

## Report

---

# Genetic Variation in Radiation-Induced Expression Phenotypes

Candace R. Correa<sup>2</sup> and Vivian G. Cheung<sup>1,2</sup>

<sup>1</sup>Departments of Pediatrics and Genetics, <sup>2</sup>School of Medicine, University of Pennsylvania, Philadelphia

Studies have demonstrated that natural variation in the expression level of genes at baseline is extensive, and the determinants of this variation can be mapped by a genetic-linkage approach. In this study, we used lymphoblastoid cells to explore the variation in radiation-induced transcriptional changes. We found that, among normal individuals, there is extensive variation in transcriptional response to radiation exposure. By studying monozygotic twins, we demonstrated that there is evidence of a heritable component to this variation. The postradiation variation in the expression level of several genes, including the ferredoxin reductase gene (*FDXR*) and the cyclin-dependent kinase inhibitor 1A gene (*CDKN1A*), is significantly greater ( $P < .001$ ) among twin pairs than within twin pairs. The induction of *FDXR* by radiation showed a bimodal distribution. Our findings have important implications for understanding the genetic basis of radiation response, which has remained largely unknown due to the lack of family material needed for genetic studies. Our approach, which uses expression phenotypes in cell lines, allows us to expose cells from family members to radiation. Similar study design can be applied to dissect the genetic basis of other complex human traits.

Humans are exposed to ionizing radiation (IR) through the environment and in medical settings during diagnostic and therapeutic procedures. Much variation in IR response has been observed among individuals. In medical settings, patients receiving the same doses of radiation were found to have different acute and long-term side effects, from dermatological changes to pneumonitis, and a lifetime increased risk of cancer. It has been suggested that there is a genetic component to variation in radiation response (reviewed by Gatti [2001]). Some of these conclusions are based on the severe response to radiation found among patients with radiosensitivity syndromes, such as ataxia telangiectasia (MIM 208900), Nijmegen breakage syndrome (MIM 251260), and Fanconi anemia (MIM 227650). In addition, polymorphic sequence variants in genes known to be involved in radiation-responsive pathways have been shown to correlate with risks of radiation-induced telangiectasia and subcutaneous fibrosis (Andreassen et al. 2003). How-

ever, the genetic basis of variation in radiation response in humans remains largely unknown (Andreassen et al. 2002).

Genetic analysis begins with measurements of phenotypic variation. Elsewhere, we have demonstrated that there is extensive variation in the expression level of genes (or in expression phenotypes [Cheung and Spielman 2002]) in lymphoblastoid cells at baseline among normal individuals. In addition, we have found evidence that there is a heritable component to this natural variation (Cheung et al. 2003), which can be mapped to specific chromosomal locations (Morley et al. 2004). We have also observed that heterozygous carriers of ataxia telangiectasia have distinct expression phenotypes that differ from those of control individuals (Watts et al. 2002). In the present study, we explore the genetic component of expression phenotypes in the setting of radiation response.

For our analysis of variation in IR response, we chose nine well-characterized genes known to be induced at least 2-fold in lymphoblastoid cells, in response to 10 Gy of IR (Jen and Cheung 2003). These genes and their functions—according to the Gene Ontology database (Ashburner et al. 2000; Harris et al. 2004)—are those listed in table 1. We chose to evaluate the transcriptional response of these genes to IR, because this allows us to investigate genes that participate in different IR-respon-

Received June 30, 2004; accepted for publication August 17, 2004; electronically published September 9, 2004.

Address for correspondence and reprints: Dr. Vivian G. Cheung, Department of Pediatrics, University of Pennsylvania, 3516 Civic Center Boulevard, ARC 516, Philadelphia, PA 19104. E-mail: vcheung@mail.med.upenn.edu

© 2004 by The American Society of Human Genetics. All rights reserved.  
0002-9297/2004/7505-00XX\$15.00

**Table 1****Nine IR-Responsive Genes Analyzed in this Study**

Gene Symbol	Gene Name	Gene Ontology Biological Process
<i>ACTA2</i>	Actin, alpha 2, smooth muscle, aorta	Cytoskeleton
<i>ATF3</i>	Activating transcription factor 3	Regulation of transcription
<i>CDKN1A</i>	Cyclin-dependent kinase inhibitor 1A (p21, Cip1)	Cell cycle
<i>CXCR4</i>	Chemokine (C-X-C motif), receptor 4 (fusin)	Cell death
<i>DDB2</i>	Damage-specific DNA binding protein 2 (48 kD)	DNA repair
<i>FDXR</i>	Ferredoxin reductase	Metabolism
<i>GADD45A</i>	Growth arrest and DNA-damage-inducible, alpha	DNA repair
<i>PPM1D</i>	Protein phosphatase 1D magnesium-dependent, delta isoform	DNA repair
<i>TP53I3</i>	Tumor protein p53 inducible protein 3	Cell death

sive pathways, such as cell-cycle regulation and DNA repair (Amundson et al. 2001, 2003; Ford et al. 2001; Jen and Cheung 2003). Results from the present study showed that the expression levels of most of these IR-responsive genes vary among normal individuals, both at baseline and in response to IR. We also found evidence for a genetic basis of these radiation-induced expression phenotypes.

Lymphoblastoid cell lines were obtained from 10 unrelated individuals who are part of the Centre d'Etude du Polymorphisme Humain (CEPH) Utah pedigrees (Dausset et al. 1990) and from 10 sets of MZ twins. The individuals are not known to have radiosensitivity syndromes or previous histories of adverse reactions to radiation. Cells were grown under identical conditions and were irradiated at 10 Gy with a cesium-137 irradiator. Cells were harvested prior to radiation (baseline) and at 2 h and 6 h postradiation. RNA was extracted, was converted to cDNA, and was used as a template for measuring the expression levels of the IR-responsive genes by RT-PCR using SYBR green assays (Applied Biosystems). All expression measurements were performed with three replicates.

First, we assessed the variation in expression level of

the IR-responsive genes at baseline and in response to radiation, among unrelated individuals. For this analysis, we used the expression levels of genes from the 10 unrelated individuals and those from 10 sets of MZ twins (equivalent to 10 additional unrelated individuals); we estimated the variance for each group separately and then combined the two estimates as the weighted average. The combined estimated variances for the nine genes at baseline range from 0.24 to 1.88, with a mean of 0.64 (table 2). *DDB2* is the most variable gene at baseline, with estimated variance between individuals of 1.88; its expression level varies by ~2-fold between the individual with the highest and the individual with the lowest expression level of this gene.

In addition to baseline variation—since the focus of this study is variability among individuals in response to radiation—we also assessed variation in changes in transcriptional response at two time points postradiation. To accomplish this, for each individual and for each gene, we plotted the expression levels at 2 h and 6 h post-IR relative to the normalized baseline level. Area under the curve (AUC) was calculated and was used as a measure of IR response (see “Material and Methods” section in appendix A [online only]) (Geara et al. 1996; Safwat et al. 2002; Jen and Cheung 2003). As in the assessment of baseline variation, we used the combined (unrelated individuals and twin sets) estimated variance of AUC to measure the variability in IR response among individuals. For the nine genes, this measurement ranged from 1.8 to 18.4, with a mean of 6.5 (table 2). The expression level of the ferredoxin reductase gene (*FDXR*) appears to be the most variable. Its expression levels at 2 h and 6 h postradiation varied by ~2 fold and ~13 fold, respectively, among the individuals (fig. 1).

To determine whether there is a genetic component to the variation in the expression level of the IR-response genes, we compared the variance within twin pairs with the variance among twin pairs by use of the intraclass correlation coefficient (ICC [Sokal and Rohlf 1994]). At baseline, the expression levels of seven of the nine genes had an ICC > 0.50, and their expression levels are sig-

**Table 2**

**Estimated Variances for Expression Levels of IR-Responsive Genes among 20 Unrelated Individuals at Baseline and in Response to Radiation**

GENE	EXPRESSION-LEVEL VARIANCE	
	At Baseline	Postradiation
<i>CXCR4</i>	.63	1.82
<i>ATF3</i>	.43	3.47
<i>TP53I3</i>	.43	3.78
<i>GADD45A</i>	.54	3.85
<i>DDB2</i>	1.88	3.95
<i>PPM1D</i>	.25	5.21
<i>ACTA2</i>	.44	6.45
<i>CDKN1A</i>	.93	11.78
<i>FDXR</i>	.24	18.35

**Table 3**

**ICC of Baseline Expression Levels of IR-Responsive Genes**

Gene	ICC <sub>Baseline</sub>	ICC 95% CI <sup>a</sup>	P <sup>b</sup>
<i>CXCR4</i>	.16	0–.69	.31
<i>ACTA2</i>	.38	0–.80	.11
<i>ATF3</i>	.54	0–.86	.04
<i>GADD45A</i>	.66	.12–.90	.01
<i>DDB2</i>	.71	.22–.92	.005
<i>PPM1D</i>	.75	.30–.93	.003
<i>FDXR</i>	.73	.26–.93	.004
<i>TP53I3</i>	.79	.38–.94	.001
<i>CDKN1A</i>	.87	.59–.97	.0001

<sup>a</sup> Negative values of lower limits are set to zero (Sokal and Rohlf 1994).

<sup>b</sup> P value from corresponding analysis of variance.

nificantly more similar ( $P < .05$ ) within twin pairs than among twin pairs (table 3). For two genes, *TP53I3* and *CDKN1A*, this difference is highly significant ( $P \leq .001$ ; table 3). Correspondingly, the post-IR expression levels of most genes had an ICC  $> 0.50$ , and, for several genes (*PPM1D*, *FDXR*, and *CDKN1A*), their expression levels are significantly more similar ( $P < .001$ ; table 4) within twin pairs than among twin pairs, supporting the idea that there is a genetic component to the variation in the expression level of these genes. These observations suggest that genetic differences among individuals contribute to variation in both the baseline and postradiation expression levels. For a few expression phenotypes, the variation within twins, at baseline and/or postradia-

**Table 4**

**ICC of Radiation Response (by AUC) of IR-Responsive Genes**

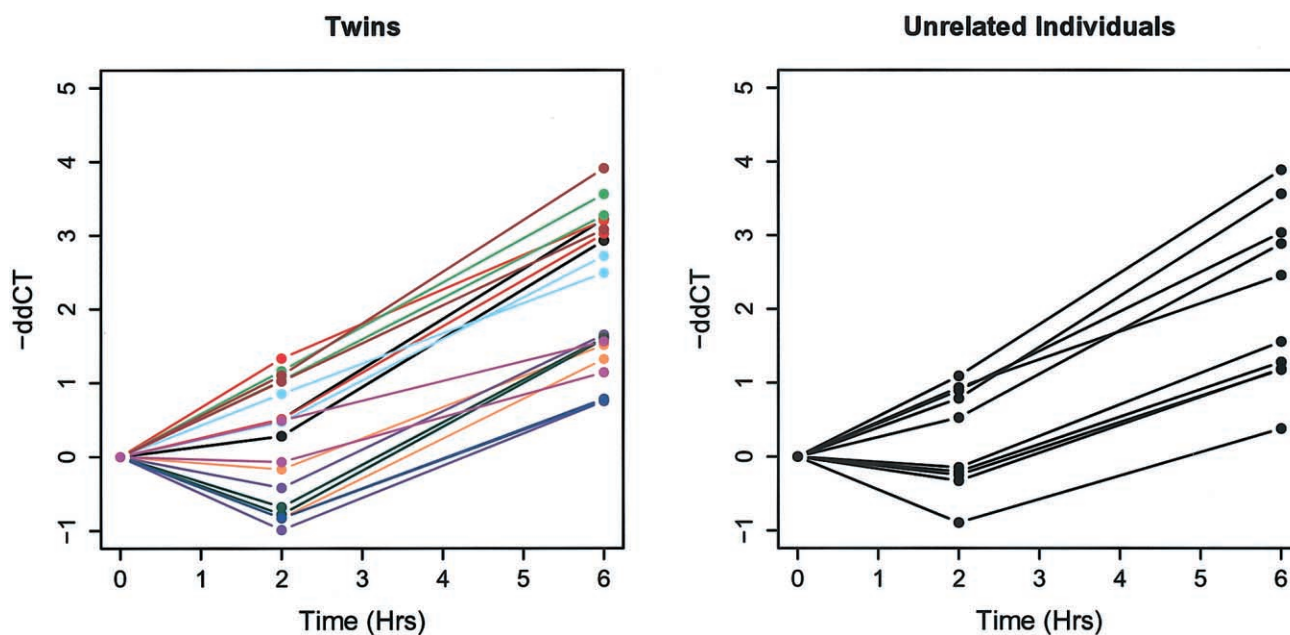
Gene	ICC <sub>Radioreponse</sub>	ICC 95% CI <sup>a</sup>	P <sup>b</sup>
<i>CXCR4</i>	.28	0–.75	.19
<i>DDB2</i>	.47	0–.83	.06
<i>GADD45A</i>	.55	0–.86	.03
<i>ACTA2</i>	.59	.01–.88	.02
<i>TP53I3</i>	.71	.22–.92	.005
<i>ATF3</i>	.73	.26–.92	.004
<i>PPM1D</i>	.87	.60–.97	.0001
<i>FDXR</i>	.90	.67–.97	.00004
<i>CDKN1A</i>	.92	.74–.98	.00001

<sup>a</sup> Negative values of lower limits are set to zero (Sokal and Rohlf 1994).

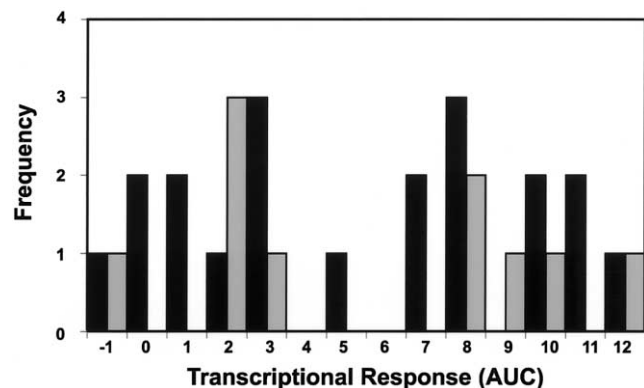
<sup>b</sup> P value from corresponding analysis of variance.

tion, is not significantly different, compared with that among twins ( $P$  values slightly greater than .05). Given our sample size (10 sets of MZ twins), these findings do not imply a lack of genetic control for these expression phenotypes. Measurements of the extent of genetic control provide possible causes for the observed variation. However, identification of the sequence variants that influence the phenotypes is the ultimate proof for genetic control. Our results suggest that transcriptional response to IR exposure is highly amenable to genetic dissection.

The IR induction of *FDXR*, as measured by AUC, showed a bimodal distribution (figs. 1 and 2), with means located at  $\sim 1.0$  and  $\sim 8.8$ . This distribution sug-



**Figure 1** Transcriptional response of *FDXR* to IR. The graphs show negative ddCT (a proxy for fold change, which equals  $2^{-ddCT}$ ) for *FDXR* versus time for 10 twin pairs and for 10 unrelated individuals. The members of each twin pair are shown with matching color lines.



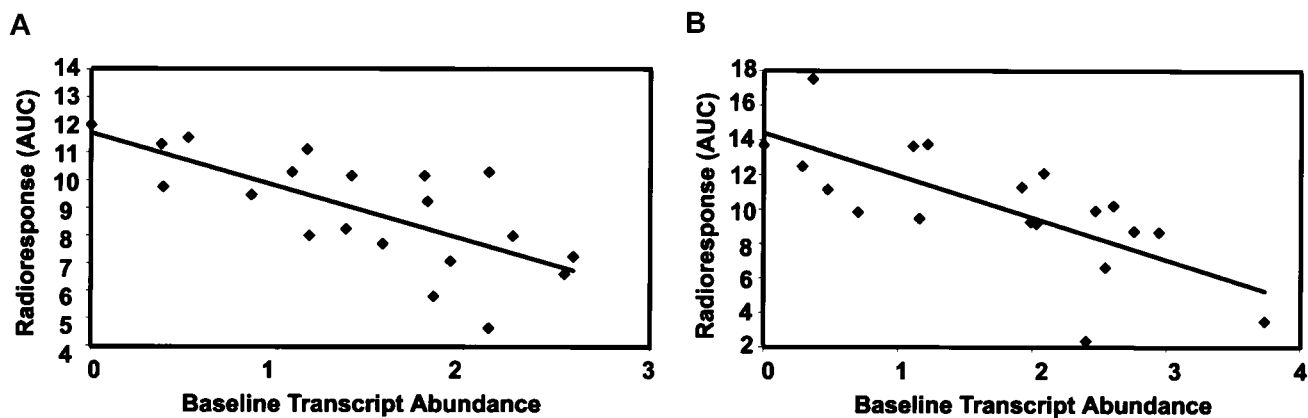
**Figure 2** Bimodal distribution of transcriptional response of *FDXR* to IR. Transcriptional response (calculated by AUC) of 10 sets of MZ twins (black bars) and 10 unrelated individuals (gray bars).

gests that individuals are either “low” or “high” responders. Among the 10 unrelated individuals, 5 were “high” responders, and 5 were “low” responders. Similarly, among the 10 twin pairs, the numbers of “high” and “low” responders are the same. Members of a twin pair always belonged to the same responder group, providing additional evidence that there is a genetic basis to the magnitude of induction of *FDXR*. To confirm the apparent bimodality, we tested for a bimodal distribution, using NOCOM software (Ott 1992), which tests for a mixture of normal distributions. The *FDXR* data are significantly better fitted by two normal distributions than by one normal distribution ( $\chi^2 = 15$ ; 2 df;  $P < 5 \times 10^{-4}$ ). We also tested the other eight genes for bimodality, using NOCOM, but none of them showed strong evidence for bimodality.

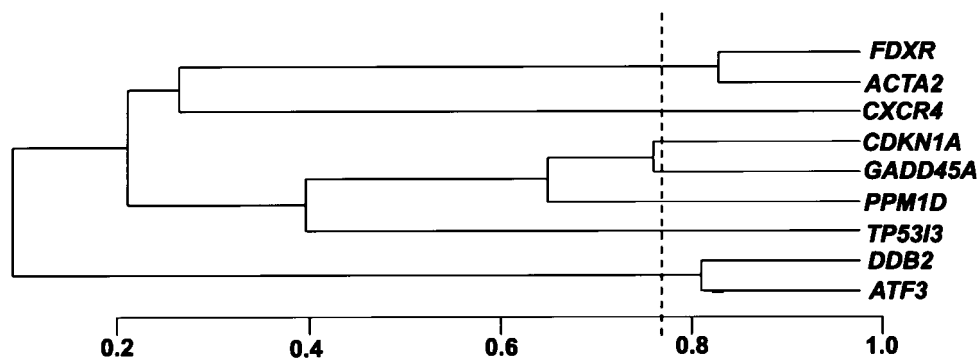
We compared the expression level at baseline with that

postradiation (as measured by AUC) to assess whether the baseline level of a gene can predict the transcriptional response to IR. For this analysis, we used data from 20 individuals, including 10 unrelated individuals and 10 individuals who are members of twin pairs (one member randomly chosen per twin pair). Among the nine genes, the highest correlations between baseline and post-IR transcript levels were found for *GADD45* and *CDKN1A*. For these two genes, the correlation coefficients of radioresponse and baseline transcript abundance were  $-0.71$  and  $-0.72$ , respectively (fig. 3) (the correlation coefficients for the other genes are shown in table A1 [online only]). For *GADD45* and *CDKN1A*, we observed a larger increase in the expression level among individuals with lower baseline transcript abundance than in the expression level among individuals with higher baseline transcript abundance. This may imply that there is a threshold effect for response. When there were more transcripts at baseline, less induction was needed to meet that threshold.

Lastly, we looked for correlation in radioresponse between the genes by examining the correlation in expression levels post-IR. The expression levels (AUCs) of 20 unrelated individuals (10 CEPH individuals and 10 members of twin sets randomly chosen, as in the analysis described above) were used. In permutation tests with 1,000 replications, the highest pairwise correlation between any two genes was 0.78. We therefore set this as the threshold for chance correlation ( $P < .001$ ). The correlations were summarized by hierarchical clustering (fig. 4). Two pairs of genes (*ACTA2* and *FDXR*; *DDB2* and *ATF3*) had the highest correlation coefficients of 0.83 and 0.81, respectively. One gene, *CXCR4*, was the least correlated with any of the other genes (the maximum correlation coefficient with any gene was 0.33).



**Figure 3** Prediction of radioresponse by use of baseline transcript abundance. Baseline transcript abundance of *GADD45A* ( $r = -0.71$ ) (A) and *CDKN1A* ( $r = -0.72$ ) (B) are negatively correlated with radioresponse (calculated by AUC) ( $n = 20$ ). Correlation coefficients between radioresponse and baseline transcript abundance for each gene are shown.



**Figure 4** Correlation of IR-responsive genes. The similarity of expression phenotypes of nine IR-responsive genes was assessed by Pearson's correlation coefficient (absolute value). The dendrogram represents hierarchical clustering of the genes by use of the average-linkage method. Expression levels of genes with branches connected to the right of the dotted line are correlated at  $P < .001$ .

This agrees with the fact that all the genes in our study except *CXCR4* are known target genes of p53; thus, it is likely that they are coregulated in the same signaling pathway.

In conclusion, by measuring the expression phenotypes of IR-responsive genes among normal individuals, we identified substantial variation and found evidence of a genetic component to this variation by studying MZ twins. Among the genes that were analyzed, the strongest evidence of a heritable component was for *FDXR* and *CDKN1A*. Their post-IR expression levels were most variable among individuals, and the variability was much smaller among genetically identical twins than among unrelated individuals. The distribution of the post-IR expression level of *FDXR* among individuals was bimodal, suggesting that a single locus may be responsible for regulating its expression.

The genetics of radiation response, like other complex quantitative human phenotypes, has been difficult to study. Part of the problem arises from the lack of phenotypes that can be precisely measured and from the difficulty in collecting family material for genetic studies. Our study design using expression phenotypes of cell lines provides some solutions to these problems. First, advances in genomic technology have made it possible to extend classical phenotypes to include expression levels of genes, which are relatively easy to obtain. In previous studies, we demonstrated that there is a heritable component to variation in baseline gene expression (Cheung et al. 2003), and genetic determinants of this variation can be mapped by linkage analysis (Morley et al. 2004). In this study, we showed that, similar to baseline expression phenotype, the expression phenotypes of response to radiation are amenable to genetic analysis. Second, for genetic studies, phenotypes and genotypes from related individuals have to be obtained. This poses a significant problem in many studies, including in the

genetic analysis of radiation response. It is not possible to expose family members to radiation, as would be required for genetic study. With gene expression as the phenotype, the starting material can be cell lines, which are easier to collect and to manipulate experimentally. In our study, cells were grown under identical conditions and were exposed to the same dose of radiation. This would have been impossible if our starting materials were patients rather than the cell lines. Our experimental design allows us to begin to identify genetic determinants that influence an individual's susceptibility to radiation. This experimental design can also be applied to study other cellular responses, as well as to study individual variation in response to other toxins, including pharmacologic agents.

## Acknowledgments

We thank Richard Spielman, for advice and comments on data analysis; Kuang-Yu Jen, for discussion on experimental design; and Josh Burdick and Weijia Yao, for technical assistance. This work is supported by National Institutes of Health grants GM070540 and HG02386 and the W. W. Smith Endowed Chair in Pediatric Genomics.

## Electronic-Database Information

URLs for data presented herein are as follows:

NOCOM and COMPMIX, <http://linkage.rockefeller.edu/ott/nocom.htm>

Online Mendelian Inheritance in Man (OMIM), <http://www.ncbi.nlm.nih.gov/Omim/> (for ataxia telangiectasia, Nijmegen breakage syndrome, and Fanconi anemia)

## References

- Amundson SA, Bittner M, Meltzer P, Trent J, Fornace AJ Jr (2001) Induction of gene expression as a monitor of exposure to ionizing radiation. *Radiat Res* 156:657–661
- Amundson SA, Lee RA, Koch-Paiz CA, Bittner ML, Meltzer P, Trent JM, Fornace AJ Jr (2003) Differential responses of stress genes to low dose-rate gamma irradiation. *Mol Cancer Res* 1:445–452
- Andreassen CN, Alsner J, Overgaard J (2002) Does variability in normal tissue reactions after radiotherapy have a genetic basis—where and how to look for it? *Radiother Oncol* 64: 131–140
- Andreassen CN, Alsner J, Overgaard M, Overgaard J (2003) Prediction of normal tissue radiosensitivity from polymorphisms in candidate genes. *Radiother Oncol* 69:127–135
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 25:25–29
- Cheung VG, Conlin LK, Weber TM, Arcaro M, Jen K-Y, Morley M, Spielman RS (2003) Natural variation in human gene expression assessed in lymphoblastoid cells. *Nat Genet* 33: 422–425
- Cheung VG, Spielman RS (2002) The genetics of variation in gene expression. *Nat Genet* 32:522–525
- Dausset J, Cann H, Cohen D, Lathrop M, Lalouel JM, White R (1990) Centre d'étude du polymorphisme humain (CEPH): collaborative genetic mapping of the human genome. *Genomics* 6:575–577
- Ford BN, Wilkinson D, Thorleifson EM, Tracy BL (2001) Gene expression responses in lymphoblastoid cells after radiation exposure. *Radiat Res* 156:668–671
- Gatti RA (2001) The inherited basis of human radiosensitivity. *Acta Oncol* 40:702–711
- Gera FB, Peters LJ, Ang KK, Garden AS, Tucker SL, Levy LB, Brown BW (1996) Comparison between normal tissue reactions and local tumor control in head and neck cancer patients treated by definitive radiotherapy. *Int J Radiat Oncol Biol Phys* 35:455–462
- Harris MA, Clark J, Ireland A, Lomax J, Ashburner M, Foulger R, Eilbeck K, et al (2004) The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res* 32: 258–261
- Jen K-Y, Cheung VG (2003) Transcriptional response of lymphoblastoid cells to ionizing radiation. *Genome Res* 13: 2092–2100
- Morley M, Molony CM, Weber TM, Devlin JL, Ewens KG, Spielman RS, Cheung VG (2004) Genetic analysis of genome-wide variation in human gene expression. *Nature* 430: 743–747
- Ott J (1992) NOCOM and COMPMIX programs release. Rockefeller University, New York
- Safwat A, Bentzen SM, Turesson I, Hendry JH (2002) Deterministic rather than stochastic factors explain most of the variation in the expression of skin telangiectasia after radiotherapy. *Int J Radiat Oncol Biol Phys* 52:198–204
- Sokal RR, Rohlf FJ (1994) *Biometry*, 3rd ed. W. H. Freeman & Company, New York, pp 213–215
- Watts JA, Morley M, Burdick JT, Fiori JL, Ewens WJ, Spielman RS, Cheung VG (2002) Gene expression phenotype in heterozygous carriers of ataxia telangiectasia. *Am J Hum Genet* 71:791–800