Gene Expression Profiling after Radiation-Induced DNA Damage Is Strongly Predictive of *BRCA1* Mutation Carrier Status

Zsofia Kote-Jarai,[1] Richard D. Williams,[2] Nicola Cattini,[3] Maria Copeland,[1] Ian Giddings,[3] Richard Wooster,[3] Robert H. tePoele,[5] Paul Workman,[5] Barry Gusterson,[6] John Peacock,[4] Gerald Gui,[7] Colin Campbell,[8] and Ros Eeles[1,7]

[1]Translational Cancer Genetics, [2]Paediatric Oncology, [3]Molecular Carcinogenesis, and [4]Academic Unit of Radiotherapy, The Institute of Cancer Research, Sutton, Surrey; [5]Cancer Research UK Centre for Cancer Therapeutics, Sutton, Surrey; [6]Division of Gene Regulation and Molecular Pathology, University of Glasgow, Glasgow; [7]Royal Marsden NHS Trust, London; and [8]Computational Intelligence Unit, Bristol University, Bristol, United Kingdom

ABSTRACT

Purpose: **The impact of the presence of a germ-line *BRCA1* mutation on gene expression in normal breast fibroblasts after radiation-induced DNA damage has been investigated.**

Experimental Design: **High-density cDNA microarray technology was used to identify differential responses to DNA damage in fibroblasts from nine heterozygous *BRCA1* mutation carriers compared with five control samples without personal or family history of any cancer. Fibroblast cultures were irradiated, and their expression profile was compared using intensity ratios of the cDNA microarrays representing 5603 IMAGE clones.**

Results: **Class comparison and class prediction analysis has shown that *BRCA1* mutation carriers can be distinguished from controls with high probability (~85%). Significance analysis of microarrays and the support vector machine classifier identified gene sets that discriminate the samples according to their mutation status. These include genes already known to interact with *BRCA1* such as *CDKN1B, ATR,* and *RAD51*.**

Conclusions: **The results of this initial study suggest that normal cells from heterozygous *BRCA1* mutation carriers display a different gene expression profile from controls in response to DNA damage. Adaptations of this pilot result to other cell types could result in the development of a functional assay for *BRCA1* mutation status.**

INTRODUCTION

It is estimated that 5–10% of breast cancer patients develop the disease because of the presence of a mutation in a breast cancer predisposition gene (1). A significant proportion of this population (just under a half) has a mutation in one of the known breast cancer predisposition genes, *BRCA1* or *BRCA2*. Besides the obviously disease-causing deleterious mutations, very often small alterations caused by a single base change (missense mutations) are found in these genes. Their functional effects are usually unknown and so they are termed variants of uncertain significance. Some of these variants of uncertain significance could also have a role in breast cancer predisposition, but it is not currently possible to establish their disease-causing effect. The available diagnostic tests for mutation analysis of *BRCA1/2* are time and labor intensive, expensive, and none of them allow the identification of all types of mutation. Our aim in this preliminary study was to investigate if gene expression profiling could be used to distinguish between heterozygous *BRCA1* mutation carriers and control samples from reduction mammoplasties with a very low chance of the presence of a *BRCA1* mutation. If this were possible, then this would have the potential to develop a method to identify the presence of a *BRCA1* defect.

The *BRCA1* gene encodes a large nuclear protein (220 kDa) that has multiple possible functions, including DNA damage signaling, DNA repair, growth inhibition, and transcription regulation (2, 3). The involvement of BRCA1 in transcriptional regulation has been revealed in several studies showing direct interaction with other transcriptional activators and repressors such as STAT1, MYC, TP53, ZBRK1, and CDKN1B (4) It appears that BRCA1 is phosphorylated as a response to various DNA damaging agents by kinases such as CHEK2, ATM, and ATR, and this phosphorylation results in changes in its protein-protein interactions, which then can lead to regulatory changes in the expression of various target genes (5).

Microarray studies have been shown to be of great value in understanding the molecular biology of many diseases, and they have been successfully used to classify various tumors based on their clinical phenotype or genetic background (reviewed in Ref. 6). This approach has enabled classification of tumors and division into prognostic groups on the basis of their global patterns of gene expression. One of the first of these studies classified myeloid and lymphoblastic leukemias using a class discovery procedure (7) demonstrating the feasibility of this technique. Similarly, it has been shown that breast epithelial cell lines have an expression profile that is distinct from breast tumors and that breast tumor samples can be divided into

Received 7/22/03; revised 10/6/03; accepted 10/13/03.
Grant support: B. G. was funded by Breakthrough Breast Cancer. Tissue collection was partly funded by a grant from the United States Department of Defense. This work also had support from Cancer Research United Kingdom and The Institute of Cancer Research.
The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
Supplementary data available at http://www.icr.ac.uk/array/array.html.
Requests for reprints: Zsofia Kote-Jarai, Translational Cancer Genetics, The Institute of Cancer Research, Sutton, Surrey, United Kingdom. Phone: 0208-661-3897; Fax: 0208-770-1489; E-mail:Zsofia.Kote-Jarai@icr.ac.uk.

different clinical outcome groups based on clustering algorithms using cDNA microarray data (8). Tumor classification using gene expression profiling has also been successfully used to distinguish breast cancer subclasses with important clinical implications (9) and to discriminate breast tumors on the basis of estrogen receptor status and lymph node status (10). *BRCA1* and *BRCA2* mutation status have also been shown to influence the gene expression profile of breast tumors (11). Various studies provide evidence that microarray analysis can be used to identify gene expression changes induced by drug treatments and radiation (12). Our aim in the present study was to assess the potential of gene expression profiling in discriminating between normal cells with *BRCA1* mutations and controls after induced DNA damage.

MATERIALS AND METHODS

Samples and cDNA Microarrays. Individuals who are heterozygous for *BRCA1* germ-line mutations were identified from the *BRCA1* and *BRCA2* predictive testing program in the Royal Marsden Hospital NHS Trust/The Institute of Cancer Research, Cancer Genetics Carrier Clinic. Prophylactic mastectomy specimens were collected from nine individuals with a germ-line *BRCA1* mutation, and short-term breast fibroblast cell lines were established. The mutation status of these samples is listed in Table 1. For our five control samples, we used similar cell lines established from reduction mammoplasty specimens. The healthy women undergoing reduction mammoplastic surgery had no family or personal history of any cancer. From the tissue specimens, short-term primary fibroblast cultures were established by standard methods. Confluent cells were irradiated with 15 Gy at a high dose rate (1 Gy/min) using a ^{60}Co source. Cells were maintained at 37 °C during irradiation. Total RNA was extracted from the cells 1 h after the treatment using an RNeasy kit (Qiagen, Valencia, CA). Reference RNA was pooled from three cancer cell lines (MDA-MB-231, A549, and DLD-1). RNA samples were reverse transcribed into cDNA fluorescently labeled with Cy3 and Cy5. Equal amounts (20 μg) of sample and reference cDNA were mixed and hybridized onto the microarrays. We used high density cDNA microarrays manufactured by the Section of Molecular Carcinogenesis, The Institute of Cancer Research, representing 5603 IMAGE cDNA clones. Details of the clone set and hybridization conditions are available online.[9] Image acquisition and analysis were performed using GenePix 4000B scanner and GenePix Pro 4 (Axon Instruments, Inc., Union City, CA), respectively. Signal intensities for Cy3 and Cy5 channels were normalized based on the assumption that on average the genes have the same overall expression level in the experimental and the reference sample.

Table 1 The range of mutations in the *BRCA1* gene in the mutation carrier samples

Mutation
185delAG
185delAG
546 T>G nonsense
1294del40
1623del5
2371delTG
2732insT
4184del4
5382insC

Data Analysis. Of the 5603 cDNA clones represented on the arrays, 1967 clones were filtered out and used in the consequent analysis by the criteria that these data points showed signal intensity in at least one of the channels (red or green) 2-fold above background in a minimum of 12 of our 14 samples.

We analyzed our data first using supervised methods of class comparison and class prediction (13). For class comparison, we have used a recently described statistical method, significance analysis of microarrays (SAM; Ref. 14), to identify genes in which the levels of expression significantly correlate with the presence of a *BRCA1* mutation. SAM uses *t* test statistics to compare the expression level of each gene in cells from *BRCA1* mutation carriers and control individuals. Our aim was to find if there is a statistically significant difference between carriers and noncarriers in the expression level of a number of genes and to identify the genes that are most likely to show consistent differences. SAM, using repeated permutation, computes a score for each gene that measures the strength of its correlation with the presence of a mutation. A threshold value can be chosen, which will determine the false positive rate (false discovery rate) as estimated by repeated permutation and counting the number of genes selected as significant.

In an independent analysis to evaluate the predictive accuracy, we have used a support vector machine (SVM) classifier (15) with a linear kernel and feature ranking using the Fisher score. Predictive accuracy was determined across a range starting with all features (data from 1967 cDNA clones as mentioned above) followed by successive removal of the least discriminative feature (according to the Fisher score) through to the top two discriminative features. The test accuracy was evaluated using leave-one-out (LOO) cross-validation (15). To provide a baseline for comparison (for null prediction), we used a permutation test, *i.e.*, the class labels were randomly shuffled 100 times, and the LOO test error was evaluated.

For hierarchical cluster analysis and principal component analysis, the Genesis software package (16) was used with the expression data of the significant genes identified by SAM and the discriminatory features identified by SVM as input.

RESULTS AND DISCUSSION

We have analyzed the expression profiles of nine normal heterozygous *BRCA1* mutation carrier breast fibroblast samples and compared these to the profiles of five reduction mammoplasty fibroblast samples with a very low probability of the presence of a *BRCA1* mutation (the controls). All these samples were short-term primary cultures, and they were irradiated (15 Gy) to induce sublethal DNA damage. In a preliminary study, we also included nonirradiated samples; however, the analysis

[9] Internet address: http://www.icr.ac.uk/array/array.html.

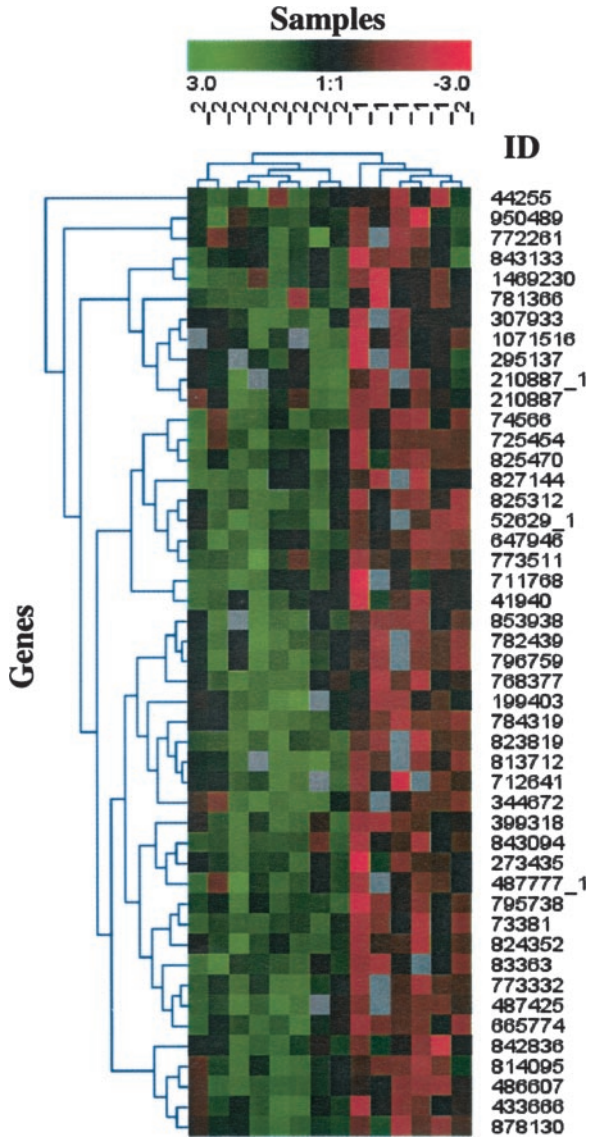

Fig. 1 Hierarchical cluster analysis of 14 normal breast fibroblast samples, 9 with *BRCA1* mutation (2), and 5 control reduction mammoplasty specimens (1). Here, we show the clustering dendogram based on expression data of the 47 significant genes identified by significance analysis of microarray with a 10% false discovery rate (see "Results"). One *BRCA1* mutation carrier sample clusters with the controls on the right hand side and this sample has the 5382insC mutation.

of this set showed no significant gene expression differences (data not shown), and we have therefore concluded to focus on the comparison of irradiated samples only. RNA samples were collected 1 h after the treatment for expression profiling. We also investigated the time course of expression profile changes on a smaller sample size at 15 min, 1 h, and 3 h after irradiation and found the most robust changes occurred at the 1-h time point and therefore chose this time point for the analysis of all samples. On the microarrays, 6503 cDNA clones were represented. Image analysis filtered the data as specified in "Methods and Materials," and expression ratios of 1967 genes were included in the final data processing. We then carried out a class comparison analysis using SAM. This statistical method showed that there are significant differences between the two classes, *BRCA1* mutation carriers and controls, and ranked the genes according to their contribution to the separation of the two classes.

Choosing a 20% false discovery rate, SAM called 113 genes significant; with a more stringent 10 or 5% false discovery rate, 47 and 22 genes, respectively, were selected as significant discriminatory features (supplementary data: Table A). Sometimes a smaller expression level change is more significant, as is shown in Table A, because the more consistent small changes mean a more persistent difference between the classes.

We then used these gene sets for unsupervised hierarchical clustering. The clustering dendogram based on 47 significant genes (10% false discovery rate) is shown in Fig. 1. Using the larger or smaller (SAM-113 or SAM-22) gene sets showed a very similar result (data not shown). This illustrates the clear separation of the two classes with the exception of one sample, which although from a mutation carrier, clusters with the controls at the extreme right hand site of the dendrogram. This individual has a substantial family history of breast cancer, so in this case, this is very likely to be a breast cancer causing mutation. However, there could be some other explanations for the clustering of this sample with the noncarriers such as an effect on gene expression levels by modifier genes in this particular family or the fact that this mutation leads to a truncation at the extreme COOH-terminal of the BRCA1 protein, which still maintains its transcription regulation function. This hypothesis, however, has to be tested with additional samples with this particular mutation.

Principal component analysis also separated the two classes with the input data of the gene sets as above. The plot shown in Fig. 2 is the result of this analysis using 47 significant genes (very similar results were obtained by the larger or smaller sets of 113 and 22 genes). This shows that samples from *BRCA1* mutation carriers and controls are well separated with the exception of the same sample mentioned above (5328insC).

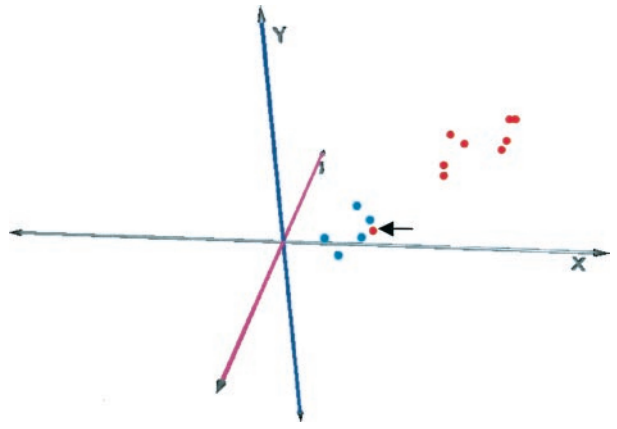

Fig. 2 Principal component analysis by experiment. The *red dots* represent the *BRCA1* mutation carrier samples, and the *blue dots* are the controls. For this analysis, we have used the data from the 47 significant genes identified by significance analysis of microarray. The outlier mutation carrier sample (*arrow*) has the 5382insC mutation.

Table 2 List of the 79 significant genes that distinguish between *BRCA1* mutation carriers and controls found by both SAM (significance analysis of microarrays) and supervised vector machine learning

Gene ID	Gene name	Gene symbol	Cytoband	SAM score	Fischer score	Fold change
823819	Small acidic protein	SMAP	11p15.1	4.7852	5.8424	1.6983
843133	Mitochondrial ribosomal protein S27	MRPS27	5q13.1	3.9985	3.7419	1.4057
827144	Mitochondrial ribosomal protein L19	MRPL19	2q11.1–q11.2	3.8322	3.3732	1.7955
814095	Leukotriene A4 hydrolase			3.7927	3.4539	1.9715
1469230	Cytochrome c oxidase subunit VIII			3.6587	3.1063	1.5867
773332	Integrin, α E			3.6425	2.2973	1.4971
795738	Guanine nucleotide binding protein 10	GNG10	9q32	3.4786	2.5591	1.6732
307933	NADH dehydrogenase (ubiquinone) 1 β subcomplex,	NDUFB5	3q27.1	3.3704	2.4379	1.5740
784319	Homosapiens cDNA FLJ11904 fis,			3.2585	2.7129	1.6436
73381	General transcription factor IIA, 2 (12kD subunit)	GTF2A2	15q21.3	3.2423	2.2717	1.3606
486607	Ubiquitin-conjugating enzyme E2E 1			3.2349	2.3016	1.7839
950489	Superoxide dismutase 1, soluble (amyotrophic lateral sclerosis 1	SOD1	21q22.11	3.2187	2.0813	1.4477
83363	Protein-L-isoaspartate (D-aspartate) O-methyltransferase	PCMT1	6q24–q25	3.1101	2.1273	1.5285
647946	Cyclin-dependent kinase inhibitor 1B (p27, Kip1)	CDKN1B	12p13.1–p12	3.0815	2.2875	1.6931
399318	Casein kinase 2, β polypeptide	CSNK2B	6p21.3	3.0208	2.1933	1.3881
825312	ATP synthase, H+ transporting, mitochondrial F0 complex,	ATP5J	21q21.1	3.0144	2.1999	1.4922
813712	ATP synthase, H+ transporting,	ATP5F1	1p13.1	2.9841	2.4326	1.9074
843094	Ubiquitin-like 1 (sentrin)	UBL1	2q33	2.9766	2.0640	1.3704
487425	Centrin, EF-hand protein, 3 (CDC31 homologue, yeast)	CETN3	5q14.3	2.9655	1.4333	2.0350
1071516	Endothelin 2	EDN2	1p34	2.9196	1.9463	1.6805
853938	Dynein, cytoplasmic, light polypeptide	DNCL1	12q24.23	2.9080	1.9238	1.8573
273435	V-yes-1 Yamaguchi sarcoma viral oncogene homologue 1	YES1	18p11.31–p11.21	2.9002	1.6008	1.5459
44255	Mitochondrial ribosomal protein L3	MRPL3	3q21–q23	2.8751	2.1402	1.2560
74566	Exportin 1 (CRM1 homologue, yeast)	XPO1	2p16	2.8219	1.7335	1.8171
842836	Serine (or cysteine) proteinase inhibitor, clade B (ovalbumin),	SERPINB1	6p25	2.7929	1.4809	1.5577
782439	ATP synthase, H+ transporting, mitochondrial F0 complex, subunit e	ATP5I	4p16.3	2.7881	2.1030	1.6593
796759	Voltage-dependent anion channel 3			2.7497	2.1154	1.7171
772261	Mitogen-activated protein kinase 14	MAPK14	6p21.3–p21.2	2.7265	1.0913	1.2858
199403	Lectin, galactoside-binding, soluble, 8 (galectin 8)	LGALS8	1q42–q43	2.7249	1.6844	1.6818
433666	Ring-box 1	RBX1	22q13.2	2.7234	1.7511	1.3774
781366	RNA binding motif protein 8A			2.6812	1.6614	1.4856
768377	Activity-dependent neuroprotector	ADNP	20q13.13–q13.2	2.6720	1.5629	1.4684
725454	CDC28 protein kinase 2	CKS2	9q22	2.6713	1.6524	2.4104
295137	Aspartoacylase (aminoacylase 2, Canavan disease)	ASPA	17pter-p13	2.6579	1.5413	1.4266
825470	Topoisomerase (DNA) II α (170kD)	TOP2A	17q21–q22	2.6457	1.4840	1.6589
824352	RAD23 homologue B (*S. cerevisiae*)	RAD23B	9q31.2	2.6148	1.5752	1.9017
665774	ESTs, Weakly similar to translation initiation factor eIF4E	EIF4E	4q21–q25	2.5891	1.5695	1.8769
210887	Suppression of tumorigenicity 13 (colon carcinoma)	ST13	22q13.2	2.5795	1.6472	1.4489
773511	Hook2 protein	HOOK2	19p13.12	2.5617	1.5728	1.7035
41940	Chromosome X open reading frame 1	CXorf1	Xq27.3	2.5111	1.3006	1.3378
878130	SMT3 suppressor of mif two 3 homologue 2 (yeast)	SMT3H2	17q25.2	2.5086	1.2273	1.5061
757144	Heterogeneous nuclear ribonucleoprotein A3			2.4734	1.3023	1.5109
767277	Peptidyl prolyl isomerase H (cyclophilin H)	PPIH	1p34.1	2.4453	1.4105	1.5530
739993	Brain and reproductive organ-expressed (TNFRSF1A modulator)	BRE	2p23.3	2.3907	1.4007	1.4335
80374	Pyruvate dehydrogenase (lipoamide) α 1			2.3775	1.6878	1.2103
788109	Ataxia telangiectasia and Rad3 related	ATR		2.3700	1.3132	1.3205
795936	Translin	TSN	2q21.1	2.3625	1.4549	1.5836
75415	Histidine triad nucleotide binding protein 1	HINT1	5q31.2	2.3432	1.1386	1.3855
327350	Heterogeneous nuclear ribonucleoprotein A2/B1	HNRPA2B1	7p15	2.3366	1.3176	1.6091
133637	Protein kinase, DNA-activated, catalytic polypeptide	PRKDC	8q11	2.3353	1.3240	1.4405
489489	Lamin B receptor	LBR	1q42.1	2.3202	1.0545	1.8520
1012799	Histone deacetylase 1	HDAC1	1p34	2.3098	1.9998	1.3305
130242	Cyclin-dependent kinase 7	CDK7	5q12.1	2.3097	1.1261	1.4370
150314	Lysophospholipase I	LYPLA1	8q11.23	2.3052	1.2361	1.5082
838744	Eukaryotic translation initiation factor 4 γ, 3	EIF4G3	1p36.12	2.3041	1.2753	3.2958
79688	Small nuclear ribonucleoprotein D2 polypeptide 16.5kDa	SNRPD2	19q13.2	2.2848	1.4725	1.2482
260303	V-ets erythroblastosis virus E26 oncogene homologue 2 (avian)			2.2813	1.2861	1.9798
897177	ESTs, Weakly similar to expressed protein			2.2637	1.5324	1.3195
785967	Erythrocyte membrane protein band 4.1-like 2	EPB41L2	6q23	2.2470	1.1712	1.5460
843426	WW45 protein	WW45	14q13–q23	2.2434	1.2059	1.5137

Table 2 Continued

Gene ID	Gene name	Gene symbol	Cytoband	SAM score	Fischer score	Fold change
813983	COP9 constitutive photomorphogenic homologue subunit 3	COPS3	17p11.2	2.2398	1.3274	1.4113
193106	ATP synthase, H+ transporting, mitochondrial F0 complex, subunit c	ATP5G3	2q31.1	2.2346	1.2695	1.4008
299388	Nuclear transport factor 2	NUTF2	16q22.1	2.2343	1.1758	1.3239
322914	Acid phosphatase 1, soluble	ACP1	2p25	2.2260	1.1579	1.4693
80410	Farnesyl diphosphate synthase (farnesyl pyrophosphate synthetase, d	FDPS	1q21.3	2.2189	1.1400	1.4935
487373	ATP synthase, H+ transporting, mitochondrial F0 complex, subunit c	ATP5G1	17q21.32	2.2181	1.2127	1.4795
1472753	Homosapiens cDNA FLJ34869 fis, clone NT2NE2014650,			2.2073	1.1529	1.9695
1476053	RAD51 homolog (RecA homologue, *E. coli*) (*S. cerevisiae*)	RAD51	15q15.1	2.2069	1.1858	1.9954
781704	High mobility group nucleosomal binding domain 3	HMGN3	6q14.2	2.1907	1.1293	1.5213
812167	Golgi associated, γ adaptin ear containing, ARF binding protein	GGA3	17q25.2	2.1775	1.1907	1.3577
795847	COP9 constitutive photomorphogenic homologue subunit 5	COPS5	8q12.3	2.1550	1.0630	1.4744
796096	Replication protein A1 (70kD)	RPA1	17p13.3	2.1511	1.2004	1.3024
342640	KIAA0101 gene product			2.1287	1.1810	3.8000
127925	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 18 (Myc-regulated)	DDX18	2q14.1	2.1050	1.0639	1.3671
47542	Small nuclear ribonucleoprotein D1 polypeptide (16kD)	SNRPD1	18q11.1	2.1029	1.0874	1.5059
814701	MAD2 mitotic arrest deficient-like 1 (yeast)	MAD2L1	4q27	2.0966	1.9901	1.6559
271744	Epiregulin	EREG	4q21.1	2.0869	1.0876	1.3443
773554	Spindlin	SPIN	9q22.1–q22.3	2.0744	1.0876	1.3261
52629	Calcium/calmodulin-dependent protein kinase 1	CAMK1	3p25.3	2.0674	1.0879	1.3139

For supervised class prediction, we used a SVM classifier and LOO cross-validation to evaluate the test error. With the use of feature selection, the top 100 discriminatory genes were selected (supplementary data: Table B). We found that class prediction was most accurate if we used only 15 of the top ranked discriminatory features (ranked by the Fisher score). In this regime, the SVM learned with zero training error and achieved a minimum of two LOO test errors from 14 for a soft margin SVM and three errors from 14 for a hard margin SVM (15). To evaluate the significance of this result, we performed a permutation test in which the class labels were randomly shuffled 100 times, and the LOO test error was evaluated. This gave a test error of 7.00 ± 2.31 for the null hypothesis (no prediction ability) that is given along with the observed test error curve (supplementary data: Fig. 1.). Therefore, finding two or three test errors from 14 samples means that this is significant, rather than occurring by chance ($P = 0.02–0.05$). Because our dataset is small, a confirmation of this effect will require more samples for an independent test set. However, the consistent drop in the test error curve below that expected from the null hypothesis does indicate that this is a real effect. When we compared the discriminatory features from SVM (100 genes) and the significant genes from SAM (113 genes), we found that ~80% of the genes (79 clones) were present in both lists, confirming their significant role in separating the two classes. In Table 2, the list of the 79 clones is presented with the SAM scores and Fischer scores (see "Methods and Materials"); the fold changes are also shown from the SAM analysis. These changes are mostly very subtle and often a smaller expression level change corresponds to a higher score if it is consistent because a more persistent small change can translate to a more significant value in this analysis.

Of these 79 cDNA clones, 5 are unknown expressed sequence tags, the remaining clones are known genes. Among them, several have a function in cell cycle regulation, DNA repair, and gene expression regulation. *RAD51* and *RAD23* were found to be significantly down-regulated in *BRCA1* mutation carriers compared with controls after radiation-induced DNA damage. Because both genes have an important role in DNA damage repair, their down-regulation in *BRCA1* mutation carriers could affect the DNA damage response in these cells. Interestingly, the *CDKN1B/p27^{Kip}* gene, which is a member of the cyclin-dependent kinase inhibitor family, was also significantly down-regulated in *BRCA1* mutation carriers compared with controls. This gene has been previously shown to be transactivated by BRCA1 (17), and the CDKN1B protein level has been shown to be decreased during breast and ovarian tumor development (18, 19). We have found that *ATR* was similarly down-regulated in mutation carriers compared with the controls. This gene has been also linked with *BRCA1* in the damage signaling pathway; there is evidence that ATR phosphorylates BRCA1 (20), but it is also possible that BRCA1 has an effect on the regulation of ATR as has been suggested in a recent study (21). It has been shown that BRCA1 facilitates the activity of ATR in its phosphorylation of various downstream elements. We propose that this scaffolding function of BRCA1 could be because of transcription activation.

In summary, we have shown that heterozygous *BRCA1*

mutation carrier fibroblasts have a distinctive gene expression phenotype after radiation-induced DNA damage. Our class prediction model was able to discriminate with 85% accuracy between normal breast fibroblasts from *BRCA1* mutation carriers and fibroblasts from reduction mammoplasty specimens from women with no family history. Furthermore, the genes that exhibit significant changes in their expression level after radiation-induced DNA damage have been shown in other studies to be part of the functional pathways through which BRCA1 operates. This supports the use of this preliminary result as the basis for the development of a functional assay to assess the downstream effects of germ-line changes in the *BRCA1* gene. Clearly, this initial study needs additional validation experiments, a larger sample set needs to be analyzed, and the prediction model will be tested on an independent sample set. However, our results provide evidence that expression profiling is a valuable approach in *BRCA1* mutation testing, which could also be applied to *BRCA2* testing. The use of fibroblasts is not the most practical as a routine screening test; however, the adaptation of these results to other cells types, *e.g.*, lymphocytes, could lead to the development of a functional assay for *BRCA1* germ-line mutation status. This would facilitate the diagnostic service in cancer genetics. Because the mutation test is expensive, samples could be prescreened by expression profiling to determine whom should be offered full mutation screening. Furthermore, it may be possible to distinguish those missense variants with functional significance from neutral polymorphisms, and it may be applicable to those families that harbor a regulatory mutation. This preliminary study has provided initial data that demonstrate that this approach is worth pursuing for clinical adaptation to functional genetic testing.

ACKNOWLEDGMENTS

We thank all of the patients and their clinicians for participating in this study. We also thank Dr. Michael Steel for providing extra information on the family of one of our carrier samples. This work was generously supported by the legacy of the late Marion Silcock. We also thank the fund "Irene Granleese for Caroline and Olive" for their tremendous support. The microarrays were produced at the Cancer Research United Kingdom (formerly Cancer Research Campaign) Microarray laboratory at The Institute of Cancer Research, United Kingdom.

REFERENCES

1. Easton, D. F., and Peto, J. The contribution of inherited predisposition to cancer incidence. Cancer Surv., *3:* 395–416, 1990.

2. Venkitaraman, A. R. Cancer susceptibility and the functions of BRCA1 and BRCA2. Cell, *108:* 171–182, 2002.

3. Tutt, A., and Ashworth, A. The relationship between the roles of BRCA genes an DNA repair and cancer predisposition. Trends Mol. Med., *8:* 571–576, 2002.

4. Welcsh, P., Lee, M. K., Gonzales-Hernandez, R. M., Black, D. J., Mahadevappa, M., Swisher, E. M., *et al.* BRCA1 transcriptionally regulates genes involved in breast tumorigenesis. Proc. Natl. Acad. Sci. USA, *99:* 7560–7565, 2002.

5. MacLachlan, T. K., Somasundaram, K., Sgagias, M., Shifman, Y., Muschel, R. J., Cowan, K. H., *et al.* BRCA1 effects on the cell cycle and the DNA damage response are linked to altered gene expression. J. Biol. Chem., *275:* 2777–2785, 2000.

6. Chung, C. H., Bernard, P. S., and Perou, C. M. Molecular portraits and the family tree of cancer. Nat. Genet., *32:* 533–540, 2002.

7. Golub, T. R., Slonim, D. K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J. P., *et al.* Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science (Wash. DC), *286:* 531–537, 1999.

8. Perou, C. M., Jeffrey, S. S., van de Rijn, R. M., Rees, C. A., Eisen, M. B., Ross, D. T., *et al.* Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. Proc. Natl. Acad. Sci. USA, *96:* 9212–9217, 1999.

9. West, M., Blanchette, C., Dressman, H., Huang, E., Ishida, S., Spang, R., *et al.* Predicting the clinical status of human breast cancer by using gene expression profiles. Proc. Natl. Acad. Sci. USA, *98:* 11462–11467, 2001.

10. van't Veer, L., Dal, H., de Vijver, M., He, Y. D., Hart, A., Mao, M., Peterse, H., Kooy, K., Marton, M., Witteveen, A., Schreiber, G., Kerhoven, R., Roberts, C., Linsley, P., Bernards, R., and Friend, S. Gene expression profiling predicts clinical outcome of breast cancer. Nature (Lond.), *415:* 530–535, 2002.

11. Hedenfalk, I., Duggan, D., Chen, Y., Radmacher, M., Bittner, M., Simon, R., *et al.* Gene-expression profiles in hereditary breast cancer. N. Engl. J. Med., *344:* 539–548, 2001.

11. Sorlie, T., Perou, C. M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., *et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc. Natl. Acad. Sci. USA, *98:* 10869–10874, 2001.

12. Clarke, P., te Poele, R. H., Wooster, R., and Workman, P. Gene expression microarray analysis in cancer biology, pharmacology and drug development. Biochem. Pharmacol., *62:* 1311–1336, 2001.

13. Ringner, M., Peterson, C., and Khan, J. Analyzing array data using supervised methods. Pharmacogenomics, *3:* 403–415, 2002.

14. Tusher, V., Tibshirani, R., and Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. Proc. Natl. Acad. Sci., *98:* 5116–5121, 2001.

15. Scholkopf, B., and Smola, A. Learning with Kernels. MIT Press 2002.

16. Sturn, A., Quackenbush, J., and Trajanoski, Z. Genesis: cluster analysis of microarray data. Bioinformatics, *18:* 207–208, 2002.

17. Williamson, E., Dadmanesh, F., and Koeffler, H. P. BRCA1 transactivates the cyclin-dependent kinase inhibitor $p27^{kip1}$. Oncogene, *21:* 3199–3206, 2002.

18. Catzavelos, C., Bhattacharya, N., Ung, Y. C., Wilson, J. A., Roncari, L., Sandhu, C., *et al.* Decreased levels of the cell-cycle inhibitor p27Kip1 protein: prognostic implications in primary breast cancer. Nat. Med., *3:* 227–230, 1997.

19. Masciullo, V., Sgamboto, A., PAcilio, C., Pucci, B., Ferrandina, G., Palazzo, J., *et al.* Frequent loss of expression of the cyclin-dependent kinase inhibitor p27 in epithelial ovarian cancer. Cancer Res., *59:* 3790–3794, 1999.

20. Chen, J. Ataxia telangiectasia-related protein is involved in the phosphorylation of BRCA1 following deoxyribonucleic acid damage. Cancer Res., *60:* 5037–5039, 2000.

21. Foray, N., Marot, D., Gabriel, A., Randrianarison, V., Carr, A. M., Perricaudet, M., Ashworth, A., and Jeggo, P. A subset of ATM- and ATR-dependent phosphorylation events requires the BRCA1 protein. EMBO J., *22:* 2860–2871, 2003.