

Classification of Sensitivity or Resistance of Cervical Cancers to Ionizing Radiation According to Expression Profiles of 62 Genes Selected by cDNA Microarray Analysis¹

Osamu Kitahara*, Toyomasa Katagiri*, Tatsuhiko Tsunoda[†], Yoko Harima[‡] and Yusuke Nakamura*

*Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, 4-6-1, Shirokanedai, Minato-ku, Tokyo 108-8639, Japan; [†]SNP Research Center, RIKEN (Institute of Physical and Chemical Research), 4-6-1, Shirokanedai, Minato-ku, Tokyo 108-8639, Japan; [‡]Department of Radiology, Kansai Medical University, 10-15 Fumizono-cho, Moriguchi, Osaka 570-8506, Japan

Abstract

To identify a set of genes related to radiosensitivity of cervical squamous cell carcinomas and to establish a predictive method, we compared expression profiles of 9 radiosensitive and 10 radioresistant tumors obtained by biopsy before treatment, on a cDNA microarray consisting of 23,040 human genes. We identified 121 genes whose expression was significantly greater in radiosensitive cells than in radioresistant cells, and 50 genes that showed higher levels of expression in radioresistant cells than in radiosensitive cells. Some of these genes had already known to be associated with the radiation response, such as aldehyde dehydrogenase 1 (*ALDH1*) and X-ray repair cross-complementing 5 (*XRCC5*) ($P < .05$, Mann-Whitney test). The validity of the total of 171 genes as radiosensitivity related genes were certified by permutation test ($P < .05$). Furthermore, we selected 62 genes on the basis of a clustering analysis, and confirmed the validity of these genes with cross-validation test. The cross-validation test also indicates the possibility of making prediction of radiosensitivity for discriminating radiation-sensitive from radiation resistant biopsy samples by predicting score (PS) values calculated from expression values of 62 genes in 19 samples, because the prediction successfully and unequivocally discriminated the radiosensitive phenotype from the radioresistant phenotype in our test panel of 19 cervical carcinomas. The extensive list of genes identified in these experiments provides a large body of potentially valuable information for studying the mechanism(s) of radiosensitivity, and selected 62 genes opens the possibility of providing appropriate and effective radiotherapy to cancer patients.

Neoplasia (2002) 4, 295–303 doi:10.1038/sj.neo.7900251

Keywords: cervical cancer, radiosensitivity, cDNA microarray, gene expression profiles, classification.

Introduction

Although proper diagnosis and effective treatments for cervical cancers are widely available now [1], this disease is still a leading cause of death for women worldwide [2].

Radiotherapy is a generally effective therapeutic method, particularly for patients with cancers at an advanced stage. However, individual patients may show quite different patterns of response against radiotherapy; some can be cured, but others cannot, and the latter may therefore suffer needlessly from severe side effects. Hence, if treatment is to become more patient specific, the molecular mechanism(s) of radiosensitivity need to be clarified.

Several molecular markers that reflect radiosensitivity have been proposed as the result of studies that have involved, for example, transfection of oncogenes such as *N-ras*, *v-myc* with *H-ras*, and *v-fos* into cultured cells to induce a radioresistant phenotype [3,4]. Activation of *c-raf-1* has been positively correlated with radioresistance of head and neck squamous cell carcinomas [5], and certain cell cycle- and apoptosis-related genes also have been correlated with radiosensitivity; e.g., loss or dysfunction of p16 renders melanoma cells resistant to ionizing radiation, whereas expression of exogenous wild-type p16 and p21 in glioblastoma cells can induce radiosensitivity [6,7]. Transfection- or radiation-induced expression of Bcl-2 proteins, which regulate apoptosis, into pro-myeloid cells has introduced a radioresistant phenotype [8,9]; furthermore, expression of Bax, when induced by gamma irradiation, confers radiosensitivity on lymphoid cells, small intestinal epithelial cells [10] and cervical cancer cells [9]. However, although such discoveries have brought partial understanding of the molecular mechanisms responsible for cellular radiosensitivity, the whole picture remains to be clarified.

Because the complex mechanism of radiosensitivity cannot be explained by a small number of genes, we need to collect genome-wide information about all the genes involved. To that end, we recruited a newly developed

Abbreviations: PS, predictive scoring; aRNA, amplified RNA; CR, complete response; NC, no change; *ALDH1*, aldehyde dehydrogenase; *XRCC5*, X-ray repair cross-complementing 5; *RBPI*, retinol-binding protein 1; DSB, double-strand break

Address all correspondence to: Dr. Yusuke Nakamura, Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan. E-mail: yusuke@ims.u-tokyo.ac.jp

¹This work was supported in part by Research for the Future Program Grant No. 00L01402 from the Japan Society for the Promotion of Science.

Received 17 December 2001; Accepted 28 January 2002.

technique, cDNA microarray [11], which provides high-throughput analysis of expression profiles by means of small-array slides spotted with cDNAs [12-15].

Here we report a genome-wide cDNA microarray analysis of 23,040 human cDNAs, in biopsy samples from 19 cervical squamous carcinomas (9 of them radiosensitive and 10 radioresistant, classified according to tumor-suppression ratios). We identified 171 genes that were differentially expressed between the two groups; of those, 121 showed elevated expression and 50 showed decreased expression in radiosensitive tumor cells relative to their expression levels in radioresistant cells. In addition, further selected 62 genes showed feasibility of predicting the radiosensitivity of cervical squamous cell carcinomas. These results not only disclose the complex nature of radiosensitivity as regards the response of cervical squamous cell carcinomas to ionizing radiation, but also provide information that should identify novel targets for efforts to expand the effectiveness of radiotherapy.

Materials and Methods

Tissue Samples

Cervical squamous cell carcinoma tissues were obtained with informed consent from 19 patients who underwent biopsy before radiotherapy at Kansai Medical University, and snap-frozen at -80°C . All cases were at stages IIB to IVB, and were papillomavirus (HPV)-positive. We observed the population of tumor cells in all biopsy samples was over 90% by hematoxylin and eosin staining. Methods for typing of HPV and establishing p53 status were described previously [9].

Radiation Treatments

All 19 patients were treated with radiotherapy after sampling. A total of 30.6 Gy was provided to the whole pelvis, plus an additional dose to parametria with central shielding to complete 52.2 Gy, along with ^{192}Ir high dose-rate intracavitary brachytherapy. Details have been described elsewhere [9]. Effects of the therapy, including local failure, were checked a month after treatment. Nine patients revealed 100% reduction in tumor size and the remaining 10 showed 0% to 40% reduction (Table 1). The former were classified as a radiosensitive group (complete response; [CR]) and the latter as a radioresistant group (no change [NC]).

RNA Extraction and Amplification

Total RNAs were extracted from each specimen by TRIZOL (Invitrogen, Life Technologies, Carlsbad, CA) according to the manufacturer's protocol. The extracted RNAs were treated for 1 hour at 37°C with 10 U of DNase I (Nippon Gene, Japan) in the presence of 1 U of RNase inhibitor (TOYOBO, Osaka, Japan), to remove any contaminating genomic DNA. After inactivation of DNase at 70°C for 10 minutes, the RNAs were purified with phenol-chloroform-isoamyl alcohol (Gibco BRL, Grand Island, NY)

and then precipitated by ethanol. Next, all DNase I-treated RNAs were subjected to T7-based RNA amplification as described previously [16]. Two rounds of amplification yielded 65 to 152 μg of amplified RNA (aRNA) from each sample. As a control for comparing gene expression profiles between the CR and NC groups on the microarray, we performed two rounds of T7-RNA amplification using a mixture of poly A⁺ RNAs from tissues of 12 normal human organs (brain, heart, liver, skeletal muscle, small intestine, spleen, placenta, thyroid, fetal brain, fetal kidney, fetal lung, and fetal liver) purchased from Clontech (Palo Alto, CA) as a control.

Microarray Design, Production, and Hybridization

We selected 23,040 cDNA clones from the UniGene database of the National Center for Biotechnology Information (Bethesda, MD) (build #131). Our cDNA microarray was constructed essentially as described previously [13]; 2.5 μg of aRNA from each cervical carcinoma was labeled with Cy5-dCTP and the control aRNA was labeled with Cy3-dCTP by a protocol described elsewhere [13]. Hybridization, washing, and scanning were carried out according to published methods [13].

Data Analysis and Selection of Differentially Expressed Genes

Signal intensities of Cy3 and Cy5 from the 23,040 spots were quantified by the Array vision software (Amersham Biosciences, Piscataway, NJ) and normalized as described previously [12]. In the quantification step, local background correction method was adopted. Because the data were unreliable when intensities fell below 2.5×10^5 relative fluorescent units or signal to noise ratios were below 3.0

Table 1. Clinical Characteristics of 19 Cervical Squamous Carcinoma Samples.

Sample No.	Stage at Diagnosis*	Initial Response [†]	Tumor Suppression Ratio in Size (%)	Age	Status at p53 [‡]
16	IIIB	CR	100	69	WT
17	IIIB	CR	100	65	WT
23	IIIB	CR	100	67	WT
47	IVA	CR	100	67	WT
74	IIIB	CR	100	77	WT
75	IIIB	CR	100	32	WT
81	IVB	CR	100	55	WT
83	IIIB	CR	100	59	WT
89	IIIB	CR	100	63	WT
31	IIIB	NC	40	53	WT
35	IIIB	NC	19	47	WT
39	IIIB	NC	0	70	WT
45	IIIB	NC	40	59	WT
52	IVB	NC	5	56	WT
53	IVB	NC	5	63	WT
55	IVB	NC	0	76	WT
85	IVB	NC	0	52	WT
87	IVB	NC	0	84	WT
96	IVB	NC	10	53	WT

*Tumors were staged according to International Federation of Gynecology and Obstetrics criteria.

[†]CR, tumor suppression ratio of 100%. NC, 50% suppression to +25% growth a month after radiotherapy.

[‡]Mutational status of p53. WT, wild type.

for both Cy3 and Cy5, genes corresponding to those spots were not investigated any further. To investigate genes that were clearly expressed differently between CR and NC tumors, the Mann-Whitney test was applied based on gene-expression values of X , where X =the Cy5/Cy3 signal intensity ratio for each gene and for each sample. U values for Mann-Whitney test were calculated for each gene. Genes with U values lower than 20 or greater than 70 were selected ($P < .05$ for comparing 9 CR samples vs. 10 NC samples). Because the U values were calculated for each sample in the CR group against each sample in the NC group for each gene based on each X value, genes that have U value lower than 20 indicate upregulated in the CR group compared to the NC group. However, genes that have U value more than 70 indicate upregulated in the NC group compared to the CR group. 297 genes were upregulated in the CR group and 132 genes were upregulated in the NC group. However, because more than half of these genes have small differences in expression level between CR and NC group, the difference might be caused by data fluctuation. Therefore, genes showing differences more than double the median expression value between the two groups ($\mu X_{CR}/\mu X_{NC} \leq 0.5$ or ≥ 2.0 , where μX_{CR} and μX_{NC} indicate median X values for the CR or NC group, respectively) were defined as radiosensitivity (or radioresistance) related genes. A total of 171 genes were selected (121 were significantly greater in radiosensitive cells than in radioresistant cells, and 50 were higher levels of expression in radioresistant cells than in radiosensitive cells).

Permutation Test

To further evaluate the validity of 171 genes selected by Mann-Whitney tests, permutation test was performed as described previously [18], and the probabilities of the genes to be correlated to group distinction, P_s , were also estimated. When each gene was represented by expression vector $v(g) = (X_1, X_2, \dots, X_{19})$, where X_i denotes the expression level of gene of the i th sample in the initial set of samples, idealized expression patterns were represented by $c = (c_1, c_2, \dots, c_{19})$, where $c_i = +1$ or 0 according to whether the i th sample belongs to the CR or NC group. The correlation between a gene and a group distinction P_{gc} was defined as follows: i.e., $P_{gc} = (\mu_{CR} - \mu_{NC}) / (\sigma_{CR} + \sigma_{NC})$, where μ_{CR} (μ_{NC}) and σ_{CR} (σ_{NC}) indicate the means and the standard deviations of $\log_2 X$ of the gene "g" for each sample in newly defined CR (NC) group. Permutation test was conducted by permuting the coordinates of c 10,000 times. During every permutation, the correlation values, P_{gcS} , were calculated. These procedures were performed 10,000 times, repeatedly. On the hypothesis that these obtained 10,000 A_g values show ideal normal distribution, P values, which imply probability of the genes to classify the two groups by chance was estimated for each selected 171 gene.

Hierarchical Clustering

These 171 genes were subjected to a hierarchical-clustering protocol using "Cluster" and "Tree view" software written by M. Eisen [17]. Before applying the clustering

algorithm, gene-expression values (X) for each gene in each of the 19 samples were log-transformed ($\log_2 X$). After that all values in each row and/or column of data were multiplied by scale factor S , so that the sum of the squares of the values in each row and column was 1.0 (a separate S is computed for each row/column). Next, row-wise and column-wise median values were subtracted from the values in each row and/or column data, so that the median value of each row and/or column is 0. Hierarchical clustering was performed using distance metrics based on Pearson correlation and adopting Average Linkage Clustering method.

Cross-Validation Test

The selected 62 genes from 171 genes by clustering experiment were subjected to cross-validation test. Among the total 19 samples tested above, one sample was withheld as test sample and the other 18 samples were used for building predictor according to the method as described previously [18]. Next, predictive score (PS) for test sample was calculated as follow; $PS = \sum V_g$, where $V_g = A_g'(X_g - B_g)$, $A_g' = (\mu_{CR}' - \mu_{NC}') / (\sigma_{CR}' + \sigma_{NC}')$, and $B_g = (\mu_{CR}' + \mu_{NC}') / 2$; μ_{CR}' (μ_{NC}') and σ_{CR}' (σ_{NC}') indicates the means and standard deviations of $\log_2 X$ of the gene "g" for each sample in the CR (NC) group defined as predictor samples. X_g denotes the $\log_2 X$ of the gene "g" for test sample. Finally, the predictor sample and test sample were changed and then PS for new test sample was calculated. This process was performed 19 times, repeatedly.

Results

Identification of Genes Responding to Radiation

We performed cDNA microarray analysis of gene expression in 19 cervical-cancer materials, of which 9 were radiation sensitive and 10 were radiation resistant on clinical grounds. By means of the Mann-Whitney test ($P < .05$) and subsequent procedures (see Materials and Methods section), we selected a total of 171 genes (including 74 ESTs) as being differently expressed between CR (complete response) and NC (no change) groups. Of those 171 genes, 121 (including 62 ESTs) revealed increased expression, and 50 (including 12 ESTs) showed decreased expression, in carcinomas belonging to the CR group compared with the NC group (Table 2, A and B). Genes involved in adipogenesis and in the MAP kinase pathway were significantly upregulated in the CR group compared to the NC group; the former included aldehyde dehydrogenase 1 (*ALDH1*), and retinol-binding protein 1 (*RBP1*; Figure 1A). The latter included mitogen-activated protein kinase kinase 2 (*MAP3K2*), G protein beta subunit-like (*GBL*), and *RAB5C* (a member of the RAS oncogene family). However, genes that are considered to be associated with repair of breaks in double-stranded DNA, including X-ray repair cross-complementing 5 (*XRCC5*; Figure 1B) were downregulated in the CR group relative to

Table 2. Genes Showing Different Expression Between CR and NC Groups.

Category	Unigene ID (build #131)	Gene Symbol	Gene Name	U	CR/NC	P	Locus
A. Genes Showing Relatively Higher Expression in Radiosensitive Carcinoma Cells than in Radioresistant Cells							
DNA repair	Hs.3248	MSH6	mutS (E. coli) homolog 6	13	2.0	0.003	2p16
Signal transduction	Hs.76578	PIAS3	protein inhibitor of activated STAT3	10	4.5	0.000	1q21
	Hs.29203	GBL	G protein beta subunit-like	11	2.8	0.044	16
	Hs.479	RAB5C	RAB5C, member RAS oncogene family	12	2.8	0.001	17q21.2
	Hs.28827	MAP3K2	mitogen-activated protein kinase kinase kinase 2	12	2.4	0.002	2
	Hs.85155	BRF1	butyrate response factor 1 (EGF-response factor 1)	14	2.4	0.003	14q22-q24
	Hs.74615	PDGFRA	platelet-derived growth factor receptor, alpha polypeptide	14.5	3.3	0.023	4q11-q13
	Hs.77439	PRKAR2B	protein kinase, cAMP-dependent, regulatory, type II, beta	15	5.9	0.009	7q22-q31.1
Transcription	Hs.83070	GRB14	growth factor receptor-bound protein 14	17	9.8	0.042	2q22-q24
	Hs.66394	RNF4	ring finger protein 4	6	5.1	0.000	4p16.3
	Hs.8858	BAZ1A	bromodomain adjacent to zinc finger domain, 1A	9	5.0	0.031	14q12-q13
	Hs.155321	SRF	serum response factor (c-fos serum response element-binding transcription factor)	9	2.0	0.038	6pter-6q15
	Hs.289068	TCF4	transcription factor 4	15	2.2	0.048	18q21.1
	Hs.760	GATA2	GATA-binding protein 2	15	2.0	0.011	3q21
	Hs.316	DDX6	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 6	16	3.3	0.009	11q23.3
	Hs.301963	HOXD8	homeo box D8	17	3.3	0.038	2q31-q37
Adipogenesis	Hs.228059	TIF1B	KRAB-associated protein 1	19	2.0	0.029	5
	Hs.76392	ALDH1	aldehyde dehydrogenase 1, soluble	15.5	14.1	0.018	9q21
	Hs.101850	RBP1	retinol-binding protein 1, cellular	19	2.1	0.018	3q23
Cytoskeleton	Hs.7645	FGB	fibrinogen, B beta polypeptide	9	4.4	0.018	4q28
	Hs.75279	LAMA2	laminin, alpha 2 (merosin, congenital muscular dystrophy)	10	5.2	0.040	6q22-q23
	Hs.75445	SPARCL1	SPARC-like 1 (mast9, hevjin)	15	4.0	0.004	7
	Hs.97266	PCDH18	protocadherin 18	16	5.9	0.006	
	Hs.11494	FBLN5	fibulin 5	16	4.0	0.008	14q32.1
	Hs.6441	TIMP2	tissue inhibitor of metalloproteinase 2	17	2.8	0.033	17q25
	Hs.79914	LUM	lumican	17	2.2	0.035	12q21.3-q22
	Hs.108896	LOC51084	lambda-crystallin	18	4.7	0.022	13cen3q14.2
	Hs.20072	MIR	myosin regulatory light-chain interacting protein	19	2.4	0.015	6p23-p22.3
Immune system	Hs.1244	CD9	CD9 antigen (p24)	8	2.2	0.001	12p13
	Hs.74631	BSG	basigin	8	2.2	0.001	19p13.3
	Hs.502	ABCB3	ATP-binding cassette, subfamily B (MDR/TAP), member 3	11	2.1	0.009	6p21.3
	Hs.24395	SCYB14	small inducible cytokine subfamily B (Cys-X-Cys), member 14 (BRAK)	17	5.0	0.007	5q31
Proteolysis	Hs.173091	UBL3	ubiquitin-like 3	15	2.5	0.016	13q12-q13
	Hs.75275	UBE4A	ubiquitination factor E4A (homologous to yeast UFD2)	19	3.8	0.047	11
Tumor related	Hs.81988	DAB2	disabled (Drosophila) homolog 2 (mitogen-responsive phosphoprotein)	16	2.4	0.003	5p13
	Hs.75462	BTG2	BTG family, member 2	16	2.1	0.028	1q32
Peptide hormone	Hs.134932	UCN	urocortin	9	11.3	0.017	2p23-p21
Others and ESTs	Hs.150926	FPGT	fucose-phosphate guanylyltransferase	10	2.1	0.009	1
	Hs.74566	DPYSL3	dihydropyrimidinase-like 3	18	2.6	0.015	5q32
	Hs.101735	DKFZP564J102	DKFZP564J102 protein	19	3.0	0.042	4
	Hs.74571	ARF1	ADP-ribosylation factor 1	19	2.9	0.015	1q42
	Hs.56874	HSPB7	heat shock 27-kDa protein family, member 7 (cardiovascular)	4	2.5	0.005	1p36.23-p34.3
	Hs.74376	NOE1	olfactomedin-related ER localized protein	7	22.4	0.028	9
	Hs.112569	GAN	giant axonal neuropathy (gigaxonin)	12	2.6	0.001	16q24.1
	Hs.7535	LOC55871	COBW-like protein	13	2.3	0.001	2
	Hs.20597	LCP	host cell factor homolog	13	2.0	0.002	
	Hs.24948	SNCAIP	synuclein, alpha interacting protein (synphilin)	14	7.0	0.003	5q23.1-q23.3
	Hs.108725	LOC51660	HSPC040 protein	15	3.0	0.005	6
	Hs.49912	LOC55895	22-kDa peroxisomal membrane protein-like	15	2.7	0.015	
	Hs.78103	NAP1L4	nucleosome assembly protein 1-like 4	17	7.0	0.009	11p15.5
	Hs.111779	SPARC	secreted protein, acidic, cysteine-rich (osteonectin)	17	2.2	0.020	5q31.3-q32
	Hs.75887	COPA	coatamer protein complex, subunit alpha	17	2.0	0.028	1q23-q25
	Hs.129872	SPAG9	sperm-associated antigen 9	17	2.9	0.001	17
	Hs.62041	NID	nidogen (enactin)	17	11.3	0.032	1q43
	Hs.83834	CYB5	cytochrome b-5	18	3.6	0.040	18q23
	Hs.29981	SLC26A2	solute carrier family 26 (sulfate transporter), member 2	19	2.8	0.001	5q31-q34

Table 2. (continued)

Category	Unigene ID (build #131)	Gene Symbol	Gene Name	U	CR/NC	P	Locus
	Hs.17930	BING4	BING4	19	2.9	0.000	6p21.3
	Hs.6113	STAU	staufen (Drosophila, RNA-binding protein)	19	2.0	0.002	20q13.1
	Hs.11951	ENPP1	ectonucleotide pyrophosphatase/phosphodiesterase 1	19.5	>50	0.047	6q22-q23
	Hs.172870		ESTs	4	>50	0.000	
	Hs.29664		Human DNA sequence from clone 682J15 on chromosome 6p11.22.3.	6	15.1	0.001	
	Hs.12365	KIAA1427	KIAA1427 protein	7	>50	0.004	
	Hs.107515		ESTs	7	3.5	0.000	
	Hs.40583		Homo sapiens clone TCBP1028 mRNA sequence	7	4.3	0.004	
	Hs.115315		ESTs	7	2.2	0.037	
	Hs.176092		ESTs, moderately similar to myosin-binding protein H [H. sapiens]	7	7.5	0.000	
	Hs.125291		ESTs	7	10.0	0.000	
	Hs.85053		<i>H. sapiens</i> clone 24440 mRNA sequence	8	11.2	0.000	
	Hs.188228		<i>H. sapiens</i> cDNA FLJ11003 fis, clone PLACE1002851	9	4.0	0.002	
	Hs.296772		Human DNA sequence from clone RP1-292B18	10	4.1	0.001	
	Hs.83724		Human clone 23773 mRNA sequence	10	6.8	0.033	
	Hs.124558		EST	10	2.8	0.008	
	Hs.120399		ESTs	10	8.2	0.016	
	Hs.116464		ESTs	11	3.4	0.014	
	Hs.281434		<i>H. sapiens</i> cDNA FLJ14028 fis, clone HEMBA1003838	11	2.9	0.031	
	Hs.191271		ESTs	11	5.5	0.003	
	Hs.26676	FLJ10850	hypothetical protein FLJ10850	11	4.8	0.004	20pter-20q12
	Hs.292162		ESTs	11	>50	0.007	
	Hs.181304	13CDNA73	putative gene product	11	>50	0.000	13
	Hs.98265		ESTs	11	10.9	0.000	
	Hs.126857		<i>H. sapiens</i> cDNA FLJ12936 fis, clone NT2RP2005018	11	20.5	0.014	
	Hs.95071		ESTs	11.5	>50	0.001	
	Hs.110373		ESTs	12	4.7	0.007	
	Hs.114453		ESTs	12	2.6	0.005	
	Hs.15725	LOC51278	hypothetical protein SBBI48	12	2.7	0.003	1p36.13q41
	Hs.291979		ESTs, Highly similar to pre-mRNA splicing SR protein ra4	12	4.6	0.019	
	Hs.113082	KIAA0443	KIAA0443 gene product	13	6.3	0.019	X
	Hs.116117		EST	13	2.2	0.003	
	Hs.24391		<i>H. sapiens</i> cDNA FLJ13612 fis, clone PLACE1010833	13	5.9	0.034	
	Hs.23120		<i>H. sapiens</i> cDNA: FLJ21421 fis, clone COL04123	13	2.1	0.007	
	Hs.116585		ESTs	13	2.0	0.014	
	Hs.49476		<i>H. sapiens</i> clone TUA8 Cri-du-chat region mRNA	14	10.8	0.010	
	Hs.112745		EST	14	2.1	0.029	
	Hs.21851		<i>H. sapiens</i> cDNA FLJ12900 fis, clone NT2RP2004321	14	2.7	0.001	
	Hs.13809		ESTs	14	2.8	0.007	
	Hs.61268		ESTs	14	>50	0.032	
	Hs.8469		ESTs	15	2.0	0.049	
	Hs.11365		<i>H. sapiens</i> cDNA FLJ12145 fis, clone MAMMA1000395	15	2.1	0.006	
	Hs.107812		ESTs, Weakly similar to SPOP [H. sapiens]	15	2.8	0.027	
	Hs.25329		ESTs	15	4.4	0.033	
	Hs.178730		ESTs	16	2.3	0.000	
	Hs.112607		ESTs	16	2.6	0.023	
	Hs.27497		<i>H. sapiens</i> cDNA FLJ11756 fis, clone HEMBA1005595	16	2.5	0.002	
	Hs.29356		ESTs	16	2.0	0.024	
	Hs.158688	IF2	KIAA0741 gene product	16	4.6	0.042	2
	Hs.23617	FLJ20531	hypothetical protein FLJ20531	17	2.8	0.041	
	Hs.44159	LOC51105	CGI-72 protein	17	13.7	0.003	8
	Hs.23650		ESTs, Weakly similar to AAB47496 NG5 [H. sapiens]	17	3.3	0.041	
	Hs.133081		ESTs, Weakly similar to hypothetical protein [H. sapiens]	17	24.6	0.011	
	Hs.191379		ESTs	17	6.1	0.003	
	Hs.72363		<i>H. sapiens</i> mRNA for FLJ00116 protein, partial cds	17.5	>50	0.038	
	Hs.173094		<i>H. sapiens</i> mRNA; cDNA DKFZp564H142 (from clone DKFZp564H142)	17.5	6.0	0.006	
	Hs.179891		ESTs, Weakly similar to prolyl 4-hydroxylase alpha subunit [H. sapiens]	18	2.5	0.005	
	Hs.22860		ESTs	18	2.3	0.011	

(continued on next page)

Table 2. (continued)

Category	Unigene ID (build #131)	Gene Symbol	Gene Name	U	CR/NC	P	Locus
	Hs.12867		ESTs	18	2.3	0.005	
	Hs.200332	FLJ20651	hypothetical protein FLJ20651	18	3.8	0.004	9p24.1-9q22.33
	Hs.273186	LOC56997	hypothetical protein, clone Telethon(Italy_B41)_Strait02270_FL142	18	2.8	0.041	1
	Hs.30643		ESTs	18	5.7	0.018	
	Hs.11805		ESTs	19	3.1	0.038	
	Hs.22505	FLJ10159	hypothetical protein FLJ10159	19	2.7	0.025	6
	Hs.127407		ESTs	19	5.6	0.011	
B. Genes Showing Relatively Higher Expression in Radioresistant Carcinoma Cells than in Radiosensitive Cells							
Category	Unigene ID (build #131)	Gene Symbol	Gene Name	U	CR/NC	P	Locus
DNA repair	Hs.84981	XRCC5	X-ray repair complementing defective repair in Chinese hamster cells 5	73	0.39	0.034	2q35
Signal transduction	Hs.155924	CREM	cAMP responsive element modulator	71	0.48	0.020	10p12.1-p11.1
	Hs.118520	LOC55970	G-protein gamma2 subunit	73	0.40	0.015	1
	Hs.34780	DCX	doublecortin; lissencephaly, X-linked (doublecortin)	74	0.48	0.012	Xq22.3-q23
	Hs.7138	CHRM3	cholinergic receptor, muscarinic 3	77	0.50	0.034	1q41-q44
	Hs.250857	CAMK2G	calcium/calmodulin-dependent protein kinase II gamma	85	0.41	0.009	10q22
Transcription	Hs.168005	TIF1GAMMA	transcriptional intermediary factor 1 gamma	72	0.38	0.000	1p13.1
	Hs.21771	WHSC2	Wolf-Hirschhorn syndrome candidate 2	73	0.47	0.021	4p16.3
	Hs.172280	SMARCC1	SWI/SNF-related, matrix-associated, actin-dependent regulator of chromatin	74	0.38	0.005	3p23-p21
	Hs.110457	WHSC1	Wolf-Hirschhorn syndrome candidate 1	78	0.45	0.005	4p16.3
	Hs.78580	DDX1	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 1	73	0.32	0.032	2p24
Translation	Hs.129673	EIF4A1	eukaryotic translation initiation factor 4A, isoform 1	78	0.43	0.001	17p13
Glycolysis	Hs.2795	LDHA	lactate dehydrogenase A	75	0.39	0.036	11p15.4
Cytoskeleton	Hs.821	BGN	biglycan	71	0.24	0.026	Xq28
	Hs.172928	COL1A1	collagen, type I, alpha 1	74	0.46	0.007	17q21.3-q22
	Hs.90408	NEO1	neogenin (chicken) homolog 1	78	0.48	0.006	15q22.3-q23
Immune system	Hs.516	CCR1	chemokine (C-C motif) receptor 1	71	0.28	0.039	3p21
	Hs.198253	HLA-DQA1	major histocompatibility complex, class II, DQ alpha 1	72.5	0.32	0.007	6p21.3
	Hs.75498	SCYA20	small inducible cytokine subfamily A (Cys-Cys), member 20	73	0.18	0.041	2q33-q37
	Hs.833	ISG15	interferon-stimulated protein, 15 kDa	78	0.25	0.042	1
Proteolysis	Hs.61153	PSMC2	prosome (prosome, macropain) 26S subunit, ATPase, 2	73	0.49	0.032	7q22.1-q22.3
Apoptosis	Hs.93213	BAK1	BCL2-antagonist/killer 1	73	0.36	0.006	6p21.3
	Hs.278602	API5	apoptosis inhibitor 5	74	0.46	0.013	11
Others and ESTs	Hs.75593	UROS	uroporphyrinogen III synthase	71	0.13	0.041	10q25.2-q26.3
	Hs.108196	LOC51659	HSPC037 protein	71	0.37	0.027	16
	Hs.64595	AASDHPPT	aminoadipate-semialdehyde dehydrogenase-phosphopantetheinyl transferase	71	0.44	0.033	11q22
	Hs.286049	PSA	phosphoserine aminotransferase	72	0.47	0.009	9
	Hs.93659	ERP70	protein disulfide isomerase-related protein	73	0.26	0.036	10
	Hs.2281	CHGB	chromogranin B (secretogranin 1)	75	0.31	0.011	20pter-p12
	Hs.110099	CBFA2T3	core-binding factor, runt domain, alpha subunit 2; translocated to, 3	75	0.33	0.009	16q24
	Hs.4747	DKC1	dyskeratosis congenita 1, dyskerin	75	0.40	0.003	Xq28
	Hs.83848	TPI1	triosephosphate isomerase 1	77	0.43	0.040	12p13
	Hs.75799	PRSS8	protease, serine, 8 (prostasin)	77	0.13	0.045	16p11.2
	Hs.143600	GPP130	type II Golgi membrane protein	77	0.35	0.024	3
	Hs.13565	T-STAR	Sam68-like phosphotyrosine protein, T-STAR	78	0.44	0.013	8q24.2
	Hs.169476	GAPD	glyceraldehyde-3-phosphate dehydrogenase	80	0.47	0.042	12p13
	Hs.114366	PYCS	pyrroline-5-carboxylate synthetase	87	0.24	0.006	10q24.3
	Hs.43445	PARN	poly(A)-specific ribonuclease (deadenylation nuclease)	90	0.35	0.003	16p13
	Hs.137556		H. sapiens mRNA; cDNA DKFZp434A132	71	0.18	0.004	
	Hs.65403	LOC51323	hypothetical protein	71	0.50	0.003	6pter-6q15
	Hs.164285		ESTs, Weakly similar to Afg1p [S. cerevisiae]	72	0.44	0.026	
	Hs.26675		ESTs	74	0.32	0.012	
	Hs.11641		H. sapiens cDNA: FLJ21432 fis, clone COL04219	74	0.50	0.000	

Table 2. (continued)

Category	Unigene ID (build #131)	Gene Symbol	Gene Name	U	CR/NC	P	Locus
	Hs.14846		<i>H. sapiens</i> mRNA; cDNA DKFZp564D016	75	0.38	0.009	
	Hs.283127		ESTs	75	0.36	0.032	
	Hs.63224		ESTs	75	0.16	0.038	
	Hs.227591		ESTs, Weakly similar to AF1488561 unknown [<i>H. sapiens</i>]	76	0.50	0.017	
	Hs.11156	LOC51255	hypothetical protein	76	0.21	0.005	2
	Hs.133207		<i>H. sapiens</i> mRNA for KIAA1230 protein, partial cds	77	0.42	0.011	
	Hs.201925		<i>H. sapiens</i> cDNA FLJ13446 fis, clone PLACE1002968	80	0.47	0.015	

U, indicates Mann-Whitney statistics. CR/NC, difference ratio between median expression values for each group. P, permutational P value. Genes used for calculating predictive scores are noted in bold type.

the NR group. In addition, a number of genes related to DNA repair, transcription, signal transduction, cell skeleton, and adhesion were among those expressed differently in the two groups.

Permutation Test

To evaluate the validity of the 171 genes selected as radiosensitivity-related genes, permutation test was performed as described in the Materials and Methods section. Expression levels of each 19 samples in both groups for each gene were permuted (randomly scrambled) 10,000 times. P_{gc} values were calculated using μ_{CR} , μ_{NC} , σ_{CR} , and σ_{NC} values derived from newly classified CR group and NC group during every permutation. Large absolute values of P_{gc} indicate a strong correlation between the gene expression and the class distinction, whereas the sign of P_{gc} being positive or negative corresponds to gene "g" being more highly expressed in the CR or NC group. After the 10,000 times permutation, the probabilities of the genes to be correlated to group distinction, P_s , were estimated under the hypothesis that these 10,000 P_{gc} values show ideal normal distribution (Table 2, A and B). As a result, all of the selected

62 genes showed P values >.05 without exception. Hence, it was proved that these selected 171 genes were to be radiosensitivity predictive gene under the confidence of $P < .05$.

Hierarchical Clustering

Of these 171 genes were subjected to hierarchical clustering as described in the Materials and Methods section. This procedure clearly separated the two groups from each other, except for tumor No. 47 (data not shown). To achieve complete separation, we selected 62 genes that showed greater than 2.0 standard deviations of expression values among all 19 samples. Cluster analysis using these 62 genes achieved complete separation of the groups (Figure 2).

Cross-Validation Test

Cross-validation test was performed to examine whether the 62 genes were crucial for classifying CR and NC groups and whether they could predict the group for test samples. Among the 19 samples, 18 samples were used for group predictor and 1 sample was used as the test sample. The

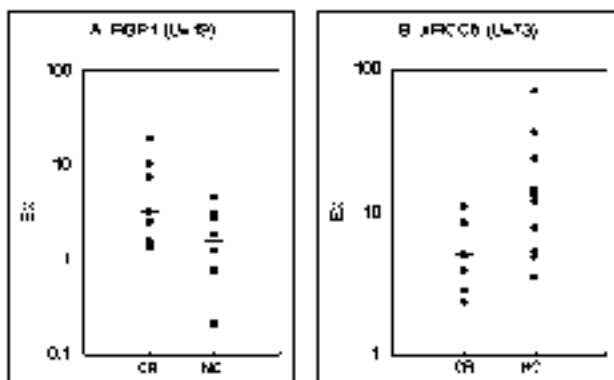


Figure 1. Differential gene expression between the radiosensitive group (CR; 9 samples) and the radioresistant group (NC; 10 samples) with significant difference ($P < .05$). U indicates the Mann-Whitney test statistic. Expression levels ($Ex = Cy5$ signal intensity from cancer sample/ $Cy3$ signal intensity from control), of two genes are plotted here. Median Ex values for each group of tumors are denoted by horizontal lines. (A) Retinol-binding protein 1 (RBP1; $U = 19$); (B) X-ray repair cross-complementing 5 (XRCC5; $U = 73$).

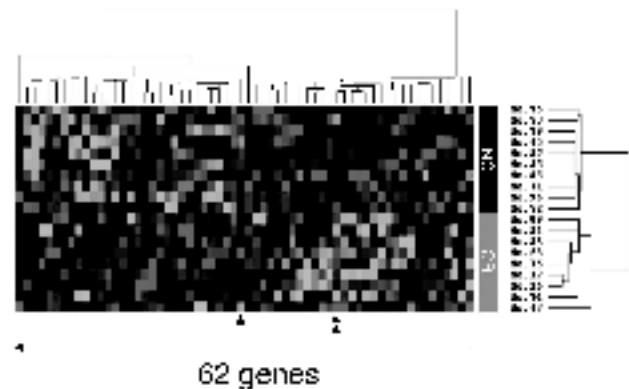


Figure 2. Expression patterns of 62 genes across 19 samples of cervical squamous cell carcinoma. Red or green colors indicate higher or lower expression, respectively, relative to the mean signal intensity of a given gene across 19 tumor samples; black, same expression level with mean value; gray, no expression detected (intensities of both $Cy3$ and $Cy5$ were below cut-off values). Each row represents each gene and each column a cervical squamous cell carcinoma sample. Single and double triangles indicate the gene-expression profiles of TCF4 and BAK1, respectively.

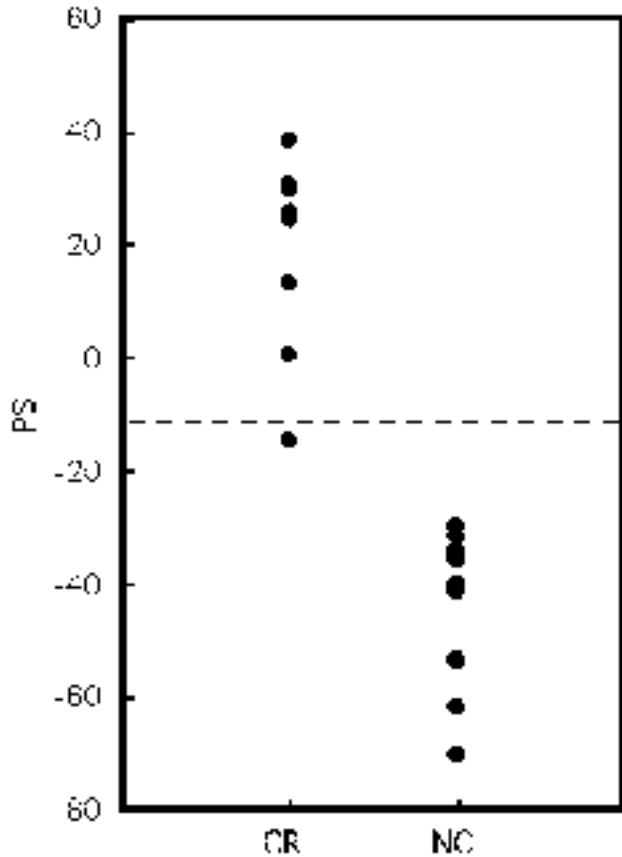


Figure 3. Predictive score (PS) for radiosensitive and radioresistant groups by cross-validation test. Details of the calculation method were noted in the Materials and Methods section. A dashed line indicates the threshold.

sample sets of predictor and test sample were changed 19 times and PS for each 19 samples was calculated as described in the Materials and Methods section (Figure 3 and Table 3). Threshold line for PSs to discriminate CR or NC group were settled at the half point between average PS values of the CR group and that of the NC group: -12 . As shown in Figure 3, PSs for samples in CR and NC groups were clearly separated, except for sample No.74 (error ratio was 5.3% (= 1 sample/19 samples)).

Discussion

cDNA Microarray analysis is a powerful tool for obtaining comprehensive information about expression of thousands of genes in cancer cells [12-15]. By combining this technology with statistical analysis, we identified 171 genes that showed different expression patterns between two distinct clinical groups and, therefore, were likely to reflect differences in the response of cervical cancer cells to radiotherapy. To examine the validity of 171 genes selected as radiosensitivity-related genes, random permutation test was performed by calculating P_{ges} , and P values for each gene were evaluated. After 10,000 times permutation test, P values for all of the 171 genes were lower than .05, indicating that these genes were significantly correlated to

radiosensitivity. Furthermore, in a clustering analysis, the expression profiles of these 171 genes were able to classify each of 19 tumor samples, except for one, to the appropriate group (radiosensitive or radioresistant). However, when the cluster analysis was limited to 62 genes having greater than 2.0 standard deviations of expression level across the 19 samples, all tumors were properly classified into their respective groups. To further evaluate the validity of 62 genes selected as radiosensitivity-related genes, we carried out cross-validation test. After 19 times cross-validation test, each PS value for each sample that belonged to the CR (or NC) group showed higher (or lower) value than threshold line. This study not only supports the feasibility of these 62 genes as radiosensitivity-related genes, but also indicates the possibility of predicting radiosensitivity for discriminating radiation-sensitive from radiation-resistant biopsy samples by PS values calculated from expression values of 62 genes. However, further study with additional tumor samples would be required to apply these genes for predicting radiosensitivity of tumors in patients before therapy begins.

Because all of the cervical cancer samples we used in this study were human papillomavirus (HPV)-positive, p53 as well as RB functions were likely to be eliminated by the viral protein [19]. Hence, the set of genes listed here may be associated with a cell-death pathway independent of p53. Radiation kills tumor cells mainly as a result of double-strand breaks (DSBs) in DNA [20]. If cells are defective in their DNA-repair systems, especially as regards DSB repair, they should be more susceptible to cell death. The gene product of *XRCC5*, Ku80, a protein that binds double-stranded DNA and a component of DNA-dependent protein kinase holoenzyme, is involved in DSB repair [21]. *XRCC5*-deficient cells and Ku80-knockout mice are hypersensitive to ionizing

Table 3. Predictive Scores to Classify Radiosensitive Group and Radio Resistant Group.

Sample No.	PS
<i>CR</i>	
16	36
17	30
23	38
47	0.6
74	-15
75	31
81	25
83	26
89	13
<i>NC</i>	
31	-48
35	-40
39	-54
45	-35
52	-62
53	-34
55	-70
85	-32
87	-42
96	-30

PS, predictive score. CR, radiosensitive group. NC, radioresistant group.

radiation [22,23]. Therefore, the higher expression of this gene we observed in radioresistant cancer cells accords with its physiological function. *XRCC5* would be one of the most crucial genes for determining the fate of cells under the genotoxic stress caused by irradiation. We also observed a relatively higher level of expression of lactate dehydrogenase (*LDHA*), which aids glycolysis under hypoxic conditions [24] in radioresistant cells. Cells in fact become radioresistant under hypoxic conditions [25], and a high level of *LDHA* expression could be an important mechanism conferring radioresistance.

In radiosensitive cells, we found elevated expression of adipogenesis-related genes including *ALDH1* and *RBP1*. The product of *ALDH1* is involved in retinoic acid (RA) synthesis [26] and *RBP1* is a transporter of retinol. Cervical carcinoma cells treated with RA before irradiation are reported to become radiosensitive [27]; furthermore, RA induces TRAIL expression and causes apoptosis [28]. Therefore, elevated expression of these genes may induce RA synthesis and, thereby, encourage apoptosis after radiation.

Our list of 171 genes should be useful not only as an aid to understanding the mechanism of radiosensitivity, but also as a means to expand the possibilities for effective radiotherapy. For example, if some novel drugs could block gene products that are involved in radioresistance, or if genes that induce apoptotic signals after radiation could be exogenously introduced, the effectiveness of radiotherapy would be increased. In addition, the 62 selected genes might prove of great benefit for diagnosing radiosensitivity of individual cervical cancers, to provide opportunities for selecting appropriate treatment (personalized medicine) for each patient.

Acknowledgements

We appreciate the help of Hiroko Bando, Noriko Nemoto, and Noriko Sudo in fabricating the cDNA microarray.

References

- Vutuc C, Haidinger G, Waldhoer T, Ahmad F, and Breitenecker G (1999). Prevalence of self-reported cervical cancer screening and impact on cervical cancer mortality in Austria. *Wien Klin Wochenschr* **111**, 354–59.
- Recio FO, Sahai Srivastava BI, Wong C, Hempling RE, Eltabakh GH, and Piver MS (1998). The clinical value of digene hybrid capture HPV DNA testing in a referral-based population with abnormal pap smears. *Eur J Gynaecol Oncol* **19**, 203–208.
- Suzuki K, Watanabe M, and Miyoshi J (1992). Differences in effects of oncogenes on resistance of gamma rays, ultraviolet light, and heat shock. *Radiat Res* **129**, 157–62.
- FitzGerald TJ, Henault S, Sakakeeny M, Santucci MA, Pierce JH, Anklesaria P, Kase K, Das I, and Greenberger JS (1990). Expression of transfected recombinant oncogenes increases radiation resistance of clonal hematopoietic and fibroblast cell lines selectively at clinical low dose rate. *Radiat Res* **122**, 44–52.
- Riva C, Lavielle JP, Reyt E, Brambilla E, Lunardi J, and Brambilla C (1995). Differential c-myc, c-jun, c-raf and p53 expression in squamous cell carcinoma of the head and neck: implication in drug and radioresistance. *Eur J Cancer Part B Oral Oncol* **31B**, 384–91.
- Matsumura Y, Yamagishi N, Miyakoshi J, Imamura S, and Takebe H (1997). Increase in radiation sensitivity of human malignant melanoma cells by expression of wild-type p16 gene. *Cancer Lett* **115**, 91–96.
- Hsiao M, Tse V, Carmel J, Costanzi E, Strauss B, Haas M, and Silverberg GD (1997). Functional expression of human p21(WAF1/CIP1) gene in rat glioma cells suppresses tumor growth in vivo and induces radiosensitivity. *Biochem Biophys Res Commun* **233**, 329–35.
- Gilbert MS, Saad AH, Rupnow BA, and Knox SJ (1996). Association of BCL-2 with membrane hyperpolarization and radioresistance. *J Cell Physiol* **168**, 114–22.
- Harima Y, Harima K, Shikata N, Oka A, Ohnishi T, and Tanaka Y (1998). Bax and Bcl-2 expressions predict response to radiotherapy in human cervical cancer. *J Cancer Res Clin Oncol* **124**, 503–10.
- Kitada S, Krajewski S, Miyashita T, Krajewska M, and Reed JC (1996). Gamma-radiation induces upregulation of Bax protein and apoptosis in radiosensitive cells *in vivo*. *Oncogene* **12**, 187–92.
- Schena M, Shalon D, Heller R, Chai A, Brown PO, and Davis RW (1996). Parallel human genome analysis: microarray-based expression monitoring of 1000 genes. *Proc Natl Acad Sci USA* **93**, 10614–19.
- Kitahara O, Furukawa Y, Tanaka T, Kihara C, Ono K, Yanagawa R, Nita M E, Takagi T, Nakamura Y, and Tsunoda T (2001). Alterations of gene expression during colorectal carcinogenesis revealed by cDNA microarrays after laser-capture microdissection of tumor tissues and normal epithelia. *Cancer Res* **61**, 3544–49.
- Ono K, Tanaka T, Tsunoda T, Kitahara O, Kihara C, Okamoto A, Ochiai K, Takagi T, and Nakamura Y (2000). Identification by cDNA microarray of genes involved in ovarian carcinogenesis. *Cancer Res* **60**, 5007–11.
- Okabe H, Satoh S, Kato T, Kitahara O, Yanagawa R, Yamaoka Y, Tsunoda T, Furukawa Y, and Nakamura Y (2001). Genome-wide analysis of gene expression in human hepatocellular carcinomas using cDNA microarray: identification of genes involved in viral carcinogenesis and tumor progression. *Cancer Res* **61**, 2129–37.
- Matsushima-Nishiu M, Unoki M, Ono K, Tsunoda T, Minaguchi T, Kuramoto H, Nishida M, Satoh T, Tanaka T, and Nakamura Y (2001). Growth and gene expression profile analyses of endometrial cancer cells expressing exogenous PTEN. *Cancer Res* **61**, 3741–49.
- Luo L, Salunga RC, Guo H, Bittner A, Joy KC, Galindo JE, Xiao H, Rogers KE, Wan JS, Jackson MR, and Erlander MG (1999). Gene expression profiles of laser-captured adjacent neuronal subtypes. *Nat Med* **5**, 117–22.
- Eisen MB, Spellman PT, Brown PO, and Botstein D (1998). Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA* **95**, 14863–68.
- Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, Bloomfield CD, and Lander ES (1999). Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* **286**, 531–37.
- Scheffner M, Werness BA, Huibregtse JM, Levine AJ, and Howley PM (1990). The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. *Cell* **63**, 1129–36.
- Khanna KK, and Jackson SP (2001). DNA double-strand breaks: signaling, repair and the cancer connection. *Nat Genet* **27**, 247–54.
- Errami A, Smider V, Rathmell WK, He DM, Hendrickson EA, Zdzienicka MZ, and Chu G (1996). Ku86 defines the genetic defect and restores X-ray resistance and V(D)J recombination to complementation group 5 hamster cell mutants. *Mol Cell Biol* **16**, 1519–26.
- Zhu C, Bogue MA, Lim DS, Hasty P, and Roth DB (1996). Ku86-deficient mice exhibit severe combined immunodeficiency and defective processing of V(D)J recombination intermediates. *Cell* **86**, 379–89.
- Nussenzweig A, Chen C, da Costa Soares V, Sanchez M, Sokol K, Nussenzweig MC, and Li GC (1996). Requirement for Ku80 in growth and immunoglobulin V(D)J recombination. *Nature* **382**, 551–55.
- Shim H, Chun YS, Lewis BC, and Dang CV (1998). A unique glucose-dependent apoptotic pathway induced by c-Myc. *Proc Natl Acad Sci USA* **95**, 15111–16.
- Chao KS, Bosch WR, Mutic S, Lewis JS, Dehdashti F, Mintun MA, Dempsey JF, Perez CA, Purdy JA, and Welch MJ (2001). A novel approach to overcome hypoxic tumor resistance: Cu-ATSM-guided intensity-modulated radiation therapy. *Int J Radiat Oncol Biol Phys* **49**, 1171–82.
- Bhat PV, Labrecque J, Boutin JM, Lacroix A, and Yoshida A (1995). Cloning of a cDNA encoding rat aldehyde dehydrogenase with high activity for retinal oxidation. *Gene* **166**, 303–306.
- Benbrook DM, Shen-Gunther J, Nunez ER, and Dynlacht JR (1997). Differential retinoic acid radiosensitization of cervical carcinoma cell lines. *Clin Cancer Res* **3**, 939–45.
- Altucci L, Rossin A, Raffelsberger W, Reitmaier A, Chomienne C, and Gronemeyer H (2001). Retinoic acid-induced apoptosis in leukemia cells is mediated by paracrine action of tumor-selective death ligand TRAIL. *Nat Med* **7**, 680–86.