Prediction of "BRCAness" in breast cancer by array comparative genomic hybridization

Cover layout "*DNA on the bench*" by SA Joosse

Prediction of "BRCAness" in breast cancer by array comparative genomic hybridization

Proefschrift ter verkrijging van de graad van Doctor aan de Universiteit Leiden, op gezag van Rector Magnificus prof.mr. P.F. van der Heijden, volgens besluit van het College voor Promoties te verdedigen op dinsdag 27 maart 2012 klokke 11:15.

door

Simon Andreas Joosse

geboren te Velsen in 1983

Promotiecommissie

Promotor
prof. dr. Peter Devilee
Copromotor
dr. Petra M. Nederlof
Overige leden
prof. dr. Marc J. van de Vijver
prof. dr. ir. Marcel J.T. Reinders
dr. Jos Jonkers

The work described in this thesis was performed at the Netherlands Cancer Institute – Antoni van Leewenhoek Hospital with financial support from the Dutch Cancer Society (KWF). For the publication of this thesis, financial support from the Dutch Cancer Society is gratefully acknowledged.

The capacity to blunder slightly is the real marvel of DNA. Without this special attribute, we would still be anaerobic bacteria and there would be no music.
- Lewis Thomas

Table of contents

Chapter 1

Introduction

Based on the article: Prediction of BRCA status

SA Joosse and J Hannemann

CML Breast Cancer. 2011 Jul; 23(2): 41-50, Leading Article.

Introduction

Based on the article: Prediction of BRCA status
Simon A Joosse and Juliane Hannemann
CML Breast Cancer, 2011 Jul; 23(2): 41-50, Leading Article

General introduction about cancer

The hallmarks of cancer

In humans, cell growth and development is a continuous process, which is strictly controlled by genes encoded by DNA of the cell nucleus. Two different types of genes regulate the mechanism of cell proliferation, the first are called proto-oncogenes that are positive regulators of cell proliferation and promote cell division and cell growth, and the second are called tumor suppressors that negatively regulate cell proliferation and suppress cell growth. Upon mutation or loss of these genetic regulators, cells may keep on dividing beyond the body's normal needs and cause harm to other tissues and body functions. This is the hallmark of cancer, a malignant growth characterized by uncontrolled, unwanted, purposeless, damaging and continuing growth of cells. These so called tumor cells differ functionally, structurally (anaplasia) and in behavior from the normal cells from which they develop; they have the potential for a limitless and uncontrolled replication independently from external stimulatory signals and uninhibited by antigrowth and apoptosis signals (1). Furthermore, not only is cancer dangerous for its direct surrounding tissue, it can be most harmful upon invasion of the body and establishment of metastasis in distant organs (2, 3). Metastasized cancer is the leading cause of death from cancer and is considered an incurable disease (4). In 2004, around 7.4 million people died of cancer - 13% of all deaths worldwide - of which breast cancer was the most common form of cancer in women (5).

Causes of cancer

Most cancer arises from a single cell that has accumulated DNA damage and genetic mutations to a number of key tumor suppressor genes and proto-oncogenes to escape programmed cell death and induce unlimited replication (1). These DNA mutations are sequentially acquired through time, which is why the likelihood for cancer increases with age. When a mutated predisposition gene is inherited, the risk for cancer is already present at a much younger age. Damage to genes is caused by both exogenous and endogenous factors. The exogenous factors are mainly environmental DNA damaging agents, such as ionizing radiation, ultraviolet rays, air pollution, and inhaled cigarette smoke. The endogenous factors include reactive oxygen species produced from cellular metabolism and replication errors of several cellular processes, such as DNA duplication and meiotic recombination. Although the risk for cancer by endogenous factors is difficult to control, over 30% of all cancer can be prevented by adopting a healthy life-style, *i.e.* not using tobacco, having a healthy diet, preventing infections that may cause cancer and being physically active (5, 6).

Classification of cancer

Cancer is initially named after the site of the body from which it originates. Next, tumors are divided into groups based on the cell of origin; carcinomas are the largest group of solid tissue malignant tumors and are of epithelial origin, such as skin, colon or mammary ducts. For that reason, tumors growing in the breast are mainly classified as "breast carcinomas". Because cancer can arise from all organs and different cell types it is a complex disease consisting of many diverse entities that all have their own unique characteristics and behavior. Since the application of radiotherapy and chemotherapy in the first half of the 20th century as anti-cancer treatment, it has become clear that different tumors can respond differently to different types of therapy (7). This has motivated researchers to find tumor markers that allow for the identification of therapeutic groups to predict prognosis and adapt therapy to the clinical situation, a process which is called tumor classification. For many years now, the most important factors in tumor classification are site, degree of local and remote invasion (staging), histological type, cell structure (grading) and site dependent tumor markers (8).

Table 1. *Breast cancer incidence per 100,000 women in Northern America, Northern and Western Europe and Australia/New Zealand. Data source: GLOBOCAN 2008, International Agency for Research on Cancer http://globocan.iarc.fr/*

Age (years)	Incidence
0-14	0.0
15-39	19.1
40-44	115.0
45-49	172.7
50-54	218.2
55-59	262.5
60-64	304.6
65-69	349.9
70-74	338.9
75+	341.3

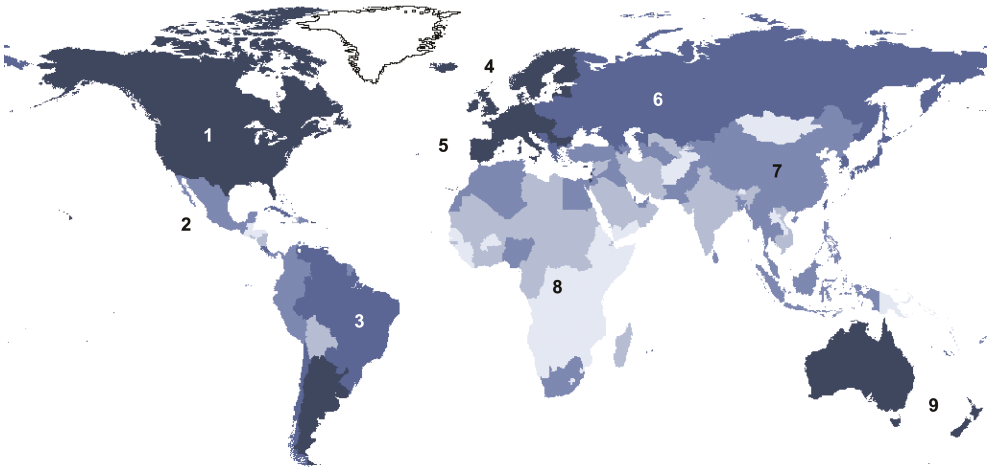
Breast cancer

Prevalence

Breast cancer accounts for approximately one-forth of all cancers in women worldwide, making it the most common female malignancy. As a result of early detection and improved treatment, death rates from breast cancer have been steadily decreasing; however, breast cancer is still the leading cause of cancer-related death in women, closely followed by lung cancer. Figure 1 depicts the worldwide breast cancer prevalence; as can be seen, the incidence of breast cancer is the highest in economically developed countries, *i.e.*, in Europe, Australia and North America, where it accounts for approximately half of all the breast cancer cases worldwide. The incidence of breast cancer in Africa or Asia is about six-fold lower than that. The disease is not common until after the age of 40 and the incidence increases with age (Table 1). The average age of women to be diagnosed with breast cancer is between 60-61 years (5, 9, 10). Less than one percent of all breast cancers occur in men (11).

Histological classification of breast cancer

A woman's breast consists of milk glands (*lobules*), tubes for transporting milk from the glands to the nipple (*ducts*), fatty and connective tissue, blood vessels, and lymph vessels (Figure 2). Pathologically, breast cancer can be divided into two main and several uncommon types. The most common type is called ductal carcinoma



	Incidence		Incidence		Incidence
1) Northern America	117.6	4) Northern and Western Europe	150.4	7) Asia	26.4
2) Central America	23.1	5) Southern Europe	116.5	8) Africa	16.8
3) South America	45.4	6) Central and Eastern Europe	73.5	9) Australia/New Zealand	126.5

Figure 1 - Breast cancer prevalence. *Age-standardized breast cancer incidence per 100,000 women by world area in 2008 (crude rate statistic). Image and data source: GLOBOCAN 2008, International Agency for Research on Cancer http://globocan.iarc.fr/*

and is thought to be derived from the breast ducts; the second type is called lobular carcinoma and is deemed to have arisen from the breast lobules at the end of the ducts. However, there is no true evidence for the site of origin (duct or lobules) and classification is made on histological parameters (12). Other, less common types of breast carcinoma are tubular, invasive cribriform, medullary, mucinous, invasive papillary, invasive micropapillary, apocrine, metaplastic, glycogen-rich clear cell, lipid-rich, adenoid cystic, acinic cell, Paget's disease of the nipple, and inflammatory carcinoma. As long as a carcinoma is still growing within the ductulo-lobular system of the breast, it is called carcinoma *in situ*. Based on the histological properties, ductal carcinoma in situ (DCIS) or lobular carcinoma in situ (LCIS) can be distinguished. If the carcinoma shows evidence of breaching the basement membrane and thereby infiltrating the adjacent stroma, the tumor is classified as invasive ductal carcinoma (IDC) or invasive lobular carcinoma (ILC). To be able to estimate prognosis of breast cancer, carcinomas are generally classified by the TNM classification system and by a grading system. In the TNM system, three different parameters are assessed (Table 2). The first parameter, designated *T*, is the size or extension of the primary tumor starting from 0 (no primary tumor) to 4 (exceeding adjacent structures). The *N* parameter concerns the rate of invasion of regional nodes scored from 0 (no metastasis) to 3 (invasion beyond regional nodes). Last, the *M* component describes the presence of metastasis where M0 classifies as no remote metastasis and M1 the presence of metastasis at distant site(s). Besides size, the primary tumor is graded as a

measure for its rate of growth and cell abnormality; the most used system for this is the Bloom-Richardson-Elston grading system. By assessing tubule formation, nuclear polymorphism and mitotic rate, a carcinoma can be graded as grade 1: well differentiated; grade 2: moderately differentiated; or grade 3: poorly differentiated (13). In the past decade, new methods have been developed to classify breast cancer based on molecular characteristics; this will be discussed further on.

Detection of breast cancer

Breast cancer can be detected by breast self-examination which, is best performed every 3 months at the end of the menstrual stage. Because breast self-examination will only detect palpable breast cancer or breast cancer in advanced stage, it does not provide early diagnosis; consequently, breast self-examination has not proven to reduce mortality in women (14). Because breast cancer has no obvious symptoms at early stage in general, the first sign of breast cancer is often an abnormality detected on a mammogram. Mammography is currently the most widely used imaging technique for early detection of breast cancer. It has been demonstrated that periodically screening by mammography can decrease breast cancer death rates; however, this technique has its limitations and does not detect all breast cancers (15). After a suspicious abnormality has been detected by mammography, additional tests such as ultra sound are performed or a needle aspiration is taken. In the presence of a highly suggestive lesion, surgery immediately follows mammography without any additional diagnostic tests. Needle biopsy is an invasive technique which can be used if needle aspiration results without a definitive diagnostic evaluation.

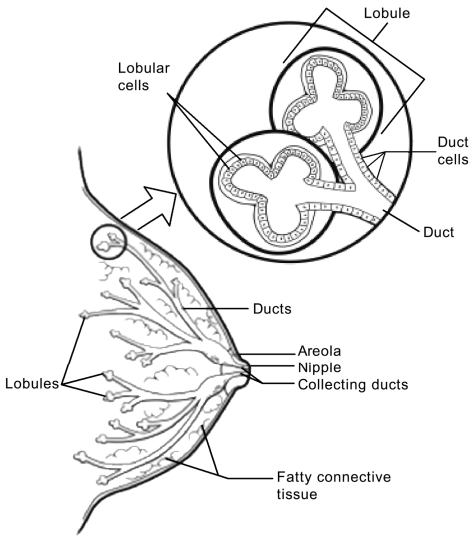


Figure 2 - Anatomy of a woman's breast. *Reprinted by the permission of the American Cancer Society, Inc. All rights reserved. From www.cancer.org*

It is taken using a wide-caliber needle and is followed by histological analysis of the biopsy, on which many different tumor markers can be examined, for the presence of tumor cells. Another imaging technique for breast cancer detection is magnetic resonance imaging (MRI) of which the sensitivity is much higher for infiltrating carcinomas and *in situ* lesions than mammography, but lacks high specificity. Therefore, it is suggested to use MRI in combination with mammography or ultrasound only (16). A technique that is fully under investigation is the detection of breast cancer by measurement of blood markers. Measurement of circulating tumor markers in breast cancer patients is currently most established in advanced disease, aimed at the diagnosis of metastasis and the evaluation of response to treatment (17).

Table 2 - *Cancer TNM classification table.*

Primary tumor (T)		Regional lymph node (N)		Remote metastasis (M)	
T0	None	N0	No lymph node metastasis	M0	None
Tis	Carcinoma *in situ*	N1	Metastasis in movable ipsilateral axillary lymph nodes	M1	Remote metastasis
T1	≤ 2 cm				
T2	2-5 cm	N2	Metastasis in ipsilateral axillary lymph nodes or in internal mammary nodes		
T3	>5 cm				
T4	Tumor of any size, extending though chest wall or skin	N3	Metastasis in axillary lymph nodes and in ipsilateral infraclavicular, supraclavicular, or ipsilateral internal mammary lymph nodes.		

Treatment

The main treatment of primary breast cancer is surgery. The goals of surgery are to remove the cancerous tissue and to analyze the size, grade and other clinical factors that are important to determine sequential therapy. Removal of the tumor can be performed by lumpectomy, a partial removal of the breast which includes the tumor and some of the surrounding tissue. Lumpectomy is breast conserving surgery (BCS) and is also referred to as wide local excision. Mastectomy is the oldest known form of breast cancer treatment and means the complete removal of the breast which may include removal of skin and muscle tissue (18). After breast cancer surgery, the tumor but also the removed surrounding breast tissue is macroscopically examined and sliced for further microscopic examination. For the preservation of cellular histological markers and long term storage, tissue blocks are prepared and are either formalin-fixed and paraffin-embedded (FFPE) or frozen in liquid nitrogen. From the paraffin blocks, slices of 3-5 μm are cut for staining with hematoxylin and eosin (H&E), and staining for estrogen receptor (ER), progesterone receptor (PR) and ERBB2 (human epidermal growth factor receptor 2). Based on the macro- and microscopic examinations, a pathologist will determine tumor type, histological grade, invasion, tumor size, pTNM stage and hormone receptor status of the breast tumor on which further treatment is based (19). Surgery might be followed by radiotherapy to destroy any remaining tumor cells in the breast, axillary tissue or chest wall. Radiotherapy can be given externally by collimated beams of radiation or internally after lumpectomy where a small pellet of radioactive material is given directly into the tumor bed (18). Next, treatment might be followed by systemic chemotherapy. The aim of adjuvant chemotherapy is to destroy or stop any tumor cells that have invaded the body. Chemotherapy interferes with the ability of rapidly growing cells to divide, including cancer cells but also cells present in the bone marrow, hair follicles and gastrointestinal tract. Different chemotherapeutic drugs are listed in Box 1 (18). Estrogen receptor positive breast tumors are dependent on estrogen for growth and proliferation. Endocrine treatment aims at estrogen starvation of the tumor cells by blocking the production of estrogen or limiting estrogen in reaching the tumor cells (18). Other targeted therapies aim at specific proliferative cell functions. These treatment options includes the targeting of ERBB2, EGFR, tyrosine kinases (IGF), and

- Surgery: lumpectomy, mastectomy.
- Radiation: internal, external.
- Chemotherapy:
- CMF (cyclophosphamide, methotrexate, and fluorouracil), AC (cyclophosphamide and doxorubicin), CAF (cyclophosphamide, doxorubicin, and fluorouracil), CEF (cyclophosphamide, epirubicin, and fluorouracil), TC (docetaxel and cyclophosphamide).
- Hormonal therapy: aromatase inhibitors (anastrazole, exemestane, letrozole, fadrozole), selective estrogen receptor modulators (tamoxifen, raloxifene, toremifene), oophorectomy.
- Targeted therapy: ERBB2 (trastuzumab, lapatinib), angiogenesis inhibitor (bevacizumab), EGFR inhibitor (gefitinib), tyrosine kinase inhibitor.

Box 1 - *Breast cancer treatment options (18, 20).*

angiogenesis (VEGF) (18, 20). Chemotherapy can also precede surgery and is then called neo-adjuvant chemotherapy. Its goal is to reduce the size of the tumor to make radical treatment intervention easier to perform and more likely to succeed (21).

Risk factors

Each year, approximately 130 per 100 000 women are diagnosed with breast cancer in Northern America, Northern and Western Europe, Australia and New Zealand (9). Based on these data, it can be estimated that approximately one in nine women will develop breast cancer during a period of > 80 years, setting the cumulative lifetime risk of breast cancer at 11% in these regions. The risk for developing breast cancer seems to be much lower in Asia or Africa as is indicated by the incidence of the disease (Figure 1). The international variations seem to disappear when Asian or African natives immigrate to regions with high breast cancer incidence (22-25). This has led to the conclusion that besides being a woman, significant risk of developing breast cancer lies in regional factors such as lifestyle and culture. The very first study to identify risk factors for breast cancer was performed by Janet Elizabeth Lane-Claypon and published in 1926 (26). Lane-Claypon identified that giving birth to a high number of children (>10) and giving birth to the first child at young age reduces the risk of breast cancer, women who have had no children at all, such as nuns, have a greater risk of breast cancer. These risk factors are examples of current cultural differences between the economically well and less developed countries. Besides reproductive behavior, other factors determined by lifestyle influence the risks for breast cancer; people in the western world tend to eat more animal products, eat less vegetable, have less physical activity and become older compared to people in other parts of the world. More factors that can influence the risk for breast cancer are birth weight, birth length, age at menarche, and age at menopause (27). Additional to the latter factors that are associated with the body's level of hormone exposure, one of the strongest risk factor for developing breast cancer is a family history of the disease (Box 2).

- Diet and diet-related factors
- Ionizing radiation
- Hormone and reproductive factors
- Benign breast disease
- Family history of breast cancer

Box 2 - *Factors involved in the risk of the development of breast cancer.*

Hereditary breast cancer

History

Families with three or more close relatives with breast cancer are commonly classified as "breast cancer families" (28). In the past, segregation analyses were performed in such families, showing an autosomal dominant mode of inheritance in most cases (29-31). By linkage analysis on a large group of families with early-onset breast cancer, the locus of a high-penetrance cancer susceptibility gene was mapped on chromosome 17q12-21 in 1990 (32). Not until 1994 a candidate gene was completely characterized and truncating mutations were linked to breast cancer (33). Because it was the first gene to be associated with hereditary breast cancer, it was called "breast cancer 1, early onset" or *BRCA1*. In the same year, the second major breast cancer susceptibility gene, *BRCA2*, could be localized on 13q12-13 and was cloned just one year later by Wooster and colleagues. The identification of loss of heterozygosity (LOH) at the *BRCA2* locus and germline mutations of this gene in breast cancer demonstrated the role of *BRCA2* as tumor suppressor gene (34, 35). After the discovery of the breast cancer susceptibility genes *BRCA1* and *BRCA2*, major changes have been made in the care of women with inherited predisposition to breast cancer such as increased screening and surveillance and risk reduction options (36). Mutations in the *BRCA1* and *BRCA2* genes are responsible for the major part of the hereditary breast cancer syndrome; however, other genes have been correlated with different forms of hereditary breast cancer syndromes. These syndromes are: Cowden disease, caused by mutation in the *PTEN* gene (37); Li-Fraumeni syndrome, caused by mutations in the *TP53* gene (38); Peutz-Jeghers syndrome, caused by mutations in the *STK11* gene (39) and Ataxia Telangiectasia, caused by mutations in the *ATM* gene (40). Furthermore, a single mutation in *CHEK2*, 1100delC, has also been reported to be associated with hereditary breast cancer (41) (Table 3). Breast cancer caused by mutations in breast cancer susceptibility genes has several distinctive clinical features such as considerably younger age at diagnosis compared to sporadic cases, the prevalence of bilateral breast cancer is higher, and associated tumors (*e.g.*, ovarian, colon, prostate, and pancreatic cancers, as well as male breast cancer) are seen in some families.

Incidence and risk

Mutations in the *BRCA1* and *BRCA2* genes can be identified in approximately 80% of families with a high number of breast cancer cases (*i.e.*, four or more) diagnosed before the age of 60 years. Germline mutations in *BRCA1*

Table 3 - *Gene associated life time risk of breast cancer in female carriers (42).*

Gene	Breast cancer risk	Syndrome
BRCA1	65%	BRCA1 syndrome
BRCA2	45%	BRCA2 syndrome
TP53	97%	Li-Fraumeni
PTEN	20-50%	Cowden
STK11	45%	Peutz-Jeghers
ATM	2-5 fold elevated risk	Ataxia Telangiectasia
*CHEK2*1100delC*	2-fold elevated risk	

have been detected in approximately half of familial breast cancer cases and in most cases of combined familial breast/ovarian cancers. *BRCA2* mutations are found in about thirty percent of the hereditary breast cancers. Clinically, carriers of a *BRCA* germline mutation present with a substantially higher risk of developing breast and ovarian cancer than the general population. By the age of 70 years, the breast cancer risk in *BRCA1* mutation carriers is 65% (95% CI 51-75%) and the ovarian cancer risk is 39% (95% CI 22-52%); in *BRCA2* mutation carriers, the corresponding risks are 45% (95% CI 33-54%) and 11% (95% CI 4-18%), respectively (43) (Table 3). The median age of diagnosis in mutation carriers is 42 years, approximately 20 years earlier than unselected breast cancer in the Western World and several years before mammographic screening is recommended in the general population (44). A small percentage of the hereditary breast cancer syndromes can be explained by other high- and low-penetrance breast cancer genes (45-48) (Table 3), but these will not be discussed further in this thesis. In total, it is estimated that 5-10% of all breast cancer cases are due to inherited mutations of which mutation in the *BRCA1* and *BRCA2* genes are the most frequent (49, 50). However, in an additional 15-20% of all breast cancer cases a positive family history of the disease is found; therefore, from all the families that are actually eligible and tested for *BRCA* germline mutations, only in approximately 25% a *BRCA1* or *BRCA2* mutation is diagnosed according to literature (51-54); in Dutch hospitals the current percentages are around 7-14% (55). About 10-25% of the cases tested for *BRCA* predisposition is diagnosed with an unclassified variant (UV) (55-57), whereas for the remaining breast cancer families the genetic test result is uninformative/inconclusive. In literature, the latter families are referred to as 'non-BRCA1/2 families' and it is likely that these people are carrier of mutations in other, still unknown, breast cancer susceptibility genes, which are collectively designated as *BRCAX* (58).

The importance to determine BRCA status

As a mutation in the *BRCA1* or *BRCA2* gene is one of the greatest risk factors for developing breast and ovarian cancer, identification of such a mutation is of significant clinical value. Mutation carriers are offered special medical care to reduce the risk of cancer development and, ultimately, mortality. First of all, providing individuals general information about *BRCA1/2* mutations by genetic counseling has been shown to reduce worrying about breast cancer, reduce anxiety and depression, and increase the likelihood of participating in genetic testing (59). Second, intensified screening for early detection of cancer by both mammography and magnetic resonance imaging (MRI) has been recommended for women with *BRCA* mutations (60, 61). Furthermore, women who have been identified with a germline mutation in *BRCA1/2* can opt for prophylactic surgery which includes prophylactic bilateral mastectomy (PBM) and prophylactic bilateral salpingo-oophorectomy (PBSO), to reduce the risk of breast and ovarian cancer by 85-100% (62-67). Lastly, women with an inherited predisposition to breast cancer can be offered chemopreventive agents such as oral contraceptives (68, 69) or tamoxifen, which has been found to reduce the incidence of breast cancer in healthy *BRCA2* mutation carriers by 62% (70). Poly(ADP-ribose) polymerase 1 (PARP1) inhibitors, a novel class of drugs which is still under investigation, have shown to be highly effective against *BRCA*$^{-/-}$ pre-cancerous

cells and might become chemopreventive agents in the future (this will be discussed in more detail further on) (71). Although these interventions might benefit women carrying a *BRCA* mutation, they should be avoided in non-carrier relatives (true negatives).

Recently, it has been shown *in vitro* that BRCA-deficient cell lines display increased sensitivity to agents causing double-strand DNA breaks such as cisplatin (72, 73). These findings may open the possibility that determination of *BRCA* status may also be used to guide therapy in the near future. However, because of the lack of prospective clinical validation, *BRCA* mutation carriers are offered similar adjuvant therapy as non-hereditary breast cancer patients at the moment (74, 75).

- Diagnosis of breast cancer at early age (<50 years)
- Triple negative breast cancer (ER, PR, ERBB2)
- Two breast cancer primaries (bilateral or ipsilateral)
- Both breast and ovarian cancer
- Ovarian/fallopian tube/primary peritoneal cancer
- One or more cases of breast and/or ovarian cancer in the family
- Clustering of breast cancer with various other cancers such as thyroid or pancreatic cancer on the same side of the family
- Presence of breast cancer in a male family member
- Known BRCA1 or BRCA2 mutation in the family
- Member of a population at risk (e.g., Ashkenazi Jewish)

Box 3 - *Family characteristics covering first-, second-, and third-degree relatives that have been described to be risk factors for hereditary breast cancer (51, 54, 56, 83). Affected individuals with one or more risk factors might be eligible for further risk evaluation.*

Eligibility for BRCA mutation testing

As a result of the large sizes of the *BRCA1* and *BRCA2* genes, mutation screening is expensive, complex and time-consuming. It would be inefficient to screen for inherited cancer susceptibility in all women diagnosed with breast cancer and it is therefore necessary to preselect eligible families for mutation testing. Several referral guidelines have been developed based on family characteristics that have been associated with increased risk of germline mutations in *BRCA1* or *BRCA2* for further risk evaluation (76-82). These risk factors include breast cancer onset at young age, ovarian cancer, two breast cancer primaries, a combination of breast and ovarian cancer, male breast cancer, or a known *BRCA* mutation in the family, ethnic group, and family history of breast or ovarian cancer (Box 3) (51, 54, 56, 83). In addition to the risk factors, several models have been developed to accurately evaluate the probability of a person carrying a *BRCA* mutation (84). Next, genetic counseling is performed with the subject and only those women with strong evidence for a germline mutation and with assumed sufficient benefit are recommended for further DNA diagnostics.

Estimation of *BRCA* carrier probability

The most important factors that determine the individual likelihood of a deleterious *BRCA* mutation in affected or cancer-free women remain family history of breast or ovarian cancer and a known family mutation (85, 86). Several models have been developed to estimate the probability that an individual person or family is a carrier of a mutation in *BRCA1* or *BRCA2* based on their family history of cancer (*i.e.*, age of onset and type of cancer in first- and second-

degree relatives), non-hereditary risk factors (*e.g.*, age at menarche and age at birth of first child), but also on the population prevalence of mutations, age-specific penetrance, and ethnic ancestry (87). The application of these tools is to select or exclude people from genetic counseling and genetic testing in order to provide a cost-efficient and clinically appropriate service. The currently available probability models include BRCAPRO (88-90), models from Myriad Genetic Laboratories (Salt Lake City, UT, USA) (56, 83, 91), the Couch model (also known as the Penn model) (54, 92), IBIS (International Breast Cancer Intervention Study) (93), and BOADICEA (Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm) (94, 95). Most models can calculate a *BRCA* mutation probability for affected as well as unaffected individuals. Table 4 describes seven of these models, noting on what their estimates are based and to whom they are applicable. The results of all of these tests should be interpreted with some caution, because each model bases its calculation of risk estimate on

Table 4 - *Models to calculate the likelihood of carrying a BRCA1 or BRCA2 mutation. References refer to original publication and electronic tool if available.*

Model	Estimates based on	Applications	References
BOADICEA (BRCA1 and BRCA2)	Extensive information on family history of cancer, male breast cancer.	Applicable for proband affected and unaffected by breast cancer.	(94-96)
BRCAPRO (BRCA1 and BRCA2)	Extensive information on family history of cancer, age at diagnosis, presence of bilateral breast cancer, male breast cancer, Ashkenazi Jewish heritage.	Applicable for individuals with or without breast or ovarian cancer.	(88-90, 97, 98)
IBIS (BRCA1 and BRCA2)	Breast/ovarian status, extensive information on family history of cancer, nonhereditary risk factors.	Applicable for proband affected and unaffected by breast cancer	(93, 99)
Myriad I (BRCA1)	Bilateral breast cancer, ovarian cancer, age at diagnosis, Ashenzi Jewish ethnicity, and family history of cancer.	Only applicable for proband affected by breast and/or ovarian cancer. Applicable to families with small numbers of affected members.	(56)
Myriad II (BRCA1 and BRCA2)	History of breast and ovarian cancer, Ashkenazi Jewish heritage, and family history of cancer.	Only applicable for proband affected by breast cancer < 50 years of age and/or ovarian cancer.	(51, 100)
Penn I/Couch (BRCA1)	Age at diagnosis, family history of cancer, Ashkenazi Jewish heritage.	Applicable for proband with or without breast or ovarian cancer.	(54)
Pen II (BRCA1 and BRCA2)	Ashkenazi Jewish heritage, family history of cancer.	Applicable for proband with or without breast or ovarian cancer but with > 2 breast cancer cases in the family.	(92, 101)

different parameters, which might mean that different results are generated for the same person (102-104). It should be well understood that these models calculate the probability of a *BRCA* mutation and not the true breast cancer risk, although some of them are able to do the latter. Other models that assess the risk of developing breast cancer include the Gail, Claus, Jonker, and extended Claus models (105-108). Additional models that have been developed to assist in selecting women for referral to genetic counseling include FHAT (Family History Assessment Tool), the Manchester scoring system, and RAGs (Risk Assessment in Genetics) (109-111).

Genetic testing

Mutations in the *BRCA1* and *BRCA2* genes are found throughout all coding regions and at splice sites, with most of these mutations being small insertions or deletions causing frameshift mutations, nonsense mutations, or splice site alterations. In order to detect these specific genetic alterations, the entire *BRCA1* and *BRCA2* genes have to be examined, which is a complex procedure. Only known founder mutations can be detected relatively easy in some high-risk families from specific ethnic groups (112). The current gold standard to determine *BRCA* mutations is direct sequencing of genomic DNA; however, since this is an expensive and time-consuming technique, many laboratories prefer the use of pre-screening techniques to detect any genetic anomalies first. Pre-screening techniques include protein truncation test (PTT), denaturing gradient gel electrophoresis (DGGE), denaturing high-performance liquid chromatography (DHPLC), single-stranded conformational polymorphism (SSCP), two-dimensional gene scanning (TDGS), fluorescent-assisted mismatch analysis (FAMA), heteroduplex analysis, and fluorescent conformational sensitive gel electrophoresis (F-CSGE) (113, 114). None of these techniques, including direct sequencing, is able to identify all cancer predisposing mutations in the *BRCA1* or *BRCA2* gene. Only by applying additional detection strategies for large deletions or duplications such as multiplex ligation dependent probe amplification (MLPA), can an estimated detection rate of up to 95% be achieved (115).

If a gene mutation is found, sequential clinical steps can be undertaken. If no mutation is found in a family member of a known *BRCA* mutation-carrying family (a "true negative"), the individual's risk of breast cancer is equal to that of the general population and no additional preventive measures are required. However, difficulties arise when no pathogenic mutation is found but the hereditary risk of breast cancer cannot be ruled out for this individual or family. Such test result is called uninformative or inconclusive which is absence of identification of a deleterious mutation in an entire family or identification of an unclassified variant (UV), a sequence variant of which the clinical significance is still unknown.

BRCA mutations

To date of writing, 1647 and 1857 unique mutations are described in the BIC database for the *BRCA1* and *BRCA2* genes, respectively. The majority of the mutations found are frame-shift or nonsense mutations (Table 5) and about 10% of the mutations are large exonic deletions or insertions (116, 117). These mutations are described as pathogenic because they result in missing, truncated or not properly functioning protein products. Splice site alterations cause incorrect splicing and may result in unstable mRNA and thus reduced levels of protein (118).

If a mutation can be traced back to a com-

Table 5 - *Source: The Breast Cancer Information Core (BIC) Database. UV, unclassified variants.*

Mutation type	*BRCA1* All mutations		*BRCA1* UV		*BRCA2* All mutations		*BRCA2* UV	
3-prime Untranslated Region (3'UTR)	2	0.1%	1	0.1%	0		0	
5-prime Untranslated Region (5'UTR)	3	0.2%	3	0.4%	10	0.5%	7	0.7%
Frameshift (F)	542	32.9%	0		544	29.3%	3	0.3%
In Frame Insertion (IFI)	1	0.1%	1	0.1%	4	0.2%	3	0.3%
In Frame Deletion (IFD)	24	1.5%	24	3.2%	23	1.2%	23	2.2%
Intervening Sequence (IVS)	264	16.0%	159	21.0%	173	9.3%	124	12.1%
Missense (M)	570	34.6%	537	71.2%	847	45.6%	812	79.2%
Nonsense (N)	194	11.8%	0		189	10.2%	1	0.1%
Splice (S)	1	0.1%	1	0.1%	2	0.1%	0	
Synonymous (Syn)	46	2.8%	29	3.8%	65	3.5%	52	5.1%
total	1647		755		1857		1025	

mon ancestor it is called a founder mutation. Such a mutation is often enriched in a certain ethnic group. The prevalence of a founder mutation can be strongly enriched among certain ethnic groups such as the Ashkenazi Jews, in which three founder mutations have been identified: the 187delAG and 5385insC in *BRCA1*, present in about 1.1 and 0.15% of the Ashkenazi Jews, and the 6174delT mutation in *BRCA2*, present in 1.5% of the Ashkenazi Jews. These three mutations account for a total *BRCA* mutation carrier frequency of 1 per 40 individuals of Ashkenazi Jewish descent, which is notably high (119-121). Another example is the founder mutation *BRCA2* 999del5 that accounts for 7-8% of female breast cancers and for 40% of male breast cancers in Iceland (122). More *BRCA* founder mutations are seen in other countries in which endogamy is a common practice among certain social or religious groups, including in the Netherlands (112, 123-125).

Unclassified variants in BRCA

Besides the pathogenic mutations found in the *BRCA1* and *BRCA2* genes, more than half of the nucleotide changes in these genes occur rarely (<1%) and their clinical significance is unknown (116). These mutations are called unclassified variants (UVs) and are generally missense (M) or intervening sequence (IVS) variants that result in substitution or loss of a single amino acid (Table 5) (118). Clinically, the identification of an UV in an individual's germline DNA is a difficult situation and to prevent unnecessary surgery, it is important to determine whether the mutation adversely affects the functions of the protein. However, determining the pathogenicity remains difficult because of the limited knowledge about the functional outcomes of such nucleotide variant; therefore, much research is currently being performed to assess the pathogenicity of each UV. Multifactorial classification models base the risk associated with a UV on combined data from variant frequency, co-segregation with cancer, and features consistent with a real pathogenic gene mutation such as family history of cancer, co-occurrence (*in trans*) of another known pathogenic mutation, tumor histopathology, loss of heterozygosity (LOH) of the wild-type allele, evolutionary conservation, and evidence from functional assays (126, 127)

(Table 6). Because multifactorial classification models are limited by the amount of families carrying the UV, complementary approaches are often required, *i.e.*, *in vitro* assays that make use of transcript and functional analyses or prediction of splicing aberrations using bioinformatics (*in silico*) (128-132). Still, determining the pathogenicity of UVs remains difficult, laborious and time consuming and new techniques are being developed to preselect for variants with high pathogenic possibility.

Non-BRCA1/2 families

Most of the women from breast cancer families do not carry a pathogenic mutation or unclassified variant in the *BRCA1* or *BRCA2* gene (53, 55). Linkage analysis mapped a third breast cancer susceptibility gene on chromosome 13q, distinct from *BRCA2* and *Rb*, but this was opposed a few years later (133). It is now suspected that an unknown number of low penetrance genes or a combination of common

Table 6 - *Types of evidence for UV classification (126, 130).*

Multifactorial classification	
Variant frequency	Frequency of the variant occurring in cases and controls provides a direct estimate of associated cancer risk but studies would need to be prohibitively large because of the rarity of the variants.
Co-segregation	Co-segregation with the disease in families allows for easily quantifiable and directly related risk.
Co-occurrence with deleterious mutation	UVs, co-occurring with a deleterious mutation in the same gene (*in trans*), can be classified as neutral if homozygotes are assumed to be embryonically lethal. However, this method exhibits less power to show causality.
Family history	Personal and family history of cancer of the carriers of the UV usually can easily be obtained; however, it is not as robust as co-segregation and the power may be low for infrequent variants.
Pathological classification	Histopathological tumor features are a potential powerful predictor for BRCA1-related tumors in which the pathological characteristics are quite distinct, however, for BRCA2-related tumors the prediction might be weak (discussed further on).
LOH	Occurrence of loss of heterozygosity in tumor DNA should be used as an adjunction to co-segregation results.
Conservation	The severity of the amino acid change and its conservation across species can be very predictive if enough evolutionary time sequence is available. Still, it is only indirectly related to disease risk.
***In vitro* assays**	
Functional analyses	Functional analyses can evaluate the effect of the variant on the protein's ability to perform (some of) its cellular functions. These assays include the determination of transcription activity, small colony phenotype, ubiquitin ligase activity, rescue of radiation resistance, embryonic stem cell-based functionality, homologous recombination, mitomycin C survival, and centrosome amplification
Transcript analyses	Transcript analysis can efficiently identify variants affecting the stability and integrity of mRNA transcripts.
In silico tools	Computation analysis of the UV can be used to predict the effect on mRNA splicing

variants (polygenic model) with multiplicative effects on risk may be responsible for this substantial proportion of hereditary breast cancer (114, 134-137). Researchers have come to this conclusion after investigation of a large cohort of 149 non-BRCA1/2 breast cancer families in which linkage analysis has not been able to provide a locus on which a third major breast cancer gene might be located with statistical significance (138). Different approaches are now being examined to decrease the genetic heterogeneity and increase the statistical power of finding a breast cancer susceptibility locus, thus far without success (139). The difficulty in obtaining genetic homogeneous groups is the lack of specific familial phenotypes such as in families carrying a *BRCA1* or *BRCA2* mutation where ovarian and male breast cancer were recognized to be common. Additionally, studying families with high breast cancer incidence at early age could provide a more genetic homogeneous group but increases the likelihood of involvement of the *BRCA1* or *BRCA2* genes (140). New ways are needed to cluster families into subgroups of single-gene disorders.

Treatment

Patients with hereditary breast cancer are offered bilateral mastectomy as treatment and to simultaneously decrease the risk for local recurrences or secondary primaries. As such, breast conserving surgery (BCS) is not the best therapy in *BRCA*-mutation carriers as these patients still have a substantial increased risk for local recurrences compared to sporadic breast cancer patients or *BRCA*-mutation carriers undergoing mastectomy; nevertheless, survival after mastectomy or BCS has not been shown to be significantly different (141). Chemotherapy after BCS has been shown to decrease the risk for local recurrence from 23.5% to 11.9% in *BRCA*-mutation carriers, however, adjuvant chemotherapy by itself does not have any additional effect (positive or negative) on survival compared to patients with sporadic breast cancer (141, 142). It is thought that *BRCA* deficient tumors might be more sensitive to poly(ADP-ribose) polymerase inhibitors, but limited clinical data are available (this will be discussed in more detail further on) (143). *In vivo* and *in vitro,* homozygote *BRCA*-mutated cells have been found to be more radiosensitive due to the lack of proper DNA repair by homologous recombination and, additionally in *BRCA1* deficient cells, due to the lack of the cell-cycle G2-M checkpoint to stop cells before mitosis upon DNA damage (144-146). It would be very interesting to exploit this in cancer treatment; on the other hand, radiotherapy could then also increase the risk for secondary cancers. However, the currently available data from clinical studies do not provide evidence of hypersensitivity for radiotherapy in breast cancer patients carrying a *BRCA*-mutation or increased cancer sensitivity, cancer recurrences are reported to be similar compared to patients with sporadic breast cancers (147). Taken together, so far the clinical treatment of hereditary breast cancer patients does not differ from the treatment given to patients with sporadic breast cancer although the roles of BRCA1 and BRCA2 in DNA repair might be a future target in cancer treatment for these patients.

Histopathological features

Breast cancer type, histological grade, invasion, estrogen receptor (ER) status, progesterone receptor (PR) status, and ERBB2 (human epidermal growth factor receptor 2) status are routinely determined histological features to guide therapy.

Table 7 - *Common histopathological features of hereditary and sporadic breast tumors. Negative status for ER, PR and ERBB2 is also referred to as "triple-negative". Data summarized from (148-151).*

	BRCA1 (%)	BRCA2 (%)	non-BRCA1/2 (%)	Sporadic (%)
Invasive lobular carcinoma	7	13	14	12
Invasive ductal carcinoma	74	71	73	69
Medullary carcinoma	18	3	2	3
Grade I	2	20	24	22
Grade II	24	42	44	42
Grade III	73	38	32	36
ER				
positive	21	65	72	66
negative	79	35	28	34
PR				
positive	20	49	60	56
negative	80	51	40	44
ERBB2				
positive	7	6	3	18
negative	93	94	97	82
Triple-negative	57	23	14	11
p53				
positive	45	27	12	27
negative	55	73	88	73
KRT5/6				
positive	65	7	13	8
negative	35	93	87	92

Hereditary breast cancers have several apparent features that separate them from sporadic breast cancers. Compared to sporadic breast tumors, *BRCA1*-associated breast tumors generally are grade III, hormone (ER and PR) and ERBB2 receptor negative (also referred to as triple-negative), often show p53 protein accumulation, much lymphocyte infiltration and a high expression of K5/6 (149-151). Most *BRCA2*-associated breast tumors are grade 2/3 and are often ER and ERBB2 positive (149). In contrast to *BRCA1*-related breast tumors, *BRCA2*-related breast tumors show much less distinctive features as compared to age-matched sporadic breast tumors (Table 7). Similar to *BRCA2*-associated breast tumors, non-BRCA1/2 breast cancer is a heterogeneous group that is comparable to sporadic breast cancer. Significant differences have been found in keratin expression, as K14 and K5/6 are higher expressed in *BRCA2*-associated and non-BRCA1/2 breast cancer (148, 149). Although several studies have been performed to classify hereditary breast cancer based on histopathological features (152), it should be noted that none of these features in itself or in combination is unique to any of the hereditary breast cancers; therefore, histopathological features are not being used to identify hereditary breast cancer cases.

Molecular biology of hereditary breast cancer

The *BRCA1* gene

The human *BRCA1* gene contains 24 exons, encoding for a protein of 1863 amino acids (33). The protein contains multiple functional domains, including a highly conserved RING finger domain in its N-terminal region, two nuclear localization signals (NLS) that are located in the *BRCA1* gene in exon 11, a DNA binding domain between amino acids 452-1079, an SQ-cluster domain (SCD) between amino acids 1280–1524, and tandem BRCT repeats in its C-terminal region (Figure 3) (153). The BRCA1 protein interacts directly or indirectly with many other molecules, including tumor suppressors, proto-oncogenes, DNA damage repair proteins, cell cycle regulators, as well as with transcriptional activators and repressors (114, 154). Normally, BRCA1 is part of a heterodimer together with BARD1 of which the interaction is mediated by alpha-helical units adjacent to the RING domain (155). Without BARD1, BRCA1 is unstable and is rapidly degraded; it would be unable to perform its tumor suppressor functions (156). The BRCA1-BARD1 complex serves as an ubiquitin ligase *in vitro* (157), however, *in vivo* it is largely unknown what its substrates are. BRCA1 is localized to the site of the double-strand break by binding Abraxas at the BRCT repeats, followed by interaction with RAP80 (158, 159). Not only is the BRCA1-Abraxas-RAP80 complex involved in DNA repair, it also regulates phosphorylation of CHK1 kinase through a yet still unknown mechanism (160). CHK1 kinase is involved in DNA damage-driven cell cycle checkpoint control and is important to arrest cells to allow them time for DNA repair before mitotic entry. Furthermore, BRCA1 plays a role in replication checkpoints that are activated in response to replicative stress such as collapsed or stalled replication forks (146, 161), but also in mitosis control where the BRCA1/BARD1 heterodimer is required for mitotic spindle pole assembly (162). Finally, the ubiquitination of topoisomerase IIa, which is involved in the decatenation of replicated DNA, is regulated by BRCA1 (163). Loss-of-function mutations of *BRCA1* would therefore result in pleiotropic phenotypes, including defective DNA damage repair, a defective G2/M cell cycle checkpoint, abnormal centrosome duplication, chromosome damage, aneuploidy, and impairment of the spindle checkpoint (164, 165).

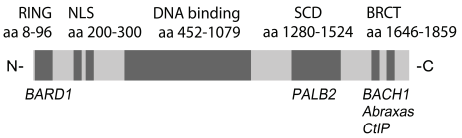


Figure 3 *- Functional domains in the BRCA1 protein (dark gray) and interacting proteins (bottom).*

The *BRCA2* gene

The human *BRCA2* gene covers 70kb of genomic DNA and has 27 exons, encoding for a protein of 3418 amino acids (35). While *BRCA1* has a wide range of functions in many different cellular processes, the primary function of *BRCA2* is limited to homologous recombination, both in meiosis and repair of double-strand breaks (154, 166). Through the interaction with PALB2, BRCA2 is located to the site of damage together with BRCA1. BRCA2 is able to bind to single strand DNA through interaction with DSS1 to the helix-rich domain (HD) (Figure 4). A tower domain emerges from the second OB (oligonucleotide binding) fold, which is topped

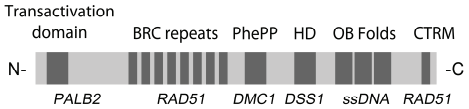


Figure 4 - *BRCA2 protein organization, domains indicated in dark gray and interacting proteins below (154, 166). HB: helical domain, OB: oligonucleotide-binding, CTRM: c-terminal RAD51 binding motive.*

by a three-helix bundle for DNA binding. Although ssDNA is preferred by BRCA2, the structure of the tower domain suggests the possibility of binding to duplex DNA too (167). Next, RAD51 is loaded onto the 3'-strand overhang, which is bound to the BRC repeats of BRCA2 to facilitate DNA repair (Figure 6). Trough the CTRM domain, BRCA2 stabilizes the resulting nucleoprotein filament. Because RAD51 is required for DNA-repair by homologous recombination, it is not surprising that BRCA2-deficient cells exhibit genetic instability (168, 169).

BRCA1 and mammary stem cell differentiation

Normal stem cells are primitive undifferentiated cells that are capable of self-renewal while maintaining the undifferentiated state but have the potency to differentiate into specialized cell types. Stem cells maintain the growth of normal proliferative tissue such as intestinal epithelium, skin, or bone marrow but also guarantee tissue regeneration after injury. Stem cells are the top of the cellular hierarchy and give rise to progenitors with more restricted lineage potential (Figure 5) (170). It is postulated that similar to normal proliferative tissue, the growth of a tumor is driven by a limited number of so called cancer stem cells (CSC) (171). Cancer stem cells maintain the growth of the neoplastic clone and give rise to rapidly proliferating and more differentiated cells that form the bulk of the tumor. One of the CSC concepts is that the tumor-initiating cell was originally an adult stem cell or a progenitor cell that has accumulated (epi)genetic damage resulting in tumorigenesis (172, 173).

The cyclical nature of mammary gland growth and involution during each pregnancy suggests the presence of stem cells in breast tissue (176, 177), but a consensus on the phenotypic definition of normal human mammary stem cells is still lacking at this point (178). Mammary adult stem cells can differentiate into two distinct cell types: luminal and myoepithelial (Figure 5). The luminal cell layer in mammary ducts is composed of progenitor luminal cells

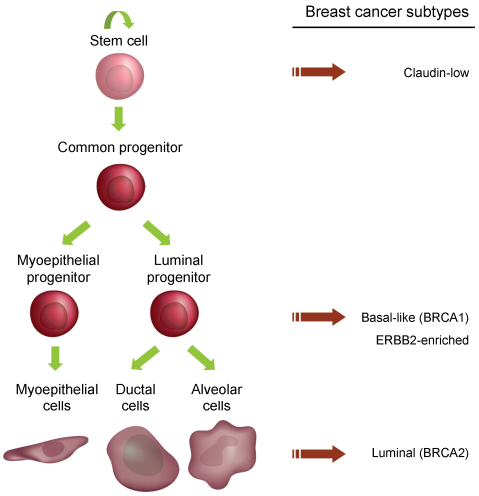
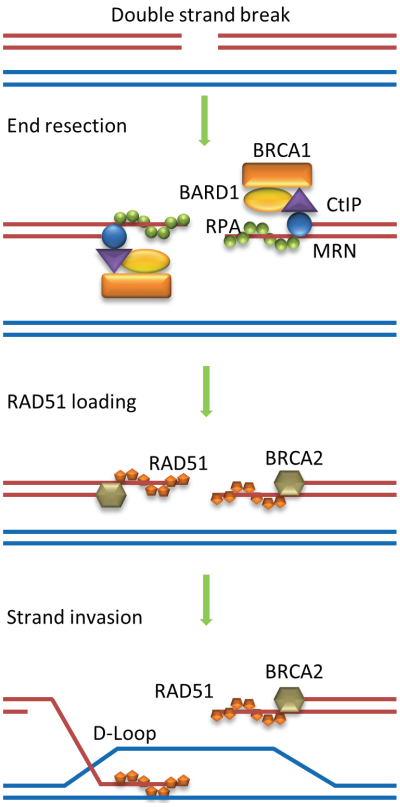


Figure 5 - *Stem cells have the ability to self-renew and give rise to more differentiated progenitor cells. The progenitor cells will further differentiate into myoepithelial, ductal, and alveolar cells and form the lobules and ducts in the breast. Depending on the cell of origin, different subtypes of breast cancer can arise. Source: (174, 175).*

lacking expression of estrogen receptor (ER) and differentiated luminal cells that express ER and progesterone receptor (PR) (179, 180). BRCA1 plays an important role in the differentiation from ER negative progenitor cells to mature ER positive luminal cells. Women with a *BRCA1*-mutation often show entire lobules in the breast tissue to be ER negative and ALDH1 (a stem cell marker) positive, although histological normal, whereas this is not seen in non-mutation carriers (181). Loss of BRCA1 function results in blocked epithelial differentiation which leads to growth of undifferentiated luminal progenitor cells. Because BRCA1 also functions in DNA repair, these progenitor cells are prime targets for further carcinogenic events (182).

The roles of BRCA1 and BRCA2 in DNA repair

DNA damage occurs continuously throughout a person's whole life, and is caused by both exogenous and endogenous stresses. Different DNA damaging sources cause different types of DNA damage and of these, double strand breaks are the severest because it affects both strands of the duplex, thus no intact complimentary strand is available as a template for repair (73). Inappropriate repair of such DNA damage in a cell can lead to either loss of viability or to chromosomal alterations that increase the likelihood of cancer development (183). Fortunately, all mammalian cells possess two enzymatic path-



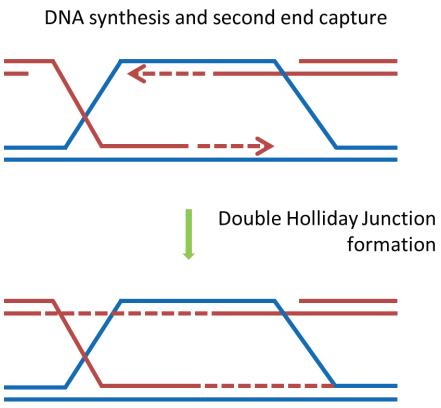


Figure 6 – *(left) Upon formation of a DNA double strand break, the MRN complex recognizes the lesion and recruits CtIP and BRCA1/BARD1. On either sites of the break, 3'-single stranded DNA overhangs are generated that bind RPA. Next, BRCA2 is recruited by BRCA1 through interaction with PALB2. BRCA2 loads RAD51 recombinase, displacing RPA. The sister chromatid is invaded by the strand overhang with RAD51 nucleoprotein and the homology search is initiated. (right) Once homology is established, DNA synthesis is started using the sister chromatid as a template for both 3'-overhangs. The resulting structure is called a double Holliday Junction and is resolved by Holliday junction resolving enzymes (figure adapted from [154]).*

ways that mediate the repair of DNA double-strand breaks (DSB): homologous recombination (HR) and non-homologous end-joining (NHEJ). The HR pathway is a very accurate repair mechanism in which a homologous stretch of DNA on a sister chromatid serves as a template to guide repair of the broken strand. It is most active in the late S-G2 phase of the cell cycle and accounts for the repair of ~10% of DSBs in mammalian cells. The role of BRCA1 is to remodel the chromatin to make the DNA damaged site become accessible to the DNA repair machinery; next BRCA2 directly translocates the DNA repair protein RAD51 to facilitate the repair (Figure 6). As both BRCA1 and BRCA2 are involved in DSB repair by homologous recombination, lack of one of these genes will result in HR defects and leaves a cell with only NHEJ to repair double strand breaks (184). In NHEJ, which can take place during the whole cell cycle, the two ends of the broken DNA molecule are processed to form compatible ends that are directly ligated. Because this repair mechanism lacks a homologous sequence control system, deletion, inversion, or any other type of abnormality in the genome could occur as a consequence (185). NHEJ is therefore recognized as a potentially less accurate form of DSB repair. The functions of BRCA1 and BRCA2 can become completely inactivated only, when both maternal and paternal genes have been silenced. The chances of losing both gene copies during a lifetime have been shown to be relatively small (186, 187); however, women carrying a germline mutation in BRCA1 or BRCA2 already have one silenced copy and the chance of losing the second copy is relatively high. The loss-of-function of the second allele, often revered to as the 'second-hit' (188) and in most cases caused by loss of heterozygosity (LOH), can lead to cancer formation (189).

Synthetic lethality in *BRCA* mutated tumors

PARP1 and PARP2 are proteins involved in the repair mechanism called base excision repair (BER), which is a key pathway for the repair of DNA single-strand breaks (SSB) (190). When a single-strand break is not repaired and encountered by a DNA replication fork, the fork will stall and eventually collapse, which will lead to DNA double-strand break formation (191). Usually, such breaks are repaired by homologous recombination (HR) or non-homologous end joining (NHEJ); however, in BRCA1- or BRCA2-deficient cells, the homologous recombination repair pathway is impaired and the cell has to resort to NHEJ only, which is error-prone. This has led to the hypothesis that homologous recombination deficient (HRD) cells, such as tumor cells in *BRCA1*-mutated breast cancer, might be hypersensitive to the inhibition of PARP and crippling the BER mechanism. In contrast, normal cells with at least one functional copy of *BRCA1/2* should be able to repair the breaks, resulting in chromosomal stability and cell viability (192). Indeed, preclinical studies have shown that PARP inhibitors are synthetic lethal in *BRCA* mutated cells (193, 194).

To date, clinical data on the use of PARP inhibitors as anticancer drugs are limited and only Phase I and II studies have been performed; however, these studies have shown promising results in terms of antitumor activity (143, 195-197). This new therapy has the potential to improve current therapy options for *BRCA*-mutation carriers, but patients with sporadic breast cancer with HRD might also benefit from this synthetic lethality. It should, however, be noted that there is currently no clinical test available to routinely investigate *BRCA* or homologous recombination status in sporadic

breast cancer. As BRCA-deficient breast tumors have not shown to be a histopathological unique entity, genetic markers found by microarray experiments might form the basis for a tool that is able to indicate HRD or BRCA-deficiency in the future.

Microarray technology

A microarray is a solid surface, generally a glass slide, on which multiple known nucleotide sequences, called probes, have been immobilized in gridded formation. The probes function as target on which fluorescently labeled DNA or cDNA can be hybridized. A quantitative measurement of hybridized sample to each probe can be made by comparison to a reference sample. In this thesis, two different microarray technologies are described, the first technique can be used to measure gene expression levels and the second to measure DNA copy number levels. The advantage of using microarrays as compared to other techniques that are able to quantitatively measure gene expression or copy number levels is its ability to perform thousands to even millions of measurements in parallel. Gene expression (GE) microarrays can be used to measure the amount of mRNA expression of basically every known gene in the human genome compared to a standard (198). The technique to study copy number levels is called comparative genomic hybridization (CGH) and as a microarray application array-CGH (aCGH). This technique can be used to measure the amount of DNA copies of most part of the genome (199). Both microarray techniques have been employed to search for markers that can be used as a target for anticancer drugs, but also to identify profiles on which breast cancer can be separated into molecular subtypes. This thesis will concentrate on the latter methodology, the identification of profiles specific for tumor subclasses.

Breast cancer subtypes

Clinically, breast cancer can be classified into two main groups: the estrogen receptor (ER) negative and ER positive breast tumors. However, an important hallmark in breast cancer classification was the identification of multiple subtypes within the ER negative and positive

Table 8 - *Data from 214 breast tumors, metastasis and normal tissue were excluded from the original study (200).*

	Luminal A (%)	Luminal B (%)	ERBB2-enriched (%)	Basal-like (%)	Claudin-low (%)	Normal-like (%)
Prevalence	31	18	13	24	11	3
Grade I	23	3	4	2	0	17
Grade II	49	26	39	10	24	17
Grade III	28	71	57	86	76	66
ER						
positive	87	86	22	13	10	67
negative	13	14	78	87	90	33
PR						
positive	73	56	21	5	30	40
negative	27	44	79	95	70	60
ERBB2						
positive	10	26	70	9	31	60
negative	90	74	30	91	69	40

tumors based on gene expression profiles (201). To date, six different molecular subtypes have been described and are known as luminal A, luminal B, ERBB2-enriched, basal-like, normal-like and claudin-low (202); the latter group was formerly classified as a subgroup within the basal-like subtype and was called basal-like B (203). The molecular subtypes are associated with differences in histopathological feature, clinical outcome and response to chemotherapy. Patients with luminal A tumors have the longest survival times, patients with basal-like or ERBB2-enriched subtypes have the shortest survival times, and patients with luminal B or claudin-low tumors have an intermediate survival (200, 204). Basal-like and claudin-low breast tumors are in general ER, PR and ERBB2 negative (triple-negative) while ERBB2-enriched generally show overexpression of ERBB2 (Table 8). It is postulated that the basal-like and ERBB2-enriched breast tumors originate from ER negative luminal progenitor cells, while claudin-low tumors originate directly from mammary stem cells (Figure 5) (174, 175, 205). Luminal breast tumors are often ER positive and originate from differentiated luminal cells. Studies investigating hereditary breast cancer have shown that the majority of breast tumors from *BRCA1* mutation carriers are of the basal-like subtype (74-90%), while breast tumors from *BRCA2* mutation carriers are more heterogeneous but predominantly of luminal B type (204, 206, 207).

Genomic instability and CGH

Genomic instability is one of the main characteristics of cancer and includes aneuploidy, polyploidy, translocations and amplification (1). Genomic changes are the causative factors in the initiation, development, and progression in breast neoplasms (208). These aberrations can be studied and characterized to better understand the evolutionary pathways a cell undergoes to ultimately grow out to cancer. Errors in chromosome duplication, segregation and telomere dysfunction in the absence of caretaker genes are examples from which chromosomal aberrations can arise (209). During tumorigenesis, DNA regions that include oncogenes are frequently amplified causing overexpression of the gene and giving the cell growth advantage; tumorsuppressor genes are often lost during the evolutionary process of cancer so cells can escape cell death. It has been noticed that several aberrations are recurrent in breast cancer such as amplifications of the genes *MYC* on chromosome 8q24 and *ERBB2* on chromosome 17q12 (210). Investigating DNA copy number alterations across a tumor's entire genome was a challenging task until the introduction of comparative genomic hybridization (CGH) technology (211). Improvements in the conventional or metaphase CGH, together with the development of microarray technology, led to the introduction of array CGH. Compared to metaphase CGH, array CGH has the advantages of being a high-throughput technology and providing a better resolution (199). The array CGH platform used in this thesis consisted of large-insert clones called BAC (bacterial artificial chromosome) clones, providing a genome wide resolution of 1 Mb on average. The procedure of performing an array CGH experiment is devised from several steps: DNA extracted from tumor material and reference DNA are differentially labeled with fluorescent dyes Cy5 and Cy3, respectively, and mixed in a 1:1 ratio. To block repetitive sequences, C_0t-1 DNA is added. Subsequently, the mixture is co-hybridized on a glass slide spotted with DNA probe sets under controlled temperature and humidity conditions. Lastly, the fluorescence of the hybridized

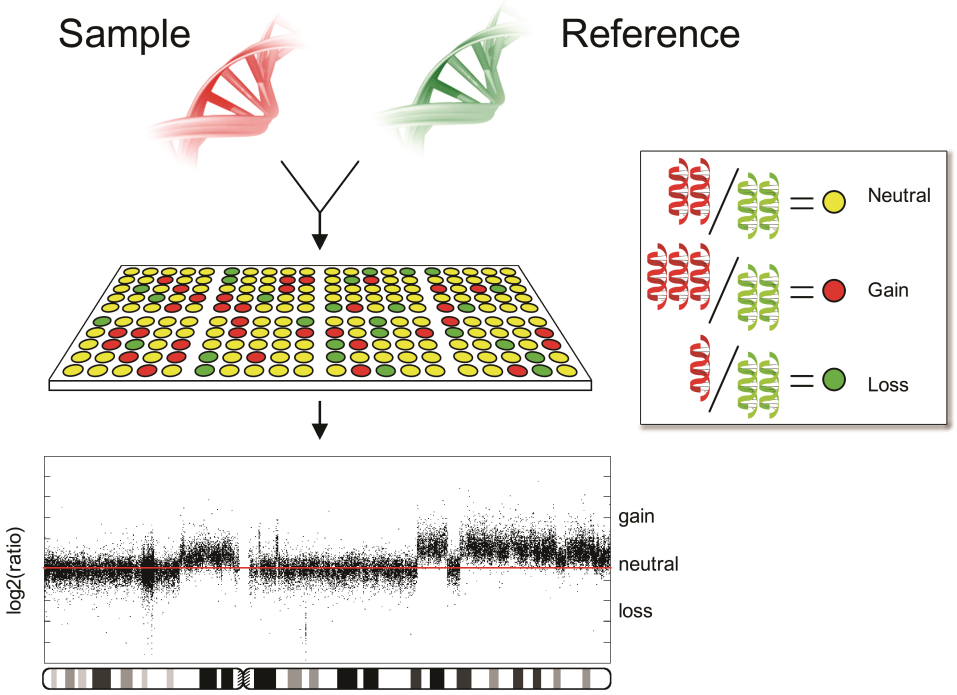


Figure 7 - *Sample DNA labeled with red fluorescent dye and reference DNA labeled with green fluorescent dye are co-hybridized to a microarray. A red-to-green ratio profile is generated from the intensity measurements. A higher red signal compared to green signal means gain/amplification of sample DNA, whereas less red signal means loss/deletion of sample DNA.*

DNA is measured and specialized software converts the fluorescent intensity data to a linear red-to-green ratio profile (called CGH profile) that correlates with the hybridization intensity, which mainly depends on the extend and size of the tumor's DNA copy number changes (Figure 7).

CGH profiles

Since the development of comparative genomic hybridization (CGH), chromosomal aberrations (*i.e.*, aneuploidy, polyploidy and amplifications) of genomic DNA in breast cancer, have been under extensive study to identify novel candidate cancer genes (212). Currently, three different types of profiles of genetic alterations detected by array-CGH are described (Figure 8). The first profile exhibits only few gains or losses of whole chromosome arms with most characteristically gain of chromosome 1q and 16p and loss of 16q. This profile is mainly associated with ER positive breast cancers and those of the Luminal A subtype. The second type of profiles contains high-level DNA amplifications with a moderate complex pattern of other gains and losses seen along the whole genome, typically for Luminal B and ERBB2-enriched subtype breast tumors. The last type of profiles is characterized by a complex pattern of numerous small aberrations including gains, losses and amplifications, and is associated with *TP53* mutated, basal-like and claudin-low

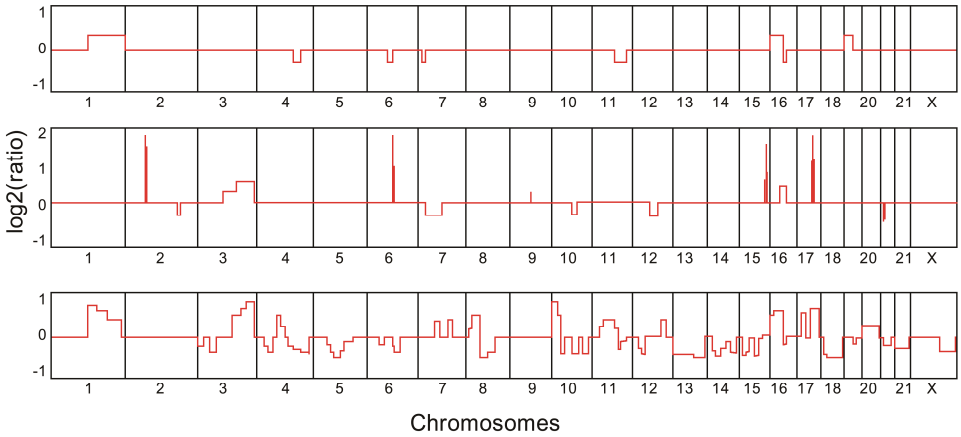


Figure 8 - *Three different types of array-CGH profiles are identified in breast cancer research. A simple profile with only a few aberrations (upper), a moderate complex profile with high-level amplifications and several other aberrations (middle), and a complex profile with many gains, losses and amplifications (bottom).*

subtype breast tumors (209). Because both *BRCA1* and *BRCA2* genes are part of DNA repair pathways, lack of function of one of these genes causes impaired DNA repair by homologous recombination, resulting in an accumulation of genetic errors and chromosomal instability (184). This thesis concentrates on characterizing the CGH profiles of *BRCA1*- and *BRCA2*-mutated breast tumors, and to extract the aberrations specific for *BRCA1*- or *BRCA2*-deficiency.

This thesis

Tumor based prediction of BRCA status

Although the inclusion criteria for *BRCA* mutation screening mentioned earlier cover many of the characteristics of typical BRCA families, mutation carriers might be missed because of lack of a family history of cancer or due to late age at breast cancer onset. In addition to pedigree-based risk assessment for *BRCA* mutations, a number of strategies exist to determine tumor-specific characteristics on which association with BRCA defects might be deduced. A recent study showed that on mammography and MRI, *BRCA*-associated lesions were more often described as rounded and with sharp margins compared with an age- and tumor type-matched control group; however, the prognostic value of this has yet to be evaluated (213). Morphological investigations show clear differences between *BRCA1*-mutated and age-matched sporadic breast tumors: the lesions are mainly of higher grade, have more pleomorphisms, a higher mitotic count, and less tubule formation. In addition, *BRCA1*-associated carcinomas are more often of the medullary type. *BRCA2*-mutated breast tumors, on the other hand, are generally more similar to sporadic breast tumors, but they show less tubule formation and the occurrence of tubular carcinoma is less common (214-218).

Histological, *BRCA1*-mutated breast tumors are in general estrogen receptor (ER), progesterone receptor (PR) and ERBB2 negative; whereas

this "triple-negative" phenotype is present in only 11% of sporadic breast tumors (152, 219-221). Other characteristics can be found at the molecular level; these include loss of heterozygosity (LOH) of the *BRCA* genes (222), frequent *TP53* mutations in *BRCA1*-mutated tumors (223), hypermethylation of *BRCA1* gene promoter CpG islands (224), and chromosomal aberrations (225). Although these characteristics are not unique to *BRCA*-mutated breast tumors and can therefore not be solely used to identify association with mutated *BRCA*, they might be of indicative value in cases of naïve inherited breast cancer susceptibility. Tumor characteristics can be informative not only for untested but also for *BRCA*-tested individuals. Members of breast cancer families may misinterpret uninformative mutation screening test results (57, 226, 227); hence, for such families, deciding for the right prevention and treatment strategies can be difficult. The reasons for an uninformative *BRCA* status test result might be defects in other genes that have thus far not been correlated with breast cancer predisposition (*i.e.*, *BRCAX*) or (epi)genetic defects located in the *BRCA1* or *BRCA2* genes missed by genetic testing (115). Besides the pedigree based risk assessment for *BRCA* mutations, so far no other strategies exist that identify association with *BRCA* defects after routine diagnostics has not been able to identify a mutation. The work described in this thesis explored the possibility to predict the likeliness of *BRCA* association in breast cancer based on the tumor's chromosomal aberrations using array CGH.

Array CGH — When and How

After a tumor has been surgically removed, the tissue will usually be fixed in formalin and subsequently embedded in paraffin for histopathological diagnosis and long-term storage. Quality of the tissue's DNA is mostly affected by variability in sample fixation time and duration of storage (228). To improve efficiency of array CGH experiments, we investigated whether we could assess the quality of the tumor's DNA prior to array CGH assays. **Chapter 2** of this thesis describes a relatively easy to perform multiplex PCR test to determine whether the DNA quality is sufficient for a successful array CGH hybridization (229). After DNA of sufficient quality has been collected, array CGH experiments can be carried out. Hybridization and slide washing was formerly performed by hand, due to technical advances this process can now also be performed in an automated fashion. Automated hybridization will reduce handling time and, most importantly, improve the reproducibility compared to manual array CGH hybridization. **Chapter 3** describes the optimization of automated array CGH for formalin-fixed, paraffin-embedded breast cancer tissues (230).

Predicting BRCAness by array CGH

The current strategy to identify *BRCA1* and *BRCA2* mutation carriers is direct sequencing preceded by pre-screening for abnormalities; however, it still remains unclear to what extent *BRCA* mutation carriers are missed by this approach. Additionally, the detection of variants of unknown clinical significance is emotionally and clinically challenging. Therefore, an additional tool that would indicate the involvement of *BRCA* in the development of the individual breast tumor would be an asset to current clinical diagnostics. Since both *BRCA1* and *BRCA2* genes are involved in DNA repair, *BRCA*-associated tumors are characterized by intensive genomic instability (231-233). This thesis describes the studies of these chromoso-

mal aberrations by array CGH, which has led to the identification of the specific aberrations of *BRCA1*- and *BRCA2*-associated breast tumors separately (234, 235). **Chapter 4** describes the usage of *BRCA1* specific chromosomal aberrations to identify *BRCA1*-associated breast tumors from a cohort in which no *BRCA1/2* mutations had been found by routine diagnostics. **Chapter 5** describes a similar process, but for *BRCA2*-association.

BRCAness in sporadic breast cancer

In contrast to other cancer predisposition genes, neither *BRCA1* nor *BRCA2* has been found to be mutated in sporadic breast cancer (186, 187). It has however, come to the attention of investigators that a small subset of sporadic breast cancers is remarkably similar in many aspects to *BRCA1*-mutated tumors, this cancer group is known as the basal-like breast cancer subtype (236, 237). Tumors of the basal phenotype are seen in 2-18% of sporadic breast tumors. They show IHC positivity for basal intermediate filaments (*e.g.*, K5, K14), are usually of high grade with large central acellular zones comprising necrosis, tissue infarction, collagen, and hyaline material, and are generally estrogen receptor (ER), progesterone receptor (PR), and ERBB2 negative (152, 220). Not only phenotypically but also genetically, sporadic basal-like tumors are similar to hereditary *BRCA1*-mutated breast tumors, as has been shown by genome wide gene expression profiling (238). **Chapter 6** in this thesis shows that also a set of characteristic cytogenetic changes in *BRCA1*-associated breast cancers can be found in sporadic basal-like breast tumors. These microarray studies suggest that similar cancer promoting pathways may lead to the development of these two tumor groups (239). It is therefore hypothesized that sporadic breast cancer displaying genomic similarities with hereditary *BRCA*-mutated tumors, exhibit dysfunctional BRCA pathways and therefore deregulated homologous recombination. Tumors with homologous recombination deficiency (HRD) are highly sensitive to agents inducing DNA double strand breaks. **Chapter 8** of this thesis will, among other subjects, discuss how prediction of *BRCA*-association can assist in clinical care of sporadic breast cancer patients by demonstrating its predictive value for therapy response and survival.

BRCAX

In many of the breast cancer families, no mutation is found in any of the known breast cancer susceptibility genes and so far, identification of a third *BRCA* gene has been unsuccessful. It is therefore likely that the non-BRCA1/2 breast tumors are a heterogeneous group consisting of a collection of low penetrance genes or a combination of common variants with multiplicative effects on breast cancer risk (114, 134-137). To be able to locate potential loci on which breast cancer susceptibility genes might be located, homogeneous groups have to be identified first; however, non-BRCA1/2 breast cancer families do not show any typical phenotypes such as the *BRCA1* or *BRCA2* families do. Because it has been shown that *BRCA1*- and *BRCA2*-mutated breast tumors display distinctive chromosomal aberrations (**Chapter 4 & 5**), it might be possible that tumors caused by other breast cancer susceptibility genes (*i.e.*, *BRCAX*) could also display such characteristic profiles. **Chapter 7** discusses the use of array CGH with the aim to describe more homogeneous groups in non-BRCA1/2 families, which could be applied for linkage studies in the future.

References

1. Hanahan D, et al. 2011. Hallmarks of cancer: the next generation. Cell 144:646-74.
2. Klein CA. 2008. Cancer. The metastasis cascade. Science 321:1785-7.
3. Chiang AC, et al. 2008. Molecular basis of metastasis. N Engl J Med 359:2814-23.
4. Eccles SA, et al. 2007. Metastasis: recent discoveries and novel treatment strategies. Lancet 369:1742-57.
5. World Health Origanization. The global burden of disease: 2004 update. WHO Press; 2008. ISBN 978 92 4 156371 0
6. World Health Origanization. 2007. Cancer prevention. Available from http://www.who.int/cancer/prevention/en/
7. Joensuu H. 2008. Systemic chemotherapy for cancer: from weapon to treatment. Lancet Oncol 9:304.
8. American Joint Committee on Cancer. 30-3-2010. http://www.cancerstaging.org
9. Ferlay J., et al. 2008. GLOBOCAN 2008, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 10. Lyon, France: International Agency for Research on Cancer; 2010 Available from: http://globocan iarc fr .
10. Jemal A, et al. 2011. Global cancer statistics. CA Cancer J Clin .
11. American Cancer Society. 2010. Cancer Facts & Figures 2010. Atlanta: American Cancer Society .
12. Tavassoli FA, Devilee P. World Health Organization: Tumours of the Breast and Female Genital Organs (IARC WHO Classification of Tumors). 1 ed. IARCPress-WHO; 2003
13. Genestie C, et al. 1998. Comparison of the prognostic value of Scarff-Bloom-Richardson and Nottingham histological grades in a series of 825 cases of breast cancer: major importance of the mitotic count as a component of both grading systems. Anticancer Res 18:571-6.
14. Weiss NS. 2003. Breast cancer mortality in relation to clinical breast examination and breast self-examination. Breast J 9 Suppl 2:S86-S89.
15. Kopans DB. 2009. Mammography: yet another challenge. Radiology 253:587-9.
16. Harris JR, Lippman ME, Osborne CK, Morrow M. Breast imaging and image-guided biopsy techniques. Diseases of the Breast, 4th Edition. Lippincott Williams & Wilkins; 2010.
17. Pantel K, et al. 2010. Circulating tumour cells in cancer patients: challenges and perspectives. Trends Mol Med 16:398-406.
18. Harris JR, Lippman ME, Osborne CK, Morrow M. Management of primary invasive breast cancer. In: Lippincott Williams & Wilkins, editor. Diseases of the Breast, 4th edition. 2010.
19. Bult P, Hoogerbrugge N. Familial breast cancer: detection of prevalent high-risk epithelial lesions. In: Hayat MA, editor. Methods of cancer diagnosis, therapy, and prognosis. Springer; 2008.
20. Harris JR, Lippman ME, Osborne CK, Morrow M. New breast cancer therapeutic approaches. In: Lippincott Williams & Wilkins, editor. Diseases of the Breast, 4th edition. 2010.
21. Chia YH, et al. 2010. Neoadjuvant endocrine therapy in primary breast cancer: indications and use as a research tool. Br J Cancer 103:759-64.
22. Maskarinec G. 2000. Breast cancer--interaction between ethnicity and environment. In Vivo 14:115-23.
23. Andreeva VA, et al. 2007. Breast cancer among immigrants: a systematic review and new research directions. J Immigr Minor Health 9:307-22.

24. Ziegler RG, et al. 1993. Migration patterns and breast cancer risk in Asian-American women. J Natl Cancer Inst 85:1819-27.
25. Kliewer EV, et al. 1995. Breast cancer mortality among immigrants in Australia and Canada. J Natl Cancer Inst 87:1154-61.
26. Lane-Claypon J.E. 1926. A further report on cancer of the breast: reports on public health and medical subjects. Ministry of Health London: Her Majesty's Stationary Office 32:1-189.
27. Velie EM, et al. 2005. Lifetime reproductive and anthropometric risk factors for breast cancer in postmenopausal women. Breast Dis 24:17-35.
28. Ready K, Arun B. Genetic Predisposition to Breast Cancer and Genetic Counseling and Testing. Breast Cancer 2nd edition (M.D. Anderson Cancer Care Series). Springer; 2008. p. 39-54.
29. Claus EB, et al. 1991. Genetic analysis of breast cancer in the cancer and steroid hormone study. Am J Hum Genet 48:232-42.
30. Go RC, et al. 1983. Genetic epidemiology of breast cancer and associated cancers in high-risk families. I. Segregation analysis. J Natl Cancer Inst 71:455-61.
31. Eccles D, et al. 1994. Genetic epidemiology of early onset breast cancer. J Med Genet 31:944-9.
32. Hall JM, et al. 1990. Linkage of early-onset familial breast cancer to chromosome 17q21. Science 250:1684-9.
33. Miki Y, et al. 1994. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. Science 266:66-71.
34. Wooster R, et al. 1994. Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13. Science 265:2088-90.
35. Wooster R, et al. 1995. Identification of the breast cancer susceptibility gene BRCA2. Nature 378:789-92.
36. Domchek SM, et al. 2006. Clinical management of BRCA1 and BRCA2 mutation carriers. Oncogene 25:5825-31.
37. Nelen MR, et al. 1996. Localization of the gene for Cowden disease to chromosome 10q22-23. Nat Genet 13:114-6.
38. Malkin D, et al. 1990. Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. Science 250:1233-8.
39. Jenne DE, et al. 1998. Peutz-Jeghers syndrome is caused by mutations in a novel serine threonine kinase. Nat Genet 18:38-43.
40. Savitsky K, et al. 1995. A single ataxia telangiectasia gene with a product similar to PI-3 kinase. Science 268:1749-53.
41. Meijers-Heijboer H, et al. 2002. Low-penetrance susceptibility to breast cancer due to CHEK2(*)1100delC in noncarriers of BRCA1 or BRCA2 mutations. Nat Genet 31:55-9.
42. Nusbaum R, et al. 2006. Susceptibility to breast cancer: hereditary syndromes and low penetrance genes. Breast Dis 27:21-50.
43. Antoniou A, et al. 2003. 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. Am J Hum Genet 72:1117-30.
44. Easton DF, et al. 1994. The genetic epidemiology of BRCA1. Breast Cancer Linkage Consortium. Lancet 344:761.
45. Chen J, et al. 1998. A study of the PTEN/MMAC1 gene in 136 breast cancer families. Hum Genet 102:124-5.
46. FitzGerald MG, et al. 1998. Germline mutations in PTEN are an infrequent cause

of genetic predisposition to breast cancer. Oncogene 17:727-31.
47. Rhei E, et al. 1997. Mutation analysis of the putative tumor suppressor gene PTEN/MMAC1 in primary breast carcinomas. Cancer Res 57:3657-9.
48. Thull DL, et al. 2004. Recognition and management of hereditary breast cancer syndromes. Oncologist 9:13-24.
49. Malone KE, et al. 2006. Prevalence and predictors of BRCA1 and BRCA2 mutations in a population-based study of breast cancer in white and black American women ages 35 to 64 years. Cancer Res 66:8297-308.
50. Newman B, et al. 1988. Inheritance of human breast cancer: evidence for autosomal dominant transmission in high-risk families. Proc Natl Acad Sci U S A 85:3044-8.
51. Frank TS, et al. 2002. Clinical characteristics of individuals with germline mutations in BRCA1 and BRCA2: analysis of 10,000 individuals. J Clin Oncol 20:1480-90.
52. Peto J, et al. 1999. Prevalence of BRCA1 and BRCA2 gene mutations in patients with early-onset breast cancer. J Natl Cancer Inst 91:943-9.
53. Easton DF. 1999. How many more breast cancer predisposition genes are there? Breast Cancer Res 1:14-7.
54. Couch FJ, et al. 1997. BRCA1 mutations in women attending clinics that evaluate the risk of breast cancer. N Engl J Med 336:1409-15.
55. Hogervorst FB, Vreeswijk MP. Percentage of BRCA mutations and UVs detected in women screened in the Dutch hospitals NKI and LUMC. Personal Communication 2011.
56. Shattuck-Eidens D, et al. 1997. BRCA1 sequence analysis in women at high risk for susceptibility mutations. Risk factor analysis and implications for genetic testing. JAMA 278:1242-50.
57. Vos J, et al. 2011. Perceiving cancer-risks and heredity-likelihood in genetic-counseling: how counselees recall and interpret BRCA 1/2-test results. Clin Genet 79:207-18.
58. Hopper JL. 2001. More breast cancer genes? Breast Cancer Res 3:154-7.
59. Meiser B, et al. 2002. What is the impact of genetic counselling in women at increased risk of developing hereditary breast cancer? A meta-analytic review. Soc Sci Med 54:1463-70.
60. Bigenwald RZ, et al. 2008. Is mammography adequate for screening women with inherited BRCA mutations and low breast density? Cancer Epidemiol Biomarkers Prev 17:706-11.
61. Warner E, et al. 2011. Prospective Study of Breast Cancer Incidence in Women With a BRCA1 or BRCA2 Mutation Under Surveillance With and Without Magnetic Resonance Imaging. J Clin Oncol 29:1664-9.
62. Hartmann LC, et al. 2001. Efficacy of bilateral prophylactic mastectomy in BRCA1 and BRCA2 gene mutation carriers. J Natl Cancer Inst 93:1633-7.
63. Meijers-Heijboer H, et al. 2001. Breast cancer after prophylactic bilateral mastectomy in women with a BRCA1 or BRCA2 mutation. N Engl J Med 345:159-64.
64. Rebbeck TR, et al. 2004. Bilateral prophylactic mastectomy reduces breast cancer risk in BRCA1 and BRCA2 mutation carriers: the PROSE Study Group. J Clin Oncol 22:1055-62.
65. Rebbeck TR, et al. 2002. Prophylactic oophorectomy in carriers of BRCA1 or BRCA2 mutations. N Engl J Med 346:1616-22.
66. Kauff ND, et al. 2002. Risk-reducing salpingo-oophorectomy in women with a

BRCA1 or BRCA2 mutation. N Engl J Med 346:1609-15.
67. van Sprundel TC, et al. 2005. Risk reduction of contralateral breast cancer and survival after contralateral prophylactic mastectomy in BRCA1 or BRCA2 mutation carriers. Br J Cancer 93:287-92.
68. Narod SA, et al. 1998. Oral contraceptives and the risk of hereditary ovarian cancer. Hereditary Ovarian Cancer Clinical Study Group. N Engl J Med 339:424-8.
69. Walker GR, et al. 2002. Family history of cancer, oral contraceptive use, and ovarian cancer risk. Am J Obstet Gynecol 186:8-14.
70. King MC, et al. 2001. Tamoxifen and breast cancer incidence among women with inherited mutations in BRCA1 and BRCA2: National Surgical Adjuvant Breast and Bowel Project (NSABP-P1) Breast Cancer Prevention Trial. JAMA 286:2251-6.
71. De Soto JA, et al. 2006. PARP-1 inhibitors: are they the long-sought genetically specific drugs for BRCA1/2-associated breast cancers? Int J Med Sci 3:117-23.
72. Cortez D, et al. 1999. Requirement of ATM-dependent phosphorylation of brca1 in the DNA damage response to double-strand breaks. Science 286:1162-6.
73. Khanna KK, et al. 2001. DNA double-strand breaks: signaling, repair and the cancer connection. Nat Genet 27:247-54.
74. Foulkes WD. 2006. BRCA1 and BRCA2: chemosensitivity, treatment outcomes and prognosis. Fam Cancer 5:135-42.
75. Yap TA, et al. 2011. Poly(ADP-ribose) polymerase (PARP) inhibitors: Exploiting a synthetic lethal strategy in the clinic. CA Cancer J Clin 61:31-49.
76. American Society of Clinical Oncology. 2003. American Society of Clinical Oncology policy statement update: genetic testing for cancer susceptibility. J Clin Oncol 21:2397-406.
77. de Bock GH, et al. 1999. The assessment of genetic risk of breast cancer: a set of GP guidelines. Fam Pract 16:71-7.
78. de Silva D, et al. 1995. Identification of women at high genetic risk of breast cancer through the National Health Service Breast Screening Programme (NHSBSP). J Med Genet 32:862-6.
79. Mouchawar J, et al. 2003. Guidelines for breast and ovarian cancer genetic counseling referral: adoption and implementation in HMOs. Genet Med 5:444-50.
80. 2011. Genetic/Familial High-Risk Assessment: Breast and Ovarian - Version 1.2011
81. Robson ME, et al. 2010. American Society of Clinical Oncology policy statement update: genetic and genomic testing for cancer susceptibility. J Clin Oncol 28:893-901.
82. 2006. The NICE clinical guideline on familial breast cancer
83. Srivastava A, et al. 2001. Risk of breast and ovarian cancer in women with strong family histories. Oncology (Williston Park) 15:889-902.
84. Nelson HD, et al. 2005. Genetic Risk Assessment and BRCA Mutation Testing for Breast and Ovarian Cancer Susceptibility. PMID: 20722133 .
85. Pharoah PD, et al. 1997. Family history and the risk of breast cancer: a systematic review and meta-analysis. Int J Cancer 71:800-9.
86. Stratton JF, et al. 1998. A systematic review and meta-analysis of family history and risk of ovarian cancer. Br J Obstet Gynaecol 105:493-9.
87. Domchek SM, et al. 2003. Application of breast cancer risk prediction models in clinical practice. J Clin Oncol 21:593-601.
88. Berry DA, et al. 1997. Probability of carrying a mutation of breast-ovarian

cancer gene BRCA1 based on family history. J Natl Cancer Inst 89:227-38.
89. Berry DA, et al. 2002. BRCAPRO validation, sensitivity of genetic testing of BRCA1/BRCA2, and prevalence of other breast cancer susceptibility genes. J Clin Oncol 20:2701-12.
90. Parmigiani G, et al. 1998. Determining carrier probabilities for breast cancer-susceptibility genes BRCA1 and BRCA2. Am J Hum Genet 62:145-58.
91. Frank TS, et al. 1998. Sequence analysis of BRCA1 and BRCA2: correlation of mutations with family history and ovarian cancer risk. J Clin Oncol 16:2417-25.
92. Blackwood MA, et al. 2001. Predicted probability of breast cancer susceptibility gene mutations. Breast Cancer Res Treat 69:223.
93. Tyrer J, et al. 2004. A breast cancer prediction model incorporating familial and personal risk factors. Stat Med 23:1111-30.
94. Antoniou AC, et al. 2004. The BOADICEA model of genetic susceptibility to breast and ovarian cancer. Br J Cancer 91:1580-90.
95. Antoniou AC, et al. 2008. The BOADICEA model of genetic susceptibility to breast and ovarian cancers: updates and extensions. Br J Cancer 98:1457-66.
96. BOADICEA . http://www.srl.cam.ac.uk/genepi/boadicea/boadicea_home.html: University of Cambridge. Computer program 2010.
97. BRCAPRO (1) . http://www4.utsouthwestern.edu/breasthealth/cagene/: CancerGene. Computer program 2004.
98. BRCAPRO (2) . http://www.cyrillicsoftware.com/: CyrillicSoftware. Computer program 2000.
99. IBIS Breast Cancer Risk Evaluation Tool . http://www.ems-trials.org/riskevaluator/: Cuzick J. Computer program 2008.
100. Myriad . https://www.myriadpro.com/brca-risk-calculator: BRCA Risk Calculator. Computer program 2010.
101. Pen II . http://www.afcri.upenn.edu/itacc/penn2: University of Pennsylvania. Computer program 2011.
102. Antoniou AC, et al. 2008. Predicting the likelihood of carrying a BRCA1 or BRCA2 mutation: validation of BOADICEA, BRCAPRO, IBIS, Myriad and the Manchester scoring system using data from UK genetics clinics. J Med Genet 45:425-31.
103. Panchal SM, et al. 2008. Selecting a BRCA risk assessment model for use in a familial cancer clinic. BMC Med Genet 9:116.
104. Lindor NM, et al. 2010. Predicting BRCA1 and BRCA2 gene mutation carriers: comparison of PENN II model to previous study. Fam Cancer 9:495-502.
105. Claus EB, et al. 1993. The calculation of breast cancer risk for women with a first degree family history of ovarian cancer. Breast Cancer Res Treat 28:115-20.
106. Gail MH, et al. 1989. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. J Natl Cancer Inst 81:1879-86.
107. Jonker MA, et al. 2003. Modeling familial clustered breast cancer using published data. Cancer Epidemiol Biomarkers Prev 12:1479-85.
108. van Asperen CJ, et al. 2004. Risk estimation for healthy women from breast cancer families: new insights and new strategies. Cancer Epidemiol Biomarkers Prev 13:87-93.
109. Gilpin CA, et al. 2000. A preliminary validation of a family history assessment form to select women at risk for breast or ovarian cancer for referral to a genetics center. Clin Genet 58:299-308.

110. Evans DG, et al. 2004. A new scoring system for the chances of identifying a BRCA1/2 mutation outperforms existing models including BRCAPRO. J Med Genet 41:474-80.
111. Emery J, et al. 1999. Computer support for recording and interpreting family histories of breast and ovarian cancer in primary care (RAGs): qualitative evaluation with simulated patients. BMJ 319:32-6.
112. Ferla R, et al. 2007. Founder mutations in BRCA1 and BRCA2 genes. Ann Oncol 18 Suppl 6:vi93-vi98.
113. Bellosillo B, et al. 2006. Pitfalls and caveats in BRCA sequencing. Ultrastruct Pathol 30:229-35.
114. Narod SA, et al. 2004. BRCA1 and BRCA2: 1994 and beyond. Nat Rev Cancer 4:665-76.
115. Evans DG, et al. 2003. Sensitivity of BRCA1/2 mutation testing in 466 breast/ovarian cancer families. J Med Genet 40:e107.
116. The Breast Cancer Information Core Database (BIC). http://research.nhgri.nih.gov/projects/bic/ . Online Source 2011.
117. Walsh T, et al. 2006. Spectrum of mutations in BRCA1, BRCA2, CHEK2, and TP53 in families at high risk of breast cancer. JAMA 295:1379-88.
118. Szabo CI, et al. 2004. Understanding germ-line mutations in BRCA1. Cancer Biol Ther 3:515-20.
119. Struewing JP, et al. 1997. The risk of cancer associated with specific mutations of BRCA1 and BRCA2 among Ashkenazi Jews. N Engl J Med 336:1401-8.
120. Struewing JP, et al. 1995. The carrier frequency of the BRCA1 185delAG mutation is approximately 1 percent in Ashkenazi Jewish individuals. Nat Genet 11:198-200.
121. Roa BB, et al. 1996. Ashkenazi Jewish population frequencies for common mutations in BRCA1 and BRCA2. Nat Genet 14:185-7.
122. Tulinius H, et al. 2002. The effect of a single BRCA2 mutation on cancer in Iceland. J Med Genet 39:457-62.
123. Petrij-Bosch A, et al. 1997. BRCA1 genomic deletions are major founder mutations in Dutch breast cancer patients. Nat Genet 17:341-5.
124. Thorlacius S, et al. 1996. A single BRCA2 mutation in male and female breast cancer families from Iceland with varied cancer phenotypes. Nat Genet 13:117-9.
125. Zeegers MP, et al. 2004. Founder mutations among the Dutch. Eur J Hum Genet 12:591-600.
126. Spurdle AB. 2010. Clinical relevance of rare germline sequence variants in cancer genes: evolution and application of classification models. Curr Opin Genet Dev 20:315-23.
127. Webb M. 2010. Developing functional assays for BRCA1 unclassified variants. Methods Mol Biol 653:281-91.
128. Tavtigian SV, et al. 2008. Classification of rare missense substitutions, using risk surfaces, with genetic- and molecular-epidemiology applications. Hum Mutat 29:1342-54.
129. Vreeswijk MP, et al. 2009. Intronic variants in BRCA1 and BRCA2 that affect RNA splicing can be reliably selected by splice-site prediction programs. Hum Mutat 30:107-14.
130. Goldgar DE, et al. 2004. Integrated evaluation of DNA sequence variants of unknown clinical significance: application to BRCA1 and BRCA2. Am J Hum Genet 75:535-44.
131. Goldgar DE, et al. 2008. Genetic evidence and integration of various data sources for

classifying uncertain variants into a single model. Hum Mutat 29:1265-72.
132. Radice P, et al. 2011. Unclassified variants in BRCA genes: guidelines for interpretation. Ann Oncol 22 Suppl 1:i18-i23.
133. Kainu T, et al. 2000. Somatic deletions in hereditary breast cancers implicate 13q21 as a putative novel breast cancer susceptibility locus. Proc Natl Acad Sci U S A 97:9603-8.
134. Oldenburg RA, et al. 2007. Genetic susceptibility for breast cancer: how many more genes to be found? Crit Rev Oncol Hematol 63:125-49.
135. Pharoah PD, et al. 2007. Association between common variation in 120 candidate genes and breast cancer risk. PLoS Genet 3:e42.
136. Turnbull C, et al. 2010. Genome-wide association study identifies five new breast cancer susceptibility loci. Nat Genet 42:504-7.
137. Easton DF, et al. 2007. Genome-wide association study identifies novel breast cancer susceptibility loci. Nature 447:1087-93.
138. Smith P, et al. 2006. A genome wide linkage search for breast cancer susceptibility genes. Genes Chromosomes Cancer 45:646-55.
139. Oldenburg RA, et al. 2006. Characterization of familial non-BRCA1/2 breast tumors by loss of heterozygosity and immunophenotyping. Clin Cancer Res 12:1693-700.
140. Ford D, et al. 1998. Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium. Am J Hum Genet 62:676-89.
141. Paradiso A, et al. 2011. Hereditary breast cancer: clinical features and risk reduction strategies. Ann Oncol 22 Suppl 1:i31-i36.
142. Rennert G, et al. 2007. Clinical outcomes of breast cancer in carriers of BRCA1 and BRCA2 mutations. N Engl J Med 357:115-23.
143. Tutt A, et al. 2010. Oral poly(ADP-ribose) polymerase inhibitor olaparib in patients with BRCA1 or BRCA2 mutations and advanced breast cancer: a proof-of-concept trial. Lancet 376:235-44.
144. Zhong Q, et al. 1999. Association of BRCA1 with the hRad50-hMre11-p95 complex and the DNA damage response. Science 285:747-50.
145. Xia F, et al. 2001. Deficiency of human BRCA2 leads to impaired homologous recombination but maintains normal nonhomologous end joining. Proc Natl Acad Sci U S A 98:8644-9.
146. Xu B, et al. 2001. Involvement of Brca1 in S-phase and G(2)-phase checkpoints after ionizing irradiation. Mol Cell Biol 21:3445-50.
147. Chistiakov DA, et al. 2008. Genetic variations in DNA repair genes, radiosensitivity to cancer and susceptibility to acute tissue reactions in radiotherapy-treated cancer patients. Acta Oncol 47:809-24.
148. Eerola H, et al. 2008. Basal cytokeratins in breast tumours among BRCA1, BRCA2 and mutation-negative breast cancer families. Breast Cancer Res 10:R17.
149. Honrado E, et al. 2005. The molecular pathology of hereditary breast cancer: genetic testing and therapeutic implications. Mod Pathol 18:1305-20.
150. Atchley DP, et al. 2008. Clinical and pathologic characteristics of patients with BRCA-positive and BRCA-negative breast cancer. J Clin Oncol 26:4282-8.
151. Dent R, et al. 2007. Triple-negative breast cancer: clinical features and patterns of recurrence. Clin Cancer Res 13:4429-34.
152. Lakhani SR, et al. 2005. Prediction of BRCA1 status in patients with breast cancer

using estrogen receptor and basal phenotype. Clin Cancer Res 11:5175-80.
153. Paterson JW. 1998. BRCA1: a review of structure and putative functions. Dis Markers 13:261-74.
154. O'Donovan PJ, et al. 2010. BRCA1 and BRCA2: breast/ovarian cancer susceptibility gene products and participants in DNA double-strand break repair. Carcinogenesis 31:961-7.
155. Brzovic PS, et al. 2001. Structure of a BRCA1-BARD1 heterodimeric RING-RING complex. Nat Struct Biol 8:833-7.
156. Joukov V, et al. 2001. Functional communication between endogenous BRCA1 and its partner, BARD1, during Xenopus laevis development. Proc Natl Acad Sci U S A 98:12078-83.
157. Ruffner H, et al. 2001. Cancer-predisposing mutations within the RING domain of BRCA1: loss of ubiquitin protein ligase activity and protection from radiation hypersensitivity. Proc Natl Acad Sci U S A 98:5134-9.
158. Wang B, et al. 2007. Abraxas and RAP80 form a BRCA1 protein complex required for the DNA damage response. Science 316:1194-8.
159. Sobhian B, et al. 2007. RAP80 targets BRCA1 to specific ubiquitin structures at DNA damage sites. Science 316:1198-202.
160. Yu X, et al. 2004. DNA damage-induced cell cycle checkpoint control requires CtIP, a phosphorylation-dependent binding partner of BRCA1 C-terminal domains. Mol Cell Biol 24:9478-86.
161. Xu B, et al. 2002. Phosphorylation of serine 1387 in Brca1 is specifically required for the Atm-mediated S-phase checkpoint after ionizing irradiation. Cancer Res 62:4588-91.
162. Joukov V, et al. 2006. The BRCA1/BARD1 heterodimer modulates ran-dependent mitotic spindle assembly. Cell 127:539-52.
163. Lou Z, et al. 2005. BRCA1 participates in DNA decatenation. Nat Struct Mol Biol 12:589-93.
164. Venkitaraman AR. 2002. Cancer susceptibility and the functions of BRCA1 and BRCA2. Cell 108:171-82.
165. Deng CX. 2006. BRCA1: cell cycle checkpoint, genetic instability, DNA damage response and cancer evolution. Nucleic Acids Res 34:1416-26.
166. Holloman WK. 2011. Unraveling the mechanism of BRCA2 in homologous recombination. Nat Struct Mol Biol 18:748-54.
167. Yang H, et al. 2002. BRCA2 function in DNA binding and recombination from a BRCA2-DSS1-ssDNA structure. Science 297:1837-48.
168. Patel KJ, et al. 1998. Involvement of Brca2 in DNA repair. Mol Cell 1:347-57.
169. Thorslund T, et al. 2007. BRCA2: a universal recombinase regulator. Oncogene 26:7720-30.
170. Rossi DJ, et al. 2008. Stems cells and the pathways to aging and cancer. Cell 132:681-96.
171. Clevers H. 2011. The cancer stem cell: premises, promises and challenges. Nat Med 17:313-9.
172. Sell S, et al. 1994. Maturation arrest of stem cell differentiation is a common pathway for the cellular origin of teratocarcinomas and epithelial cancers. Lab Invest 70:6-22.
173. Dalerba P, et al. 2007. Cancer stem cells: models and concepts. Annu Rev Med 58:267-84.
174. Groner B, et al. 2010. Stem cells of the breast and cancer therapy. Womens Health (Lond Engl) 6:205-19.
175. Lim E, et al. 2009. Aberrant luminal progenitors as the candidate target population for basal tumor development in

BRCA1 mutation carriers. Nat Med 15:907-13.
176. Smith GH, et al. 1988. A morphologically distinct candidate for an epithelial stem cell in mouse mammary gland. J Cell Sci 90 (Pt 1):173-83.
177. Lamb R, et al. 2007. Mammary development, carcinomas and progesterone: role of Wnt signalling. Ernst Schering Found Symp Proc :1-23.
178. Vivanco M. 2010. Function follows form: defining mammary stem cells. Sci Transl Med 2:31ps22.
179. Sleeman KE, et al. 2007. Dissociation of estrogen receptor expression and in vivo stem cell activity in the mammary gland. J Cell Biol 176:19-26.
180. Kendrick H, et al. 2008. Transcriptome analysis of mammary epithelial subpopulations identifies novel determinants of lineage commitment and cell fate. BMC Genomics 9:591.
181. Liu S, et al. 2008. BRCA1 regulates human mammary stem/progenitor cell fate. Proc Natl Acad Sci U S A 105:1680-5.
182. Foulkes WD. 2004. BRCA1 functions as a breast stem cell regulator. J Med Genet 41:1-5.
183. Hoeijmakers JH. 2001. Genome maintenance mechanisms for preventing cancer. Nature 411:366-74.
184. Durant ST, et al. 2005. Good timing in the cell cycle for precise DNA repair by BRCA1. Cell Cycle 4:1216-22.
185. Hakem R. 2008. DNA-damage repair; the good, the bad, and the ugly. EMBO J 27:589-605.
186. Futreal PA, et al. 1994. BRCA1 mutations in primary breast and ovarian carcinomas. Science 266:120-2.
187. Lancaster JM, et al. 1996. BRCA2 mutations in primary breast and ovarian cancers. Nat Genet 13:238-40.
188. Knudson AG, Jr. 1971. Mutation and cancer: statistical study of retinoblastoma. Proc Natl Acad Sci U S A 68:820-3.
189. Dworkin AM, et al. 2009. Methylation not a frequent "second hit" in tumors with germline BRCA mutations. Fam Cancer 8:339-46.
190. Ame JC, et al. 2004. The PARP superfamily. Bioessays 26:882-93.
191. Haber JE. 1999. DNA recombination: the replication connection. Trends Biochem Sci 24:271-5.
192. Ashworth A. 2008. A synthetic lethal therapeutic approach: poly(ADP) ribose polymerase inhibitors for the treatment of cancers deficient in DNA double-strand break repair. J Clin Oncol 26:3785-90.
193. Bryant HE, et al. 2005. Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. Nature 434:913-7.
194. Farmer H, et al. 2005. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. Nature 434:917-21.
195. Fong PC, et al. 2009. Inhibition of poly(ADP-ribose) polymerase in tumors from BRCA mutation carriers. N Engl J Med 361:123-34.
196. Fong PC, et al. 2010. Poly(ADP)-ribose polymerase inhibition: frequent durable responses in BRCA carrier ovarian cancer correlating with platinum-free interval. J Clin Oncol 28:2512-9.
197. Audeh MW, et al. 2010. Oral poly(ADP-ribose) polymerase inhibitor olaparib in patients with BRCA1 or BRCA2 mutations and recurrent ovarian cancer: a proof-of-concept trial. Lancet 376:245-51.
198. Schena M, et al. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 270:467-70.

199. Pinkel D, et al. 1998. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. Nat Genet 20:207-11.
200. Prat A, et al. 2010. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. Breast Cancer Res 12:R68.
201. Perou CM, et al. 2000. Molecular portraits of human breast tumours. Nature 406:747-52.
202. Prat A, et al. 2011. Deconstructing the molecular portraits of breast cancer. Mol Oncol 5:5-23.
203. Neve RM, et al. 2006. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. Cancer Cell 10:515-27.
204. Sorlie T, et al. 2001. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A 98:10869-74.
205. Molyneux G, et al. 2010. BRCA1 basal-like breast cancers originate from luminal epithelial progenitors and not from basal stem cells. Cell Stem Cell 7:403-17.
206. Waddell N, et al. 2010. Gene expression profiling of formalin-fixed, paraffin-embedded familial breast tumours using the whole genome-DASL assay. J Pathol 221:452-61.
207. Waddell N, et al. 2010. Subtypes of familial breast tumours revealed by expression and copy number profiling. Breast Cancer Res Treat 123:661-77.
208. Loeb LA. 2001. A mutator phenotype in cancer. Cancer Res 61:3230-9.
209. Kwei KA, et al. 2010. Genomic instability in breast cancer: pathogenesis and clinical implications. Mol Oncol 4:255-66.
210. Berns EM, et al. 1992. c-myc amplification is a better prognostic factor than HER2/neu amplification in primary breast cancer. Cancer Res 52:1107-13.
211. Kallioniemi A, et al. 1992. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. Science 258:818-21.
212. Richard F, et al. 2000. Patterns of chromosomal imbalances in invasive breast cancer. Int J Cancer 89:305-10.
213. Veltman J, et al. 2008. Breast tumor characteristics of BRCA1 and BRCA2 gene mutation carriers on MRI. Eur Radiol 18:931-8.
214. Agnarsson BA, et al. 1998. Inherited BRCA2 mutation associated with high grade breast cancer. Breast Cancer Res Treat 47:121-7.
215. Armes JE, et al. 1998. The histologic phenotypes of breast carcinoma occurring before age 40 years in women with and without BRCA1 or BRCA2 germline mutations: a population-based study. Cancer 83:2335-45.
216. Breast Cancer Linkage Consortium. 1997. Pathology of familial breast cancer: differences between breast cancers in carriers of BRCA1 or BRCA2 mutations and sporadic cases. Lancet 349:1505-10.
217. Lakhani SR, et al. 1998. Multifactorial analysis of differences between sporadic breast cancers and cancers involving BRCA1 and BRCA2 mutations. J Natl Cancer Inst 90:1138-45.
218. Lakhani SR. 1999. The pathology of familial breast cancer: Morphological aspects. Breast Cancer Res 1:31-5.
219. Abd El-Rehim DM, et al. 2005. High-throughput protein expression analysis using tissue microarray technology of a large well-characterised series identifies biologically distinct classes of breast cancer confirming recent cDNA expression analyses. Int J Cancer 116:340-50.

220. Foulkes WD, et al. 2003. Germline BRCA1 mutations and a basal epithelial phenotype in breast cancer. J Natl Cancer Inst 95:1482-5.
221. Nielsen TO, et al. 2004. Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. Clin Cancer Res 10:5367-74.
222. Osorio A, et al. 2002. Loss of heterozygosity analysis at the BRCA loci in tumor samples from patients with familial breast cancer. Int J Cancer 99:305-9.
223. Holstege H, et al. 2009. High incidence of protein-truncating TP53 mutations in BRCA1-related breast cancer. Cancer Res 69:3625-33.
224. Esteller M, et al. 2001. DNA methylation patterns in hereditary human cancers mimic sporadic tumorigenesis. Hum Mol Genet 10:3001-7.
225. Tirkkonen M, et al. 1997. Distinct somatic genetic changes associated with tumor progression in carriers of BRCA1 and BRCA2 germ-line mutations. Cancer Res 57:1222-7.
226. Bish A, et al. 2002. No news is (not necessarily) good news: impact of preliminary results for BRCA1 mutation searches. Genet Med 4:353-8.
227. Hallowell N, et al. 2002. Genetic testing for women previously diagnosed with breast/ovarian cancer: examining the impact of BRCA1 and BRCA2 mutation searching. Genet Test 6:79-87.
228. Legrand B, et al. 2002. DNA genotyping of unbuffered formalin fixed paraffin embedded tissues. Forensic Sci Int 125:205-11.
229. van Beers EH, et al. 2006. A multiplex PCR predictor for aCGH success of FFPE samples. Br J Cancer 94:333-7.
230. Joosse SA, et al. 2007. Automated array-CGH optimized for archival formalin-fixed, paraffin-embedded tumor material. BMC Cancer 7:43.
231. Wessels LF, et al. 2002. Molecular classification of breast carcinomas by comparative genomic hybridization: a specific somatic genetic profile for BRCA1 tumors. Cancer Res 62:7110-7.
232. Jonsson G, et al. 2005. Distinct genomic profiles in hereditary breast tumors identified by array-based comparative genomic hybridization. Cancer Res 65:7612-21.
233. van Beers EH, et al. 2005. Comparative genomic hybridization profiles in human BRCA1 and BRCA2 breast tumors highlight differential sets of genomic aberrations. Cancer Res 65:822-7.
234. Joosse SA, et al. 2009. Prediction of BRCA1-association in hereditary non-BRCA1/2 breast carcinomas with array-CGH. Breast Cancer Res Treat 116:479-89.
235. Joosse SA, et al. 2010. Prediction of BRCA2-association in hereditary breast carcinomas using array-CGH. Breast Cancer Res Treat .
236. Rakha EA, et al. 2008. Basal-like breast cancer: a critical review. J Clin Oncol 26:2568-81.
237. Turner NC, et al. 2006. Basal-like breast cancer and the BRCA1 phenotype. Oncogene 25:5846-53.
238. Sorlie T, et al. 2003. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proc Natl Acad Sci U S A 100:8418-23.
239. Joosse SA, et al. 2010. Genomic signature of BRCA1 deficiency in sporadic basal-like breast tumors. Genes Chromosomes Cancer

Chapter 2

A multiplex PCR predictor for aCGH success of FFPE samples.

EH van Beers, **SA Joosse**, MJ Ligtenberg, R Fles,
FB Hogervorst, S Verhoef and PM Nederlof

British Journal of Cancer. 2006 Jan; 94(2):333-7.

A multiplex PCR predictor for aCGH success of FFPE samples

Erik H van Beers, **Simon A Joosse**, Marjolijn J Ligtenberg, Renske Fles,
Frans BL Hogervorst, Senno Verhoef and Petra M Nederlof

Formalin-fixed, paraffin-embedded (FFPE) tissue archives are the largest and longest time-spanning collections of patient material in pathology archives. Methods to disclose information with molecular techniques, such as array comparative genomic hybridization (aCGH) have rapidly developed but are still not optimal. Array comparative genomic hybridization is one efficient method for finding tumor suppressors and oncogenes in solid tumors, and also for classification of tumors. The fastest way of analyzing large numbers of tumors is through the use of archival tissue samples with first, the huge advantage of larger median follow-up time of patients studied and second, the advantage of being able to locate and analyze multiple tumors, even across generations, from related individuals (families). Unfortunately, DNA from archival tissues is not always suitable for molecular analysis due to insufficient quality. Until now, this quality remained undefined. We report the optimization of a genomic-DNA isolation procedure from FFPE pathology archives in combination with a subsequent multiplex PCR-based quality-control that simply identified all samples refractory to further DNA-based analyses.

British Journal of Cancer. 2006 Jan 30; 94(2):333-7.

Background

Cancer cytogenetics has benefited greatly from the introduction of comparative genomic hybridization (CGH) for mapping chromosomal gains and losses at a genome-wide scale (1, 2). Subsequent development of the technique into array-CGH (also named matrix-CGH) has allowed increased automation, improved reproducibility and precision due to more accurate mapping of aberrations. This technology has been applied successfully to characterize congenital abnormalities at unprecedented precision (3) and to characterize and classify tumors (4, 5).

In most pathology laboratories, large archives of formalin-fixed, paraffin embedded (FFPE) material are often the only source of material for cancer research. It is our experience (5, 6) that a proportion of archival specimens appear unsuited for aCGH analysis, which is troublesome because array comparative genomic hybridization (aCGH) experiments are tedious and expensive. In the past, we have noticed that this was not solved by repeating aCGH experiments, even when DNA was isolated from new sections from the same tissue blocks (6). Nevertheless, it is possible to obtain high-quality data using archival DNA samples in array CGH experiments (Figure 1) (1, 7-11), even from 20-year-old tissue blocks, provided that robust

procedures, high-quality reagents and 'good' sample DNA quality are being used. A 'good sample quality' definition and an assay to determine this FFPE DNA sample quality would therefore be of great value.

Molecular biological assays, including aCGH on FFPE archival specimens, would be more efficient when good and bad quality samples were identified prior to aCGH assays, by a quick, cheap, simple and reliable assay. Variability, mostly in sample fixation (time), and also duration of storage affect DNA quality. Improvements in many pathology laboratories in sample handling, including shortening of the fixation duration to 24–48 hours and using buffered formalin may have contributed to the increased quality of tissue-extracted DNA (12). In an attempt to predict the success of aCGH hybridization, many laboratories have assessed DNA quality by DNA gel electrophoresis. Although such analyses provide information on the size, amount and distribution of the fragment sizes of the (partially) degraded DNA, this did not correlate well with aCGH success in our hands. Our hypothesis is that apart from the fragment length, DNA crosslinks caused by fixation are of major importance for hybridization results. We therefore focused on improvement of the DNA isolation method to reduce DNA crosslinks, and on an assay to determine the abundance of DNA crosslinks as a measure of DNA quality. This prompted us to evaluate retrospectively our good and bad aCGH experiments and devise a method that indicates DNA quality and aCGH success. This resulted in a modified DNA isolation method and a quality test using a multiplex-PCR assay for sample DNA quality control together with measurement of specific labeling of Cyanin cis-platinum-labeled nucleotides in the test DNA.

Materials and Methods

DNA isolation

Genomic DNA was isolated from 10x 10 µm-thick paraffin-embedded tissue sections. Sections were deparaffinated twice for 5 min in xylene, rehydrated in 100, 96 and 70% ethanol for 30 s each, stained with haematoxylin for 30 s, rinsed with water and incubated overnight in 1 M NaSCN at 37°C to remove crosslinks. Slides were rinsed twice 10 min in 1 x PBS at room temperature, and completely air-dried. Tumor tissue was scraped from the glass with a scalpel to obtain at least 70% tumor cells in 200 µl Qiagen ATL buffer (QIAamp® DNA extraction kit cat. 51306), transferred to eppendorf tubes and incubated with 27 µl proteinase-K (20 mg/ml stock) at 450 rpm (Eppendorf® Thermomixer R) at 55°C. Three more aliquots of 27 µl proteinase-K were added at 4, 20 and 28 h. After a total protK incubation of ~44 h, DNA isolation proceeded as in the manufacturer's protocol (Qiagen, Cat. 51306). Samples of isolated genomic DNA were analyzed by 0.8% agarose gel electrophoresis to visualize DNA concentration and size distribution. In case of tumor tissue, we scraped regions containing at least 70% tumor as indicated by an experienced breast cancer pathologist. aCGH reference DNA was isolated from peripheral blood lymphocytes from six apparently healthy female individuals. It was pooled and sonicated until its median fragment length was similar to that of the test samples.

Multiplex PCR

We analyzed 100 ng as measured by optical density at 260/280 nm of each archival genomic DNA sample by a multiplex PCR. The PCR reaction was performed with four primer sets

that produce 100, 200, 300 and 400 bp fragments from non-overlapping target sites in the GAPDH gene (chr12) in 30 µl with final concentrations: 0.133 µM of each of the following eight 5'–3' primers: 100F gttccaatatgattccaccc; 100R ctcctggaagatggtgatgg; 200F aggtggagcgaggctagc; 200R ttttgcggtggaaatgtcct; 300F aggtggacattcttgctgg; 300R tccactaaccagtcagcgtc; 400F acagtccatgccatcactgc and 400R gcttgacaaagtggtcgttg in a reaction with 10 mM Tris-HCl pH 8.8, 1.5 mM $MgCl_2$, 75 mM KCl, 0.2 mM dNTPs, 1 U Taq DNApolymerase (Invitrogen cat. 18038-26). PCR was performed in thin-wall tubes in an MJ Research PCR apparatus for 4 min 94°C, 35 cycles each of 1 min 94°C, 1 min 56°C and 3 min 72°C, followed by 7 min 72°C ending at 51°C. After addition of 6 µl (5x) loading dye, 10 µl of each sample was analyzed on a 1.5% TBE agarose ethidium bromide-stained gel. Samples were classified based on the largest of four possible PCR products detected, namely 100, 200, 300 and 400 bp. The GAPDH genomic target for amplification is more or less arbitrary but the lengths of the products were purposely chosen based on earlier experience with FFPE DNA amplification (MJL, unpublished results).

Genomic DNA labeling

All labeling reactions were performed with the Cy3 and Cy5 conjugates from the Universal Linkage System (ULS, Kreatech Biotechnology, Amsterdam the Netherlands) (13). Labeling efficiency for ULS-Cy3 and ULS-Cy5 was calculated from A260 (DNA), A280 (protein), A550 (Cy3) and A649 (Cy5) after removal of unbound ULS, on a NanoDropsND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA). The degree of labeling (DOL) was calculated from the specific molar extinction ratios for Cy3, Cy5 and DNA and must be between 1 and 4% (between 1 and 4 ULS molecules per 100 bp) for optimal hybridization signals.

Array CGH

The human 3600 BAC/PAC genomic clone set, covering the full genome at 1 Mb spacing used for the production of our arrays, was obtained from the Welcome Trust Sanger Institute (http://www.sanger.ac.uk/). Information on this clone set can be obtained at the BAC/PAC Resources Center Web Site (http://bacpac.chori.org). Degenerate oligonucleotide PCR-products from all BAC clones were prepared for spotting on CodeLink‡ Activated Slides (Amersham Biosciences, Prod. No. 300011 00) according to detailed protocols (14) with some modifications (15). All clones (three replicates for each probe) were spotted in randomized fashion across 48 sub arrays, each containing 270 spots and hybridized for 48–72 h at 37°C on an orbital shaker (300 rpm) in a humidified chamber with 2 µg tumor-DNA labeled with ULS-Cy5 and 2 µg sonicated lymphocyte control DNA labeled with ULS-Cy3. After washing, arrays were scanned on a Microarray Scanner (G2505B Agilent Technologies), and spots quantified with ImaGene® software (version 6.0.1 BioDiscovery, Marina Del Rey, CA, USA). Computation of the profiles included local background subtraction, Cy5/Cy3 ratio, log2-transformation and sub array normalization to its median. The log2-ratios for all non-flagged spots are then plotted (Figure 2D) along with the standard deviation for each triplicate as smaller dots (red) closer to the X-axis using the secondary y-scale to the right. Bad morphology or uniformity spots were flagged in ImaGene®. When flagged spots accounted for >5% of all spots, hybridizations were excluded. The BAC clones are ordered by position as assigned by NCBI-Build 35 (http://genome.ucsc.

edu/cgi-bin/hgGateway) in the genome beginning at the telomere of 1p and ending at the telomere of Yq.

Results and Discussion

In a systematic approach, we have identified and optimized the selection steps for FFPE archival material to be used in downstream applications, particularly for aCGH.

Formalin-fixed, paraffin-embedded archival tissue DNA quality

In the past, we have used size and size distribution of genomic DNA as a surrogate quality end point. The resulting aCGH profiles were sometimes inconsistent with the estimated sample quality. Figure 1A shows a typical series of 12 isolated genomic DNA samples from FFPE tissue sections. Each lane contains 5 µl (10%) of each isolate. The oldest sample was embedded and stored 26 years before DNA extraction (lane L). The amount of DNA is variable due to the variability in number of nuclei, and dependent on size of the tissue scraped. Furthermore, Figure 1A shows that genomic DNA from archival tissue is severely fragmented with an estimated median DNA fragment size often below 1 kb and varies substantially between samples (cf. lanes B *vs* J). In addition, we observed variability in the size distribution (*i.e.*, long *vs* short smear) between samples (*cf.* lanes B *vs* J).

Multiplex PCR quality assay

An unknown fraction of these samples are refractory to molecular assays including aCGH. The challenge was to identify these samples before performing aCGH. We hypothesized that FFPE samples even after de-crosslinking may still contain DNA crosslinks that prevent specific hybridization and therefore render the sample useless for aCGH. We assumed that with increasing occurrence of DNA crosslinks, the 400, 300, 200 and 100 bp PCR products would become less abundant or even disappear in that order. We thus used the relative amounts of the four possible PCR products as a reporter of DNA quality, and therefore suitability in aCGH. Our quality assay requires 100 ng genomic DNA of each sample in a single multiplex PCR reaction. Representative archival DNA preparations are shown in Figure 1. Two samples (e and l) failed to produce the 100 bp PCR fragment (Figure 1B) and were not successful in subsequent aCGH. Three samples (c, g, i) only produced the 100 bp fragment and each failed in aCGH. All seven samples with a PCR displaying fragments of 200 bp or more were successful in aCGH. Then,

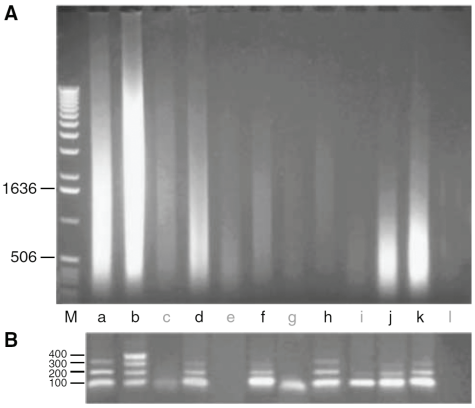


Figure 1 *- A total of 12 unselected DNA preparations from FFPE breast tumors with corresponding multiplex PCR quality controls. (**A**) DNA was isolated from archival blocks stored between 6 (lane b) and 29 (lane l) years. Lane M indicates the molecular size standard (bp). Sample a through l were fixed and stored 11, 6, 22, 20, 18, 11, 8, 7, 19, 17, 16 and 29 years ago, respectively. Lanes in bold a, b, d, f, h, j and k indicate samples with successful aCGH. The oldest samples in this panel successful in aCGH are in lanes d, k and j (20, 17 and 16 years). (**B**) Agarose gel showing multiplex PCR product sizes in bp (see Materials and Methods) for the corresponding samples above.*

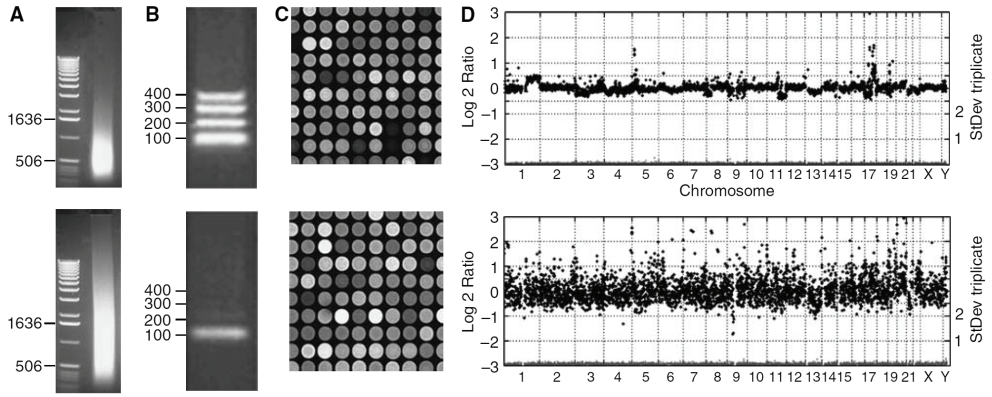


Figure 2 *- aCGH success is determined by the ability to PCR-amplify fragments of > 100 bp from the sample (FFPE) DNA template. (**A**) 0.8% agarose gel electrophoresis shows amount, size and smear-lengths of sample DNA isolated from FFPE tumor tissues. (**B**) Multiplex-PCR reveals whether a 100, 200, 300 or 400 bp fragment are amplified from 100 ng total genomic DNA. (**C**) Representative partial images of array CGH hybridizations. Array CGH was performed on 3500, DOP-amplified BAC-DNA microarrays (see Materials and Methods) printed on Codelink® slides. (**D**) Gain and loss profiles were plotted where the ordinate represents the log2 ratio for the mean of triplicates for each BAC, and abscissa the mapping on the genome (from chromosome 1 to Y, left to right). In red, the standard deviation of the triplicate measurements is plotted to a secondary Y-axis on the right.*

we tested DNA samples retrospectively for cases (N = 26) (Table 1) with known aCGH outcome. We found a good correlation between the ability to obtain PCR products and the quality of the aCGH experiment. There were 24 samples with PCR product and two without (Table 1). The two samples without PCR product as well as two out of three samples with the 100 bp PCR fragment only were not successful in aCGH (*cf.* Figure 2D lower panel). All samples with a 200 bp or greater size PCR fragments resulted in successful aCGH profiles (*cf.* Figure 2D upper panel). Then, in a prognostic approach, we used the multiplex assay outcome to decide when to perform aCGH, that is, aCGH was only performed if a sample had at least the 100 bp PCR fragment (83 of 93 samples). Only six of 37 (16%) samples that had the 100 bp as largest PCR product resulted in good aCGH results. For the samples with 200 bp as the largest product, 38 of 39 (97%) resulted in good aCGH profiles. All seven samples with 300 or 400 bp products were successful in subsequent aCGH. These results indicate that samples without a 100 bp fragment should not be used in aCGH and that DNA samples with amplification of the 200 bp fragment or larger seem to be of sufficient quality for aCGH analysis.

aCGH profiles for FFPE breast tumor samples

Figure 2 illustrates our findings on aspects of DNA quality vs. aCGH success. All four upper panels represent a good quality archival DNA sample, whereas the four lower panels represent a poor quality archival DNA sample. Both panels A show the amount and fragment size distribution for these samples after isolation without further restriction digestion. Even though the DNA fragments from the lower sample are somewhat smaller, both DNA samples theoretically consist of appropriately sized fragments for aCGH. Both panels B show

the result of multiplex quality control PCR using 100 ng of input DNA. It is here, and only here, that we detect the crucial difference between good and bad samples, defined as a minimum of 200 bp amplifiable target sequence. Both panels C show the resulting hybridization and are highly similar in quality (area shown is not the same for both arrays). Finally, panels D show gains (positive log ratios) and losses (negative log ratios) of (parts of) chromosomal material in the breast tumors. The upper panel shows a successful aCGH experiment whereas the lower panel represents a 'noisy' and therefore useless aCGH. Each black dot of the profile represents the mean of three replicates on the same array (triplicate) and the standard deviation of the replicate is plotted below to a secondary Y-scale on the right. Most standard deviations are well below 0.2 and many below 0.1, which indicate very reproducible hybridizations for the good but notably also for the bad DNA sample. The decisive difference between good and bad samples that can be easily scored is the presence of the 200 bp multiplex PCR fragment.

Table 1 *- Correlation between PCR result and subsequent successful array comparative genomic hybridization (aCGH)*

*(**A**) Retrospective correlation of 26 breast tumor formalin-fixed, paraffin-embedded (FFPE) DNA samples aCGH success with performance of their multiplex PCR*

DNA quality *vs* aCGH	**Success (%)**	**Good aCGH**	**Failed aCGH**	***N***
400 bp	100	11	0	11
300 bp	100	8	0	8
200 bp	100	2	0	2
100 bp	33	1	2	3
No product	0	0	2	2
Totals		22	4	26

*(**B**) Prospective correlation of 93 breast tumor FFPE DNA samples aCGH success with performance of their prior multiplex PCR*

Largest product in multiplex PCR	**Success (%)**	**Good aCGH**	**Failed aCGH**	***N***
400 bp	100	2	0	2
300 bp	100	5	0	5
200 bp	97	38	1	39
100 bp	16	6	31	37
No product	ND	—	—	10
Totals		51	32	93

DNA quality from three pathology archives across three decades

With our DNA isolation protocol, we were able to obtain high-quality DNA from the majority of samples from different pathology archival paraffin blocks as old as 25 years. An independent estimation of DNA quality in FFPE samples that almost entirely consisted of samples from our own institute was calculated using a different PCR, in this case generating a 157 bp fragment on 1345 samples, 1264 (94%) of which were positive in this PCR. We found no evidence for different success rates of the 157 bp PCR using samples fixed during the last 25 years studied, whereas DNA from samples fixed before 1970 was often problematic defined by the failure to produce the 157 bp PCR fragment (results not shown). There were 202 of 246 (82%) positive PCR reactions in samples fixed between 1970 and 1980, 666 of 682 (97%) samples fixed between 1980 and 1990, and 397 of 418 (95%) fixed after 1990 (M Schmidt, NKI/AvL personal communication). There appeared to be a surprisingly large difference between the archives that we sampled. We then compared the multiplex PCR quality assessment across three FFPE breast cancer sample series mentioned in this study, that is, 26 retrospective

samples, 93 prospective samples and, the independent study of 1345 breast cancer samples for which PCR success rates were 85, 55 and 94%, respectively. Although, the latter percentage (94%) in this comparison is undoubtedly an overestimation due to the fact that it is only analyzed for production of a 157 bp fragment compared to 200 bp fragment in the other two series, it seems that no *a priori* success rate can be assumed when different archives are being sampled.

*Dpn*II digestion or not?

Array CGH requires that high molecular weight genomic DNA is fragmented to an appropriate fragment size (*e.g.*, by sonication or restriction digestion). Fragmentation can be omitted for aCGH when DNA is isolated from FFPE archival material since it is already fragmented. We compared array CGH using archival DNA with and without prior *Dpn*II restriction and found similar results (data not shown). As expected, DNA gel electrophoresis of archival DNA samples clearly showed the typically fragmented DNA for FFPE samples (Figure 1) explaining why restriction digestion is unnecessary on such samples.

Conclusion

Since concentration and size distribution (as assessed by ethidium bromide agarose gel electrophoresis) of genomic DNA isolated from FFPE tissue are inadequate predictors *per se* for aCGH success, we have developed a method for DNA isolation from FFPE tissue with a subsequent simple and reliable multiplex PCR protocol that predicted successful aCGH with high accuracy. Of our archival samples, 11% (12 out of 107) proved unsuitable for any of the four PCR products and were refractory to aCGH analysis. Furthermore, when genomic DNA was re-isolated from adjacent serial sections of those paraffin tissue blocks that failed the multiplex PCR test and aCGH, both multiplex PCR and aCGH results remained unchanged indicating that DNA suitability for aCGH seems intrinsic to the embedded tissue and is probably related to tissue treatment and duration of storage. Finally, the 157 bp product PCR was used to assess the quality of a much larger set of 1345 DNA samples isolated from three independent pathology archives from samples fixed between 1970 and present. This series was positive for the 157 bp PCR in 94% of the cases, suggesting that aCGH should be widely applicable to archival samples when isolated and selected as indicated above.

Acknowledgements

Marjanka Schmidt for providing PCR results of an independent study of FFPE tumor samples. Hans Tanke, Karoly Szuhai and Jeroen Knijnenburg, LUMC for help with setting up the amplification of the Sanger BAC library and for help with the protocols. Ron Kerkhoven, Mike Heimerikx en Wim Brugman, Central Microarray Facility of the NKI for production of the array. Grant support: Dutch Cancer Society/Koningin Wilhelmina Fonds grant NKB_NKI2001-2424 (EHvB and SAJ).

References

1. Gray JW, et al. 1994. Molecular cytogenetics of human breast cancer. Cold Spring Harb Symp Quant Biol 59:645-52.
2. Kallioniemi OP, et al. 1993. Comparative genomic hybridization: a rapid new method for detecting and mapping DNA

amplification in tumors. Semin Cancer Biol 4:41-6.
3. Veltman JA, et al. 2003. Definition of a critical region on chromosome 18 for congenital aural atresia by arrayCGH. Am J Hum Genet 72:1578-84.
4. Nessling M, et al. 2005. Candidate genes in breast cancer revealed by microarray-based comparative genomic hybridization of archived tissue. Cancer Res 65:439-47.
5. Wessels LF, et al. 2002. Molecular classification of breast carcinomas by comparative genomic hybridization: a specific somatic genetic profile for BRCA1 tumors. Cancer Res 62:7110-7.
6. van Beers EH, et al. 2005. Comparative genomic hybridization profiles in human BRCA1 and BRCA2 breast tumors highlight differential sets of genomic aberrations. Cancer Res 65:822-7.
7. Albertson DG, et al. 2003. Genomic microarrays in human genetic disease and cancer. Hum Mol Genet 12 Spec No 2:R145-R152.
8. Devries S, et al. 2005. Array-based comparative genomic hybridization from formalin-fixed, paraffin-embedded breast tumors. J Mol Diagn 7:65-71.
9. Heidenblad M, et al. 2004. Genome-wide array-based comparative genomic hybridization reveals multiple amplification targets and novel homozygous deletions in pancreatic carcinoma cell lines. Cancer Res 64:3052-9.
10. Loo LW, et al. 2004. Array comparative genomic hybridization analysis of genomic alterations in breast cancer subtypes. Cancer Res 64:8541-9.
11. Ried T, et al. 1995. Comparative genomic hybridization of formalin-fixed, paraffin-embedded breast tumors reveals different patterns of chromosomal gains and losses in fibroadenomas and diploid and aneuploid carcinomas. Cancer Res 55:5415-23.
12. Legrand B, et al. 2002. DNA genotyping of unbuffered formalin fixed paraffin embedded tissues. Forensic Sci Int 125:205-11.
13. Raap AK, et al. 2004. Array comparative genomic hybridization with cyanin cis-platinum-labeled DNAs. Biotechniques 37:130-4.
14. Pinkel D, et al. 1998. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. Nat Genet 20:207-11.
15. Alers JC, et al. 1999. Universal linkage system: an improved method for labeling archival DNA for comparative genomic hybridization. Genes Chromosomes Cancer 25:301-5.

Chapter 3

Automated array-CGH optimized for archival formalin-fixed, paraffin-embedded tumor material.

SA Joosse, EH van Beers and PM Nederlof

BMC Cancer. 2007 Mar; 7:43.

Automated array-CGH optimized for archival formalin-fixed, paraffin-embedded tumor material

Simon A Joosse, Erik H van Beers and Petra M Nederlof

Array Comparative Genomic Hybridization (aCGH) is a rapidly evolving technology that still lacks complete standardization. Yet, it is of great importance to obtain robust and reproducible data to enable meaningful multiple hybridization comparisons. Special difficulties arise when aCGH is performed on archival formalin-fixed, paraffin-embedded (FFPE) tissue due to its variable DNA quality. Recently, we have developed an effective DNA quality test that predicts suitability of archival samples for BAC aCGH. In this report, we first used DNA from a cancer cell line (SKBR3) to optimize the aCGH protocol for automated hybridization, and subsequently optimized and validated the procedure for FFPE breast cancer samples. We aimed for highest throughput, accuracy, and reproducibility applicable to FFPE samples, which can also be important in future diagnostic use. Our protocol of automated array-CGH on archival FFPE ULS-labeled DNA showed very similar results compared with published data and our previous manual hybridization method. This report combines automated aCGH on unamplified archival FFPE DNA using non-enzymatic ULS labeling, and describes an optimized protocol for this combination resulting in improved quality and reproducibility.

BMC Cancer. 2007 Mar 7;7:43.

Background

Array CGH has become a successful and valuable tool for the analysis of chromosome copy-number alterations including the detection of sub-megabase alterations and has been applied to *e.g.*, cell lines, (tumor) tissues, and lymphocytes (1-5). The power of aCGH technology to detect low-level copy number changes is critically dependent on DNA quality (*e.g.*, DNA fragmentation and cross-links) and sample heterogeneity. Therefore, selection of DNA of sufficient quality, especially when using FFPE material, is of great importance for aCGH (6). Furthermore, whole genome amplification may be performed when insufficient DNA is available from a sample (7-9). In addition to sample quality, enzymatic labeling protocols decrease average DNA size further which results in increased noise due to non-specific binding (10), especially when the average PCR length of the sample template drops below 200 bp (6). As an alternative, chemical labeling protocols with cyanin cis-platinum-labeled DNA resulted in good aCGH results (11), also for FFPE archival samples (6). One of the challenges of aCGH is its lower hybridization signal-intensity compared with metaphase-CGH. Based on literature and our previous experiments we hypothesized that the hybridization improves with increasing effective concentration of labeled DNA, which is limited by the viscosity of the hybridization mixture, as well as by the duration and temperature of the hybridization. In this study we

performed in total 70 aCGH hybridizations across these parameters and report how these impact on the CGH profile quality. We first used SKBR3 DNA to explore hybridization variables that are important for aCGH and then show how this expertise can be applied to FFPE primary human tumors.

There have been earlier reports on array CGH of FFPE material (6-8, 12-14) and other reports on automated hybridization (15). This report however, is the first that combines automated hybridization of FFPE tumor material on a BAC array, using non-enzymatic labeling and provides a method without formamide in the post-hybridization washes.

Methods

DNA was isolated from the breast cancer cell line SKBR3 (obtained from ATCC) or from FFPE tumor tissue with at least 70% tumor cells as described before (6). Two micrograms of total genomic DNA were labeled with ULS-Cy5 according to the manufacturer's instructions (Kreatech Biotechnology, Amsterdam). Reference DNA was isolated from lymphocytes of six apparently healthy women, pooled, and sonicated as was done with the SKBR3 genomic DNA to obtain fragments of similar size distribution as DNA from FFPE material (approximately 300-800 bp). Two micrograms of pooled reference DNA were labeled with ULS-Cy3. Corning CodeLink® slides containing the human 3.5k BAC/PAC genomic clone set in triplicate were used as before (6). As optimization target, we used CGH profiles of six FFPE tumors containing at least 70% tumor cells and the SKBR3 cell line profile, obtained by the manual hybridization method described before (6). Automated hybridizations were done in 63.5 x 21 mm chambers in a Tecan HS4800 hybridization station, which uses liquid agitation during hybridization. Experiments involving human tissues were conducted with permission of our institute's medical ethical advisory board.

Optimal (pre-)hybridization mixture

Labeled sample and reference DNA were pooled with 125 µg C_0t-1 DNA (Roche, 1581-074) and precipitated. The pellet was dissolved in 140 µl 0.22 µm filtered hybridization buffer (50% formamide, 15% dextran sulphate (USB 14489, Mw 40-50 kDa), 0.1% Tween20, 2 x SSC, 10 mM Tris pH 7.4, and 25 mM EDTA) and 10 µl (100 µg/µl) yeast tRNA (Sigma, R-8759). The pre-hybridization solution consisted of 400 µg single stranded sheared herring sperm DNA (Sigma, D7290) and 125 µg C_0t-1 DNA dissolved in 150 µl hybridization buffer. Both hybridization and pre-hybridization mixtures were dissolved at 37°C continuously shaking at 650 rpm (Eppendorf Thermomixer) for at least one hour, denatured for 10 min at 95°C and spun for 1 min at 14000 rpm (Eppendorf centrifuge) to pellet potential particles prior to injecting 120 µl pre-hybridization mixture followed by 120 µl sample mixture into the hybridization chamber.

Optimal automated hybridization

Optimal hybridization parameters for the hybridization station: **step 1**; wet the array with 2 x SSC for 30 s at 37°C, no soak. **Step 2**; 120 µl pre-hybridization solution was slowly injected and incubated for 1 hour at 37°C, agitation set at 'high'. **Step 3**; 15 s wash at 37°C with 2 x SSC, no soak. **Step 4**; 120 µl sample mixture was injected and hybridized for 72 hours at 37°C, agitation set

at 'high'. **Step 5**; 12 x (1 min wash, 1 min soak) with 2 x SSC + 0.1% SDS at 37°C. **Step 6**; 6 x (1 min wash + 1 min soak) with 2 x SSC + 0.1% SDS at 68°C. **Step 7**; 2 x (1.5 min wash + 1 min soak) with 2 x SSC at 68°C. **Step 8**; 1.5 min wash with 0.1 x SSC at 23°C, no soak. **Step 9**; 2 min with nitrogen gas at 23°C. Slides were scanned with an Agilent DNA Microarray Scanner BA on the same day. Data processing included signal intensity measurement in ImaGene Software followed by median pintip (*c.q.* subarray) normalization and plotting in custom Matlab code as before (6).

Data analysis

Three statistics were used to determine the quality of the hybridization, the CGH profile, and to compare experiments with each other. For each CGH profile, we calculated the variance across all log2 ratios relative to the ratios of the underlying true ploidy levels as estimated by CGH-segmentation (16), secondly, we defined the dynamic range as the difference between the minimum log2 ratio and the maximum log2 ratio calculated by CGH-segmentation (16), and the average of all the standard deviations of the triplicate spot measurements of each probe was used as a third statistic. Thus, an optimal CGH profile has a low variance to give a better estimate of the copy number level, a high dynamic range to give the best resolution of copy numbers and a low average standard deviation for reproducibility.

GEO

Microarray data have been deposited in NCBIs Gene Expression Omnibus and are accessible through GEO Series accession number GSE7122.

Results and Discussion

Optimization of aCGH on a hybridization station

Further automation of CGH is indispensable to meet the demand for higher quality, higher throughput, and improved reproducibility. We here describe automated and reproducible array-CGH on FFPE material. As optimization goals we aimed to reproduce results from manual hybridizations and published results, to minimize the variance, to maximize dynamic range, to minimize standard deviation of the triplicate spot measurements, and to maximize signal-to-noise.

It is difficult to determine the quality of aCGH profiles without an independent methodology to verify gains and losses. Therefore, we chose to use the widely studied cell line SKBR3 as a model for which chromosomal aberrations have been well documented (2, 3, 17), although the existence of minor sub-clone related alterations cannot be ruled out. In a previous study (6), we performed manual hybridizations of over one hundred BAC arrays that helped us to develop the quality criteria that were now used to optimize automated hybridization. In this study we describe multiple hybridizations that were performed in synchronous pairs with one variable tested in each run, including hybridization duration of 24, 48, or 72 hours, hybridization temperature of 37, 42 or 45°C, pre- and post-hybridization wash temperatures of 37, 42, 45, 46, 65 or 68°C, viscosity of the hybridization mixture with 7, 10, 15, 17.5, or 20% dextran sulphate of 5, 10, or 50 kDa average molecular weight, pH 6, 7 or 8 of the hybridization mixture and with or without pre-hybridization. All hybridization parameters studied are relevant to

nucleic acid hybridization in general, and here optimized for the 3.5k BAC arrayCGH platform and may thus be different for other platforms. Hybridizations were done with genomic DNA isolated from the well-described SKBR3 cells or from FFPE breast tumor archival sections to optimize and validate the protocol.

Optimal conditions for automated hybridization

Using 2 µg unamplified sample DNA from FFPE tissue and 2 µg reference DNA both CyDye labeled, incubation duration was optimal at 72 hours at 37°C after pre-hybridization with a mixture of herring sperm and C_0t-1 DNA for 1 hour at 37°C. The optimal hybridization buffer contained 15% 50 kDa dextran sulphate. Washing was performed as described in Material and Methods. The steps that led to this protocol are described in detail below.

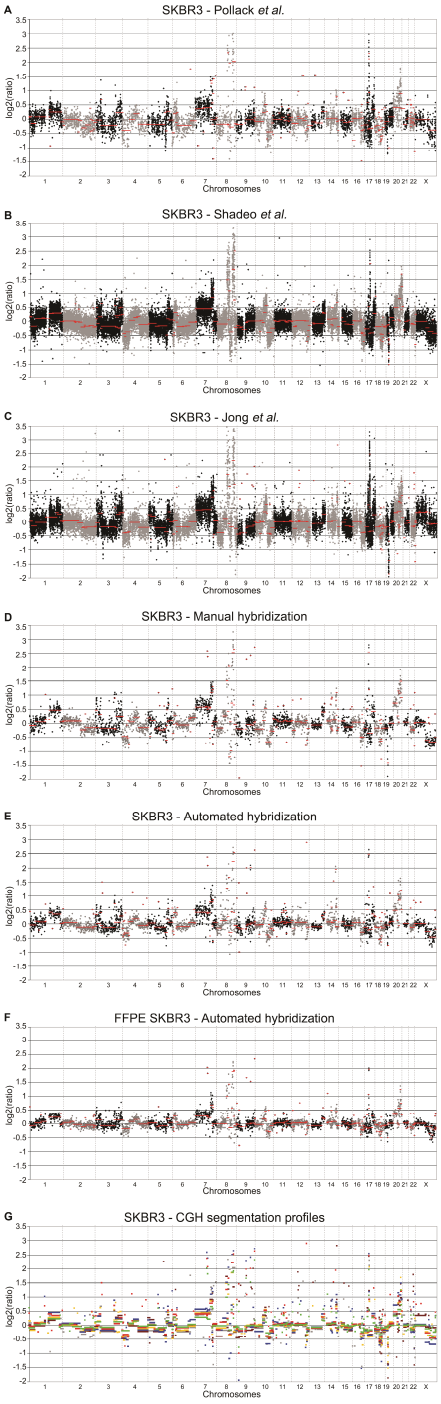


Figure 1 - SKBR3 CGH profiles obtained by various methods. *Chromosomes 1 to X (X-axis, alternate shading per chromosome) versus the log2 ratios (Y-axis) for the breast cancer cell line SKBR3, hybridized by Pollack et al. (PNAS 1999) on a 6.7k cDNA micro array, redrawn form values available on-line (**A**), hybridized by Shadeo and Lam (Breast Cancer Res 2006) on a whole-genome tiling path BAC array containing 32,433 overlapping BAC-derived DNA segments (**B**), hybridized by Jong et al. (Oncogene 2006) to a human oligonucleotide array containing 28,830 unique genes (**C**), manually hybridized to a 3.5k BAC/PAC array (**D**), and hybridized using our optimal automated method (**E**). Formalin-fixed paraffin-embedded SKBR3 cell line hybridized using our automated method (**F**). Red lines in panel A-F represent the breakpoint locations and copy number chances calculated by CGH-segmentation (BMC Bioinformatics 2005). Panel **G** contains all segmentation calls of the profiles depicted in panel A-F: Pollack* et al., *Jong* et al., *Shadeo* et al., *Manual hybridization, Automated hybridization, FFPE material.*

Table 1: Manual and automated hybridization of SKBR3

Sample	Hyb method	Variance	StDev	DR	Correlation
SKBR3	manual	0.13	0.05	3.7	
SKBR3	automated	0.11	0.04	3.9	0.85
SKBR3	automated	0.11	0.07	3.7	
FFPE SKBR3	automated	0.12	0.05	3.2	0.87

*Variance (**Variance**), standard deviation (**StDev**), dynamic range (**DR**) and Pearson correlation (**Correlation**) of SKBR3 performed using our manual and our automated hybridization method (**Hyb method**).*

Optimization model SKBR3 cell line

We used SKBR3 as a model cell line and compared its CGH profile with published (2, 3, 17) and our own manual hybridizations. Figure 1A represents the SKBR3 CGH profile published by Pollack *et al.*, hybridized to a human cDNA micro array containing 6,691 different mapped human genes (2). Figure 1B represents the SKBR3 CGH profile published by Shadeo and Lam, hybridized to a whole-genome tiling path BAC array containing 32,433 overlapping BAC-derived DNA segments (3). Figure 1C represents the SKBR3 CGH profile published by Jong *et al.*, hybridized to a human oligonucleotide array containing 28,830 unique genes (17). Figure 1D represents our manually hybridized SKBR3 CGH profile. Depicted in figure 1E is the very similar SKBR3 CGH profile hybridized with our optimal protocol for the hybridization station except for the slightly different variance and dynamic range. Figure 1F depicts the CGH profile of paraffin embedded SKBR3 (discussed later). To compare these data from different platforms and different methods we looked at the breakpoint locations and copy number estimates as is illustrated in figure 1G. This figure summarizes all the breakpoints and estimated copy number levels as plotted in red in figure 1A-1F calculated by CGH-segmentation (16). Breakpoint locations and calling of copy number levels (gain, unchanged, heterozygous loss, and homozygous loss) are provided as

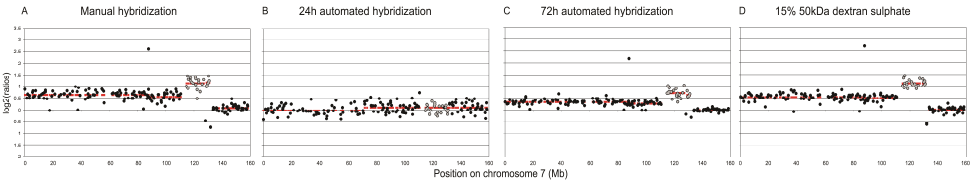


Figure 2 - CGH profiles of manual and three automated methods

*Chromosome 7 (X-axis), versus the log2 ratios (Y-axis) for the breast cancer cell line SKBR3 (alternate shading for aberration legibility); manual hybridization (**A**, detail from figure 1D); 24 hours automated hybridization (**B**); 72 hours automated hybridization (**C**); 72 hours automated hybridization using 15% 50 kDa Dextran Sulphate (**D**). With increasing hybridization duration from 24 to 72 hours, the CGH profile was more similar to the manual hybridization CGH profile. The aberration in panel **C** is detected in contrast to panel B (grey), even better with an increased concentration of dextran sulphate from 7 to 15% (**D**). Red lines in all panel represents the segmentation calls as calculated by CGH-segmentation.*

Additional File 1 and 2. Although a lower density 3.5k BAC array was used, figure 1G illustrates that nearly all aberrations and breakpoint in our results (figure 1D and 1E) are similar to the three published data sets (2, 3, 17). We concluded that the dynamic ranges of both our manual and automated hybridization protocols are adequate to detect single copy number losses and gains. Reproducibility of this automated protocol is shown by replicate hybridizations with a Pearson correlation of 0.85, dynamic ranges of 3.7 versus 3.9, the variances for both experiments 0.11, and the mean standard deviations of 0.07 versus 0.04 (table 1).

Hybridization duration and temperature

The effects of hybridization duration and temperature were measured in two experiments using SKBR3 DNA and reference DNA hybridized for 24 or 48 hours at 37°C. Hybridization mixture containing 7% 50 kDa dextran sulphate was used. Figure 2A shows the CGH profile for SKBR3 chromosome 7 (detail from figure 1B, chosen for its clear and multiple aberrations), hybridized according to our manual method. Figure 2B shows the CGH profile after 24 hours of automated hybridization at 37°C. After 24 hours, no aberrations were detected in this CGH profile. After 48 hours only large copy number changes were found and the small deletions and amplifications were not (data not shown). Also, the dynamic range was small (log2 ratios from –1.0 to 2.2) compared with our manual method (from –1.0 to 2.7). CGH profiles after 24 and 48 hours were inferior to our manual method, this is likely due to lower specific signals. To improve this, the hybridization duration was increased to 72 hours and performed at 37, 42 or 45°C. At all three temperatures, the CGH profiles were approaching the quality of the manual hybridization. Figure 2C shows the result of 37°C, as can be seen it

Table 2: Hybridization duration and temperature

Sample	Duration	Temp.	Variance	StDev	DR
SKBR3	24h	37°C	0.13	0.11	1.9
SKBR3	48h	37°C	0.12	0.06	3.1
SKBR3	72h	37°C	0.07	0.03	3.1
SKBR3	72h	42°C	0.11	0.05	3.6
SKBR3	72h	45°C	0.13	0.11	3.7

*Variance (**Variance**), standard deviation (**StDev**) and dynamic range (**DR**) of the hybridization of SKBR3 under the conditions of different hybridization duration (**Duration**) and temperature (**Temp.**).*

Table 3: Hybridization temperature and dextran sulphate concentration

Sample	Duration	Temp.	DS		Variance	StDev	DR
			conc.	Mw.			
SKBR3	72h	37°C	10%	50kDa	0.11	0.06	3.5
SKBR3	72h	37°C	15%	50kDa	0.11	0.07	3.7
SKBR3	72h	42°C	10%	50kDa	0.14	0.10	3.8
SKBR3	72h	42°C	15%	50kDa	0.12	0.09	3.8

*Variance (**Variance**), standard deviation (**StDev**) and dynamic range (**DR**) for testing the effects of different dextran sulphate concentrations (**DS conc.**) and temperature (**Temp.**) at hybridization duration (**Duration**) of 72 hours, hybridizing SKBR3.*

Table 4: Dextran sulphate concentration

Sample	Duration	Temp.	DS		Variance	StDev	DR
			conc.	Mw.			
SKBR3	72h	37°C	15%	5kDa	0.10	0.04	3.2
SKBR3	72h	37°C	17.5%	5kDa	0.10	0.05	3.7
SKBR3	72h	37°C	20%	5kDa	0.10	0.05	3.4
SKBR3	72h	37°C	15%	10kDa	0.11	0.04	3.7
SKBR3	72h	37°C	17.5%	10kDa	0.11	0.04	3.6
SKBR3	72h	37°C	20%	10kDa	0.11	0.03	3.6
SKBR3	72h	37°C	15%	50kDa	0.11	0.04	3.9
SKBR3	72h	37°C	17.5%	50kDa	0.11	0.04	3.7
SKBR3	72h	37°C	20%	50kDa	0.09	0.03	3.4

*Variance (**Variance**), standard deviation (**StDev**) and dynamic range (**DR**) for testing the effects of different dextran sulphate concentrations (**DS conc.**) and molecular weights (**DS Mw.**), and temperature (**Temp.**) at a hybridization duration (**Duration**) of 72 hours, hybridizing SKBR3.*

is quite similar to the manual hybridization (figure 2A). However, the variances and standard deviations of the triplicate spot measurements increased with hybridization temperature. Mean standard deviations were 0.03, 0.05 and 0.11 at 37, 42 and 45°C respectively. Although 45°C seemed to provide the highest dynamic-range (from -1.1 to 2.6), it was accompanied by the highest noise levels after 72 hours ($p < 0.00001$). The variances were 0.07, 0.11 and 0.13 for 37, 42 and 45°C respectively (table 2). 45°C was therefore excluded from further testing and 37 and 42°C were used to optimize dynamic range in the following experiments.

Hybridization buffer composition

A major further improvement of the hybridization was obtained by increasing the 50 kDa dextran sulphate concentration from 7 to 15%. Here we describe our results for 10 and 15%. Four hybridizations were done at 37 and 42°C each with 10 or 15% dextran sulphate. The resulting profiles were very similar to each other and to our manually hybridized aCGH. A slight systematic difference was detected in the variance, standard deviation of the triplicates and dynamic range. Hybridizing at 37°C, standard deviations were 0.06 and 0.07 at 10 and 15% dextran sulphate, respectively, and at both concentrations the variances were 0.11. At 42°C, both the variances increased to 0.14 and 0.12 and the mean standard deviations to 0.10 and 0.09 at 10 and 15% dextran sulphate, respectively (table 3). Of these four hybridizations, the best profile is shown in figure 2D, this is at 37°C using 15% dextran sulphate. In this experiment, the dynamic range was 3.7 (from –1.0 to 2.7). We chose not to hybridize at 42°C anymore because of the significant higher variance and standard deviation as a result ($p < 0.00001$). We chose to use 15% dextran sulphate in further experiments at 37°C because of its low variance and standard

Table 5: pH of the hybridization buffer

Sample	Hyb pH	Variance	StDev	DR	Max CGHseg	Correlation
FFPE tumor 2	6	0.07	0.04	1.4	0.6	0.98
FFPE tumor 2	6	0.08	0.04	1.6	0.7	
FFPE tumor 2	7	0.09	0.04	1.5	0.7	0.97
FFPE tumor 2	7	0.08	0.05	1.2	0.6	
FFPE tumor 2	8	0.08	0.04	1.3	0.6	0.94
FFPE tumor 2	8	0.10	0.05	1.7	0.8	

*Variance (**Variance**), standard deviation (**StDev**), dynamic range (**DR**), the highest log2 ratio calculated by CGH-segmentation (**Max CGHseg**), and Pearson Correlation (**Correlation**) for testing the effects of pH (**Hyb pH**) of the hybridization buffer on a CGH profile of FFPE tumor #2. The log2 ratios associated with homozygous deletions vary widely since they depend heavily on dividing very small intensities by large intensities. This can have disproportionate impact on the dynamic ranges, therefore only the highest ratio is taken as a measurement for the dynamic range (**Max CGHseg**).*

deviation and its higher dynamic range compared to using 10% dextran sulphate.

Increasing the concentration of 50 kDa dextran sulphate from 15% to 17.5 or 20% did not further improve the array results. At these concentrations the variance was 0.11 and 0.09 at 17.5 and 20% dextran sulphate respectively, notably at 20% dextran sulphate the dynamic range decreased below 3.5. Elevated concentrations of dextran sulphate render the hybridization mixture viscosity beyond the mixing capability of the hybridization station. This prompted us to evaluate the effect of lower molecular weight dextran sulphate (*i.e.*, lower viscosity at the same concentration). We used 5 kDa (Sigma) and 10 kDa (pK Chemicals, Denmark) dextran sulphate at 15, 17.5 or 20% for SKBR3 profiling. All six hybridizations showed inferior dynamic ranges compared with the 50 kDa dextran sulphate experiments, shown in table 4. Therefore, 50 kDa dextran sulphate at a concentration of 15% was used in all subsequent hybridizations.

Post-hybridization washes

Most wash protocols use large amounts of formamide to wash off non-specifically bound probe. Formamide is a toxic that we wished to exclude from all washes. The wash procedure now consists of: **step 5**; 12 x (1 min wash, 1 min soak) with 2 x SSC + 0.1% SDS at the hybridization temperature of 37, 42 or 45°C (previously discussed), **step 6**; 6 x (1 min wash, 1 min soak) with 2 x SSC + 0.1% SDS at 37, 46 or 65°C, **step 7**; 2 x (1.5 min wash, 1 min soak) with 2 x SSC at 37, 46 or 65°C, **step 8**; 15 sec wash with 0.1 x SSC at 23°C, **step 9**; dry slides for 2 minutes with nitrogen gas at 23°C.

As described before, hybridization was performed at 37, 42 or 45°C. Step 5 was done at these temperatures and results are discussed above. Step 6 and 7 were done at 37 or 46°C, both resulting in inferior profiles compared with the manual hybridization. A large proportion of the deletions and amplifications in the CGH profile could not be detected and the data were essentially as in figure 2B. Increasing the tem-

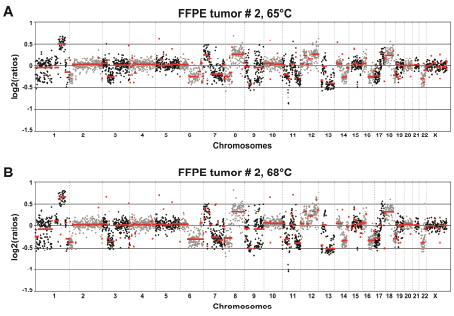


Figure 3 - CGH profiles of one FFPE tumor with post hybridization wash at different temperatures.

*CGH profiles with post-hybridization wash steps 6 and 7 at 65°C (**A**) or at 68°C (**B**) of averaged duplicates of one FFPE primary breast tumor, hybridized according to our optimized protocol for automated array-CGH. Chromosomes (X-axis, alternate shading per chromosome) versus the log2 ratios (Y-axis). At 68°C, dynamic range increased and standard deviation of the triplicate spot measurements decreased compared with 65°C, therefore 68°C was used in our optimal protocol. As can be seen in panel B, the dynamic range and the signal-to-noise are adequate to detect and to distinguish homozygous and heterozygous loss (chromosome 11p), one single-copy number gain (e.g., chromosome 7p), multiple-copy numbers gain (chromosome 1q), and unchanged chromosome copy numbers (e.g., chromosome 10). Red lines in both panel represents the segmentation calls as calculated by CGH-segmentation.*

peratures of steps 6 and 7 to 65°C resulted in good CGH profiles. We concluded that 37°C is the optimum temperature for step 5 and 65°C for steps 6 and 7 when hybridizing a cell line.

FFPE tumor tissue optimization and validation

To develop aCGH also as a diagnostic tool, it will be essential to validate its applicability on patient tumor samples and especially on archival FFPE tissue (18). Extracted DNA from this material is often heavily cross-linked, heterogeneous (*i.e.* mix of cells of different genomic composition), fragmented, and rarely composed of 100% tumor cells. Therefore, aCGH profiles of FFPE material generally have larger variances (defined as the spread around the common levels between adjacent chromosome breakpoints), lower intensities and lower dynamic range compared with hybridizations of cell line DNA.

To validate our automated hybridization method for FFPE material we compared CGH profiles from unfixed and formalin fixed paraffin embedded SKBR3 cells. Figure 1E shows the CGH profile of the formalin fixed SKBR cell line. The fresh and FFPE SKBR3 CGH profiles were highly similar and showed a Pearson correlation of 0.87 (table 1). Variances were 0.11 and 0.12, dynamic ranges 3.9 and 3.2, and mean standard deviations 0.04 and 0.05, for fresh and fixed DNA, respectively. However, the DNA quality from a paraffin embedded cell line does not necessarily represent the quality of DNA from archival tumor tissue that can be more than 25 years old and fixed under widely varying conditions.

Therefore, we validated our method on archival material. The first hybridization was done with tumor #1 DNA with or without pre-hybridization after the first wash step (**step 1**: wetting or chamber filling), for 1 hour at 37°C. The pre-hybridization mixture consisted of 400 µg single stranded sheared herring sperm DNA and 125 µg C_0t-1 DNA dissolved in 150 µl hybridization buffer. With pre-hybridization, signal intensities were almost 50% higher and the mean standard deviation of the triplicate spots 15% lower compared to the protocol without pre-hybridization resulting in good CGH profiles of FFPE material (data not shown). Although CGH profiles of SKBR3 did not improve upon adding pre-hybridization, it clearly benefited CGH profiles of DNA extracted from FFPE patient tissue (data not shown).

Because Tris is the only buffering component in the hybridization mixture, we wished to test the possibility that the formamide could

Table 6: Wash temperature

Sample	Wash	Variance	StDev	DR	Max CGHseg	Correlation
FFPE tumor 2	65°C	0.09	0.04	1.5	0.7	0.97
FFPE tumor 2	65°C	0.08	0.05	1.2	0.6	
FFPE tumor 2	68°C	0.09	0.03	1.6	0.7	0.99
FFPE tumor 2	68°C	0.09	0.03	1.7	0.7	

*Variance (**Variance**), standard deviation (**StDev**), dynamic range (**DR**), the highest log2 ratio calculated by CGH-segmentation (**Max CGHseg**), and Pearson correlation (**Correlation**) of FFPE tumor #2 washed after hybridization at 65°C and 68°C (**Wash**) in replicate.*

react with oxygen and may influence the buffer's pH during storage. To test the effect of pH on the hybridization, six hybridizations with FFPE tumor #2 DNA were done, using hybridization buffers of pH 6, 7 and 8, as measured in the final hybridization buffer. At every pH the CGH profile was very similar and highly reproducible. As can be seen in table 5, standard deviations of the triplicate spot measurements in all six hybridizations were very similar. The variances are lowest at pH 6 but not very different from the variances at pH 7 and 8. For this particular tumor, the maximal CGH-segmentation (16) value was used as dynamic range ("Max CGHseg", table 5), because the homozygous loss on chromosome 11 would have a disproportional contribution to its value (same tumor as in figure 3). These were very similar between experiments. Pearson correlation between the duplicates shows high correlations for all experiments. Therefore, we conclude that aCGH is not very sensitive to pH of the buffer between pH 6 and 8.

Subsequent experiments compared post-hybridization washing at 65°C or 68°C (step 6 and 7) both in duplicate on FFPE extracted material tumor #2. Figure 3A depicts the average profile of two hybridizations washed at 65°C, panel B shows the average CGH profile washed at 68°C, at both temperatures the CGH profiles

Table 7: Manual and automated hybridization of FFPE tumors

Sample	Hyb method	Variance	StDev	DR	Correlation
FFPE tumor 3	manual	0.11	0.06	2.0	0.82
FFPE tumor 3	automated	0.10	0.04	1.9	
FFPE tumor 4	manual	0.10	0.05	1.4	0.72
FFPE tumor 4	automated	0.08	0.03	1.3	
FFPE tumor 5	manual	0.10	0.04	1.5	0.85
FFPE tumor 5	automated	0.08	0.02	1.3	
FFPE tumor 6	manual	0.13	0.06	1.9	0.84
FFPE tumor 6	automated	0.08	0.04	1.9	

*Variance (**Variance**), standard deviation (**StDev**), dynamic range (**DR**) and Pearson correlation (**Correlation**) of four FFPE tumors hybridized with our manual and our automated method (**Hyb method**).*

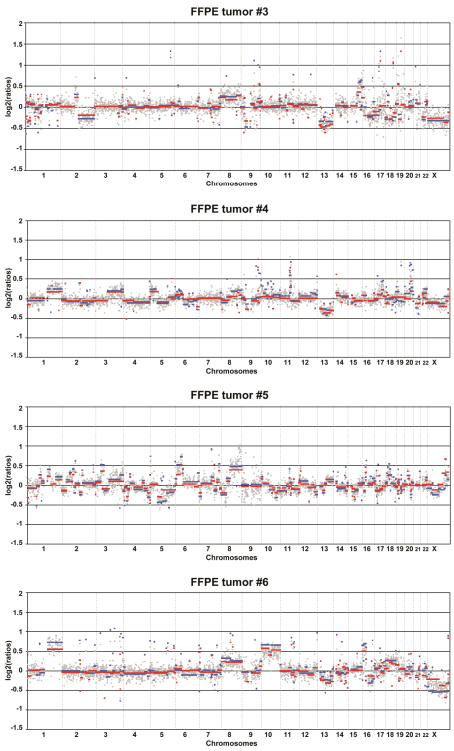


Figure 4 - CGH profiles of four FFPE tumors, hybridized manually and automated.

CGH profiles of four FFPE tumors hybridized using our manual and automated methods. Chromosomes (X-axis) versus the log2 ratios (Y-axis) averaged over the two methods. Breakpoint locations and copy number estimates calculated by CGH-segmentation of the manual hybridized tumors ■ and of the automated hybridized tumors ■ , and the average CGH profiles (log2 ratios) of the manual and the automated hybridized tumor ●.

are very similar. A very small difference could be detected in the mean standard deviation of the triplicate spot measurements as can be seen in table 6, it slightly decreased from 0.04 and 0.05 at 65°C to both 0.03 at 68°C ($p < 0.00001$). Also the dynamic range kept at similar levels (again the highest ratio calculated by CGH-segmentation (16) was used because of the homologous loss in this tumor in chromosome 11 as depicted in figure 3). Although the benefits of changing the temperature from 65°C to 68°C were small, we decided to wash at 68°C.

To validate the optimal automated hybridization described above, we hybridized four FFPE samples (tumor #3, 4, 5 and 6) that were previously hybridized using our manual method. Figure 4 shows the CGH profiles of the FFPE tumors (averaged log2 ratios of the manual and the automated hybridization), with very similar breakpoint locations and copy number estimates (16) for each hybridization method. Variance and standard deviation of the triplicate spot measurements improved slightly but significantly ($p < 0.00001$) for automated compared with manual hybridizations. The dynamic ranges between pairs of manual and automated hybridizations differed by 5%, 7%, 15%, and 0% respectively, and Pearson correlations were 0.82, 0.72, 0.85, and 0.84 (table 7). Although the dynamic ranges are slightly larger due to higher log2 ratios at high-level amplifications using the manual hybridization method (figure 4), these results show that automated and manual CGH profiles are quite similar.

So far, we performed over one hundred automated array-CGH experiments, the oldest archival material used was fixed and embedded in 1971, all with reproducible and high quality results. Figure 3B shows the average profile of one archival FFPE tumor hybridized in duplicate, performed with our optimal protocol for automated aCGH. As can be seen in figure 3, the dynamic range of the hybridizations was adequate to detect and distinguish homozygous and heterozygous loss (chromosome 11p), one single-copy number gain (*e.g.*, chromosome 7p), more then one copy number gain (chromosome 1q) and unchanged chromosome copy numbers (*e.g.*, chromosome 10) in FFPE tumor tissue.

Conclusion

To develop an automated hybridization method, we first used the breast cancer cell line SKBR3 as a model-genome and subsequently optimized and validated the protocol for FFPE breast tumors. Reproducible hybridization results for FFPE tumor tissue were obtained using ULS-labeled unamplified tumor DNA with pre-hybridization, hybridized on a hybridization station at 37°C for 72 hours with a hybridization mixture containing 15% 50 kDa dextran sulphate and post-hybridization washing steps without using formamide. Pre-hybridization did not have a detectable effect on the CGH profile of the cell line SKBR3 but did improve CGH profiles of FFPE tissue samples. All hybridization parameters studied are optimized for the 3.5k BAC array-CGH platform but may be different for other platforms. This protocol of automated array-CGH on archival FFPE ULS-labeled DNA outperformed all our manual methods with respect to accuracy, reproducibility, easy of handling, and speed.

Competing interest

The authors declare that they have no competing interests.

Authors' contributions

SJ performed all experiments, data analyses, participated in the study design and wrote the manuscript. EvB helped in writing and data analyses, and participated in the study design. PN participated in the study design and coordination and helped to draft the manuscript. All authors read and approved the final manuscript.

Additional material

Additional file 1

This file includes the aberrations calculated by CGH-segmentation for the SKBR3 CGH profiles hybridized by Pollack *et al.*, Shadeo and Lam, Jong *et al.*, manually, automatically, and hybridized FFPE SKBR3. Calling copy number changes by setting thresholds allowed comparison of different hybridization platforms and methods.
[http://www.biomedcentral.com/content/supplementary/1471-2407-7-43-S1.txt]

Additional file 2

This picture shows the copy number calls (Y-axis) from SKBR3 CGH profiles per chromosome (X-axis), hybridized by Pollack *et al.*, Shadeo and Lam, Jong *et al.*, manually, automatically, and hybridized FFPE SKBR3. Gain at 1, unchanged at 0, heterozygous loss at -1, and homozygous loss at -2.
[http://www.biomedcentral.com/content/supplementary/1471-2407-7-43-S2.jpeg]

Acknowledgements

We like to acknowledge Ron Kerkhoven, Mike Heimerikx and Wim Brugman (NKI) for their input and production of the microarrays. Renske Fles (NKI) and Hans Tanke, Karoly Szuhai and Jeroen Knijnenburg (LUMC) for help with DOP-amplification of the Sanger BAC library and for their initial Tecan hybridization protocols. Marleen Kok and Juliane Hannemann (NKI) for providing us with the SKBR3 cell line and Juliane Hannemann for critically reading the manuscript.

References

1. Kallioniemi A, et al. 1992. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. Science 258:818-21.
2. Pollack JR, et al. 2002. Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors. Proc Natl Acad Sci U S A 99:12963-8.
3. Shadeo A, et al. 2006. Comprehensive copy number profiles of breast cancer cell model genomes. Breast Cancer Res 8:R9.
4. Gunn SR, et al. 2003. Molecular characterization of a patient with central nervous system dysmyelination and cryptic unbalanced translocation between chromosomes 4q and 18q. Am J Med Genet A 120A:127-35.
5. Ghaffari SR, et al. 1998. A new strategy for cryptic telomeric translocation screening in patients with idiopathic mental retardation. J Med Genet 35:225-33.
6. van Beers EH, et al. 2006. A multiplex PCR predictor for aCGH success of FFPE samples. Br J Cancer 94:333-7.
7. Little SE, et al. 2006. Array CGH using whole genome amplification of fresh-frozen and formalin-fixed, paraffin-embedded tumor DNA. Genomics 87:298-306.
8. Johnson NA, et al. 2006. Application of array CGH on archival formalin-fixed paraffin-embedded tissues including small numbers of microdissected cells. Lab Invest 86:968-78.
9. Devries S, et al. 2005. Array-based comparative genomic hybridization from formalin-fixed, paraffin-embedded breast tumors. J Mol Diagn 7:65-71.
10. Alers JC, et al. 1999. Universal linkage system: an improved method for labeling archival DNA for comparative genomic hybridization. Genes Chromosomes Cancer 25:301-5.
11. Raap AK, et al. 2004. Array comparative genomic hybridization with cyanin cis-platinum-labeled DNAs. Biotechniques 37:130-4.
12. Reis-Filho JS, et al. 2005. Pleomorphic lobular carcinoma of the breast: role of comprehensive molecular pathology in characterization of an entity. J Pathol 207:1-13.
13. Di PS, et al. 2007. Oncocytic change in pleomorphic adenoma: molecular evidence in support of an origin in neoplastic cells. J Clin Pathol 60:492-9.
14. Lambros MB, et al. 2006. Unlocking pathology archives for molecular genetic studies: a reliable method to generate probes for chromogenic and fluorescent in situ hybridization. Lab Invest 86:398-408.
15. de Vries BB, et al. 2005. Diagnostic genome profiling in mental retardation. Am J Hum Genet 77:606-16.
16. Picard F, et al. 2005. A statistical approach for array CGH data analysis. BMC Bioinformatics 6:27.
17. Jong K, et al. 2007. Cross-platform array comparative genomic hybridization meta-analysis separates hematopoietic and mesenchymal from epithelial tumors. Oncogene 26:1499-506.
18. van Beers EH, et al. 2006. Array-CGH and breast cancer. Breast Cancer Res 8:210.

Chapter 4

Prediction of BRCA1-association in hereditary non-BRCA1/2 breast carcinomas with array-CGH.

SA Joosse, EH van Beers, IH Tielen, H Horlings, JL Peterse, N Hoogerbrugge, MJ Ligtenberg, LF Wessels, P Axwijk, S Verhoef, FB Hogervorst and PM Nederlof

Breast Cancer Res Treat. 2009 Aug; 116(3):479-89.

Prediction of BRCA1-association in hereditary non-BRCA1/2 breast carcinomas with array-CGH

Simon A Joosse, Erik H van Beers, Ivon HG Tielen, Hugo Horlings, Johannes L Peterse, Nicoline Hoogerbrugge, Marjolijn Ligtenberg, Lodewyk Wessels, Priscilla Axwijk, Senno Verhoef, Frans Hogervorst and Petra M Nederlof

While new defects in BRCA1 are still being found, it is unclear whether current breast cancer diagnostics misses many BRCA1-associated cases. A reliable test that is able to indicate the involvement of BRCA1 deficiency in cancer genesis could support decision making in genetic counseling and clinical management. To find BRCA1-specific markers and explore the effectiveness of the current diagnostic strategy, we designed a classification method, validated it and examined whether we could find BRCA1-like breast tumors in a group of patients initially diagnosed as non-BRCA1/2 mutation carriers. A classifier was built based on array-CGH profiles of 18 BRCA1-related and 32 control breast tumors, and validated on independent sets of 16 BRCA1-related and 16 control breast carcinomas. Subsequently, we applied the classifier to 48 breast tumors of patients from Hereditary Breast and Ovarian Cancer (HBOC) families in whom no germline BRCA1/BRCA2 mutations were identified. The classifier showed an accuracy of 91% when applied to the validation sets. In 48 non-BRCA1/2 patients, only two breast tumors presented a BRCA1-like CGH profile. Additional evidence for BRCA1 dysfunction was found in one of these tumors. We here describe the specific chromosomal aberrations in BRCA1-related breast carcinomas. We developed a predictive genetic test for BRCA1-association and show that BRCA1-related tumors can still be identified in HBOC families after routine DNA diagnostics.

With kind permission from Springer Science+Business Media: Breast Cancer Res Treat. 2009 Aug; 116(3):479-89.

Background

Breast cancer is the most common cancer in developed countries and one of the leading causes of death in women. One out of nine women will be affected by breast cancer (1, 2). Up to 25% of familial breast cancer cases are explained by mutations in BRCA1 and BRCA2 (3, 4). Women carrying a mutation in BRCA1/2 have a lifetime risk of up to 80% of developing breast cancer (5-8). Identification of such a mutation may not only influence the treatment of a patient or carrier (*e.g.*, chemotherapy, radiation, bilateral prophylactic mastectomy, or salphingooophorectomy (9, 10)) and surveillance, but also allows pre-symptomatic mutation screening of family members.

Based on family history and age of onset, breast cancer patients are eligible for DNA screening for pathogenic mutations in BRCA1/2. Diagnostics currently include mutation scanning and sequencing of gene fragments derived from germline DNA, however, it is possible that part of the mutations still remains undetected (4, 7, 11, 12). Additionally, the detection of variants of unknown clinical significance complicates

counseling and clinical management. Therefore, an additional tool that would indicate BRCA1 or BRCA2 involvement in breast cancer would be an asset to current clinical diagnostics.

Numerous studies show specific genetic characteristics with which tumors can be categorized into subclasses (13). For hereditary BRCA1-mutated cancer, previous publications from our research group and others show that these tumors develop distinct genetic alterations on which they can be recognized and distinguished from non-hereditary (sporadic) tumors (14-17). Various methods using expression profiling (14, 15) or comparative genomic hybridization (CGH) (16-18) show specific genetic alterations for these tumor groups. Although analysis of tumor mRNA has led to the identification of many different molecular portraits, fresh frozen tissue is often not available, especially when family screening includes deceased relatives. On the other hand, formalin-fixation and embedding in paraffin is the common procedure used to handle and archive tumor tissue for pathology based diagnosis. We have previously shown that paraffin embedded tumors can be of adequate quality to perform CGH studies (19, 20). The enhanced resolution of a microarray, compared with metaphase CGH (16), may improve the sensitivity and specificity of the detection of BRCA1- or BRCA2-related tumors using CGH technology. Additionally, it will also provide a better estimate of the location of the chromosomal breakpoints of the genetic aberrations.

To limit the already large number of individuals eligible for DNA-screening, a pre-selection procedure to detect individuals with the highest risk of carrying a mutation is desirable. Prediction models based on family history to calculate the risk for carrying a mutation can be inadequate predictors, *e.g.*, in small families (21). An independent test based on tumor characteristics that would indicate involvement of BRCA1 could help to select for those patients who may be offered more extensive mutation analysis. Studies based on clinical assessment and pathological reviewing show the limited sensitivity and specificity of predicting BRCA1-status with the currently available markers (12, 22). Genomic profiling of tumors using comparative genomic hybridization could also function as a predictive strategy to select patients with a high priori risk of a BRCA1 mutation. However, this approach has not been applied earlier in a diagnostic setting. In general, more BRCA1 mutations are being found in HBOC (Hereditary Breast and Ovarian Cancer) than in HBC (Hereditary Breast Cancer) families, and the former group would therefore be more suited for our study to identify possible missed BRCA1-associated tumors for evaluation of current diagnostics.

Patients and Methods

Patient selection

This study was performed on primary tumor samples of three breast cancer groups: *(1)* 34 breast tumors from patients with a confirmed pathogenic BRCA1 germ-line mutation, mean age at diagnosis of 38 years (range: 27– 61); *(2)* 48 sporadic breast tumors, mean age at diagnosis of 45 years (range: 32–60), no family history of breast cancer and selected from the institute's pathology archive matched for p53-status with the BRCA1-associated tumor group (Table 1); *(3)* 48 tumors from HBOC families (defined as at least two breast and one primary ovarian cancer), that were subjected to routine diagnostic testing (4) and had a negative test result for mutations in both BRCA1 and BRCA2, with a

Table 1 *– Pathological characteristics of the analyzed BRCA1 mutation carriers, sporadic, and HBOC breast carcinomas.*

	BRCA1	Sporadic	HBOC
No. analyzed	34	48	48
ER-positive	5.9% (2/34)	54.3% (25/46)	68.9% (31/45)
PR-positive	5.9% (2/34)	46.8% (22/47)	50.0% (23/46)
Her2/neu-positive	2.9% (1/34)	40.0% (17/46)	9.8% (4/41)
p53-positive	44.1% (15/34)	43.5% (20/46)	9.8% (4/41)

mean age at diagnosis of 48 years (range: 20–61). Patients' characteristics for all three groups are described in Supplementary data 1. All sample material was formalin-fixed, paraffin-embedded (FFPE) tissue and extracted DNA had to be of sufficient quality, which was tested as previously described (19). All experiments involving human tissues were conducted with the permission of the institute's medical ethical advisory board.

Immunohistochemistry

Presence of ER, PR, ERBB2 (HER2/neu), p53 and KRT5/6 was determined by immunohistochemistry using the antibodies: estrogen receptor AB-14 clone 1D5 + 6F11, titer 1:50 (Neomarkers); progesterone receptor clone PR-1, titer 1:400 (Immunologic), c-erbB-2 clone SP3, titer 1:25 (Neomarkers); p53 clone D0-7, titer 1:8000 (Dako); and keratin 5/6 clone D5/16 B4, titer 1:100 (Dako) respectively. If ≥70% of the tumor cells expressed ER, PR, p53 or CK5/6, the tumor was scored as positive (+) for the corresponding staining, in case ≤10% of the cells were stained, the tumor was scored as negative (-) and between 10 and 70% the tumor was scored as ± for the corresponding staining. ERBB2 staining was scored positive when a 3+ staining was observed, otherwise it was scored negative (only one sporadic case was IHC 2+, and was called negative).

DNA isolation

Tumor DNA was isolated from FFPE tumor tissues as follows. 10 x 10 µm slices containing at least 70% tumor cells were cleared of paraffin (2x 5 min xylene, 2x 30 s 100% ethanol, 30 s 90% ethanol, 30 s 70% ethanol, and rinsed with H_2O), treated with 1 M NaSCN at 37°C overnight, and sections of interest (>70% tumor cells) were scraped in 200 µl buffer ATL (Qiagen, cat. no. 51304). 27 µl of proteinase K (15 µg/µl, Roche, cat. no. 3115879001) was immediately added, as well as at the end of the day, and at the beginning and end of the next day; samples were constantly shaken at 37°C during the time of digestion. The following day, 40 µl RNase A (20 µg/µl, Sigma, cat. no. R5500) was added to the sample, vortexed, and incubated for 2 min at room temperature. 400 µl of buffer AL (Qiagen, cat.no. 51304) was added and incubated for 10 min at 70°C. 420 µl of 100% ethanol was added and vortexed. The sample mixture was spun on a spincolumn (Qiagen, cat. no. 51304) for 1 min at 8,000 rpm. The column was sequentially washed with 500 µl of the following reagents and spun for 1 min at 8,000 rpm: AW1, AW2, and twice with 80% ethanol. The column was spun dry for 3 min at 14,000 rpm. The sample was eluted with 50 µl of AE buffer by spinning for 1 min at 8,000 rpm. *Reference DNA* was isolated from lymphocytes from six appar-

ently healthy women and pooled. Lymphocytes were purified by adding lysis buffer (155 mM NH_4Cl, 10 mM $KHCO_3$, 1 mM EDTA) 4x the blood volume, followed by centrifugation at 3,000 rpm for 10 min at 4°C. The supernatant was removed and the cell pellet re-suspended in lysis buffer 5x the original blood volume. These steps were repeated until all erythrocytes were removed and the supernatant formed a clear solution. 1/10 of the initial blood volume DNAzol (Invitrogen, cat. no. 10503-027) was added to the cell pellet and mixed by pipetting until a clear solution remained. 1/2 of the DNAzol volume 100% ethanol was added; DNA was removed from the solution, washed in 70% ethanol and dissolved in Tris-EDTA buffer. DNA was sonicated until the average length was 300– 800 bp.

Array-CGH

As described previously (20), hybridizations were done on microarrays containing 3.5k BAC/PAC-derived DNA segments covering the whole genome with an average spacing of 1 Mb, obtained from the Welcome Trust Sanger Institute (UK). The whole library was spotted in triplicate on every slide. To prevent slide batch spotting bias, samples were hybridized in randomized order. Data processing of the scanned microarray slide included signal intensity measurement with the ImaGene software program, followed by median pin-tip (*c.q.* subarray) normalization. Intensity ratios (Cy5/Cy3) were log2-transformed and triplicate spot measurements were averaged.

Aberration detection and quantification

To analyze and visualize the chromosomal aberrations, we determined breakpoint locations and estimated copy number levels using the CGH-segmentation algorithm (23). Based on the estimated copy number levels, the frequency of gains and losses for all BAC clones was calculated using the fixed log2-ratio thresholds of 0.15 and -0.15, respectively. The association of the frequency of a clone being 'gained', 'lost' or 'unchanged' and the two tumor groups was calculated by employing 3 x 2 Fisher's exact (FE) test. A small P-value corresponds to a significant association between the observed copy number changes and the two groups. This procedure was employed to compare the whole cohort of BRCA1-related tumors to: *(1)* the whole set of sporadic tumors and *(2)* to particular subgroups in order to obtain an indication of the aberrations associated with these subgroups. First, subgroups were defined based on IHC-status of ER, PR, ERBB2 or p53 of the tumors (KRT5/6 status was not used to define a subgroup due to the small number of KRT5/6 positive tumors). IHC status can either be positive or negative (IHC-medium (±) samples were assigned to the corresponding IHC-positive group). Hence, four IHC-negative and four IHC-positive subgroups of sporadic tumors were generated. Then the FE test was employed to identify two sets of significant different aberrations. The first set, denoted as 'BRCA1/IHC-negative' includes aberrations that are significantly associated with the class label (BRCA1-related or sporadic) when comparing BRCA1-relateted and IHC-negative sporadic tumors. The second set, denoted as 'BRCA1/IHC-positive', includes aberrations that are significantly associated with the class label when comparing BRCA1-related and IHC-positive sporadic tumors. Since a IHC status (*e.g.*, ER) can strongly be associated with the class label in a dataset contacting the BRCA1-related and IHC-positive sporadic subsets, the aberrations in the 'BRCA1/IHC-positive' can be

either BRCA1 associated or IHC associated. Since BRCA1-related tumors are mostly ER, PR and ERBB2 negative, the aberrations in these 'BRCA1/IHC-negative' sets are mainly BRCA1 associated. Therefore, an aberration included in either of these subgroups is specific for the associated IHC-status. An aberration included in both sets is likely to be specific for BRCA1-associated tumors only.

Class prediction: training

To build a class predictor based on log2-ratios derived from array-CGH experiments, the shrunken centroids (SC) algorithm was employed (24) using equal priors, $\pi_k = 1/K$, where K is the number of classes. We predicted, employing the approach of Dobbin and Simon (25), that in order to detect a standardized fold change of 1.7 (1 copy number gain) amongst 3277 BAC clones at an error tolerance of 0.10 for the classifier, the minimal sample size for the training set had to be 31, equally divided over the two classes. The class predictor was built on 18 BRCA1-related and 32 sporadic breast tumors (referred to as the training sets). Since it is known that p53- and ER-status are associated with specific genomic aberrations in breast cancer (26, 27) that could influence the classification process, we stratified for p53-status in both the training sets, and for ER-status in the sporadic training set only.

Class prediction: testing

The class predictor was validated on independent sets of 16 BRCA1-mutated and 16 sporadic tumors (referred to as the validation sets). Classification of a sample using the SC algorithm results in the probability scores between 0 and 1 for each class. The sum of the two probability scores for any sample is always 1; hence, in a two-class problem, the most likely class is the class for which the probability exceeds 0.5. For legibility we only describe the highest probability. 95% reference intervals were calculated based on the class-probability distribution in the training sets and employed in the validation of the classifier and the classification of the HBOC group. Samples predicted outside the 95% reference intervals were not assigned to a class but scored as "not classified".

Methylation detection

Hypermethylation of the BRCA1 promoter for all samples (n = 130) was determined using Methylation MLPA according to the manufacture's protocol (MRC-Holland, ME001). DNA fragments were analyzed on a 3730 DNA Analyzer (AB, USA).

Loss of heterozygosity (LOH)

LOH at the BRCA1 locus was determined for the HBOC cases (n = 48) using 5 STR markers: D17S579, D17S588, D17S1322, D17S1323 and THRA1. Primers and the detailed PCR program are described in Supplementary data 2. DNA fragments were analyzed on a 3730 DNA Analyzer (AB, USA).

Results

In total, we obtained the array-CGH profiles of 34 BRCA1- related, 48 sporadic and 48 HBOC breast tumors. In this report we outline the chromosomal aberrations and their locations, the differences between the tumor groups, and the discriminating power of a class predictor based on our CGH results.

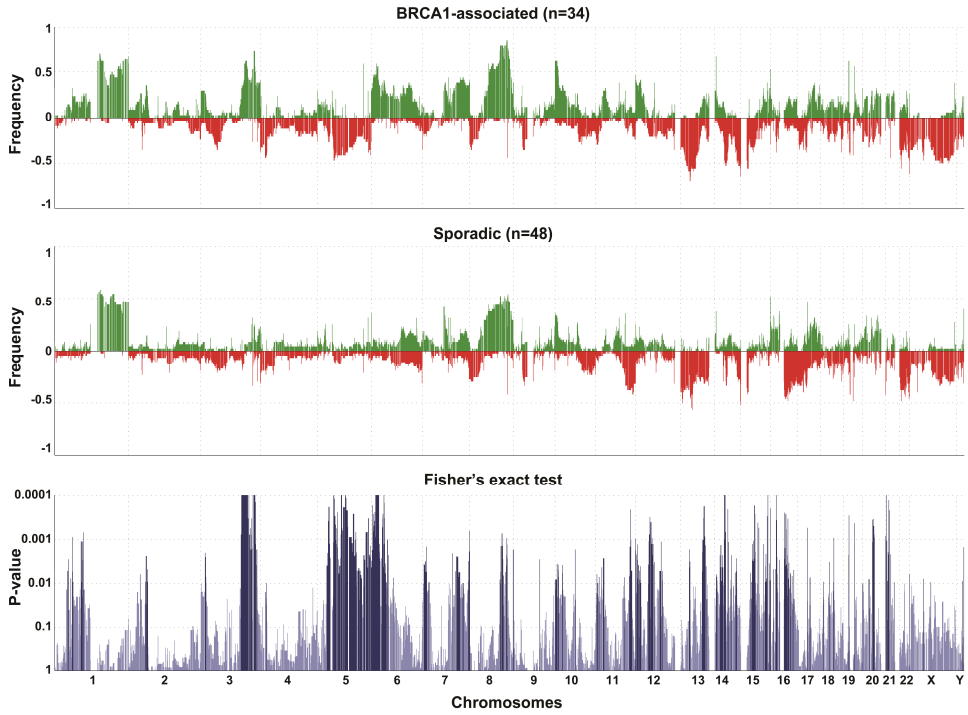


Figure 1 – Frequency plots. *The top two panels display the frequency of gain (green) and loss (red) in 34 BRCA1-associated and 48 sporadic breast carcinomas based on the estimated copy numbers as described in the Methods. The bottom panel shows the significance between the two tumor groups computed by Fisher's exact test for each clone (uncorrected for multiple testing). P-values < 0.01 are indicated dark blue.*

Chromosomal aberrations

We observed significant differences in the spectrum of aberrations with respect to the BRCA1-associated and sporadic breast tumors. The upper panels of Fig. 1 depict the frequency of gain (green) and loss (red) of the BAC clones for the BRCA1-associated and the sporadic breast tumors, respectively. The significances of the group differences are calculated by Fisher's exact test for each clone and are displayed in the bottom panel of Fig. 1.

We also found that BRCA1-related tumors have more copy number alterations (CNAs) compared with the sporadic breast tumors. Table 2 summarizes the most prominent aberrations of both tumor groups. These regions comprise several adjacent clones (at least 10 Mb in size), are aberrant in at least 30% of the tumor cases in one group, or show a significant difference between the tumor groups (average P-value for that region of < 0.01, Fisher's exact test). Gain of chromosome 1q and 8q are found in almost half of both the tumor groups which have been reported previously to be common for breast cancer (28-31). In total, the BRCA1-associated tumors showed 12 regional (> 10 Mb) gains and 11 regional losses that were observed in > 30% of the tumors. Using the same criteria, we observed gain in 2 chromosomal regions and 5 regional losses in the sporadic breast tumors.

Table 2 – Chromosomal aberrations. *Locations and average frequencies of the most prominent aberrations in 34 BRCA1-associated and 48 sporadic breast tumors together with the average P-values (FE test) for the significance in aberration difference between tumor groups.*

Chromosome	Aberration	BRCA1 (%)	Sporadic (%)	P-value
1q	Gain	53	47	0.5186
3q22-25	Gain	46	4	0.0006
3q26-29	Gain	43	13	0.0327
4p16-15	Loss	34	19	0.2466
5p13-5p12	Loss	19	1	0.0031
5q11-23	Loss	37	6	0.0017
5q31-35	Loss	24	4	0.0111
6p	Gain	37	7	0.0010
6q21-q23	Gain	32	17	0.1493
7p22-15	Loss	17	0	0.0080
7q22-36	Gain	37	13	0.0283
8p23	Loss	31	28	0.5767
8q	Gain	55	40	0.3306
9p21-13	Loss	35	28	0.7177
10p15-14	Gain	57	27	0.0129
10p14-12	Gain	32	11	0.0248
11p14-13	Gain	30	13	0.1000
11q22-25	Loss	13	35	0.1523
12p13-12	Gain	38	12	0.0165
12q12-14	Loss	20	0	0.0100
13q	Loss	36	32	0.2302
14q22-23	Loss	45	16	0.0568
14q32	Loss	49	26	0.0942
15q11-21	loss	35	13	0.0791
16q	Loss	10	36	0.0312
17p	Loss	24	32	0.5579
17q22-23	Gain	34	26	0.5522
20q11-12	Loss	26	0	0.0007
22q	Loss	34	33	0.4091
Xp22	Loss	31	15	0.1638
Xq	Loss	40	24	0.1715

BRCA1 and sporadic breast tumor class predictor

We used Shrunken Centroids (SC) (24) to discriminate between germline mutated BRCA1 and sporadic tumors. We randomly selected 18 BRCA1-associated and 32 sporadic tumors for the SC analysis; these tumors are referred to as the training set. Employing leave-one-out cross-validation (LOOCV) on the training set, $\Delta = 1.3$ was the parameter setting resulting in the smallest number of misclassifications. The training set yielded 191 discriminatory features (Supplementary data 3) which were used in this study for further classification. From these 191 features most were abundant in regions of chromosome 3q22-27 (gain), 5q12-14 (loss), 6p23-22 (gain), 12p13 (gain), 12q21-23 (loss), and 13q31-34 (gain). Based on the classification scores of the training samples, 95% reference intervals were calculated for both groups. The minimum reference interval for the BRCA1 class was a BRCA1-like probability of 0.8; the minimum reference interval for the sporadic class was a sporadic-like probability of 0.7 (*i.e.*, a BRCA1-like probability of 0.3).

The remaining samples were used as external validation for the class predictor. In the validation set, 14/16 samples of the BRCA1-related group were predicted as BRCA1-like and were inside the 95% reference interval, while the remaining two cases were outside the interval but predicted as BRCA1-like, with a lower (< 0.8) probability. One of the 16 sporadic breast cancer cases was classified as sporadic-like with a probability of 0.62 whereas all others (n = 15) classified within the 95% reference intervals. These results can be formulated as a total sensitivity of 88% and a specificity of 94% (PPP: 93%, NPP: 88%). Fig. 2a depicts the distribution

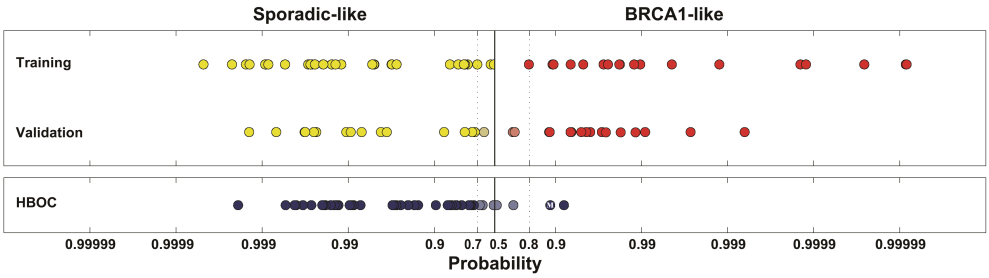


Figure 2 – Classification results. *Probability scores for the Training and Validation sets of the BRCA1-associated ● and sporadic ● tumor samples (upper panel). Samples predicted as BRCA1-like are plotted right of the 0.5 probability border and samples predicted as sporadic are plotted left. Dotted lines indicate the 95% reference intervals for both classes based on the results of their training sets. Bottom panel shows the classification of the HBOC group ● where 2 tumors were classified as BRCA1-like. Samples outside the classes' 95% reference intervals were not assigned to a class. Sample labeled 'M' showed methylation of the BRCA1 promoter.*

of the classification scores obtained on the training and validation sets; the classification scores for each individual sample are documented in Supplementary data 1.

ER, PR, ERBB2 and p53 specific aberrations

In our tumor groups, 91% of the BRCA1-mutated tumors are ER, PR and ERBB2 negative (also known as 'triple negative'), while only 19% of the sporadic cases are triple negative (Supplementary data 1). To investigate the relationship between ER, PR, ERBB2 or p53-status with chromosomal aberrations and thus the possible influence on our class predictor, the BRCA1-associated tumor group was compared to subsets of the sporadic tumors selected by their IHC phenotypes as described in the Methods section "Aberration detection and quantification". Chromosomal regions 3q22-3q26, 5p14, 6p22.3 and 14q22 were significantly more often aberrant in BRCA1-associated tumors, independent of the sporadic breast tumors' IHC phenotype (Fig. 3 upper four panels, indicated in blue). BAC clones within these BRCA1-specific regions were also represented in the classifier (Supplementary data 3, Fig. 3 bottom panel). However, the largest part of the loss in chromosome 5q, that was selected by the SC algorithm, appeared to be ER and/or ERBB2 status specific rather than BRCA1-specific as calculated by Fisher's exact test (Fig. 3, indicated in orange).

To evaluate the performance of the classifier features in discriminating BRCA1-related and sporadic tumors and the influence of the IHC profile, we performed hierarchical cluster analysis (complete linkage, Pearson correlation) to the array CGH results of the 34 BRCA1-associated and 48 sporadic breast tumors based on the 191 classifier features. The samples were separated into two large clusters, one containing most of the sporadic breast cancer cases (Fig. 4, left branch), and one containing all the BRCA1-associated tumor samples (Fig. 4, right branch). Although, some of the sporadic cases resided together with the BRCA1-associated cases, this could not be explained by association with the ER or ERBB2 status ($P = 0.24$ and $P = 0.25$, respectively; FE test). Since the basal-like phenotype is very common for BRCA1-associated breast cancer, we investigated whether the sporadic tumors clustering together with the BRCA1-related tumors were also basal-like by

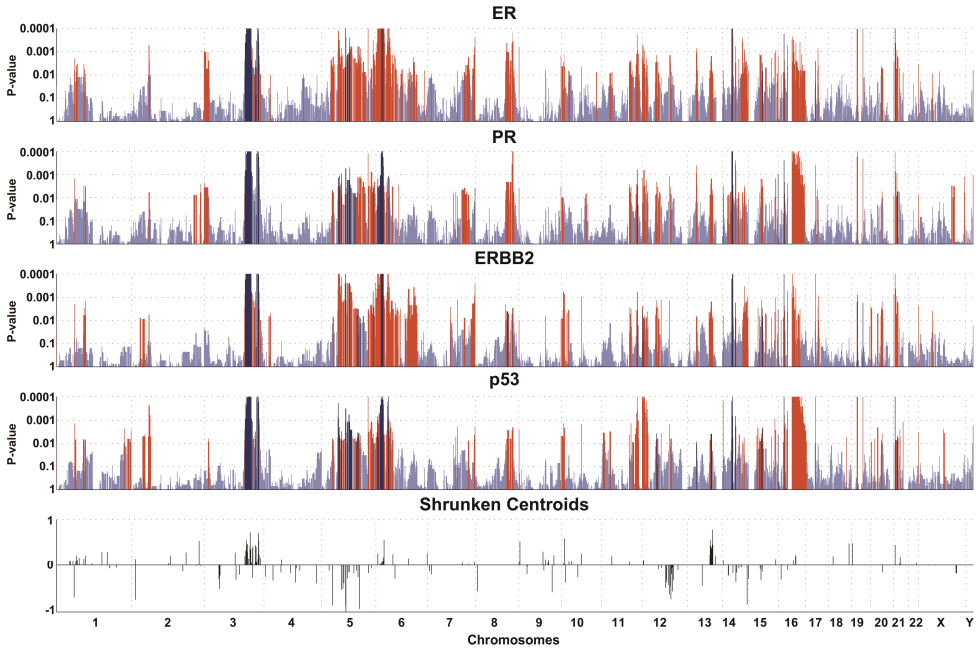


Figure 3 – Influence of IHC-status. *The four upper panels depict the significance in differences between the BRCA1-related cohort and the sporadic IHC tumor subgroups calculated by Fisher's exact test. Blue areas are BRCA1-specific aberrations while orange aberrations are specific for the corresponding IHC positive-phenotype, computed as described in the Methods. Grey are the non-significant areas (P>0.01). Lower panel shows the 191 shrunken centroids that are used for the classifier.*

performing IHC for KRT5/6 (Supplementary data 1). Only two of the five sporadic breast tumors expressing KRT5/6 clustered within the BRCA1 branch which showed not to be a significant correlation (P = 0.23; FE test).

Since ER status is highly correlated with the class label (BRAC1-related or sporadic) we compared the performance of our classifier with a classifier that predicts a tumor to be BRCA1-like when the tumor is ER negative. In the validation set, 15/16 samples of the BRCA1-related tumor group are predicted as BRCA1-like while seven of the 16 sporadic breast cancer cases were classified as sporadic-like. When employing the conventional approach in the SC classifier of assigning all samples as BRCA1-like when the probability of the BRCA1-like class exceeds 0.5 (i.e. not using rejection regions) the SC classifier has a total error rate of 0/32. The ER classifier, when also employing a single ER level cut-off, achieves an error rate of 8/32. Assuming these proportions are samples from a binomial distribution, the 95% confidence intervals for these proportions are [0; 0.1089] for the SC classifier and [0.1146; 0.4340] for the ER classifier, respectively. Employing the McNemar test for matched pairs, we can conclude that the SC classifier has a significantly better performance on the validation set than the ER classifier (P = 0.007).

Taken together, these results suggest that specific chromosomal aberrations are associated with the tumors' IHC-status (Fig. 3), nevertheless, BRCA1-related and sporadic breast tumors can be distinguished using the 191 features selected by the SC algorithm (Figs. 2, 4).

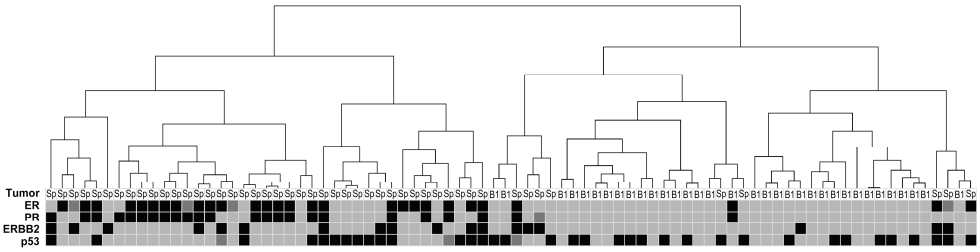


Figure 4 – Hierarchical clustering. *Complete hierarchical clustering of 34 BRCA1-related (B1) and 48 sporadic (Sp) breast carcinomas based on the log2-ratios of the 191 BAC clones (shrunken centroids) from the classifier. Shown are the IHC-status of ER, PR, ERBB2 and p53 of all samples as positive ■, negative ■, and intermediate ■ staining. The dendrogram can be divided in two main branches: BRCA1-related (right) and sporadic tumor samples (left).*

Application of the classifier on non-BRCA1/2 families

Forty-eight patients from non-BRCA1/2 HBOC families were selected and analyzed using aCGH. We found 2 samples (HBOC34 and HBOC41) to be BRCA1-like (P>0.8), 40 samples were predicted as sporadic-like, and 6 samples could not be assigned to a class with sufficient certainty as they were classified outside the 95% reference intervals. Fig. 2b shows the distribution of the clinical samples in comparison with the BRCA1-related and sporadic tumors used to build and validate our class predictor.

To find evidence for BRCA1 involvement in the two BRCA1-like breast cancer cases, we first performed LOH analysis of the BRCA1 locus. Loss of the wild type BRCA1 tumor suppressor gene is considered to be required for BRCA1-related tumor development. We observed clear LOH of BRCA1 in samples HBOC34, HBOC41 and HBOC08, as well as allelic imbalance (where one allele is diminished but still present) in HBOC03, HBOC04, HBOC07, HBOC18, HBOC29, HBOC042 and HBOC45. Allelic imbalance can be caused by trisomy of the locus, tumor heterogeneity and limited tumor cell percentage. We next performed additional tests that were not included in the original routine diagnostic setting. As BRCA1 exon 11 was analyzed for truncating mutations using the Protein Truncation Test (PTT) (32), we now sequenced exon 11 in DNA isolated from peripheral blood lymphocytes from cases HBOC34 and HBOC41 but found no mutations. The next test was to investigate somatic inactivation of BRCA1 by methylation of the BRCA1 promoter. This was determined for all BRCA1-associated, sporadic and HBOC samples using MLPA-methylation (MRC-Holland, ME001). Case HBOC34, which was classified as BRCA1-like, was the only sample that showed methylation at the BRCA1 promoter (labeled "M" in Fig. 2). This patient also presented with an ovarian carcinoma that, like the breast tumor, showed methylation at the BRCA1 promoter, interestingly, germline DNA of this patient did not show methylation at this site.

Discussion

In this study we show that BRCA1-associated breast tumors possess rearranged genomes with specific genomic aberrations that

differ significantly from sporadic breast tumors. Based on array-CGH data, we identified the most significant differences between these two tumor groups and built a class predictor with 88% sensitivity and 94% specificity using the Nearest Shrunken Centroids method (24). Compared with the BRCA1-associated tumors, aberrations are less frequent in the sporadic breast tumors. Many of the identified regions specific for the BRCA1-related tumors have been reported before (16-18, 31, 33, 34). In this study we applied the BRCA1 classifier tool on diagnostic cases in order to investigate the performance of the familial breast cancer routine mutation screening. By doing so, we identified 2 out of 48 tumors as BRCA1-like. Since all tumors were formalin-fixed and paraffin-embedded, investigation of mRNA could be problematic (35), and further analyses were performed on genomic DNA from the tumor. Since potentially any (somatic) inactivation of BRCA1 could result in a BRCA1-like phenotype (36, 37), we investigated methylation of the BRCA1 promoter. One of the BRCA1-like HBOC cases indeed showed methylation and LOH of the BRCA1 gene, strongly indicating BRCA1 dysfunction in the tumor. Cancer formation due to BRCA1 mutation is generally accompanied by the loss of the wild-type allele, i.e. LOH, which was also found in the second BRCA1-like HBOC tumor. However, no novel or described mutations in the BRCA1 gene were identified in this patient after sequencing. This particular patient's family history was atypical from that of an average BRCA1-involved family (breast and ovarian cancer), with incidence of brain cancer, colon cancer, and leukaemia. Additionally, the tumor was ER and PR positive, which is rare for BRCA1-related tumors (38). This unresolved BRCA1-like case may be analyzed more intensely when new techniques and knowledge become available. Another way to predict the involvement of a BRCA1-mutation is to use prediction models based on family characteristics. We have calculated Evans' scores (39) for all possible cases to determine the probability to find a BRCA1 mutation. Both BRCA1-like tumors showed a small probability (20 and 11.8%), which could explain why no germline BRCA1 mutations were found in these families but a somatic inactivation of BRCA1 in case HBOC34, consistent with a low Evans' score.

In an earlier study we were able to classify BRCA1-associated and control tumors using chromosomal CGH with an accuracy of 84% (16). A control group with a relatively large proportion of bilateral tumors was used to mimic the situation in high-risk breast cancer families. The disadvantage of this approach was that it resulted in many 'false' positives in the control group (specificity of 76%); however, many of those were proven to be actual BRCA1-associated tumors later on (unpublished results). We now used automated array-CGH, which is a high throughput technique and therefore suitable to be performed in specialized diagnostic laboratories. Additionally, the use of a microarray in this study localizes the significant genomic areas with increased chromosomal resolution and may help to develop a test (*e.g.*, PCR based) that can be applied in any routine diagnostic laboratory. In contrast to our previous study that contained a relatively large number of control cases later proven to be true BRCA1-mutated tumors, we now use a random control group of sporadic tumors excluding family history for breast cancer and bilateral breast cancer. Although the differences between chromosomal and array-CGH and patient selections between our previous and current studies are substantial, both loss in 5q and gain in 3q were identified as important discrimina-

tory aberrations, as confirmed by others (17, 33, 34).

While exploring chromosomal aberrations it has to be kept in mind that steroid hormone receptor status is strongly correlated with the genomic profile (26, 27). It has been reported that BRCA1-associated tumors are in general (>90%) ER, PR, and HER2/neu-negative (38). Since we chose to randomly select breast carcinomas for our control group, these tumors do not all share the triple-negative phenotype of the BRCA1-related tumors. Training on triple-negative breast tumors only could restrict our class predictor to triple-negative carcinomas while our goal was to build a general classification method to classify all non-BRCA1/2 HBOC patients. Also, selection for triple-negative tumors only would increase the likelihood for false positives since a triple-negative population is per definition enriched for unidentified BRCA1-associated cases. Instead, we investigated the extent to which specific aberrations could be associated with ER, PR, ERBB2 and p53-status to obtain an indication of their possible influence on our class predictor. Loss of chromosomal region 5q12-14, which is present in our classifier as discriminatory region (Fig. 3), was found to be specific for ER-negative tumors. Although this would suggest selection for the ER-phenotype, no false positives or false negatives were present in the validation sets. The result that all ER-negative (and triple-negative) sporadic tumors were correctly classified as sporadic-like further supports that classification is not based on ER-status alone, but is based on the combination of the BRCA1-specific regions.

Increasing evidence shows that the majority of BRCA1-related carcinomas are basal-like tumors with respect to morphology and mRNA expression level (40). This is also true for the BRCA1-specific aberrations as reported here and elsewhere (16-18, 33, 34) that show many analogies to the breast cancer basal-like subtype [(28), this article]. Similarities between these hereditary and sporadic breast cancer groups could be explained by the effect of the same deficient DNA repair pathway (*i.e.*, BRCA1). So far, our Fisher's exact test (Fig. 3), hierarchical clustering (Fig. 4), and classification results (Fig. 2) all indicate differences between the BRCA1-related and the triple-negative (basal-like) sporadic tumors that lie within the 191 discriminatory features.

In the future, it may be possible to include our profiling test in clinical genetic screening programs to select the individual in a high-risk family with the highest prior probability for finding the BRCA1 germline mutation, as an alternative or addition to screening of the youngest affected case as is currently done. Furthermore, it could help in decision making and treatment management, also when no BRCA1-like profile is found which would be an (extra) indication to rule out BRCA1 involvement. Moreover, aCGH classification of a tumor with a nucleotide variant of uncertain significance may give extra indications for the significance of the variant (41).

Conclusion

Based on aCGH data, we were able to identify BRCA1-specific aberrations that were different from those seen in sporadic breast tumors and employed this to build a class predictor. Although steroid hormone receptor status is strongly associated with genomic instability, this class-predictor distinguished BRCA1-associated tumors from sporadic breast carcinomas with increased accuracy than current screening protocols. We conclude that current BRCA1 mutation screening seems to identify

most hereditary BRCA1-associated breast tumors. However, while we could still find BRCA1-related breast tumors in a non-BRCA1/2 tumor group, our array-CGH approach may also be used as an additional tool to identify BRCA1-associated patients or families where the relation to BRCA1 is still unclear.

Acknowledgements

We like to acknowledge the Central Microarray Facility (NKI) for the production of the microarrays, Aafke Ariaens and Roelof Pruntel for technical support, Sonja Springer for management of material and Juliane Hannemann for critically reading the manuscript. Grant support Dutch Cancer Society/Koningin Wilhelmina Fonds grant NKB_NKI2005-3436.

Supplementary data for this article are online available at http://www.springerlink.com/content/e9v1174161131j5l/

References

1. American Cancer Society. 2008. Cancer Facts and Figures, 2007
2. Visser, O., Siesling, S., and van Dijck, J. A. A. M. 2003. Incidence of cancer in the Netherlands 1999/2000. Eleventh report of the Netherlands Cancer Registry
3. Szabo CI, et al. 1997. Population genetics of BRCA1 and BRCA2. Am J Hum Genet 60:1013-20.
4. van der Hout AH, et al. 2006. A DGGE system for comprehensive mutation screening of BRCA1 and BRCA2: application in a Dutch cancer clinic setting. Hum Mutat 27:654-66.
5. Antoniou A, et al. 2003. 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. Am J Hum Genet 72:1117-30.
6. Easton DF, et al. 1995. Breast and ovarian cancer incidence in BRCA1-mutation carriers. Breast Cancer Linkage Consortium. Am J Hum Genet 56:265-71.
7. Ford D, et al. 1998. Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium. Am J Hum Genet 62:676-89.
8. King MC, et al. 2003. Breast and ovarian cancer risks due to inherited mutations in BRCA1 and BRCA2. Science 302:643-6.
9. Cleator S, et al. 2007. Triple-negative breast cancer: therapeutic options. Lancet Oncol 8:235-44.
10. Tercyak KP, et al. 2007. Quality of life after contralateral prophylactic mastectomy in newly diagnosed high-risk breast cancer patients who underwent BRCA1/2 gene testing. J Clin Oncol 25:285-91.
11. Narod SA, et al. 1995. Risk modifiers in carriers of BRCA1 mutations. Int J Cancer 64:394-8.
12. van der Groep P, et al. 2006. Distinction between hereditary and sporadic breast cancer on the basis of clinicopathological data. J Clin Pathol 59:611-7.
13. Perou CM, et al. 2000. Molecular portraits of human breast tumours. Nature 406:747-52.
14. Hedenfalk I, et al. 2001. Gene-expression profiles in hereditary breast cancer. N Engl J Med 344:539-48.

15. 't Veer LJ, et al. 2002. Gene expression profiling predicts clinical outcome of breast cancer. Nature 415:530-6.
16. Wessels LF, et al. 2002. Molecular classification of breast carcinomas by comparative genomic hybridization: a specific somatic genetic profile for BRCA1 tumors. Cancer Res 62:7110-7.
17. Jonsson G, et al. 2005. Distinct genomic profiles in hereditary breast tumors identified by array-based comparative genomic hybridization. Cancer Res 65:7612-21.
18. van Beers EH, et al. 2005. Comparative genomic hybridization profiles in human BRCA1 and BRCA2 breast tumors highlight differential sets of genomic aberrations. Cancer Res 65:822-7.
19. van Beers EH, et al. 2006. A multiplex PCR predictor for aCGH success of FFPE samples. Br J Cancer 94:333-7.
20. Joosse SA, et al. 2007. Automated array-CGH optimized for archival formalin-fixed, paraffin-embedded tumor material. BMC Cancer 7:43.
21. Kang HH, et al. 2006. Evaluation of models to predict BRCA germline mutations. Br J Cancer 95:914-20.
22. Lakhani SR, et al. 2005. Prediction of BRCA1 status in patients with breast cancer using estrogen receptor and basal phenotype. Clin Cancer Res 11:5175-80.
23. Picard F, et al. 2005. A statistical approach for array CGH data analysis. BMC Bioinformatics 6:27.
24. Tibshirani R, et al. 2002. Diagnosis of multiple cancer types by shrunken centroids of gene expression. Proc Natl Acad Sci U S A 99:6567-72.
25. Dobbin KK, et al. 2008. How Large a Training Set is Needed to Develop a Classifier for Microarray Data? Clin Cancer Res 14:108-14.
26. Loo LW, et al. 2004. Array comparative genomic hybridization analysis of genomic alterations in breast cancer subtypes. Cancer Res 64:8541-9.
27. Fridlyand J, et al. 2006. Breast tumor copy number aberration phenotypes and genomic instability. BMC Cancer 6:96.
28. Bergamaschi A, et al. 2006. Distinct patterns of DNA copy number alteration are associated with different clinicopathological features and gene-expression subtypes of breast cancer. Genes Chromosomes Cancer 45:1033-40.
29. Ghazani AA, et al. 2007. Genomic alterations in sporadic synchronous primary breast cancer using array and metaphase comparative genomic hybridization. Neoplasia 9:511-20.
30. Chin SF, et al. 2007. Using array-comparative genomic hybridization to define molecular portraits of primary breast cancers. Oncogene 26:1959-70.
31. Vincent-Salomon A, et al. 2007. X inactive-specific transcript RNA coating and genetic instability of the X chromosome in BRCA1 breast tumors. Cancer Res 67:5134-40.
32. Hogervorst FB, et al. 1995. Rapid detection of BRCA1 mutations by the protein truncation test. Nat Genet 10:208-12.
33. Tirkkonen M, et al. 1997. Distinct somatic genetic changes associated with tumor progression in carriers of BRCA1 and BRCA2 germ-line mutations. Cancer Res 57:1222-7.
34. Johannsdottir HK, et al. 2006. Chromosome 5 imbalance mapping in breast tumors from BRCA1 and BRCA2 mutation carriers and sporadic breast tumors. Int J Cancer 119:1052-60.

35. Penland SK, et al. 2007. RNA expression analysis of formalin-fixed paraffin-embedded tumors. Lab Invest 87:383-91.
36. Snell C, et al. 2008. BRCA1 promoter methylation in peripheral blood DNA of mutation negative familial breast cancer patients with a BRCA1 tumour phenotype. Breast Cancer Res 10:R12.
37. Esteller M, et al. 2001. DNA methylation patterns in hereditary human cancers mimic sporadic tumorigenesis. Hum Mol Genet 10:3001-7.
38. Lakhani SR, et al. 2002. The pathology of familial breast cancer: predictive value of immunohistochemical markers estrogen receptor, progesterone receptor, HER-2, and p53 in patients with mutations in BRCA1 and BRCA2. J Clin Oncol 20:2310-8.
39. Evans DG, et al. 2004. A new scoring system for the chances of identifying a BRCA1/2 mutation outperforms existing models including BRCAPRO. J Med Genet 41:474-80.
40. Reis-Filho JS, et al. 2008. Triple negative tumours: a critical review. Histopathology 52:108-18.
41. Tischkowitz M, et al. 2008. Pathogenicity of the BRCA1 missense variant M1775K is determined by the disruption of the BRCT phosphopeptide-binding pocket: a multi-modal approach. Eur J Hum Genet 16:820-32.

Chapter 5

Prediction of BRCA2-association in hereditary breast carcinomas using array-CGH.

SA Joosse, KI Brandwijk, P Devilee, J Wesseling, FB Hogervorst, S Verhoef and PM Nederlof

Breast Cancer Res Treat. Published online: 2010 July

Prediction of BRCA2-association in hereditary breast carcinomas using array-CGH

Simon A Joosse, Kim IM Brandwijk, Peter Devilee, Jelle Wesseling, Frans BL Hogervorst, Senno Verhoef, and Petra M Nederlof

Germline mutations in BRCA1/2 increase the lifetime risk for breast and ovarian cancer dramatically. Identification of such mutations is important for optimal treatment decisions and pre-symptomatic mutation screening in family members. Although current DNA diagnostics is able to identify many different mutations, it remains unclear, how many BRCA2-associated breast cancer cases remain unidentified as such. In addition, mutation scanning detects many unclassified variants (UV) for which the clinical relevance is uncertain. Therefore, our aim was to develop a test to identify BRCA2-association in breast tumors based on the genomic signature. A BRCA2-classifier was built using array-CGH profiles of 28 BRCA2-mutated and 28 sporadic breast tumors. The classifier was validated on an independent group of 19 BRCA2-mutated and 19 sporadic breast tumors. Subsequently, we tested 89 breast tumors from suspected hereditary breast (and ovarian) cancer (HBOC) families, in which either no BRCA1/2 mutation or an UV had been found by routine diagnostics. The classifier showed a sensitivity of 89% and specificity of 84% on the validation set of known BRCA2-mutation carriers and sporadic tumor cases. Of the 89 HBOC cases, 17 presented a BRCA2-like profile. In three of these cases, additional indications for BRCA2-deficiency were found. Chromosomal aberrations that were specific for BRCA2-mutated tumors included loss on chromosome arm 13q and 14q, and gain on 17q. Since we could separate BRCA1-like, BRCA2-like, and sporadic-like tumors using our current BRCA2- and previous BRCA1-classifier, this method of breast tumor classification could be applied as additional test for current diagnostics to help clinicians in decision-making and classifying sequence variants of unknown significance.

With kind permission from Springer Science+Business Media: Breast Cancer Res Treat. Published online 2010 Jul.

Introduction

Individuals that inherit a germline mutation in *BRCA1* or *BRCA2* have a significantly increased lifetime risk of developing breast or ovarian cancer. Recent publications review the importance to identify *BRCA1/2* mutation carriers for optimal therapy and non-carriers for chemoprevention (1, 2). Defects in homologous recombination (impaired *BRCA1/2* pathway) cause high sensitivity for drugs that induce double-strand breaks (e.g., alkylating agents). However, successful mutation identification impacts not only the patient but also on the family members, since it allows for pre-symptomatic mutation screening. The current strategy to identify mutation carriers is first to select those patients eligible for mutation

screening based on prediction models that use age and family history (4). Subsequently, the mutation screening is performed; e.g., by sequencing of gene fragments in germline DNA, protein truncation test (PTT) and denaturing gradient gel electrophoresis (DGGE) (5, 6). However, it still remains unclear, to what extent mutation carriers are identified with the current diagnostic tools since many families with a history for breast cancer remain unexplained. It is known that mutation prediction models are highly dependent on the number of family members, from which information is available (4, 7); this type of information is often limited. Another clinically difficult situation is the identification of an unclassified variant (UV) in coding or non-coding regions in either one of the BRCA genes. The pathogenicity of such a nucleotide variant is often uncertain as the effect on the protein function is unknown. Therefore, its clinical significance also remains unclear. Although functional assays exist for the proteins produced by mutated *BRCA1/2* genes, these are laborious, difficult to interpret in clinical terms, limited to only a number of protein functionalities, and not yet routinely applicable in a diagnostic setting (8). Therefore, our goal was to evaluate profiling of somatic genetic changes in breast tumors as a new strategy that can provide additional information about the involvement of BRCA2 in tumorigenesis.

For *BRCA1*-mutated tumors, several molecular portraits have been generated using copy number alterations (3, 9-12) and gene expression patterns (13, 14). It has already been shown that such genetic profiling can successfully be applied to identify *BRCA1*-associated cases (3, 15) and to provide an additional indication whether an UV is pathogenic or not (16). For *BRCA2*-mutated tumors, there is much less evidence for the existence of a specific genetic signature, also the immunohistochemical phenotype is not that well defined (17). Although several studies investigated the differences between *BRCA1*-mutated, *BRCA2*-mutated and sporadic breast tumors in gene expression patterns (13) and copy number alterations (11, 12, 18, 19), these signatures have not been validated extensively and were not evaluated in a clinical setting. The number of samples was relatively small and/or the investigated tumor groups were not matched for sex, histological grade, tumor type, and estrogen receptor (ER) status, which all have been shown to have their own individual characteristics at the genomic level that could be misinterpreted as the signature of interest (20-24). This implies the need for a validated *BRCA2* signature, which is independent of tumor grade and receptor status, and which can be used in combination with a *BRCA1* signature.

Because BRCA1 and BRCA2 play important roles in DNA repair by homologous recombination, it is not surprising that breast tumors deficient in either one of the encoding genes show extensive chromosomal imbalance (3, 9). This could be exploited as the basis for molecular profiling. In this study, we have used array-CGH to investigate the copy number changes of DNA sequences extracted from formalin-fixed, paraffin-embedded (FFPE) tissue, which is readily available in pathology archives and therefore very suitable for diagnostic purposes. Additionally, using the same technique as our previous classifier, allows the combination of tests for both BRCA profiles.

Materials and methods

Patient selection

Three breast cancer groups were used which were selected from the institute's archive: (1) 47

breast carcinomas from women with a confirmed pathogenic *BRCA2* germline mutation, mean age at diagnosis of 46 years (range 26–86), referred to as *BRCA2*-mutated tumors; (2) 47 sporadic breast tumors from women with unknown *BRCA2* status, mean age at diagnosis of 45 years (range 29–78), no known family history for breast cancer and matched to the tumor group mentioned above for age, gender, ER, PR, ERBB2, and p53 immunohistochemical (IHC) status and tumor grade; (3) 89 tumors from women that were eligible for, and subjected to, routine diagnostic testing according to the HBOC criteria (25) but were negative for pathogenic BRCA1/2-mutations or carried an UV in either BRCA1/2; mean age at diagnosis of 47 years (range 29–75). This third group included 37 HBOC cases from our previous study (3), 47 new HBOC cases, and 5 cases carrying an UV (Table 1). This third group is referred to as non-BRCA1/2 tumors. All experiments involving human tissues were conducted with the permission of the institute's medical ethical advisory board. Individual sample characteristics are listed in Supplementary Table 1 including which samples were used to built or validate the classifier. The 34 CGH profiles of *BRCA1*-mutated tumors described in this manuscript are from our previous study (3).

Sample material

All sample material was formalin-fixed, paraffin-embedded (FFPE) archival tissue from invasive ductal carcinomas (IDC). 10 µm sections were cut and regions containing at least 70% tumor cells were scraped. DNA was extracted by proteinase-K digestion after deparaffination; and quality was tested by a multiplex PCR as described elsewhere (3, 26). In total, we isolated DNA of 69 BRCA2-mutated, 104 sporadic and 107 non-BRCA1/2 tumors for this study. Only those DNA samples of which PCR products of at least 200 bp could be produced, were of sufficient quality for array-CGH (data not shown).

Pathological review

Presence of ER, PR, HER2, and p53 were determined by revision of immunohistochemical staining that were previously performed in standard clinical procedure with antibodies: estrogen receptor AB-14 clone 1D5 + 6F11, titer

Table 1 – Unclassified variants. *Unclassified variants (UV) found in the HBOC tumor group. Listed are the Type and the Effect of the UVs. aCGH profiles were classified with both the 'BRCA1 classifier' and the 'BRCA2 classifier' (Classification). Case PFT2946 was diagnosed with two primary tumors.*

Case	Gene	UV	Type	Effect	Classification
PFT2946 (2x)	BRCA2	c.6842-20T>A	Intronic variant	Different splice prediction programs: no effect	Sporadic-like
PFT5737	BRCA2	c.9502-12T>G	Intronic variant	Loss of splice acceptor site, deletion of exon 26	BRCA2-like
PFT6270	BRCA2	c.1395A>C	Silent coding variant	Very likely no effect	Sporadic-like
PFT3045	BRCA1	c.81-9C>G	Intronic variant	Creation and use of novel acceptor site, frame shift	BRCA1-like

1:50 (Neomarkers); progesterone receptor clone PR-1, titer 1:400 (Immunologic); c-erbB-2 clone SP3, titer 1:25 (Neomarkers); and TP53 clone D0-7, titer 1:8,000 (DAKO), respectively.

For simplicity, IHC scoring was divided into two classes. If ≥1% of the tumor cells expressed ER, PR, or p53, the tumor was scored as positive (+), otherwise, the tumor was scored as negative (-) for the corresponding staining, according to Viale *et al.* (27). HER2 scoring was performed according to ASCO/CAP and oncoline guidelines (28, 29). A tumor was scored positive for HER2 when a 3+ staining was observed. When a 2+ staining was observed, CISH was performed to determine amplification (+ in case of 6 spots or more per nucleus) or no amplification (-). A 1+ or negative IHC staining was scored as negative (-).

Tumor grade was determined by the modified Bloom–Richardson–Elston staging system (30).

Array-CGH

Sample preparation, labeling, BAC arrays preparation, and array processing were done as previously described (31). In short, ULS-Cy5 labeled tumor DNA and ULS-Cy3 labeled reference DNA from six apparently healthy women were co-hybridized for 72 h on a microarray containing 3.5k BAC/PAC derived DNA segments covering the whole genome with an average spacing of 1 Mb spotted in triplicate. Hybridization was performed on a Tecan HS4800 hybridization station, which uses liquid agitation during hybridization. In total we performed aCGH with 57 *BRCA2*-mutated, 82 sporadic, and 77 non-BRCA1/2 tumors samples. The quality of the hybridization was assessed by calculating the standard deviations of the log2 ratios of the triplicate spots. Only aCGH profiles with a mean standard deviation < 0.1 were used. These microarray data have been deposited in NCBI's Gene Expression Omnibus (32) and are accessible through GEO Series accession numbers GSE16511 (*BRCA2*-mutated), GSE9114 (sporadic), and GSE22481 (non-BRCA1/2).

Analyses of aCGH profiles

We have employed three different methods to analyze the aCGH profiles. First, the frequency of the aberrations was calculated and plotted in a so called 'frequency plot', purely to summarize and visualize the percentage of (common) aberrations in BRCA2-mutated and sporadic tumors. As second method, a classifier was built which could discriminate between the tumor groups and assign individual tumors to a tumor class (group). Finally, the association between the individual tumors was examined using hierarchical clustering (complete linkage, Pearson correlation) based on the features used for the classifier. Details are described below.

Aberration quantification

Breakpoint locations and estimated copy number level of the chromosomal aberrations were determined by the CGH-segmentation algorithm described by Picard *et al.* (33), further referred to as the 'segmentation data'. To calculate aberration frequency and the average number of aberrations per tumor group, the segmentation data was discretized to 'neutral', 'loss', 'gain', and 'amplified' by applying thresholds as described by Chin et al. (34). Thresholds for gain/loss and amplification were defined by two and eight times the standard deviation of 50% quantile of the segmented data, respectively (34). Significant differences between the tumor groups for frequency of aberrations ('neutral', 'gained', 'lost', or 'amplified') was calculated by employing a 4 x 2 Fisher's exact (FE) test (35). P-values were not directly corrected for multiple

testing since this would be too stringent. Instead, since adjacent BAC clones are highly correlated, a genomic region was called significant when at least five adjacent BAC clones were calculated to be significant with $p < 0.01$. Using this approach, copy number variations smaller than 5 Mb could also be excluded from the results.

To calculate the association of the average number of aberrations between tumor groups, 2-tailed t-test was applied.

Shrunken centroids-based classifier

To prevent over-fitting of the classifier, the approach of Dobbin and Simon (36) was used to calculate the required sample size using a standardized fold change of 1.7. For an error tolerance of < 0.10, more than 15 samples of each class were needed. As in our previous study (3), we have used the shrunken centroids (SC) algorithm (37) to construct the 'BRCA2 classifier', however, now based on the segmentation data to eliminate technical noise. To train the 'BRCA2 classifier', a fraction of 0.6 of each group was randomly selected (28 *BRCA2*-mutated and 28 sporadic tumor profiles, total n = 56). The classifier was validated with the remaining samples of each group (19 *BRCA2*-mutated and 19 sporadic tumor profiles, total n = 38). As a result, the classification algorithm predicts the classes' likelihoods for each sample. Since the sum of the two likelihoods is always "1", we only describe the highest class probability (> 0.5). Depending on the classes' highest likelihood, the sample will be referred to as BRCA2-like or sporadic-like. Next we tested the aCGH profiles of 89 non-BRCA1/2 tumors for BRCA2 class likelihood, additionally we tested for BRCA1 likelihood with the 'BRCA1 classifier' from our previous study (3) to the same cases which uses the similar scoring method.

Additional screening for *BRCA1/2* defects

To identify defects in the *BRCA1/2* genes that could have been missed by standard diagnostics, we performed the following additional tests: BRCA2 exon deletion/duplication MLPA according to the manufacturer's protocol (MRC-Holland, The Netherlands, MLPA kit P090); mRNA sequence analysis from peripheral blood lymphocytes to determine bi/mono-allelic expression of *BRCA2*, using regions containing a single nucleotide polymorphism (SNP); loss of heterozygosity (LOH) of the *BRCA2* locus in tumor DNA using the markers D13S171, D13S260, D13S267, and D13S289 and LOH of the *BRCA1* locus as described before (3); methylation of the *BRCA1* and *BRCA2* promoters using methylation MLPA according to the manufacturer's protocol (MRC-Holland, The Netherlands, MS-MLPA kit ME001B). Moreover, we have analyzed multiple family members of four families to investigate the presence of a common CGH profile by classification.

Results

Immunohistochemistry

BRCA2-mutated tumors were predominantly ER positive (83%) with various histological tumor grade, while *BRCA1*-mutated tumors were mainly ER negative (94%) and grade III (Table 2). This is in concordance with literature which reports similar numbers (17). The distribution of tumor grade among non-BRCA1/2 HBOC tumors was similar to that of *BRCA2*-mutated tumors, although, fewer tumors were ER positive (Table 2).

Table 2 – Tumor group characteristics. *Immunohistological characteristics of the different tumor groups in this study.*

	***BRCA2*-mutated (n=47)**	**Sporadic (n=47)**	**Training B2 (n=28)**	**Training Sp (n=28)**	**Non-*BRCA1/2* (n=89)**	***BRCA1*-mutated (n=34)**
Grade						
I	15 (n= 7)	15 (n=7)	18 (n=5)	14 (n=4)	10 (n=9)	0 (n=0)
II	36 (n=17)	32 (n=15)	29 (n=8)	29 (n=8)	35 (n=31)	21 (n=7)
III	49 (n=23)	53 (n=25)	54 (n=15)	57 (n=16)	43 (n=38)	79 (n=27)
ER						
+	83 (n=39)	83 (n=39)	82 (n=23)	79 (n=22)	53 (n=47)	6 (n=2)
-	17 (n= 8)	17 (n=8)	18 (n=5)	21 (n=6)	33 (n=29)	94 (n=32)
PR						
+	45 (n=21)	57 (n=27)	54 (n=15)	57 (n=16)	40 (n=36)	6 (n=1)
-	55 (n=26)	43 (n=20)	46 (n=13)	43 (n=12)	44 (n=39)	97 (n=33)
ERBB2						
+	13 (n= 6)	19 (n= 9)	18 (n=5)	21 (n=6)	12 (n=11)	3 (n=1)
-	87 (n=41)	81 (n=38)	82 (n=23)	79 (n=22)	70 (n=62)	97 (n=33)
p53						
+	43 (n=20)	36 (n=17)	86 (n=24)	82 (n=23)	20 (n=18)	44 (n=15)
-	57 (n=27)	64 (n=30)	14 (n=4)	18 (n=5)	49 (n=44)	56 (n=19)

BRCA1-mutated tumors are from our previous study (3). Values are expressed as percentage.
Training B2 = Classifier training group BRCA2-mutated, Training Sp = Classifier training group Sporadic

Chromosomal aberrations: BRCA2 versus sporadic

Most aberrations found in the *BRCA2*-mutated tumors were also present in the sporadic tumor group, and with similar frequencies. The top two panels of Fig. 1 show the genome-wide frequency of losses (red), gains (green) and high level gains (dark green) in the *BRCA2*-mutated and the sporadic control group, respectively. Based on these numbers, 4 x 2 Fisher's exact test was employed to determine significant differences between the groups. The middle panel depicts p values with significant p-values ($p < 0.01$) indicated in dark blue. Three chromosomal aberrations were identified to be more associated with *BRCA2*-mutated tumors: loss of chromosome bands 13q12–q14, 14q23–q32 and gain of 17q11–q21.31. More associated with sporadic tumors were gain of chromosome band 16p13 and loss of 16q12 (Table 3; Supplementary Table 2). Based on the calculated breakpoints using CGH-segmentation (33), the numbers of aberrations in both tumor groups were counted. *BRCA2*-mutated tumors showed on average 75.7 ± 11.9 aberrations (range 56–109) and sporadic tumors showed a comparable average of 78.4 ± 12.3 aberrations (range 50-111), which was not significantly different ($p = 0.24$, two-sided, paired t-test).

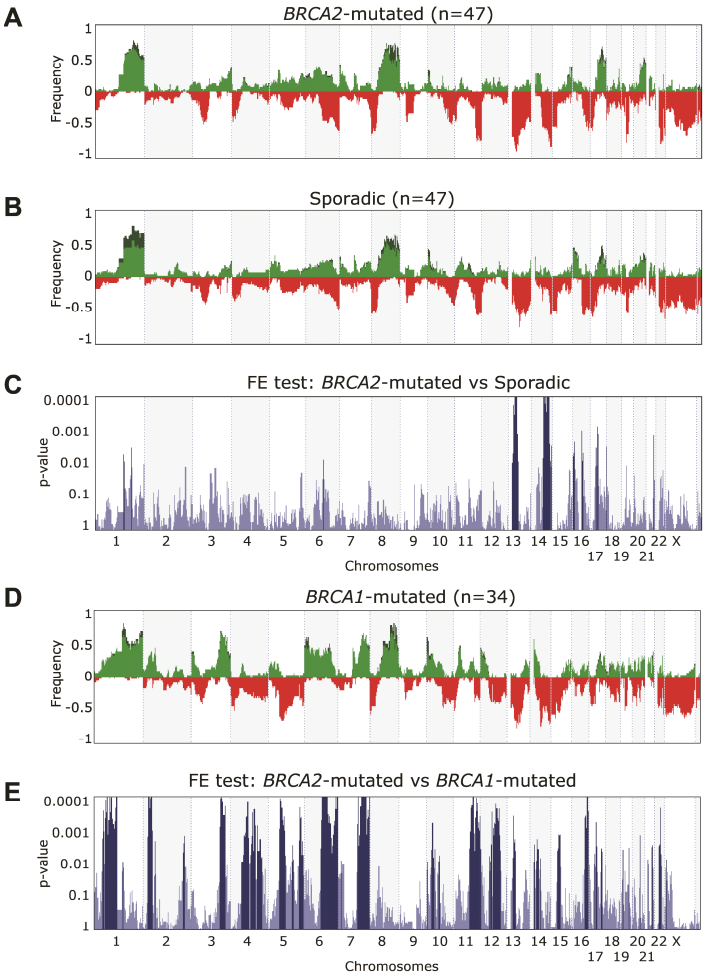


Figure 1 – Comparison of aberration frequency. *Frequency of gain (green), amplification (dark green) and loss (red) over 47 BRCA2-mutated (**A**) and 47 matched sporadic breast carcinomas (**B**) based on the estimated copy numbers as described in Material and Methods. **C**: significance between the two tumor groups computed by Fisher's exact test for each clone. P-values < 0.01 are indicated in dark blue. The bottom two panels show the aberration frequencies of 34 BRCA1-mutated breast carcinomas (**D**) and the significant regions between the BRCA2-mutated and BRCA1-mutated tumor groups (**E**), respectively. P-values are -log10 transformed.*

Chromosomal aberrations: BRCA2 versus BRCA1

Comparison of the CGH profiles of *BRCA2*- with *BRCA1*-mutated tumors revealed many significant different aberrations (Figure 1; Supplementary Table 2). The bottom two panels of Figure 1 show the genome-wide gains and losses of the *BRCA1*-mutated tumors from our previous study (3), and the p-values indicating the association of the aberration frequencies between the two hereditary breast cancer groups, respectively. The full list of aberration frequencies and p-values are documented in Supplementary Table 2. The number of aberrations differed significantly between these groups, (p = 3.75e-4). *BRCA2*-mutated tumors showed 75.7 aberrations on average, compared to 85.4 ± 11.2 aberrations (range 69-113) in *BRCA1*-mutated breast tumors.

Table 3 – BRCA2 associated chromosomal aberrations. *Five chromosomal regions (Chr.) were present in significantly different frequencies between the BRCA2-mutated and sporadic breast tumors calculated by Fisher's exact test. Given are the average percentages of gain and loss in both tumor groups of the corresponding chromosomal region and p-value (FE test).*

		BRCA2-mutated		Sporadic		FE test
Chr.	**Cytoband**	**Gain**	**Loss**	**Gain**	**Loss**	**p-value**
13	q12-q14	4%	78%	5%	44%	2.1e-3
14	q23.2-q32.2	2%	62%	9%	22%	5.7e-4
16	p13	14%	2%	41%	3%	3.7e-3
16	q12	10%	18%	5%	51%	3.0e-3
17	q11-q21.31	36%	8%	15%	32%	6.2e-3

BRCA2 classifier

Twenty-eight CGH profiles of the *BRCA2*-mutated tumor group and 28 of the sporadic tumor group were randomly selected to train the 'BRCA2 classifier'. Table 2 shows that the distribution of IHC status of the training sets is similar to the original groups, and thus also comparable with the population. Employing leave-one-out cross-validation (LOOCV), D = 0.4 led to the lowest misclassification rate. Using these 56 profiles, 703 features were selected as discriminatory by the SC algorithm (clone name and genomic location are given in Supplementary Table 3). The features selected by the SC algorithm showed a large overlap (67%) with the regions selected using the frequencies (Table 3). For the training sets, one sample of the *BRCA2*-mutated tumors and one sample of the sporadic tumors classified to the opposite class (misclassification of 4%).

The remaining 38 samples were used to validate the classifier. Figure 2 shows the distribution of the classification scores for the training as well as for the validation sets. Samples classified as sporadic-like are plotted left, while BRCA2-like samples are plotted right. In the validation sets, 17/19 BRCA2-mutated tumors (red) and 16/19 sporadic tumors (yellow) were correctly classified. Consequently, the sensitivity was 89% and specificity 84%, the positive (PPP) and negative predictive power (NPP) were 85 and 89%, respectively.

To further evaluate the chromosomal regions that were selected for the 'BRCA2 classifier', we performed hierarchical cluster analyses on the segmentation data of all the samples based on these regions only. Figure 3 depicts the result of the cluster analyses and shows that the samples are divided into three large clusters. IHC data of each sample are displayed along the cluster tree to explore whether samples of both groups clustering together would share the same IHC phenotype; this was not the case. Clusters B and C contain all except two (45/47) of the sporadic cases, cluster A contains all but two (45/47) of the BRCA2-mutated cases (Figure 3). These results indicate that the features selected for classification have indeed discriminatory power, regardless of the algorithm used and independent of IHC phenotype.

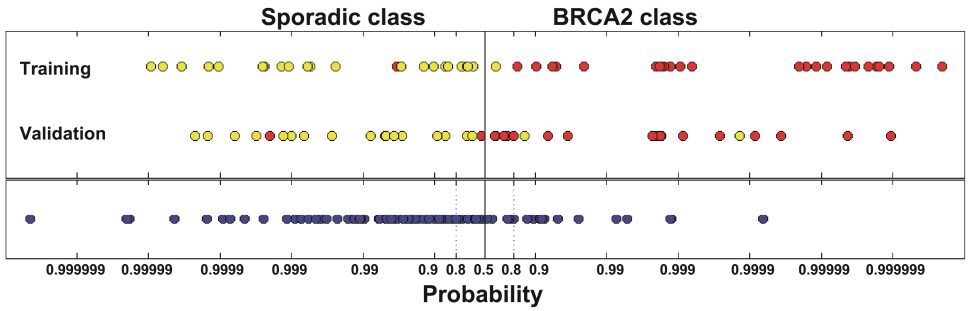


Figure 2 – Classification with the 'BRCA2 classifier'. *The top panel shows the probability scores for the training and validation sets of the BRCA2-mutated (● n=47) and sporadic (● n=47) tumor samples. Samples predicted to be BRCA2-like are plotted right and samples predicted to be sporadic are plotted left. The bottom panel depicts the classification results of the non-BRCA1/2 tumor group (● n=89), where 17 tumors were classified as BRCA2-like (probability > 0.5).*

Clinical application of the 'BRCA1/2 classifiers'

To evaluate the 'BRCA2 classifier' in clinical setting, 89 breast cancer samples from non-BRCA1/2 HBOC patients were analyzed (Figure 2, blue circles; Supplementary Table 1). Seventeen cases (19%) were classified as BRCA2-like with a BRCA2-class probability > 0.5, 13/17 with high probability > 0.8; the remaining 72 cases (81%) were classified as sporadic-like. One of the BRCA2-like cases carried the *BRCA2* UV c.9502-12T>G. The same cases were also classified using the 'BRCA1 classifier' (3), 11 samples were classified as BRCA1-like. Of these 11 tumors, one carried the *BRCA1* UV c.81-9C>G and two tumors were also classified as BRCA2-like. All 17 BRCA2-like cases, 11 BRCA1-like cases and cases carrying an UV were studied in more detail using additional molecular tests to identify possible missed BRCA1/2-associated cases (described below and listed in Supplementary Table 1).

Unclassified variants

Routine mutation analysis of germline DNA had previously revealed four unclassified variants in BRCA2 and one in BRCA1 (Table 1). To investigate whether the UVs cause aberrant mRNA molecules, mRNA was isolated from blood of these patients and analyzed by cDNA sequencing. This revealed that *BRCA2* UV c.9502-12T>G led to the deletion of exon 26. Also, *BRCA1* UV c.81-9C>G caused a splicing defect leading to a truncated protein. These results indicate that both unclassified variants are pathogenic and result in non-functional proteins. This correlates with the CGH profiles of these cases that were classified as BRCA2- and BRCA1-like, respectively. For the remaining two *BRCA2* UV cases, no indications were found for pathogenicity, also these findings were in concordance with the classifier's prediction, which was sporadic-like.

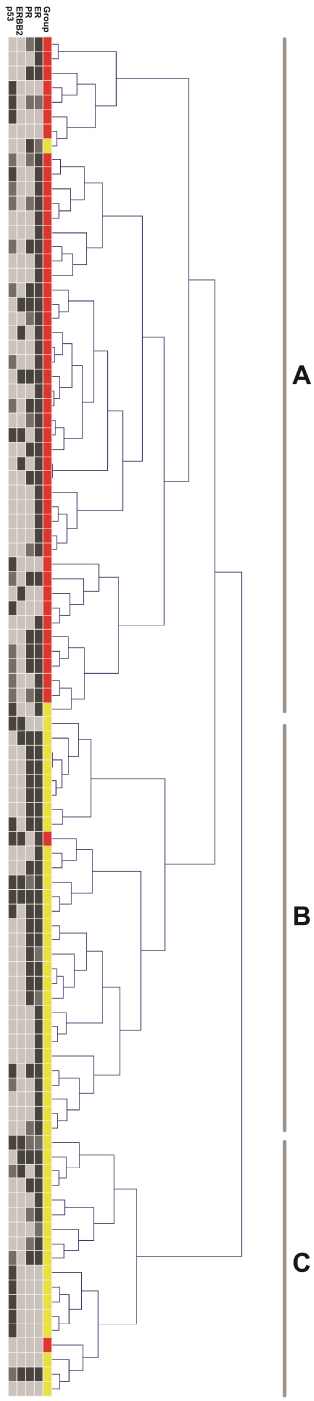


Figure 3 – Hierarchical clustering. *Complete hierarchical clustering of 47 BRCA2-mutated ■ and 47 sporadic ■ breast carcinomas based on the segmentation data of the same 704 BAC clones (shrunken centroids) that were used for the classifier. Shown are the IHC status (from left to right) of p53, ERBB2, PR and ER of all samples: IHC positive ■, negative ■, and intermediate ■ staining. The dendrogram can be divided into three main branches: one cluster of mainly BRCA2-mutated tumors (A, 47 samples) and two clusters of mainly sporadic tumor samples (B, 29 samples, and C, 18 samples).*

Mutation analysis

The *BRCA2* gene was investigated for whole-exon deletions or duplications using the P090 MLPA kit (MRC-Holland). None of the investigated cases showed such aberration.

Loss of heterozygosity (LOH)

We investigated LOH at four microsatellite markers flanking the *BRCA2* gene in the BRCA2-like cases. Most of the samples (75%) showed LOH or allelic imbalance (AI) for at least one informative (i.e., heterozygous) marker. The *BRCA1* locus was investigated using five microsatellite markers. This region showed LOH or AI in 67% of the investigated cases (Supplementary Table 1).

Promoter methylation

Methylation of the *BRCA1* and *BRCA2* promoter were investigated using the ME001 methylation MLPA kit (MRC Holland). None of the HBOC cases were found to be positive for methylation of the *BRCA2* promoter, only one BRCA1-like case showed methylation of the *BRCA1* promoter.

Allele-specific expression

Some mutations might be hidden and hard to find (e.g., intronic). In *BRCA1/2* mutation carriers, often mRNA expression of only the wild-type gene can be detected in blood. Therefore, we explored whether single allele expres-

sion of *BRCA2* could be identified, indicative of a defective gene. mRNA regions containing a SNP that was detected by routine diagnostics were sequenced to identify the ratio of expressed alleles. Eleven of the BRCA2-like cases were found to be heterozygous for a coding SNP. Only cases PFT6363 and PFT6386 appeared to express one allele of *BRCA2*, which may suggest that these patients carry a defective copy of *BRCA2* in their germline DNA.

Discussion

We investigated the chromosomal aberrations of *BRCA2*-mutated breast tumors by array-CGH to identify their molecular signature. We found that these tumors can be distinguished from sporadic tumors with an accuracy of 86.5%. To our knowledge, such accuracy has not been shown before using an array-CGH classifier. This signature can be used to give additional indications about the involvement of *BRCA2* in the tumorigenesis of a specific breast tumor case where the role of *BRCA2* is still unclear (*i.e.*, UV) or in patients in whom no mutation has been found (yet), but where a hereditary factor is suspected. In combination with our previous 'BRCA1 classifier', classification suggesting the involvement of either *BRCA1* or

BRCA2 could lead to extended diagnostics, may help clinicians in their decision making, and can lead to adjusted therapy that exploits *BRCA1/2* deficiency.

Classifier and clinical application

Using the shrunken centroids algorithm, we built a classifier with *BRCA2*-mutated and sporadic tumors resulting in a high accuracy (86.5%). For the misclassified samples, it cannot be excluded that some of the patients in the sporadic group in fact harbor a *BRCA2* germline mutation, as they were not tested for this. Based on the population frequency, this percentage will most likely not exceed 1% of all breast cancer cases. Furthermore, negative misclassification could be the result of a low tumor cell percentage, tumor heterogeneity, or an actual sporadic tumor in a germline *BRCA2* mutation family. Applying the 'BRCA2 classifier' to non-BRCA1/2 and BRCA1/2 UV cases, we found 17 tumors to be BRCA2-like. In three of these 17 cases, we have found indications for dysfunctional BRCA2. Although we also found LOH/AI of *BRCA2* in 9 tumors of the remaining 14 cases, we were unable to infer a *BRCA2* defect directly linked to tumorigenesis. Also methylation of the *BRCA2* promoter was not found, however, this is in agreement with reports suggesting that *BRCA2* promoter methylation does not occur frequently in breast cancer (38, 39). It should be noted here that based on the validation results, 16% of the samples could also be false positive. This means that of the 89 BRCA2-like cases, 14 may be false-positive sporadic tumors. Nevertheless, although these 14 BRCA2-like cases remain unsolved and could be considered false positive, these patients might benefit from the same treatment as true *BRCA2* mutation carriers, as they present similar genomic characteristics and might therefore also suffer from similar defective pathways (e.g., impaired homologous recombination, discussed below). Further investigation to support this hypothesis is needed. Applying the 'BRCA1 classifier' to the 89 non-BRCA1/2 cases, 11 were classified as BRCA1-like, which also include the two BRCA1-like cases from our previous study (3). One of the new cases was found to carry the BRCA1 UV c.81-9C>G, which led to a splicing defect. Together with LOH, this UV caused *BRCA1* deficiency.

BRCA1 and BRCA2

Several studies have investigated *BRCA1/2*-mutated tumors for chromosomal aberrations in comparison with control tumors. Most of the aberrations we have found have also been reported by others. Supplementary Table 4 describes the aberrations found on chromosomes by Tirkkonen *et al.* (9), van Beers *et al.* (11), Jonsson *et al.* (12), Stefansson *et al.* (19) and us. As has been shown by others (9, 11, 12, 19), comparison of the aCGH profiles of *BRCA2*- with *BRCA1*-mutated tumors shows a number of differences of which also many (if not most) can be related to ER status and histological grade (20-24). Since our technique makes use of the CGH profile of the whole genome, the chromosomal aberrations associated with grade and receptor status would greatly bias the groups' signatures. Therefore, we have generated two separate classifiers based on BRCA1/2 mutation status, to prevent interference of these characteristics that are not of interest. To overcome this problem, a comparison between these hereditary tumor groups using ER status and histological grade equal samples should be made. Due to the sparse occurrence of 'triple-negative', grade 3, *BRCA2*-mutated tumors, or ER-positive *BRCA1*-mutated tumors, such comparisons have not been performed yet. Interestingly, (only) two of the 89 non-BRCA1/2 cases classified as BRCA1-like and BRCA2-like, indicating that these profiles present both BRCA1- and BRCA2-specific aberrations. Whether this 'double BRCA' classification reflects a truly shared biological basis for hereditary tumors as suggested by Stefansson et al. (19), or the imperfectness of both classifiers is currently not clear yet. Since both classifiers are binomial and 'force' each sample into one class, these 'double positives' currently require a third method of classification or a pre-selection (e.g., based on grade or ER status). In addition, an additional family member could be screened. Since our method is validated on archival material, also tumor material from relatives from earlier generations could be investigated. In general, material from 1980 and onwards can be used (data not shown). Of one of these double positives, we could analyze additional family members (Supplementary Table 1, family number 2128). This one was classified as BRCA2-like, indicating that it is very likely that this family is affected by a hereditary defect in the BRCA2 pathway.

Homologous recombination deficiency (HRD)

A common genetic profile might reflect a common defect in DNA repair mechanisms also in the absence of a germline mutation as the defect may be somatic. DNA double-strand breaks (DSB) caused by DNA-damaging agents (such as alkylating chemotherapy) or the inhibition of DNA repair mechanisms (e.g., PARP inhibitors), can be lethal for cells that are deficient in homologous recombination. Homologous recombination is the only error-free repair mechanism for DNA double-strand breaks, and thus, identification of HRD tumors may lead to specifically targeting these tumors with alkylating agents or PARP inhibitors. Although both *BRCA1/2* genes are involved in homologous recombination, the histology of the *BRCA1*- and *BRCA2*-mutated tumors is quite different, as are their CGH profiles. The function of BRCA2 in DNA repair is probably restricted to HR only (40), while BRCA1 has many more functions and is involved in other DNA repair mechanisms as well, i.e., HR and NHEJ (non-homologous end joining) (41, 42). This may

explain the limited number of chromosomal breaks in BRCA2-mutated tumors where other repair mechanisms, e.g., NHEJ, may still be functional. The accumulating DNA breaks in *BRCA1*-mutated tumors may be explained by the fact that more DNA repair mechanisms are affected by the absence of BRCA1 (43). Recent studies in our institute have employed both classifiers as marker for HRD in sporadic tumors. It was shown that a BRCA1- or BRCA2-like CGH profile correlates with a higher response rate to adjuvant and neoadjuvant alkylating chemotherapy ((44), Lips *et al.* Breast Cancer Research 2011).

Future perspectives

Based on the identification of the BRCA1- and BRCA2-specific chromosomal aberrations, these regions can now be combined for the development of a simple, stable and less expensive assay, such as a PCR-based test. Such a test would be a powerful additional tool in current diagnostics routine to identify hereditary breast cancer.

Conclusion

Using archival material, we have built a classification method that is able to distinguish *BRCA2*-mutated from sporadic breast tumors based on their chromosomal aberrations with an accuracy of 86.5%. We conclude that, although current DNA diagnostics detects most *BRCA2*-mutated cases, our aCGH classifier can identify *BRCA2*-related cases in addition to those identified by current diagnostics. As such, we suggest that this new approach, together with our previous BRCA1 classifier, may be used as an additional tool to identify BRCA1/2-associated tumors, either of hereditary or sporadic origin.

Acknowledgments

We like to acknowledge the Central Microarray Facility (NKI) for the production of the microarrays and technical assistance, Roelof Pruntel, Mohamed Achahchah, and Esther Scheerman for technical support, Sonja Springer and Carla van Tiggelen for management of patient information and material, and Juliane Hannemann for critically reading the manuscript. Financial support, Dutch Cancer Society/Koningin Wilhelmina Fonds; Grant NKB_NKI2007-3749.

References

1. Foulkes WD. 2006. BRCA1 and BRCA2: chemosensitivity, treatment outcomes and prognosis. Fam Cancer 5:135-42.
2. Rubinstein WS. 2008. Hereditary breast cancer: pathobiology, clinical translation, and potential for targeted cancer therapeutics. Fam Cancer 7:83-9.
3. Joosse SA, et al. 2008. Prediction of BRCA1-association in hereditary non-BRCA1/2 breast carcinomas with array-CGH. Breast Cancer Res Treat 116:479-89.
4. Antoniou AC, et al. 2008. Predicting the likelihood of carrying a BRCA1 or BRCA2 mutation: validation of BOADICEA, BRCAPRO, IBIS, Myriad and the Manchester scoring system using data from UK genetics clinics. J Med Genet 45:425-31.
5. van der Hout AH, et al. 2006. A DGGE system for comprehensive mutation screening of BRCA1 and BRCA2: application in a Dutch cancer clinic setting. Hum Mutat 27:654-66.
6. Hogervorst FB, et al. 1995. Rapid detection of BRCA1 mutations by the protein truncation test. Nat Genet 10:208-12.

7. Kang HH, et al. 2006. Evaluation of models to predict BRCA germline mutations. Br J Cancer 95:914-20.
8. Carvalho MA, et al. 2007. Functional assays for BRCA1 and BRCA2. Int J Biochem Cell Biol 39:298-310.
9. Tirkkonen M, et al. 1997. Distinct somatic genetic changes associated with tumor progression in carriers of BRCA1 and BRCA2 germ-line mutations. Cancer Res 57:1222-7.
10. Wessels LF, et al. 2002. Molecular classification of breast carcinomas by comparative genomic hybridization: a specific somatic genetic profile for BRCA1 tumors. Cancer Res 62:7110-7.
11. van Beers EH, et al. 2005. Comparative genomic hybridization profiles in human BRCA1 and BRCA2 breast tumors highlight differential sets of genomic aberrations. Cancer Res 65:822-7.
12. Jonsson G, et al. 2005. Distinct genomic profiles in hereditary breast tumors identified by array-based comparative genomic hybridization. Cancer Res 65:7612-21.
13. Hedenfalk I, et al. 2001. Gene-expression profiles in hereditary breast cancer. N Engl J Med 344:539-48.
14. Van 't Veer LJ, et al. 2002. Gene expression profiling predicts clinical outcome of breast cancer. Nature 415:530-6.
15. van den Ouweland AM, et al. 2009. Deletion of exons 1a-2 of BRCA1: a rather frequent pathogenic abnormality. Genet Test Mol Biomarkers 13:399-406.
16. Tischkowitz M, et al. 2008. Pathogenicity of the BRCA1 missense variant M1775K is determined by the disruption of the BRCT phosphopeptide-binding pocket: a multi-modal approach. Eur J Hum Genet 16:820-32.
17. Lakhani SR, et al. 2002. The pathology of familial breast cancer: predictive value of immunohistochemical markers estrogen receptor, progesterone receptor, HER-2, and p53 in patients with mutations in BRCA1 and BRCA2. J Clin Oncol 20:2310-8.
18. Melchor L, et al. 2008. Distinct genomic aberration patterns are found in familial breast cancer associated with different immunohistochemical subtypes. Oncogene 27:3165-75.
19. Stefansson OA, et al. 2009. Genomic profiling of breast tumours in relation to BRCA abnormalities and phenotypes. Breast Cancer Res 11:R47.
20. Bergamaschi A, et al. 2006. Distinct patterns of DNA copy number alteration are associated with different clinicopathological features and gene-expression subtypes of breast cancer. Genes Chromosomes Cancer 45:1033-40.
21. Fridlyand J, et al. 2006. Breast tumor copy number aberration phenotypes and genomic instability. BMC Cancer 6:96.
22. Gruvberger S, et al. 2001. Estrogen receptor status in breast cancer is associated with remarkably distinct gene expression patterns. Cancer Res 61:5979-84.
23. Loo LW, et al. 2004. Array comparative genomic hybridization analysis of genomic alterations in breast cancer subtypes. Cancer Res 64:8541-9.
24. Melchor L, et al. 2007. Estrogen receptor status could modulate the genomic pattern in familial and sporadic breast cancer. Clin Cancer Res 13:7305-13.
25. 17-11-2009. HBOC criteria. http://www.bccancer.bc.ca/HPI/ .

26. van Beers EH, et al. 2006. A multiplex PCR predictor for aCGH success of FFPE samples. Br J Cancer 94:333-7.
27. Viale G, et al. 2007. Prognostic and predictive value of centrally reviewed expression of estrogen and progesterone receptors in a randomized trial comparing letrozole and tamoxifen adjuvant therapy for postmenopausal early breast cancer: BIG 1-98. J Clin Oncol 25:3846-52.
28. oncoline. 1-9-2008. Oncoline Guidelines
29. Wolff AC, et al. 2007. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. J Clin Oncol 25:118-45.
30. Genestie C, et al. 1998. Comparison of the prognostic value of Scarff-Bloom-Richardson and Nottingham histological grades in a series of 825 cases of breast cancer: major importance of the mitotic count as a component of both grading systems. Anticancer Res 18:571-6.
31. Joosse SA, et al. 2007. Automated array-CGH optimized for archival formalin-fixed, paraffin-embedded tumor material. BMC Cancer 7:43.
32. Edgar R, et al. 2002. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 30:207-10.
33. Picard F, et al. 2005. A statistical approach for array CGH data analysis. BMC Bioinformatics 6:27.
34. Chin SF, et al. 2007. High-resolution aCGH and expression profiling identifies a novel genomic subtype of ER negative breast cancer. Genome Biol 8:R215.
35. Joosse SA. 4x2 Fisher's exact test. Available from: http://in-silico.net/statistics/fisher_exact_test. Online Source 2010.
36. Dobbin KK, et al. 2008. How Large a Training Set is Needed to Develop a Classifier for Microarray Data? Clin Cancer Res 14:108-14.
37. Tibshirani R, et al. 2002. Diagnosis of multiple cancer types by shrunken centroids of gene expression. Proc Natl Acad Sci U S A 99:6567-72.
38. Dworkin AM, et al. 2009. Methylation not a frequent "second hit" in tumors with germline BRCA mutations. Fam Cancer 8:339-46.
39. Kontorovich T, et al. 2008. Promoter methylation patterns of ATM, ATR, BRCA1, BRCA2 and P53 as putative cancer risk modifiers in Jewish BRCA1/BRCA2 mutation carriers. Breast Cancer Res Treat 116:195-200.
40. Xia F, et al. 2001. Deficiency of human BRCA2 leads to impaired homologous recombination but maintains normal nonhomologous end joining. Proc Natl Acad Sci U S A 98:8644-9.
41. Scully R, et al. 1999. Genetic analysis of BRCA1 function in a defined tumor cell line. Mol Cell 4:1093-9.
42. Moynahan ME, et al. 1999. Brca1 controls homology-directed DNA repair. Mol Cell 4:511-8.
43. Durant ST, et al. 2005. Good timing in the cell cycle for precise DNA repair by BRCA1. Cell Cycle 4:1216-22.
44. Vollebergh MA, et al. 2009. Predicting response to alkylating chemotherapy in breast cancer patients using array comparative genomic hybridization. Cancer Research 69:361S-2S.

Chapter 6

Genomic signature of BRCA1 deficiency in sporadic basal-like breast tumors.

SA Joosse, KI Brandwijk, L Mulder, J Wesseling, J Hannemann and PM Nederlof

Genes, Chromosomes and Cancer. 2011 Feb; 50:71-81.

Genomic signature of BRCA1 deficiency in sporadic basal-like breast tumors

Simon A Joosse, Kim IM Brandwijk, Lennart Mulder, Jelle Wesseling, Juliane Hannemann and Petra M Nederlof

About 10–20% of all breast carcinomas show a basal-like phenotype, while ~90% of breast tumors from BRCA1-mutation carriers are of this subtype. There is growing evidence that BRCA1-mutated tumors are not just a specific subset of the basal-like tumors, but that (the majority of) basal-like tumors show a dysfunctional BRCA1 pathway. This has major treatment implications, because emerging regimens specifically targeting DNA repair mechanisms would then be most effective against these tumors. To further understand the involvement of BRCA1 deficiency in sporadic basal-like tumors, we investigated 41 basal-like tumors for BRCA1 mRNA expression by quantitative real-time polymerase chain reaction, BRCA1 promoter methylation, their genomic profile by array-CGH, and gene expression levels by whole genome expression arrays. Array-CGH results were compared to those of 34 proven BRCA1-mutated tumors. Basal-like tumors were subdivided into two equal groups: deficient and proficient in BRCA1 gene expression. The chromosomal makeup of BRCA1 deficient sporadic basal-like tumors was similar to that of BRCA1-mutated tumors. BRCA1 proficient sporadic basal-like tumors were more similar to nonbasal-like tumors. Only half of the basal-like breast tumors are actually deficient in BRCA1 expression. Gain of chromosome arm 3q is a marker for BRCA1 deficiency in hereditary and sporadic breast tumors.

Genes, Chromosomes and Cancer. 2011 Feb; 50:71-81.

Introduction

Breast cancer is the most frequently occurring cancer among women in the western world. It is a heterogeneous disease, consisting of several tumor subtypes. Identification and characterization of these subtypes are important to understand the pathogenesis of the disease and obtain better treatment options in the future. One of the breast cancer subtypes is called "basal-like" and describes an aggressive tumor group with poor prognosis. Basal-like breast tumors are characterized by the expression of markers often found in normal basal/myoepithelial cells (1) and the absence of the expression of estrogen receptor (ER), progesterone receptor (PR), and ERBB2 (HER2/neu) (2, 3). Because of the lack of expression of these receptors, this group is often referred to as "triple-negative". Among the sporadic breast cancers, the basal-like phenotype represents ~15% of the invasive tumors (3, 4). However, among the hereditary *BRCA1*-mutated breast tumors, this subtype accounts for 80–90% of cases (5, 6). Gene-expression profiles of *BRCA1*-mutated breast tumors show many similarities to those of sporadic basal-like tumors (7), and it has been shown that *BRCA1* mRNA expression is lower in most basal-like tumors compared to matched controls (8). This indicates that loss of function of *BRCA1* is important in basal-like tumors.

Sporadic basal-like tumors can loose *BRCA1* by various mechanisms, such as gene mutation

or promoter hypermethylation. Although in ~10% of the basal-like breast tumors a *BRCA1* mutation can be found (9), screening for gene mutations is generally not allowed in sporadic cancer cases where the patient did not give any informed consent. Furthermore, investigating promoter hypermethylation is not part of routine diagnostics for sporadic breast cancer, although it can be found in a substantial proportion of breast cancer patients (8, 10, 11). We therefore explored whether we could identify a general marker for *BRCA1* deficiency in this study.

Because *BRCA1* is involved in DNA repair by homologous recombination, loss of function will result in accumulation of DNA damage and chromosomal instability. As we and other researchers have shown, *BRCA1*-mutated tumors develop a distinct pattern of chromosomal aberrations (12-17). Some of these aberrations are similar to those of sporadic basal-like tumors (18, 19). However, the use of different detection techniques, control groups, and study designs in these studies makes direct and quantitative comparison between the sporadic basal-like and *BRCA1*-mutated tumors difficult, if not impossible. Thus far, the similarity between these two groups concerning copy number alterations remains elusive.

It would be of clinical and biological relevance to determine whether all or a fraction of the sporadic basal-like tumors are similar to the hereditary *BRCA1*-mutated tumors. The exact location of associated chromosomal aberrations and gene expression changes would result in a better understanding of tumorigenesis due to *BRCA1* deficiency in hereditary and sporadic basal-like tumors and may finally lead to the identification of common therapeutic targets. It has already been shown that breast cancer patients diagnosed with a *BRCA1*-mutation are more sensitive to DNA damage-based chemotherapy than sporadic tumors (20). Additionally, evidence for poly(ADP-ribose) polymerase (PARP) inhibitors efficiently targeting *BRCA1* deficiency is emerging quickly (21). In this study, we show the correlation between copy number alterations of basal-like sporadic breast carcinomas and their *BRCA1* mRNA expression levels.

Materials and methods

Tumor specimens

This study includes two groups of breast cancer cases that were all negative for ER, PR, and ERBB2 expression by immunohistochemistry (IHC) and scored as histological grade III. The first group consists of 41 sporadic basal-like breast tumors of invasive ductal carcinoma (IDC) type, defined as being sporadic as having no family history for any type of cancer, with a mean age at diagnosis of 48 years (range, 26–82), from which gene expression and histopathological data were available from an earlier study from our institute (2). The second group includes 34 breast carcinomas (IDC) from patients with a confirmed pathogenic *BRCA1* germ-line mutation and with a mean age at diagnosis of 38 years (range, 27–61). mRNA, and therefore gene expression data, was not available.

As an additional control, *BRCA1* gene-expression levels were measured in 83 unselected luminal sporadic tumors (IDC) by qRT–PCR, taken from an unrelated study from our institute (22), and included individuals with a mean age at diagnosis of 46 years (range, 27–78). Molecular breast cancer subtypes were determined by the subtype single sample predictor developed by Hu *et al.* (23) for both the basal-like and luminal tumor groups.

Table 1 – Median number of aberrations per tumor group.

Tumor group	Average	Range	StDev	t-test	p-value
Number of aberrations					
Basal-like (n=41)	82.3	48-129	15.1	B1 vs BL	3.0E-03
Basal-likeB1-low (n=22)	90.0	72-129	12.1	B1 vs $BL^{b1\text{-}low}$	0.39
Basal-likeB1-high (n=19)	73.4	48-95	13.4	B1 vs $BL^{b1\text{-}high}$	9.4E-07
BRCA1-mutated (n=34)	90.9	69-113	10.1	B1 vs C	2.8E-07
Non-basal-like (n=23)	75.0	58-103	10.2	$BL^{b1\text{-}low}$ vs $BL^{b1\text{-}high}$	7.5E-08
				$BL^{b1\text{-}low}$ vs C	3.3E-05
				$BL^{b1\text{-}high}$ vs C	0.33
Number of losses					
Basal-likeB1-low (n=22)	43.9	33-69	7.4	$BL^{b1\text{-}low}$ vs $BL^{b1\text{-}high}$	4.6E-4
Basal-likeB1-high (n=19)	33.1	15-52	10.1		
Number of gains					
Basal-likeB1-low (n=22)	28.5	19-39	5.9	$BL^{b1\text{-}low}$ vs $BL^{b1\text{-}high}$	0.86
Basal-likeB1-high (n=19)	28.9	19-45	7.4		
Number of amplifications					
Basal-likeB1-low (n=22)	17.8	7-34	7.6	$BL^{b1\text{-}low}$ vs $BL^{b1\text{-}high}$	0.47
Basal-likeB1-high (n=19)	15.8	2-37	9.3		

P-values are calculated between tumor groups using 2-tailed t-tests. Number of aberrations in basal-like$^{B1\text{-}low}$ and basal-like$^{B1\text{-}high}$ tumors were also separately analyzed for losses, gains, and amplifications. B1=BRCA1-mutated, BL=Basal-like, C=non-basal-like, StDev=Standard Deviation.

As a control group for chromosomal aberrations, array-CGH profiles from 23 sporadic, histological grade III, and carcinomas (IDC) were used. These tumors expressed either one or a combination of ER, PR, and ERBB2 (Supporting Information Table 1) and with a mean age at diagnosis of 45 years (range, 32–60). This group is further referred to as the nonbasal-like tumors.

All experiments involving human tissues were conducted with the permission of the institute's medical ethical advisory board.

Pathological review

The presence of ER, PR, and ERBB2 were determined by revision of immunohistochemical staining that was previously performed using a standard clinical procedure with antibodies: ER AB-14 clone 1D5 + 6F11, titer 1:50 (Neomarkers); PR clone PR-1, titer 1:400 (Immmunologic); and c-erbB-2 clone SP3, titer 1:25 (Neomarkers), respectively. For simplicity, IHC scoring was divided into two classes. If ≥1% of the tumor cells expressed ER or PR, the tumor was scored as positive (+); otherwise, the tumor was scored as negative (-) for the corresponding staining, according to Viale *et al.* (24). ERBB2 scoring was performed according to ASCO/CAP and oncoline guidelines (25, 26). A tumor was scored positive for ERBB2 when a 3+ staining was observed. When a 2+ staining was observed, CISH was performed to determine amplification (+ in case of six spots or more per nucleus) or no amplification (-). A 1+ or negative IHC staining was scored as negative (-). Tumor grade was determined using the modified Bloom–Richardson–Elston staging system (27).

DNA isolation and array-CGH

All sample material used for array-CGH experiments was formalin-fixed, paraffin-embedded tissue from the hospital's pathological archive, collected between 1985 and 2001. DNA was extracted by proteinase-K digestion after deparaffinization, and quality was tested using a multiplex PCR as previously described (13, 28). Tumor and reference DNA were labeled with Cy5 and Cy3, respectively, co-hybridized to a microarray containing 3.5k BAC/PAC-derived DNA segments covering the whole genome with an average spacing of 1 Mb and processed as already described (29). Microarray data were deposited in NCBIs Gene Expression Omnibus and are accessible through GEO Series accession number GSE22401 (basal-like), GSE9021 (*BRCA1*-mutated), and GSE9114 (nonbasal-like).

Aberration detection and quantification

To analyze and visualize chromosomal aberrations, we determined breakpoint locations and estimated copy number levels using the CGH-segmentation algorithm by Picard *et al.* (30). These data are referred to as the "segmentation data". To call the copy number level of aberrations, profile-dependent cutoffs were used that were based on the SD of the middle 50% quantile of the segmented data as described by Chin *et al.* (31). The association of the frequency of a clone being at a neutral, lost, gained, or highly gained copy number between different tumor groups was calculated by using a 2 x 4 Fisher's exact (FE) test (32). Because adjacent BAC clones are expected to be highly correlated, a genomic region was called significant when at least five adjacent BAC clones were calculated to be significant with $p < 0.01$. Using this approach, identifying a region of 5 Mb by chance is < 0.01 (Benjamini), and copy number variations smaller than 5 Mb were also excluded from the analyses.

Methylation MLPA

Methylation of the *BRCA1* promoter was investigated using a methylation-specific MLPA kit according to the manufacturer's protocol (ME001B, MRC-Holland, The Netherlands). This kit includes probes against the gene promoters of *APC*, *ATM*, *BRCA1*, *BRCA2*, *CASP8*, *CD44*, *CDH13*, *CDKN1B*, *CDKN2A*, *CDKN2B*, *CHFR*, *DAPK1*, *ESR1*, *FHIT*, *GSTP1*, *HIC1*, *IGSF4*, *MLH1*, *PTEN*, *RARB*, *RASSF1*, *TIMP3*, *TP73*, and *VHL* and includes 15 reference probes. Basal-like tumors that show *BRCA1* promoter methylation were classified as BRCA1-deficient. The *BRCA1* mRNA expression levels of these samples were used to calculate the 95% reference range. Next, the reference range was applied to *BRCA1* mRNA expression levels of basal-like samples without *BRCA1* promoter methylation. Samples with expression levels inside the reference range were included into the *BRCA1*-deficient group, and samples outside the 95% reference range were classified as *BRCA1*-proficient.

Quantitative RT–PCR

Expression levels of *BRCA1* were assessed by quantitative real-time polymerase chain reaction (qRT-PCR) as a method independent of the microarray data to prevent array-based bias in the 41 basal-like and 83 luminal breast tumors. The TaqMan Gene Expression Assay for *BRCA1* (#Hs01556193_m1, Applied Biosystems, Foster City, CA) was used for this purpose. The reactions were performed according to the manufacturer's protocol with 10 ng cDNA (2 ng/µl) for each sample, resulting in an amplification

product of 59 nucleotides. Expression levels of β-actin and GAPDH were measured as endogenous controls, and cDNA from MCF-7 cells in different dilutions was used to obtain a standard curve. qRT-PCR runs were performed on the 7500 Fast System, and analyses were conducted using 7500 Fast Real-Time PCR Software version 1.3.1. Expression levels were calculated by the relative standard curve method.

Results

BRCA1 expression in basal-like tumors

To investigate whether basal-like tumors could be subdivided into groups based on their level of *BRCA1* expression, *BRCA1*-deficient tumors were identified by measuring *BRCA1* promoter methylation. Fourteen of the basal-like tumors (34%) showed *BRCA1* promoter methylation. Promoter methylation usually results in silencing of the gene. Indeed, all cases with methylated *BRCA1* promoter showed low-*BRCA1* mRNA expression compared to the remaining samples ($p < 5.0e-5$, two-sided t-test). We defined the basal-like tumors with methylated *BRCA1* promoter as the basal-like "BRCA1-low" group, which is subsequently referred to as the basal-like$^{B1\text{-}low}$ group. Using 95% reference range on their *BRCA1* expression levels, we determined which of the basal-like samples without *BRCA1* promoter methylation could be included in the basal-like$^{B1\text{-}low}$ group. All samples outside the 95% reference range were classified as basal-like "BRCA1-high" and are subsequently referred to as the basal-like$^{B1\text{-}high}$ group. As can be seen in Supporting Information Fig. 2, a binominal distribution could be used to describe the methylation results (methylated and unmethylated), because no continuous correlation was observed between levels of methylation and mRNA expression. Twenty-one samples (51%) were included in the basal-like$^{B1\text{-}low}$ group, and 20 samples (49%) were included in the basal-like$^{B1\text{-}high}$ group (Fig. 1).

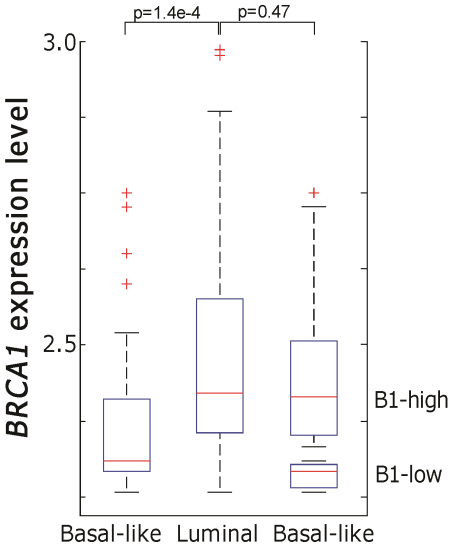


Figure 1 – BRCA1 mRNA expression. *Box plots showing relative BRCA1 mRNA expression in sporadic basal-like breast tumors (left, n=41), luminal breast tumors (middle, n=83), and the same basal-like breast tumor cohort separated on the basis of BRCA1 deficiency as described in Methods.*

Expression of *BRCA1* mRNA in the complete basal-like tumor cohort was significantly lower compared to the 83 sporadic luminal tumors ($p = 1.4e-4$, two-sided t-test) (Fig. 1). Median relative expression of *BRCA1* was 0.24 and 0.69 in basal-like and luminal breast tumors, respectively. For the basal-like$^{B1\text{-}low}$ group, the median expression level was 0.17, while it was 0.66 for the basal-like$^{B1\text{-}high}$ group, which is comparable to that of the luminal tumors ($p = 0.47$, two-sided t-test) (Fig. 1 and Support-

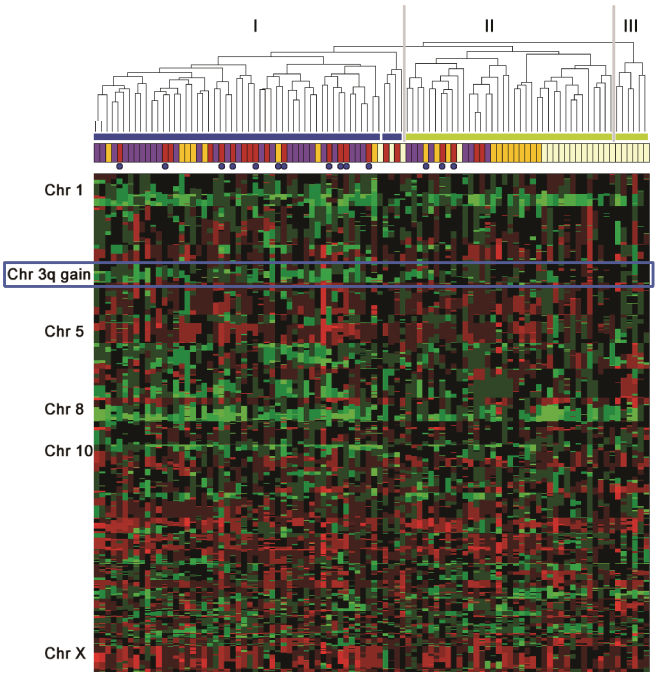


Figure 2 – Clustering of array-CGH profiles. *Unsupervised hierarchical clustering (complete linkage) of BRCA1-mutated ■, basal-like$^{B1\text{-}low}$ ■, basal-like$^{B1\text{-}high}$ ■, and nonbasal-like ■ tumors. While the BRCA1-deficient cases cluster together in cluster I ■, the BRCA1-proficient cases are located in clusters II-III ■. Shown is the heat map of the CGH segmentation data, where green is positive log2(ratio) and red negative. Here, gain on chromosome 3q has been highlighted, which has been found in BRCA1-deficient, and not BRCA1-proficient tumors, as a significant difference in later analyses. Blue circles indicate samples with methylation of the BRCA1 promoter ●.*

ing Information Table 1). These results indicate that approximately half of the basal-like breast tumors express BRCA1 at levels similar to luminal tumors and half express BRCA1 at a significantly lower level, if at all.

RASSF1 gene promoter methylation and *BRCA1* expression

Although promoter methylation of *BRCA1* was abundant in sporadic basal-like tumors, we could not detect it in nonbasal-like tumors (n = 23). Additionally, in a larger series (n > 150), we have seen that methylation of the *BRCA1* promoter in sporadic nonbasal-like tumors was rare (3%, unpublished data). Besides *BRCA1*, 23 other tumor suppressor genes were simultaneously investigated for promoter methylation including *RASSF1*. Methylation of the *RASSF1* promoter is reported to be less abundant in *BRCA1*-associated breast cancer (33). As shown in Supporting Information Table 1, a trend can be observed in our data set, such that the *RASSF1* promoter was more often methylated in basal-like$^{B1\text{-}high}$ tumors compared to basal-like$^{B1\text{-}low}$ tumors ($p = 0.034$, two-sided t-test, uncorrected for multiple testing). These data were similar to the methylation patterns of *BRCA1*-mutated tumors, in which methylation of the promoter of *RASSF1* was also absent; nonbasal-like tumors on the other hand often showed methylation of the *RASSF1* promoter (73%, unpublished data, n > 150). These results indicated that *BRCA1* expression could be correlated with methylation of the *RASSF1* promoter. Methylation states of all individual basal-like tumor samples are listed in Supporting Information Table 1.

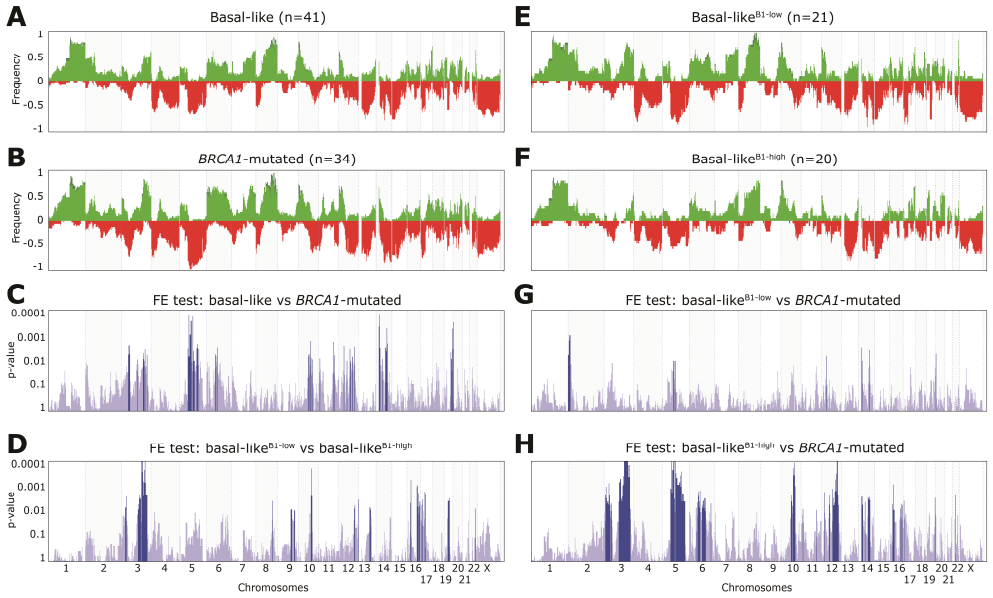


Figure 3 – Aberration frequencies. *Frequency plots of basal-like (A), BRCA1-mutated (B), basal-like$^{B1\text{-}low}$ (E) and basal-like$^{B1\text{-}high}$ (F) tumors, showing the amount of gain (green) and loss (red) along the whole genome. Fisher's exact test was used to determine the most significant regions between BRCA1-mutated and basal-like tumors (C), basal-like$^{B1\text{-}low}$ and basal-like$^{B1\text{-}high}$ tumors (D), BRCA1-mutated and basal-like$^{B1\text{-}low}$ tumors (G), and BRCA1-mutated and basal-like$^{B1\text{-}high}$ tumors (H). P-values are minus log10 transformed and depicted as light blue (p>0.01) or dark blue (p<0.01).*

Level of chromosomal imbalance is associated with level of BRCA1 expression

Histological high-grade tumors show extensive chromosomal imbalance (34). To determine whether there was a difference in the level of chromosomal imbalance among *BRCA1*-mutated, basal-like$^{B1\text{-}low}$, basal-like$^{B1\text{-}high}$, and grade III nonbasal-like control tumors, the number of aberrations (*i.e.*, calculated segments outside the profile-dependent thresholds) was counted (Table 1 and Supporting Information Fig. 3). Although all investigated tumors were histological grade III, the basal-like$^{B1\text{-}high}$ group showed significantly less aberrations compared to the *BRCA1*-mutated and basal-like$^{B1\text{-}low}$ groups ($p < 0.01$), but there was no significant difference when these tumors were compared to the nonbasal-like control tumors ($p = 0.33$). The average number of aberrations of the *BRCA1*-mutated and basal-like$^{B1\text{-}low}$ tumor groups was very similar ($p = 0.39$, Table 1). The difference in the number of aberrations was mainly caused by more losses found in basal-like$^{B1\text{-}low}$ tumors (Table 1). These results imply that the level of chromosomal imbalance is not dependent on histological grade, but on *BRCA1* status.

Copy number alterations in hereditary and sporadic breast tumors

To investigate the correlation among the genetic signatures (aberrations) of the different

tumor groups, unsupervised hierarchical clustering (complete linkage correlation) was performed on the whole-genome segmentation data of sporadic basal-like, *BRCA1*-mutated, and control tumors. Figure 2 shows that basal-like$^{B1\text{-}low}$ tumors cluster with *BRCA1*-mutated tumors, while many basal-like$^{B1\text{-}high}$ tumors cluster separately from the *BRCA1*-deficient tumors along with grade III nonbasal-like tumors. This indicates that similar aberrations are present in *BRCA1*-deficient tumors that are independent of the cause of the deficiency (*i.e.*, mutation or methylation). Additionally, tumors proficient in *BRCA1* develop a different signature of aberrations.

Next, supervised analyses were performed based on the frequency of copy-number alterations. We published previously that *BRCA1*-mutated tumors show a different spectrum of aberrations compared to the general population of sporadic breast cancer (13). In the present study, we also show that the spectrum of aberrations was very different when only compared to grade III nonbasal-like sporadic tumors (Supporting Information Fig. 1). Interestingly, the sporadic basal-like tumors were much more similar to the *BRCA1*-mutated breast tumors (Figs. 3A–3C). Panels A and B of Figure 3 display the genome-wide frequency of gain (green) and loss (red) in basal-like and *BRCA1*-mutated breast tumors, respectively. Most tumors in both groups showed the common breast cancer aberrations, namely, gain of chromosome arms 1q and 8q and loss of 8p. Moreover, previously identified aberrations specific for *BRCA1*-associated, ER negative, or basal-like tumors were also found, as represented by gains of regions in chromosome arms 3q, 6p, 10p, 12p, and 21q and losses of regions in 3p and 5q (see Supporting Information 2 for a detailed whole-genome description and exact locations). Figure 3C shows p-values calculated by Fisher's exact test based on the number of different aberrations in both groups. Several genomic regions (*e.g.*, located on 3q, 5q, 14q and 19q, see Supporting Information Table 2 for full list) were identified with significantly different frequencies between *BRCA1*-mutated and basal-like tumors ($p < 0.01$, indicated in dark blue). To determine whether these aberrations were correlated with *BRCA1* expression, the following two comparisons were performed.

First, an aberration frequency comparison was made between basal-like$^{B1\text{-}low}$ and basal-like$^{B1\text{-}high}$ tumors, which revealed several significantly different genomic regions (Fig. 3D and Supporting Information Table 2). In basal-like$^{B1\text{-}low}$ tumors, 3p24-p22.3, 3q13-q26.2, 13q22, 16p12-p11, and 16q22-q24 were more often gained, and 9q, 9q31.3-q33.1, 10q23.1-q23.31, and 12q23.3 were more often lost, compared to basal-like$^{B1\text{-}high}$ tumors.

Second, the basal-like tumor subgroups were compared to *BRCA1*-mutated tumors using a similar frequency analysis as outlined earlier. Figures 3E and 3F depict the aberration frequencies in the basal-like subgroups, whereas Figures 3G and 3H show the corresponding p-values calculated by Fisher's exact test (see also Supporting Information Table 2). Basal-like$^{B1\text{-}low}$ breast tumors were most similar to *BRCA1*-mutated tumors, and only two small genomic regions at 2p24-25 and 14q24 presented with a significantly different frequency ($P < 0.01$ at ≥ 5 adjacent BAC clones). Basal-like$^{B1\text{-}high}$ tumors showed many more aberrations with significantly different frequencies (Fig. 3H and Supporting Information Table 2), which included 3p21-p26, 3q11-26, 5q11-q33, 6q12-21, 10q21-q23, 12q13.13-q14.1, 12q21.2-q24.22, 14q11-q12, 14q23-q24, 16p12-p11, 16q22-q24, and 17p.

These results indicate that basal-like tumors

are quite similar to *BRCA1*-mutated tumors. However, differences among these groups are still present, which were only identified in basal-like$^{B1\text{-}high}$ tumors. Basal-like$^{B1\text{-}low}$ tumors are almost identical to *BRCA1*-mutated tumors.

H2B gene regulation is associated with *BRCA1* expression

Gene-expression data from Kreike and colleagues (2) were investigated to determine whether *BRCA1* mRNA expression relates to different gene-expression patterns in basal-like tumors. Differentially expressed genes in basal-like$^{B1\text{-}low}$ tumors could reveal biological processes associated with *BRCA1* deficiency. Additionally, the analysis of gene-expression patterns in basal-like$^{B1\text{-}high}$ tumors could elucidate differences within basal-like breast cancer.

To evaluate the statistical significance of gene-expression patterns between basal-like$^{B1\text{-}low}$ and basal-like$^{B1\text{-}high}$ tumors, the significance analysis of microarrays method (35) was used for all the 5830 significant genes. For a false discovery rate of 5%, delta was 0.488. Fifty-seven unique genes were found to be significantly downregulated in basal-like$^{B1\text{-}low}$ tumors when compared with basal-like$^{B1\text{-}high}$ tumors (Supporting Information 3) and none was upregulated. DAVID (36) was used to perform functional annotation clustering. From the significant genes, 7 (12.5%) were selected to be at the most significant cluster, Histone *H2B* ($p = 5.5e\text{-}9$, Benjamini), and consisted of the genes *H2BFS*, *HIST1H2BB*, *HIST1H2BD*, *HIST1H2BJ*, *HIST1H2BM*, *HIST1H2BO*, and *HIST1H2BE*.

Discussion

Breast carcinomas that are negative for ER, PR, and ERBB2, and of a basal-like subtype, are a distinct breast cancer subgroup associated with poor prognosis. Literature concerning the relationship between *BRCA1*-pathway deficiency and basal-like breast cancer has been increasing rapidly in the last few years (37-40). Because of their *BRCA1* deficiency, it is not surprising that basal-like and *BRCA1*-mutated breast tumors are similar in many aspects (20). It is of high clinical and biological interest to identify the similarities between these groups, which could lead to the identification of common therapeutic targets.

BRCA1 expression in basal-like tumors

Turner *et al.* (8) showed that basal-like breast tumors express less *BRCA1* mRNA compared to controls. Although slightly different definitions for basal-like and control cases were used, we can confirm these results and see a similar picture when comparing basal-like with luminal breast tumors (Fig. 1).

In our study, many basal-like breast tumor samples showed methylation of the *BRCA1* promoter (34%), which was significantly correlated with *BRCA1* gene downregulation. Although our methylation results accord with the findings of other studies, which showed that *BRCA1* promoter methylation is found in 32% of basal-like samples (10, 41), Turner *et al.* (8) detected a lower rate of 12% of ductal basal-like carcinomas exhibiting *BRCA1* promoter methylation. This difference might also be the result of the use of slightly different definitions for basal-like tumors. Besides promoter methylation, gene mutation can be the cause of loss of function and lowered gene expression. A recent study has shown that *BRCA1* is mutated in about 10% of sporadic basal-like breast tumors. It is therefore suggested that young women with early-onset triple-negative breast cancer are candidates for

mutation screening, regardless of family history of breast or ovarian cancer (9). Unfortunately, a limitation in our study was that permission for mutation screening was not granted for the sporadic tumors, because most samples where archived 10–20 years ago, and no indication of a hereditary mutation was present. From the remaining (non-methylated) samples, an additional 17% of cases also showed a down-regulated *BRCA1* expression, which may actually be due to a real *BRCA1* mutation. Separating basal-like$^{B1\text{-}low}$ cases based on methylation status, however, did not change any of our results (data not shown).

Forty-nine percent of all basal-like samples showed a *BRCA1* expression comparable to that of luminal tumors (Fig. 1).

BRCA1 proficiency

Although basal-like tumors are similar to *BRCA1*-mutated breast tumors in regard to gene expression profiles (7), our study, involving unsupervised analyses of genomic aberrations, showed that some basal-like tumors cluster away from *BRCA1*-mutated tumors and cluster together with nonbasal-like tumors (Fig. 2). These tumors belong primarily to the *BRCA1* expressing group (basal-like$^{B1\text{-}high}$). Analyses of aberration frequencies in basal-like$^{B1\text{-}high}$ tumors show a different pattern compared to *BRCA1*-mutated tumors. A significant difference involved a gain of chromosome arm 3q, an aberration very abundant in *BRCA1*-mutated and basal-like$^{B1\text{-}low}$ breast tumors, but almost always absent in basal-like$^{B1\text{-}high}$ and nonbasal-like breast tumors. Significant differences among the tumor groups were found not only for the frequency of specific aberrations, but also for the total number of aberrations. It can be presumed that the number of chromosomal aberrations (*i.e.*, level of genomic imbalance) is associated with the deficiency in a specific DNA repair pathway within a defined tumor population. Basal-like$^{B1\text{-}high}$ and control nonbasal-like tumors showed significantly fewer aberrations compared to basal-like$^{B1\text{-}low}$ and *BRCA1*-mutated breast tumors, indicating a difference in the handling of DNA repair. Frequent *RASSF1* promoter methylation in basal-like$^{B1\text{-}high}$ cases makes this group additionally more similar to nonbasal-like breast tumors where methylation of *RASSF1* promoter is common. *RASSF1* promoter methylation in basal-like$^{B1\text{-}low}$ and *BRCA1*-mutated breast tumors is rare, as noted by other researchers (33). Taken together, our results indicate that basal-like$^{B1\text{-}high}$ tumors are very similar to nonbasal-ike grade III breast tumors and less similar to proven *BRCA1*-deficient breast tumors.

BRCA1 deficiency

Our results indicate that basal-like breast tumors with *BRCA1* deficiency (basal-like$^{B1\text{-}low}$ tumors) show many similarities to hereditary *BRCA1*-mutated breast tumors in regard to genomic aberrations using direct comparisons. Unsupervised clustering and supervised analyses showed that the well-known BRCA1-specific aberrations located along 3p (loss), 3q (gain), 5q (loss), and 12q (loss) are shared between the two groups. Furthermore, the total number of aberrations and rare methylation of the *RASSF1* promoter is similar. When mRNA expression profiles of basal-like$^{B1\text{-}low}$ and basal-like$^{B1\text{-}high}$ tumors were compared, only a few differences were found. The most prominent gene cluster that was downregulated in basal-like$^{B1\text{-}low}$ tumors was histone *H2B*. Downregulation of histone gene expression has been shown to occur in response to DNA double-strand breaks (42). We postulate that, due to lack of functional *BRCA1*, accumulation of doublestrand breaks is high in

basal-like$^{B1\text{-}low}$ tumors, which keeps histone *H2B* downregulated. Because basal-like$^{B1\text{-}high}$ tumors show fewer aberrations and higher *BRCA1* gene expression levels, histone *H2B* is normally regulated in these tumors. However, because p53 deficiency alleviates H2B downregulation (42) and p53 is frequently mutated in basal-like and *BRCA1*-mutated breast cancer (43, 44), the correlation is counter intuitive, and the biological relevance of this observation needs to be further investigated. Nevertheless, these results indicate that the loss of *BRCA1* causes only minor and indirect gene expression changes in basal-like tumors and might explain why basal-like tumors always cluster together in gene-expression studies while greater heterogeneity is found among basal-like tumors in CGH studies.

Chromosome arm 3q as a BRCA1 deficiency marker

Several studies have investigated chromosomal aberrations in both sporadic basal-like (18, 19) and *BRCA1*-mutated tumors (12-14, 16, 17, 45). These studies show gain along 3q and 10p and loss along 5q as the most common aberrations in both groups [see also (46) Supporting Information Table 4]. This has led to the presumption that these tumor groups are similar in terms of chromosomal aberrations (40). However, in-depth analysis of these studies also reveals many discrepancies. It should be noted that the different study designs limit interstudy comparisons and hamper localization of the exact chromosomal boundaries of the aberrations that are shared by the tumor groups. In our study, we showed that only half of the basal-like tumors are very similar to *BRCA1*-mutated tumors. The genomic instability and specific aberrations that develop in basal-like tumors are strongly associated with *BRCA1* mRNA expression. Our cluster analyses even suggest that the genomic signature as a consequence of (the lack of) *BRCA1* expression could be as prominent as the dominant ER signature, because the basal-like$^{B1\text{-}high}$ tumors (ER-negative) cluster together with ER-positive, grade III, sporadic tumors instead of residing with tumors having equal ER, PR, and ERBB2 status.

In previous studies, we and other researchers have identified gain of chromosome arm 3q to be an important marker of *BRCA1*-mutated tumors when compared with sporadic tumors (12, 13, 17). In the study presented here, we showed that gain of 3q is not only present in hereditary tumors but also present as most significant and in the highest frequency in sporadic basal-like *BRCA1*-deficient tumors compared to basal-like sporadic tumors expressing *BRCA1*. Our results indicate that gain of 3q (smallest common region of gain: 3q24) could serve as a potential marker of *BRCA1* deficiency.

Conclusion

We conclude that only half of the basal-like breast tumors are actually deficient in *BRCA1* expression. Lack of *BRCA1* leads to a large amount of aberrations and accumulation of DNA damage, but not to many direct differences in gene-expression profiles. Gain of chromosome arm 3q is a marker for *BRCA1* deficiency in hereditary and sporadic breast tumors. Future research should include prescreening of basal-like tumors for gain of 3q to initiate additional *BRCA1* diagnostics (*i.e.*, mutation and promoter methylation screening) and to prove the clinical relevance of the similarity between *BRCA1*-mutated and basal-like$^{B1\text{-}low}$ breast tumors.

Acknowledgments

We acknowledge the Central Microarray Facility (NKI) for the production of the microarrays and technical assistance and Bas Kreike for kindly providing us with the gene expression microarray data.

References

1. Perou CM, et al. 2000. Molecular portraits of human breast tumours. Nature 406:747-52.
2. Kreike B, et al. 2007. Gene expression profiling and histopathological characterization of triple-negative/basal-like breast carcinomas. Breast Cancer Res 9:R65.
3. Nielsen TO, et al. 2004. Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. Clin Cancer Res 10:5367-74.
4. Abd El-Rehim DM, et al. 2005. High-throughput protein expression analysis using tissue microarray technology of a large well-characterised series identifies biologically distinct classes of breast cancer confirming recent cDNA expression analyses. Int J Cancer 116:340-50.
5. Foulkes WD, et al. 2003. Germline BRCA1 mutations and a basal epithelial phenotype in breast cancer. J Natl Cancer Inst 95:1482-5.
6. Lakhani SR, et al. 2005. Prediction of BRCA1 status in patients with breast cancer using estrogen receptor and basal phenotype. Clin Cancer Res 11:5175-80.
7. Sorlie T, et al. 2003. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proc Natl Acad Sci U S A 100:8418-23.
8. Turner NC, et al. 2007. BRCA1 dysfunction in sporadic basal-like breast cancer. Oncogene 26:2126-32.
9. Young SR, et al. 2009. The prevalence of BRCA1 mutations among young women with triple-negative breast cancer. BMC Cancer 9:86.
10. Chen Y, et al. 2009. BRCA1 promoter methylation associated with poor survival in Chinese patients with sporadic breast cancer. Cancer Sci 100:1663-7.
11. Vasilatos SN, et al. 2009. CpG island tumor suppressor promoter methylation in non-BRCA-associated early mammary carcinogenesis. Cancer Epidemiol Biomarkers Prev 18:901-14.
12. Jonsson G, et al. 2005. Distinct genomic profiles in hereditary breast tumors identified by array-based comparative genomic hybridization. Cancer Res 65:7612-21.
13. Joosse SA, et al. 2009. Prediction of BRCA1-association in hereditary non-BRCA1/2 breast carcinomas with array-CGH. Breast Cancer Res Treat 116:479-89.
14. Stefansson OA, et al. 2009. Genomic profiling of breast tumours in relation to BRCA abnormalities and phenotypes. Breast Cancer Res 11:R47.
15. Tirkkonen M, et al. 1997. Distinct somatic genetic changes associated with tumor progression in carriers of BRCA1 and BRCA2 germ-line mutations. Cancer Res 57:1222-7.
16. van Beers EH, et al. 2005. Comparative genomic hybridization profiles in human BRCA1 and BRCA2 breast tumors highlight differential sets of genomic aberrations. Cancer Res 65:822-7.

17. Wessels LF, et al. 2002. Molecular classification of breast carcinomas by comparative genomic hybridization: a specific somatic genetic profile for BRCA1 tumors. Cancer Res 62:7110-7.
18. Bergamaschi A, et al. 2006. Distinct patterns of DNA copy number alteration are associated with different clinicopathological features and gene-expression subtypes of breast cancer. Genes Chromosomes Cancer 45:1033-40.
19. Vincent-Salomon A, et al. 2007. Identification of typical medullary breast carcinoma as a genomic sub-group of basal-like carcinomas, a heterogeneous new molecular entity. Breast Cancer Res 9:R24.
20. James CR, et al. 2007. BRCA1, a potential predictive biomarker in the treatment of breast cancer. Oncologist 12:142-50.
21. Linn SC, et al. 2009. Clinical relevance of the triple-negative breast cancer concept: genetic basis and clinical utility of the concept. Eur J Cancer 45 Suppl 1:11-26.
22. de Ronde JJ, et al. 2009. Concordance of clinical and molecular breast cancer subtyping in the context of preoperative chemotherapy response. Breast Cancer Res Treat 119:119-26.
23. Hu Z, et al. 2006. The molecular portraits of breast tumors are conserved across microarray platforms. BMC Genomics 7:96.
24. Viale G, et al. 2007. Prognostic and predictive value of centrally reviewed expression of estrogen and progesterone receptors in a randomized trial comparing letrozole and tamoxifen adjuvant therapy for postmenopausal early breast cancer: BIG 1-98. J Clin Oncol 25:3846-52.
25. oncoline. 1-9-2008. Oncoline Guidelines
26. Wolff AC, et al. 2007. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. J Clin Oncol 25:118-45.
27. Genestie C, et al. 1998. Comparison of the prognostic value of Scarff-Bloom-Richardson and Nottingham histological grades in a series of 825 cases of breast cancer: major importance of the mitotic count as a component of both grading systems. Anticancer Res 18:571-6.
28. van Beers EH, et al. 2006. A multiplex PCR predictor for aCGH success of FFPE samples. Br J Cancer 94:333-7.
29. Joosse SA, et al. 2007. Automated array-CGH optimized for archival formalin-fixed, paraffin-embedded tumor material. BMC Cancer 7:43.
30. Picard F, et al. 2005. A statistical approach for array CGH data analysis. BMC Bioinformatics 6:27.
31. Chin SF, et al. 2007. High-resolution aCGH and expression profiling identifies a novel genomic subtype of ER negative breast cancer. Genome Biol 8:R215.
32. Joosse SA. 4x2 Fisher's exact test. Available from: http://in-silico.net/statistics/fisher_exact_test. Online Source 2010.
33. Suijkerbuijk KP, et al. 2008. Methylation is less abundant in BRCA1-associated compared with sporadic breast cancer. Ann Oncol 19:1870-4.
34. Reis-Filho JS, et al. 2005. The molecular genetics of breast cancer: the contribution of comparative genomic hybridization. Pathol Res Pract 201:713-25.
35. Tusher VG, et al. 2001. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A 98:5116-21.

36. Dennis G, Jr., et al. 2003. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol 4:3.
37. Melchor L, et al. 2008. An integrative hypothesis about the origin and development of sporadic and familial breast cancer subtypes. Carcinogenesis 29:1475-82.
38. Rakha EA, et al. 2008. Basal-like breast cancer: a critical review. J Clin Oncol 26:2568-81.
39. Reis-Filho JS, et al. 2008. Triple negative tumours: a critical review. Histopathology 52:108-18.
40. Turner NC, et al. 2006. Basal-like breast cancer and the BRCA1 phenotype. Oncogene 25:5846-53.
41. Galizia E, et al. 2010. BRCA1 expression in triple negative sporadic breast cancers. Anal Quant Cytol Histol 32:24-9.
42. Su C, et al. 2004. DNA damage induces downregulation of histone gene expression through the G1 checkpoint pathway. EMBO J 23:1133-43.
43. Manie E, et al. 2009. High frequency of TP53 mutation in BRCA1 and sporadic basal-like carcinomas but not in BRCA1 luminal breast tumors. Cancer Res 69:663-71.
44. Holstege H, et al. 2009. High incidence of protein-truncating TP53 mutations in BRCA1-related breast cancer. Cancer Res 69:3625-33.
45. Taylor J, et al. 2006. A tail strength measure for assessing the overall univariate significance in a dataset. Biostatistics 7:167-81.
46. Joosse SA, et al. 2010. Prediction of BRCA2-association in hereditary breast carcinomas using array-CGH. Breast Cancer Res Treat .

Chapter 7

A non-BRCA1/2 hereditary breast cancer sub-group defined by aCGH profiling of genetically related patients.

MA Didraga, EH van Beers, **SA Joosse**, KI Brandwijk, RA Oldenburg, LF Wessels, FB Hogervorst, MJ Ligtenberg, N Hoogerbrugge, S Verhoef, P Devilee and PM Nederlof

Breast Cancer Res Treat. 2011 Nov; 130(2):425-36

A non-BRCA1/2 hereditary breast cancer sub-group defined by aCGH profiling of genetically related patients

Mihaela A Didraga, Erik H van Beers, **Simon A Joosse**, Kim IM Brandwijk, Rogier A Oldenburg, Lodewijk FA Wessels, Frans BL Hogervorst, Marjolijn J Ligtenberg, Nicolin Hoogerbrugge, Senno Verhoef, Peter Devilee and Petra M Nederlof

Germline mutations in BRCA1 and BRCA2 explain approximately 25% of all familial breast cancers. Despite intense efforts to find additional high-risk breast cancer genes (BRCAx) using linkage analysis, none have been reported thus far. Here we explore the hypothesis that BRCAx breast tumors from genetically related patients share a somatic genetic etiology that might be revealed by array comparative genomic hybridization (aCGH) profiling. As BRCA1 and BRCA2 tumors can be identified on the basis of specific genomic profiles, the same may be true for a subset of BRCAx families. Analyses used aCGH to compare 58 non-BRCA1/2 familial breast tumors (designated BRCAx) to sporadic (non-familiar) controls, BRCA1 and BRCA2 tumors. The selection criteria for BRCAx families included at least three cases of breast cancer diagnosed before the age of 60 in the family, and the absence of ovarian or male breast cancer. Hierarchical cluster analysis was performed to determine sub-groups within the BRCAx tumor class and family heterogeneity. Analysis of aCGH profiles of BRCAx tumors indicated that they constitute a heterogeneous class, but are distinct from both sporadic and BRCA1/2 tumors. The BRCAx class could be divided into sub-groups. One subgroup was characterized by a gain of chromosome 22. Tumors from family members were classified within the same sub-group in agreement with the hypothesis that tumors from the same family would harbor a similar genetic background. This approach provides a method to target a sub-group of BRCAx families for further linkage analysis studies.

With kind permission from Springer Science+Business Media: Breast Cancer Res Treat. 2011 Nov; 130(2):425-36.

Introduction

According to the World Health Organization, breast cancer is the most common cancer in women worldwide, representing 16% of all female cancers (1). It is estimated that 10% of all breast cancer occurrence is familial and that known susceptibility genes account for only about 25% of familial breast cancer cases. Some families are explained by alterations in the high-penetrance breast cancer predisposition genes *BRCA1* (2) and *BRCA2* (3, 4), fewer by moderate-penetrance genes like *CHEK2* (5), *ATM* (6), *BRIP1* (7), and *PALB2* (8), while an unknown fraction is expected to be more truly polygenic, meaning that susceptibility is conferred by the joint action of several low-risk loci, each with a small effect on breast cancer risk (9-11). Other genes (*PTEN* (12), *STK11* (13), *CDH1* (14), *TP53* (15-18)) are associated with syndromes in which the incidence of breast cancer is elevated, but the actual risk is difficult to estimate (19).

Although clear progress in our understanding of the genetic susceptibility to breast cancer has occurred during the past 15 years, genetic studies have failed to identify additional high-risk breast cancer genes. The majority (75%) of families with multiple breast cancer cases do not segregate mutations in the known breast cancer genes. It is presently not known how many breast cancer susceptibility genes exist, how many will fall into the class of rare high-risk (*e.g.*, *BRCAx*) or common low-risk susceptibility genes, and if and how these factors interact with each other to cause susceptibility (a polygenic model). Thus, the remainder of familial risk might be attributable to an unknown number of genes (both recessive and dominant), combinations thereof (10, 20-25), or to common variants of genes each conferring a low risk (9, 26, 27). Linkage analysis, mutational screening of candidate genes, and association studies have been the most common methods used in the search for new genetic predisposition factors. Ponder *et al.* (28) have argued that linkage is appropriate to detect *BRCA*-like genes (rare and highly penetrant), but not common low-risk variants, while the reverse is true for association studies (28, 29). A linkage study of 149 BRCAx families, including several from the present study, failed to identify significant LOD scores that might signal a new breast cancer susceptibility gene (30), suggesting that perhaps more than one risk-conferring locus was involved. Supporting the latter notion, population-based epidemiological studies have shown that most of the excess familial clustering of breast cancer is distributed across many families, each with a small number of cases, rather than in a few very extensive families (31, 32).

Array comparative genomic hybridization (aCGH) was developed to investigate copy number changes on a genome-wide scale using one single experiment (33). However, the sensitivity of aCGH to tumor-derived DNA samples is limited by several factors, related to sample quality (*e.g.*, contamination with non-tumor cells such as lymphocytes and stromal cells, intratumoral heterogeneity) or technical parameters such as array resolution. DNA copy number may also be measured using SNP arrays (34), which measure the loss of heterozygosity (LOH) in addition to copy numbers. Since it can be applied to formalin-fixed paraffin-embedded (FFPE) material, aCGH is a valuable tool when fresh tissue is not available, which is often the case when family screening includes deceased relatives.

Considering the direct connection between genetic alterations and oncogenic pathways (35-41), we hypothesized that families with similar aCGH profiles may harbor mutations in the same susceptibility gene(s). Consequently, specific aCGH aberrations recurrent among family members or even different families may be indicators of a shared genetic defect. Indeed, we have been able to demonstrate that breast tumors with a known gene-defect in *BRCA1* or *BRCA2* display distinctive aCGH profiles, which allowed us to develop predictive genetic tests for *BRCA1*-mutation (41) and *BRCA2*-mutation (42) carriers. Motivated by our previous results from the BRCA1/2 classifiers, we compared aCGH profiles derived from BRCAx (n = 58) tumors with those of the sporadic control (n = 49), *BRCA1*-mutated (n = 28) and *BRCA2*-mutated (n = 28) tumors. Although breast tumors of the BRCAx type constitute a heterogeneous class displaying widespread genomic instability (38, 39), we identified one distinct sub-group of tumors from the BRCAx class which harbor similar genomic profiles.

Materials and methods

Patients and tumor specimens

Genomic tumor DNA of sufficient quality for automated aCGH (43) was isolated from 58 FFPE primary tissue blocks (30) belonging to 27 different families. Our criteria for classifying the samples as BRCAx tumors were as follows: at least three cases of breast cancer in the family, diagnosed before the age of 60, from which genotype could be determined (n = 216) or inferred (n = 20); families with cases of ovarian cancer or male breast cancer were excluded, and occurrences of other types of cancer were ignored (30).

Forty-nine sporadic control tumors were selected from the institute's archival tissue bank without any indication of familial risk for breast or ovarian cancer, and no history of visiting a clinical genetics center for breast cancer. These control cases were all sporadic invasive ductal breast carcinomas (IDC) from our previous study (41). Out of the 27 families contributing together 58 BRCAx tumor DNAs of sufficient quality, ten families contributed each a single aCGH profile, eight families two, five families three, three families four, and one family five aCGH profiles. Immunohistochemistry (ER, PR, ERBB2, p53) was performed as described before (44).

The 58 BRCAx tumors were also screened for *CHEK2* 1100delC germline mutation, as described by Oldenburg *et al.* (45).

For comparison, this study further included 28 aCGH profiles of *BRCA1*-mutated tumors and 28 aCGH profiles of *BRCA2*-mutated tumors, from our previous studies (41, 42). A single batch of normal reference genomic DNA was used in the array hybridizations (for its preparation see next section).

Array-CGH

Tumor DNA was extracted from ten FFPE sections, each of 10 µm thickness, as previously described (46). Reference DNA was prepared from whole blood of a reference panel composed of six healthy females and sonicated to an average fragment length comparable to that of the FFPE tumor DNA (600 bp). The control and tumor DNA were labeled with ULS-Cy3 and ULS-Cy5, respectively, and employed in ~1 Mb genome-wide aCGH using bacterial artificial chromosome (BAC) probes. Automated hybridizations were performed for 72 h, followed by automated washing and drying as described before (46). Arrays were scanned with an Agilent DNA microarray scanner (Model G2505B, Serial number US22502518). Signal intensities were determined with ImaGene Scan Control software (Version 6.0.1) and raw data processing involved median pintip (*c.q.* sub-array) and average signal normalization; for performing unsupervised hierarchical clustering (complete linkage clustering using Pearson correlation as distance metric selection) log2 ratios derived from aCGH profiles were used in Matlab software (version 7.0.1, The Mathworks, Natick MA, USA) .

The microarray data have been deposited in NCBIs Gene Expression omnibus and are accessible through GEO Series (accession number GSE18626).

Statistical methods

Raw aCGH data consisted of log2 intensity ratios for each of the 3,248 BAC probes situated at an average spacing of 1 Mb across all 22 autosomes and chromosome X. These ratios were segmented with an algorithm developed by Picard *et al.* (47). Copy number alterations (CNAs) were interpreted as gains or losses and

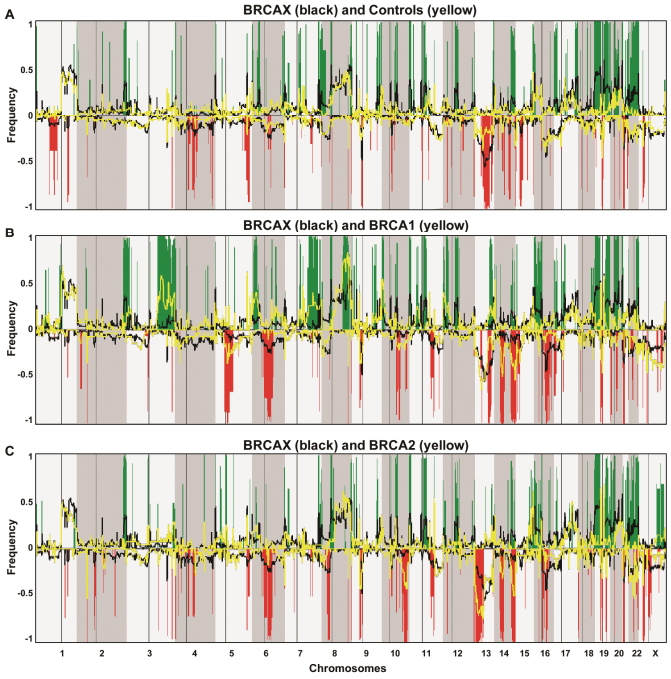


Figure 1 - Significant differential aberrations in BRCAx tumors. ***A)*** *The comparison between 58 BRCAx (black trace) and 48 control (yellow trace) tumor CGH profiles is shown. The x-axis represents all 3,248 probes on all 23 chromosomes (1 through X). Centromeres are indicated by vertical black lines. The frequencies of aberrations (|log2ratio|>0.2) are plotted on the y-axis. Vertical green bars correspond to significant differential gains as determined by a two-sided Fisher exact test. All depicted bars show p-values below 0.05; longer bars indicate higher significance, with a minimum p-value of 0.0001. Similarly, red bars indicate significance for differential losses.* ***B)*** *The comparison between BRCAx and BRCA1 tumors is shown.* ***C)*** *The comparison between BRCAx and BRCA2 tumors is shown.*

defined as segmented ratios greater than 0.2 or smaller than -0.2, respectively. Differential aCGH profiles between tumor classes (BRCAx, sporadic controls, BRCA1, and BRCA2) were defined by two-sided Fisher's exact testing.

In order to build a class predictor based on log2 ratios derived from aCGH experiments, we employed Prediction Analysis of Microarrays (PAM), which uses nearest shrunken centroids algorithm, as described by Tibshirany *et al.* (48). We used PAM 2.0 adapted to an Excel frontend, which is freely accessible from Stanford University (http://www-stat.stanford.edu/~tibs/PAM).

Results

Genomic aberrations of *BRCAx* breast tumors compared to sporadic controls, *BRCA1*- and *BRCA2*-mutated breast tumors

Array-CGH profiles were produced for 58 BRCAx breast tumors. All aCGH data were segmented as before (46, 47), and outlier frequencies were given as a fraction of each tumor class (58 BRCAx, 49 sporadic controls, 28 BRCA1, and 28 BRCA2 tumors) presenting a gain or loss for each of the 3,248 BAC probes measured (Fig. 1). BRCAx tumors were compared to the sporadic control (Fig. 1a), *BRCA1*-mutated (Fig. 1b), and *BRCA2*-mutated (Fig. 1c) breast tumors. The frequency of aberrations in the aCGH profiles of BRCAx tumors is depicted in black, whereas for the other class they are compared with the frequency depicted in yellow. Significant differential gains and losses (present at a higher frequency in one of the classes compared) are marked in each of the three

individual panels of Fig. 1 with green and red vertical bars, respectively. Chromosomal locations (including cytogenetic banding) for most significant class differences ($p < 0.01$, Fisher's exact test) between BRCAx and the other three tumor classes investigated are presented in Supplementary Table 1 (shaded boxes). While in Fig. 1 all significant differential regions are marked, the table only lists regions that consist of at least three adjacent BAC probes with the same pattern of significance.

Differential gains between BRCAx and sporadic control tumors (Fig. 1a) were prominent on chromosome arms 2q, 8p, 11p, 12q, and 19p, and along the entire chromosome 22. Differential losses were prominent on chromosome arms 1p, 4q, and 13q. Despite these differences, there was also an abundant overall similarity between aCGH profiles of BRCAx and sporadic breast tumors. For example, frequent aberrations observed in both classes were a gain of chromosome arms 1q and 8q, as well as a loss of chromosome arms 8p and 11q.

Differential gains between the aCGH profiles of BRCAx and *BRCA1*-mutated tumors (Fig. 1b) were prominent on chromosome arms 2q, 11q, 12q, and 19p, and differential losses were prominent on chromosome arms 6q and 16q. Some of the regions that were earlier identified with metaphase CGH (35) as being highly specific to *BRCA1*-mutated tumors are also visible, such as 3q (gains) and 5q (losses), as well as 3p (losses) which is less pronounced in our data set. Aberrations in these three regions were more frequent in *BRCA1*-mutated tumors than in BRCAx tumors.

Comparison of the frequency of aberrations between BRCAx and *BRCA2*-mutated tumors (Fig. 1c) also indicates significant differential regions. Differential gains include chromosome arms 2q, terminal regions of 7q, 11q, and 12q, as well as the entire chromosome 22. Differential losses are visible on chromosome arm 6q. In regions like 10q or 13p losses are more frequent in BRCA2 tumors.

To summarize, despite the fact that BRCAx tumors are heterogeneous in their aCGH profiles, they present 11 cytogenetic regions with recurrent gains (2q37, 7q36, 8p11, 10q22, 11p15-11q13, 12q13, 19p13, 19q13, 20p11, 22q13, and Xp22) and one cytogenetic region with specific recurrent losses (21q21), which differ from sporadic controls, *BRCA1*- and *BRCA2*-mutated breast tumors. These regions are indicated in Supplementary Table 1 (shaded boxes). There were few regions like 2q, 11p, 12q, 19p, and 19q, which appeared as gains in more than 30% of BRCAx tumors and were not found in the other three tumor classes.

BRCAx tumors have on average 30% (1063 vs. 816) more aberrant clones (gain or loss) than sporadic tumors. This difference was statistically significant in a two-sided, unpaired t-test ($p = 0.003$). We next investigated the genomic locations of the aberrations in BRCAx tumors, and aimed to determine whether any of the aberrations which appeared recurrent in this class were truly specific (different from controls, *BRCA1*- or *BRCA2*-mutated tumors).

BRCAx tumor profile heterogeneity

The heterogeneity among BRCAx tumors could not be seen in Fig. 1, as these graphs presented average data for the four tumor classes. Therefore, in addition to frequency plots, we visualized the genomic profiles of each chromosome per tumor per tumor class using heatmap diagrams (Fig. 2) after performing hierarchical clustering (Pearson complete correlation) of the samples. This allowed

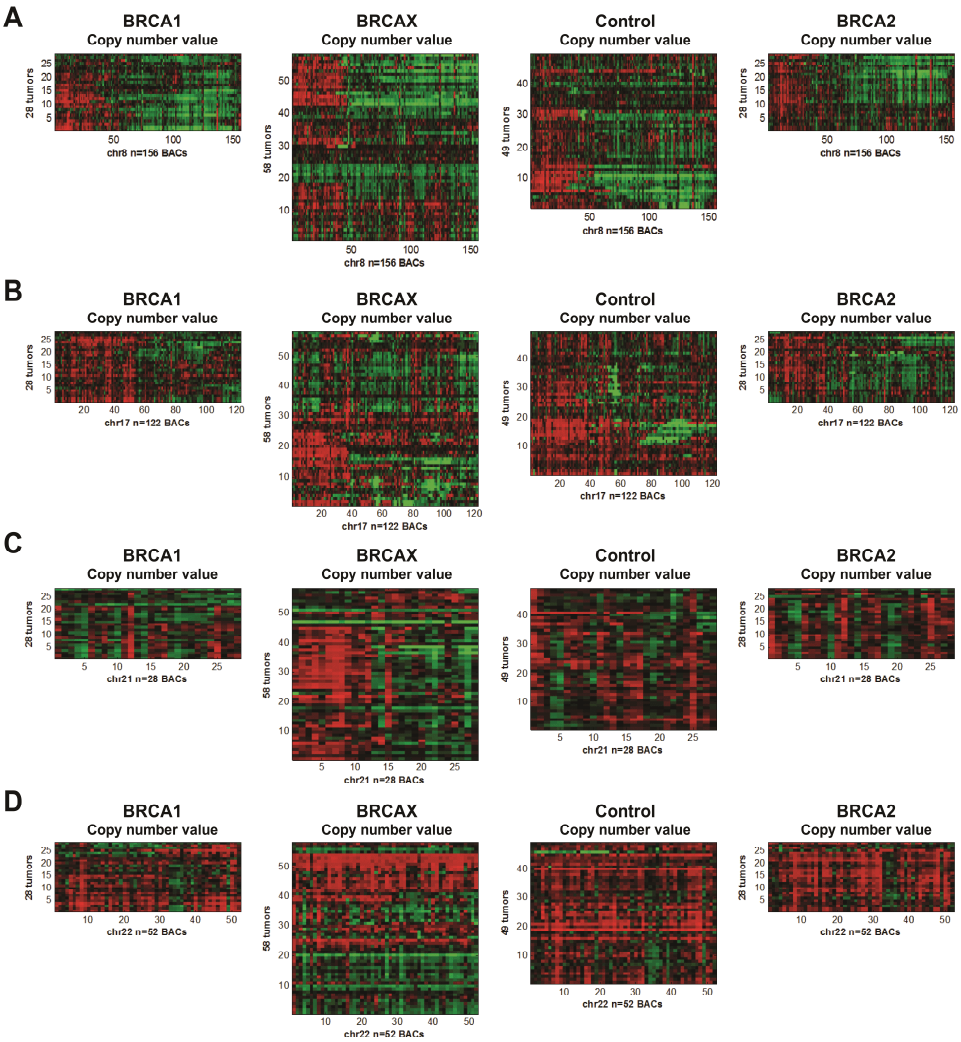


Figure 2. *Heatmap diagrams (derived from log2 ratios) for different chromosomes illustrating that BRCAx tumors are distinct from BRCA1, BRCA2, and sporadic control tumors. Diagrams for chromosomes 8, 14, 17, 21, and 22, are shown in panels a, b, c and d, respectively. Samples are sorted vertically. This vertical sorting is performed individually per panel (class), based on sample-to-sample complete Pearson correlation clustering of DNA copy number state (CGH level). Horizontal axes are not plotted to genomic scale but jump from BAC to BAC clone at the library density of approximately 1 Mb resolution. Color scales were set to saturate at -1 and 1 for relative DNA copy numbers (red–black–green). Gains are represented in green and losses in red.*

evaluation of within tumor class heterogeneity and direct genome-wide comparison of BRCAx tumors with sporadic control, *BRCA1*-, and *BRCA2*-mutated tumors. We used this approach to investigate all 22 autosomes and chromosome X. For the majority of chromosomes there were no major differences between the heatmap diagrams of BRCAx and the other tumor classes,

and there were obvious similarities between BRCAx sub-groups and at least one sub-group from the other tumor classes. Interestingly, for a number of chromosomes, the genomic profiles derived from BRCAx tumors showed distinct patterns in at least some tumor sub-groups, in contrast to all other three tumor classes. Four of the most prominent examples to illustrate these differences for chromosomes 8, 17, 21, and 22 are presented in Fig. 2a–d, respectively. Heatmap diagrams for all the other chromosomes are provided as supplementary material (Supplementary Figure 1).

For all tumor classes, chromosome 8 is characterized by frequent loss (red) of chromosome arm 8p and gain (green) of arm 8q (Fig. 2a). In contrast, a subset of BRCAx tumors showed no change, gain or loss of the entire chromosome (middle part of BRCAx panel).

For chromosome 17, there was a common pattern shared by all four classes, characterized by loss of chromosome arm 17p and gain of arm 17q. However, a large subset of BRCAx tumors displayed gain of the complete chromosome (Fig. 2b, upper half of the BRCAx panel), while a relatively small subset of BRCAx tumors presented loss of the entire chromosome (Fig. 2b, middle part of the BRCAx panel).

The majority of BRCAx tumors showed chromosome arm 21p loss and gain of arm 21q, but the reverse pattern can be discerned in a minor fraction as well (Fig. 2c, upper part of the BRCAx panel). All other tumor classes showed more scattered patches of gains and losses (Fig. 2c).

Chromosome 22 gain in a sub-group of BRCAx tumors

The profiles of BRCAx tumors for chromosome 22 were approximately equally distributed in two major classes: those with gain and those with loss of the entire chromosome. Nevertheless, a minority of tumors with more complex rearrangements were also present. While the class with loss of chromosome 22 is similar to the pattern observed in *BRCA1*-, *BRCA2*-mutated and sporadic control tumors, the class with chromosome 22 gain is specific for BRCAx tumors (Fig. 2d, bottom part of the BRCAx panel).

Given the very distinct pattern of chromosome 22, we further asked whether the BRCAx sub-groups identified based on chromosome 22 profiles were family related and whether tumors from the same family fall into the same subclass. It was our hypothesis that BRCAx subgroups that harbor a common predisposing gene might exist. In this case one would expect that the same genomic aberrations will be displayed by the tumor profiles of sub-group members and that family members would belong to the same sub-group. To test this hypothesis, we evaluated chromosome 22 profiles of all families for which we had three or more tumor cases analyzed. The most interesting were cases with a gain of chromosome 22, as that pattern seems specific for BRCAx tumors only.

Among the 27 families included in our BRCAx series, seven families contributed three or more aCGH profiles to our dataset, forming a total of 25 tumor profiles. When visualized by Unsupervised Hierarchical Clustering (Fig. 3), three of these families presented a gain of chromosome 22 in all tumors (RUL39 n = 4, RUL79 n = 5, RUL154 n = 3) and one family in two of the three tumors analyzed (RUL70 n = 3). From the remaining three families, one showed a loss of chromosome 22 in all tumors (RUL153 n = 4) and the other two families in two of the three tumors analyzed (RUL148 n = 3, RUL75 n = 3). Subsequently, the median ratio for all

BAC probes was calculated for tumors of the four 22-gain families (yellow trace in Fig. 4a), and compared to the median ratio of the remaining three BRCAx families with 22-loss (red trace in Fig. 4a) and to the median ratio of the sporadic cases (black trace in Fig. 4a). The 22-gain BRCAx subgroup had a median ratio significantly higher than that of the 22-loss BRCAx sub-group, which showed a median ratio similar to that of sporadic tumors, as expected from the heatmap diagrams (Fig. 2d).

We then considered the 22-gain group as a potential BRCAx sub-group, and aimed to determine whether other chromosome-specific patterns could also be identified. First, we selected all BAC probes that showed a large difference in median ratios between the 22-loss and the 22-gain groups. If we selected an (arbitrary) ratio difference ≥ 0.3, a total of n = 157 BAC probes were selected, which were distributed over several chromosomes (Fig. 4b). Interestingly, the sporadic group (black trace in Fig. 4b) showed a median ratio similar to that of the BRCAx 22-loss group (red trace in Fig. 4b) for nearly all BAC probes.

We further employed the nearest shrunken centroid algorithm (PAM analysis) (48) in order to build a class predictor, which could discriminate between the 22-loss and 22-gain groups. The classifier was built on ten breast tumors with 22-loss and 15 breast tumors with 22-gain (referred to as the training sets), including all 3,248 log2 measurements for all BAC probes on the array. A classifier consisting of 405 BAC probes was obtained, which was able to correctly assign 14/15 tumors from the 22-gain training set, and 9/10 tumors from the 22-loss training set (Supplementary Fig. 2). The first three chromosomes that contributed the most in the classifier were 13, 4, and 22 (see Supplementary Table 2), which contributed 12, 8, and 7% of the

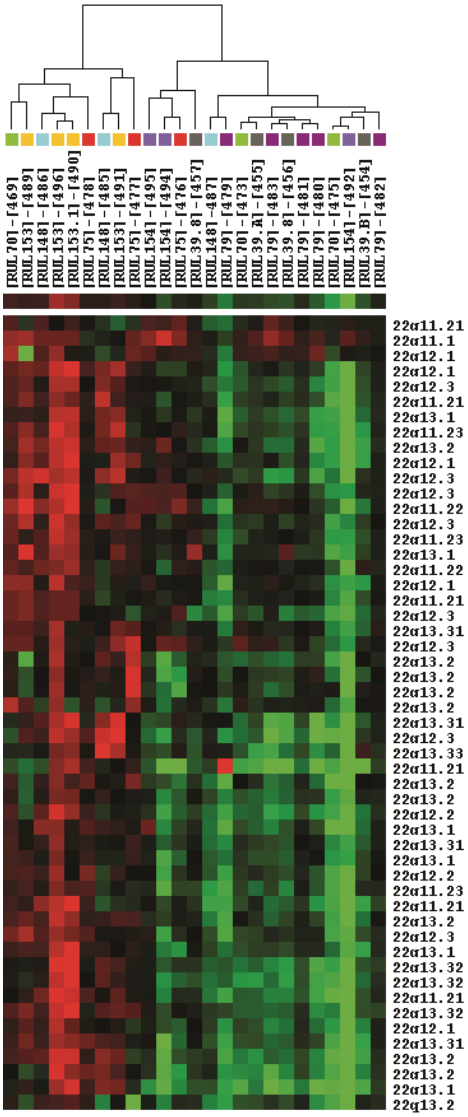


Figure 3. *Unsupervised Hierarchical Clustering for chromosome 22 of aCGH profiles derived from 25 tumors belonging to seven families from the BRCAx series. Each family contributed more than three aCGH profiles. Members of the same family are framed in rectangles of the same color. Color scales were set to saturate at -1 and 1 for relative DNA copy numbers (red–black–green). Gains are represented in green and losses in red.*

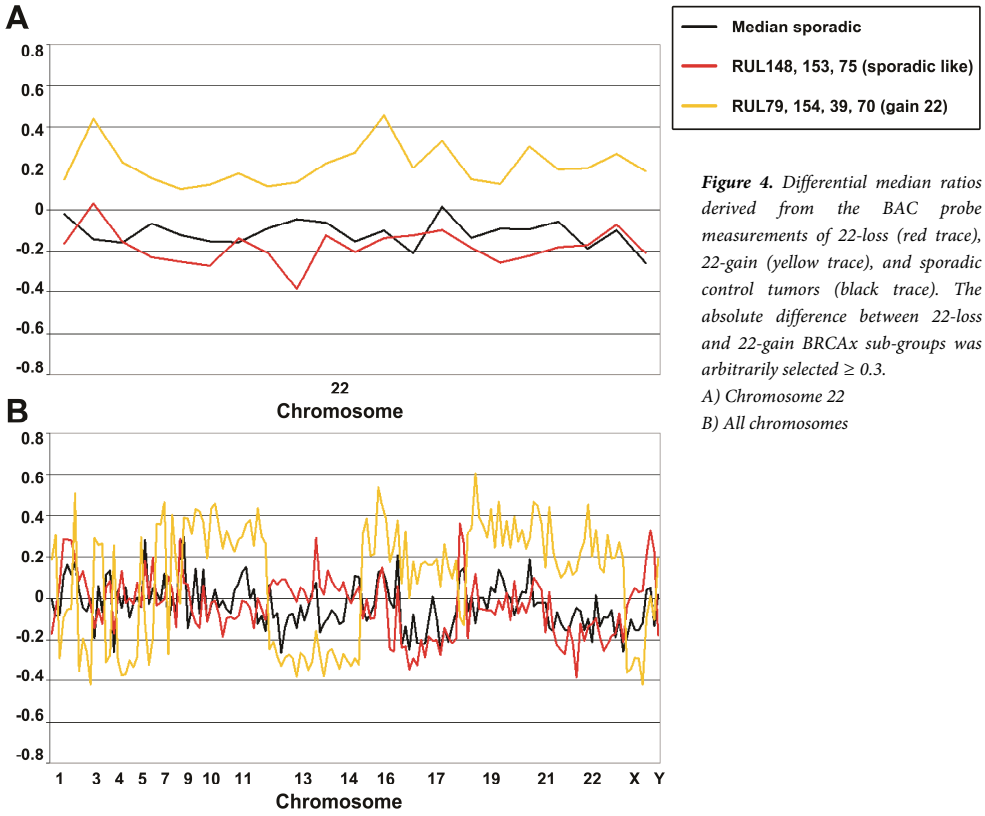


Figure 4. *Differential median ratios derived from the BAC probe measurements of 22-loss (red trace), 22-gain (yellow trace), and sporadic control tumors (black trace). The absolute difference between 22-loss and 22-gain BRCAx sub-groups was arbitrarily selected ≥ 0.3.*
A) Chromosome 22
B) All chromosomes

total number of BAC probes in the classifier, respectively. Chromosome 22 was represented by 27 BAC probes, which represented 52% of all the BAC probes of chromosome 22 on the array.

Among the remaining 33 BRCAx cases, 19 were classified as 22-gain-like (Supplementary Fig. 3). Not all cases showed the characteristic 22-gain; actually, 4/19 showed 22-loss, but all showed the characteristic loss of chromosome 13.

Most sporadic tumors showed a loss of chromosome 22 when visualized by Unsupervised Hierarchical Clustering (Fig. 2d), and resembled the BRCAx 22-loss group when the median ratios were calculated for this chromosome (Fig. 4a). These observations were also consistent with the results of the PAM classifier: 47/49 sporadic control tumors were classified as 22-loss-like (Supplementary Fig. 4).

Figure 5 presents the Unsupervised Hierarchical Clustering of the aCGH profiles derived from BRCAx and sporadic control tumors. All 3,248 log2 ratios for all 22 autosomes and chromosome X were used for the clustering. The dendrogram indicates that the 'within class' heterogeneities are comparable between sporadic and BRCAx tumors, and that there are neither major branch points nor obvious immunophenotypes across the four immunohistochemical markers ER, PR, ERBB2, and p53, suggesting an obvious stratification of BRCAx subphenotypes. On the other hand, it also seems clear that

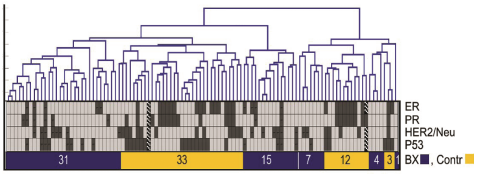


Figure 5. *Unsupervised Hierarchical Clustering of the aCGH profiles derived from BRCAx and control tumors. Clustering was performed in Matlab using log2 ratios for each of the 3248 probes from chromosome 1 to chromosome X. Vertical distances represent the similarity distance calculated acrossall 3,248 probes. BRCAx are labeled in blue, and sporadic tumors are labeled in yellow. Immunohistochemical staining scores are given as no staining (gray), positive staining (black), or missing data (hatched).*

BRCAx and sporadic control tumor profiles did not cluster at random, suggesting a possible role for an underlying class-specific biology. Both BRCAx and sporadic tumors clustered in nine sub-clusters, five of which (clusters 1, 2, 5, 6, and 9) contained only BRCAx tumors. Two other clusters (clusters 4 and 7) contained only sporadic tumors, another (cluster 3) contained 24 sporadic tumors and one BRCAx tumor, and in another cluster (cluster 8) three sporadic tumors were mixed with four BRCAx tumors.

The BRCAx tumors cluster in different branches: clusters 1 and 2 in a separate branch from cluster 5. PAM analysis was performed to determine which BACs makeup the differences between these main branches. In total, 207 BACs were selected on basis of the PAM analysis to be able to separate these two groups. These BACs are located on different chromosomes and do not seem to cluster in specific regions (Supplementary Figure 5).

Morphology and immunohistochemistry (IHC) of the 22-gain BRCAx subgroup

The majority (83%, 48/58) of the BRCAx tumors was invasive ductal carcinomas (IDC), and 16% (9/58) were lobular carcinoma (LC) or lobular carcinoma in situ (LCIS). From the 22-gain group, 73% (11/15) were IDC, which was not significantly different from the total group. Twenty-nine percent (17/58) of the BRCAx tumors were ER- and 16% (9/58) were triple negative (ER-/PR-/ERBB2-). In the 22-gain group, 20% of the BRCAx tumors were ER-, but none were triple negative. In total, 40% (6/15) of the 22-gain BRCAx tumors were p53+, but this fraction was lower in the remaining BRCAx cases (19%, 7/36).

Discussion

A number of linkage studies during the past 15 years have attempted to identify new high-risk breast cancer susceptibility gene(s), collectively called BRCAx. No major gene has been identified, apart from *PALB2* (8), *BRIP1* (7), and *RAD51C* (49), which are responsible for a small percentage of hereditary breast cancers. BRCAx families are primarily defined by the exclusion of a role for the known breast cancer susceptibility genes *BRCA1* and *BRCA2*, either by direct genetic testing, or through phenotypic (clinical) criteria such as the presence of ovarian cancer or male breast cancer cases. Shortly after the identification of *BRCA1* (2), it was noted that breast cancer families with a case of male breast cancer did not show linkage to the *BRCA1*-locus (50). A linkage search in these families readily detected a linkage signal on 13q12, the *BRCA2* locus (3). However, a feature that positively selects families that are due to an as yet unknown and rare high-risk gene is lacking. It is very likely, therefore, that a group of "non-BRCA1/2 breast cancer families" is attributable to multiple genetic defects. It is not known how many additional breast cancer susceptibility genes exist, how many will fall into the class of

rare high-risk (*i.e.*, BRCA1-like) or common low-risk susceptibility alleles, and if and how these factors interact with each other to cause susceptibility (polygenic model) (51). However, given that some of the larger linkage studies have failed to yield a significant linkage signal, it appears that even if a high-risk gene is found, it will not explain more than approximately 20% of the families. The power of linkage analysis would increase dramatically if we could select a sub-group of patients that would increase the proportion of families with this high-risk gene to over 50%. Herein, we explored whether aCGH profiles of breast tumors from patients of BRCAx families could resolve the heterogeneous genetic background of these patients. This was prompted by the fact that breast tumors from patients carrying *BRCA1* (41) and *BRCA2* (42) mutations displayed characteristic aCGH patterns.

In the present study, we used aCGH to analyze a series of 58 BRCAx tumors from selected families from a nationwide collection (30). These families had three or more members with breast cancer diagnosed before the age of 60, and no ovarian or male breast cancer. These criteria are therefore more stringent compared with earlier BRCAx reports, which did not exclude ovarian cancer (38, 39, 52), included samples with less than 60% tumor cells (39), and included families (6 of 18) with just two cases of breast cancer (53). Moreover, these studies were performed in patients from geographically distinct regions, which could increase the genetic heterogeneity, caused for example by local founder mutations. The above-mentioned differences may affect risk factor stratification and limit comparison with our study. We also included two or more tumors from the same family to evaluate the assumption that aCGH patterns of tumors from family members resemble each other more strongly than those from different families, since they probably share the same genetic predisposition.

As expected, we observed a high degree of heterogeneity among the 58 BRCAx aCGH profiles, in accordance with the failure to identify predisposing genes in previous linkage analysis in these families (25). We next examined family-specific recurrent aCGH aberrations in families contributing at least three available aCGH profiles.

We made two major observations: (1) BRCAx tumors show large heterogeneity in their genomic profiles, but overall they differ from both BRCA1/2 and sporadic tumors; (2) one distinct BRCAx sub-group (with chromosome 22-gain) can be identified on the basis of the genomic profiles. This sub-group (consisting of 4 families) is different from *BRCA1/2*-mutated and sporadic tumors, and shows an overrepresentation of tumors from the same family.

BRCAx tumor characteristic aberrations

Chromosome 22-gain among BRCAx tumors seems unique since we have not observed this high frequency gain in either sporadic, *BRCA1*- or *BRCA2*-mutated breast tumors (Fig. 2d). It will be interesting to determine whether 22-gain is a cause or consequence in a substantial proportion of BRCAx tumors. In *BRCA1*-mutated tumors, no loss of the *BRCA1* locus can be observed by aCGH in the majority of cases, although LOH analysis indicates that the wild-type allele is lost, probably due to mitotic recombination. Therefore, the predisposing gene does not necessarily need to be on chromosome 22.

Using a 1 Mb resolution BAC platform, Melchor *et al.* (54) compared 31 non-BRCA1/2

tumors with 19 *BRCA1*-, 24 *BRCA2*-mutated, and 19 sporadic breast tumors by aCGH. They distinguished common aberrations across all breast cancer groups (such as gains of 1q and 16p, and losses of 8p and 16q), which are concordant with our results. In addition, they report a set of altered regions that discriminate between tumor classes, but are not specific for only one tumor class. The most frequent alterations per tumor class are also visible in our frequency plots (*e.g.*, loss of chromosome arm 5q in *BRCA1*-mutated tumors or loss of 13q in *BRCA2*-mutated tumors). In contrast to our results, in their study population *BRCA1/2*-mutated tumors showed a higher genomic instability than BRCAx and sporadic cases. It is conceivable that the differences are partially attributable to their different tumor selection criteria, ethnicity, or both.

In a recent SNP-based linkage study, Rosa-Rosa *et al.* (55) found a single family (FAM153) with suggestive linkage signals in two different chromosomes (11q13 and 14q21). Despite the fact that mutational screening on candidate genes in 11q13 did not indicate any deleterious mutations, data from microsatellite markers (STR) confirmed 11q13 as a candidate region to contain a breast cancer susceptibility gene. This family is also included in the present study with four tumors analyzed, all classified in the 22-loss sub-group. We identified the 11q13 region as being specifically altered in BRCAx tumors, as it contains differential gains in BRCAx tumors compared to other tumor classes.

BRCAx tumor heterogeneity

An elusive but crucial aspect of BRCAx families is whether or not different risk factors were co-selected for by the clinical criteria used. Besides *BRCA1* and *BRCA2*, one other breast cancer gene variant was tested in these familial tumors, namely, CHEK2-1100delC. We observed both intra- and inter-familial heterogeneity for CHEK2 carrier status in BRCAx families and concluded that CHEK2 is not the determining factor in these cancers. Among the 15 individuals belonging to the 22-gain sub-group, only one was carrying the CHEK2-1100delC mutation.

The Unsupervised Hierarchical Clustering of the aCGH profiles derived from BRCAx and sporadic control tumors (Fig. 5) indicates comparable "within class" heterogeneities between sporadic and BRCAx tumors, without stratification of BRCAx aCGH profiles across the four immunohistochemical markers analyzed (ER, PR, ERBB2, and p53). On the other hand, tumors from both classes do not cluster at random, supporting the idea to further stratify BRCAx tumors based on "aCGH phenotypes". While our results are promising, the number of tumors analyzed is still relatively small. Accordingly, we could identify only four families with tumors that had the 22-gain aCGH-pattern. Given the limited linkage information of each of these families, this number was too small to perform meaningful linkage analyses, even under homogeneity (data not shown). We are presently validating the 22-gain aCGH-pattern in an independent set of cases, concomitantly collecting DNA for linkage analysis, with the aim of identifying additional families with this characteristic.

Acknowledgments

This study was supported by the Dutch Cancer Society/Koningin Wilhelmina Fonds, Grant NKB_NKI2007-3749.

Conflict of interest

The authors declare to have no competing interests.

References

1. World Health Organization. The global burden of disease: 2004 update. WHO Press; 2008. ISBN 978 92 4 156371 0
2. Miki Y, et al. 1994. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. Science 266:66-71.
3. Wooster R, et al. 1994. Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13. Science 265:2088-90.
4. Wooster R, et al. 1995. Identification of the breast cancer susceptibility gene BRCA2. Nature 378:789-92.
5. Meijers-Heijboer H, et al. 2002. Low-penetrance susceptibility to breast cancer due to CHEK2(*)1100delC in noncarriers of BRCA1 or BRCA2 mutations. Nat Genet 31:55-9.
6. Thompson D, et al. 2005. Cancer risks and mortality in heterozygous ATM mutation carriers. J Natl Cancer Inst 97:813-22.
7. Seal S, et al. 2006. Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles. Nat Genet 38:1239-41.
8. Rahman N, et al. 2007. PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. Nat Genet 39:165-7.
9. Cox A, et al. 2007. A common coding variant in CASP8 is associated with breast cancer risk. Nat Genet 39:352-8.
10. Easton DF, et al. 2007. Genome-wide association study identifies novel breast cancer susceptibility loci. Nature 447:1087-93.
11. Stacey SN, et al. 2008. Common variants on chromosome 5p12 confer susceptibility to estrogen receptor-positive breast cancer. Nat Genet 40:703-6.
12. Chen J, et al. 1998. A study of the PTEN/MMAC1 gene in 136 breast cancer families. Hum Genet 102:124-5.
13. Leggett BA, et al. 2003. Peutz-Jeghers syndrome: genetic screening. Expert Rev Anticancer Ther 3:518-24.
14. Pharoah PD, et al. 2001. Incidence of gastric cancer and breast cancer in CDH1 (E-cadherin) mutation carriers from hereditary diffuse gastric cancer families. Gastroenterology 121:1348-53.
15. Borresen AL, et al. 1992. Screening for germ line TP53 mutations in breast cancer patients. Cancer Res 52:3234-6.
16. Evans DG, et al. 2002. Low rate of TP53 germline mutations in breast cancer/sarcoma families not fulfilling classical criteria for Li-Fraumeni syndrome. J Med Genet 39:941-4.
17. Lalloo F, et al. 2006. BRCA1, BRCA2 and TP53 mutations in very early-onset breast cancer with associated risks to relatives. Eur J Cancer 42:1143-50.
18. Malkin D, et al. 1990. Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. Science 250:1233-8.
19. Turnbull C, et al. 2008. Genetic predisposition to breast cancer: past, present, and future. Annu Rev Genomics Hum Genet 9:321-45.
20. Walsh T, et al. 2007. Ten genes for inherited breast cancer. Cancer Cell 11:103-5.

21. Narod SA, et al. 2004. BRCA1 and BRCA2: 1994 and beyond. Nat Rev Cancer 4:665-76.
22. Kainu T, et al. 2000. Somatic deletions in hereditary breast cancers implicate 13q21 as a putative novel breast cancer susceptibility locus. Proc Natl Acad Sci U S A 97:9603-8.
23. Huusko P, et al. 2004. Genome-wide scanning for linkage in Finnish breast cancer families. Eur J Hum Genet 12:98-104.
24. Thompson D, et al. 2002. Evaluation of linkage of breast cancer to the putative BRCA3 locus on chromosome 13q21 in 128 multiple case families from the Breast Cancer Linkage Consortium. Proc Natl Acad Sci U S A 99:827-31.
25. Smith P, et al. 2006. A genome wide linkage search for breast cancer susceptibility genes. Genes Chromosomes Cancer 45:646-55.
26. Burwinkel B, et al. 2006. Transcription factor 7-like 2 (TCF7L2) variant is associated with familial breast cancer risk: a case-control study. BMC Cancer 6:268.
27. Pharoah PD, et al. 2007. Association between common variation in 120 candidate genes and breast cancer risk. PLoS Genet 3:e42.
28. Ponder BA, et al. 2005. Polygenic inherited predisposition to breast cancer. Cold Spring Harb Symp Quant Biol 70:35-41.
29. Risch N, et al. 1996. The future of genetic studies of complex human diseases. Science 273:1516-7.
30. Oldenburg RA, et al. 2008. Genome-wide linkage scan in Dutch hereditary non-BRCA1/2 breast cancer families identifies 9q21-22 as a putative breast cancer susceptibility locus. Genes Chromosomes Cancer 47:947-56.
31. Cui J, et al. 2001. After BRCA1 and BRCA2-what next? Multifactorial segregation analyses of three-generation, population-based Australian families affected by female breast cancer. Am J Hum Genet 68:420-31.
32. Antoniou AC, et al. 2002. A comprehensive model for familial breast cancer incorporating BRCA1, BRCA2 and other genes. Br J Cancer 86:76-83.
33. Pinkel D, et al. 1998. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. Nat Genet 20:207-11.
34. Curtis C, et al. 2009. The pitfalls of platform comparison: DNA copy number array technologies assessed. BMC Genomics 10:588.
35. Wessels LF, et al. 2002. Molecular classification of breast carcinomas by comparative genomic hybridization: a specific somatic genetic profile for BRCA1 tumors. Cancer Res 62:7110-7.
36. van Beers EH, et al. 2005. Comparative genomic hybridization profiles in human BRCA1 and BRCA2 breast tumors highlight differential sets of genomic aberrations. Cancer Res 65:822-7.
37. Jonsson G, et al. 2005. Distinct genomic profiles in hereditary breast tumors identified by array-based comparative genomic hybridization. Cancer Res 65:7612-21.
38. Alvarez S, et al. 2005. A predictor based on the somatic genomic changes of the BRCA1/BRCA2 breast cancer tumors identifies the non-BRCA1/BRCA2 tumors with BRCA1 promoter hypermethylation. Clin Cancer Res 11:1146-53.
39. Gronwald J, et al. 2005. Comparison of genomic abnormalities between BRCAX and sporadic breast cancers studied by comparative genomic hybridization. Int J Cancer 114:230-6.

40. Hu X, et al. 2009. Genetic alterations and oncogenic pathways associated with breast cancer subtypes. Mol Cancer Res 7:511-22.
41. Joosse SA, et al. 2009. Prediction of BRCA1-association in hereditary non-BRCA1/2 breast carcinomas with array-CGH. Breast Cancer Res Treat 116:479-89.
42. Joosse SA, et al. 2010. Prediction of BRCA2-association in hereditary breast carcinomas using array-CGH. Breast Cancer Res Treat .
43. van Beers EH, et al. 2006. A multiplex PCR predictor for aCGH success of FFPE samples. Br J Cancer 94:333-7.
44. Oldenburg RA, et al. 2006. Characterization of familial non-BRCA1/2 breast tumors by loss of heterozygosity and immunophenotyping. Clin Cancer Res 12:1693-700.
45. Oldenburg RA, et al. 2003. The CHEK2*1100delC variant acts as a breast cancer risk modifier in non-BRCA1/BRCA2 multiple-case families. Cancer Res 63:8153-7.
46. Joosse SA, et al. 2007. Automated array-CGH optimized for archival formalin-fixed, paraffin-embedded tumor material. BMC Cancer 7:43.
47. Picard F, et al. 2005. A statistical approach for array CGH data analysis. BMC Bioinformatics 6:27.
48. Tibshirani R, et al. 2002. Diagnosis of multiple cancer types by shrunken centroids of gene expression. Proc Natl Acad Sci U S A 99:6567-72.
49. Meindl A, et al. 2010. Germline mutations in breast and ovarian cancer pedigrees establish RAD51C as a human cancer susceptibility gene. Nat Genet 42:410-4.
50. Stratton MR, et al. 1994. Familial male breast cancer is not linked to the BRCA1 locus on chromosome 17q. Nat Genet 7:103-7.
51. Oldenburg RA, et al. 2007. Genetic susceptibility for breast cancer: how many more genes to be found? Crit Rev Oncol Hematol 63:125-49.
52. Rouleau E, et al. 2007. High-resolution oligonucleotide array-CGH applied to the detection and characterization of large rearrangements in the hereditary breast cancer gene BRCA1. Clin Genet 72:199-207.
53. Tirkkonen M, et al. 1999. Somatic genetic alterations in BRCA2-associated and sporadic male breast cancer. Genes Chromosomes Cancer 24:56-61.
54. Melchor L, et al. 2007. Estrogen receptor status could modulate the genomic pattern in familial and sporadic breast cancer. Clin Cancer Res 13:7305-13.
55. Rosa-Rosa JM, et al. 2009. A 7 Mb region within 11q13 may contain a high penetrance gene for breast cancer. Breast Cancer Res Treat 118:151-9.

Chapter 8

Discussion

Discussion

Materials and methods

Sample material: formalin-fixed, paraffin-embedded

In most pathology laboratories worldwide, formalin-fixed, paraffin-embedded (FFPE) samples are the only tissue specimens available for routine diagnostics. To have access to a vast amount of material, and to be able to implement our results in routine diagnostics in the future, we have limited ourselves to the use of FFPE material in the studies described in this thesis. However, it was the experience in our research laboratory from earlier studies that DNA extracted from FFPE tissue is not always suitable for array CGH analysis. Time and money could be used more efficiently if the sample quality can be assessed prior to array CGH experiments. This has motivated us to develop a multiplex PCR to determine the maximum length of DNA able to be amplified as a measure of DNA quality. Short DNA fragments due to degradation, or DNA strands that are unable to form single strands due to crosslinks, may cause unspecific binding and thus hamper hybridization efficiency which results in unreliable array data. **Chapter 2** describes the development of a multiplex PCR that is able to test the possibility of producing DNA fragments of 100, 200, 300, or 400 bp long. We postulated that the longer the fragments are that can be produced, the better the sample quality is (1). For this test, we have chosen the gene *GAPDH* as genomic target because of its importance in cancer cell survival. GAPDH plays a central role in glycolysis-dependent energy supply, and because cancer cells metabolize glucose mainly through the glycolytic pathway and depend far less on oxidative phosphorylation, tumor cells are highly dependent on GAPDH for survival and proliferation. Therefore, it is very unlikely that *GAPDH* will be lost in breast tumors, making it a reliable target for investigation. Indeed, we have not observed a successful aCGH experiment using DNA from which no PCR product could be produced *post hoc*.

The multiplex PCR has been very helpful during our studies, especially when investigating breast cancer families (see for example **Chapter 7**). Registration of families carrying a germline predisposition for breast cancer already covers several generations on many occasions in our hospital. In such cases where tumor tissue has been archived, the material can be very old and the DNA degredated beyond usage. Here, our multiplex PCR has given us insight in which samples could still be investigated by array CGH. In the past decades, fixation time, usage of buffered formalin, and storage conditions have been improved and standardized. We have seen in our studies that this has led to better conservation of sample material as DNA isolated from newer material (after 1990) is of better quality compared to older material (before 1990) in our hospital. Also other laboratories have been using our DNA quality test (2-5). And although array CGH technology has improved much in the latest couple of years to even combined SNP genotyping and copy number alteration analysis, sample quality is hard to control, which underscores selection of samples prior to analysis for even these more advanced microarray platforms (4, 5).

Array CGH: hybridization

In cancer research, array CGH has become a valuable tool for the detection of chromosome copy number alterations. Challenges of this technique were to obtain high signal-to-noise ratios to detect single copy number changes, and to obtain highly reproducible results with a minimal variance for FFPE material. In the past, we performed the hybridization procedure and washing of the arrays manually, which often led to capricious results. **Chapter 3** describes the standardization of an array CGH protocol for FFPE material, making use of a hybridization station. Deployment of a hybridization station allowed us to hybridize with a more viscose hybridization mixture, because we were no longer limited in the force of mixing by a rocking table as with the manual procedure described in **chapter 2**. A higher viscosity facilitated by doubling dextran sulphate concentration increased the effective concentration of labeled DNA and thereby increased the signal-intensities. Additionally, automation of the hybridization and washing steps of the arrays resulted in highly reproducible results and less overall variance in the CGH profiles. Another major improvement in our protocol was the elimination of formamide, which is commonly used in large amounts in most microarray wash protocols but is very toxic. Last, although DNA from FFPE tissue is generally more problematic than fresh material, our automated protocol has proven to be very successful in the use of FFPE material (6).

Further improvements in the technology could be made. In an unpublished study, we have looked at the dispersion of labeled DNA over the microarray during hybridization in a hybridization station. For this, we have injected hybridization buffer into the hybridization chamber with only the last few microliters containing labeled DNA. After hybridization and scanning, labeled DNA was found to have hybridized over the first half of the slide only (Figure 1). These results suggested that not every DNA fragment will meet its target and signal-intensities can theoretically be increased even more by applying a different mixing mechanism. The hybridization station described in this thesis uses agitating membranes, pushing the hybridization buffer back and forward, to facilitate mixing. Improvement in hybridization efficiency can be made by applying a continuous flow of the hybridization buffer over the whole slide. Nonetheless, our protocol has proven to be very efficient as it has been applied in many successful studies in our hospital as in studies by others [this thesis and *e.g.*, (7-15)].

Array CGH: the platform

When performing CGH studies, one should be aware that this technique is unable to detect copy-neutral abnormalities in the genome. Only by using more complex techniques (*e.g.*, paired-end mapping and next generation sequencing), DNA translocations, inversions and loss of heterozygosity (LOH) can be visualized on a genome wide scale in high throughput fashion (16, 17). Compared to the platform described in this thesis, other techniques and platforms with higher resolution plus the ability of genotyping (*i.e.*, SNP arrays) were available at the time we performed our experiments but were proven not to be suited for FFPE tissue. Advances in the field of high-resolution copy number analysis now allow for the use of FFPE material, but were made after the initiation of our studies (18). SNP arrays make use of single nucleotide polymorphisms (SNP) that allow for the detection of copy numbers and LOH.

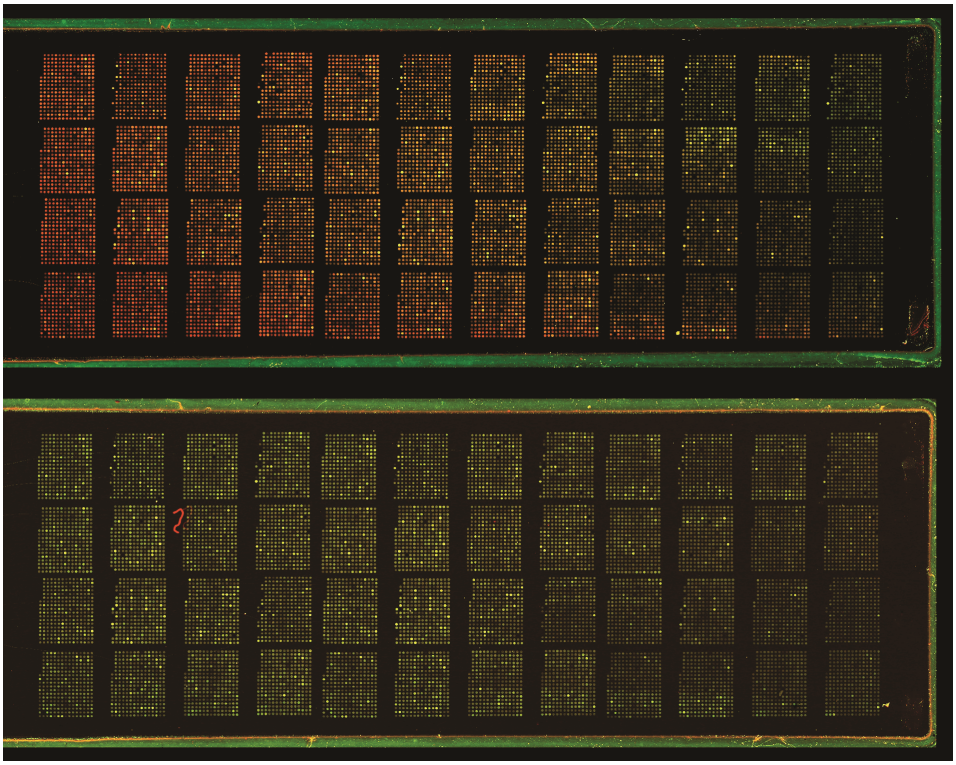


Figure 1 *- Scans of two microarrays, displaying the dispersion of labeled material during the hybridization procedure. Material was injected into the hybridization chamber on the left side.*

The BAC (bacterial artificial chromosome) array platform we used had several advantages over the other existing platforms. First, our arrays were printed "in-house", which was an enormous financial advantage. Second, the probes called BAC clones cover several hundreds of kbp each, allowing the binding of sufficient labeled DNA to produce clear signals for detection (high signal-to-noise ratio). Lastly, every probe was printed three times on a single array, thereby increasing the accuracy of every genomic measurement. The increased signal and triplicate measurements were required to produce very reliable results from FFPE material, which is a difficult source of DNA. Still, the relatively low resolution of the BAC array compared to other CGH platforms and the difficulty in array CGH to detect low copy number changes raised the question whether small, but also low, copy number aberrations are correctly called by the BAC array platform. To test this, we investigated whether we could properly detect a 2-fold gain of just 600 kbp (*ESR1* amplification) in a related study (19). For this, an *in silico* experiment was performed using array CGH data generated from a subject with four X chromosomes versus reference DNA with two X chromosomes (4X/2X) and array CGH data generated from reference versus reference DNA (2X/2X). Next, log2ratio values from

random segments of the X chromosomes, corresponding to the same number of probes spanning the 600 kbp region, were swapped from the 4X/2X data into the 2X/2X data. By doing so, we modeled a 2-fold gain of 600 kbp in the 2X/2X data. This procedure was repeated 100 times and the fraction that a gain was correctly called was calculated. We could identify the segment as gain in 71 out of 100 times, showing that the BAC array can detect 2-fold gains of small regions most of the time. Nonetheless, the genomic coverage of the BAC array is low and small aberrations between BAC arrays will be missed. Whether hereditary breast cancer shows such small aberrations and the detection of them would be of prognostic value are still unknown and have to be tested.

Other molecular techniques

A widely used microarray technique to study the behavior of cancer cells is gene expression profiling (GEP) (20). This technique has been proven to be very useful in the classification of molecular breast cancer subtypes (21) and prognostic signatures (22-24). However, the biggest disadvantage of this approach to study cancer is the requirement of messenger RNA, which is an unstable molecule and quickly degrades. Freshly frozen tumor material is needed to extract intact mRNA, which is not available in such large amounts as FFPE material. Although mRNA can be extracted from selected FFPE samples, the overall success rate of obtaining informative mRNA is much lower as for DNA (25). Because the studies in this thesis describe the investigations of hereditary breast cancer covering several generations of patients from breast cancer families, obtaining informative mRNA of a statistically sufficient number of samples would be extremely difficult. Therefore, DNA is a very suitable platform to study in our situation. Although DNA copy numbers carry a different form of information compared to mRNA gene expression patterns, it has been shown for classification problems that breast cancer can also be subtyped with array CGH as well as with GEP (26, 27).

Currently, the most elegant technique to determine chromosomal aberrations and simultaneously detect mutations and loss of heterozygosity (LOH) on a genome wide scale is next generation sequencing (NGS). It outperforms array CGH in resolution to the single nucleotide and is not limited by predefined probes. Compared with array CGH, the disadvantages of NGS are still the high costs and the computer power it requires to map sequence reads and calculate copy number variations. For diagnostic purposes, NGS will not replace CGH where such detailed knowledge about the complete genome is not required. For example, in pre-implantation genetic diagnostics only specific genomic aberrations are being investigated like monosomy for chromosome 18 and 21 or trisomy for chromosome 22. I anticipate that also for the diagnostic application for which we have used array CGH, finally not NGS but array CGH or a PCR based technique will still be used. For exploratory studies, NGS will most likely develop to become the standard of investigating chromosomal aberrations.

BRCA status prediction in breast cancer

BRCA1 status prediction in hereditary breast cancer

Routine diagnostics uses a diversity of techniques to identify as many *BRCA1* muta-

tions as possible in women in whom the suspicion of breast cancer susceptibility is high. Still, novel defects are being found and it is not known to what extent *BRCA1*-associated breast cancer patients are missed with current diagnostics (see **Chapter 1**). Therefore, we developed an array CGH based test that can discriminate between *BRCA1*-associated and sporadic breast tumors with high accuracy (28). **Chapter 4** describes the construction and validation of an array CGH based BRCA1 classifier. Using this classifier, we tested a group of 48 non-BRCA1/2 patients of whom two cases were found to be BRCA1-like. Additional analysis did not reveal a novel *BRCA1* mutation within theses cases; however, we could identify LOH of the *BRCA1* locus in both tumors (suggesting possible loss of function) and hypermethylation of the promoter of *BRCA1* in one case. Additional analysis of the latter patient also revealed *BRCA1* promoter methylation in the patent's ovarian tumor. Although recent studies have shown that hypermethylation can be inherited (29), it is questionable whether hypermethylation in this case is a germline epimutation. Although we did not investigate mRNA transcription, DNA extracted from blood did not show methylation of the *BRCA1* promoter.

Several points need to be addressed regarding the methodology of this study. First of all, the control group was a group of randomly picked sporadic tumors, only stratified on p53 status by immunohistochemistry. It has been proposed that, similar to a previous study by our group (30), breast tumors from non-BRCA1/2 patients should be used as a control group, because that will be the group on which the classifier will be employed most. However, this latter study reported on many false positive cases in their validation group that prove to be real *BRCA1*-mutated cases later on. Therefore, such control group could contain many unidentified *BRCA1*-mutated cases, weakening the classifier's power. Another possibility as control group would have been breast tumors, matched for hormone receptors, age and grade. Such a group would mainly consist of triple-negative (basal-like) breast tumors and would most likely result in a poor classifier, because approximately half of the basal-like breast tumors do not express *BRCA1* and harbor a CGH profile that is similar to *BRCA1*-mutated breast tumors (**Chapter 6**). Therefore, training of a BRCA1 classifier using basal-like tumors as control group requires at least twice as many cases as initially calculated to achieve similar power. Nevertheless, we have shown that by enriching for IHC status and comparing with our initial results, the BRCA1-specific genomic aberrations were not correlated to IHC status (**Chapter 4**).

Second, instead of using fixed thresholds for the identification of gains and losses, it would have been more appropriate to have used CGH profile and cell percentage dependent thresholds to compensate for technical variation as described elsewhere (35). The fixed thresholds were too stringent and lowered the frequency of all aberrations. Although re-analyses of the *BRCA1*-mutated breast tumors using variable thresholds showed higher frequencies of the aberrations, no differences were observed in the results in respect to the comparison of *BRCA1*-mutated and sporadic breast tumors (36, 37). Compared to other studies, many of the prominent aberrations in *BRCA1*-mutated breast tumors have also been identified by others (Table 1). Aberrations commonly identified to be specific in *BRCA1*-mutated breast cancer are gain on chromosome 3q and 10p and loss on chromosome 5q. Differences between studies can be explained by the different techniques used to detect and call aberrations and the

Table 1 *- Prominent chromosomal aberrations in BRCA1-mutated breast tumors from five different studies (28, 31-34), summarized per chromosome arm (Chr. arm).*

Chr. arm	Tirkkonen (n=21)	Van Beers (n=36)	Jönsson (n=14)	Stefansson (n=11)	Joosse (n=34)
2q	2q loss (40%)				
3p		3p gain (33%)			
3q		3q gain (67%)	3q gain (>75%)		3q gain
4p	4p loss (64%)		4p loss (>75%)	4p loss	
4q	4q loss (81%)			4q loss	
5p					5p loss
5q	5q loss (86%)	5q loss (72%)	5q loss (>75%)	5q loss	5q loss
6p	6p gain (40%)				6p gain
7p					7p loss
7q			7q gain (>75%)		7q gain
8p		*8p no gain*			
9p		9p gain (33%)			
10p	10p gain (30%)	10p gain (50%)	10p gain (>75%)	10p gain	10p gain
12p					12p gain
12q	12q loss (40%)				12q loss
13q	13q loss (55%)	13q gain (25%)			
15q			15q loss (>75%)		
16p		*16p no gain*			
16q				16q gain	*16q no loss*
17p			17p loss (>75%)		
17q	17q gain (45%)				
18p		18p gain (28%)			
20q					20q loss
Xp				Xp loss	
Xq				Xq loss	

choice of control samples to which the *BRCA1*-mutated cases were compared.

Although the classifier requires validation in a larger cohort, its usefulness was demonstrated in several subsequent studies. Van den Ouweland and colleagues showed that a novel rearrangement in *BRCA1*, deletion of exon 1a-2, was a pathogenic *BRCA1* mutation (38). In this study, breast tumors from several families were investigated by array CGH and tested for similarity to our *BRCA1* specific CGH profile. Tumors from exon 1a-2 deletion carriers were classified as being associated with deficient *BRCA1*, whereas one tumor from a family member that was diagnosed not to carry the deletion was classified as sporadic breast cancer. This study shows that deletion of exon 1a-2 in the *BRCA1* gene is not uncommon in the Dutch population, indicating that techniques such as MLPA are necessary to detect these mutations.

In a study by Tischkowitz and colleagues, the pathogenicity of the missense variant M1775K was assessed by a combination of functional, crystallographic, biophysical, molecular and evolutionary techniques, and classical genetic segregation analysis (39). These techniques included array CGH profiling of breast tumors from carriers of the M1775K variant. Our BRCA1 classifier classified the breast tumors from M1755K carriers as BRCA1-like and thus provided another line of evidence of the pathogenic effect of the missense variant. Furthermore, in one of our following publications we studied the CGH profile of the breast tumor from a subject carrying the intronic BRCA1 UV c.81-9-C>G (36). The tumor was classified as *BRCA1*-associated and as further analyses indicated, this variant results in a frame shift, thus having a pathogenic effect.

The above described results demonstrate that the BRCA1 classifier can be helpful to clinical diagnostics in providing indications of *BRCA1*-association in non-BRCA1/2 breast cancer patients, but can also aid the diagnosis of unclassified variants in providing additional indications for the pathogenicity of the variant.

BRCA2 status prediction in hereditary breast cancer

Similar to *BRCA1* mutation screening, it is unknown how many *BRCA2*-associated breast cancer patients are missed in current routine diagnostics. Furthermore, identification of unclassified variants of *BRCA2* can complicate clinical management. An additional tool, indicating the involvement of mutated *BRCA2* in cancer formation could therefore be very helpful in hereditary breast cancer diagnostics. Inspired by the success of detecting a CGH profile specific for *BRCA1*-mutated breast tumors, we investigated the chromosomal aberrations of *BRCA2*-mutated breast tumors. **Chapter 5** describes the identification of chromosomal aberrations specific for *BRCA2*-mutated breast tumors when compared to sporadic and *BRCA1*-mutated breast tumors. Table 2 describes the most prominent aberrations present in *BRCA2*-mutated breast tumors found in our and four other studies. Commonly identified were loss on chromosome 13q and gain on chromosome 17q, but also many discrepancies are present between the different studies. Based on the array CGH data, we developed and validated an array CGH BRCA2 classifier using breast tumors from proven *BRCA2*-mutation carriers and sporadic breast tumors from women without any family history for cancer. By testing a large group of 89 breast cases from non-BRCA1/2 patients, we could identify 17 cases to have a BRCA2-like CGH profile (36). Additional evidence for pathogenic mutations and non-functional BRCA2 protein was found in several cases; however, many of the BRCA2-like samples remained unexplained. During validation, we observed that the BRCA2 classifier is not very accurate in discriminating between *BRCA2*-mutated and sporadic breast tumors; this could be due to several reasons. First, the control group of the BRCA2 classifier might contain *BRCA2*-associated breast tumors and thereby diluting the aberrations of interest; however, this is highly unlikely given the prevalence of *BRCA2*-mutations in the population (see **Chapter 1**). Second, *BRCA2*-associated breast tumors might exhibit a similar pattern of chromosomal aberrations as sporadic breast cancer. Dysfunctional *BRCA2* might not lead to unique chromosomal aberrations or *BRCA2* might also become dysfunctional in sporadic breast cancer leading to similar aberrations. Third, technical limitations might be responsible

***Table 2** - Prominent chromosomal aberrations in BRCA2-mutated breast tumors from five different studies (31-34, 36), summarized per chromosome arm (Chr. arm).*

Chr. arm	Tirkkonen (n=15)	Van Beers (n=25)	Jönsson (n=12)	Stefansson (n=18)	Joosse (n=47)
1p	1p loss (45%)			1p loss	
3p	3p loss (55%)			3p loss/gain	
3q		3q gain (56%)			
6q	6q loss (60%)			6q loss	
8p				8p loss	
8q				8q gain	
11q	11q loss (65%)		11q loss (>75%)	11q loss	
13q	13q loss (73%)			13q loss	13q loss (78%)
14q				14q loss	14q loss (62%)
16p					*16p no gain*
16q	*16q no loss*			16q loss	*16q no loss*
17p				17p loss	
17q	17q gain (85%)		17q gain (>75%)	17q gain	17q gain (36%)
20q	20q gain (60%)		20q gain (>75%)		
Xp				Xp loss	

for not identifying BRCA2 specific aberrations, *i.e.*, the microarray resolution was too low or the studied cohorts were too small. Last and probably the most likely reason, because *BRCA2*-related breast tumors are also pathologically a heterogeneous group (40), the diversity among the CGH profiles might be biologically driven. It is unclear at which time point a BRCA2 deficient cell transforms to a tumor cell and what its cell of origin is. The heterogeneity among the differentiation of *BRCA2*-mutated tumors suggests different cells of origin, *i.e.*, luminal progenitor cells in diverse stages of differentiation (41). As such, *BRCA2*-mutated tumors follow different paths of development, leading to different genomic profiles (42). To be able to find BRCA2 specific characteristics in such a heterogeneous group, larger cohorts are required and subgroups might have to be defined.

BRCA1 status prediction in sporadic breast cancer

We have seen that *BRCA1*-mutated tumors harbor chromosomal aberrations that distinguish them from sporadic tumors (28). The result of classifying a tumor with *BRCA1* promoter hypermethylation as BRCA1-like suggested that the classifier selects for *BRCA1* deficiency, instead of *BRCA1* mutation status (**Chapter 4**). *BRCA* deficiency in sporadic breast cancer is also referred to as "BRCAness", and holds important implications for the clinical management of these cancers (42). The increasing evidence following our initial study that sporadic basal-like breast tumors are similar to hereditary *BRCA1*-mutated breast tumors due to *BRCA1* deficiency (43), motivated us to investigate these two breast tumor subgroups in more

detail for similarities by array CGH. As suspected, BRCA1 related aberrations were not unique to hereditary *BRCA1*-mutated breast cancer but could also be found in a subset of sporadic basal-like tumors where *BRCA1* expression has been lost (37). **Chapter 6** describes that *BRCA1* deficiency in hereditary as well as in sporadic breast cancer exhibits a common genomic profile. To define *BRCA1* deficiency, we assessed *BRCA1* mRNA expression by qRT-PCR in tumor cases with methylated *BRCA1* promoter. Subsequent cases with *BRCA1* qRT-PCR values within the 95% confidence intervals of the methylated cases were considered also to be *BRCA1* deficient. More appropriate would have been to calculate the 95 percentile using bootstrapping on the *BRCA1* qRT-PCR values of the methylated samples and use that as a cutoff for *BRCA1* deficiency and proficiency (44). Bootstrapping 1000 iterations resulted in a median 95 percentile of 0.3475 (unpublished results), and would have been more appropriate to report in the original manuscript. Nevertheless, this cutoff approaches the cutoff described in **Chapter 6** and all the subsequent results would remain identical. The results of this study show that *BRCA1* deficient sporadic basal-like tumors harbor a similar CGH profile as hereditary *BRCA1*-mutated and display the characteristic gain on chromosome 3q and 10p and the loss on chromosome 5q. These findings might allow for the identification of *BRCA1* deficiency in (sporadic) breast cancer, leading to targeted therapy in the future for such cases.

Testing the sporadic basal-like breast cancer cases with our BRCA1 classifier did not result in a clear classification of *BRCA1* deficient and proficient tumor cases as would have been expected (unpublished results). If our BRCA1 classifier would be accurate in identifying true *BRCA1*-related cases, it should be able to distinguish tumors that are not expressing *BRCA1* from those that are. Instead, many of the *BRCA1* proficient cases as defined by qRT-PCR were classified as BRCA1-like breast tumors. These results indicate that our BRCA1 classifier is suffering from underlying biologically driven noise and that the specificity is different than reported. Due to the choice of unmatched control samples, the classification might be based on triple-negative/basal-like characteristics rather than BRCA1 deficiency. A more appropriate control group for the construction of a BRCA1 classifier might had been triple-negative, basal-like breast tumors, expressing *BRCA1*.

BRCAness and homologous recombination deficiency

Both BRCA1 and BRCA2 are involved in homologous recombination to maintain chromosomal integrity. Lack of one of these genes would sensitize tumors to both DNA double-strand breaks (DSB) inducing chemotherapy and PARP inhibitors (45). Therefore, it was proposed to use the BRCA1 and BRCA2 specific CGH profiles described in this thesis as markers for homologous recombination deficiency (HRD) and to study BRCAness in sporadic breast cancer. Lips and colleagues have studied the response on neoadjuvant chemotherapy in 163 breast cancer patients in relation to the outcome of the classifiers (10). Although no association was detected between BRCA1 classification and therapy response in triple negative tumors, luminal BRCA2-like samples seem to respond significantly better to neoadjuvant chemotherapy than those classified as sporadic like. The results of this study indicate that our BRCA2 classifier has strong predictive value for neoad-

juvant chemotherapy in ER positive breast tumors. In another study, Vollebergh and colleagues investigated a cohort of tumors from breast cancer patients that were randomly assigned to adjuvant high-dose platinum-based chemotherapy or conventional anthracycline-based chemotherapy (46). The tumors were tested with our BRCA1 classifier as a synonym for HRD. Patients that received high-dose chemotherapy and of which the tumors were related to BRCA1 by array CGH lived longer than all other patients, suggesting more benefit from DSB-inducing agents. These results indicate that an interesting window of opportunity lies in the identification of *BRCA* status to guide therapy in the future.

Clinical application

The classifiers described in this thesis might function as additional diagnostic tools to screen for *BRCA* related breast cancer. It can serve as indicator for the association with *BRCA* in unclassified variants but also suggest *BRCA* deficiency in sporadic breast cancer cases (28, 36). However, the application of array CGH requires a specialized laboratory, expensive equipment, and is labor-intensive. Therefore, we have developed a PCR based method to test for *BRCA1*-association which can be performed in most diagnostic laboratories. From our array CGH based BRCA1 classifier, the most important chromosomal aberrations have been selected for which a multiplex ligation-dependent probe amplification (MLPA) assay was designed (47). This assay is able to classify *BRCA1*-mutated and sporadic breast cancer with a sensitivity and specificity of 85% and 87%, respectively, and could successfully identify BRCA1 deficient breast tumors with *BRCA1* hypermethylated promoter. Next, we studied 46 breast cancer patients with triple-negative disease that received high-dose platinum-based (HDPB) chemotherapy or conventional chemotherapy. Classifying these 46 cases with our MLPA or array CGH BRCA1 classifier clearly showed that patients with BRCA1 like breast cancer had more benefit from HDPB than conventional chemotherapy (Figure 2). These data suggest that BRCA1 deficient breast tumors can be identified by their chromosomal aberrations and that these patients might benefit from other therapy than conventional chemotherapy. It might therefore be suggested to screen every triple negative hereditary and sporadic breast tumor for BRCA1 deficiency using our BRCA1 classifier.

BRCA status unknown

BRCAX

A major part of the breast (and ovarian) cancer families, which are tested for mutations in the *BRCA1* and *BRCA2* genes, are presented with a negative test result; the genetics behind their predisposition for breast and ovarian cancer remains unknown (**Chapter 1**). Although familial breast cancer is still similarly treated as sporadic breast cancer, targeted therapy might become available in the future and identifying the genetic defects responsible for cancer formation would then be of vital importance. Therefore, much effort has been undertaken to search for undiscovered high-risk breast cancer susceptibility genes. Because linkage analysis has so far not been able to identify new high-risk genes, it is suggested that more than one risk-conferring locus is involved. To be able to identify these loci, genetically homogeneous groups have to be selected first. Using array CGH, we have seen that certain genetic mutations are correlated with specific chromosomal

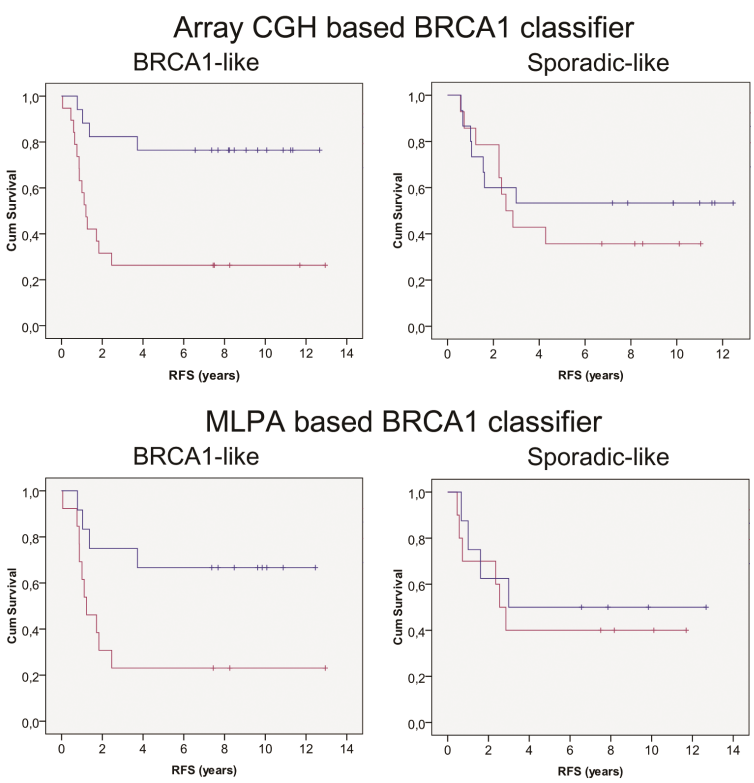


Figure 2 *- Survival curves of patients with triple-negative breast cancer, treated with high-dose platinum-based chemotherapy (blue) or conventional chemotherapy (red), separated based on BRCA1 classification using our original array CGH (upper two panels, n=64) or MLPA classifier (lower two panels, n=40) (47).*
RFS: relapse free survival. Cum Survival: cumulative survival.

aberrations (**Chapter 4 and 5**). It might therefore be possible that non-BRCA1/2 families that have the same genetic defect also present similar chromosomal aberrations. **Chapter 7** describes the array CGH study performed on tumors of non-BRCA1/2 breast cancer families (referred to as BRCAX tumors) in order to find clusters of families with similar genetic profiles and to identify their key genetic characteristics.

First, the array CGH profiles of the BRCAX tumors were compared with those of sporadic breast tumors that were taken from an earlier study (28). Although several clear differences could be distinguished between the two cohorts in respect to aberration frequency, it should be noted that the particular control group was used because of its accessibility and was not matched in any way with the BRCAX tumors. This might explain why the differences that have been found between the groups and those found in a similar study by Gronwald and colleagues (48) are not alike. In addition, because the studied groups are not matched, the results are not independent from other factors and those aberrations found to be prominent in BRCAX tumors but not in the control tumors may not be called "BRCAX-specific" and can not solely be correlated to a single hereditary factor.

Next, several of the CGH profiles of *BRCA1*- and *BRCA2*-mutated breast tumors from two of our earlier studies (28, 36) were compared with those of the BRCAX tumors. Here, as also true for the control tumor comparison, several issues of the methodology have to be dealt with. To be

able to call gains and losses in the CGH profiles, fixed threshold of 0.2 and -0.2 on the copy number alteration segment values are applied (**Chapter 7**). By this, variance between samples due to tumor cell percentage is not taken into account. Furthermore, no differentiation has been made between gain and amplification of genomic regions. It might have been more appropriate to have used CGH profile and tumor cell percentage dependent thresholds as described elsewhere (35). Subsequently, the correlations of the number of gain/no-gain and loss/no-loss between the different cohorts were studied using 2 x 2 Fisher's exact tests. By investigating only two possible observations per measurement (*e.g.*, gain versus no-gain), one assumes that a different state of the measurement is non existent; however, this is not true in respect to chromosomal aberrations. From a biological point of few, a more adequate model would have been to transfer the continuous data of copy number measurements into the following discrete parts: lost (< 2 copies), neutral (2 copies), gained (> 2 copies), and amplified (high-level gain) (35). By this, a 4x2 table containing categorical data (a contingency table) is formed and should be analyzed appropriately (*i.e.*, 4 x 2 Fisher's exact test or 4 x 2 Chi^2 test).

Another point that should be addressed is the comparison between the BRCAX and sporadic breast cancer groups by the average numbers of 'aberrant clones' (log2ratio measurements exceeding the fixed gain/loss thresholds). Such a comparison does not provide any biological information and might only reflect the level of technical noise (although other and better analyses are available for such comparisons). More relevant comparisons would have been the number of chromosomal aberrations, the sizes, and the type of CGH profiles (see **Chapter 1**) which are all missing in the study described in **Chapter 7**.

The more relevant analyses of this study were the comparisons of the BRCAX cases among each other. The main finding was that a part of the BRCAX samples showed gain of chromosome 22 while another part showed loss of chromosome 22, which was also commonly seen in sporadic breast cancer. This suggests that the studied BRCAX samples consisted of at least two subgroups. Interestingly, the chromosome 22 aberrations were consistent among the family members within the same BRCAX families, indicating that gain of chromosome 22 is a BRCAX subgroup specific aberration. Following on this finding, a classifier was built based on the samples that had gain of chromosome 22 (n = 10) and those with loss of chromosome 22 (n = 15). The classifier was built using the nearest shrunken centroid algorithm (49); however, it lacked power for reliable classification. For an error tolerance of < 0.10, more than 15 samples of each class would have been needed (50). This could explain why in subsequent testing of other BRCAX cases, samples with loss of chromosome 22 were classified within the "22-gain" class. An alternative explanation is that gain of chromosome 22 might not be a specific aberration for a subgroup of BRCAX. Therefore, it should be investigated in larger series of BRCAX cases, whether gain of chromosome 22 is a relevant marker and is not related to other factors such as *TP53* mutation status. Because of the limited linkage information of each of the "22-gain" families, a meaningful linkage analysis could not be performed.

In conclusion, because of the high heterogeneity among BRCAX tumors, much larger cohorts are needed in order to identify genetically similar subgroups to perform linkage analyses on.

BRCA status related biology

Co-occurrence of mutated *TP53* and *BRCA1*

It has been shown in mouse models that the development of mammary tumors is highly accelerated when both *Brca1* and *Trp53* have been knocked out. Because *Brca1*$^{-/-}$ mice are embryonic lethal but lethality can be partially rescued by *Trp53* knockout, it is suggested that loss of *Trp53* is required to alleviate the cell-lethal effects of loss of *Brca1*. This has raised the question whether *TP53* loss of function is also required in *BRCA1*-mutated human breast cancer. We therefore have sequenced *TP53* of *BRCA1*-mutated and sporadic breast tumors from our study described in **Chapter 4**, and found that indeed most *BRCA1*-mutated tumors have a pathogenic *TP53* mutation (51). In respect to immunohistochemistry (IHC), it has been shown that *BRCA1*-associated breast tumors are more frequently p53 positive compared to sporadic breast tumors. However, our study shows that *BRCA1*-mutated tumors, negative for p53, are carriers of *TP53* truncation mutations. These results indicate that loss of *TP53* is required in *BRCA1*-mutated tumors. Although it has been shown that *TP53* mutations can be found more often in basal-like breast cancer compared to luminal breast cancer (15), it would be interesting to investigate whether sporadic basal-like breast tumors deficient in *BRCA1* expression harbor a *TP53* mutation more frequently compared to *BRCA1* proficient basal-like breast tumors.

The above findings also open new therapeutic possibilities by targeting the BRCA1 and p53 deficiency combination. Recent research has proposed inhibitors against Cdc7 kinase as a highly specific anti-cancer drug in triple-negative breast cancers (52). Inhibition of Cdc7 activates a p53-dependent checkpoint, resulting in cell cycle arrest to avoid lethal S phase entry in normal cells. Lack of p53 results in abrogation of the Cdc7-inhibion checkpoint, which is followed by lethal S-G2-M phase progression (52).

For the training of our BRCA1 classifier described in **Chapter 4**, we have compared *BRCA1*-mutated with sporadic breast tumors, stratified for IHC based p53 status. As mentioned, mutated *TP53* is common in *BRCA1*-mutated breast tumors, even among the p53 IHC negative cases, making IHC a difficult indicator to interpret for *TP53* status. If we assume that the IHC based p53 status reflects the *TP53* mutation status in sporadic breast tumors well, it might be possible that our study was not stratified for *TP53* status after all because only 43.5% was positive for p53. Specifically, it is known that *TP53* status is associated with loss of chromosome 5q (53, 54); indeed, loss of chromosome 5q was abundantly present in *BRCA1*-mutated breast tumors and therefore part of the BRCA1 classifier (**Chapter 4**). Although *TP53* mutation is a characteristic of *BRCA1*-mutated tumors, and thus also loss of chromosome 5q is, it has to be taken into account that a higher BRCA1-like score might be given to any *TP53* mutated breast tumor *per se*. Although further investigation has to be performed, these results suggest that the CGH profile described specifically for *BRCA1*-mutated breast cancer in **Chapter 4**, probably includes several chromosomal aberrations that are not only specific for *BRCA1*-mutated status but also for other characteristics like ER negative status or *TP53* mutation that are both correlated with loss on chromosome 5q (53).

Secondary BRCA2 mutation

Tumors that are deficient in *BRCA1/2* have a decreased capability to repair DNA and are therefore sensitive to chemotherapeutic agents causing DNA cross-link such as cisplatin or carboplatin and are susceptible to synthetic lethality from PARP inhibitors (55, 56). However, *BRCA1/2* mutation carriers frequently develop recurrent disease that is resistant to platinum agents. It has been shown in *BRCA*-mutated ovarian cancer that *BRCA* function can be restored by a secondary somatic mutation in the tumor, leading to chemotherapy resistance (57-60). So far, little is known about the occurrence of secondary mutations in *BRCA* in breast cancer. To date, cell line HCC1428 is the only cell line providing evidence of secondary mutations occurring *in vivo* in breast cancer (57). HCC1428 was isolated from a woman heterozygous for 6174delT mutation in *BRCA2* after she had undergone chemotherapy. The

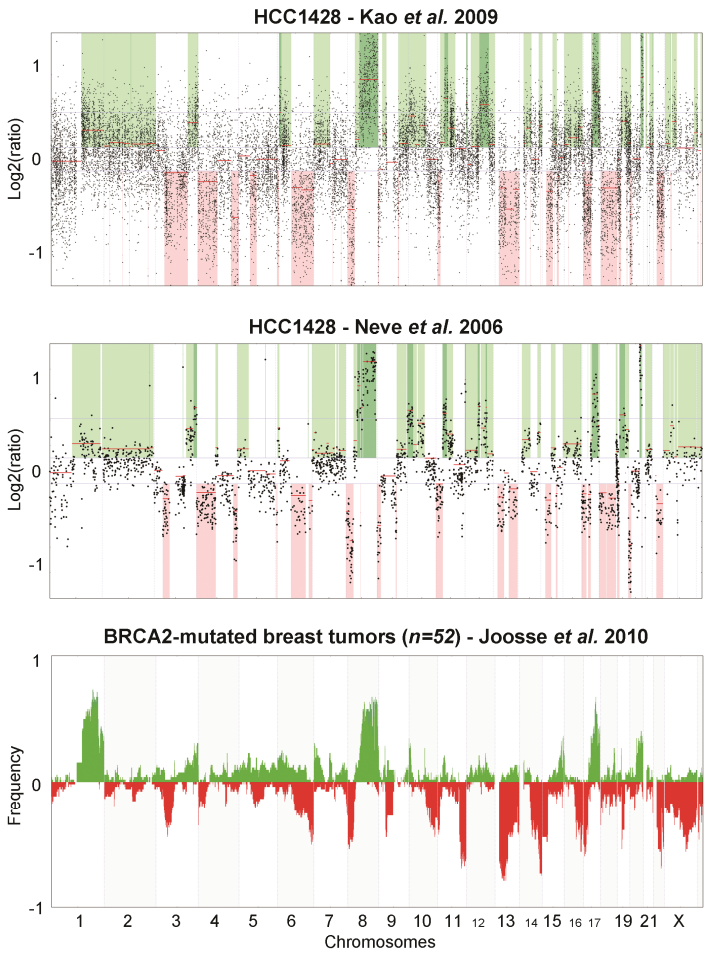


Figure 3 *- Array CGH profiles of breast cancer cell line HCC1428 by Kao et al. and Neve et al. (61 6,2) (upper two panels). Frequency plot of array CGH profiles of 52 proven BRCA2-mutated breast tumors (36). Gain of genomic material is depicted in green; loss of genomic material is depicted in red.*

BRCA2 gene in HCC1428 has a 2135 bp internal deletion that spans the original 6174delT mutation, leading to expression of functional *BRCA2* transcripts. Consequently, HCC1428 is resistant to cisplatin but could be sensitized again by BRCA2 depletion.

From a biological perspective, it would be most interesting to investigate the genomic profile of *BRCA*-mutated breast tumors that have developed a second mutation to restore *BRCA* function. Because of natural selection during the development of a tumor, certain chromosomal aberrations in breast cancer are recurrently found that (indirectly) provide growth advantage. As gain on chromosome 3q and 10p and loss on chromosome 5q are characteristic for *BRCA1*-mutated breast tumors and probably necessary for tumor survival, it can be speculated that upon restoration of *BRCA1* function, some of these aberrations might become redundant to even obnoxious. Further selection on the tumor could then change its genomic make-up. Studying *BRCA*-associated breast tumors with secondary mutations might give further insight on how selection based development of chromosomal aberrations takes place and which role *BRCA1/2* plays in these selection processes.

To test the hypothesis that the genomic makeup changes after a second mutation, I have investigated the CGH profile of the *BRCA2* double mutated cell line HCC1428, published by Kao and colleagues and Neve and colleagues (61, 62). Figure 3 shows the CGH profiles of cell line HCC1428 of both publications (upper two panels) and the frequency plot the CGH profiles of 52 *BRCA2*-mutated breast tumor specimens from our earlier study (bottom panel) (36). Although the cell line shows many similarities with *BRCA2*-mutated tumors in respect to chromosomal aberrations, our CGH BRCA2 classifier classifies the cell line as sporadic-like. This suggests that HCC1428 is not similar to *BRCA2*-mutated tumors as expected. In this specific case however, it is not certain whether the difference is due to the secondary mutation, the preservation of the cells in culture for a long period of time, or that the primary tumor never resembled the general BRCA2-associated CGH profile in the first place.

Conclusion and future perspectives

Several methods exist to predict the association of *BRCA1* or *BRCA2* in the development of cancer. In this aspect, microarray technology is a useful technique that is able to characterize breast cancer at the molecular level, linking these characteristics to *BRCA1* or *BRCA2*. The results of our classification studies and the results found by others indicate that 4-12% *BRCA* related cases, which have not been found by routine BRCA mutation screening, can be identified due to the investigation of the tumor's chromosomal aberrations (28, 36, 63). Although further validation in larger cohorts is required, prediction of *BRCA*-association based on chromosomal aberrations shows to be a promising technique. Using the predictive markers described in this thesis to develop a MLPA based assay is a logical next step to assist mutation screening in high-risk breast cancer patients, or to provide another link between *BRCA* unclassified variants and breast cancer. In addition to hereditary breast cancer diagnostics, linking *BRCA* status to sporadic breast cancer might lead to targeted therapeutic options for these patients in the future. Because BRCA1 and BRCA2 are required for DNA repair by homologous recombination, our BRCA classifiers might also be useful in the

identification of homologous recombination deficient tumors and help to guide anti-breast cancer therapy.

Future perspectives regarding the investigation of *BRCA* status specific chromosomal aberrations would include optimization of a control group for *BRCA1* deficient breast cancer. Such a control group would consist of *BRCA1* proficient basal-like breast cancer. Furthermore, because of the heterogeneity in *BRCA2*-mutated breast cancer, larger cohorts are required for the identification of possible subgroups and their specific aberrations. The same is applicable in the search for BRCAX subgroups where larger cohorts are required. For the detection of chromosomal aberrations, array CGH has been a valuable tool but newer tools such as SNP arrays and next generation sequencing are currently available and provide more detailed and accurate data.

References

1. van Beers EH, et al. 2006. A multiplex PCR predictor for aCGH success of FFPE samples. Br J Cancer 94:333-7.
2. Steidl C, et al. 2010. Genome-wide copy number analysis of Hodgkin Reed-Sternberg cells identifies recurrent imbalances with correlations to treatment outcome. Blood 116:418-27.
3. Arriola E, et al. 2007. Evaluation of Phi29-based whole-genome amplification for microarray-based comparative genomic hybridisation. Lab Invest 87:75-83.
4. Nasri S, et al. 2010. Oligonucleotide array outperforms SNP array on formalin-fixed paraffin-embedded clinical samples. Cancer Genet Cytogenet 198:1-6.
5. Geigl JB, et al. 2009. Identification of small gains and losses in single cells after whole genome amplification on tiling oligo arrays. Nucleic Acids Res 37:e105.
6. Joosse SA, et al. 2007. Automated array-CGH optimized for archival formalin-fixed, paraffin-embedded tumor material. BMC Cancer 7:43.
7. Fles R, et al. 2010. Genomic profile of endometrial tumors depends on morphological subtype, not on tamoxifen exposure. Genes Chromosomes Cancer 49:699-710.
8. Henopp T, et al. 2009. Glucagon cell adenomatosis: a newly recognized disease of the endocrine pancreas. J Clin Endocrinol Metab 94:213-7.
9. Oldenburg RA, et al. 2008. Genome-wide linkage scan in Dutch hereditary non-BRCA1/2 breast cancer families identifies 9q21-22 as a putative breast cancer susceptibility locus. Genes Chromosomes Cancer 47:947-56.

10. Lips EH, et al. 2010. Indicators of homologous recombination deficiency in breast cancer and association with response to neoadjuvant chemotherapy. Ann Oncol .
11. Holstege H, et al. 2010. Cross-species comparison of aCGH data from mouse and human B. BMC Cancer 10:455.
12. Horlings HM, et al. 2010. Integration of DNA copy number alterations and prognostic gene expression signatures in breast cancer patients. Clin Cancer Res 16:651-63.
13. Koski TA, et al. 2009. Array comparative genomic hybridization identifies a distinct DNA copy number profile in renal cell cancer associated with hereditary leiomyomatosis and renal cell cancer. Genes Chromosomes Cancer 48:544-51.
14. Bruin SC, et al. 2010. Specific genomic aberrations in primary colorectal cancer are associated with liver metastases. BMC Cancer 10:662.
15. Holstege H, et al. 2010. BRCA1-mutated and basal-like breast cancers have similar aCGH profiles and a high incidence of protein truncating TP53 mutations. BMC Cancer 10:654.
16. Nielsen R, et al. 2011. Genotype and SNP calling from next-generation sequencing data. Nat Rev Genet 12:443-51.
17. Feuk L. 2010. Inversion variants in the human genome: role in disease and genome architecture. Genome Med 2:11.
18. Oosting J, et al. 2007. High-resolution copy number analysis of paraffin-embedded archival tissue using SNP BeadArrays. Genome Res 17:368-76.
19. Horlings HM, et al. 2008. ESR1 gene amplification in breast cancer: a common phenomenon? Nat Genet 40:807-8.
20. Schena M, et al. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 270:467-70.
21. Prat A, et al. 2011. Deconstructing the molecular portraits of breast cancer. Mol Oncol 5:5-23.
22. van 't Veer LJ, et al. 2002. Gene expression profiling predicts clinical outcome of breast cancer. Nature 415:530-6.
23. Wang Y, et al. 2005. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. Lancet 365:671-9.
24. Chang HY, et al. 2005. Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival. Proc Natl Acad Sci U S A 102:3738-43.
25. Penland SK, et al. 2007. RNA expression analysis of formalin-fixed paraffin-embedded tumors. Lab Invest 87:383-91.
26. Melchor L, et al. 2008. An integrative hypothesis about the origin and development of sporadic and familial breast cancer subtypes. Carcinogenesis 29:1475-82.
27. Melchor L, et al. 2008. Distinct genomic aberration patterns are found in familial breast cancer associated with different immunohistochemical subtypes. Oncogene 27:3165-75.
28. Joosse SA, et al. 2009. Prediction of BRCA1-association in hereditary non-BRCA1/2 breast carcinomas with array-CGH. Breast Cancer Res Treat 116:479-89.
29. Venkatachalam R, et al. 2010. The epigenetics of (hereditary) colorectal cancer. Cancer Genet Cytogenet 203:1-6.
30. Wessels LF, et al. 2002. Molecular classification of breast carcinomas by comparative genomic hybridization: a

specific somatic genetic profile for BRCA1 tumors. Cancer Res 62:7110-7.

31. Jönsson G, et al. 2005. Distinct genomic profiles in hereditary breast tumors identified by array-based comparative genomic hybridization. Cancer Res 65:7612-21.
32. Stefansson OA, et al. 2009. Genomic profiling of breast tumours in relation to BRCA abnormalities and phenotypes. Breast Cancer Res 11:R47.
33. Tirkkonen M, et al. 1997. Distinct somatic genetic changes associated with tumor progression in carriers of BRCA1 and BRCA2 germ-line mutations. Cancer Res 57:1222-7.
34. van Beers EH, et al. 2005. Comparative genomic hybridization profiles in human BRCA1 and BRCA2 breast tumors highlight differential sets of genomic aberrations. Cancer Res 65:822-7.
35. Chin SF, et al. 2007. High-resolution aCGH and expression profiling identifies a novel genomic subtype of ER negative breast cancer. Genome Biol 8:R215.
36. Joosse SA, et al. 2010. Prediction of BRCA2-association in hereditary breast carcinomas using array-CGH. Breast Cancer Res Treat .
37. Joosse SA, et al. 2010. Genomic signature of BRCA1 deficiency in sporadic basal-like breast tumors. Genes Chromosomes Cancer .
38. van den Ouweland AM, et al. 2009. Deletion of exons 1a-2 of BRCA1: a rather frequent pathogenic abnormality. Genet Test Mol Biomarkers 13:399-406.
39. Tischkowitz M, et al. 2008. Pathogenicity of the BRCA1 missense variant M1775K is determined by the disruption of the BRCT phosphopeptide-binding pocket: a multi-modal approach. Eur J Hum Genet 16:820-32.
40. Lakhani SR, et al. 2002. The pathology of familial breast cancer: predictive value of immunohistochemical markers estrogen receptor, progesterone receptor, HER-2, and p53 in patients with mutations in BRCA1 and BRCA2. J Clin Oncol 20:2310-8.
41. Jeselsohn R, et al. 2010. Cyclin D1 kinase activity is required for the self-renewal of mammary stem and progenitor cells that are targets of MMTV-ErbB2 tumorigenesis. Cancer Cell 17:65-76.
42. Turner N, et al. 2004. Hallmarks of 'BRCAness' in sporadic cancers. Nat Rev Cancer 4:814-9.
43. Constantinidou A, et al. 2010. Beyond triple-negative breast cancer: the need to define new subtypes. Expert Rev Anticancer Ther 10:1197-213.
44. Good PI, Hardin JW. Part II Hypothesis testing and estimation. Common Errors in Statistics (and how to avoid them). 2 ed. Wiley-Interscience; 2006. p. 52-5.
45. Annunziata CM, et al. 2010. PARP inhibitors in BRCA1/BRCA2 germline mutation carriers with ovarian and breast cancer. F1000 Biol Rep 2.
46. Vollebergh MA, et al. 2010. An aCGH classifier derived from BRCA1-mutated breast cancer and benefit of high-dose platinum-based chemotherapy in HER2-negative breast cancer patients. Ann Oncol .
47. Lips EH, et al. 2011. Quantitative copy number analysis by multiplex ligation-dependent probe amplification (MLPA) of BRCA1-associated breast cancer regions identifies BRCAness, and as such treatment response. J Clin Oncol 29, 15_suppl:10513.

48. Gronwald J, et al. 2005. Comparison of genomic abnormalities between BRCAX and sporadic breast cancers studied by comparative genomic hybridization. Int J Cancer 114:230-6.
49. Tibshirani R, et al. 2002. Diagnosis of multiple cancer types by shrunken centroids of gene expression. Proc Natl Acad Sci U S A 99:6567-72.
50. Dobbin KK, et al. 2008. How Large a Training Set is Needed to Develop a Classifier for Microarray Data? Clin Cancer Res 14:108-14.
51. Holstege H, et al. 2009. High incidence of protein-truncating TP53 mutations in BRCA1-related breast cancer. Cancer Res 69:3625-33.
52. Rodriguez-Acebes S, et al. 2010. Targeting DNA replication before it starts: Cdc7 as a therapeutic target in p53-mutant breast cancers. Am J Pathol 177:2034-45.
53. Bergamaschi A, et al. 2006. Distinct patterns of DNA copy number alteration are associated with different clinicopathological features and gene-expression subtypes of breast cancer. Genes Chromosomes Cancer 45:1033-40.
54. Fridlyand J, et al. 2006. Breast tumor copy number aberration phenotypes and genomic instability. BMC Cancer 6:96.
55. Foulkes WD. 2006. BRCA1 and BRCA2: chemosensitivity, treatment outcomes and prognosis. Fam Cancer 5:135-42.
56. Fong PC, et al. 2009. Inhibition of poly(ADP-ribose) polymerase in tumors from BRCA mutation carriers. N Engl J Med 361:123-34.
57. Sakai W, et al. 2008. Secondary mutations as a mechanism of cisplatin resistance in BRCA2-mutated cancers. Nature 451:1116-20.
58. Swisher EM, et al. 2008. Secondary BRCA1 mutations in BRCA1-mutated ovarian carcinomas with platinum resistance. Cancer Res 68:2581-6.
59. Sakai W, et al. 2009. Functional restoration of BRCA2 protein by secondary BRCA2 mutations in BRCA2-mutated ovarian carcinoma. Cancer Res 69:6381-6.
60. Norquist B, et al. 2011. Secondary Somatic Mutations Restoring BRCA1/2 Predict Chemotherapy Resistance in Hereditary Ovarian Carcinomas. J Clin Oncol .
61. Kao J, et al. 2009. Molecular profiling of breast cancer cell lines defines relevant tumor models and provides a resource for cancer gene discovery. PLoS One 4:e6146.
62. Neve RM, et al. 2006. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. Cancer Cell 10:515-27.
63. Evans DG, et al. 2003. Sensitivity of BRCA1/2 mutation testing in 466 breast/ovarian cancer families. J Med Genet 40:e107.

Summary

Chapter 1

One of the strongest risk factors for developing breast cancer is a family history of the disease. Although in many families the cause of the high breast cancer incidence can not be identified, most of the DNA mutations that are found reside within the genes *BRCA1* and *BRCA2*. A pathogenic mutation in one of these genes increases the risk for developing breast or ovarian cancer substantially and it is therefore of high clinical importance to recognize the mutation. The gold standard to determine mutations is direct sequencing of genomic DNA. Because the prevalence of a mutation in *BRCA1* or *BRCA2* is in general low, pre-selection is performed based on family characteristics to determine the eligibility for genetic testing. Still, of all the breast cancer families that are eventually screened, it is estimated that approximately 15% of the *BRCA1* and *BRCA2* carriers are missed with standard DNA diagnostics. New techniques are being applied that are based on molecular characteristics of the tumor to identify the underlying genetic defects. Because BRCA1 and BRCA2 are both involved in DNA repair by homologous recombination, defects in one of these genes will be followed by chromosomal instability. Copy number alterations can be investigated by array comparative genomic hybridization (aCGH) which is the main technique used for studies described in this thesis.

Chapter 2

Usually, tumor material is formalin-fixed and paraffin-embedded (FFPE) after it has been surgically removed. This way, the malignant tumor can be preserved and even after long periods of time examined with newly developed techniques. However, fixation time and storage duration can affect the quality of the DNA negatively and the material might not always be suitable for array CGH. We have therefore designed a multiplex PCR that is able to produce amplification fragments of 100, 200, 300, and 400 bp depending on the quality of the DNA. Sample material, which allowed the amplification of at least the 100 and 200 bp fragments, showed to be suitable for array CGH.

Chapter 3

To standardize array CGH, and therefore be able to obtain reproducible results, we have developed an automated hybridization protocol optimized for FFPE material using a hybridization station. By studying the different hybridization durations, temperatures, washing conditions, and hybridization mixture contents, we were able to minimize noise, maximize the dynamic range, and obtain very reproducible array CGH data.

Chapter 4

To find markers specific for *BRCA1*-mutated breast cancer, we designed a classification method based on array CGH data. Chromosomal aberrations of 18 *BRCA1*-mutated and 32 sporadic breast tumors were obtained by array CGH and the two groups were compared. Gains on chromosome 3q and 6p and losses on chromosome 5p, 5q, 7p, 12q, and 20q were identified to be characteristic for *BRCA1*-mutated tumors. Based on the tumor group specific characteristics, a classifier was designed and 16 *BRCA1*-mutated and 16 sporadic breast tumors were tested in order to validate the classifier, resulting in a sensitivity and specificity of 88% and 94%,

respectively. Next, array CGH profiles of 48 non-BRCA1/2 tumors were tested using our classifier. We identified two cases to be BRCA1 like, of which we could identify BRCA1 deficiency through promoter hypermethylation in one of them.

Chapter 5

In order to find markers specific for *BRCA2*-mutated breast cancer, array CGH profiles of 28 *BRCA2*-mutated and 28 sporadic breast tumors were obtained. Chromosomal aberrations were compared and a classifier was constructed. Chromosomal regions that were characteristic for *BRCA2*-mutated breast tumors were gain on chromosome 17q and loss on chromosome 13q and 14q. Validation of the classifier was performed using 19 breast tumors of each of the two groups and resulted in a sensitivity and specificity of 89% and 84%, respectively. We tested the array CGH profiles of 89 non-BRCA1/2 breast tumor cases and found 17 cases to be BRCA2 like. Using additional analyses, clear indications were found that three of the BRCA2 like cases were indeed BRCA2 deficient.

Chapter 6

In order to investigate possible similarities between hereditary and sporadic BRCA1 deficient breast tumors, array CGH was performed on *BRCA1*-mutated and sporadic basal-like breast tumors. Within the sporadic basal-like tumor group, approximately half of the tumors were deficient for *BRCA1* gene expression and presented array CGH profiles that were similar to that of true *BRCA1*-mutated breast tumors. Sporadic basal-like tumors proficient for *BRCA1* gene expression exhibited array CGH profiles that were not similar to *BRCA1*-mutated breast cancer but rather resided to non-basal-like breast tumors using unsupervised hierarchical clustering.

Chapter 7

In more than half of the breast cancer families that are screened for mutations in *BRCA1* or *BRCA2*, no mutation is found. These non-BRCA1/2 breast cancer patients form a large heterogeneous group. To be able to find breast cancer susceptibility genes through linking analysis, more homogenous groups are required. We have therefore investigated 58 tumors from non-BRCA1/2 (BRCAX) families and obtained the array CGH profiles of the genomic make-up of the tumors. The array CGH profiles of the BRCAX tumors were dissimilar to those of *BRCA1*- or *BRCA2*-mutated breast cancer. Within the families, a subgroup was identified that is characterized by gain of chromosome 22 that might be an interesting marker for further studies.

Chapter 8

In the last chapter of this thesis, the work that is presented here is discussed in more detail and compared to current literature; in addition, follow up studies are handled. After publication of our array CGH BRCA1 and BRCA2 classifiers, several studies applied our technique to indicate the pathogenicity of unclassified variants. Moreover, the classifiers have been used in order to find homologous recombination deficient tumors, that would respond differently on certain chemotherapy compared to tumors, in which homologous recombination is still intact.

Samenvatting

Hoofdstuk 1

Eén van de grootste risico factoren voor het ontwikkelen van borstkanker is een belaste familie geschiedenis. Een reden voor hoge borstkanker incidentie kan in de meeste families niet geïdentificeerd worden, echter, de DNA mutaties die gevonden worden bevinden zich meestal in de *BRCA1* en *BRCA2* genen. Een pathogene mutatie in één van deze genen verhoogt het risico voor het ontwikkelen van borst- en ovariumkanker aanzienlijk; het is daarom klinisch van groot belang de mutatie te vinden. De gouden standaard voor het bepalen van mutaties is het sequensen van genomisch DNA. Omdat de incidentie van een *BRCA1* of *BRCA2* mutatie in het algemeen erg laag is, wordt een voorselectie uitgevoerd, die gebaseerd is op familiekarakteristieken, om te bepalen wie in aanmerking komt voor genetische bepaling. Echter, geschat wordt dat ongeveer 15% van de mutatie draagsters gemist wordt met de huidige DNA diagnostiek. Steeds meer nieuwe technieken, die gebaseerd zijn op moleculaire karakteristieken van de tumor, worden toegepast voor het vinden van de genetische defecten. Omdat BRCA1 en BRCA2 beiden betrokken zijn in het DNA reparatie mechanisme door middel van homologe recombinatie, zullen defecten in deze genen tot chromosomale instabiliteit leiden. Veranderingen in chromosoom kopie nummer kunnen onderzocht worden met behulp van array comparative genomic hybridization (aCGH), dit is de meest gebruikte techniek in de studies beschreven in dit proefschrift.

Hoofdstuk 2

Nadat een tumor chirurgisch is verwijderd, wordt het weefsel gefixeerd in formaline en ingebed in paraffine. Op deze manier kan weefsel behouden blijven en ook nog na lange tijd onderzocht worden. Echter, fixatietijd en de duur dat het weefsel bewaard is gebleven, kunnen een negatief effect hebben op de kwaliteit van het DNA wat dan niet meer geschikt is voor array CGH. We hebben een multiplex PCR ontwikkeld waarmee DNA fragmenten van 100, 200, 300 en 400 bp geamplificeerd kunnen worden, afhankelijk van de kwaliteit van het DNA. Monsters waaruit DNA fragmenten van tenminste 100 en 200 bp geproduceerd konden worden, waren geschikt voor array CGH.

Hoofdstuk 3

Om array CGH te standaardiseren en reproduceerbare resultaten te verkrijgen, hebben we een hybridisatie protocol ontwikkeld voor een geautomatiseerd hybridisatie station en geoptimaliseerd voor paraffine materiaal. Voor dit protocol hebben we verschillende tijdsduren van hybridisatie, temperaturen, was condities en hybridisatie mixen getest, waardoor we de ruis wisten te minimaliseren, het dynamische bereik te maximaliseren en zeer reproduceerbare data te verkrijgen.

Hoofdstuk 4

Om specifieke markers te vinden voor *BRCA1* gemuteerde borstkanker, hebben we een classificatie methode ontwikkeld die gebaseerd is op array CGH data. Chromosomale afwijkingen van 18 BRCA1 gemuteerde en 32 sporadische borsttumoren werden verkregen met behulp van array CGH en met elkaar vergeleken. Toename van DNA op chromosoom 3q en 6p en verlies op chromosoom 5p, 5q, 7p, 12q en 20q werden

geïdentificeerd als karakteristiek voor BRCA1 gemuteerde tumoren. Een classificatie methode werd ontwikkeld gebaseerd op de tumor groep specifieke karakteristieken en gevalideerd met behulp van 16 BRCA1 mutant en 16 sporadische borsttumoren. De classificatie methode resulteerde in een sensitiviteit van 88% en een specificiteit van 94%. Vervolgens zijn 48 array CGH profielen van niet-BRCA1/2 borsttumoren getest met onze classificatie methode. Twee zijn geclassificeerd in de BRCA1 klas, waarvan in één geval BRCA1 deficiëntie door middel van promotor hypermethylatie gevonden is.

Hoofdstuk 5

Om markers te vinden die specifiek zijn voor *BRCA2* gemuteerde borsttumoren, werden array CGH profielen verkregen van 28 *BRCA2* gemuteerde en 28 sporadische borsttumoren. Chromosomale aberraties werden vergeleken en een classificatie methode werd ontwikkeld. Chromosomale afwijkingen die karakteriserend zijn voor *BRCA2* gemuteerde tumoren zijn: toename op chromosoom 17q en verlies op 13q en 14q. Validatie van de classificatie methode werd uitgevoerd met 19 borsttumoren van elke tumor groep en resulteerde in een sensitiviteit van 89% en een specificiteit van 84%. Vervolgens zijn 89 niet BRCA1/2 tumor gevallen getest, waarvan 17 zijn geclassificeerd in de BRCA2 groep. BRCA2 deficiëntie werd gevonden in drie gevallen met behulp van extra analysen.

Hoofdstuk 6

Array CGH profielen van *BRCA1* gemuteerde en van basaal-achtige sporadische borsttumoren zijn vergeleken om mogelijke overeenkomsten te vinden. Ongeveer de helft van de sporadische groep is deficiënt in *BRCA1* gen expressie en toonde een vergelijkbaar CGH profiel met *BRCA1* gemuteerde tumoren. Sporadische basaal-achtige tumoren die *BRCA1* gen expressie vertoonde, waren niet vergelijkbaar met *BRCA1* gemuteerde tumoren maar groepeerde samen met niet basaal-achtige tumoren in een ongesuperviseerde hiërarchische analyse.

Hoofdstuk 7

In meer dan de helft van de borstkanker families wordt geen mutatie gevonden in *BRCA1* of *BRCA2*. De niet-BRCA1/2 borstkanker patiënten vormen een grote heterogene groep. Om meer borstkanker susceptibiliteitsgenen te vinden door middel van koppelings-analysen, zijn meer homologe groepen nodig. Daarom hebben we van 58 tumoren uit niet-BRCA1/2 (BRCAX) families de array CGH profielen verkregen. Deze profielen zijn anders dan die van *BRCA1* of *BRCA2* gemuteerde borsttumoren. Een subgroep kon geïdentificeerd worden in de BRCAX families, die gekarakteriseerd was met toename van kopie nummer van chromosoom 22.

Hoofdstuk 8

In het laatste hoofdstuk wordt het werk dat in dit proefschrift vermeld is, in meer detail bediscussieerd en vergeleken met andere gepubliceerde studies. Ook worden follow-up studies behandeld. Na de publicatie van onze BRCA1 en BRCA2 classificatie methoden hebben verschillende onderzoeksgroepen onze technieken toegepast voor het bepalen van de pathogeniteit van ongeclassificeerde varianten in *BRCA1/2*. Tevens zijn onze classificatie methoden gebruikt voor het vinden van homologe recombinatie deficiënte tumoren, die anders zouden reageren op bepaalde chemotherapie dan tumoren waar het homologe recombinatie mechanisme nog intact is.

List of Abbreviations

ALDH1	aldehyde dehydrogenase 1 family, member A1
APC	adenomatous polyposis coli
ATM	ataxia telangiectasia mutated
BAC	bacterical artificial chromosome
BARD1	BRCA1 associated RING domain 1
BCS	breast conserving surgery
BER	base excision repair
BIC	breast cancer information core
bp	base pair
BRCA	breast cancer, early onset
BRCT	BRCA1 C terminus (domain)
CHEK2	checkpoint kinase 2
CHK1	checkpoint kinase 1
CGH	comparative genomic hybridization
CI	confidence interval
CISH	chromogenic in situ hybridization
CNA	copy number alteration
CSC	cancer stem cell
CTRM	c-terminal RAD51 binding motive
DCIS	ductal carcinoma in situ
DGGE	denaturing gradient gel electrophoresis
DHPLC	denaturing high-performance liquid chromatography
DNA	deoxyribonucleic acid
DOL	degree of labeling
DSB	double-strand break
DSS1	deleted in split hand/foot protein 1
EDTA	ethylenediaminetetraacetic acid
EGFR	epidermal growth factor receptor
ER	estrogen receptor
ERBB2	human epidermal growth factor receptor 2
ESR1	estrogen receptor 1
FAMA	fluorescent-assisted mismatch analysis
F-CSGE	fluorescent conformational sensitive gel eletrophoresis
FE test	Fisher's exact test
FFPE	Formalin-fixed, paraffin embedded
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
GE	gene expression
GEP	gene expression profiling
GEO	gene expression omnibus
HBOC	hereditary breast and ovarian cancer
HD	helix-rich domain
HDPB	high-dose platinum-based
H&E	hematoxylin and eosin
HR	homologous recombination
HRD	homologous recombination deficient/deficiency
IDC	invasive ductal carcinoma
IGF	insulin-like growth factor
IHC	immunohistochemistry
ILC	invasive lobular carcinoma
IVS	intervening sequence
kbp	kilo base pairs
KRT	keratin
LC	lobular carcinoma
LCIS	lobular carcinoma in situ
LOD	logarithm of odds
LOH	loss of heterozygosity
LOOCV	leave-one-out cross-validation
LUMC	Leiden University Medical Center
Mbp	mega base pairs
MLPA	multiplex ligation dependent probe amplification
MRI	magnetic resonance imaging
MYC	v-myc myelocytomatosis viral oncogene homolog

NCBI	National Center for Biotechnology Information
NGS	next generation sequencing
NHEJ	non-homologous end-joining
NKI	Netherlands Cancer Institute
NLS	nuclear localization signals
NPP	negative predictive power
OB	oligonucleotide binding
PALB2	partner and localizer of BRCA2
PAM	prediciton analysis of microarrays
PARP	poly ADP ribose polymerase
PBM	prophylactic bilateral mastectomy
PBSO	prophylatic bilateral salpingo-oophorectomy
PCR	polymerase chain reaction
PPP	positive predictive power
PR	progesteron receptor
PTEN	phosphatase and tensin homolog
PTT	protein truncation test
RAD51	RAD51 homolog
RAP80	receptor-associated protein 80
RING	really interesting new gene
RNA	ribonucleic acid
RT-PCR	realtime PCR
SC	shrunken centroids
SCD	SQ-cluster domain
SCP	single-stranded conformational polymorphism
SDS	sodium dodecyl sulfate
SNP	single nucleotide polymorphism
SSB	single-strand break
SSC	saline-sodium citrate
STK11	serine/threonine kinase 11
STR	short tandem repeat
TBE	tris/borate/EDTA
TDGS	two-dimensional gene scanning
TP53	tumor protein 53
TNM	tumor/node/metestasis
ULS	universal linkage system
UV	unclassified variant
VEGF	vascular endothelial growth factor

Curriculum Vitae

Simon Joosse was born on August 6, 1983 in IJmuiden (Velsen), the Netherlands. He started studying medical biology at the faculty of Applied Sciences at Inholland University in Alkmaar in 2000. His graduation report on the subject "Gene inactivation for cell immortality" was carried out under the supervision of Dr. R.L. Beijersbergen at the department of Molecular Carcinogenesis at the Netherlands Cancer Institute, Amsterdam, in 2004. After graduation, he worked at the Netherlands Cancer Institute as a technician for two years and performed the Master Course on Experimental Oncology in 2006, organized by the Oncology school of Amsterdam. In 2007, Simon started working as a PhD student at the department of Experimental Therapy in the Netherlands Cancer Institute, under the supervision of Dr. P.M. Nederlof in the group of Dr. L.J. van 't Veer. In 2009, he moved to Hamburg, Germany, where he started working as scientist for the Department of Tumor Biology, University Medial Center Hamburg-Eppendorf, while simultaneously finishing his PhD studies performed at the Netherlands Cancer Institute. The results of this research are described in this thesis.

List of Publications

1. **Joosse SA**, Hannemann J, Spötter J, Bauche A, Andreas A, Müller V, Pantel K. *Changes in keratin expression during metastatic progression of breast cancer: impact on the detection of circulating tumor cells*. Clin Cancer Res. 2012; [Online publication ahead of print]

2. Hannemann J, Meyer-Staeckling S, Kemming D, Alpers I, **Joosse SA**, Pospisil H, Kurtz S, Görndt J, Püschel K, Riethdorf S, Pantel K, Brandt B. *Quantitative high-resolution genomic analysis of single cancer cells*. PLOS One. 2011 Nov; 6(11):e26362.

3. **Joosse SA**, Hannemann J. *Predication of BRCA status.* CML - Breast Cancer 2011 Jul; 23(2):41-50, Leading article

4. Didraga MA, van Beers EH, **Joosse SA**, Brandwijk KI, Oldenburg RA, Wessels LF, Hogervorst FB, Ligtenberg MJ, Hoogerbrugge N, Verhoef S, Devilee P, Nederlof PM. *A Non-BRCA1/2 hereditary breast cancer sub-group defined by aCGH profiling of genetically related patients*. Breast Cancer Res Treat. 2011 Nov; 130(2):425-36.

5. **Joosse SA**, Brandwijk KI, Mulder M, Wesseling J, Hannemann J, Nederlof PM. *The genomic signature of BRCA1 deficiency in sporadic basal-like breast tumors*. Genes Chromosomes Cancer. 2011 Feb; 50:71-81.

6. Bruin SC, Klijn C, Liefers GJ, Braaf LM, **Joosse SA**, van Beers EH, Verwaal VJ, Morreau H, Wessels LF, van Velthuysen ML, Tollenaar RA, van't Veer LJ. *Specific genomic aberrations in primary colorectal cancer are associated with liver metastases.* BMC Cancer. 2010 Dec; 10(1):662.

7. Holstege H, van Beers EH, Velds A, Liu X, **Joosse SA**, Klarenbeek S, Schut E, Kerkhoven R, Klijn CN, Wessels LFA, Nederlof PM, Jonkers J. *Cross-species comparison of aCGH data from mouse and human BRCA1- and BRCA2-mutated breast cancers*. BMC Cancer. 2010 Aug; 10:455.

8. **Joosse SA**, Brandwijk KI, Devilee P, Wesseling J, Hogervorst FB, Verhoef S, Nederlof PM. *Prediction of BRCA2-association in hereditary breast carcinomas using array-CGH*. Breast Cancer Res Treat. 2010 Jul; [Online publication ahead of print].

9. Horlings HM, Lai C, Nuyten DS, Halfwerk H, Kristel P, van Beers E, **Joosse SA**, Klijn C, Nederlof PM, Reinders MJ, Wessels LF, van de Vijver MJ. *Integration of DNA copy number alterations and prognostic gene expression signatures in breast cancer patients.* Clin Cancer Res. 2010 Jan; 16(2):651-63.

10. Puppe J, Drost R, Liu X, **Joosse SA**, Evers B, Cornelissen-Steijger P, Nederlof P, Yu Q, Jonkers J, van Lohuizen M, Pietersen AM. *BRCA1-deficient mammary tumor cells are dependent on EZH2 expression and sensitive to Polycomb Repressive Complex 2-inhibitor 3-deazaneplanocin A*. Breast Cancer Res. 2009 Aug; 11(4):R63.

11. van den Ouweland AM, Dinjens WN, Dorssers LC, van Veghel-Plandsoen MM, Brüggenwirth HT, Withagen-Hermans CJ, Collée JM, **Joosse SA**, Terlouw-Kromosoeto JN, Nederlof PM. *Deletion of Exons 1a-2 of BRCA1: A Rather Frequent Pathogenic Abnormality.* Genet Test Mol Biomarkers. 2009 Jun; 13(3):399-406.

12. Koski TA, Lehtonen HJ, Jee KJ, Ninomiya S, **Joosse SA**, Vahteristo P, Kiuru M, Karhu A, Sammalkorpi H, Vanharanta S, Lehtonen R, Edgren H, Nederlof PM, Hietala M, Aittomäki K, Herva R, Knuutila S, Aaltonen LA, Launonen V. *Array comparative genomic hybridization identifies a distinct DNA copy number profile in renal cell cancer associated with hereditary leiomyomatosis and renal cell cancer.* Genes Chromosomes Cancer. 2009 Jul; 48(7):544-51.

13. Holstege H, **Joosse SA**, van Oostrom CT, Nederlof PM, de Vries A, Jonkers J. *High incidence of protein-truncating TP53 mutations in BRCA1-related breast cancer.* Cancer Research. 2009 Apr; 69(8):3625-33.

14. **Joosse SA**, van Beers EH, Tielen IH, Horlings H, Peterse JL, Hoogerbrugge N, Ligtenberg MJ, Wessels LF, Axwijk P, Verhoef S, Hogervorst FB, Nederlof PM. *Prediction of BRCA1-association in hereditary non-BRCA1/2 breast carcinomas with array-CGH.* Breast Cancer Res Treat. 2009 Aug; 116(3):479-89.

15. Horlings HM, Bergamaschi A, Nordgard SH, Kim YH, Han W, Noh DY, Salari K, **Joosse SA**, Reyal F, Lingjaerde OC, Kristensen VN, Børresen-Dale AL, Pollack J, van de Vijver MJ. *ESR1 gene amplification in breast cancer: a common phenomenon?* Nat Genet. 2008 Jul; 40(7):807-8; author reply 810-2.

16. Tischkowitz M, Hamel N, Carvalho MA, Birrane G, Soni A, van Beers EH, **Joosse SA**, Wong N, Novak D, Quenneville LA, Grist SA; kConFab, Nederlof PM, Goldgar DE, Tavtigian SV, Monteiro AN, Ladias JA, Foulkes WD, *Pathogenicity of the BRCA1 missense variant M1775K is determined by the disruption of the BRCT phosphopeptide-binding pocket: a multi-modal approach.* Eur J Hum Genet. 2008 Jul; 16(7):820-32.

17. **Joosse SA**, van Beers EH, Nederlof PM. *Automated array-CGH optimized for archival formalin-fixed, paraffin-embedded tumor material.* BMC Cancer. 2007 Mar; 7:43.

18. van Beers EH, **Joosse SA**, Ligtenberg MJ, Fles R, Hogervorst FB, Verhoef S, Nederlof PM. *A multiplex PCR predictor for aCGH success of FFPE samples.* Br J Cancer. 2006 Jan; 94(2):333-7.

Acknowledgements

This work has been made possible with the help of many people, whom I would like to thank here.

First of all, Dr. Petra Nederlof, thank you for giving me the opportunity to perform a PhD project under your supervision. Although my time as a PhD student in your lab was shorter as originally planned, we were able to produce some fine papers! I have learned an enormous amount and the years in your group were most productive.

Many thanks to my promoter Prof. Dr. Peter Devilee. You were always very kind and helpful to me and made it possible that I could finish my promotion even when I moved to Germany. Your music rocks!

I am indebted to Prof. Dr. Laura van 't Veer. It has been a real pleasure working in your group and having been witness of your groundbreaking research. Thank you for supporting me, especially in my last months at the NKI.

Prof. Dr. med. Klaus Pantel is gratefully acknowledged for allowing me to finish my PhD work at the UKE in Hamburg.

Erik van Beers, Renske Fles, and Kim Brandwijk - it has been a great experience working with you in our group! Erik, you have really taught me everything about breast cancer, array CGH, how to write a manuscript (properly and verbose), what it is like being a vegetarian, saxophone, and any other aspect of life on which we had time to talk about.

Much of the work in this thesis has been made possible with the help of the histopathology department, thank you very much for the endless amount of slides you cut and stained!! Frans Hogervorst, thank you for helping me on those occasions when I didn't understand about those different mutation annotations again. Carla van Tiggelen and Sonja Springer for the administration of hundreds of samples. Hans Peterse, you have taught me much about pathology, thank you for always making time for me, spending hours behind the microscope, and being real critical about our work.

I would also like to acknowledge the people from our department C2 (formerly H6), especially: Adrian Begg, thank you for your time and for teaching me all about DNA repair; Els Wagenaar, it has been fun throwing away culture plates without names; Hans te Poele, it is a real joy knowing you, incredible how you could do your experiments, help every one with his/her (little) problems, run around though the building, and still have time for a chat; Thea Eggenhuizen, you knew the whole organization around the building, you have been a great help; and all other members of H6/C2, it has been a wonderful time.

My gratitude goes out to Ron Kerkhoven, Mike Heimerikx, Wim Brugman, and Marja Nieuwland from the microarray laboratory; Lodewyk Wessels for helping me with the bioinformatics and introducing me to MatLab, in which most calculation in this thesis have been performed; and Roderick Beijersbergen, for showing me the ins and outs of basic research during my internship at H2.

Last, but not least I would like to give special thanks to Juliane and Jannik. Thank you for proofreading everything I wrote (Juliane) and criticizing my plots (Jannik). Maar in het bijzonder bedank ik jullie dat jullie er elke dag voor mij zijn en de vreugde in mijn leven brengen.