A Predictor Based on the Somatic Genomic Changes of the BRCA1/BRCA2 Breast Cancer Tumors Identifies the Non-BRCA1/BRCA2 Tumors with BRCA1 Promoter Hypermethylation

Sara Alvarez,1 Ramon Diaz-Uriarte,3 Ana Osorio,1 Alicia Barroso,1 Lorenzo Melchor,1 Maria Fe Paz,2 Emiliano Honrado,1 Raquel Rodriguez,1 Miguel Urioste,1 Laura Valle,1 Orland Díez,5 Juan Cruz Cigudosa,4 Joaquin Dopazo,3 Manel Esteller,2 and Javier Benitez1

1Human Genetics Department and 2Cancer Epigenetics Group, Molecular Pathology Programme, and 3Bioinformatics Unit and Cancer Centre, Madrid, Spain; and 5Department of Genetics, H. Santa Creu i Sant Pau, Barcelona, Spain

ABSTRACT

The genetic changes underlying in the development and progression of familial breast cancer are poorly understood. To identify a somatic genetic signature of tumor progression for each familial group, BRCA1, BRCA2, and non-BRCA1/BRCA2 (BRCAX) tumors, by high-resolution comparative genomic hybridization, we have analyzed 77 tumors previously characterized for DNA copy number by a large number of chromosomal changes, some of which by a large number of chromosomal changes, some of which by high-resolution comparative genomic hybridization, we have analyzed 77 tumors previously characterized for DNA copy number by a large number of chromosomal changes, some of which by high-resolution comparative genomic hybridization, we have analyzed 77 tumors previously characterized for DNA copy number by a large number of chromosomal changes, some of which by high-resolution comparative genomic hybridization, we have analyzed 77 tumors previously characterized for DNA copy number by a large number of chromosomal changes, some of which by high-resolution comparative genomic hybridization, we have analyzed 77 tumors previously characterized for DNA copy number by a large number of chromosomal changes, some of which by high-resolution comparative genomic hybridization, we have analyzed 77 tumors previously characterized for DNA copy number by a large number of chromosomal changes, some of which by high-resolution comparative genomic hybridization, we have analyzed 77 tumors previously characterized for DNA copy number by a large number of chromosomal changes, some of which by high-resolution comparative genomic hybridization, we have analyzed 77 tumors previously characterized for DNA copy number by a large number of chromosomal changes, some of which by high-resolution comparative genomic hybridization, we have analyzed 77 tumors previously characterized for DNA copy number by a large number of chromosomal changes, some of which by high-resolution comparative genomic hybridization, we have analyzed 77 tumors previously characterized for DNA copy number by a large number of chromosomal changes, some of which by high-resolution comparative genomic hybridization, we have analyzed 77 tumors previously characterized for DNA copy number by a large number of chromosomal changes, some of which by high-resolution comparative genomic hybridization, we have analyzed 77 tumors previously characterized for DNA copy number by a large number of chromosomal changes, some of which by high-resolution comparative genomic hybridization, we have analyzed 77 tumors previously characterized for DNA copy number by a large number of chromosomal changes, some of which by high-resolution comparative genomic hybridization, we have analyzed 77 tumors previously characterized for DNA copy number by a large number of chromosomal changes, some of which by high-resolution comparative genomic hybridization, we have analyzed 77 tumors previously characterized for DNA copy number by a large number of chromosomal changes, some of which by high-resolution comparative genomic hybridization, we have analyzed 77 tumors previously characterized for DNA copy number by a large number of chromosomal changes, some of which by high-resolution comparative genomic hybridization, we have analyzed 77 tumors previously characterized for DNA copy number by a large number of chromosomal changes, some of which by high-resolution comparative genomic hybridization, we have analyzed 77 tumors previously characterized for DNA copy number by a large number of chromosomal changes, some of which by high-resolution comparative genomic hybridization, we have analyzed 77 tumors previously characterized for DNA copy number by a large number of chromosomal changes, some of which by high-resolution comparative genomic hybridization, we have analyzed 77 tumors previously characterized for DNA copy number by a large number of chromosomal changes, some of which by high-resolution comparative genomic hybridization, we have analyzed 77 tumors previously characterized for DNA copy number by a large number of chromosomal changes, some of which by high-resolution comparative genomic hybridization, we have analyzed 77 tumors previously characterized for DNA copy number by a large number of chromosomal changes, some of which by high-resolution comparative genomic hybridization, we have analyzed 77 tumors previously characterized for DNA copy number by a large number of chromosomal changes, some of which...
hybridization (CGH), the BRCA1 profile has been used to build a classifier that distinguished between sporadic and BRCA1 tumors (12).

In the present work, we have used a more refined genome-wide approach as high-resolution CGH (HR-CGH) that more precisely identify the somatic genetic changes associated to BRCA1, BRCA2, and non-BRCA1/BRCA2 germ line mutation carriers. Based on the distinctive profiles of tumor progression among BRCA1/BRCA2 mutation carriers, we have build a classifier that allow for the first time to differentiate BRCA1 and BRCA2 tumors and to define the differences and similarities of the BRCAX tumors compared with the BRCA1/BRCA2 classes.

**MATERIALS AND METHODS**

**Patients and Tumor Samples.** Patients were referred to the Spanish National Cancer Center from the Fundación Jimenez-Díaz in Madrid or from the Hospital San Pau in Barcelona for genetic studies. Seventy-seven breast cancer tumors from patients with BRCA1 germ line mutation (24 cases) and BRCA2 (19 cases) or without BRCA1/BRCA2 mutations (34 cases) were selected from families with at least three women affected with breast cancer, one of them <50 years old, or with ovarian cancer or with male breast cancer (13–15). The characteristics of these patients are summarized in the supplemental research data (Table 1). Mutation analysis of *BRCA1* and *BRCA2* genes of all the cases has been done as described previously (3, 15).

**CGH Analysis.** Genomic DNA was isolated from 4 × 10 μm sections of 77 paraffin-embedded tumors using the commercially available DNase mini Kit (Qiagen, Chatsworth, CA) according to the manufacturer’s recommendations. CGH was done as described previously with minor modifications. Briefly, tumor DNAs were labeled with spectrum green dUTP using the nick translation labeling kit (Vysis). Commercially available normal male DNA labeled with Texas Red dUTP was used as control reference (Vysis). Images of the hybridized normal male metaphases were evaluated with a digital image analysis system based on an Olympus AX60 epifluorescence microscope and a cooled CCD camera (Photometrics, Inc., Tucson, AZ) interfaced to a Cytovision Image Analysis System (Applied Imaging, Newcastle, United Kingdom). Three-color images were acquired from 10 to 15 metaphases per specimen. Calculation of green to red fluorescence ratios along the length of the chromosomes was done using the Cytovision system version 2.7 HR-CGH analysis software. The CGH profiles were compared with a dynamic standard reference interval based on an average of normal cases as described by Kirchhoff et al. (17). The mean ratio profile of each case with 95% confidence intervals was compared with the average ratio profile of the normal cases with similar 95% confidence intervals. Positive findings were considered those where the 95% confidence intervals of the patient profile and normal averaged profile did not overlap.

Hybridizations of normal breast tissue paraffin-embedded DNA versus reference female DNA were used as negative controls previous to the samples analysis. Because the tumor and reference DNA were not sex matched, the X and Y chromosomes were omitted from the analysis.

**Statistical Analysis of CGH Data.** A nonparametric Mann-Whitney *U* test was used to identify differences in the number of chromosomal gains and losses among the three patient groups. Differences in the frequency of involvement of individual chromosomal regions among the three familial breast cancer classes and among the BRCA1 and BRCA2 group with recurrent versus nonrecurrent mutations were tested with Fisher’s exact test. The indicated *P*s were not corrected for multiple testing and were calculated using the Stat POMELO tool available at http://pomelo.bioinfo.cnio.es (18).

**Building a Predictor Data Preprocessing.** Because a high number of variables better defined the tumor progression profile of each group, we selected 63 G-banded cytogenetic regions to refine the detected HR-CGH profile (Fig. 1). We have chosen as the most common minimal regions of involvement 50 regions including imbalances in at least 30% of the 28 BRCA1/BRCA2 cases used to build the predictor and with at least 3 cases defining the cytogenetic thresholds. To include the rest of the genome not fitting the previously defined criteria, we grouped the unselected areas on 13 chromosomal regions.

We selected the 63 previously defined chromosomal regions and some pathologic variables, such tumor grade and steroid receptor status, previously reported to behave differentially as potential predictors. A discrete value was assigned to each region for each tumor. A particular region had three possible discrete values, gain, loss, or no change. For the steroid receptor status, we recode them as positive (cases with >9% of positivity) or negative (10, 19). For grade, we impute the missing values to the most frequent variable value.

**Analyses.** We have used random forests to obtain a simple and good predictor of BRCA1 and BRCA2. Other approaches attempted without success are detailed in supporting

---

**Table 1**  Chromosomal regions with significant differences between BRCA1 and BRCA2 cases used to build the predictor (Fisher’s exact test, not corrected for multiple testing, *P* < 0.05)

<table>
<thead>
<tr>
<th>Regions</th>
<th>Percentage of affected tumors</th>
<th>BRCA1 (n = 15)</th>
<th>BRCA2 (n = 13)</th>
<th>BRCAX (n = 34)</th>
<th>BRCA1 vs. BRCA2</th>
<th>BRCA1 vs. BRCAX</th>
<th>BRCA2 vs. BRCAX</th>
</tr>
</thead>
<tbody>
<tr>
<td>8p11-12 Gain</td>
<td>13.3</td>
<td>69.2</td>
<td>29.41</td>
<td></td>
<td>0.005</td>
<td>NS</td>
<td>0.020</td>
</tr>
<tr>
<td>12q11-21 Gain</td>
<td>20</td>
<td>76.9</td>
<td>41.18</td>
<td></td>
<td>0.006</td>
<td>NS</td>
<td>0.048</td>
</tr>
<tr>
<td>2p11-21 Loss</td>
<td>0</td>
<td>30.7</td>
<td>8.8</td>
<td></td>
<td>0.042</td>
<td>NS</td>
<td>0.003</td>
</tr>
<tr>
<td>15q22-26 Loss</td>
<td>53.3</td>
<td>7.7</td>
<td>55.9</td>
<td></td>
<td>0.014</td>
<td>NS</td>
<td>0.003</td>
</tr>
<tr>
<td>18q Loss</td>
<td>60</td>
<td>15.4</td>
<td>55.9</td>
<td></td>
<td>0.020</td>
<td>NS</td>
<td>0.020</td>
</tr>
</tbody>
</table>

---

Clinical Cancer Research 1147

Downloaded from clincancerres.aacrjournals.org on October 15, 2017. © 2005 American Association for Cancer Research.
information. Random forests are essentially an ensemble of classification trees that can be used directly with several variables much larger than the number of samples and that often show very good predictive performance (20, 21).

Random forests return a prediction as the unweighed majority of predictions from a very large (e.g., 5,000) collection of classification trees (22); the important features about these trees are that each one is grown using a bootstrap sample of the data set and that, at each node, only a random subset of the original variables is examined. As part of a random forests run, we obtained measures of the importance of the variables. A large importance measure indicates that random permutation of that variable causes samples to be misclassified more often (hence, that variable is important). Although random forests provide a very good classification and return a ranking of the importance of variables, a predictor built from 66 predictor variables is cumbersome to use and very difficult to understand. Moreover, elimination of variables that do not contribute to the classification could actually result in an improved predictive performance (20, 21).

Random forests return a prediction as the unweighed majority of predictions from a very large (e.g., 5,000) collection of classification trees (22); the important features about these trees are that each one is grown using a bootstrap sample of the data set and that, at each node, only a random subset of the original variables is examined. As part of a random forests run, we obtained measures of the importance of the variables. A large importance measure indicates that random permutation of that variable causes samples to be misclassified more often (hence, that variable is important). Although random forests provide a very good classification and return a ranking of the importance of variables, a predictor built from 66 predictor variables is cumbersome to use and very difficult to understand. Moreover, elimination of variables that do not contribute to the classification could actually result in an improved predictive performance. Our approach to variable selection is very similar to that of Svetnik et al. (6); essentially, we iteratively eliminate the variable with a lower importance measure until the out-of-bag estimate of prediction error becomes larger than the out-of-bag estimate of prediction error of either the previous model or the original model. However, these out-of-bag estimates are biased down because of the variable selection process (23–25). We have thus estimated the prediction error of our simplification procedure using a leave-one-out cross-validation of the complete variable selection process; in other words, excluding each observation in turn, we run the complete variable selection procedure and then predict the left-out observation. This, therefore, yields a honest, almost unbiased estimate of prediction error, because the error rate is computed from observations that were never used for the simplification and building of the forest that is predicting an observation. For the analyses, we have used the random forests package for R (by A. Liaw and M. Wiener) that uses the Fortran random forests code by L. Breiman and A. Cutler. The R code for the variable selection is available from authors on request. The predictor will be available at http://bioinfo.cnio.es/data.

Methylation. A methylation-specific PCR assay was used to distinguish unmethylated alleles from methylated alleles of BRCA1 based on sequence changes produced by treating DNA with bisulfite, which converts unmethylated (but not methylated) cytosines to uracil followed by a PCR assay involving primers designed for either methylated or unmethylated DNA (26). The methylated product is 75 bp long and the unmethylated one is 86 bp. DNA from normal lymphocytes was used as a negative control, and in vitro methylated DNA was used as a positive control (Fig. 3E).

---

RESULTS

Overall Genetic Changes in the Three Groups of Familial Breast Cancer Tumors. The HR-CGH profiles of 15 BRCA1 (cases 1-15; supplemental research data; Table 1), 13 BRCA2 (cases 25-37; supplemental research data; Table 1), and 34 BRCAX breast carcinomas were analyzed at the 63 selected chromosomal regions (Fig. 1; see Building a Predictor). The selection of these regions was possible due to the improved sensitivity and specificity of the HR-CGH software (27, 28).

For example, we could effectively discriminate changes affecting different regions within the short arm of chromosome 8 (Fig. 2). In fact, loss of 8p21-23 region was observed in >60% of the cases in the three groups, whereas gain at 8p11-12 band was present in 69% of the BRCA2 cases and in <30% of the BRCA1 and BRCAX tumors (Fig. 1).

The mean ± SE number of genetic changes was similar among the three groups, 32.3 ± 3, 31.6 ± 3.2, and 26.5 ± 1.7 for BRCA1, BRCA2, and BRCAX, respectively. However, a higher number of DNA losses in the BRCA1 (19 ± 1.9) and BRCAX (16.53 ± 1.43) cases compared with the BRCA2 group (13.8 ± 2.7) and a lower number of gains in the BRCA1 (13.3 ± 1.9) and BRCAX tumors (10 ± 1.8) compared with the BRCA2 (17.7 ± 1.8) was observed. These differences were not statistically significant (P > 0.05, Mann-Whitney U test), except when we compare the number of gains between BRCA1 and BRCAX tumors (P = 0.0031).

The frequency and distribution of gains and losses of each group are shown in Fig. 1. In all groups, the most common changes, observed in >50% of the cases, were gain at 1q and losses at 8p21-23 and 11q22-25. Nevertheless, intergroup comparison revealed several chromosomal regions with significant differences in the frequency of involvement (Table 1). These data suggest a different pattern of genomic progression for BRCA1 and BRCAX cases.

Building a Predictor for Classifying BRCA1 and BRCA2 Tumors. The CGH profiles of a series of 15 BRCA1 (cases 1-15; supplemental research data; Table 1) and 13 BRCA2 (cases 25-37; supplemental research data; Table 1) breast carcinomas were used to develop using random forests (see Building a Predictor) a molecular classifier, which assigns the probability of belonging to the BRCA1 class >50%, was obtained in 12 of 15 (80%) BRCA1 tumors and 5 of 13 (38%) of the BRCAX tumors (Fig. 3A and B). To further validate the predictor, we run a new set of tumors. Nine BRCA1 (cases 16-24; supplemental research data; Table 1) and six BRCA2 (cases 38-43; Supplemental Research Data; Table 1) breast carcinomas were analyzed. The classification scores for this validation set are depicted in Fig. 3C. All the cases, except for one case of each group, were properly classified. The probability of being a BRCA1 tumor among the 43 BRCA1/BRCAX tumors analyzed was estimated, with an accuracy of 76.7%, a sensitivity of 83.33%, and a specificity of 68.42%.

Figure 4 depicts the a posteriori probability of these regions as they are used in the predictor. These histograms represent the probability of that a tumor is BRCA1 or BRCAX given the particular aberrations observed in the tumor and provide inside of the specific pattern of tumor progression. Therefore, the BRCA1 profile yielding the best classification performance is as follows: no change or gain at 2p11-21; no change or loss at 8p11-12 and 12q11-21; losses at 15q22-26, 18p, and 18q; and a negative estrogen receptor status. The BRCA2 profile is as follows: no change or loss at 2p11-21; no change or gain at 8p11-12, 15q22-26, 18p, and 18q; gain at 12q11-21; and a positive estrogen receptor status.

To confirm that the selected variables not exhibit statistically significant differences between the BRCA1/BRCAX carriers with or without a recurrent mutation, we compared the profile of BRCA1 tumors del185AG (six cases) and BRCA2 tumors 3036delAAAC (three cases) versus the rest of the tumors without recurrent mutations. Only slight differences between these groups were observed (Table 2). The only variable included in the predictor and with a significant difference among the BRCA1 cases with or without the del185AG recurrent mutation was the loss of 18q.

Fig. 2 Examples of HR-CGH profiles of chromosome 8 to illustrate how the genomic imbalances affect different regions within the short arm of the chromosome 8. Loss of the 8p21-23 bands are observed in all the cases. However, the 8p11-12 region shows a high heterogeneity. Gain of this region is observed on cases 27, 33 and 56 no changes on cases 44 and 5, and loss on case 8.
Effect of the BRCA1 Promoter Hypermethylation on the Somatic Genetic Profile and on the Tumoral Phenotype of BRCAX and BRCA2 Tumors. Looking for similarities and differences on the BRCAX tumor profiles, we analyzed these cases using the BRCA1/BRCA2 predictor. In 76.5% of the cases, a probability of being a BRCA1 case >50% was obtained (Fig. 3D).

To determine if the high frequency of the BRCA1 pattern of tumor progression among the BRCAX and BRCA2 cases could be due to the BRCA1 gene silencing by epigenetic factors on these two groups, we analyzed the BRCA1 promoter region for aberrant methylation. Testing (in a blinded fashion) of all specimens of BRCAX group from our study indicated that 44% of the BRCAX tumors (15 of 34 BRCAX cases) had hypermethylation of the BRCA1 promoter region. When we consider the predictor, we observed that all the cases with this epigenetic event were assigned to the BRCA1 class with a probability of >50%. Interestingly, 84.6% (11 of 13) of the cases assigned to the BRCA1 class with a probability of >80% had an aberrant methylation of the BRCA1 promoter (Fig. 3D).

These results leave out whether the BRCA1 promoter methylation events could also be occurring as germ line events. We analyzed the methylation status of the BRCA1 promoter on eight DNA obtained from five normal breast tissues and three peripheral blood leukocytes from five BRCAX cases. All these BRCAX patients showed BRCA1 promoter hypermethylation on the tumor sample (two of these cases, cases 68 and 73, are included on this study). All the DNA analyzed samples were found unmethylated. These data suggest that methylation is not a heritable trait on these families but a frequent event during tumor progression.

We also tested all the BRCA2 tumors and we found that 31.6% (6 of 19) of them were methylated. Among the methylated cases, only case 34 was misclassified as BRCA1. This case has also been classified by immunohistochemical and cDNA expression studies as a BRCA1 tumor (data not shown).

Statistical analysis comparing the family history, the grade, and the steroid receptor status of methylated versus unmethylated BRCA2 and BRCAX tumors did not show significant differences (Fisher’s exact test, $P > 0.05$).

Fig. 3 BRCA1 classification score for each individual tumor. A and B, BRCA1 probability, calculated using the leave-one-out cross-validation of the complete variable selection process, of the BRCA-positive mutation cases used to build the predictor. C and D, BRCA1 probability of the BRCA-positive mutation cases of the validation set and the BRCAX cases estimated with the predictor. A high score indicates a high probability to belong to the BRCA1 class. #, case 34, this case cluster with the BRCA1 tumors in an ongoing expression profile study (D): methylation analysis of the BRCA1 promoter region in tumor samples from three patients with breast cancer. U, unmethylated alleles; M, methylated alleles. Tumors 27 and 55 present BRCA1 promoter hypermethylation. In vitro methylated DNA is used as a positive control for methylated alleles.
DISCUSSION

Using HR-CGH, we have detected a higher number of somatic genetic changes accumulated during breast cancer development and progression in tumors of BRCA1/BRCA2 mutation carriers than in previous studies (11, 12). This novel software, due to its 2- to 3-fold improved sensitivity, has been proven to be a good system to detect chromosomal changes present in at least 50% of the analyzed cells (27, 28). This is an important feature when primary tumors that always have a variable component of normal tissue are analyzed. We found that the most frequent genetic changes, as gains at 1q and losses at 8p21-23 and 11q22-25, were present in >50% of the cases in all three patient groups. Because these abnormalities have been described also in previously reported sporadic tumors (29), they can represent a core of abnormalities common to breast cancer.

Although a similar frequency of involvement at the chromosomal regions reported previously to be significantly different between BRCA1/BRCA2 tumors and sporadic breast carcinomas were observed (supplemental research data; Table 2; refs. 11, 12), we did not find the low frequency of 16q loss and the high presence of 20q gain reported previously on the BRCA2 tumors (11). These discrepancies may be due to the differences among the studied populations, with 80% of the cases of the previously studied nordic population carrying a 999del5 BRCA2 mutation. In addition, our intergroup comparisons revealed significant differences, as gains of the 15q11-21 region, which occur in 66.6% of the BRCA2 tumors with the 3036delAAAC mutation and in none of the nonrecurrent tumors (Table 2). These facts suggest that different mutations may slightly modify the profile of somatic changes along the tumor progression.

A Classifier Based on Distinct Patterns of Tumor Progression among the BRCA Mutation Carriers. Recently, the BRCA1 profile has been used to build a classifier to prescreen high-risk patients. This predictor distinguished with a high accuracy (84%) among BRCA1 and sporadic tumors. The regions used in this classifier are located at 3p, 3q, and 5q, but these

<table>
<thead>
<tr>
<th>Regions</th>
<th>Percentage of tumors affected</th>
<th>BRCA1 recurrent mutation (del185AG), n = 6</th>
<th>BRCA1 non-recurrent mutation, n = 9</th>
<th>BRCA2 recurrent mutation (3036delAAAC), n = 3</th>
<th>BRCA2 non-recurrent mutation, n = 10</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>2q35-37 Gain</td>
<td>50</td>
<td>0</td>
<td>33.3</td>
<td>0</td>
<td></td>
<td>0.046</td>
</tr>
<tr>
<td>15q11-21 Gain</td>
<td>0</td>
<td>0</td>
<td>66.6</td>
<td>0</td>
<td></td>
<td>NS</td>
</tr>
<tr>
<td>14q11-23 Loss</td>
<td>50</td>
<td>0</td>
<td>33.3</td>
<td>10</td>
<td></td>
<td>0.044</td>
</tr>
<tr>
<td>18q Loss</td>
<td>16.6</td>
<td>88.9</td>
<td>0</td>
<td>20</td>
<td></td>
<td>0.010</td>
</tr>
</tbody>
</table>

Table 2 Chromosomal regions with a significant difference between BRCA1/BRC2 mutation carriers used to build the predictor with recurrent and nonrecurrent mutations (Fisher’s exact test, not corrected for multiple testing, P < 0.05)

Fig. 4 A posteriori probabilities for the different variables used in the classifier. Distribution of the imbalances for the seven selected variables. The probability of a tumor to be classified as BRCA1 or BRCA2 given the particular aberration observed at a variable. For example, for a tumor with 8p11-12 gain, the a posteriori distribution indicates that the probability that the tumor belongs to the BRCA2 class is high (82%) and that is low (18%) for the BRCA1 class.
regions did not show significantly different frequencies of imbalances when BRCA2 and BRCA1 tumors are compared (12).

In our study, several chromosomal regions seemed to be preferentially involved in the BRCA1 cases and yet others in the BRCA2 cases. By using random forests, we have developed for the first time a simple and good molecular classifier based on the genomic pattern of somatic changes accumulated during the tumor development that can be used to distinguish between BRCA1 and BRCA2 mutation carriers with an accuracy of 76.7%. Previously, this supervised classification method has been used for building, with similar accuracy to other methods, good prediction models (30). We obtained a high sensitivity (83.3%) identifying the BRCA1 cases, but further studies to refine the profiles, such as CGH array and a detailed analysis of the status of the BRCA1 gene among the BRCA2 tumors, are needed to improve the specificity (68.4%). The variables used to define the distinctive patterns of tumor progression in our classifier are six somatic genetic imbalances (2p11-21, 8p11-12, 12q11-21, 15q22-26, 18p, and 18q) and the estrogen receptor status. Some of these regions as 12q and 18q have also been found significantly more frequent lost in the BRCA1 cases than in the sporadic tumors (11, 31). Interestingly, we have shown previously using a familial breast cancer tissue microarray that the expression of the BCL2 gene located at 18q21.3 is more frequently lost in the BRCA1 cases (89.5%) than in the BRCA2 (57.1%) and BRCA cases (44%; ref. 10). This is in good agreement with the higher frequency of 18q loss observed by HR-CGH in the BRCA1 (60%) compared with the BRCA2 (15.4%) or the BRCA cases (56%).

We attempted several approaches to build a classifier to distinguish the three classes (see supporting information). However, we could not define a somatic genetic profile that distinguished between the BRCA1/BRCA2 mutation carriers and the BRCA cases. Therefore, we assumed that the pattern of tumor progression in our BRCA cases is not quite different to the BRCA1 and BRCA2 mutation carriers, and we analyzed the BRCA cases with the classifier. Seventy-six percent of the cases were classified as BRCA1, with a probability of >50%, confirming the somatic genetic similarities that we found by previous statistical analysis (Table 1; Fig. 3D).

Effect of DNA Methylation on the Somatic Genetic Profile of the BRCA2 and BRCA Cases. Because BRCA1 promoter hypermethylation was shown to silence the BRCA1 gene in ~13% of sporadic breast carcinomas (26, 32), we studied this phenomenon as a mechanism of BRCA1 silencing in the BRCA and the BRCA2 cases. We found this epigenetic event in 44% of the BRCA (15 of 34) and in 31% of the BRCA2 cases (6 of 19). Among the BRCA2 cases, 84.6% of the cases (11 of 13) assigned with a probability of >80% to the BRCA1 class were BRCA1 promoter hypermethylated (Fig. 3D). This epigenetic phenomenon was not observed at the germ line level.

BRCA1 does indeed seem to be implicated in sporadic breast and ovarian tumors. Avermrent methylation of BRCA1 seems to be a crucial player in the development of these tumors. Already, some genomic changes as MYC amplification have been found in a significantly higher proportion on sporadic tumors with BRCA1 dysfunction (33). In addition, it has been suggested that BRCA1 hypermethylation could influence the phenotype of the tumor on a similar way to a genetic mutation (34). Our study suggest for the first time that this epigenetic event could also influence the tumor progression genotype on a similar way to a genetic mutation in most of the BRCA cases. Further clinical studies will be needed to know if the selection of these patients is clinically relevant.

We have not found statistically significant differences in the family history or in the steroid receptor status between methylated and unmethylated tumors. However, in the analysis of a tissue array of BRCA cases, a statistically significant higher percentage of p53 and Ki-67 positivity has been observed in the BRCA1 methylated tumors, indicating a higher proliferation rate and suggesting a more aggressive behavior of these cases.7

The identification of a method to successfully subdivide the histopathologically heterogeneous group of BRCA families into recognizable groups could be of considerable value to further genetic analysis. Previously, global gene expression analysis have discover novel classes among BRCA tumors (35). Our findings illustrate that also at the genomic and epigenetic levels BRCA families can be grouped into homogeneous subsets, thereby potentially increasing the power of genetic linkage analysis. Moreover, the complete sequencing of the BRCA genes is very difficult due to the large number of exons of these genes; this tool could allow us to select cases among the BRCA patients for further genetic studies as large deletions or rearrangement analysis. At the therapeutic level, these data could point to the inclusion of demethylating agents into the treatment protocols for familial breast cancer.

Among the six BRCA2 cases with BRCA1 promoter hypermethylation, five cases were correctly assigned to the BRCA2 class and only case 34 was assigned to the BRCA1 class. This case also clusters with the BRCA1 tumors in ongoing expression studies from our laboratory (Fig. 3B). These data suggest that hypermethylation of the BRCA1 gene in the majority of the BRCA2 tumors could be a later event in the genetic progression, when the specific genetic pathway is yet established.

In summary, we have build a simple molecular classifier that distinguish with an accuracy of 76.7% between BRCA1 and BRCA2 tumors. This BRCA1 pattern of tumor progression is frequently found among the BRCA2 tumors and, in more than half of these cases, is associated with BRCA1 promoter hypermethylation, suggesting that this event can modify the profile of tumor development in most of the BRCA2 tumors. Further studies to better establish the genomic integrity of the BRCA1 gene are guaranteed. Finally, the presence of this epigenetic event among the BRCA2 cases is also frequent but not associated in the majority of the cases with a BRCA1 profile.

ACKNOWLEDGMENTS

We thank the Spanish Tumor Bank Network and the Immunohisto- logical Unit of the Spanish National Cancer Center for technical support. Adele Cutler and Andy Liaw for comments on variable selection for random forests, and Ellen Allacid for assistance in preparing the article.

7Honrado et al., unpublished data.
REFERENCES


A Predictor Based on the Somatic Genomic Changes of the BRCA1/BRCA2 Breast Cancer Tumors Identifies the Non-BRCA1/BRCA2 Tumors with BRCA1 Promoter Hypermethylation

Sara Alvarez, Ramon Diaz-Uriarte, Ana Osorio, et al.


Updated version
Access the most recent version of this article at:
http://clincancerres.aacrjournals.org/content/11/3/1146

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.