THE UNIVERSITY OF QUEENSLAND

AUSTRALIA

Intratumour heterogeneity in the development and progression of breast cancer

Jamie Rose Kutasovic

B Sc (Hons 1st class) University of Queensland (2011)

B App Sci Queensland University of Technology (2010)

A thesis submitted for the degree of Doctor of Philosophy at

The University of Queensland in 2015

School of Medicine

Abstract

The extent of intratumour heterogeneity and its clinical implications are poorly understood. This thesis addresses this theme of tumour heterogeneity with a broad focus on invasive lobular carcinomas of the breast. More specifically the focus will be on studying the various mechanisms of deregulation of E-cadherin, the morphological and phenotypic variation mixed ductal lobular carcinomas, and finally heterogeneity of metastatic progression.

Invasive lobular carcinoma (ILC) is the second most common breast cancer subtype, accounting for 10-15% of the all breast cancers, and is the most common 'special type'. Biologically, the most distinguishing feature of ILC is the loss of the cellular adhesion molecule E-cadherin (absent in up to 90% of ILC), and this is considered fundamental to development of the characteristic infiltrative growth pattern. The deregulation of E-cadherin is not fully understood, and so this thesis aims to investigate the mechanisms of E-cadherin deregulation in tumourigenesis. Co-localised with E-cadherin at the cell membrane are a series of proteins (including ECT2, RacGAP1 and N-WASP) which regulate the actin cytoskeleton; the expression of these proteins was investigated using immunohistochemistry (IHC) of breast cancer tissue microarrays (**Chapter 3**). Differences in cytoplasmic expression and localisation of these molecules were found between different histological types and clinicopathological parameters that warrant further investigation. It remains unclear whether these molecules contribute to deregulating cell adhesion *in vivo*.

Mixed ductal-lobular carcinomas (MDL) display both ductal and lobular morphology, and are a clear example of intratumour morphological heterogeneity. It is hypothesised that these different components evolve from a common ancestor and diverge following deregulation of the E-cadherin complex. To address this hypothesis, a cohort of 82 MDLs was studied for clinical, morphological and molecular features (**Chapter 4**). Key findings include: i) MDLs more frequently co-exist with ductal carcinoma *in situ* (DCIS) than lobular carcinoma *in situ* (LCIS); ii) the E-cadherin-catenin complex was often normal in the ductal component but deregulated in the lobular component; iii) E-cadherin deregulation was different to that seen in classic ILC, which are typically completely negative for this marker, not aberrant; iv) Epithelial to mesenchymal transition marker expression was not associated with E-cadherin deregulation. Exome sequencing was performed to investigate clonal relationships between the different intratumour morphologies, and identify

mechanisms underlying the change from a ‘ductal’ to a more infiltrative ‘lobular’ growth pattern. Preliminary analysis revealed that i) all morphological components within a case are clonally related; ii) divergence of the morphological components may occur early during tumour evolution (where there are both DCIS and LCIS present) or later during tumour progression (cases with only DCIS detectible); and iii) mutations were identified in genes such as *CDH1* and *ESR1*. The mechanisms driving the change in phenotype are yet to be fully elucidated, but there is significant intertumour heterogeneity and each case may utilise a unique molecular mechanism.

Breast cancers show variability in metastatic patterns depending on the tumour histology, and molecular breast cancer subtype. Of particular interest, is the unique metastatic pattern of ILC, which has the propensity to colonise less common sites such as the gynaecological organs and gastrointestinal tract. To gain a greater understanding of this complex process, we collated a unique cohort of 53 breast cancer cases with metastasis to gynaecological organs (**Chapter 5**). Analysis of clinical and pathology data found that i) patients with gynaecological metastasis (GM) presented at a young age (median of 47 years *vs.* 60 years in the general population); ii) there was a long latency to metastasis (median 6 years); iii) ILCs more frequently spread to these organs (42.6% *vs.* 15% in the general population). Interestingly, 74% of patients with GM had additional sites of metastatic deposit, most commonly gastrointestinal organs and the peritoneum. The primary tumours displayed features associated with a good prognosis: positive for ER, PR, FOXA1 and GATA3; and negative for HER2, basal markers and Ki67 (low). The expression of ER remained unchanged during metastatic progression in most cases; whereas other markers, such as FOXA1 and GATA3 were reduced, and Ki67 and p53 increased during progression. Array comparative genomic hybridisation and targeted DNA sequencing were undertaken to define the genomic changes that occur during progression to gynaecological organs. Within each case, there was evidence of both clonal similarities and divergence between the primary tumour and metastases.

The investigation of intratumour heterogeneity in primary breast cancer and its metastatic progression will lead to a better understanding of the mechanisms driving tumour progression. This thesis employs a variety of techniques to address this issue on a number of levels, including heterogeneity at molecular levels, within a primary tumour type and in the progression to metastasis. Ultimately, understanding intratumour heterogeneity will provide patients with more appropriate treatment strategies.

Declaration by author

This thesis is composed of my original work, and contains no material previously published or written by another person except where due reference has been made in the text. I have clearly stated the contribution by others to jointly-authored works that I have included in my thesis.

I have clearly stated the contribution of others to my thesis as a whole, including statistical assistance, survey design, data analysis, significant technical procedures, professional editorial advice, and any other original research work used or reported in my thesis. The content of my thesis is the result of work I have carried out since the commencement of my research higher degree candidature and does not include a substantial part of work that has been submitted to qualify for the award of any other degree or diploma in any university or other tertiary institution. I have clearly stated which parts of my thesis, if any, have been submitted to qualify for another award.

I acknowledge that an electronic copy of my thesis must be lodged with the University Library and, subject to the policy and procedures of The University of Queensland, the thesis be made available for research and study in accordance with the Copyright Act 1968 unless a period of embargo has been approved by the Dean of the Graduate School.

I acknowledge that copyright of all material contained in my thesis resides with the copyright holder(s) of that material. Where appropriate I have obtained copyright permission from the copyright holder to reproduce material in this thesis.

Publications during candidature

Peer-reviewed papers

McCart Reed AE*, **Kutasovic JR***, Vargas AC, Jayanthan J, Al-Murrani A, Reid LE, Chambers R, Da Silva L, Evans E, Porter A, Papadimos D, Lakhani SR, Simpson PT. "An Epithelial to Mesenchymal Transition programme does not drive the phenotype of Invasive Lobular Carcinomas." *J Pathol,* Under Review.
***Joint first author**

McCart Reed AE*, **Kutasovic JR***, Lakhani SR, Simpson PT. **"**Invasive Lobular Carcinoma of the breast: morphology, biomarkers and 'omics'." *Breast Cancer Research.* 2015, 17:12 (30 January 2015)
***Joint first author**

Kutasovic JR, Sim SYM, McCart Reed AE, Cummings MC, Simpson PT. **"**Intratumour heterogeneity in the progression to breast cancer metastasis." *Cancer Forum*. 2014 July;38(2).

Al-Ejeh F, Simpson PT, Saunus JM, Klein K, Kalimutho M, Shi W, Miranda M, **Kutasovic J**, Raghavendra A, Madore J, Reid L, Krause L, Chenevix-Trench G, Lakhani SR, Khanna KK. "Meta-analysis of the global gene expression profile of triple-negative breast cancer identifies genes for the prognostication and treatment of aggressive breast cancer." *Oncogenesis*. 2014 Apr 21;3:e100.

Cummings MC*, Simpson PT*, Reid LE, Jayanthan J, Skerman J, Song S, McCart Reed AE, **Kutasovic JR**, Morey AL, Marquart L, O'Rourke P, Lakhani SR. "Metastatic progression of breast cancer: insights from 50 years of autopsies." *J Pathol*. 2014 Jan;232(1): 23-31.

McCart Reed AE, Song S, **Kutasovic JR**, Reid LE, Valle JM, Vargas AC, Smart CE, Simpson PT. "Thrombospondin-4 expression is activated during the stromal response to invasive breast cancer." *Virchows Arch.* 2013 Oct;463(4):535-545.

Hosein AN, Song S, McCart Reed AE, Jayanthan J, Reid LE, **Kutasovic JR**, Cummings MC, Waddell N, Lakhani SR, Chenevix-Trench G, Simpson PT. "Evaluating the repair of DNA derived from formalin-fixed paraffin-embedded tissues prior to genomic profiling by SNP-CGH analysis." *Lab Invest*. 2013 Jun;93(6):701-10.

Conference abstracts

Kutasovic JR, McCart Reed AE, Sim SYM, Cummings MC, Lakhani SR, Simpson PT. Breast cancer metastasis to gynaecological organs: impacting young women diagnosed with a luminal tumour subtype. In: 23rd annual RBWH Healthcare Symposium. Royal Brisbane and Women's Hospital. Herston, QLD, Australia, 13-17th October 2014.

Kutasovic JR, McCart Reed AE, Sim SYM, Cummings MC, Lakhani SR, Simpson PT. Breast cancer metastasis to gynaecological organs: impacting young women diagnosed with a luminal tumour subtype. In: 29th International Association for Breast Cancer Research Conference. Manly, NSW, Australia, 14-17 September 2014.

Kutasovic JR, McCart Reed AE, Sim SYM, Cummings MC, Lakhani SR, Simpson PT. Breast cancer metastasis to gynaecological organs: impacting young women diagnosed with a luminal tumour subtype. In: Metastasis Research Society 15th International Biennial Congress. Heidelberg, Germany, 28th June – 1st July 2014.

Kutasovic JR, McCart Reed AE, Vargas AC, Jayanthan J, Reid LE, Da Silva L, Cummings MC, Porter A, Evans E, Lakhani SR, Simpson PT. Abrogated cell-cell adhesion in the phenotypic drift from ductal to lobular breast cancer types. In: Australian Breast Cancer Conference, Walter and Eliza Hall Institute Melbourne, VIC, Australia, 7th-8th November 2013.

McCart Reed AE, Song S, **Kutasovic JR**, Reid LE, Valle J, Vargas AC, Smart CE, Simpson PT. Thrombospondin-4 expression is activated during the stromal response to invasive breast cancer. In: Australia Society for Medical Research Postgraduate Student Conference, Translational Research Institute, Woolloongabba QLD, May 2013

McCart Reed AE, Song S, **Kutasovic JR**, Reid LE, Valle J, Vargas AC, Smart CE, Simpson PT. Thrombospondin-4 expression is activated during the stromal response to

invasive breast cancer. In: 25th Lorne Cancer Conference, Lorne VIC, Australia. 14-17 February 2013.

Da Silva L, McCart Reed AE, **Kutasovic JR**, Vargas AC, Jayanthan J, Reid LE, Fritz B, Cummings MC, Porter A, Evans E, Lakhani SR, Simpson PT. Abrogated cell-cell adhesion in the phenotypic drift from ductal to lobular breast cancer types. In: 24th Lorne Cancer Conference, Lorne VIC, Australia. 9-11 February 2012

Publications included in this thesis

"No publications included".

Contributions by others to the thesis

Several members of the Lakhani Lab made important contributions to the work presented in this thesis. My supervisors Dr Peter Simpson, Dr Amy McCart Reed and Prof Sunil Lakhani made significant contributions to all aspects of this thesis. My principle advisor Dr Peter Simpson greatly contributed to the conception and design of this project, alongside my associated advisors Dr Amy McCart Reed and Prof Sunil Lakhani. Prof Alpha Yap was involved in the conception of Chapter 3 and A/Prof Margaret Cummings was involved in the conception of Chapter 5.

Ethical approval was already in place before commencing my PhD. Patient samples were obtained with the help of A/Prof Margaret Cummings, Michelle Hawkins (assistant to Prof Lakhani), and the Brisbane Breast Bank Manager Kaltin Ferguson. Patients clinical information collected in this thesis was acquired with the help of the labs research nurse Colleen Niland.

This thesis would not have been possible without the help of many Pathologists. A/Prof Margaret Cummings and Drs Sarah Sim, Darina Vuong, Greg Miller, Steven Peyton, Admire Matsika and Lewis Melville provided valuable help with identifying and diagnosing tumour areas, interpretation of immunohistochemical data and answering many of my clinical and pathology related questions.

Current and previous members of the Lakhani lab built some of the TMAs utilised in this thesis. IHC for certain antibodies was performed at Pathology Queensland. Array CGH was performed at the Ramaciotti Centre for Genomics; DNA sequencing was performed at Macrogen and Queensland Centre from Clinical Genomics. Samir Lal performed array comparative genomic hybridisation data analysis. Dr Katia Nones performed the bioinformatics analysis of the exome sequencing data.

My supervisors assisted with analysis and interpretation of the research data and drafting of this thesis. My PhD Milestone Committee members Dr Michelle Hill and A/Prof Kiarash Khosrotehrani provided helpful feedback during my candidature.

Statement of parts of the thesis submitted to qualify for the award of another degree

None.

Acknowledgements

The National Breast Cancer Foundation (NBCF), Program funding from the National Health and Medical Research Council (NHMRC) and the Queensland Health Pathology Service Study Education and Research Trust Fund (QHPS SERTF) funded this work.

I would like to thank Professor Sunil Lakhani for allowing me the opportunity to work in his laboratory, lending his advice and guidance and providing access to patient samples. I would like to express many thanks to my supervisors Dr Peter Simpson and Dr Amy McCart Reed for their help, patience, advice and guidance. I am very grateful for the time and effort they both put in to help me work to the best of my ability, and allowing me to contribute to a number of publications. They have both played a pivotal role in my development as a researcher and as a person.

I thank the pathology registrars for their time reviewing all of the cases for this study, helping interpret pathology reports and teaching me breast pathology. I would like to acknowledge the patients who gave their tissue, as without them this study and all of our labs work wouldn't be possible.

I would like to thank all the members of the Lakhani lab for their assistance with all aspects of my lab work and making the lab a great environment to work in. I am grateful to have been surrounded by intelligent, friendly and supportive colleagues. I would like to thank Dr Jodi Saunus for her advice, ideas and insightful discussion relating to both my PhD and my career. I would like to thank current Lakhani Lab members Lynne Reid, Kaltin Ferguson, Margaret Cummings, Ashwini Ragavendra, Flavia Maturana, Colleen Niland, Samir Lal and previous members Mythily Mariasegaram, Rebecca Johnston, Chanel Smart, Janani Jayanthan, Debra Black and Ana Cristina Vargas who have all provided support, help and advice in a number of different ways.

Finally, the biggest thank you goes to my partner, Henry Younger, and my parents, Maxene and John Kutasovic. I am grateful for their encouragement to do what I love and constantly remind me that hard work pays off. I would have not have been able to do my PhD without their patience and support.

Keywords

Breast cancer, invasive lobular carcinoma, E-cadherin, intratumour heterogeneity, gynaecological metastasis, immunohistochemistry

Australian and New Zealand Standard Research Classifications (ANZSRC)

ANZSRC code: 111201, Cancer Cell Biology, 60%
ANZSRC code: 111203, Cancer Genetics, 20%
ANZSRC code: 111209, Solid tumours, 20%

Fields of Research (FoR) Classification

FoR code: 1119, Other Medical and Health Sciences (Molecular Pathology), 60%
FoR code: 1112, Oncology and Carcinogenesis, 20%
FoR code: 1103, Clinical Sciences, 20%

Table of Contents

List of figures

List of tables

List of appendices

List of abbreviations used in this thesis

ADH	Atypical Ductal Hyperplasia
ALH	Atypical Lobular Hyperplasia
CCL	Columnar Cell Lesion
CGH	Comparative Genomic Hybridisation
CNA	Copy Number Alteration
DCIS	Ductal Carcinoma *In Situ*
EGFR	Epidermal Growth Factor Receptor-1

EMT	Epithelial to Mesenchymal Transition
ER	Oestrogen receptor
FFPE	Formalin Fixed Paraffin Embedded
FISH	Fluorescent in situ hybridisation
GEX	Gene Expression
H&E	Haematoxylin and Eosin
HER2	Human Epidermal Growth Factor Receptor-2
IC-NST	Invasive Carcinoma – No Special Type
ICGC	International Cancer Genome Consortium
IHC	Immunohistochemistry
ILC	Invasive Lobular Carcinoma
ISH	*in situ* hybridisation
ITH	Intratumour Heterogeneity
LCIS	Lobular Carcinoma *In Situ*
LCM	Laser Capture Micro-dissection
LN	Lymph node
Lt	Left
MDL	Mixed Ductal-Lobular Carcinoma
NGS	Next Generation Sequencing
ns	not significant
PR	Progesterone receptor
PLC	Pleomorphic Lobular Carcinoma
PLCIS	Pleomorphic Lobular Carcinoma In Situ
PR	Progesterone Receptor
QC	Quality control
QCGC	Queensland Centre for Gynaecological Cancer
QIMRB	QIMR Clive Berghofer Medical Research Institute
RBWH	Royal Brisbane and Women's Hospital
Rt	Right
S & N	Sullivan Nicolaides Pathology
TCGA	The Cancer Genome Atlas
TDLU	Terminal Duct Lobular Unit
TMA	Tissue microarray
TN	Triple Negative
TNM	Tumour, Node, Metastasis staging system

WHO	World Health Organisation

CHAPTER 1

GENERAL INTRODUCTION

1. General introduction

Breast cancer is the most common cancer in women in Australia, with a total of 14,000 women diagnosed with breast cancer in 2010 and more than 1.3 million women diagnosed annually worldwide (Grayson, 2012). Breast cancer presents a significant burden to both the patient and the health care system. The risk of developing breast cancer increases with age and one in eight women will be diagnosed with breast cancer before the age of 85 in Australia (AIHW, 2012).

Breast cancer survival rates have improved dramatically over the past 50 years due to a combination of factors, including better detection and improved screening programs (such as Breast Screen Australia which provides free mammographic screening for all women over the age of 50), improved treatment and an increased understanding of the biology of the disease. In the 1950's, the 5-year overall survival was just 40%, and by 2004 overall survival had increased to 85% (Maxmen, 2012). However, breast cancer remains the second most common cause of cancer death in women and nearly half a million women still die from breast cancer each year (Grayson, 2012); 40,000 in the United States (Siegel et al., 2013) and 2700 in Australia (Cancer Australia, 2012). In spite of substantial funding for breast cancer research, which has contributed significantly to our understanding of clinical and biological aspects of the disease, there are still large numbers of women dying each year and so improvements continue to be required in clinical areas of prevention, treatment and metastasis.

1.1. Clinical management of breast cancer

Breast cancer is a difficult disease to treat because it is extremely heterogeneous with respect to morphological appearance, clinical behaviour (presentation, response to treatment, metastatic potential, outcomes) and biology (molecular profiles and activated biological pathways). To help manage this heterogeneity, various classification methods have been introduced to identify clinically and biologically relevant subtypes of disease. The diagnosis of breast cancer occurs via a number of mechanisms including palpation of the breast, radiological examination by mammography, ultrasound or magnetic resonance imaging (MRI). An abnormality identified is sampled by core needle biopsy or fine needle aspiration and assessed for malignancy. If required a lesion is removed by surgical excision, either by wide local excision or mastectomy. The tissue is assessed by a

pathologist to determine tumour stage, histological type, grade and expression of molecular biomarkers. A multidisciplinary team of surgeons, radiologists, pathologists and oncologists determine the most appropriate management strategy for the patient based on these clinical and pathological variables.

In order to predict the clinical behaviour of the tumour, the breast cancer is staged using the Tumour Node Metastasis staging system (TNM) developed by The American Joint Committee on Cancer (AJCC, 2007). The TNM classification is based on:

- The size of the primary tumour: T1, <20 mm; T2, 20-50 mm; T3, >50 mm
- The number of metastatic lymph nodes: N1, 1-3; N2, 4-9; N3, >10 nodes
- The presence (M1) or absence (M0) of distant metastasis

Based on the combined scores, clinical stage I to IV is designated, with stage IV signifying patients having the worst overall outcome.

1.2. Breast cancer classification

Breast cancers are further categorised based on cellular morphology (**Section 1.2.1**), the expression of a suite of biomarkers (**Section 1.2.2**) and also patterns of gene expression (**Section 1.2.3**).

1.2.1. Morphological classification of breast cancer

The tumour is assigned a grade using the Nottingham Combined Histological Grading system (Elston and Ellis, 1991). Morphological features of the tumour are given a score of 1 to 3 based on three different cellular features:

- The percentage of the tumour exhibiting tubule formation: 1, >75%; 2, 10-75%; 3, <10%.
- The degree of nuclear pleomorphism: 1, small regular uniform cells; 2, moderate increase in size and variability; 3 marked variation.
- Mitotic count: 1-3, with a score of 3 having the most mitoses counted in a given microscopic field area.

Taken together, breast cancers are classified as grade 1 (score of 3-5; well differentiated), grade 2 (score of 6-7; moderately differentiated) or grade 3 (score of 8-9; poorly differentiated), with grade 3 tumours having the poorest prognosis. Several independent

studies have demonstrated the prognostic value of grading in predicting tumour behaviour and patient outcome (Lakhani S. R., 2012, Rakha et al., 2010a).

Breast cancer is classified into histological types based on the internationally recognised World Health Organisation (WHO) classification of the different types of breast cancer (Lakhani S. R., 2012). Current guidelines recognise 21 histological types of breast cancer characterised by differences in cellular morphology and growth patterns. The most common type of breast cancer is Invasive Carcinoma of No Special Type (IC-NST; previously called Invasive Ductal Carcinoma IDC-NST, **Figure 1.1a**) representing up to 80% of all breast cancers. IC-NST are a morphologically heterogeneous group but they do not exhibit the specific features to be classified as a so-called 'special' subtype. In fact the remaining 20% of breast tumours are regarded as special types; Invasive Lobular Carcinoma (ILC) being the most common accounting for up to 15% of all breast cancers (**Section 1.5**; (Lakhani S. R., 2012)). ILCs are characterised by a unique pattern of growth, whereby the cells are discohesive and grow in a single file of cells or dispersed throughout the stroma (**Figure 1.1b**). There are a number of histological variants of ILC including solid, alveolar, tubulolobular and pleomorphic lobular carcinomas (**Figure 1.1f**). A less common type of invasive breast cancer that exhibits both ductal and lobular morphological features is classified as mixed ductal-lobular (MDL) carcinoma, accounting for 3-5% of breast cancers (**Section 1.6**; (Lakhani S. R., 2012, Rakha et al., 2009)). Other special types of breast cancer include metaplastic, medullary, tubular, cribriform, mucinous, and adenoid cystic carcinomas that all have unique histological features, growth patterns, biology and outcomes (**Figure 1.1**).

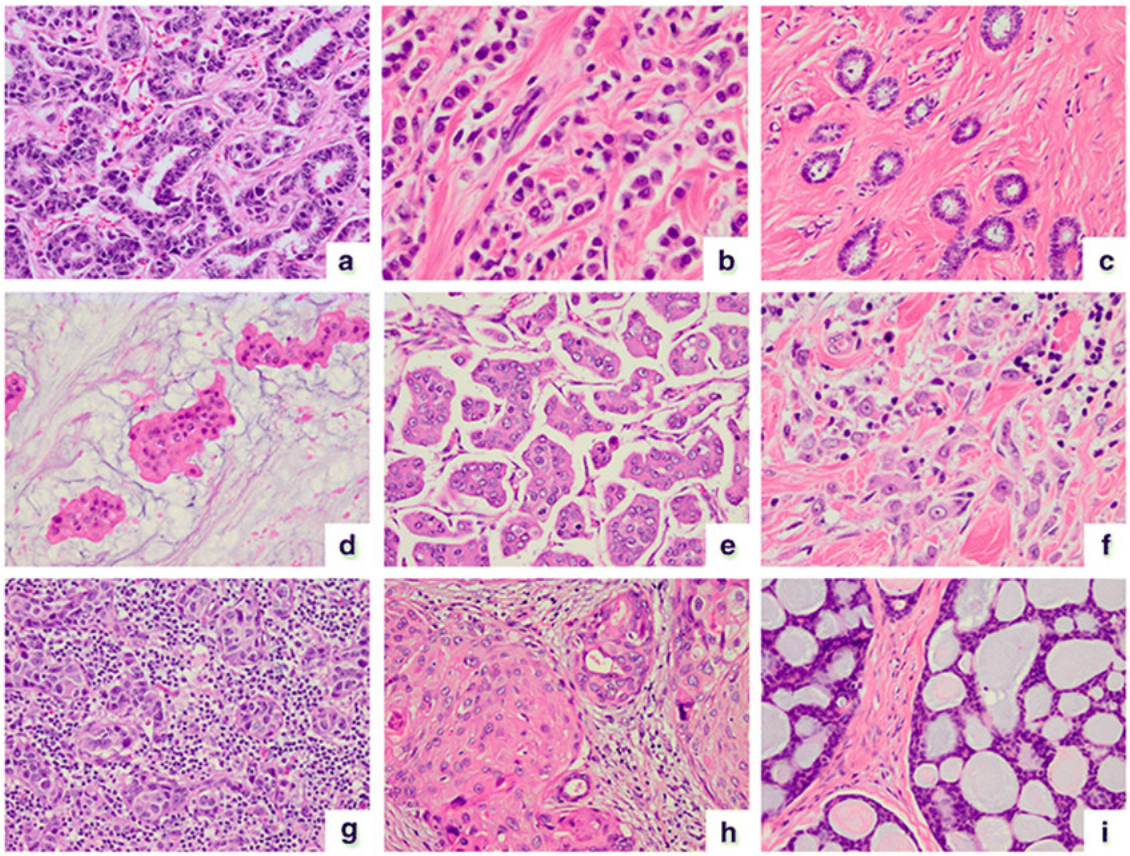


Figure 1.1: Representative images of some of the histological subtypes of breast cancer. a) invasive carcinoma – no special type, b) classic invasive lobular carcinoma, c) tubular carcinoma, d) mucinous carcinoma, e) invasive micropapillary carcinoma, f) pleomorphic lobular carcinoma, g) medullary carcinoma, h) metaplastic carcinoma, i) adenoid cystic carcinoma. (Vuong et al., 2014) with permission from Springer Science.

1.2.2. Biomarkers used in clinical practice

The tumour stage, grade and histological type are combined with an assessment of the expression of molecular biomarkers that further help predict the response to specific targeted therapies. Routinely, the expression of oestrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2) is assessed by immunohistochemistry (IHC). *In situ* hybridisation (ISH; Chromogenic- (CISH), silver (SISH) or fluorescence (FISH)) is also used to confirm the HER2/*c-ErbB2* gene copy number status (Rakha et al., 2010b). The uses of ER, PR and HER2 as biomarkers are good negative predictors for response to the targeted therapies that are currently available for breast cancer patients. For example, it is not appropriate to give endocrine-based therapy such as an ER antagonist (*e.g.* Tamoxifen) to patients with ER and PR negative tumours. Likewise, the anti-HER2 antibody trastuzumab (Herceptin®) is not appropriate for patients with HER2 negative tumours. The markers are also good positive predictors of

response to these treatments and have made a major contribution to improved outcomes for breast cancer patients (Early Breast Cancer Trialists' Collaborative, 2005). The combinations of all three markers provide more predictive value, for instance, tumours that express ER and PR are associated with a better prognosis compared to those that express HER2 or are negative for all three of these markers (triple negative) (Blows et al., 2010, Walt et al., 1976, Lakhani S. R., 2012). When these targeted therapies are used in combination with standard chemotherapy agents a greater overall patient outcome can be achieved (Rakha et al., 2010b).

ER and PR

The oestrogen and progesterone receptors are part of the nuclear receptor super family of transcription factors that regulate the expression of many genes and signalling pathways in response to steroid hormone binding (Clarke, 2003). There are two ER isoforms, ER alpha and ER beta, with ER alpha expression being the best understood in both normal and malignant breast; in the normal breast, ER and PR are expressed in the ducts and lobules. ER alpha is referred to as ER throughout the rest of this thesis.

ER and PR are activated by their ligands oestrogen and progesterone, respectively, to stimulate cell growth (Clarke, 2003). Approximately 80% of breast cancers are ER positive and up to 70% are PR positive (**Figure 1.2**, (Harvey et al., 1999)). Patients with low level ER and PR positivity have been found to respond to anti-endocrine therapy (Harvey et al., 1999) and therefore a patient is eligible to receive therapy if greater than 1% of tumour cells express ER and PR by immunohistochemistry (Hammond et al., 2010).

ER positive breast cancers benefit from anti-endocrine therapies. There are four main categories of endocrine therapies; (i) ovarian suppression by either surgical removal of the ovaries (oophorectomy) or chemically using GnRH analogues or radiation, (ii) Selective Estrogen Receptor Disruptors (SERDs) prevent the dimerization and nuclear localisation of ER *e.g.* (Fulvestrant), (iii) Selective Estrogen Receptor Modulators (SERMs) compete with oestrogen for binding to the oestrogen receptor *e.g.* Tamoxifen), and (iv) aromatase inhibitors (AIs *e.g.* Letrozole). Menopausal status is important when determining which treatment to use. For example, in post-menopausal women, oestrogen is produced peripherally by aromatisation of androgens to oestrogen, and therefore aromatase inhibitors are more appropriate for this group of patients. Tamoxifen is the gold standard of

treatment for pre- and post-menopausal women and treatment for 5 years has been shown to reduce the risk of recurrence by almost half in the first 10 years, compared to no endocrine treatment (Early Breast Cancer Trialists' Collaborative et al., 2011). Recent studies have found greater improvement with 10 years of tamoxifen administration compared to 5 years (Davies et al.). Despite the high success rates of these therapies some patients develop resistance to endocrine therapy (**Section 1.3.1**).

The progesterone receptor is an ER-regulated gene product, and expression of PR is traditionally thought to be associated with a functioning ER pathway (Clarke, 2003). However, this paradigm has been questioned by evidence that ER positive/PR negative tumours can respond to a second type of endocrine therapy after not responding to first line treatment, suggesting that the ER complex is still functional despite the lack of PR expression (Cui et al., 2005). *PR* has recently been shown to not only be a target gene of ER, but in the presence of progesterone, also form part of the protein complex that directs ER binding to DNA (Mohammed et al., 2015). The PR-ER complex altered the transcriptional program of breast cancer cells, a program of gene expression that is associated with good outcome. *In vivo* studies of xenografted human breast tumours in mice treated with a combination of tamoxifen with progesterone displayed significantly reduced tumour growth which suggests the possible utility of adding progesterone as a therapy in ER and PR positive breast cancer (Mohammed et al., 2015).

HER2

Human epidermal growth factor receptor 2 (HER2) is a cell membrane bound receptor tyrosine kinase and a member of the epidermal growth factor receptor family (Epstein et al., 2010). The HER2 gene, *ERBB2*, is located on human chromosome 17q12 and is amplified in 15% of breast cancers (Allred, 2010) (Chang and Hilsenbeck, 2010, Epstein et al., 2010). This amplification results in over-expression of HER2 protein and is considered an oncogenic driving event in HER2 positive breast cancer.

Both IHC and ISH are utilised to assess HER2 expression and amplification. HER2 over-expression is considered to be established when the tumour cells display strong and complete membrane staining in at least 30% of cells (**Figure 1.2**). ISH is used to quantify gene amplification, and is defined by the presence of 6 or more gene copies per nucleus (Wolff et al., 2007). HER2 positive breast cancer patients can benefit from three HER2

targeted therapies, Trastuzumab and Pertuzumab (monoclonal antibodies that inhibit HER2-mediated signalling) or Lapatinib (small tyrosine kinase inhibitor that blocks both EGFR and HER2) (Bedard et al., 2009, Epstein et al., 2010). Breast cancers that lack all three biomarkers (Triple negative: ER, PR and HER2) derive no benefit from currently available targeted therapy

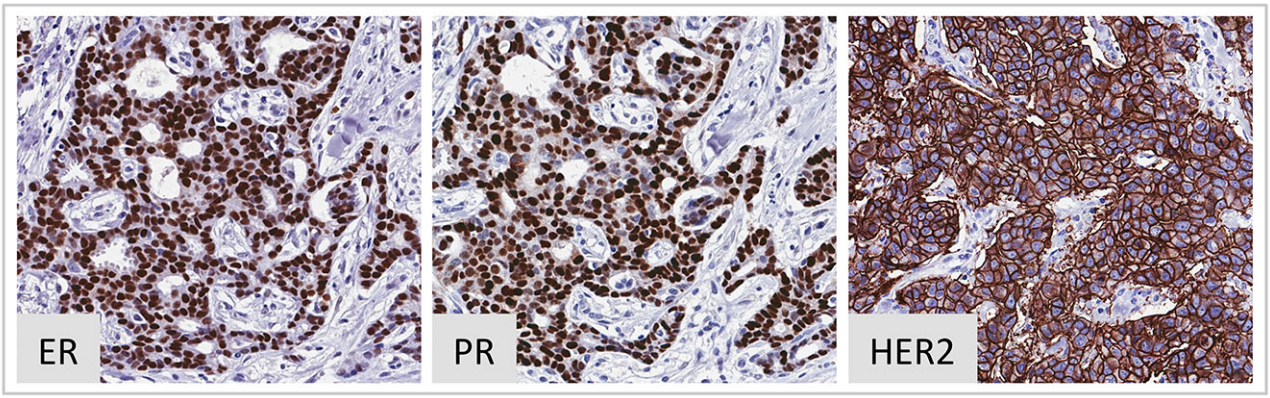


Figure 1.2: Representative images of ER, PR and HER2 positive IC-NST.

1.2.3. Intrinsic subtypes of breast cancer

Gene expression profiling studies have been utilised to further stratify breast cancer into prognostic subgroups based on distinct patterns of gene expression, revealing vast heterogeneity across all patients. Following the seminal study of Perou and Sorlie in 2000 (Perou et al., 2000), breast cancer has now been stratified into what is now known as the 'intrinsic' or molecular subtypes, of which there are broadly considered to be five; luminal A, luminal B, basal-like, HER2 and normal-like (**Figure 1.6**). The main differences driving the separation of these groups are the expression of ER and proliferation related genes. The transcriptomes of ER positive and ER negative breast cancers cluster separately, highlighting the difference in biology between these two subtypes. The distribution of these subtypes varies across array platforms, selected populations and clustering method (e.g. PAM50 (Parker et al., 2009), or Intrinsic/UNC sample predictor (Hu et al., 2006)), however the luminal A subtype is the most frequent (24%-39%), followed by basal-like (17%-37%), luminal B (10%-18%), HER2 (4%-10%) and normal-like (0-5%) (Perou et al., 2000, Sorlie et al., 2003, West et al., 2001, van de Vijver et al., 2002). The molecular classification of breast cancer has been shown to be predictive of clinical behaviour and patient survival (Sorlie et al., 2003, Paik et al., 2004). It has been shown that patients with a HER2 or basal-like breast cancer have a poorer outcome compared to the luminal breast cancers.

The luminal B subgroup has a poorer outcome compared to the luminal A subgroup (Sorlie et al., 2001). However, the analysis method used to define the molecular subtypes can produce different prognostic subgroups (Curtis et al., 2012).

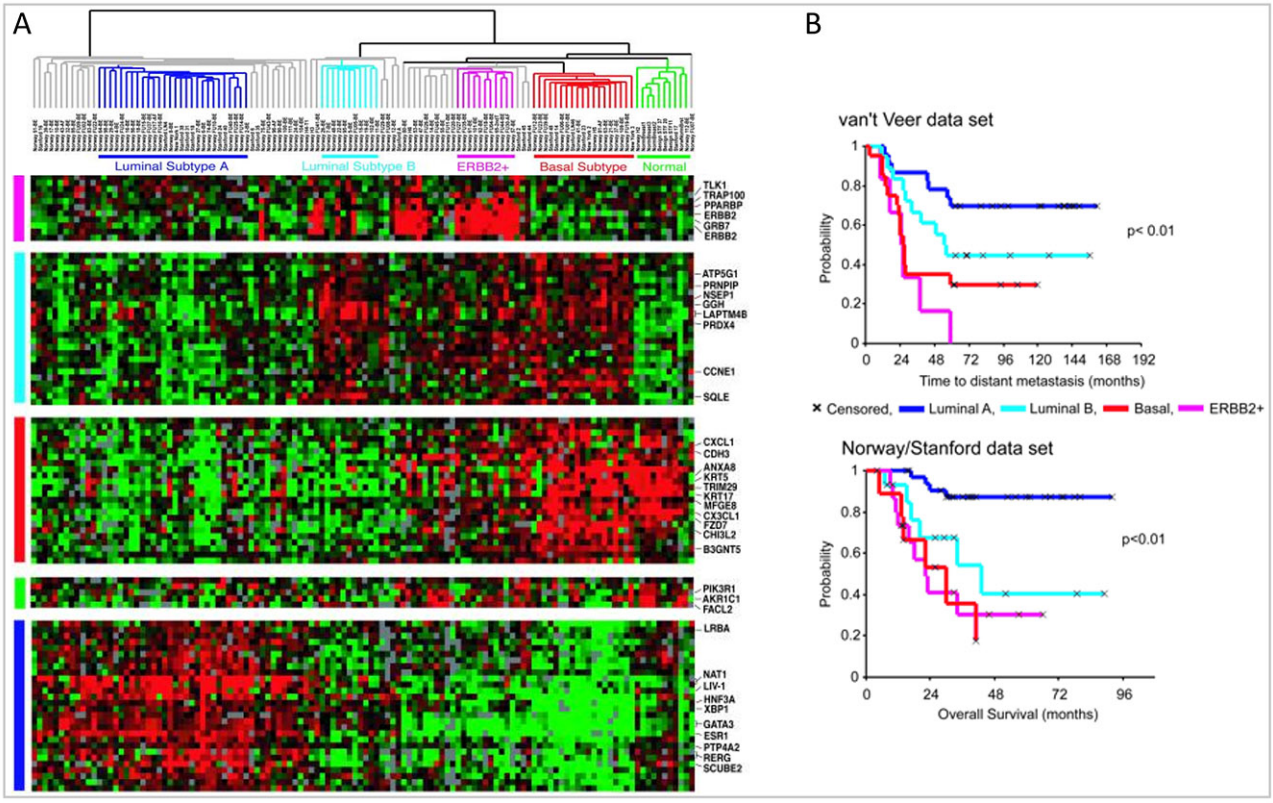


Figure 1.3: Gene expression profiling revealed 5 intrinsic breast cancer subtypes (A). (B) The intrinsic subtype classification has been associated with clinical outcome (Sorlie et al., 2003) Copyright (2003) National Academy of Sciences, USA.

The reproducibility of this classification system has been questioned and rigorously tested to find that overall approximately 75% of samples cluster within the same molecular subtype (Perou et al., 2000, Sorlie et al., 2001, Parker et al., 2009). However, Weigelt *et al*., performed a retrospective analysis using three microarray analysis methods (single sample predictors) to assess the agreement of molecular classification within individual samples (Weigelt et al., 2010). The group found that the basal-like subgroup consistently classified into same group irrespective of analysis tool used, and that there was great variability with the other subgroups dependent on which classifier was used (Weigelt et al., 2010), which has crucial implications if these were to be used to guide clinical decision making. There has been a lack of reproducibility of robustly categorising tumours in the normal-like subgroup, however this is because the original training set used true normal breast samples and the tumours that fall into this group likely represents samples with low

tumour cellularity and high stromal contamination (Prat and Perou, 2011). Also 6-36% of tumours are unable to be classified and differ between platforms (Sorlie et al., 2003, West et al., 2001, van de Vijver et al., 2002) and therefore all of these caveats need further validation.

Diagnostic platforms such as Oncotype Dx (Genomic Health, USA) and Mamma-Print (Agendia, The Netherlands) (others reviewed here (Reis-Filho and Pusztai, 2011)) have been developed utilising gene expression signatures to help guide patient management. However, these tests have not been widely adopted since added benefit of these tests compared to standard histopathology analysis is limited, as these tests mostly recapitulate the information already provided by histological grade, tumour size, ER, HER2, Ki67 and the use of EGFR, CK14 and CK5/6 as representative markers of the basal-like molecular profile (Weigelt and Reis-Filho, 2010, Cuzick et al., 2011, Nielsen et al., 2004, Bhargava et al., 2009). These tests are also expensive compared to standard IHC tests. Traditional microarray technology relies heavily on good quality frozen tissue with good quality RNA. However tests such at the Oncotype DX (real time PCR based) and Nanostring Technologies Prosigna™ Breast Cancer Prognostic Gene Signature Assay (PAM50 gene signature, using direct RNA hybridisation) can exploit the wealth of FFPE material available to generate risk of recurrence scores and validate in large cohorts with follow up information. There are also many clinical trials underway utilising the Mamma-Print platform such as PROMIS, I-SPY, MINDACT and MINT. These prospective trials have the overall goal to improve patient care based on molecular testing to predict therapy benefit, risk of recurrence and ultimately omit patients from unnecessary chemotherapy.

1.3. Biomarkers of tumour behaviour and hormonal signalling

Ki67, p53 and basal markers

Ki67 is a proliferation marker that is expressed in cells undergoing division (Gerdes et al., 1983). It is assessed to determine tumour growth rate and can be used to predict better response to chemotherapy agents (Ingolf et al., 2014). Ki67 is sometimes used in the clinical setting, however reliability is an issue, since there is no standardised IHC procedures or an agreed cut off for designating a tumour as Ki67 “high”, “low”, “positive” or “negative” (Harris et al., 2007). One study has found that a Ki67 score of 13.25% stratified luminal A and luminal B tumours (Cheang et al., 2009). The degree of value added by Ki67

is also questionable since tumour grade encompasses tumour proliferation by counting the number of mitoses. A suite of markers normally expressed in the basal/myoepithelial cells of the normal breast is used to define a group of tumours with an aggressive phenotype known as the basal-like tumours (Fulford et al., 2006, Fulford et al., 2007, Cheang et al., 2008). These markers include cytokeratin 5/6 and 14, and epidermal growth factor receptor (EGFR), and these markers highlight a poor outcome subgroup of triple negative breast cancers (Rakha et al., 2007). The tumour suppressor p53 is also typically associated with triple negative tumours (Rakha et al., 2010b). Many of these markers are of interest for research studies but are not used routinely in clinical practice.

Androgen receptor

The androgen receptor (AR) is another member of the nuclear receptor superfamily of transcription factors and is activated by the binding of testosterone. AR and testosterone have analogous roles in the prostate and prostate cancer as for ER and oestrogen in the breast. AR is expressed in the normal breast (McGhan et al., 2014) and in up to 90% of breast tumours (Allegra et al., 1979, Kuenen-Boumeester et al., 1992, Isola, 1993) (Hall et al., 1996). AR can be tumour suppressing in ER positive cells and is associated with a good prognosis (Peters et al., 2009, Castellano et al., 2010). It has been suggested that AR may be oncogenic in ER negative breast cancer cells (Ni et al., 2011). However, other studies have found that AR expression is associated with better disease free and overall survival in lymph node positive triple negative tumours (Rakha et al., 2007) and also associated with longer time to relapse in ER negative disease (Agoff et al., 2003). Cells that are ER negative and AR positive have been associated with a molecular apocrine molecular subtype (Farmer et al., 2005), apocrine differentiation (Niemeier et al., 2010) and are enriched for *ERBB2* amplification (Ni et al., 2011). Historically, breast cancers were successfully treated with testosterone (Goldenberg, 1964), yet the mechanism of action remains unknown. Testosterone treatment was stopped due to the discoveries that testosterone can be converted to oestrogen, causes adverse side effects and, ultimately, because of the successful introduction of selective oestrogen receptor modulators (Cole et al., 1971). Interest around AR being a therapeutic target in breast cancer has arisen again recently due to the high frequency of AR expression and greater understanding of the function of AR (Garay and Park, 2012). AR may be a potential target in triple negative breast cancer, where it is expressed in approximately 40% of tumours (Ogawa et al., 2008).

ER beta

ER beta (ERβ) is encoded by *ESR2*, is a transcription factor and is also part of the nuclear receptor superfamily. Although there is some structural homology with ERα, ERβ has opposing roles to ERα; ERβ is growth inhibitory and displays tumour suppressive functions (Fox et al., 2008). In fact, the two oestrogen receptors share approximately 54% homology and have similar affinity for natural oestrogens, yet differ in their transcriptional programs, response to SERMs (Katzenellenbogen and Katzenellenbogen, 2000) and the preferential binding of phytoestrogens to ERβ (Kuiper et al., 1998). There are 5 known ERβ isoforms, and due to the lack of specific antibodies available, there is a lot of contradictory data in the literature regarding the prognostic and predictive value of ERβ (Thomas and Gustafsson, 2011, Haldosen et al., 2014). For example, the expression of ERβ-1 was associated with better survival, particularly in triple negative breast cancer (Honma et al., 2008). ERβ-1 has also been shown to inhibit epithelial to mesenchymal transition in basal-like breast cancer (Thomas et al., 2012) and may explain the better outcome observed by Honma *et al*. This is in contrast to another study that found nuclear expression of ERβ-2 and ERβ-5, but not ERβ-1, was associated with better survival. Nuclear ERβ-2 was also found to be a predictor of response to endocrine therapy, however, the cytoplasmic expression of ERβ-2 was associated with poor outcomes (Shaaban et al., 2008). This study did not differentiate into intrinsic breast cancer subtypes. In a cohort of 936 breast cancer patients, 55% of the cohort expressed ERβ, only 30% of cases co-expressed both ERα and ERβ and the expression of ERβ was evenly distributed across the intrinsic subtypes (Novelli et al., 2008). This study found that ERβ predicted response to hormone therapy in luminal A lymph node negative tumours. Yet in lymph node positive luminal B tumours, ERβ positivity with PR negativity was associated with high risk of relapse (Novelli et al., 2008). In a small cohort of ER positive, HER2 negative patients (n = 81) treated with endocrine therapy, low expression of ERβ1 was associated with risk of early relapse and high expression of ERβ2 was associated with late relapse (Dhimolea et al., 2015). Overall the functional role of ERβ and its many isoforms may in fact be important, particularly with differential functions in different breast cancers subtypes. The functional role of ERβ in endocrine resistance is also an area of great interest. However much more work with validated antibodies is essential.

ER pioneer factors: FOXA1 and GATA3

A special group of transcription factors known as pioneer factors are essential for binding independently to, and unwinding, compacted chromatin to allow other transcription factors to bind. Pioneer factors facilitate rapid transcriptional responses (Hah et al., 2011) as they do not rely on other proteins for their DNA interactions. For ER, those factors are GATA3 and FOXA1, without whom ER is unable to bind to the DNA and transcribe target genes (**Figure 1.4 A**). FOXA1 and GATA3 are part of large families of highly conserved transcription factors (the forkhead box A (FOXA) and GATA families, respectively (Cirillo and Zaret, 1999, Cirillo et al., 2002)). In ER positive breast cancer cells FOXA1 was essential for all ER-DNA and PR-ER-DNA binding events (Carroll et al., 2005, Laganiere et al., 2005, Mohammed et al., 2015). Silencing of FOXA1 resulted in inhibition of ER binding and ER transcriptional activity (Hurtado et al., 2011, Carroll et al., 2005) and functional FOXA1 was required for growth of tamoxifen-resistant MCF7 cells (Hurtado et al., 2011). In prostate cancer cells, FOXA1 plays an equivalent role for AR, however, silencing of FOXA1 resulted in the reprogramming of AR binding events, and therefore changing gene expression (**Figure 1.4 B**, (Wang et al., 2011a, Sahu et al., 2011)). FOXA1 is essential for both ER and AR activity in breast and prostate cancer, respectively. FOXA1 is highly expressed in molecular apocrine breast tumours and was found to be essential for the expression of AR regulated genes (Robinson et al., 2011).

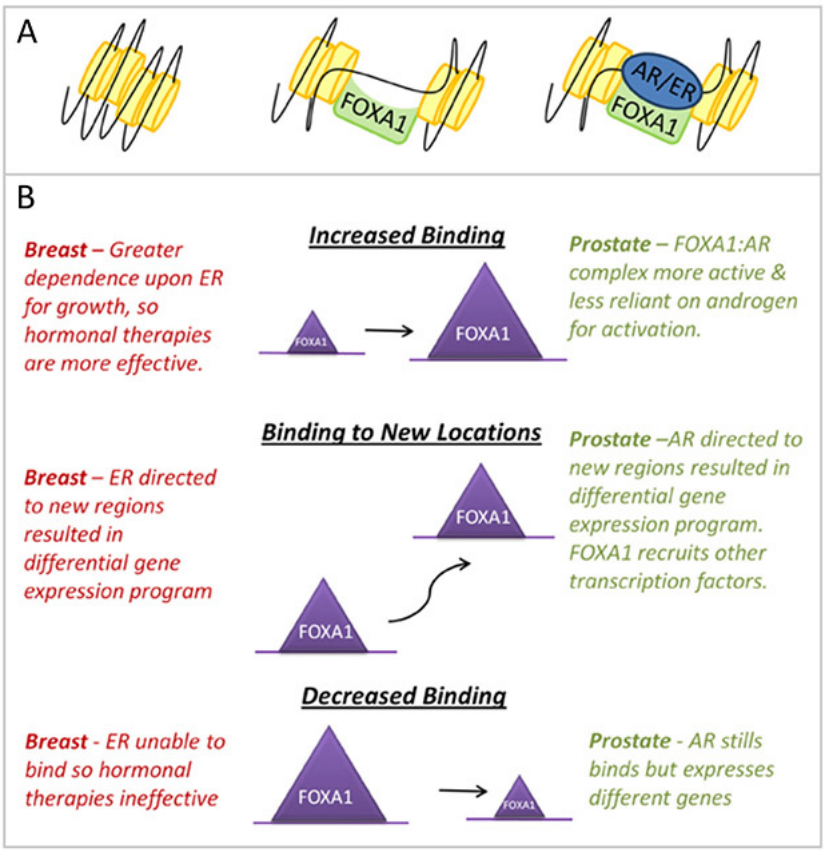


Figure 1.4: A) FOXA1 binds to compacted chromatin, allowing ER and AR to bind to the DNA. B) Altered FOXA1 binding plays different roles in breast and prostate cancer (Robinson et al., 2013b).

Interestingly, FOXA1 expression was found to mediate differential transcription programs in metastatic tumours compared to their matched primary tumour, suggesting that FOXA1 is required for ER function in a drug resistance setting (**Figure 1.4 B**, (Ross-Innes et al., 2012)). FOXA1 expression is associated with good outcome (Badve et al., 2007), as FOXA1 expression indicates that the ER is functional therefore increasing sensitivity to SERMs, and may predict also response to tamoxifen (Mehta et al., 2012). Hisamatsu et al, studied a series of ER positive, HER2 negative tumours that have high FOXA1 expression responded equally well to endocrine therapy with or without chemotherapy (Hisamatsu et al., 2012).

These studies highlight the potential utility of FOXA1 as a useful biomarker and potentially a therapeutic target, particularly in the metastatic setting, since FOXA1 was expressed alongside ER in treatment resistant cells (Ross-Innes et al., 2012, Nakshatri and Badve, 2007), but this needs to be assessed in larger metastatic cohorts.

FOXA1 is mutated in approximately 2% of all breast cancers (The Cancer Genome Atlas Network, 2012). The mutations were exclusively found in ER positive tumours, were mutually exclusive relative to *GATA3* mutations and found to cluster in the DNA binding domain and C-terminal transactivation domain (Robinson et al., 2013b). The functional consequences of these mutations have not been fully investigated, however one study has found the single nucleotide polymorphisms (SNPs) in FOXA1 binding sites modulate FOXA1-DNA interactions and change the gene expression program (Cowper-Sal lari et al., 2012). We can also extrapolate the potential functional effect of *FOXA1* mutations from *in vitro* studies. For example, silencing of *FOXA1* rendered tumour cells less sensitive to endocrine therapies, and therefore a mutation that diminishes DNA binding may allow the tumour to become independent of oestrogen and less likely to respond to therapy (Hurtado et al., 2011). On the other hand, mutation in the C-terminal domain may allow FOXA1 to bind to alternative sites in the genome and alter the transcriptional program (Robinson et al., 2013b). Mutations in the ER gene, *ESR1,* have been implicated in ER positive metastatic disease (see **Section 1.3.1**), and therefore the functional consequence of *FOXA1* mutations in metastasis is an avenue worthy of further investigation.

GATA3 is a transcription factor expressed in the luminal epithelial cells (but not the myoepithelial cells) of the breast and is essential for maintaining cellular differentiation (Asselin-Labat et al., 2007, Kouros-Mehr et al., 2006, Naylor and Ormandy, 2007). GATA3 is overexpressed in luminal breast cancers, is strongly associated with ER positive breast cancer (Sorlie et al., 2003, Hoch et al., 1999) and has been identified as one of the only frequently mutated breast cancer genes (mutated in approximately 10-15% of breast cancers (The Cancer Genome Atlas Network, 2012, Stephens et al., 2012). It has been reported that GATA3 is expressed in approximately 77-95% of ER-positive tumours and 0-24% of ER-negative tumours (Hoch et al., 1999, Voduc et al., 2008, Liu et al., 2012). GATA3 is also expressed highly in urothelial carcinomas and is a useful marker to determine whether a metastasis is of breast or urothelial origin, particularly since more than 80% of metastatic breast cancers express GATA3 (Liu et al., 2012). One study has suggested that GATA3 expression is predictive of response to hormone therapy (Parikh et al., 2005), where the absence of GATA3 expression was observed in tumours that were unresponsive to hormone therapy. It must be noted however that the sample size was small in this study. Low expression of GATA3 has been associated with poor outcome, particularly high tumour grade, positive lymph node and HER2 overexpression (Mehra et al., 2005, Liu et al., 2012, Ciocca et al., 2009, Jacquemier et al., 2009). The prognostic

role of GATA3 in breast cancer has been contradictory. The study by (Mehra et al., 2005) suggests that GATA3 is an independent marker of patient outcome in their cohort, however (Voduc et al., 2008) found that GATA3 was neither prognostic for patient outcome nor predictive of response to hormone therapy. These conflicting results are likely due to the lack of consensus for scoring cut offs, which ranged from 5-30%. The study by (Voduc et al., 2008) had the lowest frequency of GATA3 positive tumours, with only 32% of their 3119 cases showing GATA3 positivity. This is striking when GATA3 is strongly associated with ER positivity, and it is widely established that over 70% of breast cancers are ER positive.

GATA3 directly regulates the expression of *ESR1*, and reciprocally, ERα directly stimulates expression of *GATA3*. It is therefore unsurprising that GATA3 plays a crucial role in the response of ER to estrodiol in breast cancer cells (Eeckhoute et al., 2007). It has been shown that there is no physical interaction between GATA3 and ER (Eeckhoute et al., 2007), therefore other factors must be recruited to regulate expression of ER. GATA3 has been identified to act upstream of FOXA1 by directly binding to the regulatory region of *FOXA1* (Kouros-Mehr et al., 2006, Usary et al., 2004) and overexpression of GATA3 shown to up-regulate FOXA1 expression (Usary et al., 2004). Taken together, it is likely that FOXA1 mediates crosstalk between GATA3 and ER. Further work is necessary to understand the interplay between all of these factors involved in ER signalling in the mammary gland. Also the role of these transcription factors on the function of ERβ is unknown.

1.3.1. Mechanisms underlying endocrine resistance

Despite the great success of endocrine therapies in reducing patient mortality, there are a subset of patients whose tumours do not respond to endocrine therapy, specifically, one third of women treated with tamoxifen for 5 years will relapse within 15 years (Early Breast Cancer Trialists' Collaborative, 2005).

There have been many molecular mechanisms described that confer endocrine resistance (Osborne and Schiff, 2011). The most obvious example is the loss of ER expression that occurs in approximately 15-20% of metastatic breast cancers (Hoefnagel et al., 2012, Cummings et al., 2014). However, up to 85% of metastases still express ER and therefore other mechanisms of resistance must be involved (Dodwell et al., 2006). Several growth

factor signalling pathways have been implicated in driving resistance to endocrine therapy. For example the MAPK/ERK, EGF, HER2 and AKT/PI3K pathways have all been shown to inappropriately phosphorylate ER under certain conditions and many of the genes in these pathways are frequently genetically modified, such as frequent amplification of *ERBB2*, activating mutations in *PIK3CA* and loss of *PTEN* (The Cancer Genome Atlas Network, 2012). This hyper-phosphorylation causes over activation of ER function, resulting in activity in the absence of ligand and altered transcriptional activity (Martin et al., 2003, Britton et al., 2006, Jeng et al., 2000, Bhat-Nakshatri et al., 2008, Lupien et al., 2010, Miller et al., 2010, Hurtado et al., 2008). Alterations of cell cycle regulators have also been implicated in endocrine resistance due to constitutive activation of growth factor pathways, such as up-regulation of MYC (Mukherjee and Conrad, 2005) and Cyclin D1 and inactivation of RB1 (Bosco et al., 2007, Butt et al., 2005) (reviewed extensively here: (Musgrove and Sutherland, 2009)). These results are based largely around *in vitro* studies, and how these mechanisms translate into the clinical setting and the frequency of these mechanisms occurring in the patient population needs to be further explored.

The gene encoding ER alpha, *ESR1*, is rarely mutated or amplified in primary breast cancers (Karnik et al., 1994, Roodi et al., 1995, The Cancer Genome Atlas Network, 2012). Very recent reports however have found a high frequency of *ESR1* mutations in ER positive metastatic breast cancer (**Figure 1.5 A**, (Li et al., 2013, Toy et al., 2013, Robinson et al., 2013a, Merenbakh-Lamin et al., 2013, Jeselsohn et al., 2014)). Mutations have been found to cluster in the ligand binding domain of the gene and confer ligand independent, constitutive activation of the receptor (Jeselsohn et al., 2014). Genomic rearrangements are also rare but have been reported. For example, ESR1-CCDC170 and ESR1-YAP1 fusion genes have been identified and contribute to an endocrine resistant phenotype (Li et al., 2013, Veeraraghavan et al., 2014).

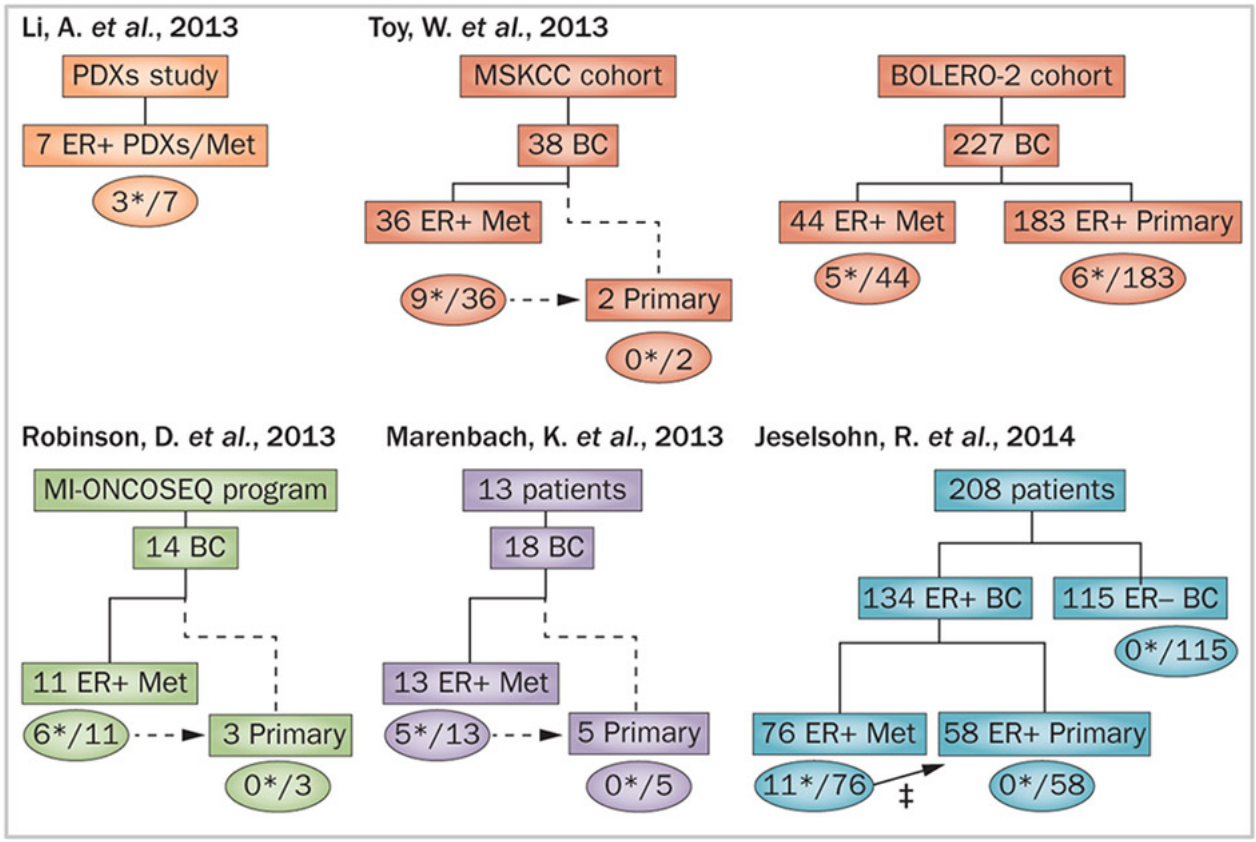


Figure 1.5: A) A summary of the 5 studies that have reported *ESR1* mutations in breast cancer metastasis. (Jeselsohn et al., 2015). * Indicates the number of *ESR1* mutations reported in each study. Matched primary tumours that were also assessed for *ESR1* mutations are indicated by the dashed line. With permission from Nature Publishing Group.

The prevalence of *ESR1* mutations in these studies ranged from 14-54%, and taken together, 29/187 (21%) of metastatic samples harboured an *ESR1* mutation in the ligand binding domain (Jeselsohn et al., 2015). Interestingly, the majority of patients were treated with an aromatase inhibitor. Of note, no *ESR1* mutations were detected in ER negative breast cancers, and *ESR1* mutations was also more prevalent in patients who had received multiple lines of endocrine therapy (Jeselsohn et al., 2014). This highlights that endocrine treatment is selecting for the outgrowth of treatment resistant subclones, or the mutation is acquired later under treatment pressure after the primary tumour has been removed. These studies found that fulvestrant and tamoxifen were still able to bind and inhibit the mutated ER, however it required much higher doses than needed for wild type ER (Robinson et al., 2013a, Toy et al., 2013, Jeselsohn et al., 2014), and cells expressing mutant ER were still able to grow in the presence of fulvestrant and tamoxifen. Mutant *ESR1* was also resistant to fulvestrant-induced degradation, revealing a molecular mechanism underlying resistance to the drug (Jeselsohn et al., 2014). Higher doses of fulvestrant have been shown to increase progression free and overall survival in a clinical

trial (Di Leo et al., 2010), thus raising the question as to the utility of higher doses of tamoxifen or the development of mutant specific SERMs and SERDs to help target ER positive metastatic disease. One such drug, GDC-0810 (ARN-810), is a much more potent SERD than fulvestrant and is currently in clinical trials (Lai et al., 2015). This data is promising as current treatment options for metastatic disease is limited.

As described earlier, FOXA1 and GATA3 also have roles in altering ER and AR function and contributing to altering gene expression and cell survival (Eeckhoute et al., 2007, Ross-Innes et al., 2012). Overexpression of AR also confers endocrine resistance in cell models (De Amicis et al., 2010). Due to the large number of putative mechanisms described in the literature, a patient with a treatment resistant recurrence will need to be managed according to the mechanism underlying the resistance.

1.4. Molecular profiling of breast cancer

Recently, international consortia such as The Cancer Genome Atlas network (TCGA) and ICGC (The Cancer Genome Atlas Network, 2012, Stephens et al., 2012, Curtis et al., 2012, Ellis et al., 2012) have generated a large molecular profiling resource of breast and other tumour types comprising of gene expression, DNA copy number, DNA methylation and gene mutation data from exome or whole genome sequencing. This has helped define the comprehensive molecular portrait of breast tumours. There are very few recurrently mutated genes considered to drive breast carcinogenesis ("driver mutations") and some tumours harbour very many mutations (mostly genomic substitutions termed "passenger mutations" that do not promote tumourigenesis) while others have very few alterations (Stephens et al., 2012, The Cancer Genome Atlas Network, 2012, Shah et al., 2012). Of 40 genes defined as a driver, seven genes contributed as much as 58% of the driver alterations; *TP53*, *PIK3CA*, *GATA3, ERBB2*, *MYC*, *FGFR1* and *CCND1,* and were altered in more than 10% of cases (**Figure 1.7**, (Stephens et al., 2012, The Cancer Genome Atlas Network, 2012)). The other 33 mutated cancer genes (responsible for 42% of the driving genetic events) were mutated relatively infrequently and include *AKT1*, *BRCA1*, *CDH1*, *PTEN*, *RB1*, *KRAS* and *SMAD4*. Surprisingly, very few new cancer genes were discovered (*e.g. MAP3K1*, *NCOR1* and *CDKN1B*)(Stephens et al., 2012). ER positive and ER negative breast tumours were demonstrated to have very different mutation profiles, with mutations in *PIK3CA*, *GATA3* and *MAP3K1* enriched in the luminal subtypes (Ellis et al., 2012, The Cancer Genome Atlas Network, 2012), while *TP53* was most frequently

mutated gene in ER negative and triple negative tumours (The Cancer Genome Atlas Network, 2012, Shah et al., 2012). These massively parallel sequencing studies have overwhelmingly exemplified breast tumour diversity within known subgroups and the complexity underlying tumourigenesis.

The signalling pathways implicated by the frequently mutated genes involve cell-cycle regulation (*e.g. TP53*, *RB1* and *CCND1*), PI3K-AKT kinase signalling (*e.g. PIK3CA*) and cell adhesion (*e.g. CDH1*). TP53 is a DNA binding protein involved in maintaining genomic stability. TP53 has many mechanisms of action, including activation of DNA repair proteins, inducing cell cycle arrest and apoptosis (Brosh and Rotter, 2009). In fact, *TP53* is the most mutated gene in human cancer, emphasising its important role as a tumour suppressor (Brosh and Rotter, 2009, Surget et al., 2013). *CCND1*, encoding cyclin D1, also regulates the progression of cell cycle by activating the cyclin-dependent kinases 4 and 6 (CDK4/6). Activated CDK4/6 promotes cell cycle entry by phosphorylating the retinoblastoma (*RB*) gene (Chau and Wang, 2003). Cyclin D1 acts as an ER cofactor, where it binds and activates ER-mediated transcription in both the presence and absence of ligand (Zwijsen et al., 1998). Due to the high prevalence of *CCND1* amplifications and its role in cell cycle regulation in breast cancer, CDK4 and 6 inhibition has been targeted for therapeutic intervention, in fact, early progress reports from ongoing clinical trials are promising (Finn et al., 2015).

PIK3CA encodes the p110α isoform of class-IA PI3-kinase. Mutations in *PIK3CA* have been found to cluster into two mutation hotspots and these mutations induce oncogenic PI3K-AKT signaling and result in increased cell proliferation and survival (Yuan and Cantley, 2008). In breast cancer, *PIK3CA* mutations are an early event in ER positive tumours; mutations in *PIK3CA* have been found in matched cases of DCIS and IC-NST (Kalinsky et al., 2011, Miron et al., 2010, Li et al., 2010b) and also in columnar cell lesions (Troxell et al., 2012) and therefore are likely to play a vital role in breast tumourigenesis.

As described above, GATA3 is a transcription factor expressed in the luminal epithelial cells of the breast and is essential for maintaining cellular differentiation (Asselin-Labat et al., 2007, Kouros-Mehr et al., 2006). Mutations in *GATA3* appear solely in ER positive breast tumours and lead to the loss of the protein's DNA binding capacity and tumour suppression activity (The Cancer Genome Atlas Network, 2012, Jiang et al., 2014, Usary et al., 2004, Si et al., 2015) (Cohen et al., 2014). *GATA3* mutations were found to be

mutually exclusive relative to *PIK3CA* and *TP53* mutations, and low protein expression is strongly correlated with markers of poor prognosis (high-grade, ER negative and HER2 overexpressing tumours (Yoon et al., 2010, Kouros-Mehr et al., 2006)). *GATA3* mutations also correlated with better patient overall survival among patients with ER positive tumours (Jiang et al., 2014), emphasising the functional role of GATA3 in the development of luminal breast cancers and may explain to some degree the heterogeneity observed in variable patient outcomes. The loss of *GATA3* function may be playing a protective role as it may not be transcribing genes associated with poor outcomes (*i.e. CCND1* and aurora kinase A have been shown to be *GATA3* target genes (Molenaar et al., 2010, Theodorou et al., 2013, Jiang et al., 2010)) and could potentially act as a biomarker of good prognosis in ER positive breast cancer patients (Cohen et al., 2014).

Integrated genomic and transcriptomic features of breast cancer

To gain greater understanding of the processes driving tumourigenesis, Curtis et al., integrated copy number profiling with gene expression in tumours with defined clinical outcome (Curtis et al., 2012). This integrated analysis revealed 10 novel subgroups (as opposed to the five identified by gene expression profiling alone (Sorlie et al., 2001)) that stratified patients based on overall survival. For example, patients with tumours classified as integrated cluster 5 have the worst outcome compared to the other nine clusters. Also of note was an ER positive subgroup (IntClust2) that has quite poor outcomes, highlighting a unique high-risk subgroup, and further accounting for the heterogeneity in clinical outcomes seen even in patient's with the hallmarks of a tumour with a 'good' prognosis. As observed in the sequencing studies, there were two groups, IntClust3 (predominately luminal A and invasive lobular carcinomas) and IntClust4 (including both ER positive and negative tumours) that had very little copy number alterations and have more favourable outcomes. This study has been seminal in gaining understanding in how copy number alterations influence gene expression in breast cancer. An integrated approach of analysing all molecular information is essential to understanding complex pathways involved in breast tumourigenesis. These analyses are still in their infancy and whilst there are large gains in our biological understanding of breast cancer, the clinical utility is largely unknown, and further investigation is warranted. In spite of there being 10 subgroups defined by (Curtis et al., 2012) no single group encompasses a pure histological special type of breast cancer.

1.5. Invasive lobular carcinoma

As mentioned above, Invasive Lobular Carcinoma (ILC) is the most important special type of breast cancer. ILC accounts for 5-15% of all invasive breast cancers (Lakhani S. R., 2012, Arpino et al., 2004, Pestalozzi et al., 2008, Rakha et al., 2008, Orvieto et al., 2008) and the incidence of ILC is increasing, whereas the frequency of IC-NST has remained stable (Li et al., 2000). There is some evidence to suggest this correlates with the increasing use of hormone replacement therapy (Daling et al., 2002, Reeves et al., 2006, Li et al., 2006) and one study suggests increase in alcohol consumption could also be responsible (Li et al., 2010a). Patients with ILC tend to present at an older age compared to IDC (average 64 and 60, respectively) (Arpino et al., 2004, Li et al., 2005, Rakha et al., 2008, McCart Reed and Kutasovic et al., 2015). ILC can be detected as a palpable mass and is evident on mammography as a spiculated mass or architectural distortion in many cases. However, mammographic screening fails to detect the tumour in as many as 57-76% of women compared to as few as 5- 15% for IDC (Evans et al., 2002, Moy et al., 2002, Soo et al., 2001). ILC also frequently presents as a multifocal or multicentric lesion (85%) and/or bilaterally (30-67%), which is also more common than for IDC (Beute et al., 1991, Newman, 1966, Foote and Stewart, 1941, Chen et al., 1999, Intra et al., 2004).

ILCs are morphologically characterised by a proliferation of non-cohesive, small neoplastic cells that are dispersed in the stroma as single cells or arranged in single files of cells. This diffuse and infiltrating growth pattern contributes to the difficulty in the detection of ILC by mammography and difficulty for surgeons sometimes to palpate the mass during surgery. As a consequence, tumours are detected and diagnosed late, are often of larger size, tumour margins are difficult to assess and patients are therefore more likely to have a mastectomy compared to patients with IDC.

Mitoses are infrequent, therefore the histological grading of ILC is typically low to moderate, with approximately 76% of ILC being grade 2 (Rakha et al., 2008, McCart Reed and Kutasovic et al., 2015). A number of histological variants of ILC have also been recognised including pleomorphic, solid, alveolar and tubulolobular types. These variants share the discohesive growth patterns with classic ILC but may tend be of higher grade and can have worse outcomes (Eusebi et al., 1992, Middleton et al., 2000, Buchanan et al., 2008, Orvieto et al., 2008). Precursor lesions for ILC are called atypical lobular hyperplasia (ALH) and lobular carcinoma *in situ* (LCIS) and are present alongside ILC in

58-98% of cases (Abdel-Fatah et al., 2007, Dixon et al., 1982, Newman, 1966, Orvieto et al., 2008, Buchanan et al., 2008).

Classic ILC are almost always ER, PR positive, and rarely exhibit HER2 amplification or express abnormal p53 or basal markers (McCart Reed and Kutasovic et al., 2015). Due to a high frequency of hormone receptor expression, patients with ILC generally have a good response to hormonal therapy (Rakha et al., 2008, Pestalozzi et al., 2008), but they have a poor response to chemotherapy and this is probably due to the low mitotic index of the tumour. Despite the presence of favourable prognostic factors, the locoregional control, disease free interval and overall survival is not necessarily better than that for patients with IDC. Some studies have shown no difference in survival (Winchester et al., 1998), whereas others have found that at five years, ILC have better survival rates than IDC, however this declines after 10 years where ILC have a significantly worse overall survival (Nagao et al., 2012, Pestalozzi et al., 2008, Rakha et al., 2008). It is not clear why this is the case but could be related to a combination of factors including the difficulty of early detection, difficulty in surgical removal, larger tumour size, poor response to chemotherapy and the nature of distant metastases.

1.6. Mixed ductal lobular carcinomas

Invasive breast cancers exhibiting both ductal and lobular morphological features are classified as mixed ductal and lobular carcinomas (MDL). MDLs account for 3-5% of all invasive breast cancers (Lakhani S. R., 2012), and are a clear example of morphological intratumour heterogeneity.

There is limited data assessing the clinical and biological significance of MDL carcinomas. MDLs are generally under reported and a small lobular component is sometimes not annotated. Small studies have found that MDLs are associated with better prognosis when compared to pure IC-NST, but poorer prognosis when compared to pure ILC, although most of these differences were lost when adjusting for tumour grade (Sastre-Garau et al., 1996, Suryadevara et al., 2010, Rakha et al., 2009). Furthermore, the metastatic pattern of mixed tumours tends to follow that of the histological type that is most prevalent in the primary tumour (Rakha et al., 2009). These tumours are therefore considered to be a distinct entity to pure IC-NST and ILC (Rakha et al., 2009, Bharat et al., 2009). There is also evidence that the incidence of MDLs is increasing (Li et al., 2003, Rakha et al., 2009).

MDL tumours may pose a problem with regard to patient management, as the different histological components may have different clinical behaviours (*i.e.* response to therapies) and metastatic patterns (*i.e.* IC-NST have a propensity to spread to the brain, liver and lung, while ILC have a propensity to spread to the gastrointestinal tract, peritoneum and gynaecological sites (Harris et al., 1984)).

To date, there is very limited understanding of the molecular evolution of MDL carcinomas. There have been a few suggested theories in relation to their origin. Firstly, are these distinct morphological entities representing two independent collision tumours or do they arise from a common clone? Secondly, if the two components are clonally related, then what is the mechanism of transition from one phenotype to the other? Small pieces of evidence indicate that the lobular and ductal components are clonally related lesions as opposed to separate collision tumours, yet this is not well characterized. LOH and CGH analyses reported the shared loss of a common allele in both DCIS and LCIS components (Wagner et al., 2009), and identified multiple sub-clones within pre-invasive lesions (Buerger et al., 2000) suggesting that these morphologically distinct entities might be related clones. However, the mechanisms underlying the change in phenotype have not been investigated. Given that the E-cadherin molecule plays such a definitive role in the phenotype of ILCs, it is likely to also be important in the biology of the MDL as well.

1.7. Molecular aspects of ILC

The frequent expression of ER and PR means ILC are typically classified as a luminal tumour phenotype (McCart Reed and Kutasovic et al., 2015), however there is heterogeneity and some lobular tumours are classified by gene expression profiling as HER2, basal-like or molecular apocrine subtypes (Weigelt et al., 2008). ILC and ER positive IDC exhibit different gene expression profiles, emphasising the different biology between the two types. Genes differentially expressed between IDC and ILC are involved in processes of cell adhesion, cell-cell signalling and actin cytoskeleton signalling (Korkola et al., 2003, Weigelt et al., 2008).

From a genomics point of view, ILC and ER positive IC-NST exhibit a different pattern of DNA copy number aberrations (CNA) than high-grade tumours, and often fewer CNAs overall. Common CNAs include gain of 1q and 16p, while loss of 16q is very common, occurring in approximately 90% of all ILCs (**Figure 1.6 A**, (Buerger et al., 1999, Etzell et

al., 2001, Mastracci et al., 2006, Hwang et al., 2004, McCart Reed and Kutasovic et al., 2015)). Amplifications of 8p12 (*FGFR*) and 11q13 (*CCND1*) are common and are considered to be potential drivers of low-grade disease and therefore potential drug targets (Reis-Filho et al., 2006). Some of these molecular features are identified in associated pre-invasive lesions, ALH and LCIS, supporting the role of these lesions as precursors for ILC (Vos et al., 1997, McCart Reed and Kutasovic et al., 2015).

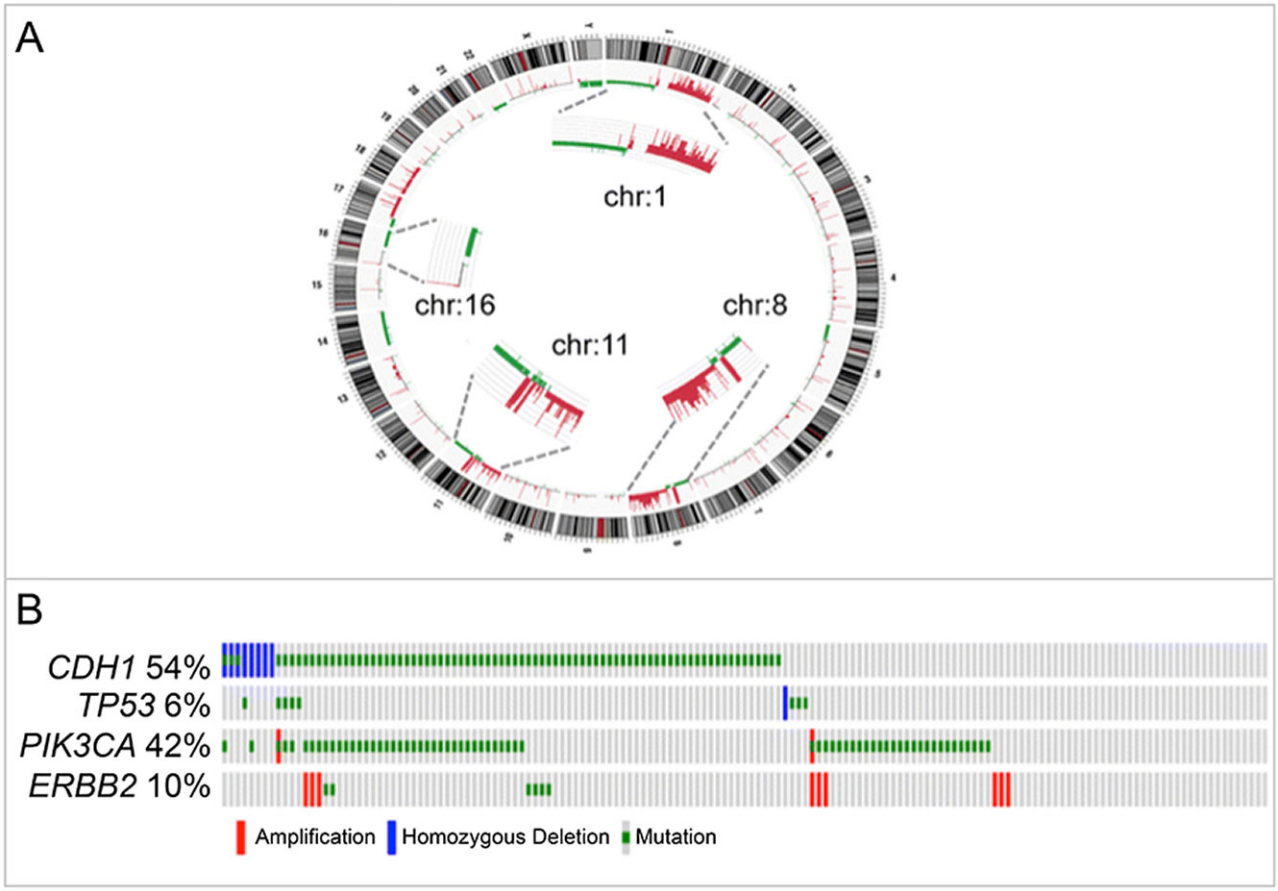


Figure 1.6: Genomic features of ILC. A) Circos plot of SNP CGH data. All chromosomes are represented on the outside. The zoomed in regions show the frequent copy number alterations identified in ILC (loss = green, gain = red). B) Across the TCGA cohort, *CDH1* and *PIK3CA* were the most frequently mutated genes. TP53 was mutated infrequently, and there was an enrichment for ERBB2 mutations in ILC compared to IC-NST. (McCart Reed and Kutasovic et al., 2015).

The TCGA has profiled the largest series of ILC to date. High resolution copy number profiling has confirmed gain of 1q and 16q loss in luminal tumours. The luminal A subtype for which ILC is part of, was found to have a low mutation rate compared to the other tumour groups, and harboured the most significant variation in mutated genes. The most frequently mutated genes in ILC were *CDH1* (54%, including both copy number loss and

mutation), followed by *PIK3CA* (42%) RUNX1 (8.4%), *MLL3* (7.7%) and *FOXA1* (6.5%). *TP53* (6%) mutations were rare (**Figure 1.6 B**, (The Cancer Genome Atlas Network, 2012, McCart Reed and Kutasovic et al., 2015)). Interestingly, enrichment for *ERBB2*/HER2 mutations was identified within lobular cancers. Of the eight somatic variants identified, all were within the kinase domain, 4/36 (11%) were ILC and 4/789 (0.5%) were IC-NST. Another recent study also identified *ERBB2* mutations in 6/22 (27%) of ILC that harboured *CDH1* mutations (Ross et al., 2013). A *CDH1* mutation was also found in the 4 ILC cases with *ERBB2* mutations from the TCGA. These mutations indicate a possible alternative mechanism for HER2 driven proliferation, as opposed to gene amplification, and that there may be potential benefit from anti-HER2 therapies in these patients.

The most significantly mutated gene in ILC however is *CDH1*, encoding the cell adhesion protein E-cadherin (see **Section 1.8**). *CDH1* mutations have been reported to occur in approximately 30% – 80% of ILC (Rakha and Ellis, 2010, Berx et al., 1996, Cleton-Jansen, 2002) yet rarely in occur in IC-NST (0/26 (Roylance et al., 2003), 0/42 (Berx et al., 1995), 0/25 (Kashiwaba et al., 1995)). The TCGA originally reported *CDH1* mutations in 30/36 ILC and the mutations also corresponded with both low mRNA and protein expression (The Cancer Genome Atlas Network, 2012). Since the TCGA data has been updated with more patient samples and is available to the public, the frequency of somatic *CDH1* mutations in ILC was identified in much fewer cases 78/155 (50%). This frequency has been confirmed by another genome sequencing study were 20/40 (50%) ILC had *CDH1* mutations (Ellis et al., 2012). The differences in reported frequencies may be attributed to differences in sequencing methods, however these recent studies highlight that the real frequency of *CDH1* mutations in ILC is likely to be lower than previously considered.

1.8. E-cadherin

Epithelial cadherin (E-cadherin), the loss of which is a defining feature of ILC is a calcium dependant transmembrane protein that mediates cell-cell adhesion and cellular polarity (Nagafuchi et al., 1987). The extracellular domain of E-cadherin binds to itself on neighbouring cells. The intracellular domain of E-cadherin associates with the actin cytoskeleton via α-, β-, and γ-catenin to form adherens junctions between non-neural epithelial cells (**Figure 1.7 A, B**, (Ozawa et al., 1989, Reynolds et al., 1994)). E-cadherin is largely regulated by its catenin-binding partners, which anchor E-cadherin to both the cell membrane and the actin cytoskeleton. E-cadherin mediated cell adhesion maintains cell

viability and when this adhesion is lost, the detached cells undergo a cell death program called anoikis (Fouquet et al., 2004).

The loss of E-cadherin occurs in up to 90% of ILC (**Figure 1.7 C**) and is considered fundamental to the development of the characteristic invasive growth pattern of lobular cancers (Berx et al., 1996). The impairment of functional cell-cell adhesion following E-cadherin loss likely contributes to the characteristic discohesive nature of tumour cells in all lobular neoplasms (Berx et al., 1996, Derksen et al., 2011, Vos et al., 1997). Knowledge is increasing as to the very dynamic regulation of this molecule, as well as the importance of E-cadherin in maintaining normal multicellular integrity and how this is disrupted in many disease states.

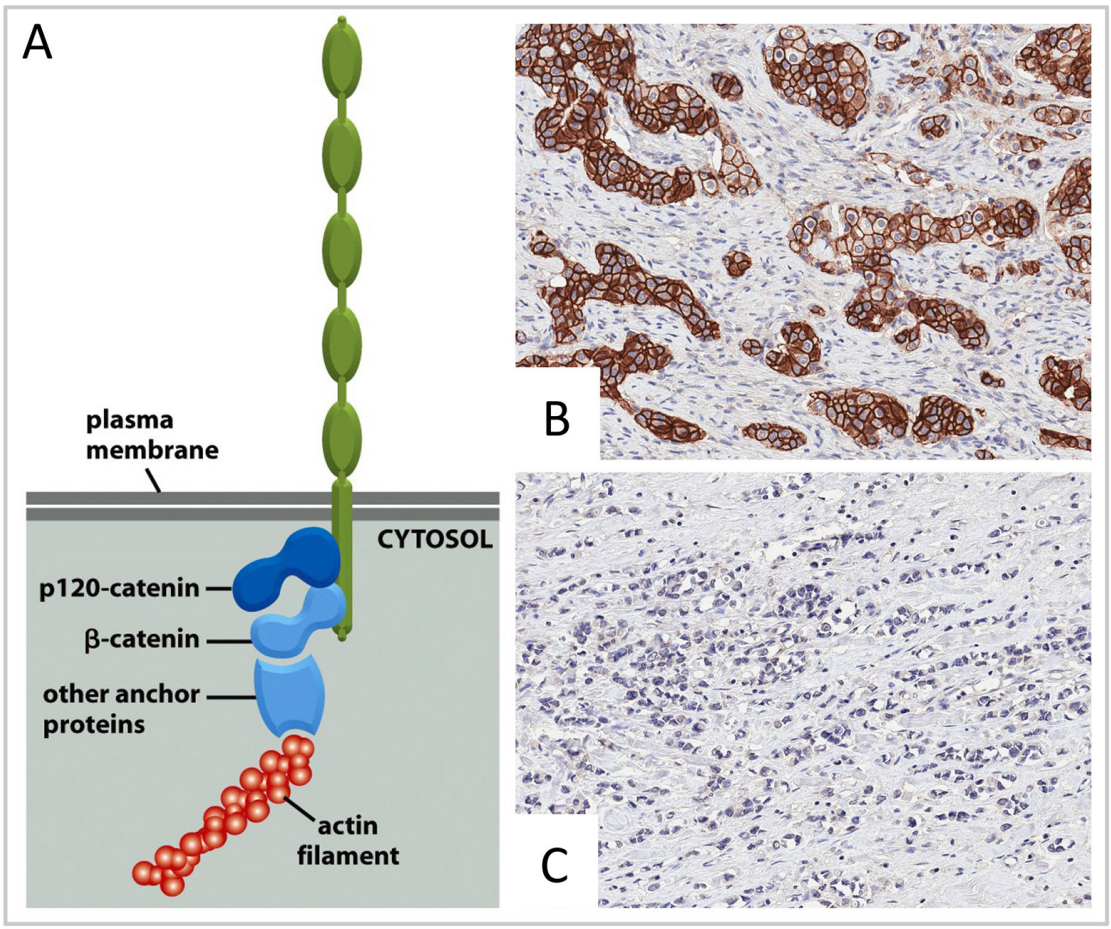


Figure 1.7: A) Schematic of the E-cadherin adhesion complex (Alberts et al., 2008); with permission from Garland Science. B) E-cadherin positive IC-NST. C) E-cadherin negative ILC.

1.8.1. Deregulation of E-cadherin in breast cancer

E-cadherin is considered a ‘master regulator’ of tumourigenesis; it is a tumour suppressor gene and a metastasis suppressor. Its role in functional adherens junctions prevents tumour invasion and migration, a fundamental step in the metastatic dissemination of tumour cells (Frixen et al., 1991, Vleminckx et al., 1991, Berx and Van Roy, 2001, Berx et al., 1995).

There is a strong genotype:phenotype correlation with the loss of *CDH1* (located on human chromosome 16q22.1) and E-cadherin protein loss in ILC (Berx et al., 1996)). There is a small of proportion of ILC tumours with detectable E-cadherin, however in these cases the IHC staining tends not to be completely ‘linear’, membranous staining. In fact, it has been shown that the E-cadherin-catenin protein complex in these cases is also probably dysfunctional (Da Silva et al., 2008), and is regarded as aberrant staining. Furthermore, the lobular component of mixed ductal-lobular carcinomas and some high-grade ductal cancers also display E-cadherin deregulation, however the mechanisms in which deregulation occurs is likely to be different. E-cadherin is used diagnostically to confirm an ILC, therefore the assessment of E-cadherin by IHC as a differential diagnosis between ILC and IDC is not always ideal, and should be based on morphology rather than E-cadherin staining (Dabbs et al., 2013).

The loss of membranous E-cadherin consequently affects the function and localisation of its membrane- binding partners. Indeed, when E-cadherin is lost, α-, β-, and γ-catenin protein expression is also frequently lost, while p120-catenin relocates to the cytoplasm. In fact this switch in cellular localisation of p120-catenin is considered a surrogate biomarker for lobular classification (Dabbs et al., 2007).

In lobular cancers, loss of E-cadherin is an early event in tumourigensis. The resulting impairment of functional cell-cell adhesion likely contributes to the characteristic discohesive nature of tumour cells in all lobular neoplasms (Mastracci et al., 2005, Da Silva et al., 2010), including the early precursor lesions atypical lobular hyperplasia (ALH) and lobular carcinoma *in situ* (LCIS) (Zou et al., 2009, Moll et al., 1993, Vos et al., 1997, De Leeuw et al., 1997, Etzell et al., 2001, Sarrio et al., 2003). This is supported by the identification of 16q loss in ALH and LCIS (Lu et al., 1998, Simpson et al., 2005b) and *CDH1* mutations in LCIS (Mastracci et al., 2005).
The various mechanisms of *CDH1* inactivation have been well documented. *CDH1* does not appear to be haploinsufficient, and inactivation tends to follow Knudson’s two hit

hypothesis, which states that a gene can only be inactivated if there is a “hit” in both alleles (Knudson, 1971). In ILC, *CDH1* is inactivated by any combination of gene mutation, promoter methylation, loss of the gene on 16q or by transcriptional repression (Berx and Van Roy, 2001). Most mutations in *CDH1* are frame-shift mutations that are predicted to yield truncated protein fragments. The mutations are scattered over the whole coding sequence and no hot spots have been identified (**Figure 1.8**, (Berx et al., 1996, The Cancer Genome Atlas Network, 2012, McCart Reed and Kutasovic et al., 2015)). *CDH1* promoter methylation is also evident in a large proportion of cases, and interestingly this has been found in non-neoplastic epithelial cells adjacent to lobular neoplasms, suggesting that these normal cells may have increased disease susceptibility and that in some cases, methylation may be the first ‘hit’ to affect the *CDH1* gene (Droufakou et al., 2001, Zou et al., 2009).

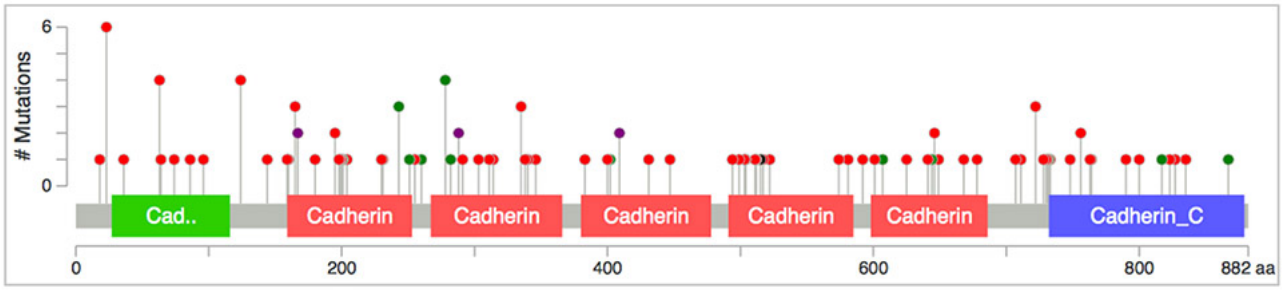


Figure 1.8: The mutation spectrum of *CDH1* (Cerami et al., 2012, Gao et al., 2013). Mutations have been found across the whole coding sequence of the gene.

Mutations in *CDH1* have also been identified in other types of epithelial cancers, most notably in diffuse gastric carcinoma (Berx et al., 1998, Hansford et al., 2015), which have a very similar diffuse growth pattern to ILC. In hereditary diffuse gastric carcinoma, 25-30% of patients inherit a germline mutation in *CDH1* and somatic methylation of the second *CDH1* allele is evident in the resulting tumours (Guilford et al., 1998, Fitzgerald et al., 2010). In families with *CDH1* germline mutation, females have a 42% chance of developing lobular breast cancer in their lifetime (Hansford et al., 2015). Genetic analysis of families with a history of ILC without gastric carcinoma, have shown that germline *CDH1* mutations in ILC are extremely infrequent, but have been reported (Petridis et al., 2014, McVeigh et al., 2014, Schrader et al., 2011), suggesting that familial ILC is driven by the germline mutation of other genes.

1.8.2. The role of the actin cytoskeleton in E-cadherin regulation

Recent evidence has emphasised the role for the actin cytoskeleton in the regulation of normal E-cadherin function, and the potential for its involvement in the deregulation of E-cadherin during tumourigenesis. The re-localisation of p120-catenin from the membrane to the cytoplasm has been shown to infer tumour cell survival by interacting with regulators of the actin cytoskeleton (Schackmann et al., 2011). In a mouse model of ILC, cytosolic p120 catenin was capable of activating Rho/Rock signalling, which activates downstream actin cytoskeleton remodelling pathways. This enabled the induction of anoikis resistance, which allows tumour cells to grow and survive independent of attachment to other cells – under normal conditions this detachment induces cell death (Schackmann et al., 2011). This pathway then facilitated invasion through the stroma and dissemination to distant organs.

Many other components of the actin cytoskeletal machinery are involved in adhesion junction maintenance including Myosin II isoforms (Smutny et al., 2010) and N-WASP (Kovacs et al., 2011). Additionally, F-actin was found to interact with α-catenin via a complex molecular pathway that both enables junctional tension between cells, and also provides a scaffold for actin polymerisation. This pathway involves the centralspindlin complex (Ratheesh et al., 2012). *In vitro* knockdown of the aforementioned molecules yielded E-cadherin deregulation and cells detached from one another, reflecting possible ways in which the actin cytoskeleton molecules regulate E-cadherin function in human cancers.

Cortactin is another actin-binding protein of interest. Located on chromosome 11q13, this Src substrate is frequently a hotspot for amplification and overexpression in cancer, particularly in breast and head neck carcinomas (Schuuring, 1995, Schuuring et al., 1998). Cortactin amplification has been found in up to 15% of breast cancers and is associated with ER/PR positive and HER2 negative tumours. Studies suggest that it may be a marker to determine risk of relapse in breast cancer (Hui et al., 1997, Hui et al., 1998). Cortactin has been found to increase cell motility by accumulating at lamellipodia (Huang et al., 1998, Bowden et al., 1999) and interact with the Arp2/3 complex to activate actin polymerisation (Uruno et al., 2001). This suggests a role for cortactin in the invasive capabilities of tumour cells and may be another potential mechanism involved in E-cadherin deregulation (Helwani et al., 2004). Disruption in any one of these molecules may therefore occur in tumours as a mechanism of destabilising the adherens junction resulting in loss of E-cadherin function. These mechanisms may contribute to the characteristic

discohesive and infiltrating nature of breast cancer cells and potentially to the metastatic progression to particular sites. The junctional cytoskeleton has been found to be greatly dynamic in its function and understanding this will give new insights into how the actin cytoskeleton regulates E-cadherin junctions in this disease state.

1.8.3. Epithelial to mesenchymal transition

E-cadherin deregulation is critical in Epithelial to Mesenchymal Transition (EMT). EMT is an essential process in embryogenesis, fibrosis and wound healing, enabling a cell to detach from its neighbours and migrate. This is essential for the arrangement of different cell types in the formation of organs and normal tissue structures (Thiery et al., 2009). In this process cells lose epithelial traits by repressing E-cadherin expression, in particular by the transcription factors TWIST, Snail and Slug. Cells take on a more mesenchymal phenotype, switching on the expression of markers such as N-cadherin and vimentin, and this enables more motile capabilities. This concept has sparked great interest as a key requirement for metastatic progression to occur, in particular enabling tumour cells to detach from the tumour mass, invade through both the local stromal tissue and the endothelial cell wall of blood or lymphatic systems. This capacity is also considered important for extravasation and invasion into the new host tissue. Since EMT is a dynamic and reversible process cells can then undergo mesenchymal to epithelial transition (MET) to enable colonisation in a new organ. These processes have been demonstrated frequently in experimental models of metastasis however it is difficult to observe in clinical samples due to the transient nature of the process. Nevertheless, markers of EMT, such as E-cadherin down-regulation, and N-cadherin and vimentin up-regulation are associated with a subgroup of high-grade ductal cancers that are triple negative (Sarrio et al., 2008, Karihtala et al., 2013, Aleskandarany et al., 2014). Unlike lobular cancers, where E-cadherin deregulation is an early and probably irreversible event driven by genomic alterations, this EMT phenomenon evident in some ductal cancers is more often a late stage and dynamic process. Infrequent expression of classic EMT markers (Fibronectin, Vimentin, N-cadherin, Smooth Muscle Actin, Osteonectin, Snail, Twist) in a large cohort also supports the suggestion that EMT does not underpin the invasive phenotype observed in ILC (McCart Reed and Kutasovic et al., 2015; under revision at *J Pathol*).

1.9. Molecular evolution of breast cancer

Like all cancers, breast cancer arises due to the accumulation of genetic alterations that manifest in the acquisition of numerous cellular features that enable the cell to sustain proliferative signalling, evade growth suppressors, resist cell death, enable replicative immortality, induce angiogenesis, activate invasion and metastasis, reprogram energy metabolism, and evade immune destruction (Hanahan and Weinberg, 2000, Hanahan and Coussens, 2012). These 'hallmark' processes provide selective advantages to the growth and development of a neoplastic clone (Stratton et al., 2009). The normal breast epithelium consists of two layers of cells - an internal layer of luminal cells surrounded by a layer of myoepithelial cells. These cells form the tree like structure of ducts and lobules that begin at the Terminal Duct Lobular Units (TDLU) and end at the nipple. The myoepithelial cells regulate normal mammary gland development and have been considered a natural tumour suppressor as they negatively regulate tumour cell growth, invasion and angiogenesis (Barsky and Karlin, 2005). Molecular abnormalities have been identified in histologically normal breast cells in women with breast cancer (Lakhani et al., 1999, Tripathi et al., 2008). Two hypotheses have been generated to explain the presence of molecular changes in the normal breast; the sick lobe hypothesis and the field cancerisation effect. The sick lobe hypothesis speculates that the genetic changes occur during early development of the breast, making the cells more susceptible to oncogenic stimuli (Tot, 2005). The latter hypothesis suggests these alterations occur as a result of acquired carcinogen events during adulthood (Dakubo et al., 2007).

Morphological evidence suggests that breast cancer arises from the normal TDLU of the breast, via a series of increasingly abnormal stages that progress into cancer over a long period of time. Intermediate stages in this multistep pathway involve the putative precursor lesions Columnar Cell lesions (CCL) (Simpson et al., 2005a) and atypical hyperplasias which can progress into an *in situ* carcinoma – a neoplastic clonal proliferation that is still confined to the ductal system (DCIS and LCIS). With the accumulation of further genetic alterations and through a dynamic interaction with the local microenvironment, malignant cells break through the basement membrane and invade into the surrounding stromal tissue to become an invasive carcinoma (**Figure 1.9**, (Simpson et al., 2005b, Lopez-Garcia et al., 2010)). These invasive cells may then acquire further attributes enabling them to invade through the stroma, intravasate into the lymph-vasculature system and metastasise to distant sites. This is a complex and variable process encompassing multiple precursor lesion types giving rise to a diverse series of invasive cancers with different biological and clinical features, and differing metastatic capabilities. The presence

of early neoplastic lesions has been associated with risk of developing invasive breast cancer. Predicting which lesions will or will not progress is difficult, and only small percentage of patients with IC-NST (20%) display ADH (Degnim et al., 2007) and approximately 50% of patients with DCIS progress to IC-NST (Collins et al., 2005) (similar findings are observed in ILC, described in more detail in **Section 1.5**). The management of patients with these early lesions can therefore be difficult and understanding of breast cancer progression is essential to improve patient management.

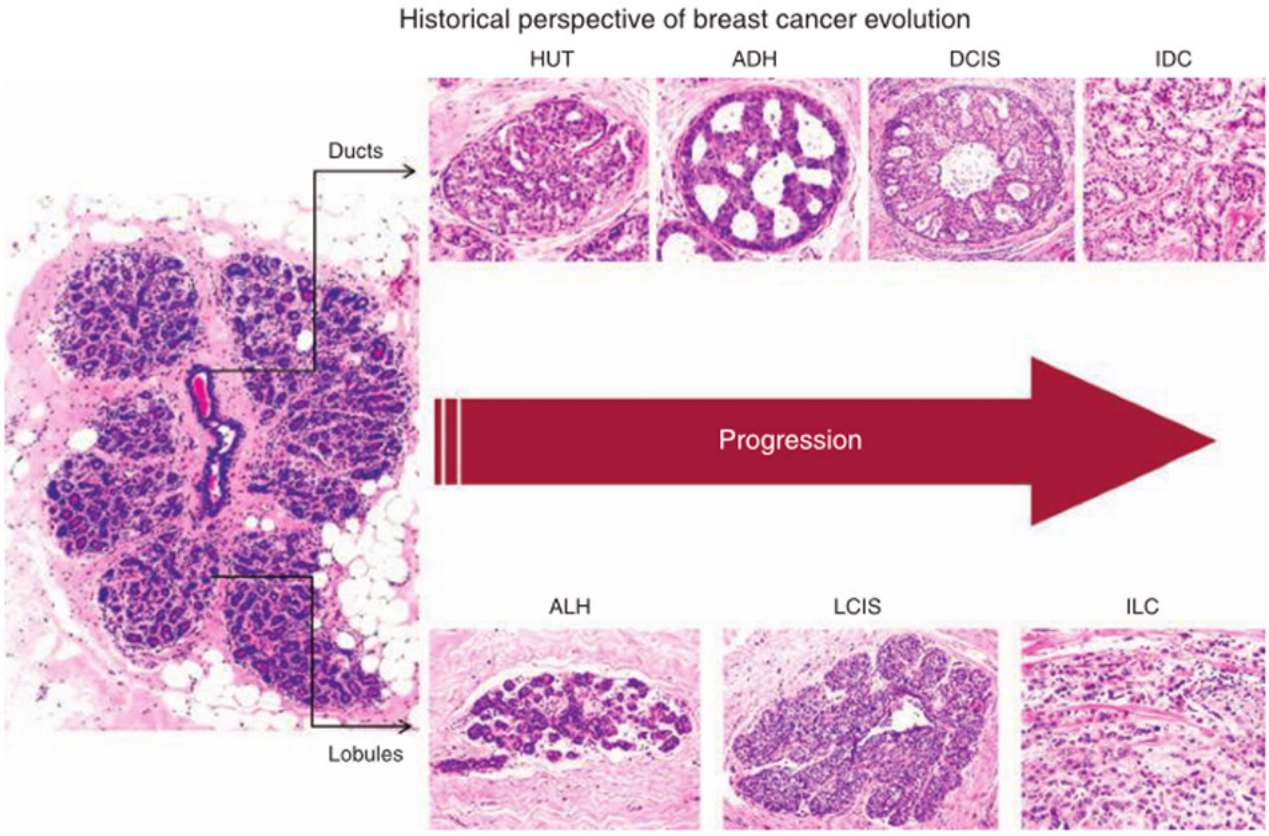


Figure 1.9: The historical model of the progression to invasive breast cancer. This model is out-dated as evidence has accumulated that shows an invasive cancer can arise from either ducts or lobules irrespective of the histological type, and hyperplasia of usual type (HUT) are no longer considered a precursor lesion. However, progression generally follows a pathway of morphologically recognisable stages over time from normal breast epithelium, to pre-invasive hyperplasia (ALH, ADH) and in situ carcinoma (LCIS, DCIS), through to the progression to invasive carcinoma (ILC, IC-NST (formally known as IDC)) (Lopez-Garcia et al., 2010) Permission from John Wiley and Sons.

Loss of Heterozygosity (LOH), Comparative Genomic Hybridisation (CGH), gene mutation analysis and gene expression studies have been utilised to understand clonal evolution during breast cancer progression. These methods have demonstrated that ER positive and

ER negative breast cancers have very distinct molecular profiles (Perou et al., 2000), and within ER positive disease, the extent of genetic alterations is associated largely with histological grade and proliferation (Roylance et al., 1999). Gene expression profiling has corroborated this by showing that lesions cluster according to histological grade and not the stage of tumour progression (Ma et al., 2003). These data have shown both that breast cancer progression is much more complex than originally suggested and that breast cancers of low- or high-grade follow distinct molecular pathways (**Figure 1.10**, (Simpson et al., 2005b, Lopez-Garcia et al., 2010)). For example, tumours of the low-grade pathway frequently express hormone receptors and display loss of 16q and gains of 1q and 16p (Abdel-Fatah et al., 2008). The high-grade group are much more heterogeneous, frequently lack the expression of hormone receptors and tumours of this group express HER2 (due to a high frequency of 17q amplification) and basal markers. The genomes of high-grade tumours are more frequently aneuploid and have very complex karyotypes (Natrajan et al., 2010). Loss of 1p, 8p and 17p and gain of 1q and 8q are the most frequently copy number changes observed in high-grade breast carcinomas (Lopez-Garcia et al., 2010).

Recently, RNA and DNA sequencing have been successfully applied to low input FFPE material investigating the molecular changes that occur in early breast neoplasias. Whole genome sequencing found shared somatic mutations and aneuploidies between early neoplasia and the matched invasive carcinoma (Newburger et al., 2013). These results suggest that the accumulation of mutational events that affect a large number of genes occurs in the early neoplasia as a result of increased ancestral cell division, and that in some cases these early events predispose the breast tissue to develop into a carcinoma. RNA sequencing was performed on a matched progression series of 25 patients to find that gene expression patterns in early neoplasia (such as ADH) were distinct from normal breast and breast cancer, with elevated transcription of *ERBB2*, *FOXA1* and *GATA3* even at the early stage of tumourigenesis (Brunner et al., 2014). This study reiterates the importance of ER signalling in tumourigenesis. These technologies are increasing our understanding of the events that occur early in the development of breast cancer and may pave the way in the ability to identify patients that may be at risk of progression.

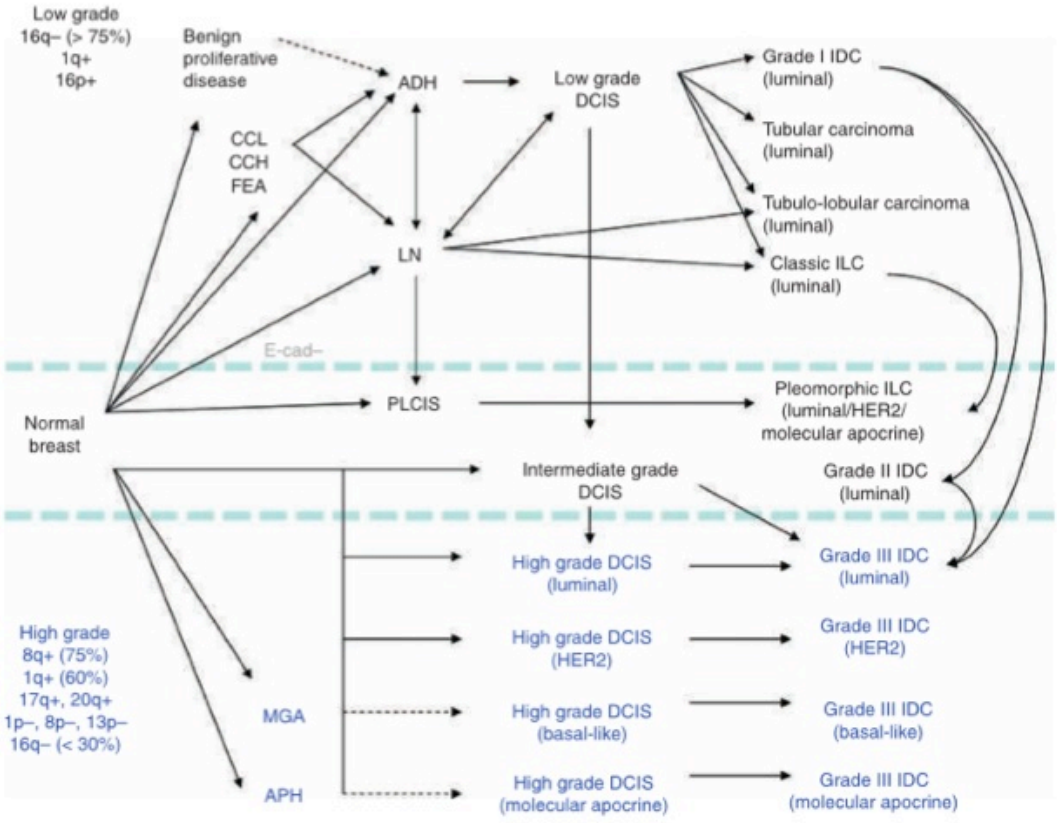


Figure 1.10: The high- and low-grade multistep model of breast cancer progression. (Lopez-Garcia et al., 2010) Permission from John Wiley and Sons.

1.10. Metastatic progression

Death from breast cancer is not commonly caused by the growth of the primary tumour, but the spread of the tumour to distant organs – the process known as metastasis. Up to 90% of all cancer deaths are the result of tumour dissemination to distant organs (Gupta and Massague, 2006). When the tumour remains localised in the breast and has not yet spread, the patient has a 98% chance of surviving more than 5 years. This decreases to 83% when the tumour has spread regionally to the lymph nodes, and decreases further to 23% once the tumour has spread to distant organs (Parise et al., 2009). Metastatic cancer is a significant burden and there is a critical need for a greater understanding of this process to improve patient survival.

1.10.1. The metastatic cascade

Metastasis is a stepwise biological process in which tumour cells hijack normal processes to facilitate the spread and colonisation of distant organs. Each step of the metastatic

cascade is complex, involving the acquisition of genetic alterations and interactions with the local microenvironment, making comprehensive understanding of this process challenging (Chiang and Massague, 2008).

For a tumour to successfully metastasise, cells must first detach from the primary tumour mass, resist anoikis, and activate pathways of local invasion. The cells degrade components of the extracellular matrix by the secretion of proteases that enable invasion through the stroma and into the vasculature. Together, this process is known as intravasation. Once cells have entered the vasculature, they must endure the harsh environment of velocity induced shear forces, lack of supporting substratum, and the presence of immune cells. The cells are now known as circulating tumour cells (CTCs), and the few that survive these conditions are largely inefficient at forming a secondary tumour, with as little as 0.01% successfully colonising a distant organ (Fidler, 1970). For colonisation to occur, circulating tumour cells escape the vasculature by arresting in endothelial layers of distant vasculature. Coagulation factors are produced by the cells to facilitate arrest in capillary beds and migration through vessel walls into the new microenvironment, a process known as extravasation. Here, the tumour cells must form its own vasculature to provide oxygen, nutrients and growth factors to survive and seed a clinical metastasis (**Figure 1.11**, (Poste G., 1979, Fidler, 2003).

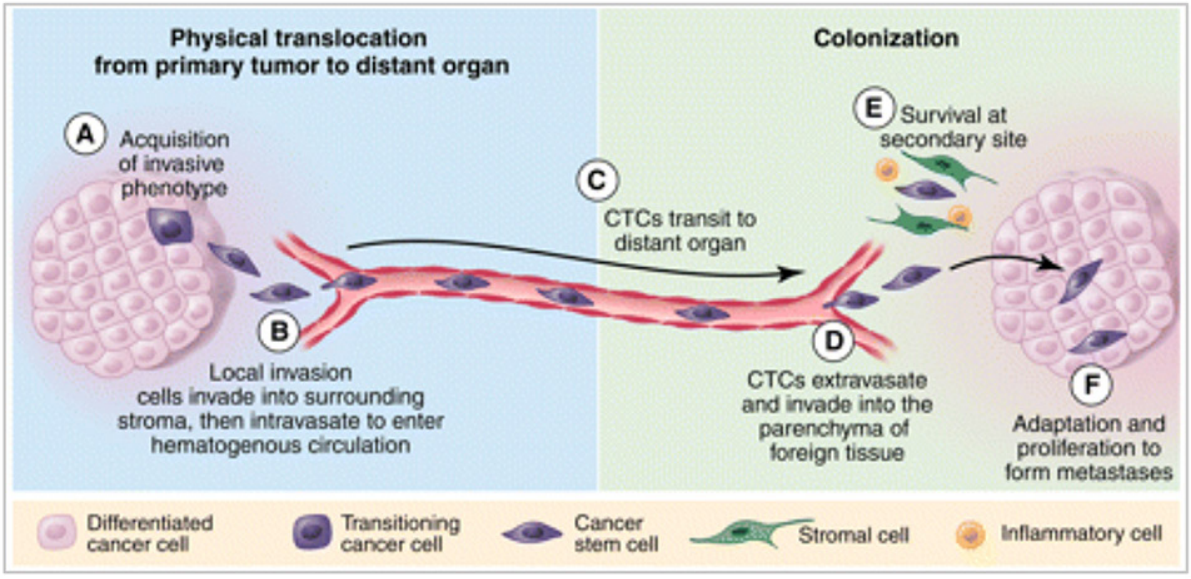


Figure 1.11: The metastatic cascade. Metastasis is a step-wise process that involves the detachment of cells from the primary tumour, the capability to invade the local microenvironment, invade to circulatory system, survive in the bloodstream and colonise and survive at a distant site. Each step in the process requires the cells to exploit many

different signalling pathways to survive (Chaffer and Weinberg, 2011). Copyright © 2011, American Association for the Advancement of Science.

1.10.2. The pre-metastatic niche

For a metastasis to develop, the circulating neoplastic cells must find an appropriate niche or adapt to a new microenvironment. Recent evidence has shown that both the primary tumour and CTCs are capable of sending out signals (through chemokines and exosomes – membrane bound vesicles that contain protein and nucleic acids (Kaplan et al., 2005, Peinado et al., 2012, Ghajar et al., 2013)) that can prime target sites and modify their microenvironments to support the successful colonisation of CTCs. This has been demonstrated in a mouse model, where human tumour exosomes were extracted from the blood of metastatic melanoma patients and injected into mice, increasing the metastatic behaviour of both mouse and human primary tumours cells. The exosomes were able to reprogram bone marrow progenitor cells toward a pro-vasculogenic phenotype and provide a favourable environment for tumour cells to colonise and increased tumour burden in these mice (Peinado et al., 2012). These primed microenvironments have been called the 'pre-metastatic niche' and greater understanding of how cells reprogram distant sites, or why some cells choose particular sites over others, may provide insights into mechanism that can block the colonisation of distant organs and ultimately preventing death from metastasis.

1.10.3. Seed and Soil Hypothesis

There are two long-standing theories defining the mechanism of cancer spread; (i) Paget's 'seed and soil hypothesis', which states that the pattern of metastasis is governed by the intrinsic properties of the primary tumour (Paget, 1889), and (ii) Ewing's proposal that metastatic dissemination is simply determined based on the mechanical flow of the blood and lymph around the body (Ewing, 1928). Metastasis was first considered to be a non-random process by Stephen Paget who in 1889 hypothesised through the analysis of a large autopsy series that metastases will only develop when the cells encounter a suitable environment (Paget, 1889). This pattern of organ preference is given the term organotropism, and successful colonisation depends on the molecular interactions between the tumour cells ("seed") and the microenvironment ("soil"). James Ewing countered this in 1928 by suggesting that tumours spread in the direction of blood flow and

hence form metastases in the organs directly in line from the site of the primary tumour (Ewing, 1928). There is evidence to support both of these theories in breast cancer metastasis.

The most common sites of breast cancer metastasis are the regional lymph nodes, lung, liver and bone (Lee, 1983). Overall, all histological types and molecular subtypes can spread to these organs, suggesting that to some degree, these organs are targeted directly by blood and lymphatic flow. However, the distribution of the spread of histological and molecular types to these organs is skewed to some extent, suggesting there may be some element of organotropism too. Through analyses of autopsy and surgical series comparing the organ specific metastatic spread of the main histological types, IC-NST and ILC it has become clear that ILC spread more frequently to the peritoneum, gynaecological organs, gastrointestinal tract, adrenal glands, central nervous system and skin, while IC-NST spread more frequently to the lung/pleura, liver and brain (Sastre-Garau et al., 1996, Jain et al., 1993, Borst and Ingold, 1993, Harris et al., 1984, Lamovec and Bracko, 1991, Arpino et al., 2004, Cummings et al., 2014). Further, a study by Porter et al., demonstrated that low-grade tumours more frequently spread to the bone whereas high-grade tumours spread to the lung and liver (Porter et al., 2004). The molecular subtype of the primary tumour also impacts the distribution of metastatic disease. For example, ER+/PR+ luminal tumours preferentially spread to the bone but not the brain, while ER-/PR- tumours tend to spread to the brain and less frequently to the bone and over widely different time periods; a much shorter latency for ER negative tumours (Maki and Grossman, 2000). HER2 positive tumours frequently colonise the brain, liver and lung, and the triple negative/basal breast cancer subtype have a propensity to spread to the brain, lung and liver (Sihto et al., 2011, Harrell et al., 2012). High-grade ductal cancers expressing the basal marker CK14 have an increased propensity to spread to the brain compared to CK14 negative tumours. Interestingly, CK14 positive tumours spread to the lung more quickly than CK14 negative tumours but the overall frequency of lung metastases was not different (Fulford et al., 2007). The data indicate that the tumour type, grade and molecular features could help inform patient management with respect to distant sites most at risk of being colonised, which might aid surveillance of those organs during clinical follow up.

1.10.4. Linear versus parallel progression models

Two fundamental conceptual models of breast cancer progression prevail. Generally, the phenotypic and genomic features of the primary tumour are reflected in the resulting metastases, supporting a stepwise, linear progression model where genomic alterations accumulate after multiple rounds of clonal selection and expansion within the primary tumour, until a cell has gained the capacity to disseminate. This has been widely accepted since patients with larger primary tumours at diagnosis have a higher risk of death from metastatic disease (Greene and Sobin, 2008). However, metastases can still arise from small primary tumours, and the linear model does not account for metastases of an unknown primary site. The model of parallel progression suggests, however, that in a polyclonal tumour mass, cells can leave and disseminate at any point in time of the development of the primary tumour. The primary tumour and metastases have increased biological variance as they continue to evolve in parallel (Yachida et al., 2010) (Weigelt et al., 2003, Liu et al., 2009) (Torres et al., 2007, Kuukasjarvi et al., 1997a, Navin et al., 2011, Cummings et al., 2014, Klein, 2009, Kutasovic et al., 2014). There is greater support for the parallel progression model due multiple lines of evidence, including the calculation of near identical tumour growth rates between primary and metastatic tumours (Friberg and Mattson, 1997, Engel et al., 2003), metastases derived from unknown primary tumours and the colonisation of multiple sites (Husemann et al., 2008). Disseminated tumour cells have been found in the bone marrow showing fewer genomic alterations than the primary tumour (Schmidt-Kittler et al., 2003) and in approximately 5% of breast cancer cases, distant metastases are identified at primary tumour diagnosis (Engel et al., 2003). However, it is likely that both linear and parallel progression can potentially occur Tumour cells have the ability to adapt to selective pressures, such as a new microenvironment and treatment, and therefore discordant features will be observed, whether the disseminated tumour cells left early or late. One study has calculated that a metastasis can be initiated 5-7 years before diagnosis of the primary tumour (Engel et al., 2003), consequently, an important area of research is now asking how tumour cells lie quiescent for many years and what breaks their dormancy (Hadfield, 1954, Sosa et al., 2014)? The distinction between these two pathways of progression has clinical implications as to how a metastasis should be treated. If the metastasis has the same biological properties as the primary tumour (linear model) it is justified to treat it based on the features of the primary tumour. If it evolved through the parallel model, the metastasis will likely have divergent features, and should be biopsied to guide more appropriate treatment.

1.10.5. Intratumour heterogeneity in breast cancer progression

The morphological and phenotypic intertumour heterogeneity among breast cancer patients has long been recognised, however the clinical implications of intratumour heterogeneity has recently come into the spotlight, especially in regards to the complex process of clonal evolution in metastatic progression. Phenotypic and genomic analyses of primary breast carcinomas and their matched metastases highlight the extent of heterogeneity within progression. The expression of tumour biomarkers has been found to change during tumour evolution and progression to metastasis, with the most frequent discordance being expression of ER and PR, which are frequently down-regulated with the spread from breast to lung, bone and liver (**Figure 1.12**, (Cummings et al., 2014, Singhakowinta et al., 1976, Wu et al., 2008, Idirisinghe et al., 2010). HER2 overexpression is generally more stable during progression, yet discordance has been reported in approximately 10% of cases (Fabi et al., 2011). There are a number of possible explanations for this phenomenon, including; evolution in response to microenvironmental changes or treatment, or the outgrowth of a minor subclone of a heterogeneous primary tumour that harboured a different genotype or phenotype.

Traditionally it has been thought that there is little difference in the molecular profiles between the primary and secondary tumours during progression, however there is evidence that this is not the case (Cummings et al., 2014, Kuukasjarvi et al., 1997a, Almendro et al., 2014, Singhi et al., 2012, Nik-Zainal et al., 2012). Elegant deep sequencing and copy number profiling studies of breast (Navin et al., 2011, Nik-Zainal et al., 2012, Shah et al., 2009, Ding et al., 2010, Yates et al., 2015), renal cell (Gerlinger et al., 2012), prostate (Liu et al., 2009), colorectal (Baldus et al., 2010), and pancreatic (Campbell et al., 2010, Yachida et al., 2010) carcinomas, as well as leukaemias (Landau et al., 2013) exemplify these findings at nucleotide resolution and demonstrate that clonal evolution during metastatic progression can be very complex in some cases. The data implies that a primary tumour can be polyclonal and that different subclones can develop metastatic capability and spread to different distant sites, thus yielding genomic diversity between different metastases in the same patient. It has been found that in some cases the primary tumour and its metastasis shared very little genomic similarity, with as few as 30% of mutations being shared among all lesions (Navin et al., 2011, Gerlinger et al., 2012, Yachida et al., 2010) (reviewed here (Kutasovic et al., 2014)). Further clonal

evolution can occur at the metastatic sites involving the alteration of key driver genes such as *KRAS, MYC, CCNE1* (Campbell et al., 2010, Singhi et al., 2012).

Recent multiregional whole-genome sequencing of breast tumours revealed that the subclonal diversity varied between cases, that driver mutations where identified in small subclones of the primary tumour and the clone that seeded the metastasis was already present in the primary tumour. The timing of clonal evolution also varied, were the divergence occurred early in some cases, while in other cases divergence occurred later during tumour progression (Yates et al., 2015). As we gain understanding of the clonal heterogeneity and phenotypic diversity in breast tumours we are beginning to understand how this is may be impacting upon patient treatment and how patient management might have to be changed in the future to address these issues. Since treatment options for metastatic breast cancer are typically based on characteristics of the primary tumour, then management is likely to be suboptimal in a number of patients.

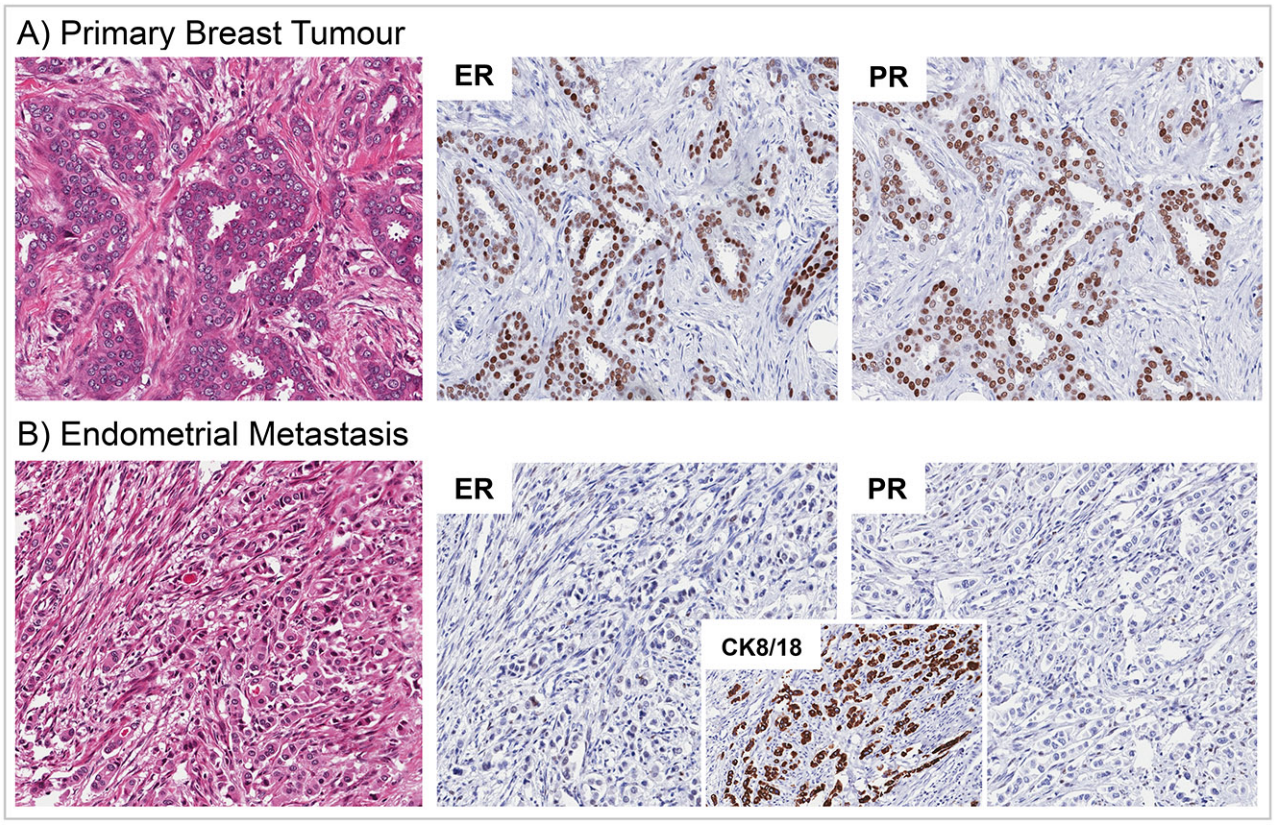


Figure 1.12: An example of heterogeneity during metastatic progression. The patient presented with an ER, PR positive IC-NST, and relapsed with a lobular-like ER, PR negative metastasis in the endometrium. CK8/18 staining demonstrates the carcinoma cells.

There are now several important examples of the selective pressure of chemotherapy or targeted therapy specifically driving the evolution of treatment resistant subclones. To receive anti-endocrine therapy, a patient's breast tumour must express ER in greater than 1% of the tumour sample tested. Therefore, it may be likely that the ER negative cells present within a tumour were never responsive to targeted therapy. As the ER positive cells die as a result of the targeted therapy, the outgrowth of the ER negative subclones can occur without restriction. Activating mutations in *ESR1* (ER alpha gene) represents a mechanism of resistance to prolonged endocrine therapy in hormone receptor positive breast cancer and accounts for the resulting development of metastatic disease (**Section 1.3.1** (Fuqua et al., 2014, Li et al., 2013)). This has also been demonstrated in tumours with heterogeneous HER2 amplification (Ng et al., 2015). Patients are eligible for anti-HER2 therapy when the tumour has a significant HER2 amplification in at least 30% of tumour cells. However, in the HER2 wild type cells within the same tumour, other genomic amplifications and HER2 somatic mutations may be compensating for lack of amplified HER2 driven tumour growth. This may account for the short time to relapse in HER2 positive patient tumours, where the HER2 negative cells did not respond to therapy.

1.10.6. Metastasis to gynaecological sites

Secondary malignancy of gynaecological organs accounts for 7-10% of all gynaecological malignancies (Young and Scully, 1991) and are hard to diagnose, as they frequently mimic a primary carcinoma (Bruls et al., 2015). Primary breast cancer is the second most common tumour to spread to gynaecological sites (Moore et al., 2004, Kondi-Pafiti et al., 2011, Yada-Hashimoto et al., 2003, Bruls et al., 2015), behind colorectal cancer(Kondi-Pafiti et al., 2011, Yada-Hashimoto et al., 2003, Moore et al., 2004)(Kondi-Pafiti et al., 2011, Yada-Hashimoto et al., 2003, Moore et al., 2004)(Kondi-Pafiti et al., 2011, Yada-Hashimoto et al., 2003, Moore et al., 2004). Mechanistically, this is interesting and indicates something peculiar about the biology of the primary tumour or the target organ that facilitate this pattern of spread. There is little data in the literature exploring this phenomenon, however it is a consistent finding. Primary colorectal carcinomas are the most frequent to spread to gynaecological sites, ranging from 39 to 57% of cases (Bruls et al., 2015, de Waal et al., 2009, Kondi-Pafiti et al., 2011, Skirnisdottir et al., 2007). This is suggestive of transcoelomic spread or lymphatic spread (Bruls et al., 2015), whereby the cells travel down through the peritoneal cavity towards the gynaecological sites. Conversely, the tendency for breast cancers to spread to bilateral ovaries (63.9% of cases

(Bruls et al., 2015)) suggests that haematogenous spread might be involved.
In up to 38% of gynaecological metastasis (GM) patients, the secondary tumour is detected before the primary (Skirnisdottir et al., 2007) and as it is inherently clinically silent, the metastatic disease is at an advanced stage when diagnosed. Patients can be asymptomatic for many years (Bigorie et al., 2010), and 75% of patients with ovarian metastases experience symptoms that are not necessarily associated with malignancy such as bloating, abdominal pain, postmenopausal bleeding and weight loss (Moore et al., 2004). Thus by the time the metastases are diagnosed, they can be widespread and hence overall survival is poor: 52% of patients survive 2 years, 24-36% survive 5 years, and 7% survive 10 years (Ayhan et al., 2005, Skirnisdottir et al., 2007, Demopoulos et al., 1987, de Waal et al., 2009, Salamalekis et al., 2004). In comparison to other primary sites, primary breast cancers have a better 5 year survival rate of 26-40% compared with 32% for colon and 12% for stomach primaries (Ayhan et al., 2005, Skirnisdottir et al., 2007, de Waal et al., 2009, Demopoulos et al., 1987).

Small studies of breast cancer patients with GM have found that patients were younger in age at diagnosis of their primary tumour (median age ranging from 46-54 years) (Bruls et al., 2015, Demopoulos et al., 1987, de Waal et al., 2009, Kondi-Pafiti et al., 2011, Salamalekis et al., 2004). This is interesting considering that in Australia the average age of sporadic breast cancer diagnosis is 60 years (Cancer Australia, 2012) and the average age of ILC diagnosis is around 64 years (Arpino et al., 2004, Borst and Ingold, 1993). There is an enrichment for lobular carcinomas to spread to these sites with 43% of the breast cancers spreading to gynaecological organs being ILC, whereas ILC accounts for up to 15% of all breast cancers in the general population (Arpino et al., 2004, Bruls et al., 2015, Bigorie et al., 2010). Metastases from ILC have a characteristic discohesive growth pattern that enables them to grow without damaging the epithelium (Lamovec and Bracko, 1991). E- cadherin dysfunction is presumably involved in the mechanisms of ILC metastasis, or at least in the diffuse growth pattern that results in these particular metastatic sites. Furthermore evidence from autopsy studies show that gynaecological metastases from mixed ductal lobular carcinomas are frequently of the lobular type (Lamovec and Bracko, 1991). ILC that spread to the peritoneum, gastrointestinal tract and gynaecological organs were also E-cadherin negative (88%; (Ferlicot et al., 2004)) and predominately ER positive and HER2 negative (Bigorie et al., 2010, St Romain et al., 2012). This suggests a possible role for hormone regulation and E-cadherin dysfunction in this unique process (Lamovec and Bracko, 1991).

In summary, it is important to understand the natural history of metastatic disease, as it will impact the management of the patient. For example, breast tumours that spread to the brain tend to be associated with the basal-like phenotype and metastasis occur relatively quickly, with patients surviving only up to 9 months after diagnosis (Lagerwaard et al., 1999). Conversely, metastasis to gynaecological organs seems to affect young women and there is enrichment for ILC and ER positive tumours. However, due to the asymptomatic nature of metastases to gynaecological sites, the interval between primary breast diagnosis and metastasis can be between 5 to 20 years (Demopoulos et al., 1987, Bigorie et al., 2010) meaning there is an opportunity for early detection and prevention.

1.11. Aims and objectives of this thesis

The work investigated in this thesis originates from our inter-related interests in ILC, E-cadherin and understanding mechanisms of invasion and metastasis in breast cancer. The general hypothesis is that tumour clones must evolve during development and progression in order to develop the necessary biological capabilities to grow and then metastasise. This clonal evolution is driven by the selective acquisition of somatic mutations and through dynamic interactions with the local microenvironment to facilitate resistance to treatment and survival.

One aspect of this is the ability of cells to alter their cell adhesion properties and acquire an invasive phenotype. Tumour cells of ILC are inherently invasive due to the dysfunction of E-cadherin regulated cell-cell adhesion. In many cases of ILC this is driven by *CDH1* mutations however other mechanisms are also likely to be involved in deregulating E-cadherin. Regulators of the actin cytoskeleton play an important role in regulating functional E-cadherin (**Section 1.8.2**, **Chapter 3**) and so we hypothesise that tumour related disruption of such molecules may contribute to destabilising the adherens junction, resulting in loss of E-cadherin function and enhanced tumour cell invasion. The following aims were undertaken in order to test this:

- Perform a meta-analysis of public molecular data to determine whether genes involved in actin cytoskeleton regulation are altered at the gene expression level in breast cancer
- Perform IHC on selected molecules of interest on tissue sections of normal breast tissue and breast cancers by tissue microarray (TMA) to assess their cellular localisation and to correlate their expression pattern with E-cadherin status.

One of the most apparent examples of intratumour heterogeneity within primary breast tumour is the mixed ductal-lobular carcinoma (MDL), which are composed of both ductal and lobular histological components (**Section 1.6**, **Chapter 4**). Mixed tumours evolve either as independent tumours that have collided or, as is more likely, from a common clonal population that diverges from ductal to lobular growth pattern due to ongoing genetic instability or cellular plasticity. We hypothesise that deregulation of the functional complex regulating cell-cell adhesion and involving E-cadherin is driving this phenotypic switch, resulting in an enhanced invasive capability of cancer cells. The following aims were undertaken to investigate this:

- Collate and characterise the clinical and morphological features of a large cohort of MDL cases.
- Perform IHC for E-cadherin, β-catenin and p120-catenin on whole sections of MDL cases to determine whether the E-cadherin complex is specifically disrupted in the lobular component of these tumours.
- Determine whether an epithelial to mesenchymal transition accounts for this apparent change to an invasive growth pattern
- Use a discovery-based exome sequencing approach to determine whether the change in morphology is driven by genomic alterations

Intriguingly, ILC tumours have a propensity to colonise less common sites such as the gynaecological organs, gastrointestinal organs and the peritoneum (**Section 1.10.6**, **Chapter 5**). We hypothesise, therefore, that there is an underlying peculiarity in the biology of the primary tumour and/or the target organ to facilitate this pattern of spread. There is little existing data exploring this phenomenon and so the following aims were undertaken to investigate this:

- Collate a series of breast cancer cases that have spread to gynaecological sites,
- Characterise the clinical and morphological features of this cohort.
- Investigate the immunohistochemical phenotype of the primary and matched metastatic tumours.
- Perform molecular analysis to investigate genomic heterogeneity during progression to this unique metastatic site.

CHAPTER 2

MATERIALS AND METHODS

2. Materials and methods

2.1. Breast cancer patient cohorts

Fresh Frozen (FF) tissue samples and matched blood samples were obtained from the Brisbane Breast Bank based in the Lakhani Lab. FF samples were used for immunofluorescence experiments in **Chapter 3**. Blood samples were used to extract normal DNA utilised in **Chapter 4**. Pathology reports and Formalin Fixed Paraffin Embedded (FFPE) tissue samples were requested from Pathology Queensland and Sullivan Nicolaides (S&N) Pathology laboratories, through the lab head Prof Sunil Lakhani and A/Prof Margaret Cummings, and were used for all IHC experiments and the molecular analyses performed in **Chapters 3**, **4** and **5**.

Several patient cohorts were used during the course of this thesis. Firstly, tissue microarrays (TMAs) representing three main subtypes of breast cancer and breast cancer progression ((i) Luminal (ML1 TMA), (ii) HER2 (HER2 TMA), (iii) Triple Negative (TN TMA) and (iv) Progression (PS1 and PS2 TMAs)) were used in **Chapter 3**. The tissue cores on these TMAs were derived from cases from the Royal Brisbane and Women's Hospital (RBWH).

Secondly, TMAs consisting of a cohort of 148 invasive lobular carcinomas (ILC) obtained from S&N were used in **Chapter 4** (TMAs named WL1-4). This cohort displays typical characteristics for this tumour type; 83.3% were grade 2, 96.5% were Oestrogen Receptor (ER) and/or Progesterone Receptor (PR) positive, 1.4% were HER2 positive, 1.4% were negative for ER, PR and HER2. E-cadherin was negative in 75.8% and aberrantly distributed in 22.8% of cases, and p120-catenin was aberrantly located to the cytoplasm in 90% of samples (McCart Reed and Kutasovic et al., manuscript under review, *J Pathol*). In both **Chapters 4** and **5**, an unselected cohort of 449 sporadic breast cancer patients diagnosed between 1987 and 1994 at the RBWH was used for comparison statistics. This cohort is named the Queensland follow up cohort (QFU TMAs) and displays features of the general breast cancer population that is recognised worldwide. For example, 57% were invasive carcinoma of no special type (IC-NST), 14% were ILC and 77% of the cohort was ER positive (Al-Ejeh et al., 2014, Junankar et al., 2015). In **Chapter 4**, pathology reports and FFPE tissue blocks from a cohort of Mixed Ductal-Lobular Carcinomas (MDL) were collected from RBWH and S&N. In order to establish the specific

gynaecological metastasis (GM) cohort of breast cancer patients for **Chapter 5**, the Queensland Centre for Gynaecological Cancer (QCGC) clinical database was accessed to identify patients that had a metastasis to gynaecological sites (in collaboration with Prof Andreas Obermair). Pathology reports for this cohort were also accessed from Pathology Queensland. Death and survival data were obtained from the Queensland Cancer Registry (QCR), where I am a visiting staff member, and treatment data were obtained via the Clinical Administrative Services (CAS, Queensland Health) database. The patient data for both the MDL and GM cohorts was then assembled into a password protected Microsoft Excel database that contained all clinical and pathology data for all patients with for analysis, including parameters such as the age of diagnosis, size and grade of the tumour, lymph node and biomarker status.

2.2. Ethics

We have approval from the RBWH and University of Queensland Human Research Ethics Committee (HREC; 2005000785) for use of patient samples and clinical data (see **Appendix 2.1** for the approval letter). All patients have completed informed consent prior to banking their samples in the Brisbane Breast Bank (BBB). Retrospective samples from archival specimens do not require consent as some patients will have died, others are lost to follow up or are no longer followed up and so it is inappropriate to now seek consent.

2.3. Statistical analysis and graphical representations

Statistical analysis was performed using GraphPad Prism version 6. Chi-square, Fisher's exact test and student's t-test was performed to evaluate differences between clinical, pathology and biomarker data between the cohorts under investigation. A p-value of <0.05 was considered significant. All Venn diagrams were created using Venny 2.0 (Oliveros, 2007-2015).

2.4. Tissue microarray construction

A TMA contains many tissue samples (up to several hundred samples) arranged in a paraffin block, allowing the assessment of any biomarker across many tissue samples. The combination of these techniques provides a high-throughput way of assessing many different molecules. The ML-1, HER2, TN, PS1-2, WL1-4 and QYFU TMAs were existing

resources generated by previous and current members of the Lakhani lab. The gynaecological metastasis cohort TMA was made specifically for this research. Each tissue block was sectioned at 4 μM thickness and stained with haematoxylin and eosin (H&E; **see 2.5.2**) for morphological analysis and tumour identification. Areas of tumour were identified and annotated on the slide in preparation for the construction of tissue microarrays (TMA). The first three TMAs (GM1, GM2, GM3) were constructed using a Beecher Instrument microarray Technology (www.beecherinstruments.com). TMAs GM4, GM5 and GM6 were built using TMArrayer (Pathology Devices, http://www.pathologydevices.com/TMArrayer.htm). Both devices used 1 mm hollow needles to extract multiple tissue cores from the donor block. One millimetre cores were chosen, as opposed to 0.6 mm cores, for this project to sample a greater area of tissue and therefore to try and capture as much tissue heterogeneity as possible. The tissue cores were placed precisely into a recipient 'empty' paraffin block in duplicate. The cores were arranged with predefined coordinates that results in a series of patient samples organised onto a grid. The coordinates for this cohort were designed with an individual case represented in a row to allow assessment of phenotypic changes across multiple tumours in a patient (see **Appendix 2.2 to 2.7** for the TMA maps). Eight or nine cases were placed on one TMA (8-9 rows), with the number of metastatic sites per case determining how many cores were in each row (12 cores being the maximum number in each TMA). A space of 1.7 mm was used in between each core. Unmatched normal breast was also included on each TMA as a quality control and normal liver tissue was included to assist with orientation of the TMA. The completed TMA block was placed in a 55°C oven for 10 minutes to set the cores into the new paraffin block, it was then allowed to reach room temperature before being placed at 4°C for storage. A 4 μM section was cut after construction of the TMA and an H&E stain was performed to verify the presence of tumour tissue in each core. Immunohistochemical staining was then performed on the individual slides.

2.5. Histological and immunohistochemical Methods

2.5.1. Sectioning and deparaffinisation

Tissue sections were cut at 4 μM from the relevant patient block or TMA using a microtome (Leica RM2135). Tissue sections were floated on water before being mounted on cation coated slides and placed in a 37°C oven overnight. The slides were

deparaffinised in 100% xylene 3 times for 5 minutes each and rehydrated for 2 minutes each in a series of graded alcohols (100%, 90%, 70%) to water.

2.5.2. Haematoxylin and Eosin staining

Haematoxylin and Eosin (H&E) staining was performed using the Leica ST5010 Autostainer XL with an automatic cover-slipping machine at the Histology Laboratory in the QIMR Berghofer Medical Research Institute (QIMRB).

2.5.3. Nuclear fast red staining

Tissue sections destined for DNA extractions were stained with nuclear fast red. After deparaffinisation, the tissue sections were incubated in nuclear fast red for 2 minutes. The sections were washed twice in water, and then dehydrated in graded ethanol (70%, 90%, 100%). The sections were allowed to air dry before placing at 4°C until ready for micro- or macro-dissection (see **Section 2.7.2**).

2.5.4. Immunohistochemistry (IHC)

The general protocol for IHC is given below. The specific Immunohistochemistry (IHC) conditions for each antibody are summarised in **Table 2.1**.
Antigen retrieval improves an antibodies ability to bind to the antigen. Antigen retrieval was performed using the decloaker method with either sodium citrate (0.01M, pH 6.0, 125°C for 5 minutes), EDTA (0.001M, pH 8.8, 105°C for 15 minutes) or chymotrypsin (37°C for 10 minutes) depending on the antibody used. All wash steps were performed using 1 X Tris Buffered Saline (TBS, pH 7.4). The slides were washed 3 times for 2 minutes each. The Mach1 Universal HRP Detection Kit (Biocare Medical, LLC. Concord CA 94520 USA) was used for detection as per the manufacturers instructions. Briefly, after deparaffinisation, antigen retrieval and washes, the slides were placed in 30% hydrogen peroxide (Sigma Aldrich, lot # 33420) for 10 minutes to block endogenous peroxidases. After washing, Sniper (Biocare Medical, LLC. Concord CA 94520 USA) was added to each section to block non-specific proteins. All antibodies were diluted in TBS, and their specific conditions optimized, including dilutions and incubation conditions (1 or 2 hours at room temperature, or 4°C overnight) are presented in **Table 2.1**. The Mach1 mouse probe was added for 15 minutes if a mouse antibody was used and the universal Horse-Radish Peroxidase (HRP)

polymer was used as the secondary antibody for 30 minutes. The Biocare Medical Rat Detection Kit was used for rat-derived antibodies. A solution of 3,3'-Diaminobenzidine (DAB substrate buffer and DAB chromagen - Biocare Medical, LLC.) was utilised for antibody visualisation by light microscopy. The slides were washed in water followed by counterstaining with haematoxylin and cover slipping at QIMRB Histology laboratory.

Table 2.1 Immunohistochemistry antibodies and conditions

Antibody (Clone)	Cat #	Species	Source	Detection method	Antigen retrieval	Antibody dilution	Inc. time	Positive control	Cellular localisation	Scoring method
ER (6F11)	NCL-L-ER-6F11	M	Novocastra	MACH1	Citrate	1/100	1 h	NB	Nucleus	Positive in the nucleus in >1% of cells
PR (16)	NCL-PGR	M	Novocastra	MACH1	Citrate	1/100	1 h	NB	Nucleus	Positive in the nucleus in >1% of cells
HER2 (A0485)*	A0485	R	Dako	MACH1	Citrate	1/100	1 h	NB	Membrane	Positive if 3+ membrane staining
HER2 (4B5)*	790-2991	R	Ventana	Ventana Ultraview DAB	Ventana CC1 20 mins	n/a	8 min	HER2+ tumour	Membrane	Positive if 3+ membrane staining
EGFR (31G7)	280005	M	Invitrogen	MACH1	Chy.	1/100	1 h	Skin	Membrane	Any positivity and percentage of positive cells
CK8/18 (5D3)	NCL-5D3	M	Novocastra	MACH1	Citrate	1/100	1 h	NB	Cytoplasm	Any positivity and percentage of positive cells
CK5/6 (D5/16B4)	MAB1620	M	Millipore	MACH1	Citrate	1/300	1 h	NB	Cytoplasm	Any positivity and percentage of positive cells
CK14 (LL002)	NCL-LL002	M	Novocastra	MACH1	Citrate	1/40	1 h	NB	Cytoplasm	Any positivity and percentage of positive cells
Ki67 (MIB-1)	M7240	M	Dako	MACH1	Citrate	1/200	1 h	NB	Nucleus	Positive in the nucleus in >1% of cells
p53 (D07)	M7001	M	Dako	MACH1	Citrate	1/200	1 h	NB	Nucleus	Positive in the nucleus in >10% of cells
Androgen receptor (MR441)	M3562	M	Dako	MACH1	Citrate	1/50	1 h	Prostate	Nucleus	Positive in the nucleus in >1% of cells

Continued next page.

Table 2.1 Immunohistochemistry antibodies and conditions

Antibody (Clone)	Cat #	Species	Source	Detection method	Antigen retrieval	Antibody dilution	Inc. time	Positive control	Cellular localisation	Scoring method
E-cadherin (HECD-1)*	HECD-1	M	Invitrogen	MACH1	Citrate	1/100	16 h	NB	Membrane	Localisation (membranous, cytoplasmic, negative)
E-cadherin (HECD-1)*	HECD-1	M	Invitrogen	Bond Polymer Refine	Bond ER2 (pH 9) 20 mins	1/100	15 min	NB	Membrane	Localisation (membranous, cytoplasmic, negative)
beta-catenin (17C2)	17C2	M	Novocastra	MACH1	Citrate	1/100	16 h	NB	Membrane	Localisation (membranous, cytoplasmic, negative)
p120-catenin (98/pp120)	98/pp120	M	BD transduction labs	MACH1	Citrate	1/200	1 h	NB	Membrane	Localisation (membranous, cytoplasmic, negative)
GATA3 (L50-823)	558686	M	BD transduction labs	Ventana Ultraview DAB	Ventana CC1 32 mins	1/500	32 min	NB	Nuclear	Positive in the nucleus in >1% of cells
FOXA1 (2F83)	ab40868	M	Abcam	MACH1	Citrate	1/100	1 h	NB	Nuclear	Positive in the nucleus in >1% of cells
N-WASP 30D10)	48485	R	Cell Signalling	MACH1	Citrate	1/50	16 h	Opt TMA	Cytoplasm	Any positivity and percentage of positive cells
ECT2 (n/a)	07-1364	R	Millipore	MACH1	Citrate	1/50	1 h	Opt TMA	Cytoplasm	Any positivity and percentage of positive cells

Continued next page.

Table 2.1 Immunohistochemistry antibodies and conditions

Antibody (Clone)	Cat #	Species	Source	Detection method	Antigen retrieval	Antibody dilution	Inc. time	Positive control	Cellular localisation	Scoring method
RacGAP1 (H-300)	sc-98617	R	Santa Cruz	MACH1	Citrate	1/100	1 h	Opt TMA	Cytoplasm	Any positivity and percentage of positive cells
N-cadherin (3B9)	X 18-0224	M	Invitrogen	MACH1	Citrate	1/150	1 h	Opt TMA	Membrane	Localisation (membranous, cytoplasmic, negative)
SNAIL (SN9H2)	4719	Rat	Cell Signalling	MACH1 – rat kit	Citrate	1/40	16 h	Kidney	Nucleus	Any positivity and percentage of positive cells
Vimentin (V9)	M0725	M	Dako	MACH1	Citrate	1/400	1 h	Opt TMA	Cytoplasm	Localisation (cytoplasmic, negative)

* IHC for this antibody was initially performed in house but was later performed at Pathology Queensland using the more reliable Ventana Autostainer. M: Mouse, R: Rabbit; Inc. time: Incubation time; 1 h: 1 hour at room temperature, 16 h:16 hours at 4°C; NB: Normal breast; Opt TMA: An optimisation TMA containing various tissues including breast, prostate, testis, ovary, pancreas, tonsil, thyroid, liver, skin and various breast tumour types; Citrate antigen retrieval conditions: pH 6.0, 125°C for 5 minutes. Chy: Chymotrypsin antigen retrieval, 0.5%, pH 7.8, 37°C for 10 minutes.

Table 2.2 Immunofluorescence antibodies and conditions

Primary antibody (Clone)	Cat #	Species	Source	Isotype	Fixation	Dilution	Inc. time	Secondary antibody	Dilution	Channel	Source	Cat #
ECT2 (n/a)	07-1364	R	Millipore	IgG	PFA	1/50	16 hrs	Goat anti-rabbit	1/400	594	Life Technologies	A11072
E-cadherin (4A2C7)	X18-0223	M	Zymed	IgG1	PFA	1/100	1 hr	Goat anti-mouse	1/400	594	Life Technologies	A21121
Collagen 4 (Coll4)	M0785	M	Dako	IgG1	PFA	1/50	1 hr	Goat anti-mouse	1/400	594	Life Technologies	A21121
CK14 (L0022)	NCL-L0022	M	Novocastra	IgG3	PFA	1/50	1 hr	Goat anti-mouse	1/400	594	Life Technologies	A21151
CK8/18 (5D3)	NCL-5D3	M	Novocastra	IgG1	PFA	1/100	1 hr	Goat anti-mouse	1/400	594	Life Technologies	A21121
E-cadherin (HECD-1)	HECD-1	M	Invitrogen	IgG1	Methanol	1/100	16 hrs	Goat anti-mouse	1/400	594	Life Technologies	A11072
N-WASP (30D10)	48485	R	Cell signalling	IgG	Methanol	1/50	16 hrs	Goat anti-rabbit	1/400	594	Life Technologies	A21121

M: Mouse, R: Rabbit. PFA: 4% paraformaldehyde. Inc. time: Incubation time; 1 hour at room temperature, 16 hours at 4°C.

Microscopy and Image analysis

Following IHC, all slides were assessed under a light microscope and if the staining was successful, the slides were scanned digitally using the Leica Aperio Scanscope XT or AT Turbo located at QIMRB. Scanned images are available to view online using the software program Spectrum. This program allows the tissue section to be reviewed and using the relevant TMA map, an image of each individual tissue core on a TMA section can be segmented and extracted as individual JPEG images. These images were analysed using a photo-viewing program. The extracted JPEG images of each individual TMA tissue core have been analysed and scored using Apple iPhoto software. In iPhoto, keywords can be assigned to individual images that are used to assign scores to each image. The images are then meta-tagged into folders of the same tag (score), and each folder can be exported and assigned a score. The scores and core coordinates (which are matched to patient data) are assembled into a Microsoft Excel. This was repeated for all antibodies to gain a comprehensive immunophenotype profile of each case.

2.6. Immunofluorescence staining of fresh frozen tissue

Fresh frozen tissue sections were cut (7 μM) with a cryostat at -28°C (Leica MICROM HM 550). The tissue was fixed on the slide after air-drying, with the optimised fixative (either 100% ice cold methanol; 3:2 100% acetone:methanol at -20°C; or, 4% paraformaldehyde in PBS – depending on the antibody used; **Table 2.2**) for 10 minutes. After air-drying, the slides were washed in TBST (0.05% Tween-20 in 1 X Tris buffered saline) and exogenous peroxidases inactivated in 3% hydrogen peroxide for 10 minutes. A blocking buffer of FBT (5% FBS, 1% BSA, 0.05% Tween-20, 10mM Tris (pH 7.5), 100mM $MgCl_2$) was used for 30 minutes to block non-specific proteins. The primary antibody was diluted in FBT at the concentration optimised for frozen tissue (see **Table 2.2**) and incubated for 1 hour at room temperature. After washing in TBST, the secondary antibody (also diluted in FBT (see **Table 2.2**)) and DAPI nuclear counter-stain (1/10000 dilution) were added to the tissue section and incubated for 30 minutes. Prolong Gold (Life Technologies, Cat # P36930) was used to mount the sections after washing and the coverslip sealed with nail polish. The sections were stored at 4°C in the dark until imaging with a fluorescent microscope (Zeiss Axio Imager M1). Negative controls (no primary antibody), were run in parallel with every sample.

2.7. DNA isolation and quantification

2.7.1. Laser Capture Micro-dissection (LCM)

Tissue samples were micro-dissected using the Leica LMD 7000 laser dissection microscope. Tissue sections were cut at 10 µM onto polyethylene naphthalate (PEN) membrane-coated slides (Leica, Cat # 11505158). Areas of interest were cut and collected into tubes containing 30 µL of tissue lysis buffer ATL (Qiagen). The tissue was digested overnight with 10 µL of proteinase K (Invitrogen, Cat # 25530049) at 55°C with shaking. DNA was extracted using Qiagen QIAamp DNA Micro kit (Qiagen, Cat # 56404) following the manufacturers instructions with the following modification - the DNA was eluted off the MinElute columns with 25 µL of 37°C molecular grade DNase and RNase free water that was incubated on the column for 30 minutes. This elution step was repeated to ensure as much DNA was eluted off the column as possible.

2.7.2. Macro-dissection using needle scraping

Tissue samples that had clear tumour margins and didn't require micro-dissection were separated from surrounding tissue under a stereomicroscope using a fine needle. A small drop of ATL buffer (Qiagen) was placed on the slide. The needle was used to scrape a small amount of buffer onto the tissue section, and the buffer allowed the tissue area of interest to be removed from the slide. The macro-dissected tissue was placed into a tube containing 180 µL of ATL buffer, with 20 µL proteinase K. The tissue was digested in a shaking incubator at 55°C for three days, with two 20 µL proteinase K additions each day to ensure complete tissue digestion. The DNA was extracted using the Qiagen QIAmp DNA Micro Kit as per the manufacturers instructions.

2.7.3. FFPE tissue cores

DNA was also extracted from cores of tissue taken using a 1 mm TMA biopsy needle. On average 4-6 cores were taken per block and these were finely minced in a sterile petrie dish and transferred to a 1.7 mL Eppendorf tube. The cores were deparaffinised in a series of xylene (3 x 5 minute treatments) and rehydrated in ethanol (100%, 90%, 70% for 2 minutes each). The samples were then treated with 1M sodium thiocyanate overnight in a 37°C oven with shaking. The sodium thiocyanate was removed, followed by two 1x

Phosphate Buffered Saline (PBS) washes for 10 minutes each. The tissue was then digested in 180 μL of ATL buffer (Qiagen) with 40 μL of Proteinase K and incubated at 55°C with shaking for 3 days, with spiking of fresh Proteinase K (20 μL) each day. DNA was extracted using the Qiagen DNeasy Blood and Tissue kit (Qiagen, Cat # 69506) according to the manufacturers instructions with the following modification: instead of using the AW2 buffer, we performed two 70% ethanol washes, as recommended by Agilent to remove impurities that interfere with the labelling protocol for array comparative genomic hybridisation (see **2.11.3**). The DNA was eluted off the Minispin columns twice with 50 μL of 37°C water following incubation on the column for 30 minutes.

2.8. Ethanol precipitation of low concentration samples

Ethanol precipitation was utilised to increase the concentration of samples that were too dilute for array CGH. Both DNA elutions were combined, measured, and the volume made up to 250 μL with nuclease free water. To precipitate the DNA, our laboratory's optimised protocol was as followed: 1/10 volume of 3M sodium acetate pH 5.2, 2.5 volumes of ice-cold 100% ethanol and 1 μL of glycogen (Invitrogen, Cat # 10814-010) was added to each tube, mixed, and incubated in an esky of ice at 4°C overnight. The next day, the tubes were centrifuged at 12,000 g for 30 minutes at 4°C. The pellet was washed twice with 70% ethanol, spinning each time at 12,000 g for 10 minutes. The pellet was air dried at room temperature and was dissolved in 12-25 μL of 37°C nuclease free water and incubated at 37°C for half an hour to help dissolve the pellet.

2.9. DNA quantification and quality control

The NanoDrop spectrophotometer (Thermo Scientific) quantifies DNA using visual light spectrophotometry. The NanoDrop provides the sample concentration, as well as both 260/280 and 260/230 reading ratios. These ratios provide an indication of sample purity, where a 260/280 ratio of 1.8 denotes pure DNA and a 260/230 ratio of between 2.0-2.2 is desired for good DNA purity. Since all nucleic acids absorb light at 260 nm, the Nanodrop tends to over-estimate sample concentration as it will detect all nucleic acids, including small fragments, single stranded DNA and contaminating RNA present. The Qubit fluorometer 2.0 (Life Technologies) more accurately quantitates DNA because the assay utilises fluorescent dyes that only emit a signal when bound to double-stranded DNA. The fluorometer generates a standard curve using standards that are provided to calculate the

sample and stock concentrations. Both the Qubit dsDNA BR (Broad Range, Cat # Q32850) and HS (High Sensitivity, Cat # Q32851) kits were used depending on the size of the sample input. The assays were carried out following the manufacturers instructions.

2.10. Illumina Infinium HD FFPE QC assay

DNA quality of samples selected for array comparative genomic hybridisation was assessed using the Illumina Infinium HD FFPE QC assay following the manufacturers protocols (Part # 15020981 Rev. C.). This qPCR based assay used SYBR Green 2x qPCR Master Mix (Life Technologies, Cat # 4309155) and the Illumina FFPE QC kit (Cat # WG-321-1001). All DNA samples were diluted in nuclease free water to 1 ng/μL and each sample was amplified in triplicate in a final reaction volume of 10 μL. The samples were analysed using the ABI StepOne Real-Time PCR machine using the Applied Biosystems StepOne software v2.1 (Applied Biosystems). The data was analysed as per manufacture's instructions. Briefly, any replicates whose C_q value diverged by more than half a cycle were removed from analysis. An average C_q for each sample was recalculated and for the standard template control provided in the FFPE QC kit. The average C_q value for the template control was subtracted from the average C_q of each sample to compute a delta C_q and all samples with a delta C_q value of less than 5 were deemed suitable for downstream applications.

2.11. Genomic analyses

2.11.1. Target enrichment DNA sequencing

Capture-based targeted sequencing was performed using the Agilent SureSelect Target Enrichment system for Illumina paired-end sequencing following the manufacturer's instructions (Manual part number G7530-90000, Version 1.7, July 2014. Agilent Technologies, Inc). The panel was designed by Dr Peter Simpson, and **Table 2.3** contains a list of the genes targeted and the regions covered. The genes were selected based on the most frequently mutated genes in breast cancer, including some important oncogenes and tumour suppressor genes that undergo amplification and deletion. Other genes were selected based on their role in hormone signalling or cell-cell adhesion. The library preparation and sequencing were performed at Queensland Centre for Medical Genomics

(QCMG; Institute of Molecular Biosciences). The samples were analysed using Agilent's SureCall software.

Table 2.3 Custom gene list for the Agilent SureSelect sequencing panel

Gene Symbol	Gene Name	Description	References
AKT1	AKT	Altered in 2% of breast cancer. Sometimes amplified.	1, 3
AR	Androgen receptor	Role in hormone signalling. Altered in 2% of breast cancers.	3
ARID1A	AT rich interactive domain 1A	Putative epigenetic tumour suppressor, altered in 3% of breast cancers.	2, 3, 9
BRCA1	Breast cancer 1, early onset	Germline predisposition.	1
BRCA2	Breast cancer 2, early onset	Germline predisposition.	1
CBFB	Core-binding factor, beta subunit	Altered in 4% of breast cancer, mutated and deleted	1, 3, 4
CCND1	Cyclin D1	Frequently amplified in breast cancer (16%).	3
CDH1	E-cadherin	Frequently mutated in breast cancer, especially in lobular breast cancer.	1, 3
CDK4	Cyclin-dependent kinase 4	Frequently amplified in breast cancer.	1, 3
CDK6	Cyclin-dependent kinase 6	Frequently amplified in breast cancer.	1, 3
CDKN1B	Cyclin-dependent kinase inhibitor 1B (p27)	Frequently altered in breast cancer.	1, 3, 9
CDKN2A	Cyclin-dependent kinase inhibitor 2A	Deleted in breast cancer. Some amplification.	1, 3
CDKN2B	Cyclin-dependent kinase inhibitor 2B (p15)	Deleted and sometimes amplified in breast cancer.	1, 3
CTNNA1	Alpha-catenin	Role in cell-cell adhesion.	
CTNNB1	Beta-catenin	Role in cell-cell adhesion.	
CTNND1	p120-catenin	Role in cell-cell adhesion.	
CTTN	Cortactin	Frequently amplified in breast cancer (14%).	3
ERBB2	Human epidermal growth factor receptor 2 (HER2)	Frequently amplified in breast cancer. Mutated in some ILC.	1, 3, 5
ERBB3	Human epidermal growth factor receptor 3 (HER3)		
ESR1	Oestrogen receptor alpha	Role in hormone signalling. Mutated in some ER+ metastases. Amplified in 4% of primary breast tumours.	3, 6
ESR2	Oestrogen receptor beta	Role in hormone signalling.	
ESRRA	Oestrogen related receptor alpha	Role in hormone signalling.	
ESRRB	Oestrogen related receptor beta	Role in hormone signalling.	

Continued next page

Table 2.3 Custom gene list for the Agilent SureSelect sequencing panel

Gene Symbol	Gene Name	Description	References
ESRRG	Oestrogen related receptor gamma	Role in hormone signalling. Frequently amplified (14%)	3
FANCA	Fanconi anaemia, complementation group A	Germline predisposition, deleted in approx. 4% of breast cancer.	3, 7
FGFR1	Fibroblast growth factor receptor 1	Frequently amplified in breast cancer (14%).	1, 3
FGFR2	Fibroblast growth factor receptor 2	Altered in 3% breast cancers. Amplified or mutated.	1, 3
FOXA1	Forkhead box A1	Frequently mutated in breast cancer, focal amplifications occur.	1, 3
GATA3	GATA binding protein 3	Mutated in >10% of all breast cancers, enriched in Luminal A subtype.	1, 3
JUP	Gamma-catenin	Role in cell-cell adhesion.	
KRAS	KRAS	Germline predisposition, amplified in some triple negative breast cancers.	1, 3, 8, 9
MAP2K4	Mitogen-activated protein kinase 4	Focal deletions and mutations.	1, 3, 9
MAP3K1	Mitogen-activated protein Kinase kinase kinase 1	Mutations enriched in Luminal A subtype.	1, 3, 9
MDM2	MDM2 oncogene, E3 Ubiquitin protein ligase	Frequently amplified in breast cancer.	1, 3
MLL3	Lysine specific methyltransferase 2C	Frequently mutated in breast cancer. Focal deletions occur.	1, 3, 9
NCOR1	Nuclear receptor corepressor 1	Altered in 5% breast cancer.	3, 9
NF1	Neurofibromin 1	Germline predisposition. Altered in 6% of breast cancer.	3, 9, 10
NR3C1	Glucocorticoid receptor	Role in hormone signalling.	
PGR	Progesterone receptor	Role in hormone signalling. Amplified or deleted in 3% of breast cancer.	3
PIK3CA	Phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit alpha	Mutated in >10% of all breast cancers. Enriched in Luminal A subtype. Focal amplifications occur.	1, 3
PTEN	Phosphatase and tensin homolog	Frequently altered in breast cancer. Germline predisposition. Focal deletions.	1, 3, 11
RB1	Retinoblastoma 1	Frequently deleted in breast cancer.	1, 3
RUNX1	Runt-related transcription factor 1	Frequently mutated in breast cancer.	1, 4
TBX3	T-box 3	Frequently mutated in breast cancer.	1, 9
TP53	p53	Mutated in >10% of all breast cancers Germline predisposition.	1, 3

[1](The Cancer Genome Atlas Network, 2012) [2](Wu and Roberts, 2013) [3](Cerami et al., 2012) and (Gao et al., 2013) [4](Banerji et al., 2012) [5](Ross et al., 2013) [6](Robinson et al., 2013a) [7](Seal et al.,

2003) [8](Paranjape et al., 2011) [9](Stephens et al., 2012) [10](Sharif et al., 2007) [11](Lynch et al., 1997). Mutation, amplification and deletion percentages were derived from the cBioportal (http://www.cbioportal.org).

2.11.2. Whole exome sequencing and data analysis

Whole exome sequencing was performed on a select number of cases of Mixed Ductal-Lobular Carcinomas and cases with gynaecological metastases that had DNA from the primary tumour and multiple metastatic sites. The exome sequencing was performed by Macrogen using the Agilent SureSelect Exome Capture kit v4 and QCMG using the Illumina Nextera Rapid Capture Exome Kit (9-plex) using 200 ng of DNA and following the manufacturers instructions. Exome sequencing was performed using the Illumina 2500 HiSEQ platform at QCMG. QCMG has automated pipelines that channel the data from the sequencer through to computational infrastructure that checks data integrity, prepares raw sequencing data for mapping and alignment, and archives and stores data. Dr Katia Nones analysed the exome data and somatic variants were called as per (Bessette et al., 2015), where the sequencing data was aligned to the reference human genome hg19 using BWA (Burrows-Wheeler alignment tool (Li et al., 2009)). Somatic variants were identified using both qSNP (Kassahn et al., 2013) and the Genome Analysis Tool Kit (GATK (McKenna et al., 2010)). The functional effects of any identified variants were determined using several online tools. To determine the pathogenicity of the mutations on the protein product, Provean (Protein Variation Effect Analyzer (Choi et al., 2012)) and PolyPhen-2 (Polymorphism Phenotyping v2 (Adzhubei et al., 2010)) were used. The online software program IntOGen (Integrative Onco Genomics (Gundem et al., 2010)) was used to determine if mutations were found in cancer driver genes and assess the functional contribution of mutations in biological pathways. Based on data generated from thousands of cancer genomes, the IntOGen pipeline uses multiple tools, including OncodriveFM (Lawrence et al., 2013), and OncodriveCLUST (Tamborero et al., 2013) to identify genes whose mutations are selected for during tumour development that are likely drivers. OncodriveFM identifies genes which accumulate mutations with a high functional impact (FM bias) and OncodriveCLUST identifies genes whose mutations cluster in particular regions of the protein sequence (CLUST bias). A functional impact score is assigned to each gene (None, Low, Medium or High) and the software determines the frequency of the mutation in each gene and/or pathway within the project.

2.11.3. Array based Comparative Genomic Hybridisation (CGH)

DNA was sent to the Ramaciotti Centre for Genomics for array Comparative Genomic Hybridisation (CGH) to measure DNA copy number changes on the Agilent SurePrint G3 Human CGH Microarray in either the 60K and 180K formats. Two array densities were used to accommodate variations in DNA concentrations obtained from the FFPE tissue samples, allowing us to include the low concentration samples on the 60K chip. A total of 68 samples were suitable for the 180K array (at a minimum of 500 ng total in a concentration of 62.5 ng/μL) and 16 samples for the 60K array (250 ng at 32 ng/μL) from 27 cases. The samples were labelled with Cy5 (and the female reference DNA labelled with Cy3) following the ULS protocol (Manual part number G4410-90020, Version 3.4, July 2012. Agilent Technologies, Inc). The ULS labelling system is a non-enzymatic assay that directly labels DNA with the fluorescent dye, and is recommended for use with DNA derived from FFPE samples. The samples were hybridised on the array for 40 hours before microarray washing and scanning.

Preliminary analysis of the data was performed by assessing the quality control documents generated for each sample. The raw data was imported into Agilent's Cytogenomics software package version 3.0 to generate copy number profiles for each sample.

Frequency plots were also generated with the help of a bioinformatics PhD student in our lab (Mr Samir Lal). The frequency plots were generated using the R programming software as follows. Background and median normalisation of the arrays was performed using the *snapCGH* package (Smith et al., 2006). Tumour sample copy number log2 ratios were segmented using circular binary segmentation (CBS) and then smoothed using the *DNAcopy* package (Olshen et al., 2004). The median of the log2 ratios + 1σ or + 4σ was computed using 50% of the central probes for gains and amplifications, respectively. For losses, the median of the log2 ratios -1.7σ and -6.8 σ was used to call heterozygous deletions and homozygous deletions, respectively. Frequency of gains, amplifications and deletions were achieved using the *copy number* package (Nilsen et al., 2012). GISTIC2.0 (Genomic Identification of Significant Targets in Cancer) was performed to identify regions from CBS smoothed segments that are significantly amplified and deleted (Mermel et al., 2011). Each region is assigned a G-score that considers the amplitude and frequency of the alteration across samples. The false discovery rate q-value is then calculated for these aberrant regions and a cut off of less than 0.05 was used. For each of the significant

regions a “peak region” is identified. These significant peak regions are comprised of the aberrant regions with the greatest amplitude and frequency of alteration. “Wide peak” regions are determined to allow for errors in the boundaries of significant peaks in a single sample. Genes in the “wide peak” boundaries are reported.
Using the human reference genome build (hg18) from the UCSC genome browser, genes were selected if they fell completely within CBS-identified copy number segment from samples using GenomicRanges package (Lawrence et al., 2013). Genes that were not found completely within a copy number segment were filtered out. A gene by sample matrix of copy number calls was generated using the plyr package (Wickham, 2011). Annotation information for each gene was obtained using the biomaRt package (Durinck et al., 2009).

2.12. Meta-analysis of publically available genomic data

The Cancer Genome Atlas (TGCA) is a coordinated project aiming to catalogue all genomic alterations among specific cancer types. The data is freely accessible to the public online providing all of the molecular data (somatic mutation, copy number alteration, gene expression, methylation, reverse phase protein arrays) that has been done on every specimen to date (as of July 2015, 21441 tumour samples from 91 cancer studies).
The genomic data is available for download through the TCGA Data Portal (National Cancer Institute; https://tcga-data.nci.nih.gov/tcga/tcgaDownload.jsp) or the cBioportal (Cerami et al., 2012, Gao et al., 2013). An analysis of gene expression patterns across different breast cancer subtypes from TCGA was performed for genes involved in actin cytoskeleton regulation (**Chapter 3**). Box and whisker plots (where the whiskers represent the 10^{th} – 90^{th} percentile) were produced using the RNASeq z-scores of expression levels for each transcript using Graph Pad Prism version 6 to compare gene expression levels between the different subtypes of breast cancer: morphological type combined with biomarker expression (*i.e.* IC-NST that are ER- HER2+, or ER+ HER2-, or ER+ HER2+, or triple negative, and lobular) and molecular subtypes (*i.e.* Luminal A, Luminal B, HER2, basal and normal-like). Ordinary one-way ANOVA tests were performed to assess if there were any significant differences in expression between the different subtypes.

All mutation data was downloaded from the TCGA for 29 Mixed Ductal Lobular carcinomas that underwent sequencing as part of the TCGA breast cancer sequencing project (**Chapter 4**). Digital copies of the tumour diagnostic slides for the TCGA samples were

accessed from the Cancer Digital Slide Archive (http://cancer.digitalslidearchive.net) to verify the diagnosis of specific tumour types.

2.13. Cells and cell culture

The human invasive lobular carcinoma cell line IPH-926 was obtained via collaboration with Matthias Christgen from the Hannover Medical School, Germany (Christgen et al., 2009). MDA-MB-134VI and MCF7 human breast cancer cell lines were obtained from the American Type Culture Collection (ATCC). IPH-926 was grown as a monolayer in RPM1-1690 media (Life Technologies, Cat # 22400089) supplemented with 20% foetal bovine serum (FBS, Gibco/Invitrogen, Cat # 10099141) and antibiotic/antimycotic (Ab/Am; 1x, Gibco/Invitrogen, Cat # 15240062), insulin (1 mg/mL, Sigma, Cat # I5500), HEPES (Gibco, Cat # 15630), glutamine (200mM, Gibco, Cat # BS003583) and sodium pyruvate (100 mM, Invitrogen, Cat # 11360). MDA-MB-134VI was grown in DMEM media supplemented with 20% FBS, glutamine (200 mM) and Ab/Am (1x). MCF7 was grown in DMEM media with 10% FBS, glutamine (200 mM), Ab/Am (1x) and insulin (1 mg/mL). All cell lines were grown in a humidified atmosphere containing 5% CO_2 at 37°C. Once the cells reached confluency, the cells were removed from the T75 flask using trypsin (TrypLE, Life Technologies, Cat # 12605010) and the cells were counted using the Countess (Life Technologies).

Fixing cells onto coverslips

IPH-926 and MDA-MB-134VI were grown on coverslips for IF analysis. Autoclaved glass coverslips were placed inside each well of a 24 well plate, upon which 20,000 to 40,000 cells were plated. Once the cells reached approximately 70-80% confluency, they were fixed using 100% ice-cold methanol for 5 minutes and stored in PBS buffer until use.

Immunofluorescent staining

Immunofluorescent staining on cells on coverslips was performed as described in **Section 2.6** with some modifications. After the cells were fixed, the coverslips were removed from the plate and IF performed on a sheet of Parafilm, whereby each reagent was spotted to form a bubble that the coverslip is then moved between. The cells were washed with PBST (0.05% TWEEN-20 in 1x Phosphate Buffered Saline) and permeabilsed using 0.1%

Triton-X in 1X PBS. The primary and secondary antibodies were diluted in FBT, the blocking agent, and incubated for the time specified in **Table 2.2**. DAPI was used to counter-stain the nuclei and was diluted 1/10000 in FBT and incubated for 5 minutes. The coverslips were mounted onto glass slides using 3 μL of Prolong Gold mounting medium. The IF was visualised using a fluorescent microscope (Zeiss Axio Imager M1).

Chapter 3

INVESTIGATING THE ROLE OF THE ACTIN CYTOSKELETON REGULATORY MOLECULES IN THE DISRUPTION OF E-CADHERIN IN BREAST CANCER

3. Investigating the role of actin cytoskeleton regulatory molecules in the disruption of E-cadherin in breast cancer

3.1. The actin cytoskeleton

Within the cytoplasm of every cell, a skeleton made up of numerous proteins forms a complex network that allows the cell to hold and change its shape, enables cellular migration and determines cellular polarity in response to extracellular stimuli. The cytoskeleton is made up of three major protein families that form three main types of filaments; (i) intermediate filaments provide mechanical strength for the cell; (ii) microtubules are responsible for directing intracellular transport and the positions of membrane-enclosed organelles; and, (iii) actin filaments are essential for determining cell shape and locomotion (**Figure 3.1**).

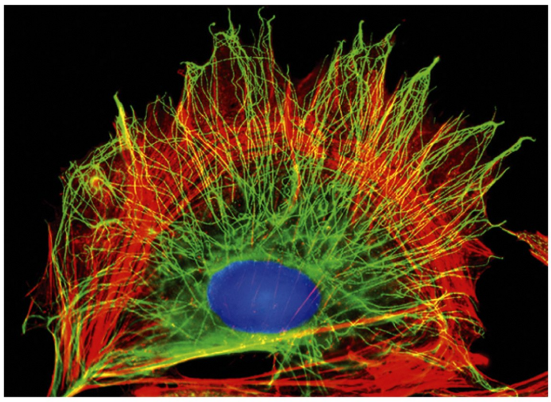


Figure 3.1: The cytoskeleton is made up of many molecules including the microtubules (stained green) and actin filaments (stained red) (Alberts et al., 2008); with permission from Garland Science.

The cytoskeleton is dynamically regulated by many accessory proteins and is continually reorganised to adapt to the cell's changing circumstances. The actin cytoskeleton in particular is a dynamic structure that undergoes continual assembly of monomeric G-actin molecules that form a polymer called filamentous (F)-actin, followed by disassembly and remodelling. Many molecules are involved in maintaining the actin filament network, including aspects of assembly, turnover and motility- for example, the Arp2/3 complex that stimulates actin assembly, and the myosin molecular motor proteins that attach to actin filaments and "walk" along them, generating contractile forces. The Rho families of small GTP-binding proteins control the regulation of actin structures (*i.e.* Rho itself for contractile acto-myosin filaments ("stress fibres"), Rac for lamellipodia and Cdc42 for filopodia (Nobes

and Hall, 1995, Hall, 1998)). In non-muscle cells, actin filaments are formed proximal to membrane surfaces (Alberts et al., 2008).

3.1.1. E-cadherin and the actin cytoskeleton

Cadherin junctions and the actin cytoskeleton are interdependent and their relationship is highly dynamic. Rac and Rho signalling are vital for this interaction (Braga et al., 1997). It has been observed *in vitro* that disruption of the actin cytoskeleton either by drugs, mutation of key actin regulators, or depletion by RNAi, that cadherin interactions are adversely effected (Angres et al., 1996, Baum and Perrimon, 2001, Mangold et al., 2011). The homophilic interactions of cadherin molecules between cells has been shown to reorganise the actin cytoskeleton and can activate various signal transduction pathways that recruit key regulators of the actin cytoskeleton to the adhesion junction (Gloushankova et al. 1997, Kovacs et al. 2002b, Vasioukhin et al. 2000).

The physical interaction of E-cadherin with the actin cytoskeleton is also dynamic and α-catenin plays a significant role in this interaction (**Figure 3.2**). This is via direct physical binding of the C-terminus of α-catenin with F-actin, or indirectly through other signalling macromolecules such as the centralspindlin complex (see below). E-cadherin concentrates in a continuous ring at the apico-lateral interface between neighbouring cells. This structure is an adherens junction known as the zonula adherens. The zonula adherens are also a site for rings of actin filaments, and coupled with E-cadherin binding, this complex of actin filaments is essential for zonula adherens integrity (Harris and Tepass, 2010).

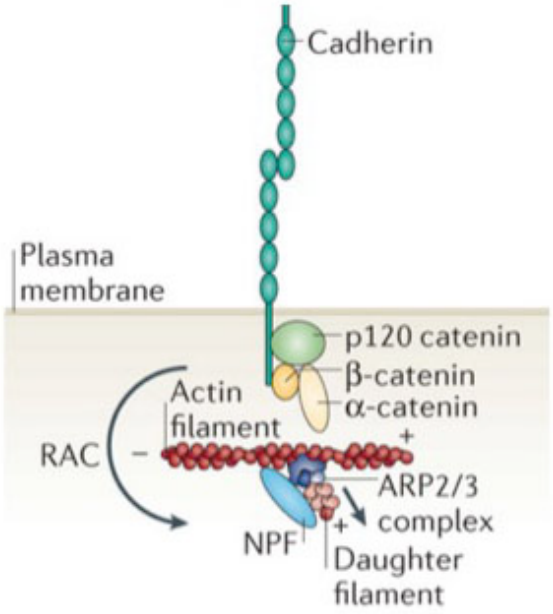


Figure 3.2: The adhesion junction. The catenin molecules (α, β, p120) are essential for E-cadherin stability at the cell membrane while simultaneously providing a scaffold for actin assembly ((Ratheesh and Yap, 2012), with permission from Nature Publishing Group).

3.1.2. E-cadherin, the actin cytoskeleton and breast cancer

A hallmark feature of a cancer cell is its ability to invade locally and systemically. In order for a cell to become invasive, it must first detach from its neighbouring cells and rearrange its cytoskeleton to produce contractile forces in order to move. Invasive lobular carcinomas are an archetypal example of a tumour cell that invades in a single file of discohesive cells throughout the breast parenchyma. The classical immunophenotype of invasive lobular carcinoma (ILC) is the loss of E-cadherin expression, alongside loss of α, β and γ-catenin and the relocalisation of p120 catenin to the cytoplasm. It was demonstrated that this relocalisation of p120 catenin was capable of activating Rho/Rock signalling which activates downstream actin cytoskeleton remodelling pathways (Schackmann et al., 2011). Of particular importance was the induction of anoikis resistance pathways, allowing the cells to survive independent of cell-adhesion by interacting with regulators of the actin cytoskeleton (Schackmann et al., 2011). This model is in the context of an E-cadherin (*CDH1*) mutation. However, many breast cancers show loss of E-cadherin in the absence of a *CDH1* mutation, such as high-grade IC-NST, which may show a transition from an epithelial to a mesenchymal phenotype (Aleskandarany et al., 2014). Aberrant E-cadherin localisation, rather than loss, has also been reported for some tumour types, including ILC (Da Silva et al., 2008) and the lobular component of mixed ductal lobular carcinomas (MDL), however the mechanisms underlying this phenomenon have not been investigated to date (See **Chapter 4**). It is plausible that there may be alterations in molecules involved in actin cytoskeleton regulation that in turn destabilise the adhesion junction in breast cancer cells and contribute to the invasive phenotype frequently observed in ILC, the lobular component of MDLs and high-grade IC-NST.

Regulators of the actin cytoskeleton involved in adherens junction integrity

Our collaborators have a strong interest in the role of the actin cytoskeleton interaction with adhesion junctions. Recent publications have highlighted two important pathways that are closely linked to E-cadherin function, including the molecules N-WASP and the Centralspindlin complex.

Model 1: N-WASP

N-WASP (Neural Wiskott-Aldrich syndrome protein – encoded by *WASL*) is a major regulatory protein of the actin cytoskeleton, and functions by stimulating the Arp2/3 complex that is responsible for actin polymerisation (Rohatgi et al., 1999).

N-WASP has been identified to localise at the zonula adhesions (and not at the lateral adhesion junctions), an active site for actin polymerisation (**Figure 3.3**). A non-canonical role for N-WASP has been found where it stabilises the cytoskeleton (after Arp2/3 complex actin nucleation) in order to maintain the integrity of the adherens junction via a pathway involving the protein WIRE. When N-WASP is removed, junctional F-actin is decreased and becomes disorganised, however, this loss did not affect actin nucleation at the adherens junction. The actin stabilising function of N-WASP was dependent upon E-cadherin adhesion (Kovacs et al., 2011).

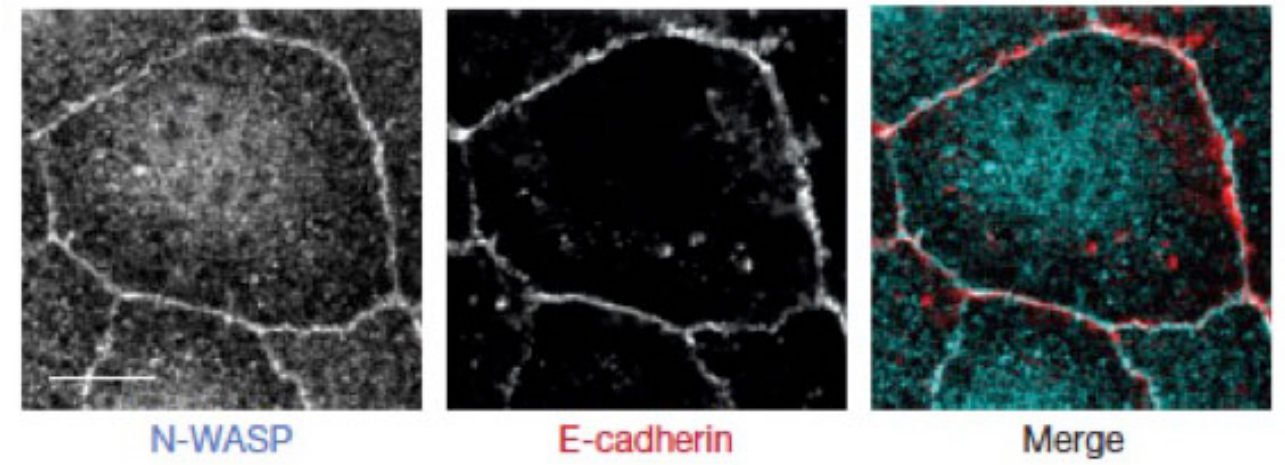


Figure 3.3: Co-immunofluorescent staining of N-WASP and E-cadherin identified co-localisation of N-WASP at the zonula adherens in Caco-2 colorectal carcinoma cells, suggesting that N-WASP plays an important role in stabilising the adherens junction ((Kovacs et al., 2011), with permission from Nature Publishing Group).

Oncogenic cell extrusion

In order to regulate the size and space limitations of the epithelia, cells under go apoptosis to control cell number and prevent overcrowding, and apoptosis is followed by cell extrusion where the cell is ejected from the epithelial layer (Andrade and Rosenblatt, 2011). Live cell extrusion has also been observed, and this is also followed by apoptosis and proposed to be a tumour-suppressive process (Eisenhoffer et al., 2012). However, if the cell is transformed by an oncogene (such as *K-Ras*, *H-Ras*, *Src* and *ERBB2*) and

surrounded by non-transformed cells, the transformed cell is extruded and is capable of surviving (termed anoikis resistance, (McFall et al., 2001, Thullberg et al., 2007)), and proliferating and initiating the outgrowth of a tumour (Hogan et al., 2009, Leung and Brugge, 2012).

Extrusion is a mechanically active process that involves both the transformed cell and its neighbours. E-cadherin and actomyosin contractility are important players in extrusion. A recent study showed that cleavage of the extracellular domain of E-cadherin is a mechanism driving cell extrusion (Grieve and Rabouille, 2014). N-WASP has also been implicated in extrusion. Extruding cells have been shown to harbour increased cortical F-actin intensity compared to neighbouring cells (Grieve and Rabouille, 2014, Wu et al., 2014). In H-Ras^{V12} transformed cells cultured among wild type cells, N-WASP was found to redistribute to areas of contact (lateral adherens) and was depleted from the zonula adherens (Wu et al., 2014). This redistribution resulted in altered junctional tension at the zonula adherens, yet increased tension at the lateral adherens. The extruding cell has more F-actin and myosin II, demonstrating N-WASP's functional role in actin stabilisation. This was reversed and extrusion inhibited when N-WASP was knocked down (Wu et al., 2014) (**Figure 3.4**). Therefore, N-WASP is necessary for extrusion to occur and may play an important role in cancer initiation or invasion.

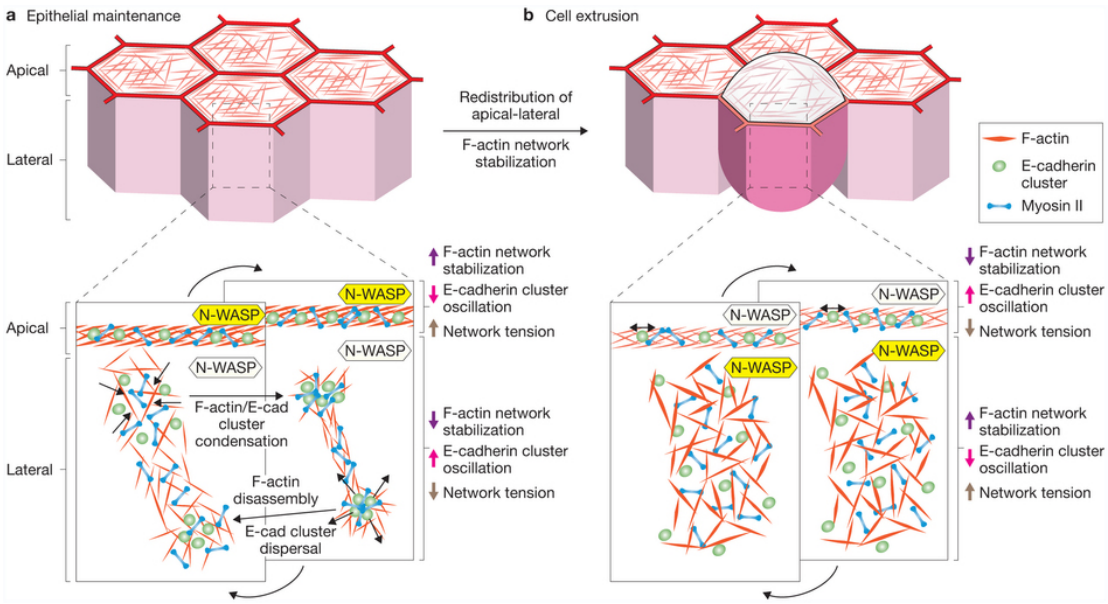


Figure 3.4: Model of epithelial cell maintenance and extrusion. The redistribution of N-WASP from the apical adherens junction to the lateral adherens junction facilitates oncogenic cell extrusion ((Behrndt and Heisenberg, 2014), with permission from Nature Publishing Group).

Model 2: The centralspindlin complex and ECT2

The centralspindlin complex and ECT2 are well-characterised molecules involved in mitosis. The centralspindlin complex is a heterodimer of RacGAP1 and MKLP1, its role in the mitotic phase of cytokinesis has been well established, where Rho accumulates at the contractile furrow and regulates the actomyosin-based processes necessary for cell division (Yoshizaki et al., 2003, Bement et al., 2006, Yuce et al., 2005). ECT2 (Epithelial cell transforming 2) is a guanine nucleotide exchange factor (which catalyses the exchange of GDP to GTP) for the Rho-family of GTPases. ECT2 is also involved in the regulation of cytokinesis. Phosphorylated ECT2 catalyses the guanine nucleotide exchange of the small GTPases, RhoA, Rac1, and Cdc42 during the G2 and M phase of mitosis. ECT2 is located in the nucleus during interphase, relocalises to the cytoplasm during prometaphase, and is condensed in the midbody during cytokinesis (Tatsumoto et al., 1999, Matthews et al., 2012).

A recent study identified both the centralspindlin complex and ECT2 at cell-cell contacts (zonula adherens) in interphase MCF7 breast cancer cells. Knockdown of ECT2 reduced junctional localisation of Rho and Rho-GTP that is necessary to support the integrity of the zonula adherens via recruitment of myosin IIA, and this was rescued when ECT2 was reintroduced. RacGAP1 was found to co-immunoprecipitate with E-cadherin and α-catenin, and knockdown of this molecule significantly reduced the amount of ECT2 that co-immunoprecipitated with α-catenin revealing that the centralspindlin complex serves as an intermediate that activates ECT2 at the zonula adherens and promotes Rho signalling to support the adherens junction. The centralspindlin complex was also found to simultaneously prevent Rho inhibition by blocking the recruitment of p190B RhoGAP (**Figure 3.5**) (Ratheesh et al., 2012).

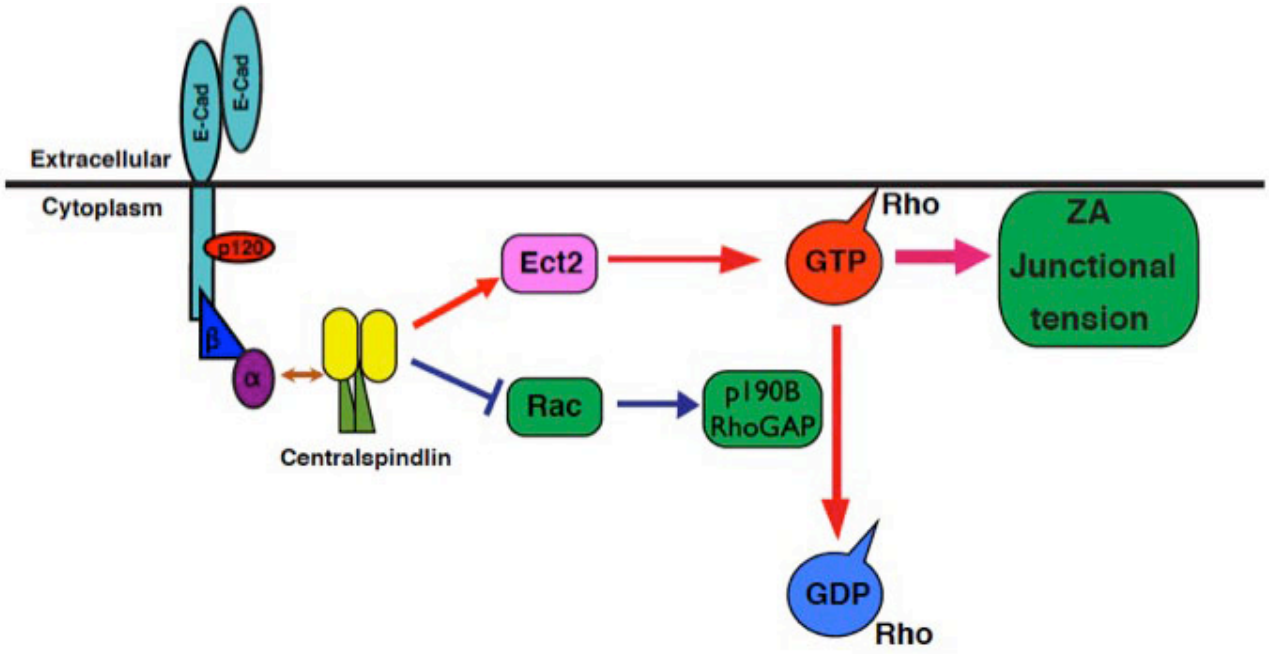


Figure 3.5: The Centralspindlin complex (a heterodimer of RacGAP1 and MKLP1) and ECT2 have been identified at the adherens junction. Both the Centralspindlin complex and ECT2 are required to maintain junctional tension (Image from (Ratheesh et al., 2012), with permission from Nature Publishing Group).

3.2. Hypotheses and aims

The actin cytoskeleton plays an integral role in maintaining adhesion junctions and we hypothesise that disruption of the pathways regulating the actin cytoskeleton may contribute to loss of cell adhesion and an increase in the invasive phenotype that is observed in some breast cancer subtypes. In particular, invasive lobular carcinomas have E-cadherin deregulation in approximately 90% of cases. Mutations in E-cadherin's gene, *CDH1*, account for some of, but not all, of the mechanisms behind E-cadherin deregulation (*CDH1* is mutated in 57% of ILC (McCart Reed and Kutasovic et al., 2015)). Aberrations of molecules that regulate the actin cytoskeleton could therefore be candidates for those cases with aberrant or lost E-cadherin expression without *CDH1* mutation.

This chapter investigates the role of the actin cytoskeleton across many different breast cancer types using a meta-analysis of publically available gene expression data, somatic mutation data, and protein analysis by IHC using breast cancer tissue microarrays. Through collaboration with Prof Alpha Yap we investigated genes that are involved in the cell adhesion-actin cytoskeleton axis to determine whether deregulation of these could contribute to the invasive phenotype frequently observed in breast cancer.

3.3. Results

3.3.1. Analysis of gene expression changes in regulators of the actin cytoskeleton

As there are many different actin cytoskeleton regulatory pathways, and since its role in cancer has only recently come to light, it is very difficult to know which molecules in particular to focus on; equally, there may be other important pathways that have not yet been discovered. Therefore we have undertaken a meta-analysis of publically available whole transcriptome gene expression profiling data, from more than 900 human breast cancer samples (**Section 2.12**). This allows us to investigate a large sample size to potentially see subtle effects of the role of the actin cytoskeleton regulators.

We have selected specific genes involved in actin cytoskeleton regulation (*e.g. Arp2/3, N-WASP*), myosin motor proteins (*e.g.* Myosin II, Myosin IV) and actin signalling molecules (*e.g. Rho, Rac, ECT2*) (**see Appendix Table 3.1**) that are of interest in the Yap laboratory (and for which they could provide antibodies) and additional genes involved in these pathways. An analysis of gene expression patterns was performed across different breast cancer subtypes from The Cancer Genome Atlas (TCGA) using the cBioportal interface (**Section 2.12**; (Cerami et al., 2012) (Gao et al., 2013)).

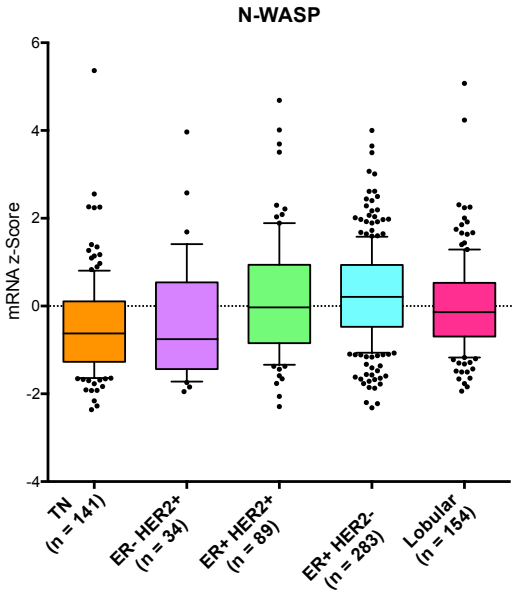
Box plots (**Figure 3.6**) for each gene were produced using GraphPad Prism 6 to compare gene expression levels between the different subtypes of breast cancer: histological type combined with biomarker expression (*i.e.* IC-NST that are ER- HER2+; ER+ HER2-; ER+ HER2+; triple negative (TN); and lobular) and molecular intrinsic subtypes (*i.e.* luminal A; luminal B; HER2; basal and normal-like).

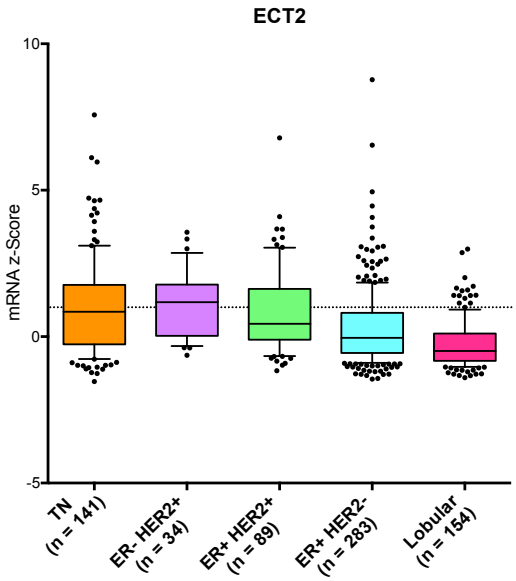
Many genes were differentially expressed between breast cancer subtypes. Through our collaboration with Prof Alpha Yap, we focused on the differential expression of *N-WASP, ECT2,* and *RacGAP1* identified between the histological and intrinsic breast cancer subtypes (**Figure 3.6**).

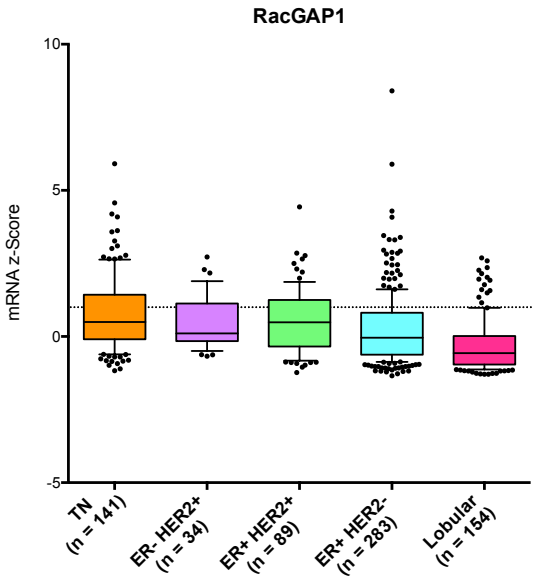
N-WASP was found to be slightly up-regulated in the luminal A intrinsic subtype and both the lobular and ER+ (irrespective of HER2 status) histological subtypes. *N-WASP* was slightly down-regulated in the ER- HER2- and triple negative histological subtypes and the

basal, HER2 and normal-like intrinsic subtypes. *ECT2* and *RacGAP1* have similar changes in gene expression within each subtype and were found to be slightly down-regulated in both the lobular and ER+ HER2- negative histological subtype, and luminal A and normal-like intrinsic subtypes. *ECT2* was slightly up-regulated in the HER2+ and triple negative histological subtypes and basal, HER2 and luminal B intrinsic subtypes, whereas *RacGAP1* was found to be up-regulated in the triple negative and ER+ HER2+ histological subtypes and luminal B intrinsic subtype (**Figure 3.6**). There are some cases however with much higher (or lower) expression compared to the average of the cohort, suggesting that a change in the expression of these genes may be important in some individual tumours. These findings suggest that cellular localisation of the protein may be more important than gene expression.

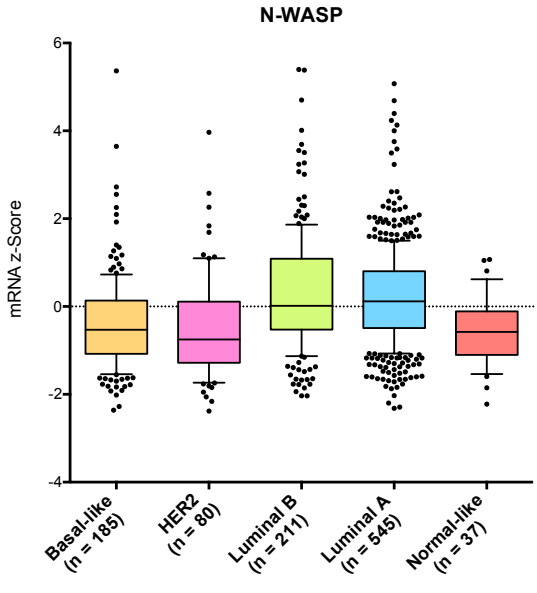
Histological subtypes

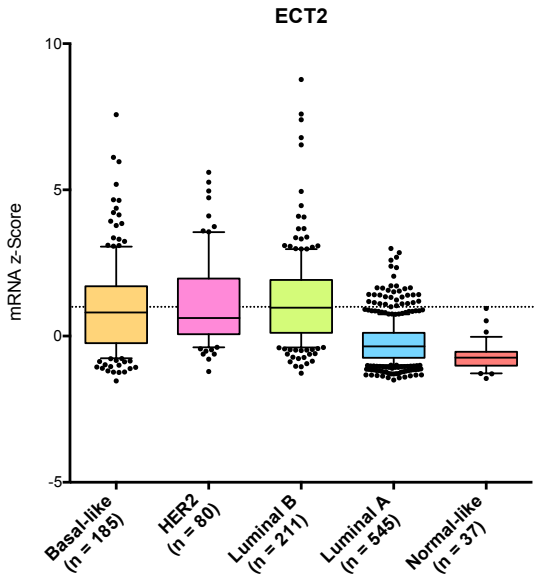






Instrinsic subtypes





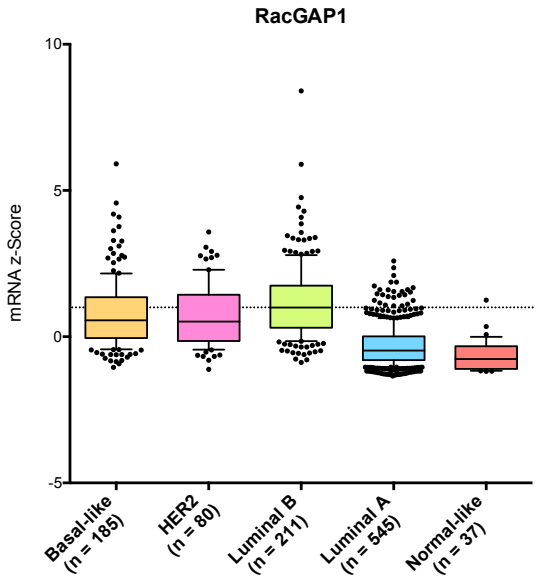


Figure 3.6: Gene expression analysis of (a) *N-WASP* (a.k.a *WASL)*, (b) *ECT2* and (c) *RacGAP1* from the TCGA across breast cancer histological and intrinsic subtypes. An RNASeq z-score is defined as a value indicating the number of standard deviations away from the mean of expression in the reference population (reference population being either all tumours that are diploid for the gene in question, or, when available, normal adjacent tissue). The RNASeq z-scores for each gene were plotted in GraphPad Prism 6; the whiskers represent the 10th – 90th percentiles.

3.3.2. Are the genes encoding N-WASP, ECT2 and RacGAP1 ever altered in breast cancer?

The cBioportal was employed to assess whether the genes of these actin cytoskeleton regulators were mutated or had copy number alterations in breast cancer (**Section 2.12**). In 962 breast cancers with both sequencing and copy number alteration data, the genes coding for N-WASP (*WASL* - located on human chromosome 7q31), *RacGAP1* (located on human chromosome 12q13) and *ECT2* (located on human chromosome 3q26) were rarely mutated or amplified (**Figure 3.7, Table 3.1**). *ECT2* was most frequently altered compared to the other genes (amplified in 4.15% of breast cancer cases). There was no correlation between genomic alteration in any of these genes and histological type. Alterations in *N-WASP*, *ECT2* and *RacGAP1* were mostly mutually exclusive and were present in cases without *CDH1* mutation or alteration (**Figure 3.7**). Across all cancer types N-WASP was mutated in 22% of pancreatic cancers, ECT2 was amplified in over 50% of lung carcinomas and over 30% of ovarian carcinomas, and RacGAP1 was lost in 15% of adenoid cystic carcinomas (**Appendix Figures 3.2, 3.3, 3.4**).

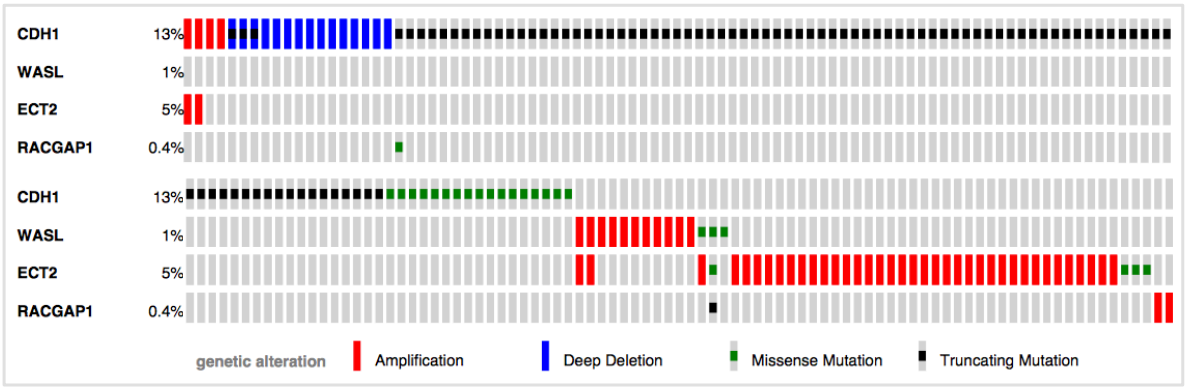


Figure 3.7: Oncoprint of *CDH1*, *WASL* (N-WASP), *ECT2* and *RacGAP1* (n = 178). Visual summary of alterations in *WASL* (N-WASP), *ECT2* and *RacGAP1* in 57/962 (6%) of breast cancer samples, as presented by the cBioportal. Each box represents an individual sample and only cases with alterations are included in the image (top panel is the first 89 patients,

bottom panel is the remaining 89 patients). Alterations in these genes were largely mutually exclusive and mostly occur in patients with wild-type *CDH1*.

Table 3.1: Summary of genomic alterations for N-WASP (*WASL*), *ECT2* and *RacGAP1* in breast cancer from the TCGA.

Gene	**Amplified n (%)**	**Mutated n (%)**
WASL	11 (1.14)	3 (0.3)
ECT2	40 (4.15)	4 (0.4)
RacGAP1	2 (0.2)	2 (0.2)

Overall, N-WASP (*WASL*), *ECT2* and *RacGAP1* are rarely mutated or have a change in copy number in breast cancer, and therefore alternative methods are likely to contribute to their changes in gene expression.

3.3.3. Protein analysis of regulators of the actin cytoskeleton

As the meta-analysis showed few genomic alterations in the genes encoding *N-WASP*, *ECT2* and *RacGAP1*, we proposed that a change in phenotype might manifest at the protein level. Therefore, the protein levels were assessed by the use of immunohistochemistry on specialised tissue microarrays (TMA) (**Sections 2.1 and 2.5.4**).

E-cadherin, N-WASP, ECT2 and RacGAP1 were investigated on four existing collections of TMAs spanning a variety of breast tumour types (ILC, HER2 positive and triple negative IC-NST, and a progression series of pre-invasive lesions with matched invasive tumours (**Section 2.1**)). The following data describes the expression of N-WASP, ECT2 and RacGAP1 in normal breast tissue and in breast cancer, followed by the correlation of expression of these markers with E-cadherin protein expression.

Each protein was scored for localisation (membranous, cytoplasmic or nuclear), the intensity of protein expression (1+, 2+, 3+) and the percentage of positive cells. Only 2+ and 3+ staining intensity was considered positive, since 1+ staining was weak. The expression was compared between histological type (ILC and IC-NST) and each histological type was further stratified by tumour grade, and surrogate intrinsic subtype based on expression of ER, PR and HER2 by immunohistochemistry (luminal-like:

ER+/PR+/HER2-; luminal/HER2: ER+/PR+/ HER2+ or ER+/PR-/HER2+; HER2: ER-/PR-/HER2+; TN: ER-/PR-/HER2-).

It should be noted that in these cohorts the majority of IC-NST are either HER2-amplified or triple negative, therefore the results are slightly biased towards a cohort of more aggressive types and cancer. All statistical analyses were derived using the Chi-square test unless otherwise specified. A cut off of $p < 0.05$ was considered significant.

3.3.3.1. N-WASP expression in normal breast and breast cancer

There were 10 normal breast samples available for IHC analysis of N-WASP expression. Five of the normal breast samples were negative for N-WASP, and 5 cases showed weak 1+ staining in the cytoplasm in over 80% of the normal breast cells (**Figure 3.8**).

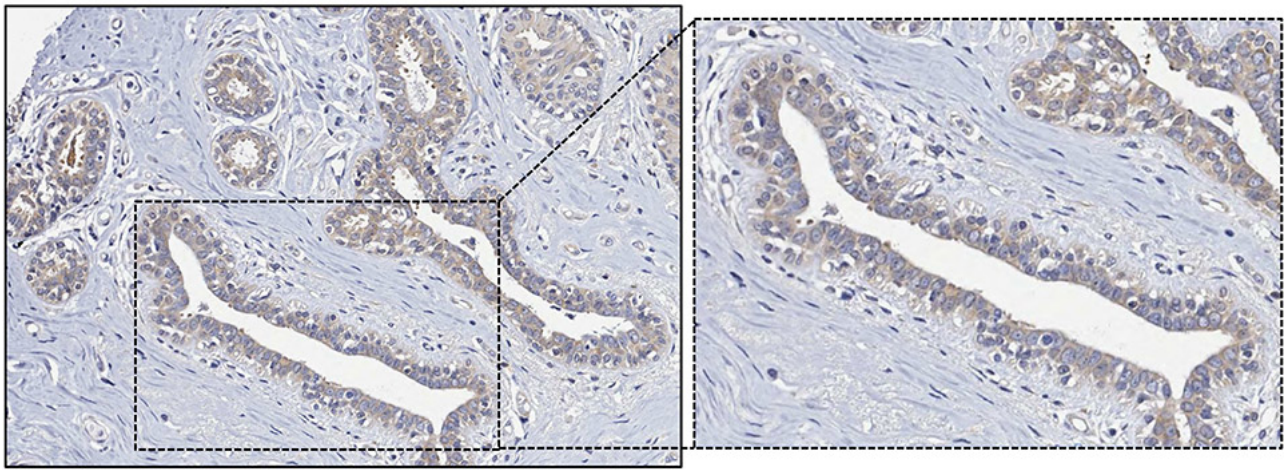


Figure 3.8: N-WASP is weakly expressed in the cytoplasm of normal breast.

There was a total of 81 IC-NST and 43 ILC available for analysis. Overall, N-WASP is more frequently expressed in ILC than IC-NST (ILC: 39/43, 90.6%; IC-NST: 43/81, 53.1%; **Table 3.2; Figure 3.9**) and this difference was found to be statistically significant ($p = 0.0001$; **Figure 3.10)**. When only considering 2+ and 3+ as positive, 55.8% of ILC were positive, compared with only 23.5% of IC-NST and this was still statistically significant ($p = 0.0003$).

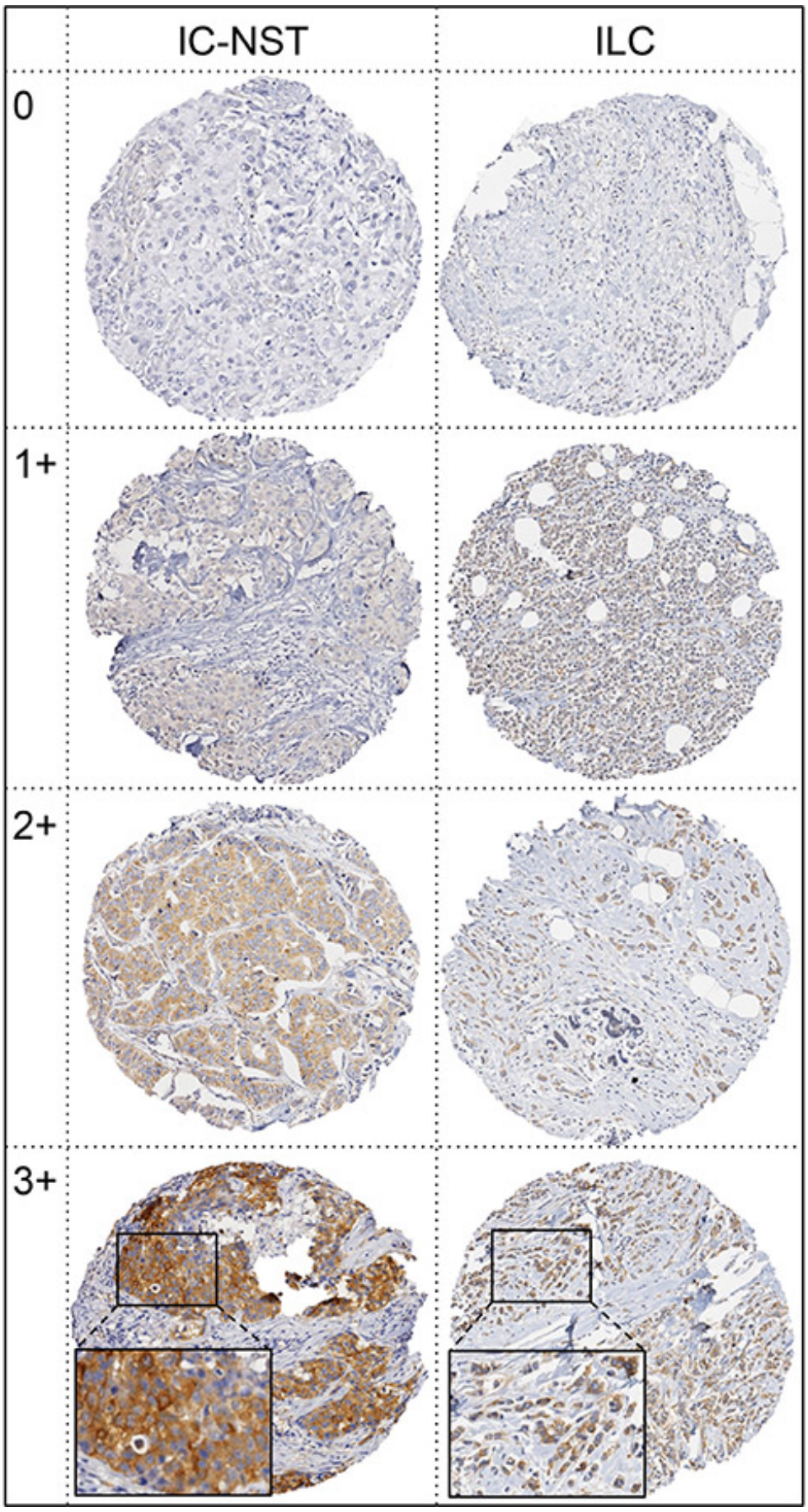


Figure 3.9: N-WASP expression in breast cancer. Representative images of staining intensity (1+, 2+, 3+) in ILC and IC-NST. Inset in the 3+ rows shows cytoplasmic localisation.

Table 3.2: IHC of N-WASP in IC-NST, ILC and their IHC-derived intrinsic subtypes.

IC-NST (n=81); n (%)					
N-WASP	**Luminal-like**	**Luminal / HER2**	**HER2**	**TN**	**Total**
Neg	5 (27.8)	11 (61.1)	13 (61.9)	9 (40.9)	38 (48.1)
1+	8 (44.4)	4 (22.2)	4 (19.0)	8 (36.4)	24 (30.4)
2+	4 (22.2)	4 (22.2)	4 (19.0)	3 (13.6)	15 (19.0)
3+	1 (5.6)	1 (5.6)	0	2 (9.10	4 (5.1)
Total	18	20	21	22	
ILC (n=43); n (%)					
N-WASP	**Luminal-like**	**Luminal / HER2**	**ER/PR neg**	**Total**	
Neg	3 (7.7)	0	1 (100)	4 (9.3)	
1+	12 (30.8)	3 (100)	0	15 (34.9)	
2+	20 (51.3)	0	0	20 (46.5)	
3+	4 (10.3)	0	0	4 (9.3)	
Total	39	3	1		

N refers to number of informative cores. TN = triple negative.

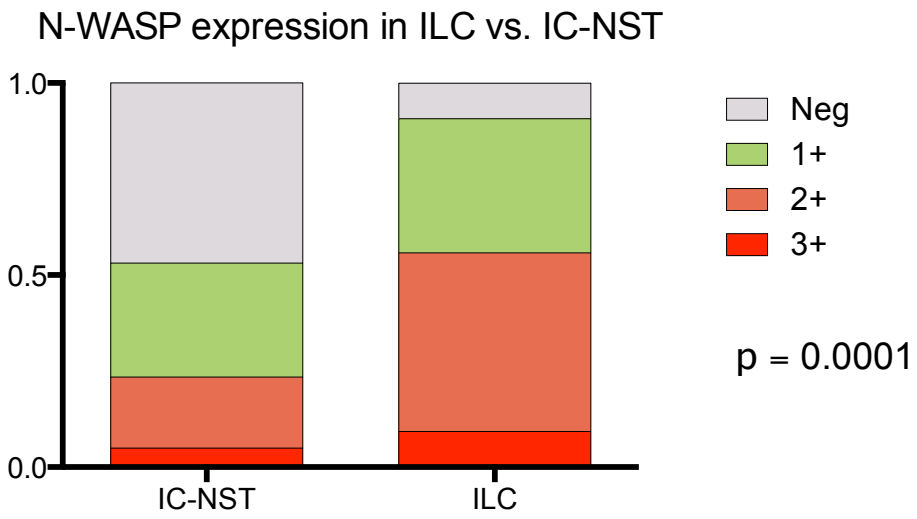


Figure 3.10: N-WASP is more frequently expressed in ILC compared to IC-NST (Chi-square test, p = 0.0001). A cut off of p < 0.05 is considered significant. Staining intensity was recorded; negative (0), weak (1+), moderate (2+) and strong (3+). When considering only 2+ and 3+ as positive, the expression of N-WASP remained statistically significant between IC-NST and ILC (p = 0.0003))

When stratified into IHC-surrogates for intrinsic molecular subtypes, the ILC primarily fell into the Luminal-like group (90%) as expected, and N-WASP was positive in 61.6% of these cases. Three ILC cases expressed HER2 and they all had 1+ weak cytoplasmic staining of N-WASP. IC-NST is a much more heterogeneous group, however, across all

intrinsic subtypes, N-WASP was frequently negative (**Table 3.2**). There was no statistically significant difference in N-WASP expression observed between the intrinsic subtypes and this might be due to the small numbers within each intrinsic subtype.

N-WASP expression was assessed across tumour grades and was frequently positive in grade 2 and 3 ILC, while IC-NST was primarily negative for N-WASP expression across all tumour grades (**Table 3.3**). Indeed, there was no statistically significant difference in N-WASP expression across tumour grade for IC-NST. Within ILC, the sample numbers are too small to analyse. There is a lack of low-grade tumours with a "good" prognosis to use as a comparison in this cohort. However, there are two grade 1ILC cases, both showing low expression of N-WASP (1+ staining intensity); five grade 1 IC-NST cases that have low to no expression of N-WASP; and, a single case with 2+ staining intensity.

Table 3.3: Analysis of N-WASP expression within tumour grades of ILC and IC-NST.

	ILC; n (%)				**IC-NST; n (%)**			
	Grade 1	**Grade 2**	**Grade 3**	**Total**	**Grade 1**	**Grade 2**	**Grade 3**	**Total**
Neg	0	3 (18.8)	1 (14.3)	4 (16)	2 (33.3)	12 (54.5)	23 (46)	37 (47.4)
1+	2 (100)	0	0	2 (8)	3 (50)	2 (9.1)	18 (36)	23 (29.5)
2+	0	10 (62.5)	5 (71.4)	15 (60)	1 (16.7)	8 (36.4)	6 (12)	15 (19.2)
3+	0	3 (18.3)	1 (14.3)	4 (16)	0	0	3 (6)	3 (3.8)
Total	2	16	7	25	6	22	50	78

Regarding alternative histologies, there were also four medullary breast carcinomas available for analysis; 4/4 tumours were grade 3 and 2/4 were positive for N-WASP (both were triple negative), therefore there is insufficient data to draw a conclusion from this rare tumour type (data not shown).

To assess if N-WASP is playing a role in the progression of breast cancer a specialised TMA consisting of 25 cases (**Section 2.1**; (Vargas et al., 2012)), each with representative cores from various stages of breast cancer progression; normal breast, *in situ* carcinoma, invasive carcinoma, and lymph node metastases was utilised. Four cases displayed an increase in N-WASP expression over the progression series, where the *in situ* and invasive carcinoma were negative, or weakly positive and the corresponding lymph node metastasis was positive. There were 5 cases with matching normal breast on this TMA and 4 of those cases had matching *in situ* carcinoma. Two cases had increased expression of N-WASP from normal breast to *in situ* (1+ to 2+), while 2 cases had no change in N-WASP expression.

A series of 8 ILC with matched lymph node metastases was also analysed and no significant change in N-WASP expression between the primary and metastatic tumour was found (data not shown). The remaining cases had varying degrees of N-WASP expression, suggesting that N-WASP does not play a major role in breast cancer progression.

3.3.3.1.1. E-cadherin and N-WASP co-expression

Since N-WASP was found localised at the membrane *in vitro* (Kovacs et al., 2011), it was of interest to correlate the expression of N-WASP with E-cadherin expression.
In the most frequent phenotype observed in ILC (46.4%), E-cadherin was negative and N-WASP was positive. In comparison, IC-NST was most frequently E-cadherin aberrant with N-WASP expression absent (44.8%; **Table 3.4; Figure 3.11**).

Table 3.4: Co-expression analysis of E-cadherin and N-WASP.

	ILC (n=28)						IC-NST (n=58)				
Immunophenotype	***n (%)***	**Luminal-like**	**Luminal/ HER2**	**HER2**	**TN**		***n (%)***	**Luminal-like**	**Luminal/ HER2**	**HER2**	**TN**
Double -	8 (28.6)	6	1	1	-		1 (1.7)	-	-	-	1
E-cad -, N-WASP +	**13 (46.4)**	13	-	-	-		0	-	-	-	-
Double +	0	-	-	-	-		15 (25.9)	-	5	9	1
E-cad +, N-WASP -	1 (3.6)	-	-	1	-		4 (6.9)	-	2	2	-
E-cad ab, N-WASP -	2 (7.1)	2	-	-	-		**26 (44.8)**	2	8	5	11
E-cad ab, N-WASP +	4 (14.3)	4	-	-	-		12 (20.7)	1	4	2	5

-: negative, +: positive, ab: aberrant.

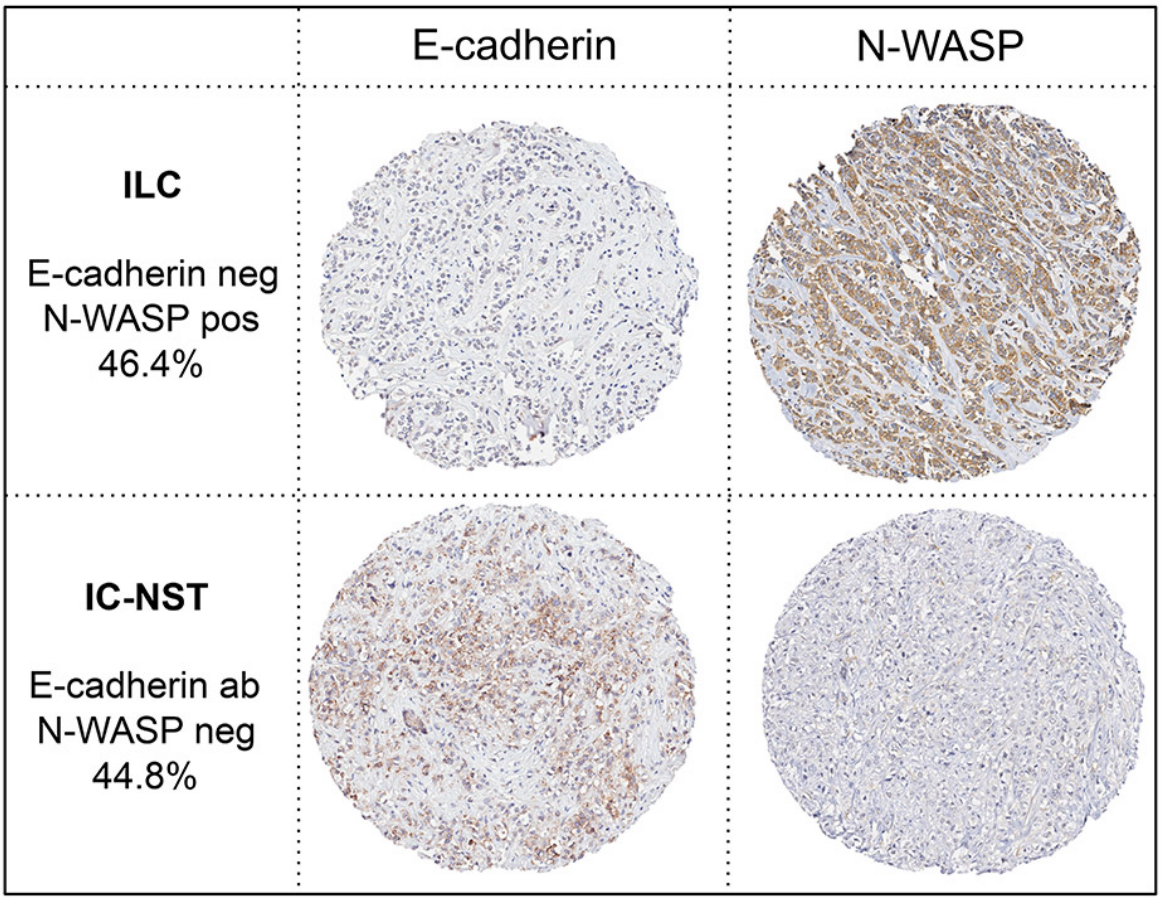


Figure 3.11: Representative images of the most frequent phenotype observed when assessing co-expression of E-cadherin and N-WASP within ILC and IC-NST.

3.3.3.1.2. E-cadherin and N-WASP expression in Mixed Ductal Lobular Carcinomas

A mixed ductal lobular carcinoma with both the ductal and lobular components present on the TMA (MDL41) was available for analysis. Interestingly, it was observed that N-WASP expression was absent in the ductal component (which was E-cadherin positive), but N-WASP was expressed in the lobular component (E-cadherin aberrant) (**Figure 3.12, top panel**). This was assessed further in the whole tissue section of this case as well as in two additional MDL cases to see if this is a common feature within MDLs and whether it may therefore represent a mechanism that drives the switch from ductal to lobular phenotype. This is explained and tested in more detail in **Chapter 4** where the hypothesis is that the lobular component arises from the ductal component due to some alteration in the integrity or regulation of the adherens junction.

In the first MDL case (MDL41), heterogeneous expression was observed (**Figure 3.12, bottom panel**). There were regions of the both the ductal and lobular components that were N-WASP positive, and others which were negative.

In the second case (MDL4), heterogeneous expression within both ductal and lobular components was also detected (data not shown). The third case (MDL6), however, was interesting in that the ductal component was for the most part negative (with some areas with weak expression) and the lobular component was positive for N-WASP. This case contains both pleomorphic lobular carcinoma *in situ* (PLCIS) and ductal carcinoma *in situ* (DCIS). N-WASP expression was heterogeneous in both the PLCIS and DCIS. Also interesting were the areas of columnar cell change (CCC), where N-WASP was highly expressed in the apical snouts (**Figure 3.13**). The staining pattern in the CCC was only observed in this one case, therefore this interesting observation will need to be explored in more cases. Overall there was no obvious correlation found between N-WASP and E-cadherin expression.

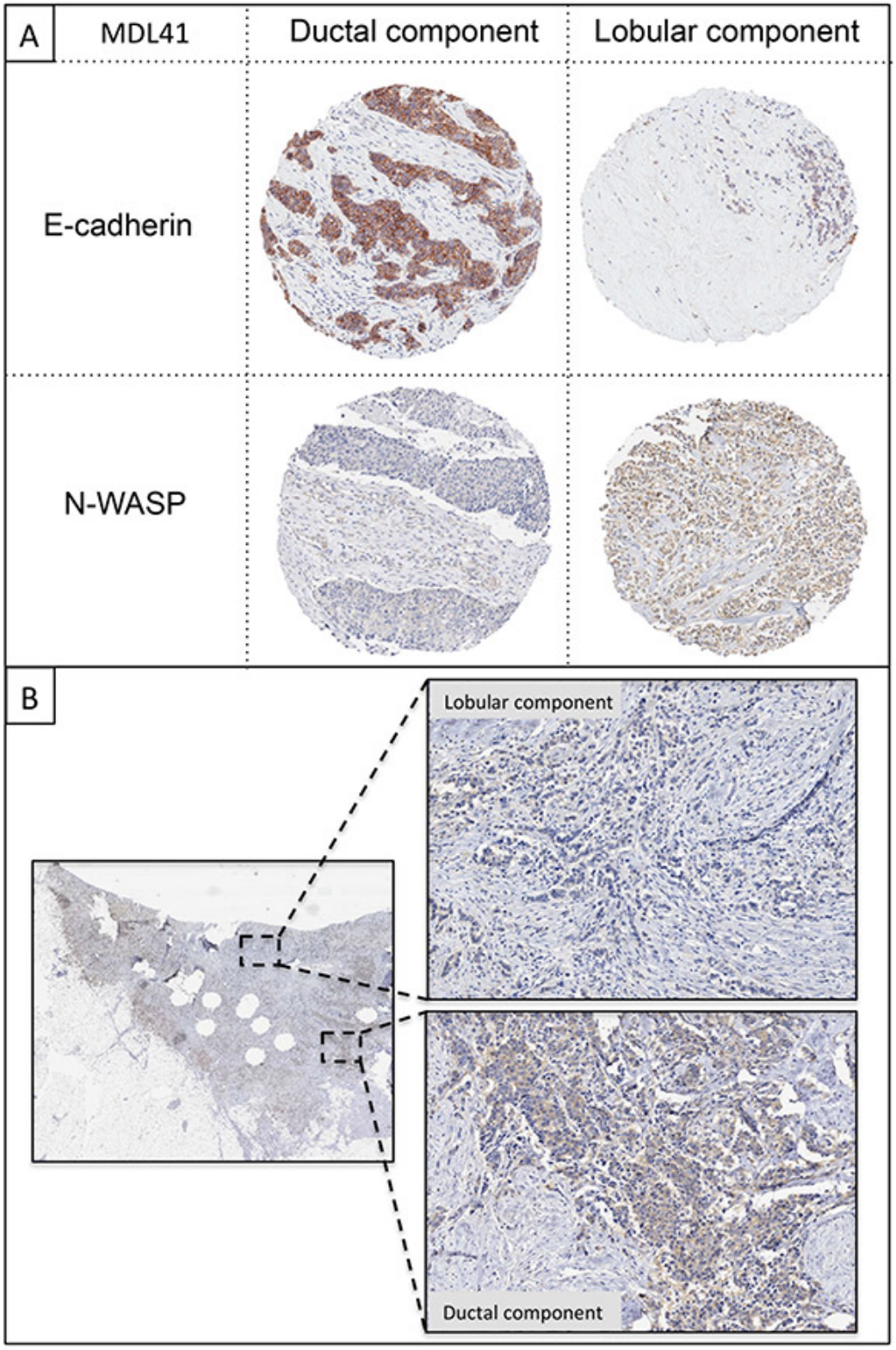


Figure 3.12: A mixed ductal lobular carcinoma (MDL41) exhibiting differences in E-cadherin expression between the ductal and lobular components. A) TMA cores: N-WASP is negative in the ductal component and positive in the lobular component. These tissue cores appear different since they were sampled from different areas of the tumour. B) The whole tissue section shows the opposite of the TMA cores. The holes in the tissue section are the areas that were sampled on the TMA.

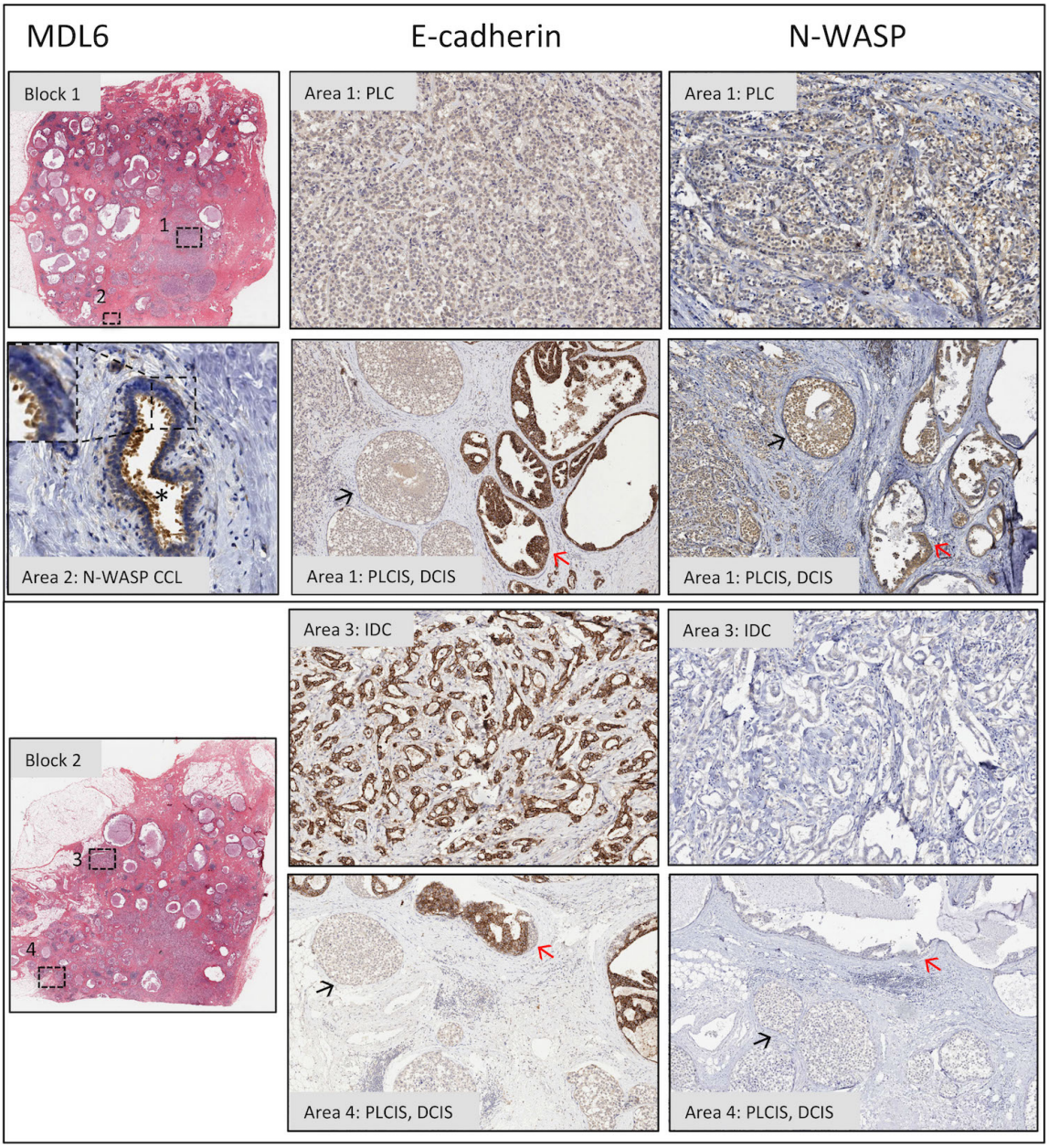


Figure 3.13: MDL6. This case shows heterogeneous expression of both E-cadherin and N-WASP in histologically different tumour regions. Two blocks from the specimen were stained. In block 1, PLC was positive for N-WASP and both the PLCIS (black arrow) and DCIS (red arrow) were positive. Interestingly, areas of columnar cell change (CCC) highly expressed N-WASP in the apical snouts of the cells and not the cytoplasm (asterisk and 20x inset). This is representative of all CCC within this tumour section. In block 2 however, the IDC, PLCIS and DCIS were all negative for N-WASP.

3.3.3.1.3. Co-immunofluorescent staining of N-WASP and E-cadherin of cells in culture

While N-WASP localises at the cell membrane in an *in vitro* model of colorectal carcinoma (Kovacs et al., 2011), our staining in breast tissue did not replicate this observation. To determine if N-WASP co-localised with E-cadherin at the membrane in breast cancer cell lines, co-immunofluorescence was performed on MCF7 cells (E-cadherin wild type IC-NST) and IPH-926 cells (E-cadherin mutated ILC – see **Section 2.13**). N-WASP was not localised to the membrane and was found in the cytoplasm in both cell lines (**Figures 3.14 and 3.15**).

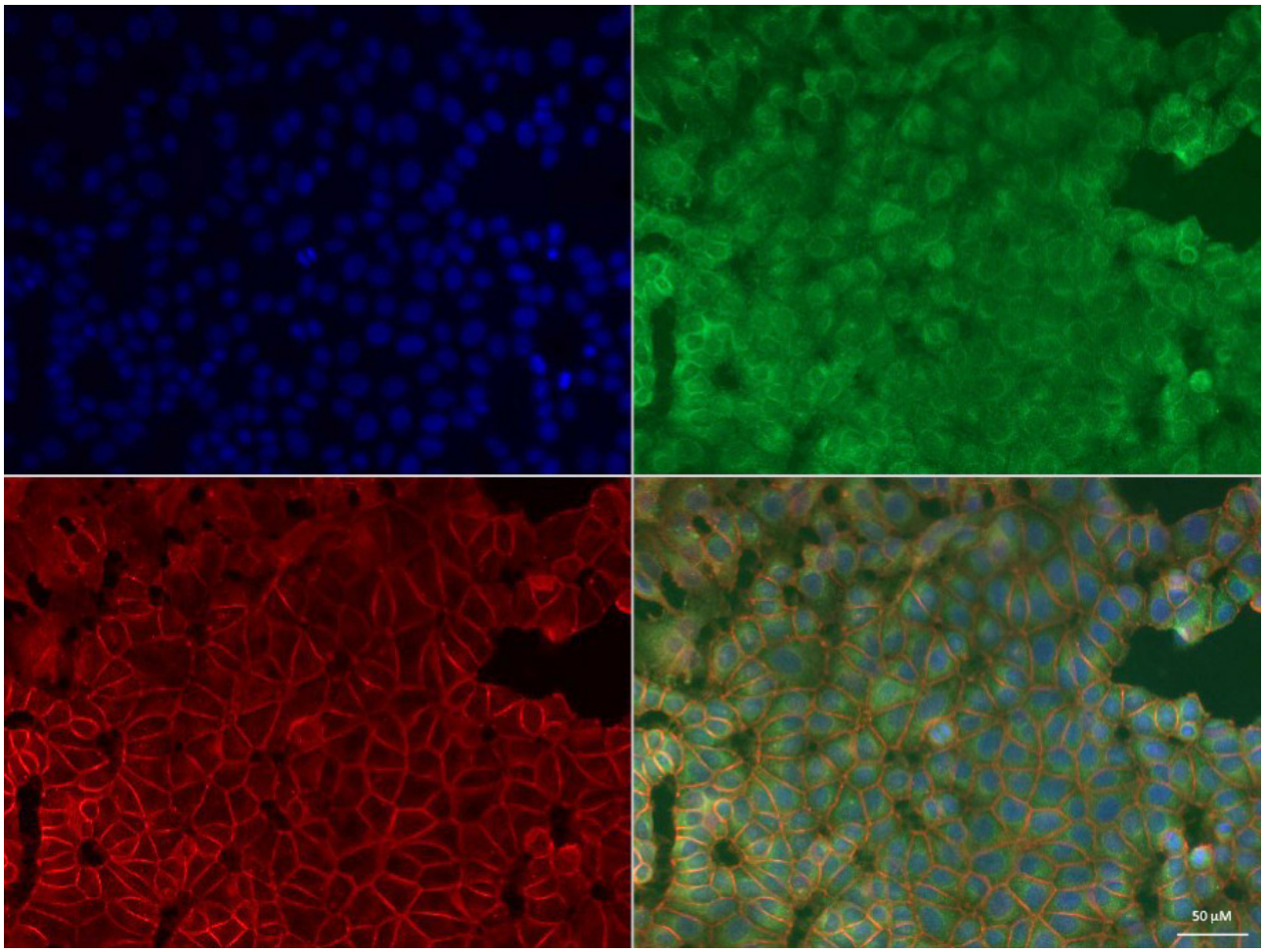


Figure 3.14: MCF-7 cells stained with N-WASP (green), E-cadherin (red) and DAPI (blue). N-WASP is localised in the cytoplasm (top right panel) with no membrane co-localisation observed between N-WASP and E-cadherin in *CDH1* wild-type IC-NST cells (bottom right panel).

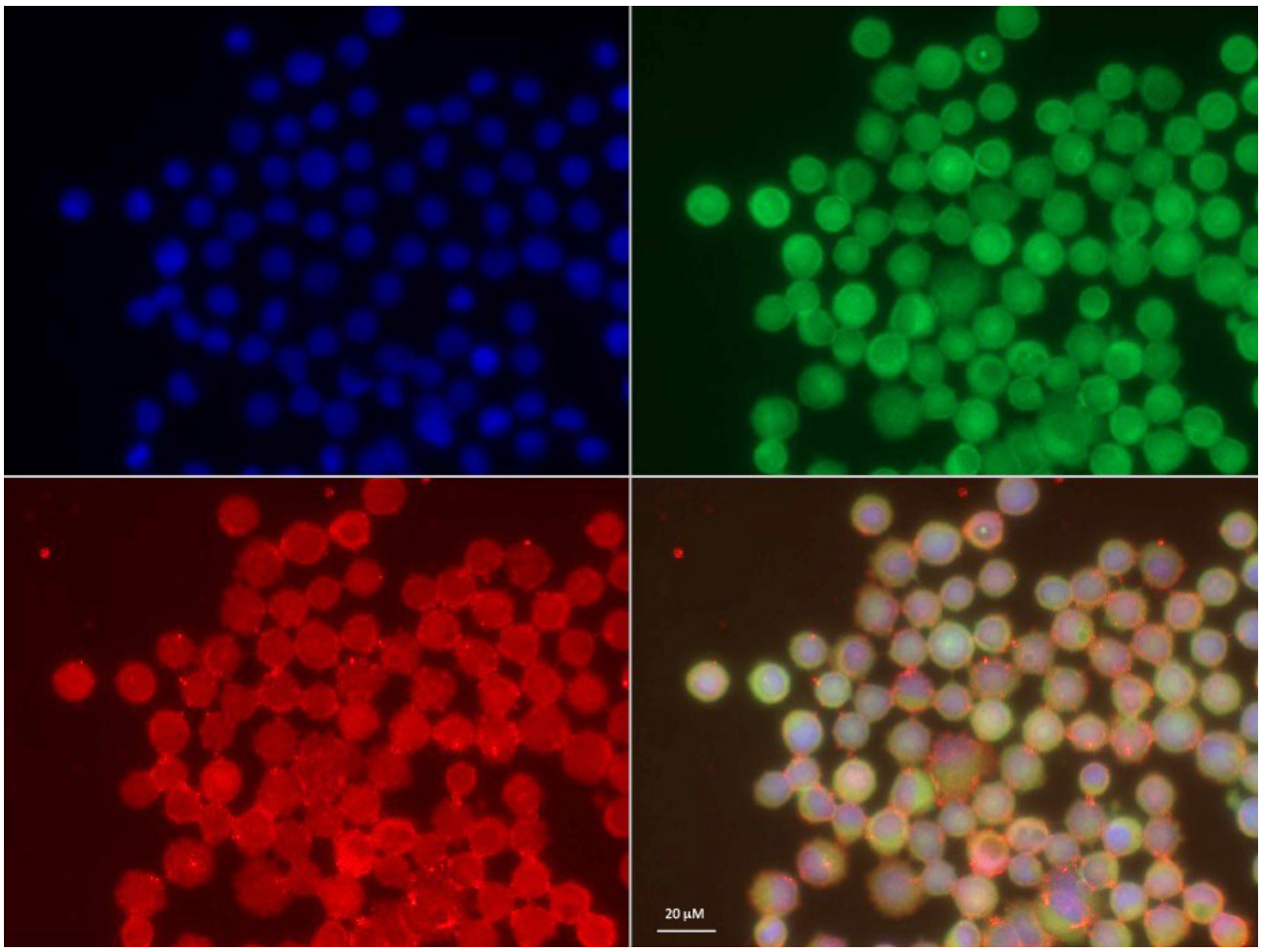


Figure 3.15: IPH-926 cells stained with N-WASP (green), E-cadherin (red) and DAPI (blue). N-WASP is localised in the cytoplasm (top right panel) with no membrane co-localisation observed between N-WASP and E-cadherin in *CDH1* mutant ILC cells (bottom right panel).

3.3.3.2. Analysis of ECT2 expression in normal breast and breast cancer

The localisation of ECT2 in the normal breast was investigated and 6 normal breast tissue cores were available for analysis. Three of 6 were negative for ECT2, however the remaining 3 cases displayed staining in the myoepithelial cell layer of the normal breast (**Figure 3.16**). To determine the localisation of ECT2 in normal breast, dual-immunofluorescence with antibodies for E-cadherin (epithelial cells, adherens junctions), collagen IV (basement membrane), CK14 (myoepithelial cell marker) and CK8/18 (luminal epithelial cell marker) was performed on fresh frozen sections. This confirmed that ECT2 is unlikely to be located at the adherens junctions but in the cytoplasm of myoepithelial cells of the normal breast (**Figure 3.17**).

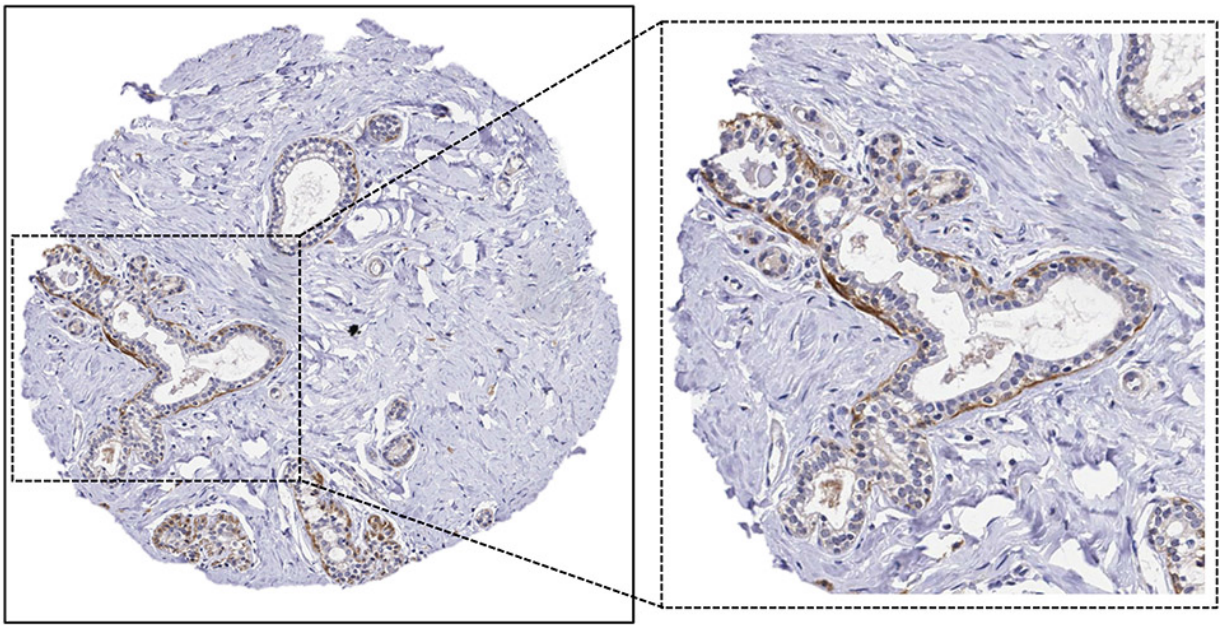


Figure 3.16: ECT2 expression in normal breast by IHC. ECT2 appears to be expressed in the myoepithelial cell layer of the normal breast.

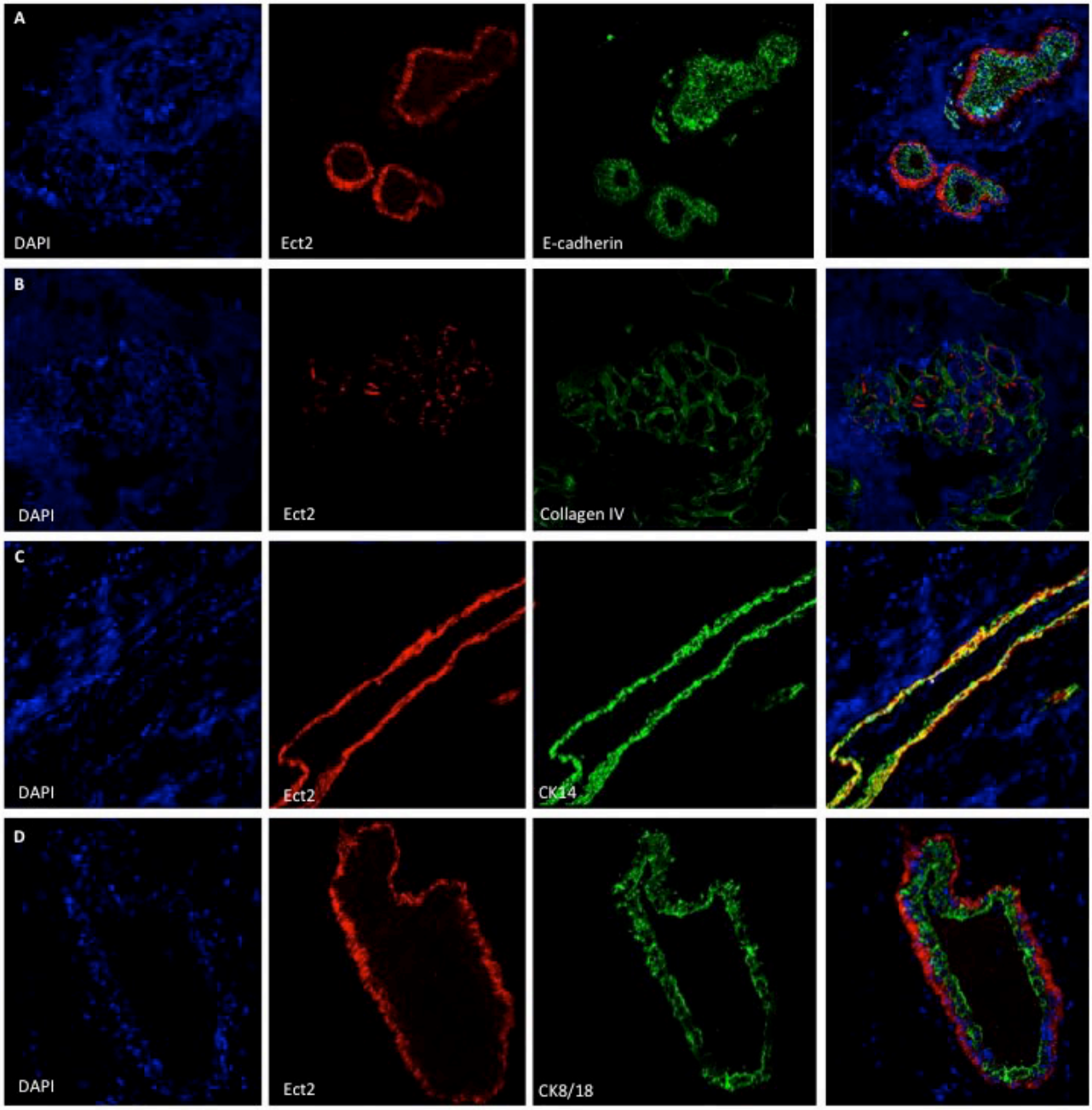


Figure 3.17: Co-immunofluorescent staining to determine the localisation of ECT2 in the normal breast. A) Co-staining of ECT2 and E-cadherin (epithelial membrane marker). B) Co-staining of ECT2 and collagen IV (a basement membrane marker). C) Co-staining of ECT2 and CK14 (a myoepithelial cell marker). D) Co-staining of ECT2 and CK8/18 (a luminal cell marker).

The expression of ECT2 in breast tumours was analysed by IHC across 96 IC-NST and 37 ILC. ECT2 protein was found localised in both the cytoplasm and nucleus. The tumour cells were scored as described above (**Section 3.3.3**). There were only 4 ILC cases with 3+ staining in the cytoplasm, no IC-NST showed 3+ staining. Cytoplasmic ECT2 was

positive in 10% of IC-NST compared with 40% of ILC (p = 0.0010, **Figure 3.18 and 3.19; Table 3.5**).

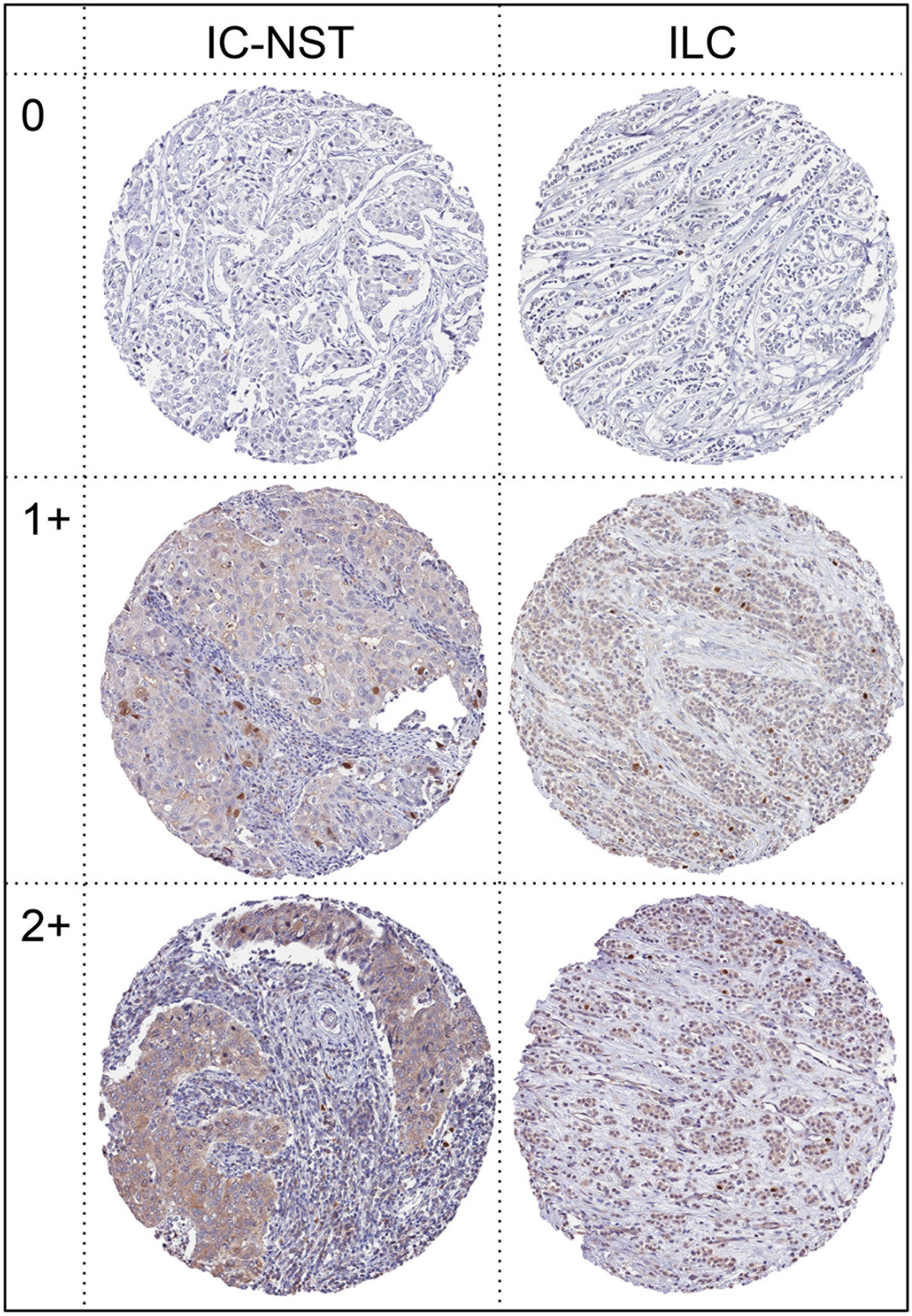


Figure 3.18: ECT2 expression in breast cancer. Representative images of staining intensity with IC-NST and ILC. 3+ cytoplasmic staining was not observed.

Table 3.5: ECT2 IHC results stratified within the intrinsic subtypes of ILC and IC-NST. Cytoplasmic (Cyto) and nuclear localisation was observed.

IC-NST (n=94); n(%)										
	Luminal (n=21)		**Luminal/HER2 (n=23)**		**HER2 (n=24)**		**TN (n=23)**		**ER-or ER-/PR- (n=3)**	
	Cyto	**Nuclear**	**Cyto**	**Nuclear**	**Cyto**	**Nuclear**	**Cyto**	**Nuclear**	**Cyto**	**Nuclear**
0	3 (14.3)	3 (14.3)	14 (60.9)	4 (17.4)	11 (45.8)	3 (13)	7 (30.4)	4 (17.4)	0	0
1+	15 (71.4)	0	9 (39.1)	1 (4.3)	13 (54.2)	0	9 (39.1)	0	3 (100)	0
2+	3 (14.3)	2 (9.5)	0	1 (4.3)	0	2 (8.7)	7 (30.4)	0	0	0
3+	0	16 (76.2)	0	17 (73.9)	0	19 (82.6)	0	19 (82.6)	0	3 (100)
Negative	**18 (85.7)**	**3 (14.3)**	**23 (100)**	**5 (21.7)**	**24 (100)**	**3 (13)**	**16 (69.6)**	**4 (17.4)**	**3 (100)**	**0**
Positive	**3 (14.3)**	**18 (85.7)**	**0**	**18 (78.3)**	**0**	**21 (91.3)**	**7 (30.4)**	**19 (82.6)**	**0**	**3 (100)**
ILC (n=36); n(%)										
	Luminal (n=32)		**Luminal/HER2 (n=1)**		**HER2 (n=1)**		**TN (n=1)**		**ER-/PR- (n=1)**	
	Cyto	**Nuclear**	**Cyto**	**Nuclear**	**Cyto**	**Nuclear**	**Cyto**	**Nuclear**	**Cyto**	**Nuclear**
0	3 (9.4)	11 (34.4)	0	0	1 (100)	0	0	1 (100)	0	1 (100)
1+	16 (50)	0	1 (100)	0	0	0	1 (100)	0	1 (100)	0
2+	9 (28.1)	3 (9.4)	0	0	0	0	0	0	0	0
3+	4 (12.5)	18(56.3)	0	1 (50)	0	1 (50)	0	0	0	0
Negative	**19 (59.4)**	**11 (34.4)**	**1 (100)**	**0**	**1 (100)**	**0**	**1 (100)**	**1 (100)**	**1 (100)**	**1 (100)**
Positive	**13 (40.6)**	**21 (65.6)**	**0**	**1 (50)**	**0**	**1 (50)**	**0**	**0**	**0**	**0**

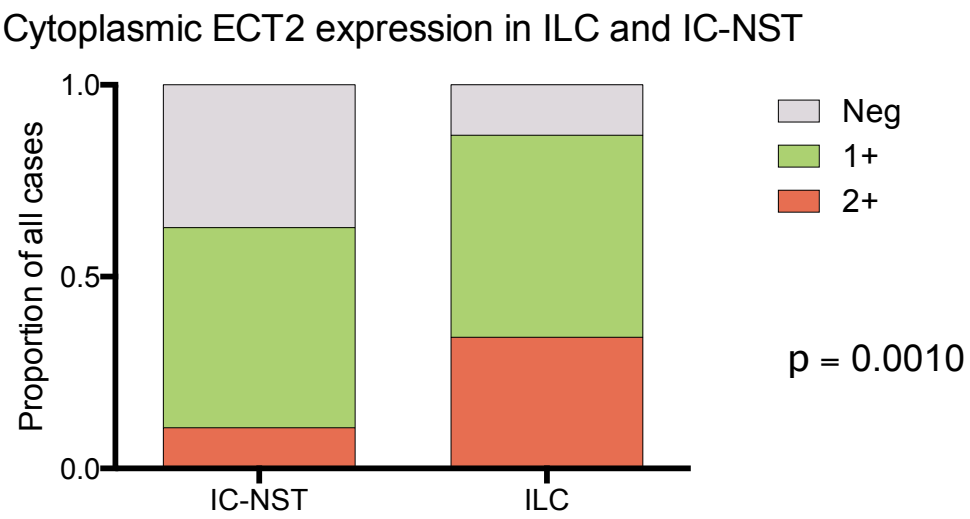


Figure 3.19: Cytoplasmic localisation of ECT2 is more frequently observed in ILC compared with IC-NST. There were no cases with 3+ staining intensity. Chi-square test, p = 0.0010.

Nuclear expression of ECT2 was higher in IC-NST (85%) than ILC (62%) (p = 0.0173, Fisher's exact test; **Figure 3.20**). As nuclear expression is an indicator of cell proliferation (as described in the introduction **Section 3.1.2** "**Model 2: The centralspindlin complex and ECT2**"), the higher nuclear expression in IC-NST is expected as the TMAs used tend to be biased towards higher-grade, proliferative tumours, and ILC are a much slower growing tumour type. This is reflected at the mRNA level, where *ECT2* mRNA is lower in the lobular and luminal A subtypes compared to the other breast cancer subtypes (**Section 3.3.1**).

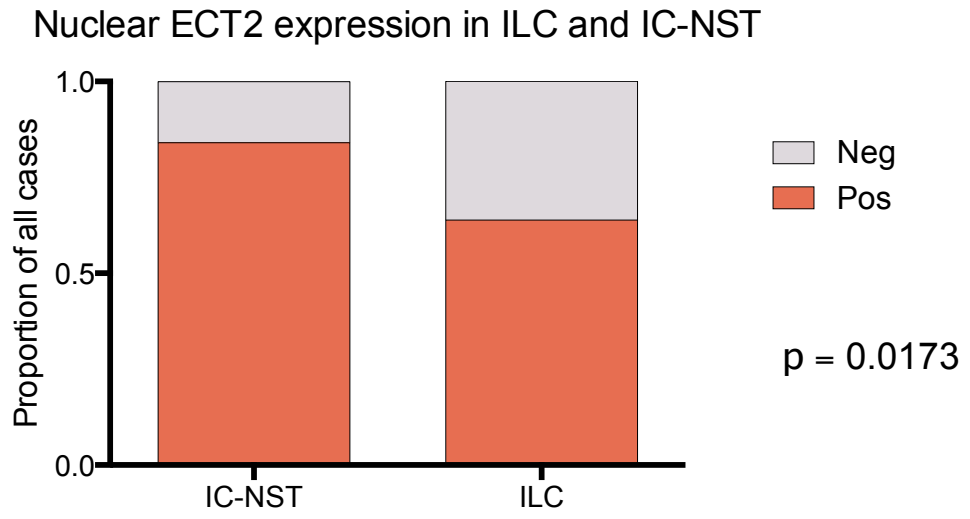


Figure 3.20: ECT2 nuclear expression is more frequent in IC-NST compared to ILC. Fisher's exact test p = 0.0173.

When stratified into IHC surrogate intrinsic subtypes, the luminal-like subtype was the most prevalent in ILC, and was frequently negative for cytoplasmic ECT2 (62.5%) and positive for nuclear ECT2 (65.6%). The same trend was observed across all IC-NST intrinsic subtypes (**Table 3.5**).

When considering tumour grade within each histological type, 61.5% of IC-NST were grade 3 (61.5%), whereas 62.2% of ILC were grade 2 (62.2%). Within IC-NST, cytoplasmic staining of ECT2 was most frequently negative (grade 1, 100%; grade 2, 92.6%; grade 3, 88.1%; p = 0.0027) and nuclear ECT2 was most frequently positive (grade 1, 60%; grade 2, 74.1%; grade 3, 89.8%; p = 0.0650). Within ILC, cytoplasmic staining of ECT2 was most frequently negative (grade 1: 67%, grade 2: 70%), although 5/9 (55%) grade 3 ILC displayed positive cytoplasmic staining (p = 0.4959, not significant). Within all grades across both IC-NST and ILC the predominant phenotype was nuclear positivity with cytoplasmic negativity (p = 0.0002; **Table 3.6**).

There were 13 cases of ILC with matched lymph node metastasis. ECT2 expression (cytoplasmic or nuclear) did not change significantly during progression to the lymph nodes in these cases (data not shown). There was no matched IC-NST lymph node metastases sampled across the TMAs used in this study.

Table 3.6: Correlation of ECT2 expression with tumour grade. Cytoplasmic expression of ECT2 was most frequently observed in high-grade ILC.

IC-NST; n (%)	Grade 1 (n=5)		Grade 2 (n=27)		Grade 3 (n=59)	
Immunophenotype	**Cyto**	**Nuclear**	**Cyto**	**Nuclear**	**Cyto**	**Nuclear**
Negative	5 (100)	2 (40)	25 (93)	7 (26)	52 (88)	6 (10)
Positive	0	3 (60)	2 (7)	20 (74)	7 (12)	53 (90)
Double pos	-	0	-	0	-	5 (8)
Double neg	-	0	-	5 (19)	-	3 (5)
Nuclear pos/Cyto neg	-	6 (60)	-	20 (74)	-	49 (83)
Cyto pos/Nuclear neg	-	2 (40)	-	2 (7)	-	2 (3)
ILC; n (%)	**Grade 1 (n=3)**		**Grade 2 (n=23)**		**Grade 3 (n=9)**	
	Cyto	**Nuclear**	**Cyto**	**Nuclear**	**Cyto**	**Nuclear**
Negative	2 (67)	1 (33)	16 (70)	9 (39)	4 (44)	2 (22)
Positive	1 (33)	2 (67)	8 (30)	14 (61)	5 (56)	7 (78)
Double pos	-	0	-	3 (13)	-	5 (56)
Double neg	-	0	-	4 (17)	-	2 (22)
Nuclear pos/Cyto neg	-	2 (67)	-	11 (48)	-	1 (11)
Cyto pos/Nuclear neg	-	1 (33)	-	5 (22)	-	1 (11)

3.3.3.2.1. E-cadherin and ECT2 co-expression analysis

The correlation of ECT2 expression with E-cadherin localisation was assessed. Only cytoplasmic ECT2 positivity was considered in this analysis. The most frequent phenotype observed in ILC was the lack of expression of both E-cadherin and ECT2 (in 50% of cases). In comparison, IC-NST was most frequently E-cadherin aberrant while ECT2 expression was absent (53.3%) (**Table 3.7; Figure 3.21**)

Table 3.7: E-cadherin and ECT2 co-expression analysis. ECT2 and E-cadherin were most frequently negative in ILC. In IC-NST, ECT2 was negative, with aberrant E-cadherin expression.

	ILC (n=32)				IC-NST (n=92)						
	n (%)	**Luminal-like**	**Luminal/ HER2**	**HER2**	***n (%)***	**Luminal-like**	**Luminal/ HER2**	**HER2**	**TN**	**ER-/PR-**	**n/a**
Double -	**16 (50)**	15	1	-	2 (2.2)	-	-	1	1	-	-
E-cad -, ECT2 +	12 (37.5)	12	-	-	0	-	-	-	-	-	-
Double +	0	-	-	-	4 (4.3)	3	-	-	1	-	-
E-cad +, ECT2 -	2 (6.3)	1	-	1	30 (32.6)	10	7	10	1	2	-
E-cad ab, ECT2 -	2 (6.3)	2	-	-	**49 (53.3)**	7	15	11	13	1	2
E-cad ab, ECT2 +	0	-	-	-	0	-	-	-	-	-	-

-: negative, +: positive, ab: aberrant

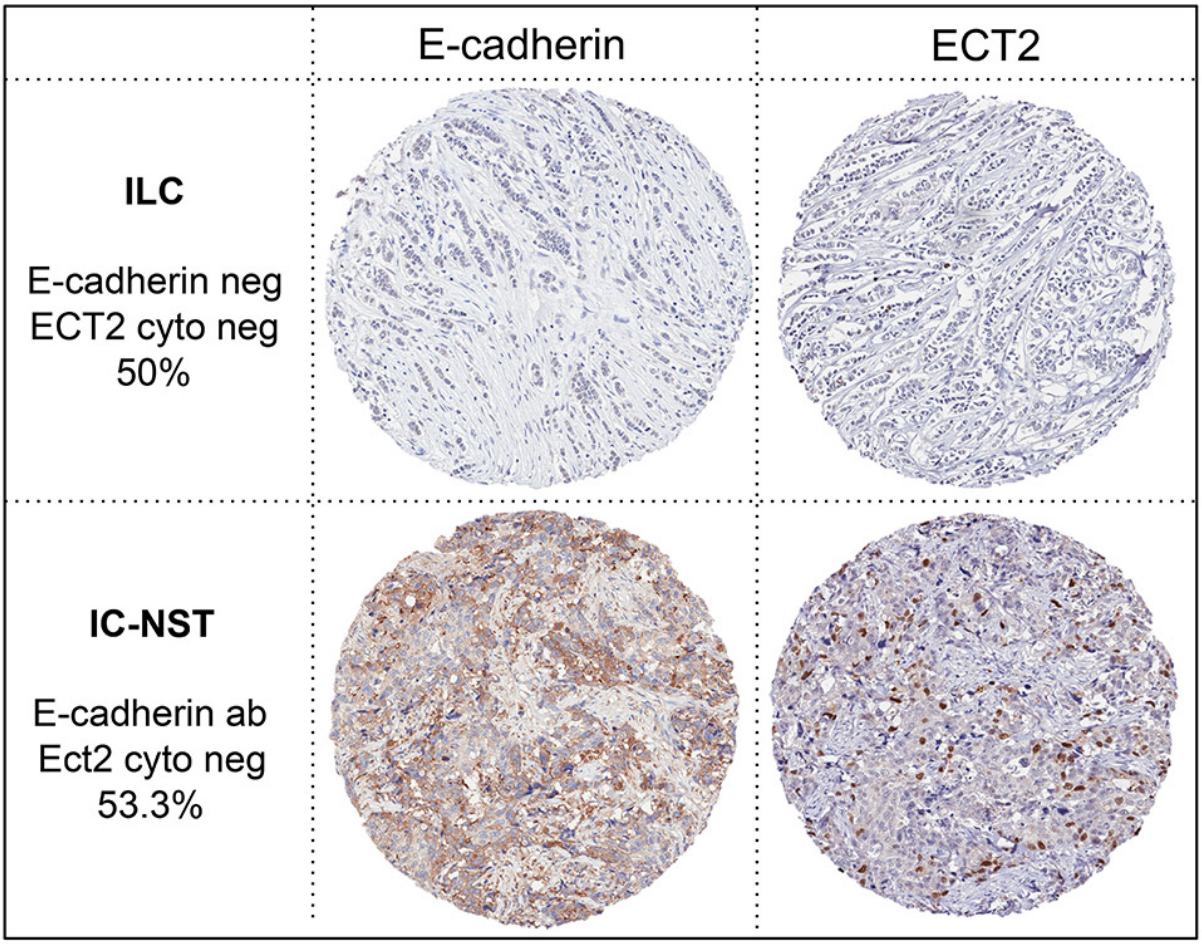


Figure 3.21: Representative images of the most frequent phenotype observed when assessing co-expression of E-cadherin and ECT2 within ILC and IC-NST. ab = aberrant.

3.3.3.3. RacGAP1 expression in normal breast and breast cancer

To investigate the localisation of RacGAP1 in the normal breast, there were 3 normal breast tissue cores available for analysis. Of the 3 cores, 1 was negative and 2 cores displayed strong cytoplasmic localisation in over 80% of the normal breast cells (**Figure 3.22**).

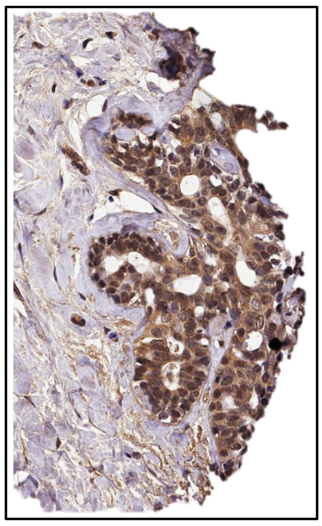


Figure 3.22: RacGAP1 expression in normal breast. Strong staining was observed in the cytoplasm of both luminal and myoepithelial cells.

IHC analysis was performed on 87 IC-NST and 39 ILC to assess the expression of RacGAP1 in breast cancer tissues. RacGAP1 was only expressed in the cytoplasm of the tumour cells. (**Table 3.8**) and each histological type was further stratified into grade and intrinsic subtype (**Section 3.3.3**).

Overall, 90% of IC-NST and 80% of ILC were positive for RacGAP1. However, when only considering 2+ and 3+ as positive (since 1+ staining was quite weak), IC-NST was positive for RacGAP1 in 60% of cases and 49% of ILC cases ($p = 0.1706$; **Table 3.8; Figure 3.23 and 3.24**). The differences observed however where not statistically significant.

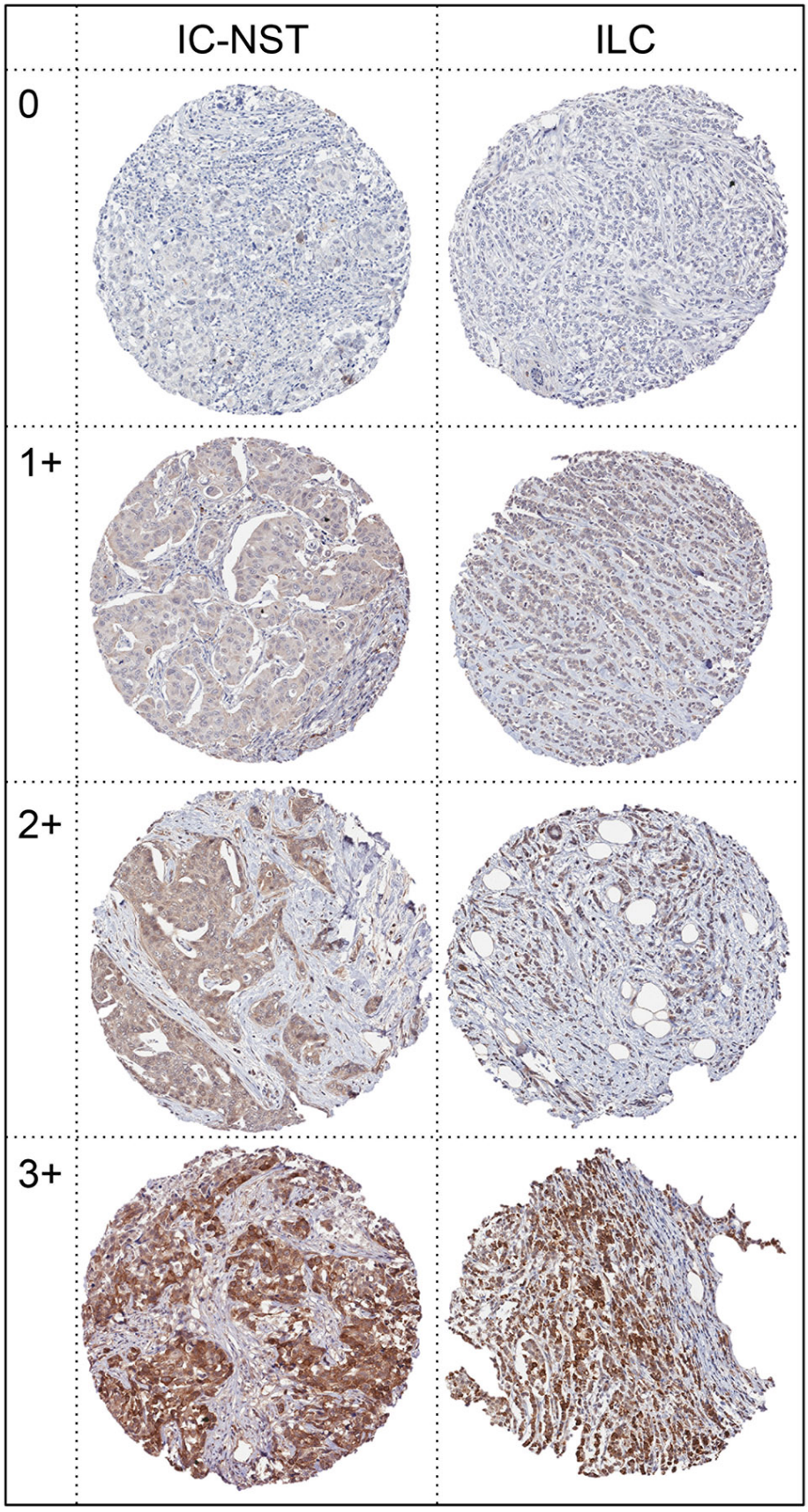


Figure 3.23: RacGAP1 expression in breast cancer. Representative images of staining intensity within IC-NST and ILC.

Table 3.8: RacGAP1 expression in breast cancer by IHC.

IC-NST (n=82)	Luminal-like (n=20)	Luminal/HER2 (n=18)	HER2 (n=21)	TN (n=23)	Total
0	4 (20)	1 (5.6)	0	3 (13)	8 (9.8)
1+	7 (35)	5 (28)	5 (5.6)	7 (30)	24 (29.3
2+	8 (40)	5 (28)	12 (57)	11 (47.8)	36 (43.9)
3+	1 (5)	7 (39)	4 (19)	2 (8.7)	14 (17.1)
Negative	**11 (55)**	**6 (33.3)**	**5 (23.8)**	**10 (43.5)**	**32 (39)**
Positive	**9 (45)**	**12 (66.7)**	**16 (76.2)**	**13 (56.5)**	**50 (61)**
ILC (n=39)	**Luminal-like (n=36)**	**Luminal/HER2 (n=1)**	**HER2 (n=2)**	**TN (n=0)**	**Total**
0	8 (22)	0	0	-	8 (20.5)
1+	11 (31)	0	1 (50)	-	12 (30.8)
2+	12 (33)	0	1 (50)	-	13 (33.3)
3+	5 (14)	1 (100)	0	-	6 (15.4)
Negative	**19 (52.8)**	**0**	**1 (50)**	-	**20 (51.3)**
Positive	**17 (47.2)**	**1 (100)**	**1 (50)**	-	**19 (48.7)**

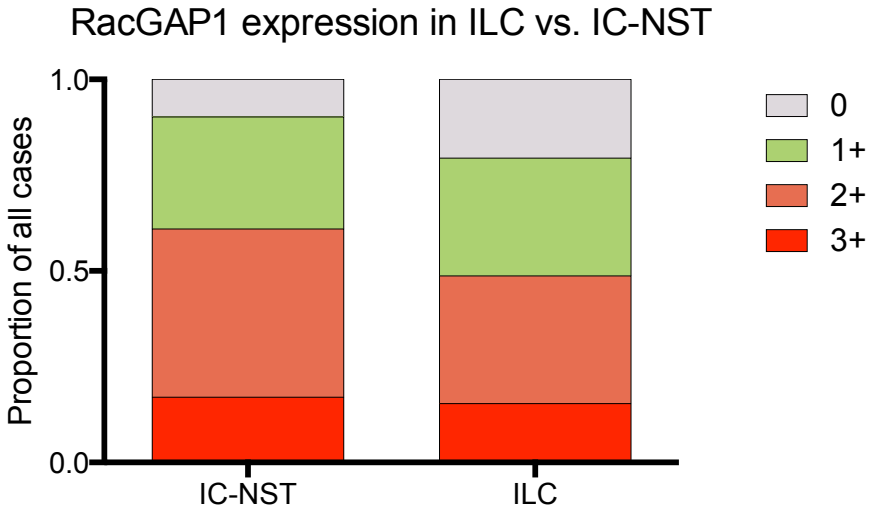


Figure 3.24: There was no statistically significant difference of RacGAP1 expression between ILC and IC-NST. Chi-square analysis p = n.s.

Within the luminal-like intrinsic subtype of ILC, RacGAP1 expression was negative in 52.8% and positive in 47.2% (p = 0.3466; **Table 3.8**). Among the IC-NST tumours, the HER2 expressing tumours were more frequently RacGAP1 positive (76.2% HER2 and 66.7% in Luminal/HER2; not statistically significant). Just over half of the triple negative tumours were RacGAP1 positive, and the luminal-like IC-NST were more frequently negative for RacGAP1. However none of the frequencies reached statistical significance.

RacGAP1 was more frequently positive in low-grade ILC tumours (grade 1, 2/3 cases; and grade 2, 14/23), however RacGAP1 was negative in 75% (9/12) of grade 3 ILC (p = 0.0743; **Table 3.9**). Conversely, within IC-NST tumours, RacGAP1 was more frequently negative in grade 1 tumours, and positive in grade 2 and 3 tumours, however the numbers of low-grade tumours were too low to make a comparison (p = 0.0964; **Table 3.9**). In summary, an inverse correlation was found between IC-NST and ILC where RacGAP1 was most frequently positive in high-grade IC-NST; in contrast, RacGAP1 was most frequently negative in high-grade ILC.

Table 3.9: RacGAP1 analysis within tumour grade.

	ILC; n (%)				IC-NST; n (%)			
	Grade 1	Grade 2	Grade 3	Total	Grade 1	Grade 2	Grade 3	Total
Neg	0	4 (17)	4 (33)	8 (21)	2 (50)	2 (8)	4 (7)	8 (10)
1+	1 (33)	5 (22)	5 (42)	11 (29)	1(25)	10 (38)	14 (26)	25 (30)
2+	0	10 (43)	3 (25)	13 (34)	1 (25)	11 (42)	26 (48)	38 (45)
3+	2 (67)	4 (17)	0	6 (16)	0	3 (12)	12 (22)	15 (18)
Total	3	23	12	38	4	26	54	84

There were 15 informative cases of matched primary tumour (ILC) with metastatic lymph node tumours . When stained for RacGAP1, a single case showed staining discordance between primary and the LN metastasis: the primary tumour was negative for RacGAP1, while the lymph node metastasis had 3+ expression in 1-10% of cells. The remaining cases, however, did not display any significant changes in RacGAP1 during metastatic progression, and it is therefore unlikely that RacGAP1 plays a role in the biology of lymph node metastasis in ILC.

3.3.3.3.1. E-cadherin and RacGAP1 co-expression analysis

The co-expression of E-cadherin and RacGAP1 was analysed. The most frequent phenotype observed in ILC was dual E-cadherin and RacGAP1 negativity (41.9% of cases). In comparison, IC-NST was most frequently E-cadherin aberrant with RacGAP1 expression was positive (31%); equally frequent was positive expression of both E-cadherin and RacGAP1 (31%) (**Table 3.10; Figure 3.25**)

Table 3.10: E-cadherin and RacGAP1 co-expression analysis..

	ILC (n=31)					IC-NST (n=87)						
Immunophenotype	*n (%)*	Luminal-like	Luminal/ HER2	HER2		*n (%)*	Luminal-like	Luminal/HER2	HER2	TN	ER-/PR-	n/a
Double -	**13 (41.9)**	12	1	-		0	-	-	-	-	-	-
E-cad -, RacGAP1 +	12 (38.7)	11	1	-		1 (1.1)	-	-	1	-	-	-
Double +	1 (3.2)	-	-	1		**27 (31)**	6	8	10	1	2	-
E-cad +, RacGAP1 -	1 (3.2)	1	-	-		8 (9.2)	6	-	1	1	-	-
E-cad ab, RacGAP1 -	0	-	-	-		24 (27.6)	5	5	3	9	-	2
E-cad ab, RacGAP1 +	4 (12.9)	4	-	-		**27 (31)**	3	6	6	11	1	-

-: negative, +: positive, ab: aberrant.

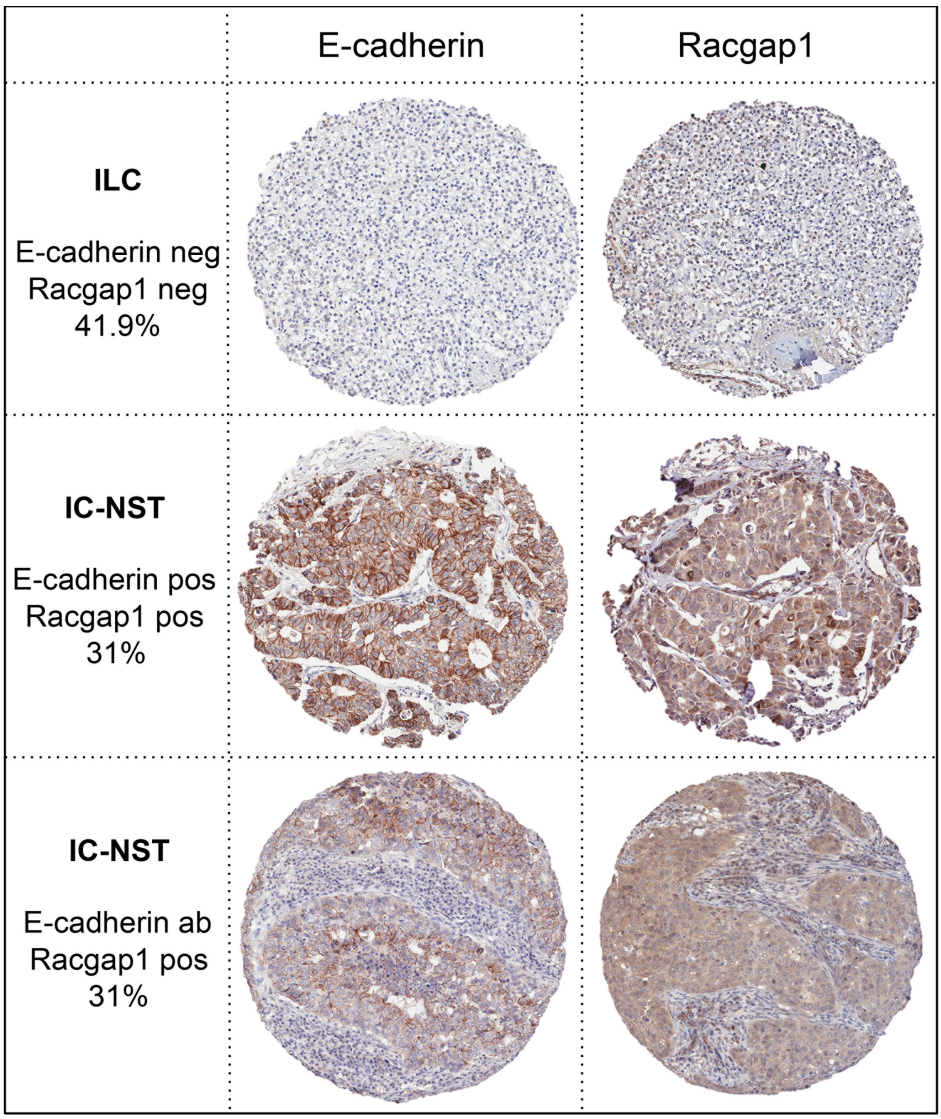


Figure 3.25: Representative images of the most frequent phenotype observed when assessing expression of E-cadherin and RacGAP1 within ILC and IC-NST samples.

3.3.3.4. E-cadherin co-expression with ECT2 and RacGAP1 (the centralspindlin complex)

As describe earlier in **Section 3.1.2** (*Model 2: The centralspindlin complex and ECT2*), both Racgap1 and ECT2 are necessary for E-cadherin integrity. Therefore the expression of both ECT2 and RacGAP1 was correlated with the expression of E-cadherin. The most frequent phenotype observed in ILC was E-cadherin negativity, while ECT2 was negative and RacGAP1 was positive (26.7%). This was followed by the lack of expression of all three proteins (23.3%) (**Table 3.11; Appendix 3.5**). In comparison, IC-NST displayed E-cadherin positivity alongside ECT2 negativity and RacGAP1 positivity (26.5%), or E-cadherin aberrant expression and negative for both ECT2 and RacGAP1 (26.5%) (**Table 3.11; Appendix 3.6**)

Table 3.11: Co-expression analysis of E-cadherin with ECT2 and RacGAP1. ILC are frequently negative for all three molecules, or only positive for RacGAP1. IC-NST are frequently E-cadherin positive or aberrant with loss of both ECT2 and RacGAP1.

	ILC (n=30)					IC-NST (n=83)						
Immunophenotype	*n (%)*	Luminal-like	Luminal/ HER2	HER2	TN	*n (%)*	Luminal-like	Luminal/ HER2	HER2	TN	ER-/PR-	n/a
Triple -	**7 (23.3)**	7	-	-	-	1 (1.2)	-	-	-	1	-	-
Triple +	0	-	-	-	-	3 (3.6)	2	-	-	1	-	-
E-cad -, ECT2 +, RacGAP1 +	4 913.3)	4	-	-	-	0	-	-	-	-	-	-
E-cad -, ECT2 -, RacGAP1 +	**8 (26.7)**	6	1	-	1	0	-	-	-	-	-	-
E-cad -, ECT2 +, RacGAP1 -	5 (16.7)	5	-	-	-	0	-	-	-	-	-	-
E-cad +, ECT2 -, RacGAP1 +	1 (3.3)	-	-	1	-	**22 (26.5)**	5	7	8	-	2	-
E-cad +, ECT2 +, RacGAP1 -	0	-	-	-	-	1 (1.2)	1	-	-	-	-	-
E-cad ab, ECT2 -, RacGAP1 -	1 (3.3)	1	-	-	-	**22 (26.5)**	5	5	6	4	-	2
E-cad ab, ECT2 -, RacGAP1 +	3 (10)	3	-	-	-	20 (24.1)	2	5	5	7	1	-
E-cad ab, ECT2 +, RacGAP1 -	0	-	-	-	-	3 (3.6)	-	-	-	3	-	-
E-cad ab, ECT2 +, RacGAP1 +	1 (3.3)	1	-	-	-	5 (6)	1	-	1	3	-	-

-: negative, +: positive, ab: aberrant.

3.4. Discussion

3.4.1. E-cadherin and the actin cytoskeleton regulation in breast cancer

Cell-cell adhesion mediated by E-cadherin is essential for epithelial cell survival. However, the loss of adhesion alongside anoikis resistance is a hallmark of a tumour cell, allowing it to survive. This is most evident in invasive lobular carcinoma cells that grow in single cell files and infiltrate the breast parenchyma and eventually distant organs. It is of vital importance to understand the molecular processes underlying tumour cell survival independent of cell adhesion in order to improve ways to target and kill tumour cells. Although the importance of the actin cytoskeleton and its regulation of E-cadherin adhesion has been researched extensively in cell culture systems, there is little research translating these findings into healthy or diseased human tissue and defining whether the molecular mechanisms found *in vitro* are corroborated *in vivo*.

3.4.2. Correlation of protein expression with gene expression data

It was difficult to correlate gene expression with protein data in this study due to the small number of samples to perform assess protein expression and the use of gene expression data from another cohort. The distribution of *N-WASP* gene expression in breast cancer was variable, although the median levels of expression were highest for luminal A and B intrinsic subtypes and the lobular and ER positive IC-NST histological subtypes compared to the ER negative breast cancer subtypes (**Figure 3.6**). Although in a small number of cases, this was not completely reflected at the protein level, with higher N-WASP protein expression found in ILC and less expression in the luminal (ER positive, HER2 negative) IC-NST (**Table 3.2**). This did not reach statistical significance and therefore needs to be investigated in a larger cohort with matching gene expression and protein data from the same patient samples.

Median *ECT2* mRNA expression levels were found to highest in the ER- HER2+ and triple negative IC-NST histological subtypes (**Figure 3.6**). This correlates with the high nuclear protein expression of ECT2 observed in the HER2+ and triple negative IC-NST (**Table 3.5**). *ECT2* mRNA was found to be down-regulated in both the lobular and ER+ HER2- IC-NST histological subtypes, and the luminal A and normal-like intrinsic subtypes. However, ECT2 was detected at the protein level in our ILC cohort, both the nucleus (65% of ILC

cases) and cytoplasm (40% of ILC cases). We were therefore unable to correlate gene expression with protein expression of ECT2 with the current ILC cohort.

The protein expression data for RacGAP1 somewhat correlates with the gene expression data presented in **Figure 3.6**, although it did not reach statistical significance. *RacGAP1* mRNA was found to be down-regulated in both the lobular and ER+ HER2- (luminal A-like) subtypes and this is validated with the luminal subtypes in both ILC and IC-NST frequently lacking RacGAP1 expression compared with the other intrinsic types (**Table 3.8**). RacGAP1 mRNA was found to be up-regulated in the triple negative, ER+ HER2+ IC-NST histological subtypes and luminal B intrinsic subtype. We did find that the HER2 expressing tumours were frequently positive for RacGAP1. There were only 23 cases of triple negative IC-NST available for assessment, and we found 56.5% to be positive for RacGAP1.

3.4.3. Membrane localisation of actin cytoskeleton regulators was not observed by IHC in clinical samples – limitations of translating *in vitro* findings *in vivo*.

The most striking observation in this study was the lack of concordance between the observed localisation of the actin cytoskeleton molecules *in vitro* (Kovacs et al., 2011, Ratheesh et al., 2012) and in clinical samples. There are a number of explanations that could account for this. Firstly, cells in a two-dimensional culture system behave very differently from cells in a three-dimensional human tissue structure. A cell culture system does not recapitulate the complicated nature of the human body such as the many different cell types present within an organ as well as the influence of the extracellular matrix, haematological and lymphovasculature systems, and the vast cocktail of molecules (such as hormones, chemokines *etc*.) that would be in supply at any given time. Cells grown in a two-dimensional culture adhere to each other in a monolayer, as well as adhering to a stiff plastic flask in a single layer of cells, forcing them into apical-basal polarity. This affects many signalling pathways such as proliferation, apoptosis and differentiation. The lack of integrin-mediated adhesion to the extracellular matrix also disrupts many signalling pathways (reviewed by (Baker and Chen, 2012)). Adhesions are quite dynamic and hard to control and manipulate *in vitro* and therefore the molecules that are recruited to the adhesion junctions in a culture system may not always be observed in a tissue section. There are also many variables within any cell culture system that can

affect data output, such as the density of cells within a flask and the composition of cell culture media, which has been shown to impact the genotype of the cell (Kim et al., 2015).

Secondly, the lack of observed membrane localisation may be due to differences between the uses of immunohistochemistry (IHC) versus immunofluorescence (IF).
IHC was applied (using the same antibodies used by the Yap Laboratory) as it allows access to many tissue samples, especially in a high throughput manner through the use of tissue microarrays. As it is has an indirect antibody detection method (the use of a secondary antibody) it amplifies the signal, making IHC very specific and sensitive. Another benefit of IHC is the ability to observe the staining in context of tissue morphology. This is more difficult using IF and a confocal microscope where a double stain with a protein of known localisation is required. A drawback of IHC is the difficulty of co-staining with more than one protein, whereas IF can use multiple fluorophores to detect multiple proteins. IF staining in breast tissue sections can also be difficult due to high autofluorescence (exacerbated by formalin fixation), although there are several methods reported to reduce autofluorescence, such as using Sudan Black, ammonia-ethanol and sodium borohydride (Baschong et al., 2001). The use of frozen tissue also improves the result; however high-throughput screening of fresh frozen tumours is not possible in the same was as an FFPE TMA analysis. Antibodies are sensitive to the type of fixation method used on FF sections (methanol *vs.* paraformaldehyde), and the fixative can also affect the levels of autofluorescence.

ECT2 co-immunofluorescence was performed successfully (after optimising with all fixation methods and without the need for chemical pre-treatment) on frozen normal human breast sections with little autofluorescence, and cytoplasmic localisation, rather than membrane, was observed (**Section 3.3.3.2**). Co-immunofluorescence was performed on cells in a monolayer culture to see if the membrane localisation of N-WASP could be reproduced (**Section 3.3.3.1.3**), however, only cytoplasmic localisation of N-WASP in both E-cadherin wild type (MCF-7) and E-cadherin mutated (IPH-926) cells lines was detected. N-WASP was found at the membrane of Caco-2 colorectal carcinoma cell line (Kovacs et al., 2011), however is important to note that the Caco-2 cell line is notoriously heterogeneous. Under different culture conditions, these cells can differentiate into cells resembling enterocytes of the small intestine, and therefore the culture conditions have a great influence on the cells' phenotypic traits (Sambuy et al., 2005). It is likely that differences in culturing conditions (the use of cytoskeleton stabilisation buffer) and

microscopy methodology are responsible for the differing results. Particularly, a spinning disk confocal microscope was used that is capable of scanning at much faster speeds to capture events that occur on the millisecond timescale in live cells (Kovacs et al., 2011). Due to restricted time and trouble optimising the antibodies, N-WASP was not assessed in tissue sections and the RacGAP1 antibody did not work for co-IF.

Overall, the use of both IHC and IF suggest that in the normal breast and breast cancer tissue these molecules are not localised at the membrane. Since human tissue samples are a snapshot in time, it is practically impossible to recapitulate *in vitro* models of such dynamic processes.

3.4.4. N-WASP may play a role in aggressive ILC

N-WASP has been found to be an essential component of invadopodia and a promoter of tumour invasion (Tang et al., 2013). In a rat model of mammary adenocarcinoma, it was found that the N-WASP containing invadopodia are essential for invasion, intravasation and metastasis to the lung (Gligorijevic et al., 2012). It was elegantly demonstrated that N-WASP promotes trafficking of membrane type 1-matrix metalloproteinase 1 (MT1-MMP) to the plasma membrane to promote extracellular matrix remodelling using human breast cancer cell lines in three-dimensional invasion assays (Yu et al., 2012). Therefore, N-WASP could potentially be a driver for greater invasive capacity *in vitro*, yet was not observed in our series of cases with metastases to the lymph nodes (**Section 3.3.3.1**).

The biology of N-WASP has been investigated in both breast and colorectal cancer, and was found to be down-regulated at both the mRNA and protein levels in tumour cells, compared to normal breast and colon epithelium (Martin TA et al., 2012, Martin et al., 2008). Patients whose tumours expressed low N-WASP had a much poorer overall survival and disease free survival. In the breast cancer cell line MDA-MB-231 and colorectal cancer cell line HRT18, it was found that the over-expression of N-WASP reduced the motility and invasive ability of the cells (Martin TA et al., 2012, Martin et al., 2008). It was hypothesised that N-WASP may be a putative tumour suppressor in breast cancer (Martin et al., 2008). This is inconsistent with the genomic data derived by the TCGA (**Figure 3.7**), where the most frequent alteration in N-WASP was gene amplification (11/962 samples) and mRNA down-regulation was not detected. It is therefore unlikely that N-WASP is a tumour suppressor. In order to confirm if N-WASP is indeed a tumour

suppressor, a larger cohort of breast cancer samples with matched normal breast needs to be assessed for both mRNA and protein expression. Functional characterisation of N-WASP in breast cancer cell line models would also help confirm this.

While our cohort does not have follow up data, we can generalise that high-grade tumours have a poorer overall survival than low-grade tumours. Our data shows that N-WASP is more likely to be highly expressed in high-grade ILC, and conversely, low-grade IC-NST. This analysis suggests that N-WASP expression may also play different roles in different histological types. The few informative cases of matched normal breast and invasive tumour were unable to confirm the results of (Martin et al., 2008). If anything, our results may be contradictory or merely demonstrate heterogeneity, whereby two cases had increased expression of N-WASP from normal breast to *in situ*, while 2 cases had no change in N-WASP expression. If N-WASP plays a role in oncogenic extrusion in a human tissue sample, we could hypothesise that N-WASP expression would increase during progression from normal to *in situ* carcinoma to invasive carcinoma. However, an increase in N-WASP expression in invasive lesions compared with its pre-invasive counterpart was only observed in 4 IC-NST cases, the remaining 21 cases were heterogeneous across the lesions. However, Yu et al, found N-WASP expression to increase in IC-NST compared to matching DCIS and normal breast (Yu et al., 2012), which also contradicts the findings of Martin et al. The differing results between these studies and ours may be due to different antibodies being used. N-WASP expression is most likely to be extremely dynamic and context dependent. It is therefore not surprising that given patient tumours are removed at various time points that cannot be controlled, we see heterogeneous protein expression within a group of tumours that are inherently heterogeneous.

As observed in **Section 3.3.3.1** N-WASP was more frequently expressed in ILC than IC-NST suggesting that maybe it is important for the invasive phenotype observed in ILC. The mechanism of which is still unknown, yet is unlikely to be driven by gene mutation or amplification as seen by the lack of genomic alterations in this gene (**Section 3.3.2)**.

If N-WASP expression is a feature of ILC, then we could hypothesis that N-WASP may be exclusively expressed in the lobular component of a mixed ductal lobular carcinoma. Across three cases of mixed ductal lobular carcinomas, heterogeneous expression of N-WASP was observed between invasive and pre-invasive lesions within the same specimen

and therefore is unlikely to drive the change in phenotype. It would be interesting however to investigate N-WASP expression in a larger cohort of MDL cases.

N-WASP is frequently mutated in pancreatic cancer (20%), and amplifications are often observed in melanoma and ovarian cancer (**Appendix Figure 3.2**) suggesting a different role for N-WASP in different cancer types. A chemical inhibitor against N-WASP has been established (Peterson et al., 2001) and if N-WASP proves to be important in driving tumourigenesis, this may pave the way towards treatments that may prevent further tumour cell invasion.

3.4.5. Could ECT2 be a potential biomarker for poor prognosis in ILC?

The role for ECT2 in cancer is emerging, particularly its oncogenic role in lung carcinoma. In normal lung epithelium, ECT2 is localised in the nucleus; however, in lung carcinomas ECT2 is found in the cytoplasm and the nucleus, with around 84% of NSCLC overexpressing ECT2 in the cytoplasm (Justilien and Fields, 2009). ECT2 is highly expressed at both the mRNA and protein levels in a variety of human tumours including brain (Salhia et al., 2008, Sano et al., 2006), lung (Hirata et al., 2009, Justilien and Fields, 2009), bladder (Saito et al., 2004), oesophageal (Hirata et al., 2009), pancreatic (Zhang et al., 2008) and ovarian tumours (Saito et al., 2004). Meta-analysis of TCGA data supports this, demonstrating multiple tumour types have genomic alterations of the *ECT2* gene, particularly gene amplifications, which are most common in lung cancer (see **Appendix Figure 3.3**).

Overexpression of *ECT2* mRNA and protein is associated with a poor prognosis in NSCLC (Hirata et al., 2009), glioblastoma (Salhia et al., 2008) (Sano et al., 2006) and oesophageal squamous cell carcinoma (Hirata et al., 2009). High ECT2 expression also correlated positively with tumour size and lymph node metastasis in oesophageal squamous cell carcinoma. High mRNA expression of *ECT2* has been found to correlate with high histological grade and poor overall survival in patients with primary gliomas. (Cheng Yung-Sheng et al., 2014). To date, ECT2 localisation in the normal breast has not been assessed and this information is vital to determine whether a protein is truly mislocalised during tumour growth. ECT2 was found in the cytoplasm of the myoepithelial cell layer of the ducts (**Section 3.3.3.2**). During normal mammary gland function the myoepithelial cells are involved in branching morphogenesis, maintaining luminal cell polarity and tumour

suppression by acting as a barrier to the stroma (Gudjonsson 2009). Breast cancers are believed to arise from the luminal cells and perhaps changes in luminal cell signalling at the beginning of tumourigenesis switches off ECT2 in the myoepithelial cells that induces their loss of polarity. In turn, ECT2 is turned on in the luminal cells, and this may provide the cells the ability to invade past the myoepithelial layer. ECT2 will need to be assessed in more cases with matched normal breast to confirm this hypothesis. It is also hypothesised that basal-like tumours may arise from progenitor cells present early in breast cell differentiation (Shehata et al., 2012). Basal-like tumours express proteins that are normally solely expressed in normal myoepithelial cells and these tumours have a poor prognosis (Badve et al., 2011). It would also be interesting, therefore, to assess the co-expression of ECT2 with IHC surrogate markers of the basal-like phenotype, such as EGFR, CK5/6 and CK14.

ECT2 was found to localise in both the nucleus and cytoplasm of breast tumour cells (refer to **section 3.3.3.2**). The role of ECT2 during mitosis is well characterised (Tatsumoto et al., 1999, Matthews et al., 2012), therefore the nuclear expression may be considered a marker of proliferating cells, analogous to Ki67. Nuclear expression also was more frequently observed in high-grade tumours, substantiating a role in proliferation. Cytoplasmic ECT2 was observed more frequently in ILC than IC-NST (35% *v.s.* 10%) and therefore it is important to assess the expression of ECT2 in a large cohort with follow up data, to assess if ECT2 expression is an indicator of poor prognosis, similar to NSCLC patients. In the analysis of ECT2 with E-cadherin expression, ECT2 is frequently lost in cases with absent or aberrant E-cadherin (50% of E-cadherin negative ILC and 53.3% of E-cadherin aberrant IC-NST), suggesting that ECT2 signalling may be play a role in maintaining E-cadherin integrity in some ILC and IC-NST. The annotation of *CDH1* mutation and methylation status is vital in understanding whether the loss of cell adhesion is due to changes in *CDH1* itself or if changes in actin cytoskeleton regulation play a role in tumours without an E-cadherin mutation.

3.4.6. RacGAP1 may be important in HER2 expressing breast tumours

RacGAP1 functions by negatively regulating Rho signalling. Its GTPase activity binds to activated RhoGTP and hydrolyses it to RhoGDP, and hence inhibiting or down regulating its function. RacGAP1 is also essential for cell division (Zhao and Fang, 2005) and embryogenesis (O'Brien et al., 2010). The role of RacGAP1 in breast cancer has been recently explored in a large series of patient samples with high-risk early breast cancer

(Pliarchopoulou et al., 2013). The expression of *RacGAP1* mRNA from 314 samples found an association between high *RacGAP1* expression and high-grade and strongly Ki67 expressing tumours, and patients had poor overall survival and poor disease free survival compared with patients with low *RacGAP1* expression. Similar results were also observed in a cohort of meningiomas (Ke et al., 2013), hepatocellular carcinoma (Wang et al., 2011b) and epithelial ovarian cancer (Lu et al., 2004). RacGAP1 is expressed in the nucleus of gastric carcinoma cells in the intestinal histological type, more so than the diffuse gastric type, and was found to be associated with poor outcome (Saigusa et al., 2015).

RacGAP1 was more frequently expressed in IC-NST compared with ILC and RacGAP1 expression was higher in tumours that expressed HER2 (see **Section 3.3.3.3**). The breast cancer study by (Pliarchopoulou et al., 2013) did not find an association with HER2 expressing tumours, however they did not look at protein expression within these samples, and the study didn't stratify histological or molecular subtypes. Equally, our results may be biased with the use of specialised TMAs as opposed to an unselected cohort.

Our meta-analysis did not find an overwhelming difference in gene expression between the breast cancer subtypes, yet protein expression was different. Other studies have found that *RacGAP1* expression is the same across all differentiated and undifferentiated embryonic stem cells, but that RacGAP1 protein is 2-fold higher in the undifferentiated cells (O'Brien et al., 2010), supporting the idea that the translation of the mRNA transcript to protein is dynamically regulated. This was also observed in a matched normal and tumour tissue from the same patient where the protein expression of RhoA, Rac1, Cdc42 was higher in tumour compared to normal, yet no changes in mRNA expression and no gene mutations were found between the tumour and normal tissue (Fritz et al., 2002). Together, this supports that there may be changes in post-transcriptional regulation that affect protein expression in different breast cancer subtypes. miRNA expression analysis would be an interesting avenue to explore in addressing these questions. There are 29 reported splice variants of *RacGAP1*, all of which have potentially different protein functions. Therefore, the antibody towards RacGAP1 may bind to many different isoforms and this may account for differing results in the literature.

Across other cancer types, *RacGAP1* alterations are most frequent in adenoid cystic carcinoma of the salivary gland and it is deleted in 15% of cases (**Appendix 3.4**).

RacGAP1 is located on 12q13, a region recurrently lost in this tumour type, and therefore RacGAP1 loss may be an incidental consequence of chromosome loss (Ho et al., 2013). There are few alterations in other cancer types, suggesting that there is little evidence to support that genomic alteration of *RacGAP1* is broadly important in tumourigenesis.

Overall, investigation of RacGAP1 in a cohort of breast cancer patients with long-term follow up cohort to validate findings of (Pliarchopoulou et al., 2013) is warranted. If RacGAP1 is involved in HER2 tumour biology, or more generally high-grade breast cancer and overall poor outcome, it would also be valid to assess RacGAP1 expression in a cohort of metastatic tumours to investigate if it is involved in metastasis to distant organs.

3.5. Conclusions

The loss of the catenin-binding complex that links the intracellular domain of E-cadherin to the actin cytoskeleton questions the role of actin cytoskeleton regulation in ILC. The role of Rho GTPases involved in actin cytoskeleton regulation in tumour development and progression has been demonstrated repeatedly in *in vitro* systems (Price and Collard, 2001); (Schmitz et al., 2000), however these data have not recapitulated the scenario in human tissue. As demonstrated in this chapter, it is challenging to validate *in vitro* findings *in vivo*; especially since human breast cancers are extremely heterogeneous in nature and are sampled at one particular point in time.

Genes encoding for these proteins are rarely altered at the genomic level, although they may still be important in a few breast cancer cases. The slight changes in gene expression observed may be the result of alterations in other signalling pathways. It is vital, therefore, to validate protein expression since the transcriptome does not reliably predict the proteome (Hack, 2004) and it is also critical to validate *in vitro* findings *in vivo* to assess if the predicted mechanisms found hold true for real life disease processes.

Considering the two models presented in the introduction (**Section 3.1.2**; *N-WASP; The Centralspindlin complex and Ect2*), it appears that they do not translate directly to human breast tissue samples. The molecules under investigation were not found in the membrane and thus their functions may be context dependent. In ILC, E-cadherin was frequently lost alongside cytoplasmic N-WASP expression. This may mean that if N-WASP is important in maintaining E-cadherin at the adherens junction *in vivo* then the loss of N-WASP from the

membrane and relocalisation to the cytoplasm may be important in some ILC, or may be a consequence of the alteration of other signalling pathways. As for the Centralspindlin complex model, ILC are frequently negative for all 3 molecules (E-cadherin, ECT2 and RacGap1), suggesting that the loss or relocalisation of these molecules may also be important in some ILC. In IC-NST however, the most frequent phenotype was E-cadherin positive with loss of both ECT2 and RacGAP1. These molecules may therefore not play a role in maintaining E-cadherin integrity in these tumours and another mechanism is likely to be involved.

Since we will always face the challenges that working with human samples brings, it would be best to study these pathways in three-dimensional culture systems and mouse models of breast cancer that are more representative of the human scenario.

Another point of future investigation is the analysis of these molecules across more normal breast epithelium samples. This is practically impossible to do in the cell culture system, as truly normal cells do not grow well in culture, and again, the challenges of interpretation of *in vitro* data are reintroduced. Co-immunofluorescence in normal breast tissue sections is a good start to identifying the localisation of these molecules in normal breast. Also, integration of mRNA and protein data is needed and the resulting data to be extended into matched samples of normal and tumour tissue.

CHAPTER 4

INVESTIGATING THE MOLECULAR EVOLUTION OF MIXED DUCTAL LOBULAR CARCINOMAS

4. Investigating the molecular evolution of mixed ductal lobular carcinomas

4.1. Mixed ductal lobular carcinomas are a distinct clinical entity

As described in **Section 1.6**, invasive breast cancers exhibiting both ductal and lobular morphological features are classified as mixed ductal lobular carcinomas (MDL) and account for 3-5% of all invasive breast cancers (Lakhani S. R., 2012). MDLs are considered to be a distinct entity to pure IC-NST and ILC (Rakha et al., 2009, Bharat et al., 2009). Throughout this chapter, the non-specialised invasive component in these mixed tumours is referred to as 'ductal', as opposed to IC-NST, since the terminology has only recently changed and to simplify terminology when comparing MDLs to pure IC-NST tumours.

4.1.1. Mixed ductal lobular carcinomas: collision tumours or clonally related?

The evolution of MDL carcinomas is not well understood. Several important questions relate to their origin; firstly, does the coincidence of these distinct morphological entities represent two independent tumours that have collided (so called 'collision tumours'), or do they arise from a common clone? Secondly, if the two components are clonally related, then what is the mechanism of transition from one growth pattern to the other? MDLs therefore represent a unique clinical model for interrogating intratumour heterogeneity, clonal evolution and the underlying mechanisms driving the acquisition of a diffuse and infiltrative growth pattern.

Historically, Loss of Heterozygosity (LOH), Comparative Genomic Hybridisation (CGH) and gene sequencing have been utilised to assess clonal relationships by assessing DNA aberrations between two (or more) lesions from the same specimen. For example, these types of analyses have demonstrated that, for example, columnar cell lesions are a non-obligate precursor lesion for low grade DCIS; DCIS is a non-obligate precursor for IC-NST; and LCIS is a non-obligate precursor for ILC (Vos et al., 1997, Simpson et al., 2005a, Kuukasjarvi et al., 1997b, Lu et al., 1998). Small pieces of evidence support the theory that when lesions with lobular and ductal morphological growth patterns are co-localised in the same specimen they are also likely to be clonally related lesions, as opposed to separate collision tumours (Wagner et al., 2009, Buerger et al., 2000), but this is not well explored. These studies reported the shared loss of a common allele (by LOH or CGH analysis) in

both the DCIS and LCIS and corresponding invasive components of either ductal or lobular morphology (Wagner et al., 2009, Buerger et al., 2000), suggesting that these morphologically distinct entities may have evolved from the same neoplastic clone. Using chromosomal CGH, Buerger et al. found identical copy number profiles in cases with co-existing DCIS and ILC and hypothesised that LCIS and low-grade DCIS are closely related lesions (Buerger et al., 2000). They also hypothesised that the loss of E-cadherin may represent the molecular switch to a more discohesive phenotype, however the mechanisms underlying this hypothesis was not explored.

During my Honours research in 2011, I studied the clonal relatedness of different morphological components of four MDL cases. Lesions with different growth patterns were laser capture microdissected and then analysed for DNA copy number alterations by chromosomal CGH (due to its capacity to utilise small amounts of DNA) (Kutasovic, 2011). DCIS, LCIS, and invasive ductal and lobular morphological components were assessed for copy number similarities and differences. All of the lesions within a case shared copy number alterations, supporting the notion that they were each derived from a common neoplastic clone (**Figure 4.1**, **Table 4.1**). There were also alterations that were unique to a given lesion (**Figure 4.1**). These alterations likely occurred during, after, or contributed to, the clonal divergence of the two components. E-cadherin's gene *CDH1* is located on 16q22.1 and E-cadherin expression is lost in about 90% of ILC (**Sections 1.5** and **1.7**). E-cadherin is therefore a candidate target gene for this loss in MDLs. To investigate whether E-cadherin deregulation plays a role in the phenotypic switch of the MDL, we examined E-cadherin subcellular localisation using IHC. We identified differential E-cadherin expression between the two phenotypic components in the 4 cases analysed. The ductal component had strong membranous staining for E-cadherin and each of its binding partners, whereas the lobular component showed aberrant staining for E-cadherin, and also for β-catenin and p-120 catenin. This aberrant staining was detected as cytoplasmic localisation, which is in contrast to most pure ILC, where E-cadherin expression is completely negative (**Figure 4.2**).

Recent advances in next generation sequencing now enable the detection of low frequency mutations and copy number alterations and are beginning to reveal greater insights into clonal progression and clonal diversity within a single tumour and between different regions of a primary tumour and subsequent metastases (Navin et al., 2011, Yachida et al., 2010, Gerlinger et al., 2012, Yates et al., 2015). The application of these

types of technology to morphological heterogeneity observed in MDL may contribute towards understanding of the molecular mechanisms that may be driving the evolution of morphological variation in these tumours (**Section 1.4** and **1.9.5**).

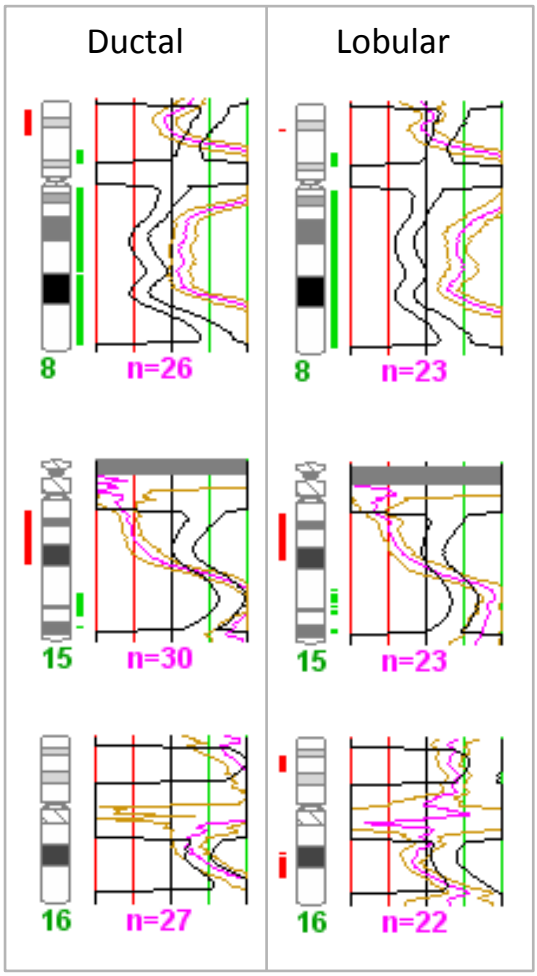


Figure 4.1: Chromosomal CGH data of invasive ductal and lobular components of case MDL4. Copy number alterations on chromosomes 8 (8p deletion, 8p-q gain) and 15 (15q deletion and gain) were shared by both morphological components. Loss on chromosome 16 was unique to the lobular component.

Table 4.1: Summary of DNA copy number alterations in 4 MDL cases.

Case #	MDL1					MDL2			MDL3			MDL4			
Component	D	L	DCIS	LCIS	Shared	D	L	Shared	D	L	Shared	D	L	Shared	
CGH – Gains		1p	2q		1q	3q	6p	15q	4q	10q	1q	2q	7p/q	1q	11q
		4p/q	8q		11q		11q	6p	9p		6p/q	9q		3q	14q
		11p	10p		16p		Xq	17q			8q	12q		5q	15q
		15q	14q					20q			10q			6q	17q
		17q	17q								17q			8p/q	20q
		18p/q									20q			10p	
Component	D	L	DCIS	LCIS	Shared	D	L	Shared	D	L	Shared	D	L	Shared	
CGH - Losses		2q	1p	6q	11q	15q	14q	6q	16p		2q	3p	7p	1p	12p
		6q	3p	8p	16q						3p	5q	16p/q	4p	13q
		8p	4q		22p/q						4p	18p/q	17q	6p	14q
		10q	11p								5p	Xp	18q	8p	15q
		13q									5q		22p/q	9p	21p/q
		14q									8p		Xp/q	10q	
											13q			11p	
											15q				

D: invasive ductal component. L: invasive lobular component. Shared: chromosomal gains and losses were identified in both the ductal and lobular components of the MDL cases. p/q: the alteration was found on both chromosome arms.

Jamie Kutasovic, BSc Hons UQ 2011. Investigating the molecular evolution of lobular breast carcinomas.

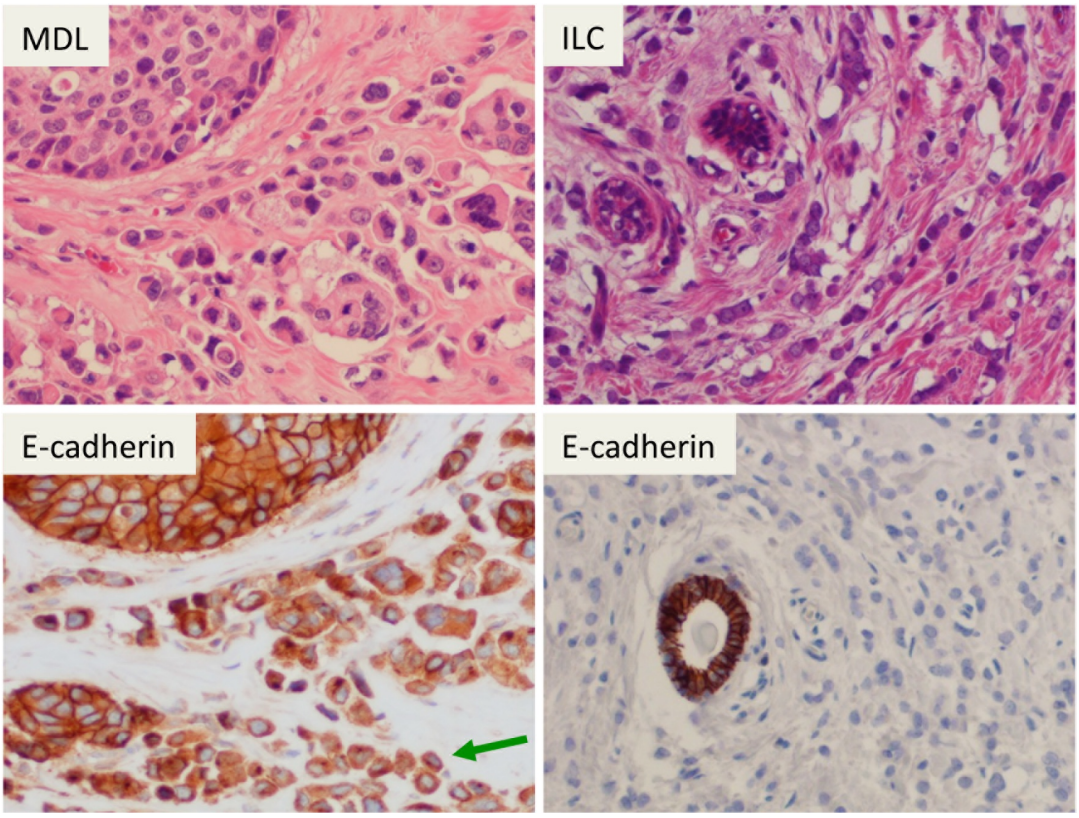


Figure 4.2: Representative images of a MDL and ILC and the differences in E-cadherin staining observed. The green arrow indicates aberrant localisation of E-cadherin in the lobular component of this MDL carcinoma.

4.1.2. Hypotheses and aims

We hypothesise that the lobular and ductal components of a MDL are growth patterns arising from a common neoplastic clone as opposed to the collision of two independent tumours. We propose that clonal progression occurs from a ductal to lobular morphological growth pattern and that deregulation of the functional E-cadherin complex that regulates epithelial cell-cell adhesion is driving this switch to a lobular-like phenotype, resulting in an enhanced invasive capability of cancer cells. We also hypothesise that the mechanisms underlying E-cadherin deregulation are different to the mechanisms known in pure ILC, since we did not find the complete loss of protein expression that is frequently observed in ILC.

To address this hypothesis, the aims of this chapter were to build on previous work by i) collating a large cohort of MDLs and characterising the clinical and morphological features; ii) interrogating the expression of the E-cadherin complex by IHC; iii) investigating whether epithelial to mesenchymal transition may play a role in the progression to the lobular-like

phenotype and iv) performing molecular analysis by sequencing the exome of selected cases to investigate clonal evolution and identify potential mutations driving progression.

4.2. Results

4.2.1. Clinical and pathology features of mixed ductal lobular carcinomas

A cohort of 82 MDL cases was accessed through Pathology Queensland and Sullivan and Nicolaides archives. **Table 4.2** describes the clinical features obtained from pathology reports of the cohort. The clinical and pathology features of this cohort were compared to those of the QFU series (specifically the IC-NST and ILC cases; **Section 2.1)**, and statistical analyses performed as described in **Section 2.3**. The average age of diagnosis was 57, which was significantly younger when compared to ILC ($p = 0.0083$), but not IC-NST ($p = n.s$). There was no significant difference in tumour size between MDLs and either IC-NST or ILC. For 6/82 cases, tumour grade was attributed to both the ductal and lobular components, and in these cases, the highest score was recorded for the tumour. MDLs were more frequently grade 2 (58.5%) and 3 (30.5%); the higher frequency of grade 2 tumours was significantly different compared to IC-NST ($p = 0.0254$). A significant difference was also found between ILC and MDLs, ILC are more frequently grade 2 ($p = 0.0006$). Compared to both IC-NST and ILC, MDL patients present more frequently with lymph node metastasis at diagnosis ($p = 0.0033$ and $p = 0.0097$, respectively). *In situ* carcinoma was diagnosed in 58/82 cases and it was interesting that 51 of 58 cases (87.9%) presented with ductal carcinoma *in situ* (DCIS). Of the 51 cases, 35 (60.3%) presented with DCIS alone, while 16 (27.6%) presented with co-existing lobular carcinoma *in situ* (LCIS) within the same tumour specimen. Only 7 of the 58 MDL cases (12.1%) had LCIS and no DCIS diagnosed. This distribution of DCIS and LCIS in the MDL cohort was significantly different to that observed in the IC-NST and ILC cases of the QFU cohort ($p < 0.0001$; see **Table 4.2**). For instance, there were significantly fewer cases with LCIS only in the MDL cohort relative to the ILC cohort ($p < 0.0001$).

MDLs more frequently expressed ER and PR compared to IC-NST ($p = 0.0102$ and $p = 0.0003$, respectively), but there was no difference in expression when compared to ILC. HER2 amplification, however, was found to be more frequent in MDLs compared to ILC ($p = 0.0196$).

Table 4.2: Clinical and pathology features of MDL carcinomas. Statistical comparisons have been made with a sporadic cohort of IC-NST (n = 256) and ILC (n = 64) from the QFU cohort.

	MDL cohort (n=82)	QFU IC-NST (n=256)	p value	QFU ILC (n=64)	p value
	n (%)	n (%)		n (%)	
Age of diagnosis					
Average	57	58	0.3238^	62	0.0083^
Range	28-86	27-88		40-85	
Median	55	58		63	
Tumour size (cm)					
<2	33 (42.9)	108 (42.2)	0.2929	21 (34.4)	0.2017
2 to 5	33 (42.9)	92 (35.9)		24 (39.3)	
>5	11 (14.3)	56 (21.9)		16 (26.2)	
Unknown	5	-		3	
Total	82	256		64	
Tumour grade*					
1	9 (11.0)	40 (15.6)	0.0254	2 (3.1)	0.0006
2	48 (58.5)	106 (41.4)		56 (87.5)	
3	25 (30.5)	110 (43.0)		6 (9.4)	
Lymph node status					
Positive	41 (68.3)	65 (41.1)	0.0033*	14 (40.0)	0.0097*
Negative	19 (31.7)	79 (54.9)		21 (60.0)	
Unknown	22	112		29	
Total	82	256		64	
Preinvasive lesions					
DCIS	35 (60.3)	119 (100)	< 0.0001	0 (0)	< 0.0001
LCIS	7 (12.1)	0 (0)		29 (93.5)	
DCIS & LCIS	16 (27.6)	0		2 (6.5)	
Not recorded	24	137		33	
Biomarker status					
Oestrogen receptor					
Positive	72 (90.0)	192 (76.8)	0.0102*	53 (91.4)	0.7842*
Negative	8 (10.0)	58 (23.3)		5 (8.6)	
Unknown	2	6		6	
Total	82	256		64	
Progesterone receptor					
Positive	67 (83.8)	154 (62.1)	0.0003*	38 (70.4)	0.0651*
Negative	13 (16.3)	94 (37.9)		16 (29.6)	
Unknown	2	8		10	
Total	82	256		64	
HER2 (IHC)					
Positive	14 (18.2)	45 (18.7)	1.0*	4 (8.0)	0.1255*
Negative	63 (81.8)	196 (81.3)		46 (92.0)	
Unknown	5	15		14	
Total	82	256		64	

Table continued over page

Table 4.2: Clinical and pathology features of MDL carcinomas. Statistical comparisons have been made with a sporadic cohort of IC-NST (n = 256) and ILC (n = 64) from the QFU cohort.

	MDL cohort (n=82)	20YFU IC-NST (n=256)	p value	20YFU ILC (n=64)	p value
	n (%)	n (%)		n (%)	
HER2 (ISH)					
Positive	8 (17.8)	27 (11.2)	0.2181*	2 (3.4)	0.0196*
Negative	37 (82.2)	215 (88.8)		56 (96.6)	
Total	45	242		58	

Chi-squared test for significance unless otherwise specified; ^ t-test; * Fisher's exact test, with a p value considered significant if < 0.05. Significant values highlighted in red.

4.2.2. Investigating the E-cadherin adhesion complex in MDLs

As the most defining feature in lobular carcinomas is the loss of E-cadherin, the integrity of E-cadherin and its binding complex partners, β-catenin and p120-catenin, was assessed in a subset of 51 cases of MDLs by IHC on whole tissue sections (**Section 2.5.4**). The expression of each protein was evaluated in each morphological component (DCIS, LCIS and invasive tumour areas of both ductal and lobular growth patterns) in each case. Each marker was scored as positive (complete membrane staining), negative (no protein expression), aberrant (either fragmented membrane staining or cytoplasmic localisation), or a mix of positive/aberrant (some cases displayed both positive membrane staining and aberrant staining), as shown in **Figure 4.3**. The distribution of staining of each molecule in each morphological component is summarised in **Table 4.3**. For the most part, the DCIS (91.7%) and invasive ductal (77.6%) components showed normal membranous protein expression for E-cadherin, β-catenin and p120-catenin. The majority of LCIS were negative for E-cadherin (70%) and β-catenin (88.9%), and cytoplasmic for p120-cateinin (62.5%). The lobular component in the majority of MDLs, exhibited aberrant cytoplasmic staining (43.1% and 35.4%) or incomplete membrane staining (27.5% and 25%) for E-cadherin and β-catenin, respectively. p120-catenin was frequently localised to the cytoplasm (60.9%). In only 9 (17.6%) cases did the lobular component exhibit archetypal E-cadherin negative staining.

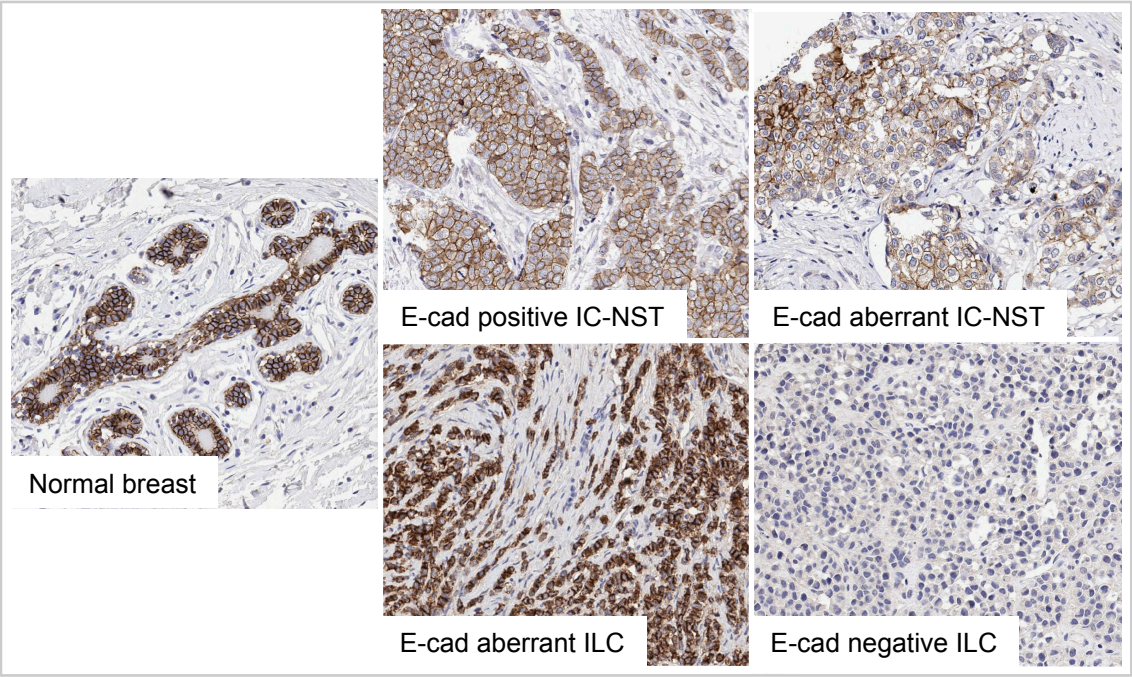
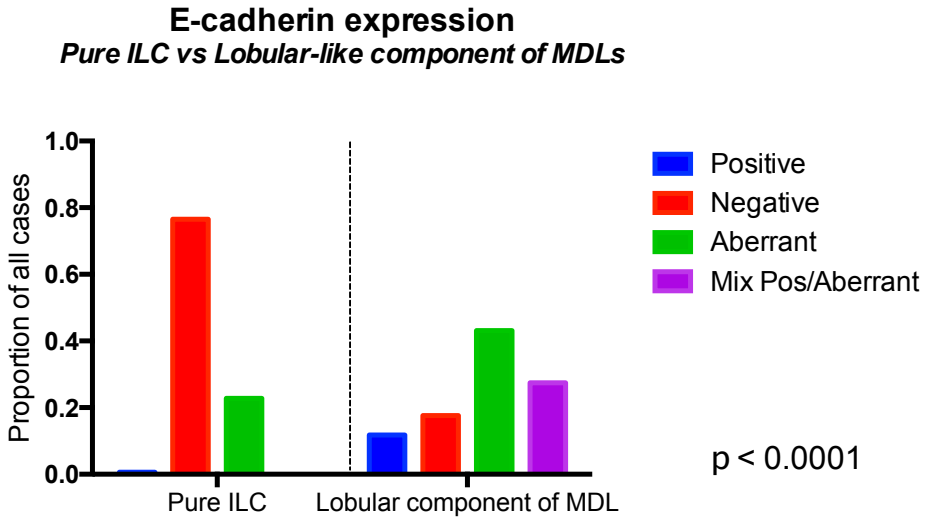


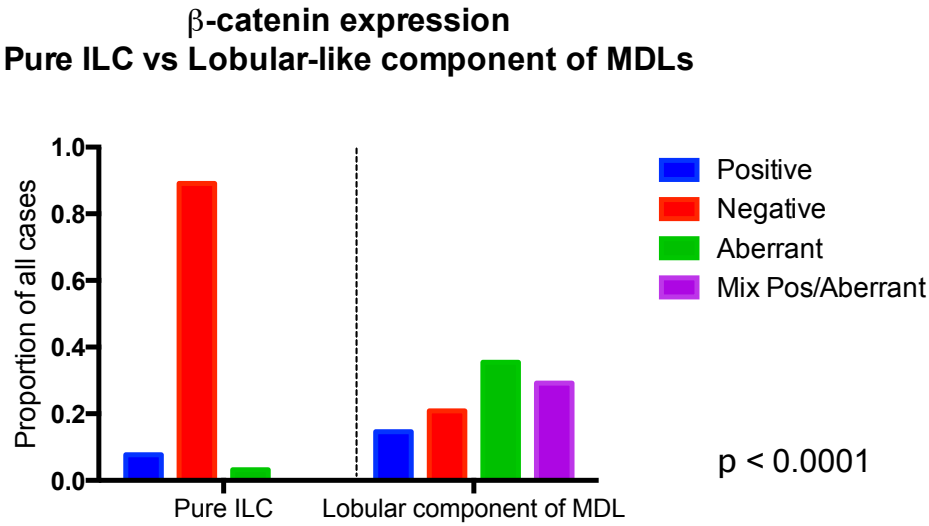
Figure 4.3: E-cadherin staining patterns in IC-NST and ILC.

Table 4.3: Expression of the E-cadherin adhesion complex in the different morphological components of mixed ductal lobular carcinomas

	DCIS (n=24)	LCIS/PLCIS (n=10)	Ductal (n=49)	Lobular (n=51)
E-cadherin; n (%)				
Pos	**22 (91.6)**	1 (10)	**38 (77.6)**	6 (11.8)
Neg	0	**7 (70)**	3 (6.1)	9 (17.6)
Aberrant	0	2 (20)	2 (4.1)	**22 (43.1)**
Mixed pos/ab	1 (4.2)	0	6 (12.2)	14 (27.5)
Mixed pos/neg	1 (4.2)	0	0	0
ND	0	0	0	0
β-catenin; n (%)				
Pos	**23 (95.8)**	1 (11.1)	**39 (79.6)**	7 (14.6)
Neg	0	**8 (88.9)**	3 (6.1)	10 (20.8)
Aberrant	1 (4.2)	0	0	**17 (35.4)**
Mixed pos/ab	0	0	3 (6.1)	12 (25.0)
Mixed pos/neg	0	0	0	2 (4.2)
ND	0	1	4 (8.2)	3
p120-catenin; n (%)				
Pos	**21 (95.5)**	2 (25.0)	**39 (84.8)**	8 (17.4)
Neg	0	1 (12.5)	0	0
Aberrant	0	**5 (62.5)**	3 (6.5)	**28 (60.9)**
Mixed pos/ab	1 (4.5)	0	4 (8.7)	10 (21.7)
Mixed pos/neg	0	0	0	0
ND	2	2	3	5

Typically, classic ILC shows complete loss of E-cadherin and β-catenin expression in up to 90% of cases, alongside cytoplasmic localisation of p120-catenin. We therefore compared the expression of the E-cadherin adhesion complex in a cohort of classic ILC (the WL TMA cohort of 148 ILC described in **Section 2.1**) to the lobular-like component of MDLs. As predicted from the small cohort assessed previously (Kutasovic, 2011), this larger cohort confirms that there was a significant difference in the localisation of all the adhesion complex markers (**Figure 4.4**; $p < 0.0001$). Most interestingly was the observation that E-cadherin protein was not lost in the lobular-like component of MDLs, but displayed fragmented membrane staining or was aberrantly expressed in the cytoplasm in most cases; the same was true for β-catenin and p120-catenin. It is likely therefore, that if disruption of the E-cadherin adhesion complex is contributing to the change in phenotype from ductal to lobular-like morphology, the mechanism of E-cadherin disruption in MDLs may be different to that seen in classical ILC, such as the genetic and epigenetic targeting of *CDH1* (reviewed in (McCart Reed and Kutasovic et al., 2015)).





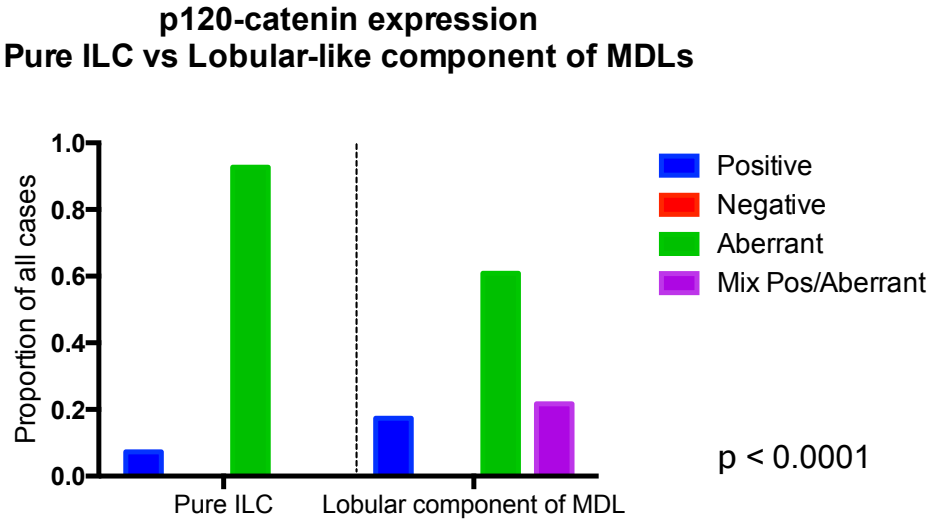


Figure 4.4: Comparison of the expression of the E-cadherin adhesion complex molecules (E-cadherin, β-catenin and p120-catenin) in pure ILC versus the lobular-like component of MDLs. Chi-square statistical test was used.

4.2.3. Does epithelial to mesenchymal transition play a role in the phenotypic switch from ductal to lobular growth pattern?

Epithelial to Mesenchymal Transition (EMT; **Section 1.8.3**) is essential for cell migration in embryogenesis, has been found to contribute to an increase in invasive capabilities *in vitro* and is speculated to therefore contribute to invasion and metastasis *in vivo*. In this process, E-cadherin expression is repressed by certain transcription factors such as TWIST, Snail and Slug. Cells undergo 'cadherin switching', down-regulating E-cadherin, expressing N-cadherin and also switching on the expression of a number of markers associated with mesenchymal cell differentiation, such as vimentin. High grade IC-NST of basal-like phenotype shows the strongest evidence of EMT in breast tissues (Aleskandarany et al., 2014). ILC and the lobular component of MDLs show a discohesive and infiltrative growth pattern and deregulated E-cadherin. We have shown that EMT rarely contributes to the invasive phenotype of ILC (McCart Reed and Kutasovic et al., manuscript under review, *J Pathol*). Here, we hypothesise that EMT may contribute to the transition of ductal to lobular-like growth pattern in MDLs.

A subset of MDL cases were assessed for the expression of the following EMT markers by IHC, N-cadherin (n = 8), SNAIL (n = 13) and Vimentin (n = 15). N-cadherin was negative in all cases; Vimentin and SNAIL were focally positive (in less than 1% of cells) in 4/15 (26.6%) and 2/13 (15.4%) of cases, respectively (**Figure 4.5**). Overall, protein analysis of these EMT markers suggests that EMT may not play a role in the progression from ductal to lobular morphology.

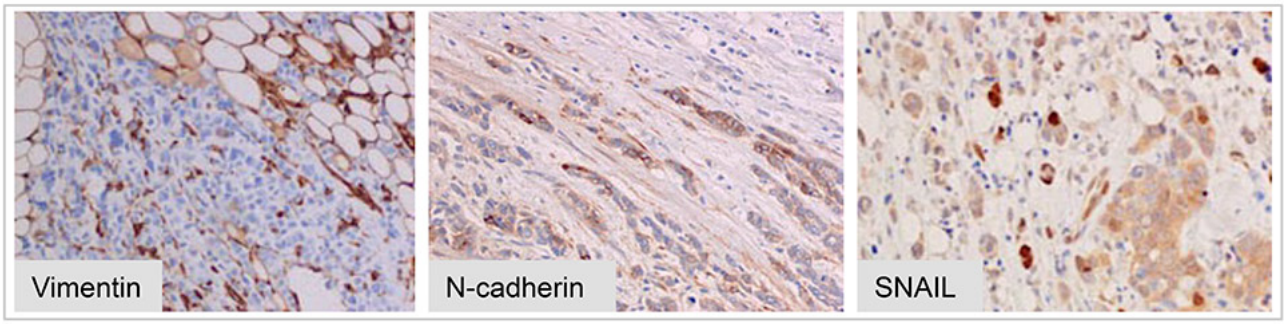


Figure 4.5: Assessment of expression of EMT markers Vimentin, N-Cadherin and SNAIL.

4.2.4. Meta-analysis of publically available gene expression and mutation data from mixed ductal lobular carcinomas

The TCGA has performed detailed molecular profiling on a large cohort of breast tumours (The Cancer Genome Atlas Network, 2012). There were 29 MDLs available according to the cBioportal user interface, where one can search for many different clinical variables alongside the publically available genomic data (Cerami et al., 2012, Gao et al., 2013). To verify the diagnosis of a MDL, we accessed the Cancer Digital Slide Archive for these 29 cases and reviewed the digital H&E slides for each samples. This also served to assess the tumour cellularity of the samples (**Section 2.12**). It is particularly important to account for tumour cellularity in this context, as the nucleic acid alterations will be representative of the most prominent cell type. There were 2 to 3 digital images provided for each case; generally 1 or 2 sections were frozen H&E sections and one was a FFPE H&E section. We have to make the assumption that the frozen H&E is representative of the tissue from which the nucleic acids were extracted. Based on this, it was found that 6 cases were enriched for the lobular component, 10 cases were enriched for the ductal component, and 13 tumours had admixed lobular and ductal components.

Meta-analysis of the gene expression of some key EMT genes (*CDH1, CDH2, TWIST1, SNAI1, VIM, ZEB1 and ZEB2*) showed that, for the most part, these genes are rarely altered in the TCGA MDL cohort (**Figure 4.6**). The case that has *CDH2* (N-cadherin) up-regulation doesn't have a hit to *CDH1*; hence it is unlikely to have undergone cadherin switching. Also, the case with concomitant over-expression of Vimentin and Zeb1, have no alteration in either *CDH1* or *CDH2*. Interestingly 19% (5/27 cases) had mRNA up-regulation of *ZEB1*. Although the numbers are small this frequency is quite high compared to both ER positive and ER negative IC-NST (*ZEB1* is up-regulated in 3% (18/532) and 0.5% (1/201) of cases, respectively) and ILC (*ZEB1* is up-regulated in 6% (9/155) of cases). Of the 5 cases with *ZEB1* up-regulation, 2 cases were predominately lobular, 2 cases were admixed, and 1 case was predominately ductal. It is also interesting to note that 3 of the 5 cases with *ZEB1* up-regulation also have up-regulation of *ZEB2* and these 3 cases show no other 'archetypal' EMT changes. The mRNA up-regulation was mutually exclusive relative to *CDH1* mutation; *ZEB1* is known to directly suppress expression of *CDH1* (Vannier et al., 2013, Wong et al., 2014). Unfortunately, these cases do not have reverse phase protein array (RPPA) data to assess whether the mRNA expression correlates with protein expression.

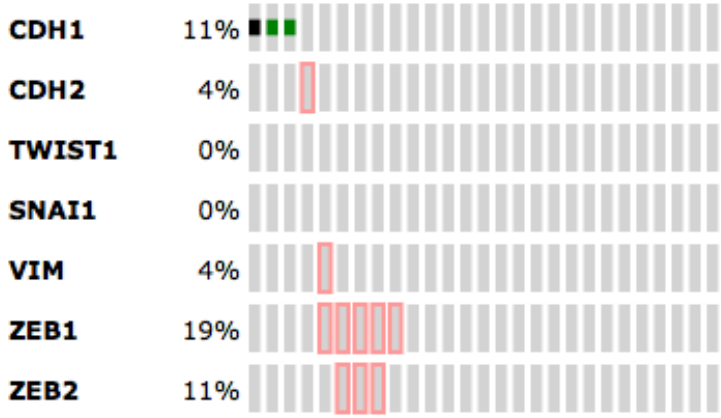


Figure 4.6: The gene expression of EMT related genes in 27 MDLs from the TCGA. The red box represents mRNA up-regulation. Each grey box represents an individual case. The black and green spots represent somatic mutation.

4.2.4.1. Somatic mutation analysis of TCGA mixed ductal lobular carcinomas

Large sequencing studies have found that there are few genes commonly mutated in breast cancer. **Section 2.11.1**, **Table 2.3**, summarises the most frequently altered genes in breast cancer. The cBioportal was utilised to assess the frequency of mutations of those genes in MDLs (data was available for 27 of the 29 cases), ILC, ER positive IC-NST and ER negative IC-NST. The MDL cases were separated out into the different components recorded from the digital slide archive (**Table 4.4**).

Although the cohort is small, there were some interesting observations to consider. Broadly, the average number of mutations was 41 (range 11 to 101) and it appears that MDLs show a mutation pattern similar to both ILC and ER positive IC-NST. MDLs have a high frequency of *PIK3CA* mutations similar to ILC and ER positive IC-NST (45.8% and 33.3% respectively *v.s.* 41.9% in the MDLs). The second most frequently mutated gene was *GATA3*, which is also the second most commonly mutated gene in ER positive IC-NST (20.8% *v.s.* 13.2%). Interestingly, the frequency of *TP53* mutations observed in MDLs (16.7%) was intermediate to that found in ILC (5.2%) and ER positive IC-NST (22.2%). This is in contrast to ER negative IC-NST were *TP53* is mutated in 65% of cases. *CDH1* mutations were more frequent in MDLs (12.5%) compared to both ER positive (2.3%) and ER negative IC-NST (1%); yet lower than observed in ILC (53.5%). Chi-square analysis was performed for all the mutated genes assessed and the difference in mutation frequency observed was significant in comparison to ER negative IC-NST (< = 0.00001)

and ILC ($p < 0.0003$), but not with ER positive IC-NST (p = n.s.). Three MDL cases did not have a mutation in any of these frequently mutated genes.

When comparing the mutation pattern between the morphology categories of MDL cases (**Section 2.12**), the main observation was that the admixed samples were most similar to samples that were predominately ductal component. There were only 6 samples that were lobular component enriched; yet the most striking difference was the higher frequency of *PIK3CA* mutations in the lobular component compared to the admixed and ductal components (**Table 4.4**).

Table 4.4: The most frequently mutated genes in MDL carcinomas compared to ER positive and ER negative IC-NST and ILC.

				MDL (n=24)			
	ER+ IC-NST (n=532)	ER- IC-NST (n=200)	ILC (n=155)	All samples	Admixed (n=10)	Ductal (n=9)	Lobular (n=5)
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
CDH1	12 (2.3)	2 (1.0)	83 (53.5)	3 (12.5)	1 (10.0)	1 (11.1)	1 (20.0)
FGFR2	6 (1.1)	2 (1.0)	1 (0.6)	1 (4.2)	1 (10.0)	0 (0)	0 (0)
FOXA1	9 (1.7)	1 (0.5)	10 (6.5)	3 (12.5)	1 (10.0)	0 (0)	2 (40.0)
GATA3	70 (13.2)	1 (0.5)	5 (3.2)	5 (20.8)	2 (20.0)	3 (33.3)	0 (0)
MLL3	40 (7.5)	8 (4.0)	12 (7.7)	1 (4.2)	0 (0)	1 (11.1)	0 (0)
MAP2K4	27 (5.1)	1 (0.5)	1 (0.6)	2 (8.3)	2 (20.0)	0 (0)	0 (0)
MAP3K1	50 (9.4)	3 (1.5)	8 (5.2)	2 (8.3)	1 (10.0)	1 (11.1)	0 (0)
NCOR1	22 (4.1)	6 (3.0)	7 (4.5)	1 (4.2)	0 (0)	1 (11.1)	0 (0)
NF1	11 (2.1)	6 (3.0)	5 (3.2)	1 (4.2)	1 (10.0)	0 (0)	0 (0)
PIK3CA	177 (33.3)	25 (12.5)	65 (41.9)	11 (45.8)	4 (40.0)	3 (33.3)	4 (80.0)
PTEN	16 (3.0)	4 (2.0)	9 (5.8)	0 (0)	0 (0)	0 (0)	0 (0)
RB1	9 (1.7)	6 (3.0)	2 (1.3)	1 (4.2)	0 (0)	1 (11.1)	0 (0)
RUNX1	13 (2.4)	1 (0.5)	13 (8.4)	1 (4.2)	1 (10.0)	0 (0)	0 (0)
TBX3	13 (2.4)	2 (1.0)	8 (5.2)	0 (0)	0 (0)	0 (0)	0 (0)
TP53	118 (22.2)	130 (65.0)	8 (5.2)	4 (16.7)	1 (10.0)	3 (33.3)	0 (0)

Highlighted in grey are the most frequently mutated genes within each subtype (mutated in over 10% of cases). Mutations in the MDL cases were pooled and then separated out based on the predominant morphological component as determined by assessing the cancer digital slide archive.

A discovery approach to variant analysis was also taken, whereby all of the mutation annotated format (MAF) files of the exome sequencing data was downloaded for each of the 27 MDL cases with data, to examine if there were other interesting mutations that may have an interesting role. All non-silent coding variations were recorded for each case and

each case was assessed individually. We were particularly interested to see if there were mutations present in genes encoding proteins involved in E-cadherin cell-cell adhesion, however, there were no reported mutations in any of the genes encoding the E-cadherin binding partners α, β, γ, or p120-catenin. However, mutations were found in other cadherin family members *CDH7*, *CDH9* and *CDH20* and in cadherin related proteins, *CDHR1* and *CDHR2*, which together with *CDH1* were all mutually exclusive (**Figure 4.7**). Mutations were also detected in protocadherins *PCDH9, PCDHA2, PCDHA5, PCDHAC1, PCDHB6, PCDHB14*, and *PCDHGA9*. When overlaying copy number and gene expression data in cBioportal, we found individual cases with increased expression of *CDHR2*, *PCDH9, PCDHA2, PCDHA5, PCDHB14*, *PCDHGA9* or *PCDHGB3*; yet none of the cadherin family members showed gene deletion or down-regulation in the MDLs. The functional consequences of these mutations are unknown.

The following genes were noted as being recurrently mutated (identified in more than 1 case). Mutations were found in genes involved in cytoskeleton regulation: 3 cases had *FMN2* (Formin 2; involved in actin cytoskeleton assembly) mutation, while 5 cases had gene amplification; 2 cases had *MACF1* (microtubule-actin crosslinking factor 1; involved in crosslinking many proteins of the cytoskeleton) mutations and 2 cases had gene up-regulation; and 2 cases had *KIF21B* (kinesin family member 21B; microtubule based motor protein involved in intracellular transport of membranous organelles) mutations and 5 cases had gene amplification which may be contributing to invasive capacity of the cancer cells. Mutations in genes involved in WNT signalling (*WNT16*, *WNT7A*, *WNT3A*) were also observed. These genes are involved in pathways associated with carcinogenesis and embryonic development, specifically in cell division and cell migration. β-catenin is also important in canonical Wnt signalling; in addition to its role in cell-cell adhesion, β-catenin can translocate to the nucleus, acting as a transcriptional co-activator for various Wnt signalling target genes (such as *MYC, CCND1, CDH1* (Klaus and Birchmeier, 2008).

The FOX (forkhead box) protein, *FOXA1* was identified as a frequently mutated gene in ILC (6.5%) and MDLs (12.5%). Mutations in other FOX proteins were also identified in individual cases: *FOXM1* (cell cycle progression) and *FOXP1* (transcription factor essential for organ development). Since these genes are essential for the regulation of many biological processes, their alterations may be important in tumour progression in these cases.

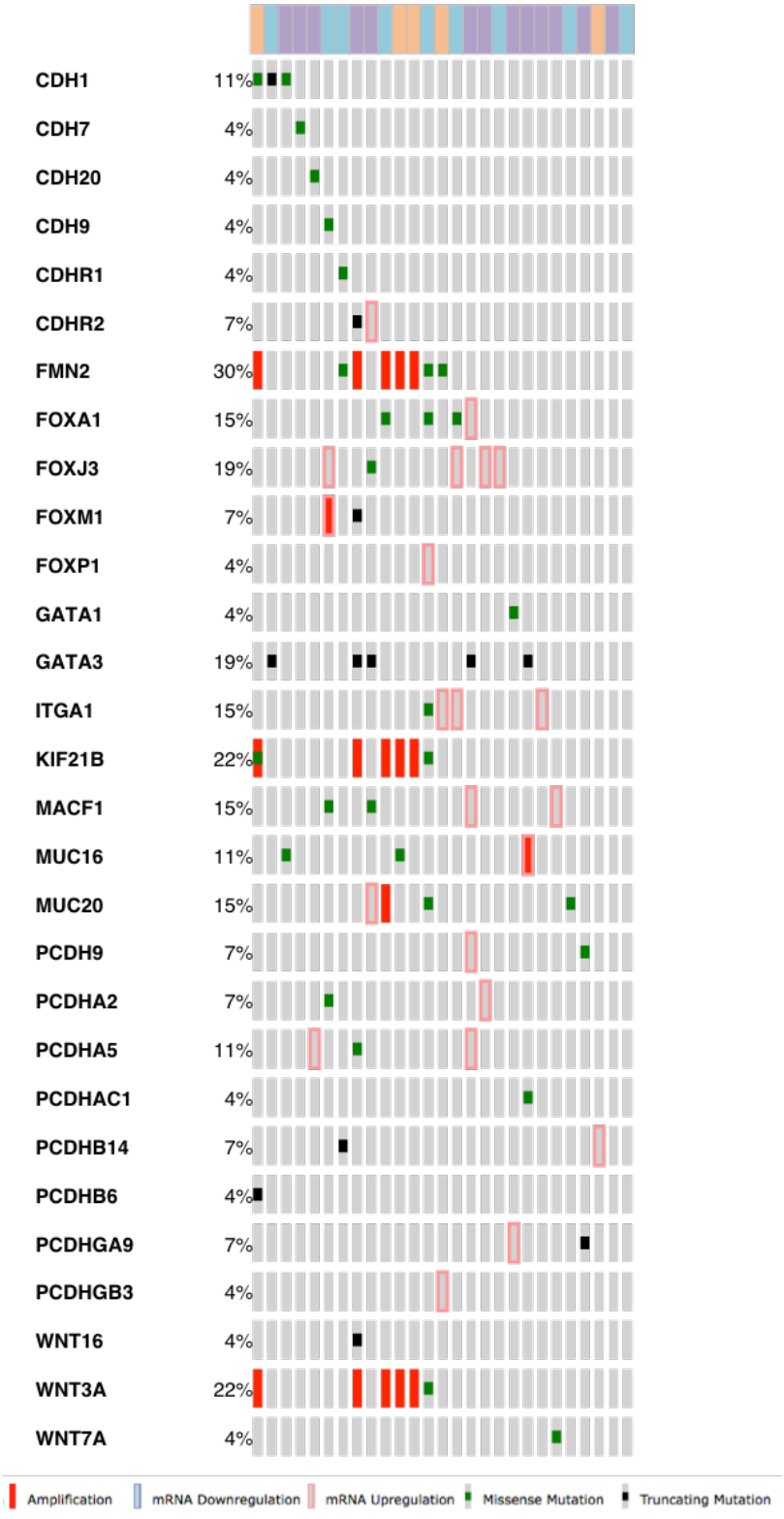


Figure 4.7: Other mutated genes of interest in MDLs. Top panel: Orange, Lobular enriched; Blue, Ductal enriched; Purple, Admixed.

4.2.5. Investigating the molecular evolution of mixed ductal lobular carcinomas

To give insight into the clonal nature of this tumour type, it is necessary to investigate the different morphological components independently using molecular techniques. To this

end, enrichment for specific cell types of interest to a high degree of purity was performed using laser capture micro-dissection (LCM) or fine needle micro-dissection. **Table 4.5** describes the morphological and phenotypic characteristics of each case, the morphological components, and the molecular analyses performed. Tumours were needle macro-dissected if the specific areas of interest were quite separate from each other, while other cases were micro-dissected using LCM for greater precision in separating tumour areas (**Section 2.7.1 and 2.7.2**).

Table 4.5: MDL cases subjected to genomic analysis.

Case ID	Component	Grade	ER	PR	HER2	E-cad	cCGH	Exome
MDL1#	Ductal	2	+	+	-	Pos/Ab	✓	
	Lobular	2				Pos/Ab	✓	
	DCIS	2				Pos	✓	
	LCIS	clas				Ab	✓	
MDL2#	Ductal	2	+	+	+	Pos	✓	
	Lobular	2				Pos/Ab	✓	
	DCIS	3				Pos	✓	
	LCIS	n/a				Neg	✓	
MDL3#	Ductal	3	+	+	+	Pos	✓	
	Lobular	n/a				Ab	✓	
	DCIS	3				Pos	✓	
MDL4*	Ductal	3	+	+	+	Pos	✓	✓
	Lobular	3				Ab	✓	✓
	DCIS	3				Pos	✓	✗
	Normal - blood					-		✓
MDL5*	Ductal	2	+	+	-	Pos		✓
	Lobular	2				Pos/Ab		✓
	DCIS	2				Pos		✓
	LCIS	clas				-		✗
	Normal LN					-		✓
MDL6#	Ductal	2	+	-	+	Pos		✓
	Lobular	3	+	-	+	Neg		✓
	PLCIS	3				Neg		✓
	DCIS	3				Pos/Neg		✓
	LN metastasis					Neg		✗
	Normal LN					-		✓
MDL7*	T1 – IC-NST	3	+	+	-	Pos		✓
	T1 – DCIS	2				Pos		✗
	T2 – ILC	2	+	+	-	Pos/Ab		✓
	T3 - MDL - ductal	2				Pos		✓
	T3 - MDL - lobular					Pos/Ab		✓
	T3 – DCIS	2				Pos		✓
	T4 – MDL - ductal	2				Pos		✓
	T4 – MDL - lobular					Pos/Ab		✓
	T4 – DCIS	2				Pos		✓
	Normal LN					-		✓

*Macro-dissected, #Laser capture micro-dissected. cCGH: Chromosomal CGH - performed previously (Kutasovic, 2011). LN; Lymph node. MDL7: T = tumour (described in more detail below in **Section 4.2.5.4**). Clas = classic LCIS type.

Four MDLs cases previously underwent chromosomal CGH (cCGH) analysis as described earlier (**Section 4.1.1**). Here we have attempted exome sequencing on four MDL cases, including one that underwent cCGH. These cases were chosen because of their interesting morphological features and the fact that each component was large enough to yield sufficient DNA for analysis. **Figure 4.8** provides a representative image of a case that was micro-dissected for different morphological regions within the same specimen with the guidance of E-cadherin staining. Whole exome sequencing was performed (as described in **Section 2.11.2**) to uncover the clonal evolution in these MDL cases and potentially reveal mechanisms driving the change in morphology. **Appendix Table 4.1** summaries the sequencing metrics for all cases. Below are detailed descriptions of each case and their preliminary analysis of the genomic data.

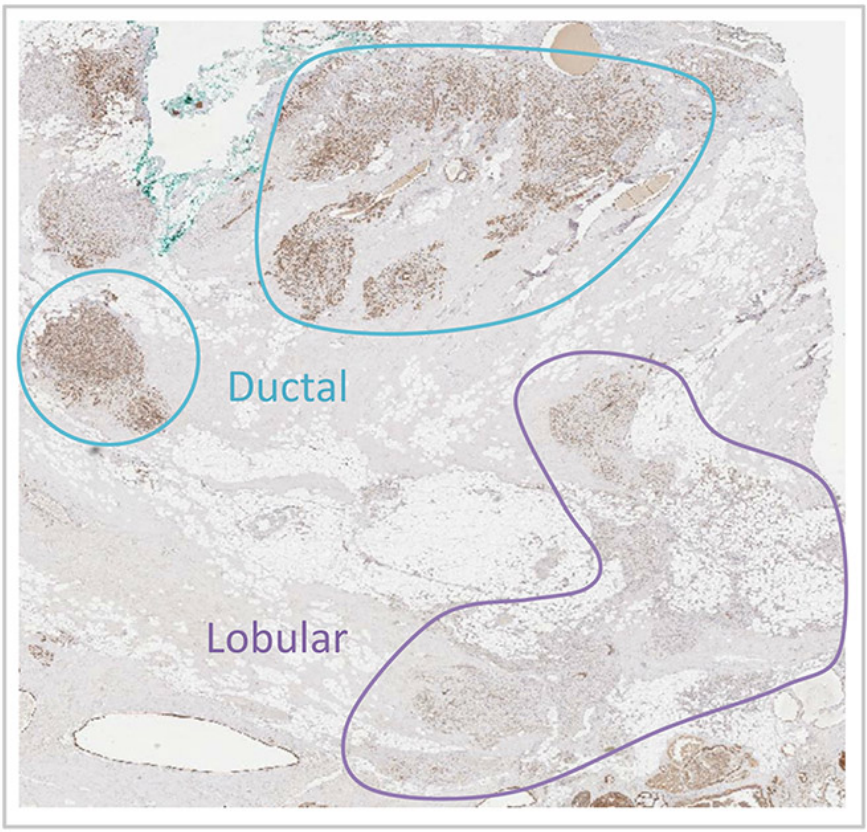


Figure 4.8: E-cadherin staining was used alongside morphological assessment to guide micro-dissection of each component prior to DNA extractions.

4.2.5.1. MDL4

A 79-year-old patient was diagnosed with a MDL carcinoma that was 20 mm in size, histological grade 3 and lymph node negative. The lobular component was of the pleomorphic variant and high-grade DCIS was present, but no LCIS was identified in the histological examination. The tumour was positive for ER and PR, and HER2 was amplified. The E-cadherin adhesion complex was positive in the DCIS and ductal component and was aberrantly localised in the lobular component (**Figure 4.9**). This case was analysed for the expression of EMT markers and was negative for vimentin, and displayed cytoplasmic localisation of Snail and N-cadherin. Positive CK8/18 staining emphasises the epithelial nature of the lobular cells.
Interestingly, from a morphological point of view, there where areas within the tumour of what appears to be a 'transition zone'- a region with a clear change from ductal to lobular-like morphology (**Figure 4.9**) that coincides with aberrant localisation of the E-cadherin adhesion complex.

Chromosomal CGH data showed that the two morphological components shared many copy number alterations (**Table 4.1** (Kutasovic, 2011)). As part of my honours project, I Sanger sequenced the exons of *CDH1* in both morphological components and no mutation was found. Both the invasive ductal and pleomorphic lobular components were macro-dissected using a fine needle and the DNA extracted was sequenced for the whole exome (together with normal DNA, **Sections 2.7** and **2.11.2**). Unfortunately the DCIS was too small for this analysis. Preliminary analysis of the exome sequencing data found that there were 77 variants shared between both the ductal and lobular components, supporting that both components are derived from a common ancestor. There were 19 and 68 private variants in the ductal and lobular components, respectively (**Figure 4.10 A**, including both silent and non-silent mutations). There were no obvious mutations in genes associated with cell adhesion (such as E-cadherin adhesion complex) in the lobular component. Of interest however, was a mutation in *ESR1* (encoding oestrogen receptor alpha) in the lobular component, and *HRAS* in the ductal component. As both components show more similar mutation profiles than different, and the lobular component shows many unique variants, we can hypothesise that the lobular component diverged late in the evolution of this tumour (**Figure 4.10 B**). Since the DCIS was not available for analysis, we hypothesise that the invasive counterparts of this tumour arose via a ductal-like pathway of progression, where the ductal and lobular components have arisen from a common DCIS.

Targeted mutation validation may be possible in the DCIS component to confirm this theory.

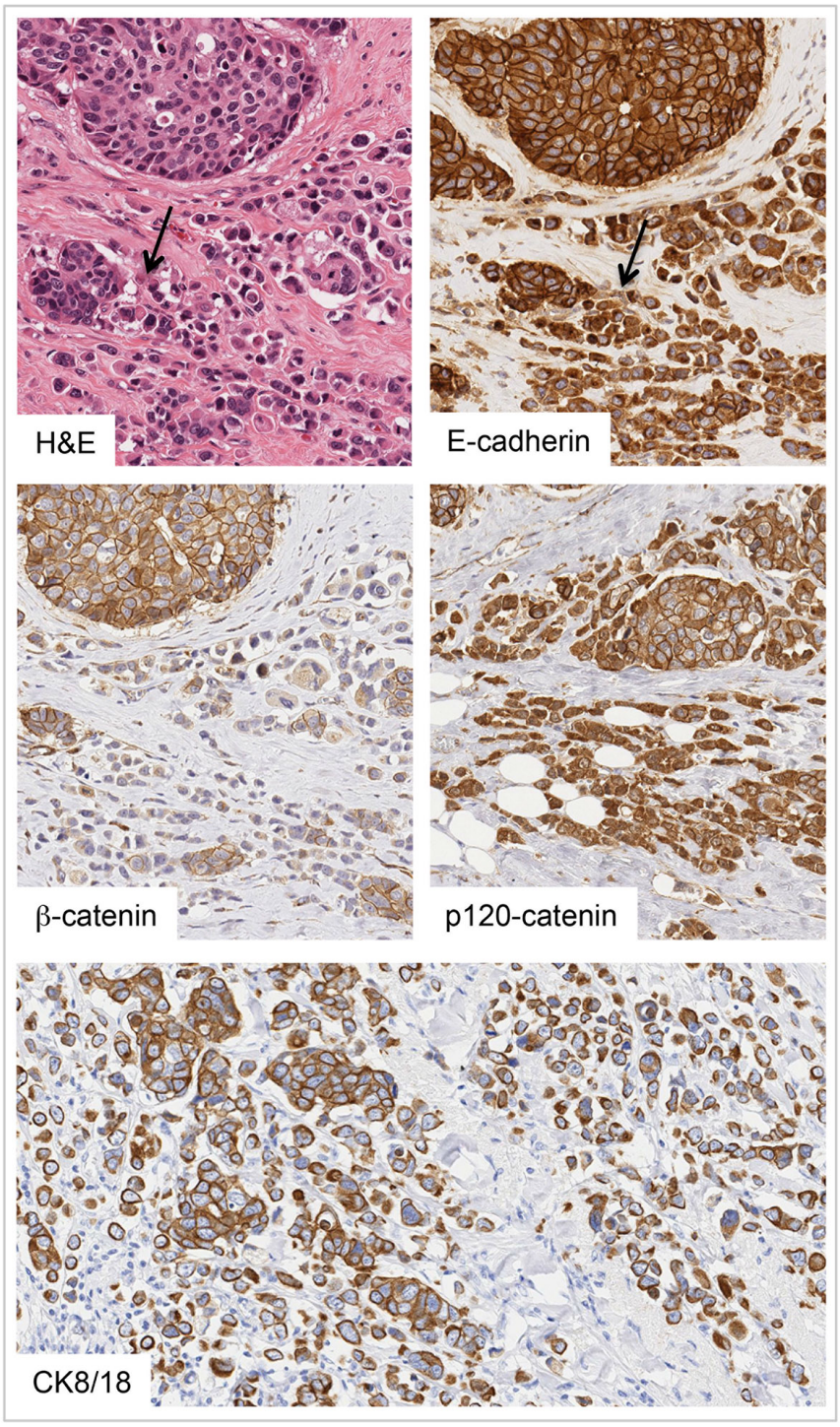


Figure 4.9: Representative images of both ductal (top left of the image) and lobular components (areas of single cells) of MDL4. A possible 'transition zone' from ductal to lobular morphology is marked by the arrow that appears to coincide with cytoplasmic E-cadherin localisation.

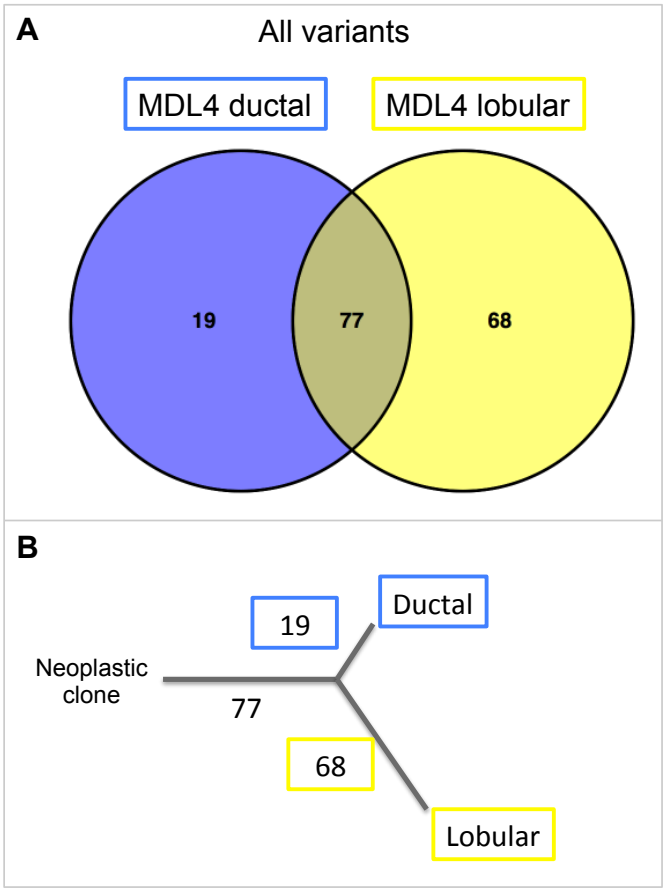


Figure 4.10: A) Summary of all the silent and non-silent variants across both morphological components. B) A hypothetical evolutionary tree. We hypothesise that the lobular component diverged late during the evolution of this tumour and arose via a "ductal-like" pathway.

4.2.5.2. MDL5

This 52-year-old patient presented with a 56 mm, grade 2 MDL carcinoma with co-existing DCIS and LCIS. This tumour was ER and PR positive, HER2 negative and had regional metastasis to the lymph nodes that was of ductal morphology. **Figure 4.11** shows the differential E-cadherin binding complex expression between the morphological components; the ductal component and DCIS were positive for the E-cadherin complex and the lobular component had aberrant expression of the E-cadherin complex. Unfortunately there was no LCIS left in the tissue sample to assess the expression of E-cadherin complex or from which to extract DNA. For exome sequencing, areas of E-cadherin positive ductal morphology, E-cadherin negative lobular morphology and DCIS were separately micro-dissected using a needle. There were 11 shared variants observed between the DCIS, E-cadherin positive and E-cadherin negative morphological regions sequenced (**Figure 4.12 A**). Unique to each component were 13 variants in the DCIS, 11

variants in the E-cadherin positive and 7 in the E-cadherin negative regions. Of the 7 genes mutated in the E-cadherin negative component (*ATH9B*, *GIGYF2*, *MRPL27*, *PREPL*, *TDG*, *VPS41*, *WDR33*), none were related to cell adhesion or provide an obvious explanation for a change in morphology or growth pattern. Overall there is evidence that all morphological variants were derived from the same neoplastic clone. However, divergence may have occurred earlier as there are both DCIS and LCIS present in this case (**Figure 4.12 B**).

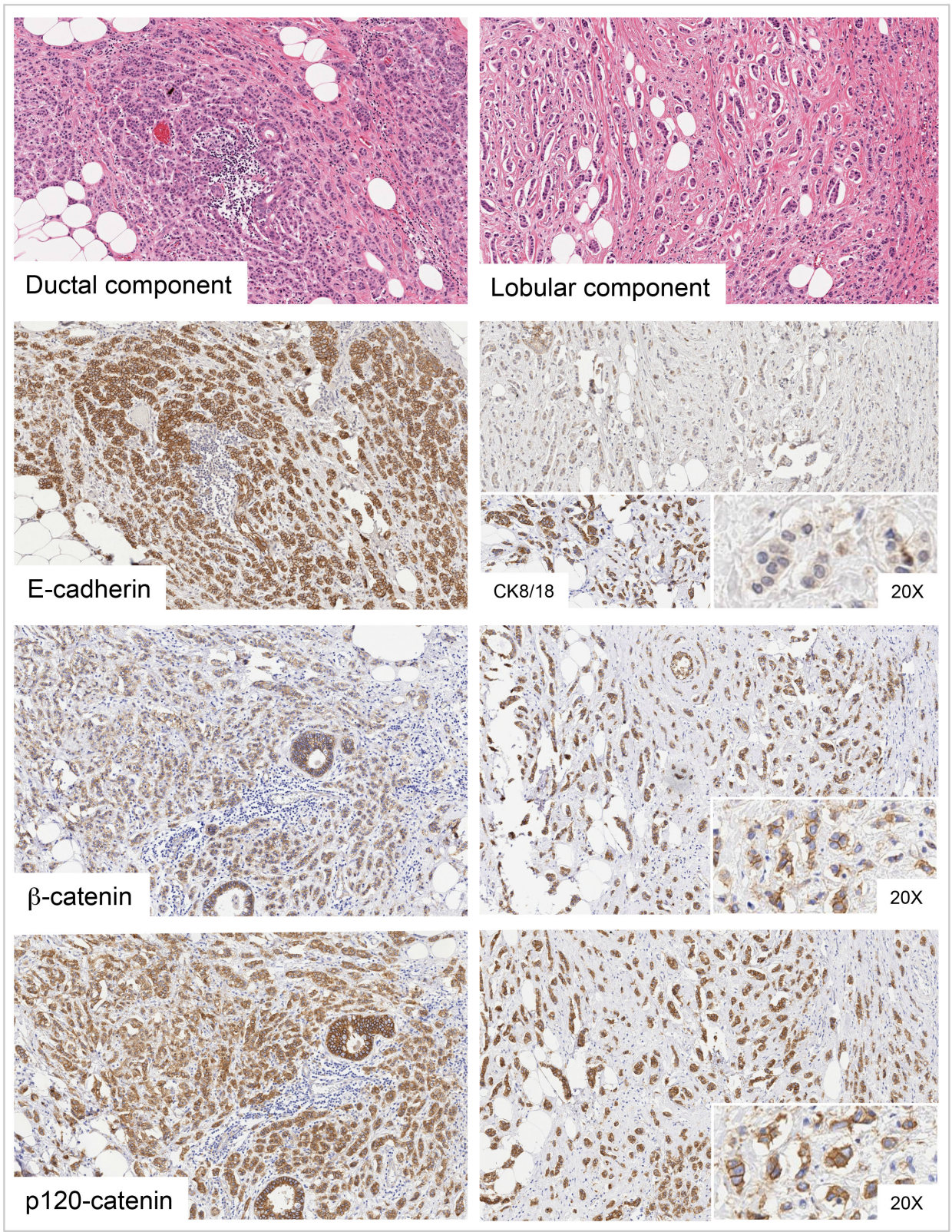


Figure 4.11: MDL5. Representative areas of ductal and lobular morphology with staining for E-cadherin, β-catenin and p120-catenin. 20X insets show the localisation of E-cadherin, β-catenin and p120 catenin in the lobular component.

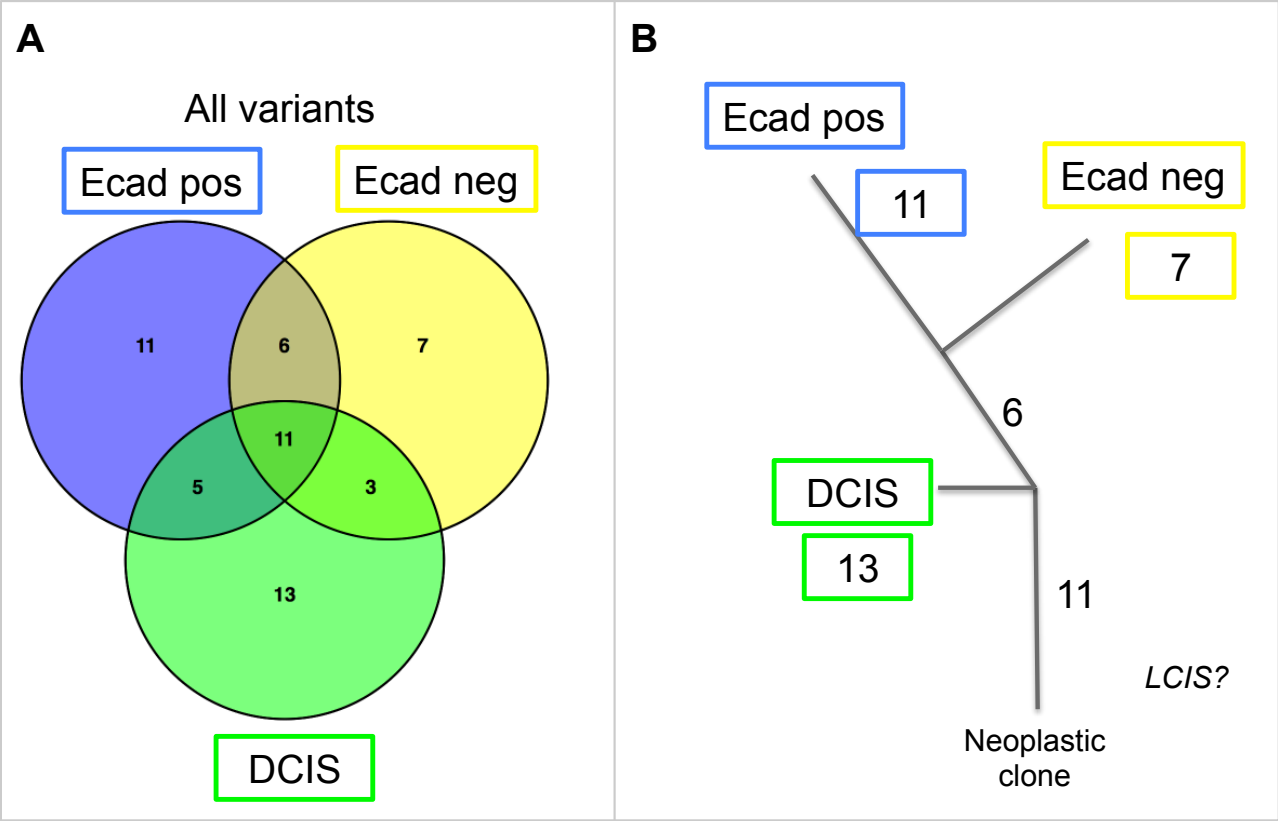


Figure 4.12: A) All silent and non-silent variants shared between the E-cadherin positive ductal component, E-cadherin negative lobular component and DCIS of MDL5. B) Hypothetical evolutionary tree. Since the LCIS was unavailable for analysis we are unable to determine the exact clonal relationships of all components.

4.2.5.3. MDL6

This 76-year-old patient was diagnosed with a MDL carcinoma that was composed of a grade 3 pleomorphic lobular component that was 40 mm alongside a grade 2 ductal component that was 14 mm. The tumour also had high-grade DCIS and PLCIS, and the regional lymph nodes were infiltrated by the PLC. Both the lobular and ductal components were positive for ER and HER2, however there was more ER positive cells (30% *v.s.* 5%) and HER2 was more highly amplified (24 copies *v.s.* 12) in the lobular component compared to the ductal component. Ki67 staining was also performed with 37% of the lobular component, and 19% of ductal component positive, indicating that the lobular component is more proliferative than the ductal component. All of the aforementioned data was provided by a very detailed pathology report. As displayed in **Figure 4.13**, each invasive and pre-invasive component has very distinct morphology and displays differential E-cadherin adhesion complex expression. The lobular component and PLCIS regions were negative for E-cadherin and β-catenin, while p120-catenin was located in the

cytoplasm. The ductal and DCIS components display membrane positivity for all adhesion complex members. Interestingly, some foci of DCIS exhibited heterogeneous E-cadherin expression with positive and negative cells co-existing in the same lesion, and could represent pagetoid spread of LCIS cells growing upward in the duct (see **Figure 4.13**, bottom panel).

This fascinating case was micro-dissected using the Leica laser capture micro-dissection system to separate out near pure populations of cells from the lobular, ductal, DCIS and PLCIS components. Exome sequencing highlighted that 56 variants were shared between all 4 components of this case, emphasising that all morphological variants within this case are clonally related (**Figure 4.14 A**). Mutated genes of interest that are shared include known tumour suppressors *BRCA2* (frame shift deletion) and *SMAD4* (missense mutation 1082G>A, Arg361His). A frame shift deletion in *TBX3* is present in all components except the DCIS. The invasive components both had the largest number of private mutations (64 in the ductal component and 266 in the lobular component). Of important note, the PLCIS and lobular components both shared a mutation in *CDH1*, which is likely to account for the lack of E-cadherin expression. The mutation in *CDH1* is predicted to cause a premature stop codon (1828C>T, Gln610*). We hypothesise that this case may be an example of early clonal divergence, due to the presence of both early pre-invasive DCIS and PLCIS lesions. (**Figure 4.14 B**)

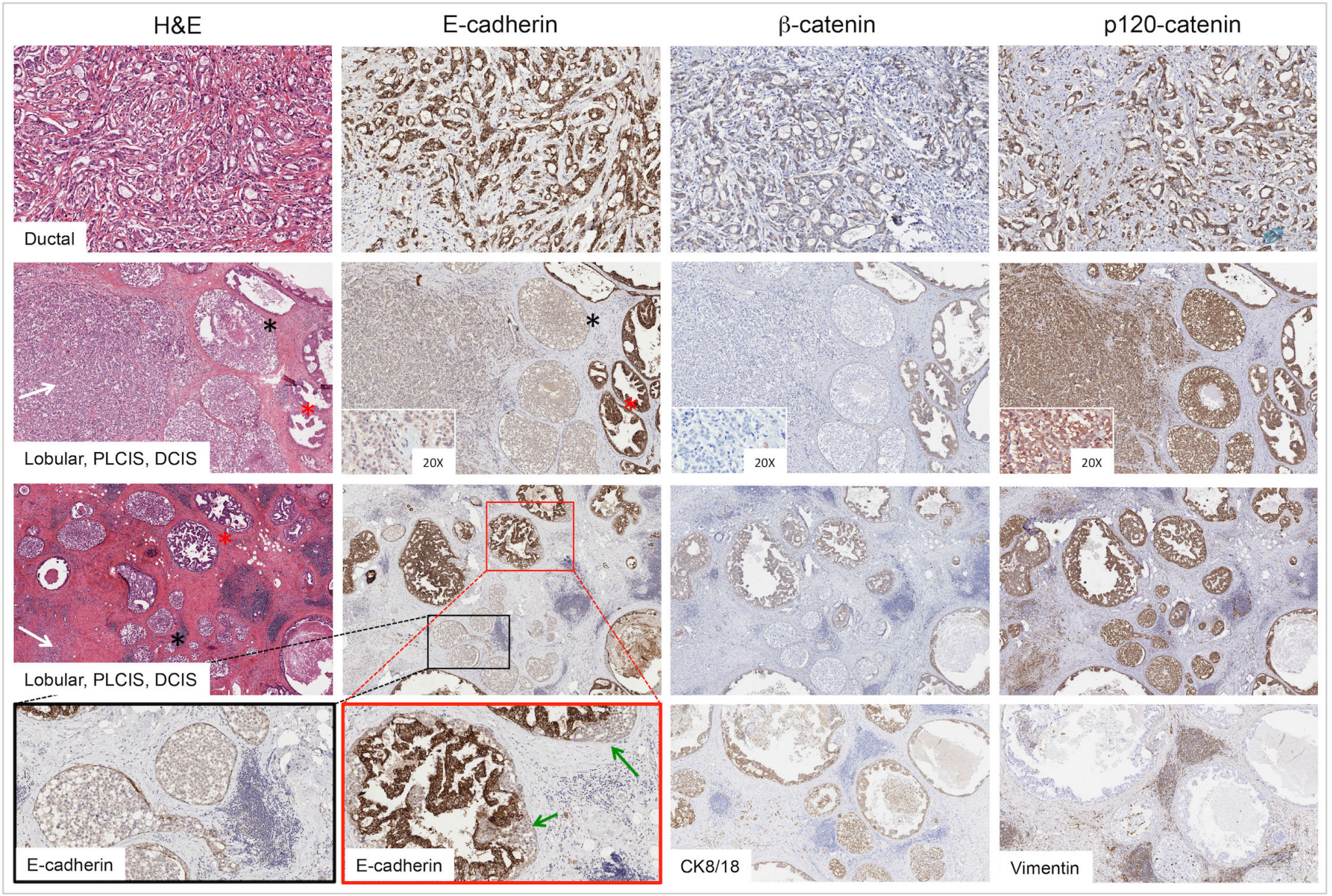

H&E
E-cadherin
β-catenin
p120-catenin
Ductal
Lobular, PLCIS, DCIS
20X
20X
20X
Lobular, PLCIS, DCIS
E-cadherin
E-cadherin
CK8/18
Vimentin

Figure 4.13 MDL6. Representative areas of the ductal component, pleomorphic lobular component (white arrow), PLCIS (black asterisk) and DCIS (red asterisk). 20X insets show the localisation of E-cadherin, β-catenin and p120 catenin in the lobular component. Black box: PLCIS that is E-cadherin negative. Red box: DCIS that is E-cadherin positive with areas of E-cadherin negative within the DCIS (green arrow). CK8/18 is positive in 100% of the tumour cells and vimentin is negative.

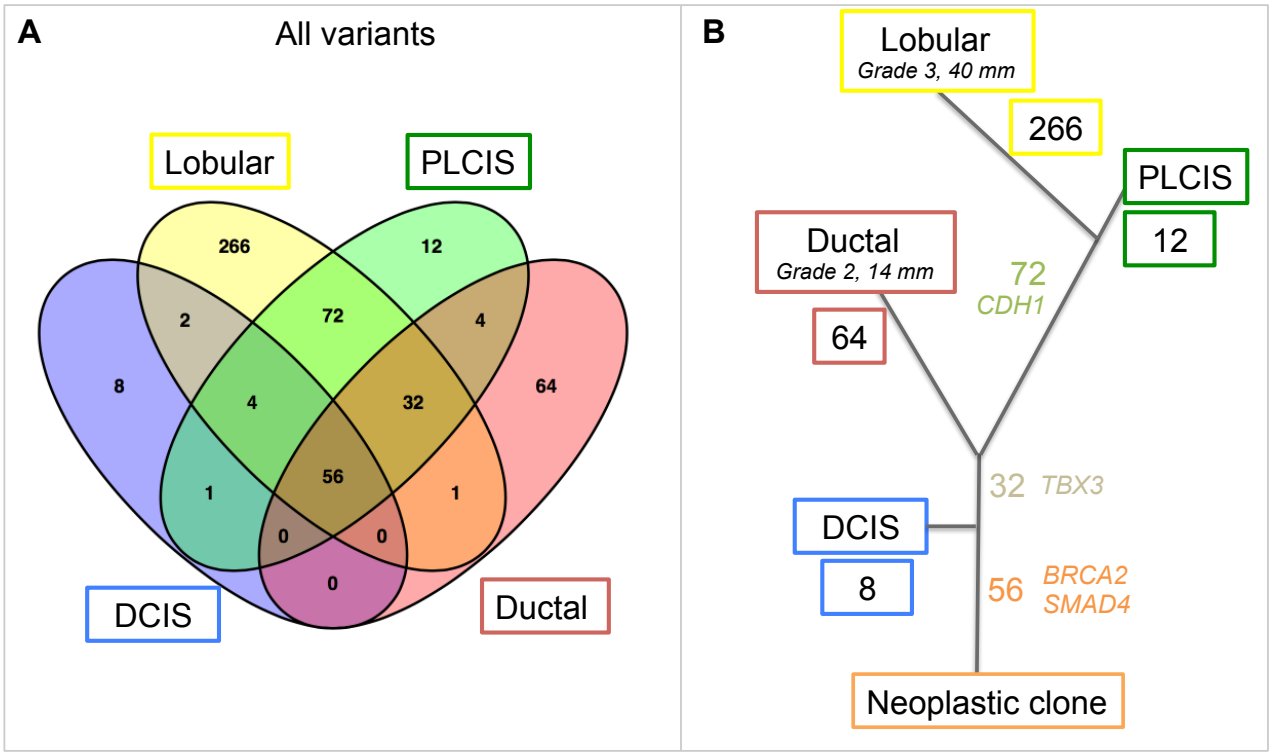


Figure 4.14: All silent and non-silent variants identified in the four morphological components. B) The construction of a hypothetical evolutionary tree assessing the pathway of tumour progression in this case.

4.2.5.4. MDL7

A 54-year-old patient underwent a mastectomy for multi-focal breast cancer. Macro- and microscopic analysis revealed 4 individual tumours of different histological types. The largest lesion was a 26 mm grade 3 IC-NST. The second lesion was a 21 mm grade 2 ILC. The other 2 lesions were grade 2 MDL carcinomas, 12 mm and 7 mm in size, respectively. **Figure 4.15** is a map outlining where each lesion was found with respect to each other, and the distances between each one. All 4 lesions were associated with DCIS, were ER positive, PR positive and HER2 negative, and no regional lymph nodes were involved at diagnosis. Each lesion was assessed for the expression of the E-cadherin adhesion complex. Overall, the IC-NST, DCIS and ductal components of the two MDL tumours were

positive for the adhesion complex members. Interestingly, the ILC and lobular-like components of the MDL tumours displayed a mix of positive and aberrant staining (**Figure 4.16**). All tumours were positive for the epithelial cell marker CK8/18 and negative for vimentin (data not shown), suggesting that EMT does not play a role in the transition from ductal to lobular phenotype. Each lesion was macro-dissected with a fine needle (see **Table 4.4**) and DNA extracted with a total of 8 tumour areas undergoing exome sequencing (**Sections 2.7** and **2.11.2**). The DCIS co-occurring with the IC-NST did not undergo exome sequencing due to low DNA concentration. This case is quite complex in that there are multiple pre invasive lesions and multiple growth patterns across multiple lesions throughout the breast (**Figure 4.15**); therefore understanding the natural history of these tumours is not so straightforward. Greater sequencing coverage is necessary to confidently call mutations and understand evolutionary relationships, therefore the samples are being re-sequenced and the data analysis is on going. However, preliminary analysis has revealed that similar to MDL5, there were no obvious mutations in genes associated with cell adhesion, such as *CDH1*.

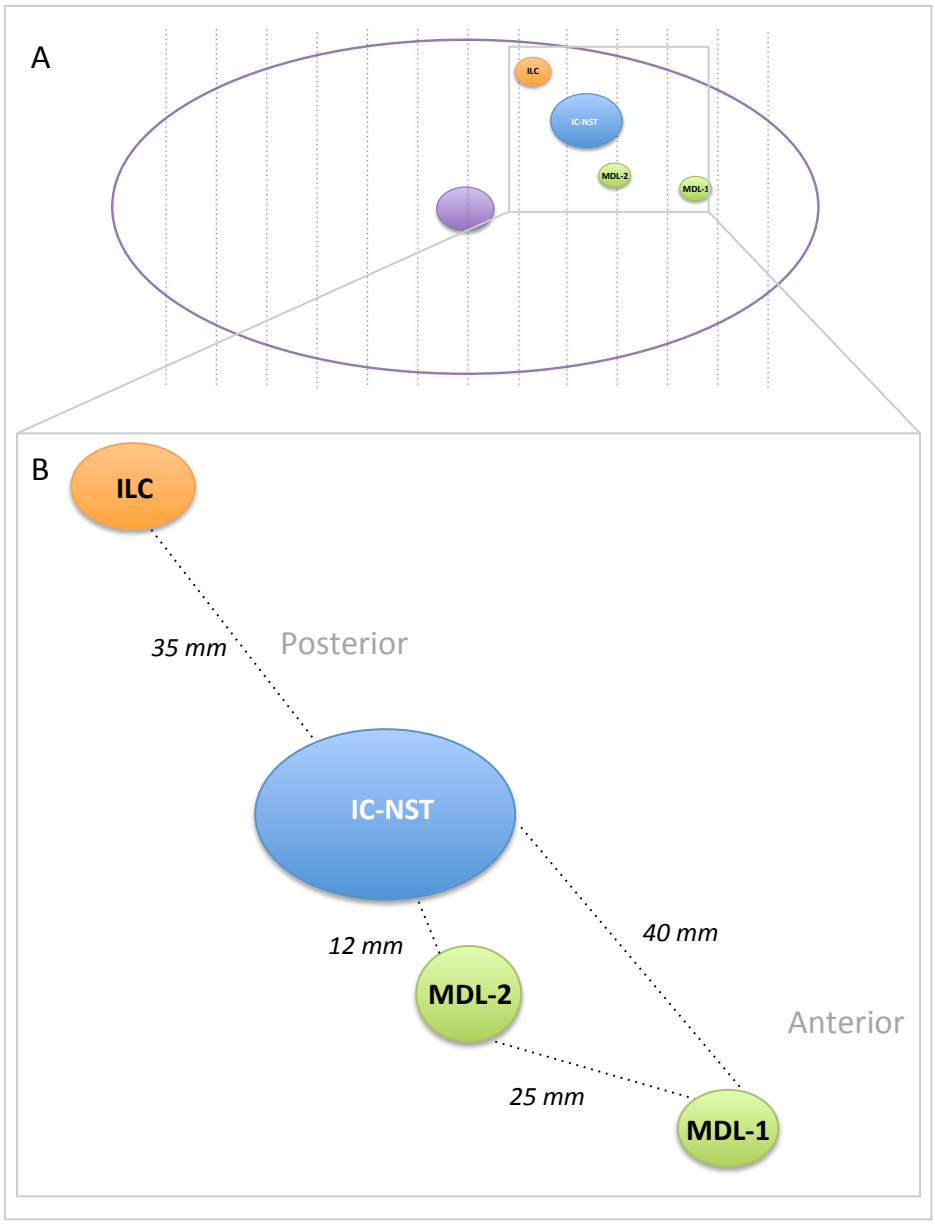


Figure 4.15 Map of all tumour foci in the left upper outer quadrant of case MDL7 (A). The dotted lines indicate the distance between each lesion (B).

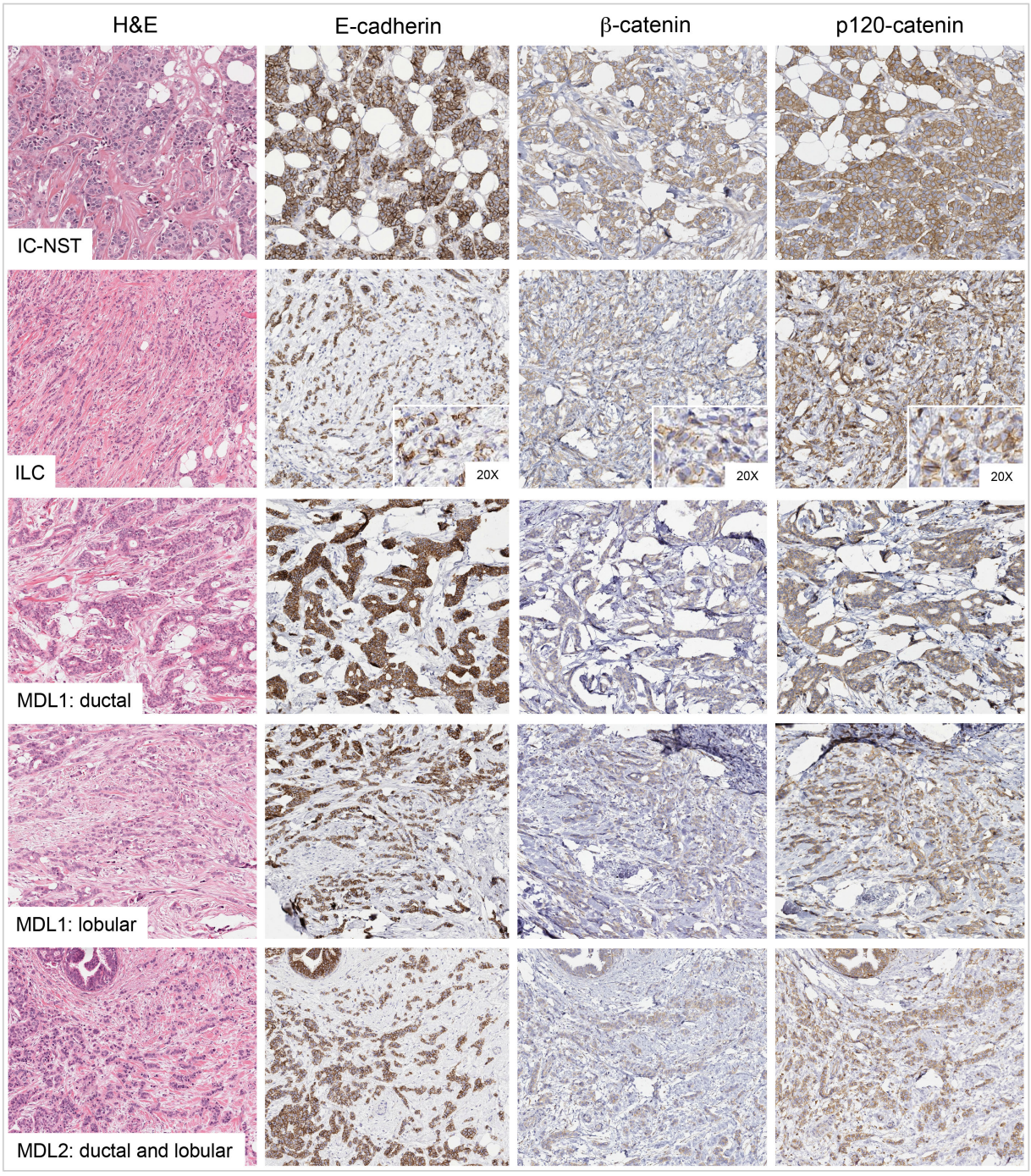


Figure 4.16 Representative images of the 4 tumour foci in case MDL7. 20X insets show the localisation of E-cadherin, β-catenin and p120 catenin in the ILC of the second lesion, which is representative of the lobular components in the two MDL lesions.

4.2.6. Identification of driver mutations and pathway analysis

All sequencing data was analysed using IntOGen ((Gundem et al., 2010), **Section 2.12**) to discover which mutations may be drivers in these MDLs, and also assess whether there are recurrently altered functional pathways. All 17 samples from the 4 cases were analysed and **Table 4.6** describes the most frequently mutated known driver genes,

candidate driver genes and the top 10 driver pathways in the cohort. The most frequently mutated genes were those identified in MDL6 (*BRCA2* and *SMAD4*) and is therefore likely to be enriched for cases with the most samples. The pathway analysis revealed that the most frequent pathways altered in over 80% of the samples were gap junction signalling, Gonadotropin-releasing hormone (GnRH) signalling and MAPK signalling.

Table 4.6: Summary of the IntOGen analysis to identify driving genes and pathways across the cohort of 4 MDL cases.

Driver genes in the cohort	Mutation frequency	IntOGen
RICTOR	0.529	
MAP3K2	0.471	
TBX3	0.471	Driver
MOK	0.412	
MEGF6	0.412	
MAD1L1	0.412	
MYLIP	0.412	
DYNC1H1	0.235	
SMAD4	0.235	Driver
Candidate driver genes	Number of samples; case	Impact
BRCA2	All samples; MDL6	High
GPS2	All samples; MDL6	High
SMAD4	All samples; MDL6	Medium
AHNAK	All samples; MDL6	Medium
TBX3	3/4 samples; MDL6	High
TBX3	All samples; MDL5	High
TP53	All samples; MDL5	High
SP3B1	All samples; MDL5	High
CDH1	2/4 samples; MDL6	High
Driver pathways	Found/studied	Mutation frequency
Gap junction	15/17	0.882
GnRH signalling pathway	15/17	0.882
MAPK signalling pathway	14/17	0.824
Cell cycle	12/17	0.706
Wnt signalling pathway	9/17	0.529
Osteoclast differentiation	8/17	0.471
Pathways in cancer	8/17	0.471
Hepatitis B	8/17	0.471
Glutamatergic synapse	7/17	0.412
Transcriptional misregulation in cancer	6/17	0.353

Found/studied = number of samples in the whole study with the alteration. Impact = the adverse impact of the mutation on the protein function

4.3. Discussion

This study investigated the clinical, pathology and molecular features of a large cohort of mixed ductal lobular carcinomas. The findings corroborate that MDLs are a distinct morphological entity (Rakha et al., 2009, Arps et al., 2013) with unique biological features that share some similarities to both IC-NST and ILC. Our findings are in agreement with previous studies stating that MDLs were mostly grade 2 (58%); were more frequently associated with DCIS than LCIS; were frequently lymph node positive (68.3%) and were more frequently ER positive compared to IC-NST ((Rakha et al., 2009, Arps et al., 2013) **Section 4.2.1**).

The current multistep model of progression considers that an IC-NST arises from a DCIS, while an ILC arises from a LCIS, via a low-grade or high-grade pathway of tumourigenesis (**Section 1.8**). The evolution of MDLs is less clear cut and less well studied. DCIS was present in 87.9% of our MDL cases, similar to that reported by Rakha et al. (89%; (Rakha et al., 2009)), where as LCIS was present in only 39% of cases (27% of cases had both DCIS and LCIS). It is possible that LCIS was present in other cases and was just not found or was completely transformed in to the lobular component. However, this different distribution between DCIS and LCIS is suggestive that the invasive lobular component of a MDL may be the result of evolution from a tumour clone of “ductal” origin in some cases.

IHC staining for E-cadherin and its catenin complex partners supports this to some extent. In almost all cases the invasive lobular component of the MDL showed deregulation of the E-cadherin adhesion complex (**Section 4.2.2**), suggesting this is important mechanistically in driving the phenotypic switch during the evolution of these tumours. In the vast majority of ductal lesions E-cadherin was expressed and located at the cell membrane. When LCIS was present, then it and the invasive lobular component were more frequently negative for E-cadherin. However, when LCIS was not present, the invasive lobular component was more frequently aberrant for E-cadherin, with cytoplasmic accumulation of the protein. This would imply that the evolution of the invasive lobular component might be different depending on whether LCIS was present or not.

This was investigated further using molecular analysis of different morphological components of MDLs by cCGH (Kutasovic, 2011) and exome sequencing (**Section 4.2.5**). This first key point was that this analysis demonstrated that all discrete lesions within the

cases study were clonally related, as opposed to being the collision of individual tumours. There was multiple DNA copy number alterations and somatic mutations that were shared between all lesions (ductal and lobular) in a case. Additional genetic alterations were identified that were not shared (private mutations); indicating further clonal evolution occurred post divergence of ductal and lobular components. cCGH data from my honours thesis (Kutasovic, 2011) and the exome data presented here suggest that when LCIS was present, it was more genetically similar to the invasive lobular component than to the 'ductal' lesions (e.g. cases MDL1 and MDL6). Although only small numbers of cases have been studied so far it is noteworthy that the only case to contain a *CDH1* mutation is the lobular component of MDL6. Molecular data in cases without LCIS suggests the invasive lobular component is closely and clonally related to the invasive ductal component; and that the mechanism underlying the phenotypic switch remains unclear.

These hypothetical concepts for the evolution of different morphological components of MDLs are summarised in **Figure 4.17:** in MDL cases with LCIS and ILC, E-cadherin is typically negative (but in some cases aberrantly localised in both the LCIS and ILC) and the divergence from ductal pathway of evolution most likely occurred early during progression. In cases with no LCIS present, the invasive lobular component is more likely to exhibit aberrantly localised E-cadherin and that divergence from ductal morphology is more likely to have occurred later during evolution.

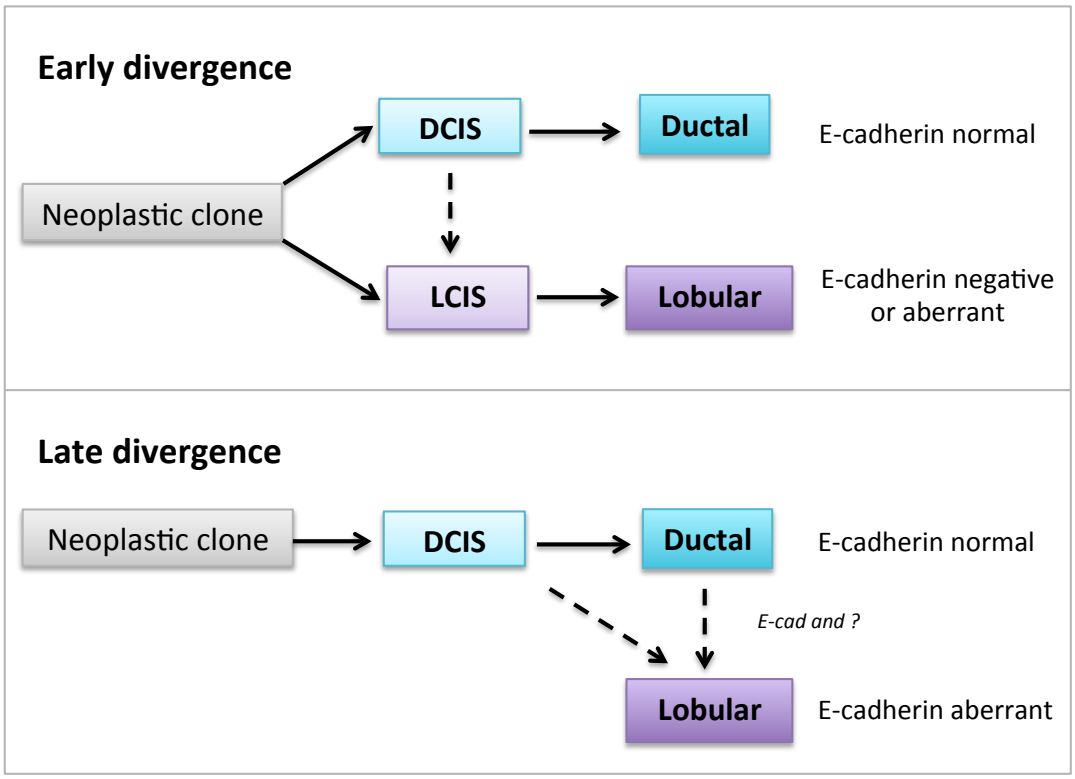


Figure 4.17: The clonal evolution of mixed ductal lobular carcinomas. We propose that in some cases early divergence from a common neoplastic clone (top panel), particularly in cases with both LCIS and DCIS present. However, in cases that only present with DCIS it is hypothesised that clonal divergence may occur later during tumour evolution (bottom panel). Dashed lines denotes our lack of understanding of the mechanisms driving the change in morphology.

It is unlikely that epithelial to mesenchymal transition is involved in the phenotypic switch to a more invasive (lobular) phenotype, as we did not see either the archetypal cadherin switch and/or activation of vimentin expression in this cohort (**Section 4.2.3**). However, the analysis of integrated protein and mRNA expression of multiple EMT markers should be performed to see if a coordinated expression of an EMT program occurs. This, in particularly should include *ZEB1*, which was up-regulated in some cases in the TCGA meta-analysis. This transcription factor is a known transcriptional repressor of *CDH1* (Wong et al., 2014) and therefore may be important in some MDL cases.

The TCGA provides a wealth of genomic data for large numbers of breast cancer samples. There were 29 MDL cases available for meta-analysis of gene expression and gene mutation data. Some caveats must be acknowledged with this approach, particularly when studying a tumour with morphological heterogeneity. For this study, it was therefore of vital importance to recognise the morphological characteristics of the specific region of the tumour sample from which DNA and RNA was extracted (lobular, ductal or admixed combination of both). We tried to address this by assessing the Cancer Digital Slide Archive (**Section 4.2.4**). There are many difficulties when looking at frozen tissue sections, particularly where freeze artefact can make the cells in a tissue section look discohesive. In most cases, two frozen section images were given and it was impossible to determine which one was used. FFPE slides were also available for most cases, and although this provides better morphology, this is unlikely to be directly representative of the frozen tissue processed for extraction. Ultimately therefore it was quite difficult to know whether the data obtained from the TCGA related to a ductal component and lobular component or an admixed component. These limitations need to be taken into account when interpreting the results.

The genes frequently mutated in the MDLs from the TCGA (**Section 4.2.4**) were those similar to ER positive IC-NST and ILC. There were no genes known to be involved in cell adhesion mutated in the TCGA data, other than three cases with *CDH1* mutations.

However, this could simply be a reflection of the proportions of either the ductal or lobular components sequenced; false negatives are a real possibility, given our morphology assumptions. E-cadherin IHC would be very useful in these cases to see if the cases without *CDH1* mutation have aberrant localisation in the lobular component.

The genes altered in the TCGA cohort could possibly affect novel pathways involved in the regulation of cell adhesion. For example, there were mutations in genes encoding proteins of the actin cytoskeleton and Wnt signalling, the alterations of which may be impacting cell adhesion through interactions with E-cadherin and β-catenin. The lack of direct hits to the cell adhesion complex may suggest that other mechanisms are at play, such as deregulation at the epigenetic or post-translational level or protein trafficking. Transcriptome profiling of a series of well-characterised, micro-dissected samples could be employed to help understand the underlying mechanisms.

Four cases were selected for micro-dissection of the various invasive and pre-invasive lesions. There were difficulties in finding cases for molecular analysis, the most common limitations being that i) one or more of the important morphological components was too small (often the invasive lobular component); ii) or too old to get enough good quality DNA; or iii) the morphological components were too admixed to have a pure population of cells from both morphological components to investigate the cause of the switch in phenotype. All of the four cases chosen were less than 5 years old, the morphological components were sufficiently discrete and large enough to yield good quality DNA and thus some success in the exome sequencing. Only a preliminary data analysis has been undertaken on 3 of the 4 cases to date, to provide data for this thesis. We are awaiting further sequencing on the same samples to increase the depth of coverage and hence the robustness of the somatic variants called. The analyses shown will then be repeated, including IntOGen analysis and together with collaborative support to formulate more formal phylogenetic modelling of tumour evolution. We will also validate somatic mutations of interest by Sanger sequencing.

Interestingly, one of the cases that displayed negative E-cadherin staining in the lobular component harboured a *CDH1* mutation. In contrast, the other E-cadherin negative case and the cases with aberrant E-cadherin localisation did not have a *CDH1* mutation, or any mutation was the obviously involved in cell-cell adhesion from this preliminary analysis. This finding supports our hypothesis that the mechanisms underlying aberrant E-cadherin

localisation are different to those in pure ILC. The deregulation of E-cadherin cell adhesion could be a result of a variety of gene expression changes, epigenetic modifications or sub-cellular localisation changes. Gene expression or methylation analysis has not been attempted on this cohort, but is an important future direction that is worth pursuing.

IntOGen is a resource developed to integrate the multitude of genome-wide data generated from the large international consortia such as the TCGA and ICGC (Gundem et al., 2010). Due to the large number of samples analysed by IntOGen, statistical analyses can be performed to detect the most likely driver mutations within a sample, and also calculate which genes contribute to similar biological pathways. In this cohort, the common driver genes across all cancer types were present at low frequency (less than 50%). The candidate driver gene analysis, however, showed that each case harboured unique driver genes. Of interest were somatic mutations in *BRCA2* and *SMAD4* in MDL6, which will be validated by Sanger sequencing. The IntOGen pathway analysis was revealing (**Section 4.2.6**), where genes in pathways involved in gap junctions, Gonadotropin-releasing hormone (GnRH) and MAPK signalling pathways were the most frequently affected. Gap junction signalling was the top hit on the pathway analysis; therefore this pathway would be the first to be investigated further. Gap junctions (made up of the protein, connexon) directly connect the cytoplasm of two cells to allow various signalling molecules to pass through neighbouring cells. There are 20 different connexon genes in the human genome and it has been reported that only connexon 26 (*GJB2*) and connexon 43 (*GJA1*) are expressed in human mammary epithelial cells (El-Saghir et al., 2011). Tumour suppressive and tumour promoting roles have been proposed for connexons, and alongside tight junctions and adhesion junctions, gap junctions are important in maintaining cellular polarity and cell-cell communication (El-Saghir et al., 2011). Taken together, this pathway is a strong candidate for future investigation.

It must be acknowledged that only a small number of cases were analysed in this study and therefore it would be of interest to also analyse the TCGA MDL cases using IntOGen to see if these same pathways are detected, and how comparable the frequencies of alterations in these pathways are in comparison to ILC, ER positive and ER negative IC-NST. Other pathway analyses such as Ingenuity Pathway Analysis (IPA) and String on the TCGA gene expression data would be very complementary to this analysis. This may reveal pathways that are unique to each histological type of breast cancer. IntOGen analysis on a case-by-case basis in our cohort would also be interesting to see whether

there are different signalling pathways altered between each morphological component. This may reveal specific pathways that contribute to a change to an invasive lobular phenotype. A comparison of altered pathways between the case with the *CDH1* mutation (MDL6, **Section 4.2.5.3**) to the cases without *CDH1* mutation would also be of great interest to begin to understand how E-cadherin is becoming mislocalised in the lobular components of MDLs.

4.4. Conclusions

This study has found that the different morphological components present within an MDL tumour are clonally related and not the result of a collision of multiple independent tumours. It appears that in some cases the divergence of the morphological components may occur early during tumour evolution (where there are both DCIS and LCIS present) or later during tumour progression (cases with only DCIS detectible; **Figure 4.5**). The cases with late occurring divergence may be arising via a ductal like pathway of progression, and this study emphasises the possibility that a lobular-like phenotype can arise via a ductal pathway. Since the gap junction signalling pathway was a top hit from pathway analysis suggests that pathways of cell-cell adhesion, other than E-cadherin based adhesion junctions, may be important in breast tumourigenesis.

CHAPTER 5

INVESTIGATING BREAST CANCER METASTASIS TO GYNAECOLOGICAL SITES

5. Investigating breast cancer metastasis to gynaecological sites

5.1. Introduction

Death from breast cancer is not commonly caused by the growth of the primary tumour, but the spread of the tumour to distant sites – the process of metastasis, which accounts for 90% of all cancer deaths. The most common sites of breast cancer metastasis include the bone, lung, liver and brain. Analyses of autopsy and surgical series have compared the organ specific metastatic spread of the two main histological types, IC-NST and ILC. IC-NST spreads more frequently to the lung/pleura, liver and brain. Whilst ILC spread to these organs too, they also have a propensity to spread to unusual sites and, relative to IC-NST they more frequently colonise the peritoneum, gynaecological sites, gastrointestinal tract, adrenal glands, central nervous system and skin. Primary tumours that are ER positive also have a different pattern of metastasis to those that are ER negative (see **Section 1.10**, (Sastre-Garau et al., 1996, Jain et al., 1993, Borst and Ingold, 1993, Lamovec and Bracko, 1991, Arpino et al., 2004, Harris et al., 1984)).

Due to our interest in ILC we decided to investigate metastasis to gynaecological sites as this has not been well characterised. Data from our own autopsy study (Cummings et al., 2014) and others have demonstrated that gynaecological metastasis occurs in quite young patients. The median age at diagnosis of the primary tumour in patients that later developed gynaecological metastases ranged from 46-54 years (Demopoulos et al., 1987, de Waal et al., 2009, Kondi-Pafiti et al., 2011, Salamalekis et al., 2004, Cummings et al., 2014). Secondary malignancies account for 7-10% of all gynaecological malignancies (Young and Scully, 1991) and patients can be asymptomatic for many years (Bigorie et al., 2010). Up to 75% of patients with ovarian metastases experience symptoms that are not necessarily associated with malignancy, such as bloating, abdominal pain, postmenopausal bleeding and weight loss (Moore et al., 2004). Thus by the time the metastases are diagnosed, they can be wide spread and therefore overall survival is poor: 52% of patients survive 2 years, 24-36% survive 5 years, and 7% survive 10 years (Ayhan et al., 2005, Skirnisdottir et al., 2007, Demopoulos et al., 1987, de Waal et al., 2009). Primary colon cancer is the most common tumour type to spread to gynaecological sites and perhaps this is not surprising given the incidence of this disease and the anatomical proximity of the tissues, where seeding may occur through the abdominal cavity. Interestingly, primary breast cancer is the second most common tumour to spread to

gynaecological organs suggesting that in this case there may be more targeted mechanisms of spread rather than by proximity (Moore et al., 2004, Kondi-Pafiti et al., 2011, Yada-Hashimoto et al., 2003). Small studies have found breast cancers that spread to gynaecological organs are predominantly ER positive and HER2 negative (Bigorie et al., 2010, St Romain et al., 2012) and as mentioned above invasive lobular carcinomas are more likely to spread to these sites, suggesting that hormonal signalling and/or a diffuse growth pattern might be important. In support of this, evidence from autopsy studies also shows that gynaecological metastases arising from mixed ductal lobular carcinomas frequently have a diffuse lobular-like growth pattern (Lamovec and Bracko, 1991).

As described in **Section 1.10.5**, morphological and phenotypic intertumour heterogeneity among breast cancer patients has long been recognised, however the clinical implications of intratumour heterogeneity has recently come into the spotlight, particularly in regards to the complex process of clonal evolution in metastatic progression. Phenotypic heterogeneity between primary tumours and metastases has been described, where the expression of tumour biomarkers has been found to change during progression. For instance, ER and PR are frequently down-regulated with progression, particularly to lung, bone and liver (Cummings et al., 2014, Singhakowinta et al., 1976, Wu et al., 2008, Idirisinghe et al., 2010). Discordance of HER2 amplification/overexpression has also been reported in approximately 10% of metastatic breast cancer cases (Fabi et al., 2011). Genomic studies have shown that different primary tumour subclones have different metastatic capabilities and can spread to different distant sites (Campbell et al., 2010, Ding et al., 2010, Gerlinger et al., 2012, Yachida et al., 2010) and the acquisition of mutations in genes such as *ESR1* may contribute to progression by conferring resistance to treatment (**Section 1.3.1** (Jeselsohn et al., 2015).

Treating metastatic breast cancer based solely on the characteristics of the primary tumour clearly has limitations and clinical practice is now recommending that metastases be biopsied to help determine treatment (Van Poznak et al., 2015). It is therefore of great importance to study metastatic progression to gain a greater understanding of the mechanisms and selective pressures that underpin tumour evolution and treatment resistance.

The extent and overall clinical significance of clonal evolution in metastatic progression remains unclear owing to the scarcity of metastatic samples available for analysis. The fact

that gynaecological sites are generally operable (*c.f.* the brain) enables the opportunity to study mechanisms of metastasis and intratumour heterogeneity in detail, within human samples.

5.1.1. Hypotheses and aims

Tumour progression to metastasis is a complex process in which tumour cells evolve and acquire advantageous characteristics based on selective pressures from therapy, immune surveillance and growing in a new microenvironment. Metastatic progression to gynaecological organs is of interest since it has been identified to be a particular haven for invasive lobular carcinomas and is associated with young women. We hypothesise that metastasis to gynaecological organs is organotropic, facilitated by something peculiar about the biology of the primary tumour and/or the target organ (e.g. deregulated E-cadherin and oestrogen signalling). The following aims were designed to try and address this hypothesis:

- Collate a cohort of samples of primary breast cancers and matched metastases.
- Characterise the cohort for clinical and pathology features.
- Characterise the immunohistochemical phenotype of breast and metastatic tumours.
- Assess the phenotypic and genomic heterogeneity between matched primary and metastatic samples.

5.2. Results

We have constructed a cohort of breast cancer patients who developed metastases to gynaecological sites. We used a variety of sources to acquire and characterise this cohort, including the identification of cases from the Queensland Centre for Gynaecological Cancer (QCGC) and Pathology Queensland (as described in **Section 2.1**).

5.2.1. Clinical characteristics of gynaecological metastatic breast cancer cohort

The QCGC database contained 1067 cases of gynaecological metastases diagnosed between 1982 and 2011 at the RBWH. The most common primary site to spread to the gynaecological sites was colorectal cancer (n = 239, 22.3%), followed by breast cancer (n = 108, 10.1%). Pathology reports and Queensland Cancer Registry reports were used to verify that the 108 cases were in deed breast cancer derived metastases. Data for 86 of

these cases were available and 54 had gynaecological metastases. There was fall out of 32 cases as the additional searches revealed that some of the cases of gynaecological tumours were primary disease, were derived from another primary site or the primary tumour was from an unknown origin.

Clinicopathological data for the 54 cases relating to the primary breast tumour (age of diagnosis, histological type, size, grade, lymph node involvement, ER, PR, HER2 status), treatment history, and characteristics of the metastases (age of diagnosis, sites involved, biomarker results) was recorded (**Table 5.1 and 5.2**). We compared the characteristics of this gynaecological metastasis cohort with our in-house 'QFU' cohort of 449 unselected sporadic primary breast cancers (described in **Section 2.1)**.

Some of the key findings are described:
The median age of primary tumour diagnosis of gynaecological metastatic breast cancer patients was 47 years old (range 30-79 years); and using a cut-off of 51 years (the average age of menopause in Australia, (Do et al., 1998)) we found that 57.4% of patients were likely to be pre-menopausal at the time of diagnosis. Both these findings were significantly different between the two cohorts (**Table 5.1**). The median age at diagnosis of metastasis was 55 years (range 35-82 years) and the median time to metastasis was 6 years (range 0-20 years). There was enrichment for ILC to spread to gynaecological organs (42.6%, compared to 17.8% in the QFU cohort; Chi-square test: $p < 0.0001$); and 83.3% of patients in the gynaecological metastasis cohort had lymph node metastasis at primary tumour diagnosis, compared with 46.2% in the QFU cohort ($p < 0.0001$; **Table 5.1**)

There were 255 metastatic deposits recorded for this cohort of patients. Of these, 137 were identified in gynaecological sites and 118 in non-gynaecological sites (**Appendix Table 5.1)**. The most common gynaecological sites involved were the ovary (46%) followed by the fallopian tube (23.7%) and uterus (20.1%; **Figure 5.1 A**). The majority (87%) of the cases also had other organs involved, most commonly the peritoneum/omentum (34.7%) and sites of the digestive system (21.2% e.g. appendix, stomach, colon). Interestingly, common sites of breast cancer metastasis, such as the lung, bone, brain and liver, were less frequently involved (9.3%, 16.9%, 5.9% and 5.9%, respectively) (**Figure 5.1 B; Appendix Table 5.1**). We also found that in 70.4% of patients, the gynaecological metastasis was the first site of distant metastasis identified clinically (based on analysis of the pathology and clinical reports, **Appendix Table 5.2**).

Table 5.1: Clinico-pathological characteristics of primary breast cancers with metastasis to gynaecological organs.

	GM		QFU	p value
Age of breast cancer diagnosis	***Years***		***Years***	
Median	47		60	< 0.0001^
Range	30-79		27-88	
Menopausal status	***n***	**%**	**%**	
Premenopausal	31	57.4	32.9	0.0004
Postmenopausal	23	42.6	67.1	
Tumour size	***n***	**%**	**%**	
<2 cm	14	43.8	46.6	0.1922
2 to 5 cm	12	37.5	44.5	
>5 cm	6	18.8	8.9	
Unknown	22			
Total	54			
Histological type	***n***	**%**	**%**	
Invasive carcinoma NST	26	48.1	71.1	< 0.0001
Invasive lobular carcinoma (incl. pleomorphic ILC)	23	42.6	17.8	
Mixed ductal-lobular carcinoma	1	1.9	11.1	
Bilateral IC-NST/ILC	2	3.7		
Medullary carcinoma	1	1.9		
Tubular carcinoma	1	1.9		
Total	54			
Tumour grade	***n***	**%**	**%**	
Grade 1	2	6.6	14.2	0.3622
Grade 2	14	46.7	49.5	
Grade 3	14	46.7	36.3	
Unknown	24			
Total	54			
Lymph node status	***n***	**%**	**%**	
Positive	30	83.3	46.2	< 0.0001
Negative	6	16.7	53.8	
Unknown	18			
Total	54			
Age of first metastasis diagnosis	***Years***			
Median	55			
Range	35-82			
Number of metastatic sites	***n***	**%**		
1-3	22	40.7		
4-11	32	59.3		
Median	4.5			
Range	1-11			
Progression free survival (n = 54)	***Years***			
Median	5			
Range	0-20			
Time until death (n=22)	***Years***			
Median	10			
Range	2-24			

^ t-test; all other p-values were derived using Chi-square analysis. All statistical analyses were performed using GraphPad Prism version 6.

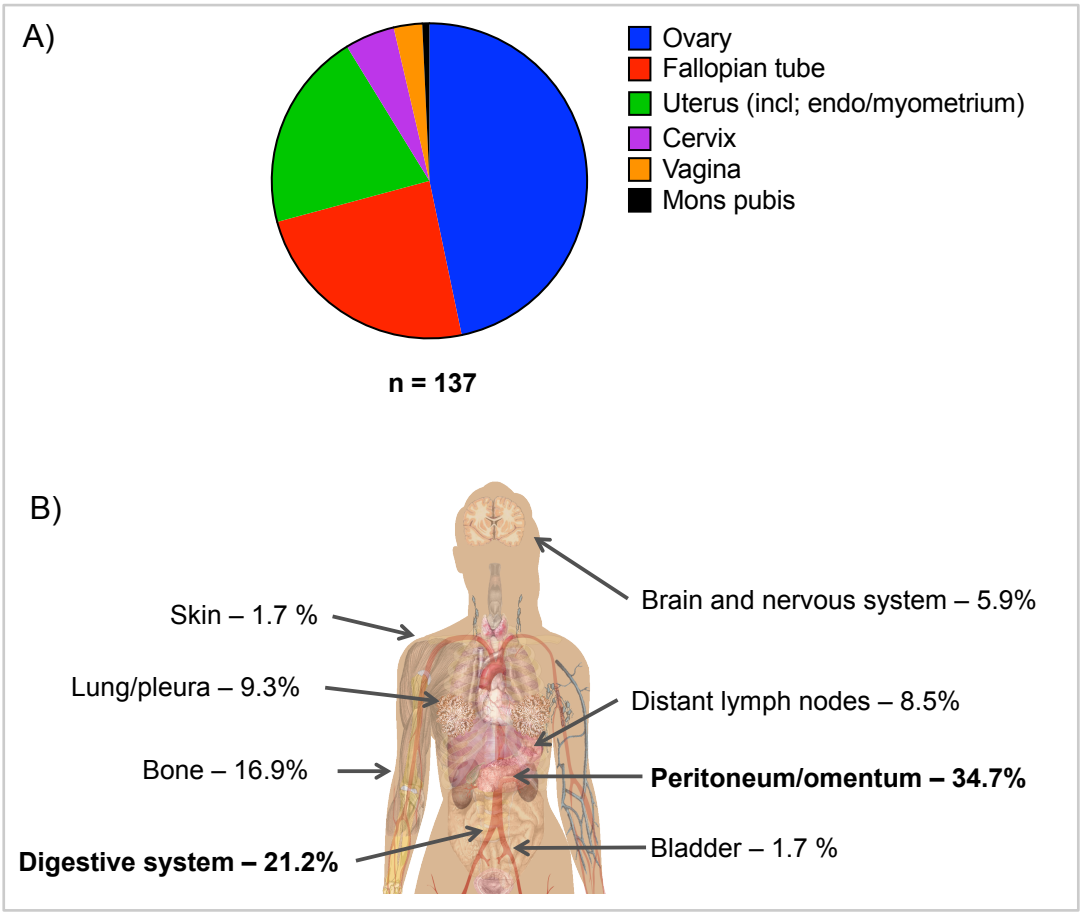


Figure 5.1: Patterns of breast cancer metastasis. A) Frequency of gynaecological sites involved. B) Sites involved alongside gynaecological metastasis.

We consulted Queensland Health clinical databases and medical records to obtain treatment information for the patients in the cohort (**Section 2.1**; summarised in **Table 5.2)**. Unfortunately, limited data was available as many patients were i) diagnosed with breast cancer over 20 years ago; ii) they were treated at hospitals not in the Brisbane region or iii) they were treated at a private hospital. Of the data available, 10 patients received both endocrine therapy and chemotherapy for treatment of their primary tumour. Seven patients received only chemotherapy and seven received only endocrine therapy. Twenty-one cases received radiation therapy as well as systemic therapy, while 4 cases only had data available for radiation treatment. It is reasonable to assume that most of the patients in this cohort received systemic therapy and radiation therapy for their primary tumour, but we were unable to locate the information.

Table 5.2: Summary of treatment received by patients with gynaecological metastases.

	Primary tumour			Metastatic disease		
	Endocrine	Chemotherapy	Radiation	Endocrine	Chemotherapy	Radiation
GM01	Y	Y	Y	N	Y	N
GM02	x	x	Y	x	Y	x
GM06	N	Y	Y	N	x	x
GM07	x	x	Y	Y	x	Y
GM09	Y	x	x	Y	Y	x
GM11	x	x	x	Y	x	Y
GM13	x	Y	N	Y	x	x
GM15	Y	x	x	Y	x	x
GM16	Y	x	Y	Y	Y	x
GM18	x	x	Y	x	x	x
GM24	x	Y	Y	x	x	x
GM29	Y	x	Y	x	x	x
GM33	Y	Y	Y	Y	x	x
GM36	x	Y	Y	Y	Y	Y
GM37	Y	x	Y	x	x	x
GM39	Y	x	Y	Y	Y	x
GM41	Y	Y	Y	Y	Y	x
GM45	x	x	x	x	Y	x
GM47	Y	Y	Y	x	x	x
GM48	x	x	x	Y	Y	x
GM49	Y	Y	Y	Y	Y	x
GM50	Y	x	Y	Y	Y	x
GM57	Y	Y	Y	Y	Y	x
GM59	Y	Y	Y	Y	Y	x
GM67	x	Y	Y	Y	x	Y
GM72	x	Y	x	Y	Y	Y
GM73	x	x	Y	x	x	x
GM74	x	Y	Y	x	x	x
GM76	Y	Y	Y	Y	Y	x
GM77	Y	Y	Y	x	x	x
GM78	x	x	x	x	Y	Y
GM87	x	x	Y	x	x	Y
GM89	Y	Y	Y	Y	Y	Y

x = data not available. Y = yes; N = no.

5.2.2. Immunophenotyping of primary breast cancers and gynaecological metastases

Tissue blocks were retrieved from various pathology laboratories around Queensland (**Section 2.1**). Blocks were available from public and private hospitals for 39 cases. Fifteen cases had matched primary tumour and metastases, while 4 cases had only the primary tumour available and 20 cases had the metastases available.
All 39 cases were sampled across 6 TMAs (**Section 2.4**, see **Appendix Figures 2.2** to **2.7** for TMA maps). Of the metastatic sites available, 107 sites were arrayed on the TMAs; 50 from ovaries, 35 from other gynaecological sites, 12 from either the peritoneum or omentum, 4 gastrointestinal sites, 5 lymph node metastases and 1 brain metastasis. In some cases we observed morphological heterogeneity, and so multiple areas were sampled (labelled as growth pattern 1, 2 *etc.*). Immunohistochemistry was performed (**Section 2.5.4**) to assess the expression of clinically relevant biomarkers ER, PR and HER2 that are routinely used in clinical practice to guide treatment options (**Section 1.2.2**). Biomarkers that are of interest in breast cancer biology (see **Sections 1.3**) were also assessed in this cohort of patients with gynaecological metastasis to further characterise the cohort, including p53, Ki67, CK5/6, CK14, EGFR, E-cadherin, AR, GATA3 and FOXA1. Data was supplemented from pathology reports where tissue was not available yet IHC had been performed for diagnostic purposes. All raw data is presented in **Appendix Table 5.3.**

Breast cancers that spread to gynaecological sites frequently expressed ER (93.5%) and PR (65.7%), but never HER2 (100% of cases were negative). This data was compared to the QFU cohort described in **Section 2.1** and the frequent expression of ER and the lack of HER2 expression was found to be significantly different in gynaecological-metastatic breast cancer compared to unselected sporadic breast tumours ($p < 0.0001$; **Figure 5.2, Figure 5.3**). The frequency of PR expression was not significantly different between the cohorts (p = ns).

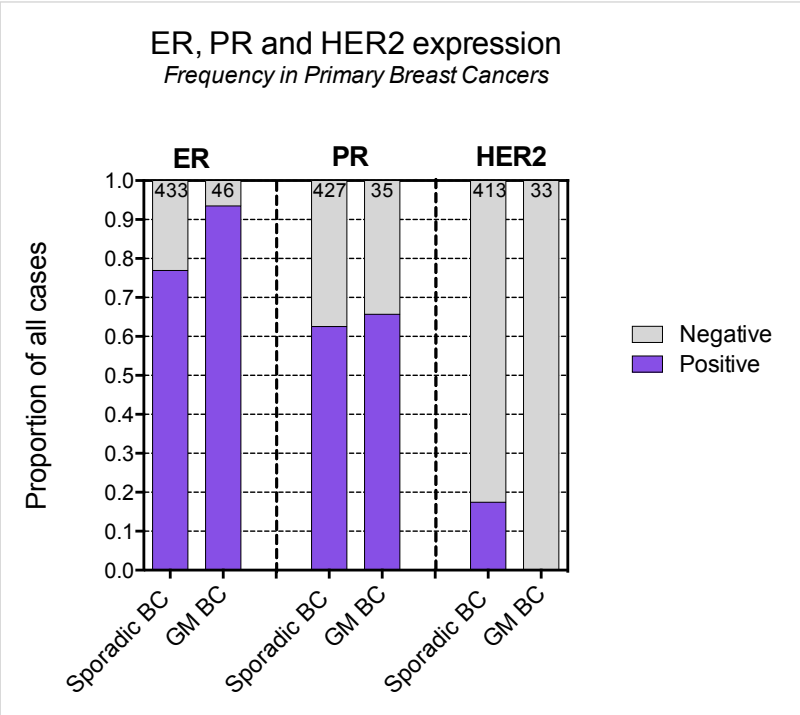


Figure 5.2: The proportion of primary breast cancers that were positive or negative for ER, PR and HER2 in our two cohorts – our in-house cohort of unselected sporadic breast cancers (the QFU cohort) and our cohort of primary breast cancers that spread to gynaecological sites (GM BC). The number at the top of each column denotes the total number of tumours analysed. The GM BC data is a combination of data from the TMA analysis and supplemented by pathology report data when the tissue was unavailable. The expression of ER and HER2 were significantly different ($p < 0.0001$) between the two cohorts.

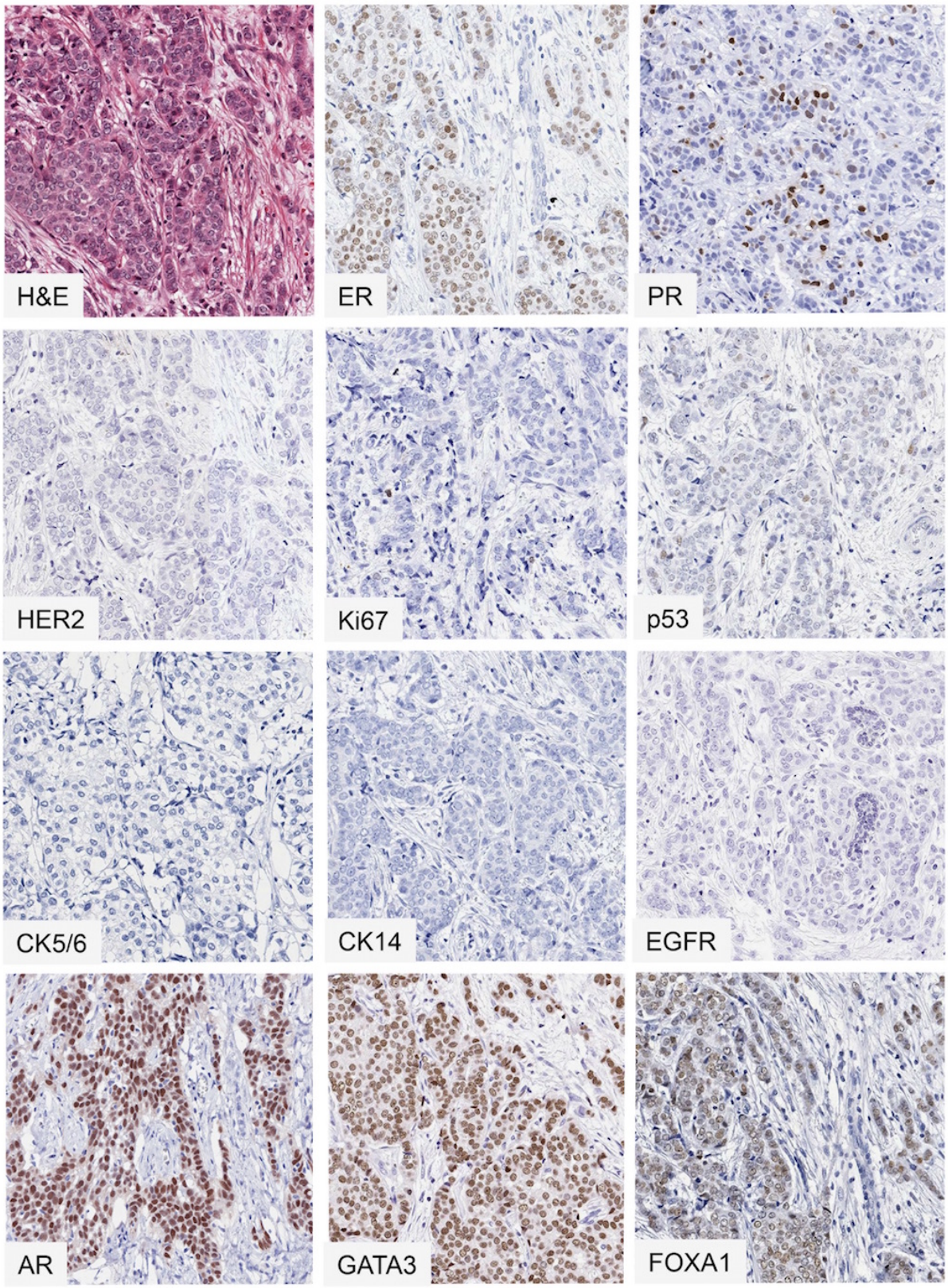


Figure 5.3: Immunophenotype of a representative gynaecological metastatic primary breast tumour. These primary tumours display features of a good prognosis (*i.e.* equivalent to a luminal A molecular phenotype) and the phenotype is largely maintained during metastatic progression.

We were interested to see if the expression of ER, PR and HER2 changed during progression to gynaecological sites. We found that 93.3% of the metastatic sites sampled expressed ER, 66% expressed PR and no metastatic sites expressed HER2. The metastatic sites were separated into discrete anatomical regions (**Figure 5.4**), and no significant differences in expression of ER and PR were observed between sites, though the omentum and peritoneum showed a slightly reduced frequency of positivity.

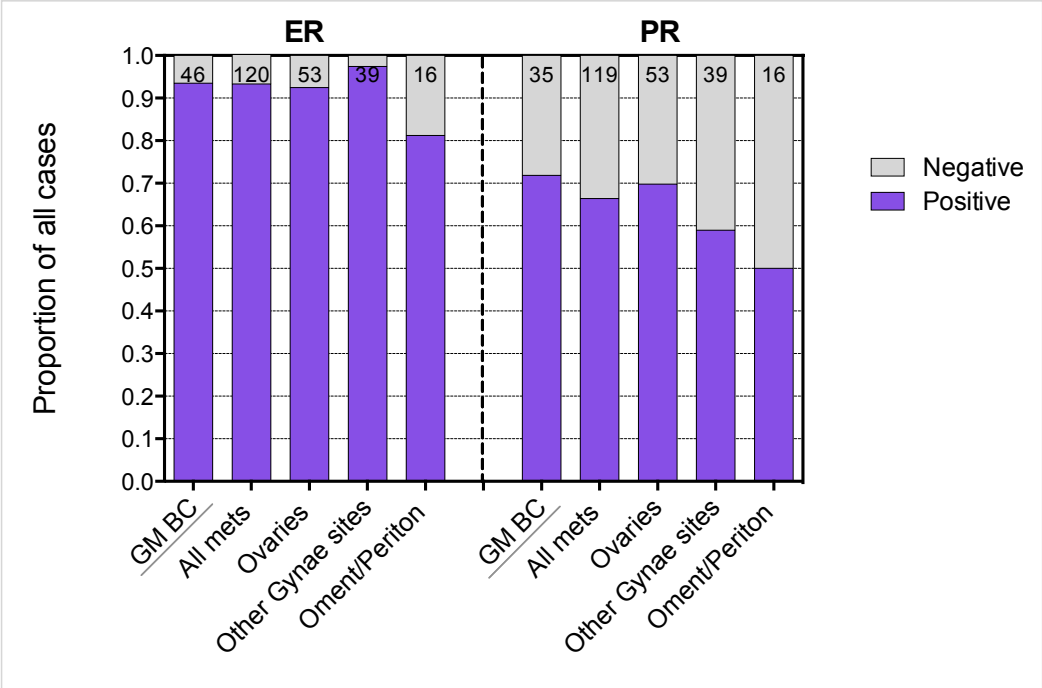


Figure 5.4: The expression of ER and PR is maintained during progression to gynaecological sites. No differences in the frequency of expression of either biomarker were observed between anatomical sites. GM BC = primary breast tumours that spread to gynaecological sites. Oment/Periton = omentum and peritoneum.

The expression of CK5/6, CK14 and EGFR are used as IHC surrogate biomarkers for tumours with a basal-like phenotype. Basal-like tumours are associated with poor overall survival and short time to metastasis (Fulford et al., 2007). None of the 10 primary tumours that had data available expressed basal markers. This was significantly different to that observed in the QFU cohort (**Figure 5.3 and 5.5 A;** CK14: $p = 0.0564$; CK5/6: $p = 0.0016$; EGFR: $p = 0.096$). All of the metastases were also negative for the basal markers. CK5/6 and CK14 were focally positive (expressed in less than 10% of cells) in one metastatic deposit from two individual cases and when using a cut off of 10% these would not be considered basal-like.

p53 IHC staining is typically associated with aggressive high-grade breast cancers (Rakha et al., 2010b). Here we scored staining in the tumour nuclei as low (10-30% of tumour nuclei positive) or high (>30% of tumour nuclei positive). Primary breast cancers that spread to gynaecological sites were most frequently negative for p53 (68.4%; 10.5% showed low expression; 21.1% showed high expression). The metastases were more frequently positive than the primary tumours (37.9% *vs.* 21.1%, respectively) and this was found to be significant (p = 0.0227; **Figure 5.5 B**). Cytoplasmic localisation of p53 was not observed.

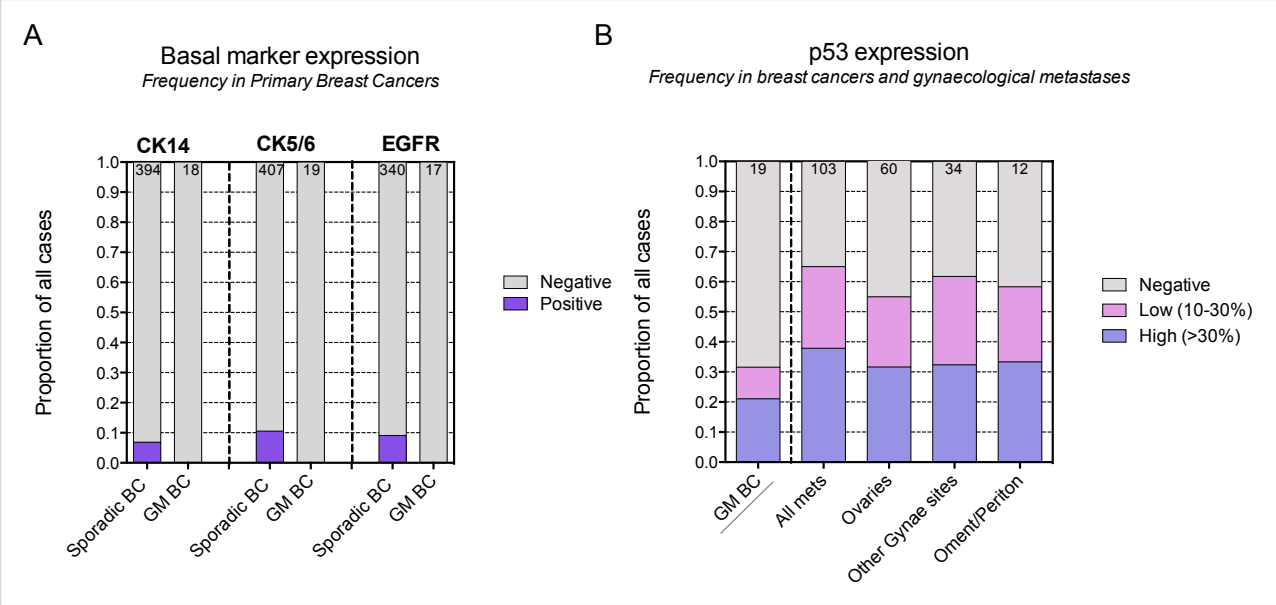


Figure 5.5: A: CK14, CK5/6 and EGFR are not expressed in breast cancers that spread to gynaecological organs. All of the metastases were also negative for CK14, CK5/6 and EGFR (not shown). B) The expression of p53 in breast cancers and gynaecological metastasis.

Ki67 is a commonly used biomarker to assess tumour cell proliferation. Tumours that express high levels of Ki67 tend to have a higher rate of proliferation and are associated with poorer patient outcome (**Section 1.3**). Ki67 has become an important surrogate biomarker for molecular subtyping. Particularly in the context of ER/PR positive tumours, where it can be used (as well as HER2 in some instances) to stratify luminal A and B subtypes (Cheang et al., 2009). In comparison to sporadic breast cancers, primary breast cancers in our gynaecological metastasis cohort had a low Ki67 proliferative index (p < 0.0001; **Figure 5.6 A, Figure 5.3**). When comparing primary tumours and metastases, Ki67 expression was higher in the metastases (p < 0.0001).

The epithelial cell specific marker CK8/18 was also assessed across the TMAs and CK8/18 was expressed in 100% of the primary tumours and metastatic samples analysed (data not shown).

In summary, this data indicates that primary tumours from patients who developed gynaecological metastasis fit with a luminal A phenotype (*i.e.* ER and PR positive, negative for HER2 and basal markers and are Ki67 low). Across the cohort expression of ER, PR, HER2 and the basal markers do not change during progression, however p53 and Ki67 shows an increase in frequency of expression in the metastases versus the primary tumours sampled.

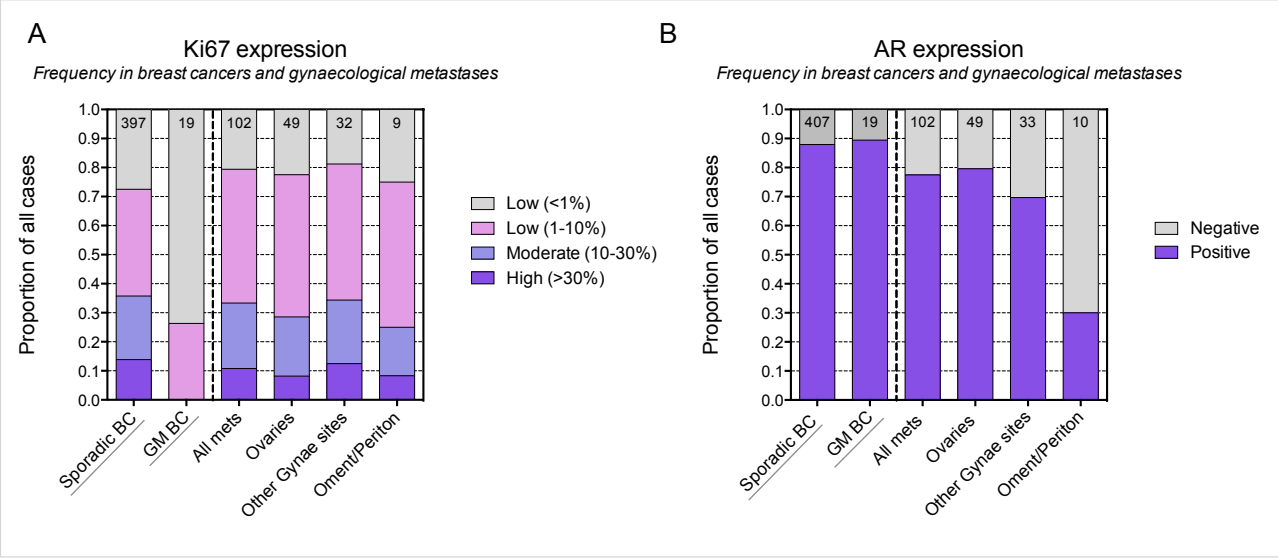


Figure 5.6: A) Ki67 expression in breast cancer and gynaecological metastases. GM BC have a lower Ki67 index *cf.* with sporadic breast cancer ($p < 0.0001$) and the metastases express more Ki67 than the primary tumours ($p < 0.0001$). B) Androgen Receptor expression in breast cancer and gynaecological metastases. No differences in AR expression were found other than lower expression of AR in the omentum/peritoneum ($p = 0.0127$) to compared to the other metastatic sites.

5.2.2.1. Expression of other hormone receptors and ER regulatory proteins in gynaecological-metastatic breast cancer

Given the luminal phenotype described of this tumour cohort, we investigated the expression of other hormone receptors and regulatory molecules. The expression of ERβ (clone 14C8) was attempted, however the antibody optimisation was unsuccessful with no staining obtained.

The androgen receptor plays an important role in hormone signalling in prostate cancer and is also frequently expressed in breast cancers; it is believed that AR plays different roles in ER positive and ER negative disease (**Section 1.3**). We found that 89.5% of gynaecological-metastatic primary tumours expressed the androgen receptor (AR), compared to 88% of sporadic breast cancers (**Figure 5.6 B**). We did not observe a significant change in AR expression in the metastatic sites (77.5% expressed AR). We did find, however, that AR expression was significantly less frequent in non-ovarian gynaecological sites (23/33; 69%) and the peritoneum/omentum (3/10; 30%), compared to all other sites was statistically significant ($p = 0.0127$).

As described in **Section 1.3**, FOXA1 and GATA3 are essential for ER transcriptional activity, and loss of these proteins has been found to redirect ER to bind to different areas of the genome and change the transcriptional program of the cell. This change of gene expression has been associated with endocrine resistance (Hurtado et al., 2011, Ross-Innes et al., 2012). FOXA1 was expressed in 77.7% (317/408) of our unselected sporadic breast cancer cohort, and this was significantly different to the 100% (19/19) FOXA1 positive primary breast tumours that metastasised to gynaecological sites (Fisher's exact test, $p = 0.0181$; **Figure 5.7 A**). FOXA1 was expressed in 73.1% of metastatic sites, and the difference in FOXA1 expression between primary tumours and all metastases was significant (Fisher's exact; $p = 0.0123$). No significant difference in FOXA1 expression between different metastatic regions was found.

All (19/19; 100%) of the primary tumours expressed GATA3, compared to 84.6% (356/421) of sporadic breast tumours ($p < 0.0001$; **Figure 5.7 B**). The expression of GATA3 in the metastatic sites (54.3%) was significantly reduced ($p < 0.0001$); there was no significant difference between metastases at different anatomical sites.

FOXA1 and GATA3 were co-expressed in the metastases in 25/39 cases (64.1%). Of the 14 cases that displayed mutually exclusive expression, 64.3% of the metastatic sites were GATA3 negative and FOXA1 positive while 35.7% of cases were the opposite (5 cases were GATA3 positive and FOXA1 negative). The number of cases is too small to correlate the changes in FOXA1 and GATA3 expression with age, menopausal status and histological type.

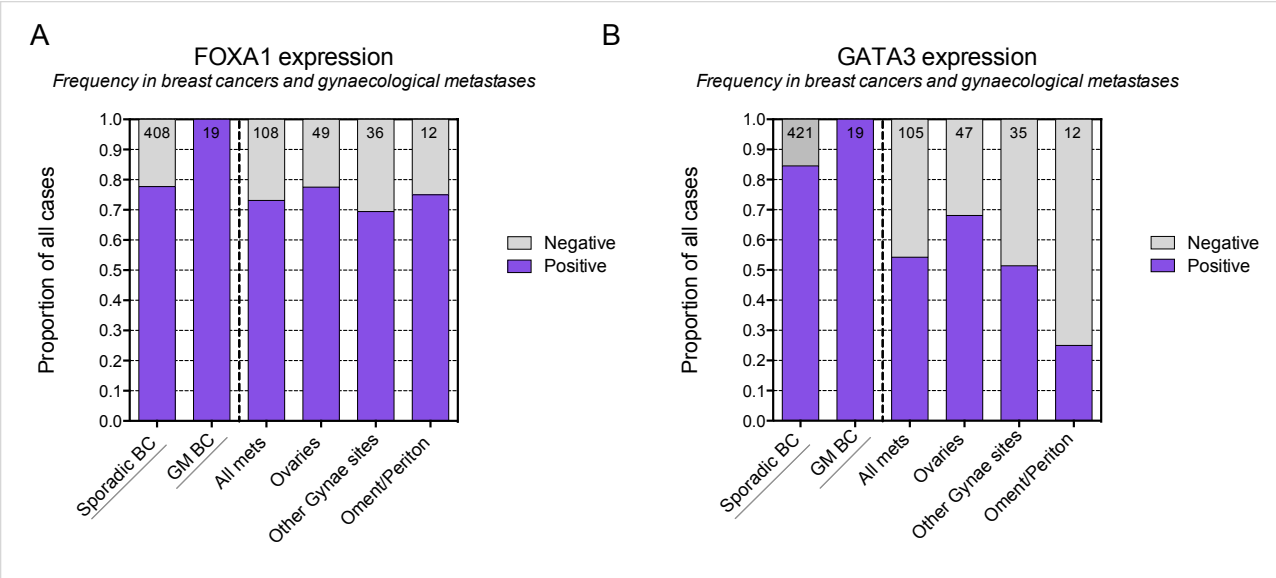


Figure 5.7: FOXA1 and GATA3 expression in primary tumours and gynaecological metastases.

A) FOXA1 was more highly expressed in gynaecological metastatic primary tumours compared to our sporadic primary tumour cohort (p = 0.0181, Fisher's exact test). Loss of FOXA1 expression in the metastases was significant more frequent compared to the primary tumours from the same cohort (Fisher's exact test, p = 0.0123).

B) Similarly, GATA3 was more highly expressed in gynaecological metastatic primary tumours compared to our sporadic primary tumour cohort and expression was lost during progression to metastasis (p < 0.0001).

To summarise, AR, FOXA1 and GATA3 were highly expressed in the primary tumours that spread to gynaecological sites. Expression of all three markers however was lost during metastatic progression in a significant number of cases.

5.2.2.2. The expression of E-cadherin during progression to gynaecological sites

The E-cadherin cell-cell adhesion complex plays an integral role in ILC biology were it is deregulated in the vast majority of cases, and can also be down regulated in high-grade IC-NST (**Section 1.8.1**).

E-cadherin data was available for 11 ILC primary tumours and 54 ILC metastases. Nine of the 11 (81.8%) primary tumours were completely negative for E-cadherin, while the remaining two cases displayed aberrant E-cadherin localisation (fragmented membrane and cytoplasmic). The most predominate phenotype in the metastases derived from an ILC was E-cadherin negative (79.6%), the remaining tumours all displayed aberrant E-

cadherin localisation. There were 9 cases with matched primary ILC and metastasises and 100% of cases were concordant.

E-cadherin IHC was performed on 12 IC-NST primary tumours and 55 IC-NST derived metastases. Positive membrane staining was observed in 8/12 (66.7%) of the primary tumours and in 19/55 (34.5%) of metastases. E-cadherin was negative in 3/12 (25%) primary tumours while 27/55 (49.1%) metastases had aberrant E-cadherin localisation (either fragmented membrane in over 50% of tumour cells or cytoplasmic). There were 6 cases with matched primary IC-NST and metastatic tumours. Half of the cases (3/6) cases were concordant, 1 case had a metastasis that was lobular like with cytoplasmic E-cadherin localisation (**Figure 5.8**, top panel GM36) and 2 cases were negative in the primary tumour but the metastases were positive (**Figure 5.8**, bottom panels GM57 and GM16). These discrepancies are interesting and can possibly be explained i) by heterogeneous staining in the tumour and therefore the tissue sampled on the TMA isn't representative of the whole tumour; or ii) a small, lobular-like subclone seeded the metastasis. This pattern of expression is unlikely to be related to experimental artefact (*e.g.* antibody failing or missing this tissue core) since these cores have worked for other antibodies and surrounding cores worked for E-cadherin.

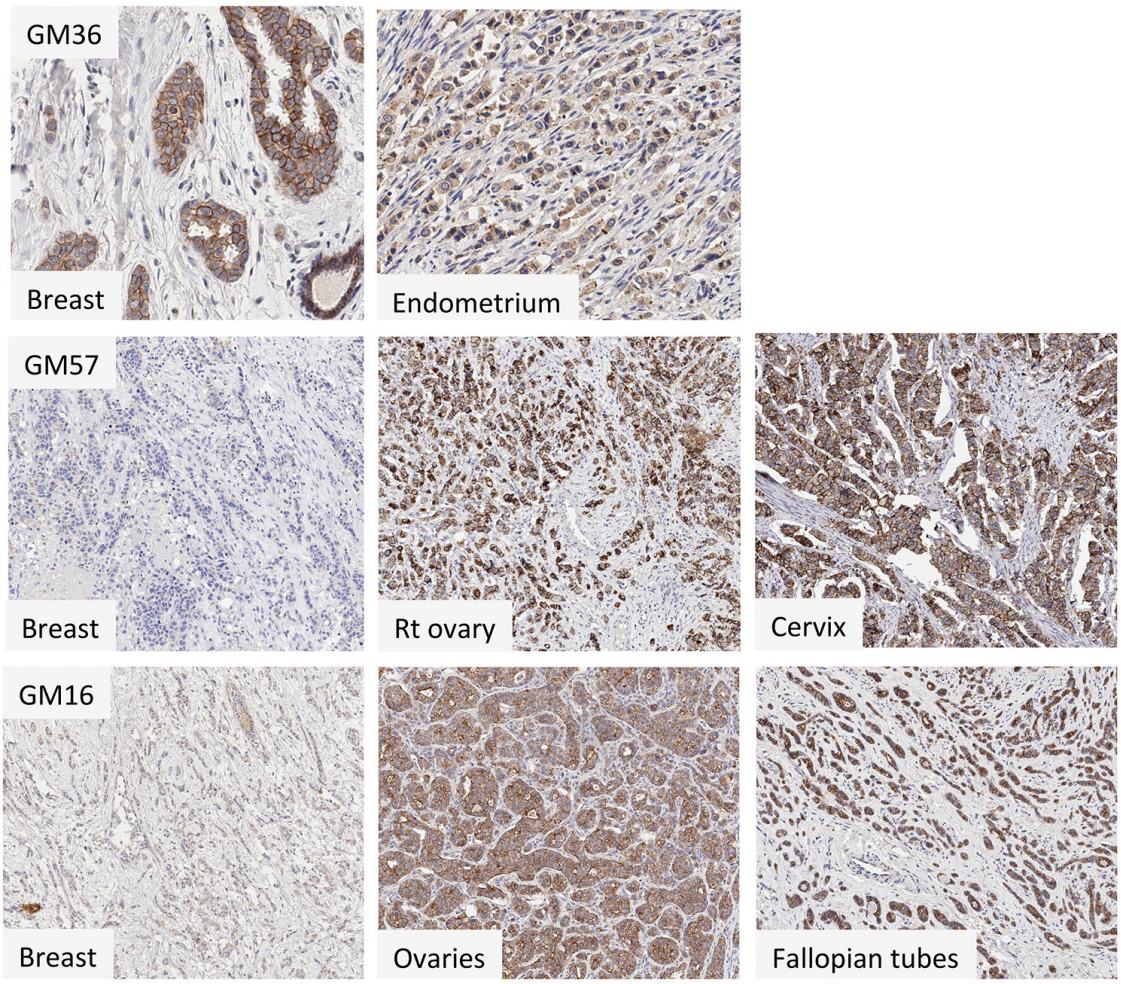


Figure 5.8: Heterogeneous E-cadherin expression in 3 cases of IC-NST with gynaecological metastases.

5.2.3. Concordance of biomarker expression during metastasis to gynaecological sites

As described in **Section 1.10.5**, phenotypic heterogeneity has been identified to occur during breast cancer progression. Particularly, ER and PR are frequently down regulated during progression to the lung, bone and liver (Cummings et al., 2014). To that end, we analysed the change of expression of biomarkers during progression to gynaecological sites on a case-by-case basis where the primary tumour and metastases were available (data summarised in **Figure 5.9**). Concordance of ER (n = 37) and PR (n = 28) expression between matched cases with data from the primary tumour and metastatic sites was analysed. Eighty-six per cent (32/37) of cases expressed ER in the primary tumour and all metastatic sites. In 4 cases, partial concordance was observed, where ER expression was the same as the primary tumour in some metastases but different in other sites (10.8%). One case was discordant and interestingly the primary tumour was ER negative and the

metastases gained ER expression. Of the 5 cases showing some degree of discordance, ER was lost in 3 and gained in 2 cases during progression (**Figure 5.9**). Concordance of PR expression between matched primary and metastatic tumours was seen in 12/28 cases (42.9%). Eight cases had partial concordance and 8 cases were discordant. Down-regulation/loss of expression of PR was also the most frequently observed change during progression (11/28, 39.2%), while in 5 cases, the metastases gained expression of PR.

AR did not change during progression in 8/13 cases (61.5%). Of the cases that were partially concordant (n = 2), or discordant (n = 3), AR was most frequently down regulated in the metastases compared to the primary tumour (n = 3). FOXA1 and GATA3 showed similar patterns during metastatic spread, with frequent concordance (GATA3: 8/13, 61.5%; FOXA1: 10/14, 71.4%). Of the cases with change of expression in at least one metastatic site, GATA3 or FOXA1 were down regulated in the metastases compared to the primary tumour (**Figure 5.9**).

For p53, 38.5% (5/13) matched cases showed complete concordance between the primary and metastases. Four cases showed partial concordance and 4 cases were discordant between the primary tumour and metastases. In the cases with discordant p53 expression it was most frequently observed that the primary tumour was negative and metastases were positive for p53 (n = 7; **Figure 5.9**). Ki67 expression in 11 matched cases showed a similar pattern to p53, where 36.4% of cases were concordant. In the cases that were partially concordant (n = 3) or discordant (n = 4), Ki67 expression was most frequently negative in the primary tumour and positive in the metastases (n = 5; **Figure 5.9**).

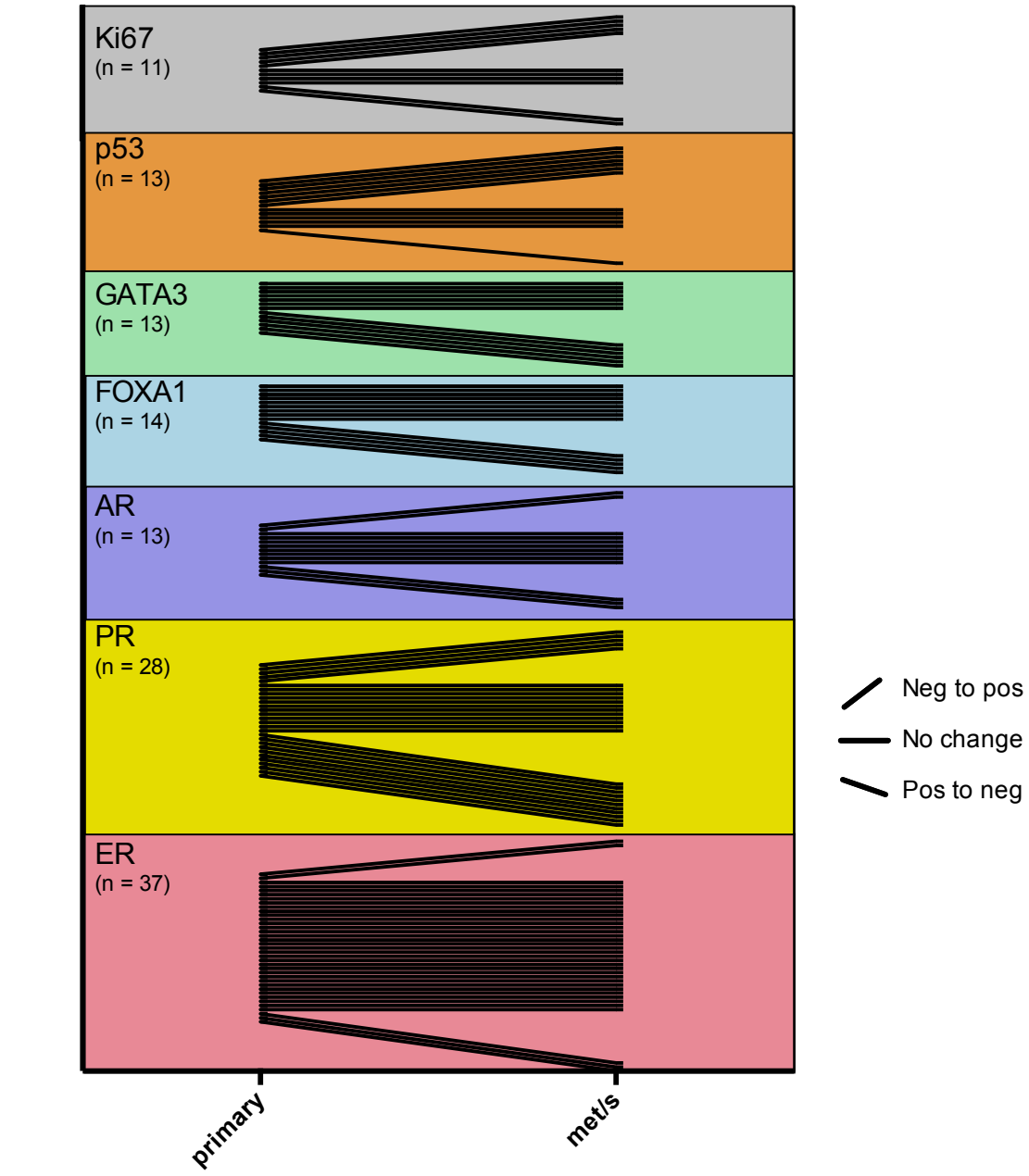


Figure 5.9 Biomarker expression during progression to gynaecological sites in matched cases. This graph includes cases with partial concordance to show if expression does change in even one site, was it up or down regulated. Each line represents an individual case.

The data (summarised in **Figure 5.9)** shows that for most cases ER, PR, AR, GATA3 and FOXA1 are all expressed in the primary tumour and expression is maintained during progression to gynaecological sites (31/99, 31.3%). As has been found for ER and PR in

other similar studies (Cummings et al., 2014), there is sometimes loss of expression of any of these markers during progression; most notably we observed loss of GATA3 (22/99, 22.2%) and PR expression (20/99, 20.2%) in metastases. All of the raw data is summarised in **Appendix Table 5.4**. In summary, most primary tumours appear to be hormone dependent and many of the metastases (31/99) maintain the expression of ER, PR, GATA3 and FOXA1. There seems to be a pattern in which all of these markers are down regulated in the metastases to some degree. PR and GATA3 loss are the most frequent markers that are down regulated (20/99 and 22/99, respectively; **Figure 5.10**).

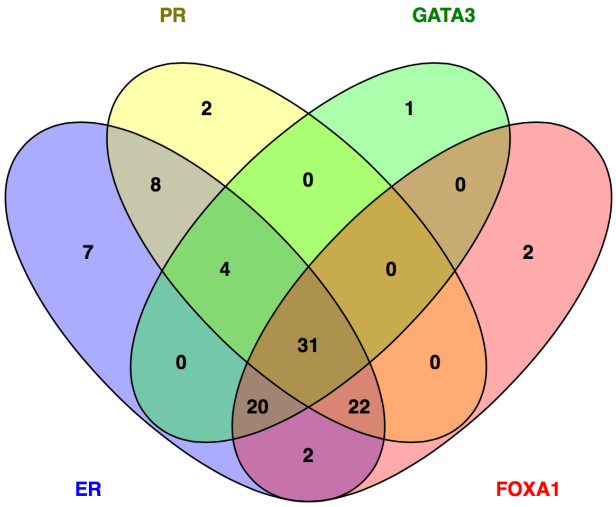


Figure 5.10 Expression relationships between ER, PR, GATA3 and FOXA1 in gynaecological metastases (a total of 99 metastatic sites from 54 patients). Venn diagram of metastatic sites with positive expression of said markers.

5.2.4. Genomic analysis of primary breast tumours and gynaecological metastases

A genomic approach was also undertaken to determine whether there are specific molecular alterations associated with metastasis to gynaecological sites. For instance, whether there are alterations in the primary tumour that might predict metastasis risk or if there are alterations in the metastases that reflect clonal selection (*e.g.* driving treatment resistance) during progression. DNA was extracted from archival FFPE tumour tissue from 27 patients and following quality control (see **Sections 2.8 to 2.10**), samples were sent to the Ramaciotti Centre for Genomics for array-based Comparative Genomic Hybridisation (aCGH) on the Agilent platform to measure genome wide copy number alterations (CNA)

(**Section 2.11.3**). A total of 84 samples underwent aCGH (**Table 5.3**); 68 samples had sufficient DNA yield and concentration to be analysed on the 180K array, and 16 lower yield samples on the 60K array (which has lower DNA requirements). We have previously used the 60k array using DNA extracted from primary breast and metastatic samples obtained from autopsies performed as far back as the 1960s (Cummings et al., 2014).

Table 5.3 Summary of the tumour sites that underwent copy number profiling.

Site	*n*	Site	*n*
Breast	10	Vagina	2
Ovary	42*	Vulva	2
Fallopian tube	10	Colon	2
Cervix	5	Omentum	2
Endometrium	5	Appendix	1
Lymph node	3		

*In 8 cases, the ovarian metastasis was sampled in two areas of the tumour

Quality control (QC) data (based on several metrics such as hybridisation intensity and background noise) generated by the Agilent software following array hybridisation initially indicated that the experiments had in general performed poorly, particularly on the 60K array, where no samples passed QC. There was also some evidence of a batch effect, with the second batch of 180K arrays having a slightly higher success rate than the first batch of 180k arrays (9/22 *v.s.* 2/20). This may suggest some technical issues with the sample labeling/array processing

We investigated whether there was any correlation between array success and QC we performed in house, such as Nanodrop and Qubit readings and the Infinium HD FFPE QC assay. All samples performed well in the Infinium QC assay and all of the Nanodrop 260/230 ratios were within the desired range (2.0-2.2), indicating that the samples were free from contaminants. Data obtained following the labelling reaction suggested incorporation of Cy3 dye was suboptimal and this showed some correlation with the hybridisation signal intensities obtained. Discussions with the Ramaciotti Centre and Technical Specialists at Agilent to resolve these issues are ongoing and we are awaiting results from trial experiments that have been performed to try and improve the dye labelling efficiency using technical repeats of samples already processed that had either worked well or had worked poorly.

5.2.4.1. Copy number profiling of breast cancers and gynaecological metastases

Although the QC data indicated the arrays had not performed well, some of the arrays were analysable and we have begun to analyse aCGH data using the Agilent's Cytogenomics software package version 3.0 and other existing CNA analysis pipelines in *R* to try and get some DNA copy number data. For instance **Figure 5.11** shows the whole genome copy number profiles obtained from the180k arrays of one case of primary tumour and multiple metastases. The figure shows some variability in data quality yet the ability to still identify copy number alterations.

All of the samples (including those on the 60K array) were analysed by the bioinformatics student in our lab. Across 10 primary breast tumours, the most frequent alterations observed (in over 75% of cases) were gains on 1q, 8p, 8q, Xp. Other frequent gains were on 4p, 11q, 16p, 17q and 19. Frequent losses occurred on 1p, 4p/q, 6p/q, 8p, 11q and 16q. Amplifications (high level gains) were detected at 1q, 8p and 22q in over 50% of cases (**Figure 5.12**. GISTIC (Genomic Identification of Significant Targets in Cancer (Mermel et al., 2011); **Section 2.11.3**) analysis was performed to identify the most significantly amplified regions and genes within the cohort. 22q11.23 amplification (*ADORA2A, SPECC1L, UPB1*) was the most significant result, followed by amplifications on 8p11.23 (*ADAM32*) and 4p14 (*APBB2, UCHL1, LIMCH1, RBM47, NSUN7*). There were no significant regions of homozygous deletion in these cases.

In total, there were 42 ovarian metastases profiled from 34 cases and copy number gains were observed in over 75% of cases on 1q, 2p, 8p, 11q, 16p and Xp. The most frequent copy number losses observed were on chromosomes 6, 13 and X. Amplifications were identified on 1q, 2q and 8q in over 60% of samples (**Figure 5.13**). The most significant amplifications as determined by GISTIC analysis were located on 2q11.2 (*PDCL3*), 11q12.3 (*POLR2G, TAF6L, ZBTB3*), 1q23.1 (22 genes, see **Appendix Table 5.5**) and 8q23.3 (*TRSP1*). There were also 21 metastases from other gynaecological sites and 11 metastases from non-gynaecological sites copy number profiled. The copy number profiles from both the gynaecological and non-gynaecological sites are similar to that seen in **Figure 5.13** (see **Appendix Figures 5.1** and **5.2**). Significant regions of amplification were shared between the primary tumours and ovarian metastases on 2p24.1, 2q11.2, 4p14, 4q31.3, 7q21.11, 8p11.23, 8q23.3, 12q24.22, 14q32.33, 22q11.23, and Xp22.33.

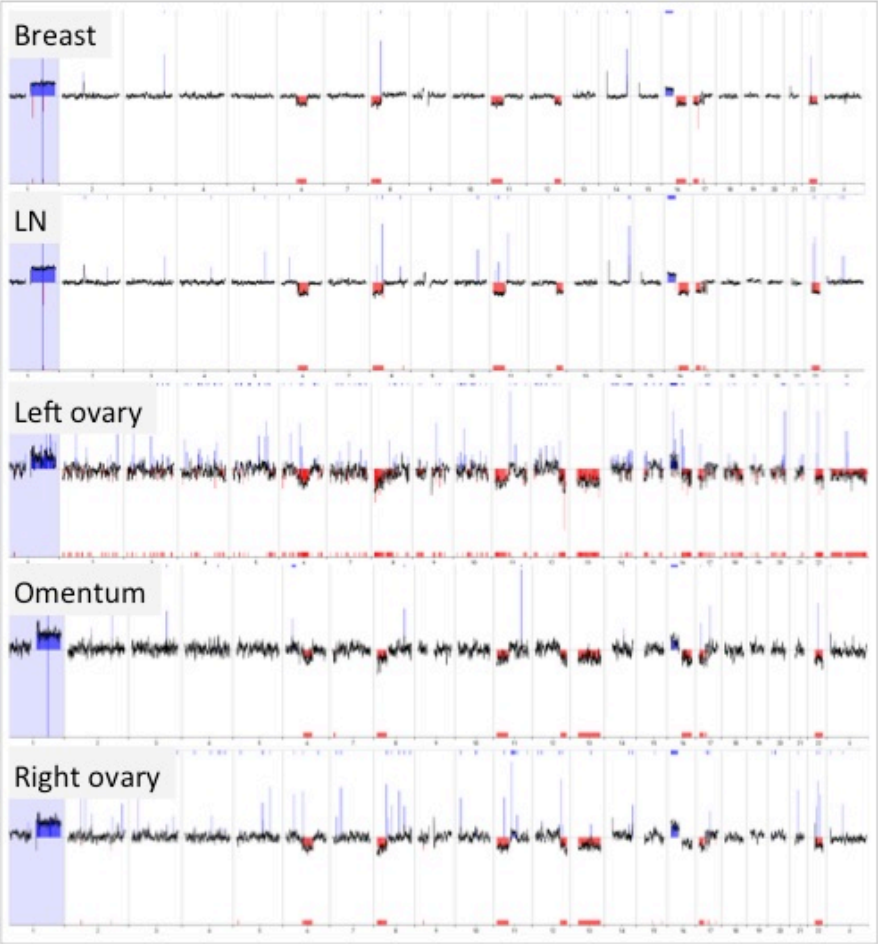


Figure 5.11: An example of the variability in data quality obtained from a case (GM63) of primary tumour and multiple metastases analysed using the 180K array. The primary tumour and lymph node metastasis was excised in 1993 and performed better then the distant metastases removed in 1999. Data was analysed using Cytogenomics software. The X axes of each plot shows chromosomes from 1 through to X; the Y axes of each plot represents the log 2 ratio of the tumour sample to the normal reference sample. The black line is a smoothed log 2 ratio of the copy number profile of each tumour. When the copy number increases above 3 copies then a copy number gain is called (+1), and greater than 4 copies is called an amplification (+2) (see highlighted blue regions in chromosome 1q and 16p of each tumour). When a copy number decrease below -1 log 2 ratio then a heterozygous deletion is called, and -2 is a homozygous deletion (see highlighted red regions in chromosomes 6q, 8p, 11p, 12q, 16q, 17p and 22 in each tumour). Despite the variable quality of array data, DNA CNAs were still identified with some confidence.

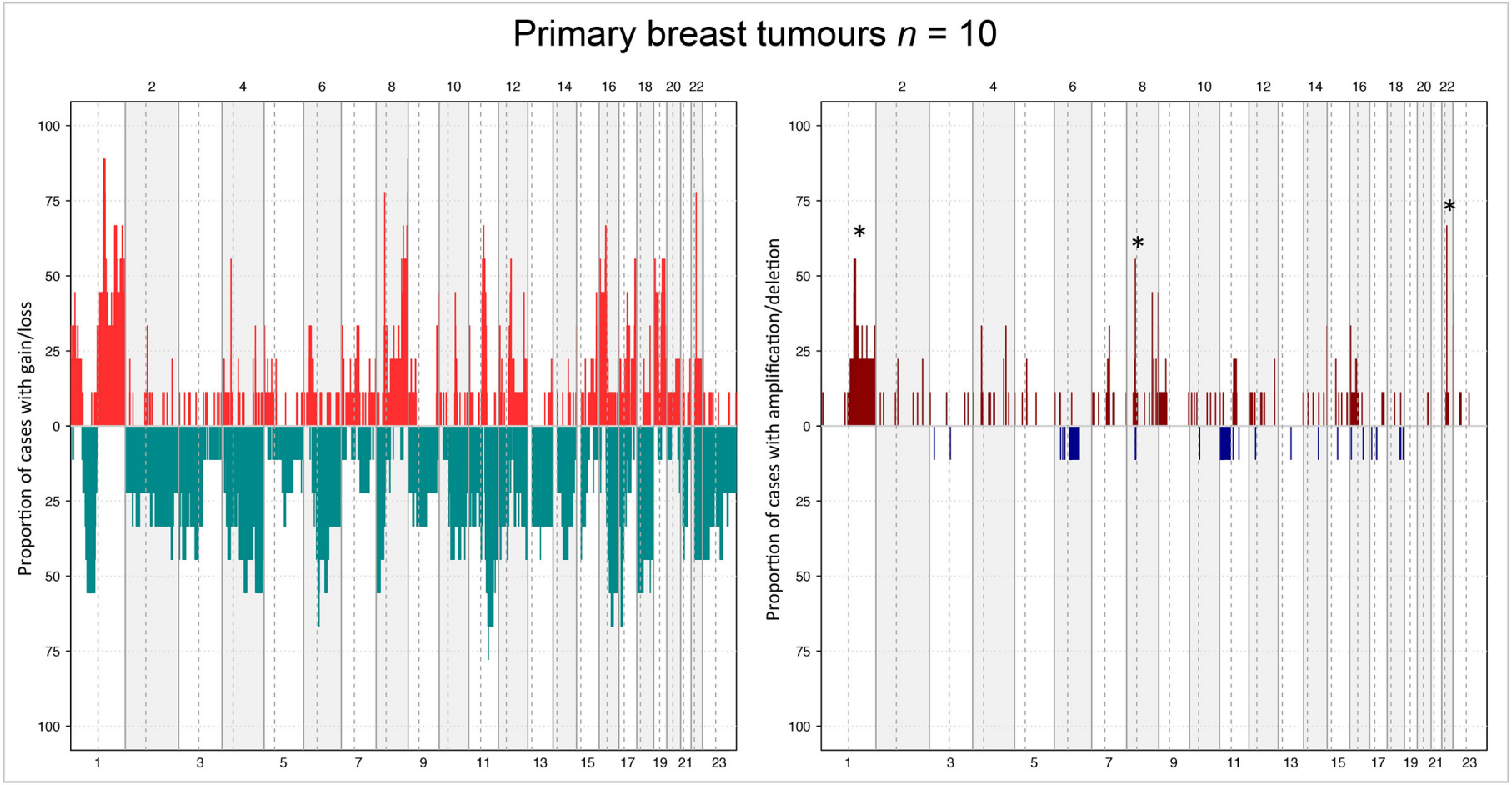


Figure 5.12 Frequency plots of DNA copy number alterations across all chromosomes in 10 primary breast tumours. Left panel displays copy number gains and loss; the Y-axis refers to the proportion of cases with a gain (shown in red) or a loss (shown in green) at a particular locus in the genome. The right panel displays the frequency of amplifications (maroon) and homozygous deletions (blue) across the genome. * Refers to the most frequent focal gene amplifications. Note chromosome '23' refers to chromosome X.

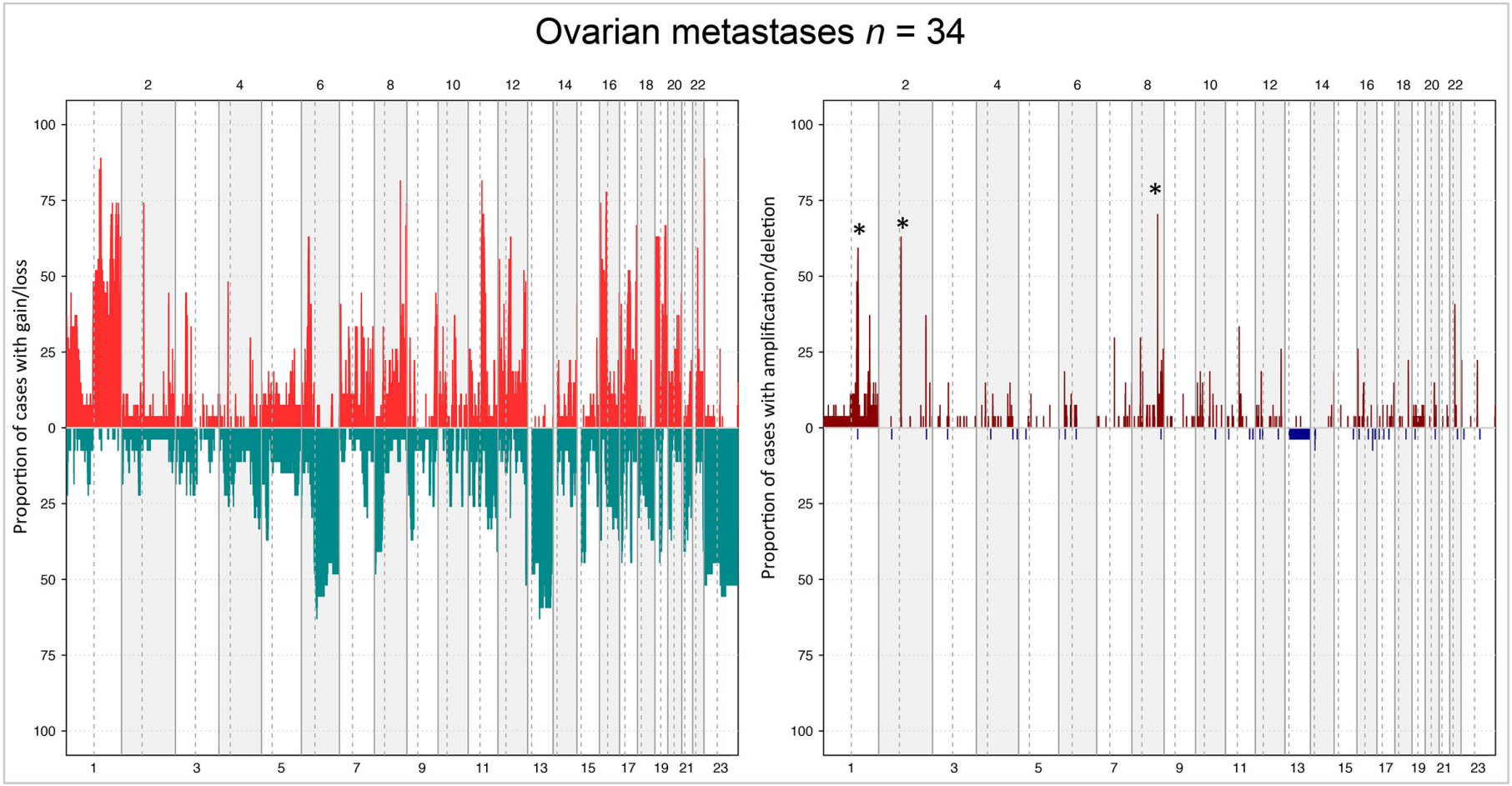


Figure 5.13 Frequency plots of DNA copy number alterations across all chromosomes in 34 ovarian metastases. Left panel displays copy number gains and losses; right panel displays amplifications and homozygous deletions.

5.2.5. Genomic evolution during metastasis to gynaecological sites

Figures 5.14 and **5.15** summarise the copy number changes across the genome in cases with a primary tumour and matched metastases. High-amplitude genomic gains (amplifications, or copy number gains of higher than 6 copies) across the 9 matched cases are presented in **Figure 5.15**. Analysis of this data starts to give some insight into tumour evolution during progression. Overall, within each case there are regions of CNAs that are shared between the primary tumour and all metastases indicating that the metastases are clonally related to the primary tumour, as expected. There are also CNAs that are unique to the primary tumour or to the metastases indicating that some level of clonal diversity has occurred during metastatic progression.

For example, there are numerous CNAs shared between the primary tumour, a left ovary metastasis and a right ovary metastasis in case GM06: gains of 8q and loss of 8p 16q and X. However, loss of 4p, focal gain and focal amplification of 1q and 11q were unique to the metastases and were not detected in the primary tumour (**Figure 5.14 and 5.15)**. Similarly there are multiple common CNAs shared between tumours in case GM63, yet loss of 13q is unique to the distant metastases and is not detected in either the primary tumour or the co-incident, regional lymph node metastasis. This data may also provide evidence of both linear and parallel evolution (**Section 1.10.4**). In cases GM06 and GM59, all of the CNAs in the primary tumour were present in the metastases, but the metastases have more CNAs, suggesting that the tumour clone that seeded the metastases left the primary site late during progression of the primary tumour (linear progression). Amplification of 11q13 (which involves the commonly amplified gene *CCND1*) in the metastases may be a driver alteration promoting progression in GM06. GM59 has gains in regions that contain the driver oncogenes *CCND1* and *MYC* that may also be important. On the other hand, cases GM29 and GM57 have evidence of shared CNAs as well as CNAs that are unique to both the primary tumour and the metastases. For example, whole arm gain of 22q and amplification of the region containing *CCND1* was present in the primary tumour but not present in the metastasis of GM29. This suggests that the clone the seeded the metastasis may have left the primary tumour early in the course of disease progression (parallel progression) and both the primary tumour and metastases continued to accumulate genomic alterations before they were detected.

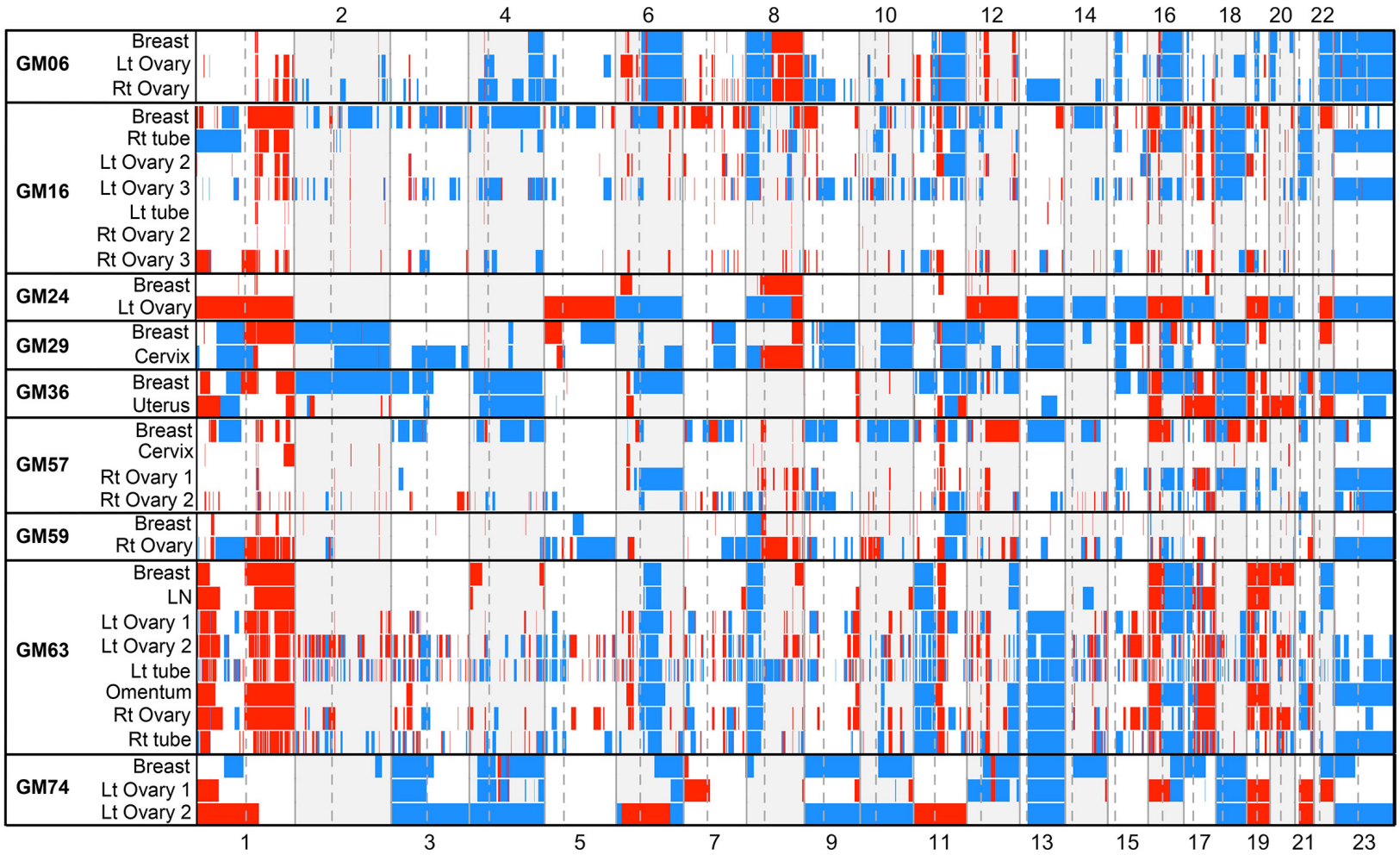


Figure 5.14 Summary of copy number alterations within 9 matched primary and metastases samples. Red areas are regions of copy number gain, while blue areas indicate copy number loss. Chromosome 23 is Chromosome X. The tenth matched primary and metastasis failed the experiment.

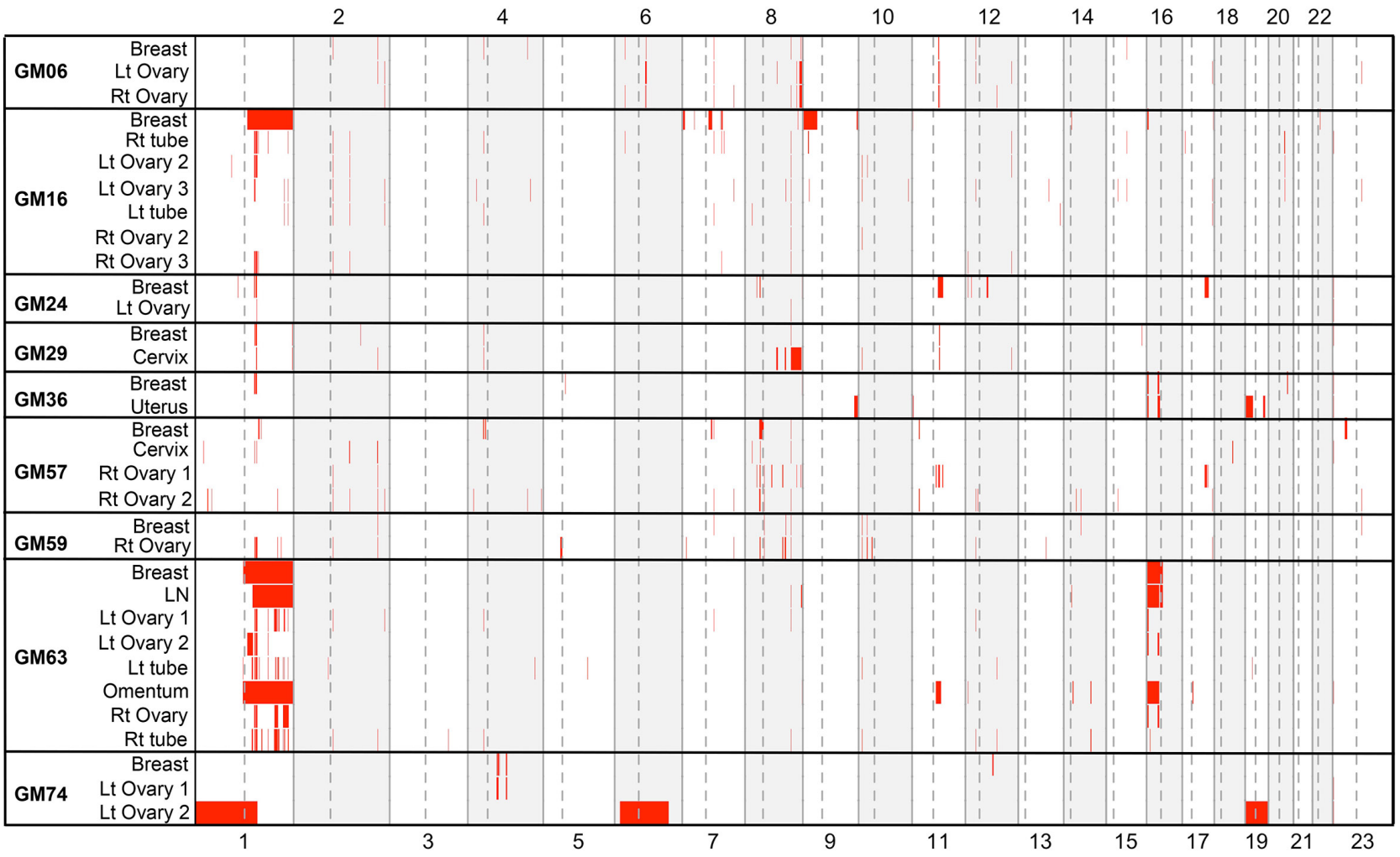


Figure 5.15 Summary of genomic amplifications within 9 matched primary and metastases samples. Red areas are regions of high-level copy number gain, known as amplifications. Chromosome 23 is Chromosome X.

5.2.5.1. Mutation analysis of breast tumours and matched gynaecological metastases

To complement the DNA copy number profiling we undertook DNA sequencing to identify gene mutations that may contribute to metastasis to gynaecological sites. Exome sequencing was attempted for six cases with primary and multiple metastatic tumours and matching normal DNA obtained from normal FFPE blocks (e.g. uninvolved lymph node). Our lab has had success in sequencing FFPE–derived tumour in a range of projects. For this project we sent the FFPE DNA to Macrogen (**Section 2.11.2**) based on competitive price and turn around time. Unfortunately all the samples failed QC metrics employed by Macrogen during library preparation; at our request a single case (GM16) was taken through to Agilent SureSelect exome library preparation, but this also failed. We then decided to trial one case (GM63) using the Illumina Nextera exome sequencing protocol at Queensland Centre for Medical Genomics (QCMG; **Section 2.11.2**). Sequencing libraries were successfully generated from the primary tumour, four metastatic sites and matched normal DNA. Preliminary data indicates that this case performed poorly, each sample had low sequencing coverage and therefore mutation calls could not be obtained accurately.

In understanding the challenges of sequencing DNA derived from FFPE material, we also designed, in parallel, an Agilent SureSelect custom targeted sequencing assay to sequence 45 genes (**Section 2.11.1, Table 2.3**). Baited capture-based targeted sequencing is a technique that can be utilised to sequence degraded FFPE DNA (**Section 2.11.1**). Both FFPE and fresh frozen DNA samples were used in a trial to compare the data quality between these types of sample and to test the assay design for gene coverage. A first pass analysis of the sequencing data revealed that the read depth of each gene was too low to confidently call any somatic mutation (**Table 5.4**); only four samples had and average of over 100 reads across all targets, therefore the samples will be sequenced again to obtain greater gene coverage.

Table 5.4 Summary of samples sequenced and the read depth in the SureSelect capture-based custom sequencing panel.

Sample #	Sample	DNA source	Average read depth	Median read depth	Range read depth
1	GM16 NLN	FFPE	158	135	57-496
2	GM16 Br	FFPE	42.96	42.5	9-57
3	GM16 LO2	FFPE	56	55	55-58
4	GM16 LT1	FFPE	61.5	58	53-100
5	GM16 RO2	FFPE	78	70.5	50-184
6	Gut met 1	FFPE	795	602	340-2845
7	Gut met 2	FFPE	69	58	39-265
8	Q330	Frozen	80	69	5-327
9	ILC-02-09-078	Frozen	359	296	45-755
11	GM59 RO	FFPE	53	48	39-81
12	GM59 Br	FFPE	47	49	11-82
10	GM57 Ovary	FFPE	673	58	19-118
13	GM57 Br	FFPE	0	0	0
14	GM13_RO	FFPE	61	63.5	5-124
15	GM74_LO	FFPE	53.9	39	6-204
16	GM74_Br	FFPE	56.9	53	6-204

NLN: normal lymph node. Br: breast. LO: left ovary. LT: left fallopian tube. RO: right ovary.

5.3. Discussion

It is important to understand the natural history of metastatic disease, as it will impact the prognosis of the patient. The most common sites of metastasis from breast cancer are the lungs, liver, bones and brain, with some differences in distribution according to the characteristics of the primary tumour (as described in **Section 1.10.3**). For example, breast tumours that spread to the brain tend to be associated with high grade IC NST of HER2 or basal-like subtypes; metastases occur relatively quickly, with a latency of less than 5 years (Lagerwaard et al., 1999). Conversely, data accumulated here and from the literature (Bigorie et al., 2010, Demopoulos et al., 1987, Ferlicot et al., 2004, Borst and Ingold, 1993, St Romain et al., 2012, Cummings et al., 2014) suggest metastasis to gynaecological organs is associated with different primary tumour characteristics and a different clinical pattern of progression to that seen for brain metastasis. Some of the key findings from this study and supporting literature are:

- Metastasis to gynaecological organs affects young women who are likely to be pre-menopausal at the time of their primary diagnosis; in our cohort the median age of primary tumour diagnosis was 47 years, and the median age of first metastasis diagnosis was 55 years.

- The interval between primary breast diagnosis and metastasis can be between 5 to 20 years (Bigorie et al., 2010, Demopoulos et al., 1987). In our cohort, the median progression free survival time was 5 years (ranging 0-20 years). Therefore, for some patients metastasis is frequently diagnosed outside the typical follow-up window of 5 years (**Section 5.2.1**).
- The primary tumour can be either IC NST (48.1%) or ILC (42.6%) histological type, which represents a significant enrichment in ILC spreading to these tissues relative to the incidence of ILC in the general breast cancer population.
- The immunophenotype of the primary tumours was associated with features typical of good prognosis disease: ER and PR positive, negative for HER2 and basal markers and low Ki67 expression (i.e. a luminal A phenotype).
- The majority of patients (83.3%) had dissemination to the regional lymph nodes at the time of primary tumour diagnosis.
- Patients with gynaecological metastasis also presented with a unique pattern of dissemination. The frequent sites of breast cancer metastasis were rarely involved (lung, liver, brain) and the high frequency of GI, omentum and peritoneum metastases suggests a unique mechanism behind metastases to these sites compared to other organs (**Figure 5.1, Appendix Table 5.1**).
- It is also interesting that the ER positivity observed in the primary tumours is maintained during progression (**Figure 5.3)**; we and others have previously shown that ER and PR are frequently down-regulated during spread to other organs such as the liver/lungs (Cummings et al., 2014) and so raises the possibility that the ER positive breast tumour cells might be specifically homing to an oestrogen rich environment of the ovaries in these young patients and colonising other tissues along the way.

Understanding the patterns of metastasis associated with different primary tumour characteristics in this way will therefore lead to the ability to recognise patients with a greater risk of relapse to particular sites and hence who may need closer surveillance.

5.3.1. Endocrine signalling in metastatic progression

Obtaining complete treatment information for the patient cohort was difficult. Where available, most patients received hormonal-based therapy and chemotherapy for treatment of their primary and secondary disease. Although we can not be sure, it is reasonable to assume that most other patients would have received anti-oestrogen therapy, since most

primary tumours were ER positive and diagnosed in the 1990s, when tamoxifen was becoming the standard of care for ER positive breast cancer patients. Progression of disease implies the development or pre-existence of a subclone of tumour cells that are resistant to endocrine and chemotherapy. This series of patients therefore represents an opportunity to study mechanisms underlying resistance to therapy. We hypothesise therefore that mutations or deregulation of genes involved in endocrine signalling may have occurred during tumour evolution and progression to gynaecological sites and this may be a result of treatment induced selection pressures.

As mentioned above ER and PR down regulation in metastasis may contribute to treatment resistance in certain organs (Cummings et al., 2014). Recent studies have also identified mutations within the *ESR1* (ER alpha) gene in metastases from ER positive disease, which was not identified within the primary tumour, contributed to therapy resistance and disease progression (Jeselsohn et al., 2015). Other studies have also suggested that deregulation of ER regulatory proteins (such as GATA3 and FOXA1) may contribute to endocrine resistance, as described in more detail in **Section 1.3.1**. Indeed both GATA3 and FOXA1 are frequently mutated in breast cancer, particularly those of a luminal/ER positive subtype (The Cancer Genome Atlas Network, 2012). It would be of interest therefore to determine if a mutation in *ESR1*, *FOXA1* or *GATA3* in the primary tumour provides a mechanism of *de novo* endocrine resistance or can give insight into the outcome of the patient.

As described **Sections 1.3 and 1.3.1**, the pioneer factors GATA3 and FOXA1 play a pivot role in the transcriptional activity of ER, and *in vitro* studies have demonstrated their role in endocrine resistance. Of note, FOXA1 was found to redirect ER binding to different areas of the genome in drug resistant tumour cells compared to the primary tumour (Ross-Innes et al., 2012). It was of interest therefore to determine the expression of these factors in our cohort, as nearly all of the tumours were ER positive. GATA3 and FOXA1 are expressed in 100% of the primary tumours in our cohort, and this was significantly higher compared to the unselected sporadic breast cancer cohort. Interestingly, only 73.1% of the metastases expressed FOXA1 and 54.3% expressed GATA3 (**Section 5.2.2.3**), suggesting the loss of these proteins occurs during metastatic progression in some cases. The interplay between these pioneer factors and ER and PR is complex, but interestingly of the 26 cases matched cases with the most complete IHC data to date (displayed in **Appendix Table 5.4**), 22/26 (84.6%) of cases had down-regulation of one or more of these four molecules

during progression.

The mechanism involved in the down-regulation of GATA3 or FOXA1 in the metastases may be the result of mutation or other regulatory mechanisms such as epigenetic or protein signalling pathway changes. It will be of interest therefore to determine the mutation status of *ESR1*, *FOXA1* and *GATA3* in this unique patient cohort and to see if the mutation status correlates with changes in protein levels between the primary tumour and metastases. This was attempted with the various sequencing approaches in this study, and will be addressed after the targeted gene-sequencing assay has been fully optimised.

Since many of the patients in this cohort maintained expression of ER, FOXA1 and GATA3, gene expression profiling is needed to determine if the ER transcriptional program between the primary tumour and metastases has changed. We may have protein based evidence of FOXA1 redirecting ER in the genome in the cases that have lost expression of PR but are positive for ER, GATA3 and FOXA1 (For examples see **Appendix Table 5.4**). *PR* is a well established target gene of ER. We could hypothesise that in these metastatic lesions that ER may be directing a unique transcriptional program.

A comparative study using chromatin immunoprecipitation followed by sequencing (ChIP Seq) of metastatic samples that are either positive for ER, FOXA1 and GATA3, or are negative for either FOXA1, GATA3, or both, would be very interesting to see whether there are different ER-DNA binding events in these metastatic samples (Hurtado et al., 2011, Ross-Innes et al., 2012, Mohammed et al., 2015). This would need to be complemented with gene expression profiling of the same groups to assess the transcription programs. This information would provide great insight into the mechanisms driving tumour cell survival during progression. However, many ChIP Seq and gene expression technologies are only useful for high quality DNA and RNA. Therefore the utility of these technologies on FFPE derived DNA and RNA may be limited until technologies are improved to handle poorer quality genetic material. One promising technology is NanoString®. However, it uses a targeted approach where a maximum of 800 genes can be analysed at one time, whereas a whole genome approach would be more appropriate for discovering changes in gene expression and remove selection biases.

5.3.2. Genomic aspects of metastatic progression

In order to gain a greater understanding of tumour evolution during progression, aCGH and sequencing was performed (**Section 5.2.4**). We hypothesised that a genomics approach would i) provide evidence of clonal relatedness of the primary tumour and matched metastases; ii) identify DNA copy number alterations or gene mutations that could drive an aggressive behaviour or treatment resistance; and iii) identify potential targetable genomic alterations to improve treatment for metastatic breast cancer patients.

There are several clinical trials ongoing, such as the AURORA study, where molecular screening programs of metastatic breast cancer is being performed to help direct targeted treatment options (Zardavas et al., 2014). A cancer panel is used utilising DNA from both archival and prospectively collected metastatic biopsy tissue. The results will guide treatment options and the patients will be followed for 10 years. The group will focus on exceptional responders and rapid progressors to try identify new markers of prognostic and predictive utility (Zardavas et al., 2014). The data generated from the AURORA study is anticipated to provide insight into many aspects of metastatic disease, particularly to improve patient outcome with more appropriate treatment. Another recent study analysed biopsy material of metastases using aCGH and Sanger sequencing for *PIK3CA* and *AKT* hotspot mutations and then implemented targeted therapy based on the findings (*e.g.* amplification of *FGFR1* and the use of FGFR inhibitor for patients with an amplification on 8p11.23) (Andre et al., 2014). Although less than 10% of patients received objective response to targeted therapy it provides proof of principle that this approach may work for some patients.

We utilised aCGH as opposed to SNP-CGH as aCGH (using the ULS labelling protocol (Alers et al., 1999)) is better suited for FFPE DNA. However this study was not as successful as our recent 'autopsy' study (Cummings et al., 2014). It remains unclear as to why this cohort has performed poorly while the autopsy samples worked well (**Section 2.4.2.1**). The DNA was extracted using the same protocol specified by Agilent. All of the autopsy samples were ethanol precipitated because the DNA concentrations were too low. In this study, only 7 samples were ethanol precipitated and those samples performed no better then those that were not concentrated. The samples analysed here ranged from 2 to 26 years in age and came from pathology departments all across the state, and hence likely underwent variations in tissue sample processing during fixation and storage (Srinivasan et al., 2002) that may contribute to the variation in aCGH success. Case GM63 may represent a good example of this, in which the primary tumour and regional lymph

node metastasis (surgery in 1993) both performed well in aCGH, yet the gynaecological metastases (surgery in 1999) performed poorly; implying tissue fixation of the metastatic samples may have been suboptimal for molecular work. There may also be technical issues with the reagents supplied by Agilent or batch effect when the arrays were hybridised, and troubleshooting is currently undergoing at the Ramaciotti Centre to investigate this. Recent studies have suggested a method to improve labelling efficiency (Salawu et al., 2012, Craig et al., 2012). We trialled one sample following these methods, performing a 95°C fragmentation step prior to labelling, and also increasing the amount of dye to 1 μL for every 0.8 μg of DNA, however there was no improvement in the degree of labelling. It will be interesting to see how the sample performs on the array as some studies suggest the different DNA fragment sizes produces noise and irreproducible results (van Beers et al., 2006, Mc Sherry et al., 2007). Although the degree of labelling was still suboptimal, the uniformity of DNA fragment sizes may still provide interpretable results.

We first analysed the frequency of alterations across the genome in the primary tumours and the metastases at different sites (**Section 5.2.4.2**, **Figures 5.15** and **5.16)**. Most metastatic deposits analysed were in the ovary (n=34, after the duplicate cases were removed) and unfortunately there were only small numbers of tumours in other groups (e.g. 10 primary breast tumours; 10 fallopian tube metastases). The numbers were therefore too small to identify a specific copy number profile of primary tumours that spread to gynaecological sites compared to that for unselected primary tumours studied from the literature. Nevertheless, frequently altered regions in ER positive tumours, including gains on 1q and 8q and losses on 1p and 16q were observed in the 10 breast tumours profiled. The CNA profile of ovarian metastases from breast cancer **(Figure 5.16)** reveals recurrent alterations in 1q, 2p, 8p, 11q, 16p and Xp and losses on chromosomes 6, 13 and X. Some of these CNAs were more frequently identified in the ovarian metastases relative to the primary tumours (e.g. 6p gain; amplifications on 2q and 8q; losses on 6, 13 and X). However, since not all of the matched primary tumours were available for profiling, we do not know if these alterations were already present in the respective primary tumour or occurred as a result of clonal evolution during progression.

We then analysed aCGH data in individual cases where the matched primary and metastases were available. Each case presented with a unique pattern of alterations and we found evidence of both clonal similarities and diversity within matched cases (**Section**

5.2.5). This highlights the clonal relatedness between the primary tumour and resulting metastases, as expected, but also inter-tumour heterogeneity during the metastatic processes. For example, the loss of chromosome 13 in all the metastases of GM63 that was not present in the primary tumour (**Figure 5.17**). Genes of interest on chromosome 13 include the critical tumour suppressor genes *RB1* (retinoblastoma 1) and *BRCA2* (breast cancer 2, early onset). RB1 is a critical regulator of cell cycle and BRCA2 plays an important role in repairing DNA damage and hence the loss of both may contribute to the aggressive metastatic tumour behaviour in this case.

It is important to validate some of these CNAs (*e.g.* this 13q deletion) using quantitative PCR on DNA already extracted or by in situ hybridisation on the TMAs or whole tissue sections. The latter approach (*i.e.* whole sections as opposed to TMAs) would be most appropriate to account for potential intratumour heterogeneity in a given tumour. It has been recognised that the primary tumour may contain multiple subclones, and that a small subclone of the primary tumour may seed a metastasis (Yachida et al., 2010, Gerlinger et al., 2012, Cummings et al., 2014, Yates et al., 2015). This was recently demonstrated, where amplification of chromosome 10 was identified by aCGH in multiple metastases in an autopsy case but was not found in the primary tumour. Florescence in situ hybridisation (FISH) confirmed the amplification was present in the metastases but also identified a small focus of cells in the primary tumour with this amplification, suggesting this small subclone spread (Cummings et al., 2014).

As described in **Section 5.2.5**, exome sequencing was attempted on 6 cases with FFPE derived DNA from matched primary tumour and metastases to try to uncover mutations that may play an important role in progression in this unique patient cohort. The exome sequencing performed at Macrogen was unsuccessful. The samples failed their QC metrics at every stage of the process. To our knowledge, Macrogen have limited experience with FFPE DNA samples. Our samples may have had more success if there was a more individualised optimisation of the sequencing protocol, as opposed to a one-size-fits-all approach of a service provider, rather than a collaborator. The SureSelect custom targeted sequencing assay utilises broadly the same approach as whole exome sequencing, however it only has baits for the genes of interest. It has been successfully applied to FFPE samples in a number of important studies (Frampton et al., 2013, Jeselsohn et al., 2014, Meric-Bernstam et al., 2014). A sequencing assay was designed to target 45 important breast cancer genes and a pilot study was performed at QCMG where

sequencing libraries were successfully generated for all 16 samples. As we have previously worked with QCMG we believe that the success was due to greater knowledge of the challenges of these FFPE DNA samples and each sample was sheared individually to get optimal fragment sizes. The first sequencing run did not provide sufficient gene coverage data and hopefully re-sequencing these cases will provide greater depth to identify mutations. Hopefully we may be able to elucidate the frequency of *ESR1*, *FOXA1* and *GATA3* mutations (among others) in the metastases and this may contribute to our understanding of the mechanisms underlying metastatic progression and of endocrine resistance.

5.3.3. Limitations of the study and future work

Working with clinical material is crucial for this type of work, yet this is associated with numerous technical challenges related to obtaining a sufficient cohort size and the quality of tissue material for molecular based research. These limitations cannot be controlled for but they certainly affect the types of analyses and statistical comparisons that are possible.

One limitation of this study was using TMA-biopsy cores for IHC profiling, as opposed to whole tissue sections. There is the possibility that tumour heterogeneity may have been under-represented. We know this can be an issue for focally expressed biomarkers (such as basal-markers) and for scoring Ki67 staining, where on whole sections the most positive area of the section is typically counted. The use of 1 mm TMA cores was chosen to capture a larger area of tissue. Conversely, performing IHC on whole sections would improve the representation of heterogeneity, however without using all the tissue blocks of the case, it is again a compromise. The use of TMAs, therefore, is a pragmatic decision and represents the most feasible option to screen many tissue samples for a large range of markers, without expending the finite patient material.

In order to understand the different biologies that drive organotropic metastasis, a comparison with metastases to other organ sites is essential. It would be of great clinical utility to be able to identify different features of a primary tumour that may predict the risk of metastasis, and if so, to which organ. However, as our understanding of tumour evolution increases, that may not be feasible without the ability to identify every subclone within a primary tumour. Next generation sequencing technology is ideally suited to resolve these issues and its cost is reducing, however data analysis bottlenecks are

becoming increasing obvious. The application of these technologies in the neoadjuvant setting is highly desirable but also limited, as sequencing a tumour biopsy is unable to define all the tumour subclones necessary to appropriately guide targeted therapy.

It would also be interesting to perform a comparison between ovarian metastases that originated from breast primary tumours with colorectal primary tumours, since breast and colorectal are the most common primary tumours to spread to the gynaecological sites. Due to the close proximity of the colon to gynaecological sites, seeding in the abdominal cavity could be a major contributing factor to this type of metastatic spread. However, the most common site of colorectal metastasis is to the liver and lung, where blood flow is likely to a driving force (Patanaphan and Salazar, 1993). Therefore we could potentially uncover common mechanisms that contribute to colonisation of the ovaries that are shared by all primary tumour types.

Some of these patients' tumours relapsed after a long time (over 10 years in some cases). This may be a result of a slow growing tumour, or suggests that mechanisms of tumour dormancy need to be investigated. Tumour cells have been found in the bone marrow of cancer patients and were associated with distant metastasis (Braun et al., 2000). This suggests that the bone marrow is a haven for tumour cells to lay dormant, and could potentially be used as a marker to predict metastasis. Studies in mouse models have found that the perivasculature in the bone marrow, lung and brain plays a role in dormancy during metastasis to these sites (Ghajar et al., 2013). The different stages of normal vasculature progression/homeostasis produced signals that either suppressed (microvasculature) or induced (sprouting neovasculature) breast cancer cell growth (Ghajar et al., 2013). These studies highlight the importance of the microenvironment in metastatic progression. For example, it has been recently shown HER2/HER3 expressing breast tumour cells may selectively grow in the brain microenvironment were the HER2/HER3 ligand neuregulin is expressed in abundance (Saunus et al., 2015, Momeny et al., 2015). Similarly, we hypothesise that the oestrogen rich environment of the ovaries in premenopausal women may be involved in homing of the ER positive breast tumour cells. Studies have shown that oestradiol treatment in hormone dependent breast cancer cell lines induces a migratory phenotype (Li et al., 2010c, Giretti et al., 2008). Therefore, an interesting future direction would be to perform *in vitro* experiments, such as transwell experiments, or co-culturing with 'normal' ovarian cells, and under a gradient of oestrogen treatment, assess if tumour cell migration is effected by different oestrogen concentrations.

These types of experiments could also be used to study the effect of mutations in genes such as *ESR1*, *FOXA1* and *GATA3*, to further tease out the role of these important proteins in the metastatic process.

This information could lead to providing knowledge to better determine which women may benefit from prophylactic oophorectomy, as it is not routine practice to add to adjuvant therapy (Griggs et al., 2011). Recent studies have shown however that ovarian suppression combined with an aromatase inhibitor does reduce risk of recurrence in ER positive patients (Pagani et al., 2014, Francis et al., 2015). However, it would be of great benefit to omit a patient from an oophorectomy if it is not required, and therefore saving a patient from unnecessary surgery and the risks associated with hormone replacement therapies.

5.4. Conclusions

In summary, this study adds to our understanding of the clinical and phenotypic characteristics of primary breast tumours and metastases to gynaecological sites. Although patients with gynaecological metastases largely maintain their immunophenotype compared to other organ sites, treating based on the characteristics of the primary tumour is still likely to not be appropriate because of changes in the complex relationships between ER and its regulatory molecules in the metastases. The tumour cells have still evaded therapy, and hopefully the genomic data collated here may provide an explanation for the different behaviour despite the similar phenotype.

CHAPTER 6

DISCUSSION

6. Discussion

Breast cancer encompasses a heterogeneous group of diseases and this thesis aimed to ask and address three major questions surrounding breast cancer biology. Firstly, are there other mechanisms involved in E-cadherin deregulation in invasive lobular carcinoma? Secondly, how does clonal evolution progress in tumours with mixed ductal and lobular morphologies? Finally, can we gain a greater understanding of clonal evolution and heterogeneity underlying breast cancer metastasis by studying a unique cohort of breast cancer patients with metastasis to gynaecological sites?

6.1. Actin cytoskeleton regulating cell-adhesion in breast cancer – challenges validating *in vitro* findings *in situ*

The loss of E-cadherin is a defining feature of invasive lobular carcinomas. Understanding the mechanisms underlying the loss of cell-cell adhesion and anoikis resistance is essential in order to begin to target this unique type of breast cancer. As described in **Section 1.8.1**, genomic alterations to *CDH1* (such as mutation, chromosome loss and methylation) contribute to a large proportion of E-cadherin loss in ILC. However there are a significant number of cases without detectable alterations to *CDH1* and therefore the mechanism of E-cadherin deregulation remains unknown. We therefore hypothesised, that deregulation of actin cytoskeletal regulatory molecules (which have been shown *in vitro* to be essential for E-cadherin function) may contribute to E-cadherin loss and the invasive phenotype observed in ILC. IHC analysis of these markers in a large cohort of breast cancer samples revealed, most importantly, the lack of membranous localisation of the molecules as identified in the *in vitro* work by (Kovacs et al., 2011) and (Ratheesh et al., 2012). This study highlights the challenges of validating *in vitro* findings *in situ*. Patient derived tissue samples are essentially a snap shot in time and therefore validating dynamic molecular mechanisms that are likely to occur on minute time scales, within heterogeneous, multi-cellular tissues are difficult. Particularly when the model is developed in a two-dimensional cell culture system, which was established within a homogeneous cell population. There were some observations that may be important to breast cancer biology (*e.g.* cytoplasmic N-WASP protein expression was identified in E-cadherin negative ILC), however, without matched normal breast tissue for each patient, we are unable to identify if the protein localisation is aberrant and contributing to the invasive phenotype observed in ILC. Our data is not sufficient to disprove the hypothesis that the

actin cytoskeleton regulators play a role in regulating E-cadherin function in breast cancer. Therefore, the next step will involve investigating these molecules in normal breast samples using dual-immunofluorescence to identify their localisation and secondly, applying this knowledge to matched normal and tumour samples to truly understand if these molecules play a role in E-cadherin deregulation during tumourigenesis.

6.2. The clonal evolution of mixed ductal-lobular carcinomas

It has been generally accepted that MDLs represent a unique biological entity (Rakha et al., 2009, Arps et al., 2013). Large consortia such as the TCGA and ICGC have contributed to our increased understanding of the genomic landscape of breast cancer (The Cancer Genome Atlas Network, 2012, Stephens et al., 2012). These studies have included thousands of samples in their studies. The TCGA have profiled the somatic mutations, copy number profiles and gene expression of 27 MDLs. The mutation profiles of the MDLs show features similar to both ER positive IC-NST and ILC; and our data also supports these findings concluding that that MDLs have overlaps with both ER positive IC-NST and ILC, but do not entirely fit into either one of the already established categories.

Small studies have suggested that the presence of different morphological components are clonally related lesions (Wagner et al., 2009, Buerger et al., 2000) and the mechanism underlying a switch in morphology has not been explored. The purpose of this study was to discover unique biological properties that explain the morphological heterogeneity observed in MDLs. We hypothesised that disruption to the E-cadherin cell adhesion complex contributes to the change from a ductal to lobular growth pattern. By studying the expression of the E-cadherin adhesion complex in 51 MDL cases, we found that E-cadherin was aberrantly localised in the lobular component. This is in contrast to the complete loss of E-cadherin expression recurrently observed in ILC. Therefore, we concluded that the mechanism underlying E-cadherin disruption is likely to be different to that seen in pure ILC.

Due to the nature of our study, we have micro-dissected discrete morphological areas and therefore have power to tease out important subtleties that would be lost when studying large numbers of cases. The TCGA samples are unable to be used in this regard due to the lack of annotation of the cellular component from which the nucleic acids have been extracted. By carefully micro-dissecting 4 MDL cases, we were able to enrich for very pure

tumour cellularity, a critical feature essential for uncovering mechanisms that drive a change in cellular morphology. Together with the previous CGH study (Kutasovic, 2011), the preliminary analysis of exome sequencing data showed that all components within each case were clonally related. It is likely that in some cases, morphological divergence occurs early in tumour progression, while in other cases divergence occurs later. Interestingly, only one of the 4 cases had a *CDH1* mutation in the lobular components of the tumour and this tumour was one of the 2 cases sequenced with complete loss of E-cadherin protein expression. These findings support our hypothesis that other mechanisms are involved in the transition from ductal to lobular morphology. The lack of co-ordinated expression of EMT protein markers in our cohort, and the lack of gene expression alterations in the TCGA cohort, suggests that EMT does not play a role in the transition between growth patterns. Pathway analysis using the exome sequencing data revealed that alterations in gap junctions signalling pathways were frequent. This is fascinating, since gap junctions are one of the three types of cell-cell adhesion machinery (including adhesion and tight junctions) and play important roles in cell-cell communication by directly transferring signalling molecules between neighbouring cells. Future research will include the validation of the mutations discovered in each case using Sanger sequencing. The discovery of alterations in gap junction signalling will also be an interesting avenue to explore. The first step will be to pinpoint the exact gap junction related genes that were altered in each case and morphological type. Then the expression of the proteins encoded by those genes and also the connexon proteins will be assessed by IHC in the MDL cohort, and compare the expression with ILC and IC-NST.

It is important to recognise different morphological components at diagnosis, particularly when it is clear that IC-NST and ILC have different propensities for colonising different organ sites. MDLs are currently treated on the basis of their worst feature. For example, if a large proportion of the tumour were a high-grade IC-NST, the patient would likely receive chemotherapy. However, the smaller lobular component present may be less likely to respond to chemotherapy, and there is the possibility that the lobular component can progress further. This has implications when it comes to patient management, and patients who present with an MDL may need to undergo a unique surveillance program. Understanding how a cell develops a discohesive and invasive phenotype may be translated into gaining a greater understanding of the same features in ILC. This may open new avenues for targeting unique properties of these single discohesive cells and ultimately increase therapeutic options for MDL and ILC patients.

6.3. Metastasis to gynaecological sites; impacting young women with tumours of a luminal phenotype

Metastatic progression is the single most significant predictor of poor outcome for patients. Tumour cells must evolve in order to develop the necessary biological capabilities to metastasise, and that this clonal evolution is driven by the selective acquisition of somatic mutations, dynamic interactions with the local microenvironment, and resistance to treatment. The extent and overall clinical significance of this diversity in metastatic progression is still unclear, owing to the scarcity of samples of metastases that are available for molecular analysis.

Previous studies, including from our lab, have found that women who develop metastasis to gynaecological sites are frequently young in age, and a high proportion of patients had an invasive lobular carcinoma (Cummings et al., 2014, Bigorie et al., 2010). The fact that gynaecological sites are frequently operable (in comparison to common metastatic sites such as the bone, lung and brain), we have the unique opportunity to characterise and investigate mechanisms underlying metastatic progression in patient samples. We amassed a cohort of 54 patients with gynaecological metastasis, and found that these patients were significantly younger than the average breast cancer patient (47 years *vs.* 60 years). This suggests that the majority of patients were premenopausal at the time of breast cancer diagnosis. We hypothesised therefore that endocrine signalling might play an important role in progression to gynaecological sites. Approximately 43% of our patient cohort had an ILC, which is a significant enrichment compared to the 10-15% of ILC reported in the general breast cancer population. Other interesting clinical features of these patients were i) the high frequency of lymph node involvement at primary tumour diagnosis, ii) involvement of gastrointestinal sites and peritoneum/omentum more frequently than common sites of breast cancer metastasis, and iii) the expression of biomarkers of associated with a good prognosis (ER and PR positive, low Ki67 and negative for HER2 and basal markers). The last point is particularly interesting considering young patients diagnosed with breast cancer are generally associated with triple negative and basal like breast cancer (Liedtke et al., 2015).

This study identified both phenotypic and genomic intratumour heterogeneity during breast cancer progression to metastasis. This is important as treatment strategies for metastatic patients are frequently based on the characteristics of the primary tumour. It is now

understood that the development of metastatic disease requires tumour cells to develop under selection pressures such as treatment and a new niche microenvironment. During progression to gynaecological sites, the expression of ER remained stable. However, the expression of other important hormone markers PR, AR, GATA3 and FOXA1 were down-regulated. Consistent with the idea that the metastases represent an aggressive biology, p53 and Ki67 were frequently up-regulated in the metastases compared to the primary tumour. Although the expression of ER in these tumours did not change during progression as compared to other metastatic sites (such as the lung and liver (Cummings et al., 2014)) there are still important implications when it comes to guiding treatment options. Where the down-regulation of ER and PR is observed in the metastatic site, it is an obvious decision to cease endocrine therapy. However, these samples highlight that even though ER is still expressed, they still have likely developed resistance to the therapy due to the fact that the primary tumour was ER positive and the patient would have received endocrine therapy. It is possible therefore that the patient is unlikely to benefit from additional anti-endocrine agents. The discovery of the role of pioneer factors in ER signalling and the identification of *ESR1* mutations conferring endocrine resistance explains why ER is still being expressed but the tumour is not responsive to therapies. It is unclear how useful the assessment of FOXA1 would be alongside ER in metastatic samples, since we found that both were generally expressed in the metastases. The work by (Hurtado et al., 2011) and (Ross-Innes et al., 2012) found that FOXA1 mediates an altered ER transcriptional program under endocrine resistant *in vitro* settings. Therefore the identification of the genes that are differentially expressed may turn out to be more appropriate biomarkers of endocrine resistance.

To further understand the biology of these tumours, several genomic analyses were performed. Array-based copy number analysis was performed on this cohort and many shared copy number alterations were identified between the primary tumour and metastases sampled. Within 9 cases with matched primary tumour and metastases there was evidence of clonal similarities and clonal diversity during progression. It was also observed that each case presented with unique genomic alterations and these need to be validated using *in situ* hybridisation. Preliminary mutation analysis was unfortunately unsuccessful, however further testing is currently underway.

This study has provided a greater understanding of the phenotypic and molecular characteristics of metastatic progression to gynaecological organs. This resource of

patient samples will be utilised in the future to give further insight into mechanisms of resistance in ER positive breast cancer. With the advent of improved molecular technologies that are capable of utilising fragmented nucleic acids from FFPE material, potential biomarkers for treatment resistance may be developed and tested using this cohort. The identification of the mechanisms underlying progression and treatment resistance in each individual case will ultimately benefit the patient with more appropriately targeted treatment options. This study supports the notion of sampling metastatic deposits to help guide treatment options (Van Poznak et al., 2015), however, new biomarkers of treatment resistance need to be identified, particularly since the expression of ER can still occur in tumours that are resistant to endocrine therapy.

6.4. Final conclusions and future perspectives

The findings of this study have shown that the regulation of E-cadherin based cell-cell adhesion is complex in ILC and MDLs. To date, no study has published a detailed investigation of the cell adhesion complex in MDLs or investigation of the clonal evolution of MDLs using whole exome sequencing. We found that in some cases morphological divergence from a common neoplastic clone occurs early in tumour progression, while in other cases divergence occurs later. I believe that this raises a few possibilities about the inherent biology of the MDL. Firstly, this raises the question of whether there are actually two ‘types’ of MDL tumours. Are the tumours with early divergence a genuine MDL, in which the lobular component arises through more classical means via early E-cadherin inactivation and progression from LCIS? In contrast, are the cases with late divergence more simply an IC-NST with lobular-like differentiation? This raises questions surrounding our definition of a lobular carcinoma, notwithstanding the well-recognised morphological features. Is it the timing, or the mechanism, of *CDH1*/E-cadherin loss that defines the lobular phenotype? For instance, does an early hit to *CDH1* form a determined commitment to a lobular lineage, or do the cases with an apparent late insult to the integrity of the E-cadherin complex represent the evolution of tumour cells not fully committed to follow the lobular lineage? I believe that these are important questions that will have an impact on pathology classification systems. In my opinion, this study emphasises the necessity of identifying and further characterising the morphological and phenotypic heterogeneity in MDLs as it can have an impact on patient management.

It is generally considered that patients diagnosed with a luminal A tumour type will have a good clinical outcome in comparison to patients with triple negative or basal-like tumours. The latter subtype is more likely to be diagnosed in young breast cancer patients and have significantly worse 5-year survival rates. This study highlights a unique group of patients who present with primary disease at a young age and with a primary tumour of “good prognosis” – ER positive, low Ki67 and negative for HER2 and basal markers – yet they have developed widespread distant metastases. In spite of improving breast cancer patient mortality, I believe that further work is needed to improve our understanding of luminal breast cancer and the associated mechanisms of treatment resistance and disease spread. ER positive disease is the most common type of breast cancer, and thousands of patients still die each year. It is paramount to be able to predict with accuracy which patients are likely to relapse. This will significantly reduce over treatment of patients who are less likely to relapse, and enable earlier detection and more appropriate management for those who need it most.

References

ABDEL-FATAH, T. M., POWE, D. G., HODI, Z., LEE, A. H., REIS-FILHO, J. S. & ELLIS, I. O. 2007. High frequency of coexistence of columnar cell lesions, lobular neoplasia, and low grade ductal carcinoma in situ with invasive tubular carcinoma and invasive lobular carcinoma. *Am J Surg Pathol,* 31**,** 417-26.

ABDEL-FATAH, T. M., POWE, D. G., HODI, Z., REIS-FILHO, J. S., LEE, A. H. & ELLIS, I. O. 2008. Morphologic and molecular evolutionary pathways of low nuclear grade invasive breast cancers and their putative precursor lesions: further evidence to support the concept of low nuclear grade breast neoplasia family. *Am J Surg Pathol,* 32**,** 513-23.

ADZHUBEI, I. A., SCHMIDT, S., PESHKIN, L., RAMENSKY, V. E., GERASIMOVA, A. & BORK, P., KONDRASHOV, A. S., SUNYAEV, S. R. 2010. A method and server for predicting damaging missense mutations. *Nat Methods*, 7, 248-9.

AGOFF, S. N., SWANSON, P. E., LINDEN, H., HAWES, S. E. & LAWTON, T. J. 2003. Androgen receptor expression in estrogen receptor-negative breast cancer. Immunohistochemical, clinical, and prognostic associations. *Am J Clin Pathol,* 120**,** 725-31.

AIHW 2012. Cancer in Australia: an overview, 2012. *Cancer series no. 71. Cat. no. CAN 67. Canberra*.

AJCC 2007. American Joint Committee on Cancer, Breast Cancer Staging.

AL-EJEH, F., SIMPSON, P. T., SANUS, J. M., KLEIN, K., KALIMUTHO, M., SHI, W., MIRANDA, M., KUTASOVIC, J., RAGHAVENDRA, A., MADORE, J., REID, L., KRAUSE, L., CHENEVIX-TRENCH, G., LAKHANI, S. R. & KHANNA, K. K. 2014. Meta-analysis of the global gene expression profile of triple-negative breast cancer identifies genes for the prognostication and treatment of aggressive breast cancer. *Oncogenesis,* 3**,** e100.

ALBERTS, B., JOHNSON, A., LEWIS, J., RAFF, M., ROBERTS, K. & WALTER, P. 2008. *Molecular Biology of the Cell,* New York, USA, Garland Science.

ALERS, J. C., ROCHAT, J., KRIJTENBURG, P. J., VAN DEKKEN, H., RAAP, A. K. & ROSENBERG, C. 1999. Universal linkage system: an improved method for labeling archival DNA for comparative genomic hybridization. *Genes Chromosomes Cancer,* 25**,** 301-5.

ALESKANDARANY, M. A., NEGM, O. H., GREEN, A. R., AHMED, M. A., NOLAN, C. C., TIGHE, P. J., ELLIS, I. O. & RAKHA, E. A. 2014. Epithelial mesenchymal transition in early invasive breast cancer: an immunohistochemical and reverse phase protein array study. *Breast Cancer Res Treat,* 145**,** 339-48.

ALLEGRA, J. C., LIPPMAN, M. E., THOMPSON, E. B., SIMON, R., BARLOCK, A., GREEN, L., HUFF, K. K., DO, H. M. & AITKEN, S. C. 1979. Distribution, frequency, and quantitative analysis of estrogen, progesterone, androgen, and glucocorticoid receptors in human breast cancer. *Cancer Res,* 39**,** 1447-54.

ALLRED, D. C. 2010. Issues and updates: evaluating estrogen receptor-alpha, progesterone receptor, and HER2 in breast cancer. *Mod Pathol,* 23 Suppl 2**,** S52-9.

ALMENDRO, V., KIM, H. J., CHENG, Y. K., GONEN, M., ITZKOVITZ, S., ARGANI, P., VAN OUDENAARDEN, A., SUKUMAR, S., MICHOR, F. & POLYAK, K. 2014. Genetic and Phenotypic Diversity in Breast Tumor Metastases. *Cancer Res*.

ANDRADE, D. & ROSENBLATT, J. 2011. Apoptotic regulation of epithelial cellular extrusion. *Apoptosis,* 16**,** 491-501.

ANDRE, F., BACHELOT, T., COMMO, F., CAMPONE, M., ARNEDOS, M., DIERAS, V., LACROIX-TRIKI, M., LACROIX, L., COHEN, P., GENTIEN, D., ADELAIDE, J., DALENC, F., GONCALVES, A., LEVY, C., FERRERO, J. M., BONNETERRE, J.,

LEFEUVRE, C., JIMENEZ, M., FILLERON, T. & BONNEFOI, H. 2014. Comparative genomic hybridisation array and DNA sequencing to direct treatment of metastatic breast cancer: a multicentre, prospective trial (SAFIR01/UNICANCER). *Lancet Oncol*.

ANGRES, B., BARTH, A. & NELSON, W. J. 1996. Mechanism for transition from initial to stable cell-cell adhesion: kinetic analysis of E-cadherin-mediated adhesion using a quantitative adhesion assay. *J Cell Biol,* 134**,** 549-57.

ARPINO, G., BARDOU, V. J., CLARK, G. M. & ELLEDGE, R. M. 2004. Infiltrating lobular carcinoma of the breast: tumor characteristics and clinical outcome. *Breast cancer research : BCR,* 6**,** R149-56.

ARPS, D. P., HEALY, P., ZHAO, L., KLEER, C. G. & PANG, J. C. 2013. Invasive ductal carcinoma with lobular features: a comparison study to invasive ductal and invasive lobular carcinomas of the breast. *Breast Cancer Res Treat,* 138**,** 719-26.

ASSELIN-LABAT, M. L., SUTHERLAND, K. D., BARKER, H., THOMAS, R., SHACKLETON, M., FORREST, N. C., HARTLEY, L., ROBB, L., GROSVELD, F. G., VAN DER WEES, J., LINDEMAN, G. J. & VISVADER, J. E. 2007. Gata-3 is an essential regulator of mammary-gland morphogenesis and luminal-cell differentiation. *Nat Cell Biol,* 9**,** 201-9.

AYHAN, A., GUVENAL, T., SALMAN, M. C., OZYUNCU, O., SAKINCI, M. & BASARAN, M. 2005. The role of cytoreductive surgery in nongenital cancers metastatic to the ovaries. *Gynecologic oncology,* 98**,** 235-41.

BADVE, S., DABBS, D. J., SCHNITT, S. J., BAEHNER, F. L., DECKER, T., EUSEBI, V., FOX, S. B., ICHIHARA, S., JACQUEMIER, J., LAKHANI, S. R., PALACIOS, J., RAKHA, E. A., RICHARDSON, A. L., SCHMITT, F. C., TAN, P. H., TSE, G. M., WEIGELT, B., ELLIS, I. O. & REIS-FILHO, J. S. 2011. Basal-like and triple-negative breast cancers: a critical review with an emphasis on the implications for pathologists and oncologists. *Mod Pathol,* 24**,** 157-67.

BADVE, S., TURBIN, D., THORAT, M. A., MORIMIYA, A., NIELSEN, T. O., PEROU, C. M., DUNN, S., HUNTSMAN, D. G. & NAKSHATRI, H. 2007. FOXA1 expression in breast cancer--correlation with luminal subtype A and survival. *Clin Cancer Res,* 13**,** 4415-21.

BAKER, B. M. & CHEN, C. S. 2012. Deconstructing the third dimension: how 3D culture microenvironments alter cellular cues. *J Cell Sci,* 125**,** 3015-24.

BALDUS, S. E., SCHAEFER, K. L., ENGERS, R., HARTLEB, D., STOECKLEIN, N. H. & GABBERT, H. E. 2010. Prevalence and heterogeneity of KRAS, BRAF, and PIK3CA mutations in primary colorectal adenocarcinomas and their corresponding metastases. *Clin Cancer Res,* 16**,** 790-9.

BANERJI, S., CIBULSKIS, K., RANGEL-ESCARENO, C., BROWN, K. K., CARTER, S. L., FREDERICK, A. M., LAWRENCE, M. S., SIVACHENKO, A. Y., SOUGNEZ, C., ZOU, L., CORTES, M. L., FERNANDEZ-LOPEZ, J. C., PENG, S., ARDLIE, K. G., AUCLAIR, D., BAUTISTA-PINA, V., DUKE, F., FRANCIS, J., JUNG, J., MAFFUZ-AZIZ, A., ONOFRIO, R. C., PARKIN, M., PHO, N. H., QUINTANAR-JURADO, V., RAMOS, A. H., REBOLLAR-VEGA, R., RODRIGUEZ-CUEVAS, S., ROMERO-CORDOBA, S. L., SCHUMACHER, S. E., STRANSKY, N., THOMPSON, K. M., URIBE-FIGUEROA, L., BASELGA, J., BEROUKHIM, R., POLYAK, K., SGROI, D. C., RICHARDSON, A. L., JIMENEZ-SANCHEZ, G., LANDER, E. S., GABRIEL, S. B., GARRAWAY, L. A., GOLUB, T. R., MELENDEZ-ZAJGLA, J., TOKER, A., GETZ, G., HIDALGO-MIRANDA, A. & MEYERSON, M. 2012. Sequence analysis of mutations and translocations across breast cancer subtypes. *Nature,* 486**,** 405-9.

BARSKY, S. H. & KARLIN, N. J. 2005. Myoepithelial cells: autocrine and paracrine suppressors of breast cancer progression. *J Mammary Gland Biol Neoplasia,* 10**,** 249-60.

BASCHONG, W., SUETTERLIN, R. & LAENG, R. H. 2001. Control of autofluorescence of archival formaldehyde-fixed, paraffin-embedded tissue in confocal laser scanning microscopy (CLSM). *J Histochem Cytochem,* 49**,** 1565-72.
BAUM, B. & PERRIMON, N. 2001. Spatial control of the actin cytoskeleton in Drosophila epithelial cells. *Nat Cell Biol,* 3**,** 883-90.
BEDARD, P. L., CARDOSO, F. & PICCART-GEBHART, M. J. 2009. Stemming resistance to HER-2 targeted therapy. *J Mammary Gland Biol Neoplasia,* 14**,** 55-66.
BEHRNDT, M. & HEISENBERG, C. P. 2014. Lateral junction dynamics lead the way out. *Nat Cell Biol,* 16**,** 127-9.
BERX, G., BECKER, K. F., HOFLER, H. & VAN ROY, F. 1998. Mutations of the human E-cadherin (CDH1) gene. *Human mutation,* 12**,** 226-37.
BERX, G., CLETON-JANSEN, A. M., NOLLET, F., DE LEEUW, W. J., VAN DE VIJVER, M., CORNELISSE, C. & VAN ROY, F. 1995. E-cadherin is a tumour/invasion suppressor gene mutated in human lobular breast cancers. *The EMBO journal,* 14**,** 6107-15.
BERX, G., CLETON-JANSEN, A. M., STRUMANE, K., DE LEEUW, W. J., NOLLET, F., VAN ROY, F. & CORNELISSE, C. 1996. E-cadherin is inactivated in a majority of invasive human lobular breast cancers by truncation mutations throughout its extracellular domain. *Oncogene,* 13**,** 1919-25.
BERX, G. & VAN ROY, F. 2001. The E-cadherin/catenin complex: an important gatekeeper in breast cancer tumorigenesis and malignant progression. *Breast cancer research : BCR,* 3**,** 289-93.
BESSETTE, D. C., TILCH, E., SEIDENS, T., QUINN, M. C., WIEGMANS, A. P., SHI, W., COCCIARDI, S., MCCART-REED, A., SAUNUS, J. M., SIMPSON, P. T., GRIMMOND, S. M., LAKHANI, S. R., KHANNA, K. K., WADDELL, N., AL-EJEH, F. & CHENEVIX-TRENCH, G. 2015. Using the MCF10A/MCF10CA1a Breast Cancer Progression Cell Line Model to Investigate the Effect of Active, Mutant Forms of EGFR in Breast Cancer Development and Treatment Using Gefitinib. *PLoS One,* 10**,** e0125232.
BEUTE, B. J., KALISHER, L. & HUTTER, R. V. 1991. Lobular carcinoma in situ of the breast: clinical, pathologic, and mammographic features. *AJR. American journal of roentgenology,* 157**,** 257-65.
BHARAT, A., GAO, F. & MARGENTHALER, J. A. 2009. Tumor characteristics and patient outcomes are similar between invasive lobular and mixed invasive ductal/lobular breast cancers but differ from pure invasive ductal breast cancers. *American journal of surgery,* 198**,** 516-9.
BHARGAVA, R., STRIEBEL, J., BERIWAL, S., FLICKINGER, J. C., ONISKO, A., AHRENDT, G. & DABBS, D. J. 2009. Prevalence, morphologic features and proliferation indices of breast carcinoma molecular classes using immunohistochemical surrogate markers. *Int J Clin Exp Pathol,* 2**,** 444-55.
BHAT-NAKSHATRI, P., WANG, G., APPAIAH, H., LUKTUKE, N., CARROLL, J. S., GEISTLINGER, T. R., BROWN, M., BADVE, S., LIU, Y. & NAKSHATRI, H. 2008. AKT alters genome-wide estrogen receptor alpha binding and impacts estrogen signaling in breast cancer. *Mol Cell Biol,* 28**,** 7487-503.
BIGORIE, V., MORICE, P., DUVILLARD, P., ANTOINE, M., CORTEZ, A., FLEJOU, J. F., UZAN, S., DARAI, E. & BARRANGER, E. 2010. Ovarian metastases from breast cancer: report of 29 cases. *Cancer,* 116**,** 799-804.
BLOWS, F. M., DRIVER, K. E., SCHMIDT, M. K., BROEKS, A., VAN LEEUWEN, F. E., WESSELING, J., CHEANG, M. C., GELMON, K., NIELSEN, T. O., BLOMQVIST, C., HEIKKILA, P., HEIKKINEN, T., NEVANLINNA, H., AKSLEN, L. A., BEGIN, L. R., FOULKES, W. D., COUCH, F. J., WANG, X., CAFOUREK, V., OLSON, J. E., BAGLIETTO, L., GILES, G. G., SEVERI, G., MCLEAN, C. A., SOUTHEY, M. C.,

RAKHA, E., GREEN, A. R., ELLIS, I. O., SHERMAN, M. E., LISSOWSKA, J., ANDERSON, W. F., COX, A., CROSS, S. S., REED, M. W., PROVENZANO, E., DAWSON, S. J., DUNNING, A. M., HUMPHREYS, M., EASTON, D. F., GARCIA-CLOSAS, M., CALDAS, C., PHAROAH, P. D. & HUNTSMAN, D. 2010. Subtyping of breast cancer by immunohistochemistry to investigate a relationship between subtype and short and long term survival: a collaborative analysis of data for 10,159 cases from 12 studies. *PLoS Med,* 7**,** e1000279.

BORST, M. J. & INGOLD, J. A. 1993. Metastatic patterns of invasive lobular versus invasive ductal carcinoma of the breast. *Surgery,* 114**,** 637-41; discussion 641-2.

BOSCO, E. E., WANG, Y., XU, H., ZILFOU, J. T., KNUDSEN, K. E., ARONOW, B. J., LOWE, S. W. & KNUDSEN, E. S. 2007. The retinoblastoma tumor suppressor modifies the therapeutic response of breast cancer. *J Clin Invest,* 117**,** 218-28.

BOWDEN, E. T., BARTH, M., THOMAS, D., GLAZER, R. I. & MUELLER, S. C. 1999. An invasion-related complex of cortactin, paxillin and PKCmu associates with invadopodia at sites of extracellular matrix degradation. *Oncogene,* 18**,** 4440-9.

BRAGA, V. M., MACHESKY, L. M., HALL, A. & HOTCHIN, N. A. 1997. The small GTPases Rho and Rac are required for the establishment of cadherin-dependent cell-cell contacts. *J Cell Biol,* 137**,** 1421-31.

BRAUN, S., PANTEL, K., MULLER, P., JANNI, W., HEPP, F., KENTENICH, C. R., GASTROPH, S., WISCHNIK, A., DIMPFL, T., KINDERMANN, G., RIETHMULLER, G. & SCHLIMOK, G. 2000. Cytokeratin-positive cells in the bone marrow and survival of patients with stage I, II, or III breast cancer. *N Engl J Med,* 342**,** 525-33.

BRITTON, D. J., HUTCHESON, I. R., KNOWLDEN, J. M., BARROW, D., GILES, M., MCCLELLAND, R. A., GEE, J. M. & NICHOLSON, R. I. 2006. Bidirectional cross talk between ERalpha and EGFR signalling pathways regulates tamoxifen-resistant growth. *Breast Cancer Res Treat,* 96**,** 131-46.

BROSH, R. & ROTTER, V. 2009. When mutants gain new powers: news from the mutant p53 field. *Nat Rev Cancer,* 9**,** 701-13.

BRULS, J., SIMONS, M., OVERBEEK, L. I., BULTEN, J., MASSUGER, L. F. & NAGTEGAAL, I. D. 2015. A national population-based study provides insight in the origin of malignancies metastatic to the ovary. *Virchows Arch,* 467**,** 79-86.

BRUNNER, A. L., LI, J., GUO, X., SWEENEY, R. T., VARMA, S., ZHU, S. X., LI, R., TIBSHIRANI, R. & WEST, R. B. 2014. A shared transcriptional program in early breast neoplasias despite genetic and clinical distinctions. *Genome Biol,* 15**,** R71.

BUCHANAN, C. L., FLYNN, L. W., MURRAY, M. P., DARVISHIAN, F., CRANOR, M. L., FEY, J. V., KING, T. A., TAN, L. K. & SCLAFANI, L. M. 2008. Is pleomorphic lobular carcinoma really a distinct clinical entity? *J Surg Oncol,* 98**,** 314-7.

BUERGER, H., OTTERBACH, F., SIMON, R., SCHAFER, K. L., POREMBA, C., DIALLO, R., BRINKSCHMIDT, C., DOCKHORN-DWORNICZAK, B. & BOECKER, W. 1999. Different genetic pathways in the evolution of invasive breast cancer are associated with distinct morphological subtypes. *The Journal of pathology,* 189**,** 521-6.

BUERGER, H., SIMON, R., SCHAFER, K. L., DIALLO, R., LITTMANN, R., POREMBA, C., VAN DIEST, P. J., DOCKHORN-DWORNICZAK, B. & BOCKER, W. 2000. Genetic relation of lobular carcinoma in situ, ductal carcinoma in situ, and associated invasive carcinoma of the breast. *Mol Pathol,* 53**,** 118-21.

BUTT, A. J., MCNEIL, C. M., MUSGROVE, E. A. & SUTHERLAND, R. L. 2005. Downstream targets of growth factor and oestrogen signalling and endocrine resistance: the potential roles of c-Myc, cyclin D1 and cyclin E. *Endocr Relat Cancer,* 12 Suppl 1**,** S47-59.

CAMPBELL, P. J., YACHIDA, S., MUDIE, L. J., STEPHENS, P. J., PLEASANCE, E. D., STEBBINGS, L. A., MORSBERGER, L. A., LATIMER, C., MCLAREN, S., LIN, M. L., MCBRIDE, D. J., VARELA, I., NIK-ZAINAL, S. A., LEROY, C., JIA, M.,

MENZIES, A., BUTLER, A. P., TEAGUE, J. W., GRIFFIN, C. A., BURTON, J., SWERDLOW, H., QUAIL, M. A., STRATTON, M. R., IACOBUZIO-DONAHUE, C. & FUTREAL, P. A. 2010. The patterns and dynamics of genomic instability in metastatic pancreatic cancer. *Nature,* 467**,** 1109-13.

CANCER AUSTRALIA 2012. *Report to the nation - breast cancer 2012,* Surry Hills, NSW,.

CARROLL, J. S., LIU, X. S., BRODSKY, A. S., LI, W., MEYER, C. A., SZARY, A. J., EECKHOUTE, J., SHAO, W., HESTERMANN, E. V., GEISTLINGER, T. R., FOX, E. A., SILVER, P. A. & BROWN, M. 2005. Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. *Cell,* 122**,** 33-43.

CASTELLANO, I., ALLIA, E., ACCORTANZO, V., VANDONE, A. M., CHIUSA, L., ARISIO, R., DURANDO, A., DONADIO, M., BUSSOLATI, G., COATES, A. S., VIALE, G. & SAPINO, A. 2010. Androgen receptor expression is a significant prognostic factor in estrogen receptor positive breast cancers. *Breast Cancer Res Treat,* 124**,** 607-17.

CERAMI, E., GAO, J., DOGRUSOZ, U., GROSS, B. E., SUMER, S. O., AKSOY, B. A., JACOBSEN, A., BYRNE, C. J., HEUER, M. L., LARSSON, E., ANTIPIN, Y., REVA, B., GOLDBERG, A. P., SANDER, C. & SCHULTZ, N. 2012. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov,* 2**,** 401-4.

CHAFFER, C. L. & WEINBERG, R. A. 2011. A perspective on cancer cell metastasis. *Science,* 331**,** 1559-64.

CHANG, J. & HILSENBECK, S. 2010. *Prognostic and predictive markers.,* Philadelphia, Wolters Kluwer Lippincott Williams & Wilkins.

CHAU, B. N. & WANG, J. Y. 2003. Coordinated regulation of life and death by RB. *Nat Rev Cancer,* 3**,** 130-8.

CHEANG, M. C., CHIA, S. K., VODUC, D., GAO, D., LEUNG, S., SNIDER, J., WATSON, M., DAVIES, S., BERNARD, P. S., PARKER, J. S., PEROU, C. M., ELLIS, M. J. & NIELSEN, T. O. 2009. Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. *J Natl Cancer Inst,* 101**,** 736-50.

CHEANG, M. C., VODUC, D., BAJDIK, C., LEUNG, S., MCKINNEY, S., CHIA, S. K., PEROU, C. M. & NIELSEN, T. O. 2008. Basal-like breast cancer defined by five biomarkers has superior prognostic value than triple-negative phenotype. *Clin Cancer Res,* 14**,** 1368-76.

CHEN, Y., THOMPSON, W., SEMENCIW, R. & MAO, Y. 1999. Epidemiology of contralateral breast cancer. *Cancer Epidemiol Biomarkers Prev,* 8**,** 855-61.

CHENG YUNG-SHENG, LIN CHIN, CHENG YEN-PO, YU YI-LIN, TANG CHI-TUN & DUENG-YUAN, H. 2014. Epithelial cell transformation sequence 2 is a potential biomarker of unfavorable survival in human gliomas. *Neurology India,* 62**,** 406-409.

CHIANG, A. C. & MASSAGUE, J. 2008. Molecular basis of metastasis. *The New England journal of medicine,* 359**,** 2814-23.

CHOI, Y., SIMS, G. E., MURPHY, S., MILLER, J. R. & CHAN, A. P. 2012. Predicting the functional effect of amino acid substitutions and indels. *PLoS One*, 7, e46688.

CHRISTGEN, M., BRUCHHARDT, H., HADAMITZKY, C., RUDOLPH, C., STEINEMANN, D., GADZICKI, D., HASEMEIER, B., ROMERMANN, D., FOCKEN, T., KRECH, T., BALLMAIER, M., SCHLEGELBERGER, B., KREIPE, H. & LEHMANN, U. 2009. Comprehensive genetic and functional characterization of IPH-926: a novel CDH1-null tumour cell line from human lobular breast cancer. *The Journal of pathology,* 217**,** 620-32.

CIOCCA, V., DASKALAKIS, C., CIOCCA, R. M., RUIZ-ORRICO, A. & PALAZZO, J. P. 2009. The significance of GATA3 expression in breast cancer: a 10-year follow-up study. *Hum Pathol,* 40**,** 489-95.

CIRILLO, L. A., LIN, F. R., CUESTA, I., FRIEDMAN, D., JARNIK, M. & ZARET, K. S. 2002. Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Mol Cell,* 9**,** 279-89.

CIRILLO, L. A. & ZARET, K. S. 1999. An early developmental transcription factor complex that is more stable on nucleosome core particles than on free DNA. *Mol Cell,* 4**,** 961-9.

CLARKE, R. B. 2003. Steroid receptors and proliferation in the human breast. *Steroids,* 68**,** 789-94.

CLETON-JANSEN, A. M. 2002. E-cadherin and loss of heterozygosity at chromosome 16 in breast carcinogenesis: different genetic pathways in ductal and lobular breast cancer? *Breast Cancer Res,* 4**,** 5-8.

COHEN, H., BEN-HAMO, R., GIDONI, M., YITZHAKI, I., KOZOL, R., ZILBERBERG, A. & EFRONI, S. 2014. Shift in GATA3 functions, and GATA3 mutations, control progression and clinical presentation in breast cancer. *Breast Cancer Res,* 16**,** 464.

COLE, M. P., JONES, C. T. & TODD, I. D. 1971. A new anti-oestrogenic agent in late breast cancer. An early clinical appraisal of ICI46474. *Br J Cancer,* 25**,** 270-5.

COLLINS, L. C., TAMIMI, R. M., BAER, H. J., CONNOLLY, J. L., COLDITZ, G. A. & SCHNITT, S. J. 2005. Outcome of patients with ductal carcinoma in situ untreated after diagnostic biopsy: results from the Nurses' Health Study. *Cancer,* 103**,** 1778-84.

COWPER-SAL LARI, R., ZHANG, X., WRIGHT, J. B., BAILEY, S. D., COLE, M. D., EECKHOUTE, J., MOORE, J. H. & LUPIEN, M. 2012. Breast cancer risk-associated SNPs modulate the affinity of chromatin for FOXA1 and alter gene expression. *Nat Genet,* 44**,** 1191-8.

CRAIG, J. M., VENA, N., RAMKISSOON, S., IDBAIH, A., FOUSE, S. D., OZEK, M., SAV, A., HILL, D. A., MARGRAF, L. R., EBERHART, C. G., KIERAN, M. W., NORDEN, A. D., WEN, P. Y., LODA, M., SANTAGATA, S., LIGON, K. L. & LIGON, A. H. 2012. DNA fragmentation simulation method (FSM) and fragment size matching improve aCGH performance of FFPE tissues. *PLoS One,* 7**,** e38881.

CUI, X., SCHIFF, R., ARPINO, G., OSBORNE, C. K. & LEE, A. V. 2005. Biology of progesterone receptor loss in breast cancer and its implications for endocrine therapy. *J Clin Oncol,* 23**,** 7721-35.

CUMMINGS, M. C., SIMPSON, P. T., REID, L. E., JAYANTHAN, J., SKERMAN, J., SONG, S., MCCART REED, A. E., KUTASOVIC, J. R., MOREY, A. L., MARQUART, L., O'ROURKE, P. & LAKHANI, S. R. 2014. Metastatic progression of breast cancer: insights from 50 years of autopsies. *J Pathol,* 232**,** 23-31.

CURTIS, C., SHAH, S. P., CHIN, S. F., TURASHVILI, G., RUEDA, O. M., DUNNING, M. J., SPEED, D., LYNCH, A. G., SAMARAJIWA, S., YUAN, Y., GRAF, S., HA, G., HAFFARI, G., BASHASHATI, A., RUSSELL, R., MCKINNEY, S., LANGEROD, A., GREEN, A., PROVENZANO, E., WISHART, G., PINDER, S., WATSON, P., MARKOWETZ, F., MURPHY, L., ELLIS, I., PURUSHOTHAM, A., BORRESEN-DALE, A. L., BRENTON, J. D., TAVARE, S., CALDAS, C. & APARICIO, S. 2012. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature,* 486**,** 346-52.

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. *J Clin Oncol,* 29**,** 4273-8.

DA SILVA, L., PARRY, S., REID, L., KEITH, P., WADDELL, N., KOSSAI, M., CLARKE, C., LAKHANI, S. R. & SIMPSON, P. T. 2008. Aberrant expression of E-cadherin in lobular carcinomas of the breast. *Am J Surg Pathol,* 32**,** 773-83.

DA SILVA, L., SIMPSON, P. T. & R., L. S. 2010. *Lobular carcinoma in situ*, Springer-Verlag Berlin Heidelberg.

DABBS, D. J., BHARGAVA, R. & CHIVUKULA, M. 2007. Lobular versus ductal breast neoplasms: the diagnostic utility of p120 catenin. *Am J Surg Pathol,* 31**,** 427-37.

DABBS, D. J., SCHNITT, S. J., GEYER, F. C., WEIGELT, B., BAEHNER, F. L., DECKER, T., EUSEBI, V., FOX, S. B., ICHIHARA, S., LAKHANI, S. R., PALACIOS, J., RAKHA, E., RICHARDSON, A. L., SCHMITT, F. C., TAN, P. H., TSE, G. M., VINCENT-SALOMON, A., ELLIS, I. O., BADVE, S. & REIS-FILHO, J. S. 2013. Lobular neoplasia of the breast revisited with emphasis on the role of E-cadherin immunohistochemistry. *Am J Surg Pathol,* 37**,** e1-11.

DAKUBO, G. D., JAKUPCIAK, J. P., BIRCH-MACHIN, M. A. & PARR, R. L. 2007. Clinical implications and utility of field cancerization. *Cancer Cell Int,* 7**,** 2.

DALING, J. R., MALONE, K. E., DOODY, D. R., VOIGT, L. F., BERNSTEIN, L., COATES, R. J., MARCHBANKS, P. A., NORMAN, S. A., WEISS, L. K., URSIN, G., BERLIN, J. A., BURKMAN, R. T., DEAPEN, D., FOLGER, S. G., MCDONALD, J. A., SIMON, M. S., STROM, B. L., WINGO, P. A. & SPIRTAS, R. 2002. Relation of regimens of combined hormone replacement therapy to lobular, ductal, and other histologic types of breast carcinoma. *Cancer,* 95**,** 2455-64.

DAVIES, C., PAN, H., GODWIN, J., GRAY, R., ARRIAGADA, R., RAINA, V., ABRAHAM, M., ALENCAR, V. H. M., BADRAN, A., BONFILL, X., BRADBURY, J., CLARKE, M., COLLINS, R., DAVIS, S. R., DELMESTRI, A., FORBES, J. F., HADDAD, P., HOU, M.-F., INBAR, M., KHALED, H., KIELANOWSKA, J., KWAN, W.-H., MATHEW, B. S., MITTRA, I., MÜLLER, B., NICOLUCCI, A., PERALTA, O., PERNAS, F., PETRUZELKA, L., PIENKOWSKI, T., RADHIKA, R., RAJAN, B., RUBACH, M. T., TORT, S., URRÚTIA, G., VALENTINI, M., WANG, Y. & PETO, R. Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years after diagnosis of oestrogen receptor-positive breast cancer: ATLAS, a randomised trial. *The Lancet,* 381**,** 805-816.

DE AMICIS, F., THIRUGNANSAMPANTHAN, J., CUI, Y., SELEVER, J., BEYER, A., PARRA, I., WEIGEL, N. L., HERYNK, M. H., TSIMELZON, A., LEWIS, M. T., CHAMNESS, G. C., HILSENBECK, S. G., ANDO, S. & FUQUA, S. A. 2010. Androgen receptor overexpression induces tamoxifen resistance in human breast cancer cells. *Breast Cancer Res Treat,* 121**,** 1-11.

DE LEEUW, W. J., BERX, G., VOS, C. B., PETERSE, J. L., VAN DE VIJVER, M. J., LITVINOV, S., VAN ROY, F., CORNELISSE, C. J. & CLETON-JANSEN, A. M. 1997. Simultaneous loss of E-cadherin and catenins in invasive lobular breast cancer and lobular carcinoma in situ. *J Pathol,* 183**,** 404-11.

DE WAAL, Y. R., THOMAS, C. M., OEI, A. L., SWEEP, F. C. & MASSUGER, L. F. 2009. Secondary ovarian malignancies: frequency, origin, and characteristics. *International journal of gynecological cancer : official journal of the International Gynecological Cancer Society,* 19**,** 1160-5.

DEGNIM, A. C., VISSCHER, D. W., BERMAN, H. K., FROST, M. H., SELLERS, T. A., VIERKANT, R. A., MALONEY, S. D., PANKRATZ, V. S., DE GROEN, P. C., LINGLE, W. L., GHOSH, K., PENHEITER, L., TLSTY, T., MELTON, L. J., 3RD, REYNOLDS, C. A. & HARTMANN, L. C. 2007. Stratification of breast cancer risk in women with atypia: a Mayo cohort study. *J Clin Oncol,* 25**,** 2671-7.

DEMOPOULOS, R. I., TOUGER, L. & DUBIN, N. 1987. Secondary ovarian carcinoma: a clinical and pathological evaluation. *International journal of gynecological pathology*

: official journal of the International Society of Gynecological Pathologists, 6**,** 166-75.
DERKSEN, P. W., BRAUMULLER, T. M., VAN DER BURG, E., HORNSVELD, M., MESMAN, E., WESSELING, J., KRIMPENFORT, P. & JONKERS, J. 2011. Mammary-specific inactivation of E-cadherin and p53 impairs functional gland development and leads to pleomorphic invasive lobular carcinoma in mice. *Disease models & mechanisms,* 4**,** 347-58.
DHIMOLEA, E., TINIAKOS, D. G., CHANTZI, N., GOUTAS, N., VASSILAROS, S. D., MITSIOU, D. J. & ALEXIS MU, N. 2015. Estrogen receptors beta1 and beta2 are associated with distinct responses of estrogen receptor alpha-positive breast carcinoma to adjuvant endocrine therapy. *Cancer Lett,* 358**,** 37-42.
DI LEO, A., JERUSALEM, G., PETRUZELKA, L., TORRES, R., BONDARENKO, I. N., KHASANOV, R., VERHOEVEN, D., PEDRINI, J. L., SMIRNOVA, I., LICHINITSER, M. R., PENDERGRASS, K., GARNETT, S., LINDEMANN, J. P., SAPUNAR, F. & MARTIN, M. 2010. Results of the CONFIRM phase III trial comparing fulvestrant 250 mg with fulvestrant 500 mg in postmenopausal women with estrogen receptor-positive advanced breast cancer. *J Clin Oncol,* 28**,** 4594-600.
DING, L., ELLIS, M. J., LI, S., LARSON, D. E., CHEN, K., WALLIS, J. W., HARRIS, C. C., MCLELLAN, M. D., FULTON, R. S., FULTON, L. L., ABBOTT, R. M., HOOG, J., DOOLING, D. J., KOBOLDT, D. C., SCHMIDT, H., KALICKI, J., ZHANG, Q., CHEN, L., LIN, L., WENDL, M. C., MCMICHAEL, J. F., MAGRINI, V. J., COOK, L., MCGRATH, S. D., VICKERY, T. L., APPELBAUM, E., DESCHRYVER, K., DAVIES, S., GUINTOLI, T., CROWDER, R., TAO, Y., SNIDER, J. E., SMITH, S. M., DUKES, A. F., SANDERSON, G. E., POHL, C. S., DELEHAUNTY, K. D., FRONICK, C. C., PAPE, K. A., REED, J. S., ROBINSON, J. S., HODGES, J. S., SCHIERDING, W., DEES, N. D., SHEN, D., LOCKE, D. P., WIECHERT, M. E., ELDRED, J. M., PECK, J. B., OBERKFELL, B. J., LOLOFIE, J. T., DU, F., HAWKINS, A. E., O'LAUGHLIN, M. D., BERNARD, K. E., CUNNINGHAM, M., ELLIOTT, G., MASON, M. D., THOMPSON, D. M., JR., IVANOVICH, J. L., GOODFELLOW, P. J., PEROU, C. M., WEINSTOCK, G. M., AFT, R., WATSON, M., LEY, T. J., WILSON, R. K. & MARDIS, E. R. 2010. Genome remodelling in a basal-like breast cancer metastasis and xenograft. *Nature,* 464**,** 999-1005.
DIXON, J. M., ANDERSON, T. J., PAGE, D. L., LEE, D. & DUFFY, S. W. 1982. Infiltrating lobular carcinoma of the breast. *Histopathology,* 6**,** 149-61.
DO, K. A., TRELOAR, S. A., PANDEYA, N., PURDIE, D., GREEN, A. C., HEATH, A. C. & MARTIN, N. G. 1998. Predictive factors of age at menopause in a large Australian twin study. *Hum Biol,* 70**,** 1073-91.
DODWELL, D., WARDLEY, A. & JOHNSTON, S. 2006. Postmenopausal advanced breast cancer: options for therapy after tamoxifen and aromatase inhibitors. *Breast,* 15**,** 584-94.
DROUFAKOU, S., DESHMANE, V., ROYLANCE, R., HANBY, A., TOMLINSON, I. & HART, I. R. 2001. Multiple ways of silencing E-cadherin gene expression in lobular carcinoma of the breast. *Int J Cancer,* 92**,** 404-8.
DURINCK, S., SPELLMAN, P. T., BIRNEY, E. & HUBER, W. 2009. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc,* 4**,** 1184-91.
EARLY BREAST CANCER TRIALISTS' COLLABORATIVE, G. 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**,** 1687-717.
EARLY BREAST CANCER TRIALISTS' COLLABORATIVE, G., DAVIES, C., GODWIN, J., GRAY, R., CLARKE, M., CUTTER, D., DARBY, S., MCGALE, P., PAN, H. C., TAYLOR, C., WANG, Y. C., DOWSETT, M., INGLE, J. & PETO, R. 2011.

Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. *Lancet,* 378**,** 771-84.
EECKHOUTE, J., KEETON, E. K., LUPIEN, M., KRUM, S. A., CARROLL, J. S. & BROWN, M. 2007. Positive cross-regulatory loop ties GATA-3 to estrogen receptor alpha expression in breast cancer. *Cancer Res,* 67**,** 6477-83.
EISENHOFFER, G. T., LOFTUS, P. D., YOSHIGI, M., OTSUNA, H., CHIEN, C. B., MORCOS, P. A. & ROSENBLATT, J. 2012. Crowding induces live cell extrusion to maintain homeostatic cell numbers in epithelia. *Nature,* 484**,** 546-9.
EL-SAGHIR, J. A., EL-HABRE, E. T., EL-SABBAN, M. E. & TALHOUK, R. S. 2011. Connexins: a junctional crossroad to breast cancer. *Int J Dev Biol,* 55**,** 773-80.
ELLIS, M. J., DING, L., SHEN, D., LUO, J., SUMAN, V. J., WALLIS, J. W., VAN TINE, B. A., HOOG, J., GOIFFON, R. J., GOLDSTEIN, T. C., NG, S., LIN, L., CROWDER, R., SNIDER, J., BALLMAN, K., WEBER, J., CHEN, K., KOBOLDT, D. C., KANDOTH, C., SCHIERDING, W. S., MCMICHAEL, J. F., MILLER, C. A., LU, C., HARRIS, C. C., MCLELLAN, M. D., WENDL, M. C., DESCHRYVER, K., ALLRED, D. C., ESSERMAN, L., UNZEITIG, G., MARGENTHALER, J., BABIERA, G. V., MARCOM, P. K., GUENTHER, J. M., LEITCH, M., HUNT, K., OLSON, J., TAO, Y., MAHER, C. A., FULTON, L. L., FULTON, R. S., HARRISON, M., OBERKFELL, B., DU, F., DEMETER, R., VICKERY, T. L., ELHAMMALI, A., PIWNICA-WORMS, H., MCDONALD, S., WATSON, M., DOOLING, D. J., OTA, D., CHANG, L. W., BOSE, R., LEY, T. J., PIWNICA-WORMS, D., STUART, J. M., WILSON, R. K. & MARDIS, E. R. 2012. Whole-genome analysis informs breast cancer response to aromatase inhibition. *Nature,* 486**,** 353-60.
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,* 19**,** 403-10.
ENGEL, J., ECKEL, R., KERR, J., SCHMIDT, M., FURSTENBERGER, G., RICHTER, R., SAUER, H., SENN, H. J. & HOLZEL, D. 2003. The process of metastasisation for breast cancer. *Eur J Cancer,* 39**,** 1794-806.
EPSTEIN, M., MA, Y. & PRESS, M. 2010. *ERBB2 testing: assessment of status for targeted therapies.,* Philadelphia, Wolters Kluwer Lippincott Williams & Wilkins.
ETZELL, J. E., DEVRIES, S., CHEW, K., FLORENDO, C., MOLINARO, A., LJUNG, B. M. & WALDMAN, F. M. 2001. Loss of chromosome 16q in lobular carcinoma in situ. *Human pathology,* 32**,** 292-6.
EUSEBI, V., MAGALHAES, F. & AZZOPARDI, J. G. 1992. Pleomorphic lobular carcinoma of the breast: an aggressive tumor showing apocrine differentiation. *Hum Pathol,* 23**,** 655-62.
EVANS, W. P., WARREN BURHENNE, L. J., LAURIE, L., O'SHAUGHNESSY, K. F. & CASTELLINO, R. A. 2002. Invasive lobular carcinoma of the breast: mammographic characteristics and computer-aided detection. *Radiology,* 225**,** 182-9.
EWING, J. 1928. *Neoplastic Diseases: A Treatise on Tumours.*, Philadelphia and London.
FABI, A., DI BENEDETTO, A., METRO, G., PERRACCHIO, L., NISTICO, C., DI FILIPPO, F., ERCOLANI, C., FERRETTI, G., MELUCCI, E., BUGLIONI, S., SPERDUTI, I., PAPALDO, P., COGNETTI, F. & MOTTOLESE, M. 2011. HER2 protein and gene variation between primary and metastatic breast cancer: significance and impact on patient care. *Clinical cancer research : an official journal of the American Association for Cancer Research,* 17**,** 2055-64.
FARMER, P., BONNEFOI, H., BECETTE, V., TUBIANA-HULIN, M., FUMOLEAU, P., LARSIMONT, D., MACGROGAN, G., BERGH, J., CAMERON, D., GOLDSTEIN, D., DUSS, S., NICOULAZ, A. L., BRISKEN, C., FICHE, M., DELORENZI, M. & IGGO,

R. 2005. Identification of molecular apocrine breast tumours by microarray analysis. *Oncogene,* 24**,** 4660-71.
FERLICOT, S., VINCENT-SALOMON, A., MEDIONI, J., GENIN, P., ROSTY, C., SIGAL-ZAFRANI, B., FRENEAUX, P., JOUVE, M., THIERY, J. P. & SASTRE-GARAU, X. 2004. Wide metastatic spreading in infiltrating lobular carcinoma of the breast. *European journal of cancer,* 40**,** 336-41.
FIDLER, I. J. 1970. Metastasis: guantitative analysis of distribution and fate of tumor embolilabeled with 125 I-5-iodo-2'-deoxyuridine. *Journal of the National Cancer Institute,* 45**,** 773-82.
FIDLER, I. J. 2003. The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited. *Nat Rev Cancer,* 3**,** 453-458.
FINN, R. S., CROWN, J. P., LANG, I., BOER, K., BONDARENKO, I. M., KULYK, S. O., ETTL, J., PATEL, R., PINTER, T., SCHMIDT, M., SHPARYK, Y., THUMMALA, A. R., VOYTKO, N. L., FOWST, C., HUANG, X., KIM, S. T., RANDOLPH, S. & SLAMON, D. J. 2015. The cyclin-dependent kinase 4/6 inhibitor palbociclib in combination with letrozole versus letrozole alone as first-line treatment of oestrogen receptor-positive, HER2-negative, advanced breast cancer (PALOMA-1/TRIO-18): a randomised phase 2 study. *Lancet Oncol,* 16**,** 25-35.
FITZGERALD, R. C., HARDWICK, R., HUNTSMAN, D., CARNEIRO, F., GUILFORD, P., BLAIR, V., CHUNG, D. C., NORTON, J., RAGUNATH, K., VAN KRIEKEN, J. H., DWERRYHOUSE, S., CALDAS, C. & INTERNATIONAL GASTRIC CANCER LINKAGE, C. 2010. Hereditary diffuse gastric cancer: updated consensus guidelines for clinical management and directions for future research. *J Med Genet,* 47**,** 436-44.
FOOTE, F. W. & STEWART, F. W. 1941. Lobular carcinoma in situ: A rare form of mammary cancer. *The American journal of pathology***,** 491-496.
FOUQUET, S., LUGO-MARTINEZ, V. H., FAUSSAT, A. M., RENAUD, F., CARDOT, P., CHAMBAZ, J., PINCON-RAYMOND, M. & THENET, S. 2004. Early loss of E-cadherin from cell-cell contacts is involved in the onset of Anoikis in enterocytes. *J Biol Chem,* 279**,** 43061-9.
FOX, E. M., DAVIS, R. J. & SHUPNIK, M. A. 2008. ERbeta in breast cancer--onlooker, passive player, or active protector? *Steroids,* 73**,** 1039-51.
FRAMPTON, G. M., FICHTENHOLTZ, A., OTTO, G. A., WANG, K., DOWNING, S. R., HE, J., SCHNALL-LEVIN, M., WHITE, J., SANFORD, E. M., AN, P., SUN, J., JUHN, F., BRENNAN, K., IWANIK, K., MAILLET, A., BUELL, J., WHITE, E., ZHAO, M., BALASUBRAMANIAN, S., TERZIC, S., RICHARDS, T., BANNING, V., GARCIA, L., MAHONEY, K., ZWIRKO, Z., DONAHUE, A., BELTRAN, H., MOSQUERA, J. M., RUBIN, M. A., DOGAN, S., HEDVAT, C. V., BERGER, M. F., PUSZTAI, L., LECHNER, M., BOSHOFF, C., JAROSZ, M., VIETZ, C., PARKER, A., MILLER, V. A., ROSS, J. S., CURRAN, J., CRONIN, M. T., STEPHENS, P. J., LIPSON, D. & YELENSKY, R. 2013. Development and validation of a clinical cancer genomic profiling test based on massively parallel DNA sequencing. *Nat Biotechnol,* 31**,** 1023-31.
FRANCIS, P. A., REGAN, M. M., FLEMING, G. F., LANG, I., CIRUELOS, E., BELLET, M., BONNEFOI, H. R., CLIMENT, M. A., DA PRADA, G. A., BURSTEIN, H. J., MARTINO, S., DAVIDSON, N. E., GEYER, C. E., JR., WALLEY, B. A., COLEMAN, R., KERBRAT, P., BUCHHOLZ, S., INGLE, J. N., WINER, E. P., RABAGLIO-PORETTI, M., MAIBACH, R., RUEPP, B., GIOBBIE-HURDER, A., PRICE, K. N., COLLEONI, M., VIALE, G., COATES, A. S., GOLDHIRSCH, A., GELBER, R. D., INVESTIGATORS, S. & INTERNATIONAL BREAST CANCER STUDY, G. 2015. Adjuvant ovarian suppression in premenopausal breast cancer. *N Engl J Med,* 372**,** 436-46.

FRIBERG, S. & MATTSON, S. 1997. On the growth rates of human malignant tumors: implications for medical decision making. *J Surg Oncol,* 65**,** 284-97.
FRITZ, G., BRACHETTI, C., BAHLMANN, F., SCHMIDT, M. & KAINA, B. 2002. Rho GTPases in human breast tumours: expression and mutation analyses and correlation with clinical parameters. *Br J Cancer,* 87**,** 635-44.
FRIXEN, U. H., BEHRENS, J., SACHS, M., EBERLE, G., VOSS, B., WARDA, A., LOCHNER, D. & BIRCHMEIER, W. 1991. E-cadherin-mediated cell-cell adhesion prevents invasiveness of human carcinoma cells. *J Cell Biol,* 113**,** 173-85.
FULFORD, L. G., EASTON, D. F., REIS-FILHO, J. S., SOFRONIS, A., GILLETT, C. E., LAKHANI, S. R. & HANBY, A. 2006. Specific morphological features predictive for the basal phenotype in grade 3 invasive ductal carcinoma of breast. *Histopathology,* 49**,** 22-34.
FULFORD, L. G., REIS-FILHO, J. S., RYDER, K., JONES, C., GILLETT, C. E., HANBY, A., EASTON, D. & LAKHANI, S. R. 2007. Basal-like grade III invasive ductal carcinoma of the breast: patterns of metastasis and long-term survival. *Breast Cancer Res,* 9**,** R4.
FUQUA, S. A., GU, G. & RECHOUM, Y. 2014. Estrogen receptor (ER) alpha mutations in breast cancer: hidden in plain sight. *Breast Cancer Res Treat,* 144**,** 11-9.
GAO, J., AKSOY, B. A., DOGRUSOZ, U., DRESDNER, G., GROSS, B., SUMER, S. O., SUN, Y., JACOBSEN, A., SINHA, R., LARSSON, E., CERAMI, E., SANDER, C. & SCHULTZ, N. 2013. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal,* 6**,** pl1.
GARAY, J. P. & PARK, B. H. 2012. Androgen receptor as a targeted therapy for breast cancer. *Am J Cancer Res,* 2**,** 434-45.
GERDES, J., SCHWAB, U., LEMKE, H. & STEIN, H. 1983. Production of a mouse monoclonal antibody reactive with a human nuclear antigen associated with cell proliferation. *Int J Cancer,* 31**,** 13-20.
GERLINGER, M., ROWAN, A. J., HORSWELL, S., LARKIN, J., ENDESFELDER, D., GRONROOS, E., MARTINEZ, P., MATTHEWS, N., STEWART, A., TARPEY, P., VARELA, I., PHILLIMORE, B., BEGUM, S., MCDONALD, N. Q., BUTLER, A., JONES, D., RAINE, K., LATIMER, C., SANTOS, C. R., NOHADANI, M., EKLUND, A. C., SPENCER-DENE, B., CLARK, G., PICKERING, L., STAMP, G., GORE, M., SZALLASI, Z., DOWNWARD, J., FUTREAL, P. A. & SWANTON, C. 2012. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *The New England journal of medicine,* 366**,** 883-92.
GHAJAR, C. M., PEINADO, H., MORI, H., MATEI, I. R., EVASON, K. J., BRAZIER, H., ALMEIDA, D., KOLLER, A., HAJJAR, K. A., STAINIER, D. Y., CHEN, E. I., LYDEN, D. & BISSELL, M. J. 2013. The perivascular niche regulates breast tumour dormancy. *Nat Cell Biol,* 15**,** 807-17.
GIRETTI, M. S., FU, X. D., DE ROSA, G., SAROTTO, I., BALDACCI, C., GARIBALDI, S., MANNELLA, P., BIGLIA, N., SISMONDI, P., GENAZZANI, A. R. & SIMONCINI, T. 2008. Extra-nuclear signalling of estrogen receptor to breast cancer cytoskeletal remodelling, migration and invasion. *PLoS One,* 3**,** e2238.
GLIGORIJEVIC, B., WYCKOFF, J., YAMAGUCHI, H., WANG, Y., ROUSSOS, E. T. & CONDEELIS, J. 2012. N-WASP-mediated invadopodium formation is involved in intravasation and lung metastasis of mammary tumors. *J Cell Sci,* 125**,** 724-34.
GOLDENBERG, I. S. 1964. Testosterone Propionate Therapy in Breast Cancer. *JAMA,* 188**,** 1069-72.
GRAYSON, M. 2012. Breast cancer. *Nature,* 485**,** S49.
GREENE, F. L. & SOBIN, L. H. 2008. The staging of cancer: a retrospective and prospective appraisal. *CA Cancer J Clin,* 58**,** 180-90.

GRIEVE, A. G. & RABOUILLE, C. 2014. Extracellular cleavage of E-cadherin promotes epithelial cell extrusion. *J Cell Sci,* 127**,** 3331-46.
GRIGGS, J. J., SOMERFIELD, M. R., ANDERSON, H., HENRY, N. L., HUDIS, C. A., KHATCHERESSIAN, J. L., PARTRIDGE, A. H., PRESTRUD, A. A. & DAVIDSON, N. E. 2011. American Society of Clinical Oncology endorsement of the cancer care Ontario practice guideline on adjuvant ovarian ablation in the treatment of premenopausal women with early-stage invasive breast cancer. *J Clin Oncol,* 29**,** 3939-42.
GUILFORD, P., HOPKINS, J., HARRAWAY, J., MCLEOD, M., MCLEOD, N., HARAWIRA, P., TAITE, H., SCOULAR, R., MILLER, A. & REEVE, A. E. 1998. E-cadherin germline mutations in familial gastric cancer. *Nature,* 392**,** 402-5.
GUNDEM, G., PEREZ-LLAMAS, C., JENE-SANZ, A., KEDZIERSKA, A., ISLAM, A., DEU-PONS, J., FURNEY, S. J. & LOPEZ-BIGAS, N. 2010. IntOGen: integration and data mining of multidimensional oncogenomic data. *Nat Methods,* 7**,** 92-3.
GUPTA, G. P. & MASSAGUE, J. 2006. Cancer metastasis: building a framework. *Cell,* 127**,** 679-95.
HACK, C. J. 2004. Integrated transcriptome and proteome data: the challenges ahead. *Brief Funct Genomic Proteomic,* 3**,** 212-9.
HADFIELD, G. 1954. The dormant cancer cell. *Br Med J,* 2**,** 607-10.
HAH, N., DANKO, C. G., CORE, L., WATERFALL, J. J., SIEPEL, A., LIS, J. T. & KRAUS, W. L. 2011. A rapid, extensive, and transient transcriptional response to estrogen signaling in breast cancer cells. *Cell,* 145**,** 622-34.
HALDOSEN, L. A., ZHAO, C. & DAHLMAN-WRIGHT, K. 2014. Estrogen receptor beta in breast cancer. *Mol Cell Endocrinol,* 382**,** 665-72.
HALL, A. 1998. Rho GTPases and the actin cytoskeleton. *Science,* 279**,** 509-14.
HALL, R. E., ASPINALL, J. O., HORSFALL, D. J., BIRRELL, S. N., BENTEL, J. M., SUTHERLAND, R. L. & TILLEY, W. D. 1996. Expression of the androgen receptor and an androgen-responsive protein, apolipoprotein D, in human breast cancer. *Br J Cancer,* 74**,** 1175-80.
HAMMOND, M. E., HAYES, D. F., DOWSETT, M., ALLRED, D. C., HAGERTY, K. L., BADVE, S., FITZGIBBONS, P. L., FRANCIS, G., GOLDSTEIN, N. S., HAYES, M., HICKS, D. G., LESTER, S., LOVE, R., MANGU, P. B., MCSHANE, L., MILLER, K., OSBORNE, C. K., PAIK, S., PERLMUTTER, J., RHODES, A., SASANO, H., SCHWARTZ, J. N., SWEEP, F. C., TAUBE, S., TORLAKOVIC, E. E., VALENSTEIN, P., VIALE, G., VISSCHER, D., WHEELER, T., WILLIAMS, R. B., WITTLIFF, J. L., WOLFF, A. C., AMERICAN SOCIETY OF CLINICAL, O. & COLLEGE OF AMERICAN, P. 2010. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer (unabridged version). *Arch Pathol Lab Med,* 134**,** e48-72.
HANAHAN, D. & COUSSENS, L. M. 2012. Accessories to the crime: functions of cells recruited to the tumor microenvironment. *Cancer cell,* 21**,** 309-22.
HANAHAN, D. & WEINBERG, R. A. 2000. The hallmarks of cancer. *Cell,* 100**,** 57-70.
HANSFORD, S., KAURAH, P., LI-CHANG, H., WOO, M., SENZ, J., PINHEIRO, H., SCHRADER, K. A., SCHAEFFER, D. F., SHUMANSKY, K., ZOGOPOULOS, G., SANTOS, T. A., CLARO, I., CARVALHO, J., NIELSEN, C., PADILLA, S., LUM, A., TALHOUK, A., BAKER-LANGE, K., RICHARDSON, S., LEWIS, I., LINDOR, N. M., PENNELL, E., MACMILLAN, A., FERNANDEZ, B., KELLER, G., LYNCH, H., SHAH, S. P., GUILFORD, P., GALLINGER, S., CORSO, G., ROVIELLO, F., CALDAS, C., OLIVEIRA, C., PHAROAH, P. D. & HUNTSMAN, D. G. 2015. Hereditary Diffuse Gastric Cancer Syndrome: CDH1 Mutations and Beyond. *JAMA Oncol,* 1**,** 23-32.

HARRELL, J. C., PRAT, A., PARKER, J. S., FAN, C., HE, X., CAREY, L., ANDERS, C., EWEND, M. & PEROU, C. M. 2012. Genomic analysis identifies unique signatures predictive of brain, lung, and liver relapse. *Breast Cancer Res Treat,* 132**,** 523-35.
HARRIS, L., FRITSCHE, H., MENNEL, R., NORTON, L., RAVDIN, P., TAUBE, S., SOMERFIELD, M. R., HAYES, D. F., BAST, R. C., JR. & AMERICAN SOCIETY OF CLINICAL, O. 2007. American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer. *J Clin Oncol,* 25**,** 5287-312.
HARRIS, M., HOWELL, A., CHRISSOHOU, M., SWINDELL, R. I., HUDSON, M. & SELLWOOD, R. A. 1984. A comparison of the metastatic pattern of infiltrating lobular carcinoma and infiltrating duct carcinoma of the breast. *Br J Cancer,* 50**,** 23-30.
HARRIS, T. J. & TEPASS, U. 2010. Adherens junctions: from molecules to morphogenesis. *Nat Rev Mol Cell Biol,* 11**,** 502-14.
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. *J Clin Oncol,* 17**,** 1474-81.
HELWANI, F. M., KOVACS, E. M., PATERSON, A. D., VERMA, S., ALI, R. G., FANNING, A. S., WEED, S. A. & YAP, A. S. 2004. Cortactin is necessary for E-cadherin-mediated contact formation and actin reorganization. *J Cell Biol,* 164**,** 899-910.
HIRATA, D., YAMABUKI, T., MIKI, D., ITO, T., TSUCHIYA, E., FUJITA, M., HOSOKAWA, M., CHAYAMA, K., NAKAMURA, Y. & DAIGO, Y. 2009. Involvement of epithelial cell transforming sequence-2 oncoantigen in lung and esophageal cancer progression. *Clin Cancer Res,* 15**,** 256-66.
HISAMATSU, Y., TOKUNAGA, E., YAMASHITA, N., AKIYOSHI, S., OKADA, S., NAKASHIMA, Y., AISHIMA, S., MORITA, M., KAKEJI, Y. & MAEHARA, Y. 2012. Impact of FOXA1 expression on the prognosis of patients with hormone receptor-positive breast cancer. *Ann Surg Oncol,* 19**,** 1145-52.
HO, A. S., KANNAN, K., ROY, D. M., MORRIS, L. G., GANLY, I., KATABI, N., RAMASWAMI, D., WALSH, L. A., ENG, S., HUSE, J. T., ZHANG, J., DOLGALEV, I., HUBERMAN, K., HEGUY, A., VIALE, A., DROBNJAK, M., LEVERSHA, M. A., RICE, C. E., SINGH, B., IYER, N. G., LEEMANS, C. R., BLOEMENA, E., FERRIS, R. L., SEETHALA, R. R., GROSS, B. E., LIANG, Y., SINHA, R., PENG, L., RAPHAEL, B. J., TURCAN, S., GONG, Y., SCHULTZ, N., KIM, S., CHIOSEA, S., SHAH, J. P., SANDER, C., LEE, W. & CHAN, T. A. 2013. The mutational landscape of adenoid cystic carcinoma. *Nat Genet,* 45**,** 791-8.
HOCH, R. V., THOMPSON, D. A., BAKER, R. J. & WEIGEL, R. J. 1999. GATA-3 is expressed in association with estrogen receptor in breast cancer. *Int J Cancer,* 84**,** 122-8.
HOEFNAGEL, L. D., MOELANS, C. B., MEIJER, S. L., VAN SLOOTEN, H. J., WESSELING, P., WESSELING, J., WESTENEND, P. J., BART, J., SELDENRIJK, C. A., NAGTEGAAL, I. D., OUDEJANS, J., VAN DER VALK, P., VAN GILS, C. H., VAN DER WALL, E. & VAN DIEST, P. J. 2012. Prognostic value of estrogen receptor alpha and progesterone receptor conversion in distant breast cancer metastases. *Cancer,* 118**,** 4929-35.
HOGAN, C., DUPRE-CROCHET, S., NORMAN, M., KAJITA, M., ZIMMERMANN, C., PELLING, A. E., PIDDINI, E., BAENA-LOPEZ, L. A., VINCENT, J. P., ITOH, Y., HOSOYA, H., PICHAUD, F. & FUJITA, Y. 2009. Characterization of the interface between normal and transformed epithelial cells. *Nat Cell Biol,* 11**,** 460-7.
HONMA, N., HORII, R., IWASE, T., SAJI, S., YOUNES, M., TAKUBO, K., MATSUURA, M., ITO, Y., AKIYAMA, F. & SAKAMOTO, G. 2008. Clinical importance of estrogen

receptor-beta evaluation in breast cancer patients treated with adjuvant tamoxifen therapy. *J Clin Oncol,* 26**,** 3727-34.
HU, Z., FAN, C., OH, D. S., MARRON, J. S., HE, X., QAQISH, B. F., LIVASY, C., CAREY, L. A., REYNOLDS, E., DRESSLER, L., NOBEL, A., PARKER, J., EWEND, M. G., SAWYER, L. R., WU, J., LIU, Y., NANDA, R., TRETIAKOVA, M., RUIZ ORRICO, A., DREHER, D., PALAZZO, J. P., PERREARD, L., NELSON, E., MONE, M., HANSEN, H., MULLINS, M., QUACKENBUSH, J. F., ELLIS, M. J., OLOPADE, O. I., BERNARD, P. S. & PEROU, C. M. 2006. The molecular portraits of breast tumors are conserved across microarray platforms. *BMC Genomics,* 7**,** 96.
HUANG, C., LIU, J., HAUDENSCHILD, C. C. & ZHAN, X. 1998. The role of tyrosine phosphorylation of cortactin in the locomotion of endothelial cells. *J Biol Chem,* 273**,** 25770-6.
HUI, R., BALL, J. R., MACMILLAN, R. D., KENNY, F. S., PRALL, O. W., CAMPBELL, D. H., CORNISH, A. L., MCCLELLAND, R. A., DALY, R. J., FORBES, J. F., BLAMEY, R. W., MUSGROVE, E. A., ROBERTSON, J. F., NICHOLSON, R. I. & SUTHERLAND, R. L. 1998. EMS1 gene expression in primary breast cancer: relationship to cyclin D1 and oestrogen receptor expression and patient survival. *Oncogene,* 17**,** 1053-9.
HUI, R., CAMPBELL, D. H., LEE, C. S., MCCAUL, K., HORSFALL, D. J., MUSGROVE, E. A., DALY, R. J., SESHADRI, R. & SUTHERLAND, R. L. 1997. EMS1 amplification can occur independently of CCND1 or INT-2 amplification at 11q13 and may identify different phenotypes in primary breast cancer. *Oncogene,* 15**,** 1617-23.
HURTADO, A., HOLMES, K. A., GEISTLINGER, T. R., HUTCHESON, I. R., NICHOLSON, R. I., BROWN, M., JIANG, J., HOWAT, W. J., ALI, S. & CARROLL, J. S. 2008. Regulation of ERBB2 by oestrogen receptor-PAX2 determines response to tamoxifen. *Nature,* 456**,** 663-6.
HURTADO, A., HOLMES, K. A., ROSS-INNES, C. S., SCHMIDT, D. & CARROLL, J. S. 2011. FOXA1 is a key determinant of estrogen receptor function and endocrine response. *Nat Genet,* 43**,** 27-33.
HUSEMANN, Y., GEIGL, J. B., SCHUBERT, F., MUSIANI, P., MEYER, M., BURGHART, E., FORNI, G., EILS, R., FEHM, T., RIETHMULLER, G. & KLEIN, C. A. 2008. Systemic spread is an early step in breast cancer. *Cancer Cell,* 13**,** 58-68.
HWANG, E. S., NYANTE, S. J., YI CHEN, Y., MOORE, D., DEVRIES, S., KORKOLA, J. E., ESSERMAN, L. J. & WALDMAN, F. M. 2004. Clonality of lobular carcinoma in situ and synchronous invasive lobular carcinoma. *Cancer,* 100**,** 2562-72.
IDIRISINGHE, P. K., THIKE, A. A., CHEOK, P. Y., TSE, G. M., LUI, P. C., FOOK-CHONG, S., WONG, N. S. & TAN, P. H. 2010. Hormone receptor and c-ERBB2 status in distant metastatic and locally recurrent breast cancer. Pathologic correlations and clinical significance. *American journal of clinical pathology,* 133**,** 416-29.
INGOLF, J. B., RUSSALINA, M., SIMONA, M., JULIA, R., GILDA, S., BOHLE, R. M., ANDREA, H., ERICH, S. & DANIEL, H. 2014. Can ki-67 play a role in prediction of breast cancer patients' response to neoadjuvant chemotherapy? *Biomed Res Int,* 2014**,** 628217.
INTRA, M., ROTMENSZ, N., VIALE, G., MARIANI, L., BONANNI, B., MASTROPASQUA, M. G., GALIMBERTI, V., GENNARI, R., VERONESI, P., COLLEONI, M., TOUSIMIS, E., GALLI, A., GOLDHIRSCH, A. & VERONESI, U. 2004. Clinicopathologic characteristics of 143 patients with synchronous bilateral invasive breast carcinomas treated in a single institution. *Cancer,* 101**,** 905-12.
ISOLA, J. J. 1993. Immunohistochemical demonstration of androgen receptor in breast cancer and its relationship to other prognostic factors. *J Pathol,* 170**,** 31-5.
JACQUEMIER, J., CHARAFE-JAUFFRET, E., MONVILLE, F., ESTERNI, B., EXTRA, J. M., HOUVENAEGHEL, G., XERRI, L., BERTUCCI, F. & BIRNBAUM, D. 2009.

Association of GATA3, P53, Ki67 status and vascular peritumoral invasion are strongly prognostic in luminal breast cancer. *Breast Cancer Res,* 11**,** R23.
JAIN, S., FISHER, C., SMITH, P., MILLIS, R. R. & RUBENS, R. D. 1993. Patterns of metastatic breast cancer in relation to histological type. *European journal of cancer,* 29A**,** 2155-7.
JENG, M. H., YUE, W., EISCHEID, A., WANG, J. P. & SANTEN, R. J. 2000. Role of MAP kinase in the enhanced cell proliferation of long term estrogen deprived human breast cancer cells. *Breast Cancer Res Treat,* 62**,** 167-75.
JESELSOHN, R., BUCHWALTER, G., ANGELIS, C., BROWN, M. & SCHIFF, R. 2015. ESR1 mutations-a mechanism for acquired endocrine resistance in breast cancer. *Nat Rev Clin Oncol*.
JESELSOHN, R., YELENSKY, R., BUCHWALTER, G., FRAMPTON, G., MERIC-BERNSTAM, F., GONZALEZ-ANGULO, A. M., FERRER-LOZANO, J., PEREZ-FIDALGO, J. A., CRISTOFANILLI, M., GOMEZ, H., ARTEAGA, C. L., GILTNANE, J., BALKO, J. M., CRONIN, M. T., JAROSZ, M., SUN, J., HAWRYLUK, M., LIPSON, D., OTTO, G., ROSS, J. S., DVIR, A., SOUSSAN-GUTMAN, L., WOLF, I., RUBINEK, T., GILMORE, L., SCHNITT, S., COME, S. E., PUSZTAI, L., STEPHENS, P., BROWN, M. & MILLER, V. A. 2014. Emergence of constitutively active estrogen receptor-alpha mutations in pretreated advanced estrogen receptor-positive breast cancer. *Clin Cancer Res,* 20**,** 1757-67.
JIANG, S., KATAYAMA, H., WANG, J., LI, S. A., HONG, Y., RADVANYI, L., LI, J. J. & SEN, S. 2010. Estrogen-induced aurora kinase-A (AURKA) gene expression is activated by GATA-3 in estrogen receptor-positive breast cancer cells. *Horm Cancer,* 1**,** 11-20.
JIANG, Y. Z., YU, K. D., ZUO, W. J., PENG, W. T. & SHAO, Z. M. 2014. GATA3 mutations define a unique subtype of luminal-like breast cancer with improved survival. *Cancer,* 120**,** 1329-37.
JUNANKAR, S., BAKER, L. A., RODEN, D. L., NAIR, R., ELSWORTH, B., GALLEGO-ORTEGA, D., LACAZE, P., CAZET, A., NIKOLIC, I., TEO, W. S., YANG, J., MCFARLAND, A., HARVEY, K., NAYLOR, M. J., LAKHANI, S. R., SIMPSON, P. T., RAGHAVENDRA, A., SAUNUS, J., MADORE, J., KAPLAN, W., ORMANDY, C., MILLAR, E. K., O'TOOLE, S., YUN, K. & SWARBRICK, A. 2015. ID4 controls mammary stem cells and marks breast cancers with a stem cell-like phenotype. *Nat Commun,* 6**,** 6548.
JUSTILIEN, V. & FIELDS, A. P. 2009. Ect2 links the PKCiota-Par6alpha complex to Rac1 activation and cellular transformation. *Oncogene,* 28**,** 3597-607.
KALINSKY, K., HEGUY, A., BHANOT, U. K., PATIL, S. & MOYNAHAN, M. E. 2011. PIK3CA mutations rarely demonstrate genotypic intratumoral heterogeneity and are selected for in breast cancer progression. *Breast Cancer Res Treat,* 129**,** 635-43.
KAPLAN, R. N., RIBA, R. D., ZACHAROULIS, S., BRAMLEY, A. H., VINCENT, L., COSTA, C., MACDONALD, D. D., JIN, D. K., SHIDO, K., KERNS, S. A., ZHU, Z., HICKLIN, D., WU, Y., PORT, J. L., ALTORKI, N., PORT, E. R., RUGGERO, D., SHMELKOV, S. V., JENSEN, K. K., RAFII, S. & LYDEN, D. 2005. VEGFR1-positive haematopoietic bone marrow progenitors initiate the pre-metastatic niche. *Nature,* 438**,** 820-7.
KARIHTALA, P., AUVINEN, P., KAUPPILA, S., HAAPASAARI, K. M., JUKKOLA-VUORINEN, A. & SOINI, Y. 2013. Vimentin, zeb1 and Sip1 are up-regulated in triple-negative and basal-like breast cancers: association with an aggressive tumour phenotype. *Breast Cancer Res Treat,* 138**,** 81-90.
KARNIK, P. S., KULKARNI, S., LIU, X. P., BUDD, G. T. & BUKOWSKI, R. M. 1994. Estrogen receptor mutations in tamoxifen-resistant breast cancer. *Cancer Res,* 54**,** 349-53.

KASHIWABA, M., TAMURA, G., SUZUKI, Y., MAESAWA, C., OGASAWARA, S., SAKATA, K. & SATODATE, R. 1995. Epithelial-cadherin gene is not mutated in ductal carcinomas of the breast. *Jpn J Cancer Res,* 86**,** 1054-9.

KASSAHN, K. S., HOLMES, O., NONES, K., PATCH, A. M., MILLER, D. K., CHRIST, A. N., HARLIWONG, I., BRUXNER, T. J., XU, Q., ANDERSON, M., WOOD, S., LEONARD, C., TAYLOR, D., NEWELL, F., SONG, S., IDRISOGLU, S., NOURSE, C., NOURBAKHSH, E., MANNING, S., WANI, S., STEPTOE, A., PAJIC, M., COWLEY, M. J., PINESE, M., CHANG, D. K., GILL, A. J., JOHNS, A. L., WU, J., WILSON, P. J., FINK, L., BIANKIN, A. V., WADDELL, N., GRIMMOND, S. M. & PEARSON, J. V. 2013. Somatic point mutation calling in low cellularity tumors. *PloS One*, 8, e74380.

KATZENELLENBOGEN, B. S. & KATZENELLENBOGEN, J. A. 2000. Estrogen receptor transcription and transactivation: Estrogen receptor alpha and estrogen receptor beta: regulation by selective estrogen receptor modulators and importance in breast cancer. *Breast Cancer Res,* 2**,** 335-44.

KE, H. L., KE, R. H., LI, S. T., LI, B., LU, H. T. & WANG, X. Q. 2013. Expression of RACGAP1 in high grade meningiomas: a potential role in cancer progression. *J Neurooncol,* 113**,** 327-32.

KIM, S. W., KIM, S. J., LANGLEY, R. R. & FIDLER, I. J. 2015. Modulation of the cancer cell transcriptome by culture media formulations and cell density. *Int J Oncol*.

KLAUS, A. & BIRCHMEIER, W. 2008. Wnt signalling and its impact on development and cancer. *Nat Rev Cancer,* 8**,** 387-98.

KLEIN, C. A. 2009. Parallel progression of primary tumours and metastases. *Nature reviews. Cancer,* 9**,** 302-12.

KNUDSON, A. G., JR. 1971. Mutation and cancer: statistical study of retinoblastoma. *Proceedings of the National Academy of Sciences of the United States of America,* 68**,** 820-3.

KONDI-PAFITI, A., KAIRI-VASILATOU, E., IAVAZZO, C., DASTAMANI, C., BAKALIANOU, K., LIAPIS, A., HASSIAKOS, D. & FOTIOU, S. 2011. Metastatic neoplasms of the ovaries: a clinicopathological study of 97 cases. *Archives of gynecology and obstetrics,* 284**,** 1283-8.

KORKOLA, J. E., DEVRIES, S., FRIDLYAND, J., HWANG, E. S., ESTEP, A. L., CHEN, Y. Y., CHEW, K. L., DAIRKEE, S. H., JENSEN, R. M. & WALDMAN, F. M. 2003. Differentiation of lobular versus ductal breast carcinomas by expression microarray analysis. *Cancer Res,* 63**,** 7167-75.

KOUROS-MEHR, H., SLORACH, E. M., STERNLICHT, M. D. & WERB, Z. 2006. GATA-3 maintains the differentiation of the luminal cell fate in the mammary gland. *Cell,* 127**,** 1041-55.

KOVACS, E. M., VERMA, S., ALI, R. G., RATHEESH, A., HAMILTON, N. A., AKHMANOVA, A. & YAP, A. S. 2011. N-WASP regulates the epithelial junctional actin cytoskeleton through a non-canonical post-nucleation pathway. *Nature cell biology,* 13**,** 934-43.

KUENEN-BOUMEESTER, V., VAN DER KWAST, T. H., VAN PUTTEN, W. L., CLAASSEN, C., VAN OOIJEN, B. & HENZEN-LOGMANS, S. C. 1992. Immunohistochemical determination of androgen receptors in relation to oestrogen and progesterone receptors in female breast cancer. *Int J Cancer,* 52**,** 581-4.

KUIPER, G. G., LEMMEN, J. G., CARLSSON, B., CORTON, J. C., SAFE, S. H., VAN DER SAAG, P. T., VAN DER BURG, B. & GUSTAFSSON, J. A. 1998. Interaction of estrogenic chemicals and phytoestrogens with estrogen receptor beta. *Endocrinology,* 139**,** 4252-63.

KUTASOVIC, J. R. 2011. *Investigating the molecular evolution of lobular breast carcinomas.* BSc Honours, University of Queensland.

KUTASOVIC, J. R., SIM, S. Y. M., MCCART REED, A. E., CUMMINGS, M. C. & SIMPSON, P. T. 2014. Intratumour heterogeneity in the progression to breast cancer metastasis. *Cancer Forum,* 38**,** 138-142.

KUUKASJARVI, T., KARHU, R., TANNER, M., KAHKONEN, M., SCHAFFER, A., NUPPONEN, N., PENNANEN, S., KALLIONIEMI, A., KALLIONIEMI, O. P. & ISOLA, J. 1997a. Genetic heterogeneity and clonal evolution underlying development of asynchronous metastasis in human breast cancer. *Cancer Res,* 57**,** 1597-604.

KUUKASJARVI, T., TANNER, M., PENNANEN, S., KARHU, R., KALLIONIEMI, O. P. & ISOLA, J. 1997b. Genetic changes in intraductal breast cancer detected by comparative genomic hybridization. *Am J Pathol,* 150**,** 1465-71.

LAGANIERE, J., DEBLOIS, G., LEFEBVRE, C., BATAILLE, A. R., ROBERT, F. & GIGUERE, V. 2005. From the Cover: Location analysis of estrogen receptor alpha target promoters reveals that FOXA1 defines a domain of the estrogen response. *Proc Natl Acad Sci U S A,* 102**,** 11651-6.

LAGERWAARD, F. J., LEVENDAG, P. C., NOWAK, P. J., EIJKENBOOM, W. M., HANSSENS, P. E. & SCHMITZ, P. I. 1999. Identification of prognostic factors in patients with brain metastases: a review of 1292 patients. *Int J Radiat Oncol Biol Phys,* 43**,** 795-803.

LAI, A., KAHRAMAN, M., GOVEK, S., NAGASAWA, J., BONNEFOUS, C., JULIEN, J., DOUGLAS, K., SENSINTAFFAR, J., LU, N., LEE, K. J., APARICIO, A., KAUFMAN, J., QIAN, J., SHAO, G., PRUDENTE, R., MOON, M. J., JOSEPH, J. D., DARIMONT, B., BRIGHAM, D., GRILLOT, K., HEYMAN, R., RIX, P. J., HAGER, J. H. & SMITH, N. D. 2015. Identification of GDC-0810 (ARN-810), an Orally Bioavailable Selective Estrogen Receptor Degrader (SERD) that Demonstrates Robust Activity in Tamoxifen-Resistant Breast Cancer Xenografts. *J Med Chem,* 58**,** 4888-904.

LAKHANI S. R., E. I. O., SCHINITT S. J., TAN .H., VAN DE VIJVER M. J. (EDS.) 2012. *WHO Classification of Tumours of the Breast*.

LAKHANI, S. R., CHAGGAR, R., DAVIES, S., JONES, C., COLLINS, N., ODEL, C., STRATTON, M. R. & O'HARE, M. J. 1999. Genetic alterations in 'normal' luminal and myoepithelial cells of the breast. *J Pathol,* 189**,** 496-503.

LAMOVEC, J. & BRACKO, M. 1991. Metastatic pattern of infiltrating lobular carcinoma of the breast: an autopsy study. *Journal of surgical oncology,* 48**,** 28-33.

LANDAU, D. A., CARTER, S. L., STOJANOV, P., MCKENNA, A., STEVENSON, K., LAWRENCE, M. S., SOUGNEZ, C., STEWART, C., SIVACHENKO, A., WANG, L., WAN, Y., ZHANG, W., SHUKLA, S. A., VARTANOV, A., FERNANDES, S. M., SAKSENA, G., CIBULSKIS, K., TESAR, B., GABRIEL, S., HACOHEN, N., MEYERSON, M., LANDER, E. S., NEUBERG, D., BROWN, J. R., GETZ, G. & WU, C. J. 2013. Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell,* 152**,** 714-26.

LAWRENCE, M., HUBER, W., PAGES, H., ABOYOUN, P., CARLSON, M., GENTLEMAN, R., MORGAN, M. T. & CAREY, V. J. 2013. Software for computing and annotating genomic ranges. *PLoS Comput Biol,* 9**,** e1003118.

LEE, Y. T. 1983. Breast carcinoma: pattern of metastasis at autopsy. *J Surg Oncol,* 23**,** 175-80.

LEUNG, C. T. & BRUGGE, J. S. 2012. Outgrowth of single oncogene-expressing cells from suppressive epithelial environments. *Nature,* 482**,** 410-3.

LI, C. I., ANDERSON, B. O., DALING, J. R. & MOE, R. E. 2003. Trends in incidence rates of invasive lobular and ductal breast carcinoma. *JAMA,* 289**,** 1421-4.

LI, C. I., ANDERSON, B. O., PORTER, P., HOLT, S. K., DALING, J. R. & MOE, R. E. 2000. Changing incidence rate of invasive lobular breast carcinoma among older women. *Cancer,* 88**,** 2561-9.

LI, C. I., CHLEBOWSKI, R. T., FREIBERG, M., JOHNSON, K. C., KULLER, L., LANE, D., LESSIN, L., O'SULLIVAN, M. J., WACTAWSKI-WENDE, J., YASMEEN, S. & PRENTICE, R. 2010a. Alcohol consumption and risk of postmenopausal breast cancer by subtype: the women's health initiative observational study. *J Natl Cancer Inst,* 102**,** 1422-31.

LI, C. I., DALING, J. R., MALONE, K. E., BERNSTEIN, L., MARCHBANKS, P. A., LIFF, J. M., STROM, B. L., SIMON, M. S., PRESS, M. F., MCDONALD, J. A., URSIN, G., BURKMAN, R. T., DEAPEN, D. & SPIRTAS, R. 2006. Relationship between established breast cancer risk factors and risk of seven different histologic types of invasive breast cancer. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology,* 15**,** 946-54.

LI, C. I., URIBE, D. J. & DALING, J. R. 2005. Clinical characteristics of different histologic types of breast cancer. *Br J Cancer,* 93**,** 1046-52.

LI, H. & DURBIN, R. 2009 Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 24, 1754-60.

LAWRENCE, M. S., STOJANOV, P., POLAK, P., KRYUKOV, G. V., CIBULSKIS, K., SIVACHENKO, A., CARTER, S. L., STEWART, C., MERMEL, C. H., ROBERTS, S. A., KIEZUN, A., HAMMERMAN, P. S., MCKENNA, A., DRIER, Y., ZOU, L., RAMOS, A. H., PUGH, T. J., STRANSKY, N., HELMAN, E., KIM, J., SOUGNEZ, C., AMBROGIO, L., NICKERSON, E., SHEFLER, E., CORTES, M. L., AUCLAIR, D., SAKSENA, G., VOET, D., NOBLE, M., DICARA, D., LIN, P., LICHTENSTEIN, L., HEIMAN, D. I., FENNELL, T., IMIELINSKI, M., HERNANDEZ, B., HODIS, E., BACA, S., DULAK, A. M., LOHR, J., LANDAU, D. A., WU, C. J., MELENDEZ-ZAJGLA, J., HIDALGO-MIRANDA, A., KOREN, A., MCCARROLL, S. A., MORA, J., LEE, R. S., CROMPTON, B., ONOFRIO, R., PARKIN, M., WINCKLER, W., ARDLIE, K., GABRIEL, S. B., ROBERTS, C. W., BIEGEL, J. A., STEGMAIER, K., BASS, A. J., GARRAWAY, L. A., MEYERSON, M., GOLUB, T. R., GORDENIN, D. A., SUNYAEV, S., LANDER, E. S. & GETZ, G. 2013. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature*, 499, 214-8.

LI, H., ZHU, R., WANG, L., ZHU, T., LI, Q., CHEN, Q., WANG, H. & ZHU, H. 2010b. PIK3CA mutations mostly begin to develop in ductal carcinoma of the breast. *Exp Mol Pathol,* 88**,** 150-5.

LI, S., SHEN, D., SHAO, J., CROWDER, R., LIU, W., PRAT, A., HE, X., LIU, S., HOOG, J., LU, C., DING, L., GRIFFITH, O. L., MILLER, C., LARSON, D., FULTON, R. S., HARRISON, M., MOONEY, T., MCMICHAEL, J. F., LUO, J., TAO, Y., GONCALVES, R., SCHLOSBERG, C., HIKEN, J. F., SAIED, L., SANCHEZ, C., GIUNTOLI, T., BUMB, C., COOPER, C., KITCHENS, R. T., LIN, A., PHOMMALY, C., DAVIES, S. R., ZHANG, J., KAVURI, M. S., MCEACHERN, D., DONG, Y. Y., MA, C., PLUARD, T., NAUGHTON, M., BOSE, R., SURESH, R., MCDOWELL, R., MICHEL, L., AFT, R., GILLANDERS, W., DESCHRYVER, K., WILSON, R. K., WANG, S., MILLS, G. B., GONZALEZ-ANGULO, A., EDWARDS, J. R., MAHER, C., PEROU, C. M., MARDIS, E. R. & ELLIS, M. J. 2013. Endocrine-therapy-resistant ESR1 variants revealed by genomic characterization of breast-cancer-derived xenografts. *Cell Rep,* 4**,** 1116-30.

LI, Y., WANG, J. P., SANTEN, R. J., KIM, T. H., PARK, H., FAN, P. & YUE, W. 2010c. Estrogen stimulation of cell migration involves multiple signaling pathway interactions. *Endocrinology,* 151**,** 5146-56.

LIEDTKE, C., RODY, A., GLUZ, O., BAUMANN, K., BEYER, D., KOHLS, E. B., LAUSEN, K., HANKER, L., HOLTRICH, U., BECKER, S. & KARN, T. 2015. The prognostic impact of age in different molecular subtypes of breast cancer. *Breast Cancer Res Treat*.

LIU, H., SHI, J., WILKERSON, M. L. & LIN, F. 2012. Immunohistochemical evaluation of GATA3 expression in tumors and normal tissues: a useful immunomarker for breast and urothelial carcinomas. *Am J Clin Pathol,* 138**,** 57-64.

LIU, W., LAITINEN, S., KHAN, S., VIHINEN, M., KOWALSKI, J., YU, G., CHEN, L., EWING, C. M., EISENBERGER, M. A., CARDUCCI, M. A., NELSON, W. G., YEGNASUBRAMANIAN, S., LUO, J., WANG, Y., XU, J., ISAACS, W. B., VISAKORPI, T. & BOVA, G. S. 2009. Copy number analysis indicates monoclonal origin of lethal metastatic prostate cancer. *Nat Med,* 15**,** 559-65.

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,* 57**,** 171-92.

LU, K. H., PATTERSON, A. P., WANG, L., MARQUEZ, R. T., ATKINSON, E. N., BAGGERLY, K. A., RAMOTH, L. R., ROSEN, D. G., LIU, J., HELLSTROM, I., SMITH, D., HARTMANN, L., FISHMAN, D., BERCHUCK, A., SCHMANDT, R., WHITAKER, R., GERSHENSON, D. M., MILLS, G. B. & BAST, R. C., JR. 2004. Selection of potential markers for epithelial ovarian cancer with gene expression arrays and recursive descent partition analysis. *Clin Cancer Res,* 10**,** 3291-300.

LU, Y. J., OSIN, P., LAKHANI, S. R., DI PALMA, S., GUSTERSON, B. A. & SHIPLEY, J. M. 1998. Comparative genomic hybridization analysis of lobular carcinoma in situ and atypical lobular hyperplasia and potential roles for gains and losses of genetic material in breast neoplasia. *Cancer Res,* 58**,** 4721-7.

LUPIEN, M., MEYER, C. A., BAILEY, S. T., EECKHOUTE, J., COOK, J., WESTERLING, T., ZHANG, X., CARROLL, J. S., RHODES, D. R., LIU, X. S. & BROWN, M. 2010. Growth factor stimulation induces a distinct ER(alpha) cistrome underlying breast cancer endocrine resistance. *Genes Dev,* 24**,** 2219-27.

LYNCH, E. D., OSTERMEYER, E. A., LEE, M. K., ARENA, J. F., JI, H., DANN, J., SWISSHELM, K., SUCHARD, D., MACLEOD, P. M., KVINNSLAND, S., GJERTSEN, B. T., HEIMDAL, K., LUBS, H., MOLLER, P. & KING, M. C. 1997. Inherited mutations in PTEN that are associated with breast cancer, cowden disease, and juvenile polyposis. *Am J Hum Genet,* 61**,** 1254-60.

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. *Proc Natl Acad Sci U S A,* 100**,** 5974-9.

MAKI, D. D. & GROSSMAN, R. I. 2000. Patterns of disease spread in metastatic breast carcinoma: influence of estrogen and progesterone receptor status. *AJNR. American journal of neuroradiology,* 21**,** 1064-6.

MANGOLD, S., WU, S. K., NORWOOD, S. J., COLLINS, B. M., HAMILTON, N. A., THORN, P. & YAP, A. S. 2011. Hepatocyte growth factor acutely perturbs actin filament anchorage at the epithelial zonula adherens. *Curr Biol,* 21**,** 503-7.

MARTIN, L. A., FARMER, I., JOHNSTON, S. R., ALI, S., MARSHALL, C. & DOWSETT, M. 2003. Enhanced estrogen receptor (ER) alpha, ERBB2, and MAPK signal transduction pathways operate during the adaptation of MCF-7 cells to long term estrogen deprivation. *J Biol Chem,* 278**,** 30458-68.

MARTIN TA, TOMS AM, DAVIES LM, CHENG S & WG, J. 2012. The clinical and biological implications of N-WASP expression in human colorectal cancer. *Transl Gastrointest Cancer,* 1**,** 10-20.

MARTIN, T. A., PEREIRA, G., WATKINS, G., MANSEL, R. E. & JIANG, W. G. 2008. N-WASP is a putative tumour suppressor in breast cancer cells, in vitro and in vivo, and is associated with clinical outcome in patients with breast cancer. *Clin Exp Metastasis,* 25**,** 97-108.
MASTRACCI, T. L., SHADEO, A., COLBY, S. M., TUCK, A. B., O'MALLEY, F. P., BULL, S. B., LAM, W. L. & ANDRULIS, I. L. 2006. Genomic alterations in lobular neoplasia: a microarray comparative genomic hybridization signature for early neoplastic proliferationin the breast. *Genes, chromosomes & cancer,* 45**,** 1007-17.
MASTRACCI, T. L., TJAN, S., BANE, A. L., O'MALLEY, F. P. & ANDRULIS, I. L. 2005. E-cadherin alterations in atypical lobular hyperplasia and lobular carcinoma in situ of the breast. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc,* 18**,** 741-51.
MATTHEWS, H. K., DELABRE, U., ROHN, J. L., GUCK, J., KUNDA, P. & BAUM, B. 2012. Changes in Ect2 localization couple actomyosin-dependent cell shape changes to mitotic progression. *Dev Cell,* 23**,** 371-83.
MAXMEN, A. 2012. The hard facts. *Nature,* 485**,** S50-S51.
MC SHERRY, E. A., MC GOLDRICK, A., KAY, E. W., HOPKINS, A. M., GALLAGHER, W. M. & DERVAN, P. A. 2007. Formalin-fixed paraffin-embedded clinical tissues show spurious copy number changes in array-CGH profiles. *Clin Genet,* 72**,** 441-7.
MCCART REED, A. E., KUTASOVIC, J. R., LAKHANI, S. R. & SIMPSON, P. T. 2015. Invasive lobular carcinoma of the breast: morphology, biomarkers and 'omics. *Breast Cancer Res,* 17**,** 519.
MCFALL, A., ULKU, A., LAMBERT, Q. T., KUSA, A., ROGERS-GRAHAM, K. & DER, C. J. 2001. Oncogenic Ras blocks anoikis by activation of a novel effector pathway independent of phosphatidylinositol 3-kinase. *Mol Cell Biol,* 21**,** 5488-99.
MCGHAN, L. J., MCCULLOUGH, A. E., PROTHEROE, C. A., DUECK, A. C., LEE, J. J., NUNEZ-NATERAS, R., CASTLE, E. P., GRAY, R. J., WASIF, N., GOETZ, M. P., HAWSE, J. R., HENRY, T. J., BARRETT, M. T., CUNLIFFE, H. E. & POCKAJ, B. A. 2014. Androgen receptor-positive triple negative breast cancer: a unique breast cancer subtype. *Ann Surg Oncol,* 21**,** 361-7.
MCKENNA, A., HANNA, M., BANKS, E., SIVACHENKO, A., CIBULSKIS, K., KERNYTSKY, A., GARIMELLA, K., ALTSHULER, D., GABRIEL, S., DALY, M. & DEPRISTO, M. A. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*, 20, 1297-303.
MCVEIGH, T. P., CHOI, J. K., MILLER, N. M., GREEN, A. J. & KERIN, M. J. 2014. Lobular breast cancer in a CDH1 splice site mutation carrier: case report and review of the literature. *Clin Breast Cancer,* 14**,** e47-51.
MEHRA, R., VARAMBALLY, S., DING, L., SHEN, R., SABEL, M. S., GHOSH, D., CHINNAIYAN, A. M. & KLEER, C. G. 2005. Identification of GATA3 as a breast cancer prognostic marker by global gene expression meta-analysis. *Cancer Res,* 65**,** 11259-64.
MEHTA, R. J., JAIN, R. K., LEUNG, S., CHOO, J., NIELSEN, T., HUNTSMAN, D., NAKSHATRI, H. & BADVE, S. 2012. FOXA1 is an independent prognostic marker for ER-positive breast cancer. *Breast Cancer Res Treat,* 131**,** 881-90.
MERENBAKH-LAMIN, K., BEN-BARUCH, N., YEHESKEL, A., DVIR, A., SOUSSAN-GUTMAN, L., JESELSOHN, R., YELENSKY, R., BROWN, M., MILLER, V. A., SARID, D., RIZEL, S., KLEIN, B., RUBINEK, T. & WOLF, I. 2013. D538G mutation in estrogen receptor-alpha: A novel mechanism for acquired endocrine resistance in breast cancer. *Cancer Res,* 73**,** 6856-64.
MERIC-BERNSTAM, F., FRAMPTON, G. M., FERRER-LOZANO, J., YELENSKY, R., PEREZ-FIDALGO, J. A., WANG, Y., PALMER, G. A., ROSS, J. S., MILLER, V. A., SU, X., EROLES, P., BARRERA, J. A., BURGUES, O., LLUCH, A. M., ZHENG, X.,

SAHIN, A., STEPHENS, P. J., MILLS, G. B., CRONIN, M. T. & GONZALEZ-ANGULO, A. M. 2014. Concordance of genomic alterations between primary and recurrent breast cancer. *Mol Cancer Ther,* 13**,** 1382-9.

MERMEL, C. H., SCHUMACHER, S. E., HILL, B., MEYERSON, M. L., BEROUKHIM, R. & GETZ, G. 2011. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol,* 12**,** R41.

MIDDLETON, L. P., PALACIOS, D. M., BRYANT, B. R., KREBS, P., OTIS, C. N. & MERINO, M. J. 2000. Pleomorphic lobular carcinoma: morphology, immunohistochemistry, and molecular analysis. *Am J Surg Pathol,* 24**,** 1650-6.

MILLER, T. W., HENNESSY, B. T., GONZALEZ-ANGULO, A. M., FOX, E. M., MILLS, G. B., CHEN, H., HIGHAM, C., GARCIA-ECHEVERRIA, C., SHYR, Y. & ARTEAGA, C. L. 2010. Hyperactivation of phosphatidylinositol-3 kinase promotes escape from hormone dependence in estrogen receptor-positive human breast cancer. *J Clin Invest,* 120**,** 2406-13.

MIRON, A., VARADI, M., CARRASCO, D., LI, H., LUONGO, L., KIM, H. J., PARK, S. Y., CHO, E. Y., LEWIS, G., KEHOE, S., IGLEHART, J. D., DILLON, D., ALLRED, D. C., MACCONAILL, L., GELMAN, R. & POLYAK, K. 2010. PIK3CA mutations in in situ and invasive breast carcinomas. *Cancer Res,* 70**,** 5674-8.

MOHAMMED, H., RUSSELL, I. A., STARK, R., RUEDA, O. M., HICKEY, T. E., TARULLI, G. A., SERANDOUR, A. A., BIRRELL, S. N., BRUNA, A., SAADI, A., MENON, S., HADFIELD, J., PUGH, M., RAJ, G. V., BROWN, G. D., D'SANTOS, C., ROBINSON, J. L., SILVA, G., LAUNCHBURY, R., PEROU, C. M., STINGL, J., CALDAS, C., TILLEY, W. D. & CARROLL, J. S. 2015. Progesterone receptor modulates ERalpha action in breast cancer. *Nature,* 523**,** 313-7.

MOLENAAR, J. J., EBUS, M. E., KOSTER, J., SANTO, E., GEERTS, D., VERSTEEG, R. & CARON, H. N. 2010. Cyclin D1 is a direct transcriptional target of GATA3 in neuroblastoma tumor cells. *Oncogene,* 29**,** 2739-45.

MOLL, R., MITZE, M., FRIXEN, U. H. & BIRCHMEIER, W. 1993. Differential loss of E-cadherin expression in infiltrating ductal and lobular breast carcinomas. *The American journal of pathology,* 143**,** 1731-42.

MOMENY, M., SAUNUS, J. M., MARTURANA, F., MCCART REED, A. E., BLACK, D., SALA, G., IACOBELLI, S., HOLLAND, J. D., YU, D., DA SILVA, L., SIMPSON, P. T., KHANNA, K. K., CHENEVIX-TRENCH, G. & LAKHANI, S. R. 2015. Heregulin-HER3-HER2 signaling promotes matrix metalloproteinase-dependent blood-brain-barrier transendothelial migration of human breast cancer cell lines. *Oncotarget,* 6**,** 3932-46.

MOORE, R. G., CHUNG, M., GRANAI, C. O., GAJEWSKI, W. & STEINHOFF, M. M. 2004. Incidence of metastasis to the ovaries from nongenital tract primary tumors. *Gynecologic oncology,* 93**,** 87-91.

MOY, L., SLANETZ, P. J., MOORE, R., SATIJA, S., YEH, E. D., MCCARTHY, K. A., HALL, D., STAFFA, M., RAFFERTY, E. A., HALPERN, E. & KOPANS, D. B. 2002. Specificity of mammography and US in the evaluation of a palpable abnormality: retrospective review. *Radiology,* 225**,** 176-81.

MUKHERJEE, S. & CONRAD, S. E. 2005. c-Myc suppresses p21WAF1/CIP1 expression during estrogen signaling and antiestrogen resistance in human breast cancer cells. *J Biol Chem,* 280**,** 17617-25.

MUSGROVE, E. A. & SUTHERLAND, R. L. 2009. Biological determinants of endocrine resistance in breast cancer. *Nat Rev Cancer,* 9**,** 631-43.

NAGAFUCHI, A., SHIRAYOSHI, Y., OKAZAKI, K., YASUDA, K. & TAKEICHI, M. 1987. Transformation of cell adhesion properties by exogenously introduced E-cadherin cDNA. *Nature,* 329**,** 341-3.

NAGAO, T., KINOSHITA, T., HOJO, T., TSUDA, H., TAMURA, K. & FUJIWARA, Y. 2012. The differences in the histological types of breast cancer and the response to neoadjuvant chemotherapy: the relationship between the outcome and the clinicopathological characteristics. *Breast,* 21**,** 289-95.
NAKSHATRI, H. & BADVE, S. 2007. FOXA1 as a therapeutic target for breast cancer. *Expert Opin Ther Targets,* 11**,** 507-14.
NATRAJAN, R., WEIGELT, B., MACKAY, A., GEYER, F. C., GRIGORIADIS, A., TAN, D. S., JONES, C., LORD, C. J., VATCHEVA, R., RODRIGUEZ-PINILLA, S. M., PALACIOS, J., ASHWORTH, A. & REIS-FILHO, J. S. 2010. An integrative genomic and transcriptomic analysis reveals molecular pathways and networks regulated by copy number aberrations in basal-like, HER2 and luminal cancers. *Breast Cancer Res Treat,* 121**,** 575-89.
NAVIN, N., KENDALL, J., TROGE, J., ANDREWS, P., RODGERS, L., MCINDOO, J., COOK, K., STEPANSKY, A., LEVY, D., ESPOSITO, D., MUTHUSWAMY, L., KRASNITZ, A., MCCOMBIE, W. R., HICKS, J. & WIGLER, M. 2011. Tumour evolution inferred by single-cell sequencing. *Nature,* 472**,** 90-4.
NAYLOR, M. J. & ORMANDY, C. J. 2007. Gata-3 and mammary cell fate. *Breast Cancer Res,* 9**,** 302.
NEWBURGER, D. E., KASHEF-HAGHIGHI, D., WENG, Z., SALARI, R., SWEENEY, R. T., BRUNNER, A. L., ZHU, S. X., GUO, X., VARMA, S., TROXELL, M. L., WEST, R. B., BATZOGLOU, S. & SIDOW, A. 2013. Genome evolution during progression to breast cancer. *Genome Res,* 23**,** 1097-108.
NEWMAN, W. 1966. Lobular carcinoma of the female breast. Report of 73 cases. *Annals of surgery,* 164**,** 305-14.
NG, C. K., MARTELOTTO, L. G., GAUTHIER, A., WEN, H. C., PISCUOGLIO, S., LIM, R. S., COWELL, C. F., WILKERSON, P. M., WAI, P., RODRIGUES, D. N., ARNOULD, L., GEYER, F. C., BROMBERG, S. E., LACROIX-TRIKI, M., PENAULT-LLORCA, F., GIARD, S., SASTRE-GARAU, X., NATRAJAN, R., NORTON, L., COTTU, P. H., WEIGELT, B., VINCENT-SALOMON, A. & REIS-FILHO, J. S. 2015. Intra-tumor genetic heterogeneity and alternative driver genetic alterations in breast cancers with heterogeneous HER2 gene amplification. *Genome Biol,* 16**,** 107.
NI, M., CHEN, Y., LIM, E., WIMBERLY, H., BAILEY, S. T., IMAI, Y., RIMM, D. L., LIU, X. S. & BROWN, M. 2011. Targeting androgen receptor in estrogen receptor-negative breast cancer. *Cancer Cell,* 20**,** 119-31.
NIELSEN, T. O., HSU, F. D., JENSEN, K., CHEANG, M., KARACA, G., HU, Z., HERNANDEZ-BOUSSARD, T., LIVASY, C., COWAN, D., DRESSLER, L., AKSLEN, L. A., RAGAZ, J., GOWN, A. M., GILKS, C. B., VAN DE RIJN, M. & PEROU, C. M. 2004. Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. *Clin Cancer Res,* 10**,** 5367-74.
NIEMEIER, L. A., DABBS, D. J., BERIWAL, S., STRIEBEL, J. M. & BHARGAVA, R. 2010. Androgen receptor in breast cancer: expression in estrogen receptor-positive tumors and in estrogen receptor-negative tumors with apocrine differentiation. *Mod Pathol,* 23**,** 205-12.
NIK-ZAINAL, S., VAN LOO, P., WEDGE, D. C., ALEXANDROV, L. B., GREENMAN, C. D., LAU, K. W., RAINE, K., JONES, D., MARSHALL, J., RAMAKRISHNA, M., SHLIEN, A., COOKE, S. L., HINTON, J., MENZIES, A., STEBBINGS, L. A., LEROY, C., JIA, M., RANCE, R., MUDIE, L. J., GAMBLE, S. J., STEPHENS, P. J., MCLAREN, S., TARPEY, P. S., PAPAEMMANUIL, E., DAVIES, H. R., VARELA, I., MCBRIDE, D. J., BIGNELL, G. R., LEUNG, K., BUTLER, A. P., TEAGUE, J. W., MARTIN, S., JONSSON, G., MARIANI, O., BOYAULT, S., MIRON, P., FATIMA, A., LANGEROD, A., APARICIO, S. A., TUTT, A., SIEUWERTS, A. M., BORG, A., THOMAS, G., SALOMON, A. V., RICHARDSON, A. L., BORRESEN-DALE, A. L., FUTREAL, P.

A., STRATTON, M. R., CAMPBELL, P. J. & BREAST CANCER WORKING GROUP OF THE INTERNATIONAL CANCER GENOME, C. 2012. The life history of 21 breast cancers. *Cell,* 149**,** 994-1007.
NILSEN, G., LIESTOL, K., VAN LOO, P., MOEN VOLLAN, H. K., EIDE, M. B., RUEDA, O. M., CHIN, S. F., RUSSELL, R., BAUMBUSCH, L. O., CALDAS, C., BORRESEN-DALE, A. L. & LINGJAERDE, O. C. 2012. Copynumber: Efficient algorithms for single- and multi-track copy number segmentation. *BMC Genomics,* 13**,** 591.
NOBES, C. D. & HALL, A. 1995. Rho, rac, and cdc42 GTPases regulate the assembly of multimolecular focal complexes associated with actin stress fibers, lamellipodia, and filopodia. *Cell,* 81**,** 53-62.
NOVELLI, F., MILELLA, M., MELUCCI, E., DI BENEDETTO, A., SPERDUTI, I., PERRONE-DONNORSO, R., PERRACCHIO, L., VENTURO, I., NISTICO, C., FABI, A., BUGLIONI, S., NATALI, P. G. & MOTTOLESE, M. 2008. A divergent role for estrogen receptor-beta in node-positive and node-negative breast cancer classified according to molecular subtypes: an observational prospective study. *Breast Cancer Res,* 10**,** R74.
O'BRIEN, R. N., SHEN, Z., TACHIKAWA, K., LEE, P. A. & BRIGGS, S. P. 2010. Quantitative proteome analysis of pluripotent cells by iTRAQ mass tagging reveals post-transcriptional regulation of proteins required for ES cell self-renewal. *Mol Cell Proteomics,* 9**,** 2238-51.
OGAWA, Y., HAI, E., MATSUMOTO, K., IKEDA, K., TOKUNAGA, S., NAGAHARA, H., SAKURAI, K., INOUE, T. & NISHIGUCHI, Y. 2008. Androgen receptor expression in breast cancer: relationship with clinicopathological factors and biomarkers. *Int J Clin Oncol,* 13**,** 431-5.
OLIVEROS, J. C. 2007-2015. Venny. An interactive tool for comparing lists with Venn's diagrams.
OLSHEN, A. B., VENKATRAMAN, E. S., LUCITO, R. & WIGLER, M. 2004. Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics,* 5**,** 557-72.
ORVIETO, E., MAIORANO, E., BOTTIGLIERI, L., MAISONNEUVE, P., ROTMENSZ, N., GALIMBERTI, V., LUINI, A., BRENELLI, F., GATTI, G. & VIALE, G. 2008. Clinicopathologic characteristics of invasive lobular carcinoma of the breast: results of an analysis of 530 cases from a single institution. *Cancer,* 113**,** 1511-20.
OSBORNE, C. K. & SCHIFF, R. 2011. Mechanisms of endocrine resistance in breast cancer. *Annu Rev Med,* 62**,** 233-47.
OZAWA, M., BARIBAULT, H. & KEMLER, R. 1989. The cytoplasmic domain of the cell adhesion molecule uvomorulin associates with three independent proteins structurally related in different species. *EMBO J,* 8**,** 1711-7.
PAGANI, O., REGAN, M. M., WALLEY, B. A., FLEMING, G. F., COLLEONI, M., LANG, I., GOMEZ, H. L., TONDINI, C., BURSTEIN, H. J., PEREZ, E. A., CIRUELOS, E., STEARNS, V., BONNEFOI, H. R., MARTINO, S., GEYER, C. E., JR., PINOTTI, G., PUGLISI, F., CRIVELLARI, D., RUHSTALLER, T., WINER, E. P., RABAGLIO-PORETTI, M., MAIBACH, R., RUEPP, B., GIOBBIE-HURDER, A., PRICE, K. N., BERNHARD, J., LUO, W., RIBI, K., VIALE, G., COATES, A. S., GELBER, R. D., GOLDHIRSCH, A., FRANCIS, P. A., TEXT, INVESTIGATORS, S. & INTERNATIONAL BREAST CANCER STUDY, G. 2014. Adjuvant exemestane with ovarian suppression in premenopausal breast cancer. *N Engl J Med,* 371**,** 107-18.
PAGET, S. 1889. The distribution of secondary growths in cancer of the breast. *Lancet,* 133**,** 571-573.
PAIK, S., SHAK, S., TANG, G., KIM, C., BAKER, J., CRONIN, M., BAEHNER, F. L., WALKER, M. G., WATSON, D., PARK, T., HILLER, W., FISHER, E. R., WICKERHAM, D. L., BRYANT, J. & WOLMARK, N. 2004. A multigene assay to

predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med,* 351**,** 2817-26.
PARANJAPE, T., HENEGHAN, H., LINDNER, R., KEANE, F. K., HOFFMAN, A., HOLLESTELLE, A., DORAIRAJ, J., GEYDA, K., PELLETIER, C., NALLUR, S., MARTENS, J. W., HOONING, M. J., KERIN, M., ZELTERMAN, D., ZHU, Y., TUCK, D., HARRIS, L., MILLER, N., SLACK, F. & WEIDHAAS, J. 2011. A 3'-untranslated region KRAS variant and triple-negative breast cancer: a case-control and genetic analysis. *Lancet Oncol,* 12**,** 377-86.
PARIKH, P., PALAZZO, J. P., ROSE, L. J., DASKALAKIS, C. & WEIGEL, R. J. 2005. GATA-3 expression as a predictor of hormone response in breast cancer. *J Am Coll Surg,* 200**,** 705-10.
PARISE, C. A., BAUER, K. R., BROWN, M. M. & CAGGIANO, V. 2009. Breast cancer subtypes as defined by the estrogen receptor (ER), progesterone receptor (PR), and the human epidermal growth factor receptor 2 (HER2) among women with invasive breast cancer in California, 1999-2004. *Breast J,* 15**,** 593-602.
PARKER, J. S., MULLINS, M., CHEANG, M. C., LEUNG, S., VODUC, D., VICKERY, T., DAVIES, S., FAURON, C., HE, X., HU, Z., QUACKENBUSH, J. F., STIJLEMAN, I. J., PALAZZO, J., MARRON, J. S., NOBEL, A. B., MARDIS, E., NIELSEN, T. O., ELLIS, M. J., PEROU, C. M. & BERNARD, P. S. 2009. Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol,* 27**,** 1160-7.
PATANAPHAN, V. & SALAZAR, O. M. 1993. Colorectal cancer: metastatic patterns and prognosis. *South Med J,* 86**,** 38-41.
PEINADO, H., ALECKOVIC, M., LAVOTSHKIN, S., MATEI, I., COSTA-SILVA, B., MORENO-BUENO, G., HERGUETA-REDONDO, M., WILLIAMS, C., GARCIA-SANTOS, G., GHAJAR, C., NITADORI-HOSHINO, A., HOFFMAN, C., BADAL, K., GARCIA, B. A., CALLAHAN, M. K., YUAN, J., MARTINS, V. R., SKOG, J., KAPLAN, R. N., BRADY, M. S., WOLCHOK, J. D., CHAPMAN, P. B., KANG, Y., BROMBERG, J. & LYDEN, D. 2012. Melanoma exosomes educate bone marrow progenitor cells toward a pro-metastatic phenotype through MET. *Nature medicine,* 18**,** 883-91.
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,* 406**,** 747-52.
PESTALOZZI, B. C., ZAHRIEH, D., MALLON, E., GUSTERSON, B. A., PRICE, K. N., GELBER, R. D., HOLMBERG, S. B., LINDTNER, J., SNYDER, R., THURLIMANN, B., MURRAY, E., VIALE, G., CASTIGLIONE-GERTSCH, M., COATES, A. S. & GOLDHIRSCH, A. 2008. Distinct clinical and prognostic features of infiltrating lobular carcinoma of the breast: combined results of 15 International Breast Cancer Study Group clinical trials. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology,* 26**,** 3006-14.
PETERS, A. A., BUCHANAN, G., RICCIARDELLI, C., BIANCO-MIOTTO, T., CENTENERA, M. M., HARRIS, J. M., JINDAL, S., SEGARA, D., JIA, L., MOORE, N. L., HENSHALL, S. M., BIRRELL, S. N., COETZEE, G. A., SUTHERLAND, R. L., BUTLER, L. M. & TILLEY, W. D. 2009. Androgen receptor inhibits estrogen receptor-alpha activity and is prognostic in breast cancer. *Cancer Res,* 69**,** 6131-40.
PETERSON, J. R., LOKEY, R. S., MITCHISON, T. J. & KIRSCHNER, M. W. 2001. A chemical inhibitor of N-WASP reveals a new mechanism for targeting protein interactions. *Proc Natl Acad Sci U S A,* 98**,** 10624-9.
PETRIDIS, C., SHINOMIYA, I., KOHUT, K., GORMAN, P., CANEPPELE, M., SHAH, V., TROY, M., PINDER, S. E., HANBY, A., TOMLINSON, I., TREMBATH, R. C.,

ROYLANCE, R., SIMPSON, M. A. & SAWYER, E. J. 2014. Germline CDH1 mutations in bilateral lobular carcinoma in situ. *Br J Cancer,* 110**,** 1053-7.
PLIARCHOPOULOU, K., KALOGERAS, K. T., KRONENWETT, R., WIRTZ, R. M., ELEFTHERAKI, A. G., BATISTATOU, A., BOBOS, M., SOUPOS, N., POLYCHRONIDOU, G., GOGAS, H., SAMANTAS, E., CHRISTODOULOU, C., MAKATSORIS, T., PAVLIDIS, N., PECTASIDES, D. & FOUNTZILAS, G. 2013. Prognostic significance of RACGAP1 mRNA expression in high-risk early breast cancer: a study in primary tumors of breast cancer patients participating in a randomized Hellenic Cooperative Oncology Group trial. *Cancer Chemother Pharmacol,* 71**,** 245-55.
PORTER, G. J., EVANS, A. J., PINDER, S. E., JAMES, J. J., CORNFORD, E. C., BURRELL, H. C., CHAN, S. Y., CHEUNG, K. L. & ROBERTSON, J. F. 2004. Patterns of metastatic breast carcinoma: influence of tumour histological grade. *Clin Radiol,* 59**,** 1094-8.
POSTE G., F. I. J. 1979. The pathogenesis of cancer metastasis. *Nature,* 283**,** 139-146.
PRAT, A. & PEROU, C. M. 2011. Deconstructing the molecular portraits of breast cancer. *Mol Oncol,* 5**,** 5-23.
PRICE, L. S. & COLLARD, J. G. 2001. Regulation of the cytoskeleton by Rho-family GTPases: implications for tumour cell invasion. *Semin Cancer Biol,* 11**,** 167-73.
RAKHA, E. A., EL-SAYED, M. E., GREEN, A. R., LEE, A. H., ROBERTSON, J. F. & ELLIS, I. O. 2007. Prognostic markers in triple-negative breast cancer. *Cancer,* 109**,** 25-32.
RAKHA, E. A., EL-SAYED, M. E., POWE, D. G., GREEN, A. R., HABASHY, H., GRAINGE, M. J., ROBERTSON, J. F., BLAMEY, R., GEE, J., NICHOLSON, R. I., LEE, A. H. & ELLIS, I. O. 2008. Invasive lobular carcinoma of the breast: response to hormonal therapy and outcomes. *European journal of cancer,* 44**,** 73-83.
RAKHA, E. A. & ELLIS, I. O. 2010. Lobular breast carcinoma and its variants. *Semin Diagn Pathol,* 27**,** 49-61.
RAKHA, E. A., GILL, M. S., EL-SAYED, M. E., KHAN, M. M., HODI, Z., BLAMEY, R. W., EVANS, A. J., LEE, A. H. & ELLIS, I. O. 2009. The biological and clinical characteristics of breast carcinoma with mixed ductal and lobular morphology. *Breast cancer research and treatment,* 114**,** 243-50.
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. 2010a. Breast cancer prognostic classification in the molecular era: the role of histological grade. *Breast Cancer Res,* 12**,** 207.
RAKHA, E. A., REIS-FILHO, J. S. & ELLIS, I. O. 2010b. Combinatorial biomarker expression in breast cancer. *Breast Cancer Res Treat,* 120**,** 293-308.
RATHEESH, A., GOMEZ, G. A., PRIYA, R., VERMA, S., KOVACS, E. M., JIANG, K., BROWN, N. H., AKHMANOVA, A., STEHBENS, S. J. & YAP, A. S. 2012. Centralspindlin and alpha-catenin regulate Rho signalling at the epithelial zonula adherens. *Nat Cell Biol,* 14**,** 818-28.
RATHEESH, A. & YAP, A. S. 2012. A bigger picture: classical cadherins and the dynamic actin cytoskeleton. *Nat Rev Mol Cell Biol,* 13**,** 673-9.
REEVES, G. K., BERAL, V., GREEN, J., GATHANI, T., BULL, D. & MILLION WOMEN STUDY, C. 2006. Hormonal therapy for menopause and breast-cancer risk by histological type: a cohort study and meta-analysis. *Lancet Oncol,* 7**,** 910-8.
REIS-FILHO, J. S. & PUSZTAI, L. 2011. Gene expression profiling in breast cancer: classification, prognostication, and prediction. *Lancet,* 378**,** 1812-23.
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,* 12**,** 6652-62.
REYNOLDS, A. B., DANIEL, J., MCCREA, P. D., WHEELOCK, M. J., WU, J. & ZHANG, Z. 1994. Identification of a new catenin: the tyrosine kinase substrate p120cas associates with E-cadherin complexes. *Mol Cell Biol,* 14**,** 8333-42.
ROBINSON, D. R., WU, Y. M., VATS, P., SU, F., LONIGRO, R. J., CAO, X., KALYANA-SUNDARAM, S., WANG, R., NING, Y., HODGES, L., GURSKY, A., SIDDIQUI, J., TOMLINS, S. A., ROYCHOWDHURY, S., PIENTA, K. J., KIM, S. Y., ROBERTS, J. S., RAE, J. M., VAN POZNAK, C. H., HAYES, D. F., CHUGH, R., KUNJU, L. P., TALPAZ, M., SCHOTT, A. F. & CHINNAIYAN, A. M. 2013a. Activating ESR1 mutations in hormone-resistant metastatic breast cancer. *Nat Genet,* 45**,** 1446-51.
ROBINSON, J. L., HOLMES, K. A. & CARROLL, J. S. 2013b. FOXA1 mutations in hormone-dependent cancers. *Front Oncol,* 3**,** 20.
ROBINSON, J. L., MACARTHUR, S., ROSS-INNES, C. S., TILLEY, W. D., NEAL, D. E., MILLS, I. G. & CARROLL, J. S. 2011. Androgen receptor driven transcription in molecular apocrine breast cancer is mediated by FoxA1. *EMBO J,* 30**,** 3019-27.
ROHATGI, R., MA, L., MIKI, H., LOPEZ, M., KIRCHHAUSEN, T., TAKENAWA, T. & KIRSCHNER, M. W. 1999. The interaction between N-WASP and the Arp2/3 complex links Cdc42-dependent signals to actin assembly. *Cell,* 97**,** 221-31.
ROODI, N., BAILEY, L. R., KAO, W. Y., VERRIER, C. S., YEE, C. J., DUPONT, W. D. & PARL, F. F. 1995. Estrogen receptor gene analysis in estrogen receptor-positive and receptor-negative primary breast cancer. *J Natl Cancer Inst,* 87**,** 446-51.
ROSS, J. S., WANG, K., SHEEHAN, C. E., BOGUNIEWICZ, A. B., OTTO, G., DOWNING, S. R., SUN, J., HE, J., CURRAN, J. A., ALI, S., YELENSKY, R., LIPSON, D., PALMER, G., MILLER, V. A. & STEPHENS, P. J. 2013. Relapsed classic E-cadherin (CDH1)-mutated invasive lobular breast cancer shows a high frequency of HER2 (ERBB2) gene mutations. *Clin Cancer Res,* 19**,** 2668-76.
ROSS-INNES, C. S., STARK, R., TESCHENDORFF, A. E., HOLMES, K. A., ALI, H. R., DUNNING, M. J., BROWN, G. D., GOJIS, O., ELLIS, I. O., GREEN, A. R., ALI, S., CHIN, S. F., PALMIERI, C., CALDAS, C. & CARROLL, J. S. 2012. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature,* 481**,** 389-93.
ROYLANCE, R., DROUFAKOU, S., GORMAN, P., GILLETT, C., HART, I. R., HANBY, A. & TOMLINSON, I. 2003. The role of E-cadherin in low-grade ductal breast tumourigenesis. *J Pathol,* 200**,** 53-8.
ROYLANCE, R., GORMAN, P., HARRIS, W., LIEBMANN, R., BARNES, D., HANBY, A. & SHEER, D. 1999. Comparative genomic hybridization of breast tumors stratified by histological grade reveals new insights into the biological progression of breast cancer. *Cancer Res,* 59**,** 1433-6.
SAHU, B., LAAKSO, M., OVASKA, K., MIRTTI, T., LUNDIN, J., RANNIKKO, A., SANKILA, A., TURUNEN, J. P., LUNDIN, M., KONSTI, J., VESTERINEN, T., NORDLING, S., KALLIONIEMI, O., HAUTANIEMI, S. & JANNE, O. A. 2011. Dual role of FoxA1 in androgen receptor binding to chromatin, androgen signalling and prostate cancer. *EMBO J,* 30**,** 3962-76.
SAIGUSA, S., TANAKA, K., MOHRI, Y., OHI, M., SHIMURA, T., KITAJIMA, T., KONDO, S., OKUGAWA, Y., TOIYAMA, Y., INOUE, Y. & KUSUNOKI, M. 2015. Clinical significance of RacGAP1 expression at the invasive front of gastric cancer. *Gastric Cancer,* 18**,** 84-92.
SAITO, S., LIU, X. F., KAMIJO, K., RAZIUDDIN, R., TATSUMOTO, T., OKAMOTO, I., CHEN, X., LEE, C. C., LORENZI, M. V., OHARA, N. & MIKI, T. 2004. Deregulation

and mislocalization of the cytokinesis regulator ECT2 activate the Rho signaling pathways leading to malignant transformation. *J Biol Chem,* 279**,** 7169-79.
SALAMALEKIS, E., BAKAS, P., SYKIOTIS, K., SALOUM, I., KONTOGIANNI, K., PABAMETO, E., PAFITI, A. & CREATSAS, G. 2004. Outcome of patients with ovarian metastatic tumors. Report of 83 cases and review. *European journal of gynaecological oncology,* 25**,** 713-5.
SALAWU, A., UL-HASSAN, A., HAMMOND, D., FERNANDO, M., REED, M. & SISLEY, K. 2012. High quality genomic copy number data from archival formalin-fixed paraffin-embedded leiomyosarcoma: optimisation of universal linkage system labelling. *PLoS One,* 7**,** e50415.
SALHIA, B., TRAN, N. L., CHAN, A., WOLF, A., NAKADA, M., RUTKA, F., ENNIS, M., MCDONOUGH, W. S., BERENS, M. E., SYMONS, M. & RUTKA, J. T. 2008. The guanine nucleotide exchange factors trio, Ect2, and Vav3 mediate the invasive behavior of glioblastoma. *Am J Pathol,* 173**,** 1828-38.
SAMBUY, Y., DE ANGELIS, I., RANALDI, G., SCARINO, M. L., STAMMATI, A. & ZUCCO, F. 2005. The Caco-2 cell line as a model of the intestinal barrier: influence of cell and culture-related factors on Caco-2 cell functional characteristics. *Cell Biol Toxicol,* 21**,** 1-26.
SANO, M., GENKAI, N., YAJIMA, N., TSUCHIYA, N., HOMMA, J., TANAKA, R., MIKI, T. & YAMANAKA, R. 2006. Expression level of ECT2 proto-oncogene correlates with prognosis in glioma patients. *Oncol Rep,* 16**,** 1093-8.
SARRIO, D., MORENO-BUENO, G., HARDISSON, D., SANCHEZ-ESTEVEZ, C., GUO, M., HERMAN, J. G., GAMALLO, C., ESTELLER, M. & PALACIOS, J. 2003. Epigenetic and genetic alterations of APC and CDH1 genes in lobular breast cancer: relationships with abnormal E-cadherin and catenin expression and microsatellite instability. *International journal of cancer. Journal international du cancer,* 106**,** 208-15.
SARRIO, D., RODRIGUEZ-PINILLA, S. M., HARDISSON, D., CANO, A., MORENO-BUENO, G. & PALACIOS, J. 2008. Epithelial-mesenchymal transition in breast cancer relates to the basal-like phenotype. *Cancer Res,* 68**,** 989-97.
SASTRE-GARAU, X., JOUVE, M., ASSELAIN, B., VINCENT-SALOMON, A., BEUZEBOC, P., DORVAL, T., DURAND, J. C., FOURQUET, A. & POUILLART, P. 1996. Infiltrating lobular carcinoma of the breast. Clinicopathologic analysis of 975 cases with reference to data on conservative therapy and metastatic patterns. *Cancer,* 77**,** 113-20.
SAUNUS, J. M., QUINN, M. C., PATCH, A. M., PEARSON, J. V., BAILEY, P. J., NONES, K., MCCART REED, A. E., MILLER, D., WILSON, P. J., AL-EJEH, F., MARIASEGARAM, M., LAU, Q., WITHERS, T., JEFFREE, R. L., REID, L. E., SILVA, L. D., MATSIKA, A., NILAND, C. M., CUMMINGS, M. C., BRUXNER, T. J., CHRIST, A. N., HARLIWONG, I., IDRISOGLU, S., MANNING, S., NOURSE, C., NOURBAKHSH, E., WANI, S., ANDERSON, M. J., LYNN FINK, J., HOLMES, O., KAZAKOFF, S., LEONARD, C., NEWELL, F., TAYLOR, D., WADDELL, N., WOOD, S., XU, Q., KASSAHN, K. S., NARAYANAN, V., TAIB, N. A., TEO, S. H., CHOW, Y. P., KCONFAB, JAT, P. S., BRANDNER, S., FLANAGAN, A. M., KHANNA, K., CHENEVIX-TRENCH, G., GRIMMOND, S. M., SIMPSON, P. T., WADDELL, N. & LAKHANI, S. R. 2015. Integrated genomic and transcriptomic analysis of human brain metastases identifies alterations of potential clinical significance. *J Pathol*.
SCHACKMANN, R. C., VAN AMERSFOORT, M., HAARHUIS, J. H., VLUG, E. J., HALIM, V. A., ROODHART, J. M., VERMAAT, J. S., VOEST, E. E., VAN DER GROEP, P., VAN DIEST, P. J., JONKERS, J. & DERKSEN, P. W. 2011. Cytosolic p120-catenin regulates growth of metastatic lobular carcinoma through Rock1-mediated anoikis resistance. *J Clin Invest,* 121**,** 3176-88.

SCHMIDT-KITTLER, O., RAGG, T., DASKALAKIS, A., GRANZOW, M., AHR, A., BLANKENSTEIN, T. J., KAUFMANN, M., DIEBOLD, J., ARNHOLDT, H., MULLER, P., BISCHOFF, J., HARICH, D., SCHLIMOK, G., RIETHMULLER, G., EILS, R. & KLEIN, C. A. 2003. From latent disseminated cells to overt metastasis: genetic analysis of systemic breast cancer progression. *Proc Natl Acad Sci U S A,* 100**,** 7737-42.

SCHMITZ, A. A., GOVEK, E. E., BOTTNER, B. & VAN AELST, L. 2000. Rho GTPases: signaling, migration, and invasion. *Exp Cell Res,* 261**,** 1-12.

SCHRADER, K. A., MASCIARI, S., BOYD, N., SALAMANCA, C., SENZ, J., SAUNDERS, D. N., YORIDA, E., MAINES-BANDIERA, S., KAURAH, P., TUNG, N., ROBSON, M. E., RYAN, P. D., OLOPADE, O. I., DOMCHEK, S. M., FORD, J., ISAACS, C., BROWN, P., BALMANA, J., RAZZAK, A. R., MIRON, P., COFFEY, K., TERRY, M. B., JOHN, E. M., ANDRULIS, I. L., KNIGHT, J. A., O'MALLEY, F. P., DALY, M., BENDER, P., KCONFAB, MOORE, R., SOUTHEY, M. C., HOPPER, J. L., GARBER, J. E. & HUNTSMAN, D. G. 2011. Germline mutations in CDH1 are infrequent in women with early-onset or familial lobular breast cancers. *J Med Genet,* 48**,** 64-8.

SCHUURING, E. 1995. The involvement of the chromosome 11q13 region in human malignancies: cyclin D1 and EMS1 are two new candidate oncogenes--a review. *Gene,* 159**,** 83-96.

SCHUURING, E., VAN DAMME, H., SCHUURING-SCHOLTES, E., VERHOEVEN, E., MICHALIDES, R., GEELEN, E., DE BOER, C., BROK, H., VAN BUUREN, V. & KLUIN, P. 1998. Characterization of the EMS1 gene and its product, human Cortactin. *Cell Adhes Commun,* 6**,** 185-209.

SEAL, S., BARFOOT, R., JAYATILAKE, H., SMITH, P., RENWICK, A., BASCOMBE, L., MCGUFFOG, L., EVANS, D. G., ECCLES, D., EASTON, D. F., STRATTON, M. R., RAHMAN, N. & BREAST CANCER SUSCEPTIBILITY, C. 2003. Evaluation of Fanconi Anemia genes in familial breast cancer predisposition. *Cancer Res,* 63**,** 8596-9.

SHAABAN, A. M., GREEN, A. R., KARTHIK, S., ALIZADEH, Y., HUGHES, T. A., HARKINS, L., ELLIS, I. O., ROBERTSON, J. F., PAISH, E. C., SAUNDERS, P. T., GROOME, N. P. & SPEIRS, V. 2008. Nuclear and cytoplasmic expression of ERbeta1, ERbeta2, and ERbeta5 identifies distinct prognostic outcome for breast cancer patients. *Clin Cancer Res,* 14**,** 5228-35.

SHAH, S. P., MORIN, R. D., KHATTRA, J., PRENTICE, L., PUGH, T., BURLEIGH, A., DELANEY, A., GELMON, K., GULIANY, R., SENZ, J., STEIDL, C., HOLT, R. A., JONES, S., SUN, M., LEUNG, G., MOORE, R., SEVERSON, T., TAYLOR, G. A., TESCHENDORFF, A. E., TSE, K., TURASHVILI, G., VARHOL, R., WARREN, R. L., WATSON, P., ZHAO, Y., CALDAS, C., HUNTSMAN, D., HIRST, M., MARRA, M. A. & APARICIO, S. 2009. Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. *Nature,* 461**,** 809-813.

SHAH, S. P., ROTH, A., GOYA, R., OLOUMI, A., HA, G., ZHAO, Y., TURASHVILI, G., DING, J., TSE, K., HAFFARI, G., BASHASHATI, A., PRENTICE, L. M., KHATTRA, J., BURLEIGH, A., YAP, D., BERNARD, V., MCPHERSON, A., SHUMANSKY, K., CRISAN, A., GIULIANY, R., HERAVI-MOUSSAVI, A., ROSNER, J., LAI, D., BIROL, I., VARHOL, R., TAM, A., DHALLA, N., ZENG, T., MA, K., CHAN, S. K., GRIFFITH, M., MORADIAN, A., CHENG, S. W., MORIN, G. B., WATSON, P., GELMON, K., CHIA, S., CHIN, S. F., CURTIS, C., RUEDA, O. M., PHAROAH, P. D., DAMARAJU, S., MACKEY, J., HOON, K., HARKINS, T., TADIGOTLA, V., SIGAROUDINIA, M., GASCARD, P., TLSTY, T., COSTELLO, J. F., MEYER, I. M., EAVES, C. J., WASSERMAN, W. W., JONES, S., HUNTSMAN, D., HIRST, M., CALDAS, C.,

MARRA, M. A. & APARICIO, S. 2012. The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature,* 486**,** 395-9.
SHARIF, S., MORAN, A., HUSON, S. M., IDDENDEN, R., SHENTON, A., HOWARD, E. & EVANS, D. G. 2007. Women with neurofibromatosis 1 are at a moderately increased risk of developing breast cancer and should be considered for early screening. *J Med Genet,* 44**,** 481-4.
SHEHATA, M., TESCHENDORFF, A., SHARP, G., NOVCIC, N., RUSSELL, I. A., AVRIL, S., PRATER, M., EIREW, P., CALDAS, C., WATSON, C. J. & STINGL, J. 2012. Phenotypic and functional characterisation of the luminal cell hierarchy of the mammary gland. *Breast Cancer Res,* 14**,** R134.
SI, W., HUANG, W., ZHENG, Y., YANG, Y., LIU, X., SHAN, L., ZHOU, X., WANG, Y., SU, D., GAO, J., YAN, R., HAN, X., LI, W., HE, L., SHI, L., XUAN, C., LIANG, J., SUN, L., WANG, Y. & SHANG, Y. 2015. Dysfunction of the Reciprocal Feedback Loop between GATA3- and ZEB2-Nucleated Repression Programs Contributes to Breast Cancer Metastasis. *Cancer Cell,* 27**,** 822-36.
SIEGEL, R., NAISHADHAM, D. & JEMAL, A. 2013. Cancer statistics, 2013. *CA Cancer J Clin,* 63**,** 11-30.
SIHTO, H., LUNDIN, J., LUNDIN, M., LEHTIMAKI, T., RISTIMAKI, A., HOLLI, K., SAILAS, L., KATAJA, V., TURPEENNIEMI-HUJANEN, T., ISOLA, J., HEIKKILA, P. & JOENSUU, H. 2011. Breast cancer biological subtypes and protein expression predict for the preferential distant metastasis sites: a nationwide cohort study. *Breast Cancer Res,* 13**,** R87.
SIMPSON, P. T., GALE, T., REIS-FILHO, J. S., JONES, C., PARRY, S., SLOANE, J. P., HANBY, A., PINDER, S. E., LEE, A. H., HUMPHREYS, S., ELLIS, I. O. & LAKHANI, S. R. 2005a. Columnar cell lesions of the breast: the missing link in breast cancer progression? A morphological and molecular analysis. *Am J Surg Pathol,* 29**,** 734-46.
SIMPSON, P. T., REIS-FILHO, J. S., GALE, T. & LAKHANI, S. R. 2005b. Molecular evolution of breast cancer. *J Pathol,* 205**,** 248-54.
SINGHAKOWINTA, A., POTTER, H. G., BUROKER, T. R., SAMAL, B., BROOKS, S. C. & VAITKEVICIUS, V. K. 1976. Estrogen receptor and natural course of breast cancer. *Ann Surg,* 183**,** 84-8.
SINGHI, A. D., CIMINO-MATHEWS, A., JENKINS, R. B., LAN, F., FINK, S. R., NASSAR, H., VANG, R., FETTING, J. H., HICKS, J., SUKUMAR, S., DE MARZO, A. M. & ARGANI, P. 2012. MYC gene amplification is often acquired in lethal distant breast cancer metastases of unamplified primary tumors. *Mod Pathol,* 25**,** 378-87.
SKIRNISDOTTIR, I., GARMO, H. & HOLMBERG, L. 2007. Non-genital tract metastases to the ovaries presented as ovarian tumors in Sweden 1990-2003: occurrence, origin and survival compared to ovarian cancer. *Gynecologic oncology,* 105**,** 166-71.
SMITH, M. L., MARIONI, J. C., HARDCASTLE, T. J. & THORNE, N. P. 2006. snapCGH: Segmentation, Normalization and Processing of aCGH Data Users' Guide.
SMUTNY, M., COX, H. L., LEERBERG, J. M., KOVACS, E. M., CONTI, M. A., FERGUSON, C., HAMILTON, N. A., PARTON, R. G., ADELSTEIN, R. S. & YAP, A. S. 2010. Myosin II isoforms identify distinct functional modules that support integrity of the epithelial zonula adherens. *Nature Cell Biology,* 12**,** 696-U147.
SOO, M. S., ROSEN, E. L., BAKER, J. A., VO, T. T. & BOYD, B. A. 2001. Negative predictive value of sonography with mammography in patients with palpable breast lesions. *AJR Am J Roentgenol,* 177**,** 1167-70.
SORLIE, T., PEROU, C. M., TIBSHIRANI, R., AAS, T., GEISLER, S., JOHNSEN, H., HASTIE, T., EISEN, M. B., VAN DE RIJN, M., JEFFREY, S. S., THORSEN, T., QUIST, H., MATESE, J. C., BROWN, P. O., BOTSTEIN, D., LONNING, P. E. & BORRESEN-DALE, A. L. 2001. Gene expression patterns of breast carcinomas

distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A,* 98**,** 10869-74.
SORLIE, T., TIBSHIRANI, R., PARKER, J., HASTIE, T., MARRON, J. S., NOBEL, A., DENG, S., JOHNSEN, H., PESICH, R., GEISLER, S., DEMETER, J., PEROU, C. M., LONNING, P. E., BROWN, P. O., BORRESEN-DALE, A. L. & BOTSTEIN, D. 2003. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci U S A,* 100**,** 8418-23.
SOSA, M. S., BRAGADO, P. & AGUIRRE-GHISO, J. A. 2014. Mechanisms of disseminated cancer cell dormancy: an awakening field. *Nat Rev Cancer,* 14**,** 611-22.
SRINIVASAN, M., SEDMAK, D. & JEWELL, S. 2002. Effect of fixatives and tissue processing on the content and integrity of nucleic acids. *Am J Pathol,* 161**,** 1961-71.
ST ROMAIN, P., MADAN, R., TAWFIK, O. W., DAMJANOV, I. & FAN, F. 2012. Organotropism and prognostic marker discordance in distant metastases of breast carcinoma: fact or fiction? A clinicopathologic analysis. *Human pathology,* 43**,** 398-404.
STEPHENS, P. J., TARPEY, P. S., DAVIES, H., VAN LOO, P., GREENMAN, C., WEDGE, D. C., NIK-ZAINAL, S., MARTIN, S., VARELA, I., BIGNELL, G. R., YATES, L. R., PAPAEMMANUIL, E., BEARE, D., BUTLER, A., CHEVERTON, A., GAMBLE, J., HINTON, J., JIA, M., JAYAKUMAR, A., JONES, D., LATIMER, C., LAU, K. W., MCLAREN, S., MCBRIDE, D. J., MENZIES, A., MUDIE, L., RAINE, K., RAD, R., CHAPMAN, M. S., TEAGUE, J., EASTON, D., LANGEROD, A., OSLO BREAST CANCER, C., LEE, M. T., SHEN, C. Y., TEE, B. T., HUIMIN, B. W., BROEKS, A., VARGAS, A. C., TURASHVILI, G., MARTENS, J., FATIMA, A., MIRON, P., CHIN, S. F., THOMAS, G., BOYAULT, S., MARIANI, O., LAKHANI, S. R., VAN DE VIJVER, M., VAN 'T VEER, L., FOEKENS, J., DESMEDT, C., SOTIRIOU, C., TUTT, A., CALDAS, C., REIS-FILHO, J. S., APARICIO, S. A., SALOMON, A. V., BORRESEN-DALE, A. L., RICHARDSON, A. L., CAMPBELL, P. J., FUTREAL, P. A. & STRATTON, M. R. 2012. The landscape of cancer genes and mutational processes in breast cancer. *Nature,* 486**,** 400-4.
STRATTON, M. R., CAMPBELL, P. J. & FUTREAL, P. A. 2009. The cancer genome. *Nature,* 458**,** 719-24.
SURGET, S., KHOURY, M. P. & BOURDON, J. C. 2013. Uncovering the role of p53 splice variants in human malignancy: a clinical perspective. *Onco Targets Ther,* 7**,** 57-68.
SURYADEVARA, A., PARUCHURI, L. P., BANISAEED, N., DUNNINGTON, G. & RAO, K. A. 2010. The clinical behavior of mixed ductal/lobular carcinoma of the breast: a clinicopathologic analysis. *World journal of surgical oncology,* 8**,** 51.
TAMBORERO, D., GONZALEZ-PEREZ, A. & LOPEZ-BIGAS, N. 2013. OncodriveCLUST: exploiting the positional clustering of somatic mutations to identify cancer genes. *Bioinformatics*, 29, 2238-44.
TANG, H., LI, A., BI, J., VELTMAN, D. M., ZECH, T., SPENCE, H. J., YU, X., TIMPSON, P., INSALL, R. H., FRAME, M. C. & MACHESKY, L. M. 2013. Loss of Scar/WAVE complex promotes N-WASP- and FAK-dependent invasion. *Curr Biol,* 23**,** 107-17.
TATSUMOTO, T., XIE, X., BLUMENTHAL, R., OKAMOTO, I. & MIKI, T. 1999. Human ECT2 is an exchange factor for Rho GTPases, phosphorylated in G2/M phases, and involved in cytokinesis. *J Cell Biol,* 147**,** 921-8.
THE CANCER GENOME ATLAS NETWORK 2012. Comprehensive molecular portraits of human breast tumours. *Nature,* 490**,** 61-70.
THEODOROU, V., STARK, R., MENON, S. & CARROLL, J. S. 2013. GATA3 acts upstream of FOXA1 in mediating ESR1 binding by shaping enhancer accessibility. *Genome Res,* 23**,** 12-22.

THIERY, J. P., ACLOQUE, H., HUANG, R. Y. & NIETO, M. A. 2009. Epithelial-mesenchymal transitions in development and disease. *Cell,* 139**,** 871-90.
THOMAS, C. & GUSTAFSSON, J. A. 2011. The different roles of ER subtypes in cancer biology and therapy. *Nat Rev Cancer,* 11**,** 597-608.
THOMAS, C., RAJAPAKSA, G., NIKOLOS, F., HAO, R., KATCHY, A., MCCOLLUM, C. W., BONDESSON, M., QUINLAN, P., THOMPSON, A., KRISHNAMURTHY, S., ESTEVA, F. J. & GUSTAFSSON, J. A. 2012. ERbeta1 represses basal breast cancer epithelial to mesenchymal transition by destabilizing EGFR. *Breast Cancer Res,* 14**,** R148.
THULLBERG, M., GAD, A., LE GUYADER, S. & STROMBLAD, S. 2007. Oncogenic H-Ras V12 promotes anchorage-independent cytokinesis in human fibroblasts. *Proc Natl Acad Sci U S A,* 104**,** 20338-43.
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 Res Treat,* 102**,** 143-55.
TOT, T. 2005. DCIS, cytokeratins, and the theory of the sick lobe. *Virchows Arch,* 447**,** 1-8.
TOY, W., SHEN, Y., WON, H., GREEN, B., SAKR, R. A., WILL, M., LI, Z., GALA, K., FANNING, S., KING, T. A., HUDIS, C., CHEN, D., TARAN, T., HORTOBAGYI, G., GREENE, G., BERGER, M., BASELGA, J. & CHANDARLAPATY, S. 2013. ESR1 ligand-binding domain mutations in hormone-resistant breast cancer. *Nat Genet,* 45**,** 1439-45.
TRIPATHI, A., KING, C., DE LA MORENAS, A., PERRY, V. K., BURKE, B., ANTOINE, G. A., HIRSCH, E. F., KAVANAH, M., MENDEZ, J., STONE, M., GERRY, N. P., LENBURG, M. E. & ROSENBERG, C. L. 2008. Gene expression abnormalities in histologically normal breast epithelium of breast cancer patients. *Int J Cancer,* 122**,** 1557-66.
TROXELL, M. L., BRUNNER, A. L., NEFF, T., WARRICK, A., BEADLING, C., MONTGOMERY, K., ZHU, S., CORLESS, C. L. & WEST, R. B. 2012. Phosphatidylinositol-3-kinase pathway mutations are common in breast columnar cell lesions. *Mod Pathol,* 25**,** 930-7.
URUNO, T., LIU, J., ZHANG, P., FAN, Y., EGILE, C., LI, R., MUELLER, S. C. & ZHAN, X. 2001. Activation of Arp2/3 complex-mediated actin polymerization by cortactin. *Nat Cell Biol,* 3**,** 259-66.
USARY, J., LLACA, V., KARACA, G., PRESSWALA, S., KARACA, M., HE, X., LANGEROD, A., KARESEN, R., OH, D. S., DRESSLER, L. G., LONNING, P. E., STRAUSBERG, R. L., CHANOCK, S., BORRESEN-DALE, A. L. & PEROU, C. M. 2004. Mutation of GATA3 in human breast tumors. *Oncogene,* 23**,** 7669-78.
VAN BEERS, E. H., JOOSSE, S. A., LIGTENBERG, M. J., FLES, R., HOGERVORST, F. B., VERHOEF, S. & NEDERLOF, P. M. 2006. A multiplex PCR predictor for aCGH success of FFPE samples. *Br J Cancer,* 94**,** 333-7.
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. *N Engl J Med,* 347**,** 1999-2009.
VAN POZNAK, C., SOMERFIELD, M. R., BAST, R. C., CRISTOFANILLI, M., GOETZ, M. P., GONZALEZ-ANGULO, A. M., HICKS, D. G., HILL, E. G., LIU, M. C., LUCAS, W., MAYER, I. A., MENNEL, R. G., SYMMANS, W. F., HAYES, D. F. & HARRIS, L. N. 2015. Use of Biomarkers to Guide Decisions on Systemic Therapy for Women

With Metastatic Breast Cancer: American Society of Clinical Oncology Clinical Practice Guideline. *J Clin Oncol*.
VANNIER, C., MOCK, K., BRABLETZ, T. & DRIEVER, W. 2013. Zeb1 regulates E-cadherin and Epcam (epithelial cell adhesion molecule) expression to control cell behavior in early zebrafish development. *J Biol Chem,* 288**,** 18643-59.
VARGAS, A. C., MCCART REED, A. E., WADDELL, N., LANE, A., REID, L. E., SMART, C. E., COCCIARDI, S., DA SILVA, L., SONG, S., CHENEVIX-TRENCH, G., SIMPSON, P. T. & LAKHANI, S. R. 2012. Gene expression profiling of tumour epithelial and stromal compartments during breast cancer progression. *Breast Cancer Res Treat,* 135**,** 153-65.
VEERARAGHAVAN, J., TAN, Y., CAO, X. X., KIM, J. A., WANG, X., CHAMNESS, G. C., MAITI, S. N., COOPER, L. J., EDWARDS, D. P., CONTRERAS, A., HILSENBECK, S. G., CHANG, E. C., SCHIFF, R. & WANG, X. S. 2014. Recurrent ESR1-CCDC170 rearrangements in an aggressive subset of oestrogen receptor-positive breast cancers. *Nat Commun,* 5**,** 4577.
VLEMINCKX, K., VAKAET, L., JR., MAREEL, M., FIERS, W. & VAN ROY, F. 1991. Genetic manipulation of E-cadherin expression by epithelial tumor cells reveals an invasion suppressor role. *Cell,* 66**,** 107-19.
VODUC, D., CHEANG, M. & NIELSEN, T. 2008. GATA-3 expression in breast cancer has a strong association with estrogen receptor but lacks independent prognostic value. *Cancer Epidemiol Biomarkers Prev,* 17**,** 365-73.
VOS, C. B., CLETON-JANSEN, A. M., BERX, G., DE LEEUW, W. J., TER HAAR, N. T., VAN ROY, F., CORNELISSE, C. J., PETERSE, J. L. & VAN DE VIJVER, M. J. 1997. E-cadherin inactivation in lobular carcinoma in situ of the breast: an early event in tumorigenesis. *British journal of cancer,* 76**,** 1131-3.
VUONG, D., SIMPSON, P. T., GREEN, B., CUMMINGS, M. C. & LAKHANI, S. R. 2014. Molecular classification of breast cancer. *Virchows Arch,* 465**,** 1-14.
WAGNER, P. L., KITABAYASHI, N., CHEN, Y. T. & SHIN, S. J. 2009. Clonal relationship between closely approximated low-grade ductal and lobular lesions in the breast: a molecular study of 10 cases. *American journal of clinical pathology,* 132**,** 871-6.
WALT, A. J., SINGHAKOWINTA, A., BROOKS, S. C. & CORTEZ, A. 1976. The surgical implications of estrophile protein estimations in carcinoma of the breast. *Surgery,* 80**,** 506-12.
WANG, D., GARCIA-BASSETS, I., BENNER, C., LI, W., SU, X., ZHOU, Y., QIU, J., LIU, W., KAIKKONEN, M. U., OHGI, K. A., GLASS, C. K., ROSENFELD, M. G. & FU, X. D. 2011a. Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. *Nature,* 474**,** 390-4.
WANG, S. M., OOI, L. L. & HUI, K. M. 2011b. Upregulation of Rac GTPase-activating protein 1 is significantly associated with the early recurrence of human hepatocellular carcinoma. *Clin Cancer Res,* 17**,** 6040-51.
WEIGELT, B., GLAS, A. M., WESSELS, L. F., WITTEVEEN, A. T., PETERSE, J. L. & VAN'T VEER, L. J. 2003. Gene expression profiles of primary breast tumors maintained in distant metastases. *Proc Natl Acad Sci U S A,* 100**,** 15901-5.
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. *J Pathol,* 216**,** 141-50.
WEIGELT, B., MACKAY, A., A'HERN, R., NATRAJAN, R., TAN, D. S., DOWSETT, M., ASHWORTH, A. & REIS-FILHO, J. S. 2010. Breast cancer molecular profiling with single sample predictors: a retrospective analysis. *Lancet Oncol,* 11**,** 339-49.

WEIGELT, B. & REIS-FILHO, J. S. 2010. Molecular profiling currently offers no more than tumour morphology and basic immunohistochemistry. *Breast Cancer Res,* 12 Suppl 4**,** S5.
WEST, M., BLANCHETTE, C., DRESSMAN, H., HUANG, E., ISHIDA, S., SPANG, R., ZUZAN, H., OLSON, J. A., JR., MARKS, J. R. & NEVINS, J. R. 2001. Predicting the clinical status of human breast cancer by using gene expression profiles. *Proc Natl Acad Sci U S A,* 98**,** 11462-7.
WICKHAM, H. 2011. The Split-Apply-Combine Strategy for Data Analysis. *Journal of Statistical Software,* 40**,** 1-29.
WINCHESTER, D. J., CHANG, H. R., GRAVES, T. A., MENCK, H. R., BLAND, K. I. & WINCHESTER, D. P. 1998. A comparative analysis of lobular and ductal carcinoma of the breast: presentation, treatment, and outcomes. *Journal of the American College of Surgeons,* 186**,** 416-22.
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, O. & COLLEGE OF AMERICAN, P. 2007. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *J Clin Oncol,* 25**,** 118-45.
WONG, T., GAO, W. & CHAN, J. 2014. Transcription Regulation of E-Cadherin by Zinc Finger E-Box Binding Homeobox Proteins in Solid Tumors,. *BioMed Research International,* 2014.
WU, J. M., FACKLER, M. J., HALUSHKA, M. K., MOLAVI, D. W., TAYLOR, M. E., TEO, W. W., GRIFFIN, C., FETTING, J., DAVIDSON, N. E., DE MARZO, A. M., HICKS, J. L., CHITALE, D., LADANYI, M., SUKUMAR, S. & ARGANI, P. 2008. Heterogeneity of breast cancer metastases: comparison of therapeutic target expression and promoter methylation between primary tumors and their multifocal metastases. *Clinical cancer research : an official journal of the American Association for Cancer Research,* 14**,** 1938-46.
WU, J. N. & ROBERTS, C. W. 2013. ARID1A mutations in cancer: another epigenetic tumor suppressor? *Cancer Discov,* 3**,** 35-43.
WU, S. K., GOMEZ, G. A., MICHAEL, M., VERMA, S., COX, H. L., LEFEVRE, J. G., PARTON, R. G., HAMILTON, N. A., NEUFELD, Z. & YAP, A. S. 2014. Cortical F-actin stabilization generates apical-lateral patterns of junctional contractility that integrate cells into epithelia. *Nat Cell Biol,* 16**,** 167-78.
YACHIDA, S., JONES, S., BOZIC, I., ANTAL, T., LEARY, R., FU, B., KAMIYAMA, M., HRUBAN, R. H., ESHLEMAN, J. R., NOWAK, M. A., VELCULESCU, V. E., KINZLER, K. W., VOGELSTEIN, B. & IACOBUZIO-DONAHUE, C. A. 2010. Distant metastasis occurs late during the genetic evolution of pancreatic cancer. *Nature,* 467**,** 1114-7.
YADA-HASHIMOTO, N., YAMAMOTO, T., KAMIURA, S., SEINO, H., OHIRA, H., SAWAI, K., KIMURA, T. & SAJI, F. 2003. Metastatic ovarian tumors: a review of 64 cases. *Gynecologic oncology,* 89**,** 314-7.
YATES, L. R., GERSTUNG, M., KNAPPSKOG, S., DESMEDT, C., GUNDEM, G., VAN LOO, P., AAS, T., ALEXANDROV, L. B., LARSIMONT, D., DAVIES, H., LI, Y., JU, Y. S., RAMAKRISHNA, M., HAUGLAND, H. K., LILLENG, P. K., NIK-ZAINAL, S., MCLAREN, S., BUTLER, A., MARTIN, S., GLODZIK, D., MENZIES, A., RAINE, K., HINTON, J., JONES, D., MUDIE, L. J., JIANG, B., VINCENT, D., GREENE-COLOZZI, A., ADNET, P. Y., FATIMA, A., MAETENS, M., IGNATIADIS, M.,

STRATTON, M. R., SOTIRIOU, C., RICHARDSON, A. L., LONNING, P. E., WEDGE, D. C. & CAMPBELL, P. J. 2015. Subclonal diversification of primary breast cancer revealed by multiregion sequencing. *Nat Med,* 21**,** 751-9.

YOON, N. K., MARESH, E. L., SHEN, D., ELSHIMALI, Y., APPLE, S., HORVATH, S., MAH, V., BOSE, S., CHIA, D., CHANG, H. R. & GOODGLICK, L. 2010. Higher levels of GATA3 predict better survival in women with breast cancer. *Hum Pathol,* 41**,** 1794-801.

YOUNG, R. H. & SCULLY, R. E. 1991. Metastatic tumors in the ovary: a problem-oriented approach and review of the recent literature. *Seminars in diagnostic pathology,* 8**,** 250-76.

YU, X., ZECH, T., MCDONALD, L., GONZALEZ, E. G., LI, A., MACPHERSON, I., SCHWARZ, J. P., SPENCE, H., FUTO, K., TIMPSON, P., NIXON, C., MA, Y., ANTON, I. M., VISEGRADY, B., INSALL, R. H., OIEN, K., BLYTH, K., NORMAN, J. C. & MACHESKY, L. M. 2012. N-WASP coordinates the delivery and F-actin-mediated capture of MT1-MMP at invasive pseudopods. *J Cell Biol,* 199**,** 527-44.

YUAN, T. L. & CANTLEY, L. C. 2008. PI3K pathway alterations in cancer: variations on a theme. *Oncogene,* 27**,** 5497-510.

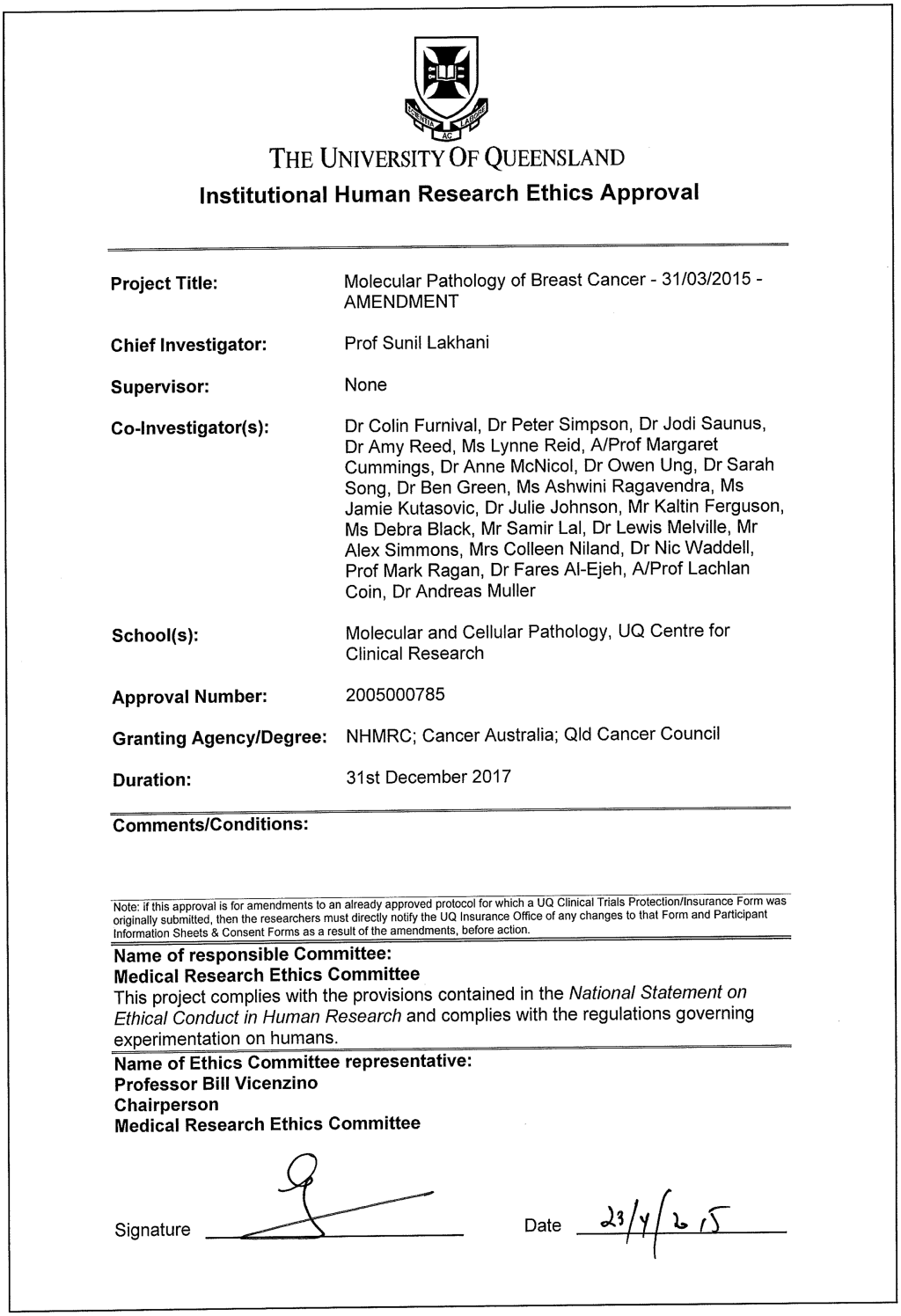
ZARDAVAS, D., MAETENS, M., IRRTHUM, A., GOULIOTI, T., ENGELEN, K., FUMAGALLI, D., SALGADO, R., AFTIMOS, P., SAINI, K. S., SOTIRIOU, C., CAMPBELL, P., DINH, P., VON MINCKWITZ, G., GELBER, R. D., DOWSETT, M., DI LEO, A., CAMERON, D., BASELGA, J., GNANT, M., GOLDHIRSCH, A., NORTON, L. & PICCART, M. 2014. The AURORA initiative for metastatic breast cancer. *Br J Cancer,* 111**,** 1881-7.

ZHAO, W. M. & FANG, G. 2005. MgcRacGAP controls the assembly of the contractile ring and the initiation of cytokinesis. *Proc Natl Acad Sci U S A,* 102**,** 13158-63.

ZOU, D., YOON, H. S., PEREZ, D., WEEKS, R. J., GUILFORD, P. & HUMAR, B. 2009. Epigenetic silencing in non-neoplastic epithelia identifies E-cadherin (CDH1) as a target for chemoprevention of lobular neoplasia. *The Journal of pathology,* 218**,** 265-72.

ZWIJSEN, R. M., BUCKLE, R. S., HIJMANS, E. M., LOOMANS, C. J. & BERNARDS, R. 1998. Ligand-independent recruitment of steroid receptor coactivators to estrogen receptor by cyclin D1. *Genes Dev,* 12**,** 3488-98.

Appendices



THE UNIVERSITY OF QUEENSLAND

Institutional Human Research Ethics Approval

Project Title:	Molecular Pathology of Breast Cancer - 31/03/2015 - AMENDMENT
Chief Investigator:	Prof Sunil Lakhani
Supervisor:	None
Co-Investigator(s):	Dr Colin Furnival, Dr Peter Simpson, Dr Jodi Saunus, Dr Amy Reed, Ms Lynne Reid, A/Prof Margaret Cummings, Dr Anne McNicol, Dr Owen Ung, Dr Sarah Song, Dr Ben Green, Ms Ashwini Ragavendra, Ms Jamie Kutasovic, Dr Julie Johnson, Mr Kaltin Ferguson, Ms Debra Black, Mr Samir Lal, Dr Lewis Melville, Mr Alex Simmons, Mrs Colleen Niland, Dr Nic Waddell, Prof Mark Ragan, Dr Fares Al-Ejeh, A/Prof Lachlan Coin, Dr Andreas Muller
School(s):	Molecular and Cellular Pathology, UQ Centre for Clinical Research
Approval Number:	2005000785
Granting Agency/Degree:	NHMRC; Cancer Australia; Qld Cancer Council
Duration:	31st December 2017

Comments/Conditions:

Note: if this approval is for amendments to an already approved protocol for which a UQ Clinical Trials Protection/Insurance Form was originally submitted, then the researchers must directly notify the UQ Insurance Office of any changes to that Form and Participant Information Sheets & Consent Forms as a result of the amendments, before action.

Name of responsible Committee:
Medical Research Ethics Committee
This project complies with the provisions contained in the *National Statement on Ethical Conduct in Human Research* and complies with the regulations governing experimentation on humans.

Name of Ethics Committee representative:
Professor Bill Vicenzino
Chairperson
Medical Research Ethics Committee

Signature _____ Date 23/4/15

Appendix Figure 2.1: University of Queensland Human Research Ethics Approval form.

GM_TMA_1	1	2	3	4	5	6	7	8	9	10	11	12
1	GM20_ RT TUBE/OVARY	GM20_ - RT TUBE/OVARY	GM20_APPENDIX	GM20_APPENDIX	GM20_ FALLOPIAN TUBE	GM20_ FALLOPIAN TUBE	GM20_ ADHERENT FAT FALLOPIAN TUBE	GM20_ ADHERENT FAT FALLOPIAN TUBE	GM20_LT OVARY	GM20_LT OVARY	GM20_OMENTUM	GM20_OMENTUM
2	GM55_bone EXCLUDED	GM55_bone EXCLUDED	GM55_LARGE BOWEL	GM55_LARGE BOWEL	GM55_ OMENTUM	GM55_ OMENTUM	GM55_PELVIS LT SIDE WALL PERI	GM55_PELVIS LT SIDE WALL PERI	GM55_BREAST	GM55_BREAST	GM55_LN	GM55_LN
3	GM49_RT OVARY	GM49_RT OVARY	GM49_RT TUBE	GM49_RT TUBE	GM49_ LT OVARY	GM49_ LT OVARY	GM49_LT TUBE	GM49_LT TUBE	GM49_OMENTUM	GM49_OMENTUM	GM49_UTERUS	GM49_UTERUS
4	GM42_RT OVARY	GM42_RT OVARY	GM42_LEFT OVARY	GM42_LEFT OVARY	GM42_PERITONEAL REFLECTIONS	GM42_PERITONEAL REFLECTIONS	GM42_ UTERINE WALL	GM42_ UTERINE WALL	GM42_OMENTUM	GM42_OMENTUM		
5	GM51_BREAST	GM51_BREAST	GM51_ DCIS	GM51_ DCIS	GM51_DUODENUM	GM51_DUODENUM	GM51_ STOMACH	GM51_ STOMACH	GM51_OMENTUM	GM51_OMENTUM		
6	GM33_CERVIX	GM33_CERVIX	GM33_ENDOMYOMETRIUM	GM33_ENDOMYOMETRIUM	GM33_LT TUBE	GM33_LT TUBE	GM33_RT TUBE	GM33_RT TUBE	GM33_RT OVARY	GM33_RT OVARY		
7	GM06_LT OVARY	GM06_LT OVARY	GM06_CERVIX	GM06_CERVIX	GM06_ENDOMYOMETRIUM	GM06_ENDOMYOMETRIUM	GM06_RT OVARY	GM06_RT OVARY				
8	GM39_LT OVARY	GM39_LT OVARY	GM39_CERVIX	GM39_CERVIX	GM39_RT OVARY	GM39_RT OVARY	GM39_BREAST	GM39_BREAST				
9	LIVER	Q541 NB		Q541 NB	Q541 NB	Q541 NB						

Appendix Figure 2.2: Gynaecological metastasis TMA 1

GM_TMA_2	1	2	3	4	5	6	7	8	9	10	11	12
1	GM78_OMENTUM	GM78_OMENTUM	GM78_COLON + SURROUNDING FAT	GM78_COLON + SURROUNDING FAT	GM78_ OVARY GROWTH PATTERN 1	GM78_ OVARY GROWTH PATTERN 1	GM78_OVARY GROWTH PATTERN 2	GM78_OVARY GROWTH PATTERN 2				
2	GM80_ENDOMYOMETRIUM	GM80_ENDOMYOMETRIUM	GM80_RT TUBE	GM80_RT TUBE	GM80_ANUS	GM80_ANUS	GM80_SIGMOID	GM80_SIGMOID	GM80_RECTAL WALL	GM80_RECTAL WALL	GM80_LT EXTERNAL ILIAC	GM80_LT EXTERNAL ILIAC
3	GM80_SMALL BOWEL	GM80_SMALL BOWEL										
4	GM53_OMENTAL CAKE	GM53_OMENTAL CAKE	GM53_OMENTUM	GM53_OMENTUM	GM53_LESSER OMENTUM	GM53_LESSER OMENTUM	GM53_RT DIAPHRAGM	GM53_RT DIAPHRAGM	GM53_DIAPHRAGMIC NODE	GM53_DIAPHRAGMIC NODE	GM53_PERITONEUM	GM53_PERITONEUM
5	GM53_BLADDER	GM53_BLADDER	GM53_LT OVARY	GM53_LT OVARY	GM53_COLON	GM53_COLON	GM53_SMALL BOWEL	GM53_SMALL BOWEL	GM53_RENAL/PERITONEAL NODE	GM53_RENAL/PERITONEAL NODE	GM53_PERITONEAL NODE	GM53_PERITONEAL NODE
6	GM53_LIVER	GM53_LIVER	GM53_APPENDIX	GM53_APPENDIX								
7	GM45_RT OVARY	GM45_RT OVARY	GM45_RT TUBE	GM45_RT TUBE	GM45_CERVIX	GM45_CERVIX	GM45_ENDOMETRIUM	GM45_NOT ENOUGH TISSUE	GM45_ LT TUBE	GM45_ LT TUBE	GM45_LT OVARY	GM45_LT OVARY
8	GM45_APPENDIX	GM45_APPENDIX	GM45_RT PELVIC LN	GM45_0RT PELVIC LN	GM45_LT PELVIC LN	GM45_LT PELVIC LN	GM45_ OMENTUM	GM45_ OMENTUM	GM45_OMENTUM LN	GM45_OMENTUM LN		
9	LIVER	LIVER	Q541 NB	Q541 NB	Q541 NB	Q541 NB						

Appendix Figure 2.3: Gynaecological metastasis TMA 2

GM_TMA_3	1	2	3	4	5	6	7	8	9	10	11	12
1	GM16_ OMENTUM	GM16_ OMENTUM	GM16_RT OVARY	GM16_RT OVARY	GM16_ RT TUBE	GM16_ RT TUBE	GM16_LT OVARY	GM16_LT OVARY	GM16_ LT TUBE	GM16_ LT TUBE	GM16_APPENDIX	GM16_APPENDIX
2	GM64_ BREAST	GM64_ BREAST	GM64_RT PELVIC RIM	GM64_RT PELVIC RIM	GM64_RT ROUND LIGAMENT	GM64_RT ROUND LIGAMENT					GM13_RT OVARY	GM13_RT OVARY
3	GM59_ BREAST	GM59_ BREAST	GM59_LN	GM59_LN	GM59_pleura	GM59_pleura					GM07_Lt ovary	GM07_Lt ovary
4	GM04_LT OVARY	GM04_LT OVARY	GM04_LT TUBE	GM04_LT TUBE	GM04_POUCH OF DOUGLAS	GM04_POUCH OF DOUGLAS				LIVER	GM84_VULVA	GM84_VULVA
5	GM85_SMALL BOWEL	GM85_SMALL BOWEL	GM85_OMENTUM	GM85_OMENTUM						LIVER	GM10_ CERVIX	GM10_ CERVIX
6	GM36_CERVIX	GM36_CERVIX	GM36_ ENDOMETRIUM	GM36_ ENDOMETRIUM					GM76_ENDOMETRIUM	GM76_ENDOMETRIUM	GM76_OVARY	GM76_OVARY
7	GM74_LT OVARY GROWTH PATTERN 1	GM74_LT OVARY GROWTH PATTERN 1	GM74_LT OVARY GROWTH PATTERN 2	GM74_LT OVARY GROWTH PATTERN 2			Q541 NB	Q541 NB	GM70_OVARY	GM70_OVARY	GM70_PERITONEUM	GM70_PERITONEUM
8	GM86_OMENTUM	GM86_OMENTUM	GM86_ABDOMINAL SCAR	GM86_ABDOMINAL SCAR			Q541 NB	Q541 NB	GM71_ LT OVARY	GM71_ LT OVARY	GM71_LT TUBE	GM71_LT TUBE

Appendix Figure 2.4: Gynaecological metastasis TMA 3

GM_TMA_4	1	2	3	4	5	6	7	8	9
1	GM63_Breast	GM63_Breast	GM63_LN	GM63_LN	GM63_Omentum	GM63_Omentum	GM63_Rt ovary	GM63_Rt ovary	GM63_LT Ovary
2	GM63_LT Ovary	GM63_Rt Fallopian tube	GM63_Rt Fallopian tube	GM63_Lt fallopian tube	GM63_Lt fallopian tube	GM25_Ovary	GM25_Ovary	GM16_Breast	GM16_Breast
3	GM01_Breast_IDC	GM01_Breast_IDC	GM01_Breast_ ILC	GM01_Ovaries	GM01_Ovaries	GM36_Breast	GM36_Breast	GM65_Breast	GM65_Breast
4	GM15_Fallopian tube	GM15_Fallopian tube	GM15_Rt Ovary	GM15_Rt Ovary	GM15_Lt Ovary	GM15_Lt Ovary	GM72_Breast	GM72_Breast	
5	GM87_Ovaries (A1)	GM87_Ovaries (A1)	GM87_Ovaries (A2)	GM87_Ovaries (A2)	GM87_Ovaries (A3)	GM87_Ovaries (A3)	GM87_Brain	GM87_Liver	
6	GM41_Ovary	GM41_Ovary	LIVER	LIVER	Q541 NB	Q541 NB	Q541 NB	Q541 NB	
7	GM06_cervix	GM06_cervix	GM45_myometrium	GM45_myometrium	GM16_appendix	GM16_appendix	GM53_appendix	GM53_appendix	
8	GM64_breast	GM64_breast	GM64_LN	empty	GM36_uterus	GM36_uterus			

Appendix Figure 2.5: Gynaecological metastasis TMA 4

GM_TMA_5	1	2	3	4	5	6	7	8	9	10	11	12
1	GM06_ Breast	GM06_ Breast	GM12_ Rt ovary	GM12_ Rt ovary	GM12_Lt ovary	GM12_Lt ovary	GM14_Fibroadipose/adnexal tissue - tube	GM14_Fibroadipose/adnexal tissue - tube	GM14_Fibroadipose/adnexal tissue - ovary	GM14_Fibroadipose/adnexal tissue - ovary	GM14_omentum	GM14_omentum
2	GM16_Pleura	GM16_Pleura	GM18_breast	GM18_breast	GM24_Breast	GM24_Breast	GM24_LT Ovary	GM24_LT Ovary	GM29_Breast	GM29_Breast	GM29_Cervix	GM29_Cervix
3	GM34_omentum	GM34_omentum	GM34_ ovary	GM34_ ovary	GM34_Fallopian tube	GM34_Fallopian tube	GM34_sigmoid fat	GM34_sigmoid fat	GM41_Breast	GM41_Breast	GM44_breast	GM44_breast
4	GM50_Lt tube p1	GM50_Lt tube p1	GM50_ Lt ovary p1	GM50_ Lt ovary p1	GM50_Lt ovary p2	GM50_Lt ovary p2	GM50_PIPE_Rt tube p2	GM50_PIPE_Rt tube p2	GM50_PIPE_Rt ovary p1	GM50_PIPE_Rt ovary p1	GM50_Rt ovary p2	GM50_Rt ovary p2
5	GM47_breast	GM47_breast	GM57_breast	GM57_breast	GM57_rt ov p1	GM57_rt ov p1	GM57_rt ov p2	GM57_rt ov p2	GM57_cervix p2	GM57_cervix p2	GM57_cervix p3	GM57_cervix p3
6	**Liver**	**Liver**	**Liver**	**NB**	**NB**	**NB**						

Appendix Figure 2.6: Gynaecological metastasis TMA 5

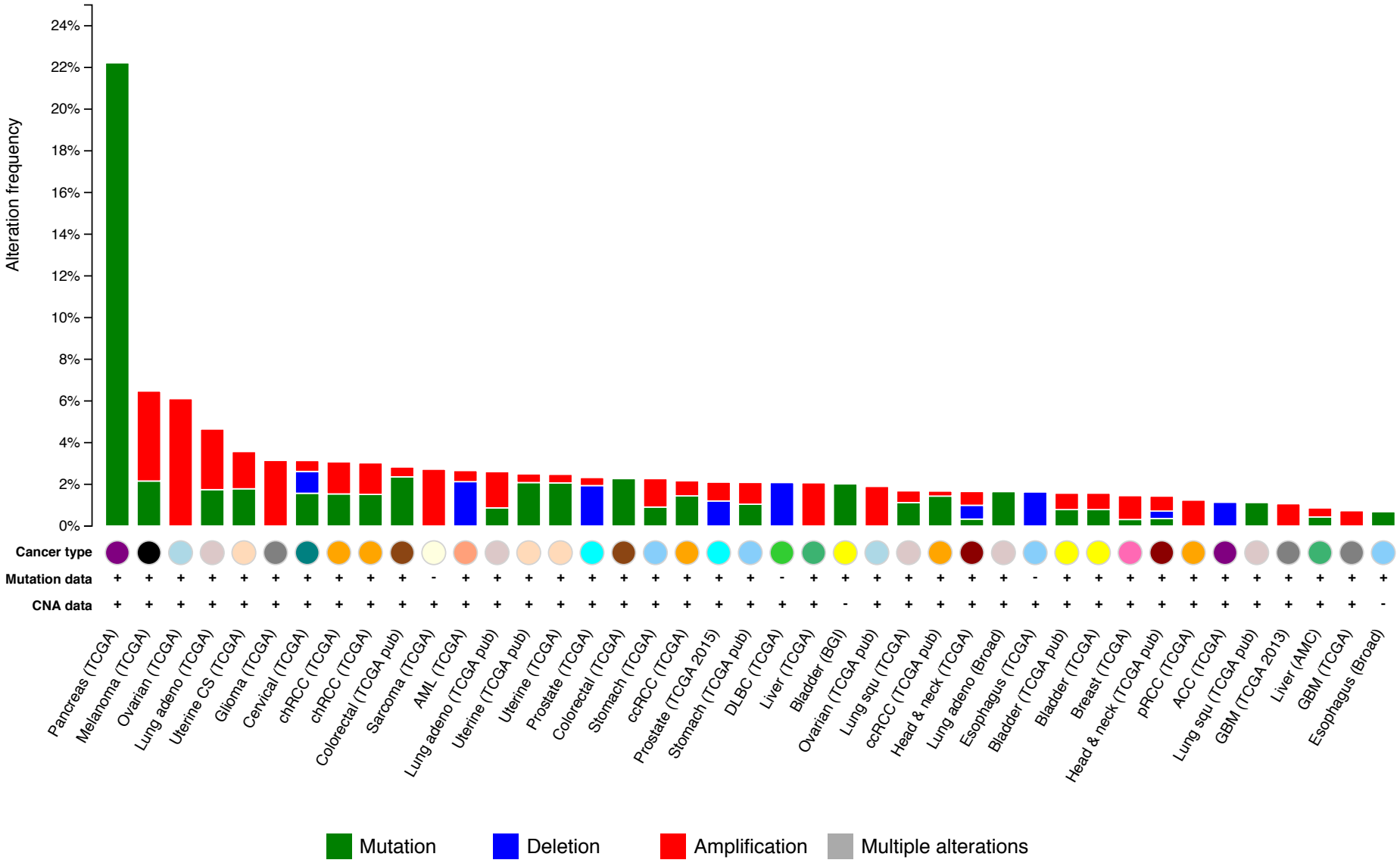
GM_TMA_6	1	2	3	4	5	6	7	8	9	10
1	GM56_ovary	GM56_ovary	GM59_Breast	GM59_Breast	GM73_vagina	GM73_vagina	GM73_cervix	GM73_cervix	GM74_Breast	GM74_Breast
2	GM78_Breast	GM78_Breast	GM77_Rt ovary	GM77_Rt ovary	GM77_Lt ovary	GM77_Lt ovary	GM77_Breast	GM77_Breast	GM88_Endometrium	GM88_Endometrium
3	GM89_Rt Ovary P1	GM89_Rt Ovary P1	GM89_Rt Ovary P2	GM89_Rt Ovary P2	GM82_Pouch of douglas	GM82_Pouch of douglas	GM82_Cervix P1	GM82_Cervix P1	GM82_Vulva P1	GM82_Vulva P1
4	GM82_Vulva P2	GM82_Vulva P2	GM82_Vagina P1	GM82_Vagina P1	GM82_Lt ovary P3	GM82_Lt ovary P3	GM82_Lt ovary P1	GM82_Lt ovary P1	GM59_Rt ovary 1K	GM59_Rt ovary 1K
5	**Liver**	**Liver**	**Liver**	**NB**	**NB**	**NB**	GM82_Rt ovary P1	GM82_Rt ovary P1	GM82_Rt ovary P3	GM82_Rt ovary P3

Appendix Figure 2.7: Gynaecological metastasis TMA 6

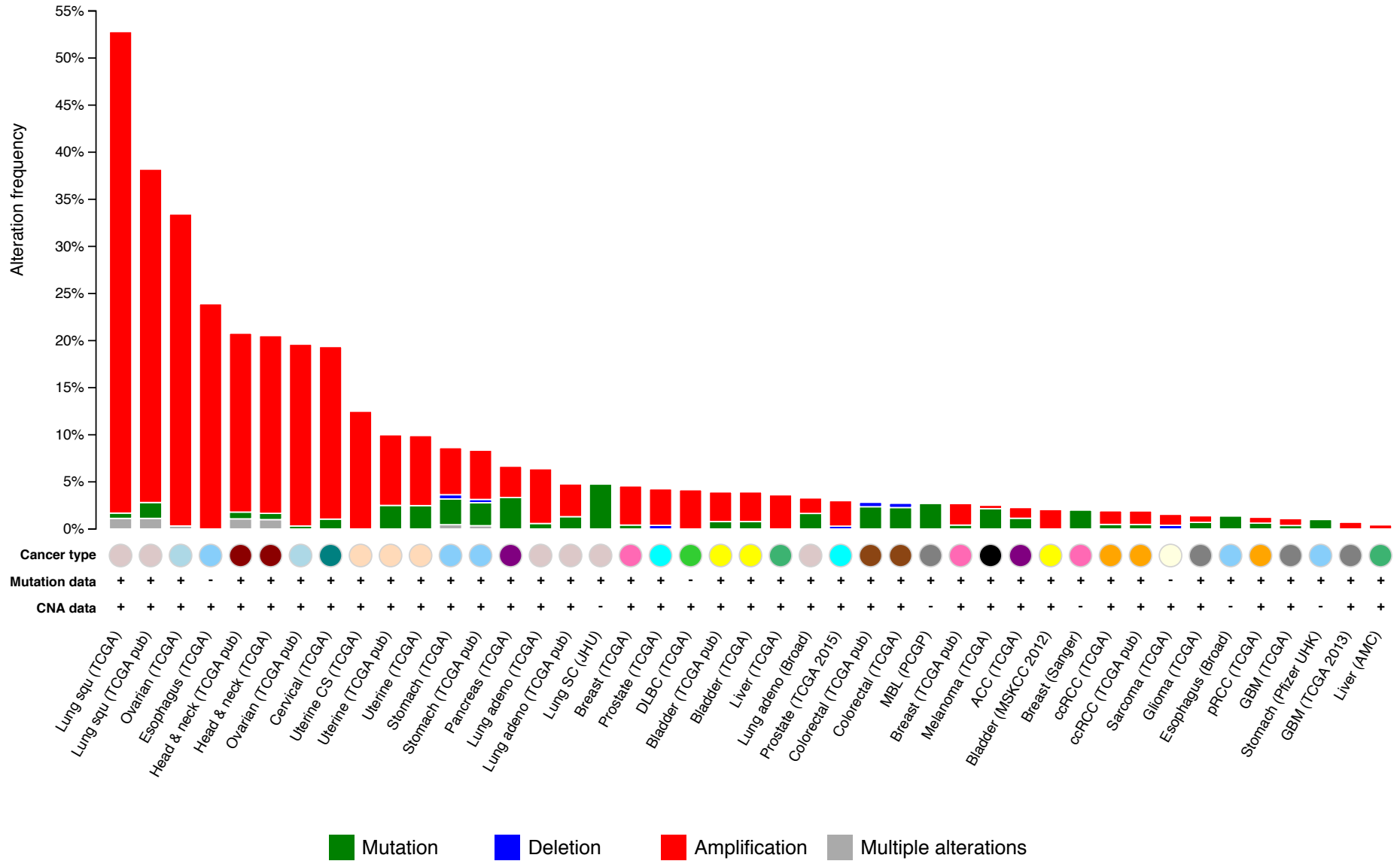
Appendix Table 3.1: List of genes relating to actin cytoskeleton regulation and the E-cadherin complex that were analysed in the TCGA meta-analysis.

Actin regulators			Signalling molecules			E-cadherin complex genes		
Protein	Gene name	Alterations*	Protein	Gene name	Alterations*	Protein	Gene name	Alterations*
Arp2/3 complex	*ARPC1A*	2%	Rho	*RND1*	0.30%	E-cadherin	*CDH1*	13%
	ARPC1B	2%		*RND2*	2%	a-catenin	*CTNNA1*	2%
	ACTR2	1%		*RND3*	0.6%	b-catenin	*CTNNB1*	0.60%
	ACTR3	0.60%	Rac	*RAC1*	0.9%	g-catenin	*JUP*	3%
	ARPC4	2%		*RAC2*	0.60%	p120	*CTNND1*	2%
	ARPC5	12%^		*RAC3*	6%			
N-WASP	*WASL*	1%	Cdc42	*CDC42*	0.90%			
WAVE2	*WASF2*	1%	Ect2	*ECT2*	5%	Myosins		
WIRE	*WIPF2*	7%	MgcRacGAP	*RACGAP1*	0.40%	Protein	Gene name	
Cortactin	*CTTN*	14%^	KIF23 (MLKP1)	*KIF23*	1%	Myosin IIA (MYO9)	*MYH9*	2%
MENA	*ENAH*	14%^	RPTPalpha	*PTPRA*	2%	Myosin IIB (MYO10)	*MYH10*	2%
VASP	*VASP*	2%	Src	*SRC*	2%	Myosin VI	*MYO6*	2%
EVL	*EVL*	1%	Rap1	*RAP1GAP*	0.90%			
Vinculin	*VCL*	3%	p130Cas	*BCAR1*	3%			
Abi1	*ABI1*	1%	ROCK	*ROCK1*	1%			
				ROCK2	3%			

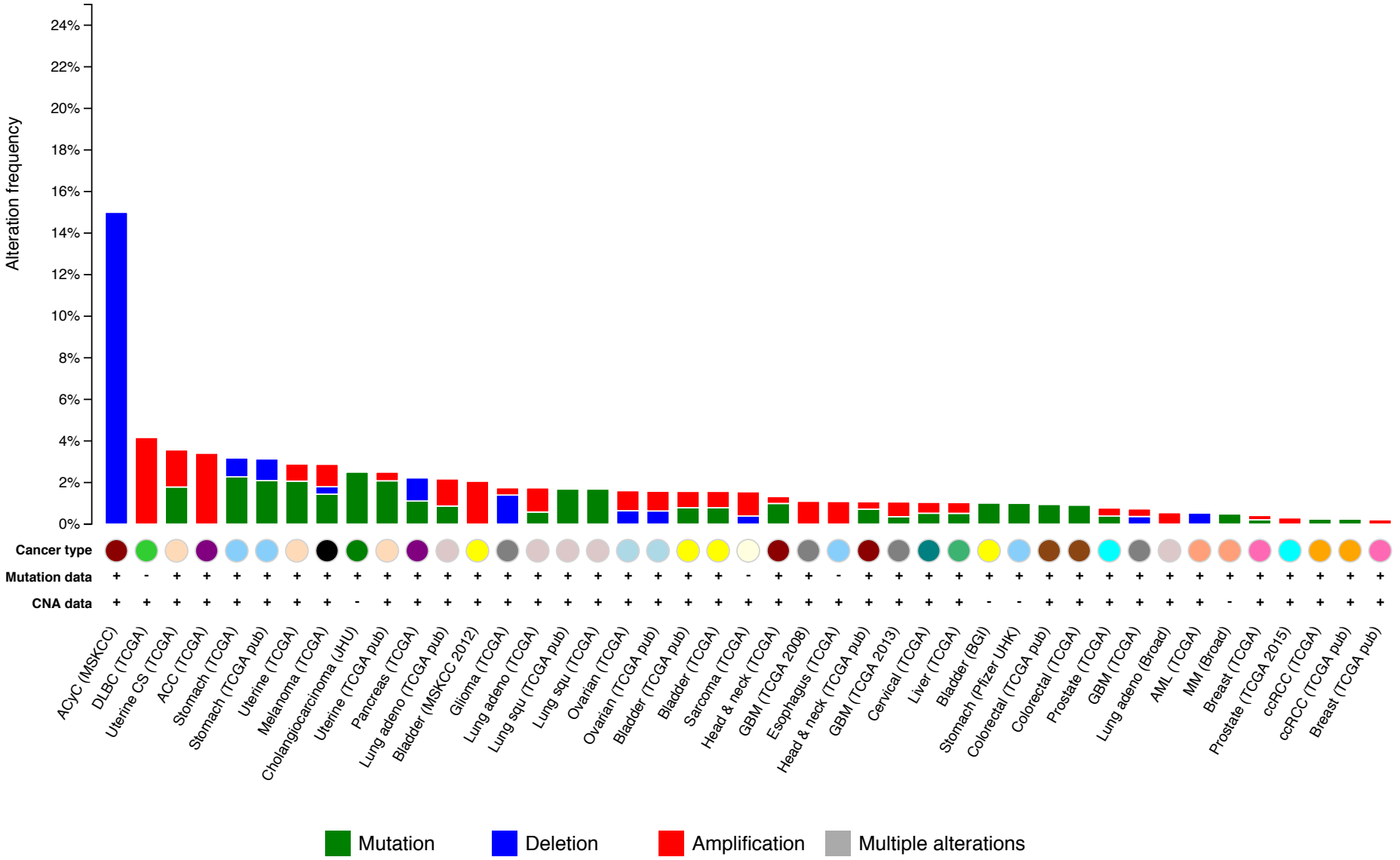
*Alterations = Using the cBioportal, gene amplification, gene deletion and nonsynonomous mutions were investigated for all of the genes across 962 breast cancer samples. The percentage indicates the number of samples with any alteration. ^ Since *ARPC5, CTTN* and *ENAH* had a high percentage of amplifications, gene expression was analysed and found there there was no mRNA overexpression alongside gene amplification. *ARPC5* and *ENAH* are located on chromosome 1q (1q25 and 1q42, respectively) and therefore may be a consequence of high level amplification of chromosome 1q that occurs commonly in breast cacner. *CTTN* is located on 11q13 and is also amplified relatively frequenly in breast cancer.



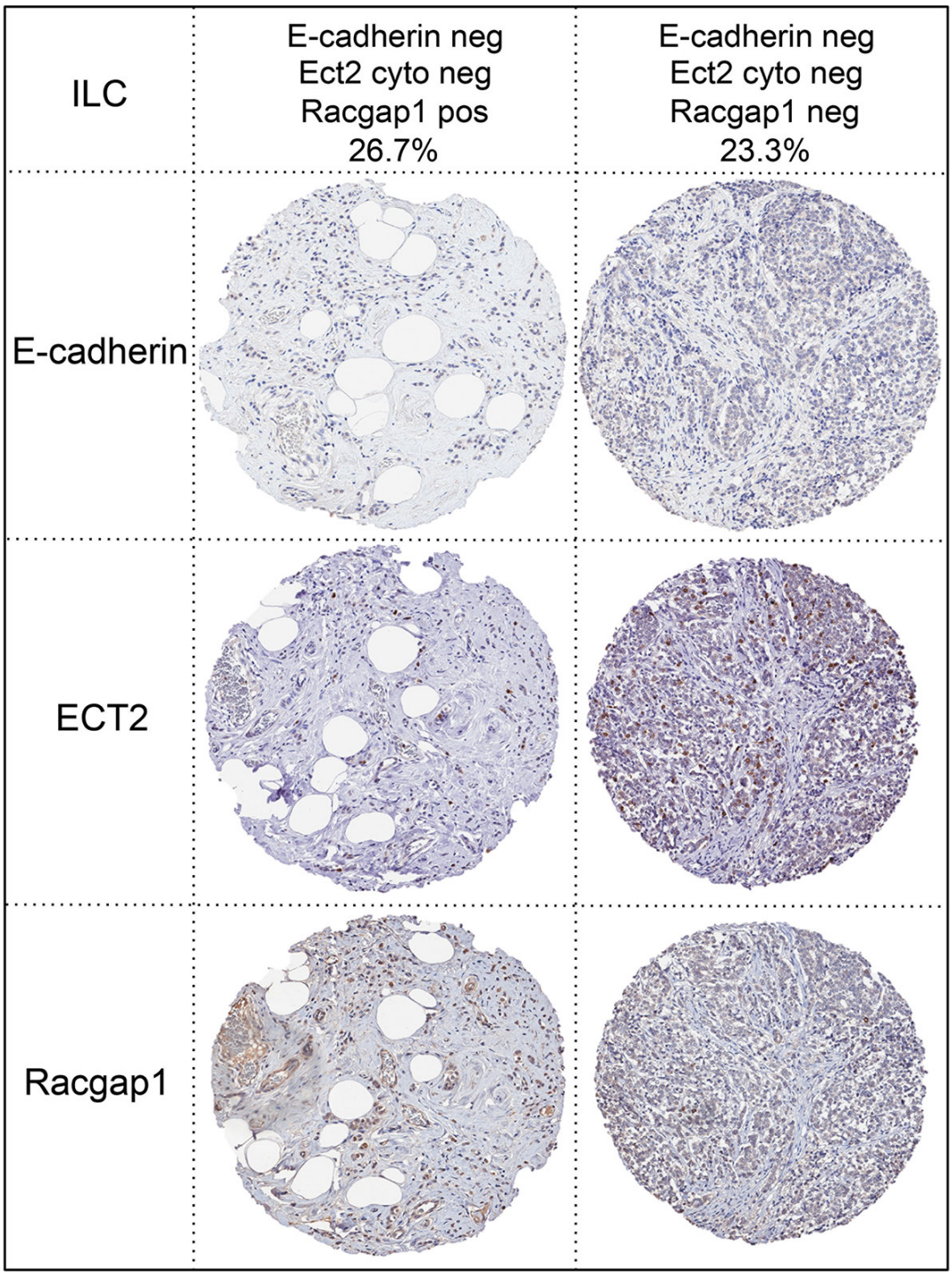
Appendix Figure 3.2: N-WASP genomic alterations across all cancer types.



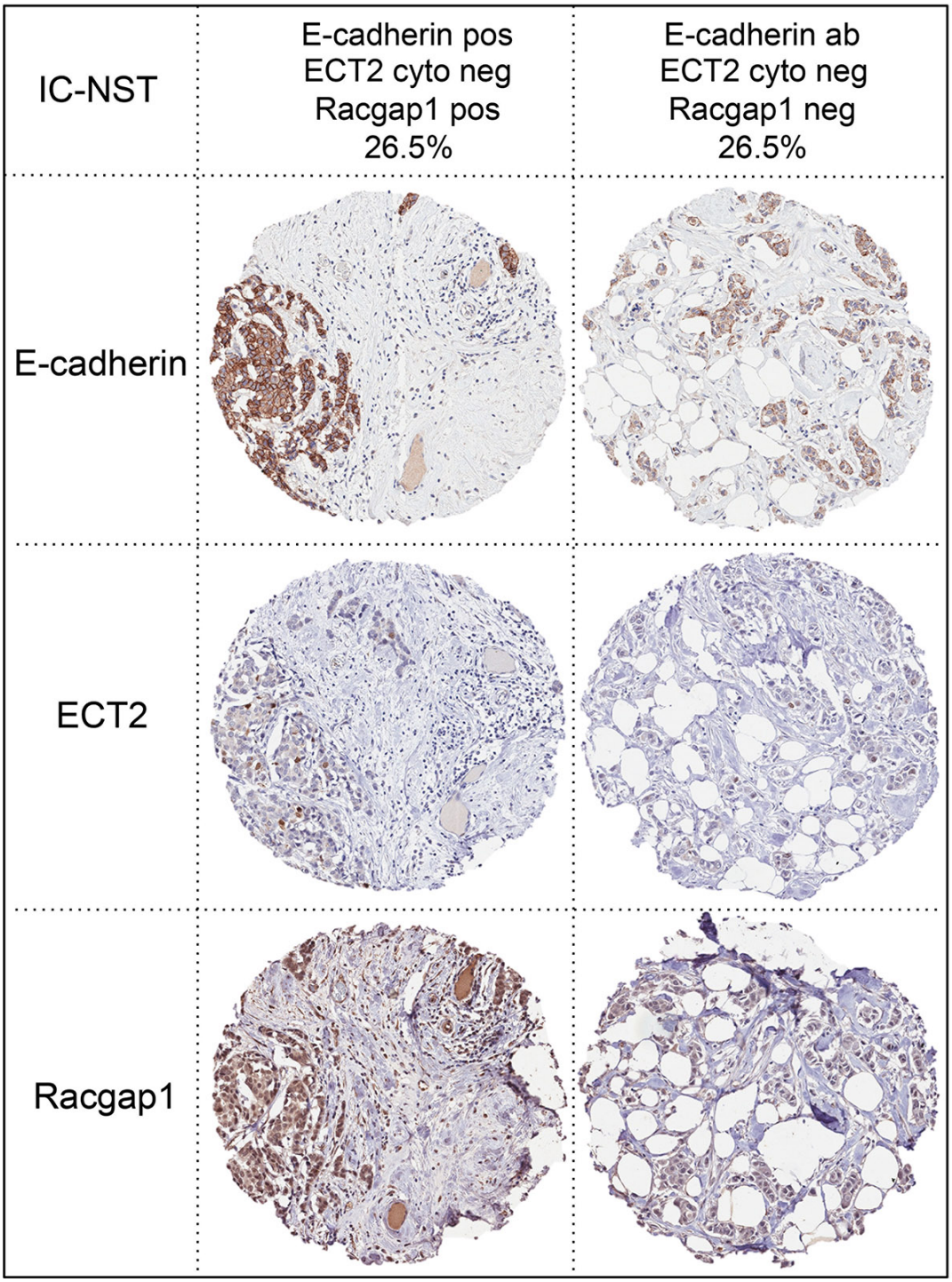
Appendix Figure 3.3: ECT2 genomic alterations across all cancer types.



Appendix Figure 3.4: RacGAP1 genomic alterations across all cancer types.



Appendix Figure 3.5: images of the most frequent phenotype observed when assessing co-expression of E-cadherin with ECT2 and RacGAP1 within invasive lobular carcinoma.



Appendix Figure 3.6: Representative images of the most frequent phenotype observed when assessing co-expression of E-cadherin with ECT2 and RacGAP1 within invasive carcinoma no special type.

Appendix Table 4.1: Whole exome sequencing metrics.

		Coverage (80% of the baits)	Average Coverage	Duplicate rate(%)	TotalMuts	Variants with rs number	% Known variants	SilentMuts	Non-silent Muts	SNPs	DNP	INDELs
MDL 4	IDC	12	59.54	16.59	98			27	71	94	2	2
	ILC	21	72.27	20.77	149			37	112	137	2	10
	Normal	36	110.39	12.72	21129*	20829	98.5					
MDL 5	Ecad pos	22	70.37	22.46	33			11	22	27	2	4
	Ecad neg	22	69.21	21.05	27			4	23	22	0	5
	DCIS	30	83.9	22.35	35			8	27	29	0	6
	Normal	23	74.1	22.26	16807*	16590	98.7					
MDL 6	IDC	34	101.73	18.47	169			41	128	161	2	6
	DCIS	32	102.79	14.01	72			18	54	67	1	4
	PLC	33	100.53	13.51	467			90	377	459	2	6
	PLCIS	37	112.67	14.31	190			40	150	183	2	5
	Normal	47	120.56	15.22	21889*	21539	98.4					
MDL 7	T1_IDC	16	76.36	21.06	22			5	17	20	0	2
	T2_ILC	16	75.1	18.23	25			6	19	24	1	0
	T3_DCIS	21	78.44	21.72	22			5	17	20	0	2
	T3_Ductal enriched	18	79.81	22.69	27			7	20	22	1	4
	T3_Admixed	16	72.33	21.78	25			6	19	22	1	2
	T4_DCIS	21	84.7	18.83	25			5	20	24	0	1
	T4_Ductal enriched	12	65.49	17.14	19			4	15	18	1	0
	T4_Admixed	13	73.32	19.05	25			8	17	24	1	0
	MDL90_Normal	21	63.82	29.12	18078*	17848	98.7					

* Germline mutations, Muts = mutations, SNPs = single nucleotide polymorphism, DNPs = dinucleotide polymorphism, INDELs = insertion or deletion.

Appendix Table 5.1: Distribution of gynaecological metastases.

	n (%)
Metastatic pattern	
Gynae only	7 (13)
Gynae & other site	47 (87)
Total	**54**
Total number of metastatic sites	
All sites	255
Gynaecological organs	137
All non-gynaecological organs	118
Gynaecological organs involved	
Ovary	64 (46.0)
Fallopian tube	33 (23.7)
Uterus (incl; endo/myometrium)	28 (20.1)
Cervix	7 (5.0)
Vagina	4 (2.9)
Mons pubis	1 (0.7)
Total	**137**
Peritoneum	
Ascitic/peritoneal fluid	7 (5.9)
Omentum	17 (14.4)
Peritoneum/retroperitoneum	11 (9.3)
Pouch of douglas	4 (3.4)
Pelvic plaque	1 (0.8)
Paracolic gutter	1 (0.8)
Total	**41 (34.7)**
Digestive system	
Appendix	6 (5.1)
Liver	7 (5.9)
Stomach	3 (2.5)
Colon	2 (1.7)
Intestine	2 (1.7)
Rectum	1 (0.8)
Anterior abdominal wall biopsy	1 (0.8)
Jejunal plaque stomach	1 (0.8)
Small bowel adhesion	1 (0.8)
Ascending colon fat	1 (0.8)
Total	**25 (21.2)**
Musculoskeletal system	
Bone	20 (16.9)
Total	**20 (16.9)**

Table continued over page

Appendix Table 5.1: Distribution of gynaecological metastases.

	n (%)
Nervous system	
Brain	4 (3.4)
Brachial plexus (nerves in neck)	2 (1.7)
Cerebrospinal fluid	1 (0.8)
Total	**7 (5.9)**
Lymphatic system	
Distant LN	3 (2.5)
Pelvic lymph node	2 (1.7)
Iliac LN	2 (1.7)
Para-aortic LN	1 (0.8)
Peritoneal lymph node	2 (1.7)
Total	**10 (8.5)**
Respiratory system	
lung	5 (4.2)
pleura/pleural fluid	5 (4.2)
Intercostal space	1 (0.8)
Total	**11**
Integumentary system	
Skin	2 (1.7)
Total	**2 (1.7)**
Urinary system	
Bladder	2 (1.7)
Total	**2 (1.7)**

Appendix Table 5.2: Distribution of gynaecological metastases.

First metastatic detected	n	%
Gynaecological organ	38	70.4
Other	16	29.6
bone	7	13.0
Pleural fluid/lung	5	9.3
Peritoneum/omentum	2	3.7
Skin	2	3.7

Appendix Table 5.3 IHC raw data

		Breast			All mets			Ovaries			Other Gynaecological organs		
Immuno marker		total	n pos	%	total	n pos	%	total	n pos	%	total	n pos	%
ER	(>1%)	46	43	93.5	120	112	93.3	53	49	92.5	39	38	97.4
PR	(>1%)	35	23	65.7	119	79	66.4	53	37	69.8	39	23	59.0
HER2	3+ M	33	0	0.0	113	0	0.0	51	0	0.0	37	0	0.0
CK5/6	(>10%)	19	0	0.0	101	1	1.0	49	0	0.0	32	1	3.1
CK14	(>10%)	18	0	0.0	101	1	1.0	48	1	2.1	33	0	0.0
EGFR	3+ M	17	0	0.0	99	1	1.0	49	1	2.0	29	0	0.0
CK8/18	(>1%)	19	19	100.0	103	103	100.0	50	50	100.0	33	33	100.0
Ki67	Low (<10%)	19	5	26.3	102	47	46.1	49	24	49.0	32	15	46.9
	Mod (10-30%)		0	0.0		23	22.5		10	20.4		7	21.9
	High (>30%)		0	0.0		11	10.8		4	8.2		4	12.5
p53	Low (10-30%)	19	2	10.5	103	28	27.2	48	14	29.2	34	10	29.4
	High (>30%)		4	21.1		39	37.9		19	39.6		11	32.4
GATA3	(>1%)	19	19	100.0	105	57	54.3	47	32	68.1	35	18	51.4
FOXA1	(>1%)	19	19	100.0	108	79	73.1	49	38	77.6	36	25	69.4
AR	(>1%)	19	17	89.5	102	79	77.5	49	39	79.6	33	23	69.7

Table continued over page.

Appendix Table 5.3 IHC raw data

		Omentum/Peritoneum			Gastrointestinal organs			Lymph nodes (distant and regional)			Brain		
Immuno marker		total	n pos	%	total	n pos	%	total	n pos	%	total	n pos	%
ER	(>1%)	16	13	81.3	5	5	100	5	5	100	1	1	100
PR	(>1%)	16	8	50.0	5	5	100	5	5	100	x	x	x
HER2	3+ M	13	0	0.0	6	0	0	5	0	0	1	0	0
CK5/6	(>10%)	11	0	0.0	4	0	0	5	0	0	x	x	x
CK14	(>10%)	11	0	0.0	2	0	0	5	0	0	x	x	x
EGFR	3+ M	12	0	0.0	4	0	0	5	0	0	x	x	x
CK8/18	(>1%)	11	11	100.0	4	4	100	5	5	100	x	x	
Ki67	Low (<10%)	12	6	50.0	4	1	25	5	1	20	x	x	x
	Mod (10-30%)		2	16.7		2	50.0		2	40			x
	High (>30%)		1	8.3		1	25.0		1	20			x
p53	Low (10-30%)	12	3	25.0	4	1	25.0	5	0	0	x	x	x
	High (>30%)		4	33.3		3	75.0		2	40		x	x
GATA3	(>1%)	12	3	25.0	4	2	50.0	5	1	20	1	1	100
FOXA1	(>1%)	12	9	75.0	5	4	80.0	5	3	60	1	0	0
AR	(>1%)	10	3	30.0	4	4	100	5	5	100	1	1	100

Appendix Table 5.4 Summary of the matched gynaecological metastatic breast cancer cases. 1 = positive, 0 = negative, x = no data.

GM#	Site	ER	PR	AR	GATA3	FOXA1	Invasive type	Pre/postmenopausal	Endocrine_primary	Chemo_primary	Rxt_primary	Endocrine_met	Chemo_met	Rxt_met	aCGH	Exome seq	Targeted seq
GM01	Breast IC-NST	1	1	1	1	1	MDL	Pre	Y	Y	Y	N	Y	N	x	x	x
GM01	Breast ILC	1	1	1	1	1									x	x	x
GM01	Ovaries	1	1	1	1	1									x	x	x
GM06	Breast	1	1	1	1	1	ILC	Pre	N	Y	Y	N	x	x	Y	x	x
GM06	Cervix	1	0	1	1	1									x	x	x
GM06	Endomyometrium	x	x	x	0	0									x	x	x
GM06	Lt ovary	1	1	1	1	1									Y	x	x
GM06	Rt ovary	1	1	1	1	1									Y	x	x
GM13	Breast	1	0	x	x	x	IC-NST	Pre	x	Y	N	Y	x	x	x	x	x
GM13	Rt ovary	1	1	0	0	0									Y	x	x
GM16	Breast	1	1	0	1	1	IC-NST	Post	Y	x	Y	Y	Y	x	Y	x	Y
GM16	Appendix	1	1	1	1	1									N	x	x
GM16	Lt ovary	1	1	1	0	1									Y	x	x
GM16	Lt fallopian tube	1	1	1	0	1									Y	x	Y
GM16	Omentum	1	1	1	0	1									x	x	Y
GM16	Pleura	1	0	1	1	1									x	x	x
GM16	Rt ovary	1	1	1	0	1									Y	x	x
GM16	Rt fallopian tube	1	1	1	0	1									Y	x	x
GM24	Breast	1	1	1	1	1	IC-NST	Pre	x	Y	Y	x	x	x	Y	x	x
GM24	Lt ovary	1	0	1	1	1									Y	x	x
GM29	Breast	1	0	1	1	1	IC-NST	Post	Y	x	Y	x	x	x	Y	x	x
GM29	Cervix	1	0	1	1	1									Y	x	x

GM#	Site	ER	PR	AR	GATA3	FOXA1	Invasive type	Pre/postmenopausal	Endocrine_primary	Chemo_primary	Rxt_primary	Endocrine_met	Chemo_met	Rxt_met	aCGH	Exome seq	Targeted seq
GM36	Breast	1	1	1	1	1	IC-NST	Pre	x	Y	Y	Y	Y	Y	Y	x	x
GM36	Uterus	1	0	1	0	1									Y	x	x
GM39	Breast	1	1	1	1	1	IC-NST	Pre	Y	x	Y	Y	Y	x	N	x	x
GM39	Cervix	1	1	1	1	1									Y	x	x
GM39	Lt ovary	1	1	1	1	1									Y	x	x
GM39	rt ovary	1	0	1	1	1									Y	x	x
GM41	Breast	1	1	1	1	1	IC-NST	Pre	Y	Y	Y	Y	Y	x	Y	x	x
GM41	Ovary	x	x	x	x	0									x	x	x
GM45	Breast	1	x	x	x	x	ILC	Pre	x	x	x	x	Y	x	x	x	x
GM45	Appendix	1	1	1	0	0									Y	x	x
GM45	Cervix	x	x	x	0	0									x	x	x
GM45	Lt ovary	1	1	1	0	1									Y	x	x
GM45	Lt pelvic lymph node	1	1	1	0	1									Y	x	x
GM45	Lt fallopian tube	1	1	1	0	1									x	x	x
GM45	Myometrium	1	1	0	1	1									x	x	x
GM45	Omentum	1	1	1	0	1									Y	x	x
GM45	Omentum lymph node	1	1	1	0	1									x	x	x
GM45	Rt ovary	1	1	1	0	1									Y	x	x
GM45	Rt pelvic lymph node	1	1	1	0	1									Y	x	x
GM45	Rt fallopian tube	1	1	1	0	1									Y	x	x

GM#	Site	ER	PR	AR	GATA3	FOXA1	Invasive type	Pre/postmenopausal	Endocrine_primary	Chemo_primary	Rxt_primary	Endocrine_met	Chemo_met	Rxt_met	aCGH	Exome seq	Targeted seq
GM49	Breast	1	1	x	x	x	ILC	Pre	Y	Y	Y	Y	Y	x	x	x	x
GM49	Lt ovary	1	1	1	1	1									Y	x	x
GM49	Lt fallopian tube	1	1	1	1	1									x	x	x
GM49	Mmentum	1	1	1	1	1									x	x	x
GM49	Rt ovary	1	1	1	1	1									Y	x	x
GM49	Rt fallopian tube	1	1	1	1	1									x	x	x
GM49	Uterus	1	1	1	1	1									x	x	x
GM50	Breast	1	1	x	x	x	ILC	Pre	Y	x	Y	Y	Y	x	x	x	x
GM50	Lt ovary P1	1	1	1	1	1									Y	x	x
GM50	Lt ovary P2	1	1	1	1	1									Y	x	x
GM50	Lt fallopian tube P1	1	0	0	1	1									Y	x	x
GM50	Rt ovaryP1	1	1	1	1	1									Y	x	x
GM50	Rt ovary P2	1	1	1	1	1									Y	x	x
GM50	Rt fallopian tube P2	1	1	0	1	1									Y	x	x
GM56	Breast	1	x	x	x	x	ILC	Post	x	x	x	x	x	x	x	x	x
GM56	Ovary	1	1	1	1	1									x	x	x
GM57	Breast	1	0	0	1	1	IC-NST	Pre	Y	Y	Y	Y	Y	x	Y	x	Y
GM57	Cervix P2	1	0	1	1	1									N	x	x
GM57	Cervix P3	1	0	1	1	1									Y	x	x
GM57	Rt ovary P1	1	1	1	1	1									Y	x	Y
GM57	Rt ovary P2	1	1	1	1	1									Y	x	x

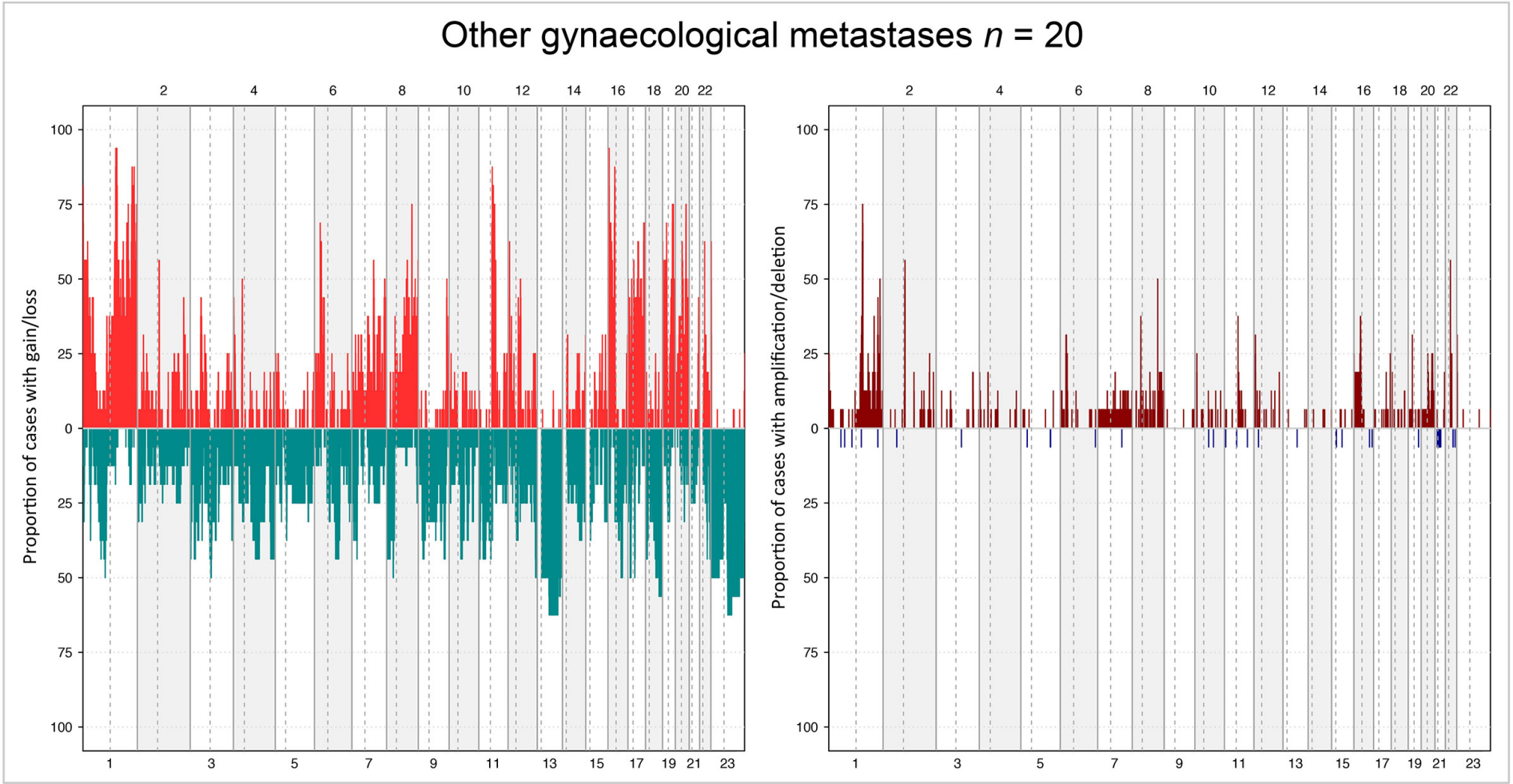
GM#	Site	ER	PR	AR	GATA3	FOXA1	Invasive type	Pre/postmenopausal	Endocrine_primary	Chemo_primary	Rxt_primary	Endocrine_met	Chemo_met	Rxt_met	aCGH	Exome seq	Targeted seq
GM59	Breast	1	1	1	1	1	IC-NST	Pre	Y	Y	Y	Y	Y	x	Y	x	Y
GM59	Lymph node	1	1	1	0	0									x	x	x
GM59	Pleura	0	0	0	0	1									x	x	x
GM59	Rt ovary	1	0	1	1	1									Y	x	Y
GM63	Breast	1	1	1	1	1	ILC	Post	x	x	x	x	x	x	Y	Y	x
GM63	Lymph node	1	1	1	1	0									Y	N	x
GM63	Lt fallopian tube	1	0	0	1	1									Y	Y	x
GM63	Lt ovary	1	0	0	1	1									Y	Y	x
GM63	Omentum	1	0	0	1	1									Y	Y	x
GM63	Rt fallopian tube	1	0	0	1	1									Y	Y	x
GM63	Rt ovary	1	1	0	1	0									Y	Y	x
GM70	Breast	1	x	x	x	x	IC-NST	Post	x	x	x	x	x	x	x	x	x
GM70	Ovary	x	0	x	x	0									x	x	x
GM70	Peritoneum	0	1	x	0	0									x	x	x
GM73	Breast	1	x	x	x	x	Tubular	Post	x	x	Y	x	x	x	x	x	x
GM73	Cervix	1	1	0	0	0									x	x	x
GM73	Vagina	1	1	0	0	0									Y	x	x
GM74	Breast	0	0	1	1	1	ILC	Pre	x	Y	Y	x	x	x	Y	x	x
GM74	Lt ovary P1	0	0	0	0	0									Y	x	x
GM74	Lt ovary P2	0	0	0	0	0									Y	x	x
GM76	Breast	1	1	x	x	x	ILC	Pre	Y	Y	Y	Y	Y	x	x	x	x
GM76	Endometrium	0	0	1	0	1									x	x	x
GM76	Ovary	1	0	1	0	1									x	x	x

GM#	Site	ER	PR	AR	GATA3	FOXA1	Invasive type	Pre/postmenopausal	Endocrine_primary	Chemo_primary	Rxt_primary	Endocrine_met	Chemo_met	Rxt_met	aCGH	Exome seq	Targeted seq
GM77	Breast	1	1	1	1	1	ILC	Pre	Y	Y	Y	x	x	x	N	x	x
GM77	Lt ovary	1	1	1	1	1									Y	x	x
GM77	Rt ovary	1	1	1	1	1									Y	x	x
GM78	Breast	1	1	1	1	1	ILC	Pre	x	x	x	x	Y	Y	Y	x	x
GM78	Colon + surrounding fat	1	1	1	0	1									Y	x	x
GM78	Omentum	1	1	1	0	1									Y	x	x
GM78	Ovary P1	1	1	1	0	1									Y	x	x
GM78	Ovary P2	1	1	1	0	1									x	x	x
GM82	Breast	1	x	x	x	x	IC-NST	Pre	x	x	x	x	x	x	x	x	x
GM82	Cervix P1	1	1	0	0	0									Y	x	x
GM82	Lt ovary P1	1	0	0	0	0									x	x	x
GM82	Lt ovary P3	1	1	0	0	0									Y	x	x
GM82	Pouch of douglas	1	0	x	0	0									Y	x	x
GM82	Rt ovary p1	1	1	0	1	0									Y	x	x
GM82	Rt ovary p3	0	1	0	0	0									x	x	x
GM82	Vagina P1	1	1	0	0	0									Y	x	x
GM82	Vulva P1	1	0	0	0	0									Y	x	x
GM82	Vulva P2	1	0	1	0	0									Y	x	x
GM87	Breast	1	1	x	x	x	ILC	Pre	x	x	Y	x	x	Y	x	x	x
GM87	Brain	1	x	1	1	0									N	x	x
GM87	Liver	x	x	x	x	0									x	x	x
GM87	Ovaries P1	1	0	1	1	1									Y	x	x
GM87	Ovaries P2	1	0	1	1	1									x	x	x
GM87	Ovaries P3	1	0	1	x	1									x	x	x

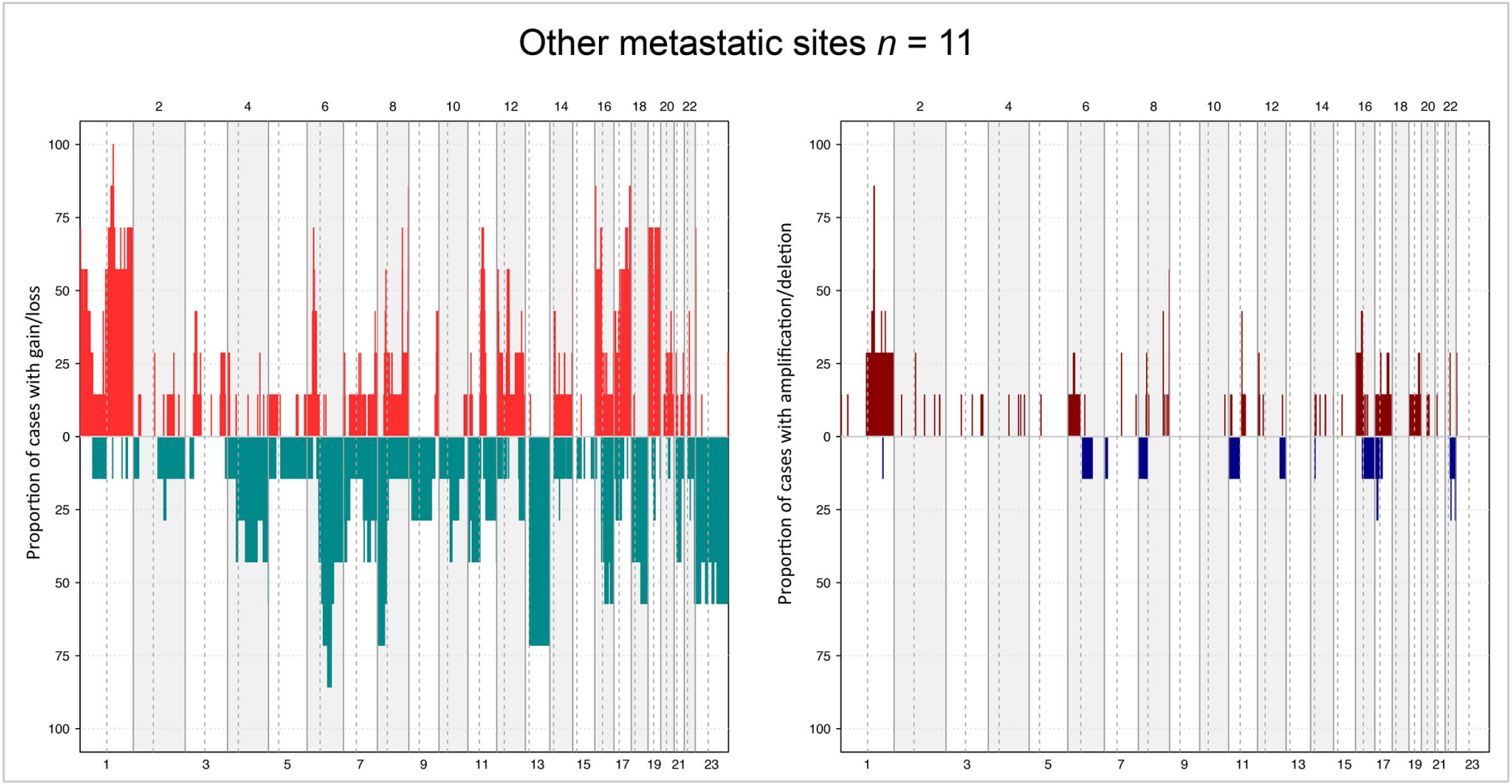
GM#	Site	ER	PR	AR	GATA3	FOXA1	Invasive type	Pre/postmenopausal	Endocrine_primary	Chemo_primary	Rxt_primary	Endocrine_met	Chemo_met	Rxt_met	aCGH	Exome seq	Targeted seq
GM88	Breast	1	1	x	x	x	ILC	Post	x	x	x	x	x	x	x	x	x
GM88	Endometrium	1	0	1	1	1									Y	x	x
GM89	Breast	1	x	x	x	x	IC-NST	Pre	Y	Y	Y	Y	Y	Y	x	x	x
GM89	Rt ovary P1	1	0	1	1	1									Y	x	x
GM89	Rt ovary P2	1	1	1	1	1									Y	x	x

Appendix Table 5.5: List of the genes located on the significantly amplified region in the ovarian metastases, 1q23.1.

ARHGEF11
BCAN
C1orf66
C1orf92
CD5L
CRABP2
ETV3
ETV3L
FCRL1
FCRL2
FCRL3
FCRL4
FCRL5
HDGF
INSRR
ISG20L2
MRPL24
NES
NTRK1
PEAR1
PRCC
SH2D2A



Appendix Figure 5.1: The frequency of genomic alterations across all chromosomes in 20 non-ovary gynaecological metastases. Left panel displays gains and loss, right panel displays amplifications and homozygous deletions.



Appendix Figure 5.2: The frequency of genomic alterations across all chromosomes in 11 non-gynaecological metastases. Left panel displays gains and loss, right panel displays amplifications and homozygous deletions.