4.2. BRCAx Breast Tumors Comprise a Heterogeneous Class, Distinct from Sporadic and BRCA1 Breast Tumors by Array-CGH, and May Reflect Multiple Etiologies

E.H. van Beers, 1 R.A. Oldenburg, 4 S.A. Joosse, 1 L.F.A. Wessels, 2,5 M.J. Ligtenberg, 7 N. Hoogerbrugge, 7 K. Kroeze-Jansema, 8 H. Meijers-Heijboer, 10 F.B. Hogervorst, 4 S. Verhoeef, 4 P. Devilee, 8,9 and P.M. Nederlof 3,11

Submitted

1 Division of Experimental Therapy, 2 Molecular Biology, 3 Molecular Pathology, and 4 Family Cancer Clinic, The Netherlands Cancer Institute, Amsterdam, The Netherlands.
5 Faculty of Electrical Engineering, Mathematics and Computer Science, University of Technology, Delft, The Netherlands.
6 Department of Clinical Genetics, Erasmus Medical Center, Rotterdam, The Netherlands.
7 Dept. of Pathology and Human Genetics, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands.
8 Dept. Of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands.
9 Department of Pathology, Leiden University Medical Center, Leiden, The Netherlands.
10 Department of Clinical Genetics, VU University Medical Center, Amsterdam, The Netherlands
11 Corresponding author P. Nederlof, e-mail: p.nederlof@nki.nl

ABSTRACT

Only about 25% of familial breast cancer is explained by mutations in BRCA1 and BRCA2, fewer by moderate penetrance genes like P53, PTEN, CHEK2, ATM and PALB2 and an unknown fraction by common variants of genes with low penetrance. Evidence suggests that additional dominant breast cancer genes exist and these are referred to as BRCAx. Clinical presentation of families with highly increased incidence of breast cancer that are non-BRCA1/BRCA2, suggests dominant inheritance of such high penetrance breast cancer genes. Because cancer genes often confer a specific clinical presentation (e.g. age of onset, sex-ratio, tissue spectrum) it seems
useful to initiate their discovery by such clinical criteria. An earlier study of BRCAx/ non-BRCA1/2 breast cancer families aimed to enrich for a common genetic defect by setting stringent inclusion criteria, failed to identify new breast cancer susceptibility loci. Motivated by results of BRCA1 and BRCA2 breast tumors that have characteristic array-CGH signatures (array-CGH ‘phenotypes’), we study BRCAx breast tumor by array-CGH and show that BRCAx tumors are distinct from sporadic controls but are otherwise still heterogeneous. This provides a possible explanation for the lack of high LOD scores in these patients and would be consistent with more than one BRCAx sub-type and therefore more than one BRCAx gene. We propose approaches that can be employed to further sub-stratify BRCAx families based on array-CGH data.

INTRODUCTION
The majority of excess familial breast cancer risk is unexplained. We now know that the underlying genetics of breast cancer susceptibility is very complex. BRCA11 and BRCA2 were the first breast cancer genes identified due to the combination of high penetrance (carriers have substantially increased risk) with large affected families for confirmative co-segregation analyses. These two genes have been widely studied and their collective contribution to hereditary breast cancer incidence is now estimated at perhaps 20-25%.3,4 The other 75% of hereditary cases are likely caused by a multitude of unknown risk factors and are attributable to recessive genes, combinations of genes or to common variants of genes conferring only slightly elevated risks.5,6 Yet, analyses of BRCA1 and BRCA2 allele frequencies and penetrance, and especially patterns of familial clustering of breast cancer suggested the existence of additional, dominant high penetrance breast cancer genes, referred to as either ‘BRCAx’ or ‘BRCA3’.7 A linkage study on 149 BRCAx families, including those in the present study, failed to identify sufficiently high LOD scores to guide positional cloning of the gene(s) in these families6 which led the authors to conclude that perhaps more than one risk conferring locus was involved. Because similar to most breast tumors, BRCAx breast tumors present with genomic instability6,10 we performed array-CGH of BRCAx breast tumors to catalogue possible distinct and recurrent CGH profiles. We compare BRCAx array-CGH profiles with those of BRCA1, BRCA2 and sporadic tumors and describe possible similarities and particularities among BRCAx breast tumors and propose an approach for further analysis of BRCAx families.
MATERIALS AND METHODS

Patients and tumor specimens

Family selection, tumor collection and immunohistochemistry have been described in Oldenburg et al. Control tumors were selected to have no evidence of familial risk for breast or ovarian cancer and were on average as young (45.5) as the BRCAX patients (52.5) years. All control samples were selected from the institute’s archival tissue bank and are described in detail elsewhere. Of 92 patients from 42 BRCAX families, we isolated high-quality DNA from 58 unique tumor samples from 27 different families to perform array-CGH as before. Immunohistochemistry was performed as described, for BRCAX and for sporadic controls.

Array-CGH

dNA from each tumor and a reference of pooled DNA from seven healthy females was used in ~1 Mb genome-wide BAC array-CGH. Automated hybridizations were performed for 72 hrs, followed by automated washes and drying as described. Arrays were scanned with an Agilent DNA microarray scanner. Signal intensities were determined with Imagene software and raw data processing involved only median pintip (c.q. sub-array) normalization. Array-CGH profile log2 ratios were used in unsupervised hierarchical clustering using Matlab software (v.7.0.1, The Mathworks, Natick MA, USA).

RESULTS

Array-CGH of BRCAX tumors

Of a previous collection of BRCAX invasive breast tumors encompassing 84 FFPE blocks, we isolated genomic tumor DNA of sufficient quality from 58 unique cases from 27 different families to perform automated array CGH. Forty-eight invasive ductal sporadic breast carcinomas (IDC) with similar age of incidence were used for comparison. Array-CGH copy number alterations (CNA’s) were interpreted as gains when log2 ratios were > 0.2 and losses when log2 ratios were < -0.2. Counts of aberrations for each of the 58 BRCAX and 48 sporadic control breast tumors are shown in figure 1. The average number of BAC’s reporting CNA in BRCAX tumors was 1063 (sd = 442) and in sporadic controls 816 (sd = 386). This difference was statistically different in a two-sided, unpaired t-test (p = 0.003). We conclude that on average BRCAX tumors have (1063/816) 130.3% aberrant clones compared with sporadic tumors.
**BRCAX array-CGH aberrations in comparison with BRCA1 and sporadic controls**

Figure 1 is a frequency plot of array-CGH gains and losses in BRCAX versus control breast tumors (figure 1a) versus BRCA1 tumors (figure 1b) and versus BRCA2 tumors (figure 1c). Continuous CGH data consisted of 3248 measurements (BAC probes) from chromosome 1p-tel to Xq-tel. All CGH data were segmented before counting outlier frequencies as a fraction of each tumor class showing a gain (log2 ratios > 0.2) or loss (log2 ratios < -0.2) for all 3248 positions measured. Gain and loss frequencies (0-1) are plotted on the left y-axis. The p-values for the differences between the tumor classes were computed by two-sided Fisher exact testing and plotted in green (differential gains) or red (differential losses). Longer sticks correspond with smaller p-values and therefore more significant regions (figure 1).

The comparison between BRCAX and sporadic tumors shows multiple regions of differential gains and losses (figure 1a). Differential gains are prominent on chromosome 2q-ter, 6p, 8p, 11p, 12p, 14q, 17p, 17q, 19p, 19q, and along the entire Chromosome 22. Differential losses are prominent on 1p, 1q, 4q, 5q, 9q, 13q, 14q, 15q, 19cen, 21p and Xp. Despite these differences there is also abundant overall similarity between BRCAX and sporadic breast tumor array-CGH profiles. For example, frequent aberrations observed in both classes are gain of chromosome 1q and 8q. Significant differential gains and losses depend highly on which tumor classes are compared (figure 1a, b, and c). A comparison of differential recurrent aberrations between BRCAX and BRCA1 (figure 1b) or BRCAX and BRCA2 (figure 1c) resulted in different sets of significant regions. Figure 1b shows for instance regions that are known to be highly specific to BRCA1 tumors such as 3p, 3q and 5cen. Aberrations in these three regions in BRCA1 tumors (figure 1a, black line) were more frequent then in BRCAX tumors (blue line), and thus more characteristic for BRCA1 tumors. Other significant differential CGH results in figure 1 include chromosome 13p loss, which is more frequent in BRCA2, while a region on 12p towards the centromere appeared as gain in approximately 30% of BRCAX tumors, which was not found in controls, BRCA1 or BRCA2 tumors.

**Array-CGH aberration banding (Pearson banding)**

Another analysis that highlights the specific gains and losses of BRCAX tumor array-CGH profiles is shown in figure 2. We calculated Pearson correlation coefficients of the log2 ratios between all possible pairs of BAC’s per tumor class per chromosome. These Pearson coefficients are plotted in the three top panels as heat-maps. The Pearson heat maps detect Pearson-stable regions ('bands') with great sensitivity (as 'oran-
Figure 1. Significantly different CGH aberrations in BRCAX tumors (blue) and control tumors (black). The top panel shows Chromosomes 1 through 8, and the bottom panel Chromosome 9 through X. The x-axis represents all 3248 probes on Chromosomes 1 through X and vertical black lines indicate centromers. On the y-axis are the frequencies of aberrations (logratio $>0.2$) in 58 BRCAX tumor CGH profiles (blue) and 48 sporadic control tumors (black). Vertical green bars correspond to between-class gain significance as determined with a two-sided Fisher exact test ($p$-value scale on right y-axis, ranging from $p=0.05$ at the X-intercept to 0.00 at the top and bottom, i.e. all depicted bars are significant at the 5% level). Similarly, red bars indicate significance for differential losses between BRCAX and sporadic tumors. (class X $>$ Class X in figure).
ge' blocks). Below each Pearson heat map, we plotted the log2 ratio heat-maps for the same chromosomes (other chromosome figures are given as supplementary data).

**Figure 2a-2e.** Pearson correlations (top panels) and CGH log2 ratios (bottom panels) in whole chromosomes 13, 16 and 22. The top panels have all BAC clones for that particular chromosome on both the x and y-axis. The bottom panels show chromosome centromers if present (arrow) while individual samples are stacked and sorted along the y-axis. This vertical sorting in the lower panels samples is performed per class, based on sample-to-sample complete correlation clustering. Color scales were set to saturate at -1 and 1 for correlation (red-blue-orange) and log2 ratios (red-white-green). Horizontal axes not plotted to scale but depend on the number of features on each particular chromosome.

**Fig. 2a Chromosome 21**

**Chromosome 21**

Figure 2A shows results of 58 BRCAX tumors, flanked by 28 BRCA1 and 49 sporadic control tumors. The majority of BRCAX tumors show loss of 21p and gain of 21q but the reverse pattern can be discerned also. Both other classes have quite different CGH profiles with more scattered patches of gains and losses. This is reflected in the three correlation panels with more distinct p and q 'blocks' of high correlation for BRCAX tumors compared with the two other classes.
Figure 2B shows chromosome 12. Here brca1 tumors seem the most homogeneous of the three classes plotted. While brcax and sporadic control cgh data (lower panels) might look somewhat similar and heterogeneous, the correlation panels indicate that brcax tumors have more off-diagonal high-correlation regions. This means similarity between dis-continuous segments of chromosome 12, for example BAC’s 1~20 with 60~80. This could suggest intra chromosomal rearrangement joining these regions in brcax tumors. The brcax cgh log2 ratios further indicate that this class is heterogeneous, with at least several different types of profiles for this chromosome.
**Chromosome 22**

Figure 2C shows the results for chromosome 22. It seems to suggest two types of BRCAx tumors, namely those with gain and those with loss of the entire chromosome, but also a few tumors with more complex rearrangements. This heterogeneity among BRCAx tumors could not be seen in Figure 1, which presents average data for whole tumor classes.

**Fig. 2d Chromosome 13**

![Graphs showing RAC similarity for BRCA1, BRCAx, and Controls](image)

**Chromosome 13**

Figure 2D shows chromosome 13. Pearson correlations are high throughout this chromosome for BRCAx samples (top middle) but vary considerably across this chromosome in the BRCA1 class (top left). We conclude that BRCAx tumors have fewer transitions from gain to loss or vice versa and thus have a more stable chromosome 13 compared with BRCA1 tumors. Interestingly, there are a few BRCAx tumors with different chromosome 13 profiles.

**Chromosome 16**

Figure 2E shows chromosome 16. Aberrations of Chromosome 16p are more variable in BRCA1 (top left, more blue) compared with either BRCAx (middle) or controls (right). The BRCAx class showed a unique recurrent loss (red in lower panel) between clone 61 and 70 that was inversely correlated to the log2 ratios (top middle, red correlations) of BAC clones 1 ~ 40 on the same chromosome. This means that a gain of 16p seems to co-occur with loss of another specific region of 16q only among BRCAx.
Overall \textit{brcax} array-cGH profiles for this chromosome are quite similar to sporadic controls. This is consistent with Figure 1 that already indicated no significant differences for chromosome 16 between \textit{brcax} and sporadic tumors.

**Heterogeneity of \textit{brcax} tumor array-cGH profiles**

We have shown that cGH profiles of \textit{brcax} tumors are different from those of sporadic controls and from \textit{brca1} and \textit{brca2} tumors. However this does not mean that \textit{brcax} tumor profiles are homogeneous. Although \textit{brcax} tumors were selected with stringent inclusion criteria, we lack understanding of the genetic factor(s) that caused their apparent familial excess risk and the possibility remains that \textit{brcax} families represent more then one risk factor that might associate with distinct \textit{brcax} array-cGH sub-phenotypes. We therefore investigated the extent of heterogeneity among these 58 \textit{brcax} tumors by unsupervised hierarchical clustering (complete linkage Pearson correlation of whole cGH profiles). Figure 3A uses all 3248 log2 ratios for all autosomes and chromosome x. Patients are never hybridized in sex mismatch since the normal reference pool is also female DNA. The hierarchical trees in Figure 3A and B show clustering of the 59 \textit{brcax} and 49 sporadic tumor samples, respectively. These dendrograms indicated that the 'within class' heterogeneities are comparable between sporadic and \textit{brcax} tumors and that there are no major branch points to suggest obvious distinct \textit{brcax} array-cGH subtypes. Then we co-clustered all sporadic and \textit{brcax} tumors in figure 3C together with their immunophenotypes, and found that individual tumors did not mix randomly. Both \textit{brcax} and sporadic
tumors remained clustered in just eight sub-clusters, five of which (I-V) contained only \textit{BRCA} tumors and three clusters contained all sporadic cases plus one \textit{BRCA} tumor.

\textbf{DISCUSSION}

\textbf{Selection of \textit{BRCA}X tumors}

Because the breast cancer gene(s) in \textit{BRCA}X families is (are) unknown, \textit{BRCA}X is solely defined by clinical criteria, including a negative test for known breast cancer genes like \textit{BRCA}1 and 2. In our study, a nation-wide collection of such \textit{BRCA}X tumors

\textbf{Figure 3. Unsupervised Hierarchical Cluster of brcax and Control tumor cgh profiles.}

(A) Unsupervised complete correlation clustering of 58 \textit{BRCA}X tumor array-cgh profiles is performed in Matlab (The Mathworks, Natick MA, USA) using log2 ratios for 3248 probes from chromosome 1-X. Vertical distances represent the similarity distance calculated across all 3248 probes. (B) Similar to A, shows 48 sporadic breast tumors. (C) \textit{BRCA}X tumors and sporadic tumors co-clustered. \textit{BRCA}X are blue, sporadics are yellow. Immunohistochemical staining scores are given as no staining (grey), positive staining (black) or missing data (hatched). The bottom legend indicates the \textit{BRCA}1-likelihood score in our \textit{BRCA}1 classifier\textsuperscript{14}. Red = ‘\textit{BRCA}1-like’, grey = undecided, yellow = ‘sporadic-like’.
from selected families\textsuperscript{8} was analyzed by array-cgh. These families had $\geq 3$ breast cancers below 60 years, and no ovarian or male breast cancer, and are more stringently selected compared with earlier \textit{BRCAX} reports that have not excluded ovarian cancer,\textsuperscript{10,17,18} that included samples with less than 70% tumor cells also,\textsuperscript{10} or included 6 (of 18) families with just 2 cases of breast cancer.\textsuperscript{19} The abovementioned differences may impact on risk factor stratification and therefore limits a comparison with our study. Approximately half of the samples (13 families) in this study were analyzed before in a linkage analysis of \textit{BRCAX} families\textsuperscript{20} and failed to map significant LOD scores leaving these families and their tumors largely uncharacterized at the genomic level. The current cgH analysis has revealed a high degree of heterogeneity among these 58 \textit{BRCAX} samples, which could explain why the previous linkage analysis was unsuccessful.

\textbf{BRCAX characteristic aberrations}

The most significant differential chromosomal losses ($< \log_2 \text{ratio} -0.2$) between \textit{BRCAX} and sporadic tumors were found on chromosomes 1, 4, 5, 13, 14, 15, 19, 21, X, and gains ($> \log_2 \text{ratio} 0.2$) on 2, 6, 8, 10, 11, 12, 14, 17, 19, 20, and 22. Notably, all these aberrations were more frequent in the \textit{BRCAX} class than in the sporadic class. This is consistent with the results in Figure 1 showing that \textit{BRCAX} tumors have $\sim 30\%$ more aberrations than sporadic breast cancers and also with results that indicate more genetic instability in hereditary compared with sporadic cases.\textsuperscript{21} Chromosome 22 gain among \textit{BRCAX} tumors seems unique since we have not observed this high frequency gain in either sporadic, \textit{BRCA1} or \textit{BRCA2} breast tumors. It will be of interest to unravel the role of chromosome 22 in \textit{BRCAX} tumors.\textsuperscript{22}

\textbf{BRCAX heterogeneity}

An elusive but crucial aspect of \textit{BRCAX} families is whether or not different risk factors were co-selected by the clinical criteria used. We realize that this ‘catch-22’ will only end with identification of the risk factor(s) and therefore we hope to facilitate their identification by providing a possibly relevant stratification based on array-cgh profiles. At this time it remains impossible to predict whether IHC, array-cgh or any other method will provide such stratification.

\textbf{Could \textit{BRCAX} tumors be false negative \textit{BRCA1} tumors?}

The possibility of false negative \textit{BRCA1} diagnoses among the \textit{BRCAX} tumors analyzed is low because these families, and individuals included in this study have tested ne-
positive for \textit{BRCA1} and \textit{BRCA2} in routine screening. Furthermore, a prognostic CGH study in HBOC families with the highest Evans’ scores (i.e predicted to have \textit{BRCA1} mutations)\textsuperscript{14} found very few \textit{BRCA1} mutations. Finally, one might argue that sensitivity of the \textit{BRCA1} CGH classifier has not been estimated in elaborate studies, but it has detected one previously unclassified \textit{BRCA1} variant M1775K.\textsuperscript{23}

\textbf{Future directions}

Heterogeneity among tumors is difficult to quantify. Figure 3C serves as a hypothesis generator with respect to the question whether one or multiple \textit{BRCA} genes exist. It seems to indicate that certain tumor array-CGH profiles are more similar to each other compared with the rest but knowing which split(s) in the dendrogram might coincide with separate risk factors can only be tested by further analysis of multiple tumors from multiple \textit{BRCA} families and is currently in progress.

Due to the patient selection, we now have a dataset comprised of multiple tumors of multiple \textit{BRCA} families that will allow us to further analyze whether breast tumor profiles are more similar within families. This has proven difficult in this preliminary analyses perhaps due to the fact that not all genomic alterations recur with the same frequency, and because array-CGH profiles have unknown contributions from random (experimental and sampling noise) and a non-random (true) CNV’s. Recent studies have estimated 26% ‘phenocopies’ in breast cancer in breast cancer families.\textsuperscript{24} Therefore, it seems that equal weighing all log2 ratios, and more importantly including all \textit{BRCA} tumors without excluding phenocopies, will be inappropriate to define such ‘family intrinsic profiles’. We hypothesize that further studies of \textit{BRCA} families based on CGH profile similarities, could contribute to the identification of the (perhaps multiple) \textit{BRCA} loci.

\textbf{REFERENCE LIST}


7. Hopper JL. More breast cancer genes? 


   (submitted)


