI components of 43 breast carcinoma tissue samples showed various appearances of NANOG immunoexpression, including negative, weak, moderate, and strong stainings. NANOG protein was always detected in the CIS and VI components, but there was no strong staining in these components. NPA = No primary antibody

3.3.2.2 (3). Intra- and inter-observer variations of immunostaining interpretation

The immunohistochemical staining results for DPPA3 (STELLA), EDR1 (PHC1), and NANOG proteins were determined by myself (a Thai board-certified pathologist) on 2 separate occasions. There was an interval of 4 weeks between assessments which were undertaken without knowledge of the first results. The ratio paired *t*-test at the 95% CI was applied for analysis of intra-observer variation (Table 3-11). This statistical analysis showed that the first evaluation of DPPA3 H-score ($X_{H\text{-}score}$ = 158.40; $SD_{H\text{-}score}$ = 68.50) was higher than the second evaluation ($X_{H\text{-}score}$ = 154.10; $SD_{H\text{-}score}$ = 67.43) ($p < 0.0001$). EDR1 and NANOG stainings had no significant variation of their H-score assessments on the different occasion ($p > 0.050$). Hence, the mean H-score of each immunoexpression was the value used for further analysis. In addition, 6 selected slides (2 cases of DCIS & IDC-NST, 2 cases of IDC-NST, a case of MC, and a case of ILC) were reviewed jointly with breast pathologist, Professor Rosemary Walker. There was an agreement about H-scores of the NME, CIS, ICa, and VI components in all 6 invasive breast carcinoma tissue samples.

Table 3-11 Intra-observer variation of immunohistochemical interpretation. The immunostaining results of expression of DPPA3 (STELLA), EDR1 (PHC1), and NANOG proteins in invasive breast carcinoma were blindly assessed by myself (a Thai board-certified pathologist) on 2 separate 4 weeks occasions. The ratio paired *t*-test at the 95% CI was applied to the determination of intra-observer variation of this interpretation.

IHC	**Evaluation of H-score**				**Ratio paired *t*-test**		
	1st		**2nd**				
	X	**SD**	**X**	**SD**	***t***	***df***	***p*-value**
DPPA3 (N = 44)	158.40	68.50	154.10	67.43	4.765	76	< 0.0001***
EDR1 (N = 43)	138.20	98.96	137.60	99.18	0.574	69	0.5681^{ns}
NANOG (N = 43)	111.80	53.01	112.70	52.74	0.964	71	0.3386^{ns}

Note: ns = No statistical significance

3.3.2.2 (4). Determination of cut-off point for positivity of DPPA3 (STELLA), EDR1 (PHC1), and NANOG immunoexpression in invasive breast carcinoma

The ICa component was identified in 43 out of 44 (97.73%) sections for DPPA3, 41 out of 43 (95.35%) sections for EDR1, and all 43 sections for NANOG. The mean H-scores of DPPA3, EDR1, and NANOG proteins in the ICa component were analysed for the appropriate cut-off values to determine positive immunoexpression by using the one-way analysis of variance (ANOVA) and the unpaired *t*-test at the 95% CI. Seven different cut-off H-scores were selected as follows: > 50, > 75, > 100, > 125, > 150, > 175, and > 200. For DPPA3, the statistical analysis excluded the H-scores of > 150, > 175, and > 200 since there were only a small number of cases (4 – 7 cases). The number of cancers with each cut-off H-score of EDR1 was adequate for statistical analysis. Cancers with NANOG H-scores of > 50, > 75, and > 100 were included in the analysis, but the other cases with H-scores of > 125, > 150, >175, and > 200 had too small number (none to one case).

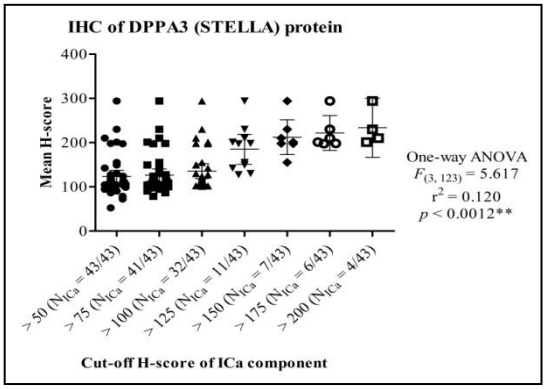
The one-way ANOVA results showed that the 7 selected cut-off values did not all have the equal mean H-score for DPPA3 ($p = 0.0012$) (Figure 3-15 A) and EDR1 ($p < 0.0001$) (Figure 3-16 A), but these values might all have the equal mean H-score for NANOG ($p = 0.2768$) (Figure 3-17 A). Hence, the unpaired *t*-test was used further to (I) confirm the one-way ANOVA results and (II) determine the most likely positive cut-off point from those 7 values for DPPA3, EDR1, and NANOG.

For DPPA3, the unpaired *t*-test showed no significant difference in the mean H-score when cut-off points of > 50, >75, and >100 were applied ($p > 0.050$) (Figure 3-15 B to D), but the cut-off point of >100 ($X_{H\text{-score}} = 135.50$) gave a lower mean H-score

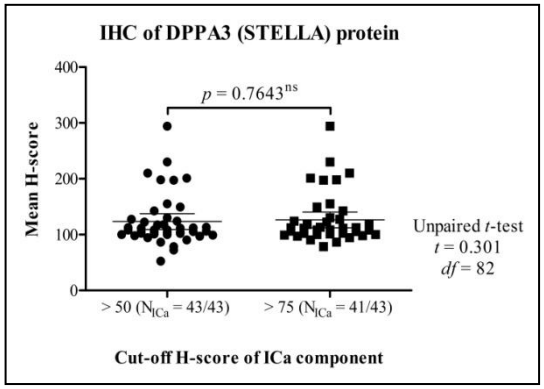
than the cut-off point of >125 ($X_{H\text{-}score}$ = 185.00) (p = 0.0049) (Figure 3-15 E). There were more cases with a DPPA3 H-score of > 100 compared to cancers with the H-score of ≤ 100 (p = 0.0020) (Figure 3-15 F).

Thirty-six out of 41 (87.80%) cancers had an EDR1 H-score of > 50 with a similar mean H-score to the 33 (80.49%) cases with H-score of > 75 (p = 0.6172) (Figure 3-16 B). Approximately 50% of samples with the H-score of > 100 showed a significantly higher EDR1 expression than both cases with H-score of > 50 (p = 0.0033) and > 75 (p = 0.0103) (Figure 3-16 C and D), but there was no difference in immunoexpression between the cut-off H-scores of >100 and > 125 (p = 0.4340) (Figure 3-16 E). Cancers with an EDR1 H-score of ≤ 50 and ≤ 75 had a statistically significant lower level of protein expression compared to cancers with H-score of > 50 (p = 0.0002) and > 75 (p < 0.0001), respectively (Figure 3-16 F and G). However, cancers were considered positive for EDR1 when the H-score was > 75 since this gave the most significant result with the unpaired *t*-test.

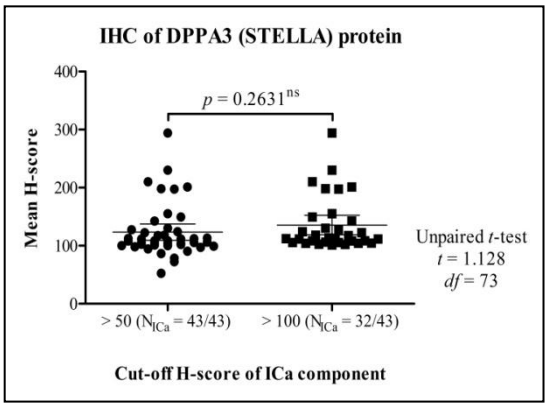
For NANOG, both one-way ANOVA and the unpaired *t*-test showed that there was no significant difference in immunoexpression of the cases with the cut-off H-scores of > 50, >75, and > 100 (p > 0.050) (Figure 3-17 B to D). The number of samples with the H-score of ≤ 50 and ≤75 were too small for statistical analysis (Figure 3-17 E and F). The twenty-two (51.16%) cancers with a NANOG H-score of ≤ 100 showed lower expression of NANOG compared to cases with the H-score of > 100 (p = 0.0038) (Figure 3-17 G).



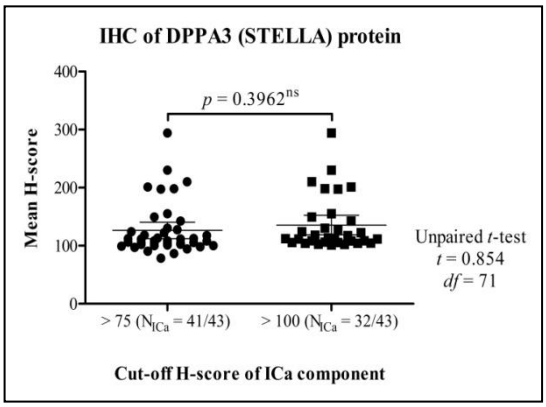
(A)



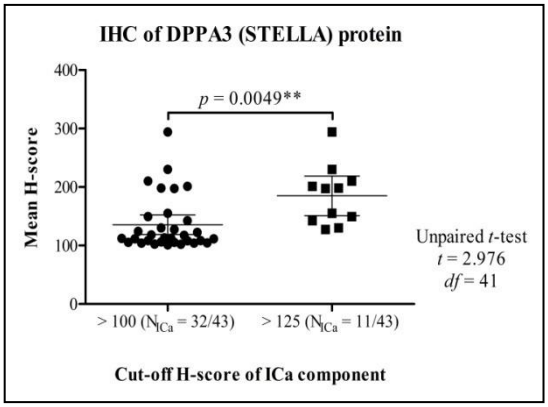
(B)



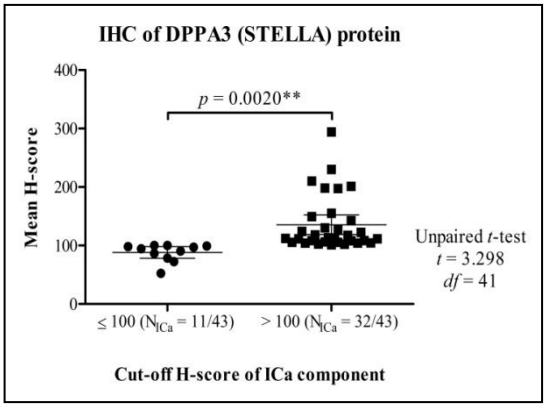
(C)



(D)

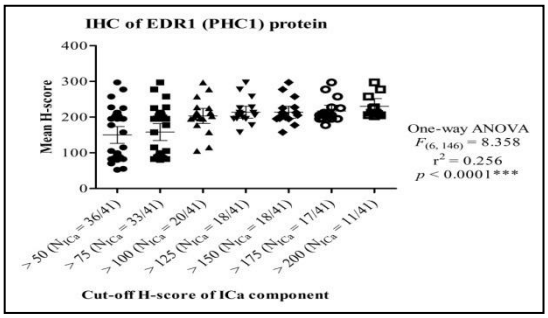


(E)

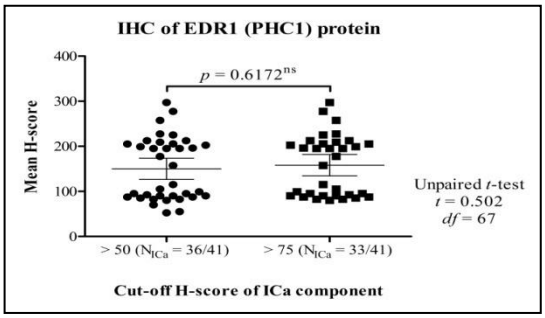


(F)

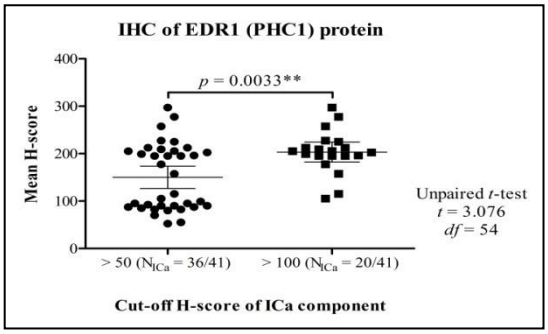
Figure 3-15 Determination of a cut-off H-score of positive DPPA3 (STELLA) immunoexpression. Scatter diagrams were plotted from the mean H-score of DPPA3 in the ICa component of 43 breast carcinoma tissue samples. The results show a horizontal line at the mean and error bars at the 95% CI. The one-way ANOVA (A) and unpaired *t*-test (B – F) were used for comparison with the cut-off H-scores. (A) All cut-off H-scores; (B) H-score of > 50 vs H-score of > 75; (C) H-score of > 50 vs H-score of > 100; (D) H-score of > 75 vs H-score of > 100; (E) H-score of > 100 vs H-score of > 125; and (F) A cut-off H-score at 100



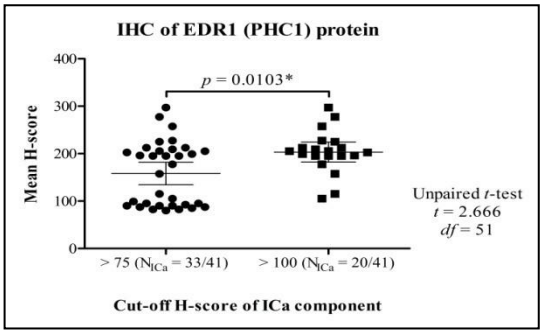
(A)



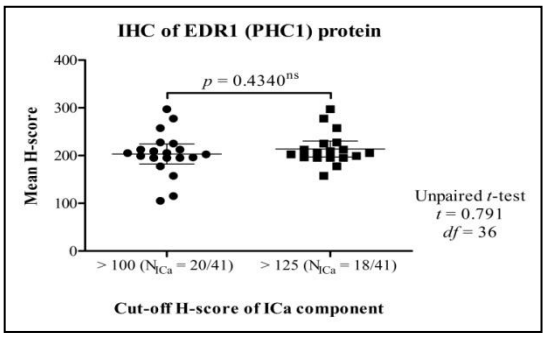
(B)



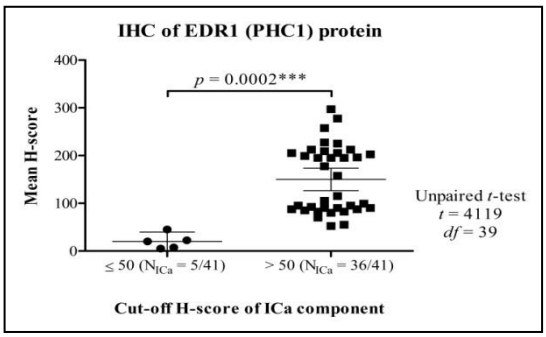
(C)



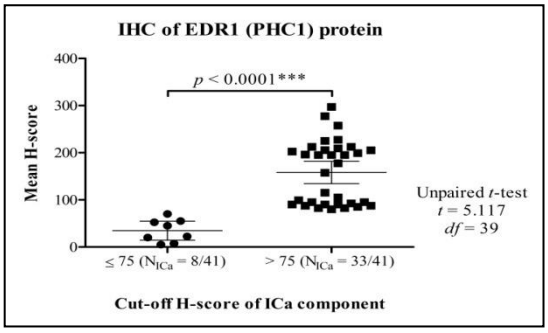
(D)



(E)

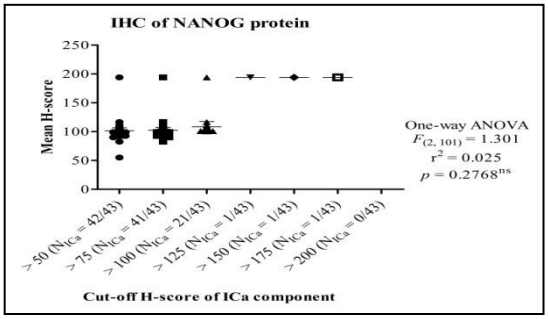


(F)

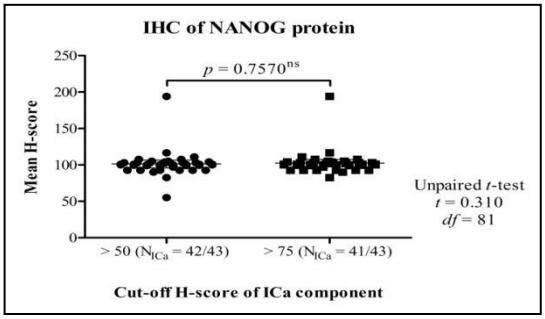


(G)

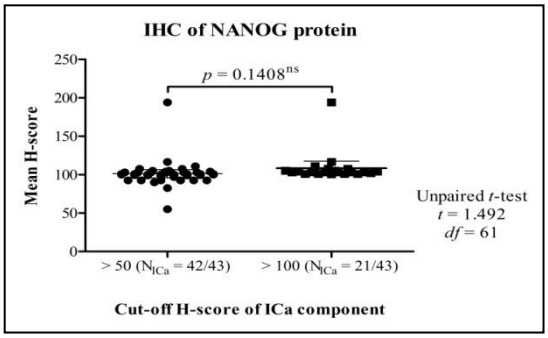
Figure 3-16 Determination of a cut-off H-score of positive EDR1 (PHC1) immunoexpression. Scatter diagrams were plotted from the mean H-score of EDR1 in the ICa component of 41 breast carcinoma tissue samples. The results show a horizontal line at the mean and error bars at the 95% CI. The one-way ANOVA (A) and the unpaired *t*-test (B – G) were used for comparison with the cut-off H-scores. (A) All cut-off H-scores; (B) H-score of > 50 vs H-score of > 75; (C) H-score of > 50 vs H-score of > 100; (D) H-score of > 75 vs H-score of > 100; (E) H-score of > 100 vs H-score of > 125; (F) A cut-off H-score of 50; and (G) A cut-off H-score of 75



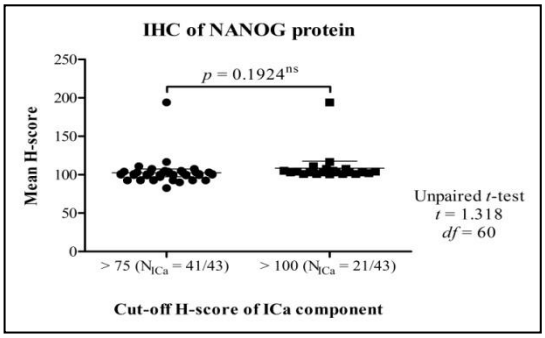
(A)



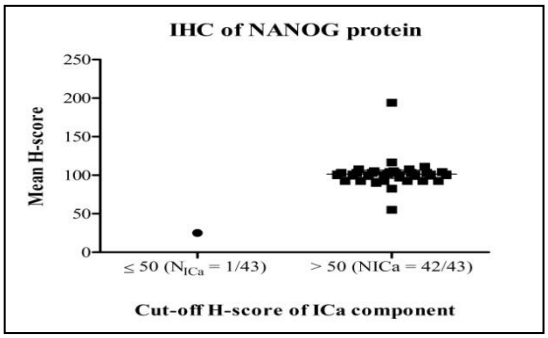
(B)



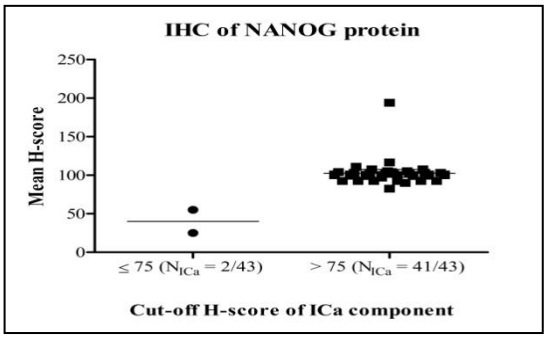
(C)



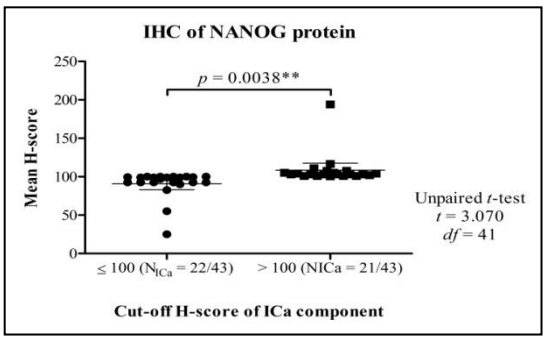
(D)



(E)



(F)

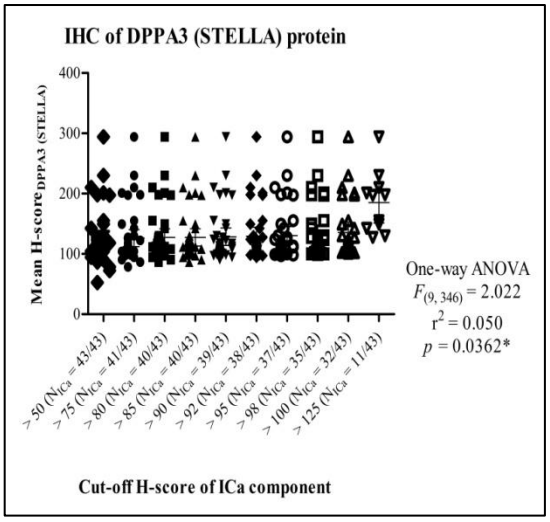


(G)

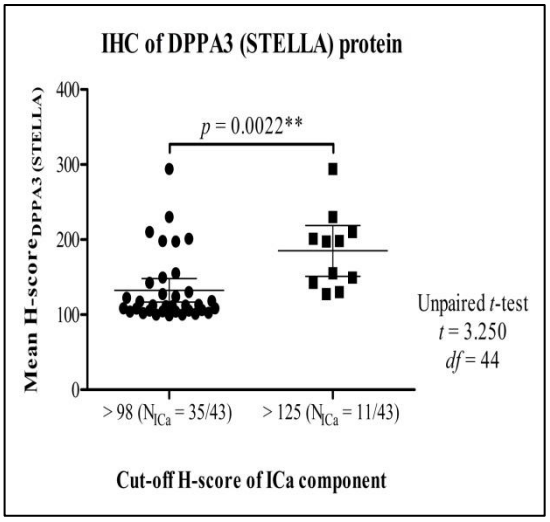
Figure 3-17 Determination of a cut-off H-score of positive NANOG immunoexpression. Scatter diagrams were plotted from the mean H-score of NANOG in the ICa component of 43 breast carcinoma tissue samples. The results show a horizontal line at the mean and error bars at the 95% CI. The one-way ANOVA (A) and unpaired *t*-test (B – D and G) were used for comparison with the cut-off H-scores. (A) All cut-off H-scores; (B) H-score of > 50 vs H-score of > 75; (C) H-score of > 50 vs H-score of > 100; (D) H-score of > 75 vs H-score of > 100; (E) A cut-off H-score of 50; (F) A cut-off H-score of 75; and (G) A cut-off H-score of 100

The positive cut-off H-score of > 100 for DPPA3 and NANOG were considered as a quite high value. Although the statistical analysis at the 95% CI showed significant difference in the expression of NANOG between cancers with a NANOG H-score of ≤ 100 and >100, but the lower limit on the H-score of > 100 was 99.44. Hence, the new cut-off H-score values were determined for DPPA3 (> 50, > 75, > 80, > 85, > 90, > 92, > 95, > 98, > 100, and > 125) and NANOG (> 50, > 75, > 80, > 85, > 90, > 92, > 95, > 98, and > 100). The one-way ANOVA at the 95% CI showed that the 10 new selected cut-off values for DPPA3 did not all have the equal mean H-score ($p = 0.0362$) (Figure 3-18 A), whereas the new 9 selected cut-off points for NANOG might all have the equal mean H-score ($p = 0.8539$) (Figure 3-19 A).

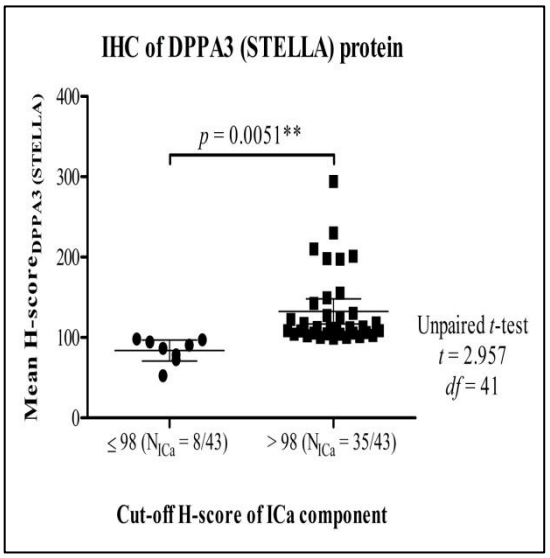
For DPPA3 immunoexpression, the unpaired *t*-test at the 95% CI showed no significant difference in the mean H-score when the new cut-off points of > 50 to >100 were applied ($p > 0.050$), but the cut-off point of > 98 ($X_{H\text{-}score} = 132.50$) gave a lower mean H-score than the cut-off point of >125 ($X_{H\text{-}score} = 185.00$) ($p = 0.0022$) (Figure 3-18 B). There were significant difference in DPPA3 expression between the cases with a H-score of ≤ 98 and > 98 ($p = 0.0051$) (Figure 3-18 C). However, the cancers with the previous positive cut-off H-score value of > 100 yielded the greater significant result ($p = 0.0020$) (Figure 3-18 D). Thus, the new statistical analysis confirmed that DPPA3 immunoexpression in the BC tissue was considered positive when the H-score value was > 100.



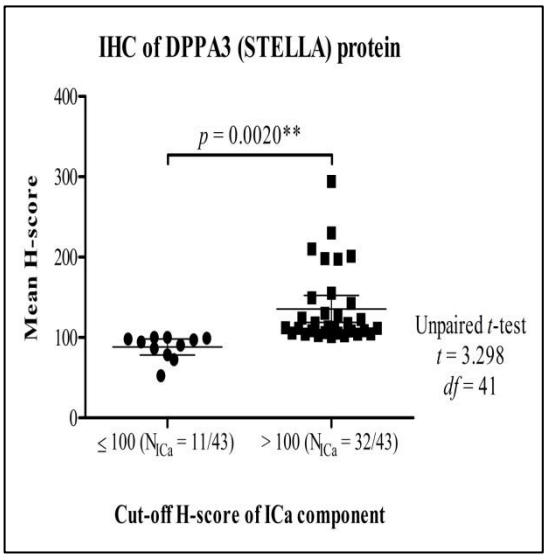
(A)



(B)



(C)

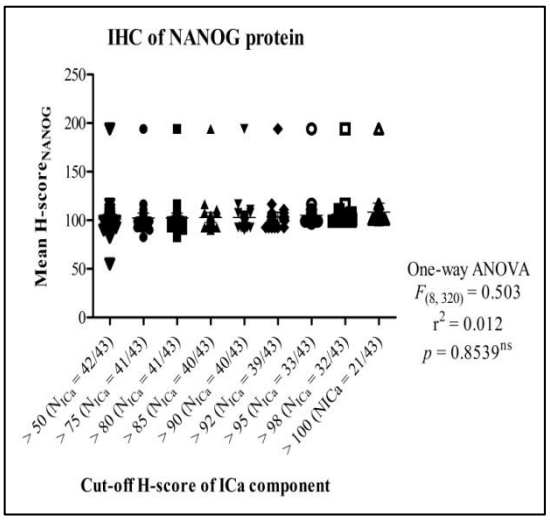


(D)

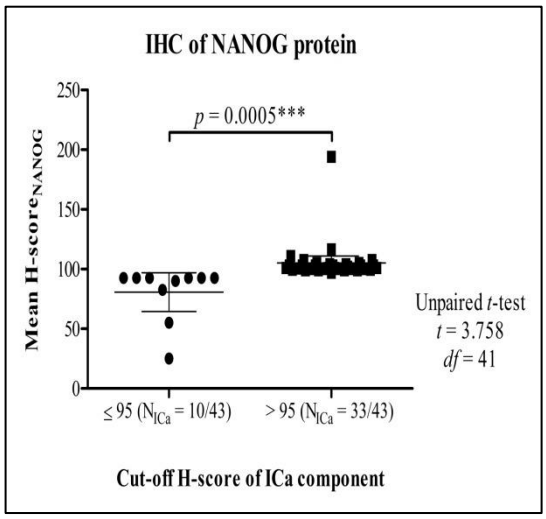
Figure 3-18 The reliability of a positive cut-off H-score of > 100 for DPPA3 (STELLA) immunoexpression. Scatter diagrams were plotted from the mean H-score of DPPA3 in the ICa component of 43 breast carcinoma (BC) tissue samples. The results show a horizontal line at the mean and error bars at the 95% CI. The one-way ANOVA (A) and the unpaired *t*-test (B – D) were used to determination of the new cut-off H-score values for positive expression of DPPA3 in the BC tissue. (A) All new cut-off H-scores; (B) H-score of > 98 vs H-score of > 125; (C) A cut-off H-score at 98; and (D) A cut-off H-score at 100

For NANOG, the unpaired *t*-test at 95% CI revealed no significant difference in immunoexpression of the cases with the new cut-off H-scores of > 50 to > 100 ($p > 0.050$), but cancers with a NANOG H-score of > 95 and > 98 showed significantly higher expression of NANOG compared to cases with the H-score of ≤ 95 ($p = 0.0005$) and cases with the H-score of ≤ 98 ($p = 0.0007$), respectively (Figure 3-19 B and C). However, the H-score of > 98 was considered as the cut-off point for positive NANOG immunoexpression in breast carcinoma tissue, since there were (I) equivocal values for the upper limit on the H-score of ≤ 95 (97.08) and the lower limit on the H-score of > 100 (99.44) and (II) appropriate values for the upper limit on the H-score of ≤ 98 (97.14) and the lower limit on the H-score of > 98 (99.45).

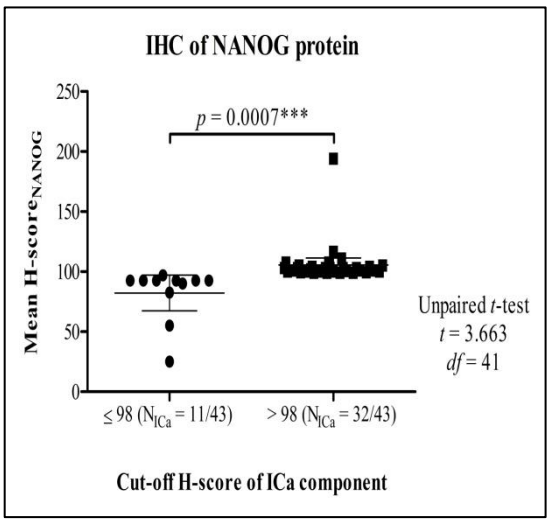
In summary, the immunoreactivity of DPPA3, EDR1, and NANOG was determined to be positive in the invasive breast carcinoma tissue at the 95% CI if their H-scores were > 100, > 75, and > 98, respectively.



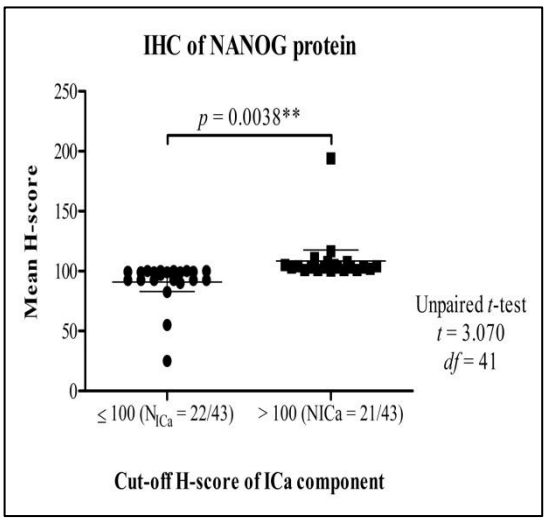
(A)



(B)



(C)



(D)

Figure 3-19 The reliability of a positive cut-off H-score of > 100 for NANOG immunoexpression. Scatter diagrams were plotted from the mean H-score of NANOG in the ICa component of 43 breast carcinoma (BC) tissue samples. The results show a horizontal line at the mean and error bars at the 95% CI. The one-way ANOVA (A) and the unpaired *t*-test (B – D) were used to determination of the new cut-off H-score values for positive expression of NANOG in the BC tissue. (A) All new cut-off H-scores; (B) A cut-off H-score at 95; (C) A cut-off H-score at 98; and (D) A cut-off H-score at 100

3.3.2.2 (5). The number and percentage of DPPA3 (STELLA)-, EDR1 (PHC1)-, and NANOG-positive breast carcinomas

Based on the cut-off H-scores at 95% CI as summarised in Section 3.3.2.2 (4), the breast carcinoma cases showed positive immunoexpression as follows: 32 out of 43 (74.42%) cases for DPPA3; 33 out of 41 (80.49%) cases for EDR1; and 32 out of 43 (74.42%) cases for NANOG. Those considered negative were 11 (25.58%) cases for DPPA3; 8 (19.51%) cases for EDR1; and 11 (25.58%) cases for NANOG (Table 3-12). The positive cases were used for: (I) the identification of the intracellular location of protein expression in Section 3.3.2.2 (6); (II) the correlation of each positive immunoexpression in Section 3.3.2.2 (7); and (III) comparison with the H-score of their corresponding NME component in Section 3.3.2.2 (8). Both positive and negative cases were analysed for the correlation with their clinicopathological information in Section 3.3.2.2 (9); and mRNA expression in Section 3.3.2.2 (10).

Table 3-12 The H-scores of negative and positive immunoexpression of DPPA3 (STELLA), EDR1 (PHC1), and NANOG proteins in breast carcinoma (BC) tissue samples. The DPPA3-, EDR1-, and NANOG-immunostained BC tissue containing invasive cancer cells were classified as negative and positive cases based on the cut-off H-scores in Section 3.3.2.2 (4). The unpaired *t*-test at the 95% CI was applied for comparison with each individual immunoexpression.

Sample (N = 43)		**Mean H-score**					
		DPPA3 (N = 43)		**EDR1 (N = 41)**		**NANOG (N = 43)**	
Tissue type	**Case number**	**Negative (N = 11/43) (25.58%)**	**Positive (N = 32/43) (74.42%)**	**Negative (N = 8/41) (19.51%)**	**Positive (N = 33/41) (80.49%)**	**Negative (N = 11/43) (51.16%)**	**Positive (N = 32/43) (48.84%)**
DCIS & IDC-NST (N = 5)	1		230.00		277.50		116.50
	2		149.50	55.00		97.00	
	3		105.50	45.00			100.50
	4		155.00		92.50		103.00
	5		101.00	52.50			100.00
IDC-NST (N = 35)	6		104.00		157.50	92.50	
	7		108.50		257.50		101.00
	8	78.50			195.00	92.50	
	9		111.50		105.00	90.00	
	10		106.00		195.00		101.00
	11		111.50	20.00		92.50	
	12		113.00		227.50		105.00
	13		124.00		99.00		104.00
	14		105.50		90.00	82.50	
	15	86.50		22.50			105.00
	16	94.50			202.50		99.00
	17		122.50		225.00		100.00
	18		108.50		212.50		103.00
	19	100.00			209.00	92.50	
	20		117.50		297.00		101.00
	21		108.00		90.00		103.00
	22	72.50			195.00	55.00	

Table 3-12 (Continued) The H-scores of negative and positive immunoexpression of DPPA3 (STELLA), EDR1 (PHC1), and NANOG proteins in breast carcinoma (BC) tissue samples. The DPPA3-, EDR1-, and NANOG-immunostained BC tissue containing invasive cancer cells were classified as negative and positive cases based on the cut-off H-scores in Section 3.3.2.2 (4). The unpaired *t*-test at the 95% CI was applied for comparison with each individual immunoexpression.

Sample (N = 43)		**Mean H-score**					
		DPPA3 (N = 43)		**EDR1 (N = 41)**		**NANOG (N = 43)**	
Tissue type	**Case number**	**Negative (N = 11/43) (25.58%)**	**Positive (N = 32/43) (74.42%)**	**Negative (N = 8/41) (19.51%)**	**Positive (N = 33/41) (80.49%)**	**Negative (N = 11/43) (25.58%)**	**Positive (N = 32/43) (74.42%)**
IDC-NST (N = 35)	23		112.00		90.00		111.00
	24		210.00	NTS			100.00
	25		104.50		87.50		101.00
	26		142.50		85.00		107.50
	27		118.00		82.50		103.00
	28		127.50		205.50		100.00
	29	100.00			115.00		99.00
	30		104.00		177.50		104.00
	31		102.00		95.00		99.00
	32	99.00			199.00		100.00
	33	97.00		7.50			99.00
	34		130.00		82.50		102.00
	35	98.00		5.00		92.50	
	36		102.50		205.00		107.50
	37		197.50		80.00		103.00
	38	90.50			87.50	92.50	
	39		201.00	NTS			99.50
	40	52.50			95.00	25.00	
MC (N = 1)	41		294.00	70.00			194.00
ILC (N = 2)	42		112.50		212.50		99.50
	43		198.00		196.00		103.00
$X_{H\text{-}score}$		**88.09**	**135.50**	**34.69**	**158.30**	**82.23**	**105.40**
$SD_{H\text{-}score}$		**14.94**	**46.58**	**24.11**	**66.71**	**22.20**	**16.60**
Unpaired *t*-test (95% CI)	***t, df***	**3.298, 41**		**5.117, 39**		**3.663, 41**	
	***p*-value**	**0.0020****		**< 0.0001*****		**0.0007*****	

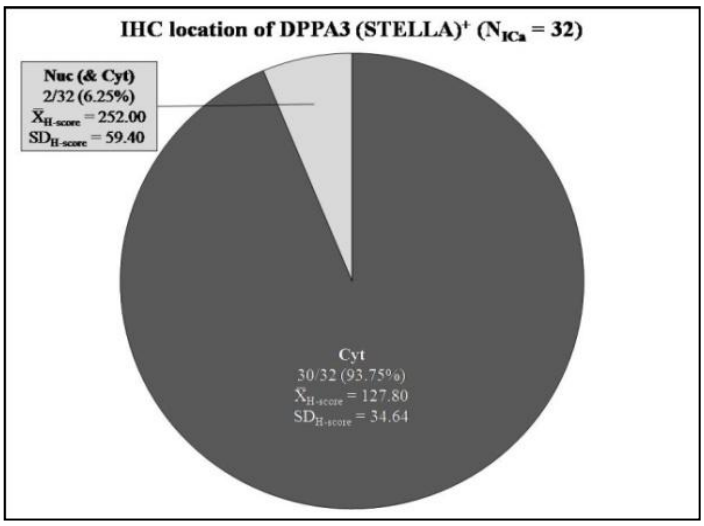
Note: NTS = No tumour seen

3.3.2.2 (6). The location of DPPA3 (STELLA), EDR1 (PHC1), and NANOG immunoexpression in invasive breast carcinomas

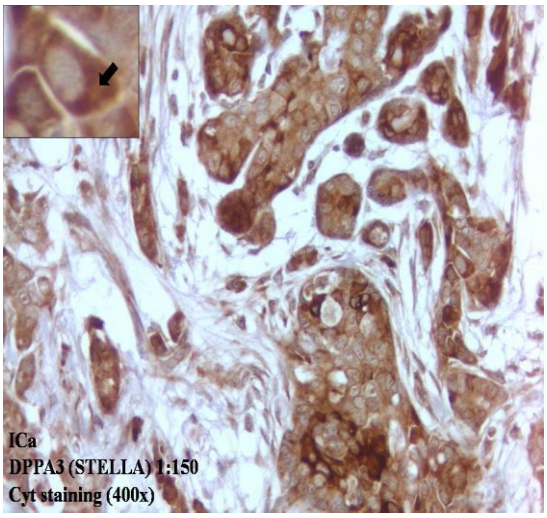
Of the 32 DPPA3-positive breast cancers, DPPA3 protein was mainly detected in the cytoplasm [N = 30/32 (93.75%)] with the other 2 cases showing nuclear and occasionally cytoplasmic staining (Figure 3-20). For 33 EDR1-positive breast cancers (Figure 3-21), there were (I) nuclear with occasionally cytoplasmic staining in 25 (75.76%) cases; (II) cytoplasmic with occasionally nuclear staining in 4 cases; (III) both nuclear and cytoplasmic staining in 3 cases; and (IV) only cytoplasmic staining in one case. Of the 32 NANOG-positive breast cancers (Figure 3-22), the most common location was the cytoplasm, with the other cases having (I) both nuclear and cytoplasmic staining in 9 cases; (II) cytoplasmic and occasionally nuclear staining in 7 cases; and (III) nuclear and occasionally cytoplasmic staining in one case.

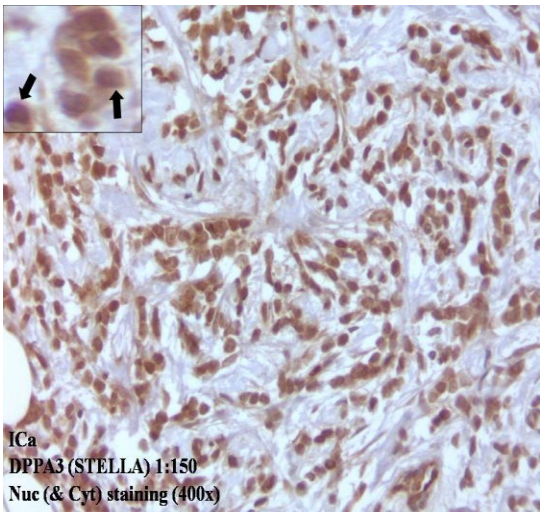
For 11 cancers with H-scores for DPPA3 and NANOG considered negative, all cases displayed only cytoplasmic staining. From 8 EDR1-negative cancers based on the cut-off H-score, there was (I) nuclear with occasionally cytoplasmic staining in one case; (II) cytoplasmic with occasionally nuclear staining in 6 cases; and (III) only cytoplasmic staining in one case.

Hence, the invasive breast carcinoma showed the positive immunostaining appearances of DPPA3, EDR1, and NANOG proteins as described in Section 3.3.2.2 (1.1): DPPA3 and NANOG proteins were mainly detected in the cytoplasm, whereas EDR1 protein was predominantly in the nucleus and occasionally in the cytoplasm.



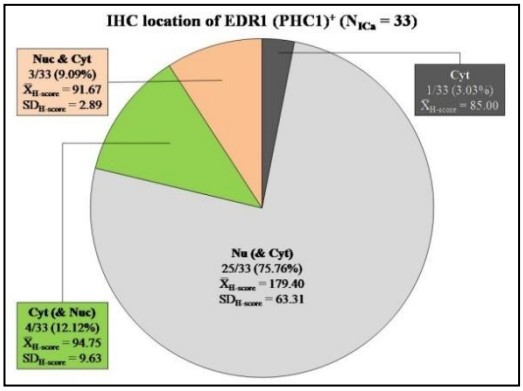
(A)





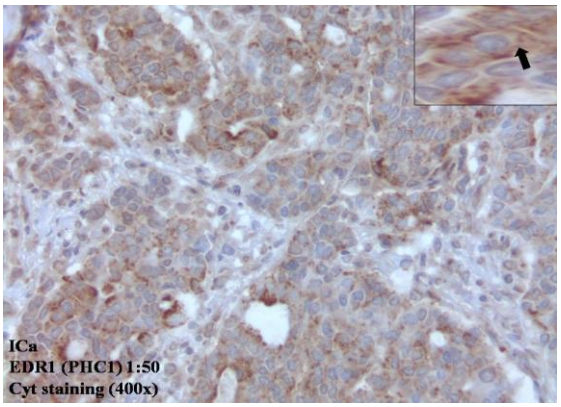
(B) (C)

Figure 3-20 The location of positive DPPA3 (STELLA) immunoexpression in the invasive breast carcinoma. (A) Pie diagram demonstrates the location of positive DPPA3 immunostaining in the invasive breast cancer, including number of cases, percentage of cases, and mean and SD of H-score. (B and C) Expression of DPPA3 was present in (I) only cytoplasm (Cyt) (B) and (II) nucleus and occasional cytoplasm [Nuc (& Cyt)] (C). The magnified images of the staining locations (Arrow) are illustrated in the insets.

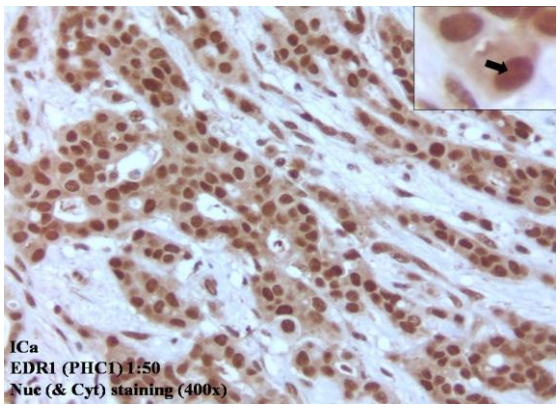


(A)

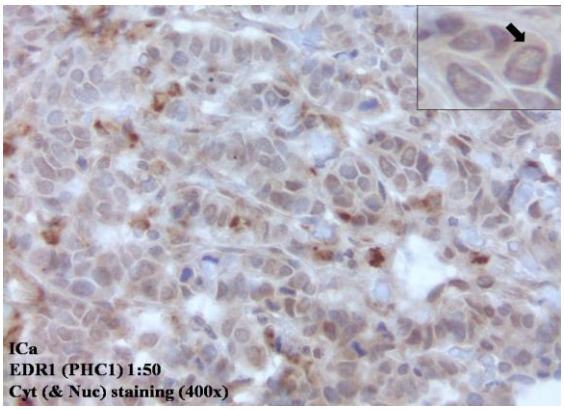
Figure 3-21 The location of positive EDR1 (PHC1) immunoexpression in the invasive breast carcinoma. (A) Pie diagram demonstrates the location of positive EDR1 immunostaining in the invasive breast cancer, including number of cases, percentage of cases, and mean and SD of H-score. (B and E) Expression of EDR1 was present in (I) only cytoplasm (Cyt) (B); (II) nucleus and occasionally cytoplasm [Nuc (& Cyt)] (C); (III) cytoplasm and occasionally nucleus [Cyt (&Nuc)] (D); and (IV) both nucleus and cytoplasm (Nuc & Cyt) (E). The magnified images of the staining locations (Arrow) are illustrated in the insets.



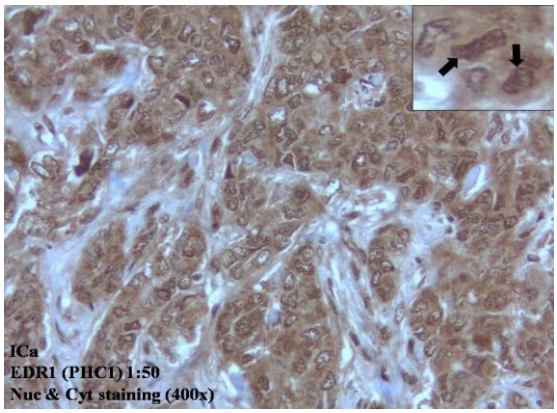
(B)



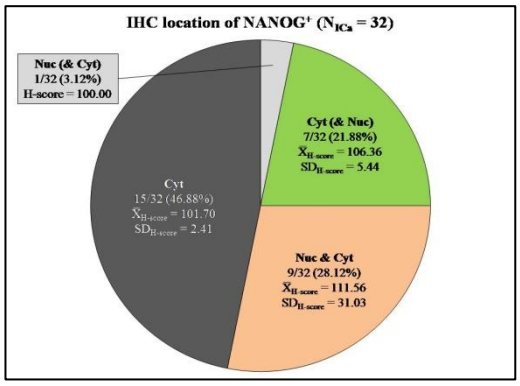
(C)



(D)

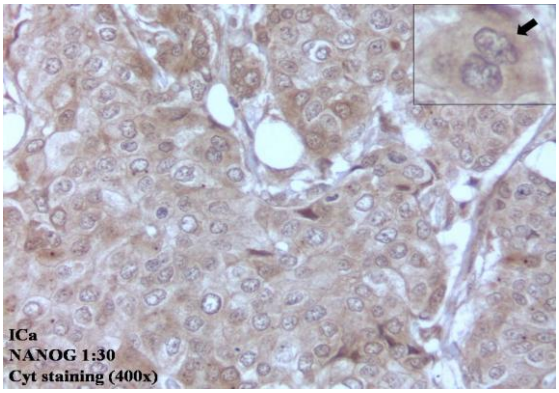


(E)

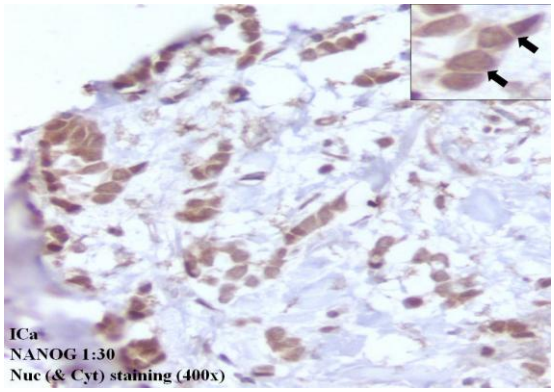


(A)

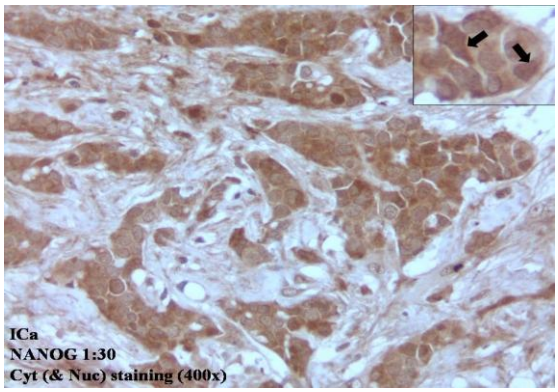
Figure 3-22 The location of positive NANOG immunoexpression in the invasive breast carcinoma. (A) Pie diagram demonstrates the location of positive NANOG immunostaining in the invasive breast cancer, including number of cases, percentage of cases, and mean and SD of H-score. (B and E) Expression of NANOG was present in (I) only cytoplasm (Cyt) (B); (II) nucleus and occasionally cytoplasm [Nuc (& Cyt)] (C); (III) cytoplasm and occasionally nucleus [Cyt (&Nuc)] (D); and (IV) both nucleus and cytoplasm (Nuc & Cyt) (E). The magnified images of the staining locations (Arrow) are illustrated in the insets.



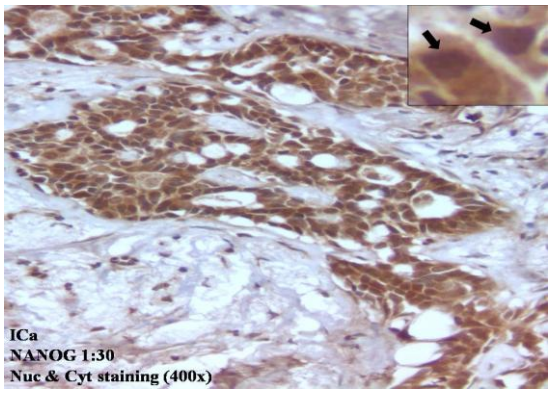
(B)



(C)



(D)

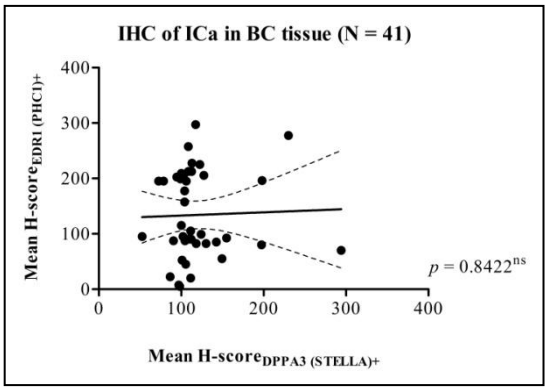


(E)

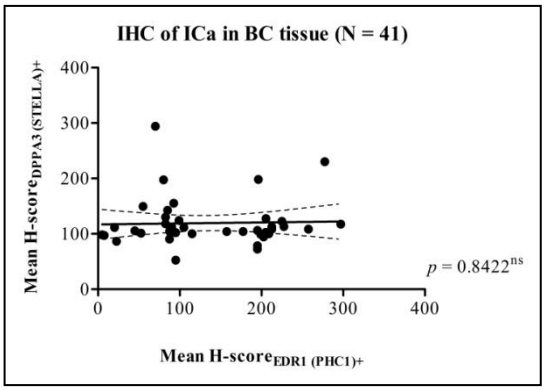
3.3.2.2 (7). The correlation between immunoexpression of DPPA3 (STELLA), EDR1 (PHC1), and NANOG proteins in the invasive breast carcinoma

Since there was no tumour present in 2 cases stained for EDR1, 41 out of 43 (95.35%) samples were analysed to determine the relationship between expression of DPPA3, EDR1, and NANOG proteins (Figure 3-23). The results showed a significant linear correlation between DPPA3 and NANOG in the cancers ($p < 0.0001$). However, there was no correlation of H-scores between DPPA3 protein and EDR1 protein or EDR1 protein and NANOG protein ($p > 0.050$).

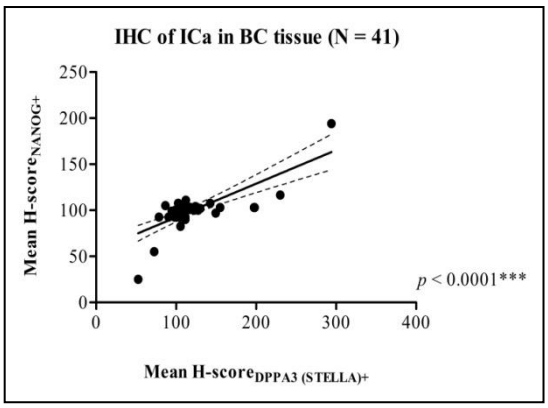
Since one case was negative for all DPPA3, EDR1, and NANOG proteins (Table 3-12), 40 out of 41 (97.56%) samples were evaluated for co-expression of these proteins (Figure 3-24). Thirty out of 40 (75.00%) cancers showed positive DPPA3 staining, 2 less than in Table 3-12 since these had been excluded. The number of positive cases for EDR1 and NANOG were the same as in Table 3-12 (33 and 32 cases, respectively). The DPPA3-positive cancers frequently expressed both EDR1 and NANOG proteins. These DPPA3-positive cases infrequently co-expressed with either EDR1 or NANOG protein. A small number of EDR1-positive cases had co-expression of NANOG protein. Independent positive expression of each protein was also present in the cancers but was infrequently detected: 2 DPPA3-positive cases, 5 EDR1-positive cases, and 2 NANOG-positive cases.



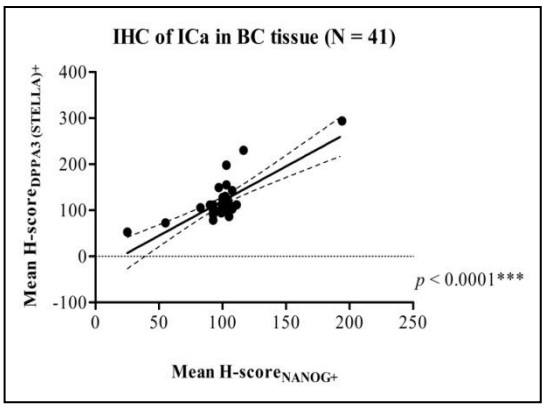
(A)



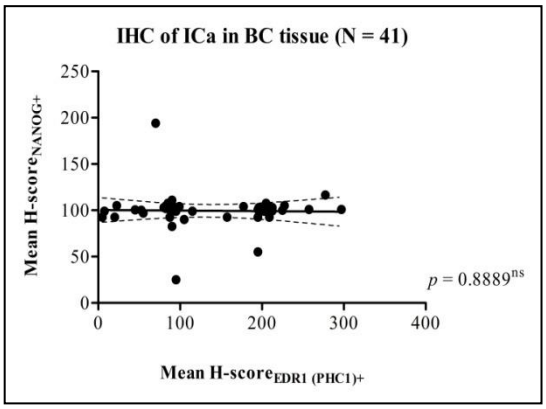
(B)



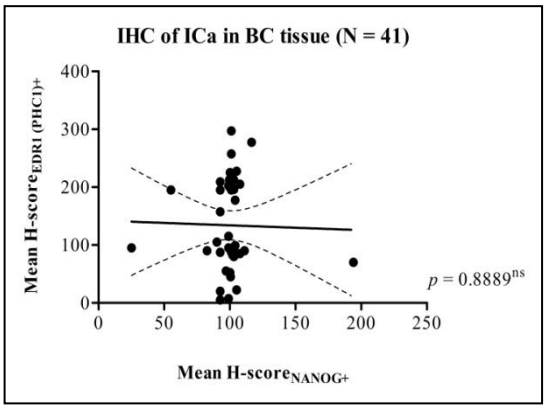
(C)



(D)



(E)



(F)

Figure 3-23 The correlation of immunoexpression between DPPA3 (STELLA), EDR1 (PHC1), and NANOG proteins in the invasive breast carcinomas. The Pearson correlation coefficient (r) at the 95% CI was used for the analysis of the correlation of H-scores between DPPA3, EDR1, and NANOG proteins in 41 breast carcinoma tissue samples. (A) and (B) DPPA3 protein versus EDR1 protein; (C) and (D) DPPA3 protein versus NANOG protein; (E) and (F) EDR1 protein versus NANOG protein

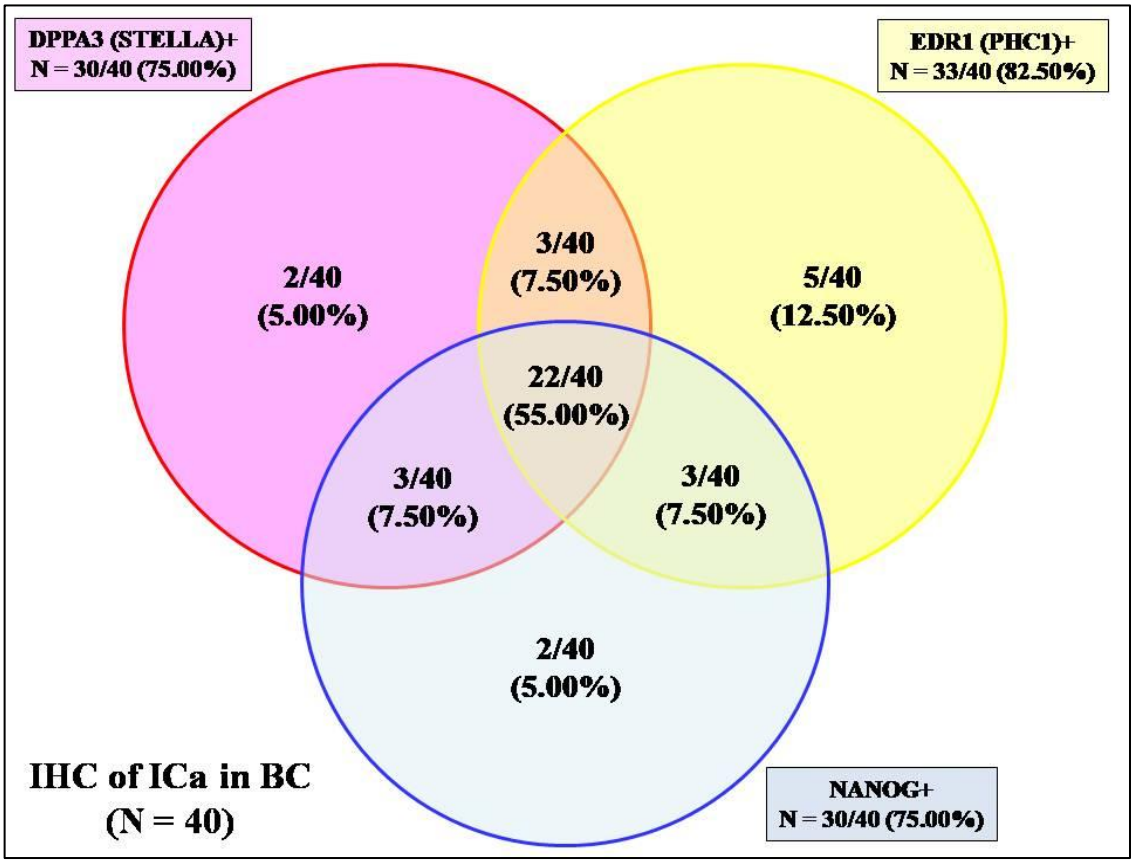
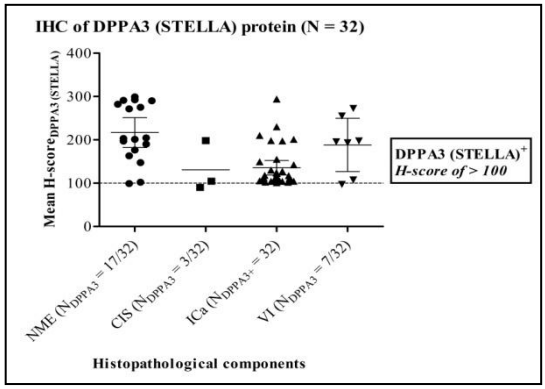


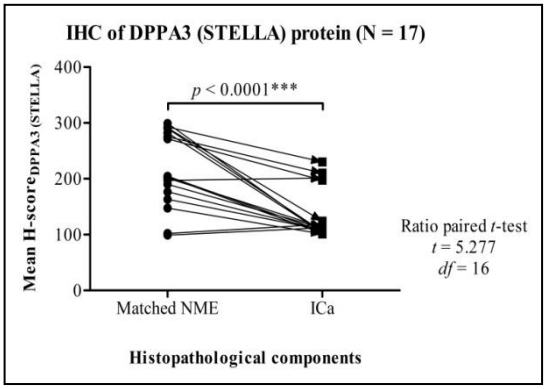
Figure 3-24 Co-expression of DPPA3 (STELLA), EDR1 (PHC1), and NANOG proteins in the invasive breast carcinoma. Venn diagram shows the number and percentage of positive expression and co-expression of DPPA3, EDR1, and NANOG proteins in 40 invasive breast cancers. Three cases were not included in this co-expression, since one case showed negative expression for all 3 proteins, and the other 2 cases had no tumour presented in the EDR1-immunostained tissue sections.

3.3.2.2 (8). Comparison of positive DPPA3 (STELLA), EDR1 (PHC1), and NANOG immunoexpression in the invasive carcinoma (ICa) component with their corresponding normal mammary epithelium (NME)

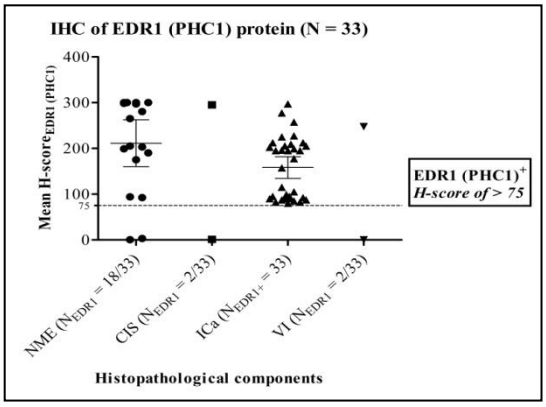
Based on the breast carcinoma (BC) tissue cases with positive DPPA3 (32 cases), EDR1 (33 cases), and NANOG (32 cases) immunoreactivity as described in Section 3.3.2.2 (5), the H-score of the ICa component was compared with that of their corresponding NME component by using the ratio paired *t*-test at the 95% CI (Figure 3-25). The carcinoma in situ (CIS) and vascular invasion (VI) were only present in 2 – 7 tissue sections and so a statistical comparison could not be undertaken. Seventeen out of 32 (53.12%) BC tissue samples had a significantly higher level of DPPA3 expression in the NME compared to the ICa ($p < 0.0001$). However, expression of EDR1 [N = 18/33 (54.54%)] and NANOG [N = 15/32 (46.88%)] showed no significant differences between NME and ICa ($p > 0.050$).



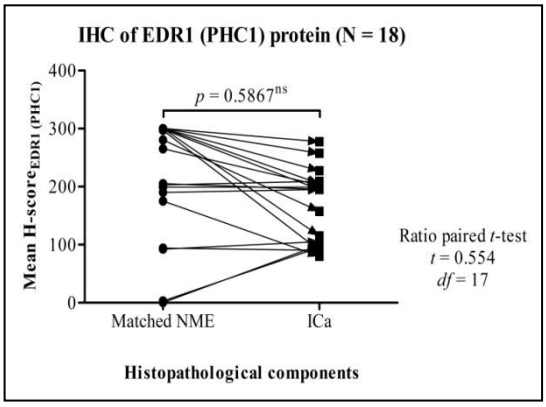
(A)



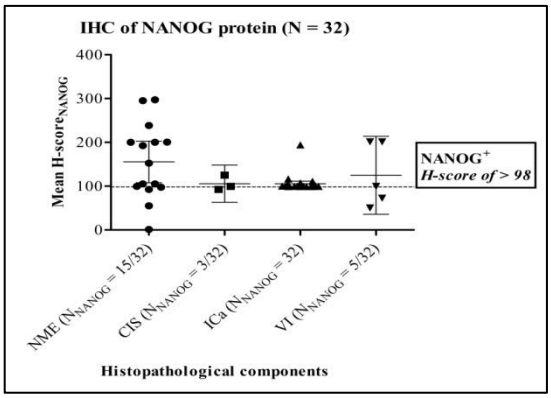
(B)



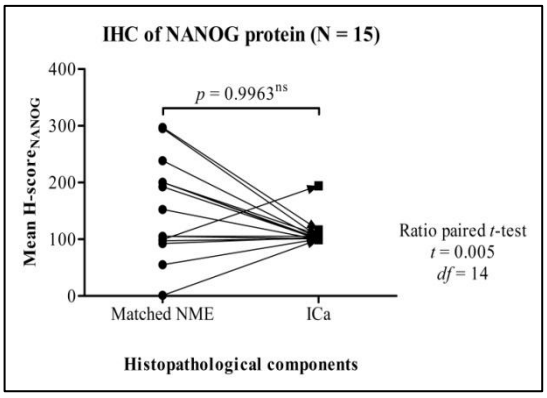
(C)



(D)



(E)



(F)

Figure 3-25 Comparison of DPPA3 (STELLA), EDR1 (PHC1), and NANOG immunoreactivity in the invasive carcinoma (ICa) component with their corresponding normal mammary epithelium (NME) component of the breast carcinoma (BC) tissue. Scatter diagrams (Left) were plotted from mean H-score of positive DPPA3 (A), EDR1 (C), and NANOG (E) immunostainings in the ICa component and their corresponding NME, carcinoma in situ (CIS), and vascular invasion (VI) components of the BC tissue samples. Results show a horizontal line at the mean and error bars at the 95% CI. The NME and ICa components was matched and compared their immunoexpression by using the ratio paired *t*-test (Right) as follows: (B) DPPA3; (D) EDR1; and (F) NANOG

3.3.2.2 (9). The correlation between DPPA3 (STELLA), EDR1 (PHC1), and NANOG immunoexpression with clinicopathological features of breast carcinomas

The immunoreactivity of DPPA3 (43 cases), EDR1 (41 cases), and NANOG (43 cases) in the breast carcinoma (BC) cases (Table 3-12) were analysed to look for any correlation with their clinicopathological features [the age (≤ 50 years old and > 50 years old); tumour size (≤ 2.0 cm and > 2 cm); tumour grade/differentiation; axillary lymph node metastasis; and the status of ER, PR, and HER2 (ERBB2)] by using the Fisher exact test at the 95% CI (Table 3-13 to 3-15). No information was identified in one sample for the status of ER and PR, so the relationship to expression of those hormonal receptors was evaluated in 42 cases for DPPA3 and NANOG and 40 cases for EDR1. The status of HER2 was not known in 29 cases for DPPA3 and NANOG and 27 cases for EDR1, thus only 14 cases of each staining were analysed for the relationship to HER2 expression.

The statistical analysis showed that there was no significant correlation between expression of DPPA3, EDR1, and NANOG and the following features: age group; tumour size; tumour grade/differentiation, axillary lymph node metastasis; and expression of ER and PR ($p > 0.050$). Five out of 14 HER2-positive cancers were associated with no expression of DPPA3 (3 cases) ($p = 0.0275$) or NANOG (4 cases) ($p = 0.0230$).

Table 3-13 The correlation between DPPA3 (STELLA) immunoexpression and clinicopathological features of the invasive breast carcinoma. The Fisher exact test at the 95% CI was applied for the analysis of the correlation between DPPA3 immunoreactivity, including negative and positive, and clinicopathological features of 43 breast cancers.

Clinicopathological features	DPPA3 immunoexpression (N = 43)			
	Negative (N = 11/43)	**Positive (N = 32/43)**	**Total N**	***p*-value**
Age (Years old)				
≤ 50	8	18	**26**	0.4801^{ns}
> 50	3	14	**17**	
Tumour size (cm)				
≤ 2.0	3	9	**12**	1.0000^{ns}
> 2.0	8	23	**31**	
Tumour grade/differentiation				
Low (I & II)/Well & Moderate	6	9	**15**	0.1504^{ns}
High (III)/Poor	5	23	**28**	
Axillary lymph node				
No metastasis	5	16	**21**	1.0000^{ns}
Metastasis	6	16	**22**	
ER				
Negative	1	8	**9**	0.4032^{ns}
Positive	10	23	**33**	
No information	0	1	**1**	
PR				
Negative	2	7	**9**	1.0000^{ns}
Positive	9	24	**33**	
No information	0	1	**1**	
HER2 (ERBB2)				
Negative	0	9	**9**	**0.0275***
Positive	3	2	**5**	
No information	8	21	**29**	

Note: ns = No statistical significance

Table 3-14 The correlation between EDR1 (PHC1) immunoexpression and clinicopathological features of the invasive breast carcinoma. The Fisher exact test at the 95% CI was applied for the analysis of the correlation between EDR1 immunoreactivity, including negative and positive, and clinicopathological features of 41 breast cancers.

Clinicopathological features	EDR1 immunoexpression (N = 41)			
	Negative (N = 8/41)	**Positive (N = 33/41)**	**Total N**	***p*-value**
Age (Years old)				
≤ 50	5	21	**26**	1.0000^{ns}
> 50	3	12	**15**	
Tumour size (cm)				
≤ 2.0	4	7	**11**	0.1777^{ns}
> 2.0	4	26	**30**	
Tumour grade/differentiation				
Low (I & II)/Well & Moderate	3	10	**13**	0.6925^{ns}
High (III)/Poor	5	23	**28**	
Axillary lymph node				
No metastasis	5	15	**20**	0.4537^{ns}
Metastasis	3	18	**21**	
ER				
Negative	2	7	**9**	1.0000^{ns}
Positive	6	25	**31**	
No information	0	1	**1**	
PR				
Negative	1	8	**9**	0.6553^{ns}
Positive	7	24	**31**	
No information	0	1	**1**	
HER2 (ERBB2)				
Negative	0	9	**9**	No statistical analysis
Positive	0	5	**5**	
No information	8	19	**27**	

Note: ns = No statistical significance

Table 3-15 The correlation between NANOG immunoexpression and clinicopathological features of the invasive breast carcinoma. The Fisher exact test at the 95% CI was applied for the analysis of the correlation between NANOG immunoreactivity, including negative and positive, and clinicopathological features of 43 breast cancers.

Clinicopathological features	NANOG immunoexpression (N = 43)			
	Negative (N = 11/43)	**Positive (N = 32/43)**	**Total N**	***p*-value**
Age (Years old)				
≤ 50	7	19	**26**	1.0000^{ns}
> 50	4	13	**17**	
Tumour size (cm)				
≤ 2.0	3	9	**12**	1.0000^{ns}
> 2.0	8	23	**31**	
Tumour grade/differentiation				
Low (I & II)/Well & Moderate	4	11	**15**	1.0000^{ns}
High (III)/Poor	7	21	**28**	
Axillary lymph node				
No metastasis	8	13	**21**	0.0883^{ns}
Metastasis	3	19	**22**	
ER				
Negative	2	7	**9**	1.0000^{ns}
Positive	9	24	**33**	
No information	0	1	**1**	
PR				
Negative	3	6	**9**	0.6756^{ns}
Positive	8	25	**33**	
No information	0	1	**1**	
HER2 (ERBB2)				
Negative	1	8	**9**	**0.0230***
Positive	4	1	**5**	
No information	6	23	**29**	

Note: ns = No statistical significance

Cancers that were positive immunoexpression for DPPA3 (32 cases), EDR1 (33 cases), and NANOG (32 cases), were analysed for any correlation between the H-score values and their clinicopathological features, using the unpaired *t*-test at the 95% CI (Table 3-16 to 3-18). There was no significant relationship between the level of protein expression for DPPA3-, EDR1-, and NANOG-positive carcinomas and the following features: the patients' age group (≤ 50 years old and > 50 years old), tumour size (≤ 2.0 cm and > 2 cm); and axillary lymph node metastasis ($p > 0.050$). For DPPA3- and EDR1-positve cancers, approximately 70% of cases that were high-grade (grade III) showed lower expression of these proteins ($p = 0.0050$ and 0.0332, respectively) than the low-grade (grade I and II) carcinomas. The NANOG-positive cancers had no difference in expression of NANOG between the low- and the high-grade tumours ($p = 0.2131$).

One cancer had no information for the status of ER and PR, thus the correlation analysis was performed on 31 DPPA3-, 32 EDR1-, and 31 NANOG-positive cases. There was no statistically different level of each protein expression in cases with and without expression of ER and PR ($p > 0.050$). The HER2 status was not known in 21 out of 32 (65.62%), 19 out of 33 (57.58%), and 23 out of 32 (71.88%) DPPA3-, EDR1-, and NANOG-positive cancers, respectively. The breast cancers that were positive for these proteins infrequently expressed HER2: 2 out of 11 DPPA3-positive cases, 5 out of 14 EDR1-positive cases, and one out of 9 NANOG-positive cases.

Table 3-16 The correlation between the H-scores and clinicopathological features of DPPA3 (STELLA)-positive breast carcinomas. The unpaired *t*-test at the 95% CI was applied for the analysis of the correlation between the H-scores and clinicopathological features of 32 DPPA3-positive invasive breast cancers.

Clinicopathological features	DPPA3^{+} (N = 32)					
	N (%)	H-score		Unpaired t-test (95% CI)		
		X	SD	*t*	*df*	*p*-value
Age (Years old)						
≤ 50	18/32 (56.25%)	126.30	35.92	1.286	30	0.2083^{ns}
> 50	14/32 (43.75%)	147.40	56.71			
Tumour size (cm)						
≤ 2.0	9/32 (28.12%)	132.40	40.85	0.232	30	0.8181^{ns}
> 2.0	23/32 (71.88%)	136.80	49.45			
Tumour grade/differentiation						
Low (I & II)/Well & Moderate	9/32 (28.12%)	171.00	65.68	3.027	30	**0.0050****
High (III)/Poor	23/32 (71.88%)	121.70	27.81			
Axillary lymph node						
No metastasis	16/32 (50.00%)	134.90	40.02	0.073	30	0.9424^{ns}
Metastasis	16/32 (50.00%)	136.20	53.68			
ER						
Negative	8/31 (25.81%)	123.30	31.14	0.730	29	0.4715^{ns}
Positive	23/31 (74.19%)	137.10	50.14			
No information	1/32 (3.12%)					
PR						
Negative	7/31 (22.58%)	126.40	32.20	0.459	29	0.6496^{ns}
Positive	24/31 (77.42%)	135.60	49.59			
No information	1/32 (3.12%)					
HER2/neu (c-erb B2)						
Negative	9/11 (81.82%)	126.90	41.86	No statistical analysis		
Positive	2/11 (18.18%)	120.80	13.08			
No information	21/32 (65.62%)					

Note: ns = No statistical significance

Table 3-17 The correlation between the H-scores and clinicopathological features of EDR1 (PHC1)-positive breast carcinomas. The unpaired *t*-test at the 95% CI was applied for the analysis of the correlation between the H-scores and clinicopathological features of 33 EDR1-positive invasive breast cancers.

Clinicopathological features	EDR1^{+} (N = 33)					
	N (%)	H-score		Unpaired t-test (95% CI)		
		X	SD	*t*	*df*	*p*-value
Age (Years old)						
≤ 50	21/33 (63.64%)	147.40	61.78	1.249	31	0.2211^{ns}
> 50	12/33 (36.36%)	177.30	73.39			
Tumour size (cm)						
≤ 2.0	7/33 (21.21%)	147.70	55.31	0.466	31	0.6445^{ns}
> 2.0	26/33 (78.79%)	161.10	70.16			
Tumour grade/differentiation						
Low (I & II)/Well & Moderate	10/33 (30.30%)	195.30	64.76	2.229	31	**0.0332***
High (III)/Poor	23/33 (69.70%)	142.20	62.16			
Axillary lymph node						
No metastasis	15/33 (45.45%)	160.10	65.12	0.139	31	0.8905^{ns}
Metastasis	18/33 (54.54%)	156.80	69.86			
ER						
Negative	7/32 (21.88%)	125.20	59.21	1.439	30	0.1604^{ns}
Positive	25/32 (78.12%)	166.00	67.95			
No information	1/33 (3.03%)					
PR						
Negative	8/32 (25.00%)	120.50	56.41	1.839	30	0.0758^{ns}
Positive	24/32 (75.00%)	169.30	67.37			
No information	1/33 (3.03%)					
HER2/neu (c-erb B2)						
Negative	9/14 (64.28%)	151.90	80.16	0.987	12	0.3433^{ns}
Positive	5/14 (35.71%)	113.00	46.62			
No information	19/33 (57.58%)					

Note: ns = No statistical significance

Table 3-18 The correlation between the H-scores and clinicopathological features of NANOG-positive breast carcinomas. The unpaired *t*-test at the 95% CI was applied for the analysis of the correlation between the H-scores and clinicopathological features of 32 NANOG-positive invasive breast cancers.

Clinicopathological features	NANOG$^+$ (33 cases)					
	N (%)	H-score		Unpaired t-test (95% CI)		
		X	SD	*t*	*df*	*p*-value
Age (Years old)						
≤ 50	19/32 (59.38%)	102.50	4.18	1.231	30	0.2279^{ns}
> 50	13/32 (40.62%)	109.80	25.53			
Tumour size (cm)						
≤ 2.0	9/32 (28.12%)	102.10	2.99	0.715	30	0.4801^{ns}
> 2.0	23/32 (71.88%)	106.80	19.46			
Tumour grade/differentiation						
Low (I & II)/Well & Moderate	11/32 (34.38%)	110.50	28.19	1.272	30	0.2131^{ns}
High (III)/Poor	21/32 (65.62%)	102.80	2.85			
Axillary lymph node						
No metastasis	13/32 (40.62%)	103.70	5.38	0.497	30	0.6231^{ns}
Metastasis	19/32 (59.38%)	106.70	21.25			
ER						
Negative	7/31 (22.58%)	102.30	1.98	0.569	29	0.5736^{ns}
Positive	24/31 (77.42%)	106.50	19.13			
No information	1/32 (3.12%)					
PR						
Negative	6/31 (19.35%)	102.70	1.86	0.455	29	0.6528^{ns}
Positive	25/31 (80.64%)	106.20	18.78			
No information	1/32 (3.12%)					
HER2 (ERBB2)						
Negative	8/9 (88.89%)	104.30	5.04	No statistical analysis		
Positive	1/9 (11.11%)	102.00	-			
No information	23/32 (71.88%)					

Note: ns = No statistical significance

3.3.2.2 (10). Summary of immunohistochemical findings for expression of DPPA3 (STELLA), EDR1 (PHC1), and NANOG protein in invasive breast carcinomas

Immunoreactivity for DPPA3, EDR1, and NANOG proteins in the invasive breast carcinomas was considered positive when their H-scores were > 100, >75, and > 98, respectively. Expression of DPPA3 (94%) and NANOG (47%) proteins was commonly present in the cytoplasm, whereas EDR1 protein (76%) expressed predominantly in the nucleus and occasionally in the cytoplasm. These proteins showed independent expression and co-expression in the breast carcinoma tissue. In addition, DPPA3 linearly correlated with only NANOG.

Approximately 53% of the breast carcinomas had higher expression of DPPA3 in the normal mammary epithelium (NME) compared to their associated invasive carcinoma (ICa) component. Both NME and ICa showed no significant difference in expression of EDR1 (54%) and NANOG (47%) proteins. Expression of DPPA3, EDR1, and NANOG proteins did not correlate with the patients' age group, tumour size, axillary lymph node metastasis, and the status of ER and PR. Breast cancers that were positive for DPPA3 and EDR1 had a low level of expression of these proteins in high graded (grade III)/poorly differentiated carcinomas (70%). Expression of NANOG protein did not relate to the tumour grade/differentiation. Expression of all 3 proteins was infrequently present in HER2-positive cancers, but numbers were small.

3.3.3. *Expression of DPPA3 (STELLA), EDR1 (PHC1), and NANOG mRNA and proteins*

3.3.3.1. The correlation between mRNA expression and immunoexpression in breast carcinomas

Expression of DPPA3 mRNA in the surrounding normal breast (SNB) and the invasive breast carcinoma (IC) tissues was similar to immunoexpression of DPPA3 protein in the normal mammary epithelium (NME) and the invasive carcinoma (ICa) component, with expression of DPPA3 mRNA and protein being higher in non-neoplastic breast parenchyma (either SNB or NME) than in breast carcinoma tissue. EDR1 and NANOG mRNA were also expressed at a higher level in the SNB compared to the IC, but there were no differences in protein levels between the NME and the ICa [Section 3.3.1.6 and Figure 3-4; Section 3.3.2.2 (8) and Figure 3-25].

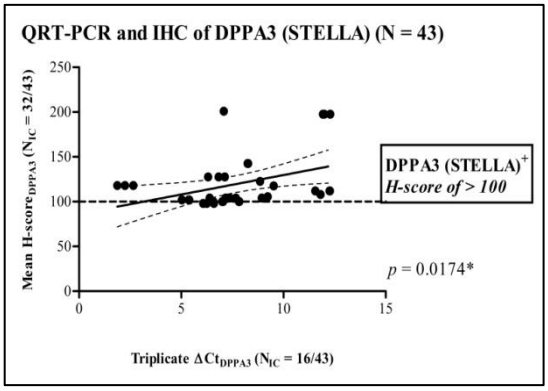
In the breast carcinomas, DPPA3 and NANOG had no independent mRNA expression but showed infrequently independent protein expression (5%). Independent expression of EDR1 mRNA was detected (4%) lesser than protein expression (12%). DPPA3 mRNA did not co-express with either EDR1 mRNA or NANOG mRNA, but co-expression of these proteins was infrequently detected (8%). Co-expression of EDR1 mRNA and NANOG mRNA was present in approximately 56% of cancers, but co-expression of their proteins was at a much lower frequency (8%). DPPA3 mRNA was expressed in combination with both EDR1 and NANOG mRNA expression (38%) was less frequent than protein co-expression (55%) [Section 3.3.1.7 and Figure 3-5; Section 3.3.2.2 (7) and Figure 3-24].

Based on Pearson correlation coefficient (r) at the 95% CI, cancers showed a significant linear correlation between DPPA3 protein and NANOG protein. EDR1 and

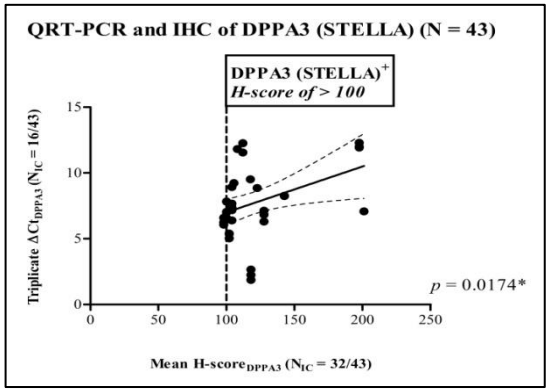
NANOG showed significant correlation between their mRNAs but had no relationship between their proteins [Section 3.3.1.7 and Figure 3-6; Section 3.3.2.2 (7) and Figure 3-23]. Expression of DPPA3 mRNA had a significant linear correlation with DPPA3 immunoexpression (p = 0.0174), whereas EDR1 and NANOG showed no statistically significant relationship between expression of mRNA and immunoreactivity (p = 0.8835 and p = 0.5503, respectively) (Figure 3-26).

3.3.3.2. Comparison of mRNA expression between negative and positive immunoexpression in breast carcinomas

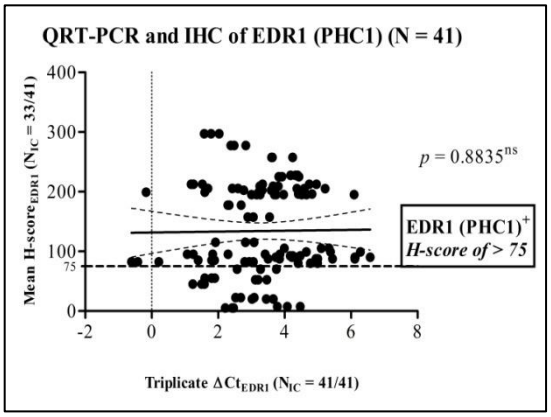
The ΔCt values for expression of DPPA3 (STELLA), EDR1 (PHC1), and NANOG mRNA in the breast carcinoma tissue (43, 41, and 43 cases, respectively) was compared between negative and positive immunoexpression by using the unpaired *t*-test at the 95% CI (Figure 3-27). For DPPA3, expression of mRNA was detectable in only 3 out of 11 DPPA3-negative cases ($X_{\Delta Ct}$ = 6.855; $SD_{\Delta Ct}$ = 0.633) and was frequently present in DPPA3-positive cases [N = 13/32 (40.62%); $X_{\Delta Ct}$ = 7.965; $SD_{\Delta Ct}$ = 3.047]. Although the statistical analysis was not undertaken due to a small number of DPPA3-negative cases, both DPPA3-negative and DPP3-positive cases showed similarity in mRNA expression. Expression of EDR1 mRNA was detected in all 8 EDR1-negative and 33 EDR1-positive cancers. EDR1 mRNA showed lower expression in EDR1-positive cancers compared to EDR1-negative cases (p = 0.0316). For NANOG, expression of mRNA was found in all 11 NANOG-negative cases ($X_{\Delta Ct}$ = 5.907; $SD_{\Delta Ct}$ = 2.677) and approximately 94% of NANOG-positive cases ($X_{\Delta Ct}$ = 5.966; $SD_{\Delta Ct}$ = 2.889). There was no significantly different level of mRNA expression between NANOG-negative and NANOG-positive cancers (p = 0.9215). However, these results were based on 8 EDR1-negative cancers and 11 NANOG-negative cancers.



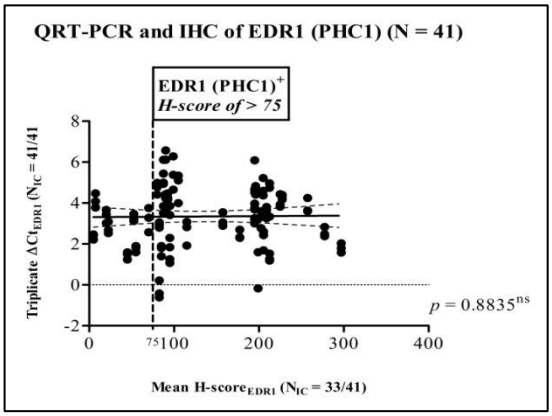
(A)



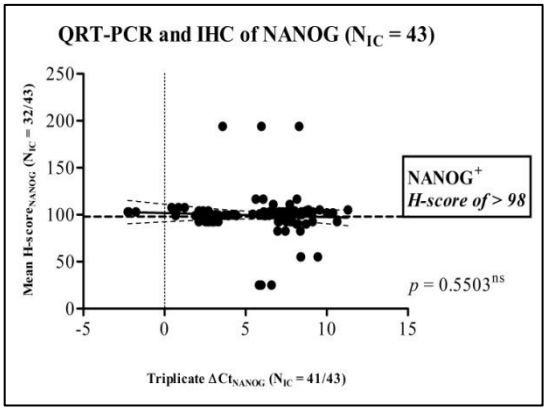
(B)



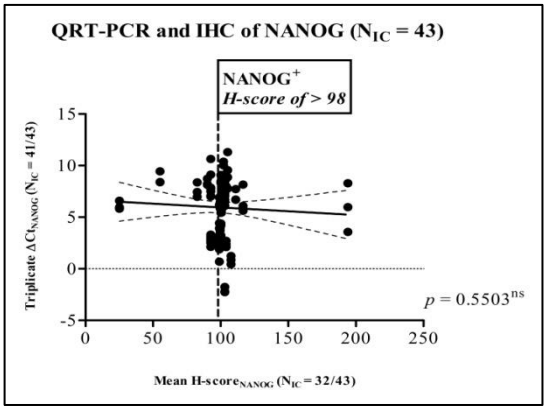
(C)



(D)

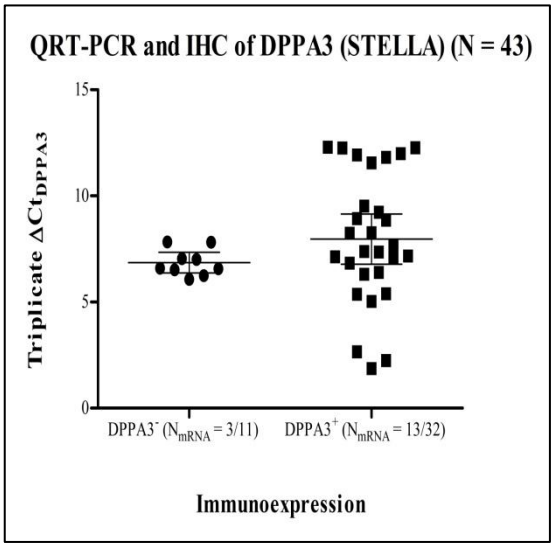


(E)

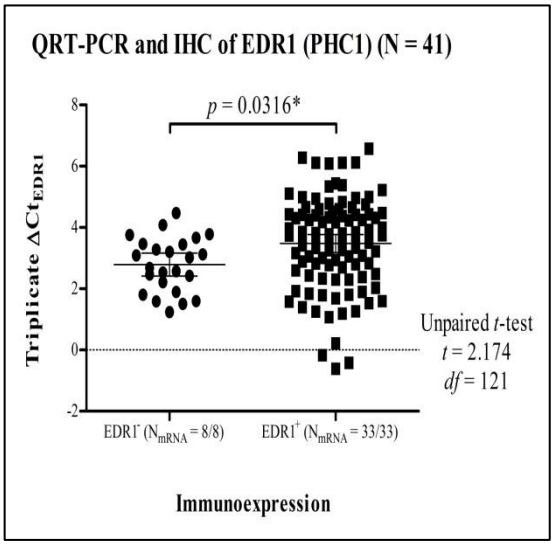


(F)

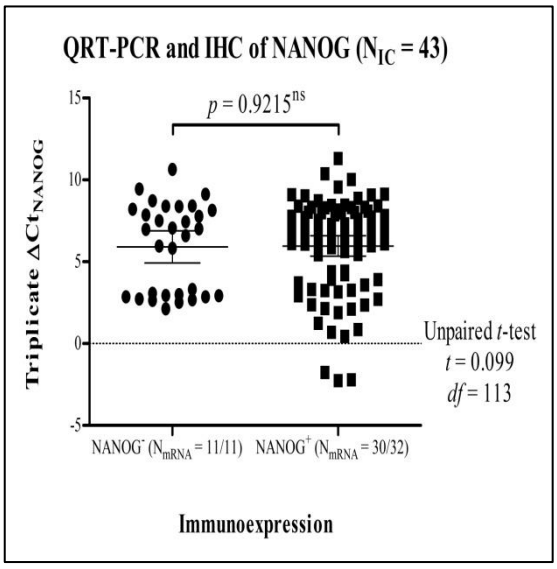
Figure 3-26 The correlation between expression of DPPA3 (STELLA), EDR1 (PHC1), and NANOG mRNA and immunoexpression in the invasive breast carcinomas. The Pearson correlation coefficient (r) at the 95% CI was used for the analysis of the correlation between the ΔCt values for expression of DPPA3, EDR1, and NANOG mRNA and their immunostaining H-scores in 43, 41, and 43 breast carcinoma tissue samples, respectively. (A) and (B) DPPA3 mRNA and protein; (C) and (D) EDR1 mRNA and protein; and (E) and (F) NANOG mRNA and protein



(A)



(B)



(C)

Figure 3-27 Comparison of expression of DPPA3 (STELLA), EDR1 (PHC1), and NANOG mRNA between negative- and positive-immunoexpressed invasive breast carcinomas. The unpaired *t*-test at the 95% CI was used for comparison of the ΔCt values for expression of DPPA3, EDR1, and NANOG mRNA between their negative and positive immunoreactivity in 43, 41, and 43 breast carcinoma tissue samples, respectively. (A) DPPA3; (B) EDR1; and (C) NANOG

3.3.3.3. Correlation of mRNA expression and immunoexpression in the breast carcinomas with clinicopathological features

As described in Section 3.3.1.9 and Section 3.3.2.2 (9), expression of DPPA3 (STELLA), EDR1 (PHC1), and NANOG mRNA and proteins did not significantly correlate with the patients' age group (≤ 50 years old and > 50 years old) and expression of ER and PR ($p > 0.050$). For tumour size (≤ 2.0 cm and > 2 cm), there was no correlation with expression of DPPA3 and EDR1 mRNA and proteins ($p > 0.050$). Approximately 65% of cancers with a size greater than 2.0 cm showed low expression of NANOG mRNA ($p = 0.0270$), but NANOG immunoreactivity had no relationship to the tumour size ($p = 1.0000$).

Expression of DPPA3 mRNA had no correlation with tumour grade/differentiation ($p = 0.7944$), but DPPA3-positive cancers showed lower expression of DPPA3 protein in high-grade (grade III)/poorly differentiated group ($p = 0.0050$). Approximately 60% of high-grade cancers had a lower level of expression of EDR1 ($p = 0.0229$) and NANOG ($p = 0.0002$) mRNA compared to the low graded cancers. The EDR1-positive cancers also showed lower expression of EDR1 protein in approximately 66% of high-grade cases ($p = 0.0332$). For NANOG-positive cancers, there was no relationship between expression of NANOG protein and tumour grade/differentiation ($p = 1.0000$).

A small number of breast carcinomas with axillary lymph node metastasis [N = 7/52 (13.46%)] showed high expression of DPPA3 mRNA ($p = 0.0029$) but there was no significant correlation with protein expression ($p = 1.0000$). Both expression of EDR1 and NANOG mRNAs and proteins did not relate to the presence or not of axillary lymph node metastasis ($p > 0.050$).

DPPA3 and NANOG mRNA and proteins were expressed infrequently in HER2-positive breast cancers. The status of HER2 did not correlate with expression of EDR1 mRNA in 16 breast carcinomas ($p = 0.0533$), nor with the 14 EDR1-positive cancers ($p = 0.3433$).

3.3.3.4. Summary of expression of DPPA3 (STELLA), EDR1 (PHC1), and NANOG mRNA and proteins in the breast carcinoma tissue

Expression of DPPA3, EDR1, and NANOG mRNA was higher in non-neoplastic breast tissue and lower in breast carcinomas. Only DPPA3 protein showed a similar pattern, whereas EDR1 and NANOG proteins showed no differences in expression between normal mammary epithelium and breast carcinomas. In breast cancers, independent expression was found only for EDR1 mRNA but was observed for DPPA3, EDR1, and NANOG proteins. Co-expression was found between EDR1 mRNA and NANOG mRNA but was present infrequently for all 3 proteins. Expression of EDR1 mRNA significantly correlated with expression of NANOG mRNA. Only expression of DPPA3 protein showed correlation with expression of NANOG protein.

Expression of DPPA3 mRNA related to its immunoexpression, but there was no correlation between expression of EDR1 and NANOG mRNA and their protein expression. There was no correlation between H-scores for DPPA3 and NANOG and the level of their mRNA expression, for both negative and positive cases, but EDR1 mRNA was expressed at higher levels in those EDR1-positive cancers.

Expression of all 3 proteins in the breast carcinoma tissue did not relate to the patient age and the status of ER and PR. Tumours with a size greater than 2.0 cm had low expression of NANOG mRNA only. High-grade (grade III)/poorly differentiated cancers showed low expression of EDR1 and NANOG mRNA, and had lower

expression of DPPA3 and EDR1 proteins even within cancers considered positive. A small number of breast carcinomas with axillary lymph node metastasis had a high level of expression of DPPA3 mRNA.

3.4. *DISCUSSION*

In breast cancer, quantitative reverse transcriptase polymerase chain reaction (QRT-PCR) and immunohistochemistry (IHC) are routinely used to evaluate expression of mRNA and protein, respectively, and to correlate results with diagnostic and prognostic data (Barberis et al. 2008, Benoy et al. 2004, Holm et al. 2008, Tischkowitz et al. 2007). Both techniques were applied in this study for the determination of mRNA and protein expression of 4 putative stem cell genes (*DPPA3*, *EDR1*, *GDF3*, and *NANOG*) in non-neoplastic and malignant breast tissues. However, each method must be first optimised for the most accurate result (Arnould, Coudert & Fumoleau 2009).

Generally, contamination is a major potential problem of the PCR process. For QRT-PCR, this is especially contamination of NTC wells of the real-time plate (Bustin, Nolan a and b, Sproul 2006). This study showed that the positions of the triplicate NTC wells on the QRT-PCR set-up plate were unlikely to give rise to cDNA contamination, in contrast to a previous study (Bustin, Nolan 2004). This study confirms that careful preparation of the master mix (a combination of all common reactants) and sterile ultrapure water was the most important factor in the elimination of contamination in the NTC wells.

The standard curves of the TaqMan® QRT-PCR assay in Section 3.3.1.2 had high values of the y-intercept which was likely to relate to degradation of mRNA and/or

cDNA template (s) during storage or from repeated thawing and freezing (Adams 2006). After reverse transcription, cDNA is degraded faster than mRNA because of the vulnerability of a cDNA-RNA complex to nucleases or buffer composition-induced degradation (Wilkening, Bader 2004). Moreover, degradation of RNA can affect the result of gene expression analysis (Ho-Pun-Cheung et al. 2009, Opitz et al. 2010, Strand et al. 2007), especially the Ct value from the QRT-PCR assay. Nevertheless, this study demonstrated the acceptable QRT-PCR efficiency because the tissue type and extraction procedure are probably the major factors affecting PCR efficiency in comparison with degraded RNA. Usually, a housekeeping/maintenance gene such as *GAPDH* or *HPRT1* is used as an endogenous reference gene to compensate for uneven cell numbers, RNA quality, and/or reverse transcription efficiency of individual samples (Fleige et al. 2006, Weisser et al. 2004). However, a single reference gene can yield inaccurate quantification of gene (s) of interest, particularly gene expression in different kinds of human tissue sample, due to variations in the level of RNA expression (Thellin et al. 1999, Tricarico et al. 2002, Vandesompele et al. 2002, Warrington et al. 2000). Hence, the mean Ct value of 3 endogenous reference genes (*GAPDH*, *HPRT1*, and *TFRC*) was applied to this TaqMan® QRT-PCR assay for more accurate determination of mRNA expression (Cai et al. 2007, Glenn et al. 2007) because individual up-regulation of any of these 3 reference genes in breast cancer may lead to misinterpretation of the results (Jiang, Elliott & Head 2010, Lin et al. 2001, Revillion et al. 2000). However, *HPTR1* has been recommended previously as an appropriate single housekeeping/maintenance gene instead of the unnecessary use of multiple reference genes for accurate normalisation of quantitative gene expression measurements in cancer research on clinical samples (de Kok et al. 2005).

In the QRT-PCR analysis of cell lines, HBL-100 was used as a non-neoplastic control since it was originally derived from normal mammary epithelial cells in an early lactation. Although HBL-100 is non-tumourigenic, this cell line might be inappropriate for being such a control due to its characteristics of immortality and in vitro transformation (Madsen, Briand 1990). Other non-tumourigenic mammary epithelial cell lines, such as MCF 10A (CRL-10317^{TM}, ATCC$^{®}$, UK) derived from fibrocystic change, could be used as a non-neoplastic control, but they were not available for this study. In addition, expression of GDF3 mRNA was undetectable in all breast cancer cell lines. This finding for MCF7 was contrary to a previous study of Ezeh and colleagues (2005) (Ezeh et al. 2005). However, co-expression of DPPA3, EDR1, and NANOG mRNA was found in breast cancer cell lines.

Western blotting of mixed germ cell tumour cell line, breast cancer cell lines, and a few breast carcinoma tissue samples showed different levels of expression of DPPA3, EDR1, and NANOG proteins with various unexpected bands at lower and higher molecular weights. The possible causes of this phenomenon are (I) post-translational modifications (PTMs) of intracellular proteins, i.e. glycosylation and phosphorylation (Borner et al. 1989, Capony et al. 1987, Liu et al. 1994); (II) the same proteins with different structures and functions [protein isoforms (splice variants)] (Sasso et al. 2011); (III) the presence of dimers, multimers, or protein-protein interaction due to incompletely reduced and denatured protein sample; (IV) cleaved or digested target protein; (V) high concentration of primary and/or secondary antibody (-ies); and (VI) a cross-reactivity of primary antibody with the same or similar epitopes on other proteins (Moore 2009). Alternative splicing (AS) of a pre-mRNA is an another fundamental mechanism for producing aforementioned protein isoforms (splice variants) from a single coding sequence of individual genes (Kalnina et al. 2005, Matlin, Clark & Smith

2005). In addition, these commercial antibodies might not be suitable for detection of DPPA3, EDR1, and NANOG proteins isolated from breast tissue by Western blotting, but might be more useful to evaluate those protein expression for the other tissue types and/or by different techniques.

Different levels of protein expression of the 3 putative stem cell genes has been previously reported in embryonic and adult cells. The antibody against DPPA3 from abcam® (ab19878) detected 17 and 37 kDa bands in mouse embryonic germ cell lysate (http://www.abcam.com/Stella-antibody-Primordial-Germ-Cell-Marker-ab19878.html). Sato and colleagues (2002) also demonstrated a 16-kDa DPPA3 protein in primordial germ cells of mice (Sato et al. 2002). EDR1 protein can be detected as both specific and non-specific bands in normal human cells and cancer cell lines with antibodies from abcam® (ab54954) (http://www.abcam.com/EDR1-PHC1-antibody-ab54954.html) and Sigma-Aldrich® (HPA006973) (http://www.proteinatlas.org/ENSG00000111752/

cancer/breast+cancer). Human NANOG protein has been found of different molecular weights of 29 – 48 kDa in various cell types from embryonic to malignant cells (Ezeh et al. 2005, Ambady et al. 2010, Bouskine et al. 2010, Busch et al. 2009, Chan et al. 2009, Chang et al. 2009, Eberle et al. 2010, Jeter et al. 2009, Salmina et al. 2010, Zhang et al. 2006, Watanabe et al. 2010).

Since, Western blotting showed different levels of expression of DPPA3, EDR1, and NANOG proteins in the breast cancer cell lines and a few breast carcinoma tissues, it was felt that expression of these proteins should be investigated further in more breast cancers. The IHC was selected to identify and specify both the cell type expressing the protein and the cellular location of the protein (Potemski et al. 2006), and because this technique is inexpensive and straightforward to perform. However, there are many

factors that can affect the immunohistochemical results, including fixation of tissue, duration and type of antigen retrieval, antibody specificity, antibody dilution, detection systems, scoring systems, and positive cut-off levels (Walker 2006).

Antigen retrieval is an important factor potentially affecting immunohistochemical staining in the formalin-fixed tissue (Bussolati, Leonardo 2008, D'Amico, Skarmoutsou & Stivala 2009, Shi, Shi & Taylor 2011, Werner, Von Wasielewski & Komminoth 1996). Formalin fixation induces the formation of cross-linking of cellular protein by methylene bridges (Fraenkel-Conrat, Brandon & Olcott 1947, Fraenkel-Conrat, Olcott 1948a and b) which leads to the masking of antigenic epitopes and an inefficient antibody binding. Heat-induced epitope antigen retrieval (HIER) method is commonly used for recovery of antigenicity by disruption to cross-linkages between formalin and protein (Bussolati, Leonardo 2008, D'Amico, Skarmoutsou & Stivala 2009, Shi, Shi & Taylor 2011, Werner, Von Wasielewski & Komminoth 1996, Shi, Key & Kalra 1991). This study showed that pressure cooker was a suitable antigen retrieval method for immunohistochemical assay by comparison with microwave antigen retrieval (Hunt, Attanoos & Jasani 1996, Pileri et al. 1997). Although Rhodes and colleagues (2001) reported that extension of heating time does improve tissue immunoreactivity, this study selected 30 seconds antigen retrieval for DPPA3 and EDR1 immunostainings because (I) prolonged heating time could be the risk of detachment of the tissue section (Lin, Shi 2011, Rhodes et al. 2001) and (II) the heating time of 60 seconds resulted in low H-score for EDR1 relating to a decrease in intensity of immunostaining. This study revealed that EDR1 antigen in the breast carcinoma tissue was possibly damaged in pressure cooker antigen retrieval with a time over 45 seconds, since there was decreased staining intensity (D'Amico, Skarmoutsou & Stivala 2009). For NANOG, the immunostaining appearance was a major factor in the selection of the optimal heating

time. The antigen retrieval time of 45 seconds yielded clearer cytoplasmic staining of NANOG compared to the heating time of 30 and 60 seconds.

The dilutions of primary antibodies in this study were optimised from the manufacturers' recommendations which gave more specific staining signal and less non-specific/background staining (Bussolati, Leonardo 2008, Lin, Shi 2011, Hsi 2001). In addition, the result of immunohistochemical staining is highly related to intra- and inter-observer subjectivity (Crowe et al. 2006, Kulka et al. 2006, Meyer et al. 2005, Ntoulia et al. 2006, Peiro et al. 2007, Rossi et al. 2006, Shin et al. 2006). Immunohistochemical assessment by myself on 2 separate occasions with an interval of 4 weeks showed some variation of H-scoring for DPPA3 immunoexpression but no variation in the H-score evaluation of EDR1 and NANOG proteins. To ensure that scoring of percentage and intensity was appropriate, a small number of breast carcinoma tissue samples were assessed with Professor Rosemary Walker. The mean H-score for each case was used for more precisely statistical analysis of expression of DPPA3, EDR1, and NANOG proteins in the breast carcinoma tissue.

Expression of DPPA3, EDR1, and NANOG mRNA and proteins was evaluated in normal breast tissue (NB) derived from reduction mammoplasties, surrounding normal breast tissue (SNB), and breast carcinoma tissue. All histopathological types of breast carcinoma was categorised as "non-invasive carcinoma (DCIS or intraductal carcinoma)" and "invasive (infiltrating) carcinoma (IC) (including combined DCIS and IDC-NST, IDC-NST, mucinous carcinoma, and ILC)". The reason for this categorisation was the tissue sections for each case derived from the large tumor mass might not accurately represent the whole histopathological appearances (Schlemmer et al. 2004). Expression of these genes in pure DCIS and the vascular invasion component in the IC tissue were not included in statistical analysis due to small numbers.

The cut-off H-scores were established for the determination of positive immunostaining of DPPA3, EDR1, and NANOG proteins as > 100, > 75, and > 98, respectively. The cancer cells mostly showed expression of these proteins in the cytoplasm for positive DPPA3 and NANOG and in the nuclei for positive EDR1. DPPA3 has previously been reported as showing both nuclear and cytoplasmic staining for different cellular stages in germ lineage cells of mice (Sato et al. 2002, Payer et al. 2003). Expression of EDR1 protein has been reported as nuclear in different kinds of normal tissue, benign tumour, and cancer (including non-haematologic and haematologic malignancies) (Sanchez-Beato et al. 2006, Dukers et al. 2004, Gunster et al. 2001, Raaphorst et al. 2004, van Leenders et al. 2007). However, Sanchez-Beato and colleagues (2006) found that normal breast tissue has no expression of EDR1 protein (Sanchez-Beato et al. 2006). For NANOG protein, there are various reported patterns of immunostaining in non-germ cell carcinomas: (I) mainly cytoplasmic staining in colorectal adenocarcinoma (Meng et al. 2010), and squamous cell carcinoma of cervix (Ye et al. 2008); (II) mainly nuclear staining in adenocarcinoma of stomach (Lin, Ding & Li 2011) and prostate gland (Jeter et al. 2009); and (III) both nuclear and cytoplasmic staining in invasive breast carcinoma (Ezeh et al. 2005). In contrast to my findings, Ezeh and colleagues (2005) showed that the NB had undetectable level of *NANOG* gene by using PCR and has no specific immunostaining. The mechanism for intracellular distribution of DPPA3, EDR1, and NANOG proteins is still unknown in cancers, thus further functional studies of these stem cell associated genes are required (See Future direction in Chapter 5 Discussion and Conclusion).

This study revealed that the NB had significantly lower mRNA expression of *DPPA3*, *EDR1*, and *NANOG* genes than the SNB. Contrary to germ cell tumours (Giuliano et al. 2005, Busch et al. 2009, Payer et al. 2003, Hart et al. 2004, Hart et al.

2005, Korkola et al. 2006, Rodriguez et al. 2003), the level of mRNA expression of these genes was lower in the IC than in the SNB. Only DPPA3 protein showed lower expression in the invasive carcinoma component compared to the normal mammary epithelium (NME), this immunoexpression was similar to expression of DPPA3 mRNA. However, there were no significant differences in the level of expression of EDR1 and NANOG proteins between the NME and the invasive cancer cells. According to the cancer field effect (Braakhuis et al. 2003, Chai, Brown 2009, Yan et al. 2006), various mRNA expression of *DPPA3*, *EDR1*, and *NANOG* genes in either SNB or NME adjacent to the primary breast cancer might be the result of expansion of a genetically abnormal clone.

It has been reported that QRT-PCR results for mRNA expression are usually in concordance with the IHC result of protein expression (Sun et al. 2011, Susini et al. 2010). The possible causes for differences between QRT-PCR and IHC in my study are the following: (I) frozen tissue was used for the QRT-PCR assay but FFPE tissue section was used for the IHC. The FFPE process can result in loss of antigenicity (D'Amico, Skarmoutsou & Stivala 2009, Otali et al. 2009, Werner et al. 2000, Xie et al. 2011). However, the frozen tissue had been used previously so there could have been degradation of RNA due to freeze/thawing; and (II) The breast cancer cells have alternative splicing (AS) of a pre-mRNA or post-translational modifications (PTMs) of intracellular proteins leading to variations in protein expression as shown in the results of Western blotting. Therefore, proteomic (Hudelist et al. 2006, Mann, Jensen 2003) and transcriptome (Grigoriadis et al. 2006, Rhodes, Chinnaiyan 2005) analyses should be required further. Interestingly, I found that the percentage of breast carcinoma cells in the tissue sections did not impact on the ΔCt values for expression of DPPA3, EDR1,

and NANOG mRNAs in contrast to a previous report for QRT-PCR study (de Kok et al. 2000).

A previous report using RT-PCR showed that the tumour lymphovascular emboli of inflammatory breast carcinoma express markers of embryonic stem cells or pluripotent stem cells such as *DPPA3* and *NANOG* (Xiao et al. 2008). My study found a relationship between up-regulation of only DPPA3 mRNA and axillary lymph node metastasis, but there was no correlation of expression of NANOG mRNA with the presence of axillary lymph node metastasis (Ben-Porath et al. 2008). High-grade (grade III)/poorly differentiated IC showed down-regulation of EDR1 and NANOG mRNA and had low expression of DPPA3 and EDR1 proteins. Previous reports have shown a correlation between an aggressive phenotype of cancers and over-expression of *DPPA3* (Xiao et al. 2008) and *NANOG* (Ben-Porath et al. 2008, Chiou et al. 2008) genes. However, expression of EDR1 protein was demonstrated in 6 out of 9 cases of low-grade astrocytoma compared to only 2 out of 10 cases of high-grade astrocytoma (Sanchez-Beato et al. 2006). This study showed that low expression of NANOG mRNA was present in breast cancers having size greater than 2.0 cm, but Ben-Porath and colleagues (2008) reported that expression of *NANOG* gene was not correlated with tumour size (Ben-Porath et al. 2008).

In summary, *DPPA3*, *EDR1*, and *NANOG* were not expressed in normal breast tissues, but there was variable expression in carcinomas, and this tended to be higher in the associated surrounding normal breast. No expression of *GDF3* was found in breast carcinomas. The lower levels of expression of *DPPA3*, *EDR1*, and *NANOG* in carcinomas was correlated with high-grade/poorly differentiated carcinoma and down-regulation of *NANOG* was associated with increased tumour size. Conversely, breast carcinomas with higher expression of *DPPA3* was associated with the presence of

axillary lymph node metastasis. Therefore, expression of these genes at the mRNA and/or protein level may provide useful prognostic information in invasive breast carcinoma.

CHAPTER 4

COPY NUMBER VARIATIONS (CNVs) OF *DPPA3* (*STELLA*), *EDR1* (*PHC1*), AND *NANOG* GENES IN BREAST CARCINOMA

4.1. *INTRODUCTION*

Normal diploid human somatic cells have 46 chromosomes containing 2 copies of each gene on each chromosome. One chromosome/gene copy is derived from each parent at fertilisation (Epstein 2003, Ross, Pawlina 2006). Differences in copy numbers of specific DNA segments are termed copy number variations (CNVs). These variations can be gain or loss (deletion) and are an underlying factor in many diseases, including mental illness, developmental disorders, and cancer (Conrad et al. 2010, Feuk, Carson & Scherer 2006, Freeman et al. 2006, Hastings et al. 2009, Kehrer-Sawatzki 2007, Redon et al. 2006, Scherer et al. 2007, Zhang et al. 2009a). In breast cancer, CNVs, especially gene amplification (multiple copies of genes), are correlated with clinical prognosis (Al-Kuraya et al. 2004, Chin et al. 2006, Hyman et al. 2002, Zhang et al. 2009b).

As discussed previously in the Introduction, genomic aberrations of mammary stem cells or progenitors have been associated with the development of different molecular phenotypes of breast cancer (Bombonati, Sgroi 2011, Lim et al. 2009, Melchor, Benitez 2008, Polyak 2007, Prat, Perou 2009). Human embryonic stem cells have been reported to have gains of chromosome 12p (Mayshar et al. 2010, Laurent et al. 2011). A number of putative stem cell-associated genes are located on chromosome 12p13, i.e. *DPPA3* (*STELLA* or *PGC7*), *EDR1* (*PHC1*, *HPH1*, or *RAE28*), *GDF3*, and *NANOG*, and these have been correlated previously with the pathogenesis of both non-

hematologic (Tanaka et al. 2009) and hematologic (Tokimasa et al. 2001) malignancies. Approximately 59% of invasive breast carcinoma has highest gene expression of chromosome 12p13 (Bertucci et al. 2006). Medullary carcinoma also shows amplification of this chromosomal region in approximately 33% (Bertucci et al. 2006, Vincent-Salomon et al. 2007). Furthermore, Yao and colleagues (2006) reported that chromosome 12p13 was amplified in metastatic breast carcinoma (Yao et al. 2006).

4.2. *AIMS*

The aims of this chapter are to:

(I) Determine CNVs of stem cell-associated genes on chromosome 12p13 (*DPPA3*, *EDR1*, and *NANOG*) in both breast cancer cell lines and tissues using TaqMan® copy number assays (CNAs) and Affymetrix® Genome-Wide Human SNP Array 6.0

(II) Relate the results of CNVs in breast carcinoma tissues to their clinicopathological features.

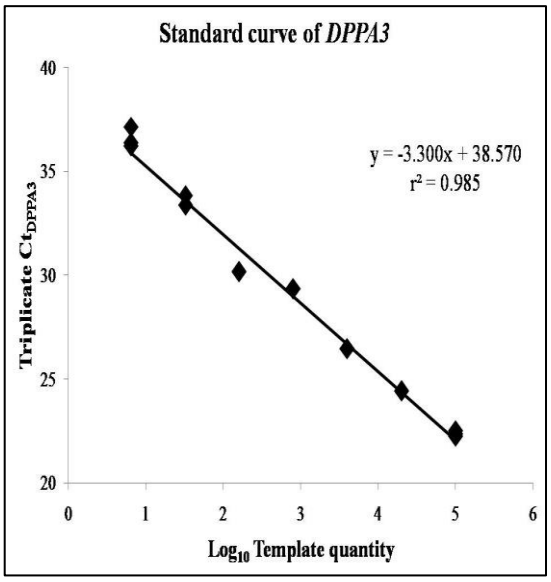
4.3. *RESULTS*

4.3.1. *Standard curves of TaqMan® copy number assays (CNAs)*

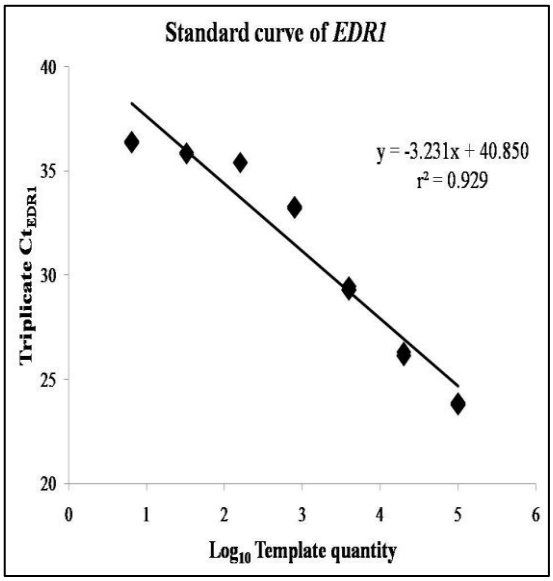
The standard curves of TaqMan® CNA for *DPPA3*, *EDR1*, and *NANOG* were generated from triplicate Ct values for 7 serial dilutions using commercial human genomic DNA (HGDNA) as a template (Figure 4-1 and Appendix 7). The calculation of the quantitative polymerase chain reaction (QPCR) efficiency (E) from the slope (m) of each standard curve yielded specific amplification of the TaqMan® CNA with an acceptable efficiency of 95% – 105% (~101% of *DPPA3*, ~104% of *EDR1*, and ~104% of *NANOG*). The coefficient of determination (r^2) of each reaction was close to 1.0, indicating the QPCR assay had the accuracy (validity) of the dilutions and precision (reproducibility) of pipetting. The y-intercept of each reaction was slightly higher than 37 cycles for acceptable values (38.570 cycles of *DPPA3*; 40.850 cycles of *EDR1*; 40.830 cycles of *NANOG*), indicating that the determination of the amount of DNA template would be slightly inaccurate at this level (Table 4-1).

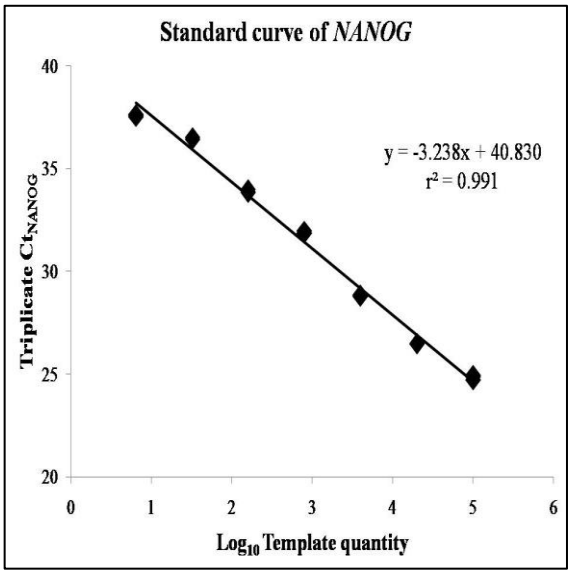
4.3.2. *Determination of the optimal endogenous reference gene*

The optimal endogenous reference gene was evaluated for each of the target genes (*DPPA3*, *EDR1*, and *NANOG*) on TaqMan® CNAs. The assay used 5 ng/µL of HGDNA as template and both *RNase P* and *TERT* as endogenous reference genes (Table 4-2). Based on the selective criteria of the mean ΔCt value close to 0 and a low standard deviation (SD) of ΔCt, *TERT* should be used as endogenous reference gene for *DPPA3*, whereas *RNase P* should be used as endogenous reference gene for *EDR1* and *NANOG*.



(A)





(B) (C)

Figure 4-1 Standard curves of TaqMan® copy number assays (CNAs). The standard curves of TaqMan® CNA for *DPPA3* (*STELLA*), *EDR1* (*PHC1*), and *NANOG* were constructed by triplicate Ct values for the 7 serial dilutions of commercial human genomic DNA (HGDNA) as a template. The assay was performed on 60 amplification cycles with manual threshold at 0.2 and automatic baseline. (A) *DPPA3*; (B) *EDR1*; and (C) *NANOG*

Table 4-1 Standard curve parameters. The calculation of the linear regression on Microsoft Excel® programme was applied to the evaluation of the slope (m), the QPCR efficiency (E), the y-intercept (b), and the coefficient of determination (r^2) for the standard curves in Figure 4-1.

Gene	Slope (m)	QPCR efficiency (E)* (%)	y-intercept (b) (Cycle)	Coefficient of determination (r^2)
DPPA3 (*STELLA*)	-3.300	100.923	38.570	0.985
EDR1 (*PHC1*)	-3.231	103.940	40.850	0.929
NANOG	-3.238	103.626	40.830	0.991

Note: * QPCR efficiency (E) = $\{[10^{(-1/\text{Slope})}] - 1\} \times 100\%$

Acceptable values: Amplification efficiency = 95% - 105%; the y-intercept = 33 – 37 cycles; and r^2 close to 1.00

Table 4-2 Comparison of the ΔCt values for different endogenous reference genes. The optimal endogenous reference gene for TaqMan® CNAs was determined on *DPPA3* (*STELLA*), *EDR1* (*PHC1*), and *NANOG* according to 5 ng/µL of commercial human genomic DNA (HGDNA) and different endogenous reference genes (*RNase P* and *TERT*). The assay was performed on 60 amplification cycles with manual threshold at 0.2 and automatic baseline. The selection was based on the mean ΔCt value close to 0 and the low standard deviation (SD) of ΔCt.

Gene	**Ct**	**X_{Ct}**	**SD_{Ct}**	**ΔCt/*RNase P***	**$X_{\Delta Ct}$**	**$SD_{\Delta Ct}$**	**ΔCt/*TERT***	**$X_{\Delta Ct}$**	**$SD_{\Delta Ct}$**
DPPA3	25.961	25.940	0.027	-0.994	-0.962	0.040	-0.695	-0.723	0.024
	25.910			-0.975			-0.734		
	25.950			-0.916			-0.740		
EDR1	27.117	27.129	0.073	0.162	0.227	0.085	0.461	0.466	0.095
	27.207			0.323			0.563		
	27.063			0.196			0.373		
NANOG	27.844	27.748	0.108	0.889	0.846	0.071	1.188	1.085	0.129
	27.770			0.885			1.126		
	27.631			0.764			0.941		
RNase P	26.955	26.902	0.047						
	26.885								
	26.867								
TERT	26.656	26.663	0.024						
	26.644								
	26.690								

4.3.3. *Precision (reproducibility) of quantitative polymerase chain reaction (QPCR) assay*

Expression of *DPPA3*, *EDR1*, and *NANOG* in HGDNA, NCCIT, and MCF7 were determined twice to assess the precision (reproducibility) of the QPCR assay. The paired assays showed no significant difference in the mean ΔCt values for these gene expressions in HGDNA, NCCIT, and MCF7 (paired *t*-tests at the 95% CI, $p > 0.050$) (Table 4-3), showing good reproducibility.

Table 4-3 Precision (reproducibility) of TaqMan® copy number assays (CNAs). TaqMan® CNA was performed twice using triplicate reaction to target genes [*DPPA3* (*STELLA*), *EDR1* (*PHC1*), and *NANOG*] and endogenous reference genes (*RNaseP* for *DPPA3* and *TERT* for *EDR1* and *NANOG*) in HGDNA, NCCIT, and MCF7. The assays were performed on 60 amplification cycles with manual threshold at 0.2 and automatic baseline. The paired *t*-test at the 95% CI was applied to the determination of precision (reproducibility) of QPCR assay.

Gene	ΔCt				Paired *t*-test at 95% CI		
	1st assay		2nd assay				
	X	SD	X	SD	*t*	*df*	*p*-value
DPPA3 (N = 3/3)	-0.898	0.760	-0.818	0.627	1.725	8	0.1228^{ns}
EDR1 (N = 3/3)	-0.759	0.201	-0.807	0.326	0.631	8	0.5458^{ns}
NANOG (N = 3/3)	-0.180	0.543	-0.199	0.563	1.476	8	0.1781^{ns}

Note: ns = No statistical significance

4.3.4. *CNVs of DPPA3 (STELLA), EDR1 (PHC1), and NANOG in the cancer cell lines*

The CNVs based on the ΔCt values for *DPPA3*, *EDR1*, and *NANOG* were evaluated in one mixed germ cell tumour cell line (NCCIT) as a positive control and 7 breast cancer cell lines (BCA CLs) (HBL-100, MCF7, MDA-MB-231, MDA-MB-436, MDA-MB-468, T47D, and ZR-75-1). One-way analysis of variance (ANOVA) showed that there was significant difference in copy number (CN) between cell lines ($p < 0.0001$) (Figure 4-2 A, C, and E). The predicted CN values were calculated from the relative quantification (RQ) values in relation to HGDNA (Figure 4-2 B, D, and F). NCCIT showed high level of amplification of *DPPA3* (CN = 8) and *EDR1* (CN = 7), but had single copy gain of *NANOG* (CN = 3). For *DPPA3* (Figure 4-2 B), heterozygous deletion [loss of heterozygosity (LOH) (CN = 1)] was observed in HBL-100 and T47D; normal diploid (CN = 2) was observed in MCF-7, MDA-MB-231, and MDA-MB-468; single copy gain was observed in ZR-75-1; and amplification was observed in MDA-MB-436 (CN = 8). For *EDR1* (Figure 4-2 D), HBL-100 showed homozygous deletion (CN = 0); MDA-MB-468 and T47D showed LOH; MCF7 and MDA-MB-231 showed normal diploid; and MDA-MB-436 and ZR-75-1 showed single copy gain. For *NANOG* (Figure 4-2 F), homozygous deletion was identified in MCF7 and MDA-MB-468; LOH was identified in HBL-100, MDA-MB-231, and T47D; normal diploid was identified in MDA-MB-436 and ZR-75-1.

Hence, according to the above findings for the cancer cell lines, further analysis of CNVs of these putative stem cell genes in human breast cancers was deemed appropriate.

4.3.5. *CNVs of DPPA3 (STELLA), EDR1 (PHC1), and NANOG in normal breast (NB), surrounding normal breast (SNB), and invasive breast carcinoma (IC)*

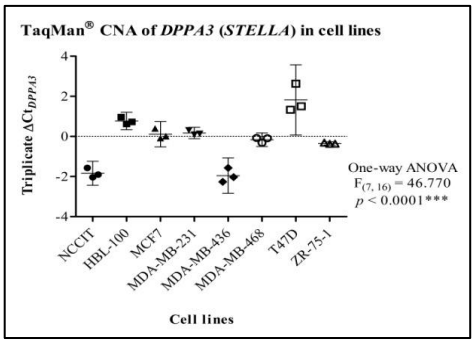
The CNVs based on the ΔCt values for *DPPA3*, *EDR1*, and *NANOG* were analysed in 10 NB, 35 SNB, and 42 IC samples. The unpaired *t*-test at the 95% CI was applied to compare the mean ΔCt value from each of samples (Figure 4-3). All 10 NB samples had a detectable CN value for *DPPA3*, but CN values for *EDR1* and *NANOG* were found in only 8 NB samples. The CN values for all 3 stem cell associated genes could be assessed in all 35 SNB and 42 IC samples.

All NB tissues and the majority of SNB tissues had ΔCt values for *DPPA3*, *EDR1*, and *NANOG* close to 0, indicating normal diploid CN of these genes. In addition, some SNB showed positive and negative ΔCt values for all 3 stem cell associated genes, indicating loss (deletion) and gain/amplification of CN, respectively. The unpaired *t*-test revealed that the CN values for *DPPA3*, *EDR1*, and *NANOG* in the NB were similar to the SNB ($p > 0.050$). The IC showed mostly gain/amplification (-ΔCt) of *DPPA3* and loss (deletion) (+ΔCt) of *EDR1* and *NANOG*. No negative ΔCt value for gain/amplification of *NANOG* was observed in the IC. For *DPPA3*, the IC had significantly higher CN values than both NB ($p = 0.0418$) and SNB ($p = 0.0005$). On the contrary, both NB and SNB had significantly higher CN values for *EDR1* and *NANOG* compared to the IC ($p \leq 0.050$). However, this statistical analysis was based on a small number of NB samples.

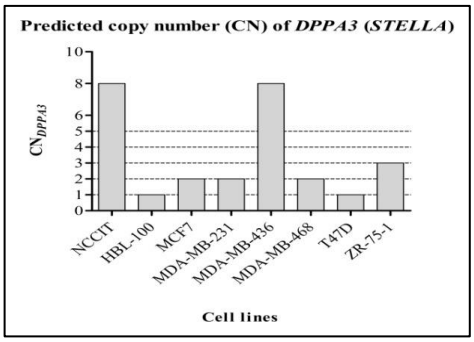
The predicted CN values for *DPPA3*, *EDR1*, and *NANOG* were calculated from their RQ values based on the ΔCt values, in relation to HGDNA (Table 4-4). The NB had normal diploid (CN = 2) of these stem cell associated genes. The majority of SNB showed normal diploid of *DPPA3* [N = 27 (77.14%)], *EDR1* [N = 25 (71.43%)], and

NANOG [N = 24 (68.57%)] genes. In addition, deletion (CN = 0 – 1) and single copy gain (CN = 3) of these genes were observed in the SNB but there was no gene amplification. Amplification of *DPPA3* and *EDR1* was detected in 9 (X_{CN} = 8; Min_{CN} = 5; Max_{CN} = 12) and one (CN = 5) out of 42 breast cancer cases, respectively. The IC also showed homozygous deletion (CN = 0), LOH (CN = 1), normal diploid, and single copy gain of both *DPPA3* and *EDR1* genes. For *NANOG*, homozygous deletion was present in 26 out of 42 (61.90%) cancer cases, whereas LOH and normal diploid were detected in 14 and 2 cases, respectively. No *NANOG* amplification was observed in the IC.

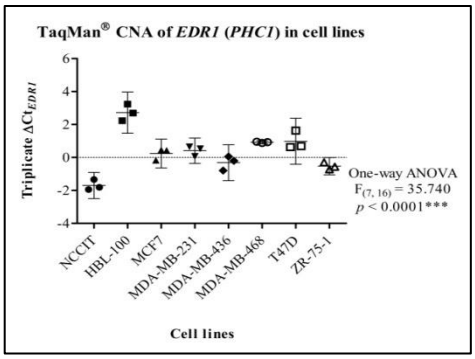
Interestingly, there was homozygous deletion of *NANOG* gene in approximately 62% of breast cancer cases, but expression of NANOG mRNA and protein could be detected in the cancer samples by QRT-PCR and IHC, respectively. These phenomena were possibly caused by the CN signal of TaqMan® CNAs derived from a single DNA segment. Therefore, the other exons of *NANOG* gene, which might have normal diploid or CN gain, still could express its mRNA and protein (See Figure 4-6 and Section 4.4 Discussion).



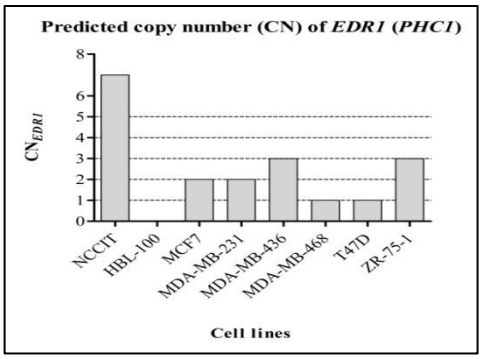
(A)



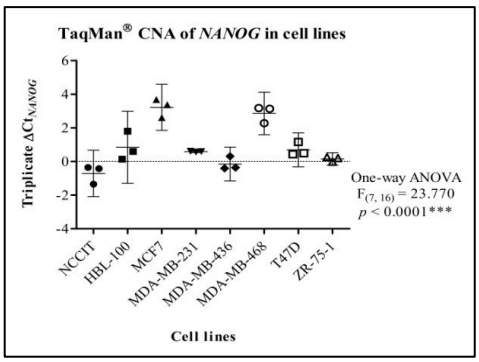
(B)



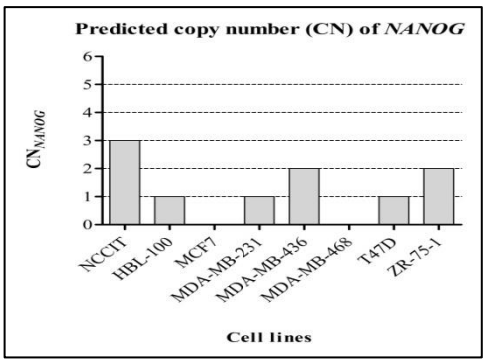
(C)



(D)

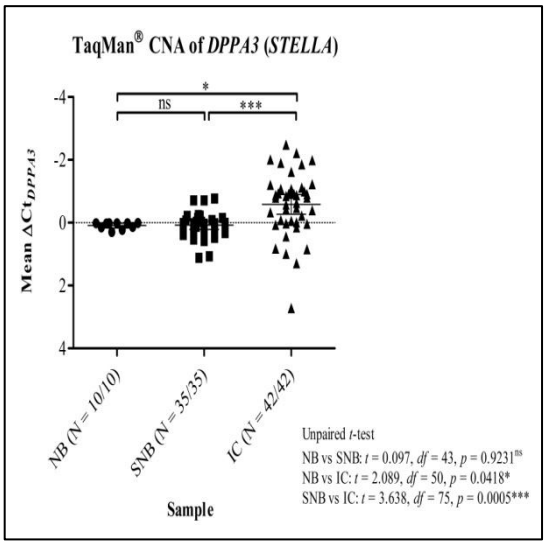


(E)

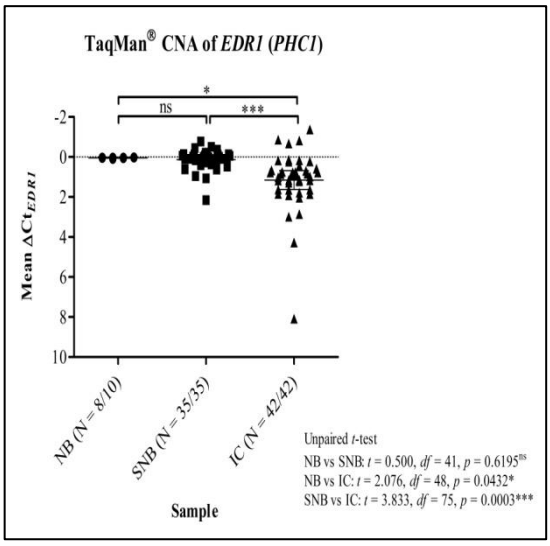


(F)

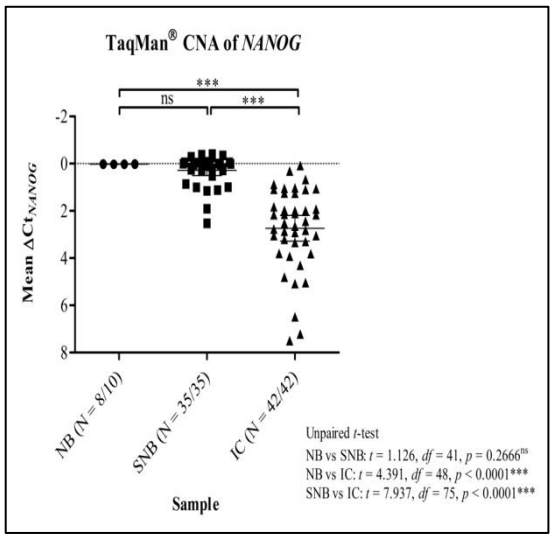
Figure 4-2 Copy number variations (CNVs) of *DPPA3* (*STELLA*), EDR1 (*PHC1*), and *NANOG* in the cancer cell lines. According to one-way analysis of variance (ANOVA), the CNVs based on the ΔCt values for *DPPA3*, *EDR1*, and *NANOG* were evaluated in one mixed germ cell tumour cell line (NCCIT) as a positive control and 7 breast cancer cell lines (BCA CLs). In relation to HGDNA, the predicted copy number (CN) values (Right) were calculated from the relative quantification (RQ) values and inferred as: CN of 0 for homozygous deletion; CN of 1 for heterozygous deletion [loss of heterozygosity (LOH)]; CN of 2 for normal diploid; CN of 3 for single copy gain; CN of 4 for two-copy gain; and CN of ≥ 5 for amplification. (A) and (B) *DPPA3*; (C) and (D) *EDR1*; and (E) and (F) *NANOG*



(A)



(B)



(C)

Figure 4-3 TaqMan® CNAs of *DPP*A3 (*STELLA*), *EDR1* (*PHC1*), and *NANOG* in breast tissue. Scatter diagrams were plotted from the mean ΔCt values for *DPPA3, EDR1*, and *NANOG* in 10 normal breast (NB), 35 surrounding normal breast (SNB), and 42 invasive breast carcinoma (IC) samples. Results show a horizontal line at the mean and error bars at the 95% CI. The unpaired *t*-test was used for comparison with the ΔCt values in each sample. (A) *DPPA3*; (B) *EDR1*; and (C) *NANOG*

Table 4-4 Predicted copy number (CN) of *DPPA3* (*STELLA*), *EDR1* (*PHC1*), and *NANOG* in breast tissue. In relation to HGDNA, the predicted CN values for *DPPA3*, *EDR1*, and *NANOG* were calculated from their relative quantification (RQ) values based on the ΔCt values in normal breast (NB), surrounding normal breast (SNB), and invasive breast carcinoma (IC).

Tissue (N)	**Predicted CN[§] (N) (%)**																	
	DPPA3						*EDR1*						*NANOG*					
	0	**1**	**2**	**3**	**4**	**≥ 5**	**0**	**1**	**2**	**3**	**4**	**≥ 5**	**0**	**1**	**2**	**3**	**4**	**≥ 5**
NB (10)	-	-	10	-	-	-	-	-	8[¶]	-	-	-	-	-	8[¶]	-	-	-
SNB (35)	-	5 (14.29%)	27 (77.14%)	3 (8.57%)	-	-	1 (2.86%)	5 (14.29%)	25 (71.43%)	4 (11.43%)	-	-	1 (2.86%)	7 (20.00%)	24 (68.57%)	3 (8.57%)	-	-
IC (42)	1 (2.38%)	4 (9.52%)	9 (21.43%)	6 (14.29%)	13 (30.95%)	9* (21.43%)	5 (11.90%)	28 (66.67%)	5 (11.90%)	1 (2.38%)	2 (4.76%)	1[#] (2.38%)	26 (61.90%)	14 (33.33%)	2 (4.76%)	-	-	-

Note:

§ CN of 0 is homozygous deletion; CN of 1 is heterozygous deletion [loss of heterozygosity (LOH)]; CN of 2 is normal diploid; CN of 3 is single copy gain; CN of 4 is two-copy gain; and CN of ≥ 5 is amplification.

¶ The CN values for *EDR1* and *NANOG* were detected in 8 out of 10 NB samples.

* The minimum, mean, and maximum CN values for *DPPA3* were 5, 8, and 12, respectively.

The CN value for *EDR1* was 5.

4.3.6. *The effect of percentage of invasive carcinoma component within the tissue section of breast cancer on the results of TaqMan® CNAs*

The percentage of invasive carcinoma within the tissue sections (≤ 50% and > 50%) was related to the ΔCt values for *DPPA3*, *EDR1*, and *NANOG* for 29 of the 42 cases (69.05%), using the unpaired *t*-test at the 95% CI (Table 4-5). There were 6 (20.69%) and 23 (79.31%) cases showing invasive carcinoma component of ≤ 50% and > 50%, respectively. Statistical analysis showed that the ΔCt values for these genes had no significant association with percentage of invasive carcinoma component using a 50% cut-off ($p > 0.050$). Nevertheless, this relationship was evaluated from a small number of samples with the invasive carcinoma component of ≤ 50%.

Table 4-5 The effect of a percentage of invasive carcinoma component in breast cancer tissue section on the results of TaqMan® copy number assays (CNAs). The unpaired *t*-test at the 95% CI was applied to the analysis of the relationship between a percentage of invasive carcinoma component in 29 BC tissue sections and the ΔCt values for *DPPA3* (*STELLA*), *EDR1* (*PHC1*), and *NANOG*.

Gene	**Percentage of invasive carcinoma component (N = 29)**				***p*-value[#]**
	≤ 50% [N = 6/29 (20.69%)]		**> 50% [N = 23/29 (79.31%)]**		
	$X_{\Delta Ct}$	**$SD_{\Delta Ct}$**	**$X_{\Delta Ct}$**	**$SD_{\Delta Ct}$**	
DPPA3	-0.776	0.726	-0.412	1.132	0.4649^{ns}
EDR1	0.983	0.481	1.593	1.772	0.4162^{ns}
NANOG	2.262	0.931	2.895	1.613	0.3688^{ns}

Note: # Unpaired *t*-test at the 95% CI

4.3.7. *The correlation of CNVs of DPPA3 (STELLA), EDR1 (PHC1), and NANOG in the IC with clinicopathological features*

The CNVs of *DPPA3*, *EDR1*, and *NANOG* in 42 invasive breast cancer cases were analysed for correlations with their clinicopathological features [the age (≤ 50 years old and > 50 years old); tumour size (≤ 2.0 cm and > 2 cm); tumour grade/differentiation; axillary lymph node metastasis; and the status of ER, PR, and HER2 (ERBB2)] by the unpaired *t*-test at the 95% CI (Table 4-6). There were only 14 cases having information on HER2 expression. The CNVs of *DPPA3* in the IC were not significantly correlated with all clinicopathological features ($p > 0.050$). For *EDR1* and *NANOG*, the IC had no significant relationship between the CNVs and the following clinicopathological features: patient's age, tumour size, axillary lymph node metastasis, and the status of ER ($p > 0.050$). There was no significant correlation between the CNVs of *EDR1* and the status of PR ($p = 0.5643$). The ICs that lacked PR [N = 8/42 (19.05%)] had significant homozygous deletion ($X_{CN} = 0$) of *NANOG* compared to cancers positive for PR [N = 34/42 (80.95%)] showing LOH ($X_{CN} = 1$) ($p = 0.0143$). Based on 14 breast cancers knowing HER2 status, HER2-negative cancers showed LOH of *EDR1* ($X_{CN} = 1$) but HER2-positive cases had normal diploid ($X_{CN} = 2$) of this gene ($p = 0.0121$). The status of HER2 was not significantly related to the CNVs of *NANOG* ($p = 0.4563$).

Table 4-6 The correlation between the ΔCt values for *DPPA3* (*STELLA*), *EDR1* (*PHC1*), and *NANOG* of invasive breast carcinoma tissues and clinicopathological features. The unpaired *t*-test at the 95% CI was applied for the analysis of the correlation between the ΔCt values for *DPPA3*, *EDR1*, and *NANOG* of 42 IC samples and their clinicopathological features.

Clinicopathological features	ΔCt value								
	DPPA3			*EDR1*			*NANOG*		
	X	SD	*p*-value[#]	X	SD	*p*-value[#]	X	SD	*p*-value[#]
Age (Years old)									
≤ 50 [N = 25/42 (59.52%)]	-0.400	1.095	0.1578[ns]	1.206	1.838	0.7962[ns]	2.908	1.620	0.4415[ns]
> 50 [N = 17/42 (40.48%)]	-0.852	0.837		1.082	0.833		2.483	1.907	
Tumour size (cm)									
≤ 2.0 [N = 14/42 (33.33%)]	-0.677	1.072	0.6755[ns]	1.368	2.179	0.5232[ns]	2.955	2.246	0.5686[ns]
> 2.0 [N = 28/42 (66.67%)]	-0.536	0.998		1.049	1.049		2.626	1.445	
Tumour grade/differentiation									
Low (I & II)/Well & Moderate [N = 16/42 (38.10%)]	-0.606	1.013	0.9104[ns]	0.940	1.392	0.4731[ns]	2.373	1.916	0.2915[ns]
High (III)/Poor [N = 26/42 (61.90%)]	-0.569	1.032		1.288	1.576		2.960	1.606	
Axillary lymph node									
No metastasis [N = 23/42 (54.76%)]	-0.541	1.056	0.7709[ns]	1.398	1.867	0.2540[ns]	3.055	2.032	0.1924[ns]
Metastasis [N = 19/42 (45.24%)]	-0.634	0.983		0.862	0.841		2.350	1.222	
ER									
Negative [N = 8/42 (19.05%)]	-0.334	1.325	0.4459[ns]	1.563	0.864	0.4006[ns]	3.740	1.544	0.0677[ns]
Positive [N = 34/42 (80.95%)]	-0.641	0.939		1.060	1.610		2.500	1.708	
PR									
Negative [N = 8/42 (19.05%)]	-0.226	1.214	0.2736[ns]	1.435	0.901	0.5643[ns]	4.060[§]	1.173	**0.0143***
Positive [N = 34/42 (80.95%)]	-0.667	0.961		1.090	1.613		2.424[¶]	1.706	
HER2 (ERBB2)									
Negative (N = 9/14)	-0.366	0.940	0.2678[ns]	1.037[£]	0.514	**0.0121***	2.898	1.487	0.4563[ns]
Positive (N = 5/14)	-0.932	0.724		0.180[¥]	0.535		3.643	2.147	
No information [N = 28/42 (66.67%)]	-0.590	1.086		1.368	1.753		2.522	1.734	

Note: # Unpaired *t*-test at the 95% CI; § Mean CN_{NANOG} of 0 for PR-negative tumours; ¶ Mean CN_{NANOG} of 1 for PR-positive tumours; £ Mean CN_{EDR1} of 1 for HER2-negative tumours; and ¥ Mean CN_{EDR1} of 2 for HER2-positive tumours;

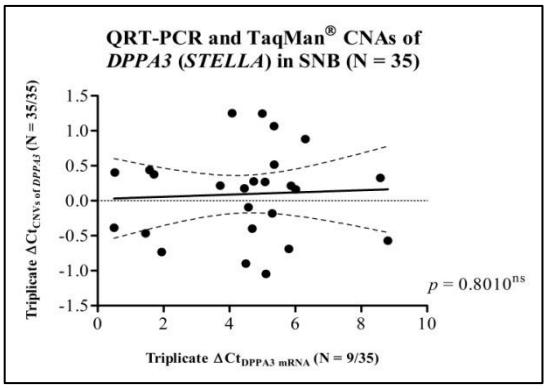
4.3.8. *Summary of findings for CNVs of DPPA3 (STELLA), EDR1 (PHC1), and NANOG in breast tissue*

The NB were all normal diploid (CN = 2) for *DPPA3*, *EDR1*, and *NANOG*. The SNB showed deletion (CN = 0 – 1) (~14 – 20%), normal diploid (~70 – 80%), and single copy gain (CN = 3) (~10%) of these genes. For *DPPA3*, the IC had variations of the CN values: deletion (~12%), normal diploid (~21%), single copy gain (~14%), two-copy gain (CN = 4) (~31%), and amplification (CN ≥ 5) (~21%). Approximately 80% of breast carcinoma cases revealed deletion of *EDR1* and *NANOG*. For *EDR1*, breast cancer also had normal diploid (~12%), gain (~7%), and amplification (~2%). A minority of breast cancers (~5%) showed normal diploid of *NANOG*. No *NANOG* gain/amplification was observed in the IC. The CNVs of *DPPA3* in the IC were not significantly correlated with clinicopathological features [patient's age (≤ 50 years old and > 50 years old); tumour size (≤ 2.0 cm and > 2 cm); tumour grade/differentiation; axillary lymph node metastasis; and the status of ER, PR, and HER2]. Although loss (deletion) of *NANOG* was present in both PR-negative and PR-positive breast carcinomas, but homozygous deletion (CN = 0) of this gene was significantly observed in PR-negative cases compared to PR-positive tumours showing LOH. A small number of breast cancers lacked HER2 also showed LOH of EDR1, whereas HER2-positive cancers had normal diploid.

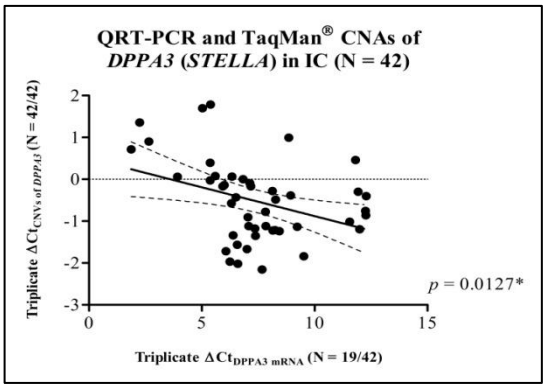
4.3.9. *The correlation between expression of mRNA and immunoexpression and CNVs of DPPA3 (STELLA), EDR1 (PHC1), and NANOG in breast tissue*

Expression of DPPA3, EDR1, and NANOG mRNAs was correlated with the CNVs in 35 SNB and 42 IC tissues using Pearson correlation coefficient at the 95% CI (Figure 4-4). The NB tissue was not analysed as these were not studied for mRNA expression. For SNB, expression of DPPA3, EDR1, and NANOG mRNAs did not significantly correlate with copy number values ($p > 0.050$). The IC showed a significant inverse relationship between expression of DPPA3 mRNA and the CN value ($p = 0.0127$), but there was no significant correlation between expression of EDR1 and NANOG mRNAs and their CN values ($p > 0.050$).

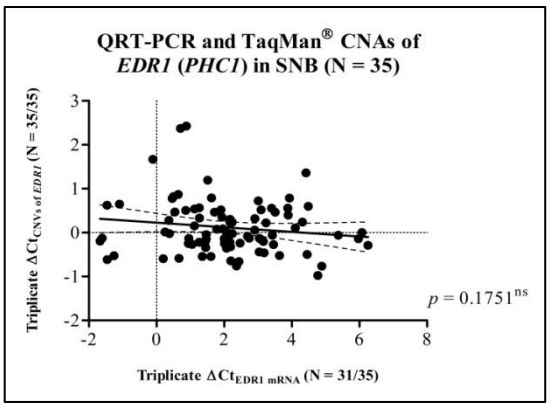
Immunoreactivity for DPPA3, EDR1, and NANOG proteins was correlated with the CNVs in 39, 38, and 39 IC cases, respectively (Figure 4-5). However, the Pearson correlation coefficient at the 95% CI showed that the IC had no significant correlation between immunoexpression of these stem cell associated genes and their CNVs ($p > 0.050$).



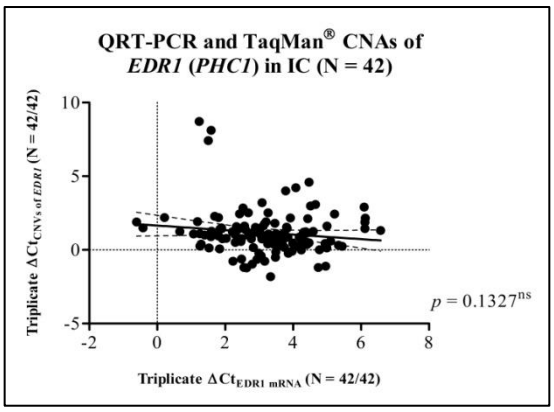
(A)



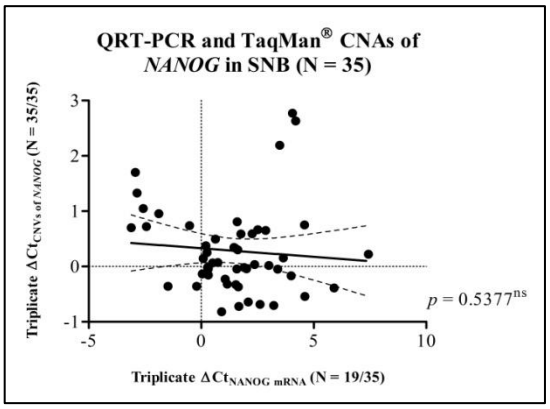
(B)



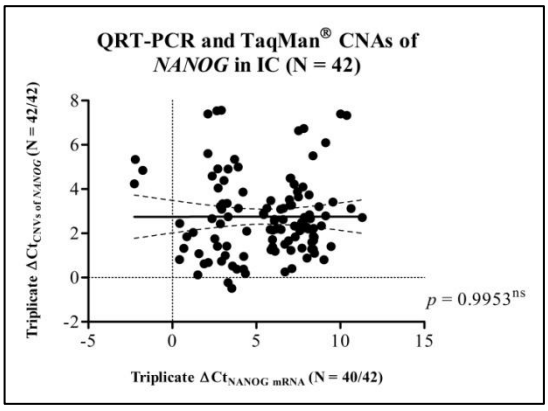
(C)



(D)

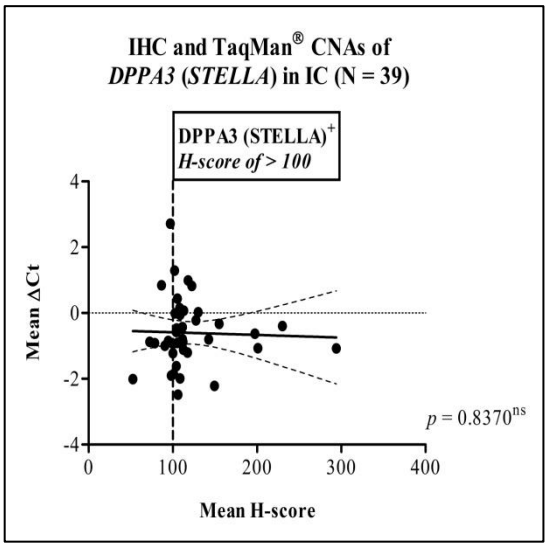


(E)

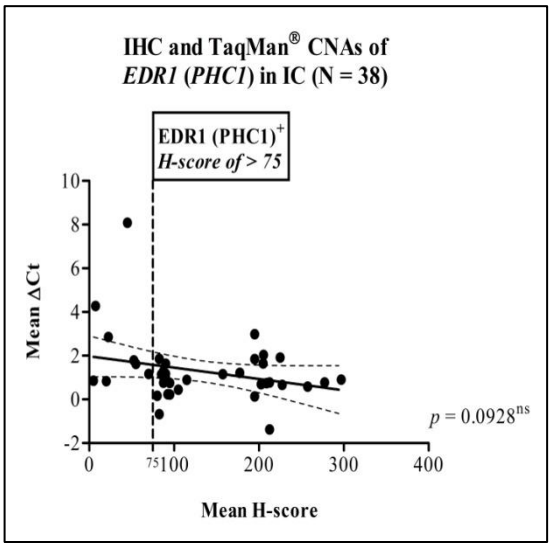


(F)

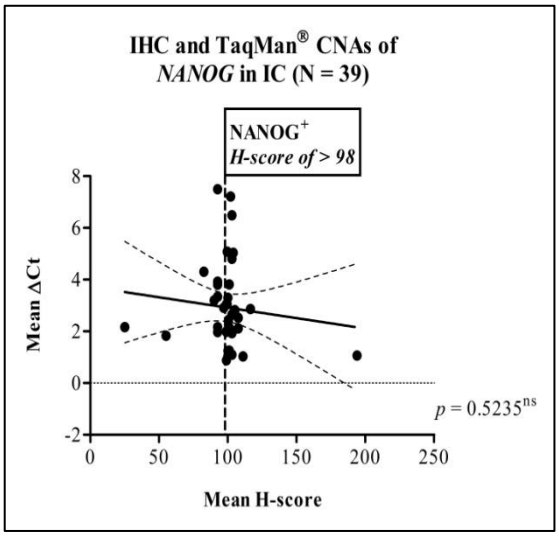
Figure 4-4 The correlation between expression of DPPA3 (STELLA), EDR1 (PHC1) and NANOG mRNA and their CNVs in surrounding normal breast tissue (SNB) and invasive breast cancer tissue (IC). The Pearson correlation coefficient (r) at the 95% CI was used for analysis of the correlation between ΔCt values for DPPA3, EDR1, and NANOG mRNAs and their CNVs in 35 SNB (Left) and 52 IC (Right) samples. (A) and (B) *DPPA3*; (C) and (D) *EDR1*; (E) and (F) *NANOG*



(A)



(B)

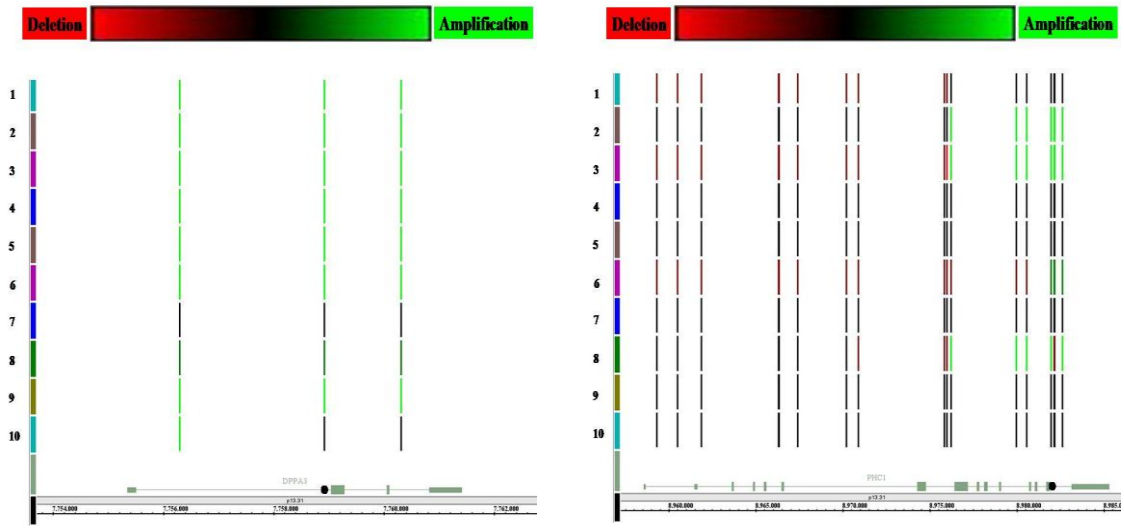


(C)

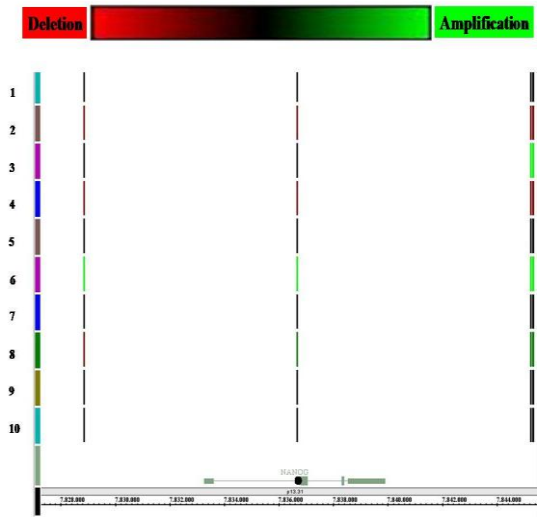
Figure 4-5 The correlation between immunoexpression of DPPA3 (STELLA), EDR1 (PHC1) and NANOG proteins and their CNVs in the IC. The Pearson correlation coefficient (r) at the 95% CI was used for analysis of the correlation between immunostaining H-scores of DPPA3, EDR1, and NANOG proteins and their CNVs in 39, 38, and 39 IC samples. (A) *DPPA3*; (B) *EDR1*; and (C) *NANOG*

4.3.10. *CNVs of DPPA3 (STELLA), EDR1 (PHC1), and NANOG in invasive breast carcinoma according to single nucleotide polymorphisms (SNPs) array*

Ten of the 42 invasive breast carcinoma samples were also profiled using Affymetrix® Genome-Wide Human SNP Array 6.0. This enables interrogation of CNVs throughout the genome. CNVs in the gene interval of *DPPA3*, *EDR1*, and *NANOG* were evaluated in these 10 tumours (Figure 4-6 and Table 4-7). For *DPPA3*, there were 3 CN probes across the gene interval. TaqMan® CNAs yielded 4 and 6 breast cancer cases having lower and higher CN values than by SNP Array, respectively. Two cancers showed CN gain (CN = 3 – 4) of *DPPA3* with both methods. Notably, 4 out of 6 breast carcinoma samples showed normal diploid (CN = 2) of *DPPA3* from the SNP Array, but they showed two-copy gain (CN = 4) in one case and amplification (CN ≥ 5) in 3 cases from TaqMan® CNAs. SNP Array had a high density of CN probes across the *EDR1* interval. Nine of 10 cases had lower CN values from TaqMan® CNAs compared to the result of the SNP Array, whereas the other case was normal diploid with both methods. Of note, 2 cases showed homozygous deletion (CN = 0) and LOH (CN = 1) from TaqMan® CNAs of *EDR1*, but SNP Array yielded normal diploid and single copy gain (CN = 3), respectively. All 10 breast carcinoma cases showed deletion (CN = 0 – 1) of *NANOG* from TaqMan® CNAs, but the result of SNP Array based on a single CN marker showed that normal diploid and single copy gain were observed in 9 and one case (s), respectively.



(A) (B)



(C)

Figure 4-6 The CNVs of *DPPA3* (*STELLA*), *EDR1* (*PHC1*), and *NANOG* in 10 breast carcinomas detected by Affymetrix® Genome-Wide Human SNP Array 6.0. Ten breast carcinoma samples were evaluated their CNVs of 3 stem cell-associated genes (*DPPA3*, *EDR1*, and *NANOG*) using Affymetrix® Genome-Wide Human SNP Array 6.0. The coloured vertical bars represent each CN marker and their CN values correspond with the top colour scale, including red for deletion and green for amplification. The exons and introns of each gene and the scale bar of gene location are shown at the bottom of the diagram. The location of TaqMan® CNA for each gene is demonstrated by black dot on gene alignment map. (A) *DPPA3*, (B) *EDR1*, and (C) *NANOG*

Table 4-7 The results of Affymetrix® Genome-Wide Human SNP Array 6.0 corresponding to TaqMan® CNAs for *DPPA3* (*STELLA*), *EDR1* (*PHC1*), and *NANOG* in 10 invasive breast carcinomas.

Sample	CN§					
	DPPA3		*EDR1*		*NANOG*	
	TaqMan® CNA	SNP 6.0 Arrays	TaqMan® CNA	SNP 6.0 Arrays	TaqMan® CNA	SNP 6.0 Arrays
1	3	4	1	2	0	3
2	5	2	1	3	1	2
3	4	3	1	2	1	2
4	1	2	1	2	0	2
5	2	3	0	2	1	2
6	4	2	1	2	0	2
7	6	2	1	2	0	2
8	1	2	1	2	1	2
9	8	2	1	2	0	2
10	3	2	2	2	0	2

Note:

§ CN of 0 is homozygous deletion; CN of 1 is heterozygous deletion [loss of heterozygosity (LOH)]; CN of 2 is normal diploid; CN of 3 is single copy gain; CN of 4 is two-copy gain; and CN of ≥ 5 is amplification.

The light grey shading represents the similar/same CN values derived from both TaqMan® CNAs and SNP 6.0 Arrays.

4.3.11. *CNVs of chromosome 12 in invasive breast carcinoma*

Ten breast cancer samples in Section 4.3.10 were also evaluated for CNVs of chromosome 12 using Affymetrix® Genome-Wide Human SNP Array 6.0 (Figure 4-7). Copy number gain was mostly observed in the following regions: p12.1, p12.3, p13.31, q12, q21.31, q21.33, and 23.1. Of note, loss of CN was present in 2 cancer cases: case 1 showed loss of q12.1, q14.1, q21.1, q21.31, and 23.1; and case 7 showed loss of q11.23, q14.1, q22, and q23.1. Additionally, deletion of q12 was identified in 4 cases.

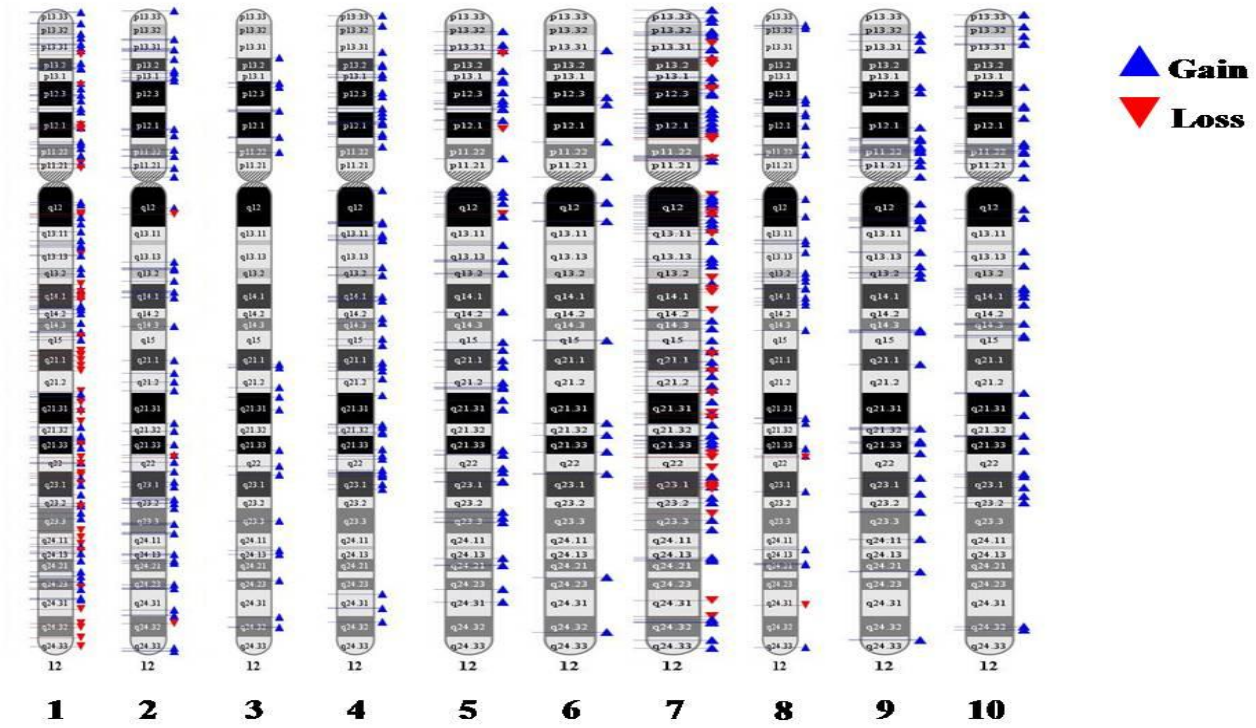


Figure 4-7 The CNVs of chromosome 12 in breast carcinoma. The CNVs of chromosome 12 were analysed in 10 breast carcinoma samples using Affymetrix® Genome-Wide Human SNP Array 6.0 based on 50 SNP and/or CN markers per segment and a minimum segment size of 50,000 base pairs (bp). Gain (a signal threshold of ≥ 3.0) and loss (a signal threshold of ≤ 1.0) of CN on chromosome arms are represented by blue and red triangles, respectively.

4.3.12. ***CNVs of BRCA1, BRCA2, Cytokeratin 5 (KRT5), Cytokeratin 6 (KRT6), EGFR, ERBB2 (HER2), MKI67, and TP53 in invasive breast carcinoma according to single nucleotide polymorphisms (SNPs) array***

Ten invasive breast carcinomas in Section 4.3.10 were evaluated for CNVs of 8 breast cancer-associated genes using Affymetrix® Genome-Wide Human SNP Array 6.0 (Figure 4-8 and Table 4-8). These genes were comprised of (I) 3 tumour suppressor genes implicated in hereditary breast cancer, including *BRCA1*, *BRCA2*, and *TP53* (for p53 protein); (II) 3 genes corresponding to markers of basal-like breast cancer, including *EGFR* and *KRT5* and *KRT6* [for cytokeratin (CK) 5/6 protein]; (III) *MKI67* gene correlated with Ki67 protein of cellular proliferation marker; and (IV) *ERBB2* (*HER2*) oncogene. The CN values for each gene interval were derived from the smoothed signal threshold of all markers. The predicted CN values were finally calculated from the mean of these CN values.

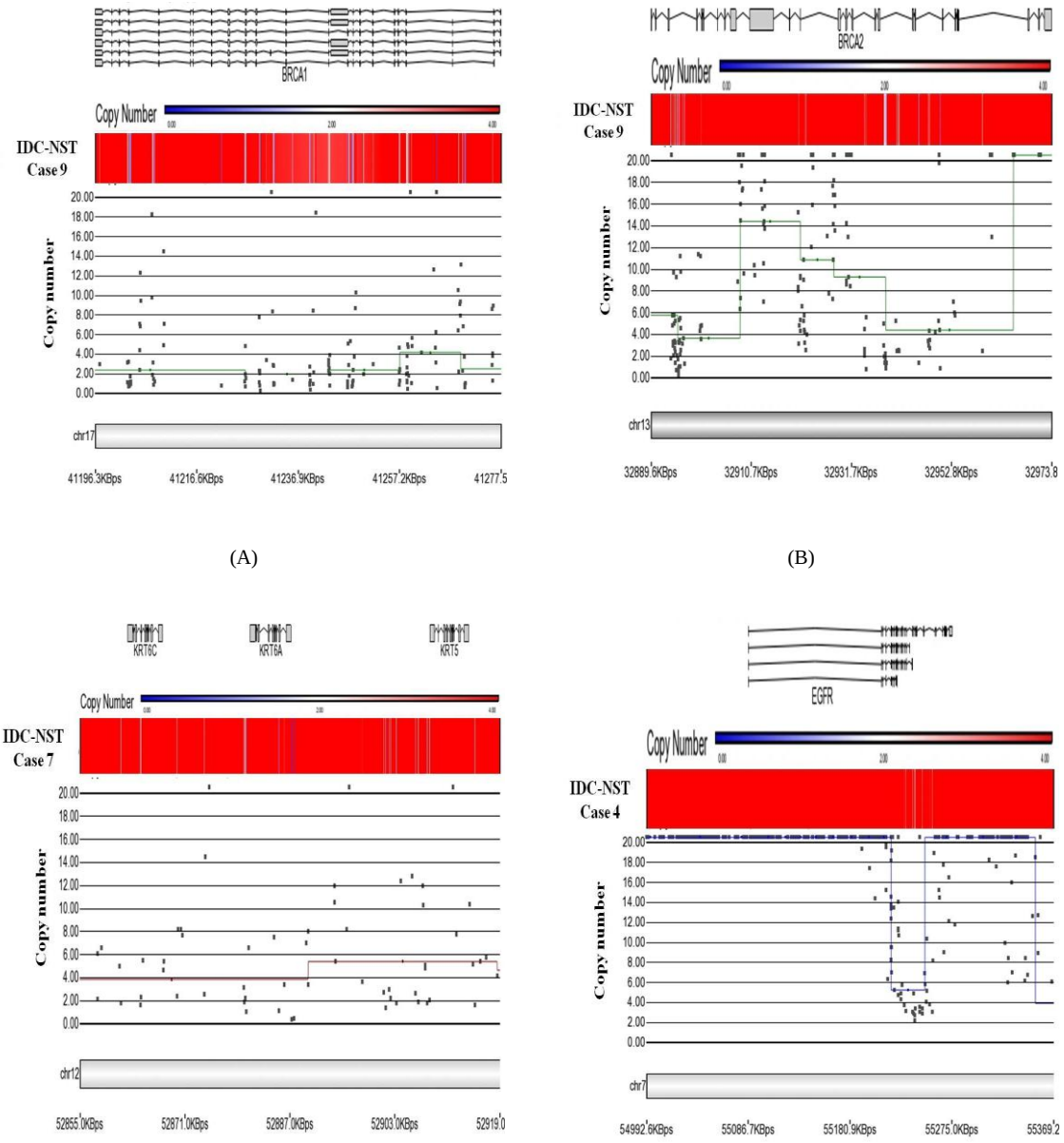
For predicted CN of *BRCA1*, 3 cases showed normal diploid (CN = 2) and 7 cases showed single copy gain (CN = 3). For *BRCA2*, normal diploid was observed in 7 cases; single copy gain was present in one case; and 2 cases had high-level amplification (CN ≥ 8) of *BRCA2*. For *TP53*, 10 breast cancers showed 6 cases of normal diploid, 3 cases of single copy gain, and one case of two-copy gain (CN = 4). Of note, all 10 breast cancers had normal diploid for *MKI67*.

The CNVs of *KRT5*, *KRT6*, and *EGFR* showed (I) heterozygous deletion [loss of heterozygosity (LOH)] (CN = 1) of *KRT5* and *KRT6* but normal diploid of EGFR in 4 cases; (II) LOH of *KRT5* but normal diploid of *KRT6* and *EGFR* in 3 cases; (III) normal diploid of *KRT5*, *KRT6*, and *EGFR* in one case; (IV) one case had LOH of *KRT5*, normal diploid of *KRT6*, and high-level amplification of *EGFR*; and (V) a single case

had *KRT5* amplification (CN = 5), single copy gain of *KRT6*, and normal diploid of *EGFR*.

According to clinicopathological information, 2 breast carcinoma samples had a lack of HER2 and 8 cases had no information on the status of HER2. The 2 cancer cases lacking HER2 by immunohistochemistry showed single copy gain and amplification of *HER2* gene, respectively. The other 8 cancers showed the CNVs of *HER2* as: normal diploid in 2 cases; single copy gain in 3 cases; and high-level amplification with maximum CN of > 20 in 3 cases.

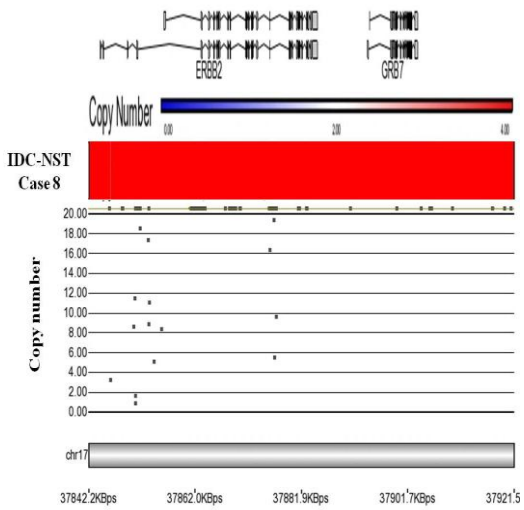
Three out of 6 cancers having normal diploid for *TP53* revealed (I) high-level amplification of *EGFR* in one case; (II) amplification of *KRT5* and *HER2* in one case; and (III) high-level amplification of *HER2* in one case. Two out of 3 cancers having single copy gain for *TP53* showed high-level amplification of *BRCA2* and *HER2*, respectively. Of note, high-level amplification of both *BRCA2* and *HER2* genes was observed in a case of two-copy gain of *TP53*.



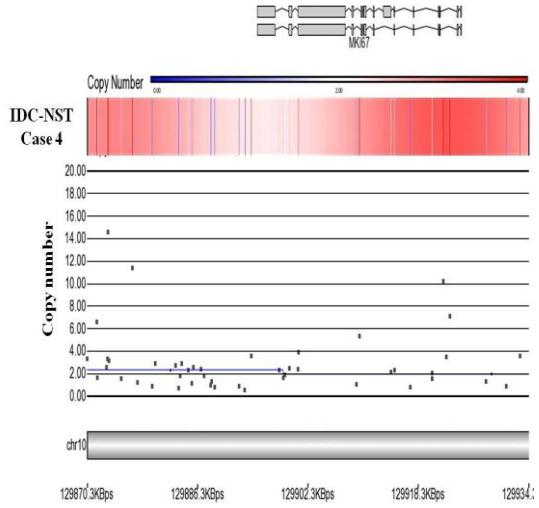
(A) (B)

(C) (D)

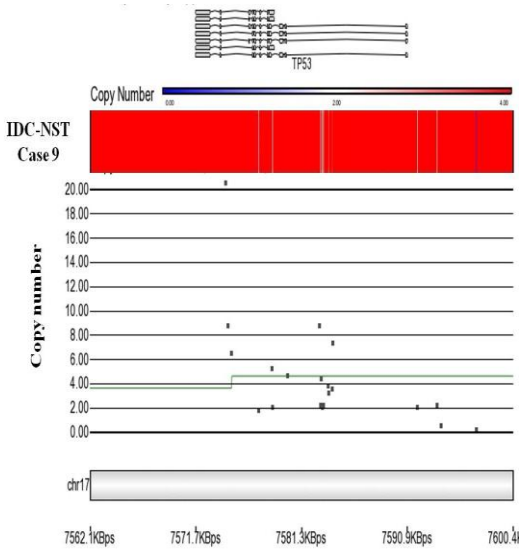
Figure 4-8 The examples of CNVs of breast cancer-associated genes based on Affymetrix® Genome-Wide Human SNP Array 6.0. The CNVs of 8 breast cancer-associated genes[*BRCA1*, *BRCA2*, *Cytokeratin 5* (*KRT5*), *Cytokeratin 6* (*KRT6*), *EGFR*, *HER2* (*ERBB2*), *MKI67*, and *TTP53*] were analysed in invasive breast carcinoma samples using Affymetrix® Genome-Wide Human SNP Array 6.0. The exons and introns of each gene and the scale bar of gene location are shown at the top and the bottom of the diagram, respectively. The coloured vertical bars and the black dots represent each CN marker and their CN values correspond with the overlying colour scale, including blue for CN of 0.00, white for CN of 2.00, and red for CN of ≥ 4.00. The smoothed signal correlated with CN was performed using Partek Genomics Suite 6.5, build 6.10.1129 (Partek). The figures illustrate the examples of CNVs of 8 breast cancer-associated genes in some cases. (A) *BRCA1* in Case 9; (B) *BRCA2* in Case 9; (C) *KRT5* and *KRT6* in Case 7; (D) *EGFR* in Case 4.



(E)



(F)



(G)

Figure 4-8 (Continued) The examples of CNVs of breast cancer-associated genes based on Affymetrix® Genome-Wide Human SNP Array 6.0. The CNVs of 8 breast cancer-associated genes[*BRCA1*, *BRCA2*, *Cytokeratin 5* (*KRT5*), *Cytokeratin 6* (*KRT6*), *EGFR*, *HER2* (*ERBB2*), *MKI67*, and *TTP53*] were analysed in invasive breast carcinoma samples using Affymetrix® Genome-Wide Human SNP Array 6.0. The exons and introns of each gene and the scale bar of gene location are shown at the top and the bottom of the diagram, respectively. The coloured vertical bars and the black dots represent each CN marker and their CN values correspond with the overlying colour scale, including blue for CN of 0.00, white for CN of 2.00, and red for CN of ≥ 4.00. The smoothed signal correlated with CN was performed using Partek Genomics Suite 6.5, build 6.10.1129 (Partek). The figures illustrate the examples of CNVs of 8 breast cancer-associated genes in some cases. (E) *ERBB2* in Case 8; (F) *MKI67* in Case 4; and (G) *TTP53* in Case 9.

Table 4-8 The copy number variations (CNVs) of breast cancer-associated genes based on Affymetrix® Genome-Wide Human SNP Array 6.0 in 10 invasive breast carcinomas. The CNVs of 8 breast cancer-associated genes[*BRCA1*, *BRCA2*, *Cytokeratin 5* (*KRT5*), *Cytokeratin 6* (*KRT6*), *EGFR*, *HER2* (*ERBB2*), *MKI67*, and *TTP53*] were analysed in 10 invasive breast carcinoma samples (Section 4.3.10) using Affymetrix® Genome-Wide Human SNP Array 6.0 with Partek Genomics Suite 6.5, build 6.10.1129 (Partek). The presumptive molecular subtypes of these cases were categorised according to the status of ER, PR, and HER2 and the CN values of 8 breast cancer-associated genes.

Case	ER	PR	HER2	CN from SNP 6.0 Arrays (Predicted CN§)								Presumptive molecular subtype
				BRCA1	***BRCA2***	***KRT5***	***KRT6***	***EGFR***	***HER2***	***MKI67***	***TP53***	
1	+ (High)	+ (High)	No information	1 to 3 (2)	3 to 19 (11)	1	1	1 to 2 (2)	1 to 3 (2)	2	2 to 4 (3)	Luminal A
2	+ (High)	+ (High)	No information	1 to 4 (3)	1 to 3 (2)	1	1	1 to 3 (2)	1 to 3 (2)	1 to 2 (2)	2 to 3 (3)	Luminal A
3	+ (High)	+ (High)	No information	2 to 3 (3)	1 to 3 (2)	1	2	1 to 3 (2)	2 to 3 (3)	1 to 2 (2)	2	Luminal B
4	-	-	-	2 to 4 (3)	1 to 2 (2)	1	2	5 to > 20	2 to 4 (3)	2	2	Basal-like
5	+ (High)	+ (High)	No information	2 to 3 (3)	1 to 5 (3)	1	1	1 to 3 (2)	3 to > 20	1 to 2 (2)	2 to 3 (3)	HER2-enriched
6	+ (High)	+ (High)	No information	1 to 4 (3)	1 to 2 (2)	2	2	1 to 3 (2)	1 to 4 (3)	1 to 2 (2)	2	Luminal B
7	+ (High)	+ (High)	-	1 to 3 (2)	1 to 2 (2)	5	3	1 to 3 (2)	2 to 8 (5)	1 to 2 (2)	2	Luminal B
8	+ (High)	+ (High)	No information	2 to 4 (3)	1 to 2 (2)	1	1	1 to 3 (2)	> 20	1 to 2 (2)	2	HER2-enriched
9	+ (High)	+ (High)	No information	2 to 4 (3)	3 to > 20	1	2	1 to 2 (2)	17 to > 20	1 to 2 (2)	3 to 4 (4)	HER2-enriched
10	-	-	No information	1 to 3 (2)	1 to 3 (2)	1	2	1 to 3 (2)	1 to 4 (3)	1 to 2 (2)	2	HER2-enriched

Note: § Predicted CN in parenthesis was derived from the mean CN value of each tumours. CN of 0 is homozygous deletion. CN of 1 is heterozygous deletion [loss of heterozygosity (LOH)]. CN of 2 is normal diploid. CN of 3 is single copy gain. CN of 4 is two-copy gain. CN of ≥ 5 is amplification.; - = Negative; and + = Positive

4.3.13. *Putative molecular subtypes of breast cancer and their CNVs*

The 10 invasive breast carcinomas investigated by SNP Array were given putative major molecular subtypes (luminal A, luminal B, HER2-enriched, basal-like, and normal breast-like) as defined previously by gene expression profiling (see Table 1-8 in Chapter 1 Introduction) (Al Tamimi et al. 2010, Correa Geyer, Reis-Filho 2009, Cuzick et al. 2011, Reis-Filho, Pusztai 2011, Weigelt, Baehner & Reis-Filho 2010, Weigelt, Geyer & Reis-Filho 2010) based on (I) the immunohistochemical status of ER, PR, HER2, and Ki67; and (II) the predicted CN values derived from the SNP Array data for *ERBB2* (*HER2*) gene and genes related to markers of basal-like breast cancer (*KRT5*, *KRT6*, and *EGFR*). The predicted CN values for *BRCA1*, *BRCA2*, and *TP53* were considered as minor categorising factors. The *MKI67* gene was not included as a normal diploid CN (CN = 2) for this gene was detected in all 10 breast cancers. Although the luminal B subtype is mainly PR-negative and HER2-negative based on immunoexpression, approximately 47% and 20% of this tumour could be positive PR and HER-2 expression, respectively. In HER2-enriched tumour, the status of ER and PR is mostly negative expression but positive immunoexpression could be observed in approximately 44% of ER and 28% of PR (Table 1-8). The tumours showing positive HER2 immunoexpression should be associated with copy number gain/amplification of *HER2* gene.

According to the above-mentioned criteria, the 10 breast carcinomas were categorised as putative luminal A (2 cases), luminal B (3 cases), HER2-enriched (4 case), and basal-like (1 case) subtypes (Table 4-8). Considering the CNVs of chromosome 12 in Section 4.3.11, all 10 tumours showed common copy number gain in the following regions: p12.1, p12.3, p13.31, q12, q21.31, q21.33, and q23.1 (Figure 4-7).

The 2 putative luminal A cancers showed single copy gain of *BRCA1* and high-level amplification of *BRCA2*, respectively. They also had LOH of *KRT5* and *KRT6* and single copy gain of *TP53*. Both also showed gain of 12p12.1, 12p13.31, 12q14.1, and 12q23.1. On the other hand, one tumour also showed loss in CN markers from 12p12.1, 12q14.1, 12q21.1, 12q21.31, and 12q23.1.

The 3 putative luminal B breast cancers showed (I) LOH of *KRT5* (1 case); (II) single copy gain of *BRCA1* (2 cases), *KRT6* (1 case), and *HER2* (2 cases); (III) amplification of *KRT5* (1 case) and *HER2* (1 case). Of note, amplification of both *KRT5* and *HER2* amplification was observed in the same case. All 3 cancers showed normal diploid copy number (CN) of *BRCA2*, *EGFR*, and *TP53*. The putative luminal B subtype mainly showed gain of 12p12.3, 12p13.31, 12q21.31, 12q21.33, 12q23.1, and 12q24.32. One of them also showed loss in CN markers from both 12p and 12q regions.

The 4 cases of putative HER2-enriched subtype had LOH of *KRT5* (all 4 cases) and *KRT6* (2 cases); single copy gain of *BRCA1* (3 cases), *BRCA2* (1 case), *HER2* (1 case), and *TP53* (1 case); two-copy number of *TP53* (1 case); and high-level amplification of *BRCA2* (1 case) and *HER2* (3 cases). All 4 cases showed normal diploid CN of *EGFR*. This tumour also showed gain of 12p11.22, 12p12.1, 12p12.3, 12p13.32, 12q12, 12q13.13, 12q15, 12q21.33, and 12q23.1.

There was a single case of putative basal-like breast cancer, which showed (I) LOH of *KRT5*; (II) normal diploid of *BRCA2*, *KRT6*, *Ki67*, and *TP53*; (III) single copy gain of *BRCA1*; and (IV) high-level amplification of *EGFR*. Although this cancer lacked immunoexpression of HER2 protein, it showed single copy gain of *HER2* gene suggesting this gene might be mutated. This putative basal-like tumour also had copy

number gain in multiple markers at 12p12.1, 12p12.2, 12p13.1, 12q12, 12q14.1, 12q21.32, 12q22, 12q23.1, and 12q24.31.

4.3.14. *Genome wide CN analysis of 10 invasive breast carcinomas*

The genome profile of the 10 invasive breast carcinomas (Section 4.3.10) was assessed from the Affymetrix® Genome-Wide Human SNP Array 6.0. The results were compared with SNP Array 6.0 data for 55 other breast cancers generated as part of another study ongoing in the group (Shaw et al. 2011). Twenty-three intervals showed high-level gain (amplification) at chromosome 1, 2, 4, 5, 7, 8, 9, 10, 14, 15, 16, and 19 based on a smoothed signal threshold of > 6.0 (Table 4-9). No gene amplification was detected on chromosome 12 at this high CN level.

Both studies, including the 10 breast cancers investigated in this thesis and the 55 breast cancers by Shaw et al. 2011, showed that the majority of breast cancers had amplification of genes on chromosome 1p36.33, 1q21.1, 4q13.2, 7q11.23, 9p11.2, 9q12, 10q11.22, 10q11.23, 14q11.2, 15q25.2, 16p12.3, and 19p13.3. Chromosome 8p23.1 and 14q32.33 regions were frequently amplified in my 10 tumours. On the other hand, Chromosome 5q13.2 and 14q11.1 revealed more amplification in the 55 breast cancers of Shaw et al. 2011 than my 10 breast cancer samples. For amplified intervals on chromosome 9q12, 14q11.1, and 14q11.2, there were no genes recorded in the HUGO Gene Nomenclature Committee (HGNC) (http://www.genenames.org/) and the Reference Sequences (RefSeq) (http://www.ncbi.nlm.nih.gov/RefSeq/) databases. The ratio paired *t*-test at the 95% confidence interval (CI) showed that breast cancers from both studies had no significant difference in frequency of amplification ($p = 0.2027$).

Table 4-9 Gene amplification based on Affymetrix® Genome-Wide Human SNP Array 6.0 in invasive breast carcinoma. Ten invasive breast carcinoma samples were analysed significant high-level gain (amplification) (a smoothed signal threshold of > 6.0) using Affymetrix® Genome-Wide Human SNP Array 6.0. These results were compared with 55 breast cancers from the study of Shaw and colleagues (2011). Gene symbols and names were based on the HUGO Gene Nomenclature Committee (HGNC) (http://www.genenames.org/). Genes with an asterisk were only present in the Reference Sequences (RefSeq) database provided by the National Centre for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/RefSeq/).

Chromosome	Amplified gene	N (%)	
		My tumours (Total N = 10)	Shaw et al, 2006 (Total N = 55)
1p36.21	*C1orf158, LOC440563* 440563*, PRAMEF2 3 4 5 7 8 9 10 12 13 14 15 16 19 20 21 22*	1 (10.00%)	4 (7.27%)
1p36.33	*LOC100132287**	10 (100.00%)	49 (89.09%)
1q21.1	*PPIAL4A, PPIAL4C*	10 (100.00%)	47 (85.45%)
2p11.2	*ANKRD36BP2, EIF2AK3, RPIA*	2 (20.00%)	13 (23.64%)
4q13.2	*TMPRSS11BNL, TMPRSS11B, UGT2B 15 17, YTHDC1*	9 (90.00%)	41 (74.54%)
5p11¥	-	1 (10.00%)	14 (25.45%)
5q13.2	*SERF1A, SERF1B, NAIP, SMA5*, SMN1 2*	4 (40.00%)	42 (76.36%)
7p11.2	*MIR4283-2*	2 (20.00%)	21 (38.18%)
7q11.21¥	-	1 (10.00%)	17 (30.91%)
7q11.23	*CLIP2, GATSL1, GTF2IRD1 2, LOC100093631, NCF1, NCF1C, PMS2P5 6, SPDYE8P, STAG3L2, WBSCR1649*	9 (90.00%)	46 (83.64%)
8p23.1	*DEFB130, DUB3, FAM66A, FAM66D, FAM85A, FAM86B1, LOC392196*, ZNF705D*	5 (50.00%)	16 (29.09%)
9p11.2	*FAM27A, LOC100132167**	10 (100.00%)	49 (89.09%)

¥ The chromosomal regions have no genes listed in both HGNC and RefSeq databases.

Table 4-9 (Continued) Gene amplification based on Affymetrix® Genome-Wide Human SNP Array 6.0 in invasive breast carcinoma. Ten invasive breast carcinoma samples were analysed significant high-level gain (amplification) (a smoothed signal threshold of > 6.0) using Affymetrix® Genome-Wide Human SNP Array 6.0. These results were compared with 55 breast cancers from the study of Shaw and colleagues (2011). Gene symbols and names were based on the HUGO Gene Nomenclature Committee (HGNC) (http://www.genenames.org/). Genes with an asterisk were only present in the Reference Sequences (RefSeq) database provided by the National Centre for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/RefSeq/).

Chromosome	Amplified gene	N (%)	
		My tumours (Total N = 10)	Shaw et al, 2006 (Total N = 55)
9q12¥	-	6 (60.00%)	48 (87.27%)
10q11.22	*BMS1P5, FRMPD2P1, GDF2 10, PTPN20B, RBP3, ZNF488*	10 (100.00%)	43 (78.18%)
10q11.23	*AGAP7 8, C10orf53, CHAT, OGDHL, PARG*	10 (100.00%)	46 (83.64%)
14q11.1¥	-	1 (10.00%)	34 (61.82%)
14q11.2¥	-	9 (90.00%)	46 (83.64%)
14q32.33	*AHNAK2, BRF1, C14orf79, CDCA4, CRIP1 2, GPR132, JAG2, KIAA0284, LOC647310*, MGC23270*, MTA1, NUDT14, PACS2*	5 (50.00%)	16 (29.09%)
15q11.2	*NF1P1, POTEB*	4 (40.00%)	24 (43.64%)
15q25.2	*GOLGA6L4 5 9, LOC440297* 727849* 80154*, RPS17*	9 (90.00%)	41 (74.54%)
16p11.2	*ALDOA, C16orf92, DOC2A, FAM57B, HIRIP3, INO80E, KCTD13, TAOK2, TMEM219*	2 (20.00%)	25 (45.45%)
16p12.3	*ABCC6P1, NOMO2*	8 (80.00%)	41 (74.54%)
19p13.3	*FAM138F, FLJ45445*, OR4F17, WASH5P*	9 (90.00%)	43 (78.18%)

¥ The chromosomal regions have no genes listed in both HGNC and RefSeq databases.

4.4. *DISCUSSION*

Copy number variations (CNVs) are commonly observed in human cancers (Shlien, Malkin 2009) and can be determined by using quantitative (real-time) polymerase chain reaction (QPCR) (D'haene, Vandesompele & Hellemans 2010, Konigshoff et al. 2003, Schaeffeler et al. 2003, Srivastava et al. 2008), fluorescence in situ hybridisation (FISH) (Beser et al. 2007, Dowsett et al. 2009), comparative genomic hybridisation (CGH) (Da Silva et al. 2009, Gao et al. 2009, Lasa et al. 2010), array comparative genomic hybridisation (aCGH) (Fridlyand et al. 2006, Li et al. 2009), single nucleotide polymorphism (SNP) array (Alkan, Coe & Eichler 2011, Carter 2007, Nowak, Hofmann & Koeffler 2009), and next-generation (massively parallel) sequencing (Meyerson, Gabriel & Getz 2010, Pasqualucci et al. 2011, Reis-Filho 2009, Robbins et al. 2011, Schweiger et al. 2011, Wood et al. 2010). This study used TaqMan® copy number assays (CNAs) and Affymetrix® Genome-Wide Human SNP Array 6.0 for evaluating the CNVs of stem cell associated genes on chromosome 12p13, including *DPPA3* (*STELLA*), *EDR1* (*PHC1*), and *NANOG*, in breast tissue.

As discussed in Chapter 3 for real-time TaqMan® quantitative reverse transcriptase polymerase chain reaction (QRT-PCR), the optimal endogenous reference gene was initially determined to ensure accurate results. Two different endogenous reference genes were used for TaqMan® CNAs in this study: (I) *TERT* for *DPPA3* and (II) *RNase P* for *EDR1* and *NANOG*. Cell lines are used as experimental models of cancers. Hence, 7 breast cancer cell lines, including HBL-100, MCF7, MDA-MB-231, MDA-MB-436, MDA-MB-468, T47D, and ZR-75-1, were initially used in this study. CNVs were identified, indicating that genetic alterations [gain and loss of copy number (CN)] could

be detected in breast cancers using these methods as shown in a previous study (Shadeo, Lam 2006).

Although the majority of SNB were normal diploid (CN = 2) for *DPPA3*, *EDR1*, and *NANOG* as were the NB, the SNB revealed loss (deletion) (CN = 0 – 1) and gain (CN ≥ 3) of DNA copy number of these genes, that were shared with the nearby cancer (Torres et al. 2007). For invasive breast carcinoma, this study showed that the majority of cancers had gain (CN = 3 – 4) and high-level gain (amplification) (CN ≥ 5) of *DPPA3* and deletion of both *EDR1* and *NANOG*. However, deletion of *DPPA3* and gain/amplification of *EDR1* were also observed in breast cancer. These stem cell-associated genes are located on chromosome 12p13 and the CNVs of this region have been reported previously as loss (van Beers et al. 2005), gain (van Beers et al. 2005, Bae et al. 2010, Han et al. 2008, Horlings et al. 2010, Joosse et al. 2009, Thompson et al. 2011, Turner et al. 2010), and amplification (Yao et al. 2006, Gao et al. 2009, Hawthorn et al. 2010, Jonsson et al. 2010, Letessier et al. 2006) in breast carcinomas. Hence, a complex picture emerges of genomic instability in these genes.

When the CNVs of chromosome 12p13 region were related to clinicopathological features, there was no relationship with the patient's age (≤ 50 years old and > 50 years old); tumour size (≤ 2.0 cm and > 2 cm); and axillary lymph node metastasis (Letessier et al. 2006). In addition, this study showed that tumour grade/differentiation and the status of ER did not correlate with the CN changes of this region. Previous studies have reported that gain/amplification of chromosome 12p13 in breast cancer was commonly observed in grade 3 (high grade) (Horlings et al. 2010, Letessier et al. 2006) and ER-negative cancers (Letessier et al. 2006). Of note, this study found the correlation between CNVs of *EDR1* and *NANOG* and the status of PR and HER2. Loss (deletion) of *NANOG* was observed in both PR-negative and PR-positive breast carcinomas, but

there was significant homozygous deletion in PR-negative tumours. In the small number of cases that could be evaluated, those positive for HER2 showed normal diploid for *EDR1* but HER2-negative cases showed LOH. On the other hand, amplification of chromosome 12p13 has previously been reported for breast cancer lacking PR and HER2 (Letessier et al. 2006, Natrajan et al. 2009). Furthermore, gain of this chromosomal region has been identified in both HER2-negative and HER2-positive breast cancers (Han et al. 2008, Staaf et al. 2010).

Interestingly, the CN values of *DPPA3* in the breast cancers studied were inversely related to mRNA expression levels. This may be because stimulation of mRNA transcription is governed by factors and/or a pathway other than the level of CN change of gene (Mrhalova et al. 2003, Williamson et al. 2005). For *EDR1* and *NANOG*, there was also no correlation between their CNVs and mRNA expression in breast cancer. These results may be associated with (I) alteration in one or more regulatory mechanism (s) at the level of gene transcription (Mrhalova et al. 2003, Bofin et al. 2004, Kang et al. 2002, Pauletti et al. 2000) and/or (II) tissue differences between the formalin-fixed paraffin-embedded tissue examined for DNA copy number changes and the fresh frozen tissue examined for mRNA expression. Additionally, the CNVs of *DPPA3*, *EDR1*, and *NANOG* did not correlate with their immunostaining H-scores for protein expression. This finding could be explained by abnormalities in a control of translational process into protein or accumulation of protein (Mrhalova et al. 2003).

The CNVs derived from TaqMan® CNAs were also different to the results of the SNP Array because TaqMan® CNAs detected the CN values for a single DNA segment (Alkan, Coe & Eichler 2011, Birrer et al. 2007) but the SNP Array gave the averaged CN signal from the whole gene interval (Nowak, Hofmann & Koeffler 2009). Furthermore, the SNP Array contains one or more CN markers for each gene interval

and the CN signals obtained from each marker were able to be loss (deletion), gain, or amplification. The TaqMan® probes correspond with any markers of the gene interval of interest on SNP Array. Hence, these probes yield the CN values for specific position on that DNA segment.

Generally, molecular classification of breast cancer is based on gene expression profiling, particularly the status of ER, PR, HER2, and Ki67 (Al Tamimi et al. 2010, Correa Geyer, Reis-Filho 2009, Cuzick et al. 2011, Reis-Filho, Pusztai 2011, Weigelt, Baehner & Reis-Filho 2010, Weigelt, Geyer & Reis-Filho 2010). However, all 10 breast cancers had complete immunoexpression data for ER and PR but information on expression of HER2 receptor was not known for 8 out of 10 breast cancer samples investigated. Additionally, I did not have information about expression of BRCA1, BRCA2, basal cell-associated markers (CK 5/6 and EGFR), tumour proliferation-associated marker (Ki67), and p53. Usually, the level of protein expression correlates with gene CN changes (Geiger, Cox & Mann 2010), especially HER2 status in breast cancer (Belengeanu et al. 2010, Lambein et al. 2011, Shah, Chen 2010).Therefore, this study had the challenge to characterise molecularly the 10 invasive breast carcinomas using a combination of (I) expression of ER, PR, and HER2 and (II) the CNVs of the following genes derived from SNP Array data: *BRCA1*, *BRCA2*, *KRT5*, *KRT6*, *EGFR*, *HER2*, *MKI67*, and *TP53*. These tumours were stratified as putative luminal A, luminal B, basal-like, and HER2-enriched. Interestingly, the single case of putative basal-like breast cancer had single copy gain of *HER2* gene but lacked expression of HER2 protein by immunohistochemistry. Triple-negative/basal-like breast cancer has been previously reported to show genomic instability, particularly CN gain (Hu et al. 2009). Although basal-like breast cancer showed mostly gain of chromosome 12p13 (Letessier

et al. 2006, Adelaide et al. 2007), chromosome 12p13.31 region was also gained in the other breast cancer subtypes (Han et al. 2008).

Alvarez and colleagues (2005) reported that gain of chromosome 12q11 – q21 was commonly identified in *BRCA2*-associated breast cancer (Alvarez et al. 2005). Of note, 2 breast cancer cases having high-level amplification of *BRCA2* were categorised as luminal A and HER2-enriched subtypes. The luminal A tumour had gain of chromosome 12q14.1 and the HER2-enriched tumour showed gain of chromosome 12q12 and 12q13.13.

In agreement with the larger SNP 6.0 Array study in my group (Shaw et al. 2011), the 10 breast carcinomas investigated in this thesis had more frequent amplification of the following chromosomal regions: 1q21.1, 4q13.2, 7q11.23, 9p11.2, 10q11.22, 10q11.23, 14q11.2, 15q25.2, and 19p13.3. These regions, excluding 14q11.2, also had several amplified genes which could be potential novel breast cancer-related CNV markers. Chromosome 14q11.2 did not have gene databases in the HUGO Gene Nomenclature Committee (HGNC) (http://www.genenames.org/) and the Reference Sequences (RefSeq) (http://www.ncbi.nlm.nih.gov/RefSeq/). Although CN gain according to SNP Array was observed on chromosome 12, there was no significant gene amplification based on a smoothed signal threshold of > 6.0 (Shaw et al. 2011). Referring to TaqMan® CNAs of these 10 breast cancers, approximately 30% of cancers had amplification of *DPPA3* gene located on chromosomal 12p13. Breast cancer showing gain of chromosome 12p13 has been correlated previously with tumour recurrence (Thompson et al. 2011) and poor prognosis (Horlings et al. 2010). Therefore, CNVs of chromosome 12 using SNP Array should be investigated further in a large number of breast carcinomas, particularly in the basal-like subtype.

In summary, genomic analysis by both QPCR and SNP 6.0 array revealed a complicated pattern of CN changes for the 3 genes, with DPPA3 predominantly showing gain and the other two showing loss. In addition, there was no clear correlation between CN state and expression at either the mRNA or protein level for each of the three genes.

CHAPTER 5

GENERAL DISCUSSION AND CONCLUSION

5.1. *DISCUSSION*

The chromosome 12p13 region contains putative stem cell-associated genes: *DPPA3* (*STELLA* or *PGC7*), *EDR1* (*PHC1*, *HPH1*, or *RAE28*), *GDF3*, and *NANOG*. In stem cells, DPPA3 (Elliman, Wu & Kemp 2006) and NANOG (Kim et al. 2008) proteins play a role in the maintenance of pluripotentiality.; The *EDR1* gene is associated with self-renewal activity (Ohta et al. 2002); and *GDF3* has the ability to maintain the undifferentiated state and also function as a regulator of cell differentiation (Levine, Brivanlou 2006). Germ cell tumours have been shown to overexpress these genes (Busch et al. 2009, Ezeh et al. 2005, Hart et al. 2005, Rodriguez et al. 2003), and these genes are also expressed in the other malignancies. The *EDR1* gene plays an important role in both children and adult cancers: loss of function of *EDR1* in B-cell maturation can lead to the development of childhood B-cell precursor acute lymphoblastic leukemia (Tokimasa et al. 2001); and shows different expression levels in various large B-cell lymphoma (Raaphorst et al. 2004); and high expression in prostatic adenocarcinomas either high-grade (Gleason score of $\geq$ 8) or with extra-prostatic extension (van Leenders et al. 2007). Overexpression of *GDF3* has been related to growth of mouse melanoma (Ehira et al. 2010). High expression of the *NANOG* gene is associated with advanced stage and worse survival prognosis in oral squamous cell carcinoma (Chiou et al. 2008).

In breast cancer, there have been only 2 studies published to date examining expression of these putative stem cell-associated genes. The first showed expression of

DPPA3, *GDF3*, and *NANOG* genes in the MCF7 breast cancer cell line and detected NANOG expression was detectable in a single case of invasive (infiltrating) ductal carcinoma stage 3 (T1N2M0) (Ezeh et al. 2005). In the second, low expression of the *EDR1* gene was observed in 2 out of 18 (11.11%) of invasive breast carcinomas (Sanchez-Beato et al. 2006). Therefore, this thesis aimed to more fully investigate the role of these putative stem cell-associated genes on chromosome 12p13, i.e. *DPPA3*, *EDR1*, *GDF3*, and *NANOG* in breast carcinomas.

This thesis found that expression of *DPPA3*, *EDR1*, and *NANOG* was rarely detected in normal breast tissues, but these genes were expressed at various levels in breast carcinomas, with higher expression typically detectable in the surrounding normal breast tissues (SNB) taken at least 4 cm distant from the primary tumour site. No expression of *GDF3* gene was detected in breast cancers at the mRNA levels and so this was not studied at the protein level. The levels of expression of *DPPA3*, *EDR1*, and *NANOG* in breast carcinomas were significantly related to in the following prognostic factors: tumour size, tumour grade/differentiation, and axillary lymph node metastasis. This study showed that lower expression of NANOG in breast cancers correlated with increased tumour size greater than 2.0 cm in contrast to a previous report that showed no correlation between this gene expression and tumour size (Ben-Porath et al. 2008). High-grade (grade III)/poorly differentiated breast carcinomas had lower expression of *DPPA3* (Xiao et al. 2008), *EDR1* (Sanchez-Beato et al. 2006), and *NANOG* (Chiou et al. 2008, Ben-Porath et al. 2008) genes. Breast carcinomas showing higher level of expression of *DPPA3* tended to have axillary lymph node metastasis (Xiao et al. 2008). Based on TaqMan® copy number assays (CNAs), the majority of breast carcinomas had gain of *DPPA3* gene and loss (deletion) of both *EDR1* and *NANOG* genes. The adjacent SNB also shared copy number (CN) changes, including gain (CN $\geq$ 3) and loss

(deletion) (CN = 0 – 1), of *DPPA3*, *EDR1*, and *NANOG* genes with the invasive breast carcinomas (Torres et al. 2007).

Although there was no significant correlation between copy number variations (CNVs), mRNA expression, and protein expression in breast carcinomas, as alterations in the levels of expression of these putative stem cell-associated genes correlated with the above-mentioned clinicopathological features of tumour progression. These could potentially provide novel prognostic biomarkers. Copy number changes and mRNA expression of these genes were observed in the SNB adjacent to the primary breast cancer, these findings might be the result of expansion of a genetically abnormal clone according to the cancer "field effect" (Braakhuis et al. 2003, Chai, Brown 2009, Yan et al. 2006). Wide local surgical excisions are routinely examined histopathologically to ensure that the breast is cancer-free at the resection margins. Therefore, the presence of copy number variations (CNVs) (Li et al. 2009) and changes of mRNA expression (Pau Ni et al. 2010) of *DPPA3*, *EDR1*, and *NANOG* genes in the nearby non-neoplastic breast parenchyma might suggest their use for monitoring the potential for development of local recurrence of breast carcinomas (Braakhuis et al. 2003, Yan et al. 2006).

As mentioned above, normal breast tissue mostly showed undetectable expression of *DPPA3*, *EDR1*, and *NANOG* genes. Thus, if the findings of this thesis are validated in more extensive cohorts, these genes could potentially provide novel pharmacologic targets for breast cancer treatment as for the successful development of *HER2*-specific drug, Trastuzumab (Herceptin®) (Iwata 2007, Kroese, Zimmern & Pinder 2007, McKeage, Lyseng-Williamson 2008, Petak et al. 2010, Shepard et al. 2008). It is possible that therapeutic antibodies targeted against these putative stem cell-associated genes could give clinical benefits by more specific elimination of cancer cells and to a reduction in chemotherapeutic side effects such as cardiac toxicity (Du et al. 2009),

myelosuppression (Shayne et al. 2009), neutropenic enterocolitis (Oehadian, Fadjari 2008), etc.

Referring to the single nucleotide polymorphism (SNP) Array 6.0 study (Shaw et al. 2011), the frequency and extent of amplification obtained in the 10 breast carcinomas investigated in this thesis were in agreement with the results of the 55 other breast cancers investigated by the group (Shaw et al. 2011). These tumours showed gain of chromosome 12 at a signal threshold of 3.0 – 4.0, but there was no significant amplification at a signal threshold of > 6.0. Of note, breast cancer patients had significantly higher frequency of amplification on chromosome 4q13.2, 7q11.23, 9p11.2, 10q11.22, 10q11.23, 15q25.2, and 16p12.3 regions in comparison with healthy female controls. Hence, gene amplification in these intervals might play a role in breast carcinogenesis.

5.2. *CONCLUSION*

This thesis showed that there was a significant difference in expression of 3 putative stem cell-associated genes (*DPPA3*, *EDR1*, and *NANOG*) on chromosome 12p13 between non-neoplastic and malignant breast tissues. Aggressive characteristics of breast carcinomas, including increased tumour size, high grade (grade III)/poorly differentiation, and axillary lymph node metastasis, were correlated with aberrations of these genes by genomic DNA copy number, mRNA transcription, and/or protein translation levels. Therefore, these alterations might be useful for novel diagnostic, prognostic, and/or therapeutic biomarkers of breast cancer.

5.3. *FUTURE DIRECTION*

In the future, it would be worth studying further the expression of the 3 putative stem cell-associated genes, including *DPPA3* (*STELLA*), *EDR1* (*PHC1*), and *NANOG*, in breast carcinoma as follows:

Firstly, the results need validating in a larger series of patients and samples including all stages of breast cancer. For example, since only 3 ductal carcinoma in situ (DCIS) samples were assessed in this study, a larger number of non-invasive breast carcinomas should be analysed for mRNA expression and copy number variations (CNVs) of these genes in early breast carcinomas using TaqMan® quantitative reverse transcriptase polymerase chain reaction (QRT-PCR) and TaqMan® copy number assays (CNAs), respectively.

Secondly, both normal and malignant breast tissues consist of a mixture of mammary epithelial cells, myoepthelial cells, and connective tissue. Breast cancer parenchyma also contains various stages of malignant cells. In this thesis, tumour tissues were manually microdissected under light microscopic examination, and so still had contamination by other cells. Therefore, ideally normal and malignant cells of interest should be isolated from fresh frozen and formalin-fixed, paraffin-embedded breast tissues by laser capture microdissection (LCM) for more accurate determination of their expression of genomic DNA, mRNA, and protein (Aaltonen et al. 2011, Morrogh et al. 2007, Xu et al. 2008, Zanni, Chan 2011).

According to possible transcriptional dysregulation and post-translational modifications (PTMs), expression of mRNA and protein should be additionally investigated by transcriptome analysis such as microarray-based gene expression profiling and proteomic analysis such as two-dimensional gel electrophoresis (2-DE or

2-D electrophoresis), respectively. Microarray-based gene expression profiling is used for the evaluation of gene transcription at mRNA expression level via cDNA (Ma et al. 2003, Rhodes, Chinnaiyan 2005). The 2-D electrophoresis should be used for separation of various modified proteins (Hudelist et al. 2006, Mann, Jensen 2003, Nimeus et al. 2007). Both "omic" approaches should be subsequently related to the genomic profiles generated by Affymetrix® Genome-Wide Human SNP Array 6.0. Since CNVs were observed in *DPPA3*, *EDR1*, and *NANOG* genes, mutation screening of these genes could also be performed on both normal and malignant breast tissues by DNA sequencing (Cavallone et al. 2010, Concin et al. 2003, Ginsburg et al. 2011, Korbel et al. 2008).

Importantly, functional study of these putative stem cell-associated genes should be investigated in breast cell lines, for example, using small interfering RNA (siRNA) transfection for gene silencing or knock-down. This procedure leads to inhibition of mRNA and protein expression. The transfected cancer cells can then be investigated in functional assay to determined effects on proliferation (Jang et al. 2009, Yamamoto et al. 2009, Zang, Pento 2008), apoptosis (Croci et al. 2008, Dougherty et al. 2008, Ge et al. 2009), and migration/invasion (Galliher, Schiemann 2006, Han et al. 2008, Karp et al. 2007).

Finally, as for other studies of circulating free DNA (cfDNA) in breast cancer carried out by the research group (Page et al. 2011), expression of *DPPA3*, *EDR1*, and *NANOG* genes could also be investigated in cfDNA using TaqMan® quantitative polymerase chain reaction (QPCR) to determine their potential as circulating biomarkers in breast cancer patients (Weigel, Dowsett 2010).

APPENDICES

Appendix 1 10% Formol saline solution [4% Formaldehyde, 150 mM Sodium chloride (NaCl)]

- 85 g of NaCl [BP3581, Fisher Scientific, the United Kingdom (UK)]
- 1 L of 40% Formaldehyde (F77P-4, Fisher Scientific, UK)
- 9 L of Ultrapure water

Appendix 2 Reagents for Western blotting

- **Gold lysis buffer working solution for extraction of protein**
  - 5 µL of Protease Inhibitor Cocktail (P8340; Sigma-Aldrich®, UK)
  - 500 µL of Gold lysis buffer
- **Preparation for Tris-glycine SDS-polyacrylamide gel**
  - ***5% Stacking gel***
    - 3.4 mL of Sterile ultrapure water
    - 5 µL of N,N,N′,N′-tetramethylethylenediamine (TEMED) (T9281; Sigma-Aldrich®, UK)
    - 50 µL of 10% Ammonium persulfate (APS) (A-3678; Sigma-Aldrich®, UK)
    - 50 µL of 10% SDS
    - 630 µL of 1 M Tris of pH 6.8
    - 830 µL of 30% (w/v) acrylamide : 0.8% (w/v) bis-acrylamide (EC-890; ProtoGel®, UK)

- ***15% Resolving gel***
    - 2.3 mL of Sterile ultrapure water
    - 2.5 mL of 1.5 M Tris of pH 8.8
    - 5 mL of 30% (w/v) acrylamide : 0.8% (w/v) bis-acrylamide
    - 4 µL of TEMED
    - 100 µL of 10% APS
    - 100 µL of 10% SDS

❖ **Buffers for Tris-glycine SDS-polyacrylamide gel electrophoresis (SDS-PAGE)**

- ***1x Western running buffer***
    - 100 mL of 10x Western running buffer stock solution
    - 900 mL of Ultrapure water
- ***10x Western running buffer stock solution***
    - 10 g of SDS
    - 30 g of Tris base
    - 144 g of Glycine
    - 1,000 mL of Ultrapure water
- ***Western loading buffer***
    - 10 µL of 2-Mercaptoethanol
    - 10 µL of Bromo phenol blue (BPB)
    - 100 µL of 5x Leammli buffer

❖ **Buffers for Western blotting transfer**

- ***1x Western transfer buffer***
    - 200 mL of 10x Western transfer buffer stock solution

- 400 mL of 20% Methanol (Methanol: M/3950/17; Fisher Scientific, UK)
- 400 mL of Ultrapure water

- ***10x Western transfer buffer stock solution***
  - 30.3 g of Tris base
  - 144.1 g of Glycine
  - 1,000 mL of Ultrapure water

❖ **1x Tris buffered saline (TBS)/0.1% Tween® 20 (TBS-T) for Western blotting detection**

- 1 mL of Tween® 20 (P1379; Sigma-Aldrich®, UK)
- 50 mL of 20x TBS
- 1,000 mL of Ultrapure water

Appendix 3 The triplicate Ct values for the standard curves of TaqMan® quantitative reverse transcriptase polymerase chain reaction (QRT-PCR). The standard curves of TaqMan® QRT-PCR assay for DPPA3 (STELLA), EDR1 (PHC1), GDF3, and NANOG mRNA were generated by triplicate Ct values for the 7 serial dilutions of the isolated cDNA isolated from NCCIT as a template. The assay was performed on 50 amplification cycles with manual threshold at 0.3 and automatic baseline.

Dilution	Quantity (ng)	Log_{10}	Target gene											
			DPPA3			*EDR1*			*GDF3*			*NANOG*		
			Ct	X	SD	Ct	X	SD	Ct	X	SD	Ct	X	SD
1:5	100,000	5.00	34.486	34.464	0.167	32.604	32.573	0.036	33.813	33.907	0.087	32.256	32.127	0.429
			34.288			32.582			33.921			31.648		
			34.619			32.534			33.986			32.476		
1:25	20,000	4.30	35.393	35.448	0.064	33.912	34.110	0.225	35.627	35.605	0.288	33.695	33.606	0.081
			35.518			34.355			35.881			33.588		
			35.434			34.064			35.306			33.535		
1:125	4,000	3.60	38.704	38.731	0.128	36.912	36.777	0.373	38.967	39.116	0.330	35.814	35.739	0.321
			38.620			36.355			38.887			36.016		
			38.871			37.064			39.494			35.388		
1:625	800	2.90	40.979	40.744	0.353	40.232	40.383	0.207	41.916	41.614	0.324	40.097	39.891	0.422
			40.913			40.297			41.653			39.405		
			40.338			40.619			41.272			40.170		
1:3,125	160	2.20	42.486	42.061	0.368	41.942	41.853	0.194	42.949	42.495	0.434	41.243	41.927	0.671
			41.854			41.986			42.083			41.954		
			41.843			41.631			42.453			42.585		
1:15,625	32	1.51	45.580	45.805	0.196	44.774	44.858	0.076	45.881	45.750	0.282	43.238	43.114	0.107
			45.928			44.921			45.426			43.053		
			45.909			44.879			45.942			43.052		
1:78,125	6.4	0.81	48.375	48.172	0.300	45.148	45.422	0.237	46.871	47.195	0.281	45.178	45.528	0.305
			48.313			45.561			47.351			45.671		
			47.828			45.556			47.364			45.737		

Note: Quantity = Template quantity (ng); Log_{10} = Log_{10} Template quantity; and Ct = Triplicate Ct values

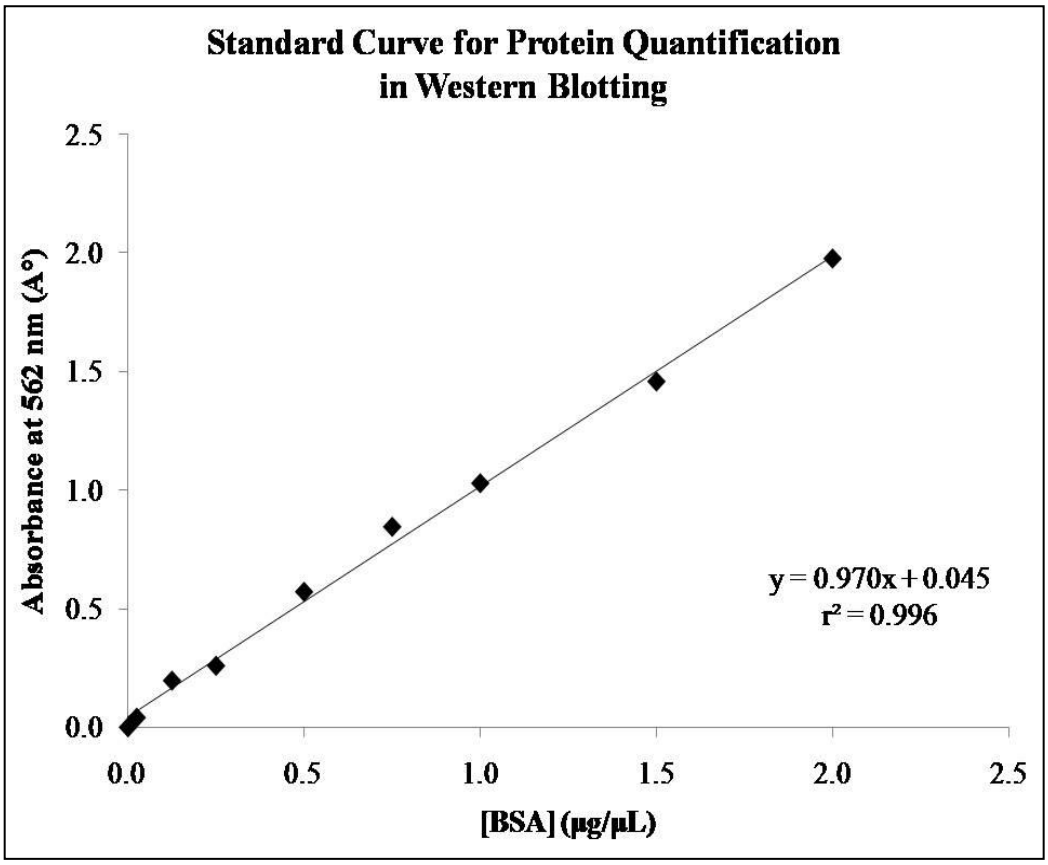
Appendix 4 The UV absorbance value at 562 nm of bovine serum albumin (BSA) standard protein. The 9 serial dilutions of BSA (Albumin Standard Ampoules, 2 mg/mL) from Pierce® BCA Protein Assay Kit was measured their UV absorbance at 562 nm twice for standard curve creation determining the concentrations of protein samples.

Vial	[BSA] (µg/µL)	UV Absorbance at 562 nm			
		1st	2nd	Mean	Corrected*
A	2.000	2.036	2.064	2.050	1.978
B	1.500	1.472	1.590	1.531	1.459
C	1.000	1.063	1.141	1.102	1.030
D	0.750	0.890	0.946	0.918	0.846
E	0.500	0.630	0.659	0.645	0.572
F	0.250	0.330	0.334	0.332	0.260
G	0.125	0.271	0.267	0.269	0.197
H	0.025	0.121	0.105	0.113	0.041
I (Blank)	0.000	0.082	0.063	0.073	0.000

Note:

* Corrected absorbance = Mean absorbance of each vial - Mean absorbance of blank vial

Appendix 5 The standard curve for determination of protein concentration. The corrected absorbance of each BSA standard and its concentration in µg/µL from Appendix 4 was applied for the construction of a standard curve by using Microsoft Office Excel program.



Appendix 6 Protein concentration for Western blotting. The concentration of protein isolated from the cell pellet of 8 cell lines (NCCIT, HBL-100, MCF7, MDA-MB-231, MDA-MB-436, MDA-MB-468, T47D, and ZR-75-1) and 3 different histopathological types of frozen breast carcinoma tissues (DCIS, DCIS & IDC-NST, and IDC-NST) were calculated on the basis of their corrected UV absorbance at 562 nm and the equation from a standard curve in Appendix 3 "y = 0.970x + 0.045", where "y" is the corrected absorbance at 562 nm and "x" is protein amount.

Sample			**UV Absorbance at 562 nm**				**[Protein] (µg/µL)**	
			1st	**2nd**	**Mean**	**Corrected* (y)**	**x**	**Mean**
Cell line	MGCT CL	NCCIT	1.739	1.790	1.765	1.692	1.697	
	BCA CL	HBL-100	1.758	1.816	1.787	1.714	1.721	1.555
		MCF7	2.058	2.128	2.093	2.020	2.036	
		MDA-MB-231	1.214	1.247	1.231	1.158	1.147	
		MDA-MB-436	1.302	1.334	1.318	1.245	1.237	
		MDA-MB-468	1.237	1.261	1.249	1.176	1.166	
		T47D	1.371	1.390	1.381	1.308	1.302	
		ZR-75-1	2.307	2.340	2.324	2.251	2.274	
Frozen tissue	DCIS	RW 2141	0.934	0.936	0.935	0.862	0.842	0.874
	DCIS & IDC-NST	RW 2175	0.923	0.924	0.924	0.851	0.830	
	IDC-NST	RW 2194	1.039	1.040	1.040	0.967	0.950	

Note: * Corrected absorbance = Mean absorbance of each sample - Mean absorbance of blank vial from Appendix 4

Appendix 7 Allred (Quick) scoring guidance on immunohistochemical staining in breast cancer (Collins, Botero & Schnitt 2005, Detre, Saclani Jotti & Dowsett 1995, Harvey et al. 1999, Leake et al. 2000, NHS Cancer Screening Programmes jointly with The Royal College of Pathologists 2005, Qureshi, Pervez 2010)

Nuclear staining		**Score**
Percentage	None	0
	0% - 1%	1
	2% - 10%	2
	11% - 33%	3
	34% - 66%	4
	67% - 100%	5
Intensity	No staining	0
	Weak	1
	Moderate	2
	Strong	3
Total score (0 – 8)	**Percentage score + Intensity score**	

Total score	**Immunoexpression**	**Chance of response to hormone (endocrine) therapy**
0 – 1	Negative	No response
2		20%
3	Positive	
4 – 6		50%
7 – 8		75%

Appendix 8 The triplicate Ct values for the standard curves of TaqMan® copy number assays (CNAs). The standard curves of TaqMan® CNA for *DPPA3*, *EDR1*, and *NANOG* were generated by triplicate Ct values for the 7 serial dilutions of HGDNA as a template. The assay was performed on 60 amplification cycles with manual threshold at 0.2 and automatic baseline.

Dilution	Template quantity (ng)	Log_{10} Template quantity	Target gene								
			DPPA3			*EDR1*			*NANOG*		
			Triplicate Ct	X	SD	Triplicate Ct	X	SD	Triplicate Ct	X	SD
1:5	100,000	5.00	22.513 22.256 22.368	22.379	0.129	23.762 23.810 23.889	23.820	0.064	24.717 24.873 24.938	24.842	0.113
1:25	20,000	4.30	24.429 24.445 24.425	24.433	0.011	26.152 26.298 26.112	26.187	0.098	26.497 26.483 26.468	26.483	0.014
1:125	4,000	3.60	26.478 26.462 26.473	26.471	0.009	29.308 29.488 29.279	29.358	0.113	28.839 28.782 28.867	28.830	0.043
1:625	800	2.90	29.338 29.345 29.360	29.348	0.012	33.202 33.270 33.314	33.262	0.056	31.875 31.865 31.987	31.909	0.067
1:3,125	160	2.20	30.189 30.166 30.191	30.182	0.014	35.386 35.412 35.453	35.417	0.034	33.983 33.845 33.876	33.901	0.072
1:15,625	32	1.51	33.381 33.833 33.371	33.528	0.264	35.907 35.911 35.811	35.876	0.057	36.442 36.470 36.550	36.487	0.056
1:78,125	6	0.81	36.392 36.223 37.156	36.590	0.497	36.348 36.415 36.412	36.391	0.038	37.538 37.632 37.587	37.586	0.047

REFERENCES

Anti-Stella antibody - Primordial Germ Cell Marker (ab19878), [Homepage of abcam], [Online]. Available: http://www.abcam.com/Stella-antibody-Primordial-Germ-Cell-Marker-ab19878.html [2011, August/30th].

Biomarkers in Cancer: An Introductory Guide for Advocates 2010, , Research Advocacy Network, USA.

Breast Cancer Facts & Figures 2011 - 2012 2011, , American Cancer Society, USA.

Breast cancer (female) - Treatment 2010, August 16th, 2010-last update [Homepage of National Health Service (NHS), UK], [Online]. Available: http://www.nhs.uk/Conditions/Cancer-of-the-breast-female/Pages/Treatment.aspx [2012, May 17th, 2012].

Breast cancer - risk factors 2010, 19th January-last update [Homepage of Cancer Research UK], [Online]. Available: http://info.cancerresearchuk.org/cancerstats/types/breast/riskfactors/#source1 [2011, December/15th].

"Breast Tumours (ICD-O C50)" 2008, in *TNM Atlas: Illustrated Guide to the TNM Classification of Malignant Tumours*, eds. C. Wittekind, F.L. Greene, R.V.P. Hutter, M. Klimpfinger & L.H. Sobin, 5th edn, Willey-Liss: John Wiley & Sons, UK, pp. 229-246.

Cell Culture Basics Companion Handbook, Invitrogen - Gibco.

EDR1 / PHC1 antibody (ab54954), [Homepage of abcam], [Online]. Available: http://www.abcam.com/EDR1-PHC1-antibody-ab54954.html [2011, August/30th].

Facts for life: Hormone Therapy 2010, , Susan G. Komen for the Cure, USA.

Facts for life: Treatment Choices - An Overview 2011a, , Susan G. Komen for the Cure, USA.

Facts for life: Targeted Therapies 2011b, , Susan G. Komen for the Cure, USA.

Fundamental Techniques in Cell Culture, The European Collection of Cell Cultures (ECACC) & Sigma-Aldrich, The United Kingdom.

*Gene: PHC1; Breast Cancer*c, [Homepage of The Swedish Human Protein Atlas Project], [Online]. Available: http://www.proteinatlas.org/ENSG00000111752/cancer/breast+cancer [2011, May/17th].

Hormonal Therapy 2012, March 19th, 2012-last update [Homepage of Breastcancer.org], [Online]. Available: http://www.breastcancer.org/treatment/hormonal/ [2012, May 17th].

Loading Controls in Western Blots 2012, May 6th, 2012-last update [Homepage of Labome.com], [Online]. Available: http://www.labome.com/method/Loading-Controls-for-Western-Blots.html [2012, May 24th].

NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines)®: Breast Cancer 2012, , Version 1 edn, National Comprehensive Cancer Network (NCCN), USA.

NCCN Guidelines for Patients™: Breast Cancer 2011, , Version 2 edn, National Comprehensive Cancer Network (NCCN), USA.

"Neoplasia" 2005a, in *Robbins and Cotran Pathologic Basis of Disease*, eds. V. Kumar, A.K. Abbas & N. Fausto, 7th edn, Elsevier Saunders, China, pp. 269-342.

"Mutations" 2005b, in *Robbins and Cotran Pathologic Basis of Disease*, eds. V. Kumar, A.K. Abbas & N. Fausto, 7th edn, Elsevier Saunders, China, pp. 147.

Ovarian ablation and breast cancer 2011, December 1st, 2011-last update [Homepage of Macmillan Cancer Support], [Online]. Available: http://www.macmillan.org.uk/Cancerinformation/Cancertypes/Breast/Treatingbreastcancer/Hormonaltherapies/Ovarianablation.aspx [2012, May 17th].

Pathology Reporting of Breast Disease: A Joint Document Incorporating the Third Edition of the NHS Breast Screening Programme's Guidelines for Pathology Reporting in Breast Cancer Screening and the Second Edition of The Royal College of Pathologists' Minimum Dataset for Breast Cancer Histopathology 2005, , 58th edn, NHS Cancer Screening Programmes jointly with The Royal College of Pathologists, UK.

*Stage Information for Breast Cancer*2011, The 14th of July-last update [Homepage of National Cancer Institute], [Online]. Available: http://www.cancer.gov/cancertopics/pdq/treatment/breast/healthprofessional/page3 [2011, 09/19].

Targeted Therapies 2012, March 19th, 2012-last update [Homepage of Breastcancer.org], [Online]. Available: http://www.breastcancer.org/treatment/targeted_therapies/ [2012, May 17th].

Tech Tip 2: Western Blotting Troubleshooting Guide Extended Version 2010, , Thermo Fisher Scientific, Belgium.

The Best Treatment: Your guide to breast cancer treatment in England and Wales 2009, , Second edn, Breakthrough Breast Cancer, UK.

Treating Breast Cancer 2012, March 6th, 2012-last update [Homepage of American Cancer Society], [Online]. Available: http://www.cancer.org/Cancer/BreastCancer/DetailedGuide/breast-cancer-treating-general-info [2012, May 17th].

Types of treatment for breast cancer 2011, October 4th, 2011-last update [Homepage of Cancer Research UK], [Online]. Available: http://cancerhelp.cancerresearchuk.org/type/breast-cancer/treatment/which-treatment-for-breast-cancer [2012, May 17th].

Aaltonen, K.E., Ebbesson, A., Wigerup, C. & Hedenfalk, I. 2011, "Laser capture microdissection (LCM) and whole genome amplification (WGA) of DNA from normal breast tissue --- optimization for genome wide array analyses", *BMC research notes*, vol. 4, pp. 69.

Adams, P.S. 2006, "Data analysis and reporting" in *Real-time PCR*, ed. M.T. Dorak, Taylor & Francis Group, The United Kingdom, pp. 39-62.

Adelaide, J., Finetti, P., Bekhouche, I., Repellini, L., Geneix, J., Sircoulomb, F., Charafe-Jauffret, E., Cervera, N., Desplans, J., Parzy, D., Schoenmakers, E., Viens, P., Jacquemier, J., Birnbaum, D., Bertucci, F. & Chaffanet, M. 2007, "Integrated profiling of basal and luminal breast cancers", *Cancer research*, vol. 67, no. 24, pp. 11565-11575.

Adriance, M.C., Inman, J.L., Petersen, O.W. & Bissell, M.J. 2005, "Myoepithelial cells: good fences make good neighbors", *Breast cancer research : BCR*, vol. 7, no. 5, pp. 190-197.

Al Tamimi, D.M., Shawarby, M.A., Ahmed, A., Hassan, A.K. & AlOdaini, A.A. 2010, "Protein expression profile and prevalence pattern of the molecular classes of breast cancer--a Saudi population based study", *BMC cancer*, vol. 10, pp. 223.

Albrektsen, G., Heuch, I. & Thoresen, S.O. 2010, "Histological type and grade of breast cancer tumors by parity, age at birth, and time since birth: a register-based study in Norway", *BMC cancer*, vol. 10, pp. 226.

Al-Hajj, M., Wicha, M.S., Benito-Hernandez, A., Morrison, S.J. & Clarke, M.F. 2003, "Prospective identification of tumorigenic breast cancer cells", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 7, pp. 3983-3988.

Alkan, C., Coe, B.P. & Eichler, E.E. 2011, "Genome structural variation discovery and genotyping", *Nature reviews.Genetics*, vol. 12, no. 5, pp. 363-376.

Al-Kuraya, K., Schraml, P., Torhorst, J., Tapia, C., Zaharieva, B., Novotny, H., Spichtin, H., Maurer, R., Mirlacher, M., Kochli, O., Zuber, M., Dieterich, H., Mross, F., Wilber, K., Simon, R. & Sauter, G. 2004, "Prognostic relevance of gene amplifications and coamplifications in breast cancer", *Cancer research*, vol. 64, no. 23, pp. 8534-8540.

Allred, D.C., Mohsin, S.K. & Fuqua, S.A. 2001, "Histological and biological evolution of human premalignant breast disease", *Endocrine-related cancer,* vol. 8, no. 1, pp. 47-61.

Alvarez, S., Diaz-Uriarte, R., Osorio, A., Barroso, A., Melchor, L., Paz, M.F., Honrado, E., Rodriguez, R., Urioste, M., Valle, L., Diez, O., Cigudosa, J.C., Dopazo, J., Esteller, M. & Benitez, J. 2005, "A predictor based on the somatic genomic changes of the BRCA1/BRCA2 breast cancer tumors identifies the non-BRCA1/BRCA2 tumors with BRCA1 promoter hypermethylation", *Clinical cancer research : an official journal of the American Association for Cancer Research,* vol. 11, no. 3, pp. 1146-1153.

Ambady, S., Malcuit, C., Kashpur, O., Kole, D., Holmes, W.F., Hedblom, E., Page, R.L. & Dominko, T. 2010, "Expression of NANOG and NANOGP8 in a variety of undifferentiated and differentiated human cells", *The International journal of developmental biology,* vol. 54, no. 11-12, pp. 1743-1754.

Amir, E., Freedman, O.C., Seruga, B. & Evans, D.G. 2010, "Assessing women at high risk of breast cancer: a review of risk assessment models", *Journal of the National Cancer Institute,* vol. 102, no. 10, pp. 680-691.

Arnould, L., Coudert, B. & Fumoleau, P. 2009, "Is immunohistochemistry less sensitive than quantitative reverse transcriptase polymerase chain reaction for hormone receptor status determination in breast cancer?", *Journal of clinical oncology : official journal of the American Society of Clinical Oncology,* vol. 27, no. 16, pp. 2733; author reply 2734-5.

Arriagada, R., Le, M.G., Dunant, A., Tubiana, M. & Contesso, G. 2006, "Twenty-five years of follow-up in patients with operable breast carcinoma: correlation between clinicopathologic factors and the risk of death in each 5-year period", *Cancer,* vol. 106, no. 4, pp. 743-750.

Ashkenazi, R., Gentry, S.N. & Jackson, T.L. 2008, "Pathways to tumorigenesis--modeling mutation acquisition in stem cells and their progeny", *Neoplasia (New York, N.Y.),* vol. 10, no. 11, pp. 1170-1182.

Bae, J.S., Choi, J.S., Baik, S.H., Park, W.C., Song, B.J., Kim, J.S., Lim, Y. & Jung, S.S. 2010, "Genomic alterations of primary tumor and blood in invasive ductal carcinoma of breast", *World journal of surgical oncology,* vol. 8, pp. 32.

Barberis, M., Pellegrini, C., Cannone, M., Arizzi, C., Coggi, G. & Bosari, S. 2008, "Quantitative PCR and HER2 testing in breast cancer: a technical and cost-effectiveness analysis", *American Journal of Clinical Pathology,* vol. 129, no. 4, pp. 563-570.

Belengeanu, A., Muresan, A., Stoicanescu, D. & Lazar, E. 2010, "Amplification of HER-2 gene in breast cancer: immunohistochemical and FISH assessment", *Romanian journal of morphology and embryology = Revue roumaine de morphologie et embryologie,* vol. 51, no. 2, pp. 321-326.

Ben Abdelkrim, S., Trabelsi, A., Missaoui, N., Beizig, N., Bdioui, A., Anjorin, A., Jomaa, W. & Mokni, M. 2010, "Distribution of molecular breast cancer subtypes among Tunisian women and correlation with histopathological parameters: a study of 194 patients", *Pathology, research and practice,* vol. 206, no. 11, pp. 772-775.

Benoy, I.H., Elst, H., Van der Auwera, I., Van Laere, S., van Dam, P., Van Marck, E., Scharpe, S., Vermeulen, P.B. & Dirix, L.Y. 2004, "Real-time RT-PCR correlates with immunocytochemistry for the detection of disseminated epithelial cells in bone marrow aspirates of patients with breast cancer", *British journal of cancer,* vol. 91, no. 10, pp. 1813-1820.

Ben-Porath, I., Thomson, M.W., Carey, V.J., Ge, R., Bell, G.W., Regev, A. & Weinberg, R.A. 2008, "An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors", *Nature genetics,* vol. 40, no. 5, pp. 499-507.

Bensalah, K., Montorsi, F. & Shariat, S.F. 2007, "Challenges of cancer biomarker profiling", *European urology,* vol. 52, no. 6, pp. 1601-1609.

Berliner, J.L., Fay, A.M. & Practice Issues Subcommittee of the National Society of Genetic Counselors' Familial Cancer Risk Counseling Special Interest Group 2007, "Risk assessment and genetic counseling for hereditary breast and ovarian cancer: recommendations of the National Society of Genetic Counselors", *Journal of genetic counseling,* vol. 16, no. 3, pp. 241-260.

Bertucci, F., Finetti, P., Cervera, N., Charafe-Jauffret, E., Mamessier, E., Adelaide, J., Debono, S., Houvenaeghel, G., Maraninchi, D., Viens, P., Charpin, C., Jacquemier, J. & Birnbaum, D. 2006, "Gene expression profiling shows medullary breast cancer is a subgroup of basal breast cancers", *Cancer research,* vol. 66, no. 9, pp. 4636-4644.

Beser, A.R., Tuzlali, S., Guzey, D., Dolek Guler, S., Hacihanefioglu, S. & Dalay, N. 2007, "HER-2, TOP2A and chromosome 17 alterations in breast cancer", *Pathology oncology research : POR*, vol. 13, no. 3, pp. 180-185.

Bhatt, A.N., Mathur, R., Farooque, A., Verma, A. & Dwarakanath, B.S. 2010, "Cancer biomarkers - current perspectives", *The Indian journal of medical research*, vol. 132, pp. 129-149.

Bhattacharya, S. & Adhikary, S. 2006, "Evaluation of risk factors, diagnosis and treatment in carcinoma breast--a retrospective study", *Kathmandu University medical journal (KUMJ)*, vol. 4, no. 1, pp. 54-60.

Biomarkers Definitions Working Group. 2001, "Biomarkers and surrogate endpoints: preferred definitions and conceptual framework", *Clinical pharmacology and therapeutics*, vol. 69, no. 3, pp. 89-95.

Birrer, M.J., Johnson, M.E., Hao, K., Wong, K.K., Park, D.C., Bell, A., Welch, W.R., Berkowitz, R.S. & Mok, S.C. 2007, "Whole genome oligonucleotide-based array comparative genomic hybridization analysis identified fibroblast growth factor 1 as a prognostic marker for advanced-stage serous ovarian adenocarcinomas", *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, vol. 25, no. 16, pp. 2281-2287.

Bocker, W., Moll, R., Poremba, C., Holland, R., Van Diest, P.J., Dervan, P., Burger, H., Wai, D., Ina Diallo, R., Brandt, B., Herbst, H., Schmidt, A., Lerch, M.M. & Buchwallow, I.B. 2002, "Common adult stem cells in the human breast give rise to glandular and myoepithelial cell lineages: a new cell biological concept", *Laboratory investigation; a journal of technical methods and pathology*, vol. 82, no. 6, pp. 737-746.

Boecker, W. & Buerger, H. 2003, "Evidence of progenitor cells of glandular and myoepithelial cell lineages in the human adult female breast epithelium: a new progenitor (adult stem) cell concept", *Cell proliferation*, vol. 36 Suppl 1, pp. 73-84.

Bofin, A.M., Ytterhus, B., Martin, C., O'Leary, J.J. & Hagmar, B.M. 2004, "Detection and quantitation of HER-2 gene amplification and protein expression in breast carcinoma", *American Journal of Clinical Pathology*, vol. 122, no. 1, pp. 110-119.

Bombonati, A. & Sgroi, D.C. 2011, "The molecular pathology of breast cancer progression", *The Journal of pathology*, vol. 223, no. 2, pp. 307-317.

Borner, C., Filipuzzi, I., Wartmann, M., Eppenberger, U. & Fabbro, D. 1989, "Biosynthesis and posttranslational modifications of protein kinase C in human breast cancer cells", *The Journal of biological chemistry*, vol. 264, no. 23, pp. 13902-13909.

Bourdon, V., Naef, F., Rao, P.H., Reuter, V., Mok, S.C., Bosl, G.J., Koul, S., Murty, V.V., Kucherlapati, R.S. & Chaganti, R.S. 2002, "Genomic and expression analysis of the 12p11-p12 amplicon using EST arrays identifies two novel amplified and overexpressed genes", *Cancer research*, vol. 62, no. 21, pp. 6218-6223.

Bouskine, A., Vega, A., Nebout, M., Benahmed, M. & Fenichel, P. 2010, "Expression of embryonic stem cell markers in cultured JKT-1, a cell line derived from a human seminoma", *International journal of andrology*, vol. 33, no. 1, pp. 54-63.

Boyd, N.F., Jensen, H.M., Cooke, G., Han, H.L., Lockwood, G.A. & Miller, A.B. 2000, "Mammographic densities and the prevalence and incidence of histological types of benign breast disease. Reference Pathologists of the Canadian National Breast Screening Study", *European journal of cancer prevention : the official journal of the European Cancer Prevention Organisation (ECP)*, vol. 9, no. 1, pp. 15-24.

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., Gifford, D.K., Melton, D.A., Jaenisch, R. & Young, R.A. 2005, "Core transcriptional regulatory circuitry in human embryonic stem cells", *Cell*, vol. 122, no. 6, pp. 947-956.

Braakhuis, B.J., Tabor, M.P., Kummer, J.A., Leemans, C.R. & Brakenhoff, R.H. 2003, "A genetic explanation of Slaughter's concept of field cancerization: evidence and clinical implications", *Cancer research*, vol. 63, no. 8, pp. 1727-1730.

Briegel, K.J. 2006, "Embryonic transcription factors in human breast cancer", *IUBMB life*, vol. 58, no. 3, pp. 123-132.

Bui, T.D., Rankin, J., Smith, K., Huguet, E.L., Ruben, S., Strachan, T., Harris, A.L. & Lindsay, S. 1997, "A novel human Wnt gene, WNT10B, maps to 12q13 and is expressed in human breast carcinomas", *Oncogene*, vol. 14, no. 10, pp. 1249-1253.

Bundred, N.J. 2001, "Prognostic and predictive factors in breast cancer", *Cancer treatment reviews*, vol. 27, no. 3, pp. 137-142.

Burstein, H.J. 2005, "The distinctive nature of HER2-positive breast cancers", *The New England journal of medicine,* vol. 353, no. 16, pp. 1652-1654.

Busch, C., Bareiss, P.M., Sinnberg, T., Just, L., Wehrmann, M., Fuchs, J., Garbe, C. & Drews, U. 2009, "Isolation of three stem cell lines from human sacrococcygeal teratomas", *The Journal of pathology,* vol. 217, no. 4, pp. 589-596.

Bussolati, G. & Leonardo, E. 2008, "Technical pitfalls potentially affecting diagnoses in immunohistochemistry", *Journal of clinical pathology,* vol. 61, no. 11, pp. 1184-1192.

Bustin, S.A. & Nolan, T. a, "The PCR step" in *A - Z of quantitative PCR*, ed. S.A. Bustin, International University Line, The United States of America, pp. 399-438.

Bustin, S.A. & Nolan, T. b, "Template Handling, preparation, and quantification" in *A - Z of quantitative PCR*, ed. S.A. Bustin, International University Line, The United States of America, pp. 143-213.

Bustin, S.A. & Nolan, T. 2004, "Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction", *Journal of biomolecular techniques : JBT,* vol. 15, no. 3, pp. 155-166.

Cai, J.H., Deng, S., Kumpf, S.W., Lee, P.A., Zagouras, P., Ryan, A. & Gallagher, D.S. 2007, "Validation of rat reference genes for improved quantitative gene expression analysis using low density arrays", *BioTechniques,* vol. 42, no. 4, pp. 503-512.

Calza, S., Hall, P., Auer, G., Bjohle, J., Klaar, S., Kronenwett, U., Liu, E.T., Miller, L., Ploner, A., Smeds, J., Bergh, J. & Pawitan, Y. 2006, "Intrinsic molecular signature of breast cancer in a population-based cohort of 412 patients", *Breast cancer research : BCR,* vol. 8, no. 4, pp. R34.

Capony, F., Morisset, M., Barrett, A.J., Capony, J.P., Broquet, P., Vignon, F., Chambon, M., Louisot, P. & Rochefort, H. 1987, "Phosphorylation, glycosylation, and proteolytic activity of the 52-kD estrogen-induced protein secreted by MCF7 cells", *The Journal of cell biology,* vol. 104, no. 2, pp. 253-262.

Cariati, M. & Purushotham, A.D. 2008, "Stem cells and breast cancer", *Histopathology,* vol. 52, no. 1, pp. 99-107.

Carter, N.P. 2007, "Methods and strategies for analyzing copy number variation using DNA microarrays", *Nature genetics,* vol. 39, no. 7 Suppl, pp. S16-21.

Cavallone, L., Arcand, S.L., Maugard, C.M., Nolet, S., Gaboury, L.A., Mes-Masson, A.M., Ghadirian, P., Provencher, D. & Tonin, P.N. 2010, "Comprehensive BRCA1 and BRCA2 mutation analyses and review of French Canadian families with at least three cases of breast cancer", *Familial cancer,* vol. 9, no. 4, pp. 507-517.

Chai, H. & Brown, R.E. 2009, "Field effect in cancer-an update", *Annals of Clinical and Laboratory Science,* vol. 39, no. 4, pp. 331-337.

Chaiwun, B., Nakrungsee, S., Sukhamwang, N. & Srisukho, S. 2010, "A study of high-nuclear-grade breast cancer in Thailand: subclassification and correlation with prognostic factors and immunohistochemical study", *Breast cancer (Tokyo, Japan),* vol. 17, no. 1, pp. 35-41.

Chan, K.K., Zhang, J., Chia, N.Y., Chan, Y.S., Sim, H.S., Tan, K.S., Oh, S.K., Ng, H.H. & Choo, A.B. 2009, "KLF4 and PBX1 directly regulate NANOG expression in human embryonic stem cells", *Stem cells (Dayton, Ohio)*, vol. 27, no. 9, pp. 2114-2125.

Chang, D.F., Tsai, S.C., Wang, X.C., Xia, P., Senadheera, D. & Lutzko, C. 2009, "Molecular characterization of the human NANOG protein", *Stem cells (Dayton, Ohio)*, vol. 27, no. 4, pp. 812-821.

Charafe-Jauffret, E., Ginestier, C. & Birnbaum, D. 2009, "Breast cancer stem cells: tools and models to rely on", *BMC cancer*, vol. 9, pp. 202.

Charafe-Jauffret, E., Monville, F., Ginestier, C., Dontu, G., Birnbaum, D. & Wicha, M.S. 2008, "Cancer stem cells in breast: current opinion and future challenges", *Pathobiology : journal of immunopathology, molecular and cellular biology*, vol. 75, no. 2, pp. 75-84.

Chen, J. & Chen, Z.L. 2010, "Technology update for the sorting and identification of breast cancer stem cells", *Chinese journal of cancer*, vol. 29, no. 3, pp. 265-269.

Chen, W., Houldsworth, J., Olshen, A.B., Nanjangud, G., Chaganti, S., Venkatraman, E.S., Halaas, J., Teruya-Feldstein, J., Zelenetz, A.D. & Chaganti, R.S. 2006, "Array comparative genomic hybridization reveals genomic copy number changes associated with outcome in diffuse large B-cell lymphomas", *Blood*, vol. 107, no. 6, pp. 2477-2485.

Chibon, F., Mariani, O., Derre, J., Mairal, A., Coindre, J.M., Guillou, L., Sastre, X., Pedeutour, F. & Aurias, A. 2004, "ASK1 (MAP3K5) as a potential therapeutic target in malignant fibrous histiocytomas with 12q14-q15 and 6q23 amplifications", *Genes, chromosomes & cancer,* vol. 40, no. 1, pp. 32-37.

Chin, K., DeVries, S., Fridlyand, J., Spellman, P.T., Roydasgupta, R., Kuo, W.L., Lapuk, A., Neve, R.M., Qian, Z., Ryder, T., Chen, F., Feiler, H., Tokuyasu, T., Kingsley, C., Dairkee, S., Meng, Z., Chew, K., Pinkel, D., Jain, A., Ljung, B.M., Esserman, L., Albertson, D.G., Waldman, F.M. & Gray, J.W. 2006, "Genomic and transcriptional aberrations linked to breast cancer pathophysiologies", *Cancer cell,* vol. 10, no. 6, pp. 529-541.

Chiou, S.H., Yu, C.C., Huang, C.Y., Lin, S.C., Liu, C.J., Tsai, T.H., Chou, S.H., Chien, C.S., Ku, H.H. & Lo, J.F. 2008, "Positive correlations of Oct-4 and Nanog in oral cancer stem-like cells and high-grade oral squamous cell carcinoma", *Clinical cancer research : an official journal of the American Association for Cancer Research,* vol. 14, no. 13, pp. 4085-4095.

Cho, R.W. & Clarke, M.F. 2008, "Recent advances in cancer stem cells", *Current opinion in genetics & development,* vol. 18, no. 1, pp. 48-53.

Collaborative Group on Hormonal Factors in Breast Cancer 2001, "Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease", *Lancet,* vol. 358, no. 9291, pp. 1389-1399.

Collins, L.C., Botero, M.L. & Schnitt, S.J. 2005, "Bimodal frequency distribution of estrogen receptor immunohistochemical staining results in breast cancer: an analysis of 825 cases", *American Journal of Clinical Pathology,* vol. 123, no. 1, pp. 16-20.

Comen, E.A., Norton, L. & Massague, J. 2011, "Breast cancer tumor size, nodal status, and prognosis: biology trumps anatomy", *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, vol. 29, no. 19, pp. 2610-2612.

Concin, N., Zeillinger, C., Tong, D., Stimpfl, M., Konig, M., Printz, D., Stonek, F., Schneeberger, C., Hefler, L., Kainz, C., Leodolter, S., Haas, O.A. & Zeillinger, R. 2003, "Comparison of p53 mutational status with mRNA and protein expression in a panel of 24 human breast carcinoma cell lines", *Breast cancer research and treatment,* vol. 79, no. 1, pp. 37-46.

Conrad, D.F., Pinto, D., Redon, R., Feuk, L., Gokcumen, O., Zhang, Y., Aerts, J., Andrews, T.D., Barnes, C., Campbell, P., Fitzgerald, T., Hu, M., Ihm, C.H., Kristiansson, K., Macarthur, D.G., Macdonald, J.R., Onyiah, I., Pang, A.W., Robson, S., Stirrups, K., Valsesia, A., Walter, K., Wei, J., Wellcome Trust Case Control Consortium, Tyler-Smith, C., Carter, N.P., Lee, C., Scherer, S.W. & Hurles, M.E. 2010, "Origins and functional impact of copy number variation in the human genome", *Nature*, vol. 464, no. 7289, pp. 704-712.

Correa Geyer, F. & Reis-Filho, J.S. 2009, "Microarray-based gene expression profiling as a clinical tool for breast cancer management: are we there yet?", *International journal of surgical pathology*, vol. 17, no. 4, pp. 285-302.

Croci, D.O., Cogno, I.S., Vittar, N.B., Salvatierra, E., Trajtenberg, F., Podhajcer, O.L., Osinaga, E., Rabinovich, G.A. & Rivarola, V.A. 2008, "Silencing survivin gene expression promotes apoptosis of human breast cancer cells through a caspase-independent pathway", *Journal of cellular biochemistry,* vol. 105, no. 2, pp. 381-390.

Crowe, J.P., Patrick, R.J., Rybicki, L.A., Escobar, P.F., Weng, D., Thomas Budd, G., Tubbs, R.R., Procop, G.W. & Hicks, D.G. 2006, "A data model to predict HER2 status in breast cancer based on the clinical and pathologic profiles of a large patient population at a single institution", *Breast (Edinburgh, Scotland),* vol. 15, no. 6, pp. 728-735.

Cuzick, J., Dowsett, M., Pineda, S., Wale, C., Salter, J., Quinn, E., Zabaglo, L., Mallon, E., Green, A.R., Ellis, I.O., Howell, A., Buzdar, A.U. & Forbes, J.F. 2011, "Prognostic value of a combined estrogen receptor, progesterone receptor, ki-67, and human epidermal growth factor receptor 2 immunohistochemical score and comparison with the genomic health recurrence score in early breast cancer", *Journal of clinical oncology : official journal of the American Society of Clinical Oncology,* vol. 29, no. 32, pp. 4273-4278.

D'Amico, F., Skarmoutsou, E. & Stivala, F. 2009, "State of the art in antigen retrieval for immunohistochemistry", *Journal of immunological methods,* vol. 341, no. 1-2, pp. 1-18.

Da Silva, L., Buck, L., Simpson, P.T., Reid, L., McCallum, N., Madigan, B.J. & Lakhani, S.R. 2009, "Molecular and morphological analysis of adenoid cystic carcinoma of the breast with synchronous tubular adenosis", *Virchows Archiv : an international journal of pathology,* vol. 454, no. 1, pp. 107-114.

Daigo, Y., Chin, S.F., Gorringe, K.L., Bobrow, L.G., Ponder, B.A., Pharoah, P.D. & Caldas, C. 2001, "Degenerate oligonucleotide primed-polymerase chain reaction-based array comparative genomic hybridization for extensive amplicon profiling of breast cancers : a new approach for the molecular analysis of paraffin-embedded cancer tissue", *The American journal of pathology,* vol. 158, no. 5, pp. 1623-1631.

Daniele, L. & Sapino, A. 2009, "Anti-HER2 treatment and breast cancer: state of the art, recent patents, and new strategies", *Recent patents on anti-cancer drug discovery,* vol. 4, no. 1, pp. 9-18.

DeGraffenried, L.A., Fulcher, L., Friedrichs, W.E., Grunwald, V., Ray, R.B. & Hidalgo, M. 2004, "Reduced PTEN expression in breast cancer cells confers susceptibility to inhibitors of the PI3 kinase/Akt pathway", *Annals of Oncology : Official Journal of the European Society for Medical Oncology / ESMO,* vol. 15, no. 10, pp. 1510-1516.

de Kok, J.B., Roelofs, R.W., Giesendorf, B.A., Pennings, J.L., Waas, E.T., Feuth, T., Swinkels, D.W. & Span, P.N. 2005, "Normalization of gene expression measurements in tumor tissues: comparison of 13 endogenous control genes", *Laboratory investigation; a journal of technical methods and pathology,* vol. 85, no. 1, pp. 154-159.

de Kok, J.B., Ruers, T.J., van Muijen, G.N., van Bokhoven, A., Willems, H.L. & Swinkels, D.W. 2000, "Real-time quantification of human telomerase reverse transcriptase mRNA in tumors and healthy tissues", *Clinical chemistry,* vol. 46, no. 3, pp. 313-318.

Detre, S., Saclani Jotti, G. & Dowsett, M. 1995, "A "quickscore" method for immunohistochemical semiquantitation: validation for oestrogen receptor in breast carcinomas", *Journal of clinical pathology,* vol. 48, no. 9, pp. 876-878.

Deugnier, M.A., Teuliere, J., Faraldo, M.M., Thiery, J.P. & Glukhova, M.A. 2002, "The importance of being a myoepithelial cell", *Breast cancer research : BCR,* vol. 4, no. 6, pp. 224-230.

D'haene, B., Vandesompele, J. & Hellemans, J. 2010, "Accurate and objective copy number profiling using real-time quantitative PCR", *Methods (San Diego, Calif.),* vol. 50, no. 4, pp. 262-270.

Didraga, M.A., van Beers, E.H., Joosse, S.A., Brandwijk, K.I., Oldenburg, R.A., Wessels, L.F., Hogervorst, F.B., Ligtenberg, M.J., Hoogerbrugge, N., Verhoef, S., Devilee, P. & Nederlof, P.M. 2011, "A non-BRCA1/2 hereditary breast cancer subgroup defined by aCGH profiling of genetically related patients", *Breast cancer research and treatment*.

Dontu, G., Abdallah, W.M., Foley, J.M., Jackson, K.W., Clarke, M.F., Kawamura, M.J. & Wicha, M.S. 2003a, "In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells", *Genes & development,* vol. 17, no. 10, pp. 1253-1270.

Dontu, G., Al-Hajj, M., Abdallah, W.M., Clarke, M.F. & Wicha, M.S. 2003b, "Stem cells in normal breast development and breast cancer", *Cell proliferation,* vol. 36 Suppl 1, pp. 59-72.

Dontu, G., El-Ashry, D. & Wicha, M.S. 2004, "Breast cancer, stem/progenitor cells and the estrogen receptor", *Trends in endocrinology and metabolism: TEM,* vol. 15, no. 5, pp. 193-197.

Dougherty, C.J., Ichim, T.E., Liu, L., Reznik, G., Min, W.P., Ghochikyan, A., Agadjanyan, M.G. & Reznik, B.N. 2008, "Selective apoptosis of breast cancer cells by siRNA targeting of BORIS", *Biochemical and biophysical research communications,* vol. 370, no. 1, pp. 109-112.

Dowsett, M., Procter, M., McCaskill-Stevens, W., de Azambuja, E., Dafni, U., Rueschoff, J., Jordan, B., Dolci, S., Abramovitz, M., Stoss, O., Viale, G., Gelber, R.D., Piccart-Gebhart, M. & Leyland-Jones, B. 2009, "Disease-free survival according to degree of HER2 amplification for patients treated with adjuvant chemotherapy with or without 1 year of trastuzumab: the HERA Trial", *Journal of clinical oncology : official journal of the American Society of Clinical Oncology,* vol. 27, no. 18, pp. 2962-2969.

Drabsch, Y., Robert, R.G. & Gonda, T.J. 2010, "MYB suppresses differentiation and apoptosis of human breast cancer cells", *Breast cancer research : BCR,* vol. 12, no. 4, pp. R55.

Du, X.L., Xia, R., Liu, C.C., Cormier, J.N., Xing, Y., Hardy, D., Chan, W. & Burau, K. 2009, "Cardiac toxicity associated with anthracycline-containing chemotherapy in older women with breast cancer", *Cancer,* vol. 115, no. 22, pp. 5296-5308.

Du, Z., Li, J., Wang, L., Bian, C., Wang, Q., Liao, L., Dou, X., Bian, X. & Zhao, R.C. 2010, "Overexpression of DeltaNp63alpha induces a stem cell phenotype in MCF7 breast carcinoma cell line through the Notch pathway", *Cancer science,* vol. 101, no. 11, pp. 2417-2424.

Duffy, M.J. 2005, "Predictive markers in breast and other cancers: a review", *Clinical chemistry,* vol. 51, no. 3, pp. 494-503.

Dukers, D.F., van Galen, J.C., Giroth, C., Jansen, P., Sewalt, R.G., Otte, A.P., Kluin-Nelemans, H.C., Meijer, C.J. & Raaphorst, F.M. 2004, "Unique polycomb gene expression pattern in Hodgkin's lymphoma and Hodgkin's lymphoma-derived cell lines", *The American journal of pathology,* vol. 164, no. 3, pp. 873-881.

Dunnwald, L.K., Rossing, M.A. & Li, C.I. 2007, "Hormone receptor status, tumor characteristics, and prognosis: a prospective cohort of breast cancer patients", *Breast cancer research : BCR,* vol. 9, no. 1, pp. R6.

Durham, S.E. & Chinnery, P.F. 2006, "Mitochondrial DNA analysis" in *Real-time PCR*, ed. M.T. Dorak, Taylor & Francis Group, UK, pp. 183-204.

Eberle, I., Pless, B., Braun, M., Dingermann, T. & Marschalek, R. 2010, "Transcriptional properties of human NANOG1 and NANOG2 in acute leukemic cells", *Nucleic acids research,* vol. 38, no. 16, pp. 5384-5395.

Ehira, N., Oshiumi, H., Matsumoto, M., Kondo, T., Asaka, M. & Seya, T. 2010, "An embryo-specific expressing TGF-beta family protein, growth-differentiation factor 3 (GDF3), augments progression of B16 melanoma", *Journal of experimental & clinical cancer research : CR,* vol. 29, pp. 135.

Elliman, S.J., Wu, I. & Kemp, D.M. 2006, "Adult tissue-specific expression of a Dppa3-derived retrogene represents a postnatal transcript of pluripotent cell origin", *The Journal of biological chemistry,* vol. 281, no. 1, pp. 16-19.

Ellison-Zelski, S.J. & Alarid, E.T. 2010, "Maximum growth and survival of estrogen receptor-alpha positive breast cancer cells requires the Sin3A transcriptional repressor", *Molecular cancer,* vol. 9, pp. 263.

Ellsworth, D.L., Ellsworth, R.E., Love, B., Deyarmin, B., Lubert, S.M., Mittal, V., Hooke, J.A. & Shriver, C.D. 2004, "Outer breast quadrants demonstrate increased levels of genomic instability", *Annals of surgical oncology,* vol. 11, no. 9, pp. 861-868.

Elston, C.W. & Ellis, I.O. 1991, "Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up", *Histopathology,* vol. 19, no. 5, pp. 403-410.

Engel, H., Friedrich, J., Kleespies, C., Kurbacher, C.M., Schondorf, T., Grecu, O., Kolhagen, H. & Mallmann, P. 1998, "Detection of chromosomal aberrations in tumor cells and tumor infiltrating lymphocytes by molecular cytogenetics in patients with gynecological cancer", *Cancer genetics and cytogenetics,* vol. 106, no. 2, pp. 159-165.

Epstein, R.J. 2003, "Replicated DNA is packaged within chromosomes" in *Human Molecular Biology: An Introduction to The Molecular Basis of Health and Disease* Cambridge University Press, The United Kingdom, pp. 62.

Esteves, V.F., Thuler, L.C., Amendola, L.C., Koifman, R.J., Koifman, S., Frankel, P.P., Vieira, R.J. & Brazilian Network of Breast and Ovarian Familial Cancer Aggregation 2009, "Prevalence of BRCA1 and BRCA2 gene mutations in families with medium and high risk of breast and ovarian cancer in Brazil", *Brazilian journal of medical and biological research = Revista brasileira de pesquisas medicas e biologicas / Sociedade Brasileira de Biofisica ...[et al.],* vol. 42, no. 5, pp. 453-457.

Evans, D.G. & Howell, A. 2007, "Breast cancer risk-assessment models", *Breast cancer research : BCR,* vol. 9, no. 5, pp. 213.

Ezeh, U.I., Turek, P.J., Reijo, R.A. & Clark, A.T. 2005, "Human embryonic stem cell genes OCT4, NANOG, STELLAR, and GDF3 are expressed in both seminoma and breast carcinoma", *Cancer,* vol. 104, no. 10, pp. 2255-2265.

Falasca, M. 2012, "Cancer biomarkers: the future challenge of cancer", *Journal of Molecular Biomarkers & Diagnosis,* , no. S2: e001.

Farabegoli, F., Santini, D., Ceccarelli, C., Taffurelli, M., Marrano, D. & Baldini, N. 2001, "Clone heterogeneity in diploid and aneuploid breast carcinomas as detected by FISH", *Cytometry,* vol. 46, no. 1, pp. 50-56.

Ferretti, G., Fabi, A., Felici, A. & Papaldo, P. 2010, "Improved prognosis by trastuzumab of women with HER2-positive breast cancer compared with those with HER2-negative disease", *Journal of clinical oncology : official journal of the American Society of Clinical Oncology,* vol. 28, no. 20, pp. e337; author reply e338-9.

Feuk, L., Carson, A.R. & Scherer, S.W. 2006, "Structural variation in the human genome", *Nature reviews.Genetics*, vol. 7, no. 2, pp. 85-97.

Filho, O.M., Ignatiadis, M. & Sotiriou, C. 2011, "Genomic Grade Index: An important tool for assessing breast cancer tumor grade and prognosis", *Critical reviews in oncology/hematology*, vol. 77, no. 1, pp. 20-29.

Fitzgibbons PL, Page DL, Weaver D, Thor AD, Allred DC, Clark GM, Ruby SG, O'Malley F, Simpson JF, Connolly JL, Hayes DF, Edge SB, Lichter A & Schnitt SJ 2000, "Prognostic Factors in Breast Cancer. College of American Pathologists Consensus Statement 1999", *Archives of Pathology & Laboratory Medicine*, vol. 124, pp. 966-978.

Fleige, S., Walf, V., Huch, S., Prgomet, C., Sehm, J. & Pfaffl, M.W. 2006, "Comparison of relative mRNA quantification models and the impact of RNA integrity in quantitative real-time RT-PCR", *Biotechnology Letters*, vol. 28, no. 19, pp. 1601-1613.

Fraenkel-Conrat, H., Brandon, B.A. & Olcott, H.S. 1947, "The reaction of formaldehyde with proteins; participation of indole groups; gramicidin", *The Journal of biological chemistry*, vol. 168, no. 1, pp. 99-118.

Fraenkel-Conrat, H. & Olcott, H.S. 1948a, "The reaction of formaldehyde with proteins; cross-linking between amino and primary amide or guanidyl groups", *Journal of the American Chemical Society*, vol. 70, no. 8, pp. 2673-2684.

Fraenkel-Conrat, H. & Olcott, H.S. 1948b, "Reaction of formaldehyde with proteins; cross-linking of amino groups with phenol, imidazole, or indole groups", *The Journal of biological chemistry,* vol. 174, no. 3, pp. 827-843.

Frank, R. & Hargreaves, R. 2003, "Clinical biomarkers in drug discovery and development", *Nature reviews.Drug discovery,* vol. 2, no. 7, pp. 566-580.

Freeman, J.L., Perry, G.H., Feuk, L., Redon, R., McCarroll, S.A., Altshuler, D.M., Aburatani, H., Jones, K.W., Tyler-Smith, C., Hurles, M.E., Carter, N.P., Scherer, S.W. & Lee, C. 2006, "Copy number variation: new insights in genome diversity", *Genome research,* vol. 16, no. 8, pp. 949-961.

Fridlyand, J., Snijders, A.M., Ylstra, B., Li, H., Olshen, A., Segraves, R., Dairkee, S., Tokuyasu, T., Ljung, B.M., Jain, A.N., McLennan, J., Ziegler, J., Chin, K., Devries, S., Feiler, H., Gray, J.W., Waldman, F., Pinkel, D. & Albertson, D.G. 2006, "Breast tumor copy number aberration phenotypes and genomic instability", *BMC cancer,* vol. 6, pp. 96.

Galliher, A.J. & Schiemann, W.P. 2006, "Beta3 integrin and Src facilitate transforming growth factor-beta mediated induction of epithelial-mesenchymal transition in mammary epithelial cells", *Breast cancer research : BCR,* vol. 8, no. 4, pp. R42.

Gao, Y., Niu, Y., Wang, X., Wei, L. & Lu, S. 2009, "Genetic changes at specific stages of breast cancer progression detected by comparative genomic hybridization", *Journal of Molecular Medicine (Berlin, Germany),* vol. 87, no. 2, pp. 145-152.

Ge, Y.L., Zhang, X., Zhang, J.Y., Hou, L. & Tian, R.H. 2009, "The mechanisms on apoptosis by inhibiting VEGF expression in human breast cancer cells", *International immunopharmacology*, vol. 9, no. 4, pp. 389-395.

Geiger, T., Cox, J. & Mann, M. 2010, "Proteomic changes resulting from gene copy number variations in cancer cells", *PLoS genetics*, vol. 6, no. 9, pp. e1001090.

Ginestier, C., Hur, M.H., Charafe-Jauffret, E., Monville, F., Dutcher, J., Brown, M., Jacquemier, J., Viens, P., Kleer, C.G., Liu, S., Schott, A., Hayes, D., Birnbaum, D., Wicha, M.S. & 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*, vol. 1, no. 5, pp. 555-567.

Ginsburg, O.M., Dinh, N.V., To, T.V., Quang, L.H., Linh, N.D., Duong, B.T., Royer, R., Llacuachaqui, M., Tulman, A., Vichodez, G., Li, S., Love, R.R. & Narod, S.A. 2011, "Family history, BRCA mutations and breast cancer in Vietnamese women", *Clinical genetics*, vol. 80, no. 1, pp. 89-92.

Giuliano, C.J., Kerley-Hamilton, J.S., Bee, T., Freemantle, S.J., Manickaratnam, R., Dmitrovsky, E. & Spinella, M.J. 2005, "Retinoic acid represses a cassette of candidate pluripotency chromosome 12p genes during induced loss of human embryonal carcinoma tumorigenicity", *Biochimica et biophysica acta*, vol. 1731, no. 1, pp. 48-56.

Glenn, S.T., Jones, C.A., Liang, P., Kaushik, D., Gross, K.W. & Kim, H.L. 2007, "Expression profiling of archival renal tumors by quantitative PCR to validate prognostic markers", *BioTechniques*, vol. 43, no. 5, pp. 639-40, 642-3, 647.

Goddard, N.C., McIntyre, A., Summersgill, B., Gilbert, D., Kitazawa, S. & Shipley, J. 2007, "KIT and RAS signalling pathways in testicular germ cell tumours: new data and a review of the literature", *International journal of andrology,* vol. 30, no. 4, pp. 337-48; discussion 349.

Gomes, L.R., Terra, L.F., Wailemann, R.A., Labriola, L. & Sogayar, M.C. 2012, "TGF-beta1 modulates the homeostasis between MMPs and MMP inhibitors through p38 MAPK and ERK1/2 in highly invasive breast cancer cells", *BMC cancer,* vol. 12, pp. 26.

Gonzalez-Suarez, E., Jacob, A.P., Jones, J., Miller, R., Roudier-Meyer, M.P., Erwert, R., Pinkas, J., Branstetter, D. & Dougall, W.C. 2010, "RANK ligand mediates progestin-induced mammary epithelial proliferation and carcinogenesis", *Nature,* vol. 468, no. 7320, pp. 103-107.

Gorringe, K.L., Boussioutas, A., Bowtell, D.D. & Melbourne Gastric Cancer Group, Peter Mac Micro Array Facility 2005, "Novel regions of chromosomal amplification at 6p21, 5p13, and 12q14 in gastric cancer identified by array comparative genomic hybridization", *Genes, chromosomes & cancer,* vol. 42, no. 3, pp. 247-259.

Grigoriadis, A., Mackay, A., Reis-Filho, J.S., Steele, D., Iseli, C., Stevenson, B.J., Jongeneel, C.V., Valgeirsson, H., Fenwick, K., Iravani, M., Leao, M., Simpson, A.J., Strausberg, R.L., Jat, P.S., Ashworth, A., Neville, A.M. & O'Hare, M.J. 2006, "Establishment of the epithelial-specific transcriptome of normal and malignant human breast cells based on MPSS and array expression data", *Breast cancer research : BCR,* vol. 8, no. 5, pp. R56.

Gudjonsson, T., Adriance, M.C., Sternlicht, M.D., Petersen, O.W. & Bissell, M.J. 2005, "Myoepithelial cells: their origin and function in breast morphogenesis and neoplasia", *Journal of mammary gland biology and neoplasia,* vol. 10, no. 3, pp. 261-272.

Gudjonsson, T. & Magnusson, M.K. 2005, "Stem cell biology and the cellular pathways of carcinogenesis", *APMIS : Acta Pathologica, Microbiologica, et Immunologica Scandinavica,* vol. 113, no. 11-12, pp. 922-929.

Gudjonsson, T., Ronnov-Jessen, L., Villadsen, R., Rank, F., Bissell, M.J. & Petersen, O.W. 2002, "Normal and tumor-derived myoepithelial cells differ in their ability to interact with luminal breast epithelial cells for polarity and basement membrane deposition", *Journal of cell science,* vol. 115, no. Pt 1, pp. 39-50.

Gunster, M.J., Raaphorst, F.M., Hamer, K.M., den Blaauwen, J.L., Fieret, E., Meijer, C.J. & Otte, A.P. 2001, "Differential expression of human Polycomb group proteins in various tissues and cell types", *Journal of cellular biochemistry.Supplement,* vol. Suppl 36, pp. 129-143.

Hammer, C., Fanning, A. & Crowe, J. 2008, "Overview of breast cancer staging and surgical treatment options", *Cleveland Clinic journal of medicine,* vol. 75 Suppl 1, pp. S10-6.

Han, G., Fan, B., Zhang, Y., Zhou, X., Wang, Y., Dong, H., Wei, Y., Sun, S., Hu, M., Zhang, J. & Wei, L. 2008, "Positive regulation of migration and invasion by vasodilator-stimulated phosphoprotein via Rac1 pathway in human breast cancer cells", *Oncology reports,* vol. 20, no. 4, pp. 929-939.

Han, W., Jung, E.M., Cho, J., Lee, J.W., Hwang, K.T., Yang, S.J., Kang, J.J., Bae, J.Y., Jeon, Y.K., Park, I.A., Nicolau, M., Jeffrey, S.S. & Noh, D.Y. 2008, "DNA copy number alterations and expression of relevant genes in triple-negative breast cancer", *Genes, chromosomes & cancer,* vol. 47, no. 6, pp. 490-499.

Hanker, L., Karn, T., Ruckhaeberle, E., Gaetje, R., Solbach, C., Schmidt, M., Engels, K., Holtrich, U., Kaufmann, M. & Rody, A. 2010, "Clinical relevance of the putative stem cell marker p63 in breast cancer", *Breast cancer research and treatment,* vol. 122, no. 3, pp. 765-775.

Hart, A.H., Hartley, L., Ibrahim, M. & Robb, L. 2004, "Identification, cloning and expression analysis of the pluripotency promoting Nanog genes in mouse and human", *Developmental dynamics : an official publication of the American Association of Anatomists,* vol. 230, no. 1, pp. 187-198.

Hart, A.H., Hartley, L., Parker, K., Ibrahim, M., Looijenga, L.H., Pauchnik, M., Chow, C.W. & Robb, L. 2005, "The pluripotency homeobox gene NANOG is expressed in human germ cell tumors", *Cancer,* vol. 104, no. 10, pp. 2092-2098.

Harvey, J.M., Clark, G.M., Osborne, C.K. & Allred, D.C. 1999, "Estrogen receptor status by immunohistochemistry is superior to the ligand-binding assay for predicting response to adjuvant endocrine therapy in breast cancer", *Journal of clinical oncology : official journal of the American Society of Clinical Oncology,* vol. 17, no. 5, pp. 1474-1481.

Hastings, P.J., Lupski, J.R., Rosenberg, S.M. & Ira, G. 2009, "Mechanisms of change in gene copy number", *Nature reviews.Genetics,* vol. 10, no. 8, pp. 551-564.

Hawthorn, L., Luce, J., Stein, L. & Rothschild, J. 2010, "Integration of transcript expression, copy number and LOH analysis of infiltrating ductal carcinoma of the breast", *BMC cancer,* vol. 10, pp. 460.

Heidenblad, M., Hallor, K.H., Staaf, J., Jonsson, G., Borg, A., Hoglund, M., Mertens, F. & Mandahl, N. 2006, "Genomic profiling of bone and soft tissue tumors with supernumerary ring chromosomes using tiling resolution bacterial artificial chromosome microarrays", *Oncogene,* vol. 25, no. 53, pp. 7106-7116.

Heidenblad, M., Jonson, T., Mahlamaki, E.H., Gorunova, L., Karhu, R., Johansson, B. & Hoglund, M. 2002, "Detailed genomic mapping and expression analyses of 12p amplifications in pancreatic carcinomas reveal a 3.5-Mb target region for amplification", *Genes, chromosomes & cancer,* vol. 34, no. 2, pp. 211-223.

Holm, M., Paaschburg, B., Balslev, E., Axelsson, C.K., Willemoe, G.L. & Flyger, H.L. 2008, "Intraoperative immunohistochemistry staining of sentinel nodes in breast cancer: clinical and economical implications", *Breast (Edinburgh, Scotland),* vol. 17, no. 4, pp. 372-375.

Ho-Pun-Cheung, A., Bascoul-Mollevi, C., Assenat, E., Boissiere-Michot, F., Bibeau, F., Cellier, D., Ychou, M. & Lopez-Crapez, E. 2009, "Reverse transcription-quantitative polymerase chain reaction: description of a RIN-based algorithm for accurate data normalization", *BMC molecular biology,* vol. 10, pp. 31.

Horlings, H.M., Lai, C., Nuyten, D.S., Halfwerk, H., Kristel, P., van Beers, E., Joosse, S.A., Klijn, C., Nederlof, P.M., Reinders, M.J., Wessels, L.F. & van de Vijver, M.J. 2010, "Integration of DNA copy number alterations and prognostic gene expression signatures in breast cancer patients", *Clinical cancer research : an official journal of the American Association for Cancer Research*, vol. 16, no. 2, pp. 651-663.

Hsi, E.D. 2001, "A practical approach for evaluating new antibodies in the clinical immunohistochemistry laboratory", *Archives of Pathology & Laboratory Medicine*, vol. 125, no. 2, pp. 289-294.

Hu, X., Stern, H.M., Ge, L., O'Brien, C., Haydu, L., Honchell, C.D., Haverty, P.M., Peters, B.A., Wu, T.D., Amler, L.C., Chant, J., Stokoe, D., Lackner, M.R. & Cavet, G. 2009, "Genetic alterations and oncogenic pathways associated with breast cancer subtypes", *Molecular cancer research : MCR*, vol. 7, no. 4, pp. 511-522.

Hudelist, G., Singer, C.F., Pischinger, K.I., Kaserer, K., Manavi, M., Kubista, E. & Czerwenka, K.F. 2006, "Proteomic analysis in human breast cancer: identification of a characteristic protein expression profile of malignant breast epithelium", *Proteomics*, vol. 6, no. 6, pp. 1989-2002.

Hunt, N.C., Attanoos, R. & Jasani, B. 1996, "High temperature antigen retrieval and loss of nuclear morphology: a comparison of microwave and autoclave techniques", *Journal of clinical pathology*, vol. 49, no. 9, pp. 767-770.

Hutson, S.W., Cowen, P.N. & Bird, C.C. 1985, "Morphometric studies of age related changes in normal human breast and their significance for evolution of mammary cancer", *Journal of clinical pathology*, vol. 38, no. 3, pp. 281-287.

Hwang-Verslues, W.W., Kuo, W.H., Chang, P.H., Pan, C.C., Wang, H.H., Tsai, S.T., Jeng, Y.M., Shew, J.Y., Kung, J.T., Chen, C.H., Lee, E.Y., Chang, K.J. & Lee, W.H. 2009, "Multiple lineages of human breast cancer stem/progenitor cells identified by profiling with stem cell markers", *PloS one,* vol. 4, no. 12, pp. e8377.

Hyman, E., Kauraniemi, P., Hautaniemi, S., Wolf, M., Mousses, S., Rozenblum, E., Ringner, M., Sauter, G., Monni, O., Elkahloun, A., Kallioniemi, O.P. & Kallioniemi, A. 2002, "Impact of DNA amplification on gene expression patterns in breast cancer", *Cancer research,* vol. 62, no. 21, pp. 6240-6245.

Ilyin, S.E., Belkowski, S.M. & Plata-Salaman, C.R. 2004, "Biomarker discovery and validation: technologies and integrative approaches", *Trends in biotechnology,* vol. 22, no. 8, pp. 411-416.

Iwasaki, M. & Tsugane, S. 2011, "Risk factors for breast cancer: epidemiological evidence from Japanese studies", *Cancer science,* vol. 102, no. 9, pp. 1607-1614.

Iwata, H. 2007, "Perspective of trastuzumab treatment", *Breast cancer (Tokyo, Japan),* vol. 14, no. 2, pp. 150-155.

Jang, S.H., Choi, C., Hong, S.G., Yarishkin, O.V., Bae, Y.M., Kim, J.G., O'Grady, S.M., Yoon, K.A., Kang, K.S., Ryu, P.D. & Lee, S.Y. 2009, "Silencing of Kv4.1 potassium channels inhibits cell proliferation of tumorigenic human mammary epithelial cells", *Biochemical and biophysical research communications,* vol. 384, no. 2, pp. 180-186.

Jemal, A., Bray, F., Center, M.M., Ferlay, J., Ward, E. & Forman, D. 2011, "Global cancer statistics", *CA: a cancer journal for clinicians,* vol. 61, no. 2, pp. 69-90.

Jeter, C.R., Badeaux, M., Choy, G., Chandra, D., Patrawala, L., Liu, C., Calhoun-Davis, T., Zaehres, H., Daley, G.Q. & Tang, D.G. 2009, "Functional evidence that the self-renewal gene NANOG regulates human tumor development", *Stem cells (Dayton, Ohio)*, vol. 27, no. 5, pp. 993-1005.

Jiang, X.P., Elliott, R.L. & Head, J.F. 2010, "Manipulation of iron transporter genes results in the suppression of human and mouse mammary adenocarcinomas", *Anticancer Research*, vol. 30, no. 3, pp. 759-765.

Jonsson, G., Staaf, J., Vallon-Christersson, J., Ringner, M., Holm, K., Hegardt, C., Gunnarsson, H., Fagerholm, R., Strand, C., Agnarsson, B.A., Kilpivaara, O., Luts, L., Heikkila, P., Aittomaki, K., Blomqvist, C., Loman, N., Malmstrom, P., Olsson, H., Johannsson, O.T., Arason, A., Nevanlinna, H., Barkardottir, R.B. & Borg, A. 2010, "Genomic subtypes of breast cancer identified by array-comparative genomic hybridization display distinct molecular and clinical characteristics", *Breast cancer research : BCR*, vol. 12, no. 3, pp. R42.

Joosse, S.A., van Beers, E.H., Tielen, I.H., Horlings, H., Peterse, J.L., Hoogerbrugge, N., Ligtenberg, M.J., Wessels, L.F., Axwijk, P., Verhoef, S., Hogervorst, F.B. & Nederlof, P.M. 2009, "Prediction of BRCA1-association in hereditary non-BRCA1/2 breast carcinomas with array-CGH", *Breast cancer research and treatment*, vol. 116, no. 3, pp. 479-489.

Kalnina, Z., Zayakin, P., Silina, K. & Line, A. 2005, "Alterations of pre-mRNA splicing in cancer", *Genes, chromosomes & cancer*, vol. 42, no. 4, pp. 342-357.

Kang, J.S., Calvo, B.F., Maygarden, S.J., Caskey, L.S., Mohler, J.L. & Ornstein, D.K. 2002, "Dysregulation of annexin I protein expression in high-grade prostatic intraepithelial neoplasia and prostate cancer", *Clinical cancer research : an official journal of the American Association for Cancer Research,* vol. 8, no. 1, pp. 117-123.

Karp, C.M., Shukla, M.N., Buckley, D.J. & Buckley, A.R. 2007, "HRPAP20: a novel calmodulin-binding protein that increases breast cancer cell invasion", *Oncogene,* vol. 26, no. 12, pp. 1780-1788.

Kehrer-Sawatzki, H. 2007, "Relevance of Copy Number Variation to Human Genetic Disease", *Encyclopedia of Life Sciences,* , pp. 1-6.

Keller, E. & Goldman, R.D. 2006, "Light microscopy" in *Basic Methods in Microscopy: Protocols and Concepts from Cells: A Laboratory Manual*, eds. D.L. Spector & R.D. Goldman, Cold Spring Harbor Laboratory Press, China, pp. 1-42.

Kenemans, P., Verstraeten, R.A. & Verheijen, R.H. 2004, "Oncogenic pathways in hereditary and sporadic breast cancer", *Maturitas,* vol. 49, no. 1, pp. 34-43.

Key, T.J., Verkasalo, P.K. & Banks, E. 2001, "Epidemiology of breast cancer", *The lancet oncology,* vol. 2, no. 3, pp. 133-140.

Killeen, A.A. 2004, "Molecular origins of cancer" in *Principles of Molecular Pathology* Humana Press, The United States of America, pp. 185-198.

Kim, J., Chu, J., Shen, X., Wang, J. & Orkin, S.H. 2008, "An extended transcriptional network for pluripotency of embryonic stem cells", *Cell,* vol. 132, no. 6, pp. 1049-1061.

Klonisch, T., Wiechec, E., Hombach-Klonisch, S., Ande, S.R., Wesselborg, S., Schulze-Osthoff, K. & Los, M. 2008, "Cancer stem cell markers in common cancers - therapeutic implications", *Trends in molecular medicine,* vol. 14, no. 10, pp. 450-460.

Komaki, K., Sano, N. & Tangoku, A. 2006, "Problems in histological grading of malignancy and its clinical significance in patients with operable breast cancer", *Breast cancer (Tokyo, Japan),* vol. 13, no. 3, pp. 249-253.

Konigshoff, M., Wilhelm, J., Bohle, R.M., Pingoud, A. & Hahn, M. 2003, "HER-2/neu gene copy number quantified by real-time PCR: comparison of gene amplification, heterozygosity, and immunohistochemical status in breast cancer tissue", *Clinical chemistry,* vol. 49, no. 2, pp. 219-229.

Korbel, J.O., Kim, P.M., Chen, X., Urban, A.E., Weissman, S., Snyder, M. & Gerstein, M.B. 2008, "The current excitement about copy-number variation: how it relates to gene duplications and protein families", *Current opinion in structural biology,* vol. 18, no. 3, pp. 366-374.

Korkola, J.E., Houldsworth, J., Chadalavada, R.S., Olshen, A.B., Dobrzynski, D., Reuter, V.E., Bosl, G.J. & Chaganti, R.S. 2006, "Down-regulation of stem cell genes, including those in a 200-kb gene cluster at 12p13.31, is associated with in vivo differentiation of human male germ cell tumors", *Cancer research,* vol. 66, no. 2, pp. 820-827.

Kroese, M., Zimmern, R.L. & Pinder, S.E. 2007, "HER2 status in breast cancer--an example of pharmacogenetic testing", *Journal of the Royal Society of Medicine,* vol. 100, no. 7, pp. 326-329.

Kulka, J., Tokes, A.M., Kaposi-Novak, P., Udvarhelyi, N., Keller, A. & Schaff, Z. 2006, "Detection of HER-2/neu gene amplification in breast carcinomas using quantitative real-time PCR - a comparison with immunohistochemical and FISH results", *Pathology oncology research : POR*, vol. 12, no. 4, pp. 197-204.

Kuroda, T., Tada, M., Kubota, H., Kimura, H., Hatano, S.Y., Suemori, H., Nakatsuji, N. & Tada, T. 2005, "Octamer and Sox elements are required for transcriptional cis regulation of Nanog gene expression", *Molecular and cellular biology*, vol. 25, no. 6, pp. 2475-2485.

Kuru, B. 2011, "Prognostic factors for locoregional recurrence and survival in stage IIIC breast carcinoma: impact of adjuvant radiotherapy", *Singapore medical journal*, vol. 52, no. 4, pp. 289-298.

Lambein, K., Praet, M., Forsyth, R., Van den Broecke, R., Braems, G., Matthys, B., Cocquyt, V., Denys, H., Pauwels, P. & Libbrecht, L. 2011, "Relationship between pathological features, HER2 protein expression and HER2 and CEP17 copy number in breast cancer: biological and methodological considerations", *Journal of clinical pathology*, vol. 64, no. 3, pp. 200-207.

Lasa, A., Ramon y Cajal, T., Llort, G., Suela, J., Cigudosa, J.C., Cornet, M., Alonso, C., Barnadas, A. & Baiget, M. 2010, "Copy number variations are not modifiers of phenotypic expression in a pair of identical twins carrying a BRCA1 mutation", *Breast cancer research and treatment*, vol. 123, no. 3, pp. 901-905.

Laurent, L.C., Ulitsky, I., Slavin, I., Tran, H., Schork, A., Morey, R., Lynch, C., Harness, J.V., Lee, S., Barrero, M.J., Ku, S., Martynova, M., Semechkin, R., Galat, V., Gottesfeld, J., Izpisua Belmonte, J.C., Murry, C., Keirstead, H.S., Park, H.S., Schmidt, U., Laslett, A.L., Muller, F.J., Nievergelt, C.M., Shamir, R. & Loring, J.F. 2011, "Dynamic changes in the copy number of pluripotency and cell proliferation genes in human ESCs and iPSCs during reprogramming and time in culture", *Cell stem cell*, vol. 8, no. 1, pp. 106-118.

Leake, R., Barnes, D., Pinder, S., Ellis, I., Anderson, L., Anderson, T., Adamson, R., Rhodes, T., Miller, K. & Walker, R. 2000, "Immunohistochemical detection of steroid receptors in breast cancer: a working protocol. UK Receptor Group, UK NEQAS, The Scottish Breast Cancer Pathology Group, and The Receptor and Biomarker Study Group of the EORTC", *Journal of clinical pathology*, vol. 53, no. 8, pp. 634-635.

Lee, A.H. 2005, "Why is carcinoma of the breast more frequent in the upper outer quadrant? A case series based on needle core biopsy diagnoses", *Breast (Edinburgh, Scotland)*, vol. 14, no. 2, pp. 151-152.

Lee, S.M., Park, J.H. & Park, H.J. 2008, "Breast cancer risk factors in Korean women: a literature review", *International nursing review*, vol. 55, no. 3, pp. 355-359.

Leis, O., Eguiara, A., Lopez-Arribillaga, E., Alberdi, M.J., Hernandez-Garcia, S., Elorriaga, K., Pandiella, A., Rezola, R. & Martin, A.G. 2012, "Sox2 expression in breast tumours and activation in breast cancer stem cells", *Oncogene*, vol. 31, no. 11, pp. 1354-1365.

Lerebours, F. & Lidereau, R. 2002, "Molecular alterations in sporadic breast cancer", *Critical reviews in oncology/hematology*, vol. 44, no. 2, pp. 121-141.

Lester, S.C. 2005, "The Breast" in *Robbins and Cotran Pathologic Basis of Disease*, eds. V. Kumar, A.K. Abbas & N. Fausto, 7th edn, Elsevier Saunders, China, pp. 1,119-1,154.

Letessier, A., Sircoulomb, F., Ginestier, C., Cervera, N., Monville, F., Gelsi-Boyer, V., Esterni, B., Geneix, J., Finetti, P., Zemmour, C., Viens, P., Charafe-Jauffret, E., Jacquemier, J., Birnbaum, D. & Chaffanet, M. 2006, "Frequency, prognostic impact, and subtype association of 8p12, 8q24, 11q13, 12p13, 17q12, and 20q13 amplifications in breast cancers", *BMC cancer,* vol. 6, pp. 245.

Levine, A.J. & Brivanlou, A.H. 2006, "GDF3, a BMP inhibitor, regulates cell fate in stem cells and early embryos", *Development (Cambridge, England),* vol. 133, no. 2, pp. 209-216.

Li, J., Wang, K., Li, S., Timmermans-Wielenga, V., Rank, F., Wiuf, C., Zhang, X., Yang, H. & Bolund, L. 2009, "DNA copy number aberrations in breast cancer by array comparative genomic hybridization", *Genomics, proteomics & bioinformatics / Beijing Genomics Institute,* vol. 7, no. 1-2, pp. 13-24.

Liang, Y., Brekken, R.A. & Hyder, S.M. 2006, "Vascular endothelial growth factor induces proliferation of breast cancer cells and inhibits the anti-proliferative activity of anti-hormones", *Endocrine-related cancer,* vol. 13, no. 3, pp. 905-919.

Lichtenstein, P., Holm, N.V., Verkasalo, P.K., Iliadou, A., Kaprio, J., Koskenvuo, M., Pukkala, E., Skytthe, A. & Hemminki, K. 2000, "Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland", *The New England journal of medicine,* vol. 343, no. 2, pp. 78-85.

Lim, E., Vaillant, F., Wu, D., Forrest, N.C., Pal, B., Hart, A.H., Asselin-Labat, M.L., Gyorki, D.E., Ward, T., Partanen, A., Feleppa, F., Huschtscha, L.I., Thorne, H.J., kConFab, Fox, S.B., Yan, M., French, J.D., Brown, M.A., Smyth, G.K., Visvader, J.E. & Lindeman, G.J. 2009, "Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers", *Nature medicine,* vol. 15, no. 8, pp. 907-913.

Lin, C., Wu, Z., Lin, X., Yu, C., Shi, T., Zeng, Y., Wang, X., Li, J. & Song, L. 2011, "Knockdown of FLOT1 impairs cell proliferation and tumorigenicity in breast cancer through upregulation of FOXO3a", *Clinical cancer research : an official journal of the American Association for Cancer Research,* vol. 17, no. 10, pp. 3089-3099.

Lin, C.H., Liau, J.Y., Lu, Y.S., Huang, C.S., Lee, W.C., Kuo, K.T., Shen, Y.C., Kuo, S.H., Lan, C., Liu, J.M., Kuo, W.H., Chang, K.J. & Cheng, A.L. 2009, "Molecular subtypes of breast cancer emerging in young women in Taiwan: evidence for more than just westernization as a reason for the disease in Asia", *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology,* vol. 18, no. 6, pp. 1807-1814.

Lin, F. & Shi, J. 2011, "Technique and troubleshooting of antibody testing" in *Handbook of Practical Immunohistochemistry: Frequently Asked Questions*, eds. F. Lin & J. Prichard, 1st edn, Springer, USA, pp. 17-21.

Lin, L., Holbro, T., Alonso, G., Gerosa, D. & Burger, M.M. 2001, "Molecular interaction between human tumor marker protein p150, the largest subunit of eIF3, and intermediate filament protein K7", *Journal of cellular biochemistry*, vol. 80, no. 4, pp. 483-490.

Lin, T., Ding, Y.Q. & Li, J.M. 2011, "Overexpression of Nanog protein is associated with poor prognosis in gastric adenocarcinoma", *Medical oncology (Northwood, London, England)*, .

Lindeman, G.J. & Visvader, J.E. 2010, "Insights into the cell of origin in breast cancer and breast cancer stem cells", *Asia-Pacific journal of clinical oncology*, vol. 6, no. 2, pp. 89-97.

Liu, J.H., Wei, S., Ussery, D., Epling-Burnette, P.K., Leonard, W.J. & Djeu, J.Y. 1994, "Expression of interleukin-2 receptor gamma chain on human neutrophils", *Blood*, vol. 84, no. 11, pp. 3870-3875.

Liu, S., Dontu, G. & Wicha, M.S. 2005, "Mammary stem cells, self-renewal pathways, and carcinogenesis", *Breast cancer research : BCR*, vol. 7, no. 3, pp. 86-95.

Longacre, T.A. & Bartow, S.A. 1986, "A correlative morphologic study of human breast and endometrium in the menstrual cycle", *The American Journal of Surgical Pathology*, vol. 10, no. 6, pp. 382-393.

Lopez-Garcia, M.A., Geyer, F.C., Lacroix-Triki, M., Marchio, C. & Reis-Filho, J.S. 2010, "Breast cancer precursors revisited: molecular features and progression pathways", *Histopathology,* vol. 57, no. 2, pp. 171-192.

Ludwig, J.A. & Weinstein, J.N. 2005, "Biomarkers in cancer staging, prognosis and treatment selection", *Nature reviews.Cancer,* vol. 5, no. 11, pp. 845-856.

Luo, J., Yin, X., Ma, T. & Lu, J. 2010, "Stem cells in normal mammary gland and breast cancer", *The American Journal of the Medical Sciences,* vol. 339, no. 4, pp. 366-370.

Lynch, H.T., Lynch, J., Conway, T., Watson, P., Feunteun, J., Lenoir, G., Narod, S. & Fitzgibbons, R.,Jr 1994, "Hereditary breast cancer and family cancer syndromes", *World journal of surgery,* vol. 18, no. 1, pp. 21-31.

Lynch, H.T. & Lynch, J.F. 1996, "Breast cancer genetics: family history, heterogeneity, molecular genetic diagnosis, and genetic counselling", *Current problems in cancer,* vol. 20, no. 6, pp. 329-365.

Lynch, M.D., Cariati, M. & Purushotham, A.D. 2006, "Breast cancer, stem cells and prospects for therapy", *Breast cancer research : BCR,* vol. 8, no. 3, pp. 211.

Ma, X.J., Salunga, R., Tuggle, J.T., Gaudet, J., Enright, E., McQuary, P., Payette, T., Pistone, M., Stecker, K., Zhang, B.M., Zhou, Y.X., Varnholt, H., Smith, B., Gadd, M., Chatfield, E., Kessler, J., Baer, T.M., Erlander, M.G. & Sgroi, D.C. 2003, "Gene expression profiles of human breast cancer progression", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 100, no. 10, pp. 5974-5979.

MacLean, J. 2004, *Breast Cancer in California: A Closer Look*, California Breast Cancer Research Program, USA.

Madsen, M.W. & Briand, P. 1990, "Relationship between tumorigenicity, in vitro invasiveness, and plasminogen activator production of human breast cell lines", *European journal of cancer (Oxford, England : 1990)*, vol. 26, no. 7, pp. 793-797.

Malkas, L.H., Herbert, B.S., Abdel-Aziz, W., Dobrolecki, L.E., Liu, Y., Agarwal, B., Hoelz, D., Badve, S., Schnaper, L., Arnold, R.J., Mechref, Y., Novotny, M.V., Loehrer, P., Goulet, R.J. & Hickey, R.J. 2006, "A cancer-associated PCNA expressed in breast cancer has implications as a potential biomarker", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 51, pp. 19472-19477.

Man, T.K., Lu, X.Y., Jaeweon, K., Perlaky, L., Harris, C.P., Shah, S., Ladanyi, M., Gorlick, R., Lau, C.C. & Rao, P.H. 2004, "Genome-wide array comparative genomic hybridization analysis reveals distinct amplifications in osteosarcoma", *BMC cancer*, vol. 4, pp. 45.

Mann, M. & Jensen, O.N. 2003, "Proteomic analysis of post-translational modifications", *Nature biotechnology*, vol. 21, no. 3, pp. 255-261.

Mark, H.F., Brown, S., Taylor, W., Bassily, N., Sun, C.L., Samy, M. & Bland, K.I. 1999, "Study of chromosome 12 copy number in breast cancer using fluorescence in situ hybridization", *Cancer genetics and cytogenetics*, vol. 108, no. 1, pp. 26-31.

Maruvada, P., Wang, W., Wagner, P.D. & Srivastava, S. 2005, "Biomarkers in molecular medicine: cancer detection and diagnosis", *BioTechniques*, vol. Suppl, pp. 9-15.

Masters, J.R. & Stacey, G.N. 2007, "Changing medium and passaging cell lines", *Nature protocols*, vol. 2, no. 9, pp. 2276-2284.

Mathis, K.L., Hoskin, T.L., Boughey, J.C., Crownhart, B.S., Brandt, K.R., Vachon, C.M., Grant, C.S. & Degnim, A.C. 2010, "Palpable presentation of breast cancer persists in the era of screening mammography", *Journal of the American College of Surgeons*, vol. 210, no. 3, pp. 314-318.

Matlin, A.J., Clark, F. & Smith, C.W. 2005, "Understanding alternative splicing: towards a cellular code", *Nature reviews.Molecular cell biology*, vol. 6, no. 5, pp. 386-398.

Mayeux, R. 2004, "Biomarkers: potential uses and limitations", *NeuroRx : the journal of the American Society for Experimental NeuroTherapeutics*, vol. 1, no. 2, pp. 182-188.

Mayshar, Y., Ben-David, U., Lavon, N., Biancotti, J.C., Yakir, B., Clark, A.T., Plath, K., Lowry, W.E. & Benvenisty, N. 2010, "Identification and classification of chromosomal aberrations in human induced pluripotent stem cells", *Cell stem cell*, vol. 7, no. 4, pp. 521-531.

McCarty, K.S.,Jr, Miller, L.S., Cox, E.B., Konrath, J. & McCarty KS, S. 1985, "Estrogen receptor analyses. Correlation of biochemical and immunohistochemical methods using monoclonal antireceptor antibodies", *Archives of Pathology & Laboratory Medicine,* vol. 109, no. 8, pp. 716-721.

McCarty, K.S.,Jr, Szabo, E., Flowers, J.L., Cox, E.B., Leight, G.S., Miller, L., Konrath, J., Soper, J.T., Budwit, D.A. & Creasman, W.T. 1986, "Use of a monoclonal anti-estrogen receptor antibody in the immunohistochemical evaluation of human tumors", *Cancer research,* vol. 46, no. 8 Suppl, pp. 4244s-4248s.

McGuire, W.L., Tandon, A.K., A.D., Chamness, G.C. & Clark, G.M. 1990, " How to use prognostic factors in axillary node-negative breast cancer patients", *Journal of The National Cancer Institute,* vol. 82, no. 12, pp. 1,006-1,015.

McKeage, K. & Lyseng-Williamson, K.A. 2008, "Trastuzumab: a pharmacoeconomic review of its use in early breast cancer", *PharmacoEconomics,* vol. 26, no. 8, pp. 699-719.

McKeon, F. 2004, "P63 and the Epithelial Stem Cell: More than Status Quo?", *Genes & development,* vol. 18, no. 5, pp. 465-469.

Melchor, L. & Benitez, J. 2008, "An integrative hypothesis about the origin and development of sporadic and familial breast cancer subtypes", *Carcinogenesis,* vol. 29, no. 8, pp. 1475-1482.

Meng, H.M., Zheng, P., Wang, X.Y., Liu, C., Sui, H.M., Wu, S.J., Zhou, J., Ding, Y.Q. & Li, J.M. 2010, "Overexpression of nanog predicts tumor progression and poor prognosis in colorectal cancer", *Cancer biology & therapy,* vol. 9, no. 4.

Meyer, J.S., Alvarez, C., Milikowski, C., Olson, N., Russo, I., Russo, J., Glass, A., Zehnbauer, B.A., Lister, K., Parwaresch, R. & Cooperative Breast Cancer Tissue Resource 2005, "Breast carcinoma malignancy grading by Bloom-Richardson system vs proliferation index: reproducibility of grade and advantages of proliferation index", *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc*, vol. 18, no. 8, pp. 1067-1078.

Meyerson, M., Gabriel, S. & Getz, G. 2010, "Advances in understanding cancer genomes through second-generation sequencing", *Nature reviews.Genetics,* vol. 11, no. 10, pp. 685-696.

Miller, C.T., Aggarwal, S., Lin, T.K., Dagenais, S.L., Contreras, J.I., Orringer, M.B., Glover, T.W., Beer, D.G. & Lin, L. 2003, "Amplification and overexpression of the dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 2 (DYRK2) gene in esophageal and lung adenocarcinomas", *Cancer research,* vol. 63, no. 14, pp. 4136-4143.

Mischel, P.S., Shai, R., Shi, T., Horvath, S., Lu, K.V., Choe, G., Seligson, D., Kremen, T.J., Palotie, A., Liau, L.M., Cloughesy, T.F. & Nelson, S.F. 2003, "Identification of molecular subtypes of glioblastoma by gene expression profiling", *Oncogene,* vol. 22, no. 15, pp. 2361-2373.

Mishra, A. & Verma, M. 2010, "Cancer biomarkers: are we ready for the prime time?", *Cancers,* vol. 2, pp. 190-208.

Moore, C. 2009, *Introduction to Western Blotting,* AbD Serotec, The United Kingdom.

Morabito, A., Magnani, E., Gion, M., Sarmiento, R., Capaccetti, B., Longo, R., Gattuso, D. & Gasparini, G. 2003, "Prognostic and predictive indicators in operable breast cancer", *Clinical breast cancer,* vol. 3, no. 6, pp. 381-390.

Morishita, Y., Tsuda, H., Fukutomi, T., Mukai, K., Shimosato, Y. & Hirohashi, S. 1997, "Clinicopathological characteristics of primary breast cancer in older geriatric women: a study of 39 Japanese patients over 80 years old", *Japanese journal of cancer research : Gann,* vol. 88, no. 7, pp. 693-699.

Morrogh, M., Olvera, N., Bogomolniy, F., Borgen, P.I. & King, T.A. 2007, "Tissue preparation for laser capture microdissection and RNA extraction from fresh frozen breast tissue", *BioTechniques,* vol. 43, no. 1, pp. 41-2, 44, 46 passim.

Mrhalova, M., Kodet, R., Kalinova, M. & Hilska, I. 2003, "Relative quantification of ERBB2 mRNA in invasive duct carcinoma of the breast: correlation with ERBB-2 protein expression and ERBB2 gene copy number", *Pathology, research and practice,* vol. 199, no. 7, pp. 453-461.

Nabi, H.A. & Zubeldia, J.M. 2002, "Clinical applications of (18)F-FDG in oncology", *Journal of nuclear medicine technology,* vol. 30, no. 1, pp. 3-9; quiz 10-1.

Nass, S.J. & Moses, H.L. (eds) 2007, *Cancer Biomarkers: The Promises and Challenges of Improving Detection and Treatment*, First edn, The National Academies Press, USA.

Natrajan, R., Lambros, M.B., Rodriguez-Pinilla, S.M., Moreno-Bueno, G., Tan, D.S., Marchio, C., Vatcheva, R., Rayter, S., Mahler-Araujo, B., Fulford, L.G., Hungermann, D., Mackay, A., Grigoriadis, A., Fenwick, K., Tamber, N., Hardisson, D., Tutt, A., Palacios, J., Lord, C.J., Buerger, H., Ashworth, A. & Reis-Filho, J.S. 2009, "Tiling path genomic profiling of grade 3 invasive ductal breast cancers", *Clinical cancer research : an official journal of the American Association for Cancer Research*, vol. 15, no. 8, pp. 2711-2722.

Nessling, M., Richter, K., Schwaenen, C., Roerig, P., Wrobel, G., Wessendorf, S., Fritz, B., Bentz, M., Sinn, H.P., Radlwimmer, B. & Lichter, P. 2005, "Candidate genes in breast cancer revealed by microarray-based comparative genomic hybridization of archived tissue", *Cancer research*, vol. 65, no. 2, pp. 439-447.

Nicholson, R.I., Bouzubar, N., Walker, K.J., McClelland, R., Dixon, A.R., Robertson, J.F., Ellis, I.O. & Blamey, R.W. 1991, "Hormone sensitivity in breast cancer: influence of heterogeneity of oestrogen receptor expression and cell proliferation", *European journal of cancer (Oxford, England : 1990)*, vol. 27, no. 7, pp. 908-913.

Nieweg, O.E., Estourgie, S.H., Deurloo, E.E., Rutgers, E.J. & Kroon, B.B. 2002, "Status of lymph node staging", *Scandinavian journal of surgery : SJS : official organ for the Finnish Surgical Society and the Scandinavian Surgical Society*, vol. 91, no. 3, pp. 263-267.

Nimeus, E., Malmstrom, J., Johnsson, A., Marko-Varga, G. & Ferno, M. 2007, "Proteomic analysis identifies candidate proteins associated with distant recurrences in breast cancer after adjuvant chemotherapy", *Journal of pharmaceutical and biomedical analysis*, vol. 43, no. 3, pp. 1086-1093.

Nowak, D., Hofmann, W.K. & Koeffler, H.P. 2009, "Genome-wide Mapping of Copy Number Variations Using SNP Arrays", *Transfusion medicine and hemotherapy : offizielles Organ der Deutschen Gesellschaft fur Transfusionsmedizin und Immunhamatologie,* vol. 36, no. 4, pp. 246-251.

Ntoulia, M., Kaklamanis, L., Valavanis, C., Kafousi, M., Stathopoulos, E., Arapantoni, P., Mavroudis, D., Georgoulias, V. & Lianidou, E.S. 2006, "HER-2 DNA quantification of paraffin-embedded breast carcinomas with LightCycler real-time PCR in comparison to immunohistochemistry and chromogenic in situ hybridization", *Clinical biochemistry,* vol. 39, no. 9, pp. 942-946.

Oehadian, A. & Fadjari, T.H. 2008, "Neutropenic enterocolitis in breast cancer patient after taxane-containing chemotherapy", *Acta medica Indonesiana,* vol. 40, no. 1, pp. 29-33.

Ohta, H., Sawada, A., Kim, J.Y., Tokimasa, S., Nishiguchi, S., Humphries, R.K., Hara, J. & Takihara, Y. 2002, "Polycomb group gene rae28 is required for sustaining activity of hematopoietic stem cells", *The Journal of experimental medicine,* vol. 195, no. 6, pp. 759-770.

O'Leary, J.J., Sheils, O., Martin, C. & Crowley, A. 2003, "TaqMan® technology and real-time polymerase chain reaction" in *Molecular Biology in Cellular Pathology*, eds. J. Crocker & P.G. Murray, 2nd edn, Wiley, Great Britain, pp. 251-265.

Olopade, O.I., Grushko, T.A., Nanda, R. & Huo, D. 2008, "Advances in breast cancer: pathways to personalized medicine", *Clinical cancer research : an official journal of the American Association for Cancer Research,* vol. 14, no. 24, pp. 7988-7999.

O'Malley, F.P. 2010, "Lobular neoplasia: morphology, biological potential and management in core biopsies", *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc,* vol. 23 Suppl 2, pp. S14-25.

Opitz, L., Salinas-Riester, G., Grade, M., Jung, K., Jo, P., Emons, G., Ghadimi, B.M., Beissbarth, T. & Gaedcke, J. 2010, "Impact of RNA degradation on gene expression profiling", *BMC medical genomics,* vol. 3, pp. 36.

O'Shaughnessy, J.A., Kelloff, G.J., Gordon, G.B., Dannenberg, A.J., Hong, W.K., Fabian, C.J., Sigman, C.C., Bertagnolli, M.M., Stratton, S.P., Lam, S., Nelson, W.G., Meyskens, F.L., Alberts, D.S., Follen, M., Rustgi, A.K., Papadimitrakopoulou, V., Scardino, P.T., Gazdar, A.F., Wattenberg, L.W., Sporn, M.B., Sakr, W.A., Lippman, S.M. & Von Hoff, D.D. 2002, "Treatment and prevention of intraepithelial neoplasia: an important target for accelerated new agent development", *Clinical cancer research : an official journal of the American Association for Cancer Research,* vol. 8, no. 2, pp. 314-346.

Otali, D., Stockard, C.R., Oelschlager, D.K., Wan, W., Manne, U., Watts, S.A. & Grizzle, W.E. 2009, "Combined effects of formalin fixation and tissue processing on immunorecognition", *Biotechnic & histochemistry : official publication of the Biological Stain Commission,* vol. 84, no. 5, pp. 223-247.

Page, K., Hava, N., Ward, B., Brown, J., Guttery, D.S., Ruangpratheep, C., Blighe, K., Sharma, A., Walker, R.A., Coombes, R.C. & Shaw, J.A. 2011, "Detection of HER2 amplification in circulating free DNA in patients with breast cancer", *British journal of cancer,* vol. 104, no. 8, pp. 1342-1348.

Pain, D., Chirn, G.W., Strassel, C. & Kemp, D.M. 2005, "Multiple retropseudogenes from pluripotent cell-specific gene expression indicates a potential signature for novel gene identification", *The Journal of biological chemistry,* vol. 280, no. 8, pp. 6265-6268.

Pandya, S. & Moore, R.G. 2011, "Breast development and anatomy", *Clinical obstetrics and gynecology,* vol. 54, no. 1, pp. 91-95.

Papelard, H., de Bock, G.H., van Eijk, R., Vliet Vlieland, T.P., Cornelisse, C.J., Devilee, P. & Tollenaar, R.A. 2000, "Prevalence of BRCA1 in a hospital-based population of Dutch breast cancer patients", *British journal of cancer,* vol. 83, no. 6, pp. 719-724.

Parham, D.M., Hagen, N. & Brown, R.A. 1992, "Simplified method of grading primary carcinomas of the breast", *Journal of clinical pathology,* vol. 45, no. 6, pp. 517-520.

Parsa, P. & Parsa, B. 2009, "Effects of reproductive factors on risk of breast cancer: a literature review", *Asian Pacific journal of cancer prevention : APJCP,* vol. 10, no. 4, pp. 545-550.

Pascal, L.E., Vencio, R.Z., Goo, Y.A., Page, L.S., Shadle, C.P. & Liu, A.Y. 2009, "Temporal expression profiling of the effects of secreted factors from prostate stromal cells on embryonal carcinoma stem cells", *The Prostate,* vol. 69, no. 12, pp. 1353-1365.

Pasqualucci, L., Trifonov, V., Fabbri, G., Ma, J., Rossi, D., Chiarenza, A., Wells, V.A., Grunn, A., Messina, M., Elliot, O., Chan, J., Bhagat, G., Chadburn, A., Gaidano, G., Mullighan, C.G., Rabadan, R. & Dalla-Favera, R. 2011, "Analysis of the coding genome of diffuse large B-cell lymphoma", *Nature genetics*, vol. 43, no. 9, pp. 830-837.

Pau Ni, I.B., Zakaria, Z., Muhammad, R., Abdullah, N., Ibrahim, N., Aina Emran, N., Hisham Abdullah, N. & Syed Hussain, S.N. 2010, "Gene expression patterns distinguish breast carcinomas from normal breast tissues: the Malaysian context", *Pathology, research and practice*, vol. 206, no. 4, pp. 223-228.

Pauletti, G., Dandekar, S., Rong, H., Ramos, L., Peng, H., Seshadri, R. & Slamon, D.J. 2000, "Assessment of methods for tissue-based detection of the HER-2/neu alteration in human breast cancer: a direct comparison of fluorescence in situ hybridization and immunohistochemistry", *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, vol. 18, no. 21, pp. 3651-3664.

Payer, B., Saitou, M., Barton, S.C., Thresher, R., Dixon, J.P., Zahn, D., Colledge, W.H., Carlton, M.B., Nakano, T. & Surani, M.A. 2003, "Stella is a maternal effect gene required for normal early development in mice", *Current biology : CB*, vol. 13, no. 23, pp. 2110-2117.

Peiro, G., Aranda, F.I., Adrover, E., Niveiro, M., Alenda, C., Paya, A. & Segui, J. 2007, "Analysis of HER2 by chromogenic in situ hybridization and immunohistochemistry in lymph node-negative breast carcinoma: Prognostic relevance", *Human pathology*, vol. 38, no. 1, pp. 26-34.

Perou, C.M., Sorlie, T., Eisen, M.B., van de Rijn, M., Jeffrey, S.S., Rees, C.A., Pollack, J.R., Ross, D.T., Johnsen, H., Akslen, L.A., Fluge, O., Pergamenschikov, A., Williams, C., Zhu, S.X., Lonning, P.E., Borresen-Dale, A.L., Brown, P.O. & Botstein, D. 2000, "Molecular portraits of human breast tumours", *Nature*, vol. 406, no. 6797, pp. 747-752.

Petak, I., Schwab, R., Orfi, L., Kopper, L. & Keri, G. 2010, "Integrating molecular diagnostics into anticancer drug discovery", *Nature reviews.Drug discovery*, vol. 9, no. 7, pp. 523-535.

Petersen, O.W. & Polyak, K. 2010, "Stem cells in the human breast", *Cold Spring Harbor perspectives in biology*, vol. 2, no. 5, pp. a003160.

Peto, J. 2001, "Cancer epidemiology in the last century and the next decade", *Nature*, vol. 411, no. 6835, pp. 390-395.

Pfaffl, M.W. 2004, "Quantification strategies in Real-Time PCR" in *A - Z of Quantitative PCR*, ed. S.A. Bustin, International University Line, The United States of America, pp. 87-120.

Pfaffl, M.W. 2006, "Relative quantification" in *Real-Time PCR*, ed. M.T. Dorak, Taylor & Francis Group, The United Kingdom, pp. 63-82.

Pileri, S.A., Roncador, G., Ceccarelli, C., Piccioli, M., Briskomatis, A., Sabattini, E., Ascani, S., Santini, D., Piccaluga, P.P., Leone, O., Damiani, S., Ercolessi, C., Sandri, F., Pieri, F., Leoncini, L. & Falini, B. 1997, "Antigen retrieval techniques in immunohistochemistry: comparison of different methods", *The Journal of pathology*, vol. 183, no. 1, pp. 116-123.

Polyak, K. 2007, "Breast cancer: origins and evolution", *The Journal of clinical investigation,* vol. 117, no. 11, pp. 3155-3163.

Polyak, K. & Hahn, W.C. 2006, "Roots and stems: stem cells in cancer", *Nature medicine,* vol. 12, no. 3, pp. 296-300.

Potemski, P., Pluciennik, E., Bednarek, A.K., Kusinska, R., Pasz-Walczak, G., Jesionek-Kupnicka, D., Watala, C. & Kordek, R. 2006, "A comparative assessment of HER2 status in operable breast cancer by real-time RT-PCR and by immunohistochemistry", *Medical science monitor : international medical journal of experimental and clinical research,* vol. 12, no. 12, pp. MT57-61.

Prat, A. & Perou, C.M. 2009, "Mammary development meets cancer genomics", *Nature medicine,* vol. 15, no. 8, pp. 842-844.

Qureshi, A. & Pervez, S. 2010, "Allred scoring for ER reporting and it's impact in clearly distinguishing ER negative from ER positive breast cancers", *JPMA.The Journal of the Pakistan Medical Association,* vol. 60, no. 5, pp. 350-353.

Raaphorst, F.M., Vermeer, M., Fieret, E., Blokzijl, T., Dukers, D., Sewalt, R.G., Otte, A.P., Willemze, R. & Meijer, C.J. 2004, "Site-specific expression of polycomb-group genes encoding the HPC-HPH/PRC1 complex in clinically defined primary nodal and cutaneous large B-cell lymphomas", *The American journal of pathology,* vol. 164, no. 2, pp. 533-542.

Rakha, E.A., Reis-Filho, J.S., Baehner, F., Dabbs, D.J., Decker, T., Eusebi, V., Fox, S.B., Ichihara, S., Jacquemier, J., Lakhani, S.R., Palacios, J., Richardson, A.L., Schnitt, S.J., Schmitt, F.C., Tan, P.H., Tse, G.M., Badve, S. & Ellis, I.O. 2010, "Breast cancer prognostic classification in the molecular era: the role of histological grade", *Breast cancer research : BCR*, vol. 12, no. 4, pp. 207.

Recareanu, F., Gruia, C., Simionescu, C. & Stepan, A. 2010, "Current research on breast carcinogenesis", *Current Health Sciences Journal*, vol. 36, no. 1.

Redon, R., Ishikawa, S., Fitch, K.R., Feuk, L., Perry, G.H., Andrews, T.D., Fiegler, H., Shapero, M.H., Carson, A.R., Chen, W., Cho, E.K., Dallaire, S., Freeman, J.L., Gonzalez, J.R., Gratacos, M., Huang, J., Kalaitzopoulos, D., Komura, D., MacDonald, J.R., Marshall, C.R., Mei, R., Montgomery, L., Nishimura, K., Okamura, K., Shen, F., Somerville, M.J., Tchinda, J., Valsesia, A., Woodwark, C., Yang, F., Zhang, J., Zerjal, T., Zhang, J., Armengol, L., Conrad, D.F., Estivill, X., Tyler-Smith, C., Carter, N.P., Aburatani, H., Lee, C., Jones, K.W., Scherer, S.W. & Hurles, M.E. 2006, "Global variation in copy number in the human genome", *Nature*, vol. 444, no. 7118, pp. 444-454.

Reis-Filho, J.S. 2009, "Next-generation sequencing", *Breast cancer research : BCR*, vol. 11 Suppl 3, pp. S12.

Reis-Filho, J.S. & Pusztai, L. 2011, "Gene expression profiling in breast cancer: classification, prognostication, and prediction", *Lancet*, vol. 378, no. 9805, pp. 1812-1823.

Reis-Filho, J.S., Simpson, P.T., Turner, N.C., Lambros, M.B., Jones, C., Mackay, A., Grigoriadis, A., Sarrio, D., Savage, K., Dexter, T., Iravani, M., Fenwick, K., Weber, B., Hardisson, D., Schmitt, F.C., Palacios, J., Lakhani, S.R. & Ashworth, A. 2006, "FGFR1 emerges as a potential therapeutic target for lobular breast carcinomas", *Clinical cancer research : an official journal of the American Association for Cancer Research,* vol. 12, no. 22, pp. 6652-6662.

Rennstam, K., Jonsson, G., Tanner, M., Bendahl, P.O., Staaf, J., Kapanen, A.I., Karhu, R., Baldetorp, B., Borg, A. & Isola, J. 2007, "Cytogenetic characterization and gene expression profiling of the trastuzumab-resistant breast cancer cell line JIMT-1", *Cancer genetics and cytogenetics,* vol. 172, no. 2, pp. 95-106.

Revillion, F., Pawlowski, V., Hornez, L. & Peyrat, J.P. 2000, "Glyceraldehyde-3-phosphate dehydrogenase gene expression in human breast cancer", *European journal of cancer (Oxford, England : 1990),* vol. 36, no. 8, pp. 1038-1042.

Reya, T., Morrison, S.J., Clarke, M.F. & Weissman, I.L. 2001, "Stem cells, cancer, and cancer stem cells", *Nature,* vol. 414, no. 6859, pp. 105-111.

Rhodes, D.R. & Chinnaiyan, A.M. 2005, "Integrative analysis of the cancer transcriptome", *Nature genetics,* vol. 37 Suppl, pp. S31-7.

Rhodes, A., Jasani, B., Balaton, A.J., Barnes, D.M., Anderson, E., Bobrow, L.G. & Miller, K.D. 2001, "Study of interlaboratory reliability and reproducibility of estrogen and progesterone receptor assays in Europe. Documentation of poor reliability and identification of insufficient microwave antigen retrieval time as a major contributory element of unreliable assays", *American Journal of Clinical Pathology,* vol. 115, no. 1, pp. 44-58.

Rifai, N., Gillette, M.A. & Carr, S.A. 2006, "Protein biomarker discovery and validation: the long and uncertain path to clinical utility", *Nature biotechnology*, vol. 24, no. 8, pp. 971-983.

Robbins, C.M., Tembe, W.A., Baker, A., Sinari, S., Moses, T.Y., Beckstrom-Sternberg, S., Beckstrom-Sternberg, J., Barrett, M., Long, J., Chinnaiyan, A., Lowey, J., Suh, E., Pearson, J.V., Craig, D.W., Agus, D.B., Pienta, K.J. & Carpten, J.D. 2011, "Copy number and targeted mutational analysis reveals novel somatic events in metastatic prostate tumors", *Genome research*, vol. 21, no. 1, pp. 47-55.

Robbins, P., Pinder, S., de Klerk, N., Dawkins, H., Harvey, J., Sterrett, G., Ellis, I. & Elston, C. 1995, "Histological grading of breast carcinomas: a study of interobserver agreement", *Human pathology*, vol. 26, no. 8, pp. 873-879.

Robinson, K.M. & Huether, S.E. 2002, "Structure and Function of the Reproductive Systems" in *Pathophysiology: The Biologic Basis for Disease in Adults & Children*, eds. K.L. McCance & S.E. Huether, 4th edn, Mosby, USA, pp. 670-704.

Rodriguez, S., Jafer, O., Goker, H., Summersgill, B.M., Zafarana, G., Gillis, A.J., van Gurp, R.J., Oosterhuis, J.W., Lu, Y.J., Huddart, R., Cooper, C.S., Clark, J., Looijenga, L.H. & Shipley, J.M. 2003, "Expression profile of genes from 12p in testicular germ cell tumors of adolescents and adults associated with i(12p) and amplification at 12p11.2-p12.1", *Oncogene*, vol. 22, no. 12, pp. 1880-1891.

Ross, M.H. & Pawlina, W. 2006, "In dividing cells, chromatin is condensed and organized into discrete bodies called chromosomes" in *Histology: A Text and Atlas with Correlated Cell and Molecular Biology*, 5th edn, Lippincott Williams & Wilkins, The United States of America, pp. 72-73.

Rossi, E., Ubiali, A., Cadei, M., Balzarini, P., Valagussa, E., Lucini, L., Alpi, F., Galletti, A., Fontana, L., Tedoldi, C. & Grigolato, P. 2006, "HER-2/neu in breast cancer: a comparative study between histology, immunohistochemistry, and molecular technique (FISH)", *Applied immunohistochemistry & Molecular Morphology: AIMM/Official Publication of The Society for Applied Immunohistochemistry,* vol. 14, no. 2, pp. 127-131.

Roy, I. & Othieno, E. 2011, "Breast carcinoma in Uganda: microscopic study and receptor profile of 45 cases", *Archives of Pathology & Laboratory Medicine,* vol. 135, no. 2, pp. 194-199.

Rozeman, L.B., Szuhai, K., Schrage, Y.M., Rosenberg, C., Tanke, H.J., Taminiau, A.H., Cleton-Jansen, A.M., Bovee, J.V. & Hogendoorn, P.C. 2006, "Array-comparative genomic hybridization of central chondrosarcoma: identification of ribosomal protein S6 and cyclin-dependent kinase 4 as candidate target genes for genomic aberrations", *Cancer,* vol. 107, no. 2, pp. 380-388.

Ryan, J. & Mariano, J. *Helpful Hints for Better Aseptic Technique*, Bionique Testing Laboratories.

Saber, R.A. 2000, "Assessment of Breast Carcinoma by 99mTC-Sestamibi Scintimammography: Effect of Lesion Site, Size and Histopathological Characteristics", *Journal of the Egyptian National Cancer Institute,* vol. 12, no. 3, pp. 183-189.

Salmina, K., Jankevics, E., Huna, A., Perminov, D., Radovica, I., Klymenko, T., Ivanov, A., Jascenko, E., Scherthan, H., Cragg, M. & Erenpreisa, J. 2010, "Up-regulation of the embryonic self-renewal network through reversible polyploidy in irradiated p53-mutant tumour cells", *Experimental cell research,* vol. 316, no. 13, pp. 2099-2112.

Sanchez-Beato, M., Sanchez, E., Gonzalez-Carrero, J., Morente, M., Diez, A., Sanchez-Verde, L., Martin, M.C., Cigudosa, J.C., Vidal, M. & Piris, M.A. 2006, "Variability in the expression of polycomb proteins in different normal and tumoral tissues. A pilot study using tissue microarrays", *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc,* vol. 19, no. 5, pp. 684-694.

Sasso, M., Bianchi, F., Ciravolo, V., Tagliabue, E. & Campiglio, M. 2011, "HER2 splice variants and their relevance in breast cancer", *Journal of Nucleic Acids Investigation,* vol. 2:e9, pp. 52-59.

Sato, M., Kimura, T., Kurokawa, K., Fujita, Y., Abe, K., Masuhara, M., Yasunaga, T., Ryo, A., Yamamoto, M. & Nakano, T. 2002, "Identification of PGC7, a new gene expressed specifically in preimplantation embryos and germ cells", *Mechanisms of development,* vol. 113, no. 1, pp. 91-94.

Sato, T., Neilson, L.M., Peck, A.R., Liu, C., Tran, T.H., Witkiewicz, A., Hyslop, T., Nevalainen, M.T., Sauter, G. & Rui, H. 2011, "Signal transducer and activator of transcription-3 and breast cancer prognosis", *American journal of cancer research,* vol. 1, no. 3, pp. 347-355.

Schaeffeler, E., Schwab, M., Eichelbaum, M. & Zanger, U.M. 2003, "CYP2D6 genotyping strategy based on gene copy number determination by TaqMan real-time PCR", *Human mutation,* vol. 22, no. 6, pp. 476-485.

Scherer, S.W., Lee, C., Birney, E., Altshuler, D.M., Eichler, E.E., Carter, N.P., Hurles, M.E. & Feuk, L. 2007, "Challenges and standards in integrating surveys of structural variation", *Nature genetics,* vol. 39, no. 7 Suppl, pp. S7-15.

Schmitt, F. 2009, "HER2+ breast cancer: how to evaluate?", *Advances in Therapy,* vol. 26 Suppl 1, pp. S1-8.

Schmittgen, T.D. 2006, "Quantitative gene expression by real-time PCR: a complete protocol" in *Real-Time PCR,* ed. M.T. Dorak, Taylor & Francis Group, The United Kingdom, pp. 127-137.

Schnitt, S.J. 2001, "Traditional and newer pathologic factors", *Journal of the National Cancer Institute.Monographs,* vol. (30), no. 30, pp. 22-26.

Schondorf, T., Engel, H., Lindemann, C., Kolhagen, H., von Rucker, A.A. & Mallmann, P. 1997, "Cellular characteristics of peripheral blood lymphocytes and tumour-infiltrating lymphocytes in patients with gynaecological tumours", *Cancer immunology, immunotherapy : CII,* vol. 44, no. 2, pp. 88-96.

Schwartz, G.F., Hughes, K.S., Lynch, H.T., Fabian, C.J., Fentiman, I.S., Robson, M.E., Domchek, S.M., Hartmann, L.C., Holland, R., Winchester, D.J. & Consensus Conference Committee The International Consensus Conference Committee 2008, "Proceedings of the international consensus conference on breast cancer risk, genetics, & risk management, April, 2007", *Cancer*, vol. 113, no. 10, pp. 2627-2637.

Schweiger, M.R., Kerick, M., Timmermann, B. & Isau, M. 2011, "The power of NGS technologies to delineate the genome organization in cancer: from mutations to structural variations and epigenetic alterations", *Cancer metastasis reviews*, vol. 30, no. 2, pp. 199-210.

Shadeo, A. & Lam, W.L. 2006, "Comprehensive copy number profiles of breast cancer cell model genomes", *Breast cancer research : BCR*, vol. 8, no. 1, pp. R9.

Shah, S. & Chen, B. 2010, "Testing for HER2 in Breast Cancer: A Continuing Evolution", *Pathology research international*, vol. 2011, pp. 903202.

Sharkey, F.E., Allred, D.C. & Valente, P.T. 1996, "Breast" in *Anderson's Pathology*, eds. I. Damjanov & J. Linder, 10th edn, Mosby, USA, pp. 2354-2385.

Shaw, J.A., Page, K., Blighe, K., Hava, N., Guttery, D., Ward, B., Brown, J., Ruangpratheep, C., Stebbing, J., Payne, R., Palmieri, C., Cleator, S., Walker, R.A. & Coombes, R.C. 2011, "Genomic analysis of circulating cell free DNA infers breast cancer dormancy", *Genome research*.

Shayne, M., Culakova, E., Wolff, D., Poniewierski, M.S., Dale, D.C., Crawford, J. & Lyman, G.H. 2009, "Dose intensity and hematologic toxicity in older breast cancer patients receiving systemic chemotherapy", *Cancer*, vol. 115, no. 22, pp. 5319-5328.

Shepard, H.M., Jin, P., Slamon, D.J., Pirot, Z. & Maneval, D.C. 2008, "Herceptin", *Handbook of Experimental Pharmacology*, vol. (181), no. 181, pp. 183-219.

Shet, T., Agrawal, A., Chinoy, R., Havaldar, R., Parmar, V. & Badwe, R. 2007, "Changes in the tumor grade and biological markers in locally advanced breast cancer after chemotherapy--implications for a pathologist", *The breast journal*, vol. 13, no. 5, pp. 457-464.

Shi, S.R., Key, M.E. & Kalra, K.L. 1991, "Antigen retrieval in formalin-fixed, paraffin-embedded tissues: an enhancement method for immunohistochemical staining based on microwave oven heating of tissue sections", *The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society*, vol. 39, no. 6, pp. 741-748.

Shi, S.R., Shi, Y. & Taylor, C.R. 2011, "Antigen retrieval immunohistochemistry: review and future prospects in research and diagnosis over two decades", *The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society*, vol. 59, no. 1, pp. 13-32.

Shin, S.J., Hyjek, E., Early, E. & Knowles, D.M. 2006, "Intratumoral heterogeneity of her-2/neu in invasive mammary carcinomas using fluorescence in-situ hybridization and tissue microarray", *International Journal of Surgical Pathology*, vol. 14, no. 4, pp. 279-284.

Shlien, A. & Malkin, D. 2009, "Copy number variations and cancer", *Genome medicine,* vol. 1, no. 6, pp. 62.

Sihto, H., Lundin, J., Lehtimaki, T., Sarlomo-Rikala, M., Butzow, R., Holli, K., Sailas, L., Kataja, V., Lundin, M., Turpeenniemi-Hujanen, T., Isola, J., Heikkila, P. & Joensuu, H. 2008, "Molecular subtypes of breast cancers detected in mammography screening and outside of screening", *Clinical cancer research : an official journal of the American Association for Cancer Research,* vol. 14, no. 13, pp. 4103-4110.

Simpson, J.F., Gray, R., Dressler, L.G., Cobau, C.D., Falkson, C.I., Gilchrist, K.W., Pandya, K.J., Page, D.L. & Robert, N.J. 2000, "Prognostic value of histologic grade and proliferative activity in axillary node-positive breast cancer: results from the Eastern Cooperative Oncology Group Companion Study, EST 4189", *Journal of clinical oncology : official journal of the American Society of Clinical Oncology,* vol. 18, no. 10, pp. 2059-2069.

Smalley, M. & Ashworth, A. 2003, "Stem cells and breast cancer: A field in transit", *Nature reviews.Cancer,* vol. 3, no. 11, pp. 832-844.

Sohn, V.Y., Arthurs, Z.M., Sebesta, J.A. & Brown, T.A. 2008, "Primary tumor location impacts breast cancer survival", *American Journal of Surgery,* vol. 195, no. 5, pp. 641-644.

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., Eystein Lonning, P. & Borresen-Dale, A.L. 2001, "Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 19, pp. 10869-10874.

Sperger, J.M., Chen, X., Draper, J.S., Antosiewicz, J.E., Chon, C.H., Jones, S.B., Brooks, J.D., Andrews, P.W., Brown, P.O. & Thomson, J.A. 2003, "Gene expression patterns in human embryonic stem cells and human pluripotent germ cell tumors", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 23, pp. 13350-13355.

Sproul, A.M. 2006, "Real-time PCR applications in hematology" in *Real-time PCR*, ed. M.T. Dorak, Taylor & Francis Group, The United Kingdom, pp. 277-289.

Srivastava, N., Gochhait, S., Gupta, P. & Bamezai, R.N. 2008, "Copy number alterations of the H2AFX gene in sporadic breast cancer patients", *Cancer genetics and cytogenetics*, vol. 180, no. 2, pp. 121-128.

Staaf, J., Jonsson, G., Ringner, M., Vallon-Christersson, J., Grabau, D., Arason, A., Gunnarsson, H., Agnarsson, B.A., Malmstrom, P.O., Johannsson, O.T., Loman, N., Barkardottir, R.B. & Borg, A. 2010, "High-resolution genomic and expression analyses of copy number alterations in HER2-amplified breast cancer", *Breast cancer research : BCR*, vol. 12, no. 3, pp. R25.

Stingl, J., Raouf, A., Emerman, J.T. & Eaves, C.J. 2005, "Epithelial progenitors in the normal human mammary gland", *Journal of mammary gland biology and neoplasia,* vol. 10, no. 1, pp. 49-59.

Strand, C., Enell, J., Hedenfalk, I. & Ferno, M. 2007, "RNA quality in frozen breast cancer samples and the influence on gene expression analysis--a comparison of three evaluation methods using microcapillary electrophoresis traces", *BMC molecular biology,* vol. 8, pp. 38.

Stratton, M. & Futreal, A. 2007, Tuesday 13th of February-last update, *The Cancer Genome Project—SNP Array Based LOH and Copy Number Analysis—Tissue: Breast—Chromosome: 12* [Homepage of Wellcome Trust Sanger Institute], [Online]. Available: http://www.sanger.ac.uk/cgi-bin/genetics/CGP/CGHviewer. cgi? tissue=breast [2007, November/Thursday 1st].

Su, Y., Zheng, Y., Zheng, W., Gu, K., Chen, Z., Li, G., Cai, Q., Lu, W. & Shu, X.O. 2011, "Distinct distribution and prognostic significance of molecular subtypes of breast cancer in Chinese women: a population-based cohort study", *BMC cancer,* vol. 11, pp. 292.

Sun, B., Zhang, F., Wu, S.K., Guo, X., Zhang, L.L., Jiang, Z.F., Wang, D.M. & Song, S.T. 2011, "Gene expression profiling for breast cancer prognosis in Chinese populations", *The breast journal,* vol. 17, no. 2, pp. 172-179.

Susini, T., Bussani, C., Marini, G., Nori, J., Olivieri, S., Molino, C., Bianchi, S., Vezzosi, V., Paglierani, M., Giachi, M., Borrani, E. & Scarselli, G. 2010, "Preoperative assessment of HER-2/neu status in breast carcinoma: the role of quantitative real-time PCR on core-biopsy specimens", *Gynecologic oncology,* vol. 116, no. 2, pp. 234-239.

Tadrous, P.J. 2007, "Breast" in *Diagnostic Criteria Handbook in HISTOPATHOLOGY: A Surgical Pathology Vade Mecum* John Wiley & Sons, England, pp. 258-266.

Tai, M.H., Chang, C.C., Kiupel, M., Webster, J.D., Olson, L.K. & Trosko, J.E. 2005, "Oct4 expression in adult human stem cells: evidence in support of the stem cell theory of carcinogenesis", *Carcinogenesis,* vol. 26, no. 2, pp. 495-502.

Tan, B.T., Park, C.Y., Ailles, L.E. & Weissman, I.L. 2006, "The cancer stem cell hypothesis: a work in progress", *Laboratory investigation; a journal of technical methods and pathology,* vol. 86, no. 12, pp. 1203-1207.

Tanaka, K., Okamoto, S., Ishikawa, Y., Tamura, H. & Hara, T. 2009, "DDX1 is required for testicular tumorigenesis, partially through the transcriptional activation of 12p stem cell genes", *Oncogene,* vol. 28, no. 21, pp. 2142-2151.

Taneja, P., Maglic, D., Kai, F., Zhu, S., Kendig, R.D., Fry, E.A. & Inoue, K. 2010, "Classical and Novel Prognostic Markers for Breast Cancer and their Clinical Significance", *Clinical Medicine Insights.Oncology,* vol. 4, pp. 15-34.

Taranger, C.K., Noer, A., Sorensen, A.L., Hakelien, A.M., Boquest, A.C. & Collas, P. 2005, "Induction of dedifferentiation, genomewide transcriptional programming, and epigenetic reprogramming by extracts of carcinoma and embryonic stem cells", *Molecular biology of the cell,* vol. 16, no. 12, pp. 5719-5735.

Thellin, O., Zorzi, W., Lakaye, B., De Borman, B., Coumans, B., Hennen, G., Grisar, T., Igout, A. & Heinen, E. 1999, "Housekeeping genes as internal standards: use and limits", *Journal of Biotechnology,* vol. 75, no. 2-3, pp. 291-295.

Thompson, P.A., Brewster, A.M., Kim-Anh, D., Baladandayuthapani, V., Broom, B.M., Edgerton, M.E., Hahn, K.M., Murray, J.L., Sahin, A., Tsavachidis, S., Wang, Y., Zhang, L., Hortobagyi, G.N., Mills, G.B. & Bondy, M.L. 2011, "Selective genomic copy number imbalances and probability of recurrence in early-stage breast cancer", *PloS one,* vol. 6, no. 8, pp. e23543.

Tischkowitz, M., Brunet, J.S., Begin, L.R., Huntsman, D.G., Cheang, M.C., Akslen, L.A., Nielsen, T.O. & Foulkes, W.D. 2007, "Use of immunohistochemical markers can refine prognosis in triple negative breast cancer", *BMC cancer,* vol. 7, pp. 134.

Tokimasa, S., Ohta, H., Sawada, A., Matsuda, Y., Kim, J.Y., Nishiguchi, S., Hara, J. & Takihara, Y. 2001, "Lack of the Polycomb-group gene rae28 causes maturation arrest at the early B-cell developmental stage", *Experimental hematology,* vol. 29, no. 1, pp. 93-103.

Tomlinson, D.C., Knowles, M.A. & Speirs, V. 2012, "Mechanisms of FGFR3 actions in endocrine resistant breast cancer", *International journal of cancer.Journal international du cancer,* vol. 130, no. 12, pp. 2857-2866.

Torres, L., Ribeiro, F.R., Pandis, N., Andersen, J.A., Heim, S. & Teixeira, M.R. 2007, "Intratumor genomic heterogeneity in breast cancer with clonal divergence between primary carcinomas and lymph node metastases", *Breast cancer research and treatment,* vol. 102, no. 2, pp. 143-155.

Tricarico, C., Pinzani, P., Bianchi, S., Paglierani, M., Distante, V., Pazzagli, M., Bustin, S.A. & Orlando, C. 2002, "Quantitative real-time reverse transcription polymerase chain reaction: normalization to rRNA or single housekeeping genes is inappropriate for human tissue biopsies", *Analytical Biochemistry,* vol. 309, no. 2, pp. 293-300.

Trichopoulos, D., Adami, H.O., Ekbom, A., Hsieh, C.C. & Lagiou, P. 2008, "Early life events and conditions and breast cancer risk: from epidemiology to etiology", *International journal of cancer.Journal international du cancer,* vol. 122, no. 3, pp. 481-485.

Trosko, J.E., Chang, C.C., Upham, B.L. & Tai, M.H. 2005, "The role of human adult stem cells and cell-cell communication in cancer chemoprevention and chemotherapy strategies", *Mutation research,* vol. 591, no. 1-2, pp. 187-197.

Tsubura, A., Yoshizawa, K., Uehara, N., Yuri, T. & Matsuoka, Y. 2007, "Multistep mouse mammary tumorigenesis through pre-neoplasia to neoplasia and acquisition of metastatic potential", *Medical molecular morphology,* vol. 40, no. 1, pp. 9-17.

Turner, N., Lambros, M.B., Horlings, H.M., Pearson, A., Sharpe, R., Natrajan, R., Geyer, F.C., van Kouwenhove, M., Kreike, B., Mackay, A., Ashworth, A., van de Vijver, M.J. & Reis-Filho, J.S. 2010, "Integrative molecular profiling of triple negative breast cancers identifies amplicon drivers and potential therapeutic targets", *Oncogene*, vol. 29, no. 14, pp. 2013-2023.

van Beers, E.H., van Welsem, T., Wessels, L.F., Li, Y., Oldenburg, R.A., Devilee, P., Cornelisse, C.J., Verhoef, S., Hogervorst, F.B., van't Veer, L.J. & Nederlof, P.M. 2005, "Comparative genomic hybridization profiles in human BRCA1 and BRCA2 breast tumors highlight differential sets of genomic aberrations", *Cancer research*, vol. 65, no. 3, pp. 822-827.

van de Vijver, M.J., He, Y.D., van't Veer, L.J., Dai, H., Hart, A.A., Voskuil, D.W., Schreiber, G.J., Peterse, J.L., Roberts, C., Marton, M.J., Parrish, M., Atsma, D., Witteveen, A., Glas, A., Delahaye, L., van der Velde, T., Bartelink, H., Rodenhuis, S., Rutgers, E.T., Friend, S.H. & Bernards, R. 2002, "A gene-expression signature as a predictor of survival in breast cancer", *The New England journal of medicine*, vol. 347, no. 25, pp. 1999-2009.

van der Groep, P., Bouter, A., van der Zanden, R., Siccama, I., Menko, F.H., Gille, J.J., van Kalken, C., van der Wall, E., Verheijen, R.H. & van Diest, P.J. 2006, "Distinction between hereditary and sporadic breast cancer on the basis of clinicopathological data", *Journal of clinical pathology*, vol. 59, no. 6, pp. 611-617.

van Leenders, G.J., Dukers, D., Hessels, D., van den Kieboom, S.W., Hulsbergen, C.A., Witjes, J.A., Otte, A.P., Meijer, C.J. & Raaphorst, F.M. 2007, "Polycomb-group oncogenes EZH2, BMI1, and RING1 are overexpressed in prostate cancer with adverse pathologic and clinical features", *European urology,* vol. 52, no. 2, pp. 455-463.

van Loon, J., van Baardwijk, A., Boersma, L., Ollers, M., Lambin, P. & De Ruysscher, D. 2011, "Therapeutic implications of molecular imaging with PET in the combined modality treatment of lung cancer", *Cancer treatment reviews,* vol. 37, no. 5, pp. 331-343.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A. & Speleman, F. 2002, "Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes", *Genome biology*, vol. 3, no. 7, pp. RESEARCH0034.

Verderio, P., Mangia, A., Ciniselli, C.M., Tagliabue, P. & Paradiso, A. 2010, "Biomarkers for Early Cancer Detection - Methodological Aspects", *Breast care (Basel, Switzerland),* vol. 5, no. 2, pp. 62-65.

Villadsen, R., Fridriksdottir, A.J., Ronnov-Jessen, L., Gudjonsson, T., Rank, F., LaBarge, M.A., Bissell, M.J. & Petersen, O.W. 2007, "Evidence for a stem cell hierarchy in the adult human breast", *The Journal of cell biology,* vol. 177, no. 1, pp. 87-101.

Vincent-Salomon, A., Gruel, N., Lucchesi, C., MacGrogan, G., Dendale, R., Sigal-Zafrani, B., Longy, M., Raynal, V., Pierron, G., de Mascarel, I., Taris, C., Stoppa-Lyonnet, D., Pierga, J.Y., Salmon, R., Sastre-Garau, X., Fourquet, A., Delattre, O., de Cremoux, P. & Aurias, A. 2007, "Identification of typical medullary breast carcinoma as a genomic sub-group of basal-like carcinomas, a heterogeneous new molecular entity", *Breast cancer research : BCR*, vol. 9, no. 2, pp. R24.

Vorgias, G., Koukouras, D., Paleogianni, V. & Tzoracoeleftherakis, E. 2001, "Prognostic significance of factors affecting disease free interval and overall survival for Stage II breast cancer in Greece. A multivariate cohort study", *European journal of obstetrics, gynecology, and reproductive biology*, vol. 95, no. 1, pp. 100-104.

Walker, R.A. 2006, "Quantification of immunohistochemistry--issues concerning methods, utility and semiquantitative assessment I", *Histopathology*, vol. 49, no. 4, pp. 406-410.

Walker, R.A. 2008, "Immunohistochemical markers as predictive tools for breast cancer", *Journal of clinical pathology*, vol. 61, no. 6, pp. 689-696.

Warren, B.S. & Devine, C.M. 2004, *Understanding Breast Cancer Risk and Risk Factors Associated With Diet and Lifestyle*, Cornell University Program on Breast Cancer and Environmental Risk Factors, USA.

Warrington, J.A., Nair, A., Mahadevappa, M. & Tsyganskaya, M. 2000, "Comparison of human adult and fetal expression and identification of 535 housekeeping/maintenance genes", *Physiological genomics*, vol. 2, no. 3, pp. 143-147.

Wasuthit, Y., Kongdan, Y., Suvikapakornkul, R., Lertsithichai, P. & Chirappapha, P. 2011, "Predictive factors of axillary lymph node metastasis in breast cancer", *Journal of the Medical Association of Thailand = Chotmaihet thangphaet,* vol. 94, no. 1, pp. 65-70.

Watanabe, K., Meyer, M.J., Strizzi, L., Lee, J.M., Gonzales, M., Bianco, C., Nagaoka, T., Farid, S.S., Margaryan, N., Hendrix, M.J., Vonderhaar, B.K. & Salomon, D.S. 2010, "Cripto-1 is a cell surface marker for a tumorigenic, undifferentiated subpopulation in human embryonal carcinoma cells", *Stem cells (Dayton, Ohio),* vol. 28, no. 8, pp. 1303-1314.

Weigel, M.T. & Dowsett, M. 2010, "Current and emerging biomarkers in breast cancer: prognosis and prediction", *Endocrine-related cancer,* vol. 17, no. 4, pp. R245-62.

Weigelt, B., Baehner, F.L. & Reis-Filho, J.S. 2010, "The contribution of gene expression profiling to breast cancer classification, prognostication and prediction: a retrospective of the last decade", *The Journal of pathology,* vol. 220, no. 2, pp. 263-280.

Weigelt, B., Geyer, F.C. & Reis-Filho, J.S. 2010, "Histological types of breast cancer: how special are they?", *Molecular oncology,* vol. 4, no. 3, pp. 192-208.

Weigelt, B., Horlings, H.M., Kreike, B., Hayes, M.M., Hauptmann, M., Wessels, L.F., de Jong, D., Van de Vijver, M.J., Van't Veer, L.J. & Peterse, J.L. 2008, "Refinement of breast cancer classification by molecular characterization of histological special types", *The Journal of pathology,* vol. 216, no. 2, pp. 141-150.

Weinberg RA 2007a, "Cancer cells are often aneuploid" in *The Biology of Cancer* Garland Science, Taylor & francis Group, The United States of America, pp. 12-13.

Weinberg, R.A. 2007b, "Multi-step tumorigenesis" in *The Biology of Cancer* Garland Science, The United States of America, pp. 399-462.

Weisser, M., Haferlach, T., Schoch, C., Hiddemann, W. & Schnittger, S. 2004, "The use of housekeeping genes for real-time PCR-based quantification of fusion gene transcripts in acute myeloid leukemia", *Leukemia : official journal of the Leukemia Society of America, Leukemia Research Fund, U.K,* vol. 18, no. 9, pp. 1551-1553.

Weng, W.H., Wejde, J., Ahlen, J., Pang, S.T., Lui, W.O. & Larsson, C. 2004, "Characterization of large chromosome markers in a malignant fibrous histiocytoma by spectral karyotyping, comparative genomic hybridization (CGH), and array CGH", *Cancer genetics and cytogenetics,* vol. 150, no. 1, pp. 27-32.

Werner, M., Chott, A., Fabiano, A. & Battifora, H. 2000, "Effect of formalin tissue fixation and processing on immunohistochemistry", *The American Journal of Surgical Pathology,* vol. 24, no. 7, pp. 1016-1019.

Werner, M., Von Wasielewski, R. & Komminoth, P. 1996, "Antigen retrieval, signal amplification and intensification in immunohistochemistry", *Histochemistry and cell biology,* vol. 105, no. 4, pp. 253-260.

Wicha, M.S., Liu, S. & Dontu, G. 2006, "Cancer stem cells: an old idea--a paradigm shift", *Cancer research,* vol. 66, no. 4, pp. 1883-90; discussion 1895-6.

Wikman, H., Nymark, P., Vayrynen, A., Jarmalaite, S., Kallioniemi, A., Salmenkivi, K., Vainio-Siukola, K., Husgafvel-Pursiainen, K., Knuutila, S., Wolf, M. & Anttila, S. 2005, "CDK4 is a probable target gene in a novel amplicon at 12q13.3-q14.1 in lung cancer", *Genes, chromosomes & cancer,* vol. 42, no. 2, pp. 193-199.

Wilkening, S. & Bader, A. 2004, "Quantitative real-time polymerase chain reaction: methodical analysis and mathematical model", *Journal of biomolecular techniques : JBT,* vol. 15, no. 2, pp. 107-111.

Williamson, D., Lu, Y.J., Gordon, T., Sciot, R., Kelsey, A., Fisher, C., Poremba, C., Anderson, J., Pritchard-Jones, K. & Shipley, J. 2005, "Relationship between MYCN copy number and expression in rhabdomyosarcomas and correlation with adverse prognosis in the alveolar subtype", *Journal of clinical oncology : official journal of the American Society of Clinical Oncology,* vol. 23, no. 4, pp. 880-888.

Wolff, A.C., Hammond, M.E., Schwartz, J.N., Hagerty, K.L., Allred, D.C., Cote, R.J., Dowsett, M., Fitzgibbons, P.L., Hanna, W.M., Langer, A., McShane, L.M., Paik, S., Pegram, M.D., Perez, E.A., Press, M.F., Rhodes, A., Sturgeon, C., Taube, S.E., Tubbs, R., Vance, G.H., van de Vijver, M., Wheeler, T.M., Hayes, D.F., American Society of Clinical Oncology & College of American Pathologists 2007, "American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer", *Journal of clinical oncology : official journal of the American Society of Clinical Oncology,* vol. 25, no. 1, pp. 118-145.

Wood, H.M., Belvedere, O., Conway, C., Daly, C., Chalkley, R., Bickerdike, M., McKinley, C., Egan, P., Ross, L., Hayward, B., Morgan, J., Davidson, L., MacLennan, K., Ong, T.K., Papagiannopoulos, K., Cook, I., Adams, D.J., Taylor, G.R. & Rabbitts, P. 2010, "Using next-generation sequencing for high resolution multiplex analysis of copy number variation from nanogram quantities of DNA from formalin-fixed paraffin-embedded specimens", *Nucleic acids research*, vol. 38, no. 14, pp. e151.

Woodward, W.A., Chen, M.S., Behbod, F. & Rosen, J.M. 2005, "On mammary stem cells", *Journal of cell science*, vol. 118, no. Pt 16, pp. 3585-3594.

Wooster, R. & Weber, B.L. 2003, "Breast and ovarian cancer", *The New England journal of medicine*, vol. 348, no. 23, pp. 2339-2347.

Wright, M.H., Calcagno, A.M., Salcido, C.D., Carlson, M.D., Ambudkar, S.V. & Varticovski, L. 2008, "Brca1 breast tumors contain distinct CD44+/CD24- and CD133+ cells with cancer stem cell characteristics", *Breast cancer research : BCR*, vol. 10, no. 1, pp. R10.

Wu, X.Z. 2008, "Origin of cancer stem cells: the role of self-renewal and differentiation", *Annals of surgical oncology*, vol. 15, no. 2, pp. 407-414.

Xiao, Y., Ye, Y., Yearsley, K., Jones, S. & Barsky, S.H. 2008, "The lymphovascular embolus of inflammatory breast cancer expresses a stem cell-like phenotype", *The American journal of pathology*, vol. 173, no. 2, pp. 561-574.

Xie, R., Chung, J.Y., Ylaya, K., Williams, R.L., Guerrero, N., Nakatsuka, N., Badie, C. & Hewitt, S.M. 2011, "Factors influencing the degradation of archival formalin-fixed paraffin-embedded tissue sections", *The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society,* vol. 59, no. 4, pp. 356-365.

Xu, C., Houck, J.R., Fan, W., Wang, P., Chen, Y., Upton, M., Futran, N.D., Schwartz, S.M., Zhao, L.P., Chen, C. & Mendez, E. 2008, "Simultaneous isolation of DNA and RNA from the same cell population obtained by laser capture microdissection for genome and transcriptome profiling", *The Journal of molecular diagnostics : JMD,* vol. 10, no. 2, pp. 129-134.

Yamamoto, H., Okumura, K., Toshima, S., Mukaisho, K., Sugihara, H., Hattori, T., Kato, M. & Asano, S. 2009, "FXYD3 protein involved in tumor cell proliferation is overproduced in human breast cancer tissues", *Biological & pharmaceutical bulletin,* vol. 32, no. 7, pp. 1148-1154.

Yan, P.S., Venkataramu, C., Ibrahim, A., Liu, J.C., Shen, R.Z., Diaz, N.M., Centeno, B., Weber, F., Leu, Y.W., Shapiro, C.L., Eng, C., Yeatman, T.J. & Huang, T.H. 2006, "Mapping geographic zones of cancer risk with epigenetic biomarkers in normal breast tissue", *Clinical cancer research : an official journal of the American Association for Cancer Research,* vol. 12, no. 22, pp. 6626-6636.

Yao, J., Weremowicz, S., Feng, B., Gentleman, R.C., Marks, J.R., Gelman, R., Brennan, C. & Polyak, K. 2006, "Combined cDNA array comparative genomic hybridization and serial analysis of gene expression analysis of breast tumor progression", *Cancer research,* vol. 66, no. 8, pp. 4065-4078.

Ye, F., Zhou, C., Cheng, Q., Shen, J. & Chen, H. 2008, "Stem-cell-abundant proteins Nanog, Nucleostemin and Musashi1 are highly expressed in malignant cervical epithelial cells", *BMC cancer*, vol. 8, pp. 108.

Yu, H.R., Kuo, H.C., Huang, H.C., Huang, L.T., Tain, Y.L., Chen, C.C., Liang, C.D., Sheen, J.M., Lin, I.C., Wu, C.C., Ou, C.Y. & Yang, K.D. 2011, "Glyceraldehyde-3-phosphate dehydrogenase is a reliable internal control in Western blot analysis of leukocyte subpopulations from children", *Analytical Biochemistry*, vol. 413, no. 1, pp. 24-29.

Yu, S.C. & Bian, X.W. 2009, "Enrichment of cancer stem cells based on heterogeneity of invasiveness", *Stem cell reviews*, vol. 5, no. 1, pp. 66-71.

Zafarana, G., Grygalewicz, B., Gillis, A.J., Vissers, L.E., van de Vliet, W., van Gurp, R.J., Stoop, H., Debiec-Rychter, M., Oosterhuis, J.W., van Kessel, A.G., Schoenmakers, E.F., Looijenga, L.H. & Veltman, J.A. 2003, "12p-amplicon structure analysis in testicular germ cell tumors of adolescents and adults by array CGH", *Oncogene*, vol. 22, no. 48, pp. 7695-7701.

Zang, X.P. & Pento, J.T. 2008, "SiRNA inhibition of ER-alpha expression reduces KGF-induced proliferation of breast cancer cells", *Anticancer Research*, vol. 28, no. 5A, pp. 2733-2735.

Zanni, K.L. & Chan, G.K. 2011, "Laser capture microdissection: understanding the techniques and implications for molecular biology in nursing research through analysis of breast cancer tumor samples", *Biological research for nursing*, vol. 13, no. 3, pp. 297-305.

Zhang, F., Gu, W., Hurles, M.E. & Lupski, J.R. 2009a, "Copy number variation in human health, disease, and evolution", *Annual review of genomics and human genetics,* vol. 10, pp. 451-481.

Zhang, J., Wang, X., Li, M., Han, J., Chen, B., Wang, B. & Dai, J. 2006, "NANOGP8 is a retrogene expressed in cancers", *The FEBS journal,* vol. 273, no. 8, pp. 1723-1730.

Zhang, Y., Martens, J.W., Yu, J.X., Jiang, J., Sieuwerts, A.M., Smid, M., Klijn, J.G., Wang, Y. & Foekens, J.A. 2009b, "Copy number alterations that predict metastatic capability of human breast cancer", *Cancer research,* vol. 69, no. 9, pp. 3795-3801.

Zhao, B., Schwartz, L.H. & Larson, S.M. 2009, "Imaging surrogates of tumor response to therapy: anatomic and functional biomarkers", *Journal of nuclear medicine : official publication, Society of Nuclear Medicine,* vol. 50, no. 2, pp. 239-249.

Zhuang, Z., Bertheau, P., Emmert-Buck, M.R., Liotta, L.A., Gnarra, J., Linehan, W.M. & Lubensky, I.A. 1995, "A microdissection technique for archival DNA analysis of specific cell populations in lesions < 1 mm in size", *The American journal of pathology,* vol. 146, no. 3, pp. 620-625.

PUBLICATIONS

1. Page, K., Hava, N., Ward, B., Brown, J., Guttery, D.S., **Ruangpratheep, C.**, Blighe, K., Sharma, A., Walker, R.A., Coombes, R.C. & Shaw, J.A. 2011, "Detection of HER2 amplification in circulating free DNA in patients with breast cancer", *British journal of cancer*, vol. 104, no. 8, pp. 1342-1348.

2. Shaw, J.A., Page, K., Blighe, K., Hava, N., Guttery, D., Ward, B., Brown, J., **Ruangpratheep, C.**, Stebbing, J., Payne, R., Palmieri, C., Cleator, S., Walker, R.A. & Coombes, R.C. 2011, "Genomic analysis of circulating cell free DNA infers breast cancer dormancy", *Genome research*.

Genomic analysis of circulating cell free DNA infers breast cancer dormancy

Jacqueline A. Shaw, Karen Page, Kevin Blighe, et al.

Genome Res. published online October 11, 2011
Access the most recent version at doi:10.1101/gr.123497.111

Genomic analysis of circulating cell free DNA infers breast cancer dormancy

Jacqueline A. Shaw,[1,3] Karen Page,[1] Kevin Blighe,[1] Natasha Hava,[2] David Guttery,[1] Becky Ward,[2] James Brown,[1] Chetana Ruangpratheep,[1] Justin Stebbing,[2] Rachel Payne,[2] Carlo Palmieri,[2] Suzy Cleator,[2] Rosemary A. Walker,[1] and R. Charles Coombes[2]

[1]*Department of Cancer Studies and Molecular Medicine, University of Leicester, Leicester LE2 7LX, United Kingdom;* [2]*Division of Cancer, Imperial College, Hammersmith Hospital, London W12 ONN, United Kingdom*

Biomarkers in breast cancer to monitor minimal residual disease have remained elusive. We hypothesized that genomic analysis of circulating free DNA (cfDNA) isolated from plasma may form the basis for a means of detecting and monitoring breast cancer. We profiled 251 genomes using Affymetrix SNP 6.0 arrays to determine copy number variations (CNVs) and loss of heterozygosity (LOH), comparing 138 cfDNA samples with matched primary tumor and normal leukocyte DNA in 65 breast cancer patients and eight healthy female controls. Concordance of SNP genotype calls in paired cfDNA and leukocyte DNA samples distinguished between breast cancer patients and healthy female controls ($P < 0.0001$) and between preoperative patients and patients on follow-up who had surgery and treatment ($P = 0.0016$). Principal component analyses of cfDNA SNP/copy number results also separated presurgical breast cancer patients from the healthy controls, suggesting specific CNVs in cfDNA have clinical significance. We identified focal high-level DNA amplification in paired tumor and cfDNA clustered in a number of chromosome arms, some of which harbor genes with oncogenic potential, including *USP17L2* (*DUB3*), *BRF1*, *MTA1*, and *JAG2*. Remarkably, in 50 patients on follow-up, specific CNVs were detected in cfDNA, mirroring the primary tumor, up to 12 yr after diagnosis despite no other evidence of disease. These data demonstrate the potential of SNP/CNV analysis of cfDNA to distinguish between patients with breast cancer and healthy controls during routine follow-up. The genomic profiles of cfDNA infer dormancy/minimal residual disease in the majority of patients on follow-up.

[Supplemental material is available for this article.]

Breast cancer is one of the most common forms of cancer in women in Western industrial countries. Although advances in diagnosis and treatment have improved survival (Early Breast Cancer Trialists' Collaborative Group 2005), it is not possible to reliably identify breast cancer patients who will relapse with metastatic disease, and relapse can occur up to 20 yr after primary treatment (Karrison et al. 1999). This potentially long period between resection and relapse is not likely to be explained by growth of secondary tumors (Meltzer 1990; Demicheli et al. 1998; Chambers and Goss 2008) but more likely suggests a period of dormancy, where there is growth restriction of unseen micrometastases (Murray 1995). Although this long latency between resection and relapse is common in breast cancer, the associated biological mechanisms are poorly understood. However, it is well established that treatment is more effective when given before overt metastatic disease develops, underscoring the need for markers of minimal disease, preferably one that also identifies a molecular target, as disclosed by gene amplification, for example.

A number of classical factors (e.g., type, grade, node status, and hormone receptor status) and prognostic and predictive markers (e.g., HER2, Ki-67) are used to determine individual risk, but these are assessed in the primary tumor removed by surgery and are not useful in monitoring minimal disease. Moreover, genetic changes can occur between metastases and the primary tumor. Therefore, the development of tests with a clinical relevance for risk estimation and monitoring is of great interest (Levenson 2007). Stroun et al. (1987) first reported that circulating DNA in cancer patients could be distinguished from other patients with non-neoplastic disease. Measurement of levels of circulating free DNA (cfDNA) were subsequently suggested for the diagnosis of breast cancers (Huang et al. 2006), but elevated levels are sometimes seen in benign disease (Zanetti-Dällenbach et al. 2008). In breast cancer, gene expression analysis has disclosed that multiple changes can occur in micrometastases in the bone marrow, compared with metastatic disease in draining lymph nodes (Gangnus et al. 2004). Thus, it would be hugely advantageous to be able to detect specific changes indicative of progression in cfDNA.

Copy number (CN) variations (CNVs) are amplified or deleted regions of the genome, of variable size, which are recognized as a major source of normal human genome variability (Iafrate et al. 2004; Sebat et al. 2004) and contribute significantly to phenotypic variation (Redon et al. 2006). Hence, specific CNVs may be characteristic of different tumor types. Loss of heterozygosity (LOH) is also common in many tumors and can reveal recessive alleles (Wang et al. 2004). The Affymetrix SNP 6.0 array contains 906,600 probes for SNPs and 946,000 probes for CNVs and represents more genetic variation on a single array than any other array platform. Analysis of SNP 6.0 array results can generate SNP genotypes, CNVs, and LOH data in a single hybridization experiment. Due to the problems inherent in obtaining sequential samples as the cancer progresses to metastatic disease, little is known of the nature of dynamic changes of the cancer

[3]Corresponding author.
E-mail js39@leicester.ac.uk.
Article published online before print. Article, supplemental material, and publication date are at http://www.genome.org/cgi/doi/10.1101/gr.123497.111. Freely available online through the *Genome Research* Open Access option.

genome over time. We hypothesized that in patients on follow-up who are otherwise disease free, evidence of tumor DNA detected in cfDNA would suggest that this is derived from or related to micrometastases in the bone marrow. Therefore, the aim of this study was to compare SNP 6.0 whole-genome profiles of the primary tumor with paired plasma cfDNA samples of breast cancer patients on follow-up and related findings to plasma cfDNA profiles of primary breast cancer patients for whom we collected presurgical blood samples and healthy female controls. This aim was achieved by the successful profiling of 251 genomes to determine CNVs and LOH in paired tumors and cfDNA and by comparison with matched normal leukocyte DNA samples from the same patients.

Results

Low levels of cfDNA were detected in all plasma samples from patients and healthy female controls, consistent with our previous studies (Page et al. 2011; Shaw et al. 2011). There was no significant difference in mean cfDNA concentration between the healthy controls and either presurgical patients or patients on follow-up as assessed by absolute quantitation of a 96-bp amplicon (Shaw et al. 2011) and by ROC curve analysis.

We surveyed 251 DNA samples, isolated from normal leukocytes, plasma, and tumor from 65 breast cancer patients and eight healthy female controls, using the Affymetrix Genome-Wide Human SNP Array 6.0. We analyzed plasma prior to any surgery or treatment in 15 breast cancer patients. The other 50 patients were on follow-up after surgical removal of their primary tumor (Table 1). We compared cfDNA in two separate plasma samples (P1 and P2) for each of these taken a mean of 6.1 and 9 yr after surgery. None of these 50 patients had any evidence of metastases or recurrent disease using standard radiologic or other clinical parameters. The 251 DNA samples were hybridized in two batches only to reduce interassay variability. We validated the approach by repeating 13 samples for the entire procedure from DNA isolation through array hybridization. The results showed excellent correlation between the replicated samples by three independent measures: quality-control (QC) call-rates ($P = 0.0001$), median of the absolute values of all pairwise differences (MAPD) ($P = 0.0005$; two tailed, paired *t*-tests), and mean Spearman correlation (0.783; range, 0.600–0.984), confirming the reproducibility of our approach. There was also high agreement for both the range and frequency of detected CNVs (Supplemental Fig. 1).

Table 1. Clinicopathologic details of 50 breast cancer patients on follow-up

	No (%)
Menopausal status	
Premenopausal	28 (56)
Post-menopausal	19 (38)
Data not available	3 (6)
Type of surgery	
Wide local excision	29 (58)
Mastectomy	20 (40)
Unknown	1 (2)
Histology	
Invasive ductal	34 (68)
Invasive lobular	7 (14)
Other/mixed invasive	6 (12)
Data not available	3 (6)
Lymph node status	
Positive	31 (62)
Negative	16 (32)
Data not available	3 (6)
Tumor size	
≥20 mm	24 (48)
<20 mm	21 (42)
Data not available	5 (10)
Tumor grade	
I	5 (10)
II	18 (36)
III	23 (46)
Data not available	4 (8)
Hormone receptor status	
ER-positive	37 (74)
ER-negative	13 (26)
ER-positive on endocrine therapy	30 (60)
ER-positive treatment unknown	1 (2)
PR-positive	24 (48)
PR-negative	23 (46)
PR data not available	3 (6)
Growth factor status	
HER2-positive	14 (28)
HER2-negative	35 (70)
Data not available (HER2)	1 (2)
Triple-negative (ER, PR, and HER2)	10 (20)
Total no. of patients	50 (100%)

Plasma SNP profiles distinguish between patients with breast cancer and healthy female controls

We first reviewed SNP call-rates for all samples as an indicator of successful array hybridization. The highest call-rates were for the normal leukocyte DNA samples (mean, 96.89%), with similar high call-rates in cfDNA from blood plasma and formalin-fixed paraffin-embedded (FFPE) tumor DNA (Supplemental Table 1). We next compared the concordance in SNP genotype calls. The normal leukocyte and plasma DNA samples from the healthy female controls showed an average of 64.23% and 63.50% concordance, respectively, with 15 female Caucasian HapMap samples (range, 62.75%–66.13% and 60.87%–65.38%; http://hapmap.ncbi.nlm.nih.gov/), underscoring the validity of our normal controls (Oldridge et al. 2010). Next, we compared SNP concordance between paired leukocyte and plasma cfDNA in all patients. The healthy controls had the highest mean concordance of SNP genotype calls (89.35%; range, 81.10%–94.08%; 95% confidence interval [CI], 0.09–2.74), and this was significantly lower for the presurgical breast cancer patients and patients on follow-up ($P < 0.0001$, one-way ANOVA), due to constitutional heterozygosity at multiple SNPs being converted to a hemizygous state in patients' plasma DNA (Fig. 1A). In the patients on follow-up, a total of 25 plasma samples (18 P1 and seven P2) showed high concordance (>80%) with their paired leukocytes, within the range observed for plasmas of the healthy controls, suggesting these plasma samples were derived largely from normal cells (Fig. 1A; Supplemental Table 2). Concordance of SNP genotype calls was low for all paired plasma and primary tumor samples (mean, 46.89%; range, 31.04%–66.20%; 95% CI, 0.12–3.78) (Fig. 1B), indicating significant differences between these.

A significant difference was also seen between the concordance of SNP genotype calls between the paired leukocyte and plasma DNA of the presurgical patients and the patients on follow-up ($P = 0.0016$, one-way ANOVA). Hence by concordance of SNP genotype calls, plasma of the presurgical breast cancer patients differs from healthy controls, and preoperative patients differ from those who have had surgery and treatment. Principal component analysis (PCA), which takes both CN and SNP markers into account, also showed clear separation between the plasma of the

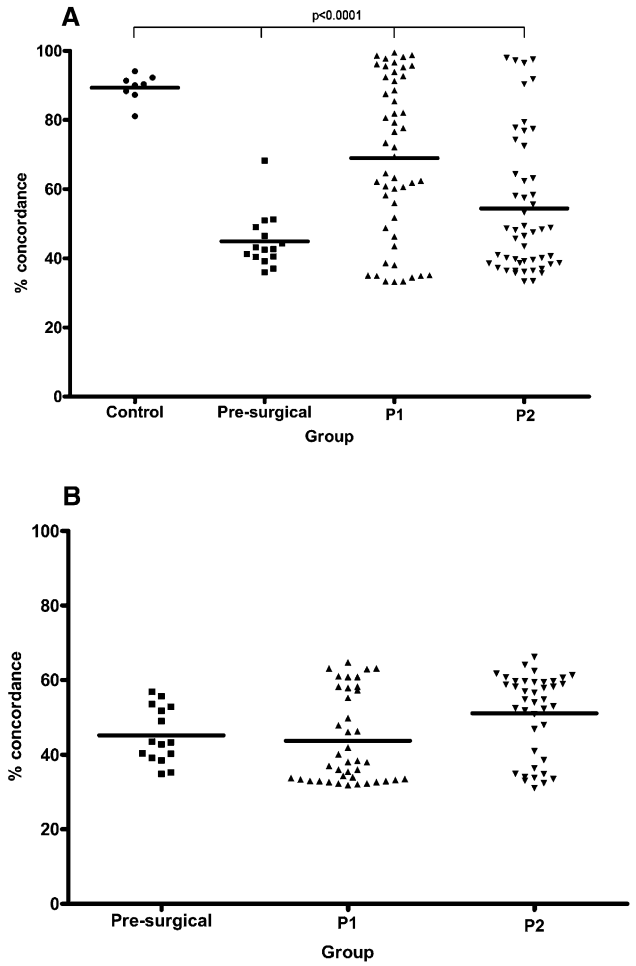


Figure 1. Plasma of breast cancer patients shows low SNP concordance with paired normal DNA. (*A*) Percentage of concordant SNP genotype calls for paired plasma and normal leukocyte DNA samples of patients and healthy controls. Percentage of concordance was significantly lower than controls in breast cancer patients ($P < 0{\cdot}0001$, one-way ANOVA). (*B*) Percentage of concordant SNP genotype calls for paired plasma and microdissected tumor (available for all presurgical patients and 40 patients on follow-up; mean 47.00%; range, 31.04%–66.20%; 95% CI, 0.07–2.28). In *A*, concordance was lowest for the 15 preoperative primary breast cancer patients (mean, 44.88%; range, 36.00%–68.27%; 95% CI, 0.13–4.02) but remained low for the 50 patients on follow-up using both P1 (mean, 69.10%; range, 33.17%–99.44%; 95% CI, 0.21–6.51) and P2 plasma samples (mean, 54.22%; range, 33.31%–97.96%; 95% CI, 0.18–5.65). Control indicates healthy female controls; presurgical, plasma of presurgical breast cancer patients; and P1 and P2, first and second plasma samples of patients on follow-up.

healthy controls and presurgical breast cancer patients (Fig. 2A). In the patients on follow-up, the plasma PCA profiles were scattered between the matched normal leukocyte and tumor DNA samples, which grouped separately (Fig. 2B). The 25 plasma samples that showed high SNP concordance with their paired leukocytes also clustered with these by PCA, suggesting a more "normal" genome profile in these samples.

We also compared the PCA profiles for P1 and P2 in the 50 patients on follow-up, based on the following sample groupings: (1) ER-positive versus ER-negative primary tumor status, (2) PR-positive versus PR-negative primary tumor status, (3) HER2-positive versus HER2-negative primary tumor status, (4) triple-negative (10 patients) versus any receptor-positive primary tumor status, (5) type of surgery (mastectomy versus wide local excision), and (6) endocrine therapy (tamoxifen/arimidex) prior to blood sampling versus none. There were no obvious trends observed in the cfDNA profiles of either the P1 or the P2 samples by PCA, for any of these variables, with samples again scattered between the matched normal leukocyte and tumor DNA samples (data not shown).

Plasma and tumor DNA show heterogeneous CNVs

We identified 7131 copy number (CN) segments in the plasma of the 15 presurgical patients and 38,560 CN segments in the plasma of the 50 patients on follow-up. Of these 55.20% completely or partially overlap with known CNVs listed in the Toronto Database of Genomic Variants (DGV) (Iafrate et al. 2004) and 44.80% were novel. The majority of CNVs detected were amplifications, with a mean of 67.25% and 58.75% in tumor and plasma, respectively (Table 2). Both the presurgical patients and patients on follow-up showed significant differences in the frequency and range of amplification and deletions detected between cfDNA and matched leukocytes, again providing evidence of genomic change in patients' cfDNA, whereas CNV results were more similar for paired cfDNA and normal leukocytes of the healthy female controls. We examined the CNV data by applying a Gaussian smoothed signal threshold of >6.0 to filter out lower-level changes, which revealed 634 CNVs common to more than one patient. Filtering these by amplification in >10% of patients identified 23 chromosomal intervals, showing amplification in plasma and tumor DNA with little or no amplification in the plasma of healthy controls (Fig. 3; Table 3). The results were reproducible across three software platforms (Affymetrix Genotyping Console, Partek Genomics Suite, and Nexus Copy Number Discovery Edition). The majority of the 23 CNVs were >50 kb in size with more than 50 markers (Supplemental Table 3): 18 have known overlapping genes, and five have none as defined by the HUGO Gene Nomenclature Committee (HGNC) gene database (http://www.genenames.org/). By applying a lower smoothed signal threshold of >4.0, seven of these intervals showed amplification in >90% of tumor and >25% of plasma samples of patients on follow-up (Supplemental Table 4). These seven CNV intervals were more frequently detected in the plasma of node-positive patients than T_1N_0 patients.

We also used linear regression analysis to compare the relationship between the presence (or not) of each of the 23 CNVs (from Table 3) in both the cfDNA and tumor DNA samples with tumor phenotype, type of surgery, and therapy. We classed each DNA sample as positive or negative at each CNV interval based on the presence or absence of a peak with a CN > 6.0 by Gaussian smooth signal. The majority of CNVs detected in cfDNA were significantly associated with breast cancer (for both the presurgical patients and 50 patients on follow-up). Of note, a number of CNVs, including 1p36.33, 1q21.1, 9p11.2, 9q12, and 19p13.3, were significantly associated with relapse. In cfDNA, 4q13.2 was associated with ER-positive cancer, and 9q12 was associated with triple-negative cancer. However, there were no significant associations with HER2 and PR (Table 4).

To validate CNVs, we developed locus-specific assays to 4q13.2 and 16p12.3 and used real-time quantitative PCR (qPCR) to analyze the unamplified tumor DNA from 37 primary breast cancers (from an independent series) and compared results with 56 normal leukocyte DNA samples. Ten of 37 tumor DNA samples (27%) showed amplification at 4q13.2, and 14 tumor DNA samples

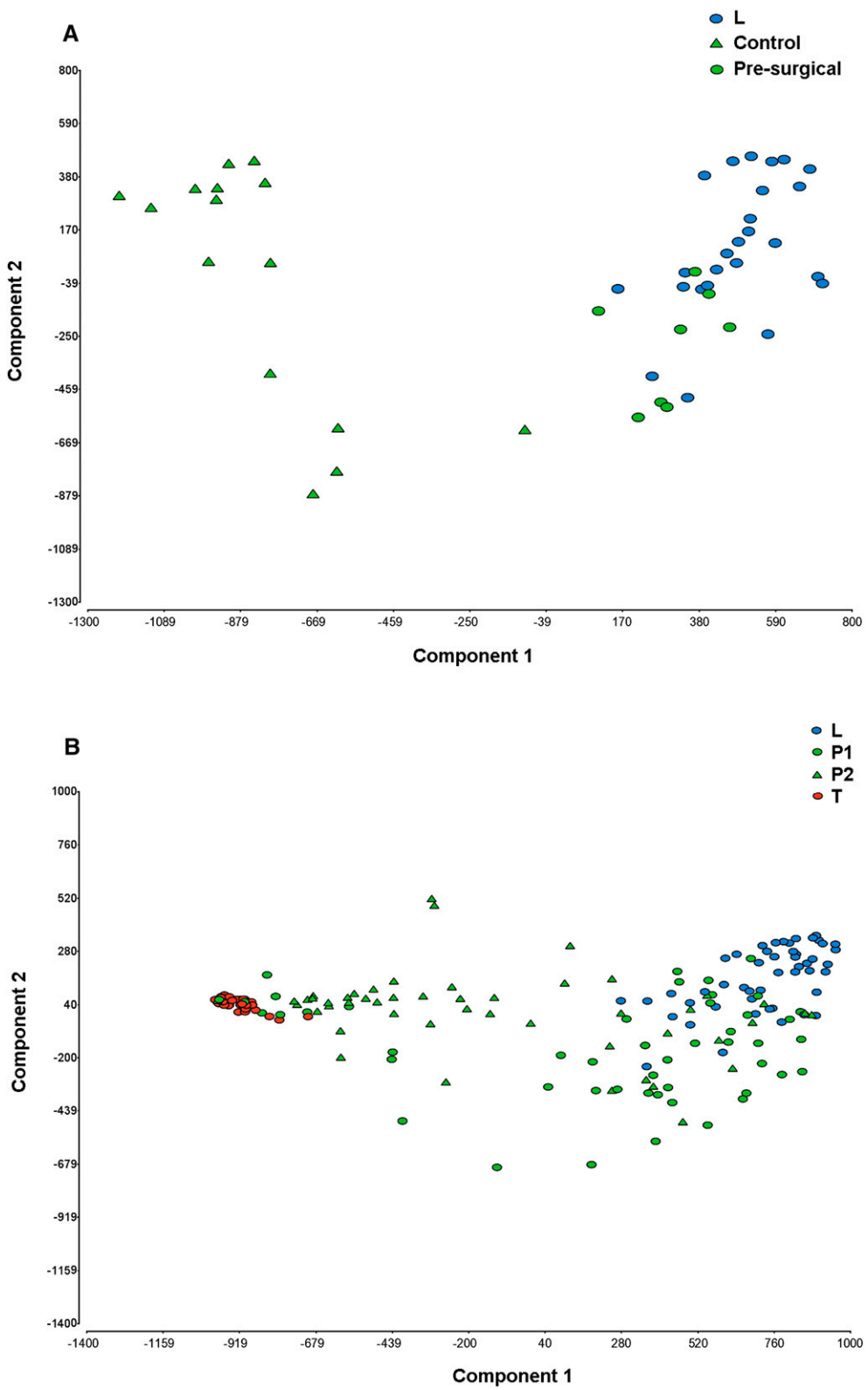


Figure 2. Principal component analysis (PCA) of SNP/CN markers separates plasma DNA of presurgical breast cancer patients from healthy female controls. (*A*) PCA profiles of 15 presurgical breast cancer patients and eight healthy controls showing clear separation of the plasma DNA profiles. The plasmas of healthy female controls clustered with normal leukocytes (blue circles). (*B*) PCA profiles of 50 patients on follow-up, showing separation of normal leukocytes and tumor DNA, with P1 and P2 samples scattered between these. Control indicates healthy female controls; presurgical, plasma of presurgical breast cancer patients; L, normal leukocyte DNA; P1 and P2, first and second plasma samples of patients on follow-up; and T, FFPE tumor DNA.

Table 2. Amplifications and deletions in plasma and tumor DNA of breast cancer patients

Patient group	Tissue (no.)	Mean total CNVs	Range	Amplification percentage	Mean total CNVs showing amplification	Amplifications range	Deletion percentage	Mean total CNVs showing deletion	Deletions range
Healthy controls	Leukocytes (8)	2922	1885–3769	67%	1961	1421–2325	33%	962	464–1686
	Plasma (8)	3619	3052–4126	59%	2134	1858–2421	41%	1485	1091–2033
Presurgical breast cancer patients	Leukocytes (15)	1549	426–3137	77%	1260	370–2382	23%	362	56–1050
	Plasma (15)	2090	1035–4272	38%	786	369–2134	62%	1304	661–2241
	Tumor (15)	1799	296–4336	70%	1229	201–3057	30%	551	95–1279
Breast cancer patients on follow-up	Leukocytes (50)	1709	134–4062	74%	1261	113–2729	26%	448	21–1333
	Plasma P1 (50)	2365	279–5110	57%	1354	177–2698	43%	1016	39–2587
	Plasma P2 (50)	2265	282–5096	45%	1136	163–2702	55%	1186	66–2789
	Tumor (40)	389	201–950	65%	311	125–1999	35%	156	38–695

Showing mean total CNVs and percentage of amplification and deletion for each patient group by sample and the range (lowest to highest values). CNVs derived based on the segmentation method, 50 consecutive markers (SNP and/or CN), *P*-value cut-off of <0.0001, and a signal-to-noise ratio of 0.5.

(38%) showed amplification at 16p12.3. In contrast, there was no amplification seen in any of the 56 normal leukocyte DNA samples, confirming the importance of the selected CNVs (Fig. 4). As the HER2 status of the primary tumor was known for many patients, we reviewed the results for the *HER2* gene interval. The normal leukocyte DNA samples showed mostly diploid CN (mean CN state = 2.0), whereas the tumor and plasma samples of HER2 3+ patients showed a mean CN state of 2.5–3.0 by Gaussian smooth signal, indicating a low level of amplification (Page et al. 2011).

Plasma SNP/CNV changes with time

There was a significant difference in SNP concordance between the first and second paired plasma samples ($P = 0{\cdot}0002$; paired *t*-test) of the 50 patients on follow-up, and all patients showed changes in CNVs between the first and second plasma samples. Thirty patients showed a decrease and 20 patients an increase in the total number of CNVs detected. Some CNVs were common between paired plasma samples (common amplification is shown in Fig. 3), but there were also many sample-specific CNVs detected (Supplemental Fig. 2). Eight patients relapsed 2–9 yr after diagnosis. For these patients, the second plasma sample surveyed was the last blood sample taken prior to relapse. These patients showed the most CNV changes with time in plasma DNA. Figure 5 illustrates the CNV gains and losses in one patient who relapsed. There was an increase in the number of CNVs between the first (1386) and second (2482) plasma sample and a change from gain to loss at multiple CNVs. Two of the eight patients who relapsed were triple-negative; the rest were ER-positive. However, there was no obvious correlation between CNVs and relapse other than for the intervals noted previously (Table 4).

Detection of LOH

There was wide heterogeneity in LOH detected both between patients and samples. The extent of the LOH overlap between paired plasma and tumor DNA also varied widely between patients, ranging from 10%–35% overlap. When we looked at LOH within exons, there were 36 LOH regions found overlapping with genes in two or more of the 15 presurgical patients' plasma samples, and 34 LOH regions found overlapping with genes in two or more plasmas of the 50 patients on follow-up (Supplemental Table 5). There was generally more LOH detected in the node-positive patients than T_1N_0 patients and an overall increase in LOH detected between P1 and P2 samples. Combining CN and LOH data showed that a small percentage of CN segments called (1.47%) exhibited copy-neutral LOH.

Discussion

We demonstrate for the first time that over a decade since diagnosis there is evidence of specific tumorigenic CNVs within cfDNA in plasma during routine follow-up of breast cancer patients.

At the present time, there are no accepted methods, using body fluids, that can reliably distinguish between patients with primary breast cancer and healthy controls, nor is there a method for monitoring patients after the completion of surgery, radiation therapy, and chemotherapy. Several groups, including ourselves, have reported that measuring circulating tumor cells (CTCs), bone marrow, or total circulating DNA can help in this regard (Meng et al. 2004; Braun et al. 2005; Slade et al. 2005; Schwarzenbach et al. 2009), but we and others only find one to two cells in 7.5 mL blood intermittently present, and other tests aimed at either increasing the number of cells detected or quantifying DNA size or other more straightforward characteristics thus far have not proved sufficiently reliable for clinical use. The results of this study suggest plasma cfDNA analysis is potentially more informative.

First, results from patients on follow-up are striking, since up to 12 yr after diagnosis many patients clearly have cfDNA in plasma with specific CNVs that mirror those in their primary cancer (Fig. 3), despite the fact that they have no clinically evident recurrent disease. Second, concordance of SNP genotype calls from whole-genome array analysis distinguished between patients with primary breast cancer and healthy controls ($P < 0{\cdot}0001$) (Figs. 1, 2) and between preoperative cancer patients and patients on follow-up who have had surgery and treatment ($P = 0{\cdot}0016$). Third, the paired plasma and leukocytes from the healthy female controls showed the highest concordance of SNP genotype calls (Fig. 1), as would be expected when the cfDNA in plasma DNA is derived from normal cells. This confirms that a representative genome sample can be obtained from plasma, even when the DNA isolated is in limiting amounts. Although whole-genome amplification (WGA) was necessary due to limiting template DNA, we pooled triplicate WGA samples to reduce the imbalance in allele ratios and differential amplification of different parts of the genome (Rook et al. 2004). In addition, we confirmed the reproducibility of the SNP array approach by QC call-rate ($P = 0.0001$), MAPD ($P = 0.0005$), and mean Spearman correlation for 13 repeated samples; hence, the results show that it is possible to reliably interrogate the entire circulating genome in a single experiment.

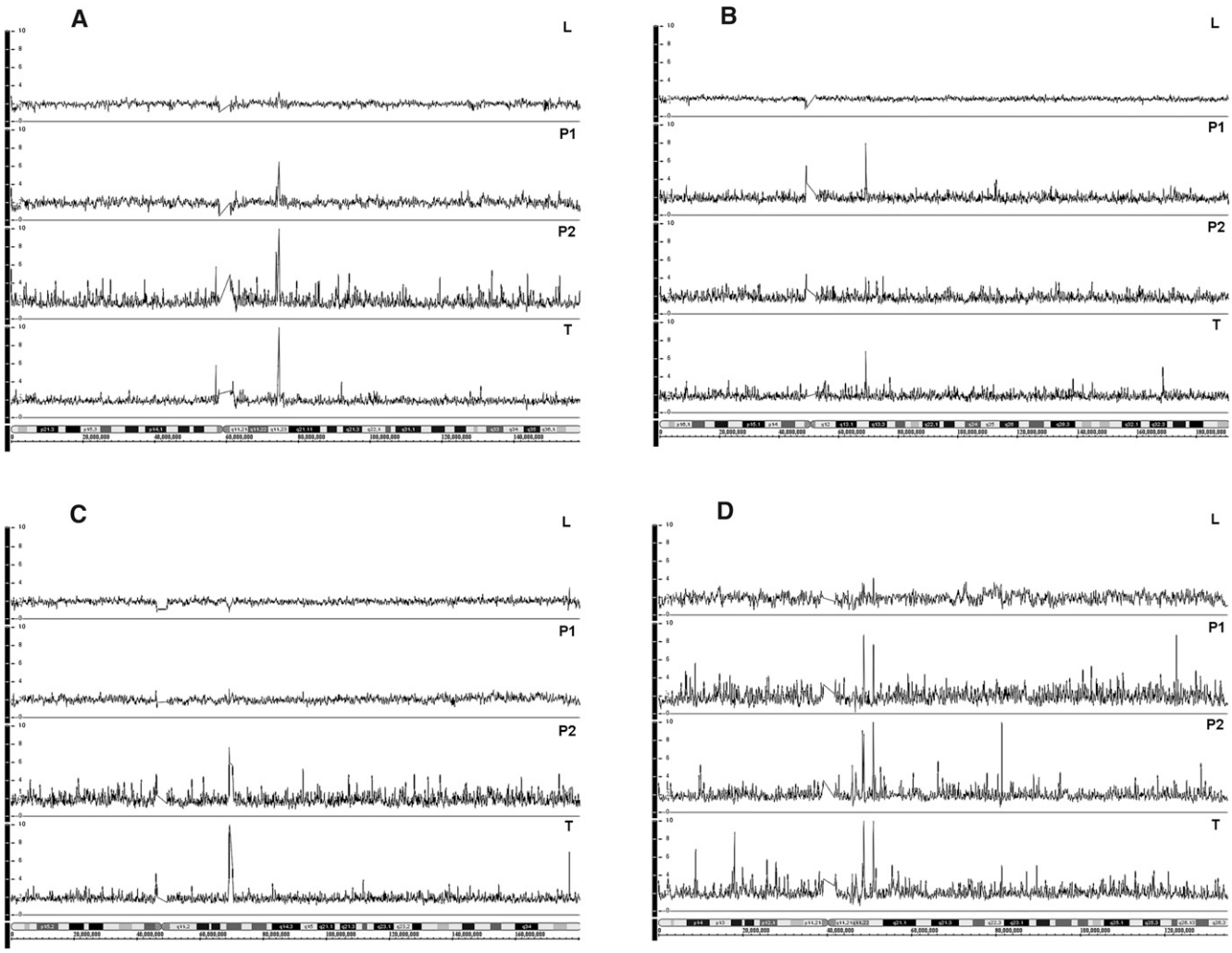


Figure 3. High-level amplification in plasma and primary tumor DNA of breast cancer patients on follow-up. (*A*) Patient 44, amplification at 7q11.23 in tumor P1 and P2; (*B*) patient 27, amplification at 4q13.2 in tumor and P1; (C) patient 35, amplification at 5q13.2 in tumor and P2; and (*D*) patient 47, amplification at 10q11, showing two clear peaks (10q11.22 and 10q11.23) in tumor, P1 and P2. *Top* to *bottom*: L indicates normal leukocyte DNA; P1 and P2, paired plasma DNA samples; and T, FFPE tumor DNA.

One important feature emerging from previous studies is the observation that tumor-specific DNA as evidenced by LOH and methylation (Levenson 2007) can persist in plasma following treatment. This finding provided the impetus for us to attempt to characterize the entire circulating genome from plasma. Compelling research, including recent parallel sequencing data, also indicates that the cancer genome can change with the evolution of metastatic disease (Gangnus et al. 2004; Ding et al. 2010), thus providing us with another reason to suppose that changes in plasma DNA might provide us with an important indicator of impending onset of life-threatening overt metastatic relapse. When we compared paired plasmas from 50 patient on follow-up, some 25 samples had an essentially normal profile, confirmed by PCA, although the remainder did not. A "normal" SNP profile would be expected if these patients are cured. Conversely, dominant oncogenes, persisting in plasma, could potentially transform stem cells in target organs and initiate metastases, as suggested by animal and in vitro cell models, the "genometastasis hypothesis" (García-Olmo et al. 1999, 2010). In support of this, we saw the most striking changes in CNVs between the P1 and P2 samples of the eight patients who had relapsed (Fig. 5), although this is too small a group to reliably identify the specific markers predictive of relapse. As with other studies concerning cfDNA, we did not try to separate DNA derived from normal cells from tumor or micrometastases prior to analysis. The CN data were supported by LOH data, which also showed an overall increase in LOH detected between P1 and P2 samples of the patients on follow-up with evidence of infrequent copy-neutral LOH. The complex CNV and LOH profiles identified from plasma suggests a mixed origin of this circulating DNA.

One other critical finding that we have made is that plasma DNA characterization may provide important information for clinicians in choosing subsequent therapies; we are able to demonstrate amplified areas of the genome, thus potentially indicating which gene products to target. There were more amplifications than deletions in most plasmas and tumors, as was found in a recent SNP 6.0 analysis of 17 different human embryonic stem cell lines (Närvä et al. 2010). In our data, by applying a Gaussian

Table 3. Chromosomal intervals with CNVs showing common amplification in plasma and tumor

		Frequency of amplification in group by sample type (%)						
		Healthy controls		Presurgical breast cancer		Breast cancer on follow-up		
		L	P	P	T	P1	P2	T
Interval	Overlapping genes	*n* = 8	*n* = 8	*n* = 15	*n* = 15	*n* = 50	*n* = 50	*n* = 40
1p36.33	*LOC100132287**	4 (50)	5 (62.5)	15 (100)	10 (66.66)	35 (70)	44 (88)	39 (98)
1p36.21	*C1orf158, LOC440563* 440563*, PRAMEF2 3 4 5 7 8 9 10 12 13 14 15 16 19 20 21 22*	0 (0)	0 (0)	8 (53)	1 (6.66)	2 (4)	12 (24)	3 (8)
1q21.1	*PPIAL4A, PPIAL4C*	3 (37.5)	4 (50)	12 (80)	9 (60)	33 (66)	37 (74)	38 (95)
2p11.2	*ANKRD36BP2, EIF2AK3, RPIA*	0 (0)	0 (0)	9 (60)	3 (20)	14 (28)	33 (66)	10 (25)
4q13.2	***TMPRSS11BNL, TMPRSS11B, UGT2B 15 17, YTHDC1***	0 (0)	0 (0)	4 (27)	4 (26.66)	9 (18)	13 (26)	37 (93)
p11	—	0 (0)	0 (0)	0 (0)	2 (13.33)	4 (8)	2 (4)	12 (30)
5q13.2	***SERF1A, SERF1B, NAIP, SMA5*, SMN1 2***	0 (0)	0 (0)	9 (60)	6 (40)	18 (36)	20 (40)	36 (90)
7p11.2	*MIR4283-2*	0 (0)	0 (0)	2 (13)	3 (20)	2 (4)	7 (14)	18 (45)
7q11.21	—	0 (0)	0 (0)	6 (40)	3 (20)	13 (26)	14 (28)	14 (35)
7q11.23	***CLIP2, GATSL1, GTF2IRD1 2, LOC100093631, NCF1, NCF1C, PMS2P5 6, SPDYE8P, STAG3L2, WBSCR16***	0 (0)	2 (25)	13 (87)	6 (40)	21 (42)	38 (78)	40 (100)
8p23.1	***DEFB130, DUB3, FAM66A, FAM66D, FAM85A, FAM86B1, LOC392196*, ZNF705D***	0 (0)	0 (0)	4 (27)	1 (6.66)	12 (24)	13 (26)	15 (38)
9p11.2	*FAM27A, LOC100132167**	2 (25)	1 (12.5)	14 (93)	10 (66.66)	26 (52)	40 (80)	39 (98)
9q12	—	1 (12.5)	1 (12.5)	13 (87)	8 (53.33)	20 (40)	37 (74)	40 (100)
10q11.22	***BMS1P5, FRMPD2P1, GDF2 10, PTPN20B, RBP3, ZNF488***	1 (12.5)	0 (0)	14 (93)	6 (40)	25 (50)	33 (66)	37 (93)
10q11.23	***AGAP7 8, C10orf53, CHAT, OGDHL, PARG***	2 (25)	3 (37.5)	14 (93)	6 (40)	30 (60)	36 (72)	40 (100)
14q11.1	—	1 (12.5)	1 (12.5)	8 (53)	6 (40)	23 (46)	27 (54)	28 (70)
14q11.2	—	0 (0)	0 (0)	15 (100)	6 (40)	0 (0)	36 (72)	40 (100)
14q32.33	*AHNAK2, BRF1, C14orf79, CDCA4, CRIP1 2, GPR132, JAG2, KIAA0284, LOC647310*, MGC23270*, MTA1, NUDT14, PACS2*	1 (12.5)	0 (0)	10 (67)	3 (20)	5 (10)	15 (30)	13 (33)
15q11.2	*NF1P1, POTEB*	0 (0)	0 (0)	5 (33)	1 (6.66)	8 (16)	16 (32)	23 (58)
15q25.2	***GOLGA6L4 5 9, LOC440297* 727849* 80154*, RPS17***	1 (12.5)	0 (0)	10 (67)	7 (46.67)	21 (42)	30 (60)	34 (85)
16p12.3	***ABCC6P1, NOMO2***	2 (25)	5 (62.5)	12 (80)	6 (40)	19 (38)	30 (60)	35 (88)
16p11.2	*ALDOA, C16orf92, DOC2A, FAM57B, HIRIP3, INO80E, KCTD13, TAOK2, TMEM219*	0 (0)	0 (0)	6 (40)	5 (33.33)	11 (22)	30 (60)	20 (50)
19p13.3	*FAM138F, FLJ45445*, OR4F17, WASH5P*	2 (25)	3 (37.5)	12 (80)	7 (46.67)	39 (76)	41 (82)	36 (90)

Frequency of amplification in group by sample type (%), defined by applying a Gaussian smoothed signal threshold of >6.0 (CNV size and number of markers are detailed in Supplemental Table 3). Only intervals showing amplification in more than one group are included. Intervals highlighted in bold showed amplification in >90% of tumor DNA samples of patients on follow-up and >25% of matched plasma samples when applying a lower smoothed signal threshold of >4.0 (details shown in Supplemental Table 4). Gene nomenclature (alphabetically for each interval) taken from HGNC. Genes with an asterisk only appear in the RefSeq (NCBI, http://www.ncbi.nlm.nih.gov/gene/) gene database. L indicates normal leukocytes; P, plasma; and T, tumor DNA.

Table 4. Linear regression analysis for 23 CNVs

	Cancer versus controls				Relapsed status			Size and nodal status			Surgery type		Endocrine therapy (tamoxifen/ arimidex)			Receptor status				
	Healthy control (P) vs. on follow-up (P1/P2/T)			Healthy control (P) vs. presurgical plasma	Relapsed vs. nonrelapsed			T > 20 mm or node-positive vs. T1N0 (T ≤ 20 mm)			Mastectomy vs. wide local excision		Therapy vs. none			HER2-positive vs. HER2-negative	ER-positive vs. ER-negative	PR-positive vs. PR-negative	Triple-negative vs. receptor-positive	
Interval	P1	P2	T	P	P1	P2	T	P1	P2	T	P1	P2	P1	P2	T	P1/P2/T	P2	P1/P2/T	P1	T
1p36.33	NS	NS	NS	0.0446	0.0019	0.0002	<0.0001	NS	NS	0.0087	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
1p36.21	NS	NS	NS	0.0088	NS	NS	NS	NS	NS	NS	NS	NS	NS	0.0322	NS		NS		NS	NS
1q21.1	NS	NS	NS	NS	NS	0.0047	<0.0001	0.0019	0.0105	0.0087	NS	NS	NS	NS	NS		NS		NS	NS
2p11.2	NS	NS	NS	0.0286	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS		NS		NS	NS
4q13.2	NS	NS	0.0002	NS	NS	NS	NS	NS	NS	NS	NS	0.0309	NS	NS	NS		NS		NS	NS
5p11	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	0.0324	NS		0.0143		NS	0.0164
5q13.2	0.0417	0.0271	<0.0001	0.0033	NS	NS	NS	NS	NS	NS	0.0457	0.0368	NS	NS	NS		NS		NS	NS
7p11.2	NS	NS	0.0338	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS		NS		NS	NS
7q11.21	NS	NS	NS	0.0386	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS		NS		NS	NS
7q11.23	NS	NS	NS	NS	NS	NS	NS	NS	0.0116	NS	NS	NS	NS	NS	NS		NS		NS	NS
8p23.1	NS	NS	0.0099	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	0.0170		NS		NS	NS
9p11.2	0.0293	<0.0001	0.0004	<0.0001	NS	0.0203	<0.0001	NS	NS	0.0031	NS	NS	NS	NS	NS		NS		NS	NS
9q12	NS	0.0088	0.0032	0.0083	0.0284	0.0173	<0.0001	NS	NS	0.0105	NS	NS	NS	NS	NS		NS		0.0306	NS
10q11.22	0.0381	0.0057	0.0004	<0.0001	NS	0.0115	NS	0.0487	NS	NS	NS	NS	NS	NS	NS		NS		NS	NS
10q11.23	NS	NS	NS	0.0150	NS	0.0287	NS	0.0109	NS	NS	NS	NS	0.0115	NS	NS		NS		NS	NS
14q11.1	NS	0.0032	<0.0001	NS	NS	0.0001	0.0372	NS	0.0105	0.0357	NS	NS	NS	NS	NS		NS		NS	NS
14q11.2	0.0271	<0.0001	<0.0001	NS	NS	0.0009	NS	NS	NS	NS	NS	NS	NS	NS	NS		NS		NS	NS
14q32.33	NS	NS	NS	0.0116	NS	NS	NS	NS	NS	NS	NS	0.0320	NS	NS	NS		NS		NS	NS
15q11.2	NS	NS	0.0039	NS	NS	0.0268	NS	NS	NS	NS	NS	NS	NS	NS	NS		NS		NS	NS
15q25.2	0.0215	0.0002	0.0002	0.0010	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS		NS		NS	NS
16p12.3	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS		NS		NS	NS
16p11.2	NS	0.0012	0.0271	NS	NS	0.0022	NS	NS	NS	NS	NS	NS	0.0275	NS	NS		NS		NS	NS
19p13.3	NS	0.0057	0.0169	0.0432	NS	0.0095	0.0376	NS	0.0347	0.0308	NS	NS	NS	0.0373	NS		NS		NS	NS

Based on the presence or absence of a peak with a CN > 6.0 by Gaussian smooth signal. Significance ($P < 0.05$) suggests an interval contributes significantly to tumor phenotype, type of surgery, and/or type of therapy.

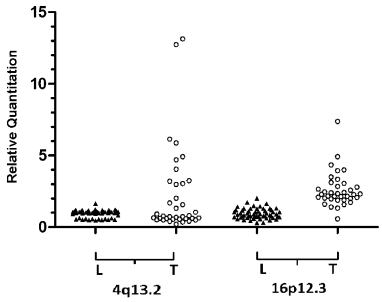


Figure 4. Detection of amplification at two CNV intervals in tumor DNA. Real-time qPCR was used to analyze locus-specific assays that map within the CNVs at 4q13.2 and 16p12.3 using unamplified template DNA. Each amplicon was measured relative to the mean of four reference loci, by relative quantitation. Unamplified tumor DNA from 37 primary breast cancers (from an independent series) was compared with 56 normal leukocyte DNA samples. Amplification (RQ > 2.5) was detected in tumor DNA only.

smoothed signal threshold of >6.0, we identified 23 chromosomal intervals (Table 3) showing common amplification in plasma and tumor of both the presurgical breast cancer patients and patients on follow-up. Some of these appear to discriminate between node-positive and node-negative patients, ER-positive cancer, triple-negative cancer and presence of relapse (Table 4) and may therefore be extremely helpful in deciding on chemotherapy. Applying a lower threshold >4.0 revealed many more CNVs and more frequent amplification in seven of the 23 chromosomal intervals (Supplemental Table 4). We also saw amplification in 69.23% of patients' plasma samples in four markers from the 10-kb interval that spans *ZNF703*, although amplification was also seen in 50.00% of plasmas from the healthy controls and 32.30% of patients' normal leukocytes. This gene has recently been shown to be a novel oncogene in Luminal B breast cancer (Holland et al. 2011). Overall, the pattern of genomic alteration seen, with focal high-level DNA amplification clustered at several chromosome arms, resembles the "amplifier" or "firestorm" type of DNA CN alterations, detected in previous genomic profiling of breast tumors (Kwei et al. 2010). The CNV of repetitive elements may be important for the five intervals identified that have no known associated gene targets, supporting the only other related study that we are aware of, which focused on repetitive elements in serum of breast cancer patients using next-generation sequencing (Beck et al. 2010). Of note, both studies have shown that there are specific breast cancer–related CNV markers, which could lead to the development of a blood-based test for breast cancer screening and monitoring.

There are many potential gene targets revealed by this genomic profiling of cfDNA (Table 3). A number are of potential interest. Expression of *UGT2B15* (at 4q13.2) has been shown to be up-regulated by 17β-estradiol in MCF-7 breast cancer cells. This gene may normally maintain steroid hormone homeostasis and prevent excessive estrogen signaling (Hu and Mackenzie 2009). Hence deregulation of *UGT2B15* by amplification might have the opposite effect. Neuronal apoptosis inhibitory protein (*NAIP*) at 5q13.2 increases in vitro and in vivo in response to androgen deprivation therapy and may be associated with enhanced survival of prostate cancers (Chiu et al. 2010). The *DUB3* gene at 8p23.1 has recently been shown to be a major regulator of *CDC25A* (Pereg et al. 2010), which is overexpressed in many human cancers. DUB3 knockdown significantly reduced growth of breast tumor xenografts in nude mice. Hence, amplification of *DUB3* might lead to *CDC25A* overexpression and increased oncogenesis.

The CNV detected at 14q32.33 contains a number of gene targets of potential interest. Amplification at this interval was found in 67% of the presurgical breast cancer patients' plasma samples but was absent from the healthy controls, which suggests this is a suitable interval for a more targeted study. The *BRF1* gene encodes a transcription factor of the RNA polymerase III complex, which, when overexpressed, can transform cells in vitro and cause tumor formation in vivo (Berns 2008). Metastasis-associated tumor antigen 1 (*MTA1*), is known to be up-regulated in several cancers and has been shown to lead to the transcriptional repression of BRCA1, with resulting abnormalities in centrosome number and chromosomal instability (Molli et al. 2008). Finally, the expression of the Notch ligand JAG2 has been correlated recently with vascular development and angiogenesis (Pietras et al. 2011). Our future studies will focus on validation of these key gene targets and intervals (Table 3) in plasma cfDNA.

The finding that tumor-specific DNA persists in plasma up to 12 yr after diagnosis, although the patient remains disease free, raises important questions regarding the issue of dormancy in breast cancer. Our own studies, as well as those of other groups, have also shown that rare disseminated tumor cells (DTCs) and CTCs can persist for many years after the end of breast cancer treatment (Slade et al. 2009; Criscitiello et al. 2010). Further, the presence of these few cells represents a balance between replication and cell death, since the half-life of these cells in the plasma is 1–2 h (Meng et al. 2004). Our findings in breast cancer may also apply to other cancers where dormancy is a feature, such as melanoma, non-Hodgkin's lymphoma, and renal cancer; all of these are characterized by the development of late recurrences, and the analysis of plasma could help in the management of these conditions. In as much as plasma DNA in part reflects the nature of dying dormant cells, the information from patient samples could help elucidate the molecular determinants of survival.

These findings now require prospective valuation, preferably as part of ongoing adjuvant studies during the follow-up of a larger group of patients. In conclusion, we have demonstrated that SNP 6.0 array analysis of plasma DNA distinguishes between patients with primary breast cancer and healthy controls and between preoperative cancer patients and those who have had surgery and treatment. We have identified focal high-level DNA amplification in paired tumor and plasma, targeting specific CNVs clustered at several chromosome arms, and have shown that these are detectable in plasma up to 12 yr after diagnosis in patients on follow-up. This finding implies dormancy/minimal residual disease in the majority of patients on follow-up. Our future studies will focus on developing high-throughput approaches to target common CNVs for screening and monitoring.

Methods

Patients and samples

The protocols were approved by the Riverside Regional Ethics Committee and conducted in accordance with the Declaration of Helsinki. All patients gave written informed consent prior to participation.

The samples were blinded for analysis, and the patients understood that the results would not be made available to them. We collected blood samples from 15 women attending a clinic who had just been diagnosed with primary breast cancer and from eight

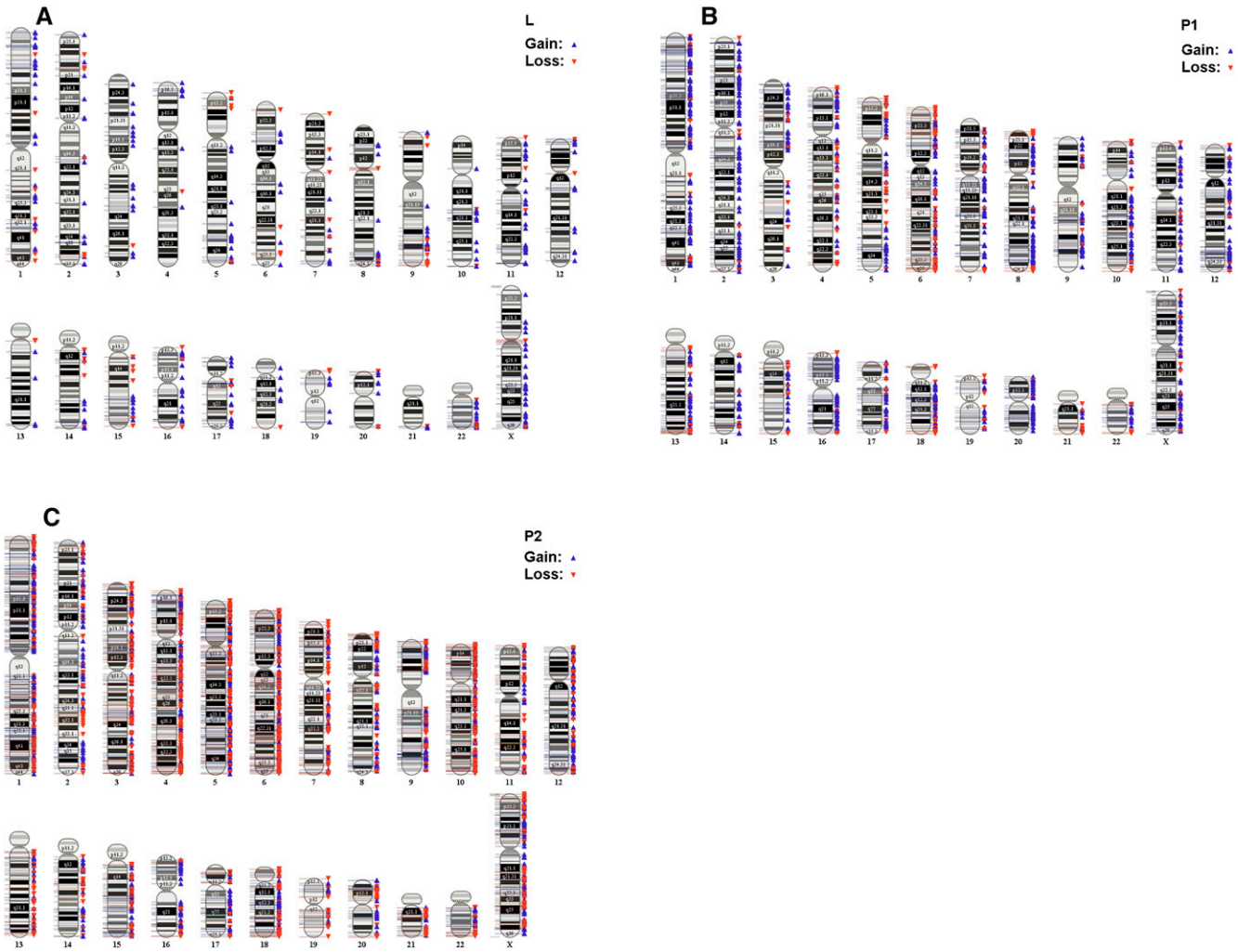


Figure 5. Chromosomal abnormalities in plasma preceding relapse. CNVs based on 50 consecutive markers (SNP and/or CN) and a minimum segment size of 50,000 bp. Example of array karyotypes of cfDNA for one patient preceding relapse: (*A*) normal leukocyte DNA sample, (*B*) P1 cfDNA sample taken 6 yr after diagnosis, and (C) P2 cfDNA taken 1 mo before the patient was diagnosed with metastatic disease. There was a significant increase in CNVs detected between P1 and P2: P1, 387 (79.08%) amplifications and 96 (20.92%) deletions; P2, 1332 (53.67%) amplifications and 1150 (46.33%) deletions.

age-matched healthy female volunteers. We also retrospectively analyzed stored plasma samples from 50 breast cancer patients who had been operated on for breast cancer at least 3 yr previously (Table 1). Eight of these patients developed recurrent disease between 2 and 9 yr after diagnosis.

Following plasma separation by centrifugation at 850*g* for 10 min (×2), plasma and cell pellets were separated and stored at −80°C. For the analysis of tumor samples, hematoxylin and eosin–stained FFPE tissue sections were reviewed, and the foci of tumor cells were isolated by manual microdissection.

DNA extraction, amplification, and SNP 6.0 arrays

DNA was extracted from blood cell pellets, 1 mL plasma, and foci of tumor cells, as described previously (Page et al. 2006; Shaw et al. 2011). WGA was performed in triplicate with the Illustra GenomiPhi V2 DNA Amplification Kit (GE Healthcare Life Sciences) and pooled (Rook et al. 2004). WGA DNA samples were hybridized on Affymetrix GeneChip Human Mapping SNP 6.0 arrays, using the Human mapping SNP 6.0 assay kit following the Genome-Wide Human SNP *Nsp/Sty* 6.0 protocol. Samples were hybridized in two batches only to reduce interassay variability.

Data processing and analysis

The analysis of raw data microarray CEL files was performed using Partek Genomics Suite 6.5, build 6.10.1129 (Partek Inc., http://www.partek.com/) with SNP and QC call-rates used as indicators of sample quality. Genotyping analysis and SNP/CN marker calls were performed using the Birdseed v2 algorithm (Broad Institute, Harvard–Massachusetts Institute of Technology, http://www.broadinstitute.org/mpg/birdsuite/index.html), incorporating regional GC correction. The International HapMap (build 270 na30 r1 a5, International HapMap Project, http://hapmap.ncbi.nlm.nih.gov/) was used as the initial reference model file.

Genomic segmentation was performed using a minimum of 50 markers per segment, *P*-value cut-off of <0.0001, and a signal-to-noise ratio of 0.5. Minimum segment sizes of 1000 bp, 50,000 bp,

100,000 bp, and 1,000,000 bp were used for viewing different-sized amplifications and deletions across different samples.

PCA was performed using Partek Genomics Suite 6.5, build 6.10.1129 (Partek). Principal components were determined using a covariance matrix method with normalized eigenvector scaling. An ANOVA *P*-value < 0.0001 (followed in some cases by a Bonferroni-corrected *P*-value < 0.0001 for multiple comparisons) was used to filter out probes of insignificance. In addition, a fold-change larger than |4| was applied to further filter data. LOH using a hidden Markov model (HMM) was also analyzed using this software on a paired basis (matched to lymphocyte) using the following parameters—genomic decay of 1 Mbp, maximum probability of 0.98, genotype error of 0.02—and was filtered using a Hardy-Weinberg Equilibrium *P*-value < 0.001 or <0.0001. The frequency analysis for CNVs was performed using Nexus Copy Number 5.1 Discovery Edition (BioDiscovery Inc., http://www.biodiscovery.com/).

Statistics

Data were analyzed using GraphPad Prism 5.0. Paired, two-tailed *t*-tests were used as appropriate. Nonparametric tests were used for further analysis; unpaired *t*-tests and one-way ANOVAs were followed by Mann-Whitney and Kruskal-Wallis tests, respectively. For all statistical analyses, the α value was set at 0.05.

Real-time qPCR

To confirm amplification at 4q13.2 and 16p12.3 identified by SNP 6.0 array, DNA samples were analyzed in triplicate by real-time qPCR using locus-specific assays designed in house in a 10 μL reaction volume. Reactions were run on an Applied Biosystems thermal cycler (Step One Plus) and analyzed with Step One v2.1 software and Microsoft Excel. The ΔCt was determined (average Ct value of the target locus minus the mean Ct value of four independent reference loci) and used to calculate the $\Delta\Delta$Ct for each DNA sample, using the mean relative quantitation (RQ) value derived from normal human genomic DNA (Roche) as the experimental calibrator. RQ values were calculated as $2^{-\Delta\Delta Ct}$ as described previously (Page et al. 2011).

Data access

All microarray raw and processed data files have been deposited at ArrayExpress (http://www.ebi.ac.uk/arrayexpress) under accession no. E-MTAB-624.

Acknowledgments

The study was supported by grants from Cancer Research UK (BIDD), the Cancer Research UK/Department of Health Experimental Cancer Medicine Centres (at Leicester University and Imperial College), and the BRC (Imperial College). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We thank Professor Sir Alec Jeffreys, Dr. J. Howard Pringle, and Dr. Tim Gant (Leicester) for helpful discussion. Probe preparation and array hybridizations were carried out as a service by Almac Diagnostics following appropriate quality-control checks.

Authors' contributions: J.A.S. jointly conceived of and supervised the study and prepared the manuscript with R.C.C.; K.P., D.G., J.B., and N.H. carried out the study with contributions from C.R., R.P., C.P., J.S., and S.C.; K.B. and K.P. performed all bioinformatics; and R.A.W. and C.R. reviewed FFPE tumor sections. All authors contributed to the final manuscript.

References

Beck J, Urnovitz HB, Mitchell WM, Schütz E. 2010. Next generation sequencing of serum circulating nucleic acids from patients with invasive ductal breast cancer reveals differences to healthy and nonmalignant controls. *Mol Cancer Res* **3:** 335–342.

Berns A. 2008. A tRNA with oncogenic capacity. *Cell* **133:** 29–30.

Braun S, Vogl FD, Naume B, Janni W, Osborne MP, Coombes RC, Schlimok G, Diel IJ, Gerber B, Gebauer G, et al. 2005. A pooled analysis of bone marrow micrometastasis in breast cancer. *N Engl J Med* **353:** 793–802.

Chambers AF, Goss PE. 2008. Putative growth characteristics of micrometastatic breast cancer. *Breast Cancer Res* **10:** 114. doi: 10.1186/bcr2197.

Chiu HH, Yong TM, Wang J, Wang Y, Vessella RL, Ueda T, Wang YZ, Sadar MD. 2010. Induction of neuronal apoptosis inhibitory protein expression in response to androgen deprivation in prostate cancer. *Cancer Lett* **292:** 176–185.

Criscitiello C, Sotiriou C, Ignatiadis M. 2010. Circulating tumor cells and emerging blood biomarkers in breast cancer. *Curr Opin Oncol* **22:** 552–558.

Demicheli R, Terenziani M, Bonadonna G. 1998. Estimate of tumor growth time for breast cancer local recurrences: rapid growth after wake-up? *Breast Cancer Res Treat* **51:** 133–137.

Ding L, Ellis MJ, Li S, Larson DE, Chen K, Wallis JW, Harris CC, McLellan MD, Fulton RS, Fulton LL, et al. 2010. Genome remodelling in a basal-like breast cancer metastasis and xenograft. *Nature* **464:** 999–1005.

Early Breast Cancer Trialists' Collaborative Group. 2005. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet* **365:** 1687–1717.

Gangnus R, Langer S, Breit E, Pantel K, Speicher MR. 2004. Genomic profiling of viable and proliferative micrometastatic cells from early-stage breast cancer patients. *Clin Cancer Res* **10:** 3457–3464.

García-Olmo D, García-Olmo DC, Ontañón J, Martinez E, Vallejo M. 1999. Tumor DNA circulating in the plasma might play a role in metastasis. The hypothesis of the genometastasis. *Histol Histopathol* **4:** 1159–1164.

García-Olmo DC, Domínguez C, García-Arranz M, Anker P, Stroun M, García-Verdugo JM, García-Olmo D. 2010. Cell-free nucleic acids circulating in the plasma of colorectal cancer patients induce the oncogenic transformation of susceptible cultured cells. *Cancer Res* **70:** 560–567.

Holland DG, Burleigh A, Git A, Goldgraben MA, Perez-Mancera PA, Chin SF, Hurtado A, Bruna A, Ali HR, Greenwood W, et al. 2011. *ZNF703* is a common Luminal B breast cancer oncogene that differentially regulates luminal and basal progenitors in human mammary epithelium. *EMBO Mol Med* **3:** 167–180.

Hu DG, Mackenzie PI. 2009. Estrogen receptor α, Fos-related antigen-2, and c-Jun coordinately regulate human UDP glucuronosyltransferase 2B15 and 2B17 expression in response to 17β-estradiol in MCF-7 cells. *Mol Pharmacol* **76:** 425–439.

Huang ZH, Li LH, Hua D. 2006. Quantitative analysis of plasma circulating DNA at diagnosis and during follow-up of breast cancer patients. *Cancer Lett* **243:** 64–70.

Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, Scherer SW, Lee C. 2004. Detection of large-scale variation in the human genome. *Nat Genet* **36:** 949–951.

Karrison TG, Ferguson DJ, Meier P. 1999. Dormancy of mammary carcinoma after mastectomy. *J Natl Cancer Inst* **91:** 80–85.

Kwei KA, Kung Y, Salari K, Holcomb IN, Pollack JR. 2010. Genomic instability in breast cancer: pathogenesis and clinical implications. *Mol Oncol* **3:** 255–266.

Levenson VV. 2007. Biomarkers for early detection of breast cancer: What, when, and where? *Biochim Biophys Acta* **1770:** 847–856.

Meltzer A. 1990. Dormancy and breast cancer. *J Surg Oncol* **43:** 181–188.

Meng S, Tripathy D, Frenkel EP, Shete S, Naftalis EZ, Huth JF, Beitsch PD, Leitch M, Hoover S, Euhus D, et al. 2004. Circulating tumor cells in patients with breast cancer dormancy. *Clin Cancer Res* **10:** 8152–8162.

Molli PR, Singh RR, Lee SW, Kumar R. 2008. MTA1-mediated transcriptional repression of BRCA1 tumor suppressor gene. *Oncogene* **27:** 1971–1980.

Murray C. 1995. Tumor dormancy: not so sleepy after all. *Nat Med* **1:** 117–118.

Närvä E, Autio R, Rahkonen N, Kong L, Harrison N, Kitsberg D, Borghese L, Itskovitz-Eldor J, Rasool O, Dvorak P, et al. 2010. High-resolution DNA analysis of human embryonic stem cell lines reveals culture-induced copy number changes and loss of heterozygosity. *Nat Biotechnol* **4:** 371–377.

Oldridge DA, Banerjee S, Setlur SR, Sboner A, Demichelis F. 2010. Optimizing copy number variation analysis using genome-wide short sequence oligonucleotide arrays. *Nucleic Acids Res* **38:** 3275–3286.

Page K, Powles T, Slade MJ, Tamburo de Bella M, Walker RA, Coombes RC, Shaw JA. 2006. The importance of careful blood processing in isolation of cell-free DNA. *Ann N Y Acad Sci* **1075:** 313–317.

Page K, Hava N, Ward B, Brown J, Guttery DS, Ruangpratheep C, Blighe K, Sharma A, Walker RA, Coombes RC, et al. 2011. Detection of *HER2* amplification in circulating free DNA in patients with breast cancer. *Br J Cancer* **104:** 1342–1348.

Pereg Y, Liu BY, O'Rourke KM, Sagolla M, Dey A, Komuves L, French DM, Dixit VM. 2010. Ubiquitin hydrolase Dub3 promotes oncogenic transformation by stabilizing Cdc25A. *Nat Cell Biol* **12:** 400–406.

Pietras A, von Stedingk K, Lindgren D, Påhlman S, Axelson H. 2011. JAG2 induction in hypoxic tumor cells alters Notch signaling and enhances endothelial cell tube formation. *Mol Cancer Res* **9:** 626–636.

Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, et al. 2006. Global variation in copy number in the human genome. *Nature* **444:** 444–454.

Rook MS, Delach SM, Deyneko G, Worlock A, Wolfe JL. 2004. Whole genome amplification of DNA from laser capture-microdissected tissue for high-throughput single nucleotide polymorphism and short tandem repeat genotyping. *Am J Pathol* **164:** 23–33.

Schwarzenbach H, Pantel K, Kemper B, Beeger C, Otterbach F, Kimmig R, Kasimir-Bauer S. 2009. Comparative evaluation of cell-free tumor DNA in blood and disseminated tumor cells in bone marrow of patients with primary breast cancer. *Breast Cancer Res* **11:** R71. doi: 10.1186/bcr2404.

Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, Månér S, Massa H, Walker M, Chi M, et al. 2004. Large-scale copy number polymorphism in the human genome. *Science* **305:** 525–528.

Shaw JA, Brown J, Coombes RC, Jacob J, Payne R, Lee B, Page K, Hava N, Stebbing J. 2011. Circulating tumor cells and plasma DNA analysis in patients with indeterminate early or metastatic breast cancer. *Biomarkers Med* **5:** 87–91.

Slade MJ, Singh A, Smith BM, Tripuraneni G, Hall E, Peckitt C, Fox S, Graham H, Lüchtenborg M, Sinnett HD, et al. 2005. Persistence of bone marrow micrometastases in patients receiving adjuvant therapy for breast cancer: Results at 4 years. *Int J Cancer* **114:** 94–100.

Slade MJ, Payne R, Riethdorf S, Ward B, Zaidi SA, Stebbing J, Palmieri C, Sinnett HD, Kulinskaya E, Pitfield T, et al. 2009. Comparison of bone marrow, disseminated tumour cells and blood-circulating tumour cells in breast cancer patients after primary treatment. *Br J Cancer* **100:** 160–166.

Stroun M, Anker P, Lyautey J. 1987. Isolation and characterization of DNA from the plasma of cancer patients. *Eur J Cancer Clin Oncol* **23:** 707–712.

Wang ZC, Lin M, Wei LJ, Li C, Miron A, Lodeiro G, Harris L, Ramaswamy S, Tanenbaum DM, Meyerson M, et al. 2004. Loss of heterozygosity and its correlation with expression profiles in subclasses of invasive breast cancers. *Cancer Res* **64:** 64–71.

Zanetti-Dällenbach R, Wight E, Fan AX, Lapaire O, Hahn S, Holzgreve W, Zhong XY. 2008. Positive correlation of cell-free DNA in plasma/serum in patients with malignant and benign breast disease. *Anticancer Res* **28:** 921–925.

Received April 20, 2011; accepted in revised form August 16, 2011.