A Genetic Basis for Luminal and Basal-Type Breast Cancer

Antoinette Hollestelle

A Genetic Basis for Luminal and Basal-Type Breast Cancer

Een genetische basis voor luminale en basale typen borstkanker

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de rector magnificus
Prof.dr. S.W.J. Lamberts
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
woensdag 7 januari 2009 om 9.45 uur

door

Antoinette Hollestelle
geboren te Nijmegen

PROMOTIECOMMISSIE

Promotor:
Prof.dr. J.G.M. Klijn

Copromotor:
Dr. M. Schutte

Overige leden:
Prof.dr. R. Fodde
Prof.dr. J.W. Oosterhuis
Prof.dr. P. Devilee

CONTENTS

General Introduction

1.1 BREAST CANCER EPIDEMIOLOGY AND ETIOLOGY

Cancer is the second leading cause of death, after heart disease, and a major health issue in the western world. In the Netherlands, 35% of women will develop cancer during their lifetime of which one third is diagnosed with breast cancer. This means that 1 out of 8 women in the Netherlands will develop breast cancer during her lifetime, making breast cancer the most frequently diagnosed cancer in women [1]. Breast cancer is not only the second leading cause of cancer death in western countries, after lung cancer, it is also the leading cause of overall mortality in women aged 35-55 years. In 2003, 11687 women were diagnosed with breast cancer in the Netherlands and 3361 women died resulting from the disease (http://www.ikcnet.nl/index.php). Currently, the 5-year survival for breast cancer in the Netherlands is 85% [1]. Fortunately, breast cancer mortality rates are declining in the Netherlands as a result of earlier diagnosis, through increased awareness and the breast cancer screening program on the one hand, and improved treatment on the other [2, 3].

Breast cancer incidence rates strongly vary around the world. The highest incidences are found in western countries of Europe and North America, whereas in the developing countries of South America and Africa incidence rates are relatively low [1, 4, 5]. Breast cancer incidence is dependent on the presence or absence of certain risk factors. The major risk factors for breast cancer include female gender, increasing age, western culture, positive family history of breast or ovarian cancer in first-degree relatives, germline mutation in a high-risk breast cancer susceptibility gene, prior diagnosis of breast cancer, benign breast disease with atypical hyperplasia, and exposure to ionizing radiation in young women. Other risk factors include early age at menarche, late age at menopause, nulliparity, late age at first child birth, small number of children, use of oral contraceptives and hormone replacement therapy, high socioeconomic status, dense tissue on mammography and postmenopausal obesity [6-8].

In the Netherlands, breast cancer incidence rates are still rising (http://www.ikcnet.nl/index.php, [9]). It is believed that this is due to the increasing average age of the population, resulting in a higher percentage of women in the age group where breast cancer is most commonly diagnosed [9]. Another reason for the increasing breast cancer incidence rates is the introduction of the national breast cancer screening program in 1989 for women aged 50 to 69 years. As a result, there was a substantial increase of breast cancer incidence in women aged 50-69 years after 1989, but not of women younger than 50 years or 70 years and older [2, 9]. Additionally, amongst women aged 50-69 years there was an increase in the rate of stage I cancers, and a decrease in stage III+ cancers [2]. As of 1998, women aged 70-74 years were therefore also invited for the national screening program [2, 9].

1.2 CLASSIFICATION OF BREAST CANCER

Fifty percent of women diagnosed with breast cancer will survive the disease without recurrence, whereas 15% of the patients will survive the disease despite a recurrence within 15 years. However, one third of breast cancer patients will die of metastases of the primary cancer within 15 years from diagnosis. It is therefore important to distinguish patients with a good prognosis that do not need additional therapy from patients with a poorer prognosis that will benefit from additional therapy. Reliable prognostic and predictive factors that classify breast tumors accurately are thus imperative for the clinician and have been a major focus in breast cancer research.

1.2.1 TRADITIONAL PROGNOSTIC AND PREDICTIVE FACTORS

TNM stage

As yet, the most powerful predictor of breast cancer recurrence is tumor stage. The TNM method for tumor staging is based on three tumor characteristics at the time of diagnosis: tumor size (T), axillary lymph node involvement (N) and the presence of metastases (M). Together these three factors define tumor stages I through IV (Table 1.1). Almost 90% of all patients with stage I cancers survive at least 5 years after diagnosis, whereas 5-year survival rates for stage II and III cancers are 60-80% and 40-50%, respectively [10]. Patients who have a stage IV cancer have a very poor 10-year survival of less than 5% [11]. The TNM stage of the tumor is thus a very strong indicator of the 5-year survival of the patient.

Axillary lymph node status

The extent of axillary lymph node involvement at the time of diagnosis is one of the most reliable independent prognostic factors for breast cancer. Patients with tumor free lymph nodes have a far better prognosis than patients with positive lymph nodes, with about 15-45% of node-negative patients having a disease recurrence compared to 50-70% of node-positive patients. Additionally, the risk of disease recurrence as well as mortality increases with an increasing number of lymph nodes involved [13-15]. The lymph node status is determined by the sentinel node procedure for staging purposes, followed by an axillary node dissection when metastases are present in sentinel nodes [16, 17].

Tumor size

The size of the tumor is a very strong prognostic factor, even after 20 years of follow-up [18, 19]. Although some pathologists measure the macroscopic size or the microscopic size of the tumor including both the invasive part and the *in situ* components, only the microscopic size of the invasive part of the tumor is clinically significant. Tumor size is directly correlated to

Table 1.1 TNM stage classification (adapted from [12])

T	Primary tumor size
T0	No evidence for primary tumor
Tis	Carcinoma in situ
T1	Tumor of 2 cm or less in greatest dimension
T2	Tumor larger than 2 cm, but not more than 5 cm in greatest dimension
T3	Tumor larger than 5 cm in greatest dimension
T4	Tumor of any size with direct extension to chest wall of skin

N	Regional lymph node involvement
N0	No regional lymph node metastases
N1	Metastases in movable ipsilateral axillary lymph node(s)
N2	Metastases in fixed ipsilateral axillary lymph node(s) or in clinically apparent ipsilateral internal mammary lymph node(s) in the absence of clinically evident axillary lymph node involvement
N3	Metastases in ipsilateral infraclavicular lymph node(s) with or without axillary lymph node involvement; or in clinically apparent ipsilateral internal mammary lymph node(s) in the presence of clinically evident axillary lymph node metastases

M	Presence of distant metastases
M0	No distant metastases
M1	Distant metastases

Stage	T	N	M
0	Tis	N0	M0
I	T1	N0	M0
IIA	T0	N1	M0
	T1	N1	M0
	T2	N0	M0
IIB	T2	N1	M0
	T3	N0	M0
IIIA	T0	N2	M0
	T1	N2	M0
	T2	N2	M0
	T3	N1,N2	M0
IIIB	T4	N0,N1,N2	M0
IIIC	Any T	N3	M0
IV	Any T	Any N	M1

The size of the primary tumor (T), involvement of regional lymph nodes (N) and the presence of distant metastases (M) together define the stage of the breast tumor at the time of diagnosis.

axillary lymph node involvement, as larger tumors frequently have more positive lymph nodes. However, larger tumor size is correlated with a worse prognosis independent of lymph node status [20-22]. This is mainly because lymph node-negative patients with a tumor smaller than 1 cm have a far better prognosis than patients with a tumor larger than 2 cm (80% versus 65%, respectively) [23].

Histological tumor grade

Although not as strong as TNM stage, lymph node status or tumor size, the histological grade of a tumor is a good prognostic marker for breast cancer patients. Tumor grade is determined by the Scarf-Bloom-Richardson Grading system, modified by Elston and Ellis [24, 25]. According this grading system 1, 2 or 3 points are given for each of the following tumor characteristics: tubule formation, nuclear pleomorphism, and mitotic count. The sum of these points forms a score of 1 to 9 that determines the grade or differentiation status of the tumor (Table 1.2). Tumors with a low grade are well differentiated and predict a more favorable prognosis for the patient than poorly differentiated tumors with a high grade. The ten-year survival of patients with the lowest grade is 90-95% as opposed to 30-80% for patients with the highest grade [26, 27]. Additionally, higher grade is associated with negative hormone receptor status and low grade with positive hormone receptor status. Therefore, histological grade is correlated with response to either endocrine therapy (low grade) or chemotherapy (high grade) [28, 29]. Importantly, the distinct patterns of chromosomal loss between grade I versus grade III tumors has led to the assumption that the majority of grade I tumors do not progress to grade III tumors, but are likely to follow distinct genetic pathways [30].

Table 1.2 Histological grade (adapted from [12])

Score	Tubule formation
1	More than 75% of the tumor has tubule formation
2	10% to 75% of the tumor has tubule formation
3	Less than 10% of the tumor has tubule fromation

Score	Nuclear pleomorphism
1	Nuclei are small and uniform in size and shape
2	Nuclei are moderate in nuclear size and variation
3	Nuclei have marked variation, are relatively large, and have prominent or multiple nucleoli

Score	Mitotic count (per 10 high power fields with field area of 0.274 mm2)
1	0-9 mitoses
2	10-19 mitoses
3	More than 20 mitoses

Grade	Differentiation status	Total score
I	Well differentiated	3-5 points
II	Moderately differentiated	6-7 points
III	Poorly differentiated	8-9 points

A score of one, two or three points is given for each of the three tumor characteristics: tubule formation, nuclear pleomorphism and mitotic count. The sum of these three scores determines the grade or differentiation status of the tumor.

ER and PR status

Estrogens and progestins are important regulators of proliferation and differentiation of the mammary gland. These nuclear transcription factors exert their function by binding to their respective receptors: the estrogen receptor (ER) and the progesterone receptor (PR). Two thirds of breast cancers express both ER and PR, 10% are ER-positive and PR-negative, 5% are ER-negative and PR-positive and 20% of all breast cancers are negative for both ER and PR [31]. Patients with ER-negative tumors initially have a shorter survival than patients with ER-positive tumors. However, from 3 years of follow-up these differences in survival diminish, thereby losing the prognostic significance of ER [32]. The preferential site of distant metastases is different for ER-positive and ER-negative tumors, as ER-positive tumors tend to metastasize to bone and soft tissue, whereas ER-negative tumors metastasize more frequently to the liver, lung and central nervous system [33]. Importantly, ER-positive tumors respond better to endocrine therapy, whereas ER-negative tumors respond better to chemotherapy [34, 35]. About 50% of ER-positive tumors respond to endocrine therapy compared to less than 10% of ER-negative tumors. This can be even further refined by combining ER and PR, with up to 80% of patients with ER and PR-positive tumors responding to endocrine therapy compared to less than 10% of patients with ER and PR-negative tumors [36]. Therefore, the real power of ER and PR lies in their ability to predict the most appropriate class of systemic therapy.

ERBB2 overexpression

The ERBB2 receptor tyrosine kinase is a member of the epidermal growth factor receptor (EGFR) family and is also referred to as HER2/neu. About 25% of invasive breast cancers have amplification of the *ERBB2* gene and/or overexpression of the ERBB2 protein [37-39]. As a prognostic factor, ERBB2 overexpression is most valuable to lymph node-positive patients and associated with a less favorable clinical outcome in the pre-trastuzumab era. Lymph node-positive patients without ERBB2 overexpression have a ten-year survival of 65%, whereas those with ERBB2 overexpression had a ten-year survival of 50% [40, 41]. For node-negative patients ERBB2 overexpression has no significant prognostic value [26, 40, 42]. However, ERBB2 overexpression is able to predict response to trastuzumab (Herceptin) antibody therapy in both patient groups, but only 50% of the advanced patients appear to respond [43-45]. Activation of the PI3K pathway in patients with ERBB2 overexpressing breast cancer appears to be a major determinant of the resistance to trastuzumab [46]. Similar to the hormone receptors, the value of ERBB2 thus lies in its prediction of targeted therapy response.

1.2.2 HISTOPATHOLOGY

1.2.2.1 HISTOLOGY OF THE NORMAL BREAST

The functional unit of the mammary gland is a hormone responsive tubulo-alveolar gland of which the functional component is the terminal ductal lobular unit (TDLU) connected to an excretory system consisting of a large duct system (Figure 1.1). These components are embedded in the stromal tissue. The TDLU is the secretory part of the gland and composed of lobules and terminal ductules that discharge into the large duct system and excrete via the nipple. The epithelium that lines the entire lobular-ductal system is composed of two layers, an inner secretory luminal epithelial layer and an outer contractile basal myoepithelial layer. The inner epithelial cells are columnar or cuboidal shaped and form a polarized continuous layer that lines the lumen. The outer myoepithelial cells are typically elongated when cut longitudinally and triangular in cross section, and reside between the luminal epithelial layer and the basement membrane. In the ducts these myoepithelial cells form a nearly continuous layer, whereas in the lobules they form a discontinuous basket-like structure around the acini [47].

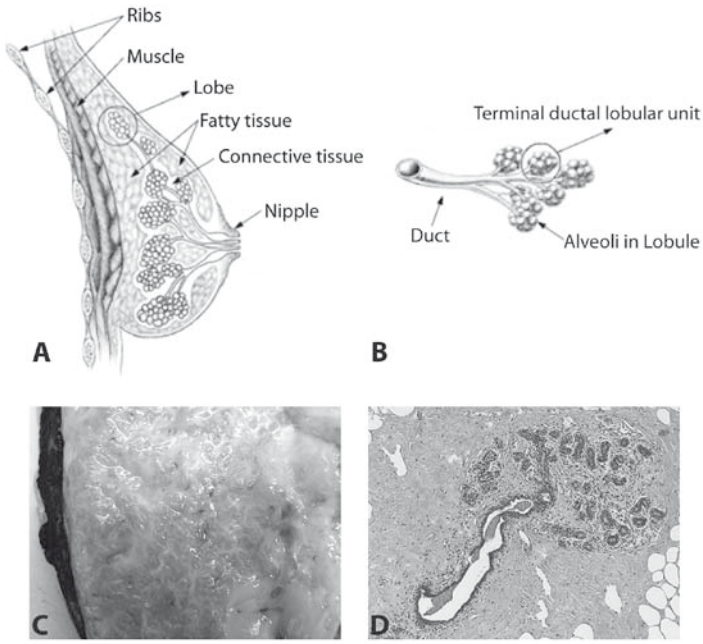


Figure 1.1 (A), schematic representation of the anatomy of the breast; (B), schematic representation of a breast lobe; (C), macroscopic view of the normal breast parenchyma; (D), microscopic view of a normal terminal ductular lobular unit of the breast. (A) and (B) were adapted from http://www.blogsforcompanies.com/TTimages/dcis_in_situ.jpg. This figure is also available in color in the appendix.

The luminal epithelial cells and the basal myoepithelial cells have distinct features and can be distinguished on the basis of expression of various cytokeratin (CK) proteins and smooth muscle actin (SMA). Luminal cytokeratins CK7, CK8, CK18, and CK19 are typically expressed in the luminal epithelial cells, but not in the basal myoepithelial cells. SMA and basal cytokeratins CK5, CK14, CK17, on the other hand, are expressed in the myoepithelial cells, but not in the luminal epithelial cells [48-50]. The two distinct epithelial cell lineages of a TDLU are clonally related and arise from a single mammary stem cell [51-57]. In addition, two types of luminal-restricted and one myoepithelial-restricted progenitor have been identified [58-60]. Still, the exact cellular hierarchy present in the breasts' epithelium is only partially understood.

1.2.2.2 HISTOPATHOLOGY OF BREAST CANCER

The vast majority of breast cancers arise in the epithelial cells of the TDLU, and are therefore classified as carcinoma. Breast carcinomas are classified pathologically on the basis of their morphology and growth pattern. The majority of the breast carcinomas (about 60%), however, can not be classified satisfactorily according to specialized pathological subtypes and are designated as invasive ductal carcinoma (IDC) not otherwise specified (NOS). Another common name for these tumors is IDC of no special type (NST). The "special type" pathological subtypes of breast carcinoma include invasive lobular carcinoma (ILC), medullary breast carcinoma (MC), mucinous breast carcinoma, tubular breast carcinoma, and metaplastic breast carcinoma (MBC). ILC is the most common of the special types, accounting for 10-15% of the breast cancers, and is thus the second most common pathological subtype after IDC NOS. The remainder of the special types of breast carcinoma each do not account for more than 5% of breast tumors, with metaplastic breast cancer being the rarest (less than 1% of breast carcinomas).

Invasive ductal carcinoma

Invasive ductal carcinoma not otherwise specified (IDC NOS) is a very heterogeneous group of tumors which includes all breast carcinomas that cannot be classified as a special pathological subtype (Figure 1.2 A and B). The tumors are classified as IDC mixed-type tumors when a special type component of more than 50% is present in addition to the IDC NOS component. In 80% of the IDC cases a precursor lesion of ductal carcinoma *in situ* (DCIS) is present, often of high grade comedo type. Although IDC NOS is generally considered to be a diverse group of breast carcinomas that can not be assigned to one of the currently-known specialized pathological subtypes, many breast pathologists would agree that there may be one or more specialized subtypes still to be defined in this subgroup of carcinomas.

Invasive lobular carcinoma

The classical pattern of invasive lobular carcinoma (ILC) is characterized by small rounded cells with scant cytoplasm, that diffusely grow through the stroma, often in strings of cells called "Indian files" (Figure 1.2 C; [12]). In most ILC cases, a lobular carcinoma *in situ* (LCIS) component is present, although DCIS has also been observed [61, 62]. In addition to classical ILC, other variants of ILC have been described including pleomorphic, alveolar, and solid lobular carcinoma [63-66]. Classical ILC tumors are frequently low grade tumors because of the morphologically uniform cells and a low mitotic index. Therefore, they have a more favorable prognosis than the ILC variants that have more marked nuclear pleomorphism and thus are of higher grade [61, 67, 68]. Complete loss of expression of the cell adhesion molecule E-cadherin has been observed for the majority (about 80%) of the ILCs [69-73], and has been associated with truncating mutations of the *E-cadherin* gene in 50% of ILC breast cancers [74-76].

Medullary carcinoma

Medullary carcinomas (MC) are poorly differentiated carcinomas with a syncitial growth pattern, absence of glandular structures, moderate to marked nuclear pleomorphism, complete histological circumscription of the tumor, and diffuse lymphocytic infiltrate (Figure 1.2 D; [12]). MCs are typically high grade tumors and are mostly ER-negative. However, the prognosis of MC is remarkably favorable and better than common IDC NOS, with 10-year survival rates of 50-90% depending on the criteria used. This probably is because less than 10% of the patients present with lymph node metastases [77-85]. Notably, 11% of MCs carry *BRCA1* germline mutations, which is about seven times more frequent than among breast cancers as a whole [86]. Reciprocally, 13-20-% of *BRCA1* mutant tumors are medullary carcinomas or carcinomas with medullary features [86-88].

Mucinous carcinoma

Mucinous or colloid carcinoma of the breast is characterized by clusters of small and uniform cells floating in a sea of extracellular mucin (Figure 1.2 E; [12]). DCIS is found to be present in 60-75% of mucinous carcinomas and may have any of the conventional patterns of DCIS (cribriform, comedo, papillary or micropapillary). Mucinous carcinomas are typically ER-positive and mostly also PR-positive [89, 90]. Mucinous carcinoma has a very good prognosis with ten-year survival rates of 80-100%, although patients with mixed variants of mucinous carcinoma tend to do worse [91-94]. The favorable prognosis probably is because pure mucinous carcinomas infrequently metastasize [90].

Tubular carcinoma

Tubular carcinomas are very well differentiated tumors that are characterized by single-layered open tubules, absence of necrosis and/or very few mitoses and minor nuclear pleomorphism (Figure 1.2 F; [12, 95]). Almost two third of these tumors have a low-grade DCIS component, usually

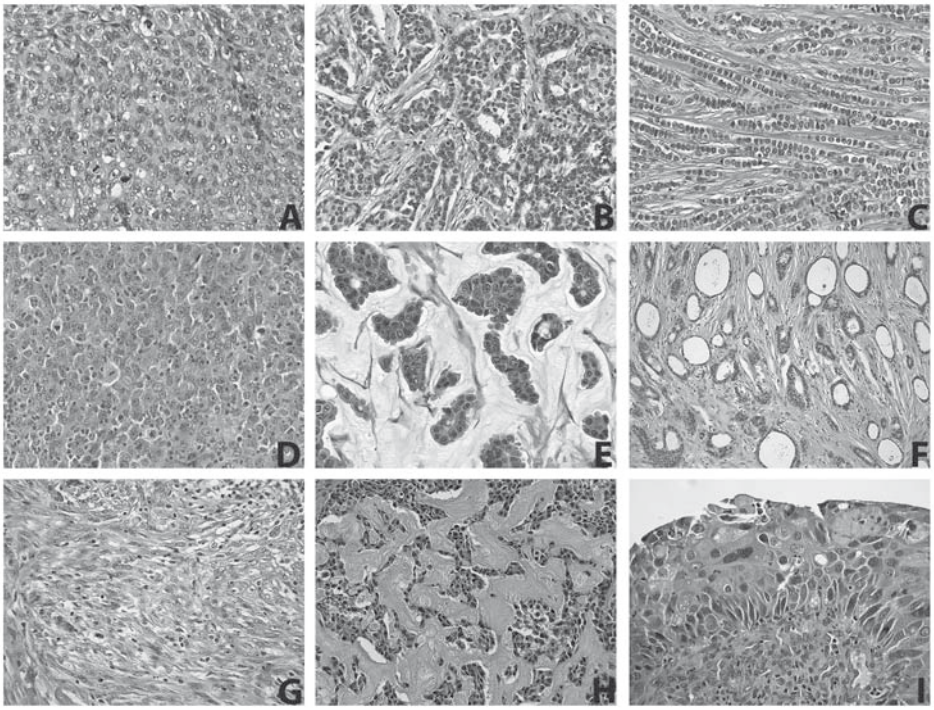


Figure 1.2 Microscopic views of histopathological subtypes of breast cancer. (A) and (B), ductal carcinoma of high and low grade, respectively; (C), lobular carcinoma with strings of cells called "Indian files"; (D), medullary carcinoma; (E), mucinous carcinoma; (F), tubular carcinoma; (G) through (I), metaplastic carcinoma of the breast with spindle, matrix-producing and squamous differentiation, respectively. This figure is also available in color in the appendix.

of cribriform or micropapillary type [96]. In comparison with IDC NOS, tubular carcinomas are more frequently ER and PR positive and EGFR and ERBB2 negative [90, 97, 98]. Similar to mucinous carcinomas, tubular carcinomas also have a particularly favorable prognosis. Although mixed variants do worse than the pure form, their prognosis still is better than IDC NOS [78, 95, 99-105]. Probably, this favorable prognosis is because tubular carcinomas tend to be smaller in size and present with less lymph node involvement at the time of diagnosis.

Metaplastic breast carcinoma

Metaplastic breast carcinoma (MBC) is a very heterogeneous group of carcinomas that are characterized by one or more prominent metaplastic components. These components can either be epithelial (squamous) or mesenchymal (matrix, spindle cell, osseous, chondroid and sarcomatous) (Figure 1.2 G through I; [12]). The mesenchymal components of MBCs frequently express vimentin and other mesenchymal markers, but in some cases epithelial characteristics are also retained, suggesting transdifferentiation rather than collision tumors [106, 107]. Importantly, genetic studies have shown that the different components in MBCs are indeed clonally

related [108-110]. MBCs are mostly ER, PR and ERBB2 negative and, as a whole, tend to have a less favorable prognosis than common IDC NOS [12]. Moreover, expression of the epidermal growth factor receptor (EGFR) is frequently observed in metaplastic breast cancers, suggesting that patients with this type of breast tumor may benefit from EGFR targeted therapy. [111].

1.2.3 GENE EXPRESSION PROFILING

Intrinsic subtypes of breast cancer

Global gene expression profiling of breast tumors has allowed additional classification of breast cancers. Large gene expression differences exist between ER-positive and ER-negative breast cancers and further molecular subclasses have been identified within these two groups [112-119]. Based on an 'intrinsic' gene set, consisting of genes that were least variably expressed between paired tumor samples from the same patient and most variably between tumors from different patients, five distinct intrinsic subtypes of breast cancer were defined. Importantly, these intrinsic subtypes were identified without supervision on the biology or clinical parameters of the tumors. The five intrinsic subtypes included the ER-negative basal-like, normal-like and ERBB2+ subtypes and the ER-positive luminal A and luminal B subtypes [115-117]. The basal-like and normal-like subtypes consist of tumors that have a high expression of the basal gene cluster (including *KRT5*, *KRT17*, *ANXA8*, *CX3CL1* and *TRIM29*) and a low expression of luminal gene cluster (including *ER*α, *GATA3*, *XBP1*, *TFF3*, *HNF3*α and *LIV1*). In addition, basal-like tumors have high expression of a novel set of genes whose coordinated function is not known (including *GGH*, *LAPTMB4*, *NSEP1* and *CCDE1*), whereas tumors of the normal-like subtype have high expression of the adipose and non-epithelial gene cluster (including *FACL2*, *AKR1C1*, *PIK3R1*). The ERBB2+ subtype is defined by tumors with high expression of genes from the *ERBB2* amplicon at chromosome 17q (including *ERBB2*, *GRB7* and *TRAP100*) and have low expression of the luminal gene cluster. Of the two luminal subtypes, tumors in the luminal A group have the highest expression of the luminal gene cluster, compared with moderate to low expression of these genes in luminal B tumors. Additionally, luminal B tumors have a relatively high expression of the novel set of genes whose coordinated function is not known, similar to basal-like tumors [116]. Importantly, these five intrinsic subtypes are conserved among different microarray platforms, different patient series and different races [120, 121].

Evidence is accumulating that the intrinsic subtypes have clinical significance (Figure 1.3). Patients with basal-like tumors had the worst overall survival, reflected by the abundance of triple negative (ER-negative, PR-negative and ERBB2-negative) tumors in this subtype, as did patients with tumors of the ERBB2+ subtype [116]. Tumors from these two intrinsic subtypes are also more sensitive to neoadjuvant chemotherapy than normal-like and luminal subtypes [114]. In addition, among the luminal subtype of tumors, the luminal B tumors had a less favorable outcome than luminal A tumors [112, 116]. The intrinsic subtypes of breast cancer also associated

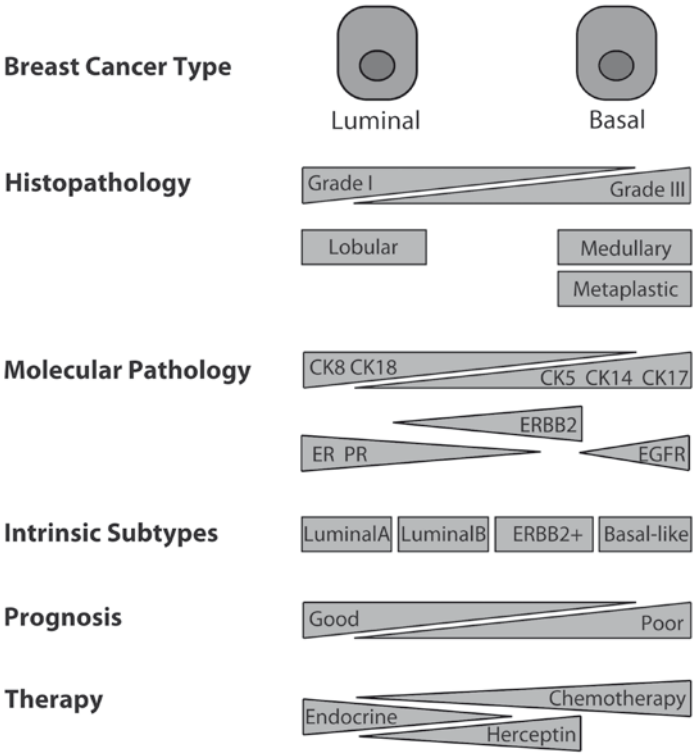


Figure 1.3 Presumed relations between breast cancer subtypes and their histopathological subtypes, expression of histological markers, intrinsic subtypes, prognosis and suggested therapy. Freely adapted from [29].

with different sites of distant metastases [122]. Breast tumors of the luminal intrinsic subtypes more frequently metastasized to bone and pleura, whereas tumors of the intrinsic basal-like and ERBB2+ subtypes more frequently metastasized to the brain. Lung metastases were most frequently observed among basal-like and luminal B breast tumors. Still, the relevance of the intrinsic gene subset lies not in its prognostic or predictive significance, but rather in its ability to capture breast cancer heterogeneity.

Prognostic gene expression profiles

Many gene expression profiles have been defined by taking clinical outcome of breast cancer patients into account. To reduce overtreatment of lymph node-negative patients with a low risk of developing metastases, a 70-gene prognostic signature was defined on tumors of lymph node-negative breast cancer patients younger than 55 years, most of whom had not received systemic treatment [123]. Partially independent validation of this gene signature showed that it was the strongest predictor of metastasis-free survival for lymph node-negative as well as lymph node-positive patients. In multivariate analysis, the 70-gene signature was independent

of the contribution of traditional prognostic factors [124]. For the same purpose, a 76-gene prognostic signature was established on tumors from lymph node- negative breast cancer patients of all age groups who had not received any adjuvant systemic treatment [125]. In contrast to the previous study, ER-positive and ER-negative tumors were analyzed separately. The 76-gene signature was a strong independent prognostic factor for metastasis-free and overall survival. Independent multicenter validation of this signature showed comparable hazard ratios to the original study [126, 127]. The 70-gene and 76-gene signatures both outperformed the NIH [128] and St. Gallen criteria [129], which are both classical clinical pathological prognostic indices, by reducing overtreatment in patients with a good prognosis [126, 130]. In addition, both signatures were as effective in selecting those high-risk patients who would be candidates for adjuvant systemic therapy as the NIH and St. Gallen criteria.

To predict metastasis in ER-positive breast cancer patients that had received adjuvant hormonal treatment, but had no local or distant metastases at the time of diagnosis, a 21-gene recurrence score (RS) was defined [131, 132]. According the RS, patients are classified in low, moderate or high-risk groups. Moreover, this signature predicted which patients in the NSABP B20 clinical trial would benefit from additional chemotherapy, showing that the prognostic signature also had relevant predictive value [133]. Currently, large prospective validation studies are underway the test the prognostic power of the 70-gene signature (MINDACT) and the 21-gene RS (TAILORx), which will provide definitive information on whether these signatures provide sufficient improvement in prognostic classification to be used for everyday clinical practice.

Prognostic gene expression signatures established by biological rather than prognostic criteria, include the genomic grade index (GGI) and the wound healing signature [134-136]. The 97-gene GGI signature was established on low versus high grade breast cancers and showed that grade II breast cancers are a mixture of grade I and III tumors rather than an intermediate grade. Importantly, the GGI signature had prognostic value for grade II breast cancers [134]. The 512-gene wound healing signature was established on the serum response of human fibroblasts but showed prognostic value in many human tumors, including breast cancer [135, 136]. The signature identified a subset of low-risk breast cancer patients among the presumed high-risk patient group and also outperformed the NIH and St. Gallen criteria in patients who had not received chemotherapy. Importantly, patients with both the wound healing signature and the poor prognosis 70-gene signature had a higher risk of metastases compared with patients without the wound healing signature and a poor prognosis 70-gene signature (47% versus 78% 10-year distant metastasis-free survival probability; [136]). This observation strongly indicates that combining various signatures may aid in risk stratification.

A valid concern regarding the various prognostic gene expression signatures was that they showed very little overlap. However, analysis of five distinct gene signatures on the same group of breast cancer patients showed a high concordance in risk stratification of patients among poor and good prognosis groups [137]. In addition, it has been shown that gene signatures that classify patients according the same clinical endpoint may include different genes but often

represent similar biological pathways, indicating that the differences among gene signatures most likely reflect differences in microarray platform and/or methodology [138].

Predictive gene expression profiles

In addition to the 21-gene recurrence score, numerous predictive gene signatures have been defined for response to hormonal treatment as well as to chemotherapy [139-147]. Patients with primary or metastatic ER-positive breast cancer are most frequently treated with the anti-estrogen tamoxifen. In the adjuvant setting, tamoxifen therapy results in a 5% and 13% improvement in survival in ER-positive lymph node-negative and lymph node-positive patients, respectively [148]. In the metastatic setting, approximately half of the patients with ER-positive tumors show intrinsic therapy resistance, while the other half shows an objective response to tamoxifen therapy. However, also almost all of the responding patients develop acquired therapy resistance at some time and eventually die of the disease. Therefore, reliable predictive factors are needed to predict the type of patients' response to tamoxifen. A 2-gene ratio of HOXB13 and IL17BR was claimed to predict the response to adjuvant tamoxifen treatment better than current clinical predictors [139]. However, this study did not have a control group of untreated patients, and the observations could have been the result of an association of the 2-gene ratio with prognosis, prediction, or both [149, 150]. Validation of this 2-gene ratio in a retrospective study then showed that it associated with both tumor aggressiveness and failure of tamoxifen treatment [149]. Similarly, a 44-gene predictor significantly correlated with the type of response to tamoxifen treatment for metastatic disease and predicted progression-free survival in multivariate analysis [140]. This tamoxifen profile was then validated in an independent series of ER-positive primary breast cancers and associated significantly with time to progression after adjusting for ER and PR [151]. For the prediction of response to chemotherapeutics, many gene expression profiles have been established. Two gene signatures have been established on tumors from patients who had received neoadjuvant docetaxel chemotherapy for primary breast cancer or locally advanced disease [141, 143]. Likewise, predictive gene signatures have been established for the response to neoadjuvant treatment with paclitaxel followed by 5-fluoracil, doxorubicin and cyclophosphamide [142, 144], neoadjuvant treatment with paclitaxel and doxorubicin [146], treatment with doxorubicin and cyclophosphamide [145] and neoadjuvant treatment with epirubicin and cyclophosphamide [147]. These predictive gene signatures are relevant for the understanding of therapy resistance and in defining the optimal treatment for breast cancer patients, thereby reducing unnecessary treatment and toxicity.

1.3 MOLECULAR GENETICS OF BREAST CANCER

Cancer is a genetic disease that involves accumulation of genetic alterations in multiple genes. These mutations enable the cell to replicate limitlessly, evade apoptosis, become insensitive

to anti-growth stimuli and self sufficient in growth signals and promote angiogenesis, invasion and metastasis [152]. Recent re-sequencing efforts of most human protein encoding genes in twenty-two breast and colorectal cancers suggested there may be as many as 15 somatic oncogenic driver mutations present in a single breast tumor [153, 154]. Importantly, the number of genes with oncogenic mutations was similar in breast cancers and colorectal cancers. Also, for both tumor types there were far more genes involved that had a low mutation frequency (gene hills) than genes with a high mutation frequency (gene mountains). However, colon cancers had more gene mountains than breast cancers [154]. This appears consistent with the fact that there is as yet no high prevalent breast cancer specific gene identified.

At least ten percent of all breast cancer cases have a family history of breast cancer. Depending on the number of affected first-degree relatives, a family history of breast cancer implies increased risk ratios of 1.5 to more than five-fold [155]. Only about 25% of the familial predisposition to breast cancer is explained by a germline mutation in one of the high-risk breast cancer susceptibility genes *BRCA1*, *BRCA2*, *p53* or *PTEN* or in the moderate-risk genes *CHEK2*, *ATM*, *BRIP1* and *PALB2* [156-163]. Recently, a genome-wide association study using single nucleotide polymorphisms identified five new breast cancer susceptibility loci, showing that some of the variation in breast cancer risk is due to common alleles [164]. As genetic linkage studies had failed to identify additional high-risk breast cancer genes, the existence of a polygenic model of breast cancer inheritance in which multiple low-risk genes act additive or multiplicative has gained much interest. Indeed, the currently known moderate-risk and low-risk susceptibility alleles all appear to operate in such a polygenic setting.

The E-cadherin/catenin complex

The *E-cadherin* tumor suppressor gene encodes a transmembrane glycoprotein that localizes to the adherence junctions of epithelial cells. Here, it mediates homophilic cell-cell adhesion between adjacent epithelial cells and thus integrity of epithelial tissues. This is established by interaction of the extracellular part of the E-cadherin molecule on one cell with the extracellular part of another E-cadherin molecule on an adjacent cell, resulting in a zipper-like structure. The C-terminal intracellular domain of E-cadherin binds to either β-catenin or γ-catenin (also known as plakoglobin), which are both proteins of the armadillo protein family and are mutually exclusive in the E-cadherin-catenin protein complex. The vinculin related α-catenin protein, in its turn, interacts with either the actin cytoskeleton or with β-catenin or γ-catenin in a dynamic fashion. The armadillo protein p120ctn also binds directly to the intracellular domain of E-cadherin, but more proximal to the cell membrane, and it stabilizes the complex [165-174].

E-cadherin reportedly is a suppressor of tumor invasion *in vitro* as well as *in vivo* and aberrant E-cadherin expression has been seen in many epithelial tumor types, including breast cancer [175-177]. The vast majority of breast cancers of the lobular subtype have lost E-cadherin protein expression, whereas in ductal breast cancers E-cadherin protein expression is mostly retained or only heterogeneously reduced [69, 70, 73, 178, 179]. Importantly, inactivating mutations of the

E-cadherin gene, located at 16q22, have only been identified in the lobular subtype of breast cancer, the diffuse subtype of gastric cancer and a small number of gynecological cancers [75, 76, 180-183]. About half of all lobular breast cancers have somatic mutations in the *E-cadherin* gene [75, 76]. In fact, mutations of *E-cadherin* in lobular breast cancer are already present in the premalignant carcinoma *in situ* stage, which makes it an early event in breast tumorigenesis [179]. The presence of *E-cadherin* somatic mutations in both lobular-type breast cancer and diffuse-type gastric cancer, which have a morphologically similar diffuse growth pattern of small rounded cells with scant cytoplasm, suggests that E-cadherin has a profound effect on cell morphology and may be causally involved in the observed characteristic histopathology. Indeed, conditional *E-cadherin* mutations in p53 knock-out mice resulted in breast carcinomas reminiscent of human invasive lobular breast cancer [184]. The causality of *E-cadherin* mutations for the lobular breast cancer phenotype makes one wonder about the absence of *E-cadherin* mutations in the other half of lobular breast cancers.

Germline mutations of *E-cadherin* have been identified in about 30% of families with predisposition to hereditary diffuse-type gastric cancer (HDGC) [185, 186]. The penetrance of autosomal dominant inherited *E-cadherin* mutations is very high and results in a cumulative life time risk of diffuse-type gastric cancer of 67% in men and 83% in women [187]. The presence of lobular breast cancers in families with HDGC suggested a role for *E-cadherin* germline mutations in families with a history of breast cancer [187-192]. However, to date only a single non-HDGC associated breast cancer family with an *E-cadherin* germline mutation has been identified [193]. Also, no *E-cadherin* germline mutations have been found in patients with LCIS, which is associated with an elevated familial breast cancer risk [194].

Also unexpected was the absence of *E-cadherin* gene mutations among carcinomas from other anatomical sites, or in the remaining breast cancers and gastric cancers. Loss of E-cadherin expression in these carcinomas was suggested to involve transcriptional silencing in association with methylation of CpG islands in the *E-cadherin* promoter region or with transcriptional repression [195-198]. The latter has been associated with expression of several transcriptional repressors of E-cadherin: Snail, Slug, SIP1, δEF1, E47 and Twist [198-205]. Specific chromatin remodeling complexes are recruited by some of these transcriptional repressors during tumorigenesis, suggesting that hypermethylation and transcriptional repression of E-cadherin is coupled [206]. The expression of E-cadherin's transcriptional repressors has been observed for various carcinoma types and has been associated with a more aggressive clinical course and with epithelial mesenchymal transition (EMT) [198]. EMT involves the conversion of polarized epithelial cells in motile cells with a mesenchymal phenotype which normally occurs during the gastrula stage of the development of an organism. In addition, this process has also been proposed to play a role in cancer metastasis [207, 208]. Loss of E-cadherin expression is considered to be one of the hallmarks of EMT, also involving (the crosstalk of) multiple pathways including the TGFβ, BMP, Wnt, RAS and PI3K pathways [207]. Interestingly, induction of EMT by ectopic expression of Snail or Twist in immortalized human mammary epithelial cells was shown to generate cells

with properties of stem cells, including acquisition of a mesenchymal phenotype, expression of stem cell markers and an increased ability to form mammospheres [209].

Besides its role in E-cadherin-mediated cell adhesion, β-catenin also plays a central role in the canonical Wnt (Wingless/INT-1) pathway. In the absence of Wnt signaling, the cytoplasmic pool of β-catenin is targeted for proteosomal degradation by a destruction complex that consists of the tumor suppressors APC en Axin and the kinases GSK3β and CK1. Activation of Wnt signaling inhibits the destruction complex. As a result, β-catenin is free to translocate to the nucleus where it interacts with TCF/LEF factors and facilitates transcription of TCF target genes [210]. In colorectal cancer, mutations of either APC, β-catenin or Axin-2 induce constitutive Wnt pathway activation, which is associated with a crypt stem cell or progenitor cell phenotype [211, 212]. The dual role of β-catenin in E-cadherin-mediated cell adhesion as well as Wnt signaling had led to the assumption that the Wnt pathway might be constitutively activated by the free β-catenin pool in E-cadherin deficient tumors [213]. However, breast cancer cell lines with either *E-cadherin* mutation or hypermethylation were shown not to have constitutive activation of the canonical Wnt signaling pathway, suggesting that aberrant activation of the canonical Wnt pathway is not of major importance in breast tumorigenesis [214].

The p53 signaling pathway

The transcription factor p53 is activated in response to DNA damage or hypoxia through phosphorylation, by among others, CHEK2 kinase. CHEK2 regulates the response to DNA damage by phosphorylating multiple substrates, including p53, CDC25C, CDC25A and BRCA1. Upon activation, p53 tetramerizes and is able to activate the transcription of downstream targets. This leads to either delayed cell cycle arrest at the G1-S cell cycle checkpoint until damage is repaired or sustained cell cycle arrest and apoptosis [215-218]. p21 is an important downstream transcription target of p53 and accumulates to levels capable of inhibiting Cyclin E and CDK2, which in turn promote progression through the G1/S checkpoint [219]. HDM2 is another protein induced by p53, which antagonizes the p53 response by binding p53 and targeting it for ubiquitination and degradation [220]. However, p14ARF is able to inhibit the interaction of HDM2 with p53 by binding HDM2, thereby stabilizing the p53 protein [221]. Cells with mutant *p53* are not capable of G1-S cell cycle arrest and its associated apoptosis, resulting in replication of damaged DNA and thus the accumulation of genetic alterations. Alternative ways of inactivating p53 are by overexpression of HDM2 or by inactivation of CHEK2 or p14ARF [221].

Mutations of the *p53* tumor suppressor gene, located on 17p13, are found in virtually every cancer type, stressing the pivotal role of p53 in different cell types and carcinogenesis. About 30-40% of clinical breast cancers have a mutation in *p53*, making *p53* the most frequently mutated gene in human breast cancer [222, 223]. Aberrant expression of p53 associates with breast cancers of the ductal subtype, a higher grade and poorer prognosis [224, 225]. In familial breast cancer patients, *p53* germline mutations are present in less than 1% of the cases [223]. Notably, germline mutations of *p53* have only been identified in families with the rare Li-Fraumeni

syndrome, that is characterized by an increased risk of breast cancer, sarcomas, brain tumors, leukemia's and adrenal tumors [226, 227].

CHEK2, located on chromosome 22q12, was first identified in association with wild-type *p53* Li-Fraumeni kindreds, suggesting that germline *CHEK2* mutations are an alternative genetic defect of Li-Fraumeni syndrome [228, 229]. As *CHEK2* founder mutations also appeared to be present at low frequency in healthy individuals it became clear that *CHEK2* could not be a susceptibility gene for Li-Fraumeni syndrome. In fact, the *CHEK2* 1100delC variant was shown to be significantly more frequent in breast cancer families than in healthy controls and it confers a modest two-fold increased breast cancer risk. *CHEK2* 1100delC thus was the first moderate-risk breast cancer susceptibility allele identified. [159, 160]. In addition, *CHEK2* 1100delC was associated with a hereditary breast and colon cancer (HBCC) phenotype, an elevated risk of bilateral breast cancer and an elevated risk for male breast cancer [159, 160, 230, 231]. *CHEK2* 1100delC breast cancers are mostly ER-positive tumors of the luminal intrinsic subtype and are of a higher grade than tumors from non-*CHEK2* 1100delC carriers (Nagel et al, submitted for publication; [232, 233]).

Amplification or overexpression of *HDM2*, located at chromosome 12q13, and mutations and deletions in *p14ARF*, located on chromosome 9q21, reportedly are uncommon in breast cancer, although hypermethylation of *p14ARF* is observed in one quarter of human breast cancers [234-236]. However, in contrast to some human cancers, inactivation of *p14ARF* frequently coincided with mutation of *p53* in breast cancer [236]. This suggests that in breast cancer *p14ARF* inactivation is not biologically similar to *p53* mutation.

The PI3K signaling pathway

The phosphatidylinositol 3-OH kinase (PI3K) pathway is pivotal for the regulation of cellular processes, including growth, proliferation and survival of cells. The main players in this pathway are PIK3CA, PTEN and the three AKT proteins. PIK3CA or p110α is released from its inhibitor p85α upon PI3K pathway activation, usually by signaling through receptor tyrosine kinases (RTKs) on the plasma membrane. Once activated, PIK3CA is able to convert phosphatidylinositol-4,5-diphosphate (PIP2) to its active form, phosphatidylinositol-3,4,5-triphosphate (PIP3). Upon production of PIP3, AKT translocates to the plasma membrane and becomes activated by phosphorylation at Thr308 and Ser473 by PDK1 and PDK2. The conversion from PIP2 to PIP3 is counteracted by the lipid phosphatase PTEN, thereby blocking the activation of AKT [237]. As constitutive signaling through the PI3K pathway is oncogenic, *PIK3CA* and *AKT* are oncogenes and *PTEN* is a tumor suppressor gene by virtue of their roles in the pathway.

PIK3CA, located at chromosome 3q26, was identified in a mutation screen of PI3K and PI3K-like genes [238] and has since been found mutated predominantly in liver, colon and breast tumors [239]. In breast cancer, *PIK3CA* is the second most frequently mutated gene after *p53*, and has been found mutated in 20-40% of breast tumors [239]. *PIK3CA* mutations have been associated with ER and PR positive tumors and tumors with ERBB2 overexpression. Importantly, mutations

of *PIK3CA* were found to be mutually exclusive with loss of PTEN expression, which is consistent with their opposing function in PI3K signaling [240].

The *PTEN* tumor suppressor gene, located on chromosome 10q23, was identified in a screen of breast cancers and glioblastomas but was found to be mutated less frequently in breast cancer (10-20%) than *PIK3CA* [241-243]. Germline mutations of *PTEN* are found in patients with Cowden disease and Bannayan-Zonana syndrome, predisposing to hamartomatous lesions and conferring an increased risk of cancer [244, 245]. Recently, loss of PTEN expression had been associated with sporadic basal-like breast cancer. In addition, large structural mutations of the *PTEN* locus, in contrast to small intragenic sequence alterations of *PTEN*, associated with *BRCA1* mutations, suggesting a role for PTEN in both sporadic as well as hereditary basal-like breast cancers [246]. In addition, a PTEN gene signature obtained from PTEN negative tumors has been associated with metastasis and poor survival of breast cancer patients [247].

AKT has three homologues in mammals of which AKT1 and AKT2 have been implicated in breast cancer. Increased AKT1 kinase activity has been observed in about 40% of human breast tumors and recently an oncogenic mutation in *AKT1* (E17K), located on chromosome 14q32, was identified in 8% of breast tumors [248, 249]. *AKT2*, located on chromosome 19q13, has been found amplified in only a minority of human breast cancers (<5%), however increased AKT2 kinase activity is present in 40% of human breast cancers [250, 251]. These data combined suggest mutational activation of the PI3K pathway in 40-70% of human breast cancers.

The RAS signaling pathway

The RAS pathway is activated by stimulation of receptor tyrosine kinases and regulates proliferation and differentiation of cells, as well as cytoskeletal rearrangements. RAS proteins are activated by guanine nucleotide exchange factors (GNEFs) that convert GDP-bound RAS to the GTP-bound state. Once activated, RAS can signal through multiple pathways, including the MAPK pathway via the kinases RAF, MEK and ERK, and the PI3K pathway via the kinase p110. Inactivation of RAS is regulated by GTPase activating proteins (GAPs) of which the most relevant for tumorigenesis is the neurofibromin1 (NF1) protein. Mutational inactivation of the *NF1* tumor suppressor gene leads to accumulation of activated GTP-bound RAS and ultimately to tumorigenesis in tissues derived from the neural crest and myeloid malignancies [252]. Germline mutations of *NF1* cause neurofibromatosis type 1, characterized by neurofibromas, iris hamartomas, café-au-lait spots and an increased risk of developing cancers, including brain tumors and leukemia's [253].

Three isoforms of human *RAS* genes exist that are implicated in human cancer: *KRAS*, *HRAS* and *NRAS*, located on chromosomes 12p12, 11p15 and 1p13 respectively. Point mutations of codons 12, 13 and 61 of the *RAS* genes have been identified in a wide variety of human tumor types, resulting in constitutively activated RAS and its downstream pathway [252, 254]. For breast cancer however, mutations in the *RAS* genes were infrequently identified (less than 10%). This

is in contrast to colon and pancreatic cancers that have mutated *KRAS* genes in 40-50% and 90-95% of the tumors, respectively [255-257].

Oncogenic hotspot mutations in *BRAF*, located on chromosome 7q34, lead to increased kinase activity, providing an alternative route to RAS pathway activation [258]. This is nicely illustrated by the mutual exclusiveness of *KRAS* and *BRAF* mutations in colorectal cancers [259]. Mutations in *BRAF* are most frequent in melanomas and colorectal tumors, but infrequent in human breast tumors (~5%, [258]). There was no evidence for mutational activation of *ARAF* and *CRAF*, the other two *RAF* isoforms in cancer [260, 261]. Interestingly, germline mutations of the RAS/MAPK pathway cause two clinically overlapping syndromes characterized by heart defects, mental retardation and distinctive facial appearances. Costello syndrome is caused by germline mutations in the *HRAS* gene, whereas germline mutations in *KRAS*, *BRAF*, *MEK1* and *MEK2* cause cardiofaciocutaneous syndrome [262-264]. Importantly, activation of the RAS pathway appears to be an infrequent (< 15%) event in breast cancer.

The RB signaling pathway

The retinoblastoma (RB) pathway is important in regulating the G1 to S-phase transition of the cell cycle. Mitogenic stimulation during the G1 phase accumulates complexes of Cyclin D1 with either CDK4 or CDK6 cyclin-dependent kinases in the cell nucleus. These complexes, together with Cyclin E/CDK2 complexes phosphorylate the RB1 gene product, inducing the subsequent release of RB1 from the E2F transcription factor. Release of RB1 from E2F initiates E2F-dependent transcription of genes, which is necessary for DNA replication and S-phase entry. Actions of the Cyclin D1/CDK complexes on RB1 are inhibited by the p16 protein. Therefore, loss of either p16 or the RB1 tumor suppressors, as well as amplification and overexpression of Cyclin D1 or CDK4 have been implicated in human tumorigenesis [265].

Germline mutations of *p16*, located on chromosome 9q21, predispose an individual to familial multiple melanoma [266, 267]. Loss of p16 is observed in about 30% of human breast cancers and occurs mainly through promoter hypermethylation, not mutation or deletion of the gene [268-271]. This is in contrast to many other tumor types that have inactivated *p16* through either deletions or point mutations, including 80% of pancreatic cancers [272-275].

Mutations of the *RB1* gene, located on chromosome 13q14, were first identified in the germline of patients with retinoblastoma, a rare childhood eye tumor. In time, somatic mutations of the gene were found in various tumor types, including osteosarcomas, small cell lung cancers and 10% of breast cancers [276-281]. In breast cancer, loss of RB1 protein expression is associated with ER-negativity and grade III ductal tumors [282].

Amplification and translocation of *Cyclin D1*, located on chromosome 11q13, has been found in various human tumors [279]. In breast cancer, *Cyclin D1* is amplified in 25% of the tumors and overexpressed in half of the tumors. Also, overexpression of Cyclin D1 is significantly associated with ER-positive breast cancers [283, 284].

Amplification of the *CDK4* gene, located on chromosome 12q13-14, is found frequently in sarcomas and gliomas and in about 15% of human breast tumors [279, 285]. An alternative mechanism of CDK4 activation is mutation of the p16 binding domain in the *CDK4* gene which is found incidentally in sporadic melanoma (less than 5%) and in melanoma-prone families (six families reported to date) [286-289].

Cyclin E and the CDK inhibitor p27kip are not known to be mutationally involved in breast cancer, but their expression (Cyclin E high/p27kip low) alone or together is associated with worse outcome of breast cancer patients [290, 291]. Overall, mutational inactivation of the RB pathway has been implicated in more than 80% of human breast cancers [292].

The BRCA1 and BRCA2 genes

BRCA1 and BRCA2 are caretakers of preserving genomic stability through their role in DNA damage signaling and repair, particularly of double strand breaks (DSBs). Both BRCA1 and BRCA2 deficient cells are unable to repair DSBs by error-free homologous replication (HR). This results in rerouting of repair by the still intact but error-prone non-homologous end-joining (NHEJ) pathway. BRCA1, however, appears to have a much broader role in DNA damage repair than BRCA2. In addition to its involvement in the BRCA1, BRCA2 and Rad51 complex, BRCA1 is also part of the BRCA1-associated genome-surveillance complex (BASC), which includes ATM, RAD50, MRE11 and NBS1 and the mismatch repair proteins MLH1, PMS2, MSH2 and MSH6. ATM, like ATR, CHEK1 and CHEK2 are able to phosphorylate BRCA1 in response to different types of DNA damage. BRCA1 is also implicated in transcription-coupled nucleotide excision repair, chromatin remodeling at DNA damage sites and ubiquitination via the BRCA1/BARD1 heterodimeric complex [293-297].

Germline mutations of the *BRCA1* gene, located on chromosome 17q21, and the *BRCA2* gene, located on chromosome 13q12, predispose a woman to an average cumulative breast cancer risk by age 70 years of 65% and 45%, respectively, when she was unselected for family history [298]. These breast cancer risk estimates are even higher in women from families with multiple breast cancer cases, as their risk may be modified by other genes. Indeed, it was shown that common breast cancer susceptibility alleles may act multiplicatively on the breast cancer risk in *BRCA1* and *BRCA2* mutation carriers [299]. Besides breast cancer, women with a *BRCA1* or *BRCA2* germline mutation have an increased risk of ovarian cancer and men have an increased risk of prostate cancer. In addition, *BRCA2* mutation carriers also have an increased risk of other cancer types such as breast cancer in men, pancreas cancer, gall bladder cancer, bile duct cancer, stomach cancer and melanoma [300]. Despite similarities in function and disease spectrum, *BRCA1* and *BRCA2* mutant tumors are very distinctive. *BRCA1* mutant tumors are frequently ER-negative, of the basal-like intrinsic subtype and more frequently of medullary histology than sporadic breast tumors [88, 117]. *BRCA2* mutant tumors, in contrast, are more frequently ER-positive. However, both *BRCA1* and *BRCA2* mutant tumors have a higher frequency of *p53* mutations than sporadic breast cancers and a higher degree of aneuploidy [87, 301]. Somatic *BRCA1* or *BRCA2*

mutations are rare, although reduced BRCA1 expression has been observed in sporadic breast cancers [302]. In these tumors, hypermethylation of CpG islands in the *BRCA1* promoter region has been observed [303, 304].

REFERENCES

1. Visser O, van Noord KJ, eds. Feiten en fabels over kanker in Nederland. Utrecht: Vereniging van Integrale Kankercentra; 2005.
2. van Dijk JAAM, Coebergh JWW, Siesling S, Visser O, eds. Trends of cancer in the Netherlands 1989-1998. Utrecht: Vereniging van Integrale Kankercentra; 2002.
3. Louwman WJ, Voogd AC, van Dijck JA, et al. On the rising trends of incidence and prognosis for breast cancer patients diagnosed 1975-2004: a long-term population-based study in southeastern Netherlands. Cancer Causes Control 2008;19(1):97-106.
4. Parkin DM, Pisani P, Ferlay J. Global cancer statistics. CA: a cancer journal for clinicians 1999;49(1):33-64, 1.
5. Parkin DM, Whelan SL, Ferlay J, Teppo L, Thomas DB, eds. Cancer Incidence in Five Continents Vol. VIII. Lyon, France: International Agency for Research on Cancer; 2002.
6. Colditz GA. Epidemiology of breast cancer. Findings from the nurses' health study. Cancer 1993;71(4 Suppl):1480-9.
7. Colditz GA, Willett WC, Hunter DJ, et al. Family history, age, and risk of breast cancer. Prospective data from the Nurses' Health Study. Jama 1993;270(3):338-43.
8. Clemons M, Goss P. Estrogen and the risk of breast cancer. The New England journal of medicine 2001;344(4):276-85.
9. Visser O, Siesling S, van Dijk JAAM, eds. Incidence of cancer in the Netherlands 1999/2000. Utrecht: Vereniging van Integrale Kankercentra; 2003.
10. Greene FL, Balch CM, Page DL, eds. American Joint Committee on Cancer (AJCC) Cancer Staging Manual. 6th edition. New York: Springer-Verlag; 2002.
11. Olivotto IA, Chua B, Allan SJ, Speers CH, Chia S, Ragaz J. Long-term survival of patients with supraclavicular metastases at diagnosis of breast cancer. J Clin Oncol 2003;21(5):851-4.
12. Tavassoli FA, Devilee P, eds. Pathology and genetics of tumours of the breast and female genital organs. Lyon: IARC press; 2003.
13. Nemoto T, Vana J, Bedwani RN, Baker HW, McGregor FH, Murphy GP. Management and survival of female breast cancer: results of a national survey by the American College of Surgeons. Cancer 1980;45(12):2917-24.
14. Fitzgibbons PL, Page DL, Weaver D, et al. Prognostic factors in breast cancer. College of American Pathologists Consensus Statement 1999. Archives of pathology & laboratory medicine 2000;124(7):966-78.
15. Weiss RB, Woolf SH, Demakos E, et al. Natural history of more than 20 years of node-positive primary breast carcinoma treated with cyclophosphamide, methotrexate, and fluorouracil-based adjuvant chemotherapy: a study by the Cancer and Leukemia Group B. J Clin Oncol 2003;21(9):1825-35.
16. Giuliano AE, Dale PS, Turner RR, Morton DL, Evans SW, Krasne DL. Improved axillary staging of breast cancer with sentinel lymphadenectomy. Annals of surgery 1995;222(3):394-9; discussion 9-401.
17. Allred DC, Elledge RM. Caution concerning micrometastatic breast carcinoma in sentinel lymph nodes. Cancer 1999;86(6):905-7.
18. Arriagada R, Le MG, Dunant A, Tubiana M, Contesso G. Twenty-five years of follow-up in patients with operable breast carcinoma: correlation between clinicopathologic factors and the risk of death in each 5-year period. Cancer 2006;106(4):743-50.
19. Warwick J, Tabar L, Vitak B, Duffy SW. Time-dependent effects on survival in breast carcinoma: results of 20 years of follow-up from the Swedish Two-County Study. Cancer 2004;100(7):1331-6.
20. Carter CL, Allen C, Henson DE. Relation of tumor size, lymph node status, and survival in 24,740 breast cancer cases. Cancer 1989;63(1):181-7.
21. Rosen PP, Groshen S, Kinne DW, Norton L. Factors influencing prognosis in node-negative breast carcinoma: analysis of 767 T1N0M0/T2N0M0 patients with long-term follow-up. J Clin Oncol 1993;11(11):2090-100.

22. Quiet CA, Ferguson DJ, Weichselbaum RR, Hellman S. Natural history of node-negative breast cancer: a study of 826 patients with long-term follow-up. J Clin Oncol 1995;13(5):1144-51.
23. Chia SK, Speers CH, Bryce CJ, Hayes MM, Olivotto IA. Ten-year outcomes in a population-based cohort of node-negative, lymphatic, and vascular invasion-negative early breast cancers without adjuvant systemic therapies. J Clin Oncol 2004;22(9):1630-7.
24. Bloom HJ, Richardson WW. Histological grading and prognosis in breast cancer; a study of 1409 cases of which 359 have been followed for 15 years. British journal of cancer 1957;11(3):359-77.
25. Elston CW, Ellis IO. 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 1991;19(5):403-10.
26. Reed W, Hannisdal E, Boehler PJ, Gundersen S, Host H, Marthin J. The prognostic value of p53 and c-erb B-2 immunostaining is overrated for patients with lymph node negative breast carcinoma: a multivariate analysis of prognostic factors in 613 patients with a follow-up of 14-30 years. Cancer 2000;88(4):804-13.
27. Frkovic-Grazio S, Bracko M. Long term prognostic value of Nottingham histological grade and its components in early (pT1N0M0) breast carcinoma. Journal of clinical pathology 2002;55(2):88-92.
28. Page DL, Jensen RA, Simpson JF. Routinely available indicators of prognosis in breast cancer. Breast cancer research and treatment 1998;51(3):195-208.
29. Sims AH, Howell A, Howell SJ, Clarke RB. Origins of breast cancer subtypes and therapeutic implications. Nature clinical practice 2007;4(9):516-25.
30. Roylance R, Gorman P, Harris W, et al. Comparative genomic hybridization of breast tumors stratified by histological grade reveals new insights into the biological progression of breast cancer. Cancer research 1999;59(7):1433-6.
31. Jonat W, Pritchard KI, Sainsbury R, Klijn JG. Trends in endocrine therapy and chemotherapy for early breast cancer: a focus on the premenopausal patient. Journal of cancer research and clinical oncology 2006;132(5):275-86.
32. Hilsenbeck SG, Ravdin PM, de Moor CA, Chamness GC, Osborne CK, Clark GM. Time-dependence of hazard ratios for prognostic factors in primary breast cancer. Breast cancer research and treatment 1998;52(1-3):227-37.
33. Alexieva-Figusch J, Van Putten WL, Blankenstein MA, Blonk-Van Der Wijst J, Klijn JG. The prognostic value and relationships of patient characteristics, estrogen and progestin receptors, and site of relapse in primary breast cancer. Cancer 1988;61(4):758-68.
34. Osborne CK. Steroid hormone receptors in breast cancer management. Breast cancer research and treatment 1998;51(3):227-38.
35. Hess KR, Pusztai L, Buzdar AU, Hortobagyi GN. Estrogen receptors and distinct patterns of breast cancer relapse. Breast cancer research and treatment 2003;78(1):105-18.
36. Steroid receptors in breast cancer: an NIH Consensus Development Conference, Bethesda, Maryland, June 27-29, 1979. Cancer 1980;46(12 Suppl):2759-963.
37. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science (New York, NY 1987;235(4785):177-82.
38. Slamon DJ, Godolphin W, Jones LA, et al. Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. Science (New York, NY 1989;244(4905):707-12.
39. Berns EM, Klijn JG, van Staveren IL, Portengen H, Noordegraaf E, Foekens JA. Prevalence of amplification of the oncogenes c-myc, HER2/neu, and int-2 in one thousand human breast tumours: correlation with steroid receptors. Eur J Cancer 1992;28(2-3):697-700.
40. Menard S, Balsari A, Casalini P, et al. HER-2-positive breast carcinomas as a particular subset with peculiar clinical behaviors. Clin Cancer Res 2002;8(2):520-5.
41. Pritchard KI, Shepherd LE, O'Malley FP, et al. HER2 and responsiveness of breast cancer to adjuvant chemotherapy. The New England journal of medicine 2006;354(20):2103-11.

42. Ferrero-Pous M, Hacene K, Bouchet C, Le Doussal V, Tubiana-Hulin M, Spyratos F. Relationship between c-erbB-2 and other tumor characteristics in breast cancer prognosis. Clin Cancer Res 2000;6(12):4745-54.
43. Cobleigh MA, Vogel CL, Tripathy D, et al. Multinational study of the efficacy and safety of humanized anti-HER2 monoclonal antibody in women who have HER2-overexpressing metastatic breast cancer that has progressed after chemotherapy for metastatic disease. J Clin Oncol 1999;17(9):2639-48.
44. Slamon DJ, Leyland-Jones B, Shak S, et al. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. The New England journal of medicine 2001;344(11):783-92.
45. Vogel CL, Cobleigh MA, Tripathy D, et al. Efficacy and safety of trastuzumab as a single agent in first-line treatment of HER2-overexpressing metastatic breast cancer. J Clin Oncol 2002;20(3):719-26.
46. Berns K, Horlings HM, Hennessy BT, et al. A functional genetic approach identifies the PI3K pathway as a major determinant of trastuzumab resistance in breast cancer. Cancer cell 2007;12(4):395-402.
47. Rosai J. Breast. In: Surgical Pathology, 9th edition: Mosby; 2004:1763-876.
48. Taylor-papdimitriou J, Lane EB. Keratin expression in the mammary gland. In: Daniel CW, ed. The mammary gland: development, regulation, and function. New York: Plenum press; 1987:181-215.
49. Debus E, Weber K, Osborn M. Monoclonal cytokeratin antibodies that distinguish simple from stratified squamous epithelia: characterization on human tissues. The EMBO journal 1982;1(12):1641-7.
50. Nagle RB, Bocker W, Davis JR, et al. Characterization of breast carcinomas by two monoclonal antibodies distinguishing myoepithelial from luminal epithelial cells. J Histochem Cytochem 1986;34(7):869-81.
51. Deng G, Lu Y, Zlotnikov G, Thor AD, Smith HS. Loss of heterozygosity in normal tissue adjacent to breast carcinomas. Science (New York, NY 1996;274(5295):2057-9.
52. Tsai YC, Lu Y, Nichols PW, Zlotnikov G, Jones PA, Smith HS. Contiguous patches of normal human mammary epithelium derived from a single stem cell: implications for breast carcinogenesis. Cancer research 1996;56(2):402-4.
53. Diallo R, Schaefer KL, Poremba C, et al. Monoclonality in normal epithelium and in hyperplastic and neoplastic lesions of the breast. The Journal of pathology 2001;193(1):27-32.
54. Deome KB, Faulkin LJ, Jr., Bern HA, Blair PB. Development of mammary tumors from hyperplastic alveolar nodules transplanted into gland-free mammary fat pads of female C3H mice. Cancer research 1959;19(5):515-20.
55. Kordon EC, Smith GH. An entire functional mammary gland may comprise the progeny from a single cell. Development (Cambridge, England) 1998;125(10):1921-30.
56. Shackleton M, Vaillant F, Simpson KJ, et al. Generation of a functional mammary gland from a single stem cell. Nature 2006;439(7072):84-8.
57. Stingl J, Eirew P, Ricketson I, et al. Purification and unique properties of mammary epithelial stem cells. Nature 2006;439(7079):993-7.
58. Stingl J, Eaves CJ, Kuusk U, Emerman JT. Phenotypic and functional characterization in vitro of a multipotent epithelial cell present in the normal adult human breast. Differentiation; research in biological diversity 1998;63(4):201-13.
59. Stingl J, Eaves CJ, Zandieh I, Emerman JT. Characterization of bipotent mammary epithelial progenitor cells in normal adult human breast tissue. Breast cancer research and treatment 2001;67(2):93-109.
60. Villadsen R, Fridriksdottir AJ, Ronnov-Jessen L, et al. Evidence for a stem cell hierarchy in the adult human breast. J Cell Biol 2007;177(1):87-101.
61. Dixon JM, Anderson TJ, Page DL, Lee D, Duffy SW. Infiltrating lobular carcinoma of the breast. Histopathology 1982;6(2):149-61.
62. Newman W. Lobular carcinoma of the female breast. Report of 73 cases. Annals of surgery 1966;164(2):305-14.
63. Eusebi V, Magalhaes F, Azzopardi JG. Pleomorphic lobular carcinoma of the breast: an aggressive tumor showing apocrine differentiation. Human pathology 1992;23(6):655-62.

64. Weidner N, Semple JP. Pleomorphic variant of invasive lobular carcinoma of the breast. Human pathology 1992;23(10):1167-71.
65. Shousha S, Backhous CM, Alaghband-Zadeh J, Burn I. Alveolar variant of invasive lobular carcinoma of the breast. A tumor rich in estrogen receptors. American journal of clinical pathology 1986;85(1):1-5.
66. Fechner RE. Histologic variants of infiltrating lobular carcinoma of the breast. Human pathology 1975;6(3):373-8.
67. DiCostanzo D, Rosen PP, Gareen I, Franklin S, Lesser M. Prognosis in infiltrating lobular carcinoma. An analysis of "classical" and variant tumors. The American journal of surgical pathology 1990;14(1):12-23.
68. du Toit RS, Locker AP, Ellis IO, Elston CW, Nicholson RI, Blamey RW. Invasive lobular carcinomas of the breast--the prognosis of histopathological subtypes. British journal of cancer 1989;60(4):605-9.
69. Gamallo C, Palacios J, Suarez A, et al. Correlation of E-cadherin expression with differentiation grade and histological type in breast carcinoma. The American journal of pathology 1993;142(4):987-93.
70. Moll R, Mitze M, Frixen UH, Birchmeier W. Differential loss of E-cadherin expression in infiltrating ductal and lobular breast carcinomas. The American journal of pathology 1993;143(6):1731-42.
71. Oka H, Shiozaki H, Kobayashi K, et al. Expression of E-cadherin cell adhesion molecules in human breast cancer tissues and its relationship to metastasis. Cancer research 1993;53(7):1696-701.
72. Palacios J, Benito N, Pizarro A, et al. Anomalous expression of P-cadherin in breast carcinoma. Correlation with E-cadherin expression and pathological features. The American journal of pathology 1995;146(3):605-12.
73. Rasbridge SA, Gillett CE, Sampson SA, Walsh FS, Millis RR. Epithelial (E-) and placental (P-) cadherin cell adhesion molecule expression in breast carcinoma. The Journal of pathology 1993;169(2):245-50.
74. Kanai Y, Oda T, Tsuda H, Ochiai A, Hirohashi S. Point mutation of the E-cadherin gene in invasive lobular carcinoma of the breast. Jpn J Cancer Res 1994;85(10):1035-9.
75. Berx G, Cleton-Jansen AM, Nollet F, et al. E-cadherin is a tumour/invasion suppressor gene mutated in human lobular breast cancers. The EMBO journal 1995;14(24):6107-15.
76. Berx G, Cleton-Jansen AM, Strumane K, et al. E-cadherin is inactivated in a majority of invasive human lobular breast cancers by truncation mutations throughout its extracellular domain. Oncogene 1996;13(9):1919-25.
77. Black CL, Morris DM, Goldman LI, McDonald JC. The significance of lymph node involvement in patients with medullary carcinoma of the breast. Surgery, gynecology & obstetrics 1983;157(6):497-9.
78. Ellis IO, Galea M, Broughton N, Locker A, Blamey RW, Elston CW. Pathological prognostic factors in breast cancer. II. Histological type. Relationship with survival in a large study with long-term follow-up. Histopathology 1992;20(6):479-89.
79. Maier WP, Rosemond GP, Goldman LI, Kaplan GF, Tyson RR. A ten year study of medullary carcinoma of the breast. Surgery, gynecology & obstetrics 1977;144(5):695-8.
80. Jensen ML, Kiaer H, Andersen J, Jensen V, Melsen F. Prognostic comparison of three classifications for medullary carcinomas of the breast. Histopathology 1997;30(6):523-32.
81. Rapin V, Contesso G, Mouriesse H, et al. Medullary breast carcinoma. A reevaluation of 95 cases of breast cancer with inflammatory stroma. Cancer 1988;61(12):2503-10.
82. Wargotz ES, Silverberg SG. Medullary carcinoma of the breast: a clinicopathologic study with appraisal of current diagnostic criteria. Human pathology 1988;19(11):1340-6.
83. Reinfuss M, Stelmach A, Mitus J, Rys J, Duda K. Typical medullary carcinoma of the breast: a clinical and pathological analysis of 52 cases. Journal of surgical oncology 1995;60(2):89-94.
84. Richardson WW. Medullary carcinoma of the breast; a distinctive tumour type with a relatively good prognosis following radical mastectomy. British journal of cancer 1956;10(3):415-23.
85. Ridolfi RL, Rosen PP, Port A, Kinne D, Mike V. Medullary carcinoma of the breast: a clinicopathologic study with 10 year follow-up. Cancer 1977;40(4):1365-85.

86. Eisinger F, Jacquemier J, Charpin C, et al. Mutations at BRCA1: the medullary breast carcinoma revisited. Cancer research 1998;58(8):1588-92.
87. Marcus JN, Watson P, Page DL, et al. Hereditary breast cancer: pathobiology, prognosis, and BRCA1 and BRCA2 gene linkage. Cancer 1996;77(4):697-709.
88. Pathology of familial breast cancer: differences between breast cancers in carriers of BRCA1 or BRCA2 mutations and sporadic cases. Breast Cancer Linkage Consortium. Lancet 1997;349(9064):1505-10.
89. Shousha S, Coady AT, Stamp T, James KR, Alaghband-Zadeh J. Oestrogen receptors in mucinous carcinoma of the breast: an immunohistological study using paraffin wax sections. Journal of clinical pathology 1989;42(9):902-5.
90. Diab SG, Clark GM, Osborne CK, Libby A, Allred DC, Elledge RM. Tumor characteristics and clinical outcome of tubular and mucinous breast carcinomas. J Clin Oncol 1999;17(5):1442-8.
91. Komaki K, Sakamoto G, Sugano H, Morimoto T, Monden Y. Mucinous carcinoma of the breast in Japan. A prognostic analysis based on morphologic features. Cancer 1988;61(5):989-96.
92. Fentiman IS, Millis RR, Smith P, Ellul JP, Lampejo O. Mucoid breast carcinomas: histology and prognosis. British journal of cancer 1997;75(7):1061-5.
93. Norris HJ, Taylor HB. Prognosis of Mucinous (Gelatinous) Carcinoma of the Breast. Cancer 1965;18:879-85.
94. Toikkanen S, Kujari H. Pure and mixed mucinous carcinomas of the breast: a clinicopathologic analysis of 61 cases with long-term follow-up. Human pathology 1989;20(8):758-64.
95. Oberman HA, Fidler WJ, Jr. Tubular carcinoma of the breast. The American journal of surgical pathology 1979;3(5):387-95.
96. van Bogaert LJ. Clinicopathologic hallmarks of mammary tubular carcinoma. Human pathology 1982;13(6):558-62.
97. Kader HA, Jackson J, Mates D, Andersen S, Hayes M, Olivotto IA. Tubular carcinoma of the breast: a population-based study of nodal metastases at presentation and of patterns of relapse. The breast journal 2001;7(1):8-13.
98. Papadatos G, Rangan AM, Psarianos T, Ung O, Taylor R, Boyages J. Probability of axillary node involvement in patients with tubular carcinoma of the breast. The British journal of surgery 2001;88(6):860-4.
99. Carstens PH, Greenberg RA, Francis D, Lyon H. Tubular carcinoma of the breast. A long term follow-up. Histopathology 1985;9(3):271-80.
100. Carstens PH, Huvos AG, Foote FW, Jr., Ashikari R. Tubular carcinoma of the breast: a clinicopathologic study of 35 cases. American journal of clinical pathology 1972;58(3):231-8.
101. Cooper HS, Patchefsky AS, Krall RA. Tubular carcinoma of the breast. Cancer 1978;42(5):2334-42.
102. McDivitt RW, Boyce W, Gersell D. Tubular carcinoma of the breast. Clinical and pathological observations concerning 135 cases. The American journal of surgical pathology 1982;6(5):401-11.
103. Peters GN, Wolff M, Haagensen CD. Tubular carcinoma of the breast. Clinical pathologic correlations based on 100 cases. Annals of surgery 1981;193(2):138-49.
104. Deos PH, Norris HJ. Well-differentiated (tubular) carcinoma of the breast. A clinicopathologic study of 145 pure and mixed cases. American journal of clinical pathology 1982;78(1):1-7.
105. Parl FF, Richardson LD. The histologic and biologic spectrum of tubular carcinoma of the breast. Human pathology 1983;14(8):694-8.
106. Santeusanio G, Pascal RR, Bisceglia M, Costantino AM, Bosman C. Metaplastic breast carcinoma with epithelial phenotype of pseudosarcomatous components. Archives of pathology & laboratory medicine 1988;112(1):82-5.
107. Eusebi V, Cattani MG, Ceccarelli C, Lamovec J. Sarcomatoid carcinomas of the breast. An immunohistochemical study of 14 cases. Progr Surg Pathol 1989;10:83-100.
108. Wada H, Enomoto T, Tsujimoto M, Nomura T, Murata Y, Shroyer KR. Carcinosarcoma of the breast: molecular-biological study for analysis of histogenesis. Human pathology 1998;29(11):1324-8.

109. Wang X, Mori I, Tang W, et al. Metaplastic carcinoma of the breast: p53 analysis identified the same point mutation in the three histologic components. Mod Pathol 2001;14(11):1183-6.
110. Zhuang Z, Lininger RA, Man YG, Albuquerque A, Merino MJ, Tavassoli FA. Identical clonality of both components of mammary carcinosarcoma with differential loss of heterozygosity. Mod Pathol 1997;10(4):354-62.
111. Leibl S, Moinfar F. Metaplastic breast carcinomas are negative for Her-2 but frequently express EGFR (Her-1): potential relevance to adjuvant treatment with EGFR tyrosine kinase inhibitors? Journal of clinical pathology 2005;58(7):700-4.
112. Sotiriou C, Neo SY, McShane LM, et al. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. Proceedings of the National Academy of Sciences of the United States of America 2003;100(18):10393-8.
113. Pusztai L, Ayers M, Stec J, et al. Gene expression profiles obtained from fine-needle aspirations of breast cancer reliably identify routine prognostic markers and reveal large-scale molecular differences between estrogen-negative and estrogen-positive tumors. Clin Cancer Res 2003;9(7):2406-15.
114. Rouzier R, Perou CM, Symmans WF, et al. Breast cancer molecular subtypes respond differently to preoperative chemotherapy. Clin Cancer Res 2005;11(16):5678-85.
115. Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. Nature 2000;406(6797):747-52.
116. Sorlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proceedings of the National Academy of Sciences of the United States of America 2001;98(19):10869-74.
117. Sorlie T, Tibshirani R, Parker J, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proceedings of the National Academy of Sciences of the United States of America 2003;100(14):8418-23.
118. Yang F, Foekens JA, Yu J, et al. Laser microdissection and microarray analysis of breast tumors reveal ER-alpha related genes and pathways. Oncogene 2006;25(9):1413-9.
119. Gruvberger S, Ringner M, Chen Y, et al. Estrogen receptor status in breast cancer is associated with remarkably distinct gene expression patterns. Cancer research 2001;61(16):5979-84.
120. Hu Z, Fan C, Oh DS, et al. The molecular portraits of breast tumors are conserved across microarray platforms. BMC genomics 2006;7:96.
121. Yu K, Lee CH, Tan PH, Tan P. Conservation of breast cancer molecular subtypes and transcriptional patterns of tumor progression across distinct ethnic populations. Clin Cancer Res 2004;10(16):5508-17.
122. Smid M, Wang Y, Zhang Y, et al. Subtypes of breast cancer show preferential site of relapse. Cancer research 2008;68(9):3108-14.
123. van 't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature 2002;415(6871):530-6.
124. van de Vijver MJ, He YD, van't Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. The New England journal of medicine 2002;347(25):1999-2009.
125. Wang Y, Klijn JG, Zhang Y, et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. Lancet 2005;365(9460):671-9.
126. Foekens JA, Atkins D, Zhang Y, et al. Multicenter validation of a gene expression-based prognostic signature in lymph node-negative primary breast cancer. J Clin Oncol 2006;24(11):1665-71.
127. Desmedt C, Piette F, Loi S, et al. Strong time dependence of the 76-gene prognostic signature for node-negative breast cancer patients in the TRANSBIG multicenter independent validation series. Clin Cancer Res 2007;13(11):3207-14.
128. Eifel P, Axelson JA, Costa J, et al. National Institutes of Health Consensus Development Conference Statement: adjuvant therapy for breast cancer, November 1-3, 2000. Journal of the National Cancer Institute 2001;93(13):979-89.

129. Goldhirsch A, Wood WC, Gelber RD, Coates AS, Thurlimann B, Senn HJ. Meeting highlights: updated international expert consensus on the primary therapy of early breast cancer. J Clin Oncol 2003;21(17):3357-65.
130. Buyse M, Loi S, van't Veer L, et al. Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. Journal of the National Cancer Institute 2006;98(17):1183-92.
131. Paik S, Shak S, Tang G, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. The New England journal of medicine 2004;351(27):2817-26.
132. Habel LA, Shak S, Jacobs MK, et al. A population-based study of tumor gene expression and risk of breast cancer death among lymph node-negative patients. Breast Cancer Res 2006;8(3):R25.
133. Paik S, Tang G, Shak S, et al. Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. J Clin Oncol 2006;24(23):3726-34.
134. Sotiriou C, Wirapati P, Loi S, et al. Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. Journal of the National Cancer Institute 2006;98(4):262-72.
135. Chang HY, Sneddon JB, Alizadeh AA, et al. Gene expression signature of fibroblast serum response predicts human cancer progression: similarities between tumors and wounds. PLoS biology 2004;2(2):E7.
136. Chang HY, Nuyten DS, Sneddon JB, et al. Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival. Proceedings of the National Academy of Sciences of the United States of America 2005;102(10):3738-43.
137. Fan C, Oh DS, Wessels L, et al. Concordance among gene-expression-based predictors for breast cancer. The New England journal of medicine 2006;355(6):560-9.
138. Yu JX, Sieuwerts AM, Zhang Y, et al. Pathway analysis of gene signatures predicting metastasis of node-negative primary breast cancer. BMC cancer 2007;7:182.
139. Ma XJ, Wang Z, Ryan PD, et al. A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen. Cancer cell 2004;5(6):607-16.
140. Jansen MP, Foekens JA, van Staveren IL, et al. Molecular classification of tamoxifen-resistant breast carcinomas by gene expression profiling. J Clin Oncol 2005;23(4):732-40.
141. Chang JC, Wooten EC, Tsimelzon A, et al. Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer. Lancet 2003;362(9381):362-9.
142. Hess KR, Anderson K, Symmans WF, et al. Pharmacogenomic predictor of sensitivity to preoperative chemotherapy with paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide in breast cancer. J Clin Oncol 2006;24(26):4236-44.
143. Iwao-Koizumi K, Matoba R, Ueno N, et al. Prediction of docetaxel response in human breast cancer by gene expression profiling. J Clin Oncol 2005;23(3):422-31.
144. Ayers M, Symmans WF, Stec J, et al. Gene expression profiles predict complete pathologic response to neoadjuvant paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide chemotherapy in breast cancer. J Clin Oncol 2004;22(12):2284-93.
145. Folgueira MA, Carraro DM, Brentani H, et al. Gene expression profile associated with response to doxorubicin-based therapy in breast cancer. Clin Cancer Res 2005;11(20):7434-43.
146. Gianni L, Zambetti M, Clark K, et al. Gene expression profiles in paraffin-embedded core biopsy tissue predict response to chemotherapy in women with locally advanced breast cancer. J Clin Oncol 2005;23(29):7265-77.
147. Modlich O, Prisack HB, Munnes M, Audretsch W, Bojar H. Predictors of primary breast cancers responsiveness to preoperative epirubicin/cyclophosphamide-based chemotherapy: translation of microarray data into clinically useful predictive signatures. Journal of translational medicine 2005;3:32.
148. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet 2005;365(9472):1687-717.

149. Jansen MP, Sieuwerts AM, Look MP, et al. HOXB13-to-IL17BR expression ratio is related with tumor aggressiveness and response to tamoxifen of recurrent breast cancer: a retrospective study. J Clin Oncol 2007;25(6):662-8.
150. Hayes DF, Trock B, Harris AL. Assessing the clinical impact of prognostic factors: when is "statistically significant" clinically useful? Breast cancer research and treatment 1998;52(1-3):305-19.
151. Kok M, Linn SC, Van Laar RK, et al. Comparison of gene expression profiles predicting progression in breast cancer patients treated with tamoxifen. Breast cancer research and treatment 2008.
152. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell 2000;100(1):57-70.
153. Sjoblom T, Jones S, Wood LD, et al. The consensus coding sequences of human breast and colorectal cancers. Science (New York, NY 2006;314(5797):268-74.
154. Wood LD, Parsons DW, Jones S, et al. The genomic landscapes of human breast and colorectal cancers. Science (New York, NY 2007;318(5853):1108-13.
155. Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. Lancet 2001;358(9291):1389-99.
156. Peto J, Collins N, Barfoot R, et al. Prevalence of BRCA1 and BRCA2 gene mutations in patients with early-onset breast cancer. Journal of the National Cancer Institute 1999;91(11):943-9.
157. Prevalence and penetrance of BRCA1 and BRCA2 mutations in a population-based series of breast cancer cases. Anglian Breast Cancer Study Group. British journal of cancer 2000;83(10):1301-8.
158. Antoniou AC, Pharoah PD, McMullan G, Day NE, Ponder BA, Easton D. Evidence for further breast cancer susceptibility genes in addition to BRCA1 and BRCA2 in a population-based study. Genetic epidemiology 2001;21(1):1-18.
159. Meijers-Heijboer H, van den Ouweland A, Klijn J, et al. Low-penetrance susceptibility to breast cancer due to CHEK2(*)1100delC in noncarriers of BRCA1 or BRCA2 mutations. Nature genetics 2002;31(1):55-9.
160. Vahteristo P, Bartkova J, Eerola H, et al. A CHEK2 genetic variant contributing to a substantial fraction of familial breast cancer. American journal of human genetics 2002;71(2):432-8.
161. Easton DF. How many more breast cancer predisposition genes are there? Breast Cancer Res 1999;1(1):14-7.
162. Seal S, Thompson D, Renwick A, et al. Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles. Nature genetics 2006;38(11):1239-41.
163. Rahman N, Seal S, Thompson D, et al. PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. Nature genetics 2007;39(2):165-7.
164. Easton DF, Pooley KA, Dunning AM, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. Nature 2007;447(7148):1087-93.
165. Takeichi M. Cadherin cell adhesion receptors as a morphogenetic regulator. Science (New York, NY 1991;251(5000):1451-5.
166. Kemler R. From cadherins to catenins: cytoplasmic protein interactions and regulation of cell adhesion. Trends Genet 1993;9(9):317-21.
167. Cowin P. Unraveling the cytoplasmic interactions of the cadherin superfamily. Proceedings of the National Academy of Sciences of the United States of America 1994;91(23):10759-61.
168. Gumbiner BM. Cell adhesion: the molecular basis of tissue architecture and morphogenesis. Cell 1996;84(3):345-57.
169. Semb H, Christofori G. The tumor-suppressor function of E-cadherin. American journal of human genetics 1998;63(6):1588-93.
170. Berx G, Van Roy F. The E-cadherin/catenin complex: an important gatekeeper in breast cancer tumorigenesis and malignant progression. Breast Cancer Res 2001;3(5):289-93.
171. Jamora C, Fuchs E. Intercellular adhesion, signalling and the cytoskeleton. Nat Cell Biol 2002;4(4):E101-8.

172. Drees F, Pokutta S, Yamada S, Nelson WJ, Weis WI. Alpha-catenin is a molecular switch that binds E-cadherin-beta-catenin and regulates actin-filament assembly. Cell 2005;123(5):903-15.
173. Yamada S, Pokutta S, Drees F, Weis WI, Nelson WJ. Deconstructing the cadherin-catenin-actin complex. Cell 2005;123(5):889-901.
174. Thoreson MA, Anastasiadis PZ, Daniel JM, et al. Selective uncoupling of p120(ctn) from E-cadherin disrupts strong adhesion. J Cell Biol 2000;148(1):189-202.
175. Frixen UH, Behrens J, Sachs M, et al. E-cadherin-mediated cell-cell adhesion prevents invasiveness of human carcinoma cells. J Cell Biol 1991;113(1):173-85.
176. Vleminckx K, Vakaet L, Jr., Mareel M, Fiers W, van Roy F. Genetic manipulation of E-cadherin expression by epithelial tumor cells reveals an invasion suppressor role. Cell 1991;66(1):107-19.
177. Perl AK, Wilgenbus P, Dahl U, Semb H, Christofori G. A causal role for E-cadherin in the transition from adenoma to carcinoma. Nature 1998;392(6672):190-3.
178. Huiping C, Sigurgeirsdottir JR, Jonasson JG, et al. Chromosome alterations and E-cadherin gene mutations in human lobular breast cancer. British journal of cancer 1999;81(7):1103-10.
179. Vos CB, Cleton-Jansen AM, Berx G, et al. E-cadherin inactivation in lobular carcinoma in situ of the breast: an early event in tumorigenesis. British journal of cancer 1997;76(9):1131-3.
180. Becker KF, Atkinson MJ, Reich U, et al. E-cadherin gene mutations provide clues to diffuse type gastric carcinomas. Cancer research 1994;54(14):3845-52.
181. Becker KF, Atkinson MJ, Reich U, et al. Exon skipping in the E-cadherin gene transcript in metastatic human gastric carcinomas. Hum Mol Genet 1993;2(6):803-4.
182. Risinger JI, Berchuck A, Kohler MF, Boyd J. Mutations of the E-cadherin gene in human gynecologic cancers. Nature genetics 1994;7(1):98-102.
183. Berx G, Becker KF, Hofler H, van Roy F. Mutations of the human E-cadherin (CDH1) gene. Human mutation 1998;12(4):226-37.
184. Derksen PW, Liu X, Saridin F, et al. Somatic inactivation of E-cadherin and p53 in mice leads to metastatic lobular mammary carcinoma through induction of anoikis resistance and angiogenesis. Cancer cell 2006;10(5):437-49.
185. Guilford P, Hopkins J, Harraway J, et al. E-cadherin germline mutations in familial gastric cancer. Nature 1998;392(6674):402-5.
186. Gayther SA, Gorringe KL, Ramus SJ, et al. Identification of germ-line E-cadherin mutations in gastric cancer families of European origin. Cancer research 1998;58(18):4086-9.
187. Caldas C, Carneiro F, Lynch HT, et al. Familial gastric cancer: overview and guidelines for management. Journal of medical genetics 1999;36(12):873-80.
188. Guilford PJ, Hopkins JB, Grady WM, et al. E-cadherin germline mutations define an inherited cancer syndrome dominated by diffuse gastric cancer. Human mutation 1999;14(3):249-55.
189. Keller G, Vogelsang H, Becker I, et al. Diffuse type gastric and lobular breast carcinoma in a familial gastric cancer patient with an E-cadherin germline mutation. The American journal of pathology 1999;155(2):337-42.
190. Brooks-Wilson AR, Kaurah P, Suriano G, et al. Germline E-cadherin mutations in hereditary diffuse gastric cancer: assessment of 42 new families and review of genetic screening criteria. Journal of medical genetics 2004;41(7):508-17.
191. Suriano G, Yew S, Ferreira P, et al. Characterization of a recurrent germ line mutation of the E-cadherin gene: implications for genetic testing and clinical management. Clin Cancer Res 2005;11(15):5401-9.
192. Pharoah PD, Guilford P, Caldas C. Incidence of gastric cancer and breast cancer in CDH1 (E-cadherin) mutation carriers from hereditary diffuse gastric cancer families. Gastroenterology 2001;121(6):1348-53.
193. Masciari S, Larsson N, Senz J, et al. Germline E-cadherin mutations in familial lobular breast cancer. Journal of medical genetics 2007;44(11):726-31.
194. Rahman N, Stone JG, Coleman G, et al. Lobular carcinoma in situ of the breast is not caused by constitutional mutations in the E-cadherin gene. British journal of cancer 2000;82(3):568-70.

195. Yoshiura K, Kanai Y, Ochiai A, Shimoyama Y, Sugimura T, Hirohashi S. Silencing of the E-cadherin invasion-suppressor gene by CpG methylation in human carcinomas. Proceedings of the National Academy of Sciences of the United States of America 1995;92(16):7416-9.
196. Graff JR, Herman JG, Lapidus RG, et al. E-cadherin expression is silenced by DNA hypermethylation in human breast and prostate carcinomas. Cancer research 1995;55(22):5195-9.
197. Graff JR, Gabrielson E, Fujii H, Baylin SB, Herman JG. Methylation patterns of the E-cadherin 5′ CpG island are unstable and reflect the dynamic, heterogeneous loss of E-cadherin expression during metastatic progression. J Biol Chem 2000;275(4):2727-32.
198. Peinado H, Olmeda D, Cano A. Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? Nat Rev Cancer 2007;7(6):415-28.
199. Batlle E, Sancho E, Franci C, et al. The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells. Nat Cell Biol 2000;2(2):84-9.
200. Cano A, Perez-Moreno MA, Rodrigo I, et al. The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. Nat Cell Biol 2000;2(2):76-83.
201. Hajra KM, Chen DY, Fearon ER. The SLUG zinc-finger protein represses E-cadherin in breast cancer. Cancer research 2002;62(6):1613-8.
202. Comijn J, Berx G, Vermassen P, et al. The two-handed E box binding zinc finger protein SIP1 down-regulates E-cadherin and induces invasion. Mol Cell 2001;7(6):1267-78.
203. Eger A, Aigner K, Sonderegger S, et al. DeltaEF1 is a transcriptional repressor of E-cadherin and regulates epithelial plasticity in breast cancer cells. Oncogene 2005;24(14):2375-85.
204. Perez-Moreno MA, Locascio A, Rodrigo I, et al. A new role for E12/E47 in the repression of E-cadherin expression and epithelial-mesenchymal transitions. J Biol Chem 2001;276(29):27424-31.
205. Yang J, Mani SA, Donaher JL, et al. Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. Cell 2004;117(7):927-39.
206. Peinado H, Cano A. Regulation of the E-cadherin cell-cell adhesion gene. In: Esteller M, ed. DNA methylation, epigenetics and metastasis: Springer; 2005:157-90.
207. Thiery JP. Epithelial-mesenchymal transitions in tumour progression. Nat Rev Cancer 2002;2(6):442-54.
208. Gupta GP, Massague J. Cancer metastasis: building a framework. Cell 2006;127(4):679-95.
209. Mani SA, Guo W, Liao MJ, et al. The epithelial-mesenchymal transition generates cells with properties of stem cells. Cell 2008;133(4):704-15.
210. Clevers H. Wnt/beta-catenin signaling in development and disease. Cell 2006;127(3):469-80.
211. van de Wetering M, Sancho E, Verweij C, et al. The beta-catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. Cell 2002;111(2):241-50.
212. Reya T, Clevers H. Wnt signalling in stem cells and cancer. Nature 2005;434(7035):843-50.
213. Behrens J. Cadherins and catenins: role in signal transduction and tumor progression. Cancer metastasis reviews 1999;18(1):15-30.
214. van de Wetering M, Barker N, Harkes IC, et al. Mutant E-cadherin breast cancer cells do not display constitutive Wnt signaling. Cancer research 2001;61(1):278-84.
215. Levine AJ. p53, the cellular gatekeeper for growth and division. Cell 1997;88(3):323-31.
216. Giaccia AJ, Kastan MB. The complexity of p53 modulation: emerging patterns from divergent signals. Genes & development 1998;12(19):2973-83.
217. Bartek J, Falck J, Lukas J. CHK2 kinase--a busy messenger. Nature reviews 2001;2(12):877-86.
218. Bartek J, Lukas J. Chk1 and Chk2 kinases in checkpoint control and cancer. Cancer cell 2003;3(5):421-9.
219. Sherr CJ, Roberts JM. CDK inhibitors: positive and negative regulators of G1-phase progression. Genes & development 1999;13(12):1501-12.
220. Juven-Gershon T, Oren M. Mdm2: the ups and downs. Molecular medicine (Cambridge, Mass 1999;5(2):71-83.
221. Gasco M, Shami S, Crook T. The p53 pathway in breast cancer. Breast Cancer Res 2002;4(2):70-6.

222. Coles C, Condie A, Chetty U, Steel CM, Evans HJ, Prosser J. p53 mutations in breast cancer. Cancer research 1992;52(19):5291-8.
223. Borresen-Dale AL. TP53 and breast cancer. Human mutation 2003;21(3):292-300.
224. Lipponen P, Ji H, Aaltomaa S, Syrjanen S, Syrjanen K. p53 protein expression in breast cancer as related to histopathological characteristics and prognosis. International journal of cancer 1993;55(1):51-6.
225. Berns EM, De Witte HH, Klijn JG, et al. Prognostic value of TP53 protein accumulation in human primary breast cancer: an analysis by luminometric immunoassay on 1491 tumor cytosols. Anticancer research 1997;17(4B):3003-6.
226. Malkin D, Li FP, Strong LC, et al. Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. Science (New York, NY 1990;250(4985):1233-8.
227. Li FP, Fraumeni JF, Jr. Rhabdomyosarcoma in children: epidemiologic study and identification of a familial cancer syndrome. Journal of the National Cancer Institute 1969;43(6):1365-73.
228. Bell DW, Varley JM, Szydlo TE, et al. Heterozygous germ line hCHK2 mutations in Li-Fraumeni syndrome. Science (New York, NY 1999;286(5449):2528-31.
229. Li FP, Fraumeni JF, Jr., Mulvihill JJ, et al. A cancer family syndrome in twenty-four kindreds. Cancer research 1988;48(18):5358-62.
230. Meijers-Heijboer H, Wijnen J, Vasen H, et al. The CHEK2 1100delC mutation identifies families with a hereditary breast and colorectal cancer phenotype. American journal of human genetics 2003;72(5):1308-14.
231. Broeks A, de Witte L, Nooijen A, et al. Excess risk for contralateral breast cancer in CHEK2*1100delC germline mutation carriers. Breast cancer research and treatment 2004;83(1):91-3.
232. de Bock GH, Schutte M, Krol-Warmerdam EM, et al. Tumour characteristics and prognosis of breast cancer patients carrying the germline CHEK2*1100delC variant. Journal of medical genetics 2004;41(10):731-5.
233. Kilpivaara O, Bartkova J, Eerola H, et al. Correlation of CHEK2 protein expression and c.1100delC mutation status with tumor characteristics among unselected breast cancer patients. International journal of cancer 2005;113(4):575-80.
234. Ho GH, Calvano JE, Bisogna M, et al. Genetic alterations of the p14ARF -hdm2-p53 regulatory pathway in breast carcinoma. Breast cancer research and treatment 2001;65(3):225-32.
235. Quesnel B, Preudhomme C, Fournier J, Fenaux P, Peyrat JP. MDM2 gene amplification in human breast cancer. Eur J Cancer 1994;30A(7):982-4.
236. Silva J, Dominguez G, Silva JM, et al. Analysis of genetic and epigenetic processes that influence p14ARF expression in breast cancer. Oncogene 2001;20(33):4586-90.
237. Vivanco I, Sawyers CL. The phosphatidylinositol 3-Kinase AKT pathway in human cancer. Nat Rev Cancer 2002;2(7):489-501.
238. Samuels Y, Wang Z, Bardelli A, et al. High frequency of mutations of the PIK3CA gene in human cancers. Science (New York, NY 2004;304(5670):554.
239. Karakas B, Bachman KE, Park BH. Mutation of the PIK3CA oncogene in human cancers. British journal of cancer 2006;94(4):455-9.
240. Saal LH, Holm K, Maurer M, et al. PIK3CA mutations correlate with hormone receptors, node metastasis, and ERBB2, and are mutually exclusive with PTEN loss in human breast carcinoma. Cancer research 2005;65(7):2554-9.
241. Li J, Simpson L, Takahashi M, et al. The PTEN/MMAC1 tumor suppressor induces cell death that is rescued by the AKT/protein kinase B oncogene. Cancer research 1998;58(24):5667-72.
242. Steck PA, Pershouse MA, Jasser SA, et al. Identification of a candidate tumour suppressor gene, MMAC1, at chromosome 10q23.3 that is mutated in multiple advanced cancers. Nature genetics 1997;15(4):356-62.
243. Parsons R. Human cancer, PTEN and the PI-3 kinase pathway. Semin Cell Dev Biol 2004;15(2):171-6.
244. Liaw D, Marsh DJ, Li J, et al. Germline mutations of the PTEN gene in Cowden disease, an inherited breast and thyroid cancer syndrome. Nature genetics 1997;16(1):64-7.

245. Marsh DJ, Dahia PL, Zheng Z, et al. Germline mutations in PTEN are present in Bannayan-Zonana syndrome. Nature genetics 1997;16(4):333-4.
246. Saal LH, Gruvberger-Saal SK, Persson C, et al. Recurrent gross mutations of the PTEN tumor suppressor gene in breast cancers with deficient DSB repair. Nature genetics 2008;40(1):102-7.
247. Saal LH, Johansson P, Holm K, et al. Poor prognosis in carcinoma is associated with a gene expression signature of aberrant PTEN tumor suppressor pathway activity. Proceedings of the National Academy of Sciences of the United States of America 2007;104(18):7564-9.
248. Sun M, Wang G, Paciga JE, et al. AKT1/PKBalpha kinase is frequently elevated in human cancers and its constitutive activation is required for oncogenic transformation in NIH3T3 cells. The American journal of pathology 2001;159(2):431-7.
249. Carpten JD, Faber AL, Horn C, et al. A transforming mutation in the pleckstrin homology domain of AKT1 in cancer. Nature 2007;448(7152):439-44.
250. Bellacosa A, de Feo D, Godwin AK, et al. Molecular alterations of the AKT2 oncogene in ovarian and breast carcinomas. International journal of cancer 1995;64(4):280-5.
251. Sun M, Paciga JE, Feldman RI, et al. Phosphatidylinositol-3-OH Kinase (PI3K)/AKT2, activated in breast cancer, regulates and is induced by estrogen receptor alpha (ERalpha) via interaction between ERalpha and PI3K. Cancer research 2001;61(16):5985-91.
252. Shaw RJ, Cantley LC. Ras, PI(3)K and mTOR signalling controls tumour cell growth. Nature 2006;441(7092):424-30.
253. Riccardi VM. Neurofibromatosis: phenotype, natural history, and pathogenesis. 2 ed. Baltimore: Johns Hopkins University Press; 1992.
254. Bos JL. ras oncogenes in human cancer: a review. Cancer research 1989;49(17):4682-9.
255. Bos JL, Fearon ER, Hamilton SR, et al. Prevalence of ras gene mutations in human colorectal cancers. Nature 1987;327(6120):293-7.
256. Almoguera C, Shibata D, Forrester K, Martin J, Arnheim N, Perucho M. Most human carcinomas of the exocrine pancreas contain mutant c-K-ras genes. Cell 1988;53(4):549-54.
257. Smit VT, Boot AJ, Smits AM, Fleuren GJ, Cornelisse CJ, Bos JL. KRAS codon 12 mutations occur very frequently in pancreatic adenocarcinomas. Nucleic acids research 1988;16(16):7773-82.
258. Davies H, Bignell GR, Cox C, et al. Mutations of the BRAF gene in human cancer. Nature 2002;417(6892):949-54.
259. Rajagopalan H, Bardelli A, Lengauer C, Kinzler KW, Vogelstein B, Velculescu VE. Tumorigenesis: RAF/RAS oncogenes and mismatch-repair status. Nature 2002;418(6901):934.
260. Emuss V, Garnett M, Mason C, Marais R. Mutations of C-RAF are rare in human cancer because C-RAF has a low basal kinase activity compared with B-RAF. Cancer research 2005;65(21):9719-26.
261. Lee JW, Soung YH, Kim SY, et al. Mutational analysis of the ARAF gene in human cancers. Apmis 2005;113(1):54-7.
262. Aoki Y, Niihori T, Kawame H, et al. Germline mutations in HRAS proto-oncogene cause Costello syndrome. Nature genetics 2005;37(10):1038-40.
263. Niihori T, Aoki Y, Narumi Y, et al. Germline KRAS and BRAF mutations in cardio-facio-cutaneous syndrome. Nature genetics 2006;38(3):294-6.
264. Rodriguez-Viciana P, Tetsu O, Tidyman WE, et al. Germline mutations in genes within the MAPK pathway cause cardio-facio-cutaneous syndrome. Science (New York, NY 2006;311(5765):1287-90.
265. Sherr CJ, McCormick F. The RB and p53 pathways in cancer. Cancer cell 2002;2(2):103-12.
266. Hussussian CJ, Struewing JP, Goldstein AM, et al. Germline p16 mutations in familial melanoma. Nature genetics 1994;8(1):15-21.
267. Kamb A, Shattuck-Eidens D, Eeles R, et al. Analysis of the p16 gene (CDKN2) as a candidate for the chromosome 9p melanoma susceptibility locus. Nature genetics 1994;8(1):23-6.
268. Silva J, Silva JM, Dominguez G, et al. Concomitant expression of p16INK4a and p14ARF in primary breast cancer and analysis of inactivation mechanisms. The Journal of pathology 2003;199(3):289-97.

269. Xu L, Sgroi D, Sterner CJ, et al. Mutational analysis of CDKN2 (MTS1/p16ink4) in human breast carcinomas. Cancer research 1994;54(20):5262-4.
270. Quesnel B, Fenaux P, Philippe N, et al. Analysis of p16 gene deletion and point mutation in breast carcinoma. British journal of cancer 1995;72(2):351-3.
271. Berns EM, Klijn JG, Smid M, van Staveren IL, Gruis NA, Foekens JA. Infrequent CDKN2 (MTS1/p16) gene alterations in human primary breast cancer. British journal of cancer 1995;72(4):964-7.
272. Pollock PM, Pearson JV, Hayward NK. Compilation of somatic mutations of the CDKN2 gene in human cancers: non-random distribution of base substitutions. Genes, chromosomes & cancer 1996;15(2):77-88.
273. Serrano M, Lee H, Chin L, Cordon-Cardo C, Beach D, DePinho RA. Role of the INK4a locus in tumor suppression and cell mortality. Cell 1996;85(1):27-37.
274. Caldas C, Hahn SA, da Costa LT, et al. Frequent somatic mutations and homozygous deletions of the p16 (MTS1) gene in pancreatic adenocarcinoma. Nature genetics 1994;8(1):27-32.
275. Schutte M, Hruban RH, Geradts J, et al. Abrogation of the Rb/p16 tumor-suppressive pathway in virtually all pancreatic carcinomas. Cancer research 1997;57(15):3126-30.
276. Friend SH, Bernards R, Rogelj S, et al. A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. Nature 1986;323(6089):643-6.
277. Lee EY, To H, Shew JY, Bookstein R, Scully P, Lee WH. Inactivation of the retinoblastoma susceptibility gene in human breast cancers. Science (New York, NY 1988;241(4862):218-21.
278. T'Ang A, Varley JM, Chakraborty S, Murphree AL, Fung YK. Structural rearrangement of the retinoblastoma gene in human breast carcinoma. Science (New York, NY 1988;242(4876):263-6.
279. Sherr CJ. Cancer cell cycles. Science (New York, NY 1996;274(5293):1672-7.
280. Hunter T. Oncoprotein networks. Cell 1997;88(3):333-46.
281. Weinberg RA. The retinoblastoma protein and cell cycle control. Cell 1995;81(3):323-30.
282. Dublin EA, Patel NK, Gillett CE, Smith P, Peters G, Barnes DM. Retinoblastoma and p16 proteins in mammary carcinoma: their relationship to cyclin D1 and histopathological parameters. International journal of cancer 1998;79(1):71-5.
283. Bartkova J, Lukas J, Muller H, Lutzhoft D, Strauss M, Bartek J. Cyclin D1 protein expression and function in human breast cancer. International journal of cancer 1994;57(3):353-61.
284. Barbareschi M, Pelosio P, Caffo O, et al. Cyclin-D1-gene amplification and expression in breast carcinoma: relation with clinicopathologic characteristics and with retinoblastoma gene product, p53 and p21WAF1 immunohistochemical expression. International journal of cancer 1997;74(2):171-4.
285. An HX, Beckmann MW, Reifenberger G, Bender HG, Niederacher D. Gene amplification and overexpression of CDK4 in sporadic breast carcinomas is associated with high tumor cell proliferation. The American journal of pathology 1999;154(1):113-8.
286. Wolfel T, Hauer M, Schneider J, et al. A p16INK4a-insensitive CDK4 mutant targeted by cytolytic T lymphocytes in a human melanoma. Science (New York, NY 1995;269(5228):1281-4.
287. Zuo L, Weger J, Yang Q, et al. Germline mutations in the p16INK4a binding domain of CDK4 in familial melanoma. Nature genetics 1996;12(1):97-9.
288. Soufir N, Avril MF, Chompret A, et al. Prevalence of p16 and CDK4 germline mutations in 48 melanoma-prone families in France. The French Familial Melanoma Study Group. Hum Mol Genet 1998;7(2):209-16.
289. Molven A, Grimstvedt MB, Steine SJ, et al. A large Norwegian family with inherited malignant melanoma, multiple atypical nevi, and CDK4 mutation. Genes, chromosomes & cancer 2005;44(1):10-8.
290. Porter PL, Malone KE, Heagerty PJ, et al. Expression of cell-cycle regulators p27Kip1 and cyclin E, alone and in combination, correlate with survival in young breast cancer patients. Nature medicine 1997;3(2):222-5.
291. Sieuwerts AM, Look MP, Meijer-van Gelder ME, et al. Which cyclin E prevails as prognostic marker for breast cancer? Results from a retrospective study involving 635 lymph node-negative breast cancer patients. Clin Cancer Res 2006;12(11 Pt 1):3319-28.

292. Malumbres M, Barbacid M. To cycle or not to cycle: a critical decision in cancer. Nat Rev Cancer 2001;1(3):222-31.
293. Tutt A, Ashworth A. The relationship between the roles of BRCA genes in DNA repair and cancer predisposition. Trends in molecular medicine 2002;8(12):571-6.
294. Venkitaraman AR. Cancer susceptibility and the functions of BRCA1 and BRCA2. Cell 2002;108(2):171-82.
295. Gowen LC, Avrutskaya AV, Latour AM, Koller BH, Leadon SA. BRCA1 required for transcription-coupled repair of oxidative DNA damage. Science (New York, NY 1998;281(5379):1009-12.
296. Wang Y, Cortez D, Yazdi P, Neff N, Elledge SJ, Qin J. BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures. Genes & development 2000;14(8):927-39.
297. Zhong Q, Chen CF, Li S, et al. Association of BRCA1 with the hRad50-hMre11-p95 complex and the DNA damage response. Science (New York, NY 1999;285(5428):747-50.
298. Antoniou A, Pharoah PD, Narod S, et al. Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case Series unselected for family history: a combined analysis of 22 studies. American journal of human genetics 2003;72(5):1117-30.
299. Antoniou AC, Spurdle AB, Sinilnikova OM, et al. Common breast cancer-predisposition alleles are associated with breast cancer risk in BRCA1 and BRCA2 mutation carriers. American journal of human genetics 2008;82(4):937-48.
300. Welcsh PL, King MC. BRCA1 and BRCA2 and the genetics of breast and ovarian cancer. Hum Mol Genet 2001;10(7):705-13.
301. Crook T, Brooks LA, Crossland S, et al. p53 mutation with frequent novel condons but not a mutator phenotype in BRCA1- and BRCA2-associated breast tumours. Oncogene 1998;17(13):1681-9.
302. Thompson ME, Jensen RA, Obermiller PS, Page DL, Holt JT. Decreased expression of BRCA1 accelerates growth and is often present during sporadic breast cancer progression. Nature genetics 1995;9(4):444-50.
303. Rice JC, Massey-Brown KS, Futscher BW. Aberrant methylation of the BRCA1 CpG island promoter is associated with decreased BRCA1 mRNA in sporadic breast cancer cells. Oncogene 1998;17(14):1807-12.
304. Esteller M, Silva JM, Dominguez G, et al. Promoter hypermethylation and BRCA1 inactivation in sporadic breast and ovarian tumors. Journal of the National Cancer Institute 2000;92(7):564-9.

Chapter 2

Aims and Outline of the Thesis

AIMS AND OUTLINE OF THE THESIS

Breast cancer not only is the most frequently diagnosed cancer in Western women, but also is the second leading cause of cancer death in the Western world [1]. Clinically, breast cancer has for long been recognized to be a heterogeneous disease. Currently, about two-thirds of breast cancer patients survive their disease, whereas, one-third of breast cancer patients will die of metastases of the primary cancer within 15 years from diagnosis. Therefore, it is important for clinicians to accurately predict the prognosis and most appropriate therapy for each breast cancer patient. Recent advances in large scale gene expression analysis have significantly improved prognostic and predictive stratification of the patients. Importantly, these analyses also identified five molecular subtypes of breast cancer, in concordance with the notion that breast cancer is a heterogeneous disease [2-26]. In this respect, the recent development of targeted trastuzumab therapy has indeed improved the survival of a subset of breast cancer patients that have tumors overexpressing the ERBB2 receptor tyrosine kinase [27-29]. However, appropriate molecular targets have as yet not been identified for most breast cancer subtypes, implying suboptimal treatment for a significant fraction of the breast cancer patients. Thus, a better understanding of the disease is needed to improve upon current methods to treat breast cancer patients.

Breast cancer is a genetic disease in which gene mutations may be inherited or acquired somatically. Recently, large scale re-sequencing efforts have suggested that there are as many as 15 oncogenic driver mutations present in a single breast tumor [30, 31]. One of the genes that is frequently inactivated in breast cancer is the tumor suppressor gene *E-cadherin* [32, 33]. The E-cadherin protein is essential in maintaining epithelial tissue integrity through intercellular cell adhesion [34]. Loss of E-cadherin in human breast cancer can be achieved by either mutation of the *E-cadherin* gene, hypermethylation of its gene promoter or transcriptional silencing by its repressors [35, 36]. Mutations of *E-cadherin* were shown to be causal for the lobular phenotype of breast cancer [37]. However, silencing of *E-cadherin* by promoter hypermethylation or by transcriptional repression has not properly been assessed. Notably, a discrepancy exists between the loss of E-cadherin expression and the presence of *E-cadherin* gene mutations in breast cancer. First, only half of lobular breast cancers have mutated the *E-cadherin* gene although most lobular breast cancers have lost E-cadherin protein expression. Second, E-cadherin protein expression is absent or aberrant in a significant fraction of breast cancers of non-lobular pathology, but no mutations of *E-cadherin* have been found in non-lobular breast cancers. It is not clear how the various inactivation mechanisms of *E-cadherin* are involved human breast carcinogenesis, and certainly not in relation with particular subtypes of breast carcinoma.

The aim of this thesis was to gain insight in the various mechanisms of E-cadherin inactivation in human breast cancer. For this purpose, we have evaluated the inactivation mechanisms of E-cadherin using human breast cancer cell lines as a model in **chapter 3**. Gene expression profiling and gene reconstitution experiments revealed that E-cadherin inactivation by gene

mutation represents a distinct biological mechanism from E-cadherin inactivation by promoter hypermethylation and transcriptional repression. The identification of two distinct modes of E-cadherin inactivation that associate with distinct (histopathological) subtypes of breast cancer *in vitro* as well as *in vivo* is important as it challenges the paradigm that genetic and epigenetic inactivation of a tumor suppressor gene are biologically similar. The results of this study may also explain recurrent controversies in E-cadherin research and calls for re-evaluation of functional E-cadherin studies as well as studies on the clinical outcome of patients with E-cadherin-negative breast cancers.

In **chapters 4-7**, we sought to identify the genetic mechanisms that underlie the breast cancer subtypes that associated with the two different modes of E-cadherin inactivation. Although mutations of *E-cadherin* were shown to be causal for the lobular subtype of breast cancer, only half of lobular breast cancers had mutated *E-cadherin* genes [32, 33, 37]. Therefore, we have pursued the identification of other E-cadherin pathway members as a tumor suppressor gene causal in lobular breast cancers without *E-cadherin* mutations. In **chapter 4**, we provide evidence that α-*catenin* is a putative new tumor suppressor gene. The results of this study underline the importance of the E-cadherin/catenin protein complex in cancer, as α-catenin is already the third member of this complex of which mutations contribute to the development of cancer.

Breast cancer cell lines of the basal-type were associated with epigenetic inactivation of E-cadherin and EMT, through expression of E-cadherin's transcriptional repressors. Notably, EMT involves (the crosstalk of) multiple signaling pathways, including the PI3K and RAS pathways [36]. Also, mutations of the *BRCA1* gene have been associated with basal breast tumors [26]. Therefore, we investigated whether mutations in genes of the PI3K and RAS pathways and the *BRCA1* gene might be associated with breast cancer cell lines that had inactivated E-cadherin through epigenetic mechanisms. In **chapters 5 and 6**, we performed mutation analysis of the genes of the PI3K and RAS pathways and the *BRCA1* gene in all breast cancer cell lines. Interestingly, we found an association of mutations of genes of the RAS pathway and the *BRCA1* gene, but not genes of the PI3K pathway with epigenetic E-cadherin inactivation and basal-type breast cancer. The results of these studies may provide some clues to the underlying genetic events of the basal-type breast cancers and EMT.

Finally, we set out to provide a genetic basis for the two major breast cancer subtypes that associated with the two different modes of E-cadherin inactivation. In **chapter 7**, we have performed protein and gene expression analyses and large scale mutation screens of 20 other known cancer genes in the breast cancer cell lines. This study resulted in the identification of two distinct mutation profiles that associated with luminal and basal-type breast cancer cell lines. The results of this study may provide a further refinement of current molecular breast cancer classification and aid the development of new treatment modalities that target the here identified potential drug targets.

The results of this thesis are summarized in **chapter 8** and further discussed in **chapter 9**.

REFERENCES

1. Visser O, van Noord KJ, eds. Feiten en fabels over kanker in Nederland. Utrecht: Vereniging van Integrale Kankercentra; 2005.
2. van 't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature 2002;415(6871):530-6.
3. van de Vijver MJ, He YD, van't Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. The New England journal of medicine 2002;347(25):1999-2009.
4. Wang Y, Klijn JG, Zhang Y, et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. Lancet 2005;365(9460):671-9.
5. Foekens JA, Atkins D, Zhang Y, et al. Multicenter validation of a gene expression-based prognostic signature in lymph node-negative primary breast cancer. J Clin Oncol 2006;24(11):1665-71.
6. Desmedt C, Piette F, Loi S, et al. Strong time dependence of the 76-gene prognostic signature for node-negative breast cancer patients in the TRANSBIG multicenter independent validation series. Clin Cancer Res 2007;13(11):3207-14.
7. Paik S, Shak S, Tang G, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. The New England journal of medicine 2004;351(27):2817-26.
8. Paik S, Tang G, Shak S, et al. Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. J Clin Oncol 2006;24(23):3726-34.
9. Habel LA, Shak S, Jacobs MK, et al. A population-based study of tumor gene expression and risk of breast cancer death among lymph node-negative patients. Breast Cancer Res 2006;8(3):R25.
10. Sotiriou C, Wirapati P, Loi S, et al. Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. Journal of the National Cancer Institute 2006;98(4):262-72.
11. Chang HY, Sneddon JB, Alizadeh AA, et al. Gene expression signature of fibroblast serum response predicts human cancer progression: similarities between tumors and wounds. PLoS biology 2004;2(2):E7.
12. Chang HY, Nuyten DS, Sneddon JB, et al. Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival. Proceedings of the National Academy of Sciences of the United States of America 2005;102(10):3738-43.
13. Ma XJ, Wang Z, Ryan PD, et al. A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen. Cancer cell 2004;5(6):607-16.
14. Jansen MP, Sieuwerts AM, Look MP, et al. HOXB13-to-IL17BR expression ratio is related with tumor aggressiveness and response to tamoxifen of recurrent breast cancer: a retrospective study. J Clin Oncol 2007;25(6):662-8.
15. Jansen MP, Foekens JA, van Staveren IL, et al. Molecular classification of tamoxifen-resistant breast carcinomas by gene expression profiling. J Clin Oncol 2005;23(4):732-40.
16. Kok M, Linn SC, Van Laar RK, et al. Comparison of gene expression profiles predicting progression in breast cancer patients treated with tamoxifen. Breast cancer research and treatment 2008.
17. Chang JC, Wooten EC, Tsimelzon A, et al. Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer. Lancet 2003;362(9381):362-9.
18. Hess KR, Anderson K, Symmans WF, et al. Pharmacogenomic predictor of sensitivity to preoperative chemotherapy with paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide in breast cancer. J Clin Oncol 2006;24(26):4236-44.
19. Iwao-Koizumi K, Matoba R, Ueno N, et al. Prediction of docetaxel response in human breast cancer by gene expression profiling. J Clin Oncol 2005;23(3):422-31.
20. Ayers M, Symmans WF, Stec J, et al. Gene expression profiles predict complete pathologic response to neoadjuvant paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide chemotherapy in breast cancer. J Clin Oncol 2004;22(12):2284-93.

21. Folgueira MA, Carraro DM, Brentani H, et al. Gene expression profile associated with response to doxorubicin-based therapy in breast cancer. Clin Cancer Res 2005;11(20):7434-43.
22. Gianni L, Zambetti M, Clark K, et al. Gene expression profiles in paraffin-embedded core biopsy tissue predict response to chemotherapy in women with locally advanced breast cancer. J Clin Oncol 2005;23(29):7265-77.
23. Modlich O, Prisack HB, Munnes M, Audretsch W, Bojar H. Predictors of primary breast cancers responsiveness to preoperative epirubicin/cyclophosphamide-based chemotherapy: translation of microarray data into clinically useful predictive signatures. Journal of translational medicine 2005;3:32.
24. Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. Nature 2000;406(6797):747-52.
25. Sorlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proceedings of the National Academy of Sciences of the United States of America 2001;98(19):10869-74.
26. Sorlie T, Tibshirani R, Parker J, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proceedings of the National Academy of Sciences of the United States of America 2003;100(14):8418-23.
27. Cobleigh MA, Vogel CL, Tripathy D, et al. Multinational study of the efficacy and safety of humanized anti-HER2 monoclonal antibody in women who have HER2-overexpressing metastatic breast cancer that has progressed after chemotherapy for metastatic disease. J Clin Oncol 1999;17(9):2639-48.
28. Slamon DJ, Leyland-Jones B, Shak S, et al. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. The New England journal of medicine 2001;344(11):783-92.
29. Vogel CL, Cobleigh MA, Tripathy D, et al. Efficacy and safety of trastuzumab as a single agent in first-line treatment of HER2-overexpressing metastatic breast cancer. J Clin Oncol 2002;20(3):719-26.
30. Sjoblom T, Jones S, Wood LD, et al. The consensus coding sequences of human breast and colorectal cancers. Science (New York, NY 2006;314(5797):268-74.
31. Wood LD, Parsons DW, Jones S, et al. The genomic landscapes of human breast and colorectal cancers. Science (New York, NY 2007;318(5853):1108-13.
32. Berx G, Cleton-Jansen AM, Nollet F, et al. E-cadherin is a tumour/invasion suppressor gene mutated in human lobular breast cancers. The EMBO journal 1995;14(24):6107-15.
33. Berx G, Cleton-Jansen AM, Strumane K, et al. E-cadherin is inactivated in a majority of invasive human lobular breast cancers by truncation mutations throughout its extracellular domain. Oncogene 1996;13(9):1919-25.
34. Takeichi M. Cadherin cell adhesion receptors as a morphogenetic regulator. Science (New York, NY 1991;251(5000):1451-5.
35. Berx G, Van Roy F. The E-cadherin/catenin complex: an important gatekeeper in breast cancer tumorigenesis and malignant progression. Breast Cancer Res 2001;3(5):289-93.
36. Peinado H, Olmeda D, Cano A. Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? Nat Rev Cancer 2007;7(6):415-28.
37. Derksen PW, Liu X, Saridin F, et al. Somatic inactivation of E-cadherin and p53 in mice leads to metastatic lobular mammary carcinoma through induction of anoikis resistance and angiogenesis. Cancer cell 2006;10(5):437-49.

Epigenetic Silencing and Mutational Inactivation of E-cadherin Associate with Distinct Breast Cancer Subtypes

Antoinette Hollestelle
Justine K. Peeters
Marcel Smid
Leon C. Verhoog
Pieter J. Westenend
Mieke Timmermans
Alan Chan
Jan G.M. Klijn
Peter J. van der Spek
John A. Foekens
Michael A. den Bakker
Mieke Schutte

Submitted for publication

ABSTRACT

Loss of E-cadherin expression has been observed for most human epithelial tumor types, but *E-cadherin* gene mutations have only been identified in half of lobular breast cancers and diffuse gastric cancers. We have investigated *E-cadherin* inactivation in a collection of 41 human breast cancer cell lines. We identified ten *E-cadherin* gene mutants and twelve cell lines with epigenetic *E-cadherin* silencing due to promoter methylation. Cell lines with silenced *E-cadherin* were quite distinct from *E-cadherin* mutants, including a vastly different intrinsic gene expression program that involved *E-cadherin* transcriptional repressors and genes associated with cancer stem cells and EMT, a spindle versus a rounded cell morphology, and resemblance to basal versus luminal-type breast cancers. Loss of wild-type *E-cadherin* expression was causative for the rounded cell morphology but not for the spindle cell morphology, further implying that the two modes of *E-cadherin* inactivation are fundamentally different. A 3-protein spindle cell signature defined on breast cancer cell lines with spindle cell morphology indeed associated with clinical breast cancers of the basal type, where E-cadherin protein loss was particularly pronounced in a pathological subtype of metaplastic breast cancer. Importantly, metaplastic breast cancers are typified by transdifferentiated components, also suggesting that E-cadherin's role in EMT may be restricted to breast cancers of the basal type. Our evidence for two biologically distinct modes of *E-cadherin* inactivation challenges the paradigm that mutational inactivation and epigenetic silencing of tumor suppressor genes are functionally similar. It also may explain recurrent controversies in E-cadherin research and calls for re-evaluation of functional E-cadherin studies as well as studies on the clinical outcome of patients with E-cadherin-negative breast cancers.

INTRODUCTION

The mammary gland consists of secretory acini connected by a tree of branching ducts, embedded in a stromal compartment. The epithelium that lines the ductal system is composed of two layers, a luminal epithelial layer and a basal myoepithelial layer. The luminal epithelial cells are cuboidal-shaped and form a polarized continuous layer that lines the lumen. The basal myoepithelial cells are typically spindle or stellate-shaped and reside between the luminal epithelial layer and the basement membrane, forming a nearly continuous layer in the ducts and a discontinuous basket-like structure around the lobular acini. Apart from their morphological appearance and localization, luminal epithelial cells are distinguished from basal myoepithelial cells by their cytokeratin (CK) protein expression profiles: luminal cells express luminal or simple cytokeratins CK7, CK8, CK18 and/or CK19 and basal myoepithelial cells express basal or stratified cytokeratins CK5, CK14 and/or CK17, although it has been noted that luminal cells may also express basal cytokeratins [1-6]. Several other proteins are differentially expressed by the two epithelial layers, such as epithelial membrane antigen (EMA, also known as MUC1), estrogen receptor alpha (ERα) and the progesterone receptor (PR) in luminal epithelial cells, and CD10/CALLA, smooth muscle actin (SMA), p63 and the epidermal growth factor receptor (EGFR) in basal myoepithelial cells [7-15]. A series of seminal studies on isolated mammary epithelial cell populations have formed the basis for an epithelial differentiation model in which CK5+ committed stem cells differentiate into bipotent precursor cells (CK5+, CK8/18+) that are positioned suprabasal in the luminal epithelium and produce fully differentiated cells of both the luminal epithelial lineage (CK8/18+) and the basal myoepithelial lineage (SMA+) [16-21]. The existence of a common precursor for luminal and basal mammary epithelia unmistakably illustrates the intricate alliance and relatedness of the epithelial and myoepithelial cell layers in the mammary gland.

Most breast cancers arise in the terminal ductal lobular unit. Pathological classification of breast cancers is based on cytological and architectural features. Ductal type breast carcinoma constitutes about two-thirds of breast cancers and lobular breast cancer accounts for 10-15%. Less prevalent pathological subtypes include tubular, mucinous (colloid), medullary and metaplastic breast cancers [22, 23]. Histological classification of breast cancers by cytokeratin expression has become another major determinant in breast cancer classification. Combined analyses revealed that some 60-80% of primary invasive breast cancers express only luminal cytokeratins, 20-40% express both luminal and basal cytokeratins, and a minority expresses only basal cytokeratins or is negative for both luminal and basal markers (each <1%) [3, 24-26]. Breast cancers can also be classified based on their gene expression profiles. Five intrinsic subtypes of breast cancer were defined based on the intrinsic gene set: luminal A, luminal B, ERBB2+, basal-like and normal-like [27-29]. The basal-like intrinsic subtype could be accurately identified by immunohistochemistry using a 4-protein signature (ERBB2, ER, CK5 and EGFR) [30]. This 4-protein signature classified four groups of breast cancers: ERBB2 overexpressing (ERBB2+), luminal (ERBB2- and ER+), basal-like

(ERBB2/ER- and CK5+ and/or EGFR+), and the negative group that lacks expression of all four proteins.

E-cadherin is a major determinant in maintaining epithelial cell integrity. The E-cadherin transmembrane protein is expressed in the adherence junctions of epithelial cells and mediates homophilic cell-cell adhesion between E-cadherin molecules on adjacent cells. With its distal part of the cytoplasmic tail, E-cadherin interacts with either β-catenin or γ-catenin, which are mutually exclusive in the E-cadherin-catenin protein complex. α-Catenin proteins, in their turn, interact with either the actin cytoskeleton or with β-catenin or γ-catenin in a dynamic fashion [31-39]. Membrane-proximal at the cytoplasmic tail of E-cadherin binds p120-catenin (p120ctn), a protein involved in stability and lateral clustering of E-cadherin at the plasma membrane [40-42]. Besides their role in E-cadherin mediated cell adhesion, catenins also function in Wnt signaling and in dynamic organization of the actin cytoskeleton through regulation of the Rho GTPases and through interaction with various actin binding molecules [43-46]. E-cadherin is a suppressor of tumor invasion and loss of E-cadherin has been noted for most human epithelial tumor types [47-49]. Although this suggests a tumor suppressor function of E-cadherin in multiple tumor types, inactivating *E-cadherin* gene mutations have only been identified in breast cancers and gastric cancers. Importantly, mutations were found in about half of lobular breast cancers and in about half of diffuse gastric cancers [50-53], but not in other subtypes of breast cancer or gastric cancer. Lobular breast cancer and diffuse gastric cancer are both characterized by a typical pathological appearance of diffusely growing, rounded cells with scant cytoplasm. *E-cadherin* gene mutations thus appear to have a profound effect on cell morphology. Unexpected was the absence of *E-cadherin* gene mutations among carcinomas from other anatomical sites, or in the remaining breast cancers and gastric cancers. Loss of E-cadherin expression in these carcinomas was suggested to involve transcriptional silencing in association with methylation of CpG islands in the *E-cadherin* promoter region [54-57]. In this respect, several transcriptional repressors of E-cadherin have been identified: SNAIL, SLUG, SIP1, δEF1, E47, and TWIST [57-64]. Expression of E-cadherin transcriptional repressors has indeed been observed for various carcinoma types and has been associated with epithelial to mesenchymal transition (EMT) and a more aggressive clinical course [57]. To gain further insight in loss of E-cadherin expression in tumorigenesis, we studied a model of 41 human breast cancer cell lines. We provide evidence that inactivation of *E-cadherin* through mutation of the gene is biologically distinct from epigenetic silencing associated with promoter methylation, where gene mutations associated with a rounded cell morphology and promoter methylation with a spindle cell morphology. We also show that these different modes of *E-cadherin* inactivation associate with clinical breast cancers of the luminal type and basal type, respectively, particularly those of lobular and metaplastic pathology.

MATERIALS AND METHODS

Breast cancer samples

The 41 human breast cancer cell lines used in this study are listed in Table 3.1. Cell lines EVSA-T, MPE600, SK-BR-5 and SK-BR-7 were kind gifts of Dr. N. de Vleesschouwer (Institut Jules Bordet, Brussels, Belgium), Dr. H.S. Smith (California Pacific Medical Center, San Francisco, CA), and Dr. E. Stockert (Sloan-Kettering Institute for Cancer Research, New York, NY), respectively. The SUM cell lines were generated in the Ethier laboratory (available at http://www.asterand.com). Cell line OCUB-F was obtained from Riken Gene Bank (Tsukuba, Japan), and all other cell lines were obtained from ATCC (Manassas, VA). All cell lines are unique and monoclonal as shown by extensive analysis of nearly 150 polymorphic microsatellite markers [65].

Formalin-fixed paraffin-embedded primary breast cancer specimens were obtained from the Department of Pathology archive at Erasmus MC in Rotterdam, selected from diagnosis years 1990 through 2005. Twenty-four ER-positive and 22 ER-negative ductal breast cancers and 28 lobular, 20 mucinous, 16 tubular and 22 medullary breast cancers were selected. The 34 metaplastic breast cancers were obtained from various hospitals from the Rotterdam/Dordrecht area. All tumor samples were re-evaluated by pathologists with a special interest in breast pathology (MdB and PJW) with metaplastic breast cancers being diagnosed according to criteria based on the WHO classification [22]. Tissue microarrays were constructed by punching three 0.6-mm cores from representative areas of each tumor and transferring them into a recipient paraffin block by using an ATA27 automated tissue microarrayer (Beecher Instruments, Sun Prairie, WI).

The Medical Ethical Committee at Erasmus MC has approved the study, which was carried out according the Code of Conduct of the Federation of Medical Scientific Societies in the Netherlands.

Gene mutation analysis

E-cadherin (*CDH1*; Genbank #Z13009) mutations were identified by direct sequencing of PCR-amplified genomic sequences of exons 2 through 16, using an ABI 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA). All mutations were confirmed by sequencing of an independently amplified DNA template. Mutation analysis of 26 breast cancer cell lines had been reported previously [66]. Allelic loss of *E-cadherin* was determined by PCR-based microsatellite analysis using markers *D16S421*, *D16S496*, *D16S2621* and *D16S2624*, where a homozygous allele pattern for all four markers was interpreted as allelic loss, with $P<0.01$ [65].

Promoter methylation analysis

For azacytidine assays, exponentially growing cells were treated for three days with 10 μM filter-sterilized 5-aza-2'-deoxycytidine (Sigma, Steinheim, Germany). On the fourth day, cells were washed with PBS at 37°C, harvested by lysis in the flask, and total RNA was isolated. As a control,

cultures without 5-aza-2'-deoxycytidine were taken along. *E-cadherin* transcript expression was determined by duplex RT-PCR with the *HPRT* housekeeper.
Methylation-specific PCR (MSP) was performed on bisulphate-treated genomic DNA using the EZ Methylation Kit (Zymed, Orange, CA). MSP for *E-cadherin* CpG islands 1 and 3 was done by using reported primers [67] under our standard PCR conditions [66].

Expression analysis

Duplex RT-PCR for *E-cadherin* and *HPRT* was done using the Qiagen (Hilden, Germany) one-step RT-PCR kit and gene-specific exonic primers.
Quantitative RT-PCR (qPCR) was performed on an ABI 7700 Taqman Analyzer (Applied Biosystems), using cDNA templates generated with oligo-dT and random hexamer primers and SYBR Green PCR Master mix (Applied Biosystems). All qPCR Ct-values were normalized according *HPRT* and *PBDG* housekeeper Ct-values. For *SNAIL*, *SLUG*, *SIP-1*, δ*EF-1* and *TWIST*, normalized Ct-values <30 were scored expressed (+), Ct-values from 30 through 35 were scored weakly expressed (+/-), and Ct-values >35 were scored not expressed (-). All primer sequences are provided in the Supplementary Data. Specificity of primers for *SNAIL, SLUG*, *SIP-1*, δ*EF-1* and *TWIST* was confirmed by sequencing of the amplification products.
Western blotting was performed as described [68], using the following antibodies: mouse anti-E-cadherin (BD Transduction Laboratories (Franklin Lakes, NJ), clone 36 and Zymed, clone HECD-1), mouse anti-β-Catenin (BD Transduction Laboratories, clone 14), mouse anti-E2A (Santa Cruz Biotechnology (Santa Cruz, CA), clone Yea), mouse anti-GAPDH (Chemicon (Billerica, MA), clone MAB374) and rabbit-anti-mouse horse radish peroxidase conjugated secondary antibodies (DAKO, Glostrup, Denmark).
Immunohistochemistry was performed by autostaining slides with Chemmate Envision+ kit (DAKO). Slides were first boiled in Tris-EDTA buffer pH=9 (Klinipath, Duiven, The Netherlands) for 20 minutes to unmask antigens. Primary monoclonal antibodies were: Caldesmon-1 (Novocastra (Newcastle upon Tyne, UK), clone TD107), Caveolin-1 (BD Transduction Laboratories, clone 2297), CDH1 (DAKO, clone NCH-38), CK5 (Novocastra, clone XM26), CK8-18 (Biogenex (San Ramon, CA), clone NCL5D3), EGFR (DAKO, EGFR PharmDx™ Kit), ER (DAKO, clone 1D5), ERBB2 (DAKO, HercepTest™), PR (DAKO, clone PgR 636) and Vimentin (DAKO, clone V9).

Transfections

Human *E-cadherin* wild-type and delEx9 cDNA's were RT-PCR amplified from breast cancer cell lines T47D and MPE600. The cDNAs were cloned into pcDNA3.0-Neo expression vector (Invitrogen, Paisley, Scotland) and inserts were verified by sequencing and restriction endonuclease digestions. Transfections were performed using Fugene-6 transfection reagent (Roche, Basel, Switzerland) and cells were grown at limiting dilutions in 96-well plates under G418 selection (Invitrogen). All transfection clones were evaluated for morphological growth pattern and for

E-cadherin and β-catenin protein expression, and checked for presence of the correct *E-cadherin* cDNA and genotype of the mother cell line.

Gene expression profiling

Breast cancer cell lines were grown to optimal cell densities. The culture medium of the cells was changed 16-20 hours before harvesting (at 37°C). RNA was isolated using the RNeasy kit (Qiagen) upon lysing the cells in the culture flask. DNAseI-treated RNA was antisense biotinylated using the MEGAScript T7 labeling kit (Ambion, Foster City, CA) and Affymetrix U133A microarrays were hybridized according Affymetrix GeneChip Manual, both performed on commercial basis by ServiceXS (Leiden, the Netherlands). Intensity values for all genes were scaled using the global normalization factor as specified by GCOS 1.1, and further normalized with Omniviz software 3.6 (Biowisdom, Maynard, MA). Intensity values <30 were set to 30. Differential gene expression was based on log2 transformed distances to the geometric mean for each probe set. Patterns of correlation were revealed by applying the Pearson matrix-ordering method that sorts samples into correlated blocks, resulting in highly similar plots and identical groupings for log2GM <-2 and >2 or log2GM <-3 and >3 probe subsets.

The Stanford intrinsic gene set for clinical breast cancers [27-29] was translated into an Affymetrix intrinsic gene set, including 451 probe sets from the Stanford list of 496 genes [69]. Breast cancer cell lines were classified for their intrinsic subtype by this Affymetrix intrinsic gene set, using average distance linkage hierarchical clustering with non-centered correlation as distance metric [70].

The transcript-based spindle cell signature was determined by Significance Analysis of Microarrays (SAM; [71]) within Omniviz software package. The criteria in identifying the top 1144 genes with significant differential expression between the spindle and non-spindle breast cancer cell lines were: falsely called median <1, false discovery rate <1, and q-values <1%. This spindle cell signature was validated by qPCR analysis of nine signature genes (*EMP3, FXYD3, SPDEF, VIM, RAB25, CLDN7, BSPRY, TACSTD1* and *ARHGAP8*) in 36 breast cancer cell lines. Affymetrix gene expression levels of these genes correlated very well with their qPCR expression levels, with average Spearman correlation coefficient of 0.85 (range 0.77-0.91; $p<0.001$). Primer sequences are provided in the Supplementary Data.

The protein-based spindle cell signature was defined by selection of 16 monoclonal antibodies based on their significance in the transcript-based spindle cell signature, reported relevance for breast cancer and availability. Nine antibodies that proved reliable in immunohistochemistry were evaluated for their ability to distinguish spindle breast cancer cell lines from non-spindle cell lines, demanding 100% specificity for non-spindle cell lines to minimize false positive rates. Five antibodies met these criteria (Supplementary Data), of which Caveolin-1, Caldesmon-1 and Vimentin were selected for their high specificity for spindle cell lines (100%, 73% and 73%, respectively).

Table 3.1 Molecular status of human breast cancer cell lines with respect to E-cadherin and related proteins.

Breast cancer cell line	Cell morphology	Allelic loss at 16q	*E-cadherin* gene sequence	*E-cadherin* methylation	*E-cadherin* transcript
HCC1937	epithelial	no	wild-type	-	+
MCF-7	epithelial	no	wild-type	-	+
BT474	epithelial	no	wild-type	-	+
UACC812	epithelial	no	wild-type	-	+
BT483	epithelial	no	wild-type	-	+
MDA-MB-175VII	epithelial	no	wild-type	-	+
MDA-MB-361	epithelial	no	wild-type	-	+
MDA-MB-415	epithelial	no	wild-type	-	+
SUM52PE	epithelial	yes	wild-type	-	+
SUM190PT	epithelial	no	wild-type	nd	+
SUM225CWN	epithelial	yes	wild-type	nd	+
SUM185PE	epithelial	yes	wild-type	-	+
T47D	epithelial	yes	wild-type	-	+
ZR75-1	epithelial	yes	wild-type	-	+
UACC893	epithelial	no	wild-type	-	+
MDA-MB-330	rounded	no	wild-type	-	+
MDA-MB-468	rounded	yes	wild-type	-	+
CAMA-1	rounded	yes	in-frame mutant	-	+
EVSA-T	rounded	yes	in-frame mutant	-	+
MPE600	rounded	yes	in-frame mutant	-	+
OCUB-F/-M	rounded	yes	truncating mutant	-	+
SUM44PE	rounded	yes	truncating mutant	nd	+
MDA-MB-134VI	rounded	yes	truncating mutant	-	+
SK-BR-5	rounded	yes	truncating mutant	-	+/-
MDA-MB-453	rounded	yes	truncating mutant	-	+/-
ZR-75-30	rounded	yes	truncating mutant	-	+/-
SK-BR-3	rounded	yes	deletion mutant	-	-
BT549	spindle	yes	wild-type	CpG1/CpG3	-
Hs578T	spindle	no	wild-type	CpG1/CpG3	-
MDA-MB-435s	spindle	yes	wild-type	CpG1/CpG3	-
MDA-MB-436	spindle	no	wild-type	CpG1/CpG3	-
SUM159PT	spindle	no	wild-type	CpG1/CpG3	-
SUM1315MO2	spindle	no	wild-type	CpG1/CpG3	-
MDA-MB-157	spindle	no	wild-type	CpG1/CpG3	+/-
MDA-MB-231	spindle	yes	wild-type	CpG1/CpG3	+/-
SK-BR-7	spindle	no	wild-type	CpG1/CpG3?	+
SUM102PT	spindle	no	wild-type	CpG1	+
SUM149PT	spindle	no	wild-type	CpG1	+
SUM229PE	spindle	no	wild-type	CpG1	+
BT20	spindle	yes	wild-type	-	+
DU4475	other	no	wild-type	-	+

Breast cancer cell lines are organized by their morphology and then by their *E-cadherin* status. *E-cadherin* gene mutations are detailed in the Supplementary Data and *E-cadherin* methylation has been indicated for CpG islands 1 and 3 in the promoter region. nd, not determined; -, negative or absent; +/-, low or barely detectable expression; +, clearly detectable expression.

E-cadherin protein	β-catenin protein	E47 protein	*SIP1* transcript	*δEF1* transcript	*SLUG* transcript	*TWIST* transcript	*SNAIL* transcript
+	+	-	-	+/-	+	+/-	+
+	+	-	-	+/-	+	+/-	+
+	+	-	-	+/-	+	+	+
+	+	-	-	+/-	+	+	+
+	+	-	-	-	+	+	+
+	+	-	-	-	+	+	+
+	+	-	-	-	+	+	+
+	+	-	-	-	+	-	+
+	+	-	-	-	+	-	+
+	+	-	nd	nd	+	nd	+
+	+	-	nd	nd	+	nd	+
+	+	-	-	-	+/-	+	+
+	+	-	-	-	-	+	+
+	+	-	-	-	-	+	+
+	+	-	-	+/-	-	-	+
+	+	-	-	+	-	+	+/-
+	+	-	-	-	+/-	+/-	+
+	+	-	-	+/-	+/-	+	+
+	+	-	-	-	-	+	+
+	+	-	-	-	+/-	+	+
-	+/-	-	-	+/-	+/-	+	+/-
-	+/-	-	-	+/-	+/-	+	+
-	-	-	-	+/-	+/-	+	+
-	-	-	-	-	+/-	+	+
-	+/-	-	-	-	+/-	-	+/-
-	-	-	-	-	+/-	+	+
-	+/-	-	-	-	+/-	+/-	+
-	+	+	+	+	+/-	+	+
-	+	+	+	+	+	+	+
-	+	+	+	+	+	+	+
-	+	+	+	+	+	+	+
-	+	+	+	+	+	+	+
-	+	+	+	+	+	+	+/-
-	+	+	+	+	+	+	+
-	+	+	+	+	+	-	+/-
+	+	+	+/-	+	+/-	+	+
+	+	+/-	+	+	+	+	+
+	+	-	+	+/-	+	+	+
+	+	+/-	+/-	+/-	+	+	+/-
+	+	-	-	+	+	-	+/-
+	+	-	-	+	+	-	+

RESULTS

A rounded cell morphology typifies *E-cadherin* mutant breast cancer cell lines

We have used a human breast cancer cell line model to investigate the biological basis of *E-cadherin* inactivation. An earlier *E-cadherin* sequence analysis [66] was now extended to all 41 breast cancer cell lines from our collection, revealing two more *E-cadherin* mutant cell lines in addition to the eight mutants that already had been identified (Table 3.1; mutations are detailed in the Supplementary Data). Together, the collection contains one cell line with deletion of the major part of the *E-cadherin* gene, six cell lines with *E-cadherin* mutations that result in premature stop codons and three mutants with an in-frame deletion. All ten mutants had lost the other *E-cadherin* allele, consistent with the tumor suppressor function of *E-cadherin* [50, 51].

Strikingly, all ten *E-cadherin* mutant cell lines grow with rounded cells with scant cytoplasm (Figure 3.1). These rounded cells may grow in clusters of cells, varying from grape-like bunches to so-called Indian files of cells, or as single cells, that may be attached to adherent cells or freely floating in the culture medium. The cell cultures typically also contain adherent cells that grow as epithelial sheets with diminished cell-cell adhesion, with a cell line-specific percentage of adherent cells that varies from less than ten percent to over ninety percent of the cell population. In addition to the ten *E-cadherin* mutant cell lines, two *E-cadherin* wild-type cell lines also have this rounded cell morphology (Table 3.1). The presence of small rounded cells with scant cytoplasm is a cytological characteristic of the lobular pathological subtype of human breast cancer. Also, *E-cadherin* gene mutations are identified in half of lobular cancers, but not in other subtypes of breast cancer [50, 51]. The twelve breast cancer cell lines with the rounded cell morphology thus appear to resemble lobular breast cancers. Consistent with this notion, four *E-cadherin* mutant cell lines from our collection were known to be derived from breast cancers with lobular characteristics (EVSA-T, MDA-MB-134VI, MDA-MB-330 and SUM44PE [72-74] and personal communication Dr. S.P. Ethier).

Apart from these rounded cell lines, two other major morphology groups were apparent among the breast cancer cell lines (Table 3.1 and Figure 3.1). First, fifteen "epithelial" cell lines grow in sheets of adherent epithelial cells or in spheroid-like cell clusters. Second, thirteen "spindle" cell lines grow rather similar to fibroblasts, yet with extensions that are less pronounced than those of fibroblasts. Cell line DU4475 could not be assigned to either of these three morphology groups. DU4475 is an atypical breast cancer cell line in that it carries an *APC* gene mutation [75] and has constitutive Wnt signaling activation [66]. All 29 non-rounded cell lines had *E-cadherin* wild-type genes, whether epithelial, spindle or unclassified.

A spindle cell morphology typifies breast cancer cell lines with *E-cadherin* promoter methylation

We have evaluated the effects of the *E-cadherin* gene mutations by expression analysis. *E-cadherin* transcripts were detected by duplex RT-PCR with the *HPRT* housekeeper, and E-cadherin proteins were detected by western blotting using HECD-1 antibody (Table 3.1 and Figure 3.2

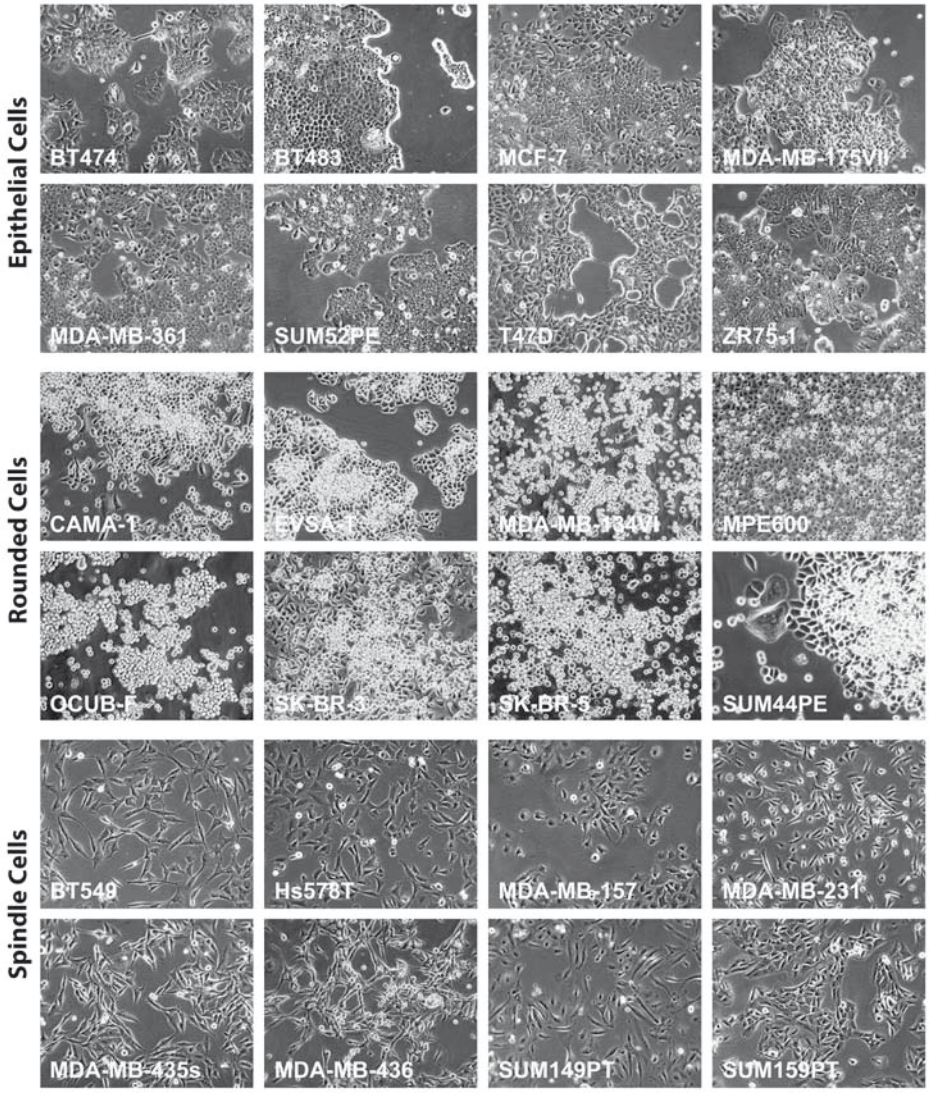


Figure 3.1 Morphology of human breast cancer cell lines. Eight examples are given for each of three morphology groups: epithelial cells that all have wild-type *E-cadherin* genes and apparently normal E-cadherin protein expression; rounded cells that all have mutant *E-cadherin* genes; and spindle cells that all have methylation at the *E-cadherin* promoter region.

A and B). Twelve breast cancer cell lines had low or no detectable *E-cadherin* transcript levels. These cell lines included four of the seven truncating *E-cadherin* mutant cell lines. As expected, none of these four cell lines, nor any of the other three truncating *E-cadherin* mutants expressed detectable levels of E-cadherin proteins. All three in-frame *E-cadherin* mutant cell lines, however, expressed *E-cadherin* transcripts and proteins at apparently normal levels. The eight other

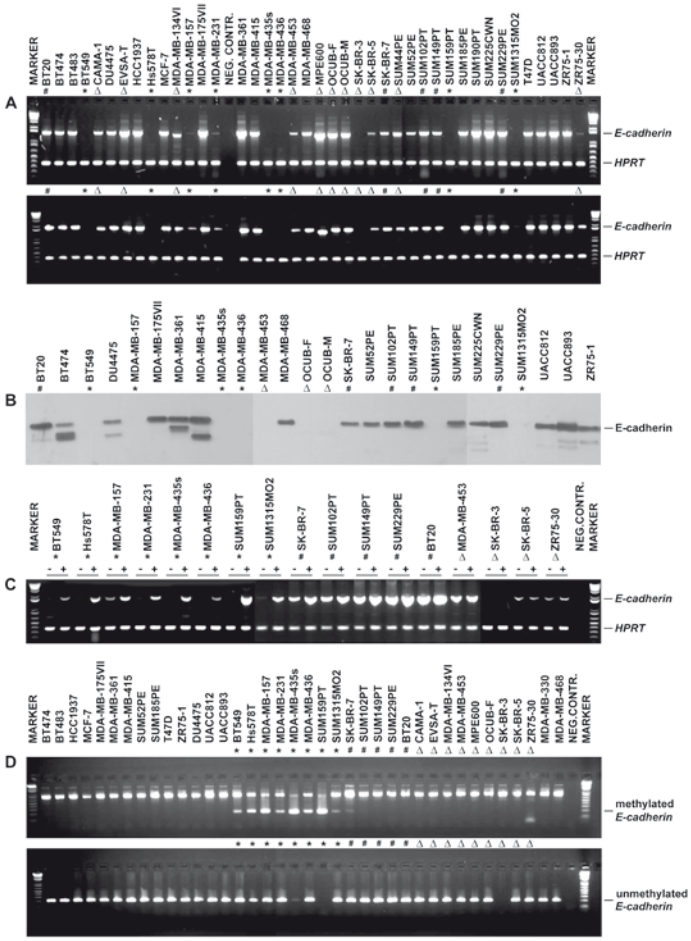


Figure 3.2 *E-Cadherin* expression and methylation analysis of human breast cancer cell lines. (A), *E-cadherin* transcript expression by duplex RT-PCR with the *HPRT* housekeeper, using primers directed at overlapping 5′ and 3′ fragments of the *E-cadherin* transcript (top and bottom panel, respectively). Low or barely detectable transcript levels were identified in cell lines with truncating *E-cadherin* mutations or in cell lines with *E-cadherin* CpG1 and CpG3 methylation. (B), E-cadherin protein expression by western blotting with HECD-1 antibody that is directed at an extracellular epitope. Spindle cell lines with *E-cadherin* methylation at CpG1 and CpG3 do not express E-cadherin proteins, whereas spindle cell lines with methylation at only CpG1 have detectable protein expression. (C), *E-cadherin* transcript expression by duplex RT-PCR with the *HPRT* housekeeper, upon azacytidine methylation assays. + and -, cells were cultured in the presence or absence of 5-aza-2-deoxycytidine. *E-cadherin* transcript expression was induced upon azacytidine treatment in *E-cadherin* methylated cell lines but not in *E-cadherin* mutant cell lines. (D), Methylation-specific PCR for CpG3 at the *E-cadherin* promoter region, using primers specific for methylated or unmethylated CpG3 (top and bottom panel, respectively). Whereas 12 of 13 spindle cell lines had methylation at CpG1, only eight or perhaps nine spindle cell lines had methylation at CpG3. Loss of E-cadherin protein expression was observed only for cell lines with methylation at both CpG islands (Table 3.1). Δ, Rounded cell lines with a mutant *E-cadherin* gene; #, Spindle cell lines with E-cadherin protein expression ∗, Spindle cell lines without E-cadherin protein expression; Neg. control, template-negative amplification reaction; Marker, 1-kb+ ladder (Invitrogen).

breast cancer cell lines with low or no detectable *E-cadherin* transcripts indeed also did not express E-cadherin proteins. Remarkably, these eight cell lines all grow with the spindle cell morphology (Table 3.1). Five other breast cancer cell lines from our collection also have the spindle cell morphology, but these five cell lines all expressed *E-cadherin* transcripts and proteins. Note that all spindle cell lines had *E-cadherin* wild-type genes. E-cadherin proteins were also expressed at apparently normal levels in all epithelial and unclassified cell lines. Together, these observations suggest that silencing of *E-cadherin* expression is not always due to mutation of the gene. In fact, the typical spindle cell morphology of the silenced *E-cadherin* wild-type cell lines suggests that their mechanism of *E-cadherin* inactivation may be biologically distinct from inactivation through *E-cadherin* gene mutation, as the latter always correlates with a rounded cell morphology.

A biological distinction between E-cadherin loss in spindle and rounded cell lines was also suggested by the β-catenin protein expression pattern that we observed for the breast cancer cell lines by western blotting (Table 3.1). Expression of β-catenin proteins was not or barely detectable in the seven rounded breast cancer cell lines with truncating *E-cadherin* mutations, suggesting that stable β-catenin protein expression requires E-cadherin protein expression. Indeed, the three *E-cadherin* mutant cell lines that had in-frame deletions — and thus had retained their intracellular β-catenin binding domain — expressed both E-cadherin and β-catenin proteins. Also, all breast cancer cell lines had apparently normal expression levels of β-*catenin* transcripts (data not shown). However, β-catenin proteins were expressed at apparently normal levels in the eight spindle cell lines that had lost E-cadherin expression. It thus appears that the β-catenin expression pattern in the rounded cell lines mirrors their E-cadherin expression pattern, whereas expression of the two proteins is uncoupled in the spindle cell lines. These observations again suggest that loss of E-cadherin expression may involve biologically distinct mechanisms in the rounded and spindle breast cancer cell lines.

Four of the spindle cell lines had reportedly silenced *E-cadherin* gene expression in association with methylation of CpG islands in its promoter region (Hs578T, MDA-MB-231 and MDA-MB-435s; [55, 76]). To investigate the extent of methylation inactivation in our breast cancer cell line model, we determined methylation-associated silencing of *E-cadherin* by azacytidine treatment of the cell lines combined with methylation-specific PCR (MSP) for CpG islands 1 and 3 (CpG1 and CpG3, respectively) [67] (Table 3.1 and Figure 3.2 C and D). MSP for CpG3, that is located nearest to the transcription start site, identified methylation in all eight E-cadherin-negative spindle cell lines and weak methylation in a single E-cadherin-expressing spindle cell line. No methylation was detectable in twenty-nine other cell lines (Figure 3.2 D). CpG1 MSP products were detected for twelve of the thirteen spindle cell lines, albeit that the methylation levels of cell lines that were also methylated at CpG3 appeared somewhat higher. Accordingly, the same twelve spindle cell lines also showed upregulation of *E-cadherin* in the azacytidine assay, whereas transcript expression was not restored in any of the four cell lines that had no or low E-cadherin expression due to mutation of the *E-cadherin* gene (Figure 3.2 C). Thus, *E-cadherin*

promoter methylation associates with a spindle cell morphology albeit that complete down regulation of *E-cadherin* transcript and protein expression involves methylation at both CpG1 and CpG3.

Ectopic E-cadherin expression restores an epithelial cell morphology in *E-cadherin* mutant cells, but not in *E-cadherin* methylated cells

We investigated causality of E-cadherin expression in determining cell morphology, by ectopic expression of human *E-cadherin* wild-type cDNA in the *E-cadherin* mutant cell line SK-BR-3 and in the *E-cadherin* methylated cell line MDA-MB-231. Cell line SK-BR-3 has a homozygous deletion of the major part of the *E-cadherin* gene and grows with rounded cells, whereas cell line MDA-MB-231 has an *E-cadherin* wild-type gene and grows with spindle cells (Table 3.1 and Figure 3.1). Neither of the two cell lines express detectable levels of *E-cadherin* proteins (Figure 3.2 A and B), rendering them suitable models for *E-cadherin* reconstitution experiments. Also, SK-BR-3 expresses very low levels of β-catenin proteins while MDA-MB-231 expresses normal levels of β-catenin proteins, even though both cell lines express β-*catenin* transcripts at apparently normal levels (Table 3.1). Cell lines SK-BR-3 and MDA-MB-231 were also reconstituted with *E-cadherin* delEx9 cDNA that contains an in-frame deletion of exon 9 of the gene, and with the empty vector. The in-frame deletions of *E-cadherin* that we and others have observed in tumor samples always locate to the extracellular calcium binding domains of the protein and are known to abolish E-cadherin's adhesion function [53, 77]. The delEx9 deletion had been identified in breast cancer cell line MPE600 (Supplementary Data) and in several clinical specimens of diffuse gastric cancer [52, 66, 78], suggesting that this deletion indeed is of functional relevance in human tumorigenesis. Importantly, MPE600 cells still express E-cadherin proteins and β-catenin proteins, rendering the delEx9 cDNA an appropriate negative control. We generated independent stable *E-cadherin* transfectants for each of four reconstitution combinations (2 cell lines x 2 *E-cadherin* constructs), by G418-selective growth under limiting dilution conditions (resulting in monoclonality with $P<0.05$). All of about 25 clones that were collected from each reconstitution combination were analyzed for cell morphology and for E-cadherin and β-catenin protein expression by immunohistochemistry using C-terminal antibodies.

Three SK-BR-3 clones that had been transfected with *E-cadherin* wild-type cDNA expressed both E-cadherin and β-catenin proteins and all three clones had converted from the rounded cell morphology to the epithelial cell morphology (Figure 3.3). Six *E-cadherin* delEx9 SK-BR-3 clones also expressed both E-cadherin and β-catenin proteins, but none of them had converted to the epithelial cell morphology, suggesting that a wild-type *E-cadherin* gene was pivotal for the morphology conversion (Figure 3.3). None of the other SK-BR-3 clones (wild-type, delEx9, and empty vector) expressed E-cadherin or β-catenin proteins, and neither had they converted to the epithelial cell morphology. These results imply that inactivation of *E-cadherin* through mutation of the gene is causal in determining the rounded cell morphology of the mutant breast cancer cell lines.

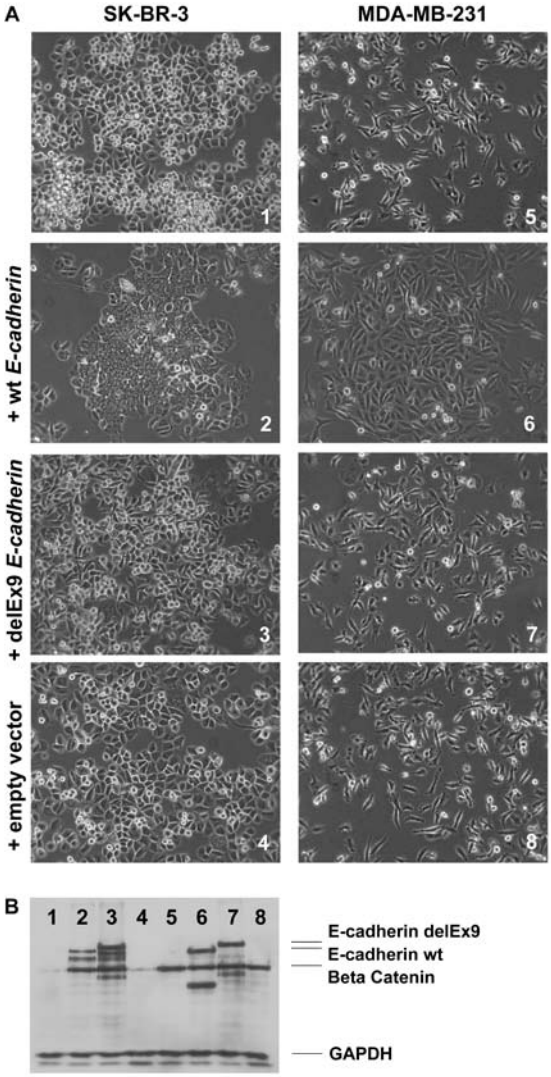


Figure 3.3 E-cadherin reconstitution experiments. Human *E-cadherin* cDNA was stably expressed in two E-cadherin null breast cancer cell lines. SK-BR-3 has lost protein expression due to a large genomic homozygous deletion of *E-cadherin* and MDA-MB-231 had lost protein expression due to *E-cadherin* promoter methylation. Expression constructs contained wild-type *E-cadherin* or mutant *E-cadherin* with an in-frame deletion of exon 9 (delEx9). (A), Conversion from rounded cell morphology to epithelial cell morphology was observed for SK-BR-3 cells transfected with the wild-type *E-cadherin* construct but not with delEx9 *E-cadherin* construct or the empty vector. MDA-MB-231 cells remained spindle-shaped irrespective of the construct used. (B), E-cadherin and β-catenin protein expression by western blotting, using GAPDH as housekeeper control. Note that expression of wild-type or delEx9 *E-cadherin* cDNA also resulted in expression of β-catenin proteins in SK-BR-3. MDA-MB-231 cells had not lost β-catenin protein expression, nor had any of the other *E-cadherin* methylated cell lines (Table 3.1). (1-4), SK-BR-3 cells untransfected, or stably transfected with wild-type *E-cadherin*, delEx9 *E-cadherin*, or empty vector; (5-8), Ibid. for MDA-MB-231 cells.

Reconstitution of the spindle-shaped *E-cadherin* methylated MDA-MB-231 cell line yielded six *E-cadherin* wild-type clones and eight *E-cadherin* delEx9 clones that expressed E-cadherin proteins (and had retained β-catenin protein expression). The *E-cadherin* wild-type clones grew less dispersed than the untransfected MDA-MB-231 cell line or the clones reconstituted with the *E-cadherin* delEx9 cDNA or empty vector. Yet, all MDA-MB-231 clones retained the spindle cell morphology, whether they were reconstituted with *E-cadherin* wild-type or delEx9 cDNA, and whether or not they expressed E-cadherin proteins (Figure 3.3). Thus, silencing of (wild-type) *E-cadherin* gene expression in association with promoter methylation is not a major determinant for the spindle cell morphology. In fact, the observation that several spindle cell lines are not methylated at CpG island 3 at the *E-cadherin* promoter and also still express E-cadherin proteins, suggests that loss of *E-cadherin* expression is secondary to the morphological differentiation status of the cells.

Expression of *SIP1* and E47 repressors, but not *SLUG, SNAIL, δEF1* and *TWIST*, is restricted to breast cancer cell lines with the spindle cell morphology

Conversion from an epithelial cell morphology to a spindle or fibroblast-like cell morphology had reportedly been associated with EMT involving transcriptional repression and downregulation of *E-cadherin* [58-64]. We therefore analyzed the breast cancer cell lines for expression of the E-cadherin transcriptional repressors *SNAIL*, *SLUG*, *SIP1*, δ*EF1* and *TWIST* by qPCR, and E47 by western blotting (Table 3.1). All six E-cadherin repressors were expressed in most of the thirteen spindle cell lines (in 11, 12 or all spindle cell lines). In contrast to *SIP1* and E47, δ*EF1*, *TWIST*, *SLUG* and *SNAIL* were also expressed in a substantial proportion of non-spindle cell lines (in 11, 21, 23 and all 28 non-spindle cell lines). However, the expression levels of δ*EF1, TWIST* and *SLUG*, but not of *SNAIL*, typically were somewhat higher in the spindle cell lines compared to the non-spindle cell lines (with average Ct values of 26 vs. 34, 25 vs. 28, 22 vs. 30, and 28 vs. 27, for δ*EF1*, *TWIST*, *SLUG* and *SNAIL*; Table 3.1). Thus, expression of *SIP1* and E47 repressors, and to a lesser extent the other E-cadherin repressors, strongly associates with the spindle cell morphology, albeit that this does not always result in significant downregulation of *E-cadherin*.

Genetic *E-cadherin* inactivation occurs in luminal-type breast cancer cell lines and methylation-associated silencing in basal-type breast cancer cell lines

Our analyses thus far strongly suggested that genetic mutation of *E-cadherin* is fundamentally distinct from transcriptional silencing of *E-cadherin* through promoter methylation, where the former is typified by the rounded cell morphology and the latter by the spindle cell morphology. To conclusively resolve whether these two mechanisms of *E-cadherin* inactivation indeed involve different biological pathways, we have determined gene expression profiles of 39 breast cancer cell lines using Affymetrix U133A microarrays. Unsupervised Pearson correlation, in which samples are positioned according to their overall similarity in gene expression profiles, revealed two main clusters of cell lines, whether the correlation was calculated from a log2GM

<-2 and >2 probe subset or log2GM <-3 and >3 probe subset (data shown for log2GM <-2 and >2 in Figure 3.4 A). The lower cluster included all 13 spindle cell lines, a single epithelial cell line and a single rounded cell line (HCC1937 and MDA-MB-468, respectively). The upper cluster included all other epithelial and rounded cell lines intermingled, but none of the spindle cell lines (Figure 3.4). Cell line DU4475 was atypical as it did not belong to either of the two clusters. The lower cluster could be further subdivided into a major subgroup of ten spindle cell lines

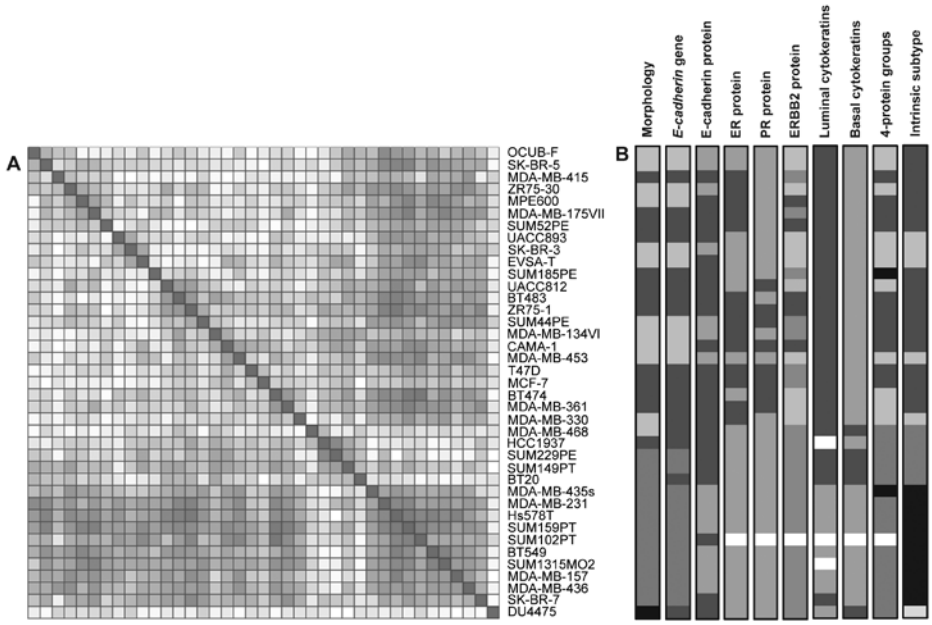


Figure 3.4 Pearson correlation and expression of breast cancer associated proteins in human breast cancer cell lines. (A), Pearson correlation plot based on the log2GM <-2 and >2 subset (5490 probe sets). The Pearson correlation coefficient algorithm positions samples according their overall similarity in gene expression, where red indicates high overall similarity (positive correlation coefficient) and blue indicates low similarity (negative correlation coefficient). (B), Various characterizations of the cell lines indicated that the upper cluster in the Pearson correlation plot contains the epithelial and rounded cell lines intermingled whereas the lower cluster contains all spindle cell lines and a single epithelial cell line and a single rounded cell line (HCC1937 and MDA-MB-468). One cell line was atypical as it did not belong to either of the two clusters (DU4475). The lower cluster included two subgroups that by the intrinsic gene set classified as basal-like and normal-like intrinsic subtypes, where all E-cadherin-negative spindle cell lines classified as normal-like. The lower cluster classified as basal breast cancers by the 4-protein signature of ERBB2, ER, CK5 and EGFR. Color coding morphology column: green, epithelial morphology; yellow, rounded cell morphology; orange, spindle cell morphology; pink, other cell morphology. *E-cadherin* gene column: green, wild-type *E-cadherin* gene; yellow, mutant *E-cadherin* gene; orange, methylated *E-cadherin* gene. E-cadherin protein, ER protein, PR protein, ERBB2 protein, luminal cytokeratins and basal cytokeratins columns: red, protein expression; blue, no protein expression; brown, protein overexpression. 4-protein groups column: green, luminal group; brown, ERBB2+ group; black, negative group; orange, basal-like group. Intrinsic subtypes column: green, luminal subtype; brown, ERBB2+ subtype; orange, basal-like subtype; black, normal-like subtype; pink, not of any subtype. This figure is also available in color in the appendix.

(8 E-cadherin negative and 2 E-cadherin positive), and a minor subgroup of three spindle cell lines and the HCC1937 and MDA-MB-468 cell lines (all E-cadherin positive). It is important to note that morphological classification of HCC1937 was somewhat ambiguous, as this cell line grows with spindle features at low cell densities but has an epithelial cell morphology at higher cell densities. The unambiguous division of the spindle cell lines from the non-spindle cell lines, without any supervision on the samples, indicates that the spindle cell lines have a gene expression program that is very different from that of the non-spindle cell lines. This implies that the spindle cell lines have a distinct differentiation status that first of all is characterized by its morphological appearance and only secondary by epigenetic silencing of *E-cadherin* expression.

Gene expression profiling of clinical breast cancers had defined five intrinsic subtypes of breast cancer [27-29]. When we classified the breast cancer cell lines based on expression of the intrinsic gene set, the minor spindle cell subgroup was identified as the basal-like intrinsic subtype and the major spindle cell subgroup was identified as normal-like (Figure 3.4 B). The epithelial/rounded cell line cluster classified as luminal, with a minor ERBB2+ subgroup that included five of eleven cell lines with ERBB2 overexpression. Cell line DU4475 again was atypical as it could not be assigned to any of the intrinsic breast cancer subtypes. Since the intrinsic gene set had been defined on clinical breast cancers, classification of the cell lines according the intrinsic molecular subtypes not only implied that these subtypes were determined by gene expression of the tumor cells in the clinical specimens (in stead of non-malignant stromal cells) but also that breast cancer cell lines are a relevant model to study human breast cancer.

Perou and colleagues recently also defined a 4-protein signature of ERBB2, ER, CK5 and EGFR to identify breast cancers of the basal-like intrinsic subtype by immunohistochemistry [30]. Analysis of the breast cancer cell lines with this 4-protein signature revealed a strong correlation between classification based on the 4-protein signature and that based on the intrinsic gene set, although the intrinsic normal-like cell lines and DU4475 were classified as basal-like, six more cell lines were classified as ERBB2+, and two cell lines classified as negative group (SUM185PE and MDA-MB-435s; Figure 3.4 B). Yet, we were able to distinguish the intrinsic normal-like from intrinsic basal-like breast cancer cell lines by their absence of basal and luminal cytokeratins and E-cadherin protein expression (Figure 3.4 B). Apart from these minor discrepancies, the three classification methods generated highly similar results, suggesting two major subtypes of breast cancer cell lines: the luminal and basal types. Most important, our data indicated that methylation-associated loss of E-cadherin protein expression resides in basal-type breast cancer cell lines whereas loss of E-cadherin protein expression due to mutation of the gene is restricted to luminal-type breast cancer cell lines, implying that these two mechanisms of E-cadherin inactivation are biologically distinct.

In contrast to lobular breast cancers, frequent loss of E-cadherin protein expression in metaplastic breast cancers is not associated with mutational inactivation of the gene

E-cadherin gene mutations reportedly were identified in breast cancers with lobular pathology [50, 51]. Consistent with our cell line data, lobular breast cancers are luminal-type breast cancers. Here, we report methylation-associated silencing of *E-cadherin* in basal-type breast cancer cell lines, suggestive for another pathological subtype of clinical breast cancers with frequent loss of E-cadherin protein expression. To search for this putative breast cancer subtype in an unbiased manner, we analyzed 166 primary breast cancers of a variety of pathological subtypes by immunohistochemistry. Tissue microarrays included ductal breast cancers (24 ER+ and 22 ER-) and lobular, mucinous, tubular and medullary breast cancers (28, 20, 16 and 22 cases, respectively). Whole tissue sections were analyzed for 34 metaplastic breast tumors because of their characteristic morphological heterogeneity. As expected, loss of E-cadherin protein expression was frequently observed among lobular breast cancers (82%, Figure 3.5 B). Interestingly, loss of E-cadherin protein expression was also observed for half of metaplastic breast cancers, whereas protein loss was found for only 20% of mucinous and 14% of medullary breast cancers and never exceeded 10% of breast cancers from other pathological subtypes. Evaluation of the clinical breast cancers for the 4-protein signature of ERBB2, ER, CK5 and EGFR [30] indicated that all 28 lobular breast cancers indeed were of the luminal group whereas the vast majority of metaplastic breast cancers were basal-like (94%, and 6% negative group; Figures 3.5 A and 3.6). The duality that we observed for E-cadherin loss among the breast cancer cell lines was thus reproduced in clinical breast cancers from the lobular and metaplastic pathological subtypes.
To confirm that mutational inactivation of *E-cadherin* associates with lobular breast cancer and epigenetic silencing with metaplastic breast cancer, we screened all 41 E-cadherin-negative

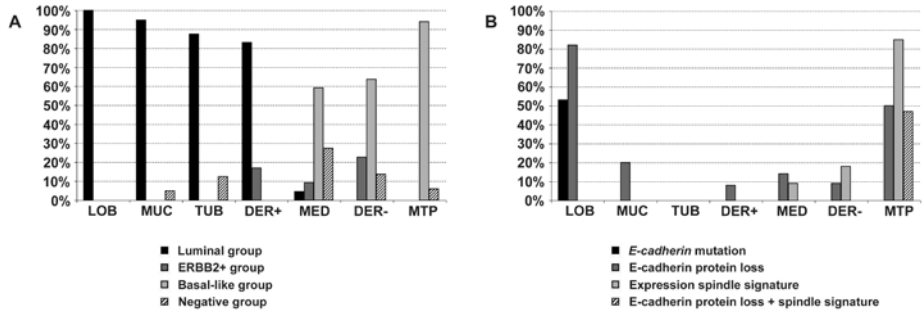


Figure 3.5 Molecular characterization of clinical breast cancers. Classification of seven pathological subtypes of clinical breast cancers by (A), the 4-protein signature [30] and (B), *E-cadherin* status and our 3-protein spindle cell signature. Pathological breast cancer subtypes: LOB, lobular; MUC, mucinous; TUB, tubular; DER+, ductal ER-positive; MED, medullary; DER-, ductal ER-negative; MTP, metaplastic. Similar to breast cancer cell lines, the spindle cell signature associated with basal type breast cancers, particularly metaplastic breast cancers. Mutational inactivation of *E-cadherin* was detected in lobular breast cancers only.

clinical breast cancers with DNA available for mutations in the *E-cadherin* gene by direct sequencing (Figure 3.5 B). We identified ten mutants among the 23 luminal breast cancers and all mutations predicted premature terminations of the encoded E-cadherin proteins (detailed in the Supplementary Data). All mutations were identified in breast cancers with lobular histology (Figure 3.5 B). Importantly, no *E-cadherin* mutations were detected among 17 basal-like and a single negative group breast cancer. We also addressed *E-cadherin* promoter methylation by MSP of CpG1 and CpG3, but methylation was observed in all samples, irrespective of their pathological subtype. Evaluation of *E-cadherin* promoter methylation in primary cancer specimens was reported to be severely hampered by the inevitable presence of leukocytes with *E-cadherin* methylation [79]. Indeed, we even detected *E-cadherin* methylation in all of eight dissected primary breast cancer samples (with <5% leukocytes), effectively precluding analysis of *E-cadherin* promoter methylation in clinical breast cancers. We therefore took advantage of the characteristic gene expression program that we had observed for the spindle cell lines, as an indirect but more comprehensive measure for the differentiation program that involves epigenetic *E-cadherin* silencing. A spindle cell gene signature of 1144 probe sets was determined by significance analysis of microarrays (SAM) [71] of the thirteen spindle cell lines versus 21 epithelial and rounded cell lines, with <1 falsely called positive probe set. Remarkably, the list of differentially expressed genes from this spindle cell gene signature included many genes known to be associated with EMT and/or the putative CD44+ breast cancer stem cell, such as *N-cadherin*, *Vimentin*, *SLUG*, *δEF1, CD44, IGFBP7, CALD1* and TGFβ pathway members ([80-82]; The gene list is provided in the Supplementary Data). This transcript-based spindle cell signature was translated into a protein-based signature to allow screening of clinical breast cancers by immunohistochemistry. A 3-protein spindle cell signature of Caldesmon-1, Caveolin-1 and Vimentin (over)expression correctly classified all 25 non-spindle breast cancer cell lines and 75% of 13 spindle cell lines, thus validating the transcript-to-protein translation (Supplementary Data). From the 166 clinical breast cancers, 35 classified with the spindle cell protein signature and these included 33 basal-like and 2 negative group breast cancers (Figures 3.5 and 3.6). Importantly, 29 of the 35 spindle cell cases were of metaplastic pathology, and the 16 spindle cell cases with loss of E-cadherin protein expression were all metaplastic breast cancers. We thus have identified metaplastic breast cancers as a second pathological subtype with frequent loss of E-cadherin protein expression that, in contrast to lobular breast cancers, is not characterized by *E-cadherin* gene mutations.

DISCUSSION

Mutational inactivation of *E-cadherin* is distinct from epigenetic silencing

By studying a model of 41 human breast cancer cell lines, we have shown that cell lines with *E-cadherin* gene mutations resemble breast cancers of the luminal type, whereas cell lines

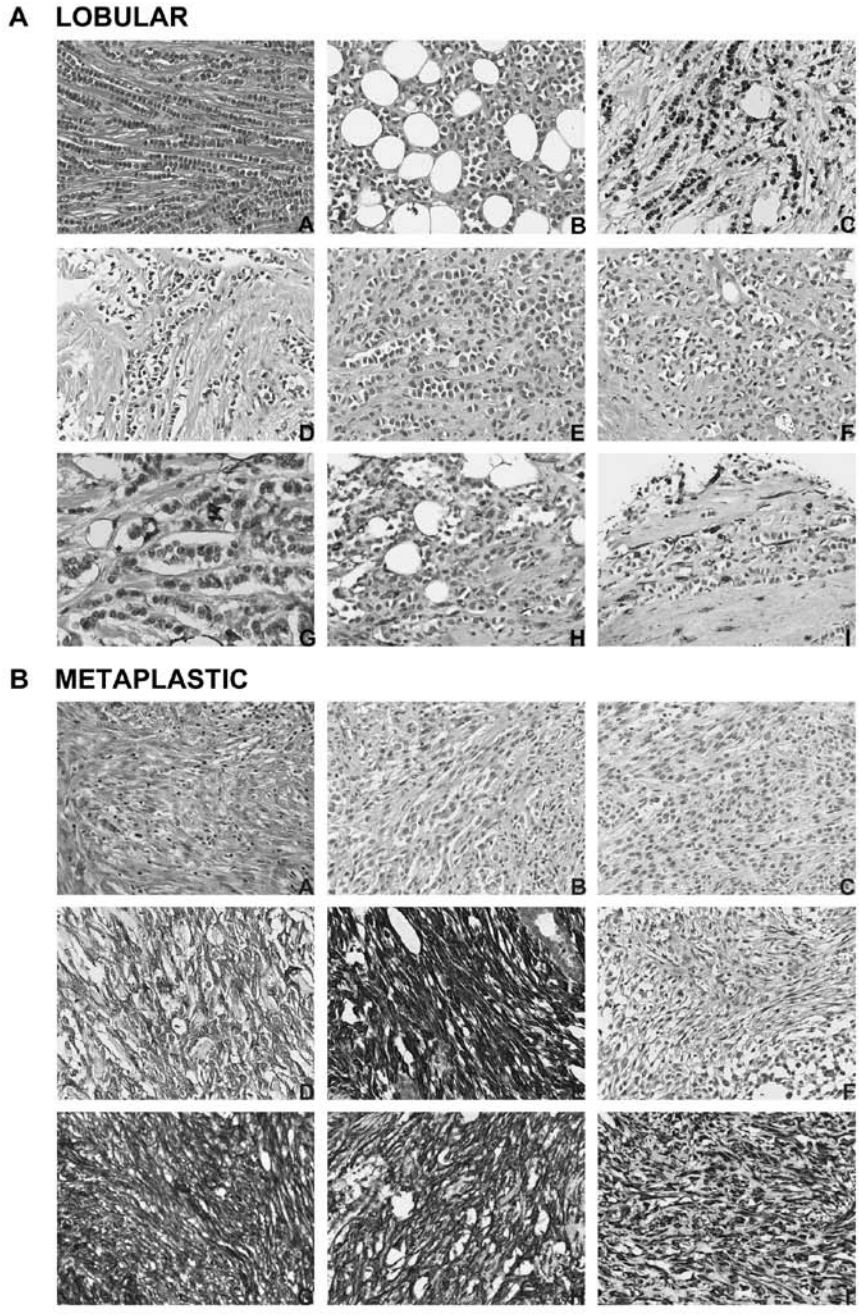


Figure 3.6 Protein expression in clinical breast cancers. Examples of immunohistochemical analysis of (A), lobular breast cancers and (B,) metaplastic breast cancers. Microscopic views: (A), HE-staining; (B), ERBB2; (C), ER; (D), EGFR; (E), CK5; (F), E-cadherin; (G), Caveolin-1; (H), Caldesmon; (I), Vimentin. This figure is also available in color in the appendix.

with epigenetic silencing of *E-cadherin* in association with promoter methylation resemble basal-type breast cancers. The differences in differentiation status between both subtypes of breast cancer cell lines were reflected by their different morphological appearance (rounded cells versus spindle cells) and by their very distinct intrinsic gene expression program. This duality in E-cadherin loss was also observed in clinical breast cancers. E-cadherin loss of protein expression was frequent among lobular breast cancers and metaplastic breast cancers (that are of the luminal and basal types, respectively), but inactivating *E-cadherin* gene mutations were only identified in lobular tumors whereas the 3-protein spindle cell signature was pronounced among metaplastic tumors. In cancer, genetic mutation and epigenetic silencing in association with promoter methylation are widely accepted as two mechanistic means to the same end: inactivation of a tumor suppressor gene. Here we have shown that, at least for *E-cadherin*, this may not always be true.

The identification of two distinct modes of *E-cadherin* inactivation may very well explain recurrent discrepancies in E-cadherin scientific literature. E-cadherin "null" cell lines may not always generate consistent results in functional studies when defined simply by loss of E-cadherin protein expression. E-cadherin's acclaimed role in tumor invasion, for example, has for long been based on the papers by Frixen *et al.* and Vleminckx *et al.* [47, 48] in which tumor invasion was being investigated by experimental manipulation of *E-cadherin* methylated cell lines and the spindle-shaped *v-ras* transformed MDCK cell line. In contrast, when *E-cadherin* was inactivated by targeted intragenic deletion in mammary epithelium, mice developed invasive breast cancers only upon concurrent *p53* inactivation [83]. Importantly, these *E-cadherin* mutant tumors were of the lobular pathological subtype and thus represent the very first murine model that faithfully replicates human *E-cadherin* mutational inactivation. In humans, it has been shown that *E-cadherin* mutations already arise in premalignant carcinoma *in situ* lesions of the breast, also implying that mutational inactivation of *E-cadherin* is an early event that does not yet bring about tumor invasion [84]. Yet, loss of E-cadherin protein expression has been associated with a worse clinical outcome of cancer patients [85]. Although apparently contradictory, all of these observations are likely to be correct. The data reported here imply that the flaw lies in the interpretation of results. It is now imperative to revisit functional E-cadherin studies as well as clinical studies on the prognosis of patients with E-cadherin-negative breast cancers, with careful dissection of the mode of E-cadherin inactivation. Our 3-protein spindle cell signature makes this both feasible and doable.

Breast cancer cell lines are a valid model to study human breast cancer

Cancer cell lines provide the unique opportunity to perform sensitive and in-depth characterizations on an unlimited source of purely tumor cells. For example, expression of E-cadherin transcriptional repressors could be analyzed in detail, revealing widespread expression in spindle cell lines, but unexpectedly, also some expression in non-spindle cell lines. It thus appears that the EMT differentiation program in spindle cell lines involves a rather complex concerted

action of E-cadherin repressors. Also, determination of *E-cadherin* promoter methylation is essentially precluded in clinical cancer samples due to the inevitable presence of methylated leukocytes in these specimens [79]. But most important, the breast cancer cell lines allowed us to comprehensively investigate their gene expression program. We identified four of the intrinsic subtypes among the cell lines [27-29]. Our results were highly similar to those obtained by Gray and colleagues albeit that they designated the basal-like and normal-like intrinsic subtypes as "basal A" and "basal B" [86]. In fact, our data suggest that their nomenclature may indeed be more appropriate. It should be kept in mind that our breast cancer cell line collection and Gray's partially-overlapping collection both have an overrepresentation of normal-like/basal B cell lines and perhaps also of cell lines with lobular characteristics. Nevertheless, classification of the cell lines according the established intrinsic subtypes of clinical breast cancers shows that breast cancer cell lines are indeed a valid model to study human breast cancer.

Epigenetic silencing of *E-cadherin* and its role in EMT is restricted to basal-type breast cancers, particularly metaplastic breast cancers

Our reconstitution experiments revealed that mutation of the *E-cadherin* gene was causative in the conversion from an epithelial to rounded cell morphology, but epigenetic silencing of *E-cadherin* expression was not causative in spindle cell morphology. This latter observation was rather unexpected since loss of E-cadherin expression is considered a hallmark of EMT [57]. In fact, the spindle cell gene signature of the breast cancer cell lines with epigenetically silenced *E-cadherin* had remarkable similarities with gene signatures that previously have been associated with both EMT and breast cancer stem cells, among others including *N-cadherin*, *vimentin, CD44 and* TGFβ pathway members [80-82]. In contrast, the gene expression program of the rounded *E-cadherin* mutant cell lines and the epithelial *E-cadherin* wild-type cell lines had no resemblance with EMT nor cancer stem cells. It thus appears that genetic inactivation of *E-cadherin* solely targets E-cadherin's adhesion function, whereas epigenetic inactivation of *E-cadherin* involves EMT with widespread gene expression changes also including downregulation of E-cadherin.

A causative role for E-cadherin in EMT was also challenged by its expression in five of the thirteen spindle cell lines from our collection, where all E-cadherin-negative cell lines were of the normal-like intrinsic subtype and all basal-like intrinsic subtype cell lines were E-cadherin-positive. Extensive characterization of our collection of breast cancer cell lines revealed that the basal-like and normal-like cell lines do share many characteristics, including their spindle cell morphology, widespread expression of E-cadherin transcriptional repressors and its associated gene expression program, the triple-negative phenotype (i.e., ER/PR/ERBB2-negative), and a gene mutation spectrum that includes *p16/p14ARF* deletion, *RB1* and *BRCA1* mutation, and mutational activation of the RAS pathway (chapter 7). Yet, loss of expression of both luminal and basal cytokeratins distinguished normal-like from basal-like intrinsic subtype breast cancer cell lines, apart from their differential E-cadherin protein expression. It may be that breast cancers

of the basal-like intrinsic subtype arise from luminal breast cancers by an EMT-like transdifferentiation, and that normal-like intrinsic subtype breast cancers then represent cancers with fully completed EMT. Alternatively, luminal breast cancers may arise from the luminal ductal epithelium in the breast whereas basal-like intrinsic subtype breast cancers arise from basal ductal epithelium. Normal-like intrinsic subtype breast cancers then again appear transdifferentiated basal-like intrinsic subtype breast cancers. Although somewhat irrelevant for this line of thought, it may also be that the phenotypic subtype of breast cancers is dictated by their gene mutation profile in stead of their cellular origin (chapter 7). Either way, breast cancers of the basal-like and normal-like intrinsic subtypes appear to constitute two ends of a spectrum of basal type breast cancers — with complete EMT being signified by loss of expression of luminal and basal cytokeratins as well as E-cadherin proteins. Similar to the cell lines, we observed loss of E-cadherin protein expression in a substantial fraction of clinical breast cancers of the basal-like group, particularly among those of the metaplastic pathological subtype (20 of 59 basal-like tumors, of which 16 were metaplastic). And again there was a strong but not exclusive association of the spindle cell signature with E-cadherin-negative basal-like group breast cancers (Figure 3.5). It appears that the spindle cell signature in clinical breast cancers also heralds an EMT-like transdifferentiation even before these cancers loose E-cadherin protein expression. The question arises whether all basal-type breast cancers are susceptible to EMT or whether this is restricted to basal-type breast cancers with the spindle cell signature. Irrespective, our data strongly suggest that E-cadherin's role in EMT is restricted to basal type breast cancers.

ACKNOWLEDGEMENTS

We appreciate the technical assistance of Dr. Anieta Sieuwerts, Thierry van de Wetering, and members of the pathology immunohistochemistry labs. We thank Prof. Wolter Oosterhuis and Dr. John Martens for insightful discussions. This work was supported by grants from KWF Dutch Cancer Society (DDHK 2002-2687), Erasmus MC Mrace 2005 and Netherlands Genomics Initiative (NGI)/Netherlands Organization for Scientific Research (NWO).

REFERENCES

1. Moll R, Franke WW, Schiller DL, Geiger B, Krepler R. The catalog of human cytokeratins: patterns of expression in normal epithelia, tumors and cultured cells. Cell 1982;31(1):11-24.
2. Debus E, Weber K, Osborn M. Monoclonal cytokeratin antibodies that distinguish simple from stratified squamous epithelia: characterization on human tissues. Embo J 1982;1(12):1641-7.
3. Nagle RB, Bocker W, Davis JR, et al. Characterization of breast carcinomas by two monoclonal antibodies distinguishing myoepithelial from luminal epithelial cells. J Histochem Cytochem 1986;34(7):869-81.
4. Smith GH, Mehrel T, Roop DR. Differential keratin gene expression in developing, differentiating, preneoplastic, and neoplastic mouse mammary epithelium. Cell Growth Differ 1990;1(4):161-70.
5. Su L, Morgan PR, Lane EB. Expression of cytokeratin messenger RNA versus protein in the normal mammary gland and in breast cancer. Hum Pathol 1996;27(8):800-6.
6. Page MJ, Amess B, Townsend RR, et al. Proteomic definition of normal human luminal and myoepithelial breast cells purified from reduction mammoplasties. Proc Natl Acad Sci U S A 1999;96(22):12589-94.
7. Petersen OW, van Deurs B. Characterization of epithelial membrane antigen expression in human mammary epithelium by ultrastructural immunoperoxidase cytochemistry. J Histochem Cytochem 1986;34(6):801-9.
8. Petersen OW, Hoyer PE, van Deurs B. Frequency and distribution of estrogen receptor-positive cells in normal, nonlactating human breast tissue. Cancer Res 1987;47(21):5748-51.
9. Gusterson BA, Monaghan P, Mahendran R, Ellis J, O'Hare MJ. Identification of myoepithelial cells in human and rat breasts by anti-common acute lymphoblastic leukemia antigen antibody A12. J Natl Cancer Inst 1986;77(2):343-9.
10. Skalli O, Ropraz P, Trzeciak A, Benzonana G, Gillessen D, Gabbiani G. A monoclonal antibody against alpha-smooth muscle actin: a new probe for smooth muscle differentiation. J Cell Biol 1986;103(6 Pt 2):2787-96.
11. Zhang RR, Man YG, Vang R, et al. A subset of morphologically distinct mammary myoepithelial cells lacks corresponding immunophenotypic markers. Breast Cancer Res 2003;5(5):R151-6.
12. Barbareschi M, Pecciarini L, Cangi MG, et al. p63, a p53 homologue, is a selective nuclear marker of myoepithelial cells of the human breast. Am J Surg Pathol 2001;25(8):1054-60.
13. Coleman S, Silberstein GB, Daniel CW. Ductal morphogenesis in the mouse mammary gland: evidence supporting a role for epidermal growth factor. Dev Biol 1988;127(2):304-15.
14. Moller P, Mechtersheimer G, Kaufmann M, et al. Expression of epidermal growth factor receptor in benign and malignant primary tumours of the breast. Virchows Arch A Pathol Anat Histopathol 1989;414(2):157-64.
15. Santini D, Ceccarelli C, Tardio ML, Taffurelli M, Marrano D. Immunocytochemical expression of epidermal growth factor receptor in myoepithelial cells of the breast. Appl Immunohistochem Mol Morphol 2002;10(1):29-33.
16. O'Hare MJ, Ormerod MG, Monaghan P, Lane EB, Gusterson BA. Characterization in vitro of luminal and myoepithelial cells isolated from the human mammary gland by cell sorting. Differentiation 1991;46(3):209-21.
17. Pechoux C, Gudjonsson T, Ronnov-Jessen L, Bissell MJ, Petersen OW. Human mammary luminal epithelial cells contain progenitors to myoepithelial cells. Dev Biol 1999;206(1):88-99.
18. Bocker W, Moll R, Poremba C, et al. Common adult stem cells in the human breast give rise to glandular and myoepithelial cell lineages: a new cell biological concept. Lab Invest 2002;82(6):737-46.
19. Petersen OW, Gudjonsson T, Villadsen R, Bissell MJ, Ronnov-Jessen L. Epithelial progenitor cell lines as models of normal breast morphogenesis and neoplasia. Cell Prolif 2003;36 Suppl 1:33-44.

20. Gudjonsson T, Villadsen R, Nielsen HL, Ronnov-Jessen L, Bissell MJ, Petersen OW. Isolation, immortalization, and characterization of a human breast epithelial cell line with stem cell properties. Genes Dev 2002;16(6):693-706.
21. Boecker W, Buerger H. Evidence of progenitor cells of glandular and myoepithelial cell lineages in the human adult female breast epithelium: a new progenitor (adult stem) cell concept. Cell Prolif 2003;36 Suppl 1:73-84.
22. Tavassoli FA, Devilee P, International Agency for Research on Cancer., World Health Organization. Pathology and genetics of tumours of the breast and female genital organs. Lyon
Oxford: International Agency for Research on Cancer ;
Oxford University Press (distributor); 2003.
23. Rosen PP. Rosen's breast pathology. 2nd ed. Philadelphia: Lippincott Williams & Wilkins; 2001.
24. Wetzels RH, Kuijpers HJ, Lane EB, et al. Basal cell-specific and hyperproliferation-related keratins in human breast cancer. Am J Pathol 1991;138(3):751-63.
25. Malzahn K, Mitze M, Thoenes M, Moll R. Biological and prognostic significance of stratified epithelial cytokeratins in infiltrating ductal breast carcinomas. Virchows Arch 1998;433(2):119-29.
26. Abd El-Rehim DM, Pinder SE, Paish CE, et al. Expression of luminal and basal cytokeratins in human breast carcinoma. J Pathol 2004;203(2):661-71.
27. Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. Nature 2000;406(6797):747-52.
28. Sorlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A 2001;98(19):10869-74.
29. Sorlie T, Tibshirani R, Parker J, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proc Natl Acad Sci U S A 2003;100(14):8418-23.
30. Nielsen TO, Hsu FD, Jensen K, et al. Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. Clin Cancer Res 2004;10(16):5367-74.
31. Takeichi M. Cadherin cell adhesion receptors as a morphogenetic regulator. Science 1991;251(5000):1451-5.
32. Kemler R. From cadherins to catenins: cytoplasmic protein interactions and regulation of cell adhesion. Trends Genet 1993;9(9):317-21.
33. Cowin P. Unraveling the cytoplasmic interactions of the cadherin superfamily. Proc Natl Acad Sci U S A 1994;91(23):10759-61.
34. Gumbiner BM. Cell adhesion: the molecular basis of tissue architecture and morphogenesis. Cell 1996;84(3):345-57.
35. Semb H, Christofori G. The tumor-suppressor function of E-cadherin. American journal of human genetics 1998;63(6):1588-93.
36. Berx G, Van Roy F. The E-cadherin/catenin complex: an important gatekeeper in breast cancer tumorigenesis and malignant progression. Breast Cancer Res 2001;3(5):289-93.
37. Jamora C, Fuchs E. Intercellular adhesion, signalling and the cytoskeleton. Nat Cell Biol 2002;4(4):E101-8.
38. Yamada S, Pokutta S, Drees F, Weis WI, Nelson WJ. Deconstructing the cadherin-catenin-actin complex. Cell 2005;123(5):889-901.
39. Drees F, Pokutta S, Yamada S, Nelson WJ, Weis WI. Alpha-catenin is a molecular switch that binds E-cadherin-beta-catenin and regulates actin-filament assembly. Cell 2005;123(5):903-15.
40. Davis MA, Ireton RC, Reynolds AB. A core function for p120-catenin in cadherin turnover. J Cell Biol 2003;163(3):525-34.
41. Ireton RC, Davis MA, van Hengel J, et al. A novel role for p120 catenin in E-cadherin function. J Cell Biol 2002;159(3):465-76.
42. Yap AS, Niessen CM, Gumbiner BM. The juxtamembrane region of the cadherin cytoplasmic tail supports lateral clustering, adhesive strengthening, and interaction with p120ctn. J Cell Biol 1998;141(3):779-89.

43. Rubinfeld B, Souza B, Albert I, et al. Association of the APC gene product with beta-catenin. Science 1993;262(5140):1731-4.
44. Su LK, Vogelstein B, Kinzler KW. Association of the APC tumor suppressor protein with catenins. Science 1993;262(5140):1734-7.
45. Perez-Moreno M, Fuchs E. Catenins: keeping cells from getting their signals crossed. Developmental cell 2006;11(5):601-12.
46. Scott JA, Yap AS. Cinderella no longer: alpha-catenin steps out of cadherin's shadow. Journal of cell science 2006;119(Pt 22):4599-605.
47. Frixen UH, Behrens J, Sachs M, et al. E-cadherin-mediated cell-cell adhesion prevents invasiveness of human carcinoma cells. J Cell Biol 1991;113(1):173-85.
48. Vleminckx K, Vakaet L, Jr., Mareel M, Fiers W, van Roy F. Genetic manipulation of E-cadherin expression by epithelial tumor cells reveals an invasion suppressor role. Cell 1991;66(1):107-19.
49. Perl AK, Wilgenbus P, Dahl U, Semb H, Christofori G. A causal role for E-cadherin in the transition from adenoma to carcinoma. Nature 1998;392(6672):190-3.
50. Berx G, Cleton-Jansen AM, Nollet F, et al. E-cadherin is a tumour/invasion suppressor gene mutated in human lobular breast cancers. Embo J 1995;14(24):6107-15.
51. Berx G, Cleton-Jansen AM, Strumane K, et al. E-cadherin is inactivated in a majority of invasive human lobular breast cancers by truncation mutations throughout its extracellular domain. Oncogene 1996;13(9):1919-25.
52. Becker KF, Atkinson MJ, Reich U, et al. Exon skipping in the E-cadherin gene transcript in metastatic human gastric carcinomas. Hum Mol Genet 1993;2(6):803-4.
53. Becker KF, Atkinson MJ, Reich U, et al. E-cadherin gene mutations provide clues to diffuse type gastric carcinomas. Cancer Res 1994;54(14):3845-52.
54. Yoshiura K, Kanai Y, Ochiai A, Shimoyama Y, Sugimura T, Hirohashi S. Silencing of the E-cadherin invasion-suppressor gene by CpG methylation in human carcinomas. Proc Natl Acad Sci U S A 1995;92(16):7416-9.
55. Graff JR, Herman JG, Lapidus RG, et al. E-cadherin expression is silenced by DNA hypermethylation in human breast and prostate carcinomas. Cancer Res 1995;55(22):5195-9.
56. Graff JR, Gabrielson E, Fujii H, Baylin SB, Herman JG. Methylation patterns of the E-cadherin 5′ CpG island are unstable and reflect the dynamic, heterogeneous loss of E-cadherin expression during metastatic progression. J Biol Chem 2000;275(4):2727-32.
57. Peinado H, Olmeda D, Cano A. Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? Nat Rev Cancer 2007;7(6):415-28.
58. Batlle E, Sancho E, Franci C, et al. The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells. Nat Cell Biol 2000;2(2):84-9.
59. Cano A, Perez-Moreno MA, Rodrigo I, et al. The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. Nat Cell Biol 2000;2(2):76-83.
60. Hajra KM, Chen DY, Fearon ER. The SLUG zinc-finger protein represses E-cadherin in breast cancer. Cancer Res 2002;62(6):1613-8.
61. Comijn J, Berx G, Vermassen P, et al. The two-handed E box binding zinc finger protein SIP1 downregulates E-cadherin and induces invasion. Mol Cell 2001;7(6):1267-78.
62. Eger A, Aigner K, Sonderegger S, et al. DeltaEF1 is a transcriptional repressor of E-cadherin and regulates epithelial plasticity in breast cancer cells. Oncogene 2005;24(14):2375-85.
63. Perez-Moreno MA, Locascio A, Rodrigo I, et al. A new role for E12/E47 in the repression of E-cadherin expression and epithelial-mesenchymal transitions. J Biol Chem 2001;276(29):27424-31.
64. Yang J, Mani SA, Donaher JL, et al. Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. Cell 2004;117(7):927-39.
65. Harkes IC, Elstrodt F, Dinjens WN, et al. Allelotype of 28 human breast cancer cell lines and xenografts. Br J Cancer 2003;89(12):2289-92.

66. van de Wetering M, Barker N, Harkes IC, et al. Mutant E-cadherin breast cancer cells do not display constitutive Wnt signaling. Cancer Res 2001;61(1):278-84.
67. Graff JR, Herman JG, Myohanen S, Baylin SB, Vertino PM. Mapping patterns of CpG island methylation in normal and neoplastic cells implicates both upstream and downstream regions in de novo methylation. J Biol Chem 1997;272(35):22322-9.
68. Wasielewski M, Elstrodt F, Klijn JG, Berns EM, Schutte M. Thirteen new p53 gene mutants identified among 41 human breast cancer cell lines. Breast Cancer Res Treat 2006;99(1):97-101.
69. Smid M, Wang Y, Zhang Y, et al. Subtypes of breast cancer show preferential site of relapse. Cancer Research 2008;68(9):3108-14.
70. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 1998;95(25):14863-8.
71. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A 2001;98(9):5116-21.
72. Engel LW, Young NA. Human breast carcinoma cells in continuous culture: a review. Cancer Res 1978;38(11 Pt 2):4327-39.
73. Cailleau R, Young R, Olive M, Reeves WJ, Jr. Breast tumor cell lines from pleural effusions. J Natl Cancer Inst 1974;53(3):661-74.
74. Cailleau R, Olive M, Cruciger QV. Long-term human breast carcinoma cell lines of metastatic origin: preliminary characterization. In Vitro 1978;14(11):911-5.
75. Schlosshauer PW, Brown SA, Eisinger K, et al. APC truncation and increased beta-catenin levels in a human breast cancer cell line. Carcinogenesis 2000;21(7):1453-6.
76. Lombaerts M, van Wezel T, Philippo K, et al. E-cadherin transcriptional downregulation by promoter methylation but not mutation is related to epithelial-to-mesenchymal transition in breast cancer cell lines. Br J Cancer 2006;94(5):661-71.
77. Ozawa M, Engel J, Kemler R. Single amino acid substitutions in one Ca2+ binding site of uvomorulin abolish the adhesive function. Cell 1990;63(5):1033-8.
78. Hiraguri S, Godfrey T, Nakamura H, et al. Mechanisms of inactivation of E-cadherin in breast cancer cell lines. Cancer Res 1998;58(9):1972-7.
79. Lombaerts M, Middeldorp JW, van der Weide E, et al. Infiltrating leukocytes confound the detection of E-cadherin promoter methylation in tumors. Biochem Biophys Res Commun 2004;319(2):697-704.
80. Onder TT, Gupta PB, Mani SA, Yang J, Lander ES, Weinberg RA. Loss of E-cadherin promotes metastasis via multiple downstream transcriptional pathways. Cancer Res 2008;68(10):3645-54.
81. Mani SA, Guo W, Liao MJ, et al. The epithelial-mesenchymal transition generates cells with properties of stem cells. Cell 2008;133(4):704-15.
82. Shipitsin M, Campbell LL, Argani P, et al. Molecular definition of breast tumor heterogeneity. Cancer Cell 2007;11(3):259-73.
83. Derksen PW, Liu X, Saridin F, et al. Somatic inactivation of E-cadherin and p53 in mice leads to metastatic lobular mammary carcinoma through induction of anoikis resistance and angiogenesis. Cancer Cell 2006;10(5):437-49.
84. Vos CB, Cleton-Jansen AM, Berx G, et al. E-cadherin inactivation in lobular carcinoma in situ of the breast: an early event in tumorigenesis. Br J Cancer 1997;76(9):1131-3.
85. Siitonen SM, Kononen JT, Helin HJ, Rantala IS, Holli KA, Isola JJ. Reduced E-cadherin expression is associated with invasiveness and unfavorable prognosis in breast cancer. American journal of clinical pathology 1996;105(4):394-402.
86. Neve RM, Chin K, Fridlyand J, et al. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. Cancer Cell 2006;10(6):515-27.

Chapter 4

α-Catenin is a Putative New Tumor Suppressor Gene

Antoinette Hollestelle
Fons Elstrodt
Mieke Timmermans
Anieta M. Sieuwerts
Justine K. Peeters
Jan G.M. Klijn
Peter J. van der Spek
John A. Foekens
Michael A. den Bakker
Mieke Schutte

In preparation

ABSTRACT

Mutations of *E-cadherin* have been identified in half of lobular breast cancers and diffuse-type gastric cancers, two tumor subtypes with remarkably similar pathological appearances. A causal role for *E-cadherin* gene mutations in the lobular breast cancer subtype was recently demonstrated in *E-cadherin* knock-out mice. Similarly, we have shown that *E-cadherin* mutation also was causal for the characteristic rounded cell morphology of *E-cadherin* mutant human breast cancer cell lines. Here, we have investigated whether another gene in the E-cadherin tumor suppressor pathway might be mutated in wild-type *E-cadherin* breast cancer cell lines with the rounded cell morphology. Using gene expression and gene mutation analyses, we have identified four bi-allelic inactivating α-*catenin* mutations among 55 human breast cancer cell lines. All four α-*catenin* mutations predicted premature termination of the encoded proteins, and concordantly, none of the four mutant cell lines expressed α-catenin proteins. Importantly, three of the α-*catenin* mutant cell lines had the rounded cell morphology and all fourteen cell lines with the rounded cell morphology had mutations of either *E-cadherin* or α-*catenin*. As anticipated, loss of α-catenin protein expression was associated with the lobular pathological subtype in primary breast cancers. Together, our observations suggest that α-*catenin* is a new tumor suppressor gene that operates in the E-cadherin tumor suppressor pathway.

INTRODUCTION

The E-cadherin/catenin protein complex consists of the cytoplasmic proteins α-, β-, γ- and p120-catenin and the transmembrane protein E-cadherin. The armadillo proteins β- and γ-catenin bind directly to the intracellular carboxy-terminal tail of the E-cadherin protein and p120-catenin (p120ctn) binds more membrane-proximal to E-cadherin. The vinculin-related protein α-catenin acts as a molecular switch and binds either to β- or γ-catenin as a monomer or to the actin cytoskeleton as a homodimer. Extracellular, E-cadherin forms homodimers with E-cadherin proteins on adjacent epithelial cells in a calcium-dependent fashion. The E-cadherin/catenin protein complex thus maintains the integrity of epithelial tissues through cell-cell adhesion [1-8].

Inactivation of the *E-cadherin* tumor suppressor gene in human cancer may involve mutational inactivation or transcriptional silencing by transcriptional repressors and/or promoter hypermethylation [9-11]. We recently have shown that transcriptional silencing of *E-cadherin* is biologically distinct from mutational inactivation of *E-cadherin*, where transcriptional silencing predominantly associated with metaplastic breast cancers and mutational inactivation with lobular breast cancers (chapter 3). Mutations of the *E-cadherin* gene had indeed only been found in lobular breast cancers and in diffuse-type gastric cancers, two tumor subtypes that have remarkable similarities in their histopathological appearance [12-15]. Also, germline mutations of *E-cadherin* predispose an individual to hereditary diffuse-type gastric carcinoma (HDGC) [16, 17]. Although *E-cadherin* germline mutations have as yet not been associated with familial lobular breast cancer, cases of lobular breast cancer have been observed in HDGC families [18-20]. Interestingly, we previously have found mutations of *E-cadherin* solely in breast cancer cell lines with a rounded cell morphology — a classifying characteristic of lobular breast cancer and diffuse-type gastric cancer (chapter 3). The strong association of *E-cadherin* gene mutations with the typical morphology of rounded cells had suggested causality for *E-cadherin* mutations. Indeed, mutational inactivation of *E-cadherin* has been shown to be causal for a lobular cancer phenotype in conditional knock-out mice [21] and we have shown causality for the rounded cell morphology by genetic manipulation of human breast cancer cell lines (chapter 3).

Inactivating mutations of *E-cadherin* are found in only half of lobular breast cancers and half of diffuse-type gastric cancers ([12-15] and chapter 3). Similarly, we did not detect mutations in *E-cadherin* in three of fourteen breast cancer cell lines with the rounded cell morphology. Given the causal role of *E-cadherin* gene mutations in the rounded cell morphology, we hypothesized that another gene in the E-cadherin tumor suppressor pathway might be mutated in these three rounded *E-cadherin* wild-type cell lines. Here, we report the identification of bi-allelic inactivating α-*catenin* mutations in human breast cancer cell lines, classifying α-*catenin* as a putative new tumor suppressor gene in human breast cancer.

MATERIALS AND METHODS

Breast cancer cell lines

The 55 human breast cancer cell lines used in this study are listed in Table 4.1. Cell lines EVSA-T, MPE600, and SK-BR-5/7 were kind gifts of Dr. N. de Vleesschouwer (Institut Jules Bordet, Brussels, Belgium), Dr. H.S. Smith (California Pacific Medical Center, San Francisco, CA) and Dr. E. Stockert (Sloan-Kettering Institute for Cancer Research, New York, NY), respectively. The SUM cell lines were generated in the Ethier laboratory (available at http://www.asterand.com). Cell lines OCUB-F and -M had been generated from the same tumor, and were obtained from Riken Gene Bank (Tsukuba, Japan). The other cell lines were obtained from ATCC (Manassas, VA). The cell lines are unique and monoclonal as shown by extensive analysis of nearly 150 polymorphic microsatellite markers [22].

Clinical breast cancers

Formalin-fixed paraffin-embedded primary breast cancers included 33 lobular, 48 ductal, 22 medullary, 22 mucinous and 14 tubular breast cancers. All 139 breast cancers were obtained from the archive of the Pathology department at Erasmus MC, randomly selected from diagnosis years 1990 through 2005. Tissue microarrays were constructed by punching three 0.6-mm cores from representative areas of each tumor and transferring them into a recipient paraffin block using an ATA-27 automated tissue microarrayer (Beecher Instruments, Sun Prairie, WI).
One hundred and seven fresh-frozen primary lobular breast cancers were obtained from the liquid nitrogen tumor bank of the Medical Oncology department at Erasmus MC, randomly selected from diagnosis years 1978 through 1995.
The Medical Ethical Committee at Erasmus MC has approved the study (MEC 02-953), which was carried out according the Code of Conduct of the Federation of Medical Scientific Societies in the Netherlands (http://www.fmwv.nl).

Gene expression profiling

Breast cancer cell lines were grown to optimal cell densities in RPMI 1640 supplemented with 10% fetal bovine serum. The culture medium (at 37°C) was changed 16-20 hours before harvesting. RNA was isolated using the RNeasy kit (Qiagen, Hilden, Germany) upon lysing the cells in the culture flask. DNAseI-treated RNA was antisense biotinylated using the MEGAScript T7 labeling kit (Ambion, Foster City, CA) and Affymetrix U133A microarrays were hybridized according the Affymetrix GeneChip Manual, both performed on commercial basis by ServiceXS (Leiden, the Netherlands). Intensity values for all probe sets were scaled using the global normalization factor as specified by GCOS 1.1 (Affymetrix, Santa Clara, CA) and intensity values <30 were set to 30. For each probe set, the geometric mean (GM) of the intensity values was calculated among the cell lines. The expression level of each probe set in the cell lines was determined relative to its GM and log2-transformed to ascribe equal weight to gene expression levels with similar

relative distances to the GM. Deviation from this log2-transformed GM reflected differential gene expression.

Significance Analysis of Microarrays (SAM; [23]) within the Omniviz software package 3.6 (Biowisdom, Maynard, MA) was used to determine differentially expressed genes between mutant and wild-type *E-cadherin* cell lines.

Gene mutation analysis

E-cadherin and α-*catenin* (ENSG00000039068 and ENSG00000044115) coding sequences were analyzed by direct sequencing of PCR-amplified fragments from tumor-derived genomic DNA, using an ABI 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA) and intronic primers (Supplementary Data). All identified mutations were confirmed by sequencing of an independently amplified DNA template and all deletions were confirmed by performing a duplex PCR including primers for the *ZEB2* gene (ENSG00000169554) that served as positive control. Regrettably, original (uncultured) tumor material was not available for any of the four α-*catenin* mutant breast cancer cell lines, precluding confirmation that the mutations had been present in the original tumors.

Expression analysis

Reverse transcriptase (RT)-PCR for *E-cadherin,* α-*catenin,* β-*catenin,* γ-*catenin* and *p120ctn* was done using the Qiagen one-step RT-PCR kit and gene-specific exonic primers (Supplementary Data).

Immunohistochemistry was performed using the EnVision+ kit (DAKO, Glostrup, Denmark). To unmask antigens, slides were boiled for 20 minutes in Citrate buffer pH=6 (Klinipath, Duiven, the Netherlands) for α-catenin or Tris-EDTA buffer pH=9 (Klinipath) for the other antigens. Primary monoclonal antibodies and their dilutions were: E-cadherin 1:25 (DAKO, clone NCH-38), α-catenin 1:40 (Santa Cruz Biotechnology (Santa Cruz, CA), clone G-11), β-catenin 1:100 (BD Transduction Laboratories (Franklin Lakes, NJ), clone 14), γ-catenin 1:120 (BD Transduction Laboratories, clone 15) and p120ctn 1:150 (BD Transduction Laboratories, clone 98).

RESULTS

Catenin protein expression associates with *E-cadherin* gene status in breast cancer cell lines

By studying 41 human breast cancer cell lines, we recently have shown that *E-cadherin* gene mutation was causally associated with a rounded cell morphology (chapter 3). We now have extended our *E-cadherin* mutation analysis to 55 breast cancer cell lines, and identified one additional *E-cadherin* mutant breast cancer cell line that had a rounded cell morphology (Table 4.1). Cell line HCC2218 had a large homozygous deletion of the first ten exons of the *E-cadherin* gene and did not express E-cadherin proteins (c.1_1565del1565). Additionally, we identified

Table 4.1 *E-cadherin* and α-*catenin* gene status and E-cadherin/catenin expression patterns in 55 breast cancer cell lines.

Breast cancer cell line	Cell morphology	*E-cadherin* gene status	*E-cadherin* transcript	E-cadherin protein	α-*catenin* gene status
HCC202	epithelial	wild-type*	nd	nd	wild-type
HCC1008	epithelial	wild-type*	nd	nd	wild-type
HCC1419	epithelial	wild-type*	nd	nd	wild-type
HCC1500	epithelial	wild-type*	nd	nd	wild-type
HCC1569	epithelial	wild-type*	nd	nd	wild-type
HCC1599	epithelial	wild-type*	nd	nd	wild-type
HCC1806	epithelial	wild-type*	nd	nd	wild-type
HCC1954	epithelial	wild-type*	nd	nd	wild-type
BT474	epithelial	wild-type	+	+	wild-type
BT483	epithelial	wild-type	+	+	wild-type
HCC1937	epithelial	wild-type	+	+	wild-type
MCF-7	epithelial	wild-type	+	+	wild-type
MDA-MB-175 VII	epithelial	wild-type	+	+	wild-type
MDA-MB-361	epithelial	wild-type	+	+	wild-type
MDA-MB-415	epithelial	wild-type	+	+	wild-type
SUM52PE	epithelial	wild-type	+	+	wild-type
SUM185PE	epithelial	wild-type	+	+	wild-type
SUM190PT	epithelial	wild-type	+	+	wild-type
SUM225CWN	epithelial	wild-type	+	+	wild-type
T47D	epithelial	wild-type	+	+	wild-type
UACC812	epithelial	wild-type	+	+	wild-type
UACC893	epithelial	wild-type	+	+	wild-type
ZR75-1	epithelial	wild-type	+	+	wild-type
CAMA-1	rounded	IF mutant	+	+	wild-type
EVSA-T	rounded	IF mutant	+	+	wild-type
MPE600	rounded	IF mutant	+	+	wild-type
MDA-MB-134 VI	rounded	PT mutant	+	-	wild-type
SUM44PE	rounded	PT mutant	+	-	nd
OCUB-F,-M	rounded	PT mutant	+	+/-	wild-type
MDA-MB-453	rounded	PT mutant	+/-	-	wild-type
SK-BR-5	rounded	PT mutant	+/-	-	wild-type
ZR-75-30	rounded	PT mutant	+/-	-	wild-type
SK-BR-3	rounded	PT mutant	-	-	wild-type
HCC2218	rounded	PT mutant*	nd	-	wild-type
HCC1187	rounded	wild-type*	nd	nd	PT mutant
MDA-MB-330	rounded	wild-type	+	+	PT mutant
MDA-MB-468	rounded	wild-type	+	+	PT mutant
MDA-MB-157	spindle	methylated	+/-	-	PT mutant
BT549	spindle	methylated	-	-	wild-type
Hs578T	spindle	methylated	-	-	wild-type
MDA-MB-435s	spindle	methylated	-	-	wild-type
MDA-MB-436	spindle	methylated	-	-	wild-type
SUM159PT	spindle	methylated	-	-	wild-type
SUM1315MO2	spindle	methylated	-	-	wild-type
MDA-MB-231	spindle	methylated	+/-	-	wild-type
SK-BR-7	spindle	methylated	+	+	wild-type
SUM102PT	spindle	methylated	+	nd	wild-type
SUM149PT	spindle	methylated	+	+	wild-type
SUM229PE	spindle	methylated	+	+	wild-type
BT20	spindle	wild-type	+	+	wild-type
HCC38	spindle	wild-type*	nd	nd	wild-type
HCC70	spindle	wild-type*	nd	nd	wild-type
HCC1143	spindle	wild-type*	nd	nd	wild-type
HCC1395	spindle	wild-type*	nd	nd	wild-type
DU4475	other	wild-type	+	+/-	wild-type

Breast cancer cell lines are organized by their morphology and by their *E-cadherin* status. *E-cadherin* and α-*catenin* gene status was determined by mutation analysis of all coding exons and for *E-cadherin* also by methylation specific PCR and azacytidine assays. E-cadherin and catenin transcript and protein expression was determined by RT-PCR and immunohistochemistry. IF, in-frame; PT, protein truncation; *, methylation status has not been evaluated; **, two shortened α-*catenin* transcripts; ***, transcript is only detectable up to exon 7; -, not detectable; +/-, low or barely detectable expression; +, clearly detectable expression; nd, not determined.

α-*catenin* transcript	α-catenin protein	β-*catenin* transcript	β-catenin protein	γ-*catenin* transcript	γ-catenin protein	*p120ctn* transcript	p120ctn protein	Protein expression pattern
nd	nd	nd	nd	nd	nd	nd	nd	nd
nd	nd	nd	nd	nd	nd	nd	nd	nd
nd	nd	nd	nd	nd	nd	nd	nd	nd
nd	nd	nd	nd	nd	nd	nd	nd	nd
nd	nd	nd	nd	nd	nd	nd	nd	nd
nd	nd	nd	nd	nd	nd	nd	nd	nd
nd	nd	nd	nd	nd	nd	nd	nd	nd
nd	nd	nd	nd	nd	nd	nd	nd	nd
+	+	+	+	+	+	+	+	A
+	+	+	+	+	+	+	+	A
+	+	+	+	+	+	+	+	A
+	+	+	+	+	+	+	+	A
+	+	+	+	+	+	+	+	A
+	+	+	+	+	+	+	+	A
+	+	+	+	+	+	+	+	A
+	+	+	+	+	+	+	+	A
+	+	+	+	+	+	+	+	A
+	+	+	+	+	+	+	+	A
+	+	+	+	+	+	+	+	A
+	+	+	+	+	+	+	+	A
+	+	+	+	+	+	+	+	A
+	+	+	+	+	+	+	+	A
+	+	+	+	+	+	+	+	A
+	+	+	+	+	+	+	+	A
+	+	+	+	+	+	+	+	A
+	+	+	+	+	+	+	+	A
+	+/-	+	-	+	+/-	+	+	D
+	+/-	+	-	+	+/-	+	+	D
+	+	+	+/-	+	+/-	+	+	D
+	+/-	+	-	+	+/-	+	+	D
+	+	+	-	+	+/-	+	+	D
+	+	+	+-	+	+/-	+	+	D
+	+	+	-	+	+/-	+	+	D
nd	nd	nd	nd	nd	nd	nd	nd	nd
nd	-	nd	nd	nd	nd	nd	nd	nd
+	-	+	+	+	+	+	+	E
+**	-	+	+	+	+	+	+	E
+***	-	+	+	+	+/-	+	+	C
+	+	+	+	+/-	+/-	+	+	C
+	+	+	+	+	+/-	+	+	C
+	+	+	+	+	-	+	+	C
+	+	+	+	+	+/-	+	+	C
+	+	+	+	+	+	+	+	C
+	+/-	+	+	+/-	-	+	+	C
+	+	+	+	+	-	+	+	C
+	+	+	+	+	+	+	+	A
+	nd	+	nd	+	nd	+	nd	A
+	+	+	+	+	+	+	+	A
+	+	+	+	+	+	+	+	A
+	+	+	+	+	+	+	+	A
nd	nd	nd	nd	nd	nd	nd	nd	nd
nd	nd	nd	nd	nd	nd	nd	nd	nd
nd	nd	nd	nd	nd	nd	nd	nd	nd
nd	nd	nd	nd	nd	nd	nd	nd	nd
+	+	+	++	+	+/-	+	+/-	B

Chapter 4

four single nucleotide polymorphisms (SNPs) and one silent mutation in the *E-cadherin* gene sequence of ten cell lines. In total, the 55 cell lines included 14 rounded cell lines and eleven of these had a mutation of the *E-cadherin* gene (Table 4.1). Here, we sought to identify the genetic basis for the rounded cell morphology of the three *E-cadherin* wild-type cell lines from our collection: HCC1187, MDA-MB-330 and MDA-MB-468. It was our premise that these three cell lines had mutated another gene from the E-cadherin tumor suppressor pathway.

To search for this putative cancer gene, we have evaluated Affymetrix U133A gene expression profiles of ten *E-cadherin* mutant breast cancer cell lines and nine *E-cadherin* wild-type cell lines from our collection. *E-cadherin* mutant cell lines were: CAMA-1, EVSA-T, MDA-MB-134VI, MDA-MB-453, MPE600, OCUB-F, SK-BR-3, SK-BR-5, SUM44PE and ZR-75-30; and *E-cadherin* wild-type cell lines were: BT474, BT483, MCF-7, MDA-MB-175VII, MDA-MB-361, SUM52PE, SUM185PE, T47D and ZR75-1. All ten *E-cadherin* mutant cell lines had the rounded cell morphology and all nine *E-cadherin* wild-type cell lines expressed E-cadherin at apparently normal levels and had the epithelial cell morphology that we previously had associated with this genotype (Table 4.1 and chapter 3). Significant analysis of microarrays (SAM; [23]) was used to identify differentially expressed genes between the two cell line groups, but no differentially expressed genes were identified at q-values up to 20%. This result was similar to that of a reported study on another, partially overlapping cohort of breast cancer cell lines [24]. It thus appears that mutational inactivation of the *E-cadherin* gene has no significant distinguishing effect on the gene expression program of mutant breast cancer cells.

A more directed search for the putative cancer gene in the E-cadherin tumor suppressor pathway involved expression analysis of the genes from the E-cadherin/catenin protein complex in 41 breast cancer cell lines. These included ten cell lines with *E-cadherin* gene mutations (7 truncation mutants and 3 mutants with in-frame deletions), twelve cell lines with *E-cadherin* promoter methylation and 19 cell lines with wild-type *E-cadherin* genes (Table 4.1 and Figure 4.1). Expression analysis of the four catenin proteins by RT-PCR and immunohistochemistry showed that p120ctn expression levels were non-informative. In contrast, expression levels of the other three catenins revealed five expression patterns that related to the *E-cadherin* gene status Twenty-three cell lines expressed E-cadherin, α-catenin, β-catenin and γ-catenin at apparently normal levels, whether determined at transcript level or at protein level (Pattern A in Table 4.1 and Figure 4.1). These cell lines included all fifteen cell lines with wild-type *E-cadherin* genes, five cell lines with methylated *E-cadherin* genes, and the three cell lines with mutant *E-cadherin* genes that caused in-frame deletions in the encoded proteins. Sixteen cell lines had decreased or no detectable E-cadherin protein expression. One of these, DU4475, had increased β-catenin protein expression, decreased γ-catenin protein expression and apparently normal α-catenin protein expression (Pattern B). It is likely that this protein expression pattern, particularly that of β-catenin, is related to the mutant *APC* gene and aberrant Wnt pathway activation in this cell line ([25, 26] and chapter 7). Eight other cell lines with aberrant E-cadherin protein expression had mostly decreased γ-catenin protein expression and mostly

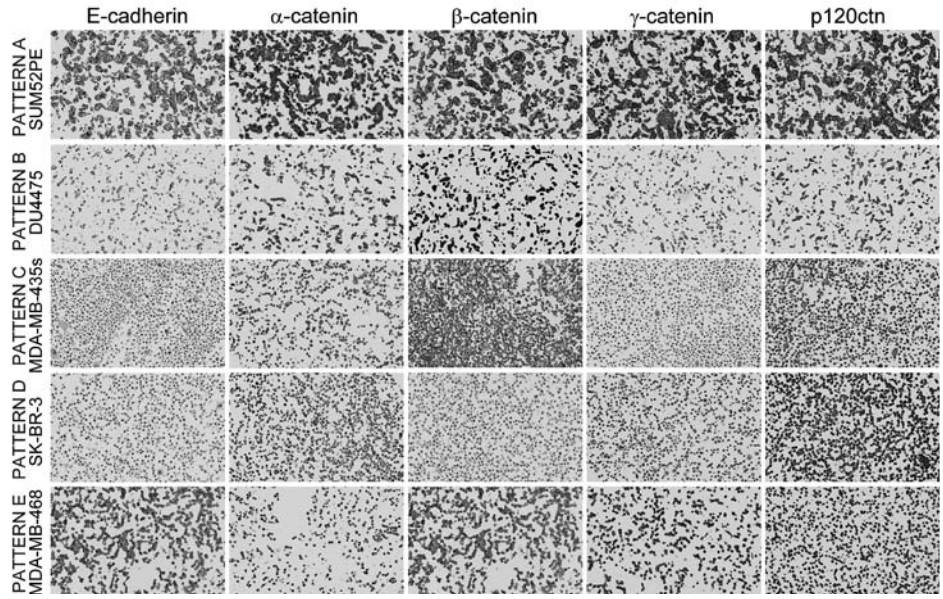


Figure 4.1 Expression analysis of the E-cadherin/catenin protein complex in 41 human breast cancer cell lines. Pattern A entails expression of E-cadherin and all four catenins at apparently normal levels. This pattern is represented by SUM52PE and included fifteen cell lines with wild-type *E-cadherin* genes, five cell lines with (partially) methylated *E-cadherin* genes, and all three cell lines with mutant *E-cadherin* genes that caused in-frame deletions in the encoded proteins. Pattern B is identified for DU4475 only, which had increased β-catenin protein expression, decreased E-cadherin and γ-catenin protein expression and apparently normal α-catenin and p120ctn protein expression. Pattern C entails loss of E-cadherin protein expression, mostly decreased γ-catenin protein expression and mostly normal α-catenin, β-catenin and p120ctn protein expression. Pattern C is represented by MDA-MB-435s and included eight cell lines that had a methylated *E-cadherin* gene promoter. Pattern D entails decreased or absent E-cadherin, β-catenin and γ-catenin protein expression and mostly normal α-catenin and p120ctn protein expression. This pattern is represented by SK-BR-3 and included seven cell lines that had a mutant *E-cadherin* gene that caused a premature termination in the encoded proteins. Pattern E is represented by MDA-MB-468 and included both MDA-MB-468 and MDA-MB-330, which had wild-type *E-cadherin* genes and apparently normal expression levels for all proteins, except for complete absence of α-catenin proteins. This figure is also available in color in the appendix.

normal α-catenin and β-catenin protein expression (Pattern C). All of these eight cell lines had a methylated *E-cadherin* gene status. The remaining seven cell lines with aberrant E-cadherin protein expression had decreased β-catenin and γ-catenin protein expression and mostly normal α-catenin protein expression (Pattern D). All of these latter seven cell lines had a mutant *E-cadherin* gene that caused a premature termination in the encoded proteins. Finally, two cell lines had a remarkable protein expression pattern in that all proteins had apparently normal expression levels except for a complete absence of α-catenin proteins: MDA-MB-330 and MDA-MB-468 (Pattern E). Interestingly, these two α-catenin null cell lines belonged to the set of cell lines for which we had postulated mutation in another gene in the E-cadherin tumor suppressor pathway: the rounded cell lines with wild-type *E-cadherin* genes. Most promising, RT-PCR

analysis revealed that two breast cancer cell lines expressed shortened α-*catenin* transcripts (MDA-MB-468 and MDA-MB-157), suggestive for a mutation or deletion in the α-*catenin* gene. Thus, α-*catenin* appeared a good candidate for the cancer gene that we had postulated in the E-cadherin tumor suppressor pathway.

Truncating *α-catenin* mutations in human breast cancer cell lines

To seek evidence that α-*catenin* is a putative cancer gene in the E-cadherin tumor suppressor pathway, we have analyzed the complete coding sequence of α-*catenin* for genetic alterations in the two cell lines with shortened transcripts. In cell line MDA-MB-468, α-*catenin* exons 4 and 5 were homozygously deleted from the genome, resulting in two aberrant transcript lengths (c.302_588del287; Figure 4.2 A). One transcript lacked exons 4 and 5, predicting a

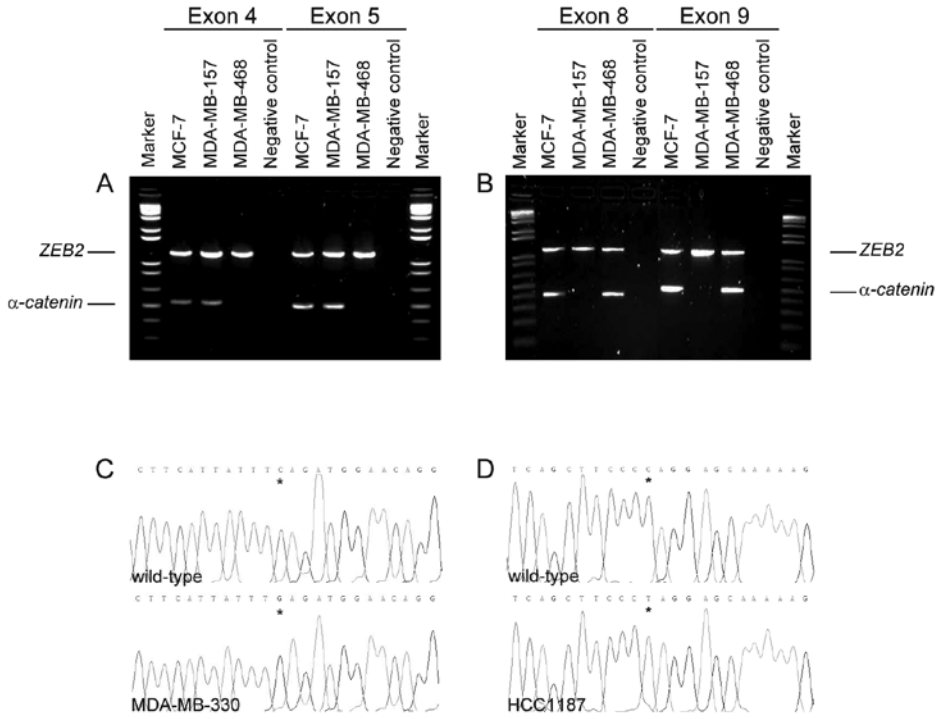


Figure 4.2 Identification of α-*catenin* gene mutations among 55 breast cancer cell lines. (A), homozygous deletion of exons 4 and 5 of α-*catenin* in cell line MDA-MB-468 and (B), homozygous deletion of exons 8 and 9 of α-*catenin* in cell line MDA-MB-157. Duplex PCR amplification products are shown for three breast cancer cell lines, using primers specific for *ZEB2* and α-*catenin* (upper and lower fragments, respectively). Negative control, template negative control; Marker, size marker 1 kb+ DNA ladder (Invitrogen). (C), α-*catenin* nonsense mutation identified in MDA-MB-330 and (D), α-*catenin* nonsense mutation identified in HCC1187. Mutations were identified by PCR amplification and sequencing of genomic DNA (lower electropherograms). The wild-type α-*catenin* gene sequence is shown for comparison (top electropherograms). This figure is also available in color in the appendix.

premature termination of the encoded α-catenin protein (r.302_588del287; p.D102IfsX8). The other transcript lacked exons 3 through 5, predicting an in frame deletion of 161 amino acids (r.106_588del483; p.V36_Q196del161). As we did not detect α-catenin protein expression in MDA-MB-468 with a C-terminal antibody, it appears that the second transcript is not being translated. In cell line MDA-MB-157, we were not able to amplify α-*catenin* exons 8 and 9 from the genome (Figure 4.2 B). Using various combinations of intronic and exonic primers for both exons 8 and 9, we never observed PCR amplification products for exon 8 and, depending on the primer combination, shorter or no amplification products for exon 9. Transcript analysis by RT-PCR allowed amplification of exons 1 through 7, but none of the more 3' located exons. Together, these results suggested a complex genomic rearrangement or translocation in MDA-MB-157, involving deletion of α-*catenin* exon 8 and partial deletion of exon 9. Sequence analysis of α-*catenin* in the other breast cancer cell lines from our collection then revealed homozygous nonsense mutations in HCC1187 and MDA-MB-330 (c.2032C>T / p.Q678X and c.1322C>G / p.S441X, respectively; Figure 4.2 C and D), but no mutations in any of the other cell lines. Additionally, we identified two SNPs and two silent mutations in 22 cell lines (Supplementary Data). The homozygous mutation patterns observed for all four α-*catenin* mutants indicated that they all had lost the other α-*catenin* allele and concordantly, none of the four mutants had α-catenin protein expression, suggesting a tumor suppressor function of α-*catenin*.

Importantly, fourteen of the 55 breast cancer cell lines from our collection had the rounded cell morphology, with three of these carrying bi-allelic inactivating mutations in α-*catenin* and the other eleven carrying bi-allelic inactivating mutations in *E-cadherin*. One additional cell line harbored a bi-allelic inactivating α-*catenin* mutation, totaling to four oncogenic α-*catenin* mutants among 55 breast cancer cell lines. The mutation analysis thus supports α-*catenin* as a putative new tumor suppressor gene in human breast cancer.

Loss of α-catenin protein expression associates with clinical lobular breast cancers

To substantiate our premise that α-*catenin* operates in the E-cadherin tumor suppressor pathway, we have evaluated α-catenin protein expression in clinical breast cancers. Therefore, we have constructed a tissue microarray of 139 formalin-fixed paraffin-embedded (FFPE) breast cancers that included five pathological subtypes of breast carcinoma. Protein expression analysis by immunohistochemistry revealed absence of α-catenin protein expression in fourteen of 139 tumors (10%) and absence of E-cadherin protein expression in 37 of 139 tumors (27%), with thirteen tumors that had lost expression of both proteins (Table 4.2). As expected, loss of E-cadherin protein expression was particularly pronounced among lobular breast cancers (79% of 33 tumors) and did not exceed 20% of tumors from other pathological subtypes. Most important, loss of α-catenin protein expression was also particularly pronounced among lobular breast cancers (33% of 33 tumors) and did not exceed 5% in the other subtypes (Table 4.2). Loss of α-catenin protein expression thus was significantly associated with breast cancers

Table 4.2 E-cadherin and α-catenin protein expression and gene mutations among 139 FFPE clinical breast cancers.

Pathological subtype	No of tumors	Loss of α-catenin protein expression	*α-catenin* gene mutations	Loss of E-cadherin protein expression	*E-cadherin* gene mutations	Loss of α-catenin and E-cadherin protein expression
Lobular	33	11/33 (33%)	0/9 (0%)	26/33 (79%)	10/20 (50%)	11/33 (33%)
Ductal	48	2/48 (4%)	0/2 (0%)	4/48 (8%)	0/2 (0%)	1/48 (2%)
Mucinous	22	1/22 (4%)	0/1 (0%)	4/22 (18%)	0/2 (0%)	1/22 (5%)
Medullary	22	0/22 (0%)	-	3/22 (14%)	0/1 (0%)	0/22 (0%)
Tubular	14	0/14 (0%)	-	0/14 (0%)	-	0/14 (0%)
Total	139	14/139 (10%)	0/12 (0%)	37/139 (27%)	10/25 (42%)	13/139 (9%)

of the lobular subtype ($\chi 2$ $P<0.001$), again suggesting that α-*catenin* is a tumor suppressor that operates in the E-cadherin pathway.

Conclusive evidence that α-*catenin* is a new tumor suppressor gene requires identification of inactivating mutations in (uncultured) clinical cancer samples. We have searched for mutations in the α-*catenin* gene by sequence analysis of twelve available FFPE breast cancers with loss of α-catenin protein expression. No alterations in the α-*catenin* coding sequence were observed in any of the twelve breast cancers. Yet, analysis of the *E-cadherin* gene sequence in a partially overlapping cohort of 25 available FFPE breast cancers with loss of E-cadherin protein expression revealed ten tumors with truncating *E-cadherin* mutations (chapter 3 and Table 4.2). Three of these *E-cadherin* mutant tumors also had lost α-catenin protein expression. Thus, *E-cadherin* mutation analysis identified inactivating mutations in half of lobular breast cancers with loss of E-cadherin and/or α-catenin protein expression, leaving the genetic basis of almost half of lobular breast cancers unidentified — 60% of which had lost α-catenin protein expression.

A cohort of twelve clinical breast cancers may be too small to identify α-*catenin* mutations. We therefore also have analyzed a cohort of 107 fresh-frozen lobular breast cancers for α-*catenin* mutations. Sequence alterations were identified in two breast cancers. One tumor had a thymine to cytosine substitution at five bases before the intron/exon boundary of exon 6, possibly leading to splicing artifacts (c.589-5T>C). However, transcript analysis by RT-PCR and sequencing revealed expression of wild-type α-*catenin* transcripts. Yet, clinical breast cancer samples inevitably also include normal, non-neoplastic cells. We therefore can not exclude that the α-*catenin* transcript in the tumor cells had been downregulated due to a splicing artifact and that we have detected a wild-type transcript derived from the normal tissue present in the sample. However, splice site prediction software did not predict aberrant splicing related to the mutation (http://www.cbs.dtu.dk./services/NetGene2/; [27]). We also identified a missense mutation in α-*catenin* exon 4 in another tumor, predicting the substitution of arginine to glycine at codon 129 (c.385C>G; p.R129G). Although this missense mutation represents a non-conserved amino acid change at an evolutionary conserved amino acid in the encoded protein, it presently is unclear whether the mutation has oncogenic significance. Together, we have identified two α-*catenin* mutations of unclear oncogenic significance among 107 fresh frozen lobular

breast cancers and 12 α-*catenin* negative paraffin embedded breast cancers, momentarily keeping us from conclusively designating α-*catenin* a new tumor suppressor gene in human breast cancer.

DISCUSSION

We have identified four human breast cancer cell lines with bi-allelic truncating α-*catenin* mutations among 55 human breast cancer cell lines, suggesting that α-*catenin* is a new tumor suppressor gene. Similar to the *E-cadherin* mutant breast cancer cell lines, three of the α-*catenin* mutant cell lines had a rounded cell morphology. Importantly, all fourteen rounded breast cancer cell lines in our collection had either a mutation of α-*catenin* or *E-cadherin*, strongly suggesting that the two genes operate in the same tumor suppressor pathway. All four α-*catenin* mutant breast cancer cell lines had lost α-catenin protein expression. Again consistent with a similar biological function, loss of α-catenin and E-cadherin protein expression was most pronounced among clinical breast cancers of lobular pathology. It was therefore quite surprising that we detected only two α-*catenin* gene mutations among 107 fresh-frozen lobular breast cancers and 12 α-*catenin* negative paraffin-embedded breast cancers, both of which were of unclear oncogenic significance. The question arises whether we have failed to detect α-*catenin* mutations in the clinical cancer samples or that we have not analyzed the appropriate sample cohort.

Mutation detection may be severely hampered by the inevitable presence of normal, non-neoplastic cells in clinical tumor samples. We have used PCR and direct sequencing to detect α-*catenin* gene mutations. This approach did allow us to detect *E-cadherin* gene mutations in clinical breast tumors albeit that these mutations all were intragenic deletions or insertions of a few nucleotides. Larger homozygous deletions of one or multiple exons are essentially not detectable by this approach. Enrichment for tumor cells by laser-capture technology still is rather labor-intensive and thus allows analysis of small tumor cohorts only. Fluorescence in situ hybridization (FISH) would allow detection of relatively large homozygous deletions in clinical tumors as-they-are, whereas multiplex ligation-dependent probe amplification (MLPA) allows detection of relatively small homozygous deletions. Two of the four α-*catenin* mutations that we have identified among the breast cancer cell lines were intragenic homozygous deletions, suggesting that MLPA would be the most appropriate alternative method to screen clinical tumors for α-*catenin* mutations. MLPA would also be most suitable to detect germline deletions in blood-derived DNA samples from breast cancer families with lobular breast cancer cases. However, there currently is no α-*catenin* MLPA kit commercially available, precluding us to use this mutation screening approach.

Many tumor suppressor genes have a characteristic spectrum of mutations. Homozygous deletions, for example, are rare for the *p53* tumor suppressor gene but are the most commonly

identified mutation of the *p16* gene. The α-*catenin* mutations that we have identified in the breast cancer cell lines included two intragenic homozygous deletions of each two exons and two nonsense mutations. If these four mutants were typical for the α-*catenin* tumor suppressor gene, we should have been able to detect nonsense mutations in the clinical breast tumors, similar to the *E-cadherin* mutations that we had detected in these samples. However, it may be that the mutation spectrum among cell lines is different from that among clinical tumors. We have observed such a phenomenon for *E-cadherin* gene mutations, where only three of the eleven mutations in breast cancer cell lines were intragenic deletions or insertions of a few nucleotides compared to all of nineteen mutations in clinical breast cancers (chapter 3). One could argue that these results reflect limitations in mutation detection, similar as discussed above. Yet, *E-cadherin* mutation detection by MLPA did not identify homozygous deletions among 40 clinical breast cancers with E-cadherin protein loss (Hollestelle and Schutte, unpublished results). It is important to realize that *E-cadherin* mutations in clinical diffuse-type gastric cancers typically were intragenic homozygous deletions, similar to the *E-cadherin* mutations that we have found among breast cancer cell lines. It thus appears that the intragenic homozygous deletions of *E-cadherin* in the cell lines do not represent mutations that artifactually had been introduced during their *in vitro* propagation. More likely, there is a bias for breast cancers that are more easily established as cell lines, perhaps also including those with α-*catenin* gene mutations. If so, it might be that we will only identify α-*catenin* gene mutations upon screening sufficiently large clinical tumor cohorts.

Many tumor suppressor genes have a characteristic specificity for a particular tumor type or types. For example, somatic *E-cadherin* mutations have been identified in about half of lobular breast cancers and in about half of diffuse-type gastric cancers. Yet, germline *E-cadherin* mutations have been found exclusively in families with diffuse-type gastric cancer but not in breast cancer families with lobular breast cancers ([16, 17] and Hollestelle and Schutte, unpublished results). Also, BRCA1 and BRCA2 mutations rarely have been acquired somatically whereas p53 mutations predominantly have been acquired somatically. It indeed still is unclear what determines tissue specificity of tumor suppressor genes or what determines their preferred somatic or germline occurrence. It may be of interest that incidental α-*catenin* gene mutations had previously been reported for an ovarian cancer cell line, a colon cancer cell line, a prostate cancer cell line and a lung cancer cell line [28-32]. α-*catenin* might be a tumor suppressor gene for multiple tumor types, and identification of mutations in clinical tumors may require mutation screening of a wide variety of tumor types.

Pending identification of α-*catenin* mutations in clinical tumor samples, the argument for α-*catenin* as a new tumor suppressor gene is compelling. The four α-*catenin* mutant breast cancer cell lines that we have identified each had bi-allelic mutations that predicted premature termination of the encoded proteins – the classical hallmark of a tumor suppressor gene. But even more convincing was the mutual exclusive occurrence of α-*catenin* and *E-cadherin* gene mutations in breast cancer cell lines with the rounded cell morphology. Rounded cells with

scant cytoplasm is a defining feature of lobular breast cancers as well as diffuse-type gastric cancers. In concordance, we have shown that E-cadherin mutation was causative for the rounded cell morphology in breast cancer cell lines (chapter 3) and others have shown that α-*catenin* mutations were similarly causative for the diminished cell adhesion in α-*catenin* mutant cell lines [28, 33-35]. Together, these results strongly suggest that α-*catenin* and *E-cadherin* operate in the same tumor suppressor pathway. Our observation that α-catenin and E-cadherin protein expression was lost predominantly among lobular subtype breast cancers further substantiates the functional interrelationship between the two proteins. Whatever the mutation detection method, the tumor type or subtype, it seems a matter of time and effort that α-*catenin* mutations are found in clinical tumors. Till then, α-*catenin* stays a most promising but putative new tumor suppressor gene.

ACKNOWLEDGEMENTS

We thank Adrian Mombrun for technical assistance. This work was supported by grants from KWF Dutch Cancer Society (DDHK 2002-2687), Erasmus MC Mrace 2005 and Netherlands Genomics Initiative (NGI)/Netherlands Organization for Scientific Research (NWO).

REFERENCES

1. Takeichi M. Cadherin cell adhesion receptors as a morphogenetic regulator. Science (New York, NY 1991;251(5000):1451-5.
2. Cowin P. Unraveling the cytoplasmic interactions of the cadherin superfamily. Proc Natl Acad Sci U S A 1994;91(23):10759-61.
3. Jamora C, Fuchs E. Intercellular adhesion, signalling and the cytoskeleton. Nat Cell Biol 2002;4(4):E101-8.
4. Kemler R. From cadherins to catenins: cytoplasmic protein interactions and regulation of cell adhesion. Trends Genet 1993;9(9):317-21.
5. Semb H, Christofori G. The tumor-suppressor function of E-cadherin. American journal of human genetics 1998;63(6):1588-93.
6. Nollet F, Berx G, van Roy F. The role of the E-cadherin/catenin adhesion complex in the development and progression of cancer. Mol Cell Biol Res Commun 1999;2(2):77-85.
7. Drees F, Pokutta S, Yamada S, Nelson WJ, Weis WI. Alpha-catenin is a molecular switch that binds E-cadherin-beta-catenin and regulates actin-filament assembly. Cell 2005;123(5):903-15.
8. Yamada S, Pokutta S, Drees F, Weis WI, Nelson WJ. Deconstructing the cadherin-catenin-actin complex. Cell 2005;123(5):889-901.
9. Berx G, Van Roy F. The E-cadherin/catenin complex: an important gatekeeper in breast cancer tumorigenesis and malignant progression. Breast Cancer Res 2001;3(5):289-93.
10. Hajra KM, Fearon ER. Cadherin and catenin alterations in human cancer. Genes, chromosomes & cancer 2002;34(3):255-68.
11. Peinado H, Olmeda D, Cano A. Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? Nat Rev Cancer 2007;7(6):415-28.
12. Berx G, Cleton-Jansen AM, Nollet F, et al. E-cadherin is a tumour/invasion suppressor gene mutated in human lobular breast cancers. The EMBO journal 1995;14(24):6107-15.
13. Berx G, Cleton-Jansen AM, Strumane K, et al. E-cadherin is inactivated in a majority of invasive human lobular breast cancers by truncation mutations throughout its extracellular domain. Oncogene 1996;13(9):1919-25.
14. Becker KF, Atkinson MJ, Reich U, et al. Exon skipping in the E-cadherin gene transcript in metastatic human gastric carcinomas. Hum Mol Genet 1993;2(6):803-4.
15. Becker KF, Atkinson MJ, Reich U, et al. E-cadherin gene mutations provide clues to diffuse type gastric carcinomas. Cancer research 1994;54(14):3845-52.
16. Guilford P, Hopkins J, Harraway J, et al. E-cadherin germline mutations in familial gastric cancer. Nature 1998;392(6674):402-5.
17. Gayther SA, Gorringe KL, Ramus SJ, et al. Identification of germ-line E-cadherin mutations in gastric cancer families of European origin. Cancer research 1998;58(18):4086-9.
18. Guilford PJ, Hopkins JB, Grady WM, et al. E-cadherin germline mutations define an inherited cancer syndrome dominated by diffuse gastric cancer. Human mutation 1999;14(3):249-55.
19. Keller G, Vogelsang H, Becker I, et al. Diffuse type gastric and lobular breast carcinoma in a familial gastric cancer patient with an E-cadherin germline mutation. The American journal of pathology 1999;155(2):337-42.
20. Pharoah PD, Guilford P, Caldas C. Incidence of gastric cancer and breast cancer in CDH1 (E-cadherin) mutation carriers from hereditary diffuse gastric cancer families. Gastroenterology 2001;121(6):1348-53.
21. Derksen PW, Liu X, Saridin F, et al. Somatic inactivation of E-cadherin and p53 in mice leads to metastatic lobular mammary carcinoma through induction of anoikis resistance and angiogenesis. Cancer Cell 2006;10(5):437-49.
22. Harkes IC, Elstrodt F, Dinjens WN, et al. Allelotype of 28 human breast cancer cell lines and xenografts. Br J Cancer 2003;89(12):2289-92.

23. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A 2001;98(9):5116-21.
24. Lombaerts M, van Wezel T, Philippo K, et al. E-cadherin transcriptional downregulation by promoter methylation but not mutation is related to epithelial-to-mesenchymal transition in breast cancer cell lines. Br J Cancer 2006;94(5):661-71.
25. Schlosshauer PW, Brown SA, Eisinger K, et al. APC truncation and increased beta-catenin levels in a human breast cancer cell line. Carcinogenesis 2000;21(7):1453-6.
26. van de Wetering M, Barker N, Harkes IC, et al. Mutant E-cadherin breast cancer cells do not display constitutive Wnt signaling. Cancer research 2001;61(1):278-84.
27. Brunak S, Engelbrecht J, Knudsen S. Prediction of human mRNA donor and acceptor sites from the DNA sequence. Journal of molecular biology 1991;220(1):49-65.
28. Bullions LC, Notterman DA, Chung LS, Levine AJ. Expression of wild-type alpha-catenin protein in cells with a mutant alpha-catenin gene restores both growth regulation and tumor suppressor activities. Molecular and cellular biology 1997;17(8):4501-8.
29. Vermeulen SJ, Nollet F, Teugels E, et al. The alphaE-catenin gene (CTNNA1) acts as an invasion-suppressor gene in human colon cancer cells. Oncogene 1999;18(4):905-15.
30. Morton RA, Ewing CM, Nagafuchi A, Tsukita S, Isaacs WB. Reduction of E-cadherin levels and deletion of the alpha-catenin gene in human prostate cancer cells. Cancer research 1993;53(15):3585-90.
31. Oda T, Kanai Y, Shimoyama Y, Nagafuchi A, Tsukita S, Hirohashi S. Cloning of the human alpha-catenin cDNA and its aberrant mRNA in a human cancer cell line. Biochemical and biophysical research communications 1993;193(3):897-904.
32. Shimoyama Y, Nagafuchi A, Fujita S, et al. Cadherin dysfunction in a human cancer cell line: possible involvement of loss of alpha-catenin expression in reduced cell-cell adhesiveness. Cancer research 1992;52(20):5770-4.
33. Watabe M, Nagafuchi A, Tsukita S, Takeichi M. Induction of polarized cell-cell association and retardation of growth by activation of the E-cadherin-catenin adhesion system in a dispersed carcinoma line. The Journal of cell biology 1994;127(1):247-56.
34. Ewing CM, Ru N, Morton RA, et al. Chromosome 5 suppresses tumorigenicity of PC3 prostate cancer cells: correlation with re-expression of alpha-catenin and restoration of E-cadherin function. Cancer research 1995;55(21):4813-7.
35. Maeno Y, Moroi S, Nagashima H, et al. alpha-catenin-deficient F9 cells differentiate into signet ring cells. The American journal of pathology 1999;154(5):1323-8.

Chapter 5

Phosphatidylinositol-3-OH Kinase or RAS Pathway Mutations in Human Breast Cancer Cell Lines

Antoinette Hollestelle
Fons Elstrodt
Jord H.A. Nagel
Wouter W. Kallemeijn
Mieke Schutte

Mol Cancer Res 2007; 5(2):195-201

ABSTRACT

Constitutive activation of the PI3K and RAS signaling pathways are important events in tumor formation. This is illustrated by the frequent genetic alteration of several key players from these pathways in a wide variety of human cancers. Here, we report a detailed sequence analysis of the *PTEN*, *PIK3CA*, *KRAS*, *HRAS*, *NRAS* and *BRAF* genes in a collection of 40 human breast cancer cell lines. We identified a surprisingly large proportion of cell lines with mutations in the PI3K or RAS pathways (54% and 25%, respectively), with mutants for each of the six genes. The *PIK3CA*, *KRAS* and *BRAF* mutation spectra of the breast cancer cell lines were similar to those of colorectal cancers. Unlike in colorectal cancers, however, mutational activation of the PI3K pathway was mutually exclusive with mutational activation of the RAS pathway in all but one of 30 mutant breast cancer cell lines (p=0.001). These results suggest that there is a fine distinction between the signaling activators and downstream effectors of the oncogenic PI3K and RAS pathways in breast epithelium and those in other tissues.

INTRODUCTION

The phosphatidylinositol-3-OH kinase (PI3K) and RAS signaling pathways are pivotal to the transduction of extracellular signals to intracellular targets. Both signaling pathways may be activated by growth factors or nutrients in the cell's environment. The subsequent signaling events regulate cell metabolism, cell survival, cell cycle progression and cell growth. Upon activation, usually via receptor tyrosine kinases, PI3K converts phosphatidylinositol-4,5-diphosphate (PIP_2) to its active form, phosphatidylinositol-3,4,5-triphosphate (PIP_3). This lipid second messenger then transduces the activation signal to downstream targets, most notably members of the AKT family of serine/threonine kinases. The PIP_2 to PIP_3 conversion is counteracted by PTEN phosphatase, thus serving a negative feedback for PI3K signaling (reviewed in [1-4]). The RAS proteins are also major effectors of growth factor signaling through RTKs. Ligand-induced activation of receptor tyrosine kinases generates a cascade of signaling events, during which the RAS GTPase proteins are converted from the inactive GDP-bound state to the active GTP-bound state. Activated RAS proteins confer signals to downstream effectors, including members of the RAF family protein kinases, through interaction with their RAS binding domain. RAF kinases, in turn, further transduce the signals upon the mitogen-activated protein kinase pathway or a number of other possible effectors (reviewed in [5-7]).

Cross-talk between the PI3K and RAS signaling pathways may occur at several stages. GTP-bound RAS proteins may directly activate PI3K [8]. Further downstream, activation of the AKT pathway, through PI3K signaling, may converge with signals from the mitogen-activated protein kinase pathway, through RAS signaling, on mammalian target of rapamycin kinase [5, 9]. There are ample downstream effectors of the PI3K and/or RAS pathways, with a variety of signaling routes. Specificity of the signal transduction is determined by the activating extracellular signaling molecules, with an apparent additional specificity related to cell type and cell activation status. Particularly the unraveling of the regulation of this specificity within the PI3K and RAS signaling pathways is currently a major research challenge.

The importance of the PI3K and RAS signaling pathways for cellular processes is illustrated by their frequent mutational activation in human cancers. Cancer is a genetic disease driven by the accumulation of genetic abrogations in pathways that regulate the growth of cells, their survival and their integrity. After the *p53* tumor suppressor, members of the PI3K pathway are most frequently mutated in human cancers. Most prevalent are activating mutations in the *PIK3CA* gene, which encodes the p110α catalytic subunit of PI3K, and inactivating mutations in the *PTEN* tumor suppressor gene. *PIK3CA* amplification is found in ovarian, cervical and thyroid carcinoma [10-12], while mutations are found predominantly in liver, colon and breast tumors [13-15]. Most *PIK3CA* mutations are located in three mutational hot-spot regions in the gene sequence, which result in increased kinase activity of PI3K [13, 16]. The *PTEN* tumor suppressor gene was originally identified by genetic screens of breast cancers and glioblastomas [17, 18], but it soon became apparent that its mutational involvement also includes many other tumor types [2].

Importantly, germline *PTEN* mutations were identified in patients with Cowden Disease [19], and in patients with Bannayan-Zonana syndrome [20] (OMIM #158350 and #153480), two cancer predisposition syndromes that share clinical symptoms such as benign hamartomatous lesions. Similar symptoms are characteristic for the Tuberous Sclerosis and Peutz-Jeghers syndromes, which have been associated with germline mutations in the *TSC1*, *TSC2* and *LKB1* genes [21-24]. Each of these genes encodes downstream effectors from the PI3K signaling pathway, illustrating both the ubiquitous involvement of this pathway and its tissue specificity.

Mutational activation of the RAS signaling pathway in human cancers is mainly achieved by mutations in the *RAS* and *BRAF* genes. Although many RAS GTPases have been identified, activating oncogenic mutations have been reported for only three *RAS* isoforms: *KRAS*, *HRAS* and *NRAS*. Oncogenic *RAS* mutations appear restricted to codons 12, 13 and 61 of the proteins, resulting in constitutive active RAS GTPase. *RAS* mutations have been identified in a wide variety of human tumor types, and display tissue specificity [25]. *KRAS* is frequently mutated in pancreatic cancers and colorectal cancers, whereas mutations in *NRAS* appear to be more pronounced in melanoma and hematological cancers. Activating *BRAF* mutations are also found in many different tumor types, but their mutational involvement is particularly pronounced in melanoma [6, 26]. Oncogenic *BRAF* mutations are restricted mainly to exons 11 and 15 of the gene, and hotspot mutations have been shown to result in increased kinase activity of BRAF [26].

Oncogenic mutations in the PI3K and RAS signaling pathways have been instrumental in deciphering the biology of these pathways. Conversely, knowledge of the functional implications of oncogenic mutations has increased our understanding of human carcinogenesis, through the commonalities as well as the differences between tumor types. Few studies, however, have addressed the mutational activation of both the PI3K pathway and the RAS pathway in a single cohort of human tumor samples. Here, we report a detailed sequence analysis of six genes (*PTEN*, *PIK3CA*, *KRAS*, *HRAS*, *NRAS* and *BRAF*) that are of major importance for the PI3K and RAS signaling pathways in a collection of 40 human breast cancer cell lines.

MATERIALS AND METHODS

Breast cancer cell lines

The 40 human breast cancer cell lines used in this study are listed in Table 5.2. Cell lines EVSA-T, MPE600, and SK-BR-5/7 were kind gifts of Dr. N. de Vleesschouwer (Institut Jules Bordet, Brussels, Belgium), Dr. H.S. Smith (California Pacific Medical Center, San Francisco, CA), and Dr. E. Stockert (Sloan-Kettering Institute for Cancer Research, New York, NY), respectively. The SUM cell lines were generated in the Ethier laboratory (available at http://www.asterand.com). Cell line OCUB-F was obtained from Riken Gene Bank (Tsukuba, Japan), and all other cell lines were obtained from ATCC (Manassas, VA). All cell lines are unique and monoclonal as shown by extensive analysis of nearly 150 polymorphic microsatellite markers [27]. We were unsuccessful in

obtaining constitutional normal tissues or tumor blocks from the cell lines, precluding assessment of the somatic or germline nature of mutations.

Mutational analysis

The complete coding sequences and intron-exon boundaries of *PTEN* (ENSG00000171862) and *PIK3CA* (ENSG00000121879), as well as exons 2 and 3 of the *RAS* genes (ENSG00000133703, ENSG00000174775, ENS00000168638) and exons 7, 11 and 15 of *BRAF* (ENSG00000157764) were analyzed for genetic alterations. For each of the six genes, intronic primers were used to PCR-amplify gene-specific fragments from genomic DNA. *PTEN* transcripts were amplified from total RNA, using the Qiagen (Hilden, Germany) one-step reverse transcription-PCR kit and gene-specific exonic primers (with or without inclusion of gene-specific *HPRT* primers). For sequence analysis, amplification products were incubated with Shrimp Alkaline Phosphatase and Exonuclease-I enzymes, and subsequently sequenced with the Big Dye Terminator Cycle Sequencing kit (Applied Biosystems, Foster City, CA) using an ABI3100 Genetic Analyzer. All sequence variants identified were validated by sequencing an independently amplified PCR product, and for *PTEN* mutants also by transcript sequencing. Allelic loss at the *PTEN* chromosomal locus was determined by microsatellite analysis, using markers *D10S1765, D10S1687* and *D10S1744*. Forward microsatellite primers contained a M13 sequence at their 5' end. Amplification products were obtained by using both the microsatellite primers and a FAM-labeled complementary M13 sequence in a single reaction. Product lengths were determined on an ABI3100 Genetic Analyzer. Primer sequences are available as Supplementary Data. Amplification of the *PIK3CA* locus at chromosome 3q and the *AKT2* locus at chromosome 19q was established from SNP array data that were available for 19 cell lines at www.sanger.ac.uk/cgi, with an intensity ratio cut-off of 1.5 for low-level amplification (equivalent to 3 allele copies).

Gene cloning

PTEN transcripts of cell line CAMA-1 were amplified with the Qiagen one-step reverse transcription-PCR kit, using gene-specific primers designed to include either a *Bam*HI or *Eco*RI restriction site and to span both mutations in CAMA-1 (Supplementary Data). The RT-PCR products were digested with these restriction enzymes and subsequently cloned in the multiple cloning site of the pcDNA3.0 vector (Invitrogen, Paisley, Scotland). Inserts from 14 single colonies were PCR amplified and sequenced using vector-specific primers.

Methylation analysis

Exponentially growing cells were seeded at a density of approximately 1 million cells per T75 flask, in RPMI 1640 with 10% FCS. On each of the following three days, 10 μM filter-sterilized 5-aza-2'-deoxycytidine (Sigma, Steinhein, Germany) was added to the cell cultures. On the fourth day, cells were washed with PBS at 37°C, harvested by lysis in the flask, and total RNA was isolated. As a control, cultures untreated with 5-aza-2'-deoxycytidine were included.

RESULTS

We analyzed forty human breast cancer cell lines for mutations in the *PTEN*, *PIK3CA*, *RAS* and *BRAF* genes, by direct sequencing of PCR-amplified genomic DNA fragments. Mutational analysis of all nine exons of the *PTEN* tumor suppressor gene revealed eight mutant cell lines (Table 5.1). One cell line had a homozygous deletion of exons 1 through 9 of *PTEN*, three cell lines had truncating mutations (IVS4+1G>T, 821delG, 951delACTT), and four cell lines had missense mutations (D92H, L108R, C136Y, E307K). The IVS4+1G>T splice site mutation resulted in the exact deletion of exon 4 from the encoded transcript, predicting a change in the protein sequence after codon 71 with four additional amino acids followed by a stop codon. This splice site mutation has also been identified in the germline of a patient with Cowden Disease, in two endometrial carcinomas, and in a glioblastoma [28], rendering it highly likely that this mutation is relevant for tumorigenesis. The 821delG mutation is also presumed to be oncogenic, as it resulted in a premature stop at codon 275 that has been identified in eight endometrial carcinomas [28]. The 951delACTT mutation resulted in a premature stop at codon 319 that was also found in the germline of a patient with Cowden Disease and in seven endometrial carcinomas, three glioblastomas and a prostate carcinoma [28]. The D92H and C136Y missense mutations are both presumed oncogenic, as a mutation at codon 92 was found in an endometrial carcinoma and C136Y was found in the germline of a patient with Cowden Disease [28]. The L108R mutation has never been reported in clinical cancer samples but is likely oncogenic, as it is located in the phosphatase domain of PTEN which is frequently mutated in Cowden Disease patients, Bannayan-Zonana patients, and in endometrial carcinoma [28]. The E307K mutation also has not been reported, but it is located in the C2 domain of PTEN, and neighboring codons have been found mutated in a Cowden Disease patient and in two endometrial carcinomas [28]. However, the functional significance of the E307K mutation in cell line MDA-MB-453 is unclear, as this mutation is heterozygous and we did not identify additional *PTEN* sequence alterations in this cell line. All other *PTEN* mutant breast cancer cell lines had lost the other *PTEN* allele, except for cell line CAMA-1. CAMA-1 carried the D92H mutation at one allele and had a second mutation at the other allele, where an insertion of four base pairs at position 802 was followed by a deletion of four base pairs at position 834, predicting the exchange of twelve amino acids within the PTEN protein sequence (D268_F279delins12; Figure 5.1). The biallelic nature of the mutations in CAMA-1 was confirmed by transcript analysis and by cloning and sequencing of transcript fragments, both only identifying the D92H mutation. We also identified a possible primer site polymorphism in cell line UACC893, as we were unable to PCR amplify exon 2 from genomic DNA even though sequence analysis revealed expression of the wild-type *PTEN* transcript. Analysis of *PTEN* transcript expression by RT-PCR revealed that cell lines HCC1937, MDA-MB-436 and SUM149PT did not express *PTEN* transcripts (Figure 5.2). Whereas cell line HCC1937 had a homozygous deletion of the *PTEN* gene, both MDA-MB-436 and SUM149PT had a wild-type *PTEN* gene sequence (Table 5.2). We excluded transcriptional silencing through

Table 5.1 Mutations identified in the *PTEN, PIK3CA, RAS and BRAF* genes in human breast cancer cell lines

Breast Cancer Cell Line	Affected Gene	Gene sequence	Transcript sequence	Predicted protein effect*	Oncogenic
HCC1937†	*PTEN*	HD Ex. 1-9‡	not detectable	no expression	yes
MDA-MB-468†	*PTEN*	IVS4+1G>T‡	c.del210_253 (Ex. 4)	A72fsX5	yes
BT549†	*PTEN*	821delG‡	821delG	V275X	yes
EVSA-T	*PTEN*	951delACTT‡	951delACTT	T319X	yes
CAMA-1	*PTEN*	274G>C §	274G>C	D92H	yes
		802insTAGG/ 834delCTTC §	not detectable	no expression	yes
ZR-75-1†	*PTEN*	323T>G‡	323T>G	L108R	likely
MDA-MB-415†	*PTEN*	407G>A‡	407G>A	C136Y	yes
MDA-MB-453	*PTEN*	919G>A	919G>A	E307K	likely
BT474†	*PIK3CA*	333G>C	na	K111N	yes
BT20†	*PIK3CA*	1616C>G	na	P539R	yes
BT483†	*PIK3CA*	1624G>A	na	E542K	yes
MCF-7†	*PIK3CA*	1633G>A	na	E545K	yes
MDA-MB-361†	*PIK3CA*	1633G>A	na	E545K	yes
MDA-MB-361	*PIK3CA*	1700A>G	na	K567R	likely
BT20†	*PIK3CA*	3140A>G	na	H1047R	yes
MDA-MB-453†	*PIK3CA*	3140A>G	na	H1047R	yes
OCUB-F	*PIK3CA*	3140A>G‡	na	H1047R	yes
SK-BR-5	*PIK3CA*	3140A>G	na	H1047R	yes
SUM102PT†	*PIK3CA*	3140A>G	na	H1047R	yes
SUM185PE†	*PIK3CA*	3140A>G‡	na	H1047R	yes
SUM190PT†	*PIK3CA*	3140A>G	na	H1047R	yes
T47D†	*PIK3CA*	3140A>G	na	H1047R	yes
UACC893†	*PIK3CA*	3140A>G	na	H1047R	yes
SUM159PT†	*PIK3CA*	3140A>T	na	H1047L	yes
MDA-MB-134VI†	*KRAS*	34G>C	na	G12R	yes
SK-BR-7	*KRAS*	34G>T	na	G12C	yes
SUM229PE	*KRAS*	35G>A	na	G12D	yes
MPE600	*KRAS*	35G>T	na	G12V	yes
MDA-MB-231†	*KRAS*	38G>A	na	G13D	yes
Hs578T†	*HRAS*	35G>A	na	G12D	yes
SUM159PT	*HRAS*	35G>A	na	G12D	yes
SK-BR-7	*NRAS*	182A>G	na	Q61R	yes
ZR-75-30†	*BRAF*	977T>C	na	I326T	unknown
MDA-MB-231	*BRAF*	1391G>T	na	G464V	yes
DU4475	*BRAF*	1799T>A	na	V600E	yes
MDA-MB-435s	*BRAF*	1799T>A	na	V600E	yes

Abbreviations: HD, homozygous deletion; Ex, exon; IVS, intervening sequence; del, deletion; ins, insertion; NA, not analyzed.

* Frameshift mutations are indicated by the first changed codon and the number of newly encoded codons, including premature termination codon X.

† cell lines were reported to be mutated in [17, 26, 38, 45-49].

‡ Mutations were homozygous based on sequence analysis and confirmed with polymorphic markers.

§ Mutations are heterozygous, but located on different alleles.

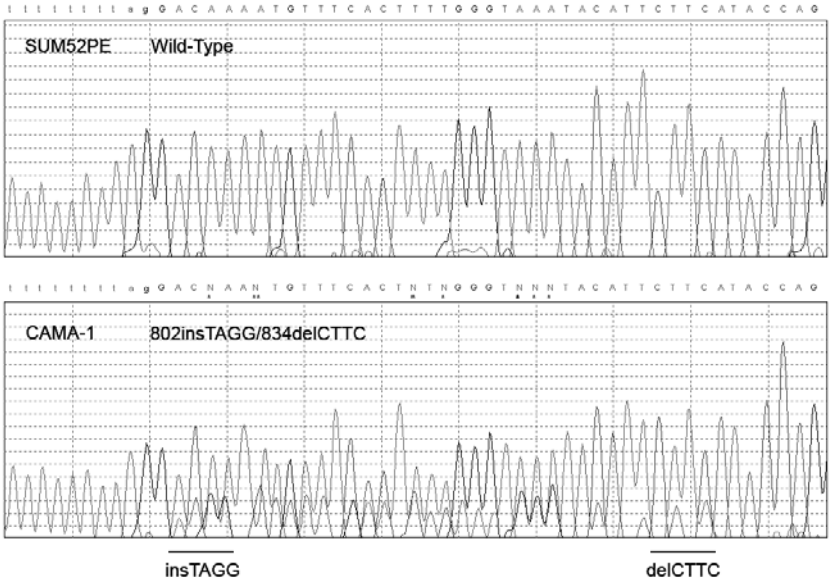


Figure 5.1 Identification of the *PTEN* 802insTAGG/834delCTTC mutation in cell line CAMA-1 by PCR amplification and sequencing of genomic DNA (bottom electropherogram). The wild-type *PTEN* gene sequence is shown for comparison (top electropherogram). This figure is also available in color in the appendix.

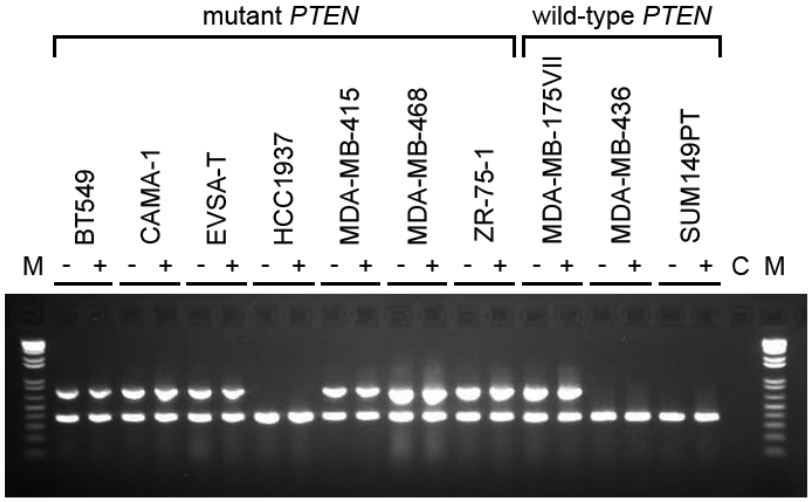


Figure 5.2 Analysis of *PTEN* transcriptional silencing through promoter methylation by cell culture in the presence (+) or absence (-) of 5-azacytidine. Reverse transcription-PCR amplification products are shown from seven mutant and three wild-type *PTEN* breast cancer cell lines, using primers specific for *PTEN* (top fragments) and the *HPRT* housekeeper (bottom fragments). These cell lines included the three cell lines without detectable *PTEN* expression, but there was no indication of *PTEN* promoter methylation. C, template negative control; M, size marker (1 kb+ DNA ladder, Invitrogen, Paisley, Scotland).

hypermethylation of the *PTEN* promoter region as a probable cause, by culturing the cell lines in the presence of the demethylating agent 5-azacytidine. As a result, neither MDA-MB-436 nor SUM149PT re-expressed *PTEN* transcripts even though 5-azacytidine did induce expression of *E-cadherin* transcripts in both cell lines (Figure 5.2 and chapter 3). Together, seven of 38 (18%) breast cancer cell lines had biallelic inactivating *PTEN* mutations, one cell line had a monoallelic missense mutation and two cell lines did not express *PTEN* transcripts for reasons unknown.

Mutational analysis of the complete coding sequence of the *PIK3CA* oncogene revealed 16 missense mutations in 14 breast cancer cell lines (Table 5.1). The mutations K111N, P539R, E542K, K567R and H1047L were each identified in one cell line, the E545K mutation was found in two cell lines, and the H1047R mutation was found in nine cell lines. All mutations except for K111N and K567R were previously identified in colon carcinomas [13], and functional analysis of the E542K, E545K and H1047R mutations had shown that these mutations were oncogenic [16, 29]. Although the K111N mutation was not previously reported in a primary cancer, this codon was found deleted in a colon carcinoma [13], suggesting that the K111N mutation is oncogenic. The K567R mutation has also not been reported, but its location in the helicase domain of PIK3CA suggests that it may have functional implications. Notably, we identified the K567R mutation in cell line MDA-MB-361, that also carried the oncogenic E545K mutation. Similarly, cell line BT20 carried both the P539R and H1047R mutations. All *PIK3CA* mutations were heterozygous, except for the H1047R mutations in cell lines OCUB-F and SUM185PE. In addition, we identified the as yet unreported synonymous 363C>T alteration in cell line MDA-MB-231, and the 1173A>G single nucleotide polymorphism in five cell lines (SNP rs3729680; heterozygous in SUM52PE, T47D and ZR-75-30, and homozygous in MDA-MB-231 and SUM149PT). Available SNP array data for nineteen cell lines revealed a single low-level amplification of 4 copies at the *PIK3CA* locus for the mutant cell line T47D and no amplifications at the *AKT2* locus, suggesting that *PIK3CA* and *AKT2* amplification is uncommon in breast cancer (average intensity ratio for *PIK3CA* was 1.1, range 0.7 to 1.9; average intensity ratio for *AKT2* was 1.0, range 0.7 to 1.4; www.sanger.ac.uk/cgi). Together, we identified activating *PIK3CA* mutations in 14 of 39 (36%) breast cancer cell lines.

Mutational analysis of exons 2 and 3 of the three human *RAS* oncogenes revealed eight heterozygous *RAS* mutations in seven of 40 breast cancer cell lines (18%; Table 5.1). We identified five cell lines with each a different *KRAS* mutation (G12C, G12D, G12R, G12V and G13D). The *HRAS* G12D mutation was found in two cell lines and the *NRAS* Q61R mutation was found once. The latter mutation was identified in cell line SK-BR-7, that also carried the *KRAS* G12C mutation. In addition to these well described oncogenic *RAS* mutations, we identified the synonymous *HRAS* 81T>C SNP in 15 cell lines (SNP rs12628; heterozygous in BT483, MDA-MB-175VII, MDA-MB-415 and SK-BR-3, and homozygous in BT20, BT474, CAMA-1, HCC1937, MDA-MB-453, MPE600, SK-BR-5, SK-BR-7, SUM149PT, SUM159PT and T47D).

Mutational analysis of exons 7, 11 and 15 of the *BRAF* oncogene revealed four of 40 breast cancer cell lines with a heterozygous *BRAF* mutation (10%; Table 5.1). We identified the I326T

Table 5.2 Mutational activation of the PI3K and RAS pathways is mutually exclusive in human breast cancer cell lines

Breast Cancer Cell Lines	PTEN	PIK3CA	KRAS	BRAF	HRAS	NRAS
BT549	V275X					
CAMA-1	D92H					
EVSA-T	T319X					
HCC1937	no protein					
MDA-MB-415	C136Y					
MDA-MB-468	A72fsX5					
ZR-75-1	L108R					
MDA-MB-453	E307K	H1047R				
BT20		P539R / H1047R				
MDA-MB-361		E545K / K567R				
BT474		K111N				
BT483		E542K				
MCF-7		E545K				
OCUB-F		H1047R				
SK-BR-5		H1047R				
SUM102PT		H1047R				
SUM185PE		H1047R				
SUM190PT	na	H1047R				
T47D		H1047R				
UACC893		H1047R				
SUM159PT		H1047L			G12D	
Hs578T					G12D	
SK-BR-7			G12C			Q61R
MDA-MB-134VI			G12R			
MPE600			G12V			
SUM229PE			G12D			
MDA-MB-231			G13D	G464V		
MDA-MB-435s				V600E		
DU4475				V600E		
ZR-75-30				I326T		
MDA-MB-157						
MDA-MB-175VII						
MDA-MB-330						
MDA-MB-436						
SK-BR-3						
SUM149PT						
SUM225CWN	na	na				
SUM1315MO2						
SUM52PE						
UACC812						
Mutation Rate	8 of 38 (21%)	14 of 39 (36%)	5 of 40 (13%)	4 of 40 (10%)	2 of 40 (5%)	1 of 40 (3%)

Note: overview of mutations that were identified in 40 human breast cancer cell lines.
The mutations are detailed in Table 5.1.
Abbreviation: NA, not analyzed.

and G464V mutations each in a single cell line and the V600E mutation was found in two cell lines. The V600E mutation is the most frequently identified oncogenic mutation in the *BRAF* gene. The G464V mutation is less frequently identified, but also considered to be oncogenic as it is located within the highly conserved G loop region [26]. Importantly, the G464V and V600E mutations both resulted in an increased activity of BRAF kinase [26]. So far, the I326T variant has only been identified in the ZR-75-30 breast cancer cell line and its functional effect is yet unknown [26]. It is important to note that the *BRAF* mutant MDA-MB-435s cell line was recently shown to be genetically identical to the M14 melanoma cell line, although it had not conclusively been investigated which of the two cell lines was correct ([30] and references therein). Since *BRAF* mutations typically associate with melanoma, one could perhaps also wonder on the origin of the other three *BRAF* mutant breast cancer cell lines. Based on gene expression and methylation profiles, there is no reason to doubt the breast origin of MDA-MB-231 [31-33]. No profiles have been reported for ZR-75-30 and DU4475, but our recent identification of a truncating *E-cadherin* mutation in cell line ZR-75-30 renders it likely that this cell line indeed is of breast origin (chapter 3). We can not be certain on DU4475, as we have as yet not identified breast-specific mutations in this cell line. But then, one never can be sure about the origin of a cancer cell line. Even so, we identified four *BRAF* mutant breast cancer cell lines or, when MDA-MB-435s and DU4475 would turn out not to be of breast origin, two *BRAF* mutants were identified.

DISCUSSION

We performed a mutational analysis of six major cancer genes from the PI3K and RAS signaling pathways in a collection of 40 human breast cancer cell lines. We identified 26 unique mutations: nine mutations in *PTEN*, seven mutations in *PIK3CA*, five in *KRAS*, one each in *HRAS* and *NRAS*, and three in *BRAF*. Four of these mutations have not yet been described in the literature (Table 5.1). In total, 30 of the 40 breast cancer cell lines had mutations in any of these six genes, 40% of which had not yet been reported (Table 5.1). This detailed mutational analysis of the PI3K and RAS pathway genes is complemented by our previously reported mutational analyses of the *E-cadherin*, *MKK4*, *p53* and *BRCA1* genes, rendering this collection of breast cancer cell lines a valuable model for functional and pharmacological studies [34-37].

Mutational activation of the PI3K signaling pathway was detected in 21 breast cancer cell lines (Table 5.2). Two cell lines were *PIK3CA* double mutants. Cell line BT20 carried the P539R and H1047R mutations, for which kinase assays had shown that the H1047R mutation resulted in a substantially higher PI3K activity [13, 16]. Cell line MDA-MB-361 carried the E545K and K567R mutations, of which only the E545K mutation had been previously identified and had been shown to increase PI3K activity [16]. It is conceivable that these *PIK3CA* double mutants reflect a progression of tumorigenesis through further mutational activation of PI3K. In this scenario,

the more oncogenic H1047R and E545K mutations would have been the second hit of the PI3K pathway in the original breast cancers. Indeed, *PIK3CA* double mutant tumors have previously been reported for three primary breast cancers and a gastric cancer [38, 39], suggesting that a two-hit mutational activation of the PI3K pathway may not be uncommon. Similarly, we identified the highly oncogenic *PIK3CA* H1047R mutation together with the *PTEN* E307K mutation in cell line MDA-MB-453. Importantly, MDA-MB-453 had retained a wild-type *PTEN* allele. As the *PTEN* E307K mutation is located in a mutational hot-spot domain [28], it appears that *PTEN* is haploinsufficient in cell line MDA-MB-453. Mutation of *PIK3CA* at its critical H1047 residue would then have been the second hit to full activation of the PI3K pathway in cell line MDA-MB-453. Of course, a two-hit activation of the PI3K signaling pathway awaits further confirmation in primary cancer specimens, allowing dissection of tumor progression by mutational analysis of the earlier premalignant tumor lesions. Either way, our observation of mutational activation of the PI3K pathway in half of human breast cancer cell lines suggests that this signaling pathway may be more important for breast carcinogenesis than currently perceived.

Mutational activation of the RAS signaling pathway was detected in 10 breast cancer cell lines (Table 5.2). We were somewhat surprised by the 13% *KRAS* mutation frequency among the breast cancer cell lines, given the general conviction that *KRAS* mutations are relatively rare in human breast cancers [25]. Two *RAS* double mutant cell lines were identified. Cell line SK-BR-7 carried the *KRAS* G12C mutation and the *NRAS* Q61R mutation, whereas cell line MDA-MB-231 carried the *KRAS* G13D mutation and the *BRAF* G464V mutation. The *BRAF* G464V mutation was shown to be a less potent activator of BRAF kinase than the more prevalent *BRAF* V600E mutation (2 and 10 times wild-type kinase activity, respectively) [26]. One again can conceive a two-hit activation of the RAS pathway, through the *BRAF* G464V mutation and subsequent mutation of *KRAS* G13D. In agreement, only *KRAS* and *BRAF* V600E mutations were reportedly mutually exclusive in colorectal cancers, and one of the four reported double mutants harbored the same combination of *KRAS* G13D with *BRAF* G464V [26, 40].

We identified an unexpected large proportion of breast cancer cell lines with mutational activation of the PI3K and RAS signaling pathways (54% and 25%, respectively). Perhaps even more surprising was that only one of the 30 mutant cell lines had mutations in both pathways (*PIK3CA* H1047L and *HRAS* G12D; Table 5.2), suggesting that mutational activation of the PI3K pathway is essentially mutually exclusive with mutational activation of the RAS pathway in breast cancer (χ^2 p=0.0012, with exclusion of SUM225CWN from the analysis, and p=0.0043 when MDA-MB-435s and DU4475 were also excluded; Table 5.2). This could imply that signals critical for breast carcinogenesis converge through the PI3K and RAS pathways, targeting a single downstream effector. Concurrent mutational activation of both the PI3K and RAS pathways would then not be observed, as double mutants would not have a selective growth advantage over single mutants. In this respect, it is of interest that the mutation spectra of genes from the PI3K and RAS pathway may differ among tumor types. For example, the *BRAF* mutation spectra of breast cancers, colorectal cancers and melanomas are each dominated by the V600E mutation.

However, these three tumor types differ in that activating *RAS* mutations occur predominantly in the *NRAS* gene in melanomas and in the *KRAS* gene in breast cancers and colorectal cancers [25]. Similarly, breast cancers and colorectal cancers share a *PIK3CA* mutation spectrum that is dominated by the H1047R, E545K and E542K mutations, whereas *PIK3CA* mutations are rare in melanomas [41]. Breast cancers and colorectal cancers thus have similar *PIK3CA*, *BRAF* and *KRAS* mutation spectra. Yet, *PIK3CA* mutations are coincident with RAS pathway mutations in colorectal cancers [42], whereas we found that in breast cancers mutational activation of the PI3K pathway was mutually exclusive with mutational activation of the RAS pathway. In melanoma on the other hand, *PTEN* mutations are coincident with *BRAF* mutations, but not with mutations of *NRAS* [43, 44]. Such differences in PI3K and RAS pathway mutations among human tumor types suggest that there is a tissue-specific distinction in the activation and transduction of signals through these oncogenic pathways, at the very least for the skin and epithelia of the colon and breast.

REFERENCES

1. Vivanco I, Sawyers CL. The phosphatidylinositol 3-Kinase AKT pathway in human cancer. Nat Rev Cancer 2002;2(7):489-501.
2. Parsons R. Human cancer, PTEN and the PI-3 kinase pathway. Semin Cell Dev Biol 2004;15(2):171-6.
3. Hay N. The Akt-mTOR tango and its relevance to cancer. Cancer Cell 2005;8(3):179-83.
4. Bader AG, Kang S, Zhao L, Vogt PK. Oncogenic PI3K deregulates transcription and translation. Nat Rev Cancer 2005;5(12):921-9.
5. Shaw RJ, Cantley LC. Ras, PI(3)K and mTOR signalling controls tumour cell growth. Nature 2006;441(7092):424-30.
6. Garnett MJ, Marais R. Guilty as charged: B-RAF is a human oncogene. Cancer Cell 2004;6(4):313-9.
7. Rapp UR, Gotz R, Albert S. BuCy RAFs drive cells into MEK addiction. Cancer Cell 2006;9(1):9-12.
8. Rodriguez-Viciana P, Warne PH, Dhand R, et al. Phosphatidylinositol-3-OH kinase as a direct target of Ras. Nature 1994;370(6490):527-32.
9. Cully M, You H, Levine AJ, Mak TW. Beyond PTEN mutations: the PI3K pathway as an integrator of multiple inputs during tumorigenesis. Nat Rev Cancer 2006;6(3):184-92.
10. Shayesteh L, Lu Y, Kuo WL, et al. PIK3CA is implicated as an oncogene in ovarian cancer. Nat Genet 1999;21(1):99-102.
11. Ma YY, Wei SJ, Lin YC, et al. PIK3CA as an oncogene in cervical cancer. Oncogene 2000;19(23):2739-44.
12. Wu G, Mambo E, Guo Z, et al. Uncommon mutation, but common amplifications, of the PIK3CA gene in thyroid tumors. J Clin Endocrinol Metab 2005;90(8):4688-93.
13. Samuels Y, Wang Z, Bardelli A, et al. High frequency of mutations of the PIK3CA gene in human cancers. Science 2004;304(5670):554.
14. Samuels Y, Velculescu VE. Oncogenic mutations of PIK3CA in human cancers. Cell Cycle 2004;3(10):1221-4.
15. Samuels Y, Ericson K. Oncogenic PI3K and its role in cancer. Curr Opin Oncol 2006;18(1):77-82.
16. Kang S, Bader AG, Vogt PK. Phosphatidylinositol 3-kinase mutations identified in human cancer are oncogenic. Proc Natl Acad Sci U S A 2005;102(3):802-7.
17. Li J, Yen C, Liaw D, et al. PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer. Science 1997;275(5308):1943-7.
18. Steck PA, Pershouse MA, Jasser SA, et al. Identification of a candidate tumour suppressor gene, MMAC1, at chromosome 10q23.3 that is mutated in multiple advanced cancers. Nat Genet 1997;15(4):356-62.
19. Liaw D, Marsh DJ, Li J, et al. Germline mutations of the PTEN gene in Cowden disease, an inherited breast and thyroid cancer syndrome. Nat Genet 1997;16(1):64-7.
20. Marsh DJ, Dahia PL, Zheng Z, et al. Germline mutations in PTEN are present in Bannayan-Zonana syndrome. Nat Genet 1997;16(4):333-4.
21. van Slegtenhorst M, de Hoogt R, Hermans C, et al. Identification of the tuberous sclerosis gene TSC1 on chromosome 9q34. Science 1997;277(5327):805-8.
22. Identification and characterization of the tuberous sclerosis gene on chromosome 16. The European Chromosome 16 Tuberous Sclerosis Consortium. Cell 1993;75(7):1305-15.
23. Hemminki A, Markie D, Tomlinson I, et al. A serine/threonine kinase gene defective in Peutz-Jeghers syndrome. Nature 1998;391(6663):184-7.
24. Jenne DE, Reimann H, Nezu J, et al. Peutz-Jeghers syndrome is caused by mutations in a novel serine threonine kinase. Nat Genet 1998;18(1):38-43.
25. Bos JL. ras oncogenes in human cancer: a review. Cancer Res 1989;49(17):4682-9.
26. Davies H, Bignell GR, Cox C, et al. Mutations of the BRAF gene in human cancer. Nature 2002;417(6892):949-54.
27. Harkes IC, Elstrodt F, Dinjens WN, et al. Allelotype of 28 human breast cancer cell lines and xenografts. Br J Cancer 2003;89(12):2289-92.

28. Bonneau D, Longy M. Mutations of the human PTEN gene. Hum Mutat 2000;16(2):109-22.
29. Bader AG, Kang S, Vogt PK. Cancer-specific mutations in PIK3CA are oncogenic in vivo. Proc Natl Acad Sci U S A 2006;103(5):1475-9.
30. Rae JM, Creighton CJ, Meck JM, Haddad BR, Johnson MD. MDA-MB-435 cells are derived from M14 Melanoma cells--a loss for breast cancer, but a boon for melanoma research. Breast Cancer Res Treat 2007;104(1):13-9.
31. Ross DT, Scherf U, Eisen MB, et al. Systematic variation in gene expression patterns in human cancer cell lines. Nat Genet 2000;24(3):227-35.
32. Scherf U, Ross DT, Waltham M, et al. A gene expression database for the molecular pharmacology of cancer. Nat Genet 2000;24(3):236-44.
33. Paz MF, Fraga MF, Avila S, et al. A systematic profile of DNA methylation in human cancer cell lines. Cancer Res 2003;63(5):1114-21.
34. van de Wetering M, Barker N, Harkes IC, et al. Mutant E-cadherin breast cancer cells do not display constitutive Wnt signaling. Cancer Res 2001;61(1):278-84.
35. Su GH, Song JJ, Repasky EA, Schutte M, Kern SE. Mutation rate of MAP2K4/MKK4 in breast carcinoma. Hum Mutat 2002;19(1):81.
36. Wasielewski M, Elstrodt F, Klijn JG, Berns EM, Schutte M. Thirteen new p53 gene mutants identified among 41 human breast cancer cell lines. Breast Cancer Res Treat 2006;99:97-101.
37. Elstrodt F, Hollestelle A, Nagel JH, et al. BRCA1 mutation analysis of 41 human breast cancer cell lines reveals three new deleterious mutants. Cancer Res 2006;66(1):41-5.
38. Saal LH, Holm K, Maurer M, et al. PIK3CA mutations correlate with hormone receptors, node metastasis, and ERBB2, and are mutually exclusive with PTEN loss in human breast carcinoma. Cancer Res 2005;65(7):2554-9.
39. Lee JW, Soung YH, Kim SY, et al. PIK3CA gene is frequently mutated in breast carcinomas and hepatocellular carcinomas. Oncogene 2005;24(8):1477-80.
40. Rajagopalan H, Bardelli A, Lengauer C, Kinzler KW, Vogelstein B, Velculescu VE. Tumorigenesis: RAF/RAS oncogenes and mismatch-repair status. Nature 2002;418(6901):934.
41. Omholt K, Krockel D, Ringborg U, Hansson J. Mutations of PIK3CA are rare in cutaneous melanoma. Melanoma Res 2006;16(2):197-200.
42. Velho S, Oliveira C, Ferreira A, et al. The prevalence of PIK3CA mutations in gastric and colon cancer. Eur J Cancer 2005;41(11):1649-54.
43. Haluska FG, Tsao H, Wu H, Haluska FS, Lazar A, Goel V. Genetic alterations in signaling pathways in melanoma. Clin Cancer Res 2006;12(7 Pt 2):2301s-7s.
44. Curtin JA, Fridlyand J, Kageshita T, et al. Distinct sets of genetic alterations in melanoma. N Engl J Med 2005;353(20):2135-47.
45. Tomlinson GE, Chen TT, Stastny VA, et al. Characterization of a breast cancer cell line derived from a germ-line BRCA1 mutation carrier. Cancer Res 1998;58:3237-42.
46. Li J, Simpson L, Takahashi M, et al. The PTEN/MMAC1 tumor suppressor induces cell death that is rescued by the AKT/protein kinase B oncogene. Cancer Res 1998;58:5667-72.
47. Prosperi MT, Dupre G, Lidereau R, Goubin G. Point mutation at codon 12 of the Ki-ras gene in a primary breast carcinoma and the MDA-MB-134 human mammary carcinoma cell line. Cancer Lett 1990;51:169-74.
48. Kozma SC, Bogaard ME, Buser K, et al. The human c-Kirsten ras gene is activated by a novel mutation in codon 13 in the breast carcinoma cell line MDA-MB231. Nucleic Acids Res 1987;15:5963-71.
49. Kraus MH, Yuasa Y, Aaronson SA. A position 12-activated H-ras oncogene in all HS578T mammary carcinosarcoma cells but not normal mammary cells of the same patient. Proc Natl Acad Sci U S A 1984;81:5384-8.

Chapter 6

BRCA1 Mutation Analysis of 41 Human Breast Cancer Cell Lines Reveals Three New Deleterious Mutants

Fons Elstrodt
Antoinette Hollestelle
Jord H.A. Nagel
Michael Gorin
Marijke Wasielewski
Ans van den Ouweland
Sofia D. Merajver
Stephen P. Ethier
Mieke Schutte

Cancer Res 2006; 66(1):41-45

ABSTRACT

Germline mutations of the *BRCA1* gene confer a high risk of breast cancer and ovarian cancer to female mutation carriers. The BRCA1 protein is involved in the regulation of DNA repair. How specific tumor-associated mutations affect the molecular function of BRCA1, however, awaits further elucidation. Cell lines that harbor *BRCA1* gene mutations are invaluable tools for such functional studies. Up to now, the HCC1937 cell line was the only human breast cancer cell line with an identified *BRCA1* mutation. In this study, we identified three other *BRCA1* mutants from among 41 human breast cancer cell lines by sequencing of the complete coding sequence of *BRCA1*. Cell line MDA-MB-436 had the 5396+1G>A mutation in the splice donor site of exon 20. Cell line SUM149PT carried the 2288delT mutation and SUM1315MO2 carried the 185delAG mutation. All three mutations were accompanied by loss of the other *BRCA1* allele. The 185delAG and 5396+1G>A mutations are both classified as pathogenic mutations. In contrast with wild-type cell lines, none of the *BRCA1* mutants expressed nuclear BRCA1 proteins as detected with Ab-1 and Ab-2 anti-BRCA1 monoclonal antibodies. These three new human *BRCA1* mutant cell lines thus seem to be representative breast cancer models that could aid in further unraveling of the function of BRCA1.

INTRODUCTION

Germline mutations of the *BRCA1* breast cancer susceptibility gene predispose female carriers to develop breast cancer and ovarian cancer (OMIM 113705; http://www.ncbi.nlm.nih.gov/omim/). The BRCA1 protein normally resides in a nuclear multiprotein complex, including BRCA2, BARD1, and RAD51, and the DNA damage repair proteins MSH2, MLH1, MSH6, ATM, NBS1, MRE11, RAD50, BLM, and RFC. This BRCA1-associated genome surveillance complex functions as a sensor of abnormal DNA structures, such as double-strand breaks and base pair mismatches. BRCA1 has been suggested to have a pivotal function within BRCA1-associated genome surveillance complex by coordinating the actions of damage-sensing proteins and executive repair proteins. BRCA1 may also act as a transcriptional regulator of genes involved in checkpoint reinforcement and, in complexes with BARD1, as a ubiquitin ligase (reviewed in [1-4]). Thus, mutations of *BRCA1* likely impair the repair of damaged DNA, thereby rendering the mutant cells prone to malignant transformation. To fully unravel the function of BRCA1 in DNA damage responses, cell lines with naturally occurring mutations of the gene provide invaluable research tools as they allow extensive analyses and in vitro manipulation. Only a single human *BRCA1* mutant breast cancer cell line had thus far been described (HCC1937; [5]). To identify additional mutants, we screened 41 human breast cancer cell lines for alterations in the *BRCA1* gene sequence.

MATERIALS AND METHODS

Breast cancer cell lines

The 41 human breast cancer cell lines used in this study are listed in Table 6.1. The SUM-series were generated in the Ethier laboratory (available at www.asterand.com). Cell lines EVSA-T, MPE600 and SK-BR-5/7 were kind gifts of Dr. N. de Vleesschouwer (Institut Jules Brodet, Brussels, Belgium), Dr. H.S. Smith (California Pacific Medical Center, San Francisco, CA) and Dr. E. Stockert (Sloan-Kettering Institute for Cancer Research, New York, NY), respectively. Cell line OCUB-F was obtained from Riken Gene Bank. All other cell lines were obtained from American Type Culture Collection. Extensive analysis of near 150 polymorphic microsatellite markers had shown that all cell lines are unique and monoclonal [6].

Mutation analysis

The complete coding sequence and exon-intron boundaries of *BRCA1* (Genbank U14680) were analyzed for genetic alterations in all cell lines, except for SUM44PE and ZR-75-30 (only exons 11-15 and exons 3-7 and 11-15 were analyzed, respectively). Exons 1a through 11 and 16 through 24 were PCR-amplified from genomic DNA templates and exons 12 through 15 were amplified from RNA templates, as described [7]. Amplification products were then analyzed for sequence

alterations with the Big Dye Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, CA), using an ABI 3100 Genetic Analyzer. All unique sequence alterations were confirmed by sequencing of an independently amplified template. This approach may allow mutations to go undetected in cell lines without allelic loss, specifically deletions when analyzing DNA and truncating mutations that result in down-regulated transcripts when analyzing RNA. Allelic loss of the *BRCA1* gene was determined by PCR-based microsatellite analysis, as described [7]. *BRCA1* and *HPRT* transcripts were concurrently amplified from RNA templates, using the Qiagen Onestep reverse transcription-PCR (RT-PCR) kit and gene-specific primers. Primer sequences are available upon request.

Immunocytochemistry

Cell lines were cultured to optimal cell density in eight T162 flasks and medium was refreshed 24 hours before harvesting. Cells were harvested by scraping, washed twice with PBS, and fixed in PBS with 2% formalin for 12 to 72 hours. Cells were then washed once with PBS, resuspended in liquidized PBS with 2% agarose, and embedded in paraffin by routine diagnostic procedures. Paraffin sections (4 μm) on Starfrost microscope slides (Knittel Gläser, Braunschweig, Germany) were routinely deparaffinized and dehydrated, and epitopes were retrieved in Tris-EDTA (pH 9.0) for 30 minutes at 100°C in a microwave oven. Slides were blocked with 2% BSA in PBS for 30 minutes at room temperature and then incubated overnight at 4°C with antibodies diluted in Normal Antibody Diluent (Scytek Laboratories, Logan, UT). Anti-BRCA1 mouse monoclonal antibodies Ab-1 (Clone MS110; 1:100 or 1 μg/ml) and Ab-2 (Clone MS13; 1:320 or 0.6 μg/ml) were both purchased from Calbiochem (Darmstadt, Germany) and isotype-matched control monoclonal antibody X0931 (1:100 or 1 μg/ml) from Dako (Glostrup, Denmark). Slides were developed using the DakoCytomation Envision System horseradish peroxidase (3,3'-diaminobenzidine) kit, with omission of the antiperoxidase treatment. Slides were counterstained for 5 seconds with hematoxylin. Both anti-BRCA1 antibodies were titrated in two-step serial dilutions on *BRCA1* wild-type cell lines. At the presumed optimal antibody dilution, both antibodies showed distinct nuclear staining and Ab-1 also gave slight cytoplasmic staining. More diluted antibodies showed only nuclear staining for both antibodies and less diluted antibodies were aspecific (examples of wild-type and mutant cell lines at several antibody dilutions are provided as Supplementary Data).

RESULTS AND DISCUSSION

Sequencing of *BRCA1* revealed eighteen different alterations in the gene sequence among 41 human breast cancer cell lines (Tables 6.1 and 6.2). Alterations were presumed to be non-pathogenic polymorphisms when they were described as such in the Breast Cancer Information Core (BIC) mutation database (http://research.nhgri.nih.gov/bic/). Together, 11 *BRCA1*

Table 6.1 *BRCA1* mutation analysis of 41 human breast cancer cell lines.

Breast cancer cell line	*BRCA1* allelic loss	*BRCA1* gene variants*	*BRCA1* mutation status	*BRCA1* transcript expression†	
BT20	Loss	-	Wild-type	+	Unmethylated
BT474	Loss	6,7,9,10,11,13,14,15	Wild-type	++	Unmethylated
BT483	No loss	-	Wild-type	++	
BT549	Loss	-	Wild-type	++	Unmethylated
CAMA-1	No loss	3,5,6,7,9,10,11,13,14,15	Wild-type	++	Unmethylated
DU4475	No loss	6,7,9,10,11,13	Wild-type	+	Unmethylated
EVSA-T	Loss	-	Wild-type	+	Unmethylated
HCC1937	Loss	17	5382insC	++	
HS578T	Loss	3	Wild-type	++	Unmethylated
MCF-7	Loss	-	Wild-type	+/-	Unmethylated
MDA-MB-134VI	No loss	3	Wild-type	++	
MDA-MB-157	Loss	10	Wild-type	++	Unmethylated
MDA-MB-175VII	No loss	7,9,10,11,13,15	Wild-type	++	Unmethylated
MDA-MB-231	Loss	3,5	Wild-type	++	Unmethylated
MDA-MB-330	No loss	3,4,5	Wild-type	++	
MDA-MB-361	Loss	7,9,10,11,13,14,15	Wild-type	++	Unmethylated
MDA-MB-415	Loss	3	Wild-type	++	
MDA-MB-435S	Loss	-	Wild-type	++	Unmethylated
MDA-MB-436	Loss	7,9,10,11,13,14,15,18	5396+1G>A	++‡	
MDA-MB-453	Loss	-	Wild-type	++	Unmethylated
MDA-MB-468	Loss	10	Wild-type	++	Unmethylated
MPE600	No loss	3,5	Wild-type	++	Unmethylated
OCUB-F	Loss	2	Wild-type	++	
SK-BR-3	Loss	7,9,10,11,13,14,15	Wild-type	++	Unmethylated
SK-BR-5	Loss	3	Wild-type	++	
SK-BR-7	No loss	3	Wild-type	++	
SUM44PE	Loss	14	Wild-type	++	
SUM52PE	Loss	7,9,10,11,13,14,15	Wild-type	++	
SUM102PT	No loss	7,9,10,11,13,15	Wild-type	+/-	
SUM149PT	Loss	8,10	2288delT	+	
SUM159PT	No loss	3,5	Wild-type	++	
SUM185PE	Loss	3	Wild-type	+	
SUM190PT	Loss	-	Wild-type	++	
SUM225CWN	Loss	7,9,10,11,13,14,15	Wild-type	++	
SUM229PE	No loss	7,9,10,11,13,14,15	Wild-type	++	
SUM1315MO2	Loss	1,7,9,10,11,13,14,15	185delAG	+	
T47D	No loss	-	Wild-type	++	Unmethylated
UACC812	Loss	10,12,16	Wild-type	++	Unmethylated
UACC893	Loss	-	Wild-type	++	Unmethylated
ZR75-1	No loss	-	Wild-type	++	Unmethylated
ZR-75-30	Loss	7,9,10,11,13,14	Wild-type	+	

* Identified *BRCA1* sequence variants are detailed in Table 6.2.

† Transcript expression based on five experiments (see text): ++, Normal transcript levels; +, Low transcript levels; +/-, Barely detectable transcripts.

‡ Two transcript lengths that both differ from the wild-type sequence (see text).

Unmethylated, no hypermethylation at the *BRCA1* promoter region, as reported elsewhere [8]

Table 6.2 *BRCA1* sequence variants among 41 human breast cancer cell lines.

Variant	Nucleotide change*	Exon	Predicted protein effect†	Type of variant‡	No. in cell lines§	No. in BIC db§
1	185delAG	2	E23fsX17	Path	1	1642
2	233G>A	3	K38K	Poly	1	5
3	561–34C>T			Poly	11	18
4	788+3G>A			UV	1	1
5	1186A>G	11	Q356R	Poly	5	57
6	2196G>A	11	D693N	Poly	3	16
7	2201C>T	11	S694S	Poly	13	25
8	2288delT	11	N723fsX13	Mut	1	0
9	2430T>C	11	L771L	Poly	13	39
10	2731C>T	11	P871L	Poly	17	38
11	3232A>G	11	E1038G	Poly	13	48
12	3537A>G	11	S1140G	UV	1	27
13	3667A>G	11	K1183R	Poly	13	41
14	4427T>C	13	S1436S	Poly	11	49
15	4956A>G	16	S1613G	Poly	11	51
16	5106–20A>G			UV	1	17
17	5382insC	20	Q1756fsX74	Path	1	1676
18	5396+1G>A		E1731del28 I1760insX8	Path	1	46

* Numbering of nucleotide changes according *BRCA1* Genbank sequence U14680 and nomenclature according the BIC mutation database (http://research.nhgri.nih.gov/bic/).

† Frame shift and insertion mutations are indicated by the first changed codon and the number of newly encoded codons, including premature termination codon X. Predicted effect of variant 18 is based on sequence analysis from both transcript lengths (see text).

‡ Path, pathogenic variant according the BIC mutation database; Poly, polymorphism or non-pathogenic variant according the BIC; UV, unclassified variant according the BIC; Mut, variant not previously reported but presumed mutant as it generates a frame shift with a premature termination codon in the transcripts and because the cell line does not express nuclear BRCA1 proteins.

§ Number of cell lines with a particular *BRCA1* variant and number of citations of the variant in the BIC mutation database by July 2005.

polymorphisms were identified among 29 of the breast cancer cell lines. Three other *BRCA1* variants had been described as unclassified variant in the BIC mutation database and were each detected once in the cell lines (788+3G>A in MDA-MB-330, and S1140G and 5106-20A>G both in UACC812). Deleterious *BRCA1* mutations were identified in four breast cancer cell lines (Tables 6.1 and 6.2). The insertion of a cytosine residue at position 5382 of *BRCA1* in cell line HCC1937 had been reported [5]. In cell line MDA-MB-436, we identified the 5396+1G>A mutation in the splice donor site of exon 20 (Figure 6.1). Analysis of *BRCA1* transcripts from MDA-MB-436 identified two transcript lengths. Sequencing revealed that one transcript had skipped exon 20, predicting an in-frame deletion of 28 amino acids in the encoded BRCA1 proteins, whereas the other transcript had spliced at a cryptic splice site in intron 20 (5396+88/89), predicting an insertion of seven amino acids encoded by intron sequences followed by a termination codon. The patient from whom MDA-MB-436 was generated had been diagnosed with adenocarcinoma

of the breast at age 39 [9], an early onset that is suggestive for hereditary breast cancer. The original tumor was not available for analysis, but the 5396+1G>A mutation has been reported 46 times in the BIC mutation database and is classified as pathogenic. In cell line SUM149PT, we identified the deletion of a thymine residue at position 2288 of *BRCA1* (Figure 6.1). The 2288delT mutation predicts a shift in the *BRCA1* reading frame with an insertion of 12 new amino acids after codon 723 followed by a termination codon. The patient was diagnosed at age 35 years with inflammatory breast carcinoma and she had a single known second-degree relative with postmenopausal breast cancer. The 2288delT mutation was not present in the germline of the patient as we did not detect the mutation in a DNA sample from her blood. Of note, the identity of the donor was confirmed by analysis of 10 microsatellite markers from three chromosomes, with heterozygosity ratios of $\geq$ 0.80 for all markers ($P < 10^{-7}$). The original tumor was not available for analysis, but the 2288delT mutation was detected in all available passages of the SUM149PT cell line. It is thus unclear whether the mutation was somatically acquired during tumorigenesis in the patient or *in vitro* during establishment or propagation of the SUM149PT cell line. Importantly, we detected the 2288delT mutation in the earliest available passage P16 and cells were only distributed to other laboratories after this passage. We identified an AG dinucleotide deletion at position 185 of *BRCA1* in cell line SUM1315MO2, predicting a shift in the *BRCA1* reading frame with an insertion of 16 new amino acids after codon 22 followed by a termination codon (Figure 6.1). The patient was diagnosed with invasive ductal carcinoma of the breast but the age at diagnosis nor the cancer history of her family is known. The original tumor was not available for analysis, but the 185delAG mutation is a well-described pathogenic *BRCA1* mutation that is prevalent in the Ashkenazi Jewish population (http://research.nhgri.nih.gov/bic/).

Allelic loss of the *BRCA1* gene was determined by PCR-amplification of microsatellite markers *D17S1321*, *D17S932*, *D17S855*, *D17S1327* and *D17S1325*. These markers are located within a 0.7-Mb chromosomal region encompassing the *BRCA1* gene at 17q21. Analysis of the markers

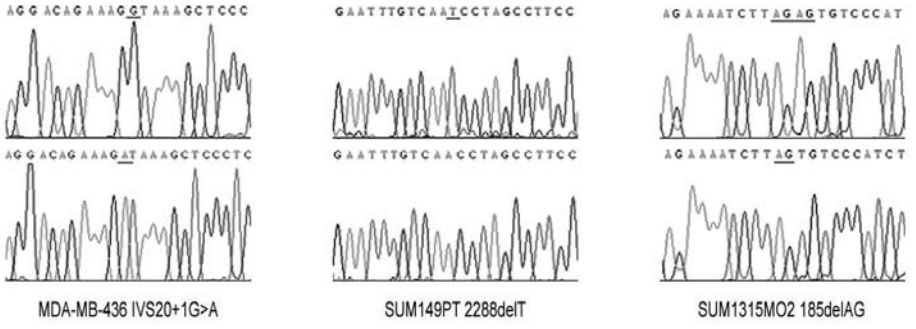


Figure 6.1 Identification of three new *BRCA1* mutant breast cancer cell lines by PCR amplification and direct sequencing. Top, electropherograms displaying the wild-type sequence. Bottom, electropherograms displaying the mutations. This figure is also available in color in the appendix.

on germline DNA's from 25 randomly selected Dutch individuals revealed heterozygosity ratios of 0.61, 0.76, 0.60, 0.55 and 0.88, respectively. Allelic loss of the *BRCA1* locus was presumed when each of the five markers had a single allele size, resulting in a reliability of $P = 0.002$ [6]. None of the 25 control DNA's had a homozygous allele pattern at the *BRCA1* locus, thus validating this statistical approach. Of the 41 breast cancer cell lines, 28 (68%) had allelic loss of the *BRCA1* locus, including the four *BRCA1* mutants (Table 6.1). Similar allelic loss frequencies have been reported for primary breast cancer specimens [10-12]. It is important to note that several regions at 17q are frequently amplified in human breast cancers. Allelic losses at 17q are therefore often underestimated, as karyotype-based methods do not detect loss when the retained allele is amplified or reduplicated [6, 13]. Indeed, we did not identify loss of the *BRCA1* locus in three *BRCA1* mutant cell lines that we analyzed by array CGH (data not shown), whereas our microsatellite analysis revealed allelic loss in all of them. Conclusively, all *BRCA1* mutants were homozygous in the sequence analysis (Figure 6.1).

BRCA1 transcript expression was analyzed by semi-quantitative RT-PCR using five overlapping PCR fragments (Table 6.1). Cell lines HCC1937 and MDA-MB-436 had *BRCA1* transcript expression levels that were comparable to those of most other cell lines, SUM149PT had variable but always lower expression levels and SUM1315MO2 had consistently low expression of *BRCA1* transcripts. In contrast with wild-type cell lines, nuclear BRCA1 protein expression was not detectable in any of the four mutant cell lines, as determined by immunocytochemistry on paraffin-embedded cells using anti-BRCA1 monoclonal antibodies Ab-1/MS110 and Ab-2/MS13 (Figure 6.2 and Supplementary Data).

We thus describe four cell lines with a *BRCA1* mutation from among 41 human breast cancer cell lines, three of which had not previously been reported. All *BRCA1* mutants had lost the other *BRCA1* allele, in accordance with the tumor suppressor function of BRCA1. Three mutations generated a premature termination codon in the *BRCA1* transcript, whereas the fourth mutation resulted in two transcripts of which one had an in-frame deletion and the other generated a premature termination codon. Three of the *BRCA1* mutations have been classified as pathogenic mutations and none of the *BRCA1* mutant cell lines expressed nuclear BRCA1 proteins. In an ongoing effort to characterize our panel of breast cancer cell lines, we identified biallelic mutations of the *p53* tumor suppressor gene in each of the four *BRCA1* mutants [14], consistent with the notion that *BRCA1* mutant tumors frequently harbor *p53* mutations (reviewed in [15]). Pending further mutational data, these *BRCA1* mutant breast cancer cell lines already are a valuable asset in pinpointing the BRCA1 functions that are critical in the suppression of breast tumorigenesis.

ACKNOWLEDGEMENTS

We thank Hans Stoop and Mieke Timmermans for technical advise regarding BRCA1 immunocytochemistry. Funding was provided by the Dutch Cancer Society.

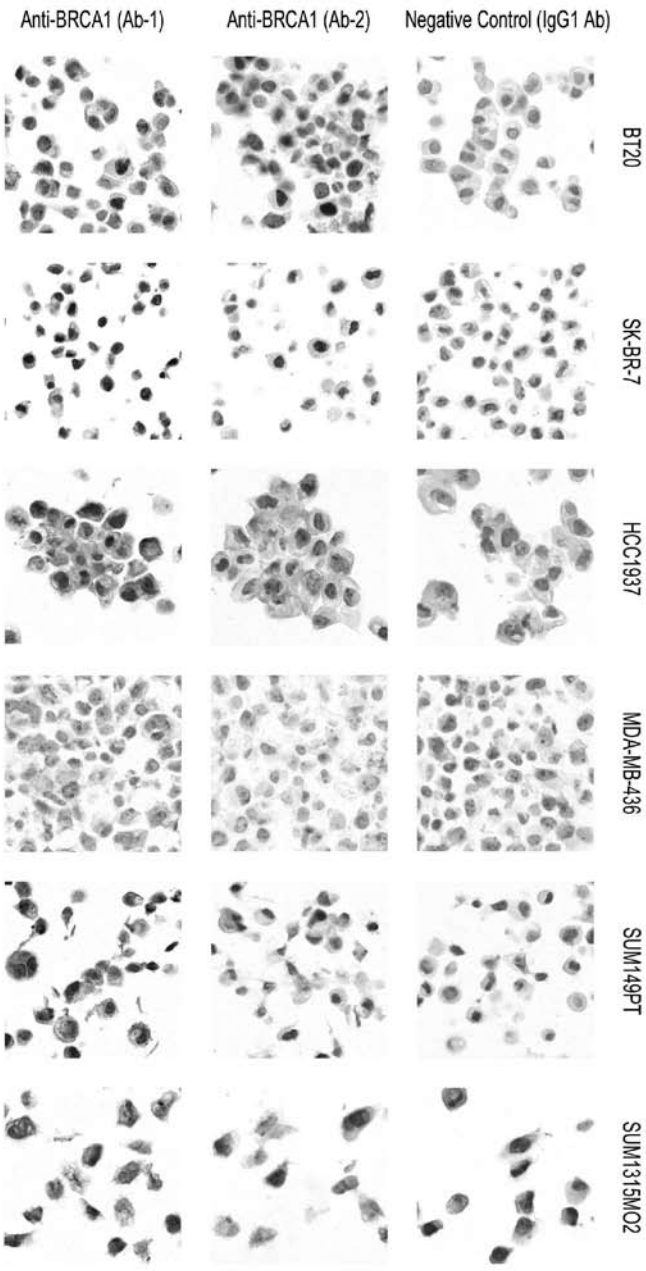


Figure 6.2 BRCA1 immunocytochemistry in *BRCA1* mutant and wild-type breast cancer cell lines. In contrast with the two wild-type cell lines (BT20 and SK-BR-7), none of the four *BRCA1* mutants had nuclear BRCA1 staining with either of the two anti-BRCA1 monoclonal antibodies Ab-1 and Ab-2. There is some cytoplasmic staining of unclear significance in all samples with Ab-1, which is not observed with more diluted Ab-1 antibodies nor with Ab-2 (see also Supplementary Data). The negative control antibody is an IgG1 isotype-matched antibody. Magnification 40X. This figure is also available in color in the appendix.

REFERENCES

1. Scully R, Livingston DM. In search of the tumour-suppressor functions of BRCA1 and BRCA2. Nature 2000;408(6811):429-32.
2. Venkitaraman AR. Cancer susceptibility and the functions of BRCA1 and BRCA2. Cell 2002;108(2):171-82.
3. Tutt A, Ashworth A. The relationship between the roles of BRCA genes in DNA repair and cancer predisposition. Trends in molecular medicine 2002;8(12):571-6.
4. Starita LM, Parvin JD. The multiple nuclear functions of BRCA1: transcription, ubiquitination and DNA repair. Current opinion in cell biology 2003;15(3):345-50.
5. Tomlinson GE, Chen TT, Stastny VA, et al. Characterization of a breast cancer cell line derived from a germ-line BRCA1 mutation carrier. Cancer research 1998;58(15):3237-42.
6. Harkes IC, Elstrodt F, Dinjens WN, et al. Allelotype of 28 human breast cancer cell lines and xenografts. British journal of cancer 2003;89(12):2289-92.
7. van de Wetering M, Barker N, Harkes IC, et al. Mutant E-cadherin breast cancer cells do not display constitutive Wnt signaling. Cancer research 2001;61(1):278-84.
8. Esteller M, Silva JM, Dominguez G, et al. Promoter hypermethylation and BRCA1 inactivation in sporadic breast and ovarian tumors. Journal of the National Cancer Institute 2000;92(7):564-9.
9. Brinkley BR, Beall PT, Wible LJ, Mace ML, Turner DS, Cailleau RM. Variations in cell form and cytoskeleton in human breast carcinoma cells in vitro. Cancer research 1980;40(9):3118-29.
10. Fujii H, Szumel R, Marsh C, Zhou W, Gabrielson E. Genetic progression, histological grade, and allelic loss in ductal carcinoma in situ of the breast. Cancer research 1996;56(22):5260-5.
11. Fukino K, Iido A, Teramoto A, et al. Frequent allelic loss at the TOC locus on 17q25.1 in primary breast cancers. Genes, chromosomes & cancer 1999;24(4):345-50.
12. Shen CY, Yu JC, Lo YL, et al. Genome-wide search for loss of heterozygosity using laser capture microdissected tissue of breast carcinoma: an implication for mutator phenotype and breast cancer pathogenesis. Cancer research 2000;60(14):3884-92.
13. Lengauer C, Kinzler KW, Vogelstein B. Genetic instabilities in human cancers. Nature 1998;396(6712):643-9.
14. Wasielewski M, Elstrodt F, Klijn JG, Berns EM, Schutte M. Thirteen new p53 gene mutants identified among 41 human breast cancer cell lines. Breast cancer research and treatment 2006;99(1):97-101.
15. Gasco M, Yulug IG, Crook T. TP53 mutations in familial breast cancer: functional aspects. Human mutation 2003;21(3):301-6.

Chapter 7

Distinct Gene Mutation Profiles Among Luminal and Basal-Type Breast Cancer Cell Lines

Antoinette Hollestelle*
Jord H.A. Nagel*
Marcel Smid
Suzanne Lam
Fons Elstrodt
Marijke Wasielewski
Ser Sue Ng
Pim J. French
Justine K. Peeters
Marieke J. Rozendaal
Muhammad Riaz
Ellen C. Zwarthoff
Amina Teunisse
Peter J. van der Spek
Jan G.M. Klijn
Winand N.M. Dinjens
Stephen P. Ethier
Hans Clevers
Aart G. Jochemsen
Michael A. den Bakker
John A. Foekens
John W.M. Martens
Mieke Schutte

Submitted for publication

ABSTRACT

Our knowledge of the genetics of human breast cancers is still fairly limited. We here report an extensive molecular characterization of a collection of 41 human breast cancer cell lines. Protein and gene expression analyses indicated that the collection of breast cancer cell lines has retained most, if not all, molecular characteristics that are typical for clinical breast cancers. Gene mutation analyses identified 146 oncogenic mutations among 27 well-known cancer genes, amounting to an average of 3.6 mutations per cell line. Mutations in genes from the p53, RB and PI3K tumor suppressor pathways were widespread among all breast cancer cell lines. Most important, we have identified two gene mutation profiles that specifically associated with luminal-type and basal-type breast cancer cell lines. The luminal mutation profile involved *E-cadherin* and *MAP2K4* gene mutations and amplifications of *Cyclin D1*, *ERBB2* and *HDM2*, whereas the basal mutation profile involved *BRCA1*, *RB1*, *RAS* and *BRAF* gene mutations and deletions of *p16* and *p14ARF*. These subtype-specific gene mutation profiles constitute a genetic basis for the heterogeneity observed among human breast cancers, providing clues for their underlying biology and providing guidance for targeted pharmacogenetic intervention in breast cancer patients.

INTRODUCTION

Breast cancer is a heterogeneous disease with respect to prognosis and treatment response of patients. Clinical outcome of cancer patients is to a large extent driven by the biology of their tumors. Accurate classification of breast cancers therefore has been a major objective in breast cancer research. Protein expression of the estrogen receptor (ER), progesterone receptor (PR) and/or ERBB2 (also known as HER2 or NEU) has for long guided breast cancer classification. Over recent years, this classification has been refined by additional histological criteria, in particular cytokeratin protein expression [1-4]. Luminal phenotype breast cancers express luminal cytokeratins only, representing some 60-80% of primary breast cancers. Basal phenotype breast cancers express both luminal and basal cytokeratins and represent 20-40% of tumors. Two minor subtypes of breast cancer express only basal cytokeratins or are negative for both luminal and basal markers, each representing less than 1% of tumors (stem cell and null phenotypes, respectively). Luminal breast cancers are more often positive for ER and/or PR proteins or have overexpression of ERBB2. The so-called triple-negative phenotype of tumors that express neither ER, PR nor ERBB2 is observed primarily among basal breast cancers. Basal breast cancers also more often have EGFR expression. Histological classification of breast cancers has clinical relevance in that patients with basal breast cancers more often have a worse prognosis — often related to their inherent resistance to chemotherapy — and patients with luminal breast cancers tend to fare better [5, 6]. There is however no clear consensus on the definition (and nomenclature) of basal-type breast cancer (reviewed in [7]). Some have argued that expression of any basal cytokeratin (CK) defines basal breast cancers [8], whereas others suggested expression of basal CK14 as being diagnostic [9]. Perou and colleagues have defined a 4-protein signature that defined four groups of breast cancers: ERBB2 overexpressing (ERBB2+), luminal (ERBB2- and ER+), basal-like (ERBB2/ER- and CK5+ and/or EGFR+), and the negative group that lacks expression of all four proteins [10]. A major discrepancy among these and other definitions lies in breast cancers that express basal cytokeratins as well as ER, which may be as much as one-third of all basal cytokeratin expressing breast cancers [8]. A consensus on the definition of basal-type breast cancers may only be reached once such definition proves clinically relevant.

Recent research also has focused on gene expression profiling to classify breast cancers. Five subtypes of breast cancer were defined based on the intrinsic gene set [11-13]. ER expression was a major classifier, including the luminal A and luminal B subtypes of ER-positive breast cancers. ER-negative breast cancers included the basal-like and normal-like subtypes and the ERBB2+ subtype that was characterized by overexpression of this membrane protein. The intrinsic subtypes of breast cancer have clinical relevance [11, 12, 14, 15], as have gene expression signatures that had been defined based on clinical outcome of breast cancer patients [16-19]. Importantly, the success of gene expression profiling in breast cancer classification implies a relatively restricted set of recurrent expression patterns among breast cancer subtypes.

Cancer is a genetic disease. Mutations in a cancer gene often induce constitutive changes in the expression of its downstream pathway members or targets. The subtype-specific gene expression patterns among breast cancers were therefore likely to reflect, at least in part, the accumulation of mutations in subtype-specific cancer genes. Here, we have evaluated this concept by extensive molecular characterization of 41 human breast cancer cell lines at protein, transcript and gene level. We have performed mutation screens for 27 well-known cancer genes and identified a total of 146 oncogenic mutations. Two gene mutation profiles were apparent that associated with each of the two major breast cancer subtypes: a luminal mutation profile and a basal mutation profile.

MATERIALS AND METHODS

Cell lines

The 41 human breast cancer cell lines used in this study are listed in Figure 1. Cell lines EVSA-T, MPE600, and SK-BR-5/7 were kind gifts of Dr. N. de Vleesschouwer (Institut Jules Bordet, Brussels, Belgium), Dr. H.S. Smith (California Pacific Medical Center, San Francisco, CA) and Dr. E. Stockert (Sloan-Kettering Institute for Cancer Research, New York, NY), respectively. The SUM cell lines were generated in the Ethier laboratory (available at http://www.asterand.com). Cell lines OCUB-F and -M had been generated from the same tumor, and were obtained from Riken Gene Bank (Tsukuba, Japan). The other cell lines were obtained from ATCC (Manassas, VA). All cell lines are unique and monoclonal as shown by extensive analysis of nearly 150 polymorphic microsatellite markers [20].

Melanoma cell line M14 was obtained from its provider at Micromet Inc. (Munich, Germany; [21]) at passage-10 (M14 is not commercially available). Breast cancer cell line MDA-MB-435s was obtained from ATCC, at passage-239 (MDA-MB-435s was not available from its provider). M14 and MDA-MB-435s genomic DNA was genotyped by using eight polymorphic microsatellite markers, revealing genetic identity.

Expression analyses

Protein expression was determined by western blotting or by immunohistochemistry on a tissue microarray of paraffin-embedded cell line samples, transcript expression was determined by quantitative real-time PCR (qPCR), and gene expression profiling was performed by using Affymetrix U133A microarrays, all described in detail in chapter 3. Antibodies for western blotting of RB and p53 pathway members were: CDK4 (Santa Cruz Biotechnology (Santa Cruz, CA), clone C-22); Cyclin D1 (Upstate Biotechnology (Lake Placid, NY), 06-137); HDM2 (a mixture of 4B2, kind gift of Dr. Arnold Levine, and clone SMP14 from Santa Cruz Biotechnology); p16 (JC8, kind gift of Dr. Gordon Peters); p21 (Upstate Biotechnology, clone CP74); p53 (Santa Cruz Biotechnology, clone DO-1); and RB1 (BD Pharmingen (Franklin Lakes, NJ), clone

G3-245). Antibodies for immunohistochemistry were: Calponin (DAKO (Glostrup, Denmark), clone CALP), CDH1/E-cadherin (DAKO, clone NCH-38), CDH2/N-cadherin (DAKO, clone 6G11), CHEK2 (Sigma (Steinheim, Germany), clone DCS-270), CK5 (Novocastra (Newcastle upon Tyne, UK), clone XM26), CK8-18 (Biogenex (San Ramon, CA), clone NCL5D3), CK14 (Neomarkers (Fremont, CA), clone LL002), CK19 (DAKO, clone RCK108), EGFR (DAKO, EGFR PharmDx™ Kit), ER (DAKO, clone 1D5), ERBB2 (DAKO, HercepTest™ kit), Ki-67 (Zymed (Orange, CA), clone 7B11), p63 (DAKO, clone 4A4), PR (DAKO, clone PgR 636), S100 (DAKO, Z0311), SMA (DAKO, clone 1A4) and Vimentin (DAKO, clone V9).

Classification by intrinsic gene expression

The Stanford intrinsic gene set for clinical breast cancers [11-13] was translated into an Affymetrix intrinsic gene set, including 451 probe sets from the Stanford list of 496 genes [15]. Breast cancer cell lines were classified for their intrinsic subtype by this Affymetrix intrinsic gene set, using average distance linkage hierarchical clustering with non-centered correlation as distance metric [22]. Details on gene expression profiling methodology are described in chapter 3.

Gene mutation analyses

Gene sequence alterations typically were identified by direct sequencing of PCR-amplified genomic templates that included exon sequences and intron/exon boundaries, using an ABI 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA). For *RB1*, exons 3 through 27 were analyzed by sequencing of transcript templates, amplified by RT-PCR from RNA of emetine-treated cell lines. The complete coding sequences have been analyzed for *BRCA1* [23], *CHEK2* [24], *E-cadherin/CDH1* [25], *MAP2K4* [26], *MSH6*, *p14ARF*, *p16*, *p53* [27], *PIK3CA* [28], *PTEN* [28] and *RB1*, whereas only known mutational hotspots have been analyzed for *AKT1* (codon 17; [29]), *APC* (mutation E1577X; [30]), *BRAF* (exons 7, 11 and 15; [28]), *CDK4* (p16-binding domain in exon 2; [31]), *FGFR3* (codons 248, 249, 372, 375, 393 and 652; [32]), and codons 12, 13 and 61 for *HRAS*, *KRAS* and *NRAS* [28]. All oncogenic mutations have been confirmed by sequencing of an independently amplified DNA and/or RNA template.

Amplifications *of AKT1, AKT2, AKT3, CDK4*, *c-MYC*, *Cyclin D1*, *EGFR*, *ERBB2* and *HDM2* were defined by genomic gain of the gene locus with concurrent transcript and/or protein overexpression. Genomic gains were determined by screening of publicly available CGH and SNP data ([33] and www.sanger.ac.uk), requiring gain of at least four allele copies. Genomic gain for the *ERBB2* and *EGFR* loci were determined by chromogenic in situ hybridization (CISH) using the SPoT-Light® HER2 CISH™ kit and the SPoT-Light® EGFR probe with the CISH™ Tissue Pretreatment and CISH™ Polymer Detection Kits, respectively, all according the manufacturers recommendations (Zymed).

Promoter hypermethylation was determined for *BRCA1* [34], *E-cadherin*, *p16*, *p14ARF* and *RB1*, by azacytidine assays and/or by methylation specific PCR as described in chapter 3. Allelic loss was determined by microsatellite analysis, where a homozygous allele pattern for several

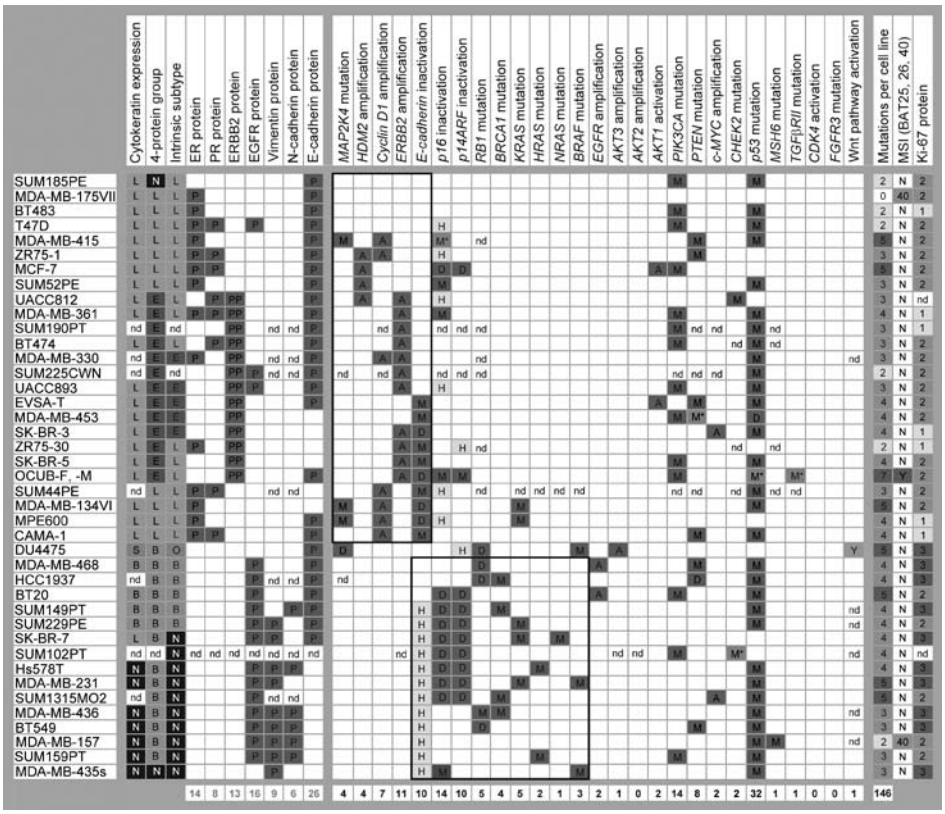


Figure 7.1 Molecular characterization of 41 human breast cancer cell lines. (Left panel) Classification of the cell lines by expression analyses based on cytokeratin proteins, the 4-protein signature of ERBB2, ER, CK5 and EGFR, the intrinsic gene set, and several proteins relevant for breast cancer. Classification by cytokeratin expression: luminal (L), CK8/18+ and/or CK19+, CK5-; basal (B), CK8/18+ and/or CK19+, CK5+; null (N), CK8/18 low, CK19-, CK5-; stem cell (S), CK8/18 low, CK19-, CK5+. Classification by the 4-protein signature: ERBB2 overexpression (E); luminal (L), ERBB2-, ER+; basal (B), ERBB2/ER-, CK5+ and/or EGFR+; negative (N), ERBB2/ER/CK5/EGFR-. Classification by intrinsic gene expression (see also Figure 7.2): ERBB2+ (E); luminal (L); basal-like (B); normal-like (N); other subtype (O). For individual proteins, expression (P) and overexpression (PP) is indicated in blue and absence of expression in white. The protein expression profiles confirmed the remarkably concordant classification by histological criteria or by intrinsic gene expression, and suggested that basal/basal-like and null/negative/normal-like cell lines represent two related subtypes of basal-type breast cancers. (Middle panel) Cancer gene mutation analysis of the cell lines. Genes are indicated at the top and the number of oncogenic mutations identified in each gene at the bottom. Oncogenic mutations (M), sizeable deletions (D) and amplifications (A) are in red; heterozygous oncogenic mutations in tumor suppressor genes (M*) are in pink; and wild-type genes are in white. Promoter hypermethylation (H) is in yellow and constitutive Wnt pathway activation (Y) is in green. nd, not determined. The observed dichotomy among the breast cancer cell lines by protein and gene expression analyses was further supported by the two distinct gene mutation profiles among luminal-type and basal-type breast cancer cell lines (areas with black borders). (Right panel) Number of oncogenic mutations identified in each cell line, microsatellite instability (MSI) and Ki-67 proliferation analysis. MSI: no (N); yes (Y), MSI with BAT 25, 26 and 40; (40), MSI with BAT 40 only. Ki-67 immunohistochemistry: 1, less than 33% of cells positive; 2, 33-66% of cells positive; 3, more than 66% of cells positive. There was a correlation of proliferation rate with breast cancer type (χ^2 $P<0.001$), which was not associated with the number of mutations identified per cell line. This figure is also available in color in the appendix.

consecutive markers at P<0.05 was interpreted as allelic loss [20]. Wnt pathway activation was determined by reporter construct assays, using the TOP/FOP constructs [25].

RESULTS

Expression analysis of breast cancer cell lines

We have evaluated 41 human breast cancer cell lines for expression of a variety of proteins that are known to be relevant for breast cancer, using immunohistochemistry (Figure 7.1; data are detailed in the Supplementary Data). ER and PR protein expression was observed in fourteen and eight breast cancer cell lines, respectively, and thirteen cell lines had overexpression of ERBB2. CK protein expression patterns classified the breast cancer cell lines as those of the histological luminal phenotype (luminal CK8/18+ and/or CK19+ and basal CK5-; n=22), basal phenotype (CK8/18+ and/or CK19+ and CK5+; n=4), null phenotype (CK8/18 low, CK19- and CK5-; n=7) and stem cell phenotype (CK8/18 low, CK19- and CK5+; n=1). None of the cell lines expressed basal CK14. All ER, PR and/or ERBB2 (over)expressing cell lines were of the luminal phenotype according their CK protein expression. Sixteen cell lines had the so-called triple-negative phenotype (ER/PR/ERBB2-negative), most of which were of the basal or null phenotypes. Expression of EGFR and particularly expression of the epithelial-mesenchymal transition (EMT) markers Vimentin and N-cadherin was essentially restricted to basal and null phenotype cell lines. Interestingly, null phenotype cell lines typically had lost E-cadherin protein expression, albeit that loss of E-cadherin protein expression was also observed in a subset of *E-cadherin* mutant luminal cell lines. No obvious expression patterns were revealed for several myoepithelial markers, including Smooth Muscle Actin (SMA), p63, S100 and Calponin. As a whole, the protein expression profiles of the breast cancer cell lines were highly consistent with those observed for clinical breast cancers, suggesting two major subtypes of breast cancers.

The dichotomy among the breast cancer cell lines was substantiated by gene expression profiling, using Affymetrix U133A microarrays (Figure 7.2 A and chapter 3). Two main clusters of cell lines were identified when samples were positioned according to their overall similarity in gene expression, whether the correlation was calculated from a log2GM <-2 and >2 probe subset or log2GM <-3 and >3 probe subset (data shown for log2GM <-2 and >2 in Figure 7.2 A). The major cluster contained all cell lines that (over)expressed ER, PR and/or ERBB2 and all cell lines were luminal as defined by CK protein expression (Figure 7.1 and Supplementary Data). The minor cluster contained triple-negative cell lines only, with a subdivision in basal and null phenotype cell lines. A similar clustering of breast cancer cell lines was obtained by hierarchical clustering based on the intrinsic gene set (Figure 7.2 B; [11-13, 15]). Luminal phenotype cell lines by CK expression classified either as luminal or ERBB2+ by intrinsic gene expression. Similarly, basal and null phenotype cell lines were basal-like and normal-like, respectively, by intrinsic gene expression. Cell line DU4475 could not be assigned to any of the intrinsic breast cancer

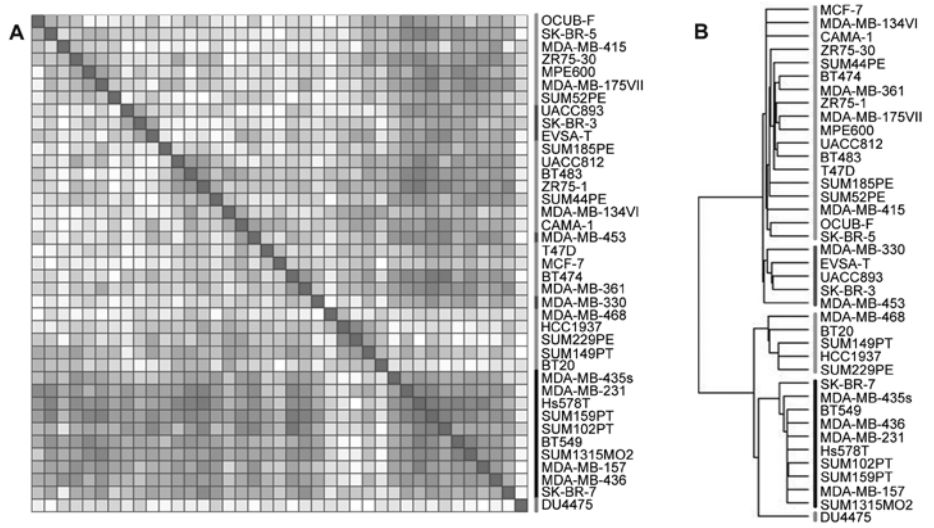


Figure 7.2 Global gene expression and intrinsic subtypes of breast cancer cell lines. (A), Pearson correlation plot of 39 cell lines based on the subset of 5,490 log2GM <-2 and >2 differentially expressed probe sets by Affymetrix U133A microarrays. The Pearson correlation coefficient algorithm positions samples according their overall similarity in gene expression, where red indicates high overall similarity and blue indicates low similarity. Two main clusters of 23 and 15 cell lines were apparent, whereas the DU4475 cell line did not belong to either cluster. (B), Dendrogram of hierarchical clustering of 39 cell lines based on the intrinsic gene set and U133A microarray data. The luminal and ERBB2+ intrinsic subtypes coincided with the major cluster of cell lines in Figure 1A, and the basal-like and normal-like subtypes coincided with the minor cluster of cell lines. Cell line DU4475 did not classify for any of the intrinsic breast cancer subtypes. Color coding of intrinsic subtypes (see also Figure 7.1): green, luminal; brown, ERBB2+; orange, basal-like; black, normal-like; pink, not of any subtype. This figure is also available in color in the appendix.

subtypes. DU4475 was triple-negative and, interestingly, it was the only cell line that classified for the rare stem cell phenotype by CK expression. Thus, classification of the breast cancer cell lines by protein expression or by gene expression was very similar, albeit that the intrinsic ERBB2+ subtype had not been defined as a separate phenotype by CK protein expression (Figure 7.1). In contrast, classification according the 4-protein signature of ERBB2, ER, CK5 and EGFR [10] identified all thirteen cell lines with ERBB2 overexpression. Yet, the 4-protein signature was not able to distinguish basal or basal-like cell lines from null or normal-like cell lines (Figure 7.1). It appears that the correct classification of breast cancers has not yet been defined. The highly concordant results of the various classifications that we have evaluated, however, suggest that it is reasonable to designate the two major subtypes of breast cancer cell lines as "luminal-type" and "basal-type" breast cancers. Also, classification of the cell lines by intrinsic gene expression — that had been defined on clinical breast cancer samples — implied that breast cancer cell lines are a relevant model to study human breast cancer.

Mutation analysis of breast cancer cell lines

In an ongoing effort to characterize the breast cancer cell lines, we have performed mutation analyses of 27 well-known cancer genes (results of 16 genes are first reported here and the results of the remaining genes were reported in references [23-28, 34] and Ng and Clevers, manuscript in preparation; all mutations are detailed in the Supplementary Data). Mutations typically were identified by PCR amplification and sequencing of the complete coding sequences of the genes or their known mutational hotspots. Oncogenic mutations included those mutations that predicted a premature termination or a significant deletion/insertion of the encoded proteins, splice site mutations and missense mutations that previously had been identified in clinical cancers or had been associated with hereditary cancer syndromes. For oncogenes that reportedly are amplified in clinical cancers, we considered chromosomal gains with concurrent transcript and/or protein overexpression also as oncogenic mutations or amplifications. Together, we have identified 146 oncogenic mutations among 24 cancer genes (Figure 7.1). Oncogenic mutations in tumor suppressor genes included 30 sizeable deletions, 26 nonsense mutations and small deletions/insertions causing premature terminations, 7 splice site mutations and 29 missense mutations, totaling to 92 mutations. In concordance with the recessive nature of tumor suppressor gene mutations, 87 (95%) of these mutations were bi-allelic. Oncogenic mutations in oncogenes included 29 amplifications and 25 missense mutations, totaling to 54 mutations. Three tumor suppressor genes had transcriptional silencing by promoter hypermethylation, together in 20 cell lines. The 146 oncogenic mutations were identified among 40 of the 41 breast cancer cell lines, resulting in an average of 3.6 mutations per cell line and ranging from no to seven mutations each identified in a single cell line (Figure 7.1). Our results thus confirm that breast cancers, like other carcinomas, arise through the accumulation of mutations in several cancer genes [35].

For most cancer genes, we also determined transcript and protein expression patterns, mostly by qPCR and western blotting. As anticipated, transcript expression of mutant genes with premature terminations typically was low or absent. Missense mutations resulted in relatively normal transcript expression levels albeit that their expression range often was wider than for wild-type genes. Protein expression patterns mostly followed transcript expression patterns and/or were in concordance with those reported for clinical cancers. For example, *p53* mutants with premature terminations typically had low transcript and protein expression whereas *p53* missense mutants had relatively normal transcript expression but high protein expression (Table 7.1). For some mutant genes, we observed expression modulation of other proteins in the same tumor suppressor pathway. In the RB pathway, for example, most *RB1* mutants had high p16 protein expression and very low Cyclin D1 protein expression and *Cyclin D1* mutants had low p16 protein expression (Table 7.2 and Figure 7.3). The transcript and protein expression patterns of mutant cancer genes thus further confirmed the functional significance of the identified oncogenic gene mutations.

Table 7.1 Mutation analysis of p53 pathway genes.

Breast cancer cell lines	*p14ARF* gene sequence	*p14ARF* transcript expression	*p53* gene sequence	*p53* transcript expression	p53 protein expression	*CHEK2* gene sequence
ZR-75-30	methylated	+/-	wild-type	+	+	nd
DU4475	methylated	+	wild-type	+	+	wild-type
MCF-7	deleted	-	wild-type	+	+	wild-type
SK-BR-7	deleted	-	wild-type	+	+	wild-type
SUM102PT	deleted	-	wild-type	+	nd	mutant*
BT20	deleted	-	mutant AA	+	++	wild-type
SUM149PT	deleted	-	mutant AA	+	++	wild-type
SUM229PE	deleted	-	mutant AA	+	++	wild-type
SUM1315M02	deleted	-	mutant AA	+	++	wild-type
Hs578T	deleted	-	mutant AA	++	++	wild-type
MDA-MB-231	deleted	-	mutant AA	++	++	wild-type
OCUB-F	mutant	+	mutant AA*	+	nd	nd
MDA-MB-361	mutant?	+/-	mutant TR	+	+/-	wild-type
SUM159PT	wild-type	-	mutant AA	+	++	wild-type
BT483	wild-type	+	mutant AA	+	nd	wild-type
MDA-MB-330	wild-type	+	mutant AA	+	++	wild-type
MDA-MB-435S	wild-type	+	mutant AA	+	++	wild-type
MDA-MB-468	wild-type	+	mutant AA	+	++	wild-type
T47D	wild-type	+	mutant AA	+	++	wild-type
SK-BR-5	wild-type	+	mutant AA	+	++	wild-type
BT474	wild-type	+	mutant AA	++	++	nd
EVSA-T	wild-type	+	mutant AA	++	++	wild-type
SK-BR-3	wild-type	+	mutant AA	++	++	wild-type
MDA-MB-415	wild-type	+	mutant AA	++	+	wild-type
BT549	wild-type	++	mutant AA	++	++	wild-type
CAMA-1	wild-type	++	mutant AA	+	++	wild-type
MDA-MB-134VI	wild-type	++	mutant AA	+	+/-	wild-type
SUM225CWN	nd	nd	mutant AA	nd	nd	wild-type
SUM190PT	nd	nd	mutant TR	nd	+	wild-type
HCC1937	wild-type	++	mutant TR	+/-	+	wild-type
MDA-MB-157	wild-type	++	mutant TR	+/-	+/-	wild-type
MDA-MB-436	wild-type	+	mutant TR	+/-	+/-	wild-type
SUM44PE	wild-type	+	mutant TR	+/-	+/-	nd
SUM185PE	wild-type	+	mutant TR	+/-	+/-	wild-type
UACC893	wild-type	+	mutant TR	+/-	+	wild-type
SUM52PE	wild-type	+	mutant TR	+/-	+/-	wild-type
MDA-MB-453	wild-type	+	deleted	nd	+	wild-type
UACC812	wild-type	+/-	wild-type	+	+	mutant
ZR75-1	wild-type	+/-	wild-type	+	+	wild-type
MDA-MB-175VII	wild-type	+	wild-type	+	+	wild-type
MPE600	wild-type	+	wild-type	+	+	wild-type

Breast cancer cell lines are organized by their gene mutation status. Protein and transcript expression (determined by Western blotting and qPCR): -, no detectable expression; +/-, barely detectable; +, expression at apparently normal level; ++, overexpression. Presumed oncogenic amplifications are defined by chromosomal gain of the gene locus with concurrent protein and/or transcript overexpression; >>

CHEK2 protein expression	*c-MYC* gene locus	*c-MYC* transcript expression	*HDM2* gene locus	HDM2 protein expression	*p21* transcript expression	p21 protein expression
+	gain	+	balanced	+/-	+	+
+	balanced	+	balanced	++	+	+
+	gain	+	gain^{amp}	++	+	++
+	nd	+	nd	+	++	+
-	balanced	+	nd	nd	++	nd
+	balanced	+	balanced	+/-	+/-	-
+	balanced	+	balanced	+/-	+	+/-
+	balanced	+	nd	+/-	++	+
+	gain^{amp}	++	balanced	+/-	+	+/-
+	gain	+	balanced	+	+	+/-
+	balanced	+	balanced	+	+	+/-
+	gain	+	balanced	nd	+	nd
+	balanced	+	balanced	+/-	+	+/-
+	gain	+	balanced	+/-	+	+/-
+	gain	+	balanced	nd	+	nd
+	nd	+/-	nd	+/-	+	+/-
+	balanced	+	balanced	+/-	+/-	-
+	balanced	+	balanced	+/-	+	+/-
+	balanced	+	balanced	+/-	+	+
+	nd	+/-	nd	+	+	+/-
+	balanced	+	balanced	+	+	+/-
+	gain	+	balanced	+	+/-	-
+	gain^{amp}	++	balanced	+	+	+/-
+	balanced	+	balanced	+/-	+	+/-
+	balanced	+	balanced	+/-	+/-	-
+	gain	+	balanced	+	+	-
+	balanced	+	balanced	+	+	++
+	nd	nd	balanced	nd	nd	nd
+	gain	nd	balanced	+/-	nd	+
+	balanced	+	balanced	+/-	+/-	+/-
+	gain	+	balanced	+/-	+	+/-
+	balanced	+	balanced	+/-	+/-	-
+	nd	+	balanced	+/-	++	+
+	balanced	+/-	balanced	+/-	+	+
+	gain	+	balanced	+/-	+	+/-
+	gain	+	gain^{amp}	++	+/-	+/-
+	balanced	+	balanced	+/-	+/-	+
-	balanced	+	gain^{amp}	++	+	+
+	balanced	+	gain^{amp}	++	+	+
+	gain	+/-	balanced	++	++	++
+	balanced	+	balanced	+	++	+

mutant AA, missense mutant; mutant TR, truncating mutant; *, heterozygous oncogenic mutation in a tumor suppressor gene; ?, missense mutation of unclear significance; amp, amplified; nd, not determined. Mutational inactivation of the p53 pathway was identified in 37 of 41 (90%) breast cancer cell lines.

Table 7.2 Mutation analysis of RB pathway genes.

Breast cancer cell lines	Allelic loss at 13q	*RB1* gene sequence	*RB1* transcript expression	RB1 protein expression	Allelic loss at 9p	*p16* gene sequence	*p16* transcript expression
DU4475	yes	deleted	-	-	no	wild-type	+
MDA-MB-468	yes	deleted	-	-	no	wild-type	++
BT549	yes	deleted	-	-	no	wild-type	++
MDA-MB-436	yes	mutant	+/-	-	yes	wild-type	++
HCC1937	yes	deleted	+	-	nd	wild-type	++
BT20	yes	mutant?	+	+	yes	deleted	-
MCF-7	yes	wild-type	+	+	yes	deleted	-
SUM229PE	yes	wild-type	+	+	yes	deleted	-
MDA-MB-231	yes	wild-type	+	++	yes	deleted	-
Hs578T	yes	wild-type	+/-	+	yes	deleted	-
SUM149PT	yes	wild-type	+/-	+	yes	deleted	-
SUM1315MO2	no	wild-type	+	+	yes	deleted	-
SK-BR-7	no	wild-type	+	+	yes	deleted	-
SUM102PT	no	wild-type	+	nd	yes	deleted	-
MDA-MB-435s	yes	wild-type	+	+	yes	mutant	++
SUM52PE	yes	wild-type	+	+	yes	mutant	+
MDA-MB-361	yes	wild-type	+	+	yes	mutant	+
OCUB-F	no	wild-type	+	nd	yes	mutant	+
UACC812	yes	wild-type	+	+	no	methylated	+/-
UACC893	yes	wild-type	+/-	+	no	methylated	+/-
T47D	yes	wild-type	+	+	yes	methylated	-
ZR75-1	yes	wild-type	+	+	no	methylated	+/-
MPE600	no	wild-type	+	+	yes	methylated	+/-
SUM44PE	nd	nd	+	+	no	methylated	+/-
MDA-MB-415	no	nd	++	+	no	mutant*	++
CAMA-1	no	wild-type	+	+	no	wild-type	+
MDA-MB-134VI	no	wild-type	+	+	no	wild-type	+
MDA-MB-330	nd	nd	+	+	no	wild-type	+
BT474	yes	wild-type	+	+	yes	wild-type	+
MDA-MB-175VII	no	wild-type	+	+	no	wild-type	+
SK-BR-3	yes	wild-type	+	+	no	wild-type	+
SK-BR-5	yes	wild-type	+	+	yes	wild-type	+
EVSA-T	yes	wild-type	+	+	no	wild-type	+
MDA-MB-157	yes	wild-type	+/-	+	yes	wild-type	++
ZR-75-30	nd	nd	+	+	no	wild-type	+
BT483	no	wild-type	+	nd	no	wild-type	+
MDA-MB-453	no	wild-type	+	+	yes	wild-type	+
SUM185PE	no	wild-type	+	++	nd	wild-type	+/-
SUM159PT	no	wild-type	+	++	no	wild-type	-

Breast cancer cell lines are organized by their gene mutation status. Expression levels and other items are explained in the legend of Table 7.1. Mutational activation of the RB pathway was identified in 25 of 39 (64%) breast cancer cell lines.

p16 protein expression	*Cyclin D1* gene locus	*Cyclin D1* transcript expression	Cyclin D1 protein expression	*CDK4* p16 BD sequence	*CDK4* gene locus	*CDK4* transcript expression	CDK4 protein expression
-	balanced	+/-	+/-	wild-type	balanced	+	+
++	balanced	+/-	+/-	wild-type	balanced	+	+
++	balanced	+/-	+/-	wild-type	balanced	+	+
++	balanced	+/-	+/-	nd	nd	+	+
++	balanced	+	+	nd	balanced	+	+
-	balanced	+	+/-	wild-type	balanced	+	+
-	balanced	+	+	wild-type	balanced	+	+
-	balanced	+	+	wild-type	nd	+	+
-	balanced	+	+	wild-type	balanced	+	+
-	balanced	+	+	wild-type	balanced	+	+
-	balanced	+	+/-	wild-type	balanced	+	+
-	balanced	+	+/-	wild-type	balanced	+	+
-	balanced	+	+	nd	nd	+	+
nd	balanced	++	nd	nd	nd	+	nd
+	balanced	+/-	+/-	wild-type	balanced	+	+
+	balanced	+	+	nd	balanced	+	+
+/-	balanced	+	+	wild-type	balanced	+	+
nd	balanced	+	nd	nd	balanced	+	nd
+	balanced	+	+	wild-type	balanced	+	+
+	balanced	+	+	wild-type	balanced	+	+
-	balanced	+	+	wild-type	balanced	+	+
+	gain^{amp}	++	++	wild-type	balanced	+	+
+	gain^{amp}	++	++	nd	balanced	+	+
+/-	gain^{amp}	++	++	nd	balanced	+	+
+/-	gain^{amp}	++	++	wild-type	balanced	+	+
+	gain^{amp}	++	++	wild-type	balanced	+	+
-	gain^{amp}	++	++	nd	balanced	+	+
-	gain^{amp}	++	++	wild-type	nd	+	+
+	gain	+	+	wild-type	balanced	+	+
+/-	gain	+	+	wild-type	balanced	+	+
+	balanced	+	+	nd	balanced	+	+
+	nd	+	+/-	wild-type	nd	+	+
+	balanced	+/-	+/-	wild-type	balanced	+	+
++	balanced	++	+	wild-type	balanced	+	+
+	balanced	+	++	nd	balanced	+	+
nd	balanced	+	nd	nd	balanced	+	nd
+/-	balanced	+	+	wild-type	balanced	+	+
+/-	balanced	+	+	wild-type	balanced	+	+
+/-	balanced	+	+	wild-type	balanced	+	+

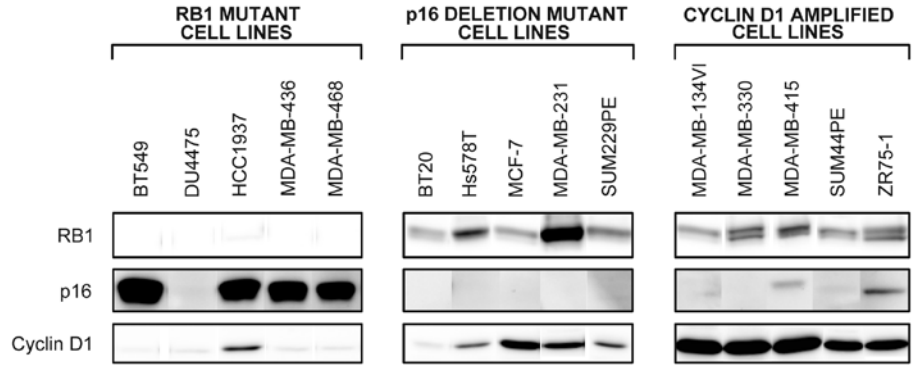


Figure 7.3 Protein expression of RB pathway genes. Western blotting of RB1, p16 and Cyclin D1 proteins. (Left panel) *RB1* mutant cell lines typically had lost RB1 expression and had high expression of p16 and low expression of Cyclin D1. (Middle panel) *p16* deletion mutant cell lines typically had lost p16 expression and had apparently normal expression of RB1 and low to normal expression of Cyclin D1. (Right panel) *Cyclin D1* amplified cell lines had Cyclin D1 overexpression, apparently normal expression of RB1 and low expression of p16.

Gene mutation profiles among breast cancer cell lines associate with breast cancer subtypes

Compilation of the mutation data revealed two prevalent gene mutation patterns among the breast cancer cell lines (Figure 7.1). The first pattern involved frequent mutations among the cell lines in genes from the same tumor suppressor pathway. These included the p53 pathway in 90% of the cell lines (*p53*, *CHEK2*, *HDM2*, *c-MYC* and *p14ARF*; Table 7.1; [36-38]), the RB pathway in 64% (*RB1*, *p16*, *Cyclin D1* and *CDK4*; Table 7.2; [37, 38]) and the PI3K pathway in 56% (*PTEN*, *PIK3CA*, *AKT1*, *AKT2* and *AKT3*; [39, 40]). In concordance with their functional relationship, oncogenic mutations within these three tumor suppressor pathways were mostly mutually exclusive among cell lines. Somewhat surprising, however, was the frequent concurrent mutation of *p53* or *CHEK2* mutations with mutations in *p14ARF*, *C-MYC* or *HDM2*. It may be of interest that all *p53/p14ARF* and *p53/c-MYC* double mutants had *p53* missense mutations, suggesting some residual activity of these p53 mutants (Fisher's exact P=0.03; Table 7.1 and Supplementary Data). Whatever the functional significance of these double mutations [38], the frequent mutational abrogation of the p53, RB and PI3K tumor suppressor pathways among the cell lines suggests that their functionality is of major importance in breast carcinogenesis.

Most exciting, the second gene mutation pattern discerned among the breast cancer cell lines involved oncogenic mutations that cluster in a particular subset of cell lines (Figure 7.1). These included a "luminal mutation profile" involving *E-cadherin* and *MAP2K4* gene mutations and amplifications of *Cyclin D1*, *ERBB2* and *HDM2*, and a "basal mutation profile" involving *BRCA1*, *RB1*, *RAS* and *BRAF* gene mutations and deletions of *p16* and *p14ARF*. Specifically, the cohort of 25 luminal-type breast cancer cell lines included all ten *E-cadherin* mutants, all seven *Cyclin D1* mutants, all eleven *ERBB2* mutants, all four *HDM2* mutants, and three of four *MAP2K4* mutants (Fisher's exact P<0.0001). The cohort of 15 basal-type breast cancer cell lines included all four

BRCA1 mutants, four of five *RB1* mutants, eight of nine *p16/p14ARF* deletion mutants, and six of nine cell lines with mutational activation of the RAS pathway (through *KRAS*, *HRAS*, *NRAS* or *BRAF*; P<0.0001). The existence of specific gene mutation profiles for each of the two major subtypes of breast cancer cell lines suggests a dichotomy in the genetic basis of human breast cancer.

The functional relationship among luminal mutation profile genes or among basal mutation profile genes is as yet unclear. It appeared that mutations in *E-cadherin* and *MAP2K4* were mutually exclusive with *HDM2* amplifications. Similarly, amplifications of *ERBB2* appeared essentially mutually exclusive with *Cyclin D1* amplifications and with *MAP2K4* mutations. Yet, three *MAP2K4* mutants had concurrent *Cyclin D1* amplifications. It may be of interest that the DU4475 cell line is the single *MAP2K4* mutant without Cyclin D1 amplification. DU4475 has mutational activation of the Wnt pathway, of which *Cyclin D1* is a known transcriptional target [41]. Among basal mutation profile genes, *BRCA1* mutations appeared mutually exclusive with RAS pathway activation and mutations in *p16* and *RB1* also were mutually exclusive. But the four *BRCA1* mutants had concurrent mutation of either *p16* or *RB1*. Although the numbers are still small, these data may suggest functional cross talk between biological pathways that are mutated within a breast cancer subtype.

DISCUSSION

Molecular characterization suggests two major types of breast cancer cell lines

Classification of breast cancer cell lines was remarkably consistent, whether they were classified by classical breast cancer parameters (ER, PR and ERBB2), by histological criteria (CK expression or the 4-protein signature), or by gene expression profiling (intrinsic subtypes; Figure 7.1). All 24 breast cancer cell lines with (over)expression of ER, PR and/or ERBB2 classified as intrinsic luminal or ERBB2+ subtype breast cancers, and all of these cell lines also classified as luminal phenotype breast cancers by CK expression. It is important to realize that clinical breast cancers of the intrinsic ERBB2+ subtype typically do not express ER and that ER/ERBB2-positive breast cancers are mostly classified as intrinsic luminal subtype breast cancers. The reported subdivision of ERBB2-overexpressing clinical breast cancers by their gene expression program had been significant [13], but our data suggest that it might be more appropriate to consider breast cancers of the intrinsic ERBB2+ subtype (ERBB2+, ER-) as a subgroup of luminal-type breast cancers.

All 15 breast cancer cell lines that classified as basal-like or normal-like breast cancers by intrinsic gene expression profiling and all basal or null phenotype cell lines by CK expression were triple-negative (ER/PR/ERBB2-negative). Our mutation analyses strongly suggested that these two subtypes of breast cancer cell lines represent two ends of a spectrum of basal-type breast cancers (Figure 7.1). This indeed was supported by the protein expression profiling data,

specifically expression of E-cadherin, N-cadherin, Vimentin and EGFR. Most notable, a similar interrelationship has been proposed for clinical breast cancers [9] and, accordingly, Neve *et al.* had designated these two subtypes of breast cancer cell lines as basal A and basal B [33]. Our molecular characterization of human breast cancer cell lines thus supports the concept that the two major types of luminal and basal breast cancers may be subdivided in the luminal A, luminal B and ERBB2+ subtypes, and in the basal A and basal B (intrinsic normal-like or CK null) subtypes, respectively.

Accumulation of mutations in many well-known cancer genes in breast cancer cell lines

We have identified 146 oncogenic mutations among 41 breast cancer cell lines. This mutation frequency was much higher than we had anticipated. For example, 78% of cell lines carried a *p53* mutation, whereas 30-50% of clinical breast cancers reportedly carry *p53* mutations [42]. Almost one-quarter of cell lines had homozygous deletions of *p16/p14ARF* and another one-quarter had *E-cadherin* mutations versus about 15% and 10% of clinical breast cancers, respectively [43, 44] The high mutation frequency among cell lines likely reflects a higher rate of mutation detection in these samples. Importantly, the inevitable presence of normal, non-neoplastic cells in clinical breast cancer samples essentially precludes identification of homozygous deletions by standard PCR, whereas these represent 20% of 146 oncogenic mutations that we have identified in the cell lines. The Cancer Genome Project at the Sanger Institute also has screened known cancer genes in breast cancer cell lines, with 13 genes and 22 cell lines overlapping with our study ([45] and www.sanger.ac.uk). It is notable that the Sanger Institute failed to identify 34% of 64 oncogenic mutations that we had detected and that the Sanger Institute did not identify additional mutations. The reason for this discrepancy is unclear, but we expect that our mutation screens may involve more detailed primer design and PCR optimization and more elaborate analysis of sequence data. With regard to the higher mutation frequency in cell lines compared to clinical samples, we consider it unlikely that a significant proportion of mutations has been acquired during *in vitro* propagation of the cell lines, although there is no hard evidence to support this notion nor, for that matter, to reject it. Most important, the majority of gene alterations identified in the cell lines appeared to be of functional relevance as they had previously been reported in clinical cancer samples, in the germline of patients with a hereditary cancer syndrome, or were sizeable deletions or intragenic mutations that predicted a premature protein termination (and thus are less likely to be recurrent). Silent mutations or missense mutations of unclear significance represented less than 10% of all identified deviations from the reference coding gene sequences (5 and 7 mutations, respectively; [23-28, 34]). It also is significant that mutations in several genes follow patterns observed for clinical breast cancers, such as the exclusive presence of *Cyclin D1* amplifications in ER-positive cell lines, *RB1* mutations in ER-negative cell lines and the concurrent mutation of *p53* and *BRCA1* in ER-negative cell lines [46-49]. A more likely explanation for the high mutation frequency among cell lines is a bias for breast cancers that are more easily propagated *in vitro*, such as *p53* mutant breast cancers and

likely also those of the null phenotype — both being cancers that are known to have a more aggressive clinical course. This might, however, not necessarily be disadvantageous as it allows one to study the more clinically-needy breast cancers.

Gene mutation profiles provide a genetic basis for luminal and basal-type breast cancers

We have identified two subtype-specific gene mutation profiles among the breast cancer cell lines. The luminal mutation profile, involving *E-cadherin* and *MAP2K4* gene mutations and amplifications of *Cyclin D1*, *ERBB2* and *HDM2*, was identified among 21 of 25 luminal-type breast cancer cell lines and 97% of 36 mutations in luminal mutation profile genes had been identified among the luminal cell lines. The DU4475 cell line was the single non-luminal cell line that had a mutation in a luminal mutation profile gene: *MAP2K4*. DU4475 also was the only cell line that did not classify for any of the intrinsic subtypes of breast cancer and it was the only cell line that classified for the rare stem cell phenotype by cytokeratin expression (CK8/18 low, CK19- and CK5+, representing <1% of clinical breast cancers [4]). DU4475 also is an atypical breast cancer cell line because of its aberrant Wnt pathway activation, due to a nonsense *APC* gene mutation (Supplementary Data). Wnt signaling has been associated with stem cells in colonic and other epithelia [41, 50]. It is tempting to speculate that DU4475 represents a malignant counterpart of the committed stem cell in normal breast epithelium [51].

The basal mutation profile involved *BRCA1*, *RB1* and *RAS* pathway gene mutations and deletions of *p16* and *p14ARF*. The basal mutation profile was identified among 14 of 15 basal-type breast cancers. Together, the basal breast cancer cell lines carried 33 of the 39 (85%) mutations that we had identified in basal mutation profile genes. DU4475 again was atypical because of its mutations in the basal mutation profile genes *RB1* and *BRAF*. The MDA-MB-435s cell line was the only basal-type cell line that did not express EGFR proteins. MDA-MB-435s was recently shown to be genetically identical to the M14 melanoma cell line from the NCI-60 cell line panel (several MDA-MB-435s vials versus a single M14 vial; [52]). We have obtained M14 cells from several laboratories, including a vial of passage-10 cells from its originator, and MDA-MB-435s was obtained from ATCC at passage-239. Indeed, we confirmed by microsatellite analysis that the two cell lines were genetically identical. It thus remains unclear which is the correct origin of the cell lines. Be as it may, it is not likely that the other basal-type breast cancer cell lines are of melanocytic origin. Four of the basal cell lines carry *BRCA1* mutations and one carries the *CHEK2* 1100delC mutation, both well-known breast cancer susceptibility genes [53-57]. In fact, the consistent histological and intrinsic breast cancer classification of the basal cell lines strongly suggests that a melanoma-like gene mutation profile (that includes mutations in *p16*, *RB1* and *RAS* pathway genes; [58]) is characteristic for basal-type breast cancers. Comparable gene mutation profiles among basal-type breast cancers and melanomas, and likely also pancreatic cancers [59], might be conceivable given the typically aggressive clinical course of these tumor types and their frequent inherent therapy resistance. In this respect, distinct gene mutation

profiles between luminal and basal-type breast cancers could very well explain the significant heterogeneity in clinical outcome of breast cancer patients.

Tumor biology reflects the inherent differentiation status of the tumor cells, but also their genetic make-up of germline and somatic gene mutations. It is of interest that recent large-scale sequencing efforts in breast cancer and colorectal cancer have identified a similar number of somatic oncogenic mutations among breast cancers and colorectal cancers, but that the spectrum of mutated biological pathways was far more diverse among breast cancers than colorectal cancers (an average of 14 and 15 mutations per tumor among 108 and 38 pathways for breast cancer and colorectal cancer, respectively; [35, 60]). These differences in mutation spectrum between breast cancers and colorectal cancers had suggested genetic heterogeneity among breast cancers, and are thus in line with our finding of two distinct gene mutations profiles among breast cancer cell lines. After all, the existence of two gene mutation profiles among breast cancers implies more mutant pathways when considering breast cancers at large. Similarly, mutations in subtype-specific cancer genes will then be less prevalent. Our study has focused on mutations in well-known cancer genes and the identified mutations therefore are more likely to represent drivers of breast carcinogenesis. Although this approach does not draw the full genomic landscape of breast cancers, it did allow us to decipher specific patterns of gene mutations among the two major breast cancer types. Of course, breast cancer cell lines are not clinical breast cancers. As a proof of principle, we have shown in chapter 3 that *E-cadherin* mutation and *E-cadherin* promoter methylation indeed also associate with distinct subtypes of clinical breast cancers. If the evidence is to be extended to other genes from the two gene mutation profiles, as genome copy number analysis suggested [61], we may at last have begun to elucidate the genetic basis for the inherent biological and clinical heterogeneity among human breast cancers.

ACKNOWLEDGEMENTS

We appreciate the technical assistance of Michel Molier, Sofia Zuñiga, Mieke Timmermans, Hein Sleddens and members of the Pathology Immunohistochemistry laboratories at Erasmus MC. We thank Dr. Jörg Volkland, Dr. Nelleke Gruis, Dr. Goos van Muijen and Dr. Jan Verheijen for kindly providing us with M14 cells. Members of the Foekens, Berns and Dorssers laboratories are kindly acknowledged for useful discussions. This work was supported by grants from the Dutch Cancer Society (DDHK 2002-2687 and 2003-2862), Susan G. Komen Breast Cancer Foundation (BCTR0601309), Association for International Cancer Research (05-273), Netherlands Genomics Initiative (NGI)/Netherlands Organization for Scientific Research (NWO) and Erasmus MC Mrace 2005.

REFERENCES

1. Nagle RB, Bocker W, Davis JR, et al. Characterization of breast carcinomas by two monoclonal antibodies distinguishing myoepithelial from luminal epithelial cells. J Histochem Cytochem 1986;34(7):869-81.
2. Wetzels RH, Kuijpers HJ, Lane EB, et al. Basal cell-specific and hyperproliferation-related keratins in human breast cancer. Am J Pathol 1991;138(3):751-63.
3. Malzahn K, Mitze M, Thoenes M, Moll R. Biological and prognostic significance of stratified epithelial cytokeratins in infiltrating ductal breast carcinomas. Virchows Arch 1998;433(2):119-29.
4. Abd El-Rehim DM, Pinder SE, Paish CE, et al. Expression of luminal and basal cytokeratins in human breast carcinoma. J Pathol 2004;203(2):661-71.
5. Jones C, Nonni AV, Fulford L, et al. CGH analysis of ductal carcinoma of the breast with basaloid/myoepithelial cell differentiation. British journal of cancer 2001;85(3):422-7.
6. Korsching E, Packeisen J, Agelopoulos K, et al. Cytogenetic alterations and cytokeratin expression patterns in breast cancer: integrating a new model of breast differentiation into cytogenetic pathways of breast carcinogenesis. Lab Invest 2002;82(11):1525-33.
7. Gusterson BA, Ross DT, Heath VJ, Stein T. Basal cytokeratins and their relationship to the cellular origin and functional classification of breast cancer. Breast Cancer Res 2005;7(4):143-8.
8. Rakha EA, El-Sayed ME, Green AR, Paish EC, Lee AH, Ellis IO. Breast carcinoma with basal differentiation: a proposal for pathology definition based on basal cytokeratin expression. Histopathology 2007;50(4):434-8.
9. Fulford LG, Easton DF, Reis-Filho JS, et al. Specific morphological features predictive for the basal phenotype in grade 3 invasive ductal carcinoma of breast. Histopathology 2006;49(1):22-34.
10. Nielsen TO, Hsu FD, Jensen K, et al. Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. Clin Cancer Res 2004;10(16):5367-74.
11. Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. Nature 2000;406(6797):747-52.
12. Sorlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A 2001;98(19):10869-74.
13. Sorlie T, Tibshirani R, Parker J, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proc Natl Acad Sci U S A 2003;100(14):8418-23.
14. Rouzier R, Perou CM, Symmans WF, et al. Breast cancer molecular subtypes respond differently to preoperative chemotherapy. Clin Cancer Res 2005;11(16):5678-85.
15. Smid M, Wang Y, Zhang Y, et al. Subtypes of breast cancer show preferential site of relapse. Cancer Research 2008;68(9):3108-14.
16. van 't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature 2002;415(6871):530-6.
17. Sotiriou C, Wirapati P, Loi S, et al. Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. Journal of the National Cancer Institute 2006;98(4):262-72.
18. Wang Y, Klijn JG, Zhang Y, et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. Lancet 2005;365(9460):671-9.
19. Foekens JA, Wang Y, Martens JWM, Berns EMJJ, Klijn JGM. The use of genomic tools for the molecular understanding of breast cancer and to guide personalized medicine. Drug Discovery Today 2008;13(11-12):481-7.
20. Harkes IC, Elstrodt F, Dinjens WN, et al. Allelotype of 28 human breast cancer cell lines and xenografts. British journal of cancer 2003;89(12):2289-92.
21. Chee DO, Boddie AW, Roth JA, Holmes EC, Morton DL. Production of melanoma-associated antigen(s) by a defined malignant melanoma cell strain grown in chemically defined medium. Cancer Res 1976;36(4):1503-9.

22. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 1998;95(25):14863-8.
23. Elstrodt F, Hollestelle A, Nagel JH, et al. BRCA1 mutation analysis of 41 human breast cancer cell lines reveals three new deleterious mutants. Cancer Res 2006;66(1):41-5.
24. Wasielewski M, Hanifi-Moghaddam P, Hollestelle A, et al. Deleterious CHEK2 1100delC and L303X mutants identified among 38 human breast cancer cell lines. Breast cancer research and treatment 2008;In press.
25. van de Wetering M, Barker N, Harkes IC, et al. Mutant E-cadherin breast cancer cells do not display constitutive Wnt signaling. Cancer Res 2001;61(1):278-84.
26. Su GH, Song JJ, Repasky EA, Schutte M, Kern SE. Mutation rate of MAP2K4/MKK4 in breast carcinoma. Human mutation 2002;19(1):81.
27. Wasielewski M, Elstrodt F, Klijn JG, Berns EM, Schutte M. Thirteen new p53 gene mutants identified among 41 human breast cancer cell lines. Breast cancer research and treatment 2006;99(1):97-101.
28. Hollestelle A, Elstrodt F, Nagel JH, Kallemeijn WW, Schutte M. Phosphatidylinositol-3-OH kinase or RAS pathway mutations in human breast cancer cell lines. Mol Cancer Res 2007;5(2):195-201.
29. Carpten JD, Faber AL, Horn C, et al. A transforming mutation in the pleckstrin homology domain of AKT1 in cancer. Nature 2007;448(7152):439-44.
30. Schlosshauer PW, Brown SA, Eisinger K, et al. APC truncation and increased beta-catenin levels in a human breast cancer cell line. Carcinogenesis 2000;21(7):1453-6.
31. Zuo L, Weger J, Yang Q, et al. Germline mutations in the p16INK4a binding domain of CDK4 in familial melanoma. Nature genetics 1996;12(1):97-9.
32. van Oers JM, Lurkin I, van Exsel AJ, et al. A simple and fast method for the simultaneous detection of nine fibroblast growth factor receptor 3 mutations in bladder cancer and voided urine. Clin Cancer Res 2005;11(21):7743-8.
33. Neve RM, Chin K, Fridlyand J, et al. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. Cancer Cell 2006;10(6):515-27.
34. Esteller M, Silva JM, Dominguez G, et al. Promoter hypermethylation and BRCA1 inactivation in sporadic breast and ovarian tumors. Journal of the National Cancer Institute 2000;92(7):564-9.
35. Wood LD, Parsons DW, Jones S, et al. The genomic landscapes of human breast and colorectal cancers. Science (New York, NY 2007;318(5853):1108-13.
36. Vogelstein B, Lane D, Levine AJ. Surfing the p53 network. Nature 2000;408(6810):307-10.
37. Sherr CJ, McCormick F. The RB and p53 pathways in cancer. Cancer Cell 2002;2(2):103-12.
38. Sherr CJ. Divorcing ARF and p53: an unsettled case. Nature reviews 2006;6(9):663-73.
39. Manning BD, Cantley LC. AKT/PKB signaling: navigating downstream. Cell 2007;129(7):1261-74.
40. Vivanco I, Sawyers CL. The phosphatidylinositol 3-Kinase AKT pathway in human cancer. Nature reviews 2002;2(7):489-501.
41. Tetsu O, McCormick F. Beta-catenin regulates expression of cyclin D1 in colon carcinoma cells. Nature 1999;398(6726):422-6.
42. Borresen-Dale AL. TP53 and breast cancer. Human mutation 2003;21(3):292-300.
43. Cairns P, Polascik TJ, Eby Y, et al. Frequency of homozygous deletion at p16/CDKN2 in primary human tumours. Nature genetics 1995;11(2):210-2.
44. Berx G, Becker KF, Hofler H, van Roy F. Mutations of the human E-cadherin (CDH1) gene. Human mutation 1998;12(4):226-37.
45. Ikediobi ON, Davies H, Bignell G, et al. Mutation analysis of 24 known cancer genes in the NCI-60 cell line set. Molecular cancer therapeutics 2006;5(11):2606-12.
46. Bieche I, Lidereau R. Loss of heterozygosity at 13q14 correlates with RB1 gene underexpression in human breast cancer. Molecular carcinogenesis 2000;29(3):151-8.
47. Barnes DM, Gillett CE. Cyclin D1 in breast cancer. Breast cancer research and treatment 1998;52(1-3):1-15.

48. Gasco M, Yulug IG, Crook T. TP53 mutations in familial breast cancer: functional aspects. Human mutation 2003;21(3):301-6.
49. Turner NC, Reis-Filho JS. Basal-like breast cancer and the BRCA1 phenotype. Oncogene 2006;25(43):5846-53.
50. Clevers H. Wnt/beta-catenin signaling in development and disease. Cell 2006;127(3):469-80.
51. Gudjonsson T, Villadsen R, Nielsen HL, Ronnov-Jessen L, Bissell MJ, Petersen OW. Isolation, immortalization, and characterization of a human breast epithelial cell line with stem cell properties. Genes & development 2002;16(6):693-706.
52. Rae JM, Creighton CJ, Meck JM, Haddad BR, Johnson MD. MDA-MB-435 cells are derived from M14 melanoma cells--a loss for breast cancer, but a boon for melanoma research. Breast cancer research and treatment 2007;104(1):13-9.
53. Miki Y, Swensen J, Shattuck-Eidens D, et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. Science (New York, NY 1994;266(5182):66-71.
54. Meijers-Heijboer H, van den Ouweland A, Klijn J, et al. Low-penetrance susceptibility to breast cancer due to CHEK2(*)1100delC in noncarriers of BRCA1 or BRCA2 mutations. Nature genetics 2002;31(1):55-9.
55. Vahteristo P, Bartkova J, Eerola H, et al. A CHEK2 genetic variant contributing to a substantial fraction of familial breast cancer. American journal of human genetics 2002;71(2):432-8.
56. Thompson D, Easton D. The genetic epidemiology of breast cancer genes. Journal of mammary gland biology and neoplasia 2004;9(3):221-36.
57. Walsh T, King MC. Ten genes for inherited breast cancer. Cancer Cell 2007;11(2):103-5.
58. Gray-Schopfer V, Wellbrock C, Marais R. Melanoma biology and new targeted therapy. Nature 2007;445(7130):851-7.
59. Jaffee EM, Hruban RH, Canto M, Kern SE. Focus on pancreas cancer. Cancer Cell 2002;2(1):25-8.
60. Sjoblom T, Jones S, Wood LD, et al. The consensus coding sequences of human breast and colorectal cancers. Science (New York, NY 2006;314(5797):268-74.
61. Adelaide J, Finetti P, Bekhouche I, et al. Integrated profiling of basal and luminal breast cancers. Cancer Res 2007;67(24):11565-75.

Summary and Samenvatting

SUMMARY

E-cadherin is a well-known tumor suppressor gene that encodes a cell adhesion molecule involved in maintaining integrity of epithelial tissues. In breast cancer, E-cadherin was found to be inactivated by mutation of the gene in half of breast cancers of the lobular subtype. Alternatively, E-cadherin may be inactivated by hypermethylation of CpG islands in the *E-cadherin* promoter region and transcriptional repression through its repressor proteins, but this had not been associated with a particular breast cancer subtype. The studies described in this thesis were aimed at improving our understanding of E-cadherin inactivation mechanisms in breast carcinogenesis, by using a model of 41 human breast cancer cell lines. Furthermore, we sought to identify the genetic basis of the breast cancer subtypes that associated with the two different modes of E-cadherin inactivation.

Chapter 1 comprises a general introduction on the epidemiology, classification and molecular genetics of breast cancer and **chapter 2** describes the aim and outline of this thesis.

In **chapter 3** we have studied inactivation mechanisms of *E-cadherin* in the breast cancer cell lines. Mutations of *E-cadherin* were found solely in cell lines with a rounded cell morphology. In contrast, hypermethylation of *E-cadherin* and expression of its transcriptional repressors associated with cell lines with a spindle cell morphology. Gene reconstitution experiments revealed that inactivation of *E-cadherin* was causal for the rounded cell morphology but not for the spindle cell morphology. Concordantly, transcriptional profiling of the cell lines showed that cell lines with genetic and epigenetic inactivation of *E-cadherin* had vastly different gene expression programs, implying that the two inactivation mechanisms involve distinct biological pathways. Indeed, we observed mutations of *E-cadherin* only in lobular breast cancers, whereas epigenetic inactivation of *E-cadherin* associated with clinical breast cancers of the metaplastic pathological subtype. These results challenge the paradigm that inactivation of a tumor suppressor gene by promoter hypermethylation is biologically similar to mutational inactivation of the gene.

Next, we have evaluated the genetic basis of the two inactivation modes of E-cadherin in the breast cancer cell lines. In **chapter 4** we have searched for another cancer gene in the E-cadherin mutational tumor suppressor pathway. Like in lobular breast cancers, not all breast cancer cell lines with the rounded cell morphology had a mutant *E-cadherin* gene. Because we had shown that E-cadherin was causal for the rounded cell phenotype, we hypothesized that another gene in the E-cadherin pathway was mutated in the rounded cell lines with wild-type *E-cadherin*. Expression and mutation analyses of the E-cadherin-associated proteins α-, β-, γ-catenin and p120ctn in the cell lines revealed that α-*catenin* was genetically inactivated in the rounded cell lines with wild-type *E-cadherin*, suggesting that α-*catenin* is a new tumor suppressor gene involved in lobular breast cancer. Indeed, loss of α-catenin protein expression specifically associated with clinical breast cancers of the lobular subtype. These results thus suggest α-*catenin* as a putative new tumor suppressor gene.

In **chapter 5** we have analyzed the breast cancer cell lines for mutations of PI3K and RAS pathway genes. Inactivation of the PI3K pathway, through mutations of *PIK3CA* and *PTEN*, was identified in 21 cell lines, but the mutations were equally prevalent among *E-cadherin* mutant and methylated cell lines. RAS pathway mutations, in *KRAS*, *HRAS*, *NRAS* and *BRAF*, were identified in nine cell lines. Interestingly, eight of the eleven RAS pathway mutations were found in *E-cadherin* methylated cell lines, suggesting an association with this inactivation mechanism of E-cadherin. Moreover, we found that unlike in colorectal cancers, mutational activation of the PI3K pathway was mutually exclusive with mutational activation of the RAS pathway. This suggests that there is a fine distinction between the signaling activators and downstream effectors of the oncogenic PI3K and RAS pathways in breast epithelium and colonic epithelium.

In **chapter 6** all breast cancers cell lines were analyzed for mutations in the *BRCA1* gene. We identified three new human *BRCA1* mutant cell lines that seem to be representative breast cancer models that could aid in further unraveling of the function of BRCA1. Importantly, three of the four cell lines with deleterious *BRCA1* mutations had epigenetic inactivation of E-cadherin, suggesting an association of *BRCA1* mutation with *E-cadherin* methylation.

In **chapter 7** we have performed a large scale molecular characterization of the breast cancer cell lines. Protein and transcript expression analysis revealed that the cell lines resemble two major subtypes of luminal and basal-type breast cancers. Mutation analysis of 27 well-known cancer genes identified 146 oncogenic mutations, including 92 inactivating mutations among 12 tumor suppressor genes and 52 activating mutations among 12 oncogenes. The mutational data combined revealed two gene mutation patterns among the cell lines. First, we identified frequent mutations among genes of the p53, PI3K and RB pathways. Second, we identified distinct mutational profiles specific for the luminal and basal-types of breast cancer. The luminal mutation profile was observed among 21 of 25 luminal-type cell lines and involved the mutation of *E-cadherin* and *MAP2K4*, as well as amplification of *ERBB2*, *CyclinD1* and *HDM2*. The basal mutation profile was observed among 14 of the 15 basal-type cell lines and involved *RB1*, *BRCA1*, *RAS* and *BRAF* gene mutations and *p16/p14ARF* deletions. The identification of two subtype-specific mutation profiles among the breast cancer cell lines provides a genetic basis for luminal and basal-type breast cancer.

SAMENVATTING

E-cadherine is een bekend tumorsuppressorgen dat codeert voor een celadhesiemolecuul dat de integriteit van epitheliale weefsels bewaakt. In borstkanker wordt E-cadherine geïnactiveerd door mutatie in de helft van de borstkankers van het lobulaire subtype. Als alternatief kan E-cadherine geïnactiveerd worden door hypermethylatie van CpG-eilanden in de promotorregio van het *E-cadherine* gen of door transcriptionele onderdrukking door repressoreiwitten, maar dit was nog niet geassocieerd met een bepaald subtype van borstkanker. De studies die zijn beschreven in dit proefschrift beogen het verbeteren van ons begrip van deze E-cadherine inactivatiemechanismen in borstkanker door onderzoek te doen aan 41 humane borstkankercellijnen, die model staan voor borstkanker bij de mens. Daarnaast hebben we getracht andere genetische veranderingen te identificeren in de borstkankersubtypen die associeerden met de twee verschillende manieren van E-cadherine inactivatie.

Hoofdstuk 1 omvat een algemene inleiding in de epidemiologie, classificatie en moleculaire genetica van borstkanker en in **hoofdstuk 2** worden de doelstellingen en de opzet van dit proefschrift beschreven.

In **hoofdstuk 3** wordt verslag gedaan van onderzoek naar de inactivatiemechanismen van *E-cadherine* in de borstkankercellijnen. Mutaties in het *E-cadherine* gen werden enkel gevonden in borstkankercellijnen met een rondcellige morfologie. Dit in tegenstelling tot hypermethylatie van *E-cadherine* en expressie van de transcriptionele repressoreiwitten, die associeerden met borstkankercellijnen met een spoelcellige morfologie. Genreconstitutie experimenten lieten zien dat inactivatie van *E-cadherine* oorzakelijk was voor de rondcellige morfologie, maar niet voor de spoelcellige morfologie. In overeenstemming met deze bevinding werd gevonden dat de genexpressieprofielen van de cellijnen met genetische en epigenetische inactivatie van *E-cadherine* immens verschillen. Dit impliceert dat de twee inactivatiemechanismen verschillende biologische routes betreffen. We vonden mutaties van *E-cadherine* inderdaad alleen in lobulaire borstkankers, terwijl epigenetische inactivatie van *E-cadherine* associeerde met borstkanker van het metaplastische subtype. Deze resultaten ondermijnen het paradigma dat inactivatie van een tumorsuppressorgen door promotorhypermethylatie biologisch gelijk is aan inactivatie door mutatie van het gen.

Vervolgens hebben wij de genetische basis van de twee manieren van E-cadherine inactivatie in de borstkankercellijnen geëvalueerd. In **hoofdstuk 4** wordt de zoektocht naar andere kankergenen in de E-cadherine tumorsuppressorroute beschreven. Net als in lobulaire borsttumoren hadden niet alle borstkankercellijnen met een rondcellige morfologie een mutatie in het *E-cadherine* gen. Omdat wij hebben laten zien dat E-cadherine oorzakelijk is voor de rondcellige morfologie was onze hypothese dat een ander gen in de E-cadherine signaaltransductieroute gemuteerd moest zijn in cellijnen met rondcellige morfologie zonder een mutatie in het *E-cadherine* gen. Expressie en mutatie analyse van de met E-cadherine geassocieerde eiwitten α-, β-, γ-catenine en p120ctn in de cellijnen liet zien dat α-*catenine* juist genetisch

geïnactiveerd is in de cellijnen met rondcellige morfologie zonder een gemuteerd *E-cadherine* gen. Dit suggereert dat α-*catenine* een tumorsuppressorgen is dat genetisch is geïnactiveerd in lobulair borstkanker. Inderdaad, verlies van α-catenine eiwitexpressie werd specifiek gevonden in borstkanker van het lobulaire subtype. Deze resultaten suggereren dat α-*catenine* vermoedelijk een nieuw tumorsuppressorgen is.

In **hoofdstuk 5** rapporteren wij de analyse van mutaties in genen van de PI3K en RAS signaaltransductieroutes in borstkankercellijnen. Inactivatie van de PI3K signaaltransductieroute, door mutaties in de genen *PIK3CA* en *PTEN*, werd geïdentificeerd in 21 cellijnen, maar de mutaties kwamen even vaak voor in *E-cadherine*-mutante en gemethyleerde cellijnen. Mutaties in genen van de RAS signaaltransductieroute werden gevonden in negen cellijnen. Interessant was dat acht van de elf mutaties in genen van de RAS signaaltransductie route gevonden werden in *E-cadherine* gemethyleerde cellijnen, wat een associatie met dit E-cadherine inactivatiemechanisme suggereert. Daarnaast hebben we gevonden dat, in tegenstelling tot darmtumoren, activatie door mutatie van de PI3K en de RAS signaaltransductieroutes elkaar uitsluit. Dit suggereert dat er een subtiel verschil is tussen de signaalactiveerders en de stroomafwaartse effectors van de oncogene PI3K en RAS signaaltransductieroutes in borstepitheel en die in darmepitheel.

In **hoofdstuk 6** is de analyse beschreven van mutaties in het *BRCA1* gen in borstkankercellijnen. We hebben drie nieuwe humane *BRCA1* mutante cellijnen geïdentificeerd. Deze cellijnen lijken representatieve modellen die verder kunnen bijdragen aan het ontrafelen van de functie van BRCA1. Belangrijk is dat drie van de vier cellijnen met een mutatie in het *BRCA1* gen ook epigenetische inactivatie van *E-cadherine* hadden. Dit suggereert een associatie van *BRCA1* mutatie met *E-cadherine* methylatie.

In **hoofdstuk 7** rapporteren wij een grootschalige moleculaire karakterisatie van de borstkankercellijnen. Eiwit- en transcriptanalyses brachten aan het licht dat de cellijnen onderverdeeld kunnen worden in twee hoofdsubtypen van borstkanker, namelijk het luminale en het basale type borstkanker. Mutatieanalyse van 27 bekende kankergenen identificeerde 146 oncogene mutaties, waarvan 92 inactiverende mutaties in 12 tumorsuppressorgenen en 52 activerende mutaties in 12 oncogenen. Alle mutatiedata gecombineerd maakten twee genmutatiepatronen onder de borstkankercellijnen zichtbaar. Als eerste vonden we dat genen uit de p53, PI3K and RB signaaltransductieroutes vaak gemuteerd waren. Ten tweede identificeerden we verschillende mutatieprofielen die specifiek waren voor de basale en luminale typen van borstkanker. Het luminale mutatieprofiel was aanwezig in 21 van de 25 luminale borstkankercellijnen en omvatte zowel mutaties van *E-cadherine* en *MAP2K4*, als amplificatie van *ERBB2*, *CyclinD1* en *HDM2*. Het basale mutatieprofiel was aanwezig in 14 van de 15 basale borstkankercellijnen en omvatte *RB1*, *BRCA1*, *RAS* en *BRAF* genmutaties en *p16/p14ARF* deleties. De identificatie van twee subtypenspecifieke mutatieprofielen onder de borstkankercellijnen voegt een genetische dimensie toe aan de huidige, expressiegebaseerde indeling in luminale en basale typen van borstkanker.

Chapter 9

General Discussion

GENERAL DISCUSSION

In this thesis we have evaluated E-cadherin inactivation mechanisms by using human breast cancer cell lines as a model. We found that genetic inactivation of E-cadherin was biologically distinct from epigenetic inactivation and expression of E-cadherin's transcriptional repressors. Mutational inactivation of E-cadherin associated with a rounded cell morphology in breast cancer cell lines and with clinical breast cancers of the lobular subtype, whereas epigenetic silencing by promoter hypermethylation associated with a spindle cell morphology in cell lines and with clinical metaplastic breast cancers. Importantly, the two inactivation mechanisms of E-cadherin associated with vastly different transcriptional programs that we linked to luminal-type and basal-type breast cancers. Mutation analysis of 27 cancer genes then revealed that the two inactivation mechanisms also associated with distinct gene mutation profiles. The luminal mutation profile included mutation of *E-cadherin* and *MAP2K4* and amplification of *ERBB2*, *Cyclin D1* and *HDM2*. The basal mutation profile included hypermethylation of *E-cadherin*, mutation of *BRCA1*, *RB1*, and *RAS* pathway genes and deletion of *p16/p14ARF*.

Genetic versus epigenetic E-cadherin inactivation in breast cancer

The observation that genetic and epigenetic inactivation of *E-cadherin* involve distinct biological pathways explains recurrent controversies regarding both the functional role and prognostic value of E-cadherin. Whereas functional studies had suggested that inactivation of the *E-cadherin* tumor suppressor gene is involved in the invasion stage of carcinogenesis [1, 2], *E-cadherin* gene mutations were found to be already present in the premalignant carcinoma *in situ* stage of human breast carcinoma [3]. These studies initially appear contradictory. However, Frixen *et al.* and Vleminckx *et al.* had based their conclusions on experiments involving manipulation of an *E-cadherin* hypermethylated cell line and spindle-shaped v-*ras*-transformed MDCK cells, respectively. This suggests that epigenetic inactivation mechanisms are involved in the invasion steps of breast tumorigenesis, whereas mutation of *E-cadherin* is an earlier event in tumorigenesis. The association of the distinct modes of *E-cadherin* inactivation with their involvement at different time points in breast tumorigenesis thus strengthens our observed association with distinct biological pathways.

Recurrent discrepancies also exist among the prognostic value of E-cadherin as these studies are mostly based on loss of E-cadherin protein expression alone. However, the results of this thesis suggest that E-cadherin protein expression is not a good marker to assess the prognostic role of E-cadherin inactivation. First of all, loss of E-cadherin protein expression does not distinguish inactivation by mutation from inactivation by hypermethylation. Second, tumors that have hypermethylated *E-cadherin* genes without loss of E-cadherin protein expression will not correctly be assigned to the '*E-cadherin* hypermethylation' group. And third, tumors with in-frame mutations of *E-cadherin* will not correctly be assigned to the '*E-cadherin* mutant' group. Proper assessment of the prognostic value of E-cadherin would require analysis of both inactivation

modes separately and it is therefore necessary to also include the 4-protein signature of Nielsen *et al.*, (ERBB2, ER, CK5, EGFR) as well as our 3-protein spindle signature (CAV1, VIM, CALD1) to aid correct classification of the *E-cadherin* hypermethylated tumors ([4] and chapter 3). The correct classification of mutant *E-cadherin* breast tumors still poses a problem as tumors with in-frame mutations of *E-cadherin* can not be detected by immuno staining for E-cadherin and the two aforementioned protein signatures. Recently, p120ctn protein expression was shown to have diagnostic utility in discriminating lobular from ductal breast cancers [5]. In lobular breast cancers p120ctn was present in the cytoplasm only, whereas in ductal carcinomas p120ctn was present at the cell membrane. Immuno staining of our breast cancer cell lines for p120ctn showed that E-cadherin expressing cell lines with in-frame mutations had cytoplasmic staining, as did cell lines with truncating mutations of *E-cadherin* and E-cadherin protein loss (unpublished results), thereby making the correct classification of *E-cadherin* mutated cell lines possible. The ability of p120ctn to identify not only truncating *E-cadherin* mutants, but also in-frame *E-cadherin* mutants makes p120ctn a better protein marker for the classification of lobular breast cancer than E-cadherin. Therefore, pathologists in doubt of classifying a breast carcinoma as either lobular or ductal should strongly consider using p120ctn as a routine marker in the classification of lobular breast cancer.

The identification of distinct modes of E-cadherin inactivation challenges the paradigm that mutation and hypermethylation of a tumor suppressor gene are two means to the same end. Hypermethylation of *E-cadherin*, but not mutation of *E-cadherin*, associated with expression of the transcriptional repressors of E-cadherin and with a spindle cell morphology (chapter 3). Expression of the transcriptional repressors and a spindle cell morphology has been associated with epithelial to mesenchymal transition (EMT) and loss of E-cadherin is regarded as one of the hallmarks of EMT [6]. EMT involves the transdifferentiation of epithelial cells that express E-cadherin to cells with a mesenchymal phenotype that no longer express E-cadherin. Recently, induction of EMT by ectopic expression of Snail or Twist in immortalized human mammary epithelial cells was shown to generate cells with properties of stem cells, including acquisition of a mesenchymal phenotype, expression of stem cell markers and an increased ability to form mammospheres [7]. Our results that *E-cadherin* hypermethylation associated with expression of the transcriptional repressors and with a spindle cell morphology would therefore imply that EMT is associated with hypermethylation of the *E-cadherin* promoter, but not mutation of the *E-cadherin* gene. Moreover, we found that loss of E-cadherin through epigenetic mechanisms was not causal for the spindle cell morphology, nor was E-cadherin protein expression lost in all *E-cadherin* hypermethylated spindle cell lines (chapter 3). It is likely that inactivation of *E-cadherin* by hypermethylation is only a secondary effect of the EMT process, that may rather reflect the differentiation state of the cell than an oncogenic event that drives tumorigenesis. Importantly, our results question whether loss of E-cadherin expression is a hallmark of EMT.

The differential β-catenin protein expression pattern of genetically and epigenetically-inactivated *E-cadherin* breast cancer cell lines also is of interest. β-Catenin protein expression,

but not transcript expression, was absent or decreased in all *E-cadherin* mutant cell lines that had deleted E-cadherin's β-catenin binding domain (chapter 3). Most likely, failure of β-catenin proteins to interact with E-cadherin results in phosphorylation and ubiquitylation of the β-catenin pool and subsequent degradation by the APC/GSK3β destruction complex. Consistent with this notion, we observed restoration of β-catenin protein expression in SK-BR-3 cells upon reconstitution of E-cadherin's β-catenin binding domain. In contrast, β-catenin proteins were expressed in all breast cancer cell lines with epigenetic silencing of *E-cadherin* (chapter 3). Indeed, it recently was shown that several aspects of EMT required β-catenin expression, and that this β-catenin pool was largely unphosphorylated [8]. It appears that β-catenin has a dual role in cell adhesion and in EMT, similar to E-cadherin. The recent identification of frequent mutations in the Wnt pathway members β-*catenin*, *APC* and *WISP3* among metaplastic breast cancers — the pathological breast cancer subtype that we have associated with EMT — may provide a genetic basis for β-catenin's role in EMT [9]. Interestingly, we have found no evidence for canonical Wnt pathway activation in the spindle breast cancer cell lines [10]. However, it is still possible that non-canonical Wnt signaling is involved in EMT through expression regulation of E-cadherin's repressor Snail by GSK3β [11, 12].

Genetic inactivation of *E-cadherin* has been observed only in lobular breast cancers and has been shown to be causal for this subtype of breast cancer (chapter 3 and [13-15]). This implies that mutation of *E-cadherin* evokes a cell morphology change which we had expected to involve downstream signaling. However, we could not measure the biological effect of *E-cadherin* gene mutation by gene expression analysis of *E-cadherin* wild-type epithelial cell lines versus *E-cadherin* mutant rounded cell lines (chapter 4). This suggests that either gene expression changes upon mutation of *E-cadherin* are very small, the downstream pathway affects only protein expression, kinase activity and/or cellular localization, or the tumor suppressive effect of *E-cadherin* mutations is solely based on loss of cell-cell adhesion. Alternatively, genes in a putative downstream signaling pathway of E-cadherin could be mutated in the wild-type *E-cadherin* cell lines with an epithelial phenotype. In this scenario, mutation of this particular gene (or genes) would have no impact on cell morphology as it signals downstream of the cytoskeleton in the E-cadherin pathway. If this premise is accurate, gene expression profiling of a mutant *E-cadherin* breast cancer cell line versus its isogenic cell line reconstituted with wild-type *E-cadherin* cDNA should reveal differential gene expression. Importantly, it would render the E-cadherin pathway far more important in breast cancer than currently thought.

Mutation profiles among cancer signaling pathways

According the pathway theory [16], mutations of genes from the same signaling pathway are expected to occur in a mutually exclusive fashion because the tumorigenic clone would not acquire a selective survival advantage upon mutation of a second gene in a pathway. In chapter 4 we have identified mutations of α-*catenin* in four breast cancer cell lines, suggesting that α-*catenin* is a new tumor suppressor gene. Strikingly, all cell lines with the rounded cell

morphology had mutations in either α-*catenin* or *E-cadherin*. This mutually exclusive mutation pattern was consistent with our observation that *E-cadherin* gene mutations were causal for the rounded cell morphology of breast cancer cell lines (chapter 3). However, one α-*catenin* mutant (MDA-MB-157) did not have a rounded cell morphology but a spindle cell morphology and had loss of E-cadherin through promoter hypermethylation. Thus, MDA-MB-157 has inactivated both E-cadherin pathways: the E-cadherin mutation pathway by α-*catenin* gene mutation and the E-cadherin epigenetic pathway in association with *E-cadherin* promoter hypermethylation. The dominant morphological phenotype of the E-cadherin epigenetic pathway was consistent with E-cadherin reconstitution experiments that we had performed, where the spindle cell morphology of the *E-cadherin* hypermethylated cell line MDA-MB-231 was not affected by reconstitution with wild-type *E-cadherin* cDNA (chapter 3). The expression of E-cadherin proteins in some of the *E-cadherin* hypermethylated spindle cell lines also suggested that inactivation of *E-cadherin* by hypermethylation is secondary to the spindle cell morphology and the distinct differentiation program that goes along with this genotype. The presence of an α-*catenin* mutation in a cell line with spindle cell morphology and hypermethylation of *E-cadherin* thus only further substantiates our finding that genetic inactivation and epigenetic inactivation of the E-cadherin pathway are biologically distinct.

Mutual exclusiveness of gene mutations among different cancer signaling pathways may also reflect functional cross talk between these pathways. For example, we found that *RAS* and *BRAF* mutations were mutually exclusive with *PTEN* and *PIK3CA* mutations in breast cancer cell lines (chapter 5). However, in melanomas and colorectal cancers, mutations of the RAS and PI3K pathways were not mutually exclusive. In melanomas, mutations of *PTEN* coincided with *BRAF* mutations, but not *NRAS* mutations [17], whereas *PIK3CA* mutations coincided with *KRAS* or *BRAF* mutations in colorectal cancers [18]. Together, these observations suggest that RAS and BRAF signal through the PI3K pathway in breast cancers but not in melanomas and colorectal cancers. It thus seems that there is a distinction in upstream activators and downstream effectors of the PI3K signaling pathway that is inherent to the different tissues.

Mutual exclusiveness of gene mutations in genes from the same signaling pathway not necessarily implies similar biological effects. Consistent with the pathway theory, we found that mutations of the RB pathway genes *RB1*, *p16* and *Cyclin D1* were mutually exclusive (chapter 7). However, we observed a duality among cell lines with mutational activation of the RB pathway by *RB1* or *Cyclin D1* mutations. *RB1* inactivation was found in ER-negative cell lines while *Cyclin D1* amplification was found in ER-positive cell lines. Importantly, this association with ER status had also been observed in primary breast cancers [19, 20]. The preference of these two RB pathway genes for either ER-negative or ER-positive breast cancers suggests that there is a fine functional distinction between mutation of each gene. This has not only been observed for different genes within a signaling pathway, but also for distinct inactivation mechanisms within a single gene. For example, *E-cadherin* mutation and hypermethylation, although occurring mutually exclusive, associate with distinct subtypes of breast cancer, implying distinct

biological pathways (chapter 3). Similarly, *p16* mutation appears to be mutual exclusive with mutation of *Cyclin D1* and *RB1*, whereas hypermethylation of *p16* concurrently occurs in *Cyclin D1* mutants (chapter 7). Importantly, the pathway theory does not hold true if mutual exclusive mutation patterns associate with distinct biological pathways.

Concurrent occurrence of mutations in genes from the same biological pathway also may imply distinct biological functions of the genes. By mutational analysis of the p53 pathway members, we found that mutations of *CHEK2* and *p53* were not present in the same breast cancer cell lines, although they often occurred concurrent with mutations of *p14ARF*, *c-MYC* or *HDM2* (chapter 7). These results suggested that *p14ARF*, *c-MYC* or *HDM2* mutations have a distinct or additional biological function from *p53* or *CHEK2* mutations. This is supported by the finding that triple knock-out mice lacking functional p53, HDM2 and p14ARF proteins developed multiple tumors at a greater frequency than mice lacking functional p53 and HDM2 or p53 alone [21]. It is possible that mutation of multiple genes within the same pathway has an additive effect in carcinogenesis. In this respect, it is of interest that all nine *p53/c-MYC* and *p53/p14ARF* double mutants had *p53* missense mutations. One could envision that *p53* missense mutations are less deleterious than truncating or deletion mutations and that the concurrent occurrence of mutations amongst genes from the p53 pathway reflects an additive functional effect instead of distinct biological functions of the genes. Similarly, mutations of *p53* are more frequently observed among *BRCA1* and *BRCA2* mutant tumors. Interestingly, the functional effects of *p53* mutations observed in *BRCA1* and *BRCA2* mutant tumors tend to be distinct from those observed in sporadic cases, suggesting an additive effect of *BRCA1* and *BRCA2* mutations on a *p53* mutation [22]. Indeed, all four *BRCA1* mutant breast cancer cell lines also had mutations of the *p53* gene (chapter 6). We also found double mutants among the PI3K, RAS and RB pathways that harbor a strong oncogenic mutation together with a less oncogenic mutation, likely as a result from either heterozygosity or the functional effect of the mutation (chapters 5 and 7). We observed double mutants having two mutations in one single gene, as well as double mutants harboring mutations in two different genes from the same tumorigenic pathway. Of course, the assumption then is that the less oncogenic mutation arose first in the tumor. Concurrent occurrence of mutations in genes within the same signaling pathway may thus not necessarily imply distinct functions of these genes.

Thus, the pathway theory of mutual exclusive mutation patterns as a result of similar biological effects may not always hold true. Genes that function in the same tumorigenic pathway may have (slightly) different biological effects when these genes specifically associate with distinct cancer subtypes. This does not only apply to different genes from a signaling pathway but also to different inactivation mechanisms within a single gene. In addition, concurrent occurrence of mutations in two genes from a signaling pathway may not always imply distinct functions of these genes. Large scale cancer gene analysis as the one conducted in chapter 7 of this thesis may therefore contribute to the unraveling of molecular mechanisms of tumorigenesis.

Mutation profiles among luminal and basal-type breast cancers

ER expression has for years been the major classifier among breast cancers and breast cancer patients are treated according their tumor's ER status. In concordance, ER also was a major discriminator in gene expression profiling [23-31]. ER-positive tumors were mainly luminal tumors, whereas ER-negative tumors were mostly of the normal-like, basal-like or ERBB2+ intrinsic subtypes. However, the subdivision among ER was not perfect, with a minority of ER-positive tumors being of the "ER-negative" intrinsic subtypes, and vice versa. Our Pearson correlation based on the gene expression profiles of the breast cancer cell lines identified two clusters of cell lines that had an immense difference in their differentiation programs, largely associating with ER status. Additional characterizations revealed that the two clusters resembled normal luminal and basal epithelial cells. Importantly, mutations of *E-cadherin* were present in luminal-type breast cancer cell lines, whereas hypermethylation of *E-cadherin* associated with basal-type breast cancers. To be precise, all *E-cadherin* hypermethylated cell lines with loss of E-cadherin protein expression were of the normal-like intrinsic subtype, whereas those of the basal-like intrinsic subtype had retained expression of E-cadherin proteins. Apart from their differential E-cadherin protein expression, we could also distinguish normal-like from basal-like cell lines by their loss of expression of both luminal and basal cytokeratins. However, expression of vimentin, N-cadherin and the epithelial growth factor receptor (EGFR) was found in both basal-like and normal-like cell lines. In addition, we identified a so-called basal mutation profile for both normal-like and basal-like cell lines, involving deletion of *p16* and *p14ARF* and mutation of *RB1*, *BRCA1*, *RAS* and *BRAF*. Together, these results suggested that breast cancers of the basal-like and normal-like intrinsic subtypes constitute two ends of a spectrum of basal-type breast cancers. We also identified a luminal mutation profile for luminal and ERBB2+ intrinsic subtype breast cancer cell lines. The luminal mutation profile involved mutation of *E-cadherin* and *MAP2K4*, and amplification of *ERBB2*, *Cyclin D1* and *HDM2*. Thus, the ERBB2+ and luminal intrinsic subtypes of breast cancer also appear to constitute two ends of a spectrum, with amplification of *HDM2* and mutation of *MAP2K4* being the discriminatory events. We therefore propose that there exist two major subtypes of breast cancer: the luminal and basal-types, and that these are subdivided in the luminal and ERBB2+ intrinsic subtypes and the basal-like and normal-like intrinsic subtypes, respectively.

The identification of two distinct mutation profiles provides a genetic basis for luminal and basal-type breast cancers and aids our understanding of breast tumorigenesis. Importantly, the mutation profiles suggest three plausible scenarios for breast tumorigenesis. In the first "transdifferentiation" scenario, all breast cancers arise from a luminal lineage-restricted progenitor cell, initially all as luminal-type breast cancer. Breast cancers of the basal-like intrinsic subtype arise from luminal-type breast cancers by an EMT-like transdifferentiation. Normal-like breast cancers then represent cancers with fully completed EMT, defined by loss of E-cadherin expression and loss of luminal and basal cytokeratin expression. We found that mutations of *p53*, *PIK3CA* and *PTEN* are not associated with a particular breast cancer subtype and they

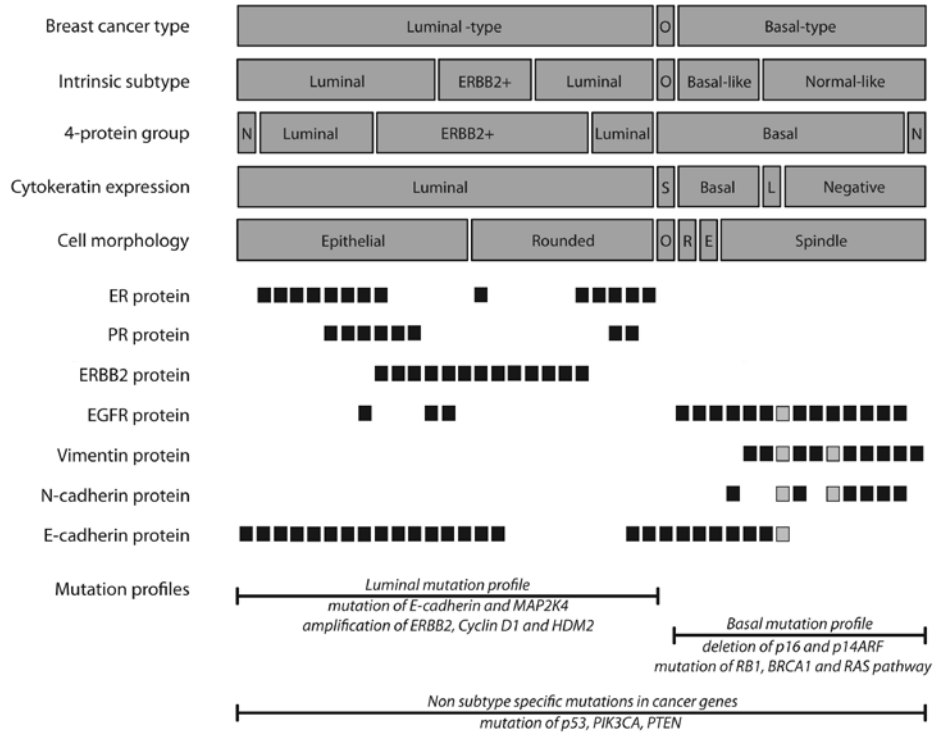


Figure 9.1 Schematic representation of the molecular characterization of 41 breast cancer cell lines. Breast cancer cell lines were classified according breast cancer type, intrinsic subtype, 4-protein group, cytokeratin expression and cell morphology. O, other; N, negative; S, stem cell-like; L, luminal; R, rounded; E, epithelial. Protein expression of ER, PR, ERBB2, EGFR, Vimentin, N-cadherin and E-cadherin is indicated by black boxes. Gray boxes represent cell lines that have not been tested but are anticipated to express a protein.

would therefore represent early events in breast tumorigenesis, whereas mutations of subtype-specific genes necessarily occur after the tumor becomes dedicated either to remain luminal or to transdifferentiate to the basal-type. Alternatively, subtype-specific mutations may be causal in determining whether the tumor remains luminal-type or becomes basal-type. EMT would then be driven by mutations of genes from the basal mutation profile, implying that these mutations regulate the activation and inactivation of EMT-associated pathways. In the second "distinct cell of origin" scenario, luminal and basal-type breast cancers arise from different cell lineages: luminal-type breast cancers arise from normal luminal epithelial cells or from a lineage-restricted luminal progenitor cell in the breast, whereas basal-like intrinsic subtype breast cancers arise from normal basal epithelial cells or from a lineage-restricted basal progenitor cell in the breast. Normal-like intrinsic subtype breast cancers then again are dedifferentiated basal-like breast cancers. In the "distinct cell of origin" scenario, the subtype-specific mutations may reflect necessities for malignant transformation in each of the different cell lineages. In contrast, mutations of *p53*, *PTEN* and *PIK3CA* would be necessary for tumorigenesis in all cell

lineages. In the third "common progenitor cell" scenario, both luminal and basal-type breast cancers arise directly from a single cell, being either a bi-potential progenitor cell or a stem cell in the breast. In this scenario, subtype-specific mutations would be causal in determining whether the tumor becomes luminal-type or basal-type. In contrast, mutations of *p53*, *PTEN* and *PIK3CA* would represent driving events all breast cancers.

Experimental evidence for any of these three scenarios for human breast tumorigenesis is limited. We have shown that metaplastic breast cancers are of the basal-type (chapter 3), even though they frequently also contain a luminal component. Importantly, the distinct components within metaplastic breast tumors have been shown to be clonally related by gene mutation analysis [32-34]. The coexistence of both luminal and basal components within a single breast tumor seems to argue against the "distinct cell of origin" and "common progenitor cell" scenarios and to favor the "transdifferentiation" scenario. However, the rarity of metaplastic breast cancers (about 1% of all breast cancers) does not seem to support transdifferentiation as a general mechanism. In this respect, it should also be realized that it is possible in the "common progenitor cell" scenario that the tumorigenic clone undergoes expansion before committing to luminal or basal-type breast cancer. A recent SAGE profiling study of clinical breast cancers is noteworthy because it showed that both CD24+ and CD44+ components may be present within single breast tumors [35]. Since the CD44 versus CD24 SAGE signature was highly reminiscent to our 1144-gene spindle cell signature, it might be inferred that CD24+ cells are luminal-type cells and CD44+ cells are basal-type. Shipitsin *et al*. thus have provided evidence that luminal and basal components also coexist in other breast cancer subtypes than metaplastic breast cancer, favoring the "transdifferentiation" scenario. However, they also determined that the CD24+ and CD44+ components were genetically identical in some breast tumors but that in other breast tumors the CD24+ component contained additional genetic aberrations, suggesting that these cells had undergone further clonal evolution and rendering the "transdifferentiation" scenario less likely. However, it had not been specified what proportion of breast tumors contained both CD24+ and CD44+ components, nor what proportion of breast tumors had genetically distinct components [35]. To distinguish between the "transdifferentiation" and "common progenitor cell" scenarios, further research should thus focus on determining how often luminal and basal components coexist within breast tumors and whether or not coexistence is restricted to particular pathological subtypes of breast cancer. Either way, based on these data, the "distinct cell of origin" scenario seems an unlikely mechanism in breast tumorigenesis.

In order to distinguish between the "transdifferentiation" and "common progenitor cell" scenarios, it would be helpful to gain insight in the timing of the mutations in the subtype-specific genes as well as the non-subtype-specific genes. For example, if mutations in basal subtype-specific genes prove to be early events in breast tumorigenesis, it would argue against the "transdifferentiation" scenario. The cell of origin of the luminal and basal-types of breast cancer could be revealed through genetic manipulation of the distinct cell lineages in normal breast epithelium, by introducing mutant oncogenes or siRNA-silencing of wild-type tumor suppressor

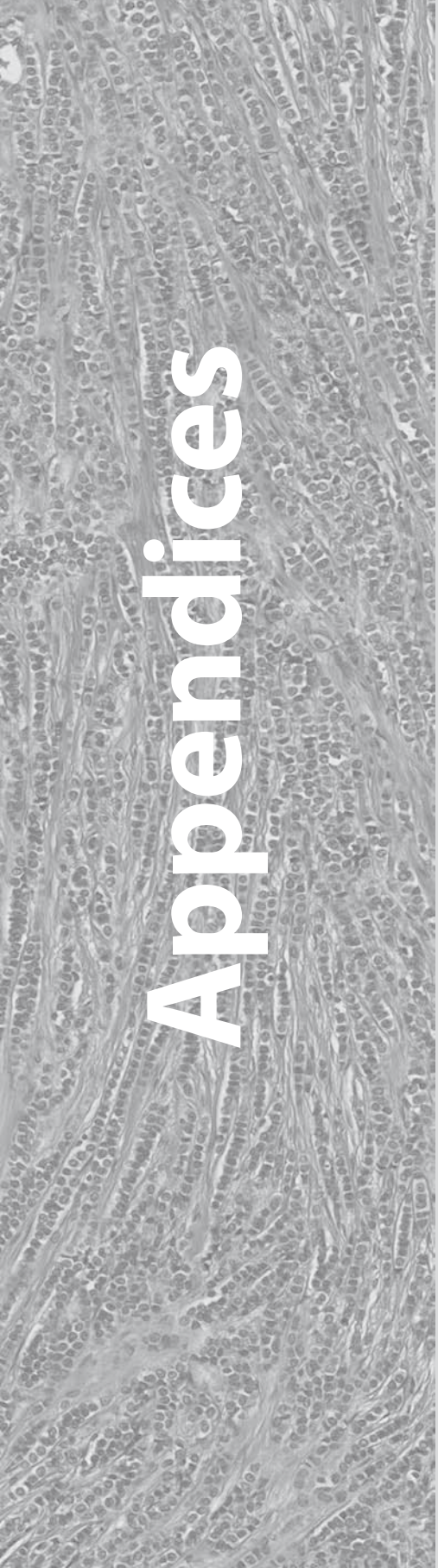
genes from the basal and luminal mutation profiles. However, such experiments might require further characterization of the different cell types present in normal breast epithelium and their cellular hierarchy. Similarly, introducing mutant oncogenes or siRNA-silencing of wild-type tumor suppressor genes of the basal and luminal mutation profiles in either luminal or basal-type breast cancer cell lines, respectively, may reveal whether the "transdifferentation" scenario is involved in breast tumorigenesis. Such experiments may also provide clues as to which of the genes from the basal mutation profile would be causally involved in the transdifferentiation process and whether the ability to transdifferentiate is a property of all or only a subset of luminal-type breast cancer cell lines.

Even without complete understanding of how breast cancers evolve, the mutation profiles identified in this thesis may already allow further refinement of current molecular breast cancer classification and aid the development of new treatment modalities that target the here identified potential drug targets. However, our mutation profiles include only a proportion of the genes mutated in breast cancer and extension of the number of subtype-specific cancer genes, for example with those genes found in recent whole genome screens [36, 37], may draw a more complete landscape of the breast cancer genome. But most importantly, the here identified mutation profiles will need to be confirmed in uncultured, clinical breast cancers.

REFERENCES

1. Frixen UH, Behrens J, Sachs M, et al. E-cadherin-mediated cell-cell adhesion prevents invasiveness of human carcinoma cells. J Cell Biol 1991;113(1):173-85.
2. Vleminckx K, Vakaet L, Jr., Mareel M, Fiers W, van Roy F. Genetic manipulation of E-cadherin expression by epithelial tumor cells reveals an invasion suppressor role. Cell 1991;66(1):107-19.
3. Vos CB, Cleton-Jansen AM, Berx G, et al. E-cadherin inactivation in lobular carcinoma in situ of the breast: an early event in tumorigenesis. British journal of cancer 1997;76(9):1131-3.
4. Nielsen TO, Hsu FD, Jensen K, et al. Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. Clin Cancer Res 2004;10(16):5367-74.
5. Dabbs DJ, Bhargava R, Chivukula M. Lobular versus ductal breast neoplasms: the diagnostic utility of p120 catenin. The American journal of surgical pathology 2007;31(3):427-37.
6. Peinado H, Olmeda D, Cano A. Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? Nat Rev Cancer 2007;7(6):415-28.
7. Mani SA, Guo W, Liao MJ, et al. The epithelial-mesenchymal transition generates cells with properties of stem cells. Cell 2008;133(4):704-15.
8. Onder TT, Gupta PB, Mani SA, Yang J, Lander ES, Weinberg RA. Loss of E-cadherin promotes metastasis via multiple downstream transcriptional pathways. Cancer Res 2008;68(10):3645-54.
9. Hayes MJ, Thomas D, Emmons A, Giordano TJ, Kleer CG. Genetic changes of Wnt pathway genes are common events in metaplastic carcinomas of the breast. Clin Cancer Res 2008;14(13):4038-44.
10. van de Wetering M, Barker N, Harkes IC, et al. Mutant E-cadherin breast cancer cells do not display constitutive Wnt signaling. Cancer Res 2001;61(1):278-84.
11. Zhou BP, Deng J, Xia W, et al. Dual regulation of Snail by GSK-3beta-mediated phosphorylation in control of epithelial-mesenchymal transition. Nature cell biology 2004;6(10):931-40.
12. Yook JI, Li XY, Ota I, et al. A Wnt-Axin2-GSK3beta cascade regulates Snail1 activity in breast cancer cells. Nature cell biology 2006;8(12):1398-406.
13. Berx G, Cleton-Jansen AM, Nollet F, et al. E-cadherin is a tumour/invasion suppressor gene mutated in human lobular breast cancers. The EMBO journal 1995;14(24):6107-15.
14. Berx G, Cleton-Jansen AM, Strumane K, et al. E-cadherin is inactivated in a majority of invasive human lobular breast cancers by truncation mutations throughout its extracellular domain. Oncogene 1996;13(9):1919-25.
15. Derksen PW, Liu X, Saridin F, et al. Somatic inactivation of E-cadherin and p53 in mice leads to metastatic lobular mammary carcinoma through induction of anoikis resistance and angiogenesis. Cancer Cell 2006;10(5):437-49.
16. Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. Nat Med 2004;10(8):789-99.
17. Tsao H, Goel V, Wu H, Yang G, Haluska FG. Genetic interaction between NRAS and BRAF mutations and PTEN/MMAC1 inactivation in melanoma. J Invest Dermatol 2004;122(2):337-41.
18. Velho S, Oliveira C, Ferreira A, et al. The prevalence of PIK3CA mutations in gastric and colon cancer. Eur J Cancer 2005;41(11):1649-54.
19. Bieche I, Lidereau R. Loss of heterozygosity at 13q14 correlates with RB1 gene underexpression in human breast cancer. Molecular carcinogenesis 2000;29(3):151-8.
20. Barnes DM, Gillett CE. Cyclin D1 in breast cancer. Breast cancer research and treatment 1998;52(1-3):1-15.
21. Weber JD, Jeffers JR, Rehg JE, et al. p53-independent functions of the p19(ARF) tumor suppressor. Genes & development 2000;14(18):2358-65.
22. Gasco M, Yulug IG, Crook T. TP53 mutations in familial breast cancer: functional aspects. Human mutation 2003;21(3):301-6.
23. Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. Nature 2000;406(6797):747-52.

24. Gruvberger S, Ringner M, Chen Y, et al. Estrogen receptor status in breast cancer is associated with remarkably distinct gene expression patterns. Cancer Res 2001;61(16):5979-84.
25. Sorlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proceedings of the National Academy of Sciences of the United States of America 2001;98(19):10869-74.
26. Sorlie T, Tibshirani R, Parker J, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proceedings of the National Academy of Sciences of the United States of America 2003;100(14):8418-23.
27. Pusztai L, Ayers M, Stec J, et al. Gene expression profiles obtained from fine-needle aspirations of breast cancer reliably identify routine prognostic markers and reveal large-scale molecular differences between estrogen-negative and estrogen-positive tumors. Clin Cancer Res 2003;9(7):2406-15.
28. Sotiriou C, Neo SY, McShane LM, et al. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. Proceedings of the National Academy of Sciences of the United States of America 2003;100(18):10393-8.
29. Rouzier R, Perou CM, Symmans WF, et al. Breast cancer molecular subtypes respond differently to preoperative chemotherapy. Clin Cancer Res 2005;11(16):5678-85.
30. Yang F, Foekens JA, Yu J, et al. Laser microdissection and microarray analysis of breast tumors reveal ER-alpha related genes and pathways. Oncogene 2006;25(9):1413-9.
31. Wang Y, Klijn JG, Zhang Y, et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. Lancet 2005;365(9460):671-9.
32. Zhuang Z, Lininger RA, Man YG, Albuquerque A, Merino MJ, Tavassoli FA. Identical clonality of both components of mammary carcinosarcoma with differential loss of heterozygosity. Mod Pathol 1997;10(4):354-62.
33. Wada H, Enomoto T, Tsujimoto M, Nomura T, Murata Y, Shroyer KR. Carcinosarcoma of the breast: molecular-biological study for analysis of histogenesis. Human pathology 1998;29(11):1324-8.
34. Wang X, Mori I, Tang W, et al. Metaplastic carcinoma of the breast: p53 analysis identified the same point mutation in the three histologic components. Mod Pathol 2001;14(11):1183-6.
35. Shipitsin M, Campbell LL, Argani P, et al. Molecular definition of breast tumor heterogeneity. Cancer Cell 2007;11(3):259-73.
36. Sjoblom T, Jones S, Wood LD, et al. The consensus coding sequences of human breast and colorectal cancers. Science (New York, NY 2006;314(5797):268-74.
37. Wood LD, Parsons DW, Jones S, et al. The genomic landscapes of human breast and colorectal cancers. Science (New York, NY 2007;318(5853):1108-13.



Appendices

Dankwoord
List of Publications
PhD Portfolio
Color Figures

DANKWOORD

Graag zou ik van deze gelegenheid gebruik willen maken om iedereen te bedanken die op enige wijze heeft bijgedragen aan het tot stand komen van dit proefschrift. Alle collega's van het JNI met wie ik de afgelopen jaren heb mogen samenwerken wil ik graag bedanken voor alle hulp, adviezen en de prettige samenwerking. Vrienden en familie, dank voor jullie belangstelling, steun en gezelligheid. Er zijn echter een aantal mensen die ik graag in het bijzonder wil bedanken.

Mijn copromotor Mieke Schutte, jouw enthousiasme voor wetenschappelijk onderzoek, jouw uitstekende begeleiding en jouw vertrouwen in mij, hebben mij niet alleen op wetenschappelijk vlak, maar ook mij als persoon gevormd tot wie ik vandaag ben. Bedankt voor jouw geduld en de vrijheid die je mij hebt gegeven om mezelf te kunnen ontplooien. Ik vind het een enorme eer dat jij jouw kennis en ervaring met mij hebt willen delen. Jouw inspanning en toewijding maakt dat ik mij geen betere copromotor had kunnen wensen. Dank je wel voor alles!

Beste Professor Klijn, ik wil u hartelijk danken voor uw bereidheid mijn promotor te zijn en voor alle tijd die u geïnvesteerd heeft in het tot stand brengen van dit proefschrift. Ik vind het een eer bij u te mogen promoveren.

Alle leden van de kleine commissie, Prof.Dr. Riccardo Fodde, Prof.Dr. Oosterhuis en Prof.Dr. Devilee, maar ook Prof.Dr. John Foekens, Dr. Els Berns en Dr. Michael den Bakker wil ik hartelijk danken voor het kritisch lezen van het manuscript en het waardevolle commentaar.

Alle leden van de grote commissie, Prof.Dr. Clevers, Dr. Jonkers, Prof.Dr. John Foekens en Prof.Dr. Looijenga wil ik bedanken voor de bereidheid plaats te nemen in de commissie.

De artikelen in dit proefschrift hadden natuurlijk nooit kunnen worden zoals ze zijn zonder de bijdrage van alle co-auteurs en collega's.

Fons Elstrodt en Marijke Wasielewski, we werken al heel wat jaren samen en omdat ieder aan zijn eigen project werkte, hadden wij zo onze eigen taken en specialiteiten in het lab. Hierdoor waren we een goed team, altijd bereid elkaar te helpen. Ik heb met heel veel plezier met jullie samengewerkt. Fons, bedankt voor de gezelligheid. Ik wens jou een hele succesvolle toekomst. Marijke, beide zijn we als stagiair van dezelfde HLO bij Mieke terechtgekomen en uiteindelijk gaan promoveren. Heel veel succes met jouw laatste loodjes en verdere toekomst! Jord Nagel, dank je wel voor alle wetenschappelijke discussies die we samen hebben gevoerd. Wouter Kallemeijn, jij werkte als stagiair aan het sequencen van PTEN, maar ondertussen ben je bijna afgestudeerd en ga je beginnen met een promotieonderzoek. Succes met jouw carrière!

In een lab is het altijd een komen en gaan van mensen. Ik wil iedereen die deel heeft uitgemaakt van het lab bedanken voor de prettige samenwerking.

Anieta Sieuwerts, bij jou kon ik terecht voor al mijn vragen over qPCR, celkweek en hoe jij dat allemaal gedaan hebt toen jij promoveerde. Dank voor het delen van al je kennis en je bereidheid mij te helpen. Mieke Timmermans, als expert op het gebied van immunohistochemie en histologie kon ik altijd bij jou aankloppen voor vragen en advies. Dank je wel! Berthe Bijl, bedankt voor al jouw hulp bij het verzamelen van artikelen. John Martens en John Foekens, bedankt voor jullie wetenschappelijke bijdrage aan de artikelen in dit proefschrift. Ook alle andere collega's van het Interne Oncologie lab wil ik danken voor de prettige samenwerking.

Justine Peeters and Marcel Smid, I want to thank you both for sharing your bio-informatics knowledge with me and the numerous analysis you have performed, which have become an essential part of this thesis.

Michael den Bakker, ook jouw bijdrage aan de inhoud van dit proefschrift mag zeker niet onopgemerkt blijven. Door jouw hulp hebben we uiteindelijk onze bevindingen in cellijnen kunnen valideren in tumoren. Graag wil ik jou bedanken voor het uitvoeren van de vele zoekopdrachten, het zetten van de vele cirkeltjes, het scoren van vele coupes en TMA's en het beantwoorden van de vele vragen die ik had. Ik heb veel van jou geleerd over borstpathologie.

Mijn paranimf Denice Tjon A Fat, bedankt voor jouw gezelligheid, belangstelling en enthousiasme. Fijn om een "grote zus" erbij te hebben. Ik vind het *mááásterlijk* dat jij bereid bent naast mij te staan op deze bijzondere dag in mijn leven.

Mijn ouders. Lieve pap en mam, jullie hebben het vast niet makkelijk gehad met zo'n eigenwijze dochter zoals ik, die alles anders wil doen. Desondanks zijn jullie er altijd voor me geweest en hebben jullie mij altijd gesteund. Bedankt voor alle mogelijkheden die jullie me hebben gegeven. Ik hou van jullie!

Lieve Grace, jouw steun, betrokkenheid, vertrouwen en liefde geven mij de kracht om dingen te doen die voor mij in de eerste instantie te hoog gegrepen lijken. Wij vullen elkaar precies goed aan en samen zijn wij een geweldig "team". Ik ben zo dankbaar dat ik jou in mijn leven mag hebben. Op naar het volgende hoogtepunt in ons leven.

LIST OF PUBLICATIONS

Publications related to this thesis

α-*Catenin* is a putative new tumor suppressor gene.
Antoinette Hollestelle, Fons Elstrodt, Mieke Timmermans, Anieta Sieuwerts, Justine K. Peeters, Jan G.M. Klijn, Peter J. van der Spek, John A. Foekens, Michael A. den Bakker, and Mieke Schutte.
In preparation

Distinct gene mutation profiles among luminal and basal type breast cancer cell lines.
Antoinette Hollestelle*, Jord H.A. Nagel*, Marcel Smid, Suzanne Lam, Fons Elstrodt, Marijke Wasielewski, Ser Sue Ng, Pim J. French, Justine K. Peeters, Marieke J. Roozendaal, Muhammad Riaz, Ellen C. Zwarthoff, Amina Teunisse, Peter J. van der Spek, Jan G.M. Klijn, Stephen P. Ethier, Hans Clevers, Aart G. Jochemsen, Michael A. den Bakker, John A. Foekens, John W.M. Martens, and Mieke Schutte.
Submitted for publication

Epigenetic silencing and mutational inactivation of *E-cadherin* associate with distinct breast cancer subtypes.
Antoinette Hollestelle, Justine K. Peeters, Marcel Smid, Leon C. Verhoog, Pieter J. Westenend, Mieke Timmermans, Alan Chan, Jan G.M. Klijn, Peter J. van der Spek, John A. Foekens, Michael A. den Bakker, and Mieke Schutte.
Submitted for publication

Phosphatidylinositol-3-OH kinase or RAS pathway mutations in human breast cancer cell lines.
Antoinette Hollestelle, Fons Elstrodt, Jord H.A. Nagel, Wouter W. Kallemeijn, and Mieke Schutte.
Mol Cancer Res 2007; 5(2): 195-201

BRCA1 mutation analysis of 41 human breast cancer cell lines reveals three new deleterious mutants.
Fons Elstrodt, Antoinette Hollestelle, Jord H.A. Nagel, Michael Gorin, Marijke Wasielewski, Ans van den Ouweland, Sofia D. Merajver, Stephen P. Ethier, and Mieke Schutte.
Cancer Res 2006; 66(1): 41-45

Other publications

Low-risk susceptibility alleles in 40 human breast cancer cell lines.
Muhammad Riaz, Fons Elstrodt, Antoinette Hollestelle, Jan G.M. Klijn, Mieke Schutte.
Submitted for publication

Exon expression arrays as a tool to identify new cancer genes.
Mieke Schutte, Fons Elstrodt, Linda B.C. Bralten, Jord H.A. Nagel, Elza Duijm, Antoinette Hollestelle, Maartje Vuerhard, Marijke Wasielewski, Justine K. Peeters, Peter van der Spek, Peter A. Sillevis Smitt, Pim J. French.
PloS ONE 2008; 3(8): e3007

Deleterious *CHEK2* 1100delC and L303X mutants identified among 38 human breast cancer cell lines.
Marijke Wasielewski, Pejman Hanifi-Moghaddam, Antoinette Hollestelle, Sofia D. Merajver, Ans van den Ouwenland, Jan G.M. Klijn, Stephen P. Ethier, and Mieke Schutte
Breast Cancer Res Treat 2008; DOI 10.1007/310549-008-9942-3

The *CHEK2* 1100delC mutation identifies families with a hereditary breast and colorectal cancer phenotype.
Hanne Meijers-Heijboer, Juul Wijnen, Hans Vasen, Marijke Wasielewski, Anja Wagner, Antoinette Hollestelle, Fons Elstrodt, Renate van den Bos, Anja de Snoo, Grace Tjon A Fat, Cecile Brekelmans, Shantie Jagmohan, Patrick Franken, Paul Verkuijlen, Ans van den Ouwenland, Pamela Chapman, Carli Tops, Gabriela Moslein, John Burn, Henry Lynch, Jan Klijn, Riccardo Fodde, and Mieke Schutte.
Am J Hum Genet 2003; 72(5): 1308-1314

Low penetrance susceptibility to breast cancer due to *CHEK2* 1100delC in noncarriers of *BRCA1* or *BRCA2* mutations
Hanne Meijers-Heijboer, Ans van den Ouwenland, Jan G.M. Klijn, Marijke Wasielewski, Anja de Snoo, Rogier Oldenburg, Antoinette Hollestelle, Mark Houben, Ellen Crepin, Monique van Veghel-Plandsoen, Fons Elstrodt, Cornelia van Duijn, Carina Bartels, Carel Meijers, Mieke Schutte, Lesley McGuffog, Deborah Thompson, Douglas F. Easton, Nayanta Sodha, Sheila Seal, Rita Barfoot, Jon Mangion, Jenny Chang-Claude, Diana Eccles, Rosalind Eeles, D. Gareth Evans, Richard Houlston, Victoria Murday, Steven Narod, Tamara Peretz, Julian Peto, Catherine Phelan, Hong Xiang Zhang, Csilla Szabo, Peter Devilee, David Goldgar, P. Andrew Futreal, Katherine L. Nathanson, Barbara L. Weber, Nazneen Rahman, and Michael R. Stratton.
Nat Genet 2002; 31(1): 55-59

Mutant *E-cadherin* breast cancer cells do not display constitutive Wnt signaling
Marc van de Wetering, Nick Barker, I. Clara Harkes, Marcel van der Heyden, Nicolette J. Dijk, Antoinette Hollestelle, Jan G.M. Klijn, Hans Clevers, and Mieke Schutte.
Cancer Res 2001; 61(1): 278-284

Book chapters

Representational difference analysis as a tool in the search for new tumor suppressor genes.
Antoinette Hollestelle and Mieke Schutte.
In: Methods in Molecular Medicine, Pancreatic Cancer Methods and Protocols. Humana Press, Totowa, New Jersey, 2004

PHD PORTFOLIO

A summary of PhD training and teaching activities

Name PhD student:	Antoinette Hollestelle
Erasmus MC department:	Medical Oncology
Research school:	Postgraduate school Molecular Medicine
PhD period:	February 2005-December 2008
Promotor:	Prof.Dr. J.G.M. Klijn
Supervisor:	Dr. M. Schutte

1. PhD TRAINING

	Year	Workload
General academic skills		
-Biomedical English writing and communication	2006	3.0 ECTS
Research skills		
-Microsoft Access	2002	0.9 ECTS
-Classical methods for data analysis	2005	5.7 ECTS
-Risk and quality management in laboratories	2005	0.3 ECTS
In-depth courses		
-Molecular medicine	2000	1.9 ECTS
-Oncogenesis and tumor biology	2003	1.5 ECTS
-From development to disease	2003	0.9 ECTS
-Cellular responses to DNA damage and cancer risk	2004	1.5 ECTS
-Analysis of microarray gene expression data	2004	0.6 ECTS
-SNPs and human diseases	2005	1.5 ECTS
-Prognostic research	2005	1.4 ECTS
-Clinical trials and drug risk assessment	2005	1.4 ECTS
-In vivo imaging: from cell to organism	2005	1.5 ECTS
-Comparative gene analysis: from yeast to man	2006	1.5 ECTS

Presentations

-Annual oral presentation at the JNI Scientific Lab meetings	2004-2008	1.5 ECTS
-Oral presentation at the annual Tumor Cell Biology meeting of the Dutch Cancer Society in Lunteren	2005	0.3 ECTS
-Poster presentation at the annual Molecular Medicine Day in Rotterdam	2005	0.3 ECTS
-Poster presentation at the annual meeting of the American Association for Cancer Research (AACR) in Los Angeles, CA, USA	2007	0.3 ECTS

International conferences

-Annual meeting of the American Association for Cancer Research (AACR) in Los Angeles, CA, USA	2007	1.5 ECTS

Seminars and workshops

-Monthly JNI Oncology Lectures	2005-2008	0.9 ECTS
-Annual Tumor Cell Biology meeting of the Dutch Cancer Society in Lunteren	2005	0.6 ECTS
-Annual Molecular Medicine Day in Rotterdam	2005	0.3 ECTS
-Master class with Prof.dr. Mary-Claire King	2006	0.3 ECTS

2. TEACHING ACTIVITIES

	Year	Workload
Supervising practicals and bachelor's thesis		
-Internship of HLO student Wouter Kallemeijn May 2005–February 2006	2005-2006	10.8 ECTS

COLOR FIGURES

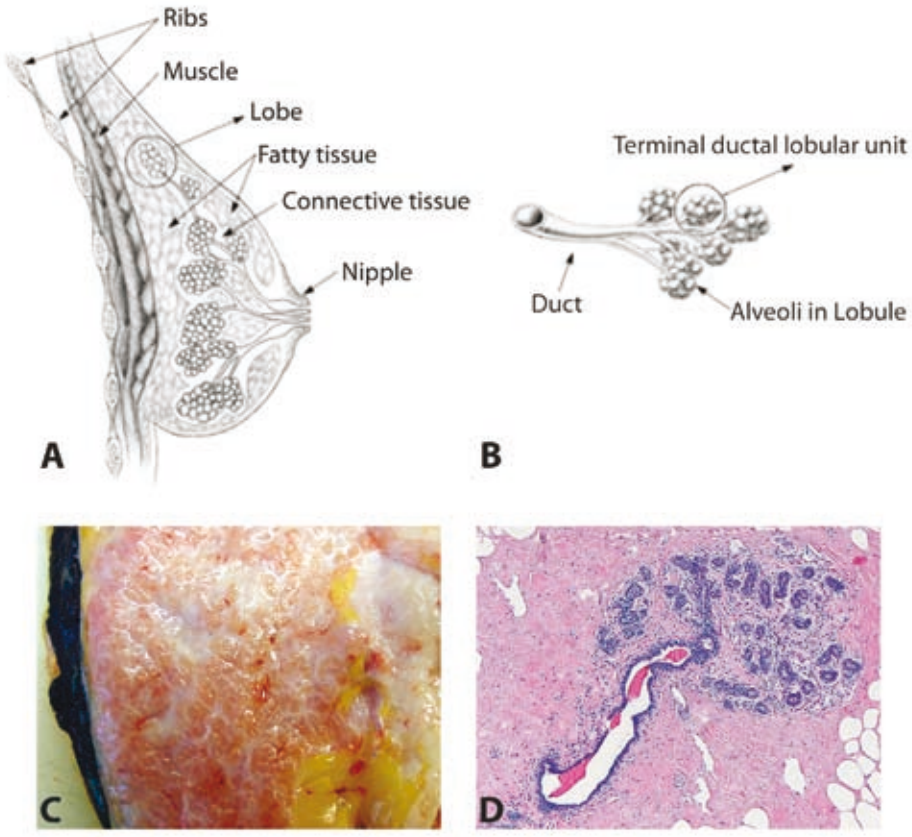


Figure 1.1 (A), schematic representation of the anatomy of the breast; (B), schematic representation of a breast lobe; (C), macroscopic view of the normal breast parenchyma; (D), microscopic view of a normal terminal ductular lobular unit of the breast. (A) and (B) were adapted from http://www.blogsforcompanies.com/TTimages/dcis_in_situ.jpg.

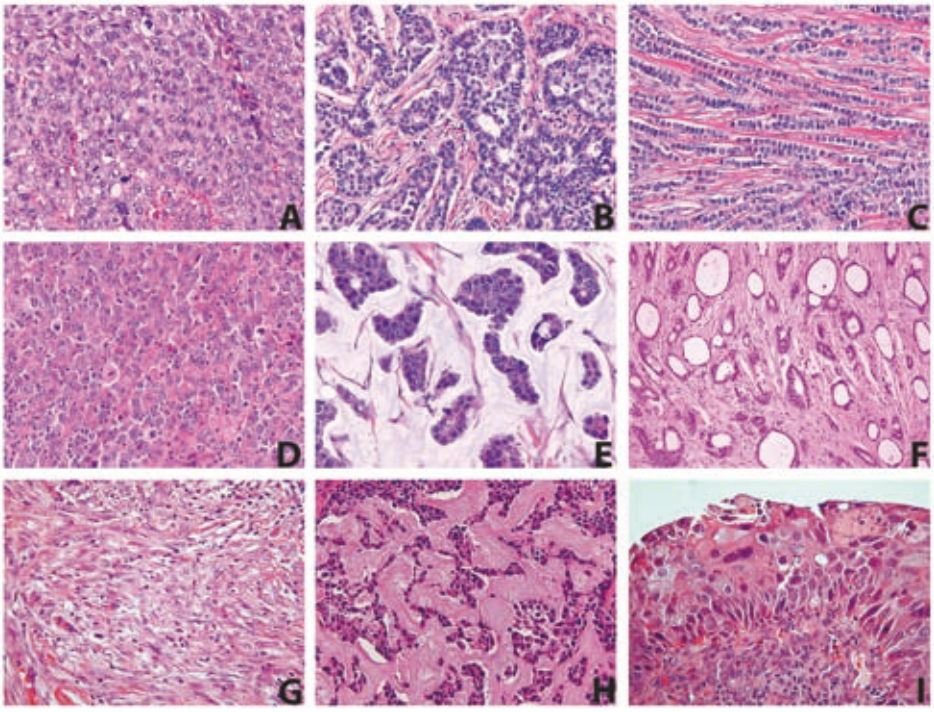


Figure 1.2 Microscopic views of histopathological subtypes of breast cancer. (A) and (B), ductal carcinoma of high and low grade, respectively; (C), lobular carcinoma with strings of cells called "Indian files"; (D), medullary carcinoma; (E), mucinous carcinoma; (F), tubular carcinoma; (G) through (I), metaplastic carcinoma of the breast with spindle, matrix-producing and squamous differentiation, respectively.

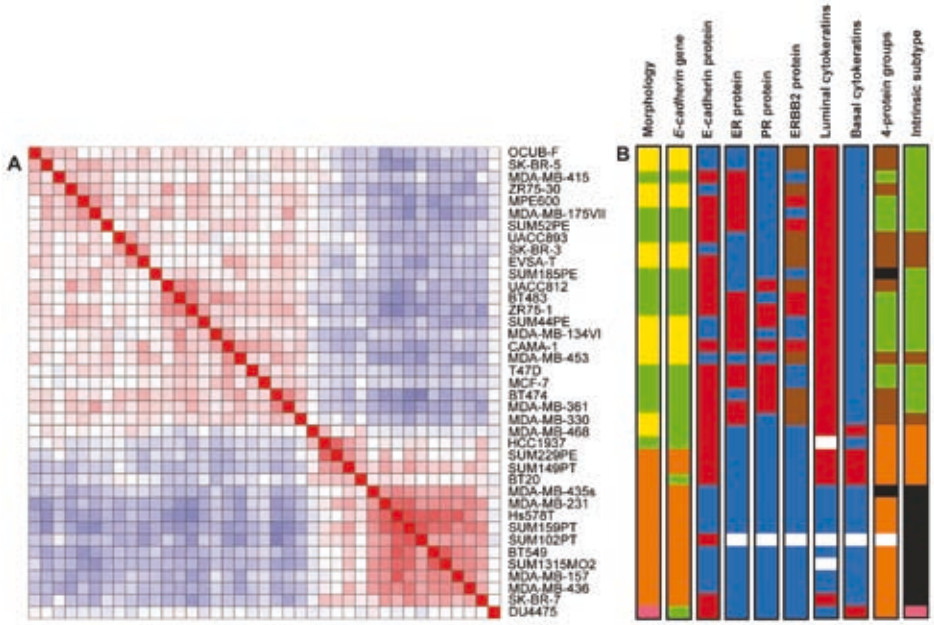


Figure 3.4 Pearson correlation and expression of breast cancer associated proteins in human breast cancer cell lines. (A), Pearson correlation plot based on the log2GM <-2 and >2 subset (5490 probe sets). The Pearson correlation coefficient algorithm positions samples according their overall similarity in gene expression, where red indicates high overall similarity (positive correlation coefficient) and blue indicates low similarity (negative correlation coefficient). (B), Various characterizations of the cell lines indicated that the upper cluster in the Pearson correlation plot contains the epithelial and rounded cell lines intermingled whereas the lower cluster contains all spindle cell lines and a single epithelial cell line and a single rounded cell line (HCC1937 and MDA-MB-468). One cell line was atypical as it did not belong to either of the two clusters (DU4475). The lower cluster included two subgroups that by the intrinsic gene set classified as basal-like and normal-like intrinsic subtypes, where all E-cadherin-negative spindle cell lines classified as normal-like. The lower cluster classified as basal breast cancers by the 4-protein signature of ERBB2, ER, CK5 and EGFR. Color coding morphology column: green, epithelial morphology; yellow, rounded cell morphology; orange, spindle cell morphology; pink, other cell morphology. E-cadherin gene column: green, wild-type E-cadherin gene; yellow, mutant E-cadherin gene; orange, methylated E-cadherin gene. E-cadherin protein, ER protein, PR protein, ERBB2 protein, luminal cytokeratins and basal cytokeratins columns: red, protein expression; blue, no protein expression; brown, protein overexpression. 4-protein groups column: green, luminal group; brown, ERBB2+ group; black, negative group; orange, basal-like group. Intrinsic subtypes column: green, luminal subtype; brown, ERBB2+ subtype; orange, basal-like subtype; black, normal-like subtype; pink, not of any subtype.

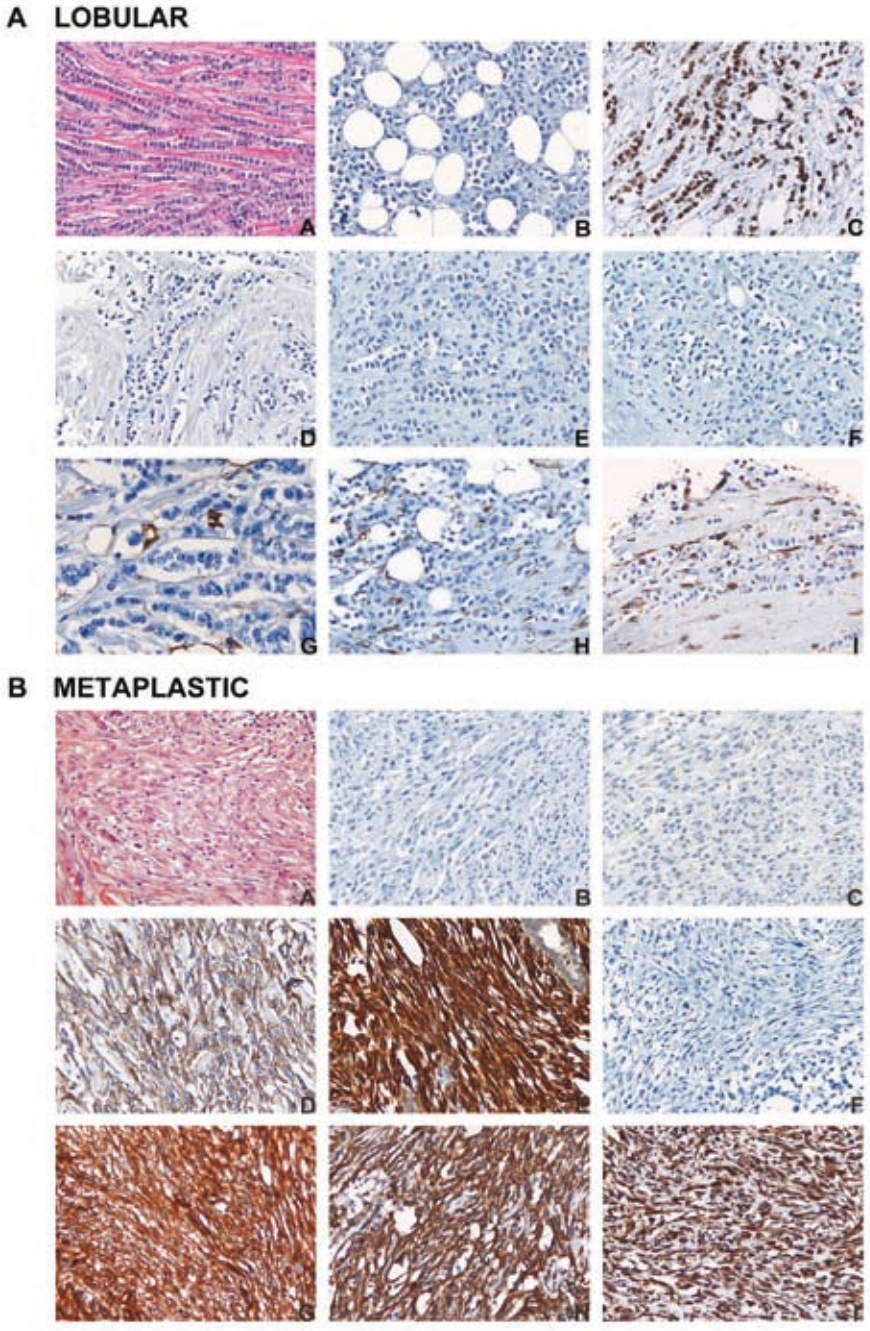


Figure 3.6 Protein expression in clinical breast cancers. Examples of immunohistochemical analysis of (A), lobular breast cancers and (B,) metaplastic breast cancers. Microscopic views: (A), HE-staining; (B), ERBB2; (C), ER; (D), EGFR; (E), CK5; (F), E-cadherin; (G), Caveolin-1; (H), Caldesmon; (I), Vimentin.

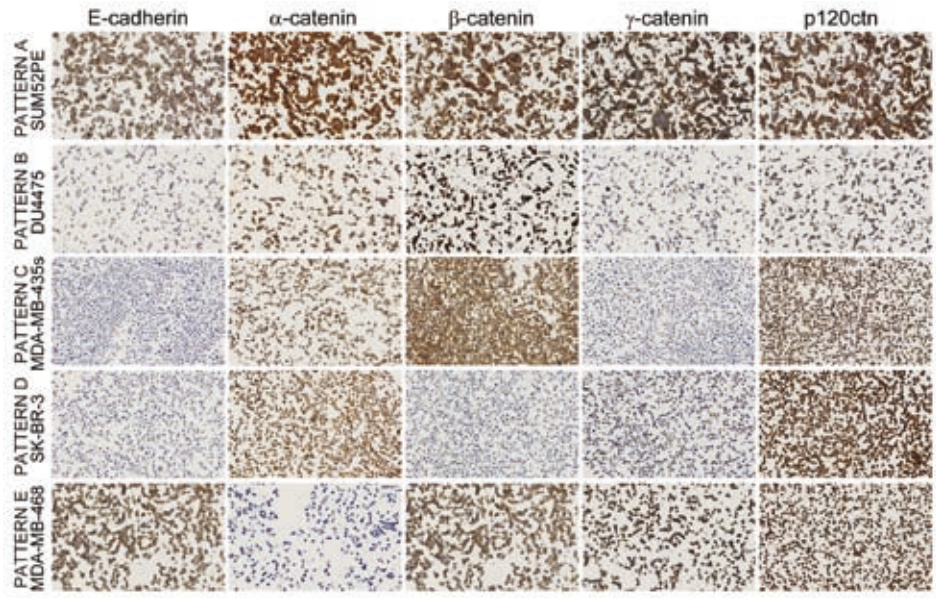


Figure 4.1 Expression analysis of the E-cadherin/catenin protein complex in 41 human breast cancer cell lines. Pattern A entails expression of E-cadherin and all four catenins at apparently normal levels. This pattern is represented by SUM52PE and included fifteen cell lines with wild-type *E-cadherin* genes, five cell lines with (partially) methylated *E-cadherin* genes, and all three cell lines with mutant *E-cadherin* genes that caused in-frame deletions in the encoded proteins. Pattern B is identified for DU4475 only, which had increased β-catenin protein expression, decreased E-cadherin and γ-catenin protein expression and apparently normal α-catenin and p120ctn protein expression. Pattern C entails loss of E-cadherin protein expression, mostly decreased γ-catenin protein expression and mostly normal α-catenin, β-catenin and p120ctn protein expression. Pattern C is represented by MDA-MB-435s and included eight cell lines that had a methylated *E-cadherin* gene promoter. Pattern D entails decreased or absent E-cadherin, β-catenin and γ-catenin protein expression and mostly normal α-catenin and p120ctn protein expression. This pattern is represented by SK-BR-3 and included seven cell lines that had a mutant *E-cadherin* gene that caused a premature termination in the encoded proteins. Pattern E is represented by MDA-MB-468 and included both MDA-MB-468 and MDA-MB-330, which had wild-type *E-cadherin* genes and apparently normal expression levels for all proteins, except for complete absence of α-catenin proteins.

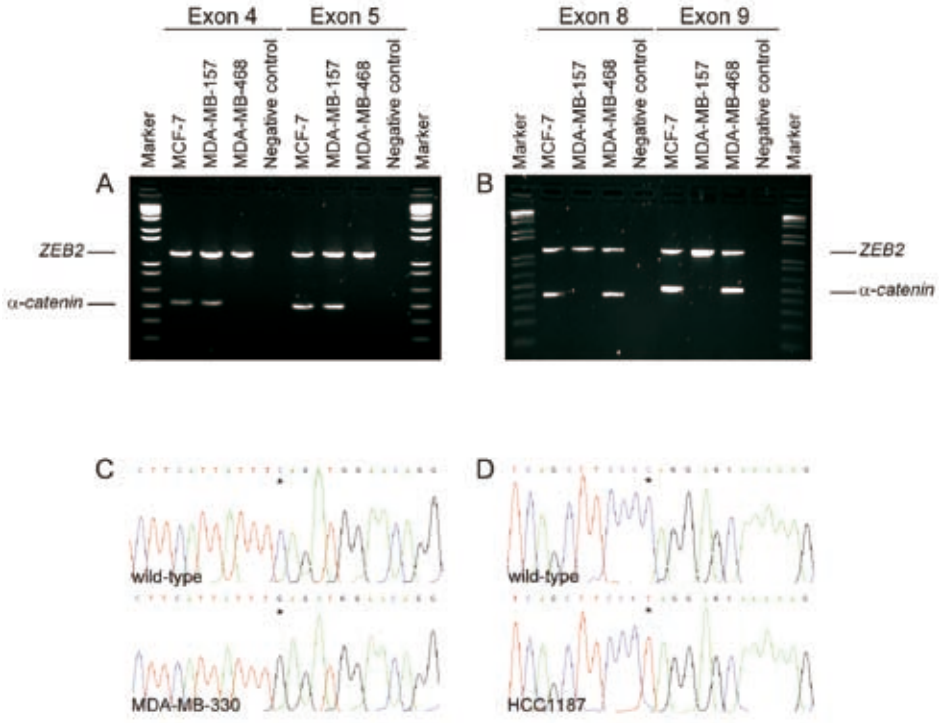


Figure 4.2 Identification of α-*catenin* gene mutations among 55 breast cancer cell lines. (A), homozygous deletion of exons 4 and 5 of α-*catenin* in cell line MDA-MB-468 and (B), homozygous deletion of exons 8 and 9 of α-*catenin* in cell line MDA-MB-157. Duplex PCR amplification products are shown for three breast cancer cell lines, using primers specific for *ZEB2* and α-*catenin* (upper and lower fragments, respectively). Negative control, template negative control; Marker, size marker 1 kb+ DNA ladder (Invitrogen). (C), α-*catenin* nonsense mutation identified in MDA-MB-330 and (D), α-*catenin* nonsense mutation identified in HCC1187. Mutations were identified by PCR amplification and sequencing of genomic DNA (lower electropherograms). The wild-type α-*catenin* gene sequence is shown for comparison (top electropherograms).

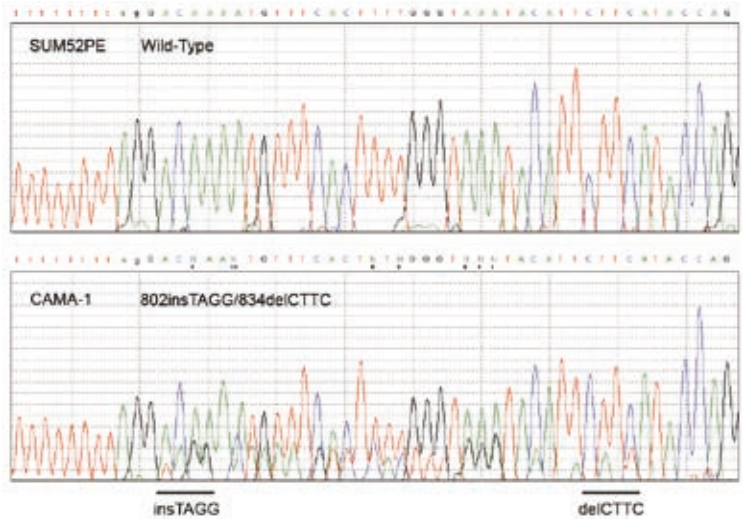


Figure 5.1 Identification of the *PTEN* 802insTAGG/834delCTTC mutation in cell line CAMA-1 by PCR amplification and sequencing of genomic DNA (bottom electropherogram). The wild-type *PTEN* gene sequence is shown for comparison (top electropherogram).

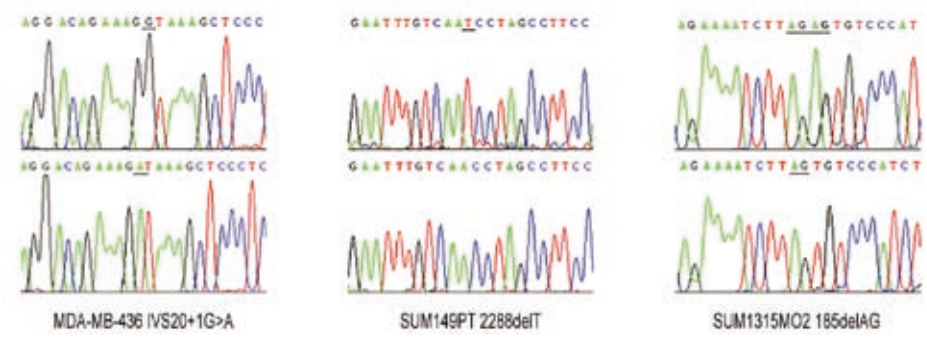


Figure 6.1 Identification of three new *BRCA1* mutant breast cancer cell lines by PCR amplification and direct sequencing. Top, electropherograms displaying the wild-type sequence. Bottom, electropherograms displaying the mutations.

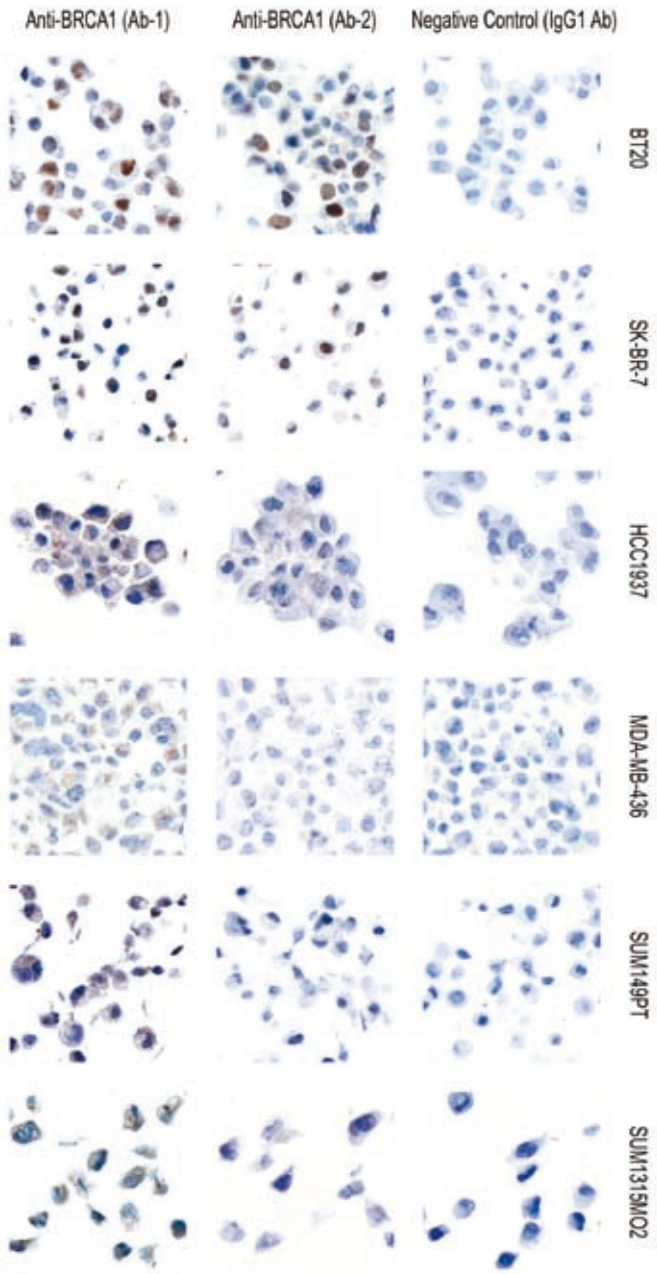


Figure 6.2 BRCA1 immunocytochemistry in *BRCA1* mutant and wild-type breast cancer cell lines. In contrast with the two wild-type cell lines (BT20 and SK-BR-7), none of the four *BRCA1* mutants had nuclear BRCA1 staining with either of the two anti-BRCA1 monoclonal antibodies Ab-1 and Ab-2. There is some cytoplasmic staining of unclear significance in all samples with Ab-1, which is not observed with more diluted Ab-1 antibodies nor with Ab-2 (see also Supplementary Data). The negative control antibody is an IgG1 isotype-matched antibody. Magnification 40X.

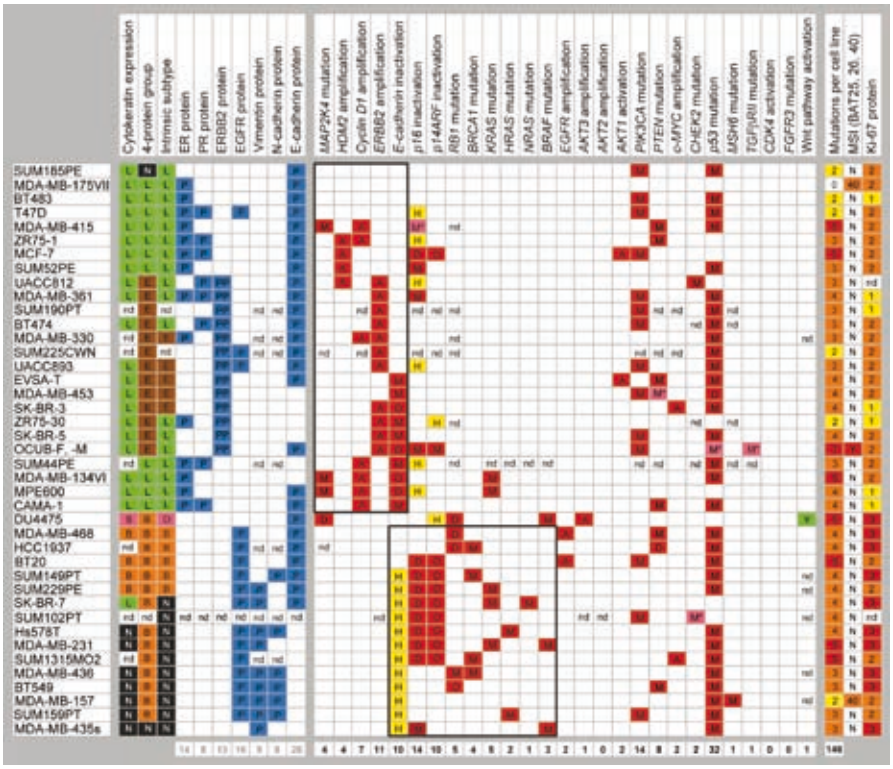


Figure 7.1 Molecular characterization of 41 human breast cancer cell lines. (Left panel) Classification of the cell lines by expression analyses based on cytokeratin proteins, the 4-protein signature of ERBB2, ER, CK5 and EGFR, the intrinsic gene set, and several proteins relevant for breast cancer. Classification by cytokeratin expression: luminal (L), CK8/18+ and/or CK19+, CK5-; basal (B), CK8/18+ and/or CK19+, CK5+; null (N), CK8/18 low, CK19-, CK5-; stem cell (S), CK8/18 low, CK19-, CK5+. Classification by the 4-protein signature: ERBB2 overexpression (E); luminal (L), ERBB2-, ER+; basal (B), ERBB2/ER-, CK5+ and/or EGFR+; negative (N), ERBB2/ER/CK5/EGFR-. Classification by intrinsic gene expression (see also Figure 7.2): ERBB2+ (E); luminal (L); basal-like (B); normal-like (N); other subtype (O). For individual proteins, expression (P) and overexpression (PP) is indicated in blue and absence of expression in white. The protein expression profiles confirmed the remarkably concordant classification by histological criteria or by intrinsic gene expression, and suggested that basal/basal-like and null/negative/normal-like cell lines represent two related subtypes of basal-type breast cancers. (Middle panel) Cancer gene mutation analysis of the cell lines. Genes are indicated at the top and the number of oncogenic mutations identified in each gene at the bottom. Oncogenic mutations (M), sizeable deletions (D) and amplifications (A) are in red; heterozygous oncogenic mutations in tumor suppressor genes (M*) are in pink; and wild-type genes are in white. Promoter hypermethylation (H) is in yellow and constitutive Wnt pathway activation (Y) is in green. nd, not determined. The observed dichotomy among the breast cancer cell lines by protein and gene expression analyses was further supported by the two distinct gene mutation profiles among luminal-type and basal-type breast cancer cell lines (areas with black borders). (Right panel) Number of oncogenic mutations identified in each cell line, microsatellite instability (MSI) and Ki-67 proliferation analysis. MSI: no (N); yes (Y), MSI with BAT 25, 26 and 40; (40), MSI with BAT 40 only. Ki-67 immunohistochemistry: 1, less than 33% of cells positive; 2, 33-66% of cells positive; 3, more than 66% of cells positive. There was a correlation of proliferation rate with breast cancer type (χ^2 $P<0.001$), which was not associated with the number of mutations identified per cell line.

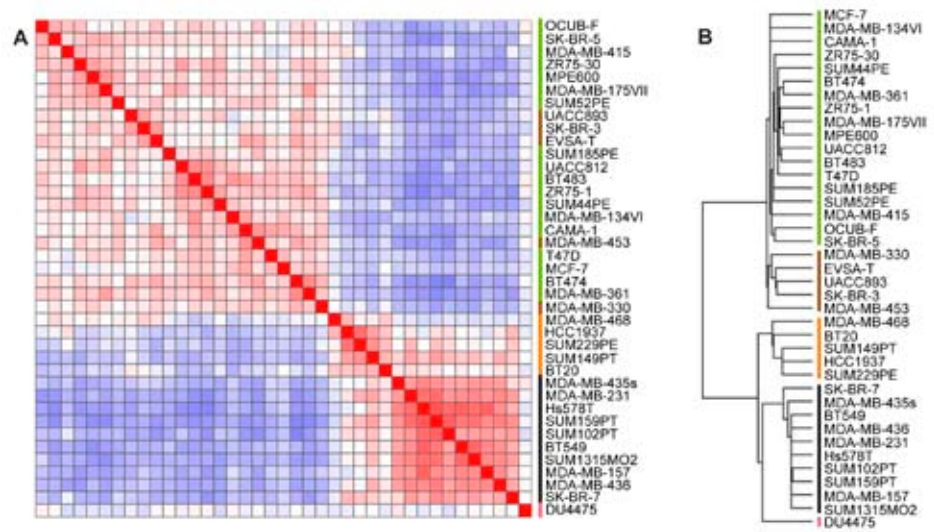


Figure 7.2 Global gene expression and intrinsic subtypes of breast cancer cell lines. (A), Pearson correlation plot of 39 cell lines based on the subset of 5,490 log2GM <-2 and >2 differentially expressed probe sets by Affymetrix U133A microarrays. The Pearson correlation coefficient algorithm positions samples according their overall similarity in gene expression, where red indicates high overall similarity and blue indicates low similarity. Two main clusters of 23 and 15 cell lines were apparent, whereas the DU4475 cell line did not belong to either cluster. (B), Dendrogram of hierarchical clustering of 39 cell lines based on the intrinsic gene set and U133A microarray data. The luminal and ERBB2+ intrinsic subtypes coincided with the major cluster of cell lines in Figure 1A, and the basal-like and normal-like subtypes coincided with the minor cluster of cell lines. Cell line DU4475 did not classify for any of the intrinsic breast cancer subtypes. Color coding of intrinsic subtypes (see also Figure 7.1): green, luminal; brown, ERBB2+; orange, basal-like; black, normal-like; pink, not of any subtype.