SCISPACE formerly Typeset

Open access • Journal Article • DOI:10.1007/S10549-008-0117-Z

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

Simon A. Joosse, Erik H. van Beers, Ivon H. G. Tielen, Hugo M. Horlings ...+8 more authors

Institutions: Netherlands Cancer Institute, Radboud University Nijmegen Medical Centre

Published on: 01 Aug 2009 - Breast Cancer Research and Treatment (Springer US)

Topics: Breast disease, Breast cancer, Breast carcinoma, Cancer and Genetic counseling

Related papers:

- Hallmarks of 'BRCAness' in sporadic cancers
- An aCGH classifier derived from BRCA1-mutated breast cancer and benefit of high-dose platinum-based chemotherapy in HER2-negative breast cancer patients
- Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy
- Molecular Classification of Breast Carcinomas by Comparative Genomic Hybridization: a Specific Somatic Genetic Profile for BRCA1 Tumors
- Inhibition of Poly(ADP-Ribose) Polymerase in Tumors from BRCA Mutation Carriers

Share this paper:

View more about this paper here: https://typeset.io/papers/prediction-of-brca1-association-in-hereditary-non-brca1-2-sks4qcynl8

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

Simon A. Joosse, Erik H. Beers, Ivon H. G. Tielen, Hugo Horlings, Johannes L. Peterse, Nicoline Hoogerbrugge, Marjolijn J. Ligtenberg, Lodewyk F. A. Wessels, Priscilla Axwijk, Senno Verhoef, et al.

► **To cite this version:**

Simon A. Joosse, Erik H. Beers, Ivon H. G. Tielen, Hugo Horlings, Johannes L. Peterse, et al.. Prediction of -association in hereditary non- breast carcinomas with array-CGH. Breast Cancer Research and Treatment, Springer Verlag, 2008, 116 (3), pp.479-489. 10.1007/s10549-008-0117-z . hal-00478249

HAL Id: hal-00478249

https://hal.archives-ouvertes.fr/hal-00478249

Submitted on 30 Apr 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Breast Cancer Res Treat (2009) 116:479–489
DOI 10.1007/s10549-008-0117-z

PRECLINICAL STUDY

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

Simon A. Joosse · Erik H. van Beers · Ivon H. G. Tielen · Hugo Horlings · Johannes L. Peterse · Nicoline Hoogerbrugge · Marjolijn J. Ligtenberg · Lodewyk F. A. Wessels · Priscilla Axwijk · Senno Verhoef · Frans B. L. Hogervorst · Petra M. Nederlof

Received: 5 March 2008 / Accepted: 1 July 2008 / Published online: 14 August 2008
© Springer Science+Business Media, LLC. 2008

Abstract *Background* 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 counselling 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 tumours in a group of patients initially diagnosed as non-*BRCA1/2* mutation carriers. *Methods* A classifier was built based on array-CGH profiles of 18 *BRCA1*-related and 32 control breast tumours, and validated on independent sets of 16 *BRCA1*-related and 16 control breast carcinomas. Subsequently, we applied the classifier to 48 breast tumours of patients from Hereditary Breast and Ovarian Cancer (HBOC) families in whom no germ line *BRCA1/BRCA2* mutations were identified. *Results* The classifier showed an accuracy of 91% when applied to the validation sets. In 48 non-*BRCA1/2* patients, only two breast tumours presented a BRCA1-like CGH profile. Additional evidence for *BRCA1* dysfunction was found in one of these tumours. *Conclusion* We here

S. A. Joosse and E. H. van Beers are shared first authors.

Johannes L. Peterse—deceased.

Electronic supplementary material The online version of this article (doi:10.1007/s10549-008-0117-z) contains supplementary material, which is available to authorized users.

S. A. Joosse · E. H. van Beers · H. Horlings
Division of Experimental Therapy, The Netherlands Cancer Institute NKI/AvL, Plesmanlaan 121, 1066CX Amsterdam, The Netherlands
e-mail: s.joosse@nki.nl

S. A. Joosse
e-mail: s.joosse@nki.nl

Present Address:
E. H. van Beers
Skyline-Diagnostics B.V, Rotterdam, The Netherlands

I. H. G. Tielen · J. L. Peterse · F. B. L. Hogervorst · P. M. Nederlof
Department of Pathology, The Netherlands Cancer Institute NKI/AvL, Plesmanlaan 121, 1066CX Amsterdam, The Netherlands

N. Hoogerbrugge · M. J. Ligtenberg
Department of Human Genetics, Radboud University Nijmegen Medical Centre, P. O. Box 9101, 6500HB Nijmegen, The Netherlands

M. J. Ligtenberg
Department of Pathology, Radboud University Nijmegen Medical Centre, P. O. Box 9101, 6500HB Nijmegen, The Netherlands

L. F. A. Wessels
Department of Molecular Biology, The Netherlands Cancer Institute NKI/AvL, Plesmanlaan 121, 1066CX Amsterdam, The Netherlands

P. Axwijk · S. Verhoef
Family Cancer Clinic, The Netherlands Cancer Institute NKI/AvL, Plesmanlaan 121, 1066CX, Amsterdam, The Netherlands

P. M. Nederlof (✉)
Postbus 90203, 1006BE Amsterdam, The Netherlands
e-mail: p.nederlof@nki.nl

describe the specific chromosomal aberrations in *BRCA1*-related breast carcinomas. We developed a predictive genetic test for *BRCA1*-association and show that *BRCA1*-related tumours can still be identified in HBOC families after routine DNA diagnostics.

Keywords BRCA1 · Hereditary breast cancer · Array CGH · Classification · Comparative genomic hybridisation

Abbreviations

CGH	Comparative genomic hybridisation
FE	Fisher's exact
FFPE	Formalin-fixed, paraffin-embedded
HBOC	Hereditary breast and ovarian cancer
LOH	Loss of heterozygosity

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 salphingo-oophorectomy [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 germ line DNA, however, it is possible that part of the mutations still remains undetected [4, 6, 11, 12]. Additionally, the detection of variants of unknown clinical significance complicates counselling 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 tumours can be categorised into subclasses [13]. For hereditary *BRCA1*-mutated cancer, previous publications from our research group and others show that these tumours develop distinct genetic alterations on which they can be recognised and distinguished from non-hereditary (sporadic) tumours [14–17]. Various methods using expression profiling [14, 15] or comparative genomic hybridisation (CGH) [16–18] show specific genetic alterations for these tumour groups. Although analysis of tumour mRNA has led to the identification of many different molecular portraits, freshly 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 tumour tissue for pathology based diagnosis. We have previously shown that paraffin embedded tumours 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 tumours 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 tumour 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 tumours using comparative genomic hybridisation 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 tumours for evaluation of current diagnostics.

Patients and methods

Patient selection

This study was performed on primary tumour samples of three breast cancer groups: (1) 34 breast tumours from patients with a confirmed pathogenic *BRCA1* germ-line mutation, mean age at diagnosis of 38 years (range: 27–61); (2) 48 sporadic breast tumours, 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 tumour group (Table 1); (3) 48 tumours 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

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

	BRCA1	Sporadic	HBOC
No. analysed	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)

BRCA1 and *BRCA2*, with a mean age at diagnosis of 48 years (range: 20–61). Patient's 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 CK5/6 was determined by immunohistochemistry staining using the antibodies: estrogen receptor AB-14 clone 1D5 + 6F11, titre 1:50 (Neomarkers); progesterone receptor clone PR-1 titre 1:400 (Immunologic), c-erbB-2 clone SP3, titre 1:25 (Neomarkers); p53 clone D0-7, titre 1:8000 (Dako); and cytokeratin 5/6 clone D5/16 B4, titre 1:100 (Dako) respectively. If $\geq$70% of the tumour cells expressed ER, PR, p53 or CK5/6, the tumour was scored as positive (+) for the corresponding staining, in case $\leq$10% of the cells were stained, the tumour was scored as negative (−) and between 10 and 70% the tumour was scored as ± for the corresponding staining. HER2/neu 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

Tumour DNA was isolated from FFPE tumour tissues as follows. 10 × 10 μm slices containing at least 70% tumour cells were cleared of paraffin (2 × 5 min xylene, 2 × 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% tumour 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 apparently healthy women and pooled. Lymphocytes were purified by adding lysis buffer (155 mM NH_4Cl, 10 mM $KHCO_3$, 1 mM EDTA) 4× 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 5× 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], hybridisations were done on microarrays containing 3.5 k 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 hybridised in randomised 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 $\log_2$-transformed and triplicate spot measurements were averaged.

Aberration detection and quantification

To analyse 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 $\log_2$-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 tumour groups was calculated by employing a 3 × 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 tumours to: (1) the whole set of sporadic tumours 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, HER2/neu or p53 of the tumours (CK5/6 was not used to define a subgroup due to the small number of CK5/6 positive tumours). 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 tumours 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 tumours. 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 tumours. 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 tumours are mostly ER, PR and HER2/neu 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 tumours only.

Class prediction: training

To build a class predictor based on $\log_2$-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 2.32 (1 copy number change) amongst 3277 BAC clones at an error tolerance of 0.05 for the classifier, the minimal sample size for the training set had to be 26, equally divided over the two classes. The class predictor was built on 18 *BRCA1*-related and 32 sporadic breast tumours (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 tumours (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 is 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 analysed 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 analysed 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 tumours. In this report we outline the chromosomal aberrations and their locations, the differences between the tumour groups, and the discriminating power of a class predictor based on our CGH results.

Chromosomal aberrations

We observed significant differences in the spectrum of aberrations with respect to the *BRCA1*-associated and sporadic breast tumours. The upper panels of Fig. 1 depict

Fig. 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 tumour groups computed by Fisher's exact test for each clone (no multiple testing correction). *P*-values <0.01 are indicated dark blue

the frequency of gain (green) and loss (red) of the BAC clones for the *BRCA1*-associated and the sporadic breast tumours, 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 tumours have more copy number alterations (CNAs) compared with the sporadic breast tumours. Table 2 summarizes the most prominent aberrations of both tumour groups. These regions comprise several adjacent clones (at least 10 Mb in size), are aberrant in at least 30% of the tumour cases in one group, or show a significant difference between the tumour 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 tumour groups which have been reported previously to be common for breast cancer [28–31]. In total, the *BRCA1*-associated tumours showed 12 regional (>10 Mb) gains and 11 regional losses that were observed in >30% of the tumours. Using the same criteria, we observed gain in 2 chromosomal regions and 5 regional losses in the sporadic breast tumours.

BRCA1 and sporadic breast tumour class predictor

We used Shrunken Centroids (SC) [24] to discriminate between germ-line-mutated *BRCA1* and sporadic tumours. We randomly selected 18 *BRCA1*-associated and 32 sporadic tumours for the SC analysis; these tumours 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),

Table 2 Chromosomal aberrations

Chromosome	Aberration	BRCA1 percentage	Sporadic percentage	*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

Locations and average frequencies of the most prominent aberrations in 34 *BRCA1*-associated and 48 sporadic breast tumours together with the average *P*-values (FE test) for the significance in aberration difference between tumour groups

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%). Figure 2a depicts the distribution 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, HER2/neu and p53 specific aberrations

In our tumour groups, 91% of the *BRCA1*-mutated tumours are ER, PR and HER2/neu-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, HER2/neu or p53-status with chromosomal aberrations and thus the possible influence on our class predictor, the *BRCA1*-associated tumour group was compared to subsets of the sporadic tumours 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 tumours, independent of the sporadic breast tumours' 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 HER2/neu 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 tumours 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 tumours 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 tumour 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 HER2/neu-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 tumours clustering together with the *BRCA1*-related tumours were also basal-like by performing IHC for CK5/6 (Supplementary data 1). Only, two of the five sporadic breast tumours expressing CK5/6 clustered within the BRCA1-

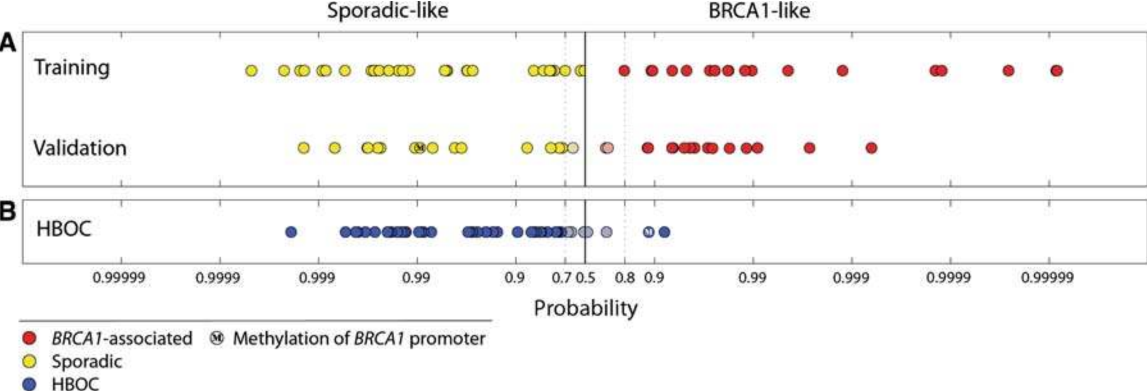

Fig. 2 Classification results. Probability scores for the Training and Validation sets of the *BRCA1*-associated (red) and sporadic (yellow) tumour samples (**a**). Samples predicted as BRCA1-like are plotted right of the 0.5 probability border and samples predicted as sporadic are plotted left. Dotted lines indicate the 95% reference intervals for both classes based on the results of their training sets. Panel **b** shows the classification of the HBOC group (blue) where 2 tumours were classified as BRCA1-like. Samples outside the classes' 95% reference intervals were not assigned to a class (greyed). Samples labelled 'M' show methylation of the *BRCA1* promoter

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 tumour to be BRCA1-like when the tumour is ER negative. In the validation set, 15/16 samples of the *BRCA1*-related tumour 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 tumours' IHC-status (Fig. 3), nevertheless, *BRCA1*-related and sporadic breast tumours can be distinguished using the 191 features selected by the SC algorithm (Figs. 2, 4).

Application of the classifier on non-BRCA1/2 families

Forty-eight patients from non-BRCA1/2 HBOC families were selected and analysed 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. Figure 2b shows the distribution of the clinical samples in comparison with the *BRCA1*-related and sporadic tumours 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* tumour suppressor gene is considered to be required for *BRCA1*-related tumour 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 locos, tumour heterogeneity and limited tumour cell percentage. We next performed additional tests that were not included in the original routine diagnostic setting. As *BRCA1* exon 11 was analysed for truncating mutations using the Protein Truncation Test (PTT) [26], 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, and sporadic tumour C048, which was present in the validation set of the classifier, were the only samples that showed methylation at the *BRCA1* promoter (labelled "M" in Fig. 2). Case HBOC34 also presented with an ovarian carcinoma that, like the breast tumour, showed methylation at the *BRCA1* promoter, interestingly, germ-line DNA of this patient did not show methylation at this site.

Fig. 3 Influence of IHC-status. The four upper panels depict the significance in differences between the *BRCA1*-related cohort and the sporadic IHC tumour 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

Discussion

In this study we show that *BRCA1*-associated breast tumours possess rearranged genomes with specific genomic aberrations that differ significantly from sporadic breast tumours. Based on array-CGH data, we identified the most significant differences between these two tumour groups and built a class predictor with 88% sensitivity and 94% specificity using the Nearest Shrunken Centroids method [24]. Compared with the *BRCA1*-associated tumours, aberrations are less frequent in the sporadic breast tumours. Many of the identified regions specific for the *BRCA1*-related tumours have been reported before [16–18, 32–34]. 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 tumours as BRCA1-like. Since all tumours were formalin-fixed and paraffin-embedded, investigation of mRNA could be problematic [35], and further analyses were performed on genomic DNA from the tumour. 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 tumour. 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 tumour. However, no novel or described mutations in the *BRCA1* gene were

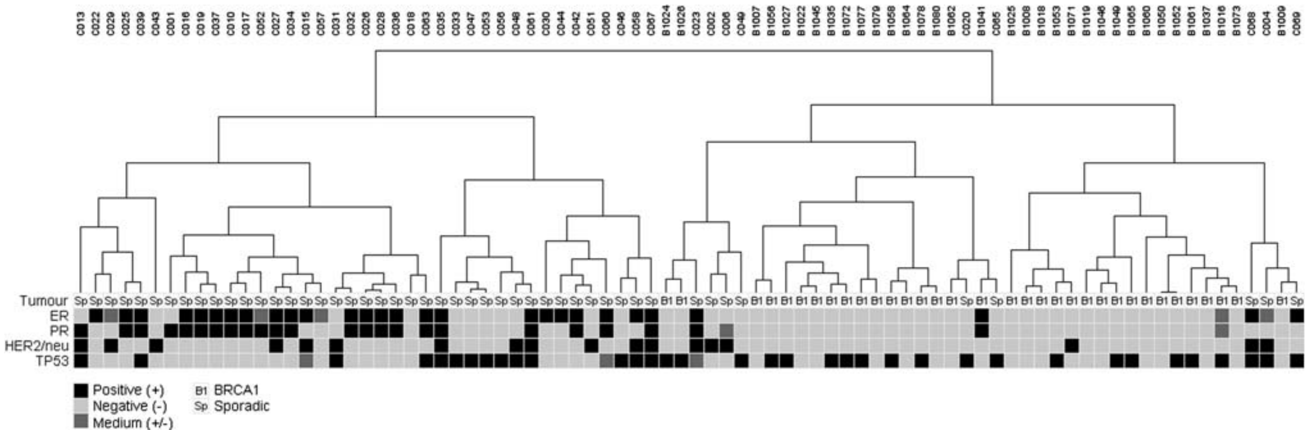

Fig. 4 Hierarchical clustering. Complete hierarchical clustering of 34 *BRCA1*-related and 48 sporadic breast carcinomas based on the $\log_2$-ratios of the 191 BAC clones (shrunken centroids) from the classifier. Shown are the IHC-status of ER, PR, HER2/neu and p53 of all samples. The dendrogram can be divided in two main branches: *BRCA1*-related (right) and sporadic tumour samples (left)

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 tumour was ER and PR-positive which is rare for *BRCA1*-related tumours [38]. This unresolved BRCA1-like case may be analysed 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 tumours showed a small probability (20 and 11.8%) which could explain why no germ-line *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 tumours using chromosomal CGH with an accuracy of 84% [16]. A control group with a relatively large proportion of bilateral tumours 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 tumours 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 micro array in this study localises 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 tumours, we now use a random control group of sporadic tumours 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 discriminatory aberrations, as confirmed by others [17, 32, 33].

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 tumours are in general (>90%) ER, PR, and HER2/neu-negative [38]. Since we chose to randomly select breast carcinomas for our control group, these tumours do not all share the triple-negative phenotype of the *BRCA1*-related tumours. Training on triple-negative breast tumours 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 tumours 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, HER2/neu 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 tumours. 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 tumours 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 tumours 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, 32, 33] 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 tumours 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* germ-line 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 tumour 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 tumours 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 tumours 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 tumours. However, while we could still find *BRCA1*-related breast tumours in a non-BRCA1/2 tumour 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.

References

1. Visser O, Siesling S, van Dijck JAAM (2003) Incidence of cancer in the Netherlands 1999/2000 eleventh report of the Netherlands cancer registry. Available from http://www.ikcnet.nl
2. American Cancer Society (2008) Cancer facts and figures, 2007. Available from http://www.cancer.org/
3. Szabo CI, King MC (1997) Population genetics of BRCA1 and BRCA2. Am J Hum Genet 60:1013–1020
4. van der Hout AH, van den Ouweland AM, van der Luijt RB, Gille HJ, Bodmer D, Bruggenwirth H et al (2006) A DGGE system for comprehensive mutation screening of BRCA1 and BRCA2: application in a Dutch cancer clinic setting. Hum Mutat 27:654–666. doi:10.1002/humu.20340
5. Easton DF, Ford D, Bishop DT (1995) Breast and ovarian cancer incidence in BRCA1-mutation carriers. Breast Cancer Linkage Consortium. Am J Hum Genet 56:265–271
6. Ford D, Easton DF, Stratton M, Narod S, Goldgar D, Devilee P 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 62:676–689. doi:10.1086/301749
7. Antoniou A, Pharoah PD, Narod S, Risch HA, Eyfjord JE, Hopper JL 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 72:1117–1130. doi:10.1086/375033
8. King MC, Marks JH, Mandell JB (2003) Breast and ovarian cancer risks due to inherited mutations in BRCA1 and BRCA2. Science 302:643–646. doi:10.1126/science.1088759
9. Tercyak KP, Peshkin BN, Brogan BM, DeMarco T, Pennanen MF, Willey SC 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 25:285–291. doi:10.1200/JCO.2006.07.3890
10. Cleator S, Heller W, Coombes RC (2007) Triple-negative breast cancer: therapeutic options. Lancet Oncol 8:235–244. doi:10.1016/S1470-2045(07)70074-8
11. Narod SA, Goldgar D, Cannon-Albright L, Weber B, Moslehi R, Ives E et al (1995) Risk modifiers in carriers of BRCA1 mutations. Int J Cancer 64:394–398. doi:10.1002/ijc.2910640608
12. van der Groep P, Bouter A, van der Zanden R, Siccama I, Menko FH, Gille JJ et al (2006) Distinction between hereditary and sporadic breast cancer on the basis of clinicopathological data. J Clin Pathol 59:611–617. doi:10.1136/jcp. 2005.032151
13. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA et al (2000) Molecular portraits of human breast tumours. Nature 406:747–752. doi:10.1038/35021093
14. Hedenfalk I, Duggan D, Chen Y, Radmacher M, Bittner M, Simon R et al (2001) Gene-expression profiles in hereditary breast cancer. N Engl J Med 344:539–548. doi:10.1056/NEJM200102223440801
15. van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AAM, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R, Friend SH (2002) Gene expression profiling of breast cancer accurately predicts clinical outcome of disease. Nature 415:530–536. doi:10.1038/415530a
16. Wessels LF, van Welsem T, Hart AA, van't Veer LJ, Reinders MJ, Nederlof PM (2002) Molecular classification of breast carcinomas by comparative genomic hybridization: a specific somatic genetic profile for BRCA1 tumors. Cancer Res 62:7110–7117
17. Jonsson G, Naylor TL, Vallon-Christersson J, Staaf J, Huang J, Ward MR et al (2005) Distinct genomic profiles in hereditary breast tumors identified by array-based comparative genomic hybridization. Cancer Res 65:7612–7621
18. van Beers EH, van Welsem T, Wessels LF, Li Y, Oldenburg RA, Devilee P et al (2005) Comparative genomic hybridization profiles in human BRCA1 and BRCA2 breast tumors highlight differential sets of genomic aberrations. Cancer Res 65:822–827

19. van Beers EH, Joosse SA, Ligtenberg MJ, Fles R, Hogervorst FB, Verhoef S et al (2006) A multiplex PCR predictor for aCGH success of FFPE samples. Br J Cancer 94:333–337. doi:10.1038/sj.bjc.6602889
20. Joosse SA, van Beers EH, Nederlof PM (2007) Automated array-CGH optimized for archival formalin-fixed, paraffin-embedded tumor material. BMC Cancer 7:43. doi:10.1186/1471-2407-7-43
21. Kang HH, Williams R, Leary J, Ringland C, Kirk J, Ward R (2006) Evaluation of models to predict BRCA germ-line mutations. Br J Cancer 95:914–920. doi:10.1038/sj.bjc.6603358
22. Lakhani SR, Reis-Filho JS, Fulford L, Penault-Llorca F, van der Vijver M, Parry S et al (2005) Prediction of BRCA1 status in patients with breast cancer using estrogen receptor and basal phenotype. Clin Cancer Res 11:5175–5180. doi:10.1158/1078-0432.CCR-04-2424
23. Picard F, Robin S, Lavielle M, Vaisse C, Daudin JJ (2005) A statistical approach for array CGH data analysis. BMC Bioinformatics 6:27. doi:10.1186/1471-2105-6-27
24. Tibshirani R, Hastie T, Narasimhan B, Chu G (2002) Diagnosis of multiple cancer types by shrunken centroids of gene expression. Proc Natl Acad Sci USA 99:6567–6572. doi:10.1073/pnas.082099299
25. Dobbin KK, Zhao Y, Simon RM (2008) How large a training set is needed to develop a classifier for microarray data? Clin Cancer Res 14:108–114. doi:10.1158/1078-0432.CCR-07-0443
26. Loo LW, Grove DI, Williams EM, Neal CL, Cousens LA, Schubert EL et al (2004) Array comparative genomic hybridization analysis of genomic alterations in breast cancer subtypes. Cancer Res 64:8541–8549. doi:10.1158/0008-5472.CAN-04-1992
27. Fridly J, Snijders AM, Ylstra B, Li H, Olshen A, Segraves R et al (2006) Breast tumor copy number aberration phenotypes and genomic instability. BMC Cancer 6:96. doi:10.1186/1471-2407-6-96
28. Bergamaschi A, Kim YH, Wang P, Sorlie T, Hernandez-Boussard T, Lonning PE 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 45:1033–1040. doi:10.1002/gcc.20366
29. Chin SF, Wang Y, Thorne NP, Teschendorff AE, Pinder SE, Vias M et al (2007) Using array-comparative genomic hybridization to define molecular portraits of primary breast cancers. Oncogene 26:1959–1970. doi:10.1038/sj.onc.1209985
30. Ghazani AA, Arneson N, Warren K, Pintilie M, Bayani J, Squire JA et al (2007) Genomic alterations in sporadic synchronous primary breast cancer using array and metaphase comparative genomic hybridization. Neoplasia 9:511–520. doi:10.1593/neo.07301
31. Vincent-Salomon A, Gruel N, Lucchesi C, MacGrogan G, Dendale R, Sigal-Zafrani B 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. doi:10.1186/bcr1666
32. Tirkkonen M, Johannsson O, Agnarsson BA, Olsson H, Ingvarsson S, Karhu R et al (1997) Distinct somatic genetic changes associated with tumor progression in carriers of BRCA1 and BRCA2 germ-line mutations. Cancer Res 57:1222–1227
33. Johannsdottir HK, Jonsson G, Johannesdottir G, Agnarsson BA, Eerola H, Arason A et al (2006) Chromosome 5 imbalance mapping in breast tumors from BRCA1 and BRCA2 mutation carriers and sporadic breast tumors. Int J Cancer 119:1052–1060. doi:10.1002/ijc.21934
34. Vincent-Salomon A, Ganem-Elbaz C, Manie E, Raynal V, Sastre-Garau X, Stoppa-Lyonnet D et al (2007) X inactive-specific transcript RNA coating and genetic instability of the X chromosome in BRCA1 breast tumors. Cancer Res 67:5134–5140. doi:10.1158/0008-5472.CAN-07-0465
35. Penland SK, Keku TO, Torrice C, He X, Krishnamurthy J, Hoadley KA et al (2007) RNA expression analysis of formalin-fixed paraffin-embedded tumors. Lab Invest 87:383–391
36. Snell C, Krypuy M, Wong EM, Loughrey MB, Dobrovic A (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. doi:10.1186/bcr1858
37. Esteller M, Fraga MF, Guo M, Garcia-Foncillas J, Hedenfalk I, Godwin AK et al (2001) DNA methylation patterns in hereditary human cancers mimic sporadic tumorigenesis. Hum Mol Genet 10:3001–3007. doi:10.1093/hmg/10.26.3001
38. Lakhani SR, van de Vijver MJ, Jacquemier J, Anderson TJ, Osin PP, McGuffog L 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 20:2310–2318. doi:10.1200/JCO.2002.09.023
39. Evans DG, Eccles DM, Rahman N, Young K, Bulman M, Amir E et al (2004) A new scoring system for the chances of identifying a BRCA1/2 mutation outperforms existing models including BRCAPRO. J Med Genet 41:474–480. doi:10.1136/jmg.2003.017996
40. Reis-Filho JS, Tutt AN (2008) Triple negative tumours: a critical review. Histopathology 52:108–118. doi:10.1111/j.1365-2559.2008.03046.x
41. Tischkowitz M, Hamel N, Carvalho MA, Birrane G, Soni A, van Beers EH, Joosse SA, Wong N, Novak D, Quenneville LA, Grist SA, Nederlof PM, Goldgar DE, Tavtigian SV, Monteiro AN, Ladias JA, Foulkes WD (2008) Pathogenicity of the BRCA1 missense variant M1775 K is determined by the disruption of the BRCT phosphopeptide-binding pocket: a multi-modal approach. Eur J Hum Genet 16:820–832