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).

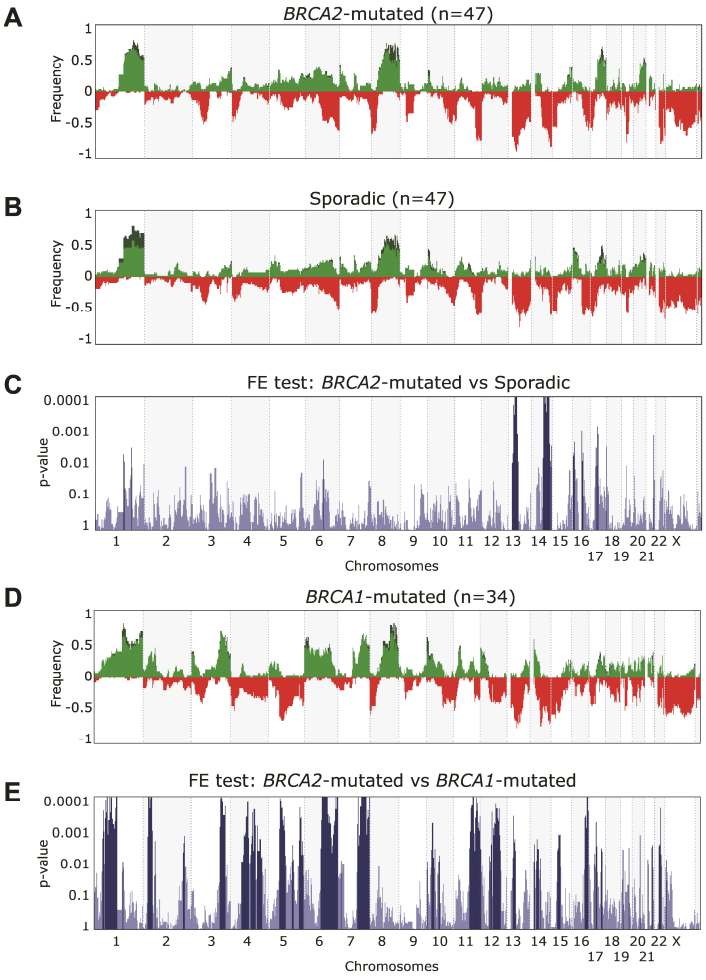

Figure 1 – Comparison of aberration frequency. *Frequency of gain (green), amplification (dark green) and loss (red) over 47 BRCA2-mutated (**A**) and 47 matched sporadic breast carcinomas (**B**) based on the estimated copy numbers as described in Material and Methods. **C**: significance between the 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.

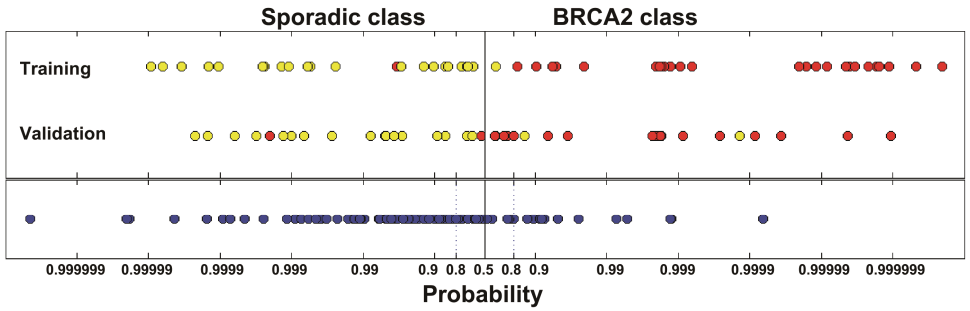

Figure 2 – Classification with the 'BRCA2 classifier'. *The top panel shows the probability scores for the training and validation sets of the BRCA2-mutated (● n=47) and sporadic (● n=47) tumor samples. Samples predicted to be BRCA2-like are plotted right and samples predicted to be sporadic are plotted 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.

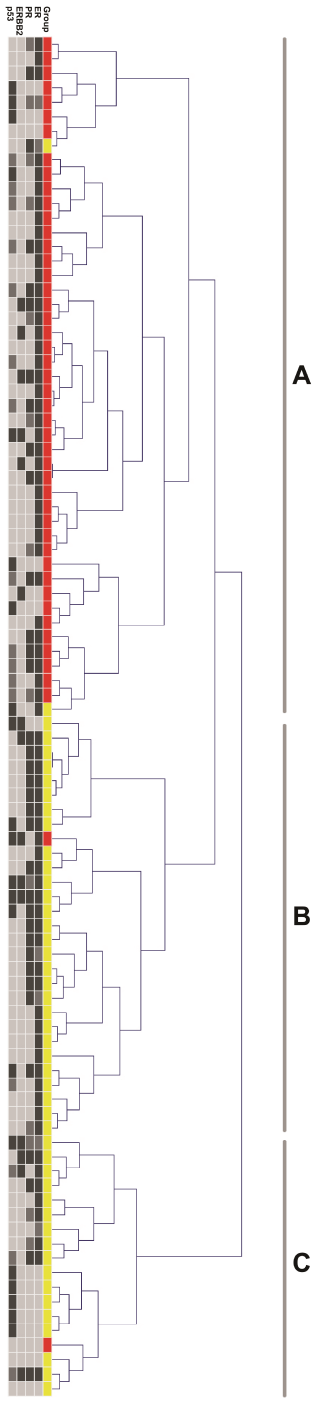

Figure 3 – Hierarchical clustering. *Complete hierarchical clustering of 47 BRCA2-mutated ■ and 47 sporadic ■ breast carcinomas based on the segmentation data of the same 704 BAC clones (shrunken centroids) that were used for the classifier. Shown are the IHC status (from left to right) of p53, ERBB2, 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.