NEOPLASIA
www.neoplasia.com

Volume 10 Number 11 November 2008 pp. 1222–1230 **1222**

Chromosomal Instability Is Associated with Higher Expression of Genes Implicated in Epithelial-Mesenchymal Transition, Cancer Invasiveness, and Metastasis and with Lower Expression of Genes Involved in Cell Cycle Checkpoints, DNA Repair, and Chromatin Maintenance[1]

Anna V. Roschke*, Oleg K. Glebov*, Samir Lababidi†, Kristen S. Gehlhaus*, John N. Weinstein† and Ilan R. Kirsch*

*Genetics Branch, Center for Cancer Research, National Cancer Institute, Bethesda, MD 20889-5105, USA; †Laboratory of Molecular Pharmacology, Center for Cancer Research, National Cancer Institute, Bethesda, MD 20889-5105, USA

Abstract

Chromosomal instability—a hallmark of epithelial cancers—is an ongoing process that results in aneuploidy and karyotypic heterogeneity of a cancer cell population. Previously, we stratified cancer cell lines in the NCI-60 drug discovery panel based on their karyotypic complexity and heterogeneity. Using this stratification in conjunction with drug response data for the cell lines allowed us to identify classes of chemical compounds whose growth-inhibitory activity correlates with karyotypic complexity and chromosomal instability. In this article, we asked the question: What are the biological processes, pathways, or genes associated with chromosomal instability of cancer cells? We found that increased instability of the chromosomal content in a cancer cell population, particularly, persistent gains and losses of chromosomes, is associated with elevated expression of genes involved with aggressive cellular behavior, including invasion- and metastasis-associated changes in cell communication, adhesion, motility, and migration. These same karyotypic features are negatively correlated with the expression of genes involved in cell cycle checkpoints, DNA repair, and chromatin maintenance.

Neoplasia (2008) 10, 1222–1230

Introduction

Most cancers have an abnormal chromosomal content characterized by changes in chromosomal structure and number. Chromosomal aberrations tend to be more numerous in malignant tumors than in benign ones [1,2], and karyotypic complexity is associated with aggressive clinical behavior of tumors and poor prognoses [3–5]. Cancers contain cells that not only possess an abnormal number of chromosomes but also often show population heterogeneity with regard to the exact chromosomal complement. That heterogeneity is a marker of ongoing chromosomal instability in cancers—accelerated rates of gains and losses of whole chromosomes or large portions of cancer cell genomes [6,7]. Chromosomal instability is a process that facilitates cancer cell evolution under selection pressure, and because of instability and selection, tumors most frequently acquire an aneuploid chromosomal content.

The NCI-60 cancer cell line panel was assembled by the Developmental Therapeutics Program of the National Cancer Institute for *in vitro* anticancer drug screening. It includes human cancer cell lines

Abbreviations: EMT, epithelial-mesenchymal transition; NH, numerical chromosomal heterogeneity; NC, numerical chromosomal complexity; SH, structural chromosomal heterogeneity; SC, structural chromosomal complexity; GO, Gene Ontology; SCC, Spearman correlation coefficient

Address all correspondence to: Anna V. Roschke, Genetics Branch, Center for Cancer Research, National Cancer Institute, NNMC, 8901 Wisconsin Avenue, Building 8, Room 5101, Bethesda, MD 20889-5105. E-mail: roschkea@mail.nih.gov

[1]This article refers to supplementary materials, which are designated by Tables W1 to W4 and Figures W1 to W4 and are available online at www.neoplasia.com.

Received 12 June 2008; Revised 25 August 2008; Accepted 26 August 2008

DOI 10.1593/neo.08682

of lung, renal, colorectal, ovarian, breast, prostate, central nervous system (CNS), melanocyte, and hematological origins [8]. Since 1990, the NCI-60 cells have been exposed to more than 100,000 compounds in short-term cytotoxicity assays (http://dtp.nci.nih.gov) [9], and they have been profiled more extensively at the DNA, RNA, protein, and functional levels than any other set of cell lines [10]. Previously, we defined and studied karyotypic complexity and heterogeneity (as markers of ongoing instability) in the NCI-60 [11] and used this stratification to identify groups and classes of chemical compounds with higher growth-inhibitory activity in the more karyotypically complex and chromosomally unstable lines [12–14]. In the present work, we ask the question: What biological processes, pathways, and genes are associated with karyotypic instability of cancer cells? To address that question, we analyzed the correlation between karyotypic parameters and expression of genes across the NCI-60 panel.

Materials and Methods

Karyotypic Parameters of the NCI-60 Panel of Cancer Cell Lines

The NCI-60 panel of cancer cell lines was developed by the National Cancer Institute for *in vitro* anticancer drug screening, and it included cell lineages derived from different tissues (lung, renal, colorectal, ovarian, breast, prostate, CNS, melanoma, and hematological malignancies). Every cell line in the NCI-60 panel carries karyotypic abnormalities with notable individual variations among the cell lines at the level of karyotypic complexity and heterogeneity [11]. Complexity of the karyotype was defined on the basis of numerical abnormalities and structural rearrangements. Numerical chromosomal complexity (NC) was expressed in relation to the cell line ploidy level, according to the International System for Human Cytogenetic Nomenclature convention, and was calculated as a sum of the number of deviations of each specific chromosome from the designated ploidy level. Structural chromosomal complexity (SC) was expressed as the number of different structurally rearranged chromosomes present in two or more metaphases. Numerical chromosomal heterogeneity (NH) corresponds to the number of cell-to-cell variations of similar chromosomes. Loss of a chromosome in only one or two cells or gain in only one cell was not considered in the calculation of numerical heterogeneity because of the possibility of mechanical loss or gain during preparation of the metaphase spreads. Any specific chromosome displaying a higher number of gains or losses was considered to show variability and tallied as "1 point" in the numerical heterogeneity index. Structural chromosomal heterogeneity (SH) was estimated as the number of nonclonal (i.e., present in only one metaphase, according to the International System for Human Cytogenetic Nomenclature convention) structurally abnormal chromosomes per metaphase.

NCI-60 Gene Expression Data

We used mRNA expression databases for the NCI-60 cell lines from the HG-U95 and the HG-U133 Affymetrix sets. HG-U95A data set includes ~12,000 array features corresponding to 8978 genes. HG-U133 data set has ~22,000 array features corresponding to 13,032 genes [15]. Gene expression data for NCI-60 cell lines were generated by hybridization to U133A and U133B arrays (Affymetrix, Santa Clara, CA), and CEL files were downloaded from http://discover.nci.nih.gov/cellminer. Three cell lines were excluded from analysis for different reasons (MDA-N was not available for karyotypic analysis, NCI/ADR-RES was a derivative of OVCAR-8, and lung cancer cell line H23 had no gene expression data available).

Data Analysis

CEL files for 57 cell lines were imported into R-2.4.0 language and statistical computing environment [16] using the Affy package of BioConductor-1.8 [17]. Probe set summarization, convolution background correction, and quantile normalization were conducted using a robust multichip average method [18] separately for U133A and U133B array sets. Data were filtered to have expression measurements to be above 16 in at least 25% of the samples and the interquartile range across the samples to be at least 0.5 on $\log_2$ scale. This nonspecific filtering leaves 7500 probe sets for the U133A array set and 4036 probe sets for the U133B array set. Probe set expression measurements ($\log_2$-transformed) were combined into one data matrix (11,536 probe sets) that was transposed and aligned with karyotypic characteristics of cell lines for analysis of Spearman's rank correlation between gene expression and karyotypic parameters.

To calculate the false discovery rate (FDR) for each Spearman correlation coefficients (SCCs), 1000 random permutations of the karyotypic characteristic value (NC, NH, SC, or SH) were performed, and SCCs were recalculated for all 11,536 probe sets for each permutation. The number of probe sets having an SCC with a P value ≤ .001 per random permutation was used to estimate FDR as the number of such probe sets at 0.95 or 0.99 quantiles (corresponding percentiles were considered as confidence levels for FDR estimates).

Genes differentially expressed in cell lines depending on tissue of origin were identified separately for U133A and U133B arrays using univariate F test as it is implemented in BRB-Array Tools [19]. Genes were considered differentially expressed if their F test parametric P value was less than .001.

To identify the most relevant GO terms associated with gene lists, gene-GO term enrichment analysis was performed using the DAVID2007 Functional Annotation Tool with customized gene backgrounds [20,21].

Results

Correlation of Expression Profiles with Karyotypic Parameters

We obtained SCCs between karyotypic parameters (SC, SH, NC, and NH) and each of the 11,536 probe sets from HG-U133A and HG-U133B data set obtained after filtering. The highest number of statistically significant (P < .001) correlations was identified for NH (454 array elements corresponding to 360 genes), followed by SH (318 array elements corresponding to 264 genes), NC (278 array elements corresponding to 224 genes), and SC (54 array elements corresponding to 50 genes; Table W1). To estimate the proportion of false-positives, 1000 random permutations of NH, NC, SH, and SC values were performed, and for each permutation, SCCs were calculated for all 11,536 probe sets. The density distributions of the number of SCCs with P value ≤ .001 for permuted sample were then analyzed. With 99% confidence, among 454 array elements correlated with NH for which SCCs have P values .001 or less, there were no more than 73 false-positive probe sets for which a positive or negative correlation of expression measurements with NH may be due to random variation in gene expression or NH (Table W1).

Because NH correlates with the highest number of gene expression profiles, and the number of possible false-positive results at $P < .001$ is the lowest for this parameter, correlations with gene expression profiles obtained for NH were chosen for further investigation.

Data for 454 probe sets for which the SCCs between NH and gene expression had P value .001, are presented in Table W2. Variability in SCCs for these 454 array elements was estimated by drawing 1000 bootstrap replicates and calculating the average bootstrap SCCs and corresponding quantiles as confidence intervals. For all 454 array elements, 99% bootstrap confidence intervals of SCCs do not include zero, remaining either positive or negative throughout.

Top Genes Associated with NH

Table 1 shows the top 50 Affy IDs positively associated with NH. These 50 Affy IDs correspond to only 39 gene transcripts because 8 genes (*ABL2*, *ATP6V0E*, *CRTAP*, *FN1*, *GNG12*, *IL13RA1*, *TIMP2*, and *TNPO1*) had more than one Affy ID selected in the top 50 positive list. On the basis of the SCC and P value, the three top genes with expression positively associated with NH are *VEGFC*, *leprecan-like 1*, and *Ras-related GTP-binding C* (Table 1). The list of 50 positive top genes is significantly enriched with genes that control cell communication (15 genes) and signal transduction (14 genes), response to wounding and cell motility (6 genes), cytokine–cytokine receptor interaction (5 genes), and cell migration (4 genes).

The top 50 Affy IDs (corresponding to 46 genes) expression of which negatively correlated with NH are shown in the Table 1*B*. In this top 50 negative list, *NUP210* is represented by four Affy IDs and *BTBD16* is represented by two. Both are among the three top genes negatively associated with NH. The third is *MLLT6* (Table 1). Among all 46 gene transcripts, 26 are significantly associated with intracellular organelles, 22 have membrane-bound products, and 18 are related to the nucleus.

Identification of Origin-Independent Transcriptional Changes Associated with NH

The NCI-60 panel includes cell lines derived from different tissues. It seems possible, therefore, that the differences in gene expression profiles that correlate with karyotype are actually determined by tissue of origin and that correlations with karyotypic parameters simply reflect that fact epiphenomenally. Indeed, cluster analysis (Supplementary Materials and Methods) shows that the main factor that affects pattern of gene expression in the NCI-60 cell lines is the tissue of origin of the cell line (Figures W1–W3). Even when we selected for this analysis the 454 probe sets whose expression correlates with NH, cell lines still tend to cluster according to the tissue of origin (Figure W4). However, in this case, there is also a correlation of the clustering pattern with NH (Supplementary Results).

To address this question of origin-independent transcriptional changes associated with NH, we used two approaches. First, to estimate possible biases imposed on karyotype correlations by tissue-of-origin groupings, we used a "jackknife" procedure: each tissue-of-origin was omitted in turn, and SCCs for the 454 probe sets that had SCCs between NH and gene expression with P values ≤ .001 were calculated for the excluded subset of cell lines. For all 454, 99% jackknife confidence intervals of SCCs excluded zero and remained positive or negative throughout, confirming that correlation of the expression of these transcripts with NH is not associated with any specific tissue-of-origin (data not shown).

Second, univariate F test identifies 1579 and 1098 probe sets in U133A and U133B arrays, respectively, as differentially expressed between cell lines grouped according to their 9 tissues of origin. Among them, 241 array features (corresponding to 172 genes) overlap with the previously identified ones significantly associated with NH ($P < .001$). Consequently, among the 454 Affymetrix IDs associated with NH at $P < .001$, 213 IDs corresponding to 188 genes are not differentially expressed based on their tissue of origin (129 features/121 genes negative, 84 features/67 genes positive; Table W3).

Discussion

Aneuploidy and chromosomal instability are common conditions for most epithelial cancer cells, but the relationships between those factors and cellular functions are not clear. We, therefore, used correlations between NH and gene expression profiles of the NCI-60 cell lines, followed by Gene Ontology (GO) categorization, to identify genes and cellular processes associated with the increase of chromosomal instability in cancer cells.

Gene Ontology analysis of the distribution of 360 genes correlated with NH ($P < .001$) indicated that cell communication and signal transduction, cell adhesion, motility, and migration, response to wounding and inflammatory response, negative regulation of cell proliferation, and DNA replication are the main biological processes associated with numerical heterogeneity of the chromosomal content in the cancer cells (Table 2). Moreover, when these genes were divided into two groups based on their positive or negative correlation coefficients, we saw a striking difference between these two groups. Genes, expression of which was positively correlated with NH, fell into GO categories such as cell communication and signal transduction, including cell surface receptor–linked signal transduction, cell adhesion, locomotion, motility, and migration, development, morphogenesis, and differentiation, response to wounding, and inflammatory response (Figure 1*A*). Products of these genes were associated with extracellular matrix and extracellular space, plasma membrane, and cytoskeleton and were involved in the focal adhesion pathway (HSA04510), cytokine–cytokine receptor interaction (HSA04060), regulation of actin cytoskeleton (HSA04810), JAK-STAT signaling pathways (HSA04630), cell communication (HSA01430), and ECM-receptor interaction pathways (HSA04512; Table 2).

Genes whose expression negatively correlated with NH fell into totally different GO categories: cellular metabolism, nucleic acid metabolism, regulation of transcription, DNA replication, response to DNA damage stimulus, DNA repair, chromosome organization and biogenesis, DNA packaging, unwinding and replication initiation, and base-excision repair (Figure 1*B* and Table 2). Products of those genes serve as transcription regulators, involved in nucleic acid binding, linked to ATP-ase activity, and associated with the cell cycle regulation pathway (HSA04110). They localize on intracellular organelles and are, for the most part, found in the nucleus, chromosome/chromatin, or nuclear envelope.

Among the 454 Affy IDs associated with NH at $P < .001$, 213 of them (corresponding to 188 genes) are not differentially expressed based on their tissue of origin (see the Results section). When we examined distribution of 188 genes across GO categories, genes whose expression is negatively associated with NH (121 genes) again were significantly enriched within such categories as intracellular membrane–bound organelle, nuclear protein, nucleic acid, or protein binding, regulation of transcription, DNA replication initiation, cell

Table 1. Top 50 Genes with Expression Profiles Correlated Positively (A) or Negatively (B) with NH across the NCI-60 Panel of Cancer Cell Lines.

Affy ID	Gene Symbol	Gene Name	ρ	P
(A)				
209946_at	*VEGFC*	Vascular endothelial growth factor C	0.5817	2.08e–06
218717_s_at	*LEPREL1*	Leprecan-like 1	0.5763	2.72e–06
218088_s_at	*RRAGC*	Ras-related GTP binding C	0.5750	2.89e–06
204140_at	*TPST1*	Tyrosylprotein sulfotransferase 1	0.5549	7.48e–06
226656_at	*CRTAP*	Cartilage-associated protein	0.5501	9.31e–06
235086_at	*THBS1*	Thrombospondin 1	0.5460	1.12e–05
212464_s_at	*FN1*	Fibronectin 1	0.5428	1.29e–05
220092_s_at	*ANTXR1*	Hypothetical protein FLJ10601	0.5425	1.30e–05
201426_s_at	*VIM*	Vimentin	0.5377	1.61e–05
202859_x_at	*IL8*	Interleukin 8	0.5349	1.82e–05
201172_x_at	*ATP6V0E*	ATPase, H+ transporting, lysosomal 9 kDa, V0 subunit E	0.5341	1.88e–05
211719_x_at	*FN1*	Fibronectin 1	0.5331	1.96e–05
210495_x_at	*FN1*	Fibronectin 1	0.5330	1.97e–05
201828_x_at	*CXX1*	CAAX BOX 1	0.5306	2.18e–05
201380_at	*CRTAP*	Cartilage-associated protein	0.5301	2.22e–05
237444_at		Unknown	0.5288	2.36e–05
205743_at	*STAC*	SH3 and cysteine-rich domain	0.5265	2.60e–05
201105_at	*LGALS1*	Lectin, galactoside-binding, soluble, 1 (galectin 1)	0.5257	2.67e–05
200096_s_at	*ATP6V0E*	ATPase, H+ transporting, lysosomal 9 kDa, V0 subunit E	0.5257	2.68e–05
204214_s_at	*RAB32*	RAB32, member *ras* oncogene family	0.5248	2.78e–05
200885_at	*RHOC*	Ras homolog gene family, member C	0.5242	2.85e–05
231907_at	*ABL2*	V-*abl* Abelson murine leukemia viral oncogene homolog 2 (ARG, Abelson-related gene)	0.5225	3.05e–05
226955_at	*FLJ36748*	Hypothetical protein FLJ36748	0.5195	3.46e–05
216442_x_at	*FN1*	Fibronectin 1	0.5172	3.79e–05
212509_s_at	*MXRA7*	FLJ46603 protein	0.5144	4.26e–05
222834_s_at	*GNG12*	Guanine nucleotide binding protein (G protein), gamma 12	0.5130	4.50e–05
224560_at	*TIMP2*	TIMP metallopeptidase inhibitor 2	0.5127	4.56e–05
211612_s_at	*IL13RA1*	Interleukin 13 receptor, alpha 1	0.5114	4.80e–05
226939_at	*CPEB2*	Cytoplasmic polyadenylation element binding protein 2	0.5110	4.88e–05
202378_s_at	*LEPROT*	Leptin receptor overlapping transcript	0.5107	4.93e–05
235072_s_at		Unknown	0.5104	4.99e–05
209226_s_at	*TNPO1*	Transportin 1	0.5086	5.36e–05
202733_at	*P4HA2*	Procollagen-proline, 2-oxoglutarate 4-dioxygenase (proline 4-hydroxylase), alpha polypeptide II	0.5083	5.43e–05
220407_s_at	*TGFB2*	Transforming growth factor, beta 2	0.5074	5.62e–05
212294_at	*GNG12*	Guanine nucleotide binding protein (G protein), gamma 12	0.5073	5.64e–05
202377_at	*LEPR*	Leptin receptor	0.5064	5.85e–05
209278_s_at	*TFPI2*	Tissue factor pathway inhibitor 2	0.5056	6.02e–05
201887_at	*IL13RA1*	Interleukin 13 receptor, alpha 1	0.5048	6.23e–05
209225_x_at	*TNPO1*	Transportin 1	0.5047	6.23e–05
208924_at	*RNF11*	Ring finger protein 11	0.5044	6.31e–05
212658_at	*LHFPL2*	Lipoma HMGIC fusion partner-like 2	0.5044	6.31e–05
213696_s_at	*MED8*	Mediator of RNA polymerase II transcription, subunit 8 homolog (yeast)	0.5043	6.33e–05
229465_s_at	*PTPRD*	Protein tyrosine phosphatase, receptor type, D	0.5043	6.34e–05
203262_s_at	*FAM50A*	Family with sequence similarity 50, member A	0.5037	6.48e–05
209013_x_at	*TRIO*	Triple functional domain (PTPRF interacting)	0.5034	6.56e–05
207657_x_at	*TNPO1*	Transportin 1	0.5026	6.78e–05
231579_s_at	*TIMP2*	TIMP metallopeptidase inhibitor 2	0.5013	7.12e–05
200998_s_at	*CKAP4*	Cytoskeleton-associated protein 4	0.5009	7.22e–05
229307_at	*ANKRD28*	Ankyrin repeat domain 28	0.5001	7.45e–05
226893_at	*ABL2*	V-*abl* Abelson murine leukemia viral oncogene homolog 2 (ARG, Abelson-related gene)	0.5000	7.48e–05
(B)				
226148_at	*BTBD15*	BTB (POZ) domain containing 15	–0.6236	2.20e–07
212316_at	*NUP210*	Nucleoporin 210 kDa	–0.6040	6.56e–07
224784_at	*MLLT6*	Myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, *Drosophila*); translocated to, 6	–0.5733	3.15e–06
200069_at	*SART3*	Squamous cell carcinoma antigen recognized by T cells 3	–0.5686	3.94e–06
212315_s_at	*NUP210*	Nucleoporin 210 kDa	–0.5666	4.35e–06
226482_s_at	*F11R*	F11 receptor	–0.5620	5.38e–06
206687_s_at	*PTPN6*	Protein tyrosine phosphatase, non–receptor type 6	–0.5518	8.60e–06
213947_s_at	*NUP210*	Nucleoporin 210 kDa	–0.5473	1.06e–05
212482_at	*FLJ13910*	Hypothetical protein FLJ13910	–0.5423	1.32e–05
227134_at	*SYTL1*	Synaptotagmin-like 1	–0.5405	1.42e–05
204142_at	*ENOSF1*	Enolase superfamily member 1	–0.5376	1.62e–05
224428_s_at	*CDCA7*	Cell division cycle–associated 7	–0.5354	1.78e–05
201969_at	*NASP*	Nuclear autoantigenic sperm protein (histone-binding)	–0.5354	1.78e–05
225887_at		Chromosome 13 open reading frame 23	–0.5305	2.19e–05
218491_s_at	*THYN1*	Thymocyte nuclear protein 1	–0.5300	2.24e–05
227560_at	*SFXN2*	Sideroflexin 2	–0.5288	2.35e–05
204798_at	*MYB*	V-*myb* myeloblastosis viral oncogene homolog (avian)	–0.5250	2.76e–05
212978_at	*LRRC8B*	Leucine-rich repeat containing 8 family, member B	–0.5218	3.15e–05
212446_s_at	*LASS6*	LAG1 longevity assurance homolog 6 (*Saccharomyces cerevisiae*)	–0.5215	3.18e–05
202107_s_at	*MCM2*	MCM2 minichromosome maintenance deficient 2, mitotin (*S. cerevisiae*)	–0.5202	3.36e–05

Table 1. (*continued*)

Affy ID	Gene Symbol	Gene Name	ρ	P
212873_at	*HMHA1*	Histocompatibility (minor) HA-1	−0.5197	3.43e−05
227378_x_at	*MGC13114*	Hypothetical protein MGC13114	−0.5192	3.49e−05
213149_at	*DLAT*	Dihydrolipoamide *S*-acetyltransferase (E_2 component of pyruvate dehydrogenase complex)	−0.5189	3.55e−05
227586_at	*LOC124491*	LOC124491	−0.5188	3.55e−05
204767_s_at	*FEN1*	Flap structure–specific endonuclease 1	−0.5170	3.83e−05
201038_s_at	*ANP32A*	Acidic (leucine-rich) nuclear phosphoprotein 32 family, member A	−0.5166	3.90e−05
208901_s_at	*TOP1*	Topoisomerase (DNA) I	−0.5163	3.94e−05
219378_at	*NARG1L*	NMDA receptor–regulated 1–like	−0.5155	4.07e−05
225179_at	*HIP2*	Huntingtin interacting protein 2	−0.5151	4.13e−05
225716_at	*BRI3BP*	BRI3 binding protein	−0.5139	4.34e−05
217980_s_at	*MRPL16*	Mitochondrial ribosomal protein L16	−0.5125	4.59e−05
225845_at	*BTBD15*	BTB (POZ) domain containing 15	−0.5116	4.75e−05
213251_at	*LOC441046*	Hypothetical LOC 441046	−0.5101	5.04e−05
220035_at	*NUP210*	Nucleoporin 210 kDa	−0.5091	5.26e−05
223268_at	*C11ORF54*	Chromosome 11 open reading frame 54	−0.5091	5.26e−05
210206_s_at	*DDX11*	Dead/H (Asp-Glu-Ala-Asp/His) box polypeptide 11 (CHL1-like helicase homolog, *S. cerevisiae*)	−0.5079	5.52e−05
202778_s_at	*ZNF198*	Zinc finger protein 198	−0.5058	5.99e−05
201202_at	*PCNA*	Proliferating cell nuclear antigen	−0.5053	6.09e−05
212943_at	*KIAA0528*	KIAA0528	−0.5048	6.21e−05
219067_s_at	*C10ORF86*	Chromosome 10 open reading frame 86	−0.5044	6.31e−05
219188_s_at	*LRP16*	LRP16 protein	−0.5020	6.94e−05
201401_s_at	*ADRBK1*	Adrenergic, beta, receptor kinase 1	−0.5014	7.10e−05
203375_s_at	*TPP2*	Tripeptidyl peptidase II	−0.4994	7.65e−05
200091_s_at	*RPS25*	Ribosomal protein S25	−0.4992	7.72e−05
224944_at	*TMPO*	Thymopoietin	−0.4982	8.01e−05
228992_at	*MED28*	Mediator of RNA polymerase II transcription, subunit 28 homolog (yeast)	−0.4970	8.39e−05
213626_at	*CBR4*	Carbonic reductase 4	−0.4966	8.54e−05
231887_s_at	*KIAA1274*	KIAA1274	−0.4926	9.91e−05
202163_s_at	*CNOT8*	CCR4-NOT transcription complex, subunit 8	−0.4906	1.07e−04
202163_s_at	*CNOT8*	CCR4-NOT transcription complex, subunit 8	−0.4906	1.07e−04

Table 2. Gene Ontology Categories Associated with Genes Whose Expression Profiles Correlated with NH ($P < .001$) Across the NCI-60 Panel of Cancer Cell Lines.

Category	GO Term	GO Terms Associated with 360 Gene Transcripts Correlated with NH, $P < .001$			GO Terms Associated with 189 Gene Transcripts Positively Correlated with NH, $P < .001$			GO Terms Associated with 171 Gene Transcripts Negatively Correlated with NH, $P < .001$		
		Count	%	P	Count	%	P	Count	%	P
GOTERM_BP_ALL	homophilic cell adhesion	22	5.76	1.14e−13	22	10.48	9.44e−19			
GOTERM_BP_ALL	cell-cell adhesion	23	6.02	4.25e−11	23	10.95	3.08e−16			
GOTERM_BP_ALL	cell adhesion	36	9.69	1.16e−08	36	17.14	2.35e−15			
GOTERM_BP_ALL	cell surface receptor–linked signal transduction	32	8.38	9.20e−04	27	12.86	3.65e−06			
GOTERM_BP_ALL	cell communication	70	18.32	4.48e−02	54	25.71	8.16e−05			
GOTERM_BP_ALL	signal transduction	65	17.02	6.08e−02	50	23.81	2.30e−04			
GOTERM_BP_ALL	response to wounding	13	3.40	7.31e−03	11	5.24	9.33e−04			
GOTERM_BP_ALL	response to external stimulus	16	4.19	7.84e−03	13	6.19	1.11e−03			
GOTERM_BP_ALL	development	41	10.80	2.80e−01	36	17.14	1.14e−03			
GOTERM_BP_ALL	locomotion	15	3.93	1.39e−03	11	5.24	1.33e−03			
GOTERM_BP_ALL	localization of cell	15	3.93	1.39e−03	11	5.24	1.33e−03			
GOTERM_BP_ALL	cell motility	15	3.93	1.39e−03	11	5.24	1.33e−03			
GOTERM_BP_ALL	cell migration	8	2.09	1.20e−02	7	3.33	2.88e−03			
GOTERM_BP_ALL	G-protein–coupled receptor protein signaling pathway	14	3.66	3.38e−03	10	4.76	4.18e−03			
GOTERM_BP_ALL	inflammatory response	9	2.36	5.25e−03	7	3.33	4.26e−03			
GOTERM_BP_ALL	cellular morphogenesis				10	4.76	1.75e−02			
GOTERM_BP_ALL	cell differentiation				12	5.71	2.68e−02			
GOTERM_BP_ALL	response to other organism	13	3.40	8.48e−02	10	4.76	3.16e−02			
GOTERM_BP_ALL	neurogenesis				5	2.38	3.17e−02			
GOTERM_BP_ALL	acute-phase response	3	0.79	8.75e−02	3	1.43	3.18e−02			
GOTERM_BP_ALL	morphogenesis	15	3.90	4.70e−01	14	6.67	4.30e−02			
GOTERM_BP_ALL	negative regulation of cell proliferation	10	2.62	2.49e−02	6	2.86	9.36e−02			
GOTERM_BP_ALL	neutrophil chemotaxis	3	0.79	2.30e−02						
GOTERM_BP_ALL	regulation of chemotaxis	3	0.79	2.30e−02						
GOTERM_BP_ALL	positive regulation of chemotaxis	3	0.79	2.30e−02						
GOTERM_BP_ALL	immune cell migration	3	0.79	3.33e−02						
GOTERM_BP_ALL	immune cell chemotaxis	3	0.79	3.33e−02						
GOTERM_CC_ALL	membrane	108	28.42	5.18e−02	85	40.48	1.83e−06			
GOTERM_CC_ALL	intrinsic to membrane	78	20.53	7.70e−02	64	30.48	1.77e−05			
GOTERM_CC_ALL	plasma membrane				29	13.81	6.32e−03			

Table 2. (*continued*)

Category	GO Term	GO Terms Associated with 360 Gene Transcripts Correlated with NH, $P < .001$			GO Terms Associated with 189 Gene Transcripts Positively Correlated with NH, $P < .001$			GO Terms Associated with 171 Gene Transcripts Negatively Correlated with NH, $P < .001$		
		Count	%	*P*	Count	%	*P*	Count	%	*P*
GOTERM_CC_ALL	extracellular region	25	6.58	6.30e–03	25	11.90	1.75e–06			
GOTERM_CC_ALL	intrinsic to plasma membrane				20	9.52	1.24e–02			
GOTERM_CC_ALL	cytoskeleton				19	9.05	5.25e–02			
GOTERM_CC_ALL	extracellular matrix	10	2.63	4.36e–02	10	4.76	1.66e–03			
GOTERM_CC_ALL	extracellular space				8	3.81	3.35e–02			
GOTERM_CC_ALL	heterotrimeric G-protein complex				3	1.43	4.23e–02			
GOTERM_CC_ALL	extrinsic to plasma membrane	4	1.05	2.63e–02	3	1.43	5.77e–02			
GOTERM_MF_ALL	calcium ion binding	35	9.21	6.86e–06	35	16.67	8.96e–13			
GOTERM_MF_ALL	metal ion binding				66	31.43	2.57e–06			
GOTERM_MF_ALL	cation binding				59	28.10	1.48e–05			
GOTERM_MF_ALL	transmembrane receptor activity	16	4.21	2.62e–02	16	7.62	7.32e–05			
GOTERM_MF_ALL	signal transducer activity				39	18.57	1.52e–04			
GOTERM_MF_ALL	protein binding	135	35.53	3.29e–03	81	38.57	1.45e–03			
GOTERM_MF_ALL	receptor activity				22	10.48	2.28e–03			
GOTERM_MF_ALL	actin binding	12	3.16	6.81e–02	10	4.76	1.06e–02			
GOTERM_MF_ALL	growth factor binding				4	1.90	2.60e–02			
GOTERM_MF_ALL	carbohydrate binding				6	2.86	3.63e–02			
GOTERM_MF_ALL	L-ascorbic acid binding				3	1.43	4.02e–02			
GOTERM_MF_ALL	enzyme regulator activity				15	7.14	4.06e–02			
GOTERM_MF_ALL	cytoskeletal protein binding				11	5.24	4.17e–02			
GOTERM_MF_ALL	extracellular matrix structural constituent				4	1.90	4.64e–02			
GOTERM_MF_ALL	oncostatin-M receptor activity	2	0.53	9.61e–02	2	0.95	5.34e–02			
KEGG_PATHWAY	HSA04060: CYTOKINE–CYTOKINE RECEPTOR INTERACTION	10	2.63	2.17e–03	10	4.76	6.75e–05			
KEGG_PATHWAY	HSA04510: FOCAL ADHESION	12	3.16	3.43e–02	11	5.24	3.78e–03			
KEGG_PATHWAY	HSA04630: JAK-STAT SIGNALING PATHWAY	8	2.11	1.68e–02	7	3.33	6.57e–03			
KEGG_PATHWAY	HSA01430: CELL COMMUNICATION	6	1.58	4.99e–02	5	2.38	3.85e–02			
KEGG_PATHWAY	HSA04810: REGULATION OF ACTIN CYTOSKELETON				8	3.81	4.37e–02			
KEGG_PATHWAY	HSA04512: ECM-RECEPTOR INTERACTION				5	2.38	5.13e–02			
GOTERM_BP_ALL	DNA-dependent DNA replication	12	3.14	1.67e–04				12	7.02	3.70e–08
GOTERM_BP_ALL	DNA replication	15	3.93	1.80e–03				15	8.77	1.08e–07
GOTERM_BP_ALL	DNA replication initiation	7	1.83	1.74e–04				7	4.09	1.25e–06
GOTERM_BP_ALL	nucleobase, nucleoside, nucleotide and nucleic acid metabolism							57	33.33	3.65e–06
GOTERM_BP_ALL	DNA metabolism							22	12.87	1.09e–05
GOTERM_BP_ALL	regulation of cellular metabolism							38	22.22	1.03e–04
GOTERM_BP_ALL	regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism							36	21.05	1.21e–04
GOTERM_BP_ALL	transcription							36	21.05	2.14e–04
GOTERM_BP_ALL	regulation of transcription							34	19.88	3.73e–04
GOTERM_BP_ALL	macromolecule metabolism							61	35.67	4.92e–03
GOTERM_BP_ALL	cellular metabolism							85	49.71	5.44e–03
GOTERM_BP_ALL	response to DNA damage stimulus							10	5.85	6.34e–03
GOTERM_BP_ALL	response to endogenous stimulus							10	5.85	7.73e–03
GOTERM_BP_ALL	cell cycle	24	N/A	N/A				18	10.53	8.62e–03
GOTERM_BP_ALL	DNA unwinding during replication	3	0.79	4.51e–02				3	1.75	8.85e–03
GOTERM_BP_ALL	DNA repair							9	5.26	9.74e–03
GOTERM_BP_ALL	DNA packaging							7	4.09	3.44e–02
GOTERM_BP_ALL	phosphoinositide-mediated signaling							4	2.34	3.78e–02
GOTERM_BP_ALL	base-excision repair							3	1.75	3.96e–02
GOTERM_BP_ALL	chromosome organization and biogenesis							8	4.68	4.19e–02
GOTERM_BP_ALL	DNA ligation during DNA repair	2	0.52	9.85e–02				2	1.17	4.23e–02
GOTERM_BP_ALL	base-excision repair, DNA ligation	2	0.52	9.85e–02				2	1.17	4.23e–02
GOTERM_CC_ALL	nucleus							71	41.52	6.28e–10
GOTERM_CC_ALL	intracellular organelle							92	53.80	2.32e–07
GOTERM_CC_ALL	organelle							92	53.80	2.37e–07
GOTERM_CC_ALL	intracellular membrane-bound organelle							85	49.71	4.34e–07
GOTERM_CC_ALL	intracellular							100	58.48	1.88e–06
GOTERM_CC_ALL	chromosome	15	3.68	4.65e–02				15	8.77	2.62e–06
GOTERM_CC_ALL	chromatin							8	4.68	4.68e–04
GOTERM_CC_ALL	intracellular non–membrane-bound organelle							25	14.62	1.11e–02
GOTERM_CC_ALL	synaptic vesicle							3	1.75	2.53e–02
GOTERM_CC_ALL	nuclear envelope							5	2.92	3.50e–02
GOTERM_CC_ALL	condensed chromosome							3	1.75	4.36e–02
GOTERM_MF_ALL	nucleic acid binding							55	32.16	4.48e–06
GOTERM_MF_ALL	DNA binding							37	21.64	7.16e–05
GOTERM_MF_ALL	DNA-dependent ATPase activity	6	1.32	2.09e–02				6	3.51	9.65e–05
GOTERM_MF_ALL	RNA binding							16	9.36	9.44e–03
GOTERM_MF_ALL	transcription regulator activity							22	12.87	9.78e–03
GOTERM_MF_ALL	ATPase activity							9	5.26	2.13e–02
GOTERM_MF_ALL	ATPase activity, coupled							8	4.68	3.19e–02
KEGG_PATHWAY	HSA04110: CELL CYCLE	10	2.63	3.13e–02				10	5.85	1.64e–05

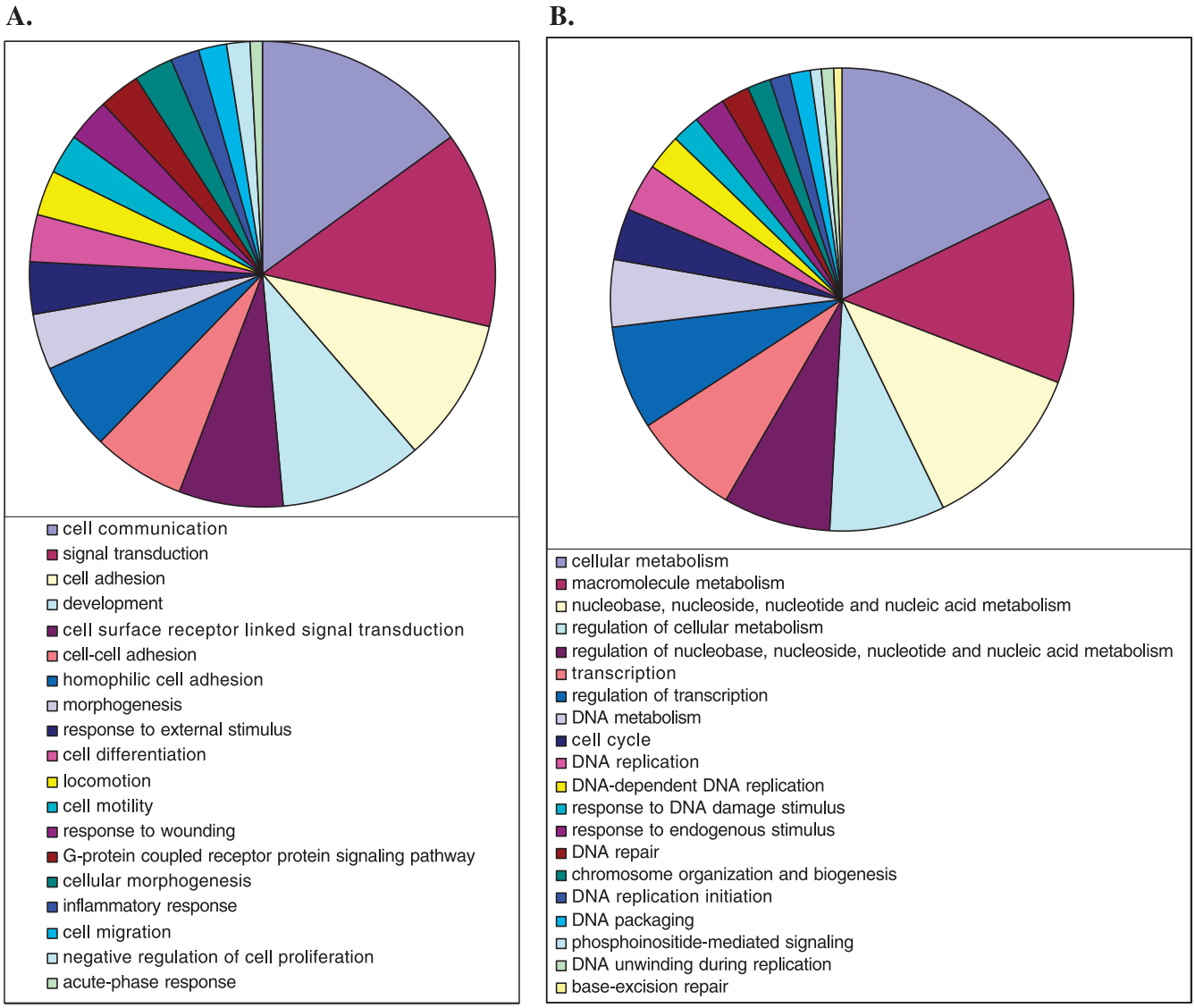

Figure 1. Biological processes associated with expression of genes positively (A) and negatively (B) correlated with NH.

cycle, chromatin-related protein, and DNA repair. Genes whose expression correlated positively with NH (67 genes) again showed significant enrichment in such categories as cell communication and signal transduction, extracellular region, development and morphogenesis, response to wounding, cell migration, and motility. In summary, then, the GO classifications of genes positively and negatively associated with NH did not seem to depend on tissue of origin of the cancer cell.

Evaluation of 1122 genes whose expression correlated with NH at $P < .01$ (Table W4) revealed further enrichment of the same GO categories identified under stricter conditions ($P < .001$; data shown in the Table 2). That observation suggests that the essential findings are not sensitive to the P value cutoff used. Also reassuring with respect to robustness of the findings, analysis of the HG-U95A data showed similar GO category enrichments (data not shown).

Focusing on the negative correlations of expression with NH, we found enrichment of the following biological processes:

1) Cellular metabolism
2) Nucleic acid metabolism
3) Regulation of transcription
4) DNA replication
5) Response to DNA damage stimulus
6) DNA repair
7) Chromosome organization and biogenesis
8) DNA replication initiation
9) DNA packaging
10) Cell cycle regulation

The CIN phenotype has been associated with very rare mutations in the checkpoint genes [22–27] and with decreased protein levels of mitotic checkpoint components [25,27,28]. Deletion in mice of one allele of *Mad2*, *Bub1*, or *Bub3* compromises the mitotic checkpoint, yielding higher rates of chromosome missegregation [29–32]. In our study, regulators of mitotic cell cycle checkpoint (*MAD2* and *BUB3*), as well as a component of APC/C (*APC4*), are found among the genes whose expression is negatively correlated with NH. Correlations do not imply causative relationships; however, it would not be unreasonable to suggest that the decreased level of mitotic checkpoint components could be the basis of mitotic checkpoint relaxation

leading to increased gains and losses of chromosomes. This supports already existing assumptions that a compromised mitotic checkpoint leads to accelerated rates of chromosomal instability in cancer cells [33].

The expression of genes involved in DNA damage checkpoints (*CHK1*, *CHK2*, *H2AX*, *RAD21*, *XRCC5*, *DDB1*) and DNA rereplication prevention (*BCCIP*, *BRCA2*, *CDT1*, *MCM2-7*, *cyclin B2*) negatively correlates with NH as well. The expression levels of genes involved in DNA packaging, chromosome condensation, and kinetochore formation (*H3 histone*, *H1FX*, *H2AX*, *H2AZ*, *TOP1*, *RCC1*, *RCC2*, *SMARCA5*, *RCBTB1*, *CENPC1*, *ZWINT*) are also relatively down-regulated in cancer cells with higher level of chromosomal instability compared to cancer cells with a lower level of instability.

Compromised cell cycle checkpoints give cancer cells an advantage in that they may be able to proliferate in a stressful environment with uncompleted DNA repair, perhaps with some level of unfinished chromatin condensation and/or individualization of chromatids, and perhaps with defects of mitotic chromosome organization. As a result, on-going gains and losses of chromosomes or their fragments can occur while cells are proliferating. We also found that chromosomal instability is associated with less effective cellular metabolism, DNA replication and transcription, DNA repair and packaging, weakness in proper chromatin condensation, and mitotic chromosome structural organization possibly owing to extensive imbalances in cellular protein composition of cells that undergo continuous gains and losses of parts of genome.

Gene Ontology's biological processes positively correlated with higher NH include the following:

1) Cell communication and signal transduction
2) Cell adhesion, including cell-cell adhesion
3) Cell surface receptor–linked signal transduction
4) Development and differentiation
5) Cell motility and migration
6) Response to wounding
7) Inflammatory response
8) G-protein–coupled receptor protein signaling
9) Cellular morphogenesis

A collective molecular portrait of numerical chromosomal heterogeneity in cancer cells includes relative up-regulation of genes that are associated with increased motility and migration, epithelial-mesenchymal transition (EMT), and are critical for tumor invasion and metastasis: *RhoC*, *fibronectin*, *LOX*, *TWIST*, *SNAI2*, *EGFR*, *laminins*, *integrins*, *collagens*, *CDC42* effector protein (Rho GTPase binding), Rho family GTPase 3, *RAB*, *CXCL2*, *TGF-b2*, *VEGFC*, *IL-6*, *IL-8*, *CTGF*, *vimentin*, *N-cadherin*, *CD44*, *BCAR3*, *protocadherins*, *MMP2* and *MMP14*, *NOTCH2*, *SERPINE1*, *2*, and *8*, *IGFBP3* and *7*, *TNFAIP3*, *TNFRSF12A* and *19*, *PLAUR*, and *SPARC*. Expression of several genes that promote cell proliferation and G_1 entry into cell cycle (*CCD1*, *EGFR*, *VEGFC*) correlate positively with the higher NH as well.

Advances in the molecular profiling of cancer using genome-wide approaches have revealed genes whose expression levels in primary tumors correlate strongly with the likelihood of metastatic recurrence [34–37]. In particular, genes that are involved in physiological programs of cellular response to tissue damage (inflammation, wound healing, tissue remodeling, and regeneration) and genes that participate in differentiation and morphogenesis during epithelial tissue development are involved in tumor invasiveness and metastasis [38–41]. Important to those processes is an EMT, which enables cells to undergo major changes in morphology, lose cellular contacts, and acquire motility and ability to migrate. Also an important part of those programs is stimulation of proliferation and differentiation of epithelial cells themselves, as well as angiogenesis and lymphangiogenesis. Our findings suggest a link, at least indirectly, between chromosomal instability and cancer proliferation, invasion, and metastasis.

Chromosomal instability can create both advantages and disadvantages for cancer cells. For instance, severe genomic damage due to losses or gains of whole chromosomes or their essential fragments can prevent cancer cells from further proliferation. At the same time, the program of response to tissue damage represents an advantageous feature for overall cancer progression. Release of chemokines and cytokines promotes a response from surrounding tumor and stromal cells leading to the EMT in cancer cells, changes in cell adhesion and motility, promoting epithelial cell proliferation, as well as angiogenesis and lymphangiogenesis. We suggest that gross instability of cancer cell genomes leads to evolution of cancer cell populations, which "internalize" various inducers of a tissue damage response, gradually making cancer cells more and more independent of environmental stimuli. These changes in the gene expression pattern persist in cancer-derived cell lines.

Acknowledgments

The authors thank W. Michael Kuehl for helpful discussion and suggestions.

References

[1] Mitelman F, Johansson B, Mandahl N, and Mertens F (1997). Clinical significance of cytogenetic findings in solid tumors. *Cancer Genet Cytogenet* **95**, 1–8.
[2] Ried T, Heselmeyer-Haddad K, Blegen H, Schrock E, and Auer G (1999). Genomic changes defining the genesis, progression, and malignancy potential in solid human tumors: a phenotype/genotype correlation. *Genes Chromosomes Cancer* **25**, 195–204.
[3] Pandis N, Idvall I, Bardi G, Jin Y, Gorunova L, Mertens F, Olsson H, Ingvar C, Beroukas K, Mitelman F, et al. (1996). Correlation between karyotypic pattern and clincopathologic features in 125 breast cancer cases. *Int J Cancer* **66**, 191–196.
[4] Bardi G, Fenger C, Johansson B, Mitelman F, and Heim S (2004). Tumor karyotype predicts clinical outcome in colorectal cancer patients. *J Clin Oncol* **22**, 2623–2634.
[5] Duesberg P, Li R, and Rasnick D (2004). Aneuploidy approaching a perfect score in predicting and preventing cancer: highlights from a conference held in Oakland, CA in January, 2004. *Cell Cycle* **3**, 823–828.
[6] Lengauer C, Kinzler KW, and Vogelstein B (1997). Genetic instability in colorectal cancers. *Nature* **386**, 623–627.
[7] Lengauer C, Kinzler KW, and Vogelstein B (1998). Genetic instabilities in human cancers. *Nature* **396**, 643–649.
[8] Shoemaker RH, Monks A, Alley MC, Scudiero DA, Fine DL, McLemore TL, Abbott BJ, Paull KD, Mayo JG, and Boyd MR (1988). Development of human tumor cell line panels for use in disease-oriented drug screening. *Prog Clin Biol Res* **276**, 265–286.
[9] Weinstein JN and Pommier Y (2006). Connecting genes, drugs and diseases. *Nat Biotechnol* **24**, 1365–1366.
[10] Weinstein JN (2006). Spotlight on molecular profiling: "integromic" analysis of the NCI-60 cancer cell lines. *Mol Cancer Ther* **5**, 2601–2605.
[11] Roschke AV, Tonon G, Gehlhaus KS, McTyre N, Bussey KJ, Lababidi S, Scudiero DA, Weinstein JN, and Kirsch IR (2003). Karyotypic complexity of the NCI-60 drug-screening panel. *Cancer Res* **63**, 8634–8647.
[12] Roschke AV, Lababidi S, Tonon G, Gehlhaus KS, Bussey K, Weinstein JN, and Kirsch IR (2005). Karyotypic "state" as a potential determinant for anticancer drug discovery. *Proc Natl Acad Sci USA* **102**, 2964–2969.

[13] Wallqvist A, Huang R, Covell DG, Roschke AV, Gelhaus KS, and Kirsch IR (2005). Drugs aimed at targeting characteristic karyotypic phenotypes of cancer cells. *Mol Cancer Ther* **4**, 1559–1568.

[14] Roschke AV and Kirsch IR (2005). Targeting cancer cells by exploiting karyotypic complexity and chromosomal instability. *Cell Cycle* **4**, 679–682.

[15] Shankavaram UT, Reinhold WC, Nishizuka S, Major S, Morita D, Chary KK, Reimers MA, Scherf U, Kahn A, Dolginow D, et al. (2007). Transcript and protein expression profiles of the NCI-60 cancer cell panel: an integromic microarray study. *Mol Cancer Ther* **6**, 820–832.

[16] Ihaka R and Gentleman R (1996). R: a language for data analysis and graphics. *J Comput Graph Stat* **5**, 299–314.

[17] Gentleman R, Carey V, Huber W, Irizarry R, and Dudoit S (2005). *Bioinformatics and Computational Biology Solutions Using R and Bioconductor.* New York, NY: Springer.

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

[19] Simon R and Lam A (2007). *BRB-ArrayTools. Version 3.6.* http://linus.nci.nih.gov/brb.

[20] Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, and Lempicki RA (2003). DAVID: database for annotation, visualization, and integrated discovery. *Genome Biol* **4**, P4.

[21] Hosack DA, Dennis G Jr, Sherman BT, Lane HC, and Lempicki RA (2003). Identifying biological themes within lists of genes with EASE. *Genome Biol* **4**, R70.

[22] Cahill DP, Lengauer C, Yu J, Riggins GJ, Willson JK, Markowitz SD, Kinzler KW, and Vogelstein B (1998). Mutations of mitotic checkpoint genes in human cancers. *Nature* **392**, 300–303.

[23] Imai Y, Shiratori Y, Kato N, Inoue T, and Omata M (1999). Mutational inactivation of mitotic checkpoint genes, *hsMAD2* and *hBUB1*, is rare in sporadic digestive tract cancers. *Jpn J Cancer Res* **90**, 837–840.

[24] Sato M, Sekido Y, Horio Y, Takahashi M, Saito H, Minna JD, Shimokata K, and Hasegawa Y (2000). Infrequent mutation of the *hBUB1* and *hBUBR1* genes in human lung cancer. *Jpn J Cancer Res* **91**, 504–509.

[25] Shichiri M, Yoshinaga K, Hisatomi H, Sugihara K, and Hirata Y (2002). Genetic and epigenetic inactivation of mitotic checkpoint genes *hBUB1* and *hBUBR1* and their relationship to survival. *Cancer Res* **62**, 13–17.

[26] Ru HY, Chen RL, Lu WC, and Chen JH (2002). *hBUB1* defects in leukemia and lymphoma cells. *Oncogene* **21**, 4673–4679.

[27] Wang X, Jin DY, Ng RW, Feng H, Wong YC, Cheung AL, and Tsao SW (2002). Significance of MAD2 expression to mitotic checkpoint control in ovarian cancer cells. *Cancer Res* **62**, 1662–1668.

[28] Wang X, Jin DY, Wong YC, Cheung AL, Chun AC, Lo AK, Liu Y, and Tsao SW (2000). Correlation of defective mitotic checkpoint with aberrantly reduced expression of MAD2 protein in nasopharyngeal carcinoma cells. *Carcinogenesis* **21**, 2293–2297.

[29] Michel LS, Liberal V, Chatterjee A, Kirchwegger R, Pasche B, Gerald W, Dobles M, Sorger PK, Murty VV, and Benezra R (2001). MAD2 haplo-insufficiency causes premature anaphase and chromosome instability in mammalian cells. *Nature* **409**, 355–359.

[30] Michel L, Benezra R, and Diaz-Rodriguez E (2004). MAD2 dependent mitotic checkpoint defects in tumorigenesis and tumor cell death: a double edged sword. *Cell Cycle* **3**, 990–992.

[31] Dai W, Wang Q, Liu T, Swamy M, Fang Y, Xie S, Mahmood R, Yang YM, Xu M, and Rao CV (2004). Slippage of mitotic arrest and enhanced tumor development in mice with BubR1 haploinsufficiency. *Cancer Res* **64**, 440–445.

[32] Kalitsis P, Fowler KJ, Griffiths B, Earle E, Chow CW, Jamsen K, and Choo KH (2005). Increased chromosome instability but not cancer predisposition in haploinsufficient Bub3 mice. *Genes Chromosomes Cancer* **44**, 29–36.

[33] Schmidt M and Medema RH (2006). Exploiting the compromised spindle assembly checkpoint function of tumor cells: dawn on the horizon? *Cell Cycle* **5**, 159–163.

[34] Chen JJ, Peck K, Hong TM, Yang SC, Sher YP, Shih JY, Wu R, Cheng JL, Roffler SR, Wu CW, et al. (2001). Global analysis of gene expression in invasion by a lung cancer model. *Cancer Res* **61**, 5223–5230.

[35] van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, et al. (2002). Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **415**, 530–536.

[36] Kluger HM, Kluger Y, Gilmore-Hebert M, DiVito K, Chang JT, Rodov S, Mironenko O, Kacinski BM, Perkins AS, and Sapi E (2004). cDNA microarray analysis of invasive and tumorigenic phenotypes in a breast cancer model. *Lab Invest* **84**, 320–331.

[37] Rhodes DR, Yu J, Shanker K, Deshpande N, Varambally R, Ghosh D, Barrette T, Pandey A, and Chinnaiyan AM (2004). Large-scale meta-analysis of cancer microarray data identifies common transcriptional profiles of neoplastic transformation and progression. *Proc Natl Acad Sci USA* **101**, 9309–9314.

[38] Gupta PB, Kuperwasser C, Brunet JP, Ramaswamy S, Kuo WL, Gray JW, Naber SP, and Weinberg RA (2005). The melanocyte differentiation program predisposes to metastasis after neoplastic transformation. *Nat Genet* **37**, 1047–1054.

[39] Gupta GP and Massague J (2006). Cancer metastasis: building a framework. *Cell* **127**, 679–695.

[40] Huber MA, Kraut N, and Beug H (2005). Molecular requirements for epithelial-mesenchymal transition during tumor progression. *Curr Opin Cell Biol* **17**, 548–558.

[41] Kopfstein L and Christofori G (2006). Metastasis: cell-autonomous mechanisms *versus* contributions by the tumor microenvironment. *Cell Mol Life Sci* **63**, 449–468.

Supplement: Cluster Analysis

Materials and Methods

For cluster analysis, U133A and U133B array data were processed separately as described (Materials and Methods, main text) with modifications. First, robust multichip average–transformed data were filtered to have expression measures higher than 64 at least in two cell lines, which is the minimum number of cell lines in a group. Two gene sets were selected: one for which the interquartile range in expression values was higher than 0.5 on $\log_2$ scale (11,848 probe sets) and the second for which the interquartile range in expression values was higher than 1 on $\log_2$ scale (this left 1667 and 682 probe sets for U133A and U133B array data sets, respectively). Gene expression matrixes were then additionally normalized by subtracting $\log_2$(median) expression value for each probe set. Gene expression matrixes for differentially expressed genes selected according to the *F* test (Materials and Methods, main text) and for 454 probe sets that show correlation with numerical heterogeneity (NH) level were also normalized by subtracting $\log_2$(median) expression value for each probe set. This normalization makes results more comparable when different distance measures are used. For Figures W2 to W4, clustering was done by Hierarchical Ordered Partitioning And Collapsing Hybrid (HOPACH) method as implemented in R package *hopach* [1] using cosine angle (uncentered correlation) and Euclidean distances for genes and arrays, respectively. For Figure W1, hierarchical clustering was performed by using function *hclust* (R package *cluster*), Euclidean distance, and average linkage. In the HOPACH algorithm, minimization of the median split silhouette value is used to choose the cluster number and cluster ordering. Compared with other clustering algorithms, the HOPACH method produces consistent cluster ordering regardless of the original ordering of the arrays in gene expression matrix [2]. Bootstrap resampling was done by selecting 10,000 sets of 57 arrays one array at a time with replacement. The proportion of resampled data sets in which each cell line falls into each of the clusters serves as an estimate of the membership of that cell line in each cluster. Bigger proportion of cell lines with consistent membership in a cluster indicates that cluster is well differentiated from other clusters and stable. Because we wanted to analyze mainly cell line characteristics (particularly karyotypic instability), the results are presented as ordered distance matrixes for cell lines together with bar plots of the bootstrap reappearance proportions for each cell line and each cluster. Additional details and results of cluster analysis are available on request.

Results

Hierarchical clustering of the 57 cell lines from NCI-60 collection shows that cell lines are mostly grouped according to the presumptive tissue of origin when using 11,848 probe sets with moderate level in expression variability (Figure W1).

We observed such correspondence after using for hierarchical clustering of cell lines 2349 genes that were selected also nonspecifically as displaying a higher (more than twofold difference between 25th and 75th percentiles) level of expression variability across cell lines (Figure W2). However, it can be seen that there are some similarities between cell lines unrelated to common tissue of origin: colon cell lines are similar to some prostate, renal, and ovarian cell lines and some colon cell lines are also similar to leukemia cell lines (Figure W2). Thus, though the tissue of origin seems to be a critical factor that influences gene expression pattern in the NCI-60 cell lines, there are additional variables that can be used to stratify cell lines if the gene set is selected appropriately (see also Ross et al. [3]).

For example, the correspondence of cell line gene expression patterns to tissue of origin can be enhanced if clustering is performed using a subset of genes that were selected according to an *F* test as differentially expressed between cell lines of different tissues of origin (Figure W3). Cell lines derived from leukemias, melanomas, colon tumors, renal tumors, CNS tumors, and, to a lesser extent, lung tumors form clusters that can be seen as separate blocks on the diagonal of the ordered dissimilarity matrix. This confirms the results obtained using spotted cDNA arrays [3] and illustrates that selection of genes changes the cluster analysis results: the clusters defined by differentially expressed genes are not only better separated but also are more stable than clusters obtained using nonspecifically selected genes, as follows from a bootstrap resampling analysis (compare Figures W2 and W3).

Impact of the tissue of origin of cell lines on gene expression pattern can still be traced after clustering using 454 probe sets whose expression correlates with NH index. However, the ordered dissimilarity matrix is quite different from the one based on genes differentially expressed among cell lines of different tissues of origin (Figure W4). Particularly, all cell lines derived from leukemia and colon tumors (together with two cell lines from breast tumors and one cell line from a lung tumor) form a tight cluster that is remarkably stable upon bootstrap resampling. Cell lines in this cluster have an average NH index of 0.27, whereas 21 cell lines in the biggest cluster that show maximal dissimilarity (denoted by the presence of a white color in the matrix) with the cluster of leukemia and colon cell lines have an average NH index 0.65.

Clustering analysis demonstrates that gene expression is influenced mainly by the tissue of origin, and at the same time, clustering of genes selected based on their correlation with NH is different, demonstrating an influence by NH as well.

References

[1] Pollard KS and van der Laan MJ (2005). Cluster analysis of genomic data. In: *Bioinformatics and Computational Biology Solutions Using R and Bioconductor.* Gentlemen R, et al. (Eds). Springer, New York, NY. pp. 209–228.

[2] van der Laan M and Pollard K (2003). A new algorithm for hybrid hierarchical clustering with visualization and bootstrap. *J Stat Plan Inference* **117**, 275–303.

[3] Ross DT, Scherf U, Eisen MB, Perou CM, Rees C, Spellman P, Iyer V, Jeffrey SS, Van de Rijn M, Waltham M, et al (2000). Systematic variation in gene expression patterns in human cancer cell lines. *Nat Genet* **24** (3), 227–235.

Table W1. Number of Gene Expression Array Elements and Corresponding Genes Correlated with Structural Complexity, Structural Heterogeneity, Numerical Complexity, and Numerical Heterogeneity with $P < .01$ and $P < .001$.

	$P < .01$		$P < .001$	
	Array Elements	Genes	Array Elements	Genes
SC-positive	190	165	20	18
SC-negative	187	166	34	32
SC total	377	331	54	50
FP 0.95	279		38	
FP 0.99	430		75	
SH-positive	733	560	215	168
SH-negative	398	352	103	96
SH total	1131	912	318	264
FP 0.95	271		36	
FP 0.99	446		82	
NC-positive	568	419	148	106
NC-negative	479	407	130	118
NC total	1047	826	278	224
FP 0.95	306		44	
FP 0.99	484		98	
NH-positive	834	571	261	189
NH-negative	640	551	193	171
NH total	1474	1122	454	360
FP 0.95	290		43	
FP 0.99	450		73	

FP indicates false-positive results.

Table W2. List of 454 Array Features and Names of Corresponding Genes with Expression Profiles Correlated with NH ($P < .001$) Across the NCI-60 Panel of Cancer Cell Lines.

Affy ID	Gene Symbol	Description	ρ	P
226148_at	*BTBD15*	BTB (POZ) domain containing 15	−0.6236	2.20e−07
212316_at	*NUP210*	nucleoporin 210 kDa	−0.6040	6.56e−07
209946_at	*VEGFC*	vascular endothelial growth factor C	0.5817	2.08e−06
218717_s_at	*LEPREL1*	leprecan-like 1	0.5763	2.72e−06
218088_s_at	*RRAGC*	Ras-related GTP binding C	0.5750	2.89e−06
224784_at	*MLLT6*	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, *Drosophila*); translocated to, 6	−0.5733	3.15e−06
200069_at	*SART3*	squamous cell carcinoma antigen recognized by T cells 3	−0.5686	3.94e−06
212315_s_at	*NUP210*	nucleoporin 210 kDa	−0.5666	4.35e−06
226482_s_at	*F11R*	F11 receptor	−0.5620	5.38e−06
204140_at	*TPST1*	tyrosylprotein sulfotransferase 1	0.5549	7.48e−06
206687_s_at	*PTPN6*	protein tyrosine phosphatase, non–receptor type 6	−0.5518	8.60e−06
226656_at	*CRTAP*	cartilage associated protein	0.5501	9.31e−06
213947_s_at	*NUP210*	nucleoporin 210 kDa	−0.5473	1.06e−05
235086_at	*THBS1*	Thrombospondin 1	0.5460	1.12e−05
212464_s_at	*FN1*	fibronectin 1	0.5428	1.29e−05
220092_s_at	*ANTXR1*	anthrax toxin receptor 1	0.5425	1.30e−05
212482_at	*RMND5A*	hypothetical protein FLJ13910	−0.5423	1.32e−05
227134_at	*SYTL1*	synaptotagmin-like 1	−0.5405	1.42e−05
201426_s_at	*VIM*	vimentin	0.5377	1.61e−05
204142_at	*ENOSF1*	enolase superfamily member 1	−0.5376	1.62e−05
224428_s_at	*CDCA7*	cell division cycle–associated 7	−0.5354	1.78e−05
201969_at	*NASP*	nuclear autoantigenic sperm protein (histone-binding)	−0.5354	1.78e−05
202859_x_at	*IL8*	interleukin 8	0.5349	1.82e−05
201172_x_at	*ATP6V0E1*	ATPase, H+ transporting, lysosomal 9 kDa, V0 subunit e	0.5341	1.88e−05
211719_x_at	*FN1*	fibronectin 1	0.5331	1.96e−05
210495_x_at	*FN1*	fibronectin 1	0.5330	1.97e−05
201828_x_at		CAAX box 1	0.5306	2.18e−05
225887_at	*C13orf23*	chromosome 13 open reading frame 23	−0.5305	2.19e−05
201380_at	*CRTAP*	cartilage-associated protein	0.5301	2.22e−05
218491_s_at	*THYN1*	thymocyte protein Thy28	−0.5300	2.24e−05
227560_at	*SFXN2*	sideroflexin 2	−0.5288	2.35e−05
237444_at		Transcribed locus	0.5288	2.36e−05
205743_at	*STAC*	SH3 and cysteine–rich domain	0.5265	2.60e−05
201105_at	*LGALS1*	lectin, galactoside-binding, soluble, 1 (galectin 1)	0.5257	2.67e−05
200096_s_at	*ATP6V0E*	ATPase, H+ transporting, lysosomal 9 kDa, V0 subunit e	0.5257	2.68e−05
204798_at	*MYB*	v-*myb* myeloblastosis viral oncogene homolog (avian)	−0.5250	2.76e−05
204214_s_at	*RAB32*	RAB32, member RAS oncogene family	0.5248	2.78e−05
200885_at	*RHOC*	ras homolog gene family, member C	0.5242	2.85e−05
231907_at	*ABL2*	V-*abl* Abelson murine leukemia viral oncogene homolog 2 (ARG, Abelson-related gene)	0.5225	3.05e−05
212978_at	*LRRC8B*	Leucine-rich repeat containing 8 family, member B	−0.5218	3.15e−05
212446_s_at	*LASS6*	LAG1 longevity assurance homolog 6 (*S. cerevisiae*)	−0.5215	3.18e−05

Table W2. (*continued*)

Affy ID	Gene Symbol	Description	ρ	P
202107_s_at	*MCM2*	MCM2 minichromosome maintenance deficient 2, mitotin (*S. cerevisiae*)	−0.5202	3.36e−05
212873_at	*HMHA1*	minor histocompatibility antigen HA-1	−0.5197	3.43e−05
226955_at	*FLJ36748*	hypothetical protein FLJ36748	0.5195	3.46e−05
227378_x_at	*MGC13114*	hypothetical protein MGC13114	−0.5192	3.49e−05
213149_at	*DLAT*	dihydrolipoamide *S*-acetyltransferase (E_2 component of pyruvate dehydrogenase complex)	−0.5189	3.55e−05
227586_at	*LOC124491*	LOC124491	−0.5188	3.55e−05
216442_x_at	*FN1*	fibronectin 1	0.5172	3.79e−05
204767_s_at	*FEN1*	flap structure-specific endonuclease 1	−0.5170	3.83e−05
201038_s_at	*ANP32A*	acidic (leucine-rich) nuclear phosphoprotein 32 family, member A	−0.5166	3.90e−05
208901_s_at	*TOP1*	topoisomerase (DNA) I	−0.5163	3.94e−05
219378_at	*NARG1L*	NMDA receptor regulated 1–like	−0.5155	4.07e−05
225179_at	*HIP2*	Huntingtin interacting protein 2	−0.5151	4.13e−05
212509_s_at	*MXRA7*	matrix-remodeling associated 7	0.5144	4.26e−05
225716_at		full-length cDNA clone CS0DK008YI09 of HeLa cells Cot 25-normalized of *Homo sapiens* (human)	−0.5139	4.34e−05
222834_s_at	*GNG12*	guanine nucleotide binding protein (G protein), gamma 12	0.5130	4.50e−05
224560_at	*TIMP2*	TIMP metallopeptidase inhibitor 2	0.5127	4.56e−05
217980_s_at	*MRPL16*	mitochondrial ribosomal protein L16	−0.5125	4.59e−05
225845_at	*BTBD15*	BTB (POZ) domain containing 15	−0.5116	4.75e−05
211612_s_at	*IL13RA1*	interleukin 13 receptor, alpha 1	0.5114	4.80e−05
226939_at	*CPEB2*	cytoplasmic polyadenylation element binding protein 2	0.5110	4.88e−05
202378_s_at	*LEPROT*	leptin receptor overlapping transcript	0.5107	4.93e−05
209153_s_at	*TCF3*	transcription factor 3 (E2A immunoglobulin enhancer binding factors E12/E47)	−0.5104	4.99e−05
235072_s_at		Transcribed locus	0.5104	4.99e−05
213251_at		Hypothetical LOC 441046	−0.5101	5.04e−05
220035_at	*NUP210*	nucleoporin 210 kDa	−0.5091	5.26e−05
223268_at	*PTD012*	PTD012 protein	−0.5091	5.26e−05
209226_s_at	*TNPO1*	transportin 1	0.5086	5.36e−05
202733_at	*P4HA2*	procollagen-proline, 2-oxoglutarate 4-dioxygenase (proline 4-hydroxylase), alpha polypeptide II	0.5083	5.43e−05
210206_s_at	*DDX11*	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 11 (CHL1-like helicase homolog, *S. cerevisiae*)	−0.5079	5.52e−05
220407_s_at	*TGFB2*	transforming growth factor, beta 2	0.5074	5.62e−05
212294_at	*GNG12*	guanine nucleotide binding protein (G protein), gamma 12	0.5073	5.64e−05
202377_at	*LEPR*; *LEPROT*	leptin receptor; leptin receptor overlapping transcript	0.5064	5.85e−05
202778_s_at	*ZMYM2*	zinc finger protein 198	−0.5058	5.99e−05
209278_s_at	*TFPI2*	tissue factor pathway inhibitor 2	0.5056	6.02e−05
201202_at	*PCNA*	proliferating cell nuclear antigen	−0.5053	6.09e−05
212943_at	*KIAA0528*	KIAA0528 gene product	−0.5048	6.21e−05
201887_at	*IL13RA1*	interleukin 13 receptor, alpha 1	0.5048	6.23e−05
209225_x_at	*TNPO1*	transportin 1	0.5047	6.23e−05
219067_s_at	*NSMCE4A*	chromosome 10 open reading frame 86	−0.5044	6.31e−05
208924_at	*RNF11*	ring finger protein 11	0.5044	6.31e−05
212658_at	*LHFPL2*	lipoma HMGIC fusion partner–like 2	0.5044	6.31e−05
213696_s_at	*MED8*	mediator of RNA polymerase II transcription, subunit 8 homolog (yeast)	0.5043	6.33e−05
229465_s_at	*PTPRS*	protein tyrosine phosphatase, receptor type, S	0.5043	6.34e−05
203262_s_at	*FAM50A*	family with sequence similarity 50, member A	0.5037	6.48e−05
209013_x_at	*TRIO*	triple functional domain (PTPRF interacting)	0.5034	6.56e−05
207657_x_at	*TNPO1*	transportin 1	0.5026	6.78e−05
219188_s_at		LRP16 protein	−0.5020	6.94e−05
201401_s_at	*ADRBK1*	adrenergic, beta, receptor kinase 1	−0.5014	7.10e−05
231579_s_at	*TIMP2*	TIMP metallopeptidase inhibitor 2	0.5013	7.12e−05
200998_s_at	*CKAP4*	cytoskeleton-associated protein 4	0.5009	7.22e−05
229307_at	*ANKRD28*	ankyrin repeat domain 28	0.5001	7.45e−05
226893_at	*ABL2*	V-*abl* Abelson murine leukemia viral oncogene homolog 2 (ARG, Abelson-related gene)	0.5000	7.48e−05
224796_at	*DDEF1*	development and differentiation enhancing factor 1	0.5000	7.49e−05
203375_s_at	*TPP2*	tripeptidyl peptidase II	−0.4994	7.65e−05
238447_at	*RBMS3*	RNA binding motif, single-stranded interacting protein	0.4994	7.66e−05
200091_s_at	*RPS25*	ribosomal protein S25	−0.4992	7.72e−05
224944_at	*TMPO*	Thymopoietin	−0.4982	8.01e−05
228234_at	*TICAM2*	toll-like receptor adaptor molecule 2	0.4974	8.27e−05
209019_s_at	*PINK1*	PTEN-induced putative kinase 1	0.4974	8.28e−05
200757_s_at	*CALU*	calumenin	0.4973	8.29e−05
228992_at	*MED28*	Mediator of RNA polymerase II transcription, subunit 28 homolog (yeast)	−0.4970	8.39e−05
213626_at	*CBR4*	carbonic reductase 4	−0.4966	8.54e−05
209263_x_at	*TSPAN4*	tetraspanin 4	0.4964	8.61e−05
225765_at	*TNPO1*	Transportin 1	0.4955	8.89e−05
203695_s_at	*DFNA5*	deafness, autosomal dominant 5	0.4952	8.99e−05
212470_at	*SPAG9*	sperm-associated antigen 9	0.4948	9.12e−05
213455_at		hypothetical LOC283677	0.4947	9.16e−05
204359_at	*FLRT2*	fibronectin leucine rich transmembrane protein 2	0.4945	9.25e−05
214696_at		hypothetical protein MGC14376	0.4937	9.53e−05
231887_s_at	*KIAA1274*	KIAA1274	−0.4926	9.91e−05
203773_x_at	*BLVRA*	biliverdin reductase A	0.4921	1.01055e−
200756_x_at	*CALU*	calumenin	0.4916	1.03041e−
209264_s_at	*TSPAN4*	tetraspanin 4	0.4909	1.05831e−

Table W2. (*continued*)

Affy ID	Gene Symbol	Description	ρ	P
224790_at	*DDEF1*	development and differentiation enhancing factor 1	0.4908	1.06088e–
202163_s_at	*CNOT8*	CCR4-NOT transcription complex, subunit 8	–0.4906	1.07122e–
221771_s_at		M-phase phosphoprotein, mpp8	–0.4903	1.08198e–
218458_at	*GMCL1*	germ cell–less homolog 1 (*Drosophila*)	–0.4900	1.09481e–
212973_at	*RPIA*	ribose 5-phosphate isomerase A (ribose 5-phosphate epimerase)	–0.4900	1.09481e–
221737_at	*GNA12*	guanine nucleotide binding protein (G protein) alpha 12	0.4896	1.10947e–
201976_s_at	*MYO10*	myosin X	0.4896	1.10947e–
221753_at	*SSH1*	slingshot homolog 1 (*Drosophila*)	0.4895	1.11215e–
212430_at	*RBM38*	RNA-binding region (RNP1, RRM) containing 1	–0.4884	1.15975e–
200943_at	*HMGN1*	high-mobility group nucleosome binding domain 1	–0.4884	1.16009e–
222514_at	*RRAGC*	Ras-related GTP binding C	0.4881	1.17272e–
233888_s_at	*SRGAP1*	SLIT-ROBO Rho GTPase activating protein 1	0.4881	1.17555e–
223199_at	*MKNK2*	MAP kinase interacting serine/threonine kinase 2	–0.4877	1.19155e–
200990_at	*TRIM28*	tripartite motif-containing 28	–0.4871	1.21574e–
218980_at	*FHOD3*	formin homology 2 domain containing 3	0.4871	1.21719e–
209832_s_at	*CDT1*	DNA replication factor	–0.4865	1.24670e–
227228_s_at	*KIAA1509*	KIAA1509	–0.4863	1.25418e–
224504_s_at	*MGC13125*	hypothetical protein MGC13125	–0.4862	1.25718e–
217785_s_at	*YKT6*	SNARE protein Ykt6	0.4861	1.26170e–
212607_at	*AKT3*	v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma)	0.4861	1.26321e–
200859_x_at	*FLNA*	filamin A, alpha (actin binding protein 280)	0.4860	1.26774e–
210053_at	*TAF5*	TAF5 RNA polymerase II, TATA box binding protein (TBP)–associated factor, 100 kDa	–0.4853	1.29954e–
224791_at	*DDEF1*	development and differentiation enhancing factor 1	0.4848	1.32651e–
202555_s_at	*MYLK*	myosin, light polypeptide kinase	0.4845	1.34239e–
226651_at	*HOMER1*	homer homolog 1 (*Drosophila*)	–0.4844	1.34718e–
201885_s_at	*CYB5R3*	cytochrome b5 reductase 3	0.4843	1.3520e–0
208549_x_at	*PTMA*	prothymosin, alpha (gene sequence 28); similar to prothymosin alpha; hypothetical gene supported by BC013859; hypothetical gene supported by BC013859; BC070480	–0.4839	1.36815e–
223625_at	*DRCTNNB1A*	down-regulated by Ctnnb1, a	0.4838	1.37629e–
205083_at	*AOX1*	aldehyde oxidase 1	0.4832	1.40764e–
204141_at	*TUBB2A*	tubulin, beta 2	0.4831	1.41265e–
202087_s_at	*CTSL1*	cathepsin L	0.4825	1.44305e–
208838_at	*CAND1*	cullin-associated and neddylation-dissociated 1	–0.4822	1.45847e–
224677_x_at	*C11orf31*	chromosome 11 open reading frame 31	–0.4820	1.46756e–
206157_at	*PTX3*	pentraxin-related gene, rapidly induced by IL-1 beta	0.4807	1.54150e–
209420_s_at	*SMPD1*	sphingomyelin phosphodiesterase 1, acid lysosomal (acid sphingomyelinase)	0.4806	1.54512e–
211066_x_at	*PCDHG*	protocadherin gamma subfamily	0.4806	1.54694e–
213746_s_at	*FLNA*	filamin A, alpha (actin binding protein 280)	0.4802	1.56889e–
215501_s_at	*DUSP10*	dual-specificity phosphatase 10	0.4797	1.59722e–
205717_x_at	*PCDHG*	protocadherin gamma subfamily	0.4796	1.59868e–
209012_at	*TRIO*	triple functional domain (PTPRF interacting)	0.4786	1.66008e–
209083_at	*CORO1A*	coronin, actin binding protein, 1A	–0.4782	1.68691e–
204143_s_at	*ENOSF1*	enolase superfamily member 1	–0.4778	1.70668e–
225878_at	*KIF1B*	kinesin family member 1B	0.4773	1.74079e–
224610_at		full-length cDNA clone CL0BB014ZH04 of neuroblastoma of *Homo sapiens* (human)	–0.4772	1.74688e–
201930_at	*MCM6*	MCM6 minichromosome maintenance deficient 6 (MIS5 homolog, *Schizosaccharomyces pombe*) (*S. cerevisiae*)	–0.4768	1.76935e–
201110_s_at	*THBS1*	thrombospondin 1	0.4755	1.85546e–
226099_at	*ELL2*	elongation factor, RNA polymerase II, 2	0.4749	1.89665e–
210547_x_at	*ICA1*	islet cell autoantigen 1, 69 kDa	–0.4746	1.91645e–
201051_at	*ANP32A*	acidic (leucine-rich) nuclear phosphoprotein 32 family, member A	–0.4744	1.93198e–
201185_at	*HTRA1*	HtrA serine peptidase 1	0.4742	1.94315e–
224663_s_at	*CFL2*	cofilin 2 (muscle)	0.4735	1.99526e–
217843_s_at	*MED4*	mediator of RNA polymerase II transcription, subunit 4 homolog (yeast)	–0.4724	2.06753e–
224823_at	*MYLK*	myosin, light polypeptide kinase	0.4722	2.08417e–
222036_s_at	*MCM4*	MCM4 minichromosome maintenance deficient 4 (*S. cerevisiae*)	–0.4720	2.10334e–
214752_x_at	*FLNA*	filamin A, alpha (actin binding protein 280)	0.4719	2.11121e–
200668_s_at	*UBE2D3*	ubiquitin-conjugating enzyme E2D 3 (UBC4/5 homolog, yeast)	–0.4718	2.11299e–
205207_at	*IL6*	interleukin 6 (interferon, beta 2)	0.4717	2.12268e–
211160_x_at	*ACTN1*	actinin, alpha 1	0.4713	2.15198e–
222880_at	*AKT3*	v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma)	0.4707	2.19914e–
229067_at		similar to Formin binding protein 2 (srGAP2)	0.4702	2.23957e–
218420_s_at	*C13orf23*	chromosome 13 open reading frame 23	–0.4697	2.27551e–
226574_at	*PSPC1*; *LOC374491*	paraspeckle component 1; TPTE and PTEN homologous inositol lipid phosphatase pseudogene	–0.4697	2.27551e–
221563_at	*DUSP10*	dual-specificity phosphatase 10	0.4690	.0002335
224773_at	*NAV1*	neuron navigator 1	0.4688	2.34631e–
202074_s_at	*OPTN*	optineurin	0.4686	2.3650e–0
211981_at	*COL4A1*	collagen, type IV, alpha 1	0.4682	2.40005e–
202830_s_at	*SLC37A4*	solute carrier family 37 (glycerol-6-phosphate transporter), member 4	–0.4681	2.40548e–
200788_s_at	*PEA15*	phosphoprotein enriched in astrocytes 15	0.4680	2.41912e–
205129_at	*NPM3*	nucleophosmin/nucleoplasmin, 3	–0.4678	2.43558e–
221829_s_at	*TNPO1*	transportin 1	0.4675	2.45842e–
233496_s_at	*CFL2*	cofilin 2 (muscle)	0.4661	2.58543e–

Table W2. (*continued*)

Affy ID	Gene Symbol	Description	ρ	P
205130_at	*RAGE*	renal tumor antigen	0.4659	2.59708e–
209897_s_at	*SLIT2*	slit homolog 2 (*Drosophila*)	0.4658	2.60584e–
217872_at	*PIH1D1*	hypothetical protein FLJ20643	–0.4656	2.62935e–
225464_at	*FRMD6*	FERM domain containing 6	0.4653	2.65007e–
225398_at	*RPUSD4*	RNA pseudouridylate synthase domain containing 4	–0.4653	2.65304e–
225703_at	*KIAA1545*	KIAA1545 protein	–0.4653	2.65602e–
200663_at	*CD63*	CD63 antigen (melanoma 1 antigen)	0.4650	2.68074e–
235706_at	*CPM*	carboxypeptidase M	0.4650	2.68294e–
208808_s_at	*HMGB2*	high-mobility group box 2	–0.4648	2.6980e–0
220750_s_at	*LEPRE1*	leucine-proline–enriched proteoglycan (leprecan) 1	0.4648	2.70102e–
226621_at	*OSMR*	oncostatin M receptor	0.4639	2.78689e–
200099_s_at	*RPS3A*	ribosomal protein S3A	–0.4636	2.81501e–
204948_s_at	*FST*	follistatin transformer-2 alpha	0.4635	2.82529e–
204658_at			–0.4627	2.89772e–
213069_at	*HEG1*	HEG homolog 1 (zebrafish)	0.4625	2.92361e–
235570_at	*RBMS3*	RNA binding motif, single-stranded interacting protein	0.4624	2.93337e–
235138_at	*PUM2*	Pumilio homolog 2 (*Drosophila*)	–0.4624	2.93663e–
235705_at			0.4621	2.96284e–
200073_s_at	*HNRPD*	heterogeneous nuclear ribonucleoprotein D (AU-rich element RNA binding protein 1, 37 kDa)	–0.4621	2.96613e–
204425_at	*ARHGAP4*	Rho GTPase activating protein 4	–0.4619	2.97933e–
221096_s_at	*TMCO6*	hypothetical protein PRO1580	–0.4619	2.98263e–
214247_s_at	*DKK3*	dickkopf homolog 3 (*Xenopus laevis*)	0.4619	2.98594e–
222222_s_at	*HOMER3*	homer homolog 3 (*Drosophila*)	0.4615	3.01950e–
231045_x_at	*C11orf31*	chromosome 11 open reading frame 31	–0.4614	3.02929e–
214212_x_at	*PLEKHC1*	pleckstrin homology domain containing, family C (with FERM domain) member 1	0.4613	3.03937e–
217882_at	*TMEM111*	30 kDa protein	0.4613	3.04274e–
216232_s_at	*GCN1L1*	GCN1 general control of amino acid synthesis 1–like 1 (yeast)	–0.4612	3.05286e–
226084_at	*MAP1B*	microtubule-associated protein 1B	0.4612	3.05625e–
226982_at	*ELL2*	elongation factor, RNA polymerase II, 2	0.4611	3.05963e–
204032_at	*BCAR3*	breast cancer anti–estrogen resistance 3	0.4611	3.06981e–
224694_at	*ANTXR1*	anthrax toxin receptor 1	0.4608	3.10052e–
210416_s_at	*CHEK2*	CHK2 checkpoint homolog (*S. pombe*)	–0.4605	3.13244e–
202052_s_at	*RAI14*	retinoic acid–induced 14	0.4601	3.16977e–
229663_at	*LONPL*	peroxisomal LON protease like	0.4600	3.17677e–
208637_x_at	*ACTN1*	actinin, alpha 1	0.4597	3.21908e–
212457_at	*TFE3*	transcription factor binding to IGHM enhancer 3	0.4596	3.22618e–
226697_at	*LOC92689*	hypothetical protein BC001096	0.4595	3.23329e–
200755_s_at	*CALU*	calumenin	0.4591	3.27629e–
225146_at	*C9orf25*	chromosome 9 open reading frame 25	0.4588	3.31615e–
213713_s_at		hypothetical protein BC008326	–0.4587	.0003327
213052_at	*PRKAR2A*	protein kinase, cAMP-dependent, regulatory, type II, alpha	0.4587	3.33076e–
201471_s_at	*SQSTM1*	sequestosome 1	0.4585	3.34542e–
208973_at	*PRNPIP*	prion protein interacting protein	0.4585	3.34910e–
204009_s_at	*KRAS*	v-Ki-*ras2* Kirsten rat sarcoma viral oncogene homolog	–0.4582	3.38605e–
223206_s_at	*HSCARG*	HSCARG protein	–0.4581	3.39721e–
224774_s_at	*NAV1*	neuron navigator 1	0.4580	3.40094e–
218848_at	*THOC6*	WD repeat domain 58	–0.4580	3.40840e–
200772_x_at	*PTMA*	prothymosin, alpha (gene sequence 28)	–0.4579	3.41877e–
218448_at	*C20orf11*	chromosome 20 open reading frame 11	–0.4573	3.48387e–
212330_at	*TFDP1*	transcription factor Dp-1	–0.4570	3.51833e–
235308_at	*ZBTB20*	zinc finger and BTB domain containing 20	0.4570	3.52603e–
211729_x_at	*BLVRA*	biliverdin reductase A	0.4569	3.53761e–
200047_s_at	*YY1*	YY1 transcription factor	–0.4566	3.56971e–
211921_x_at	*PTMA*	prothymosin, alpha (gene sequence 28)	–0.4562	3.61173e–
201108_s_at	*THBS1*	thrombospondin 1	0.4559	3.65131e–
224416_s_at	*MED28*	mediator of RNA polymerase II transcription, subunit 28 homolog (yeast)	–0.4558	3.66725e–
232816_s_at	*DDX11*	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 11 (CHL1-like helicase homolog, *S. cerevisiae*)	–0.4554	3.71858e–
214150_x_at	*ATP6V0E1*	ATPase, H+ transporting, lysosomal 9 kDa, V0 subunit e	0.4553	3.72760e–
218407_x_at	*NENF*	neuron-derived neurotrophic factor	0.4545	3.82509e–
226777_at	*ADAM12*	ADAM metallopeptidase domain 12 (meltrin alpha)	0.4545	3.83018e–
221641_s_at	*ACOT9*	acyl-CoA thioesterase 9	0.4539	3.90986e–
214845_s_at	*CALU*	calumenin	0.4537	3.93959e–
208804_s_at	*SFRS6*	splicing factor, arginine/serine–rich 6	–0.4530	4.02132e–
200753_x_at	*SFRS2*	splicing factor, arginine/serine–rich 2	–0.4530	4.03002e–
225688_s_at	*PHLDB2*	pleckstrin homology–like domain, family B, member 2	0.4529	4.04309e–
202330_s_at	*UNG*	uracil-DNA glycosylase	–0.4529	4.04309e–
202164_s_at	*CNOT8*	CCR4-NOT transcription complex, subunit 8	–0.4527	4.06934e–
204490_s_at	*CD44*	CD44 antigen (homing function and Indian blood group system)	0.4526	4.07374e–
224840_at	*FKBP5*	FK506 binding protein 5	–0.4525	4.09697e–
223245_at	*STRBP*	spermatid perinuclear RNA binding protein	–0.4525	4.10017e–
208896_at	*DDX18*	DEAD (Asp-Glu-Ala-Asp) box polypeptide 18	–0.4523	4.11789e–
204768_s_at	*FEN1*	flap structure-specific endonuclease 1	–0.4521	4.15352e–
212593_s_at	*PDCD4*	programmed cell death 4 (neoplastic transformation inhibitor)	–0.4515	4.22562e–

Table W2. (*continued*)

Affy ID	Gene Symbol	Description	ρ	P
215446_s_at	*LOX*	lysyl oxidase	0.4514	4.24383e–
203167_at	*TIMP2*	TIMP metallopeptidase inhibitor 2	0.4512	4.27126e–
212419_at	*C10orf56*	chromosome 10 open reading frame 56	0.4512	4.27585e–
227771_at	*LIFR*	leukemia inhibitory factor receptor	0.4511	4.28504e–
214257_s_at	*SEC22B*	SEC22 vesicle trafficking protein–like 1 (*S. cerevisiae*)	0.4510	4.30348e–
225798_at	*JAZF1*	juxtaposed with another zinc finger gene 1	0.4509	4.31735e–
218069_at		XTP3-transactivated protein A	–0.4504	4.39203e–
221881_s_at	*CLIC4*	chloride intracellular channel 4	0.4499	4.45834e–
202303_x_at	*SMARCA5*	SWI/SNF–related, matrix-associated, actin-dependent regulator of chromatin, subfamily a, member 5	–0.4496	4.50625e–
201493_s_at	*PUM2*	pumilio homolog 2 (*Drosophila*)	–0.4495	4.52072e–
213610_s_at	*KLHL23*	kelch-like 23 (*Drosophila*)	–0.4494	4.54007e–
226751_at	*C2orf32*	chromosome 2 open reading frame 32	0.4493	4.54978e–
209079_x_at	*PCDHG*	protocadherin gamma subfamily	0.4491	4.57900e–
219944_at	*CLIP4*	restin-like 2	0.4491	4.58389e–
227740_at	*UHMK1*	U2AF homology motif (UHM) kinase 1	0.4489	4.61332e–
212195_at	*IL6ST*	interleukin 6 signal transducer (gp130, oncostatin M receptor)	0.4488	4.62316e–
201376_s_at	*HNRPF*	heterogeneous nuclear ribonucleoprotein F	–0.4488	4.63303e–
210139_s_at	*PMP22*	peripheral myelin protein 22	0.4488	4.63303e–
201559_s_at	*CLIC4*	chloride intracellular channel 4	0.4486	4.65282e–
218067_s_at		hypothetical protein FLJ10154	–0.4485	4.66771e–
208682_s_at	*MAGED2*	melanoma antigen family D, 2	0.4484	4.68763e–
57163_at	*ELOVL1*	elongation of very long chain fatty acids (FEN1/Elo2, SUR4/Elo3, yeast)–like 1	0.4483	4.70262e–
225563_at	*PAN3*	PABP1-dependent poly A–specific ribonuclease subunit PAN3	–0.4483	4.70262e–
202391_at	*BASP1*	brain-abundant, membrane-attached signal protein 1	0.4480	4.75291e–
200956_s_at	*SSRP1*	structure-specific recognition protein 1	–0.4479	4.76303e–
213329_at	*SRGAP2*	SLIT-ROBO Rho GTPase activating protein 2	0.4476	4.80369e–
210371_s_at	*RBBP4*	retinoblastoma binding protein 4	–0.4475	4.82414e–
209754_s_at	*TMPO*	thymopoietin	–0.4475	4.83439e–
225391_at	*LOC93622*	hypothetical protein BC006130	–0.4474	4.84981e–
217784_at	*YKT6*	SNARE protein Ykt6	0.4471	4.89633e–
44702_at	*SYDE1*	synapse defective 1, Rho GTPase, homolog 1 (*Caenorhabditis elegans*)	0.4466	4.97476e–
216262_s_at	*TGIF2*	TGFB-induced factor 2 (TALE family homeobox)	–0.4466	4.97476e–
202259_s_at		phosphonoformate immunoassociated protein 5	–0.4465	4.98003e–
205909_at	*POLE2*	polymerase (DNA-directed), epsilon 2 (p59 subunit)	–0.4465	4.98531e–
204819_at	*FGD1*	FYVE, RhoGEF and PH domain containing 1 (faciogenital dysplasia)	0.4464	4.99587e–
225381_at	*LOC399959*	hypothetical gene supported by BX647608	0.4463	5.02768e–
218298_s_at	*C14orf159*	chromosome 14 open reading frame 159	–0.4462	5.03833e–
215836_s_at	*PCDHG*	protocadherin gamma subfamily	0.4461	5.04899e–
209750_at	*NR1D2*	nuclear receptor subfamily 1, group D, member 2	0.4459	5.08111e–
201035_s_at	*HADH*	L-3-hydroxyacyl–Coenzyme A dehydrogenase, short chain	–0.4458	5.10802e–
226342_at	*SPTBN1*	Spectrin, beta, non–erythrocytic 1	0.4458	5.10802e–
226934_at	*CPSF6*	Cleavage and polyadenylation specific factor 6, 68 kDa	–0.4449	5.25556e–
231810_at	*BRI3BP*	BRI3 binding protein	–0.4445	5.31677e–
219779_at	*ZFHX4*	zinc finger homeodomain 4	0.4443	5.36169e–
210904_s_at	*IL13RA1*	interleukin 13 receptor, alpha 1	0.4442	5.36733e–
202258_s_at		phosphonoformate-immunoassociated protein 5	–0.4441	5.38994e–
226912_at	*ZDHHC23*	zinc finger, DHHC-type containing 23	–0.4438	5.43542e–
218168_s_at	*CABC1*	chaperone, ABC1 activity of bc1 complex like (*S. pombe*)	–0.4436	5.48125e–
203632_s_at	*GPRC5B*	G protein–coupled receptor, family C, group 5, member B	0.4435	5.49852e–
241879_at		Transcribed locus	0.4435	5.50429e–
200041_s_at	*BAT1*	HLA-B–associated transcript 1	–0.4437	5.53484e–
228174_at	*GOLGA1*	Golgi autoantigen, golgin subfamily a, 1	–0.4428	5.62083e–
219520_s_at	*WWC3*	KIAA1280 protein	0.4428	5.62083e–
228121_at	*TGFB2*	Transforming growth factor, beta 2	0.4423	5.70970e–
200957_s_at	*SSRP1*	structure-specific recognition protein 1	–0.4422	5.72763e–
201616_s_at	*CALD1*	caldesmon 1	0.4421	5.75762e–
232113_at		Hypothetical gene supported by BX647608	0.4421	5.75762e–
217828_at	*SLTM*	modulator of estrogen induced transcription	–0.4420	5.76363e–
219557_s_at	*NRIP3*	nuclear receptor interacting protein 3	0.4418	5.79984e–
223711_s_at	*THY28*	thymocyte protein Thy28	–0.4415	5.86065e–
218656_s_at	*LHFP*	lipoma HMGIC fusion partner	0.4415	5.87153e–
230130_at		Transcribed locus	0.4414	5.88513e–
212709_at	*NUP160*	nucleoporin 160 kDa	–0.4411	5.94056e–
202125_s_at	*TRAK2*	amyotrophic lateral sclerosis 2 (juvenile) chromosome region, candidate 3	0.4410	5.95294e–
200679_x_at	*HMGB1*	high-mobility group box 1	–0.4409	5.98399e–
227484_at		CDNA FLJ41690 fis, clone HCASM2009405	0.4405	6.04654e–
222204_s_at	*RRN3*	RRN3 RNA polymerase I transcription factor homolog (yeast)	–0.4403	6.08436e–
214882_s_at	*SFRS2*	splicing factor, arginine/serine–rich 2	–0.4403	6.08436e–
213836_s_at	*WIPI1*	WD40 repeat protein interacting with phosphoinositides of 49 kDa	0.4401	6.12874e–
208178_x_at	*TRIO*	triple functional domain (PTPRF interacting)	0.4401	6.14008e–
224856_at	*FKBP5*	FK506 binding protein 5	–0.4401	6.14148e–
213238_at	*ATP10D*	ATPase, class V, type 10D	0.4399	6.17983e–
201450_s_at	*TIA1*	TIA1 cytotoxic granule–associated RNA binding protein	–0.4398	6.18624e–

Table W2. (*continued*)

Affy ID	Gene Symbol	Description	ρ	P
201712_s_at	*RANBP2*	RAN binding protein 2	–0.4395	6.24281e–
200977_s_at	*TAX1BP1*	Tax1 (human T-cell leukemia virus type I) binding protein 1	0.4393	6.28315e–
225685_at	*CDC42EP3*	CDC42 effector protein (Rho GTPase binding) 3	0.4390	6.34196e–
213139_at	*SNAI2*	snail homolog 2 (*Drosophila*)	0.4390	6.34196e–
202948_at	*IL1R1*	interleukin 1 receptor, type I	0.4390	6.34853e–
203317_at	*PSD4*	pleckstrin and Sec7 domain containing 4	–0.4387	6.40974e–
202238_s_at	*NNMT*	nicotinamide *N*-methyltransferase	0.4386	6.42778e–
208795_s_at	*MCM7*	MCM7 minichromosome maintenance deficient 7 (*S. cerevisiae*)	–0.4386	6.43443e–
233364_s_at		Hypothetical gene supported by BX647608	0.4386	6.43443e–
201848_s_at	*BNIP3*	BCL2/adenovirus E1B 19 kDa interacting protein 3	0.4384	6.48111e–
203231_s_at	*ATXN1*	ataxin 1	0.4382	6.50793e–
212233_at	*MAP1B*	Microtubule-associated protein 1B; *Homo sapiens*, clone IMAGE:5535936, mRNA	0.4381	6.54159e–
208747_s_at	*C1S*	complement component 1, s subcomponent	0.4380	6.56186e–
200999_s_at	*CKAP4*	cytoskeleton-associated protein 4	0.4379	6.57541e–
214109_at	*LRBA*	LPS-responsive vesicle trafficking, beach and anchor containing	–0.4377	6.62302e–
213338_at	*TMEM158*	Ras-induced senescence 1	0.4374	6.67782e–
224940_s_at	*PAPPA*	pregnancy-associated plasma protein A, pappalysin 1	0.4374	6.69158e–
207749_s_at	*PPP2R3A*	protein phosphatase 2 (formerly 2A), regulatory subunit B″, alpha	0.4372	6.72805e–
203856_at	*VRK1*	vaccinia-related kinase 1	–0.4371	6.74689e–
212603_at	*MRPS31*	mitochondrial ribosomal protein S31	–0.4368	6.80261e–
236533_at	*DDEF1*	development and differentiation enhancing factor 1	0.4367	6.82909e–
212789_at	*NCAPD3*	KIAA0056 protein	–0.4366	6.85170e–
204647_at	*HOMER3*	homer homolog 3 (*Drosophila*)	0.4366	6.85874e–
203276_at	*LMNB1*	lamin B1	–0.4366	6.85874e–
201170_s_at	*BHLHB2*	basic helix-loop-helix domain containing, class B, 2	0.4364	6.90112e–
202302_s_at	*RSRC2*	similar to splicing factor, arginine/serine–rich 4	–0.4363	6.92239e–
212443_at	*NBEAL2*	neurobeachin-like 2	–0.4362	6.94373e–
212418_at	*ELF1*	E74-like factor 1 (ets domain transcription factor)	–0.4361	6.95891e–
207949_s_at	*ICA1*	islet cell autoantigen 1, 69 kDa	–0.4360	6.99374e–
209421_at	*MSH2*	mutS homolog 2, colon cancer, nonpolyposis type 1 (*Escherichia coli*)	–0.4359	7.00092e–
218443_s_at	*DAZAP1*	DAZ-associated protein 1	–0.4359	7.00809e–
225273_at	*KIAA1280*	KIAA1280 protein	0.4359	7.00809e–
203229_s_at	*CLK2*	CDC-like kinase 2	–0.4356	7.0730e–0
213798_s_at	*CAP1*	CAP, adenylate cyclase–associated protein 1 (yeast)	0.4356	7.07504e–
216969_s_at	*KIF22*	kinesin family member 22	–0.4355	7.10202e–
204489_s_at	*CD44*	CD44 antigen (homing function and Indian blood group system)	0.4354	7.10929e–
208908_s_at	*CAST*	calpastatin	0.4353	7.13844e–
225481_at	*FRMD6*	FERM domain containing 6	0.4351	7.18237e–
203388_at	*ARRB2*	arrestin, beta 2	–0.4351	7.19707e–
212014_x_at	*CD44*	CD44 antigen (homing function and Indian blood group system)	0.4348	7.26355e–
218295_s_at	*NUP50*	nucleoporin 50 kDa	–0.4344	7.34556e–
214663_at	*RIPK5*	receptor interacting protein kinase 5	0.4343	7.36056e–
200770_s_at	*LAMC1*	laminin, gamma 1 (formerly LAMB2)	0.4343	7.36807e–
218714_at	*PRR14*	hypothetical protein MGC3121	–0.4342	7.39818e–
212822_at	*HEG1*	HEG homolog 1 (zebrafish)	0.4340	7.44950e–
212920_at			–0.4338	7.49516e–
217892_s_at	*LIMA1*	epithelial protein lost in neoplasm beta	0.4337	7.51209e–
226425_at	*RSNL2*	restin-like 2	0.4337	7.51209e–
203358_s_at	*EZH2*	enhancer of zeste homolog 2 (*Drosophila*)	–0.4337	7.51974e–
200787_s_at	*PEA15*	phosphoprotein enriched in astrocytes 15	0.4334	7.57348e–
215947_s_at		hypothetical protein FLJ14668	–0.4332	7.62202e–
232524_x_at	*ANAPC4*	anaphase promoting complex subunit 4	–0.4330	7.67422e–
39729_at	*PRDX2*	peroxiredoxin 2	–0.4328	7.71328e–
201849_at	*BNIP3*	BCL2/adenovirus E1B 19 kDa interacting protein 3	0.4327	7.75252e–
241418_at		Hypothetical LOC344887	0.4326	7.78405e–
200625_s_at	*CAP1*	CAP, adenylate cyclase–associated protein 1 (yeast)	0.4324	7.81569e–
203729_at	*EMP3*	epithelial membrane protein 3	0.4323	7.85540e–
227846_at	*GPR*	putative G protein–coupled receptor	0.4322	7.87134e–
222872_x_at	*FLJ22833*	hypothetical protein FLJ22833	0.4321	7.89357e–
201430_s_at	*DPYSL3*	dihydropyrimidinase-like 3	0.4318	7.97564e–
205780_at	*BIK*	BCL2-interacting killer (apoptosis-inducing)	–0.4316	8.01609e–
231271_x_at	*HSCARG*	HSCARG protein	–0.4315	8.04276e–
226405_s_at	*ARRDC1*	arrestin domain containing 1	–0.4315	8.04858e–
201560_at	*CLIC4*	chloride intracellular channel 4	0.4313	8.09753e–
223154_at	*MRPL1*	mitochondrial ribosomal protein L1	–0.4313	8.10572e–
211506_s_at	*IL8*	interleukin 8	0.4311	8.14676e–
221081_s_at	*DENND2D*	DENN/MADD domain containing 2D	–0.4310	8.16323e–
226281_at	*DNER*	delta-notch–like EGF repeat-containing transmembrane	0.4310	8.16323e–
211376_s_at	*NSMCE4A*	chromosome 10 open reading frame 86	–0.4308	8.21282e–
221191_at		DKFZp434A0131 protein	–0.4308	8.21282e–
205308_at	*C8orf70*	chromosome 8 open reading frame 70	0.4306	8.26267e–
206237_s_at	*NRG1*	neuregulin 1	0.4306	8.27101e–
200606_at	*DSP*	desmoplakin	–0.4305	8.30443e–

Table W2. (*continued*)

Affy ID	Gene Symbol	Description	ρ	P
242673_at			−0.4303	8.35479e–
201865_x_at	*NR3C1*	nuclear receptor subfamily 3, group C, member 1 (glucocorticoid receptor)	0.4301	8.41389e–
226760_at	*LOC203411*	hypothetical protein LOC203411	0.4300	8.43933e–
212056_at	*KIAA0182*	KIAA0182 protein	−0.4296	8.54180e–
217788_s_at	*GALNT2*	UDP–*N*-acetyl-α-D-galactosamine: polypeptide *N*-acetylgalactosaminyltransferase 2 (GalNAc-T2)	0.4295	8.56575e–
214744_s_at	*RPL23*	ribosomal protein L23	−0.4294	8.58483e–
214446_at	*ELL2*	elongation factor, RNA polymerase II, 2	0.4291	8.67147e–
212267_at	*WAPAL*	KIAA0261	−0.4290	8.68017e–
228009_x_at	*ZNRD1*	zinc ribbon domain containing, 1	−0.4290	8.68889e–
211057_at	*ROR1*	receptor tyrosine kinase–like orphan receptor 1	0.4290	8.69574e–
214927_at	*ITGBL1*	integrin, beta-like 1 (with EGF-like repeat domains)	0.4290	8.69761e–
233167_at	*SELO*	selenoprotein O	−0.4288	8.75889e–
210757_x_at	*DAB2*	disabled homolog 2, mitogen-responsive phosphoprotein (*Drosophila*)	0.4287	8.78528e–
210095_s_at	*IGFBP3*	insulin-like growth factor binding protein 3	0.4286	8.79409e–
204160_s_at	*ENPP4*	ectonucleotide pyrophosphatase/phosphodiesterase 4 (putative function)	−0.4285	8.82056e–
234074_at		CDNA FLJ10946 fis, clone PLACE1000005	0.4284	8.84711e–
202013_s_at	*EXT2*	exostoses (multiple) 2	0.4281	8.92975e–
224886_at	*LOC339123*	hypothetical LOC339123	−0.4279	8.98993e–
202828_s_at	*MMP14*	matrix metallopeptidase 14 (membrane-inserted)	0.4279	8.98993e–
224942_at	*PAPPA*	pregnancy-associated plasma protein A, pappalysin 1	0.4279	8.99893e–
210212_x_at	*MTCP1*	mature T-cell proliferation 1	0.4278	9.03499e–
204115_at	*GNG11*	guanine nucleotide binding protein (G protein), gamma 11	0.4277	9.04402e–
202532_s_at	*DHFR*	dihydrofolate reductase	−0.4276	9.08024e–
202326_at	*EHMT2*	euchromatic histone-lysine *N*-methyltransferase 2	−0.4275	9.09646e–
202998_s_at	*LOXL2*	lysyl oxidase–like 2	0.4274	9.12571e–
202822_at	*LPP*	LIM domain containing preferred translocation partner in lipoma	0.4274	9.13744e–
217762_s_at	*RAB31*	RAB31, member RAS oncogene family	0.4273	9.15308e–
203441_s_at	*CDH2*	cadherin 2, type 1, N-cadherin (neuronal)	0.4273	9.16222e–
204853_at	*ORC2L*	origin recognition complex, subunit 2–like (yeast)	−0.4273	9.16222e–
227561_at	*DDR2*	Discoidin domain receptor family, member 2	0.4271	9.22645e–
212077_at	*CALD1*	caldesmon 1	0.4268	9.30961e–
218170_at	*ISOC1*	isochorismatase domain containing 1	−0.4266	9.36543e–
241353_s_at	*LOC202775*; *LOC402617*	hypothetical LOC202775; hypothetical LOC402617	0.4265	9.40281e–
232254_at	*FBXO25*	F-box protein 25	0.4260	9.54419e–
218688_at	*DAK*	dihydroxyacetone kinase 2 homolog (yeast)	−0.4259	9.57411e–
218028_at	*ELOVL1*	elongation of very long chain fatty acids (FEN1/Elo2, SUR4/Elo3, yeast)–like 1	0.4258	9.59174e–
212723_at	*JMJD6*	phosphatidylserine receptor	0.4257	9.62037e–
209127_s_at	*SART3*	squamous cell carcinoma antigen recognized by T cells 3	−0.4257	9.62037e–
202162_s_at	*CNOT8*	CCR4-NOT transcription complex, subunit 8	−0.4256	9.63950e–
201069_at	*MMP2*	matrix metallopeptidase 2 (gelatinase A, 72-kDa gelatinase, 72-kDa type IV collagenase)	0.4256	9.65866e–
241763_s_at			0.4255	9.66826e–
227129_x_at	*LOC284701*	hypothetical protein LOC284701	−0.4253	9.73566e–
209143_s_at	*CLNS1A*; CLDND1	chloride channel, nucleotide-sensitive, 1A; chromosome 3 open reading frame 4	−0.4253	9.74533e–
212097_at	*CAV1*	caveolin 1, caveolae protein, 22 kDa	0.4251	9.79378e–
214075_at	*NENF*	neuron-derived neurotrophic factor	0.4249	9.85220e–
204341_at	*TRIM16*	tripartite motif-containing 16; similar to tripartite motif-containing 16; estrogen-responsive B box protein	0.4249	9.87175e–
215489_x_at	*HOMER3*	homer homolog 3 (*Drosophila*)	0.4248	9.88153e–
222900_at	*NRIP3*	nuclear receptor interacting protein 3	0.4248	9.89905e–
213793_s_at	*HOMER1*	homer homolog 1 (*Drosophila*)	−0.4246	9.94043e–
228297_at	*CNN3*	calponin 3, acidic	0.4245	9.97987e–

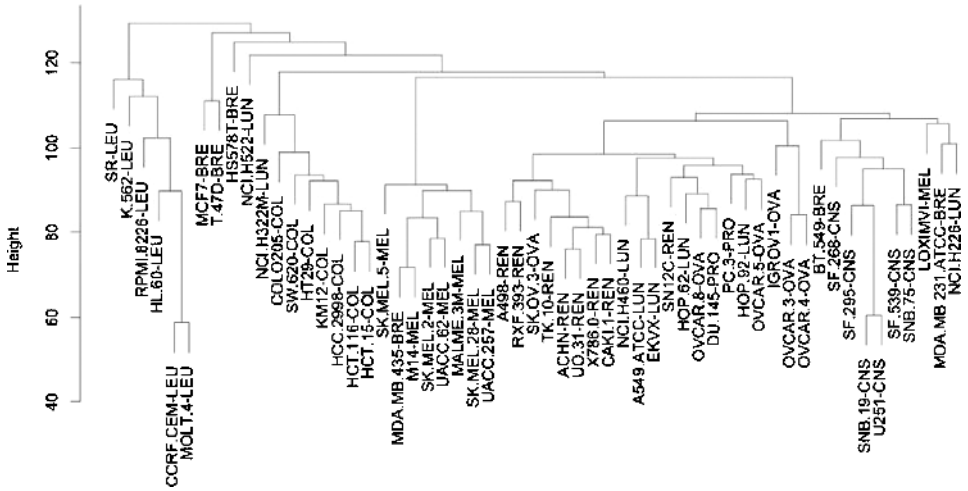

Figure W1. Hierarchical clustering of NCI-60 cell lines using Euclidean distance and average linkage (according to the expression of 11,848 genes).

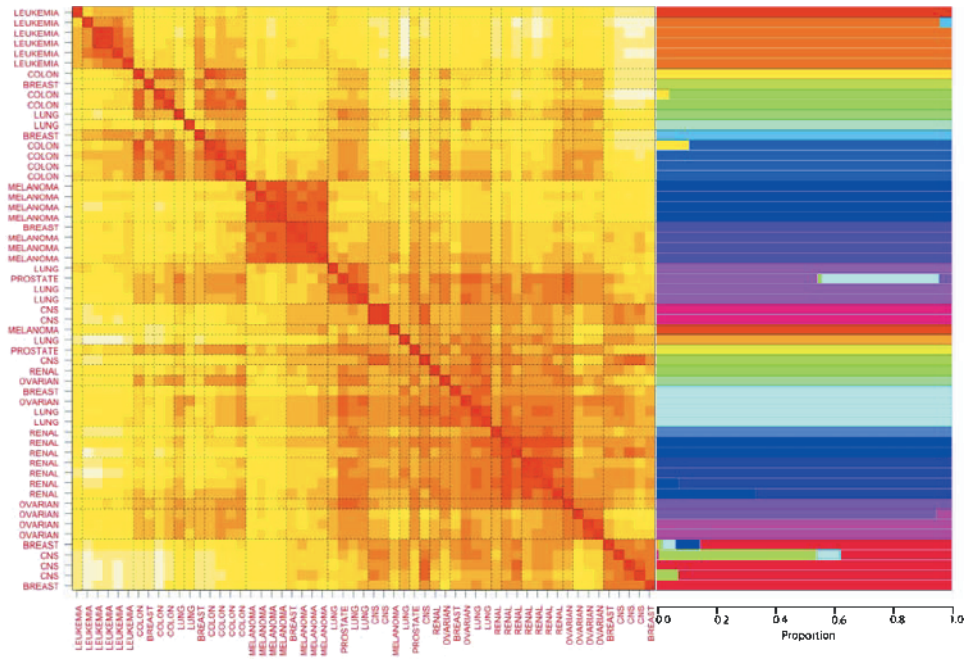

Figure W2. HOPACH clustering of NCI-60 cell lines according to expression of 2349 genes selected unspecifically. Red color indicates small distance; white, large distance. Cell lines are ordered according to the final level of the cluster tree. The proportion of the estimated membership within a cluster is plotted on the right part of figure. Each cluster is represented by a different color. Dotted lines indicate the cluster boundaries at the tree level with minimum median split silhouette.

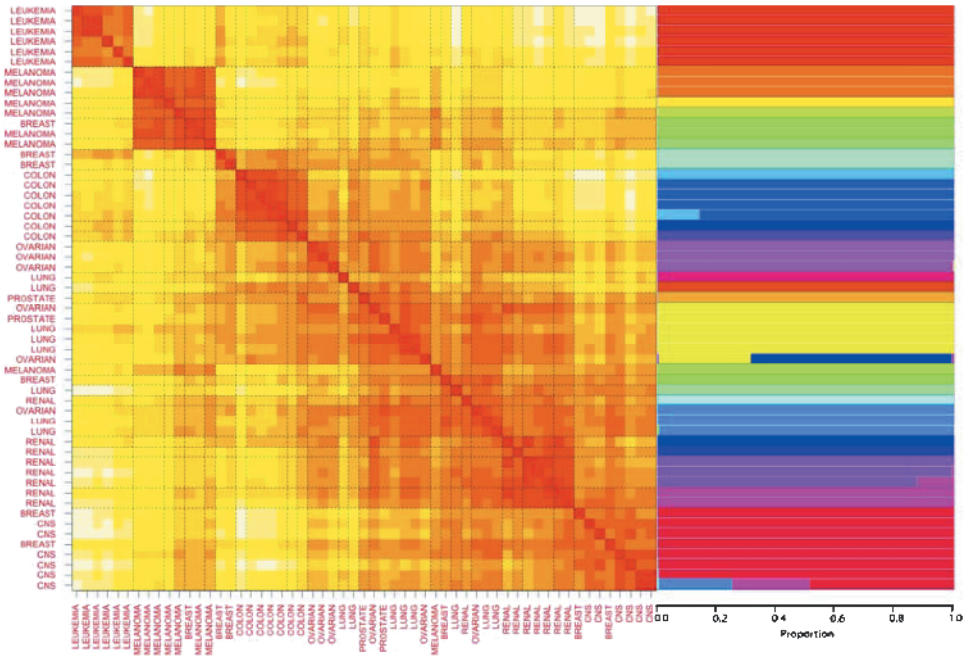

Figure W3. HOPACH clustering of NCI-60 cell lines according to genes differentially expressed between cell lines derived from different tissues. Red color indicates small distance; white, large distance. Cell lines are ordered according to the final level of the cluster tree. The proportion of the estimated membership within a cluster is plotted on the right part of figure. Each cluster is represented by a different color. Dotted lines indicate the cluster boundaries at the tree level with minimum median split silhouette.

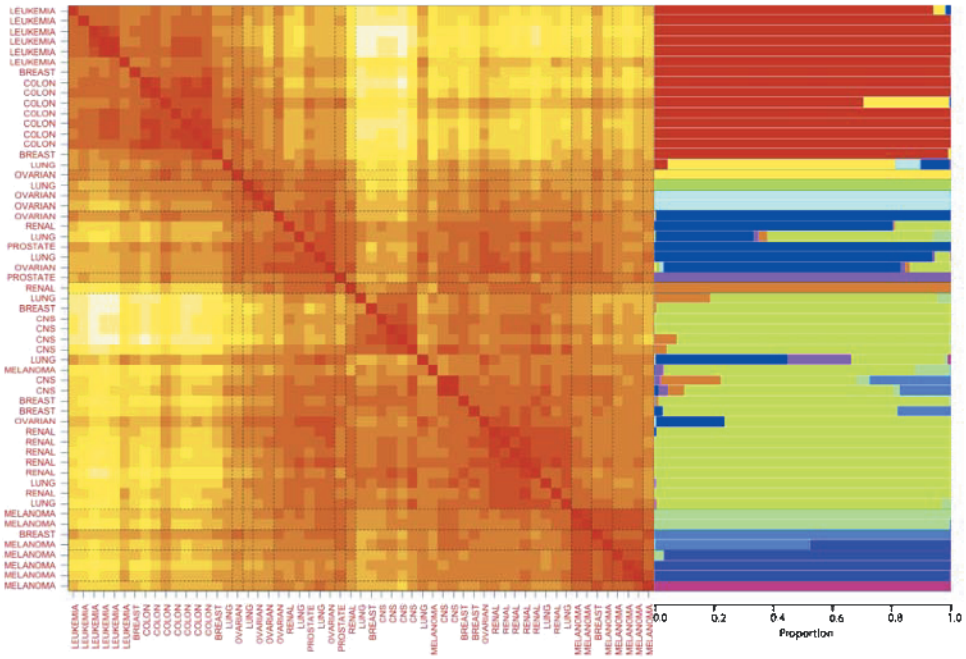

Figure W4. HOPACH clustering of NCI-60 cell lines according to genes whose expression correlated with numerical heterogeneity index. Red color indicates small distance; white, large distance. Cell lines are ordered according to the final level of the cluster tree. The proportion of the estimated membership within a cluster is plotted on the right part of figure. Each cluster is represented by a different color. Dotted lines indicate the cluster boundaries at the tree level with minimum median split silhouette.

Table W3. List of 213 Array Features and Names of Corresponding Genes with Expression Profiles Correlated with NH ($P < .001$) Across the NCI-60 Panel of Cancer Cell Lines, After Exclusion of Array Elements Showing Differential Expression Based on Their Tissue of Origin.

Affy ID	Gene Symbol	Description	ρ	P
226148_at	*BTBD15*	BTB (POZ) domain containing 15	−0.6236	2.2e−07
218088_s_at	*RRAGC*	Ras-related GTP binding C	0.5750	2.89e−06
200069_at	*SART3*	squamous cell carcinoma antigen recognized by T cells 3	−0.5686	3.94e−06
226482_s_at	*F11R*	F11 receptor	−0.5620	5.38e−06
220092_s_at	*ANTXR1*	anthrax toxin receptor 1	0.5425	1.3e−05
202859_x_at	*IL8*	interleukin 8	0.5349	1.82e−05
201172_x_at	*ATP6V0E1*	ATPase, H+ transporting, lysosomal 9 kDa, V0 subunit e	0.5341	1.88e−05
218491_s_at	*THYN1*	thymocyte protein Thy28	−0.5300	2.24e−05
227560_at	*SFXN2*	sideroflexin 2	−0.5288	2.35e−05
212978_at	*LRRC8B*	leucine-rich repeat containing 8 family, member B	−0.5218	3.15e−05
212446_s_at	*LASS6*	LAG1 longevity assurance homolog 6 (*S.cerevisiae*)	−0.5215	3.18e−05
213149_at	*DLAT*	dihydrolipoamide *S*-acetyltransferase (E_2 component of pyruvate dehydrogenase complex)	−0.5189	3.55e−05
204767_s_at	*FEN1*	flap structure–specific endonuclease 1	−0.5170	3.83e−05
208901_s_at	*TOP1*	topoisomerase (DNA) I	−0.5163	3.94e−05
219378_at	*NARG1L*	NMDA receptor regulated 1–like	−0.5155	4.07e−05
225179_at	*HIP2*	Huntingtin interacting protein 2	−0.5151	4.13e−05
225845_at	*BTBD15*	BTB (POZ) domain containing 15	−0.5116	4.75e−05
211612_s_at	*IL13RA1*	interleukin 13 receptor, alpha 1	0.5114	4.8e−05
213251_at		Hypothetical LOC 441046	−0.5101	5.04e−05
223268_at	*PTD012*	PTD012 protein	−0.5091	5.26e−05
202778_s_at	*ZMYM2*	zinc finger protein 198	−0.5058	5.99e−05
212943_at	*KIAA0528*	KIAA0528 gene product	−0.5048	6.21e−05
201887_at	*IL13RA1*	interleukin 13 receptor, alpha 1	0.5048	6.23e−05
209225_x_at	*TNPO1*	transportin 1	0.5047	6.23e−05
212658_at	*LHFPL2*	lipoma HMGIC fusion partner–like 2	0.5044	6.31e−05
213696_s_at	*MED8*	mediator of RNA polymerase II transcription, subunit 8 homolog (yeast)	0.5043	6.33e−05
209013_x_at	*TRIO*	triple functional domain (PTPRF interacting)	0.5034	6.56e−05
224944_at	*TMPO*	Thymopoietin	−0.4982	8.01e−05
209019_s_at	*PINK1*	PTEN-induced putative kinase 1	0.4974	8.28e−05
213626_at	*CBR4*	carbonic reductase 4	−0.4966	8.54e−05
212470_at	*SPAG9*	sperm-associated antigen 9	0.4948	9.12e−05
204359_at	*FLRT2*	fibronectin leucine-rich transmembrane protein 2	0.4945	9.25e−05
203773_x_at	*BLVRA*	biliverdin reductase A	0.4921	.000101
202163_s_at	*CNOT8*	CCR4-NOT transcription complex, subunit 8	−0.4906	.000107
218458_at	*GMCL1*	germ cell–less homolog 1 (*Drosophila*)	−0.4900	.000109
221737_at	*GNA12*	guanine nucleotide binding protein (G protein) alpha12	0.4896	.000111
221753_at	*SSH1*	slingshot homolog 1 (*Drosophila*)	0.4895	.000111
212430_at	*RBM38*	RNA-binding region (RNP1, RRM) containing 1	−0.4884	.000116
200943_at	*HMGN1*	high-mobility group nucleosome binding domain 1	−0.4884	.000116
222514_at	*RRAGC*	Ras-related GTP binding C	0.4881	.000117
200990_at	*TRIM28*	tripartite motif-containing 28	−0.4871	.000122
209832_s_at	*CDT1*	DNA replication factor	−0.4865	.000125
224504_s_at	*MGC13125*	hypothetical protein MGC13125	−0.4862	.000126
217785_s_at	*YKT6*	SNARE protein Ykt6	0.4861	.000126
210053_at	*TAF5*	TAF5 RNA polymerase II, TATA box binding protein (TBP)–associated factor, 100 kDa	−0.4853	.00013
226651_at	*HOMER1*	homer homolog 1 (*Drosophila*)	−0.4844	.000135
201885_s_at	*CYB5R3*	cytochrome b5 reductase 3	0.4843	.000135
208549_x_at	*PTMA*	prothymosin, alpha (gene sequence 28); similar toprothymosin alpha; hypothetical gene supported by BC013859; hypothetical gene supported by BC013859; BC070480	−0.4839	.000137
205083_at	*AOX1*	aldehyde oxidase 1	0.4832	.000141
208838_at	*CAND1*	cullin-associated and neddylation-dissociated 1	−0.4822	.000146
224677_x_at	*C11orf31*	chromosome 11 open reading frame 31	−0.4820	.000147
206157_at	*PTX3*	pentraxin-related gene, rapidly induced by IL-1 beta	0.4807	.000154
209420_s_at	*SMPD1*	sphingomyelin phosphodiesterase 1, acid lysosomal (acid sphingomyelinase)	0.4806	.000155
224610_at		full-length cDNA clone CL0BB014ZH04 of neuroblastoma of *Homo sapiens* (human)	−0.4772	.000175
201930_at	*MCM6*	MCM6 minichromosome maintenance deficient 6 (MIS5 homolog, *S. pombe*) (*S. cerevisiae*)	−0.4768	.000177
226099_at	*ELL2*	elongation factor, RNA polymerase II, 2	0.4749	.00019
201051_at	*ANP32A*	acidic (leucine-rich) nuclear phosphoprotein 32 family, member A	−0.4744	.000193
201185_at	*HTRA1*	HtrA serine peptidase 1	0.4742	.000194
217843_s_at	*MED4*	mediator of RNA polymerase II transcription, subunit 4 homolog (yeast)	−0.4724	.000207
222036_s_at	*MCM4*	MCM4 minichromosome maintenance deficient 4 (*S. cerevisiae*)	−0.4720	.00021
200668_s_at	*UBE2D3*	ubiquitin-conjugating enzyme E2D 3 (UBC4/5 homolog, yeast)	−0.4718	.000211
205207_at	*IL6*	interleukin 6 (interferon, beta 2)	0.4717	.000212
218420_s_at	*C13orf23*	chromosome 13 open reading frame 23	−0.4697	.000228
226574_at	*PSPC1*; *LOC374491*	paraspeckle component 1; TPTE and PTEN homologous inositol lipid phosphatase pseudogene	−0.4697	.000228
202830_s_at	*SLC37A4*	solute carrier family 37 (glycerol-6-phosphatetransporter), member 4	−0.4681	.000241
205129_at	*NPM3*	nucleophosmin/nucleoplasmin, 3	−0.4678	.000244
209897_s_at	*SLIT2*	slit homolog 2 (*Drosophila*)	0.4658	.000261
217872_at	*PIH1D1*	hypothetical protein FLJ20643	−0.4656	.000263
225398_at	*RPUSD4*	RNA pseudouridylate synthase domain containing 4	−0.4653	.000265
235706_at	*CPM*	carboxypeptidase M	0.4650	.000268
220750_s_at	*LEPRE1*	leucine proline–enriched proteoglycan (leprecan) 1	0.4648	.00027

Table W3. (*continued*)

Affy ID	Gene Symbol	Description	ρ	P
200099_s_at	*RPS3A*	ribosomal protein S3A	–0.4636	.000282
204948_s_at	*FST*	follistatin	0.4635	.000283
204658_at		transformer-2 alpha	–0.4627	.00029
235138_at	*PUM2*	Pumilio homolog 2 (*Drosophila*)	–0.4624	.000294
221096_s_at	*TMCO6*	hypothetical protein PRO1580	–0.4619	.000298
222222_s_at	*HOMER3*	homer homolog 3 (*Drosophila*)	0.4615	.000302
231045_x_at	*C11orf31*	chromosome 11 open reading frame 31	–0.4614	.000303
214212_x_at	*PLEKHC1*	pleckstrin homology domain containing, family C (with FERM domain) member 1	0.4613	.000304
217882_at	*TMEM111*	30-kDa protein	0.4613	.000304
216232_s_at	*GCN1L1*	GCN1 general control of amino acid synthesis 1–like 1 (yeast)	–0.4612	.000305
226982_at	*ELL2*	elongation factor, RNA polymerase II, 2	0.4611	.000306
210416_s_at	*CHEK2*	CHK2 checkpoint homolog (*S. pombe*)	–0.4605	.000313
229663_at	*LONPL*	peroxisomal LON protease like	0.4600	.000318
213052_at	*PRKAR2A*	protein kinase, cAMP-dependent, regulatory, type II, alpha	0.4587	.000333
208973_at	*PRNPIP*	prion protein interacting protein	0.4585	.000335
204009_s_at	*KRAS*	v-Ki-*ras2* Kirsten rat sarcoma viral oncogene homolog	–0.4582	.000339
223206_s_at	*HSCARG*	HSCARG protein	–0.4581	.00034
218848_at	*THOC6*	WD repeat domain 58	–0.4580	.000341
218448_at	*C20orf11*	chromosome 20 open reading frame 11	–0.4573	.000348
211729_x_at	*BLVRA*	biliverdin reductase A	0.4569	.000354
232816_s_at	*DDX11*	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 11 (CHL1-like helicase homolog, *S. cerevisiae*)	–0.4554	.000372
214150_x_at	*ATP6V0E1*	ATPase, H+ transporting, lysosomal 9 kDa, V0 subunit e	0.4553	.000373
218407_x_at	*NENF*	neuron-derived neurotrophic factor	0.4545	.000383
221641_s_at	*ACOT9*	acyl-CoA thioesterase 9	0.4539	.000391
208804_s_at	*SFRS6*	splicing factor, arginine/serine–rich 6	–0.4530	.000402
200753_x_at	*SFRS2*	splicing factor, arginine/serine–rich 2	–0.4530	.000403
202330_s_at	*UNG*	uracil-DNA glycosylase	–0.4529	.000404
202164_s_at	*CNOT8*	CCR4-NOT transcription complex, subunit 8	–0.4527	.000407
224840_at	*FKBP5*	FK506 binding protein 5	–0.4525	.00041
208896_at	*DDX18*	DEAD (Asp-Glu-Ala-Asp) box polypeptide 18	–0.4523	.000412
204768_s_at	*FEN1*	flap structure–specific endonuclease 1	–0.4521	.000415
212593_s_at	*PDCD4*	programmed cell death 4 (neoplastic transformation inhibitor)	–0.4515	.000423
218069_at		XTP3-transactivated protein A	–0.4504	.000439
221881_s_at	*CLIC4*	chloride intracellular channel 4	0.4499	.000446
202303_x_at	*SMARCA5*	SWI/SNF–related, matrix-associated, actin-dependent regulator of chromatin, subfamily a, member 5	–0.4496	.000451
201493_s_at	*PUM2*	pumilio homolog 2 (*Drosophila*)	–0.4495	.000452
213610_s_at	*KLHL23*	kelch-like 23 (*Drosophila*)	–0.4494	.000454
227740_at	*UHMK1*	U2AF homology motif (UHM) kinase 1	0.4489	.000461
201376_s_at	*HNRPF*	heterogeneous nuclear ribonucleoprotein F	–0.4488	.000463
201559_s_at	*CLIC4*	chloride intracellular channel 4	0.4486	.000465
57163_at	*ELOVL1*	elongation of very long chain fatty acids (FEN1/Elo2, SUR4/Elo3, yeast)–like 1	0.4483	.00047
225563_at	*PAN3*	PABP1-dependent poly A–specific ribonuclease subunit PAN3	–0.4483	.00047
200956_s_at	*SSRP1*	structure-specific recognition protein 1	–0.4479	.000476
213329_at	*SRGAP2*	SLIT-ROBO Rho GTPase activating protein 2	0.4476	.00048
225391_at	*LOC93622*	hypothetical protein BC006130	–0.4474	.000485
217784_at	*YKT6*	SNARE protein Ykt6	0.4471	.00049
216262_s_at	*TGIF2*	TGFB-induced factor 2 (TALE family homeobox)	–0.4466	.000497
202259_s_at		phosphonoformate immuno-associated protein 5	–0.4465	.000498
205909_at	*POLE2*	polymerase (DNA-directed), epsilon 2 (p59 subunit)	–0.4465	.000499
204819_at	*FGD1*	FYVE, RhoGEF and PH domain containing 1 (faciogenital dysplasia)	0.4464	.0005
218298_s_at	*C14orf159*	chromosome 14 open reading frame 159	–0.4462	.000504
209750_at	*NR1D2*	nuclear receptor subfamily 1, group D, member 2	0.4459	.000508
201035_s_at	*HADH*	L-3-hydroxyacyl–Coenzyme A dehydrogenase, short chain	–0.4458	.000511
210904_s_at	*IL13RA1*	interleukin 13 receptor, alpha 1	0.4442	.000537
202258_s_at		phosphonoformate immuno-associated protein 5	–0.4441	.000539
226912_at	*ZDHHC23*	zinc finger, DHHC-type containing 23	–0.4438	.000544
218168_s_at	*CABC1*	chaperone, ABC1 activity of bc1 complex like (*S. pombe*)	–0.4436	.000548
228174_at	*GOLGA1*	Golgi autoantigen, golgin subfamily a, 1	–0.4428	.000562
200957_s_at	*SSRP1*	structure-specific recognition protein 1	–0.4422	.000573
232113_at		Hypothetical gene supported by BX647608	0.4421	.000576
219557_s_at	*NRIP3*	nuclear receptor interacting protein 3	0.4418	.00058
223711_s_at	*THY28*	thymocyte protein Thy28	–0.4415	.000586
212709_at	*NUP160*	nucleoporin 160 kDa	–0.4411	.000594
200679_x_at	*HMGB1*	high-mobility group box 1	–0.4409	.000598
222204_s_at	*RRN3*	RRN3 RNA polymerase I transcription factor homolog (yeast)	–0.4403	.000608
214882_s_at	*SFRS2*	splicing factor, arginine/serine–rich 2	–0.4403	.000608
208178_x_at	*TRIO*	triple functional domain (PTPRF interacting)	0.4401	.000614
213238_at	*ATP10D*	ATPase, class V, type 10D	0.4399	.000618
201450_s_at	*TIA1*	TIA1 cytotoxic granule-associated RNA binding protein	–0.4398	.000619
201712_s_at	*RANBP2*	RAN binding protein 2	–0.4395	.000624
200977_s_at	*TAX1BP1*	Tax1 (human T-cell leukemia virus type I) binding protein 1	0.4393	.000628
202948_at	*IL1R1*	interleukin 1 receptor, type I	0.4390	.000635
203317_at	*PSD4*	pleckstrin and Sec7 domain containing 4	–0.4387	.000641

Table W3. (*continued*)

Affy ID	Gene Symbol	Description	ρ	P
233364_s_at		Hypothetical gene supported by BX647608	0.4386	.000643
208795_s_at	*MCM7*	MCM7 minichromosome maintenance deficient 7 (*S. cerevisiae*)	−0.4386	.000643
203231_s_at	*ATXN1*	ataxin 1	0.4382	.000651
208747_s_at	*C1S*	complement component 1, s subcomponent	0.4380	.000656
214109_at	*LRBA*	LPS-responsive vesicle trafficking, beach and anchor containing	−0.4377	.000662
213338_at	*TMEM158*	Ras-induced senescence 1	0.4374	.000668
224940_s_at	*PAPPA*	pregnancy-associated plasma protein A, pappalysin 1	0.4374	.000669
207749_s_at	*PPP2R3A*	protein phosphatase 2 (formerly 2A), regulatory subunit B″, alpha	0.4372	.000673
203856_at	*VRK1*	vaccinia-related kinase 1	−0.4371	.000675
212603_at	*MRPS31*	mitochondrial ribosomal protein S31	−0.4368	.00068
212789_at	*NCAPD3*	KIAA0056 protein	−0.4366	.000685
201170_s_at	*BHLHB2*	basic helix-loop-helix domain containing, class B, 2	0.4364	.00069
202302_s_at	*RSRC2*	similar to splicing factor, arginine/serine–rich 4	−0.4363	.000692
212443_at	*NBEAL2*	neurobeachin-like 2	−0.4362	.000694
207949_s_at	*ICA1*	islet cell autoantigen 1, 69 kDa	−0.4360	.000699
209421_at	*MSH2*	mutS homolog 2, colon cancer, nonpolyposis type 1 (*E. coli*)	−0.4359	.0007
218443_s_at	*DAZAP1*	DAZ-associated protein 1	−0.4359	.000701
203229_s_at	*CLK2*	CDC-like kinase 2	−0.4356	.000707
213798_s_at	*CAP1*	CAP, adenylate cyclase–associated protein 1 (yeast)	0.4356	.000708
216969_s_at	*KIF22*	kinesin family member 22	−0.4355	.00071
208908_s_at	*CAST*	calpastatin	0.4353	.000714
203388_at	*ARRB2*	arrestin, beta 2	−0.4351	.00072
218714_at	*PRR14*	hypothetical protein MGC3121	−0.4342	.00074
212920_at			−0.4338	.00075
217892_s_at	*LIMA1*	epithelial protein lost in neoplasm beta	0.4337	.000751
215947_s_at		hypothetical protein FLJ14668	−0.4332	.000762
232524_x_at	*ANAPC4*	anaphase promoting complex subunit 4	−0.4330	.000767
39729_at	*PRDX2*	peroxiredoxin 2	−0.4328	.000771
241418_at		Hypothetical LOC344887	0.4326	.000778
200625_s_at	*CAP1*	CAP, adenylate cyclase–associated protein 1 (yeast)	0.4324	.000782
203729_at	*EMP3*	epithelial membrane protein 3	0.4323	.000786
227846_at	*GPR*	putative G protein coupled receptor	0.4322	.000787
231271_x_at	*HSCARG*	HSCARG protein	−0.4315	.000804
226405_s_at	*ARRDC1*	arrestin domain containing 1	−0.4315	.000805
223154_at	*MRPL1*	mitochondrial ribosomal protein L1	−0.4313	.000811
211506_s_at	*IL8*	interleukin 8	0.4311	.000815
226281_at	*DNER*	delta-notch–like EGF repeat-containing transmembrane	0.4310	.000816
221191_at		DKFZp434A0131 protein	−0.4308	.000821
211376_s_at	*NSMCE4A*	chromosome 10 open reading frame 86	−0.4308	.000821
205308_at	*C8orf70*	chromosome 8 open reading frame 70	0.4306	.000826
206237_s_at	*NRG1*	neuregulin 1	0.4306	.000827
242673_at			−0.4303	.000835
212056_at	*KIAA0182*	KIAA0182 protein	−0.4296	.000854
214744_s_at	*RPL23*	ribosomal protein L23	−0.4294	.000858
214446_at	*ELL2*	elongation factor, RNA polymerase II, 2	0.4291	.000867
212267_at	*WAPAL*	KIAA0261	−0.4290	.000868
211057_at	*ROR1*	receptor tyrosine kinase–like orphan receptor 1	0.4290	.00087
214927_at	*ITGBL1*	integrin, beta-like 1 (with EGF-like repeat domains)	0.4290	.00087
233167_at	*SELO*	selenoprotein O	−0.4288	.000876
210095_s_at	*IGFBP3*	insulin-like growth factor binding protein 3	0.4286	.000879
204160_s_at	*ENPP4*	ectonucleotide pyrophosphatase/phosphodiesterase 4 (putative function)	−0.4285	.000882
202013_s_at	*EXT2*	exostoses (multiple) 2	0.4281	.000893
224886_at	*LOC339123*	hypothetical LOC339123	−0.4279	.000899
224942_at	*PAPPA*	pregnancy-associated plasma protein A, pappalysin 1	0.4279	.0009
202532_s_at	*DHFR*	dihydrofolate reductase	−0.4276	.000908
202326_at	*EHMT2*	euchromatic histone-lysine *N*-methyltransferase 2	−0.4275	.00091
204853_at	*ORC2L*	origin recognition complex, subunit 2–like (yeast)	−0.4273	.000916
218170_at	*ISOC1*	isochorismatase domain containing 1	−0.4266	.000937
232254_at	*FBXO25*	F-box protein 25	0.4260	.000954
218688_at	*DAK*	dihydroxyacetone kinase 2 homolog (yeast)	−0.4259	.000957
218028_at	*ELOVL1*	elongation of very long chain fatty acids (FEN1/Elo2, SUR4/Elo3, yeast)–like 1	0.4258	.000959
212723_at	*JMJD6*	phosphatidylserine receptor	0.4257	.000962
209127_s_at	*SART3*	squamous cell carcinoma antigen recognized by T cells 3	−0.4257	.000962
202162_s_at	*CNOT8*	CCR4-NOT transcription complex, subunit 8	−0.4256	.000964
227129_x_at	*LOC284701*	hypothetical protein LOC284701	−0.4253	.000974
209143_s_at	*CLNS1A*; *CLDND1*	chloride channel, nucleotide-sensitive, 1A; chromosome 3 open reading frame 4	−0.4253	.000975
214075_at	*NENF*	neuron-derived neurotrophic factor	0.4249	.000985
222900_at	*NRIP3*	nuclear receptor interacting protein 3	0.4248	.00099
213793_s_at	*HOMER1*	homer homolog 1 (*Drosophila*)	−0.4246	.000994

Table W4. Gene Ontology Categories Associated with Genes Whose Expression Profiles Correlated with NH ($P < .01$) Across the NCI-60 Panel of Cancer Cell Lines.

Categories Associated with Expression of Genes Negatively Correlated with NH ($P < .01$)				
Category	GO Term	Count	%	*P*
GOTERM_BP_ALL	cellular metabolism	275	49.73	4.80e–08
GOTERM_BP_ALL	macromolecule metabolism	197	35.62	8.32e–08
GOTERM_BP_ALL	nucleobase, nucleoside, nucleotide and nucleic acid metabolism	185	33.45	2.96e–19
GOTERM_BP_ALL	biopolymier metabolism	149	26.94	4.01e–09
GOTERM_BP_ALL	regulation of biological process	131	23.69	.0104873
GOTERM_BP_ALL	regulation of cellular process	128	23.15	.0027890
GOTERM_BP_ALL	regulation of physiological process	128	23.15	5.97e–04
GOTERM_BP_ALL	regulation of cellular physiological process	126	22.78	5.81e–04
GOTERM_BP_ALL	regulation of metabolism	96	17.36	1.95e–04
GOTERM_BP_ALL	regulation of cellular metabolism	95	17.18	6.59e–05
GOTERM_BP_ALL	transcription	92	16.64	5.53e–05
GOTERM_BP_ALL	regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism	90	16.27	5.56e–05
GOTERM_BP_ALL	transcription, DNA-dependent	87	15.73	3.80e–05
GOTERM_BP_ALL	regulation of transcription	86	15.55	1.68e–04
GOTERM_BP_ALL	regulation of transcription, DNA-dependent	83	15.01	7.26e–05
GOTERM_BP_ALL	DNA metabolism	58	10.49	1.87e–10
GOTERM_BP_ALL	RNA metabolism	50	9.04	1.14e–08
GOTERM_BP_ALL	cell cycle	48	8.68	6.20e–04
GOTERM_BP_ALL	RNA processing	47	8.50	2.27e–10
GOTERM_BP_ALL	mRNA metabolism	36	6.51	4.95e–10
GOTERM_BP_ALL	mRNA processing	31	5.61	1.13e–08
GOTERM_BP_ALL	DNA replication	26	4.70	1.46e–07
GOTERM_BP_ALL	RNA splicing	25	4.52	4.15e–07
GOTERM_BP_ALL	response to endogenous stimulus	24	4.34	6.50e–04
GOTERM_BP_ALL	response to DNA damage stimulus	24	4.34	4.14e–04
GOTERM_BP_ALL	DNA repair	23	4.16	1.82e–04
GOTERM_BP_ALL	chromosome organization and biogenesis	22	3.98	.0013376
GOTERM_BP_ALL	chromosome organization and biogenesis (*sensu* Eukaryota)	20	3.62	.0023155
GOTERM_BP_ALL	nuclear mRNA splicing, through spliceosome	20	3.62	1.71e–05
GOTERM_BP_ALL	RNA splicing, through transesterification reactions with bulged adenosine as nucleophile	20	3.62	1.71e–05
GOTERM_BP_ALL	RNA splicing, through transesterification reactions	20	3.62	1.71e–05
GOTERM_BP_ALL	DNA packaging	18	3.25	.0017667
GOTERM_BP_ALL	protein complex assembly	16	2.89	.0468127
GOTERM_BP_ALL	M phase	16	2.89	.0229671
GOTERM_BP_ALL	establishment and/or maintenance of chromatin architecture	16	2.89	.0075817
GOTERM_BP_ALL	DNA-dependent DNA replication	15	2.71	3.74e–05
GOTERM_BP_ALL	chromatin assembly or disassembly	9	1.63	.0425323
GOTERM_BP_ALL	ribosome biogenesis and assembly	9	1.63	.0425323
GOTERM_BP_ALL	DNA recombination	8	1.45	.0060408
GOTERM_BP_ALL	transcription initiation	7	1.27	.0226955
GOTERM_BP_ALL	DNA replication initiation	7	1.27	9.09e–04
GOTERM_BP_ALL	M phase of meiotic cell cycle	5	0.90	.0438018
GOTERM_BP_ALL	mRNA catabolism	5	0.90	.0308043
GOTERM_BP_ALL	DNA unwinding during replication	4	0.72	.0090748
GOTERM_CC_ALL	intracellular	304	54.97	2.02e–12
GOTERM_CC_ALL	organelle	282	50.99	2.02e–17
GOTERM_CC_ALL	intracellular organelle	282	50.99	1.90e–17
GOTERM_CC_ALL	membrane-bound organelle	259	46.84	2.70e–16
GOTERM_CC_ALL	intracellular membrane–bound organelle	259	46.84	2.61e–16
GOTERM_CC_ALL	nucleus	202	36.53	2.54e–19
GOTERM_CC_ALL	protein complex	93	16.82	9.34e–05
GOTERM_CC_ALL	intracellular non–membrane-bound organelle	68	12.30	.0014360
GOTERM_CC_ALL	non–membrane-bound organelle	68	12.30	.0014360
GOTERM_CC_ALL	mitochondrion	39	7.05	.0217935
GOTERM_CC_ALL	ribonucleoprotein complex	35	6.33	1.77e–05
GOTERM_CC_ALL	membrane-enclosed lumen	31	5.61	.0266555
GOTERM_CC_ALL	organelle lumen	31	5.61	.0266555
GOTERM_CC_ALL	chromosome	31	5.61	3.74e–08
GOTERM_CC_ALL	nuclear lumen	27	4.88	.005185
GOTERM_CC_ALL	nucleoplasm	18	3.25	.0267792
GOTERM_CC_ALL	ribosome	14	2.53	.0304751
GOTERM_CC_ALL	chromatin	14	2.53	3.90e–04
GOTERM_CC_ALL	nuclear chromosome	7	1.27	.0300376
GOTERM_CC_ALL	condensed chromosome	5	0.90	.0186952
GOTERM_CC_ALL	synaptic vesicle	4	0.72	.0413124
GOTERM_CC_ALL	obsolete cellular component	4	0.72	.0413124
GOTERM_CC_ALL	delta-DNA polymerase cofactor complex	3	0.54	.0243645
GOTERM_MF_ALL	nucleic acid binding	153	27.67	8.89e–11
GOTERM_MF_ALL	nucleotide binding	92	16.64	.0023895
GOTERM_MF_ALL	DNA binding	92	16.64	7.93e–06
GOTERM_MF_ALL	transcription regulator activity	55	9.95	.0049911
GOTERM_MF_ALL	RNA binding	55	9.95	6.23e–09

Table W4. (*continued*)

Categories Associated with Expression of Genes Negatively Correlated with NH ($P < .01$)

Category	GO Term	Count	%	*P*
GOTERM_MF_ALL	ATPase activity	20	3.62	.0098998
GOTERM_MF_ALL	transcriptional activator activity	17	3.07	.0134395
GOTERM_MF_ALL	ATPase activity, coupled	16	2.89	.0496740
GOTERM_MF_ALL	structural constituent of ribosome	15	2.71	.0339051
GOTERM_MF_ALL	mRNA binding	8	1.45	1.73e–04
GOTERM_MF_ALL	acetyltransferase activity	7	1.27	.0203796
GOTERM_MF_ALL	chromatin binding	7	1.27	.0149992
GOTERM_MF_ALL	DNA-dependent ATPase activity	7	1.27	.0030113
GOTERM_MF_ALL	double-stranded DNA binding	5	0.90	.0259470
BIOCARTA	h_prc2Pathway: The PRC2 Complex Sets Long-term Gene Silencing Through Modification of Histone Tails	4	0.72	.0301009

Categories Associated with Expression of Genes Positively Correlated with NH ($P < .01$)

Category	Term	Count	%	*P*
GOTERM_BP_ALL	cell communication	136	23.01	3.55e–05
GOTERM_BP_ALL	signal transduction	127	21.49	4.00e–05
GOTERM_BP_ALL	localization	117	19.80	3.79e–04
GOTERM_BP_ALL	establishment of localization	117	19.80	3.08e–04
GOTERM_BP_ALL	development	97	16.41	1.75e–09
GOTERM_BP_ALL	cell organization and biogenesis	73	12.35	1.07e–05
GOTERM_BP_ALL	cell adhesion	68	11.51	1.39e–19
GOTERM_BP_ALL	intracellular signaling cascade	52	8.80	5.60e–04
GOTERM_BP_ALL	negative regulation of biological process	46	7.78	3.08e–06
GOTERM_BP_ALL	negative regulation of cellular process	45	7.61	1.03e–06
GOTERM_BP_ALL	phosphorus metabolism	42	7.11	.0049447
GOTERM_BP_ALL	phosphate metabolism	42	7.11	.0049447
GOTERM_BP_ALL	negative regulation of physiological process	41	6.94	6.83e–06
GOTERM_BP_ALL	negative regulation of cellular physiological process	40	6.77	7.04e–06
GOTERM_BP_ALL	morphogenesis	39	6.60	3.21e–06
GOTERM_BP_ALL	positive regulation of biological process	35	5.92	2.68e–04
GOTERM_BP_ALL	death	32	5.41	6.70e–04
GOTERM_BP_ALL	organ development	32	5.41	1.34e–04
GOTERM_BP_ALL	cell death	31	5.25	.0012428
GOTERM_BP_ALL	positive regulation of cellular process	31	5.25	3.58e–04
GOTERM_BP_ALL	cellular localization	30	5.08	.0056816
GOTERM_BP_ALL	programmed cell death	30	5.08	.0014543
GOTERM_BP_ALL	apoptosis	30	5.08	.0013914
GOTERM_BP_ALL	establishment of cellular localization	29	4.91	.0092904
GOTERM_BP_ALL	intracellular transport	29	4.91	.0078798
GOTERM_BP_ALL	response to external stimulus	29	4.91	8.90e–04
GOTERM_BP_ALL	cell proliferation	28	4.74	.0032481
GOTERM_BP_ALL	locomotion	28	4.74	8.90e–10
GOTERM_BP_ALL	localization of cell	28	4.74	8.90e–10
GOTERM_BP_ALL	cell motility	28	4.74	8.90e–10
GOTERM_BP_ALL	cell-cell adhesion	28	4.74	2.52e–10
GOTERM_BP_ALL	cell differentiation	26	4.40	.0054730
GOTERM_BP_ALL	cellular morphogenesis	25	4.23	1.16e–06
GOTERM_BP_ALL	homophilic cell adhesion	25	4.23	7.28e–13
GOTERM_BP_ALL	system development	22	3.72	.0328183
GOTERM_BP_ALL	nervous system development	22	3.72	.0306327
GOTERM_BP_ALL	regulation of cell cycle	21	3.55	.0464284
GOTERM_BP_ALL	regulation of progression through cell cycle	21	3.55	.0456837
GOTERM_BP_ALL	response to abiotic stimulus	21	3.55	.0336106
GOTERM_BP_ALL	vesicle-mediated transport	21	3.55	.0040294
GOTERM_BP_ALL	protein kinase cascade	21	3.55	1.70e–04
GOTERM_BP_ALL	cytoskeleton organization and biogenesis	20	3.38	.0207670
GOTERM_BP_ALL	response to wounding	20	3.38	.0189762
GOTERM_BP_ALL	regulation of programmed cell death	20	3.38	.0066011
GOTERM_BP_ALL	regulation of apoptosis	20	3.38	.0062443
GOTERM_BP_ALL	regulation of cell proliferation	20	3.38	9.79e–04
GOTERM_BP_ALL	regulation of signal transduction	20	3.38	2.33e–05
GOTERM_BP_ALL	response to chemical stimulus	19	3.21	.0315919
GOTERM_BP_ALL	cell migration	16	2.71	2.07e–07
GOTERM_BP_ALL	small GTPase-mediated signal transduction	15	2.54	.0288176
GOTERM_BP_ALL	growth	15	2.54	.0013231
GOTERM_BP_ALL	cell growth	14	2.37	5.31e–04
GOTERM_BP_ALL	regulation of cell size	14	2.37	5.31e–04
GOTERM_BP_ALL	organ morphogenesis	13	2.20	.0215043
GOTERM_BP_ALL	enzyme-linked receptor protein signaling pathway	13	2.20	.0070186
GOTERM_BP_ALL	I-κB kinase/NF-κB cascade	13	2.20	1.12e–04
GOTERM_BP_ALL	homeostasis	12	2.03	.0261415

Table W4. (*continued*)

Categories Associated with Expression of Genes Positively Correlated with NH ($P < .01$)

Category	Term	Count	%	*P*
GOTERM_BP_ALL	regulation of I-κB kinase/NF-κB cascade	12	2.03	2.82e–05
GOTERM_BP_ALL	secretory pathway	11	1.86	.0439977
GOTERM_BP_ALL	actin filament–based process	11	1.86	.0229747
GOTERM_BP_ALL	inorganic anion transport	11	1.86	.0206056
GOTERM_BP_ALL	actin cytoskeleton organization and biogenesis	11	1.86	.0151636
GOTERM_BP_ALL	negative regulation of progression through cell cycle	11	1.86	.0145675
GOTERM_BP_ALL	negative regulation of cell proliferation	11	1.86	.0123569
GOTERM_BP_ALL	dephosphorylation	11	1.86	.0104091
GOTERM_BP_ALL	negative regulation of programmed cell death	11	1.86	.0079367
GOTERM_BP_ALL	negative regulation of apoptosis	11	1.86	.0075730
GOTERM_BP_ALL	protein amino acid dephosphorylation	11	1.86	.0072224
GOTERM_BP_ALL	regulation of growth	11	1.86	.0041368
GOTERM_BP_ALL	antiapoptosis	11	1.86	.0027945
GOTERM_BP_ALL	positive regulation of signal transduction	11	1.86	4.31e–04
GOTERM_BP_ALL	cell homeostasis	10	1.69	.0347712
GOTERM_BP_ALL	positive regulation of cell proliferation	10	1.69	.0154140
GOTERM_BP_ALL	regulation of cell growth	10	1.69	.0075542
GOTERM_BP_ALL	wound healing	10	1.69	.0023127
GOTERM_BP_ALL	positive regulation of I-κB kinase/NF-κB cascade	10	1.69	4.02e–04
GOTERM_BP_ALL	skeletal development	9	1.52	.0398567
GOTERM_BP_ALL	transmembrane receptor protein tyrosine kinase signaling pathway	9	1.52	.0370367
GOTERM_BP_ALL	locomotory behavior	9	1.52	.0318068
GOTERM_BP_ALL	chemotaxis	9	1.52	.0260108
GOTERM_BP_ALL	taxis	9	1.52	.0260108
GOTERM_BP_ALL	cell development	9	1.52	.0191954
GOTERM_BP_ALL	phosphate transport	9	1.52	.0095473
GOTERM_BP_ALL	neurogenesis	9	1.52	.0095473
GOTERM_BP_ALL	muscle development	9	1.52	.0076198
GOTERM_BP_ALL	regulation of body fluids	8	1.35	.0342186
GOTERM_BP_ALL	neuron differentiation	8	1.35	.0227081
GOTERM_BP_ALL	hemostasis	8	1.35	.0205718
GOTERM_BP_ALL	Golgi vesicle transport	8	1.35	.0195580
GOTERM_BP_ALL	coagulation	8	1.35	.0167270
GOTERM_BP_ALL	blood coagulation	8	1.35	.0150089
GOTERM_BP_ALL	regulation of development	8	1.35	.0082556
GOTERM_BP_ALL	cell cycle arrest	8	1.35	.0029397
GOTERM_BP_ALL	ER to Golgi vesicle–mediated transport	8	1.35	.0022683
GOTERM_BP_ALL	neuron development	7	1.18	.0309592
GOTERM_BP_ALL	integrin-mediated signaling pathway	7	1.18	.0167194
GOTERM_BP_ALL	blood vessel morphogenesis	7	1.18	.0157055
GOTERM_BP_ALL	blood vessel development	7	1.18	.0157055
GOTERM_BP_ALL	vasculature development	7	1.18	.0157055
GOTERM_BP_ALL	angiogenesis	7	1.18	.0129206
GOTERM_BP_ALL	aminoglycan metabolism	5	0.85	.0275782
GOTERM_BP_ALL	glycosaminoglycan metabolism	5	0.85	.0254349
GOTERM_BP_ALL	regulation of locomotion	5	0.85	.0094927
GOTERM_BP_ALL	regulation of cell motility	5	0.85	.0094927
GOTERM_BP_ALL	regulation of cell migration	5	0.85	.0094927
GOTERM_BP_ALL	positive regulation of development	5	0.85	.0064945
GOTERM_BP_ALL	regulation of axonogenesis	3	0.51	.0367302
GOTERM_BP_ALL	regulation of Ras protein signal transduction	3	0.51	.0367302
GOTERM_BP_ALL	regulation of epithelial cell proliferation	3	0.51	.0109926
GOTERM_BP_ALL	epithelial cell proliferation	3	0.51	.0109926
GOTERM_BP_ALL	copper ion homeostasis	3	0.51	.0074668
GOTERM_BP_ALL	positive regulation of epithelial cell proliferation	3	0.51	.0074668
GOTERM_CC_ALL	cytoplasm	128	21.66	.0153518
GOTERM_CC_ALL	plasma membrane	79	13.37	9.37e–05
GOTERM_CC_ALL	extracellular region	70	11.84	6.77e–07
GOTERM_CC_ALL	cytoskeleton	58	9.81	6.81e–08
GOTERM_CC_ALL	intrinsic to plasma membrane	49	8.29	.0269904
GOTERM_CC_ALL	integral to plasma membrane	49	8.29	.0246202
GOTERM_CC_ALL	extracellular matrix	35	5.92	9.43e–11
GOTERM_CC_ALL	extracellular matrix (*sensu* Metazoa)	35	5.92	5.21e–11
GOTERM_CC_ALL	endoplasmic reticulum	33	5.58	1.66e–04
GOTERM_CC_ALL	Golgi apparatus	24	4.06	.0121383
GOTERM_CC_ALL	extracellular space	24	4.06	.0076828
GOTERM_CC_ALL	actin cytoskeleton	18	3.05	3.01e–04
GOTERM_CC_ALL	microtubule cytoskeleton	17	2.88	.0054881
GOTERM_CC_ALL	microtubule	13	2.20	.0076041
GOTERM_CC_ALL	vacuole	10	1.69	.0220081
GOTERM_CC_ALL	basement membrane	10	1.69	4.10e–05
GOTERM_CC_ALL	collagen	8	1.35	1.15e–04

Table W4. (*continued*)

Categories Associated with Expression of Genes Positively Correlated with NH ($P < .01$)

Category	Term	Count	%	*P*
GOTERM_CC_ALL	extrinsic to membrane	6	1.02	.0104410
GOTERM_CC_ALL	basal lamina	5	0.85	.0106105
GOTERM_CC_ALL	heterotrimeric G-protein complex	4	0.68	.0425953
GOTERM_CC_ALL	laminin complex	3	0.51	.0265177
GOTERM_CC_ALL	anchoring collagen	3	0.51	.0265177
GOTERM_CC_ALL	sheet-forming collagen	3	0.51	.0160747
GOTERM_CC_ALL	laminin-1	3	0.51	.0160747
GOTERM_MF_ALL	cation binding	121	20.47	1.18e–04
GOTERM_MF_ALL	hydrolase activity	73	12.35	.0283996
GOTERM_MF_ALL	nucleotide binding	72	12.18	.0088437
GOTERM_MF_ALL	calcium ion binding	68	11.51	1.91e–15
GOTERM_MF_ALL	purine nucleotide binding	67	11.34	.0019492
GOTERM_MF_ALL	adenyl nucleotide binding	50	8.46	.0218535
GOTERM_MF_ALL	ATP binding	47	7.95	.0382664
GOTERM_MF_ALL	structural molecule activity	43	7.28	7.94e–04
GOTERM_MF_ALL	enzyme regulator activity	40	6.77	1.49e–06
GOTERM_MF_ALL	kinase activity	38	6.43	.0151332
GOTERM_MF_ALL	phosphotransferase activity, alcohol group as acceptor	31	5.25	.0185969
GOTERM_MF_ALL	cytoskeletal protein binding	29	4.91	1.21e–07
GOTERM_MF_ALL	protein kinase activity	28	4.74	.0117125
GOTERM_MF_ALL	hydrolase activity, acting on acid anhydrides	26	4.40	.0143863
GOTERM_MF_ALL	hydrolase activity, acting on acid anhydrides, in phosphorus-containing anhydrides	26	4.40	.0119770
GOTERM_MF_ALL	pyrophosphatase activity	26	4.40	.0112973
GOTERM_MF_ALL	nucleoside-triphosphatase activity	25	4.23	.0120011
GOTERM_MF_ALL	actin binding	22	3.72	1.61e–06
GOTERM_MF_ALL	protein serine/threonine kinase activity	21	3.55	.0095445
GOTERM_MF_ALL	phosphoric ester hydrolase activity	19	3.21	.0024685
GOTERM_MF_ALL	guanyl nucleotide binding	18	3.05	.0328025
GOTERM_MF_ALL	phosphoric monoester hydrolase activity	17	2.88	6.99e–04
GOTERM_MF_ALL	GTPase regulator activity	16	2.71	.0068181
GOTERM_MF_ALL	carbohydrate binding	16	2.71	.0035107
GOTERM_MF_ALL	obsolete molecular function	15	2.54	.0188299
GOTERM_MF_ALL	magnesium ion binding	15	2.54	.0145196
GOTERM_MF_ALL	enzyme inhibitor activity	15	2.54	.0075281
GOTERM_MF_ALL	phosphoprotein phosphatase activity	14	2.37	6.78e–04
GOTERM_MF_ALL	GTPase activity	13	2.20	.0014018
GOTERM_MF_ALL	extracellular matrix structural constituent	13	2.20	1.44e–05
GOTERM_MF_ALL	small GTPase regulator activity	12	2.03	.0093352
GOTERM_MF_ALL	pattern binding	12	2.03	1.62e–04
GOTERM_MF_ALL	polysaccharide binding	12	2.03	6.74e–05
GOTERM_MF_ALL	glycosaminoglycan binding	12	2.03	4.76e–05
GOTERM_MF_ALL	cAMP-dependent protein kinase activity	11	1.86	.0325455
GOTERM_MF_ALL	cyclic nucleotide–dependent protein kinase activity	11	1.86	.0325455
GOTERM_MF_ALL	protein kinase CK2 activity	11	1.86	.0288109
GOTERM_MF_ALL	protease inhibitor activity	10	1.69	.0189656
GOTERM_MF_ALL	endopeptidase inhibitor activity	10	1.69	.0182981
GOTERM_MF_ALL	calmodulin binding	10	1.69	.0046835
GOTERM_MF_ALL	protein tyrosine phosphatase activity	10	1.69	.0010162
GOTERM_MF_ALL	heparin binding	10	1.69	1.43e–04
GOTERM_MF_ALL	phospholipid binding	9	1.52	.0194455
GOTERM_MF_ALL	structural constituent of cytoskeleton	9	1.52	.0021281
GOTERM_MF_ALL	copper ion binding	7	1.18	.0135518
GOTERM_MF_ALL	prenylated protein tyrosine phosphatase activity	7	1.18	.0067538
GOTERM_MF_ALL	growth factor binding	7	1.18	.0039932
GOTERM_MF_ALL	kinase regulator activity	6	1.02	.0347643
GOTERM_MF_ALL	tubulin binding	5	0.85	.0403221
GOTERM_MF_ALL	dioxygenase activity	5	0.85	.0134921
GOTERM_MF_ALL	insulin-like growth factor binding	5	0.85	.0029651
GOTERM_MF_ALL	L-ascorbic acid binding	5	0.85	5.08e–04
GOTERM_MF_ALL	integrin binding	4	0.68	.0426353
GOTERM_MF_ALL	actin filament binding	4	0.68	.0043206
GOTERM_MF_ALL	low-density lipoprotein binding	3	0.51	.0283400
GOTERM_MF_ALL	Ras GTPase activator activity	3	0.51	.0283400
GOTERM_MF_ALL	lipoprotein receptor activity	3	0.51	.0151426
GOTERM_MF_ALL	low-density lipoprotein receptor activity	3	0.51	.0115423
GOTERM_MF_ALL	oncostatin-M receptor activity	2	0.34	.0482958
BBID	12.IL-6_type_cytok-signal-transduct	4	0.68	.0203501
BIOCARTA	h_prionPathway: Prion Pathway	4	0.68	.0103901
BIOCARTA	h_il10Pathway: IL-10 Anti-inflammatory Signaling Pathway	4	0.68	.0080196
BIOCARTA	h_reckPathway: Inhibition of Matrix Metalloproteinases	4	0.68	.0060027
BIOCARTA	h_akap13Pathway: Rho-Selective Guanine Exchange Factor AKAP13 Mediates Stress Fiber Formation	3	0.51	.0290874
COG_KOG_ONTO LOGY	Cell division and chromosome partitioning	12	2.03	.0107010