Int. J. Radiat. Biol., 2007, pp. 1 – 12, i*First* article

TPD52 and NFKB1 gene expression levels correlate with G2 chromosomal radiosensitivity in lymphocytes of women with and at risk of hereditary breast cancer

A. H. SIMS[1], P. FINNON[2], C. J. MILLER[3], S. D. BOUFFLER[2], A. HOWELL[1], D. SCOTT[4],* & R. B. CLARKE[1]

[1]*Breast Biology Group, Paterson Institute for Cancer Research, University of Manchester,* [2]*Radiation Effects Department, Radiation Protection Division, Health Protection Agency, Centre for Radiation, Chemical and Environmental Hazards, Chilton, Didcot, Oxfordshire,* [3]*Cancer Research UK, Bioinformatics Group, Paterson Institute for Cancer Research, University of Manchester, and* [4]*Paterson Institute for Cancer Research, Manchester, UK*

(Received 24 October 2006; revised 19 February 2007; accepted 27 February 2007)

Abstract

Purpose: To evaluate a transcriptomic approach to identify healthy women at increased risk of breast cancer due to G2-radiosensitivity and look at transcripts that are differentially expressed between individuals.

Materials and methods: We perform the first study to assess the association of G2 radiosensitivity with basal gene expression in cultured T-lymphocytes from 11 women with breast cancer and 12 healthy female relatives using Affymetrix GeneChips.

Results: Transcripts associated with radiosensitivity and breast cancer risk were predominantly involved in innate immunity and inflammation, such as interleukins and chemokines. Genes differentially expressed in radiosensitive individuals were more similarly expressed in close family members than in un-related individuals, suggesting heritability of the trait. The expression of tumour protein D52 (*TPD52*), a gene implicated in cell proliferation, apoptosis, and vesicle trafficking was the most strongly correlated with G2 score while nuclear factor (kappa) – B (*NFKB1*) was highly inversely correlated with G2 score. *NFKB1* is known to be activated by irradiation and its inhibition has been previously shown to increase radiosensitivity.

Conclusions: Gene expression analysis of lymphocytes may provide a quantitative measure of radiation response potential and is a promising marker of breast cancer susceptibility.

Keywords: *Radiosensitivity, microarray, lymphocytes, breast cancer, NFKB1, G2 assay, TPD52*

Introduction

Breast epithelial cells become genetically unstable during early pre-malignant stages of breast cancer development. Due to the difficulty of measuring this change in the normal human breast *in situ*, blood lymphocytes have been used as surrogates of genetic change in the breast (Howell et al. 2005). Several groups have demonstrated ionising radiation-induced chromosome damage in the G2 phase of the cell cycle is higher in the lymphocytes of individuals with a number of different cancer types, including breast cancer (reviewed in Scott (2004), see also Howe et al. [2005]) compared to healthy controls. Pedigree analysis of blood relatives of individuals with increased chromosome damage demonstrated that G2-radiosensitivity may represent a heritable predisposition marker for breast cancer (Roberts et al. 1999). Previous analysis of gene expression in lymphoblastoid cells of the general population has demonstrated a familial aggregation of expression phenotypes (Cheung et al. 2003).

Whilst the G2 assay has been utilized successfully by several groups to assess radiosensitivity (Scott 2004), it is clear that highly stringent standardized laboratory procedures are needed to ensure assay reproducibility, which makes general applicability of the test difficult (Bryant et al. 2002,

---

Correspondence: Andrew H. Sims, Breast Biology Group, Paterson Institute for Cancer Research, University of Manchester, Wilmslow Road, Manchester M20 4BX, UK. Tel: +44(0) 161 446 3212. Fax: +44(0) 161 446 3218. E-mail: Andrew.H.Sims@manchester.ac.uk

*D. Scott has retired.

ISSN 0955-3002 print/ISSN 1362-3095 online
DOI: 10.1080/09553000701317366

Smart et al. 2003). There have also been reports of significant intra-individual variability in a few healthy normal donors (Smart et al. 2003, Vral et al. 2004, Howe et al. 2005). A more robust method of measuring an individual's radiosensitivity is required in order to exploit these findings for estimating cancer risk. Gene expression analysis offers a possible alternative approach. The identities of the gene or genes underlying radiosensitivity are unknown, although potential candidates might be those involved in the repair and processing of DNA damage. Putative mutants and polymorphic variants of these genes may be relatively common within the population and may be present at higher frequencies amongst patients with breast cancer (Roberts et al. 1999). The highly penetrant breast cancer genes, breast cancer 1, early onset (BRCA1) and breast cancer 2, early onset (BRCA2), responsible for more than 80% of families with six or more cases of early onset cancer, are known to have a role in DNA repair after exposure to ionizing radiation (Antoniou et al. 2001). Genes associated with G2-radiosensitivity may, in part account for inherited predisposition to breast cancer in families with fewer cases of the disease, as population studies have shown that BRCA genes only account for a minority of the overall familial risk of breast cancer (Antoniou et al. 2001).

The ability to identify women at increased risk of developing breast cancer from gene expression analysis of lymphocytes would be a highly attractive and minimally invasive option. Data from a recent study of gene expression in peripheral blood cells suggested that a blood-based gene expression test could be developed to detect breast cancer in asymptomatic patients (Sharma et al. 2005). By combining analysis of lymphocyte gene expression patterns that associate with G2 radiosensitivity and breast cancer in families, improved risk prediction may be achieved.

In this pilot study, we look at differential basal gene expression in cultured T-lymphocytes between radiosensitive (G2 high) and control (G2 low) women and also between breast cancer cases and healthy female relatives. We evaluate the feasibility of the approach, possible confounding factors and challenges towards using gene expression analysis of lymphocytes as a method of quantifying risk of breast cancer due to heritable G2 radiosensitivity.

Materials and methods

Sample details

Patient consent and approval by the local research ethics committee was obtained to use a subset of frozen lymphocyte samples from a previous study (Roberts et al. 1999) for further research. The current study is based upon peripheral blood lymphocytes taken from 23 women either diagnosed with breast cancer and close female blood relatives (Figure 1) who had previously been tested for radiosensitivity at this institute (Scott et al. 1999).

G2 assay

Full details of the G2 assay previously performed on these samples were given in Scott et al. (1999). The assay measures the number of chromosomal aberrations observed per 100 metaphase cells induced by 0.5 Gy X-ray exposure. Radiosensitive patients are defined as those with a G2 score greater than the 90th percentile of the distribution in a population of apparently normal individuals (Scott et al. 1999), which has been suggested to be and used as the cut-off for the radiosensitivity phenotype (Roberts et al. 1999).

Tissue culture

Frozen lymphocytes were cultured as described previously (O'Donovan et al. 1995). Cultures of thawed mononuclear cells were initiated by seeding $1-2\times10^5$ cells/ml in 10 ml SR10 containing $1-2\times10^5$ x-irradiated feeder cells/ml (GM1899A lymphoblastoid cell line) in 25 cm^2 flasks. SR10 consists of RPMI (Roswell Park Memorial Institute)-1640 medium (Dutch modification, Sigma-Aldrich, St Louis, MO, USA) with 10% foetal bovine serum, standard concentrations of penicillin/streptomycin, L-glutamine and sodium pyruvate, 50 μM 2-mercaptoethanol (all Invitrogen, Carlsbad, CA, USA), 0.4 μg/ml HA16 phytohaemagglutinin (PHA, Remel Inc., Lenexa, KS, USA) and 250 IU/ml Interleukin-2 (Novartis Vaccines and Diagnostics Inc., Emeryville, CA, USA). Seeded flasks were incubated at 37°C/5% CO_2 until day 4 of culture when, the cells were disaggregated and counted, then counted daily until cell concentrations reached 1×10^6/ml, cells were then diluted 1:1 with GR10 (SR10 without PHA). On day 7–8 cells were re-suspended at 2×10^6 cells/ml in Trizol® (Invitrogen, Carlsbad, CA, USA) and stored at −70°C before courier shipment on dry ice from the Health Protection Agency to the Paterson Institute for Cancer Research. Samples from four of the individuals were re-cultured from the original frozen lymphocytes to investigate inter-culture variation in gene expression.

Array processing and normalization

RNA was isolated from the cells suspended in Trizol® according to the manufacturer's instructions, purified using Qiagen RNeasy columns (Qiagen, Valencia, CA, USA) and quantified using a

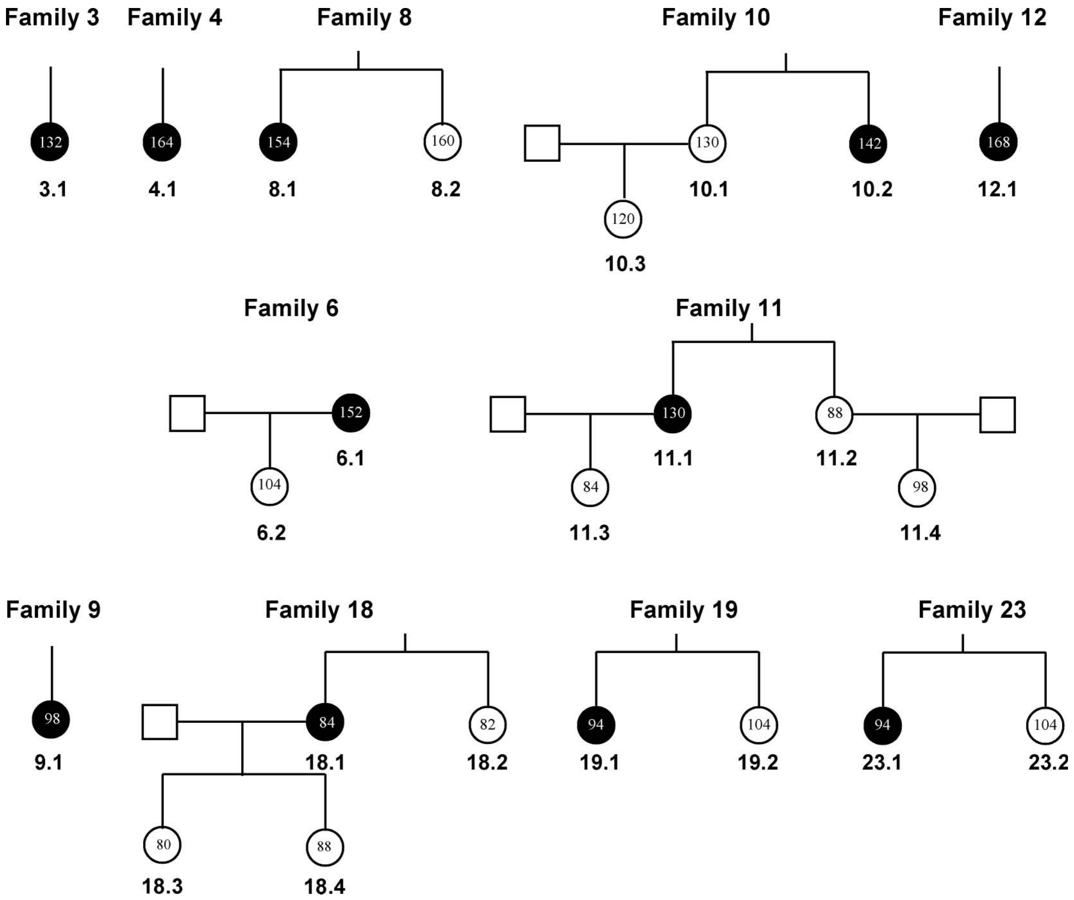

Figure 1. Pedigrees of the subset of samples for 23 women used in this pilot study. Filled symbols represent breast cancer patients. Number within symbol is the G2 assay score (number of chromosome aberrations per 100 cells) of the lymphocyte sample used for gene expression analysis. Radiosensitive individuals are those with a G2 assay score greater than 110 chromosome aberrations per 100 cells. The numbers below the symbols are individual identifiers and the family number relates to the original pedigree study (Roberts et al. 1999).

Nanodrop spectrophotometer (Labtech, Wilmington, DE, USA). The quality and amount of starting RNA was confirmed with an Agilent Bioanalyzer 2100 (Agilent, Santa Clara, CA, USA) prior to labelling and hybridization to Affymetrix HG-U133 plus 2.0 GeneChips (Affymetrix, Santa Clara, CA, USA) using standard protocols, (http://bioinformatics.picr.man.ac.uk/mbcf/protocols.shtml). After scanning, array images were assessed by eye to confirm scanner alignment and the absence of significant bubbles or scratches on the chip surface. Ratios (3′/5′) for glyceraldehyde-3-phosphate dehydrogenase and ß-actin were confirmed to be within acceptable limits (0.86 – 1.38 and 0.70 – 1.24, respectively), and BioB spike controls were found to be present on all chips, with BioC, BioD, and CreX also present in increasing intensity. Raw spot readings were imported into BioConductor (http://www.bioconductor.org) using R (Ihaka & Gentleman 1996) and Simpleaffy (Wilson & Miller 2005) with robust multi-chip average normalization (RMA) (Irizarry et al. 2003). All microarray data are MIAME compliant and accessible via MIAME VICE (http://bioinformatics.picr.man.ac.uk/mvice/index.jsp). Hierarchical clustering of samples was performed using average linkage clustering with a Pearson correlation measure of similarity using the MADE4 package (Culhane et al. 2005).

Rank product, prediction analysis of microarrays and pathway analysis

Due to the low number of samples, a rank product analysis method (Breitling et al. 2004) was used to determine the significance of differentially expressed probesets. This method has been shown to reliably and consistently outperform a non-parametric t-test, even with noisy data (Breitling et al. 2004). The RankProd BioConductor module within R was set to a false discovery rate (FDR) of 10%. The prediction analysis of microarrays program was used as described previously (Tibshirani et al. 2002) to look for possible gene classifiers based upon breast cancer and radiosensitivity status or G2 score. The GoMiner program (Zeeberg et al. 2005) was used to identify pathways that were over-represented by differentially expressed genes; *p*-values were calculated from a one-sided Fisher's exact test. Cut-offs of

$p < 0.02$ and a FDR of 0.2% were used to identify the most significant GO categories.

Results

Interculture variation

Lymphocyte cultures were assessed for gene expression in three batches, involving 8, 9 and 10 women, respectively. The third batch included four repeat cultures from the previous batches, two of which were positive and two negative for previously defined G2-radiosensitivity. These repeats provided a test of inter-culture variability for individual samples. Unsupervised hierarchical clustering of all 54,675 probesets represented on the HGU133 plus 2.0 microarray following RMA normalization (Irizarry et al. 2003) revealed that the 27 expression profiles separated into three distinct groups reflecting the three batches of lymphocytes. However, a simple supervised correction method was used to scale each batch of GeneChips relative to the whole dataset. A valuable consequence of this correction method is that it leaves fold-changes within an experiment-set unchanged. This means that the list of differentially expressed genes found, not only by a simple fold-change cut-off, but also by statistical techniques, remains identical to that identified prior to correction. The method has also been employed to reconcile other protocol effects by reducing systematic differences occurring due to different amplification protocols, array types and even scanner hardware (paper in preparation):

$$\text{Corrected expression} = \frac{\text{Uncorrected Expression}}{\text{Mean Batch expression}} \times \text{Mean overall expression}$$

Following correction and when considering the top 5% (2734) of probesets with the greatest variance, repeat cultures from the same women were seen to cluster together. These probesets with the highest variance were not clustered by genomic location, mapping widely across all chromosomes (data not shown), as seen previously (Cheung et al. 2003).

Differentially expressed genes

Gene expression levels were compared in 10 women (12 GeneChips) with high G2 scores (>110 aberrations per 100 cells = radiosensitive) with 13 women (15 GeneChips) with low G2 scores (controls). In addition to looking for genes that associate with radiosensitivity, we also identified genes that were differentially expressed in the lymphocytes from the 11 women diagnosed with breast cancer (13 GeneChips) compared with the 12 healthy controls (14 GeneChips). Approximately half of the probesets that associated with G2 radiosensitivity were also differentially expressed in the breast cancer cases compared to healthy controls. Similarly, two-thirds of the breast cancer associated probesets were also differentially expressed between women with high G2 scores and women with low G2 scores (Table I). Although radiosensitivity has been associated with breast cancer, it is not clear whether the two variables can be considered as mutually exclusive in our gene expression analysis. Identifying genes that were differentially expressed in association with both variables may highlight genes that most strongly correlated with breast cancer risk. The numbers of individuals in this pilot study were too few to strictly separate the radiosensitivity and breast cancer variables by looking for differentially expressed genes only in healthy individuals with a normal and radiosensitive responses or by looking at genes differentially expressed in breast cancer cases by considering only breast cancer patients and controls that have a normal radiation response. However, several previous studies have shown that these variables are associated (reviewed in Scott [2004]).

The Gene Ontology (GO) categories most significantly overrepresented in radiosensitive individuals compared to non-sensitive controls included cytokines, cytokine receptors and chemokine ligands (Table II). Transcripts representing interleukin 8 (*IL8*), interleukin 5 (*IL5*) and interleukin 17 receptor B (*IL17RB*) were upregulated in radiosensitive women compared to controls, whereas transcripts representing chemokine (C motif) ligand 1 (*XCL1*) and chemokine (C motif) ligand 2 (*XCL2*) were down regulated in radiosensitive women compared to controls, see Table I. *IL17RB* is a proinflammatory cytokine receptor, which induces the expression of mediators of inflammation. In a previous study, *IL17RB* was identified in a two-gene ratio (along with homeobox 13) expression signature, predictive of disease-free survival of breast cancer in tamoxifen-treated women, which outperformed existing biomarkers (Ma et al. 2004).

Correlation of G2 radiosensitivity with gene expression

An alternative analytical approach sought to identify genes that were quantitatively associated with radiosensitivity. This would remove any bias of the numbers of individuals classed as radiosensitive or controls and the strict 110 aberrations per 100 cells cut off. Of the 2734 probesets with the greatest variance, 261 had a correlation coefficient (r) of greater than 0.4 or lower than -0.4 ($p < 0.05$, see supplementary Table III, online version only). Three of the genes that were most strongly correlated with G2 score are also known to be associated with breast

Table I. The most differentially expressed genes associated with radiosensitivity and breast cancer. ‘Radiosensitive genes’ were identified by comparing the 10 women with G2 assay scores greater than 110 chromosomal aberrations per cell with the 13 women with lower G2 scores. The ‘breast cancer genes’ were identified by comparing 11 the breast cancer cases with 12 the healthy female relatives.

				Radiosensitive		Breast cancer	
Gene	Description	Affymetrix probeset ID	Chromosome location	FCR	PFP %	FCBC	PFP %
IL5	interleukin 5	207952_at	5q31.1	4.35	0	2.33	0
IL17RB	interleukin 17 receptor B	219255_x_at	3p21.1	3.85	0	2.78	0
IL17RB	interleukin 17 receptor B	224156_x_at	3p21.1	3.85	0	2.63	0
IL17RB	interleukin 17 receptor B	224361_s_at	3p21.1	3.45	0	2.50	0
IL8	interleukin 8	202859_x_at	4q13-q21	2.56	0	2.04	1.0
MAOA	monoamine oxidase A	212741_at	Xp11.3	2.56	0	2.08	2.9
KCNK1	potassium channel, subfamily K, member 1	204679_at	1q42-q43	2.38	0	1.92	4.9
IL8	interleukin 8	211506_s_at	4q13-q21	2.33	0	1.92	2.4
PROK2	prokineticin 2	232629_at	3p21.1	2.22	0	2.33	0
GPR27	G protein-coupled receptor 27	227769_at	3p21-p14	2.00	0.3	2.08	0.8
CCL3	chemokine (C-C motif) ligand 3	205114_s_at	17q11-q21	1.79	2.8	2.50	0
MGC11324	hypothetical protein MGC11324	224480_s_at	4q21.23	1.69	0.3	1.69	2.0
EOMES	eomesodermin homolog (Xenopus laevis)	231776_at	3p21.3-p21.2	1.45	1.7	2.17	0
unknown	Similar to RIKEN cDNA 1700028P14	243610_at	9q21.11	1.35	0.2	0.75	0
HLA-DQA1	MHC, class II, DQ alpha 1	213831_at	6p21.3	1.09	0.3	1.27	0.2
unknown	unknown	209480_at	6p21.3	1.02	1.5	1.37	0.2
HLA-DRB4	MHC, class II, DR beta 4	209728_at	6p21.3	0.86	2.4	1.85	0
MS4A1	Membrane-spanning 4-domains, subfamily A, member 1	228592_at	11q12	0.73	0.9	0.64	0.6
unknown	unknown	M10098_5_at		0.67	3.3	0.68	2.4
IGFBP2	insulin-like growth factor binding protein 2	202718_at	2q33-q34	0.58	1.0	0.47	0
IFI44L	interferon-induced protein 44-like	204439_at	1p31.1	0.57	0	0.68	0
XCL1	chemokine (C motif) ligand 1	214567_s_at	1q23	0.52	2.5	0.47	0.8
XCL2	chemokine (C motif) ligand 2	206366_x_at	1q23-q25	0.52	1.7	0.47	0.5
XCL1	chemokine (C motif) ligand 1	206365_at	1q23	0.50	1.3	0.53	1.6
PGDS	prostaglandin D2 synthase, hematopoietic	206726_at	4q22.3	2.27	0	–	–
MAOA	monoamine oxidase A	204389_at	Xp11.3	2.22	0.3	–	–
KIAA1211	KIAA1211 protein	227230_s_at	4q12	2.13	0.5	–	–
MAOA	monoamine oxidase A	204388_s_at	Xp11.3	2.08	1.3	–	–
CHDH	Choline dehydrogenase	229954_at	3p21.1	2.08	0.4	–	–
PLA2G4A	phospholipase A2, group IVA (cytosolic, calcium-dependent)	210145_at	1q25	2.00	0.8	–	–
HLA-DRB4	major histocompatibility complex, class II, DR beta 4	215666_at	6p21.3	1.96	2.1	–	–
XDH	Xanthine dehydrogenase	241994_at	2p23-p22	1.96	0.7	–	–
OSBPL10	oxysterol binding protein-like 10	219073_s_at	3p22.3	1.92	0.4	–	–
FLJ41238	FLJ41238 protein	229764_at	3q28	1.92	2.2	–	–
MRPL43	mitochondrial ribosomal protein L43	230026_at	10q24.31	1.92	1.0	–	–
ATP12A	ATPase, H+/K+ transporting, nongastric, alpha polypeptide	207367_at	13q12.12	1.89	3.5	–	–
PCNX	pecanex homolog (Drosophila)	215175_at	14q24.2	1.79	0.4	–	–
NTRK2	neurotrophic tyrosine kinase, receptor, type 2	221795_at	9q22.1	1.69	4.1	–	–
NTRK2	neurotrophic tyrosine kinase, receptor, type 2	221796_at	9q22.1	1.69	0.9	–	–

(continued)

Table I. (*Continued*).

Gene	Description	Affymetrix probeset ID	Chromosome location	Radiosensitive		Breast cancer	
				FCR	PFP %	FCBC	PFP %
GSTM3	glutathione S-transferase M3 (brain)	202554_s_at	1p13.3	1.54	0.6	–	–
KLRC1	killer cell lectin-like receptor subfamily C, member 1	206785_s_at	12p13	1.49	2.7	–	–
HLA-DRB4	Major histocompatibility complex, class II, DR beta 4	221491_x_at	6p21.3	1.41	4.4	–	–
HLA-DRB1	major histocompatibility complex, class II, DR beta 1	238900_at	6p21.3	0.63	4.6	–	–
PSPH	phosphoserine phosphatase	205048_s_at	7p15.2-p15.1	0.52	0.9	–	–
unknown	unknown	1558048_x_at		0.51	1.5	–	–
KSP37	Ksp37 protein	223836_at	4p16	0.50	1.2	–	–
LOC441057	hypothetical gene	226558_at		0.47	1.2	–	–
SP192	Hypothetical protein SP192	1560263_at	1p34.1	0.34	0	–	–
RBPMS	RNA binding protein with multiple splicing	209487_at	8p12-p11	–	–	2.08	5.4
TGFBR3	transforming growth factor, beta receptor III	204731_at	1p33-p32	–	–	1.92	5.0
CCL4	chemokine (C-C motif) ligand 4	204103_at	17q12	–	–	1.89	4.2
TGFBR3	Transforming growth factor, beta receptor III	226625_at	1p33-p32	–	–	1.89	4.1
PLEK	pleckstrin	203471_s_at	2p13.3	–	–	1.79	5.2
FCGR3A	Fc fragment of IgG, low affinity IIIa, receptor (CD16a)	204006_s_at	1q23	–	–	1.79	3.3
IFNG	Interferon, gamma	210354_at	12q14	–	–	1.75	4.4
GNG8	guanine nucleotide binding protein (G protein), gamma 8	234284_at	19q13.2-q13.3	–	–	0.68	4.9
EGR3	early growth response 3	206115_at	8p23-p21	–	–	0.66	2.3
ANXA3	annexin A3	209369_at	4q13-q22	–	–	0.64	2.4
LOC143903	layilin	228080_at	11q23.1	–	–	0.56	2.7
SDK2	Sidekick homolog 2 (chicken)	242064_at	17q25.1	–	–	0.55	1.3

FCR = Gene expression fold change between the radiosensitive and control women; FCBC = Gene expression fold change between the breast cancer cases and healthy female relatives; PFP = The estimated percentage chance that the gene is a false positive.

Table II. Gene Ontology (GO) pathways that contained significant numbers of differentially expressed genes in radiosensitive women compared to controls.

GO Classifications and pathways	Total[a]	Changed[b]	*p*-value	FDR	GO ID
Chemokine activity	10	3	0.0006	0.15	GO:0008009
Chemokine receptor binding	10	3	0.0006	0.15	GO:0042379
Behavior	62	6	0.0007	0.12	GO:0007610
Vasculature development	26	4	0.0011	0.09	GO:0001944
Blood vessel development	26	4	0.0011	0.09	GO:0001568
Potassium ion binding	5	2	0.0033	0.13	GO:0030955
Angiogenesis	18	3	0.0040	0.11	GO:0001525
Chemotaxis	39	4	0.0051	0.11	GO:0006935
Taxis	39	4	0.0051	0.11	GO:0042330
Locomotory behavior	45	4	0.0086	0.15	GO:0007626
Potassium ion transport	8	2	0.0089	0.14	GO:0006813
Defense response	247	10	0.0094	0.14	GO:0006952
Biogenic amine metabolism	9	2	0.0113	0.15	GO:0006576
Antigen processing, exogenous antigen via MHC class II	9	2	0.0113	0.15	GO:0019886
MHC class II receptor activity	9	2	0.0113	0.15	GO:0045012
Amino acid and derivative metabolism	50	4	0.0124	0.15	GO:0006519
Cytokine activity	50	4	0.0124	0.15	GO:0005125
Response to biotic stimulus	260	10	0.0135	0.15	GO:0009607
Antigen presentation, exogenous antigen	10	2	0.0139	0.17	GO:0019884
Alkali metal ion binding	10	2	0.0139	0.17	GO:0031420
Electron transport	29	3	0.0155	0.17	GO:0006118
Response to chemical stimulus	82	5	0.0156	0.16	GO:0042221
Amine metabolism	56	4	0.0184	0.19	GO:0009308

[a]Total is the number of probesets representing the given GO classifications within the top 5% most variable probesets; [b]Changed is the subset of probesets that were significantly differentially expressed (Table I). FDR is the percentage false discovery rate.

cancer (Figure 2). Tumour protein D52 (TPD52) was first identified in breast cancer cells and has been implicated in cell proliferation, apoptosis, and vesicle trafficking (Cao et al. 2006). GATA binding protein 3 (GATA3) is a known prognostic factor that may play a role in mechanisms controlling the response to estrogen. Annexin A1 has been shown to be both up (Ahn et al. 1997) and down-regulated (Shen et al. 2005) in breast cancer, with greater expression in invasive breast cancer cell lines (MDA-MB-231) than in non-invasive cell lines (MCF7s) and has anti-inflammatory properties.

G2-score was most strongly inversely correlated ($r = -0.67$, $p < 0.001$) with the expression of a probeset representing *NFKB1*, a transcription regulator with a central role in inflammation and innate immunity (Figure 2A). G2 score was also inversely correlated with the gene expression level of transcripts representing *NFKBIA* (I-kappa-B-alpha, $R = -0.48$, $p = 0.01$), which inactivates *NFKB1* by trapping it in the cytoplasm and *REL*, which binds to *NFKB1* to form the NFKB-complex (significant on a log2 scale, $R = -0.42$, not logged $R = -0.34$). The expression levels of Cycloxygenase 2 (*COX2*) and the signalling receptor for stromal cell-derived factor 1 (*CXCR4*) that are known to have increased expression in breast carcinomas (Soslow et al. 2000, Allinen et al. 2004) were also inversely correlated with G2 score (Figure 2A). However, the expression of other chemokines and cytokines were seen to directly correlate with G2 score, including *IL8*, *IL13*, *IL17RB* and *CCR2* (see supplementary Table III, online version only).

Although the expression of a number of genes that correlated with G2 score have previously been associated with breast cancer, there was higher statistical association of the expression of all of these genes (annexin A1 (*ANXA1*), *NFKB1*, chemokine (C-X-C motif) receptor 4 (*CXCR4*), cycloxygenase 2 (*COX2*), GATA binding protein 3 (*GATA3*) and tumour protein D52 (*TPD52*); Figure 2A) in radiosensitive compared to control women (Student's *t*-test $p = 0.002$, 0.008, 0.006, 0.01, 0.04 and 0.04), than there was with breast cancer patients compared to healthy individuals ($p = 0.2$, 0.03, 0.06, 0.2, 0.08 and 0.5, respectively), suggesting that overall gene expression in the radiosensitive individuals is not simply influenced by the breast cancer cases.

Heritability of radiosensitive gene expression

The variance in gene expression level between a pair of individuals can be used to assess whether genes are likely to have an inherited pattern of expression (Cheung et al. 2003). The greater the difference between the average variability in gene expression of more closely related individuals and that of those less/un-related, the greater the likelihood that gene

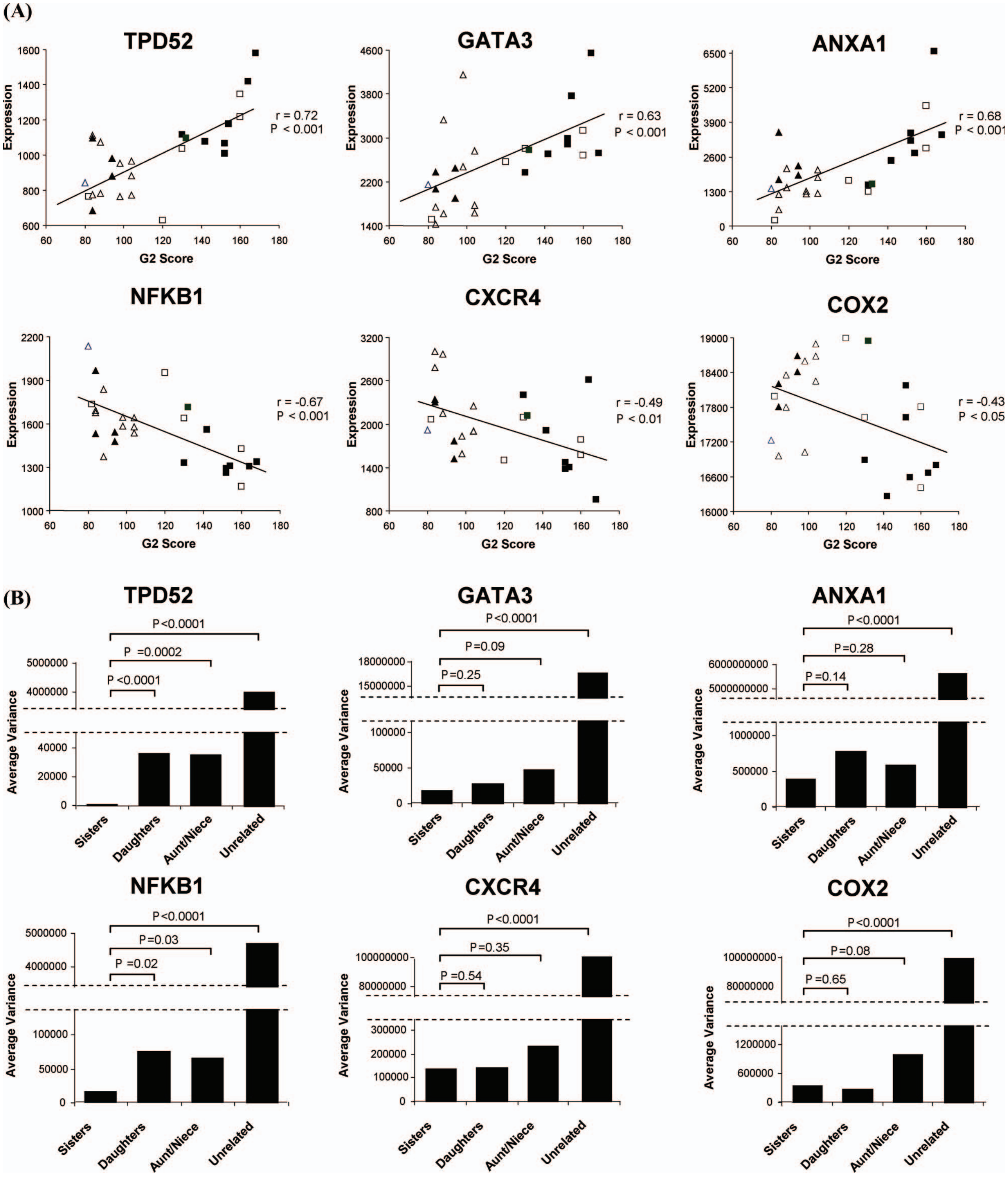

Figure 2. Correlation between gene expression and radiosensitivity, which may represent an inherited phenotype. (A) The G2 assay score represents the number of chromosome aberrations per 100 cells. Radiosensitive individuals are those with a G2-score greater than 110 aberrations per 100 cells. Squares represent radiosensitive women; triangles represent women with normal responses. Filled symbols represent the breast cancer cases. r = correlation coefficient, p = probability based upon the correlation coefficient. (B) Variation in transcript levels between pairs of related and unrelated individuals for genes whose expression correlates with G2 score. p-values are a measure of the significance of the difference in variance measured by the F-test.

expression is not just similar by chance, but actually represents an inherited trait. A total of 9, 10 and 6 pairs of expression values were used to calculate the average variance for sisters, daughter/mothers and aunts/nieces, respectively (Figure 1). The average variance for unrelated individuals for each probeset was calculated from all possible paired combinations of women from different families (321), assuming that none of the families were related to each other. The difference in variance between relatives and unrelated women for the highlighted genes in Figure 2A was highly significant using the F-test (Figure 2B). A trend for reduced variance with relatedness was observed, which was significant for *NFKB1* ($p = 0.02$ or $p = 0.03$) and highly significant for *TPD52* ($p < 0.0001$ or $p = 0.0002$) between sisters and daughters or aunts/nieces (Figure 2B).

Discussion

Genes differentially expressed in the blood lymphocytes of individuals who were radiosensitive and/or were diagnosed with breast cancer were predominantly involved with signal transduction pathways that control the expression of genes encoding cytokines and chemokines or enzymes that regulate innate and active immune responses. These genes (Table I) may provide clues to the molecular nature of radiosensitivity and breast cancer susceptibility.

It was noted that normal healthy women who have a radiosensitive sister with breast cancer (individuals 8.2, 10.1 and 11.2, Figure 1) have a gene expression profile that is more similar to that of breast cancer patients than normal women who have a sister with breast cancer that is not radiosensitive (individuals 18.2, 19.2 and 23.2). However, in order to fully evaluate the predictive capacity of a gene classifier derived from this analysis, larger numbers samples would be required to generate separate 'training' and 'test' sets. While the numbers of individuals in the current study were too few to evaluate the predictive capacity of a 'radiosensitive breast cancer prediction profile', it seems likely that such a profile could be evaluated in a larger study.

Segregation analysis in the previous study (Roberts et al. 1999) of inheritance of G2-radiosensitivity in breast cancer suggested that it is dominated by one or more major genes. A single gene was thought to account for 82% of the variance observed between members of 20 families, although a model with a second rarer gene fitted the data best (Roberts et al. 1999). A number of epidemiological studies have led to polygenic risk models (Antoniou et al. 2001, Houlston & Peto 2004, Pharoah et al. 2002) for breast cancer. Given the relatively high prevalence of radiosensitivity, observed in 40% of breast cancer patients and 10% of healthy controls (Scott et al. 1999), a polygenic model may better account for the radiosensitive phenotype. A number of the genes highlighted in this study were most similarly expressed in the closest related family members, which is consistent with the findings of the previous study (Roberts et al. 1999), that G2 radiosensitivity represents an inherited predisposition to breast cancer.

The expression of the transcript representing TPD52 was most strongly directly correlated with G2 score and demonstrated highly statistically significant heritability (Figure 2). This gene was first identified as being overexpressed in human breast tumours (Byrne et al. 1995) and has subsequently been identified as a candidate target gene for events resulting in amplification of chromosome band 8q21, a region frequently gained in breast and other human cancers (Balleine et al. 2000). Tumour protein D52-like 1 (TPD52L1) is a related gene that is known to interact with TPD52, these proteins have an amino terminal coiled-coil domain that allows members to form homo- and heterodimers with each other (Cao et al. 2006). TPD52L1 has been implicated in G2/M transit (Boutros & Byrne 2005), which may be consistent with a role in G2-radiosensitivity. Some of the processes in which TPD52 and related proteins have been implicated, particularly proliferation, cell cycle transit and apoptosis, may be of functional relevance to G2-radiosensitivity and breast cancer risk.

The *GATA3* transcript is also highly correlated to G2-sensitivity and is co-expressed with *TPD52* gene in the luminal sub-type of breast tumours (Mehra et al. 2005). This is relevant because it has been recently demonstrated that GATA3 regulates luminal epithelial cell differentiation in normal breast development (Kouros-Mehr et al. 2006, Asselin-Labat et al. 2007), and has previously been shown to regulate lymphocyte differentiation into T helper 2 cells (Murphy & Reiner 2002). One might speculate that higher gene expression of *GATA3* indicates greater levels of T helper 2 cell differentiation in our lymphocyte samples and may be a surrogate for more luminal cell differentiation in the breast. How this would influence radiosensitivity and susceptibility to breast cancer is currently unknown.

The transcript most inversely correlated with G2-radiosensitivity in this study was NFKB1, a major transcription regulator that is known to be activated in response to clinically relevant (≤ 2 Gy) doses of ionizing radiation (Criswell et al. 2003). Previous studies have demonstrated that gamma irradiation of rats *in vivo* induces a cascade of inflammatory responses involving NFKB1 (Linard et al. 2004) and deficient NFKB1-induced gene expression has been implicated in Fanconi's anemia (Kontou et al. 2003), a cancer-prone human disease marked by chromosomal instability. Mendonca et al. (2003) demonstrated increased radiosensitivity and an increased percentage of cells in G2/M following inhibition of NFKB1 with parthenolide. These and other observations have led to the suggestion of using NFKB1-inhibitors to enhance radiosensitivity in patients undergoing radiation treatment (Magne et al. 2006).

Exactly how NFKB1 responds to irradiation is not clearly understood, although it has been shown that the ataxia telangiectasia mutated (ATM) protein is required for sustained activation of NFKB1 following DNA damage (Piret et al. 1999), which would be consistent with previous work demonstrating that cells from ATM homozygotes and heterozygotes have increased G2 radiosensitivity (Scott et al. 1999) and an impaired NFKB1 response to ionizing radiation (Jung et al. 1995). Alternatively, induction of inflammation and immune pathways via NFKB1 may involve interferon gamma (*IFGN*), which responds to irradiation (Wasserman et al. 1991) and was up-regulated in the breast cancer patients in this study (Table I).

In the current study, the transcript level of the inhibitor of NFKB1 (NFKBIA) was also seen to inversely correlate with G2-score, which is consistent with a study showing that inhibition of NFKBIA phosphorylation causes radiosensitization of glioma cells (Ding et al. 2003). Exactly why the expression of the transcription factor and its inhibitor are both inversely correlated with G2-score is not clear. It may simply be another example of the conflicting differences between transcripts and proteins that are 'induced' (requiring synthesis) and those which are 'activated' (not requiring synthesis) by varying doses of ionizing radiation in different cell types (Criswell et al. 2003). On the other hand, it may be indicative of the dual roles of NFKB1 in cancer, where it has been seen to be constituently active in many types of cancer associated with chronic inflammation and yet NFKB1 inhibition has been implicated in carcinogen-induced tumourogenesis (Pikarsky & Ben-Neriah 2006). The definable mechanisms that explain whether *NFKB1* plays a positive, negative or neutral role in modulation of radiosensitivity *in vivo*, remain to be resolved (Criswell et al. 2003).

At this stage, it is not possible to predict whether the results observed in our study would be replicated by taking freshly collected blood or whether storing and culturing frozen lymphocytes introduces an additional bias. Care was taken to ensure that all samples were treated identically, to minimize any processing effects. However, Burrill et al. (2000) demonstrated more variability in cryopreserved than fresh lymphocytes using the MN assay and similar findings have previously been observed for the G2 assay (Scott, unpublished). Blood samples are currently being prospectively collected from a large cohort of women for the Norwegian Women and Cancer Study, which will constitute a valuable resource for identifying risk factors (Borresen-Dale & Sharma 2005). An initial report from this study has shown that early detection of breast cancer is possible using gene expression patterns in peripheral blood cells (Sharma et al. 2005). However, there was no clear overlap between the genes identified in our study with those seen by Sharma et al. (2005). Gene expression analysis of irradiation-induced DNA damage in cultured skin fibroblasts distinguishes between *BRCA1/2* mutation carriers and controls (Kote-Jarai et al. 2004, 2006). Therefore, it would be very interesting to analyze and compare gene expression of the DNA damage response to irradiation in lymphocytes from radiosensitive and control women.

It was perhaps surprising that more genes involved in DNA repair were not identified in the present study. A mechanistic study of G2 chromosomal radiosensitivity elucidated that cdk1/cyclin B activity levels during G2 to M transition play a major role in the yield of chromatin breaks induced after G2 irradiation of synchronized breast cancer cell lines (Terzoudi et al. 2000). In the current study of unsynchronized, unirradiated lymphocytes from both normal and cancer patients we observed higher cyclin-B expression levels in radiosensitive individuals, although this was not statistically significant.

The current study highlights the potential to investigate known phenotypic heritable markers, such as G2-radiosensitivity from blood lymphocytes using a high throughput technology in order to better understand the risk of developing breast cancer. Clearly, further work is needed in order to determine whether the differentially expressed genes identified represent efficient markers for risk prediction. Examining the expression of genes and proteins within the breast itself may ultimately enable improved risk prediction (Howell et al. 2005), however advocating breast biopsies on a population basis is unrealistic. Differential gene expression due to the radiosensitive phenotype (as determined by the G2 assay) may represent a predisposition to ionizing radiation. Our results showing increased levels of TPD52 and reduced levels of *NFKB1* associated with radiosensitivity may represent molecular markers of non-BRCA heritable forms of breast cancer.

Acknowledgements

We thank all the patients and family members for their invaluable contribution to this study. We are grateful for the excellent technical assistance of Stuart Pepper and Yvonne Hey of the Cancer Research UK Affymetrix core facility. RBC is supported by Cancer Research UK and Breast Cancer Campaign. AHS is supported by the Genetic Innovation network and Breakthrough Breast Cancer.

References

Ahn SH, Sawada H, Ro JY, Nicolson GL. 1997. Differential expression of annexin I in human mammary ductal epithelial cells in normal and benign and malignant breast tissues. Clinical and Experimental Metastasis 15:151–156.

Allinen M, Beroukhim R, Cai L, Brennan C, Lahti-Domenici J, Huang H, Porter D, Hu M, Chin L, Richardson A, Schnitt S, Sellers WR, Polyak K. 2004. Molecular characterization of the tumor microenvironment in breast cancer. Cancer Cell 6:17 – 32.

Antoniou AC, Pharoah PD, McMullan G, Day NE, Ponder BA, Easton D. 2001. Evidence for further breast cancer susceptibility genes in addition to BRCA1 and BRCA2 in a population-based study. Genetic Epidemiology 21:1 – 18.

Asselin-Labat ML, Sutherland KD, Barker H, Thomas R, Shackleton M, Forrest NC, Hartley L, Robb L, Grosveld FG, van der Wees J, Lindeman GJ, Visvader JE. 2007. Gata-3 is an essential regulator of mammary-gland morphogenesis and luminal-cell differentiation. Nature Cell Biology 9:201 – 209.

Balleine RL, Fejzo MS, Sathasivam P, Basset P, Clarke CL, Byrne JA. 2000. The hD52 (TPD52) gene is a candidate target gene for events resulting in increased 8q21 copy number in human breast carcinoma. Genes Chromosomes Cancer 29: 48 – 57.

Borresen-Dale AL, Sharma P. 2005. Expression profiling of peripheral blood cells for early detection. Breast Cancer Research 7(Supp2):S8.

Boutros R, Byrne JA. 2005. D53 (TPD52L1) is a cell cycle-regulated protein maximally expressed at the G2-M transition in breast cancer cells. Experimental Cell Research 310:152 – 165.

Breitling R, Armengaud P, Amtmann A, Herzyk P. 2004. Rank products: A simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. FEBS Letters 573:83 – 92.

Bryant PE, Gray L, Riches AC, Steel CM, Finnon P, Howe O, Kesterton I, Vral A, Curwen GB, Smart V, Tawn EJ, Whitehouse CA. 2002. The G2 chromosomal radiosensitivity assay: A technical report. International Journal of Radiation Biology 78:863 – 866.

Burrill W, Levine EL, Hindocha P, Roberts SA, Scott D. 2000. The use of cryopreserved lymphocytes in assessing inter-individual radiosensitivity with the micronucleus assay. International Journal of Radiation Biology 76:375 – 382.

Byrne JA, Tomasetto C, Garnier JM, Rouyer N, Mattei MG, Bellocq JP, Rio MC, Basset P. 1995. A screening method to identify genes commonly overexpressed in carcinomas and the identification of a novel complementary DNA sequence. Cancer Research 55:2896 – 2903.

Cao Q, Chen J, Zhu L, Liu Y, Zhou Z, Sha J, Wang S, Li J. 2006. A testis-specific and testis developmentally regulated tumor protein D52 (TPD52)-like protein TPD52L3/hD55 interacts with TPD52 family proteins. Biochemical and Biophysical Research Communications 344:798 – 806.

Cheung VG, Conlin LK, Weber TM, Arcaro M, Jen KY, Morley M, Spielman RS. 2003. Natural variation in human gene expression assessed in lymphoblastoid cells. Nature Genetics 33:422 – 425.

Criswell T, Leskov K, Miyamoto S, Luo G, Boothman DA. 2003. Transcription factors activated in mammalian cells after clinically relevant doses of ionizing radiation. Oncogene 22:5813 – 5827.

Culhane AC, Thioulouse J, Perriere G, Higgins DG. 2005. MADE4: An R package for multivariate analysis of gene expression data. Bioinformatics 21:2789 – 2790.

Ding GR, Honda N, Nakahara T, Tian F, Yoshida M, Hirose H, Miyakoshi J. 2003. Radiosensitization by inhibition of IkappaB-alpha phosphorylation in human glioma cells. Radiation Research 160:232 – 237.

Houlston RS, Peto J. 2004. The search for low-penetrance cancer susceptibility alleles. Oncogene 23:6471 – 6476.

Howe OL, Daly PA, Seymour C, Ormiston W, Nolan C, Mothersill C. 2005. Elevated G2 chromosomal radiosensitivity in Irish breast cancer patients: A comparison with other studies. International Journal of Radiation Biology 81:373 – 378.

Howell A, Sims AH, Ong KR, Harvie M, Evans DG, Clarke RB. 2005. Mechanisms of disease: prediction and prevention of breast cancer – cellular and molecular interactions. Nature Clinical Practice Oncology 2:635 – 646.

Ihaka R, Gentleman R. 1996. R: A language for data analysis and graphics. Journal of Computational and Graphical Statistics 5:299 – 314.

Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP. 2003. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 4:249 – 264.

Jung M, Zhang Y, Lee S, Dritschilo A. 1995. Correction of radiation sensitivity in ataxia telangiectasia cells by a truncated I kappa B-alpha. Science 268:1619 – 1621.

Kontou M, Adelfalk C, Hirsch-Kauffmann M, Schweiger M. 2003. Suboptimal action of NF-kappaB in Fanconi anemia cells results from low levels of thioredoxin. Biological Chemistry 384:1501 – 1507.

Kote-Jarai Z, Matthews L, Osorio A, Shanley S, Giddings I, Moreews F, Locke I, Evans DG, Eccles D, Williams RD, Girolami M, Campbell C, Eeles R. 2006. Accurate prediction of BRCA1 and BRCA2 heterozygous genotype using expression profiling after induced DNA damage. Clinical Cancer Research 12:3896 – 3901.

Kote-Jarai Z, Williams RD, Cattini N, Copeland M, Giddings I, Wooster R, tePoele RH, Workman P, Gusterson B, Peacock J, Gui G, Campbell C, Eeles R. 2004. Gene expression profiling after radiation-induced DNA damage is strongly predictive of BRCA1 mutation carrier status. Clinical Cancer Research 10:958 – 963.

Kouros-Mehr H, Slorach EM, Sternlicht MD, Werb Z. 2006. GATA-3 maintains the differentiation of the luminal cell fate in the mammary gland. Cell 127:1041 – 1055.

Linard C, Marquette C, Mathieu J, Pennequin A, Clarencon D, Mathe D. 2004. Acute induction of inflammatory cytokine expression after gamma-irradiation in the rat: Effect of an NF-kappaB inhibitor. International Journal of Radiation Oncology Biology Physics 58:427 – 434.

Ma XJ, Wang Z, Ryan PD, Isakoff SJ, Barmettler A, Fuller A, Muir B, Mohapatra G, Salunga R, Tuggle JT, Tran Y, Tran D, Tassin A, Amon P, Wang W, Wang W, Enright E, Stecker K, Estepa-Sabal E, Smith B, Younger J, Balis U, Michaelson J, Bhan A, Habin K, Baer TM, Brugge J, Haber DA, Erlander MG, Sgroi DC. 2004. A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen. Cancer Cell 5:607 – 616.

Magne N, Toillon RA, Bottero V, Didelot C, Houtte PV, Gerard JP, Peyron JF. 2006. NF-kappaB modulation and ionizing radiation: Mechanisms and future directions for cancer treatment. Cancer Letters 231:158 – 168.

Mehra R, Varambally S, Ding L, Shen R, Sabel MS, Ghosh D, Chinnaiyan AM, Kleer CG. 2005. Identification of GATA3 as a breast cancer prognostic marker by global gene expression meta-analysis. Cancer Research 65:11259 – 11264.

Mendonca MS, Hardacre M, Datzman N, Comerford K, Chin-Sinex H, Sweeny C. 2003. Inhibition of NFKappaB activity by the anti-inflammatory sesquiterpene, parthenolide slows cell growth and increases radiation sensitivity. International Journal of Radiation Oncology-Biology-Physics 57: S354.

Murphy KM, Reiner SL. 2002. The lineage decisions of helper T cells. National Review of Immunology 2:933 – 944.

O'Donovan MR, Freemantle MR, Hull G, Bell DA, Arlett CF, Cole J. 1995. Extended-term cultures of human T-lymphocytes: A practical alternative to primary human lymphocytes for use in genotoxicity testing. Mutagenesis 10:189 – 201.

Pharoah PD, Antoniou A, Bobrow M, Zimmern RL, Easton DF, Ponder BA. 2002. Polygenic susceptibility to breast cancer and implications for prevention. Nature Genetics 31:33 – 36.

Pikarsky E, Ben-Neriah Y. 2006. NF-kappaB inhibition: A double-edged sword in cancer? European Journal of Cancer 42:779 – 784.

Piret B, Schoonbroodt S, Piette J. 1999. The ATM protein is required for sustained activation of NF-kappaB following DNA damage. Oncogene 18:2261 – 2271.

Roberts SA, Spreadborough AR, Bulman B, Barber JB, Evans DG, Scott D. 1999. Heritability of cellular radiosensitivity: A marker of low-penetrance predisposition genes in breast cancer? American Journal of Human Genetics 65:784 – 794.

Scott D. 2004. Chromosomal radiosensitivity and low penetrance predisposition to cancer. Cytogenetics and Genome Research 104:365 – 370.

Scott D, Barber JB, Spreadborough AR, Burrill W, Roberts SA. 1999. Increased chromosomal radiosensitivity in breast cancer patients: A comparison of two assays. International Journal of Radiation Biology 75:1 – 10.

Sharma P, Sahni NS, Tibshirani R, Skaane P, Urdal P, Berghagen H, Jensen M, Kristiansen L, Moen C, Zaka A, Arnes J, Sauer T, Akslen LA, Schlichting E, Borresen-Dale AL, Lonneborg A. 2005. Early detection of breast cancer based on gene-expression patterns in peripheral blood cells. Breast Cancer Research 7:R634 – 644.

Shen D, Chang HR, Chen Z, He J, Lonsberry V, Elshimali Y, Chia D, Seligson D, Goodglick L, Nelson SF, Gornbein JA. 2005. Loss of annexin A1 expression in human breast cancer detected by multiple high-throughput analyses. Biochemical and Biophysical Research Communications 326:218 – 227.

Smart V, Curwen GB, Whitehouse CA, Edwards A, Tawn EJ. 2003. Chromosomal radiosensitivity: A study of the chromosomal G(2) assay in human blood lymphocytes indicating significant inter-individual variability. Mutation Research 528: 105 – 110.

Soslow RA, Dannenberg AJ, Rush D, Woerner BM, Khan KN, Masferrer J, Koki AT. 2000. COX-2 is expressed in human pulmonary, colonic, and mammary tumors. Cancer 89:2637 – 2645.

Terzoudi GI, Jung T, Hain J, Vrouvas J, Margaritis K, Donta-Bakoyianni C, Makropoulos V, Angelakis P, Pantelias GE. 2000. Increased G2 chromosomal radiosensitivity in cancer patients: The role of cdk1/cyclin-B activity level in the mechanisms involved. International Journal of Radiation Biology 76:607 – 615.

Tibshirani R, Hastie T, Narasimhan B, Chu G. 2002. Diagnosis of multiple cancer types by shrunken centroids of gene expression. Proceedings of the National Academy of Sciences USA 99:6567 – 6572.

Vral A, Thierens H, Baeyens A, De Ridder L. 2004. Chromosomal aberrations and *in vitro* radiosensitivity: Intra-individual versus inter-individual variability. Toxicology Letters 149:345 – 352.

Wasserman J, Petrini B, Wolk G, Vedin I, Glas U, Blomgren H, Ekre HP, Strannegard O. 1991. Cytokine release from mononuclear cells in patients irradiated for breast cancer. Anticancer Research 11:461 – 464.

Wilson CL, Miller CJ. 2005. Simpleaffy: A BioConductor package for Affymetrix Quality Control and data analysis. Bioinformatics 21:3683 – 3685.

Zeeberg BR, Qin H, Narasimhan S, Sunshine M, Cao H, Kane DW, Reimers M, Stephens RM, Bryant D, Burt SK, Elnekave E, Hari DM, Wynn TA, Cunningham-Rundles C, Stewart DM, Nelson D, Weinstein JN. 2005. High-Throughput GoMiner, an 'industrial-strength' integrative gene ontology tool for interpretation of multiple-microarray experiments, with application to studies of Common Variable Immune Deficiency (CVID). BMC Bioinformatics 6:168.