Gene Expression Signatures for Compound Response in Cancers

Ningning He and Sukjoon Yoon*

Department of Biological Sciences, Sookmyung Women's University, Seoul 140-742, Korea

Abstract

Recent trends in generating multiple, large-scale datasets provide new challenges to manipulating the relationship of different types of components, such as gene expression and drug response data. Integrative analysis of compound response and gene expression datasets generates an opportunity to capture the possible mechanism of compounds by using signature genes on diverse types of cancer cell lines. Here, we integrated datasets of compound response and gene expression profiles on NCI60 cell lines and constructed a network, revealing the relationship for 801 compounds and 341 gene probes. As examples, obtusol, which shows an exclusive sensitivity on a small number of colon cell lines, is related to a set of gene probes that have unique overexpression in colon cell lines. We also found that the SLC7A11 gene, a direct target of miR-26b, might be a key element in understanding the action of many diverse classes of anticancer compounds. We demonstrated that this network might be useful for studying the mechanisms of varied compound response on diverse cancer cell lines.

Keywords: NCI60, drug response, gene expression, integrating, network

Introduction

The interest in studying mechanisms of compounds and finding new compounds with better efficacy than known anticancer drugs has recently increased due to the decreasing number and increasing cost of newly released anticancer drugs (Xu and Cote, 2011). NCI60 project as a predictive model of drug response and gene expression was developed by the US National Cancer Institute (NCI) (Shoemaker, 2006). This panel of 60 cell lines includes nine distinct tumor types: leukemia (LN), colon

(CO), lung (non-small cell lung cancer) (NSCLC), CNS, renal (REN), melanoma (ME), ovarian (OVR), breast (BR) and prostate (PRO) (Table 1) (Scudiero *et al.*, 1988). From the Developmental Therapeutics Program (DTP) of NCI/NIH, multiple high-throughput screening data are available, including drug response data (Weinstein, 2006), gene expression data (Blower *et al.*, 2007), DNA copy number change data (Ikediobi *et al.*, 2006), protein analysis (Nishizuka *et al.*, 2003), DNA methylation (Ehrich *et al.*, 2008), functional target analysis (Lee *et al.*, 1994), and reverse phase protein array (RPPA) analysis (Shankavaram *et al.*, 2007) and so on. The NCI drug stock is made of more than 100,000 compounds, and the screening data for more than 40,000 compounds is publicly

Table 1. List of cell lines in NCI60 dataset

Cell line	Lineage	Cell line	Lineage
CCRF-CEM	Leukemia/Lymphoma	M14	Skin/Melanoma
HL-60	Leukemia/Lymphoma	MDA-MB-435	Breast
K-562	Leukemia/Lymphoma	SK-MEL-2	Skin/Melanoma
MOLT-4	Leukemia/Lymphoma	SK-MEL-28	Skin/Melanoma
RPMI-8226	Leukemia/Lymphoma	SK-MEL-5	Skin/Melanoma
SR	Leukemia/Lymphoma	UACC-257	Skin/Melanoma
A549	Lung	UACC-62	Skin/Melanoma
EKVX	Lung	IGROV-1	GYN/Ovarian
HOP-62	Lung	OVCAR-3	GYN/Ovarian
HOP-92	Lung	OVCAR-4	GYN/Ovarian
NCI-H226	Lung	OVCAR-5	GYN/Ovarian
NCI-H23	Lung	OVCAR-8	GYN/Ovarian
NCI-H322M	Lung	NCI-ADR-RES	GYN/Ovarian
NCI-H460	Lung	SK-OV-3	GYN/Ovarian
NCI-H522	Lung	786-0	Renal
COLO-205	GI/Colon	A498	Renal
HCC2998	GI/Colon	ACHN	Renal
HCT-116	GI/Colon	CAKI-1	Renal
HCT-15	GI/Colon	RXF-393	Renal
HT-29	GI/Colon	SN12C	Renal
KM12	GI/Colon	TK-10	Renal
SW620	GI/Colon	UO-31	Renal
SF-268	CNS	PC-3	GU/Prostate
SF-295	CNS	DU-145	GU/Prostate
SF-539	CNS	MCF7	Breast
SNB-19	CNS	MDA-MB-231	Breast
SNB-75	CNS	MDA-MB-468	Breast
U251	CNS	HS-578-T	Breast
LOXIMVI	Skin/Melanoma	BT-549	Breast
MALME-3M	Skin/Melanoma	T-47D	Breast

A total of 60 diverse cell lines were used for compound screening and DNA microarray experiment.

*Corresponding author: E-mail yoonsj@sookmyung.ac.kr

Tel +82-2-710-9415, Fax +82-2-2077-7322

Received 19 November 2011, Revised 4 December 2011,

Accepted 8 December 2011

available (<http://dtp.cancer.gov/>). The cell line response to the compounds is quantified by GI50 value, the concentration required to inhibit the growth of an exposed cell line culture to 50% relative to the untreated control.

Recently, the construction of biophysical network models of numerous components has contributed to our understanding of the relationship between different types of biological responses (Pe'er and Hachohen, 2011). A critical question remaining is how to interpret the network and take out the new biological insights or knowledge. Our approach to this question is to calculate the simple Pearson correlation coefficient (PCC) (Liu *et al.*, 2007) (Jobson *et al.*, 2007) of drug response and gene expression profile over diverse NCI60 cell lines. It is possible to obtain a certain number of significant correlations between gene expression and compound response on subsets of cancer samples. In this research, our goal is to create a global network including all known compound data and gene expression profiles, and to subtract specific networks of genes and compounds revealing unique mechanisms of compound action.

Methods

Acquisition and filtration of compound response data

Data of compound response (GI50) on NCI60 cell lines was downloaded from the NCI Developmental Therapeutics Program (DTP) site (www.dtp.nci.nih.gov). The GI50 data released in August 2004 was used in this study. They provided more than 40,000 compounds with a GI50 value to characterize compound sensitivity to diverse NCI60 cancer cell lines. GI50 is defined as the concentration of a compound required to inhibit cell growth by 50% after exposed to the compound for 48 hours in comparison with untreated cells. GI50 values were further transformed to $-\log_{10}GI50$ to linearly profile the compounds' sensitivity. The 2D structures and chemical names of compounds were also obtained through the DTP website of NCI.

Compound data were further filtered based upon several criteria; the number of missing data on 60 cell lines < 5 , the standard deviation (SD) of compound response > 0.1 , and $-\log_{10}GI50 > 5$ in at least one cell line. In order to remove compounds with many missing data (eg, GI50 value), we chose compounds with missing data on only less than 5 cell lines. We also selected compounds with relatively varied response on diverse cell lines by removing compounds with consistent response on all 60 cell lines (ie, $SD < 0.1$) (Deeken *et al.*, 2009). Furthermore, compounds with no significant potency in any cell

lines were removed. Thus, we selected compounds with $-\log_{10}GI50 > 5$ (i.e., $GI50 < 10 \mu M$) on at least one cell line.

Analysis of DNA microarray data

NCI60 gene expression data was also downloaded from the DTP website. NCI60 gene expression data generated using Affymetrix u133A/B microarrays and processed using the MAS5 normalization algorithm (Zhou and Rocke, 2005). Data for a total of 44,928 gene probes screened on 59 cell lines (data missing on NCI-H23 cell line) were retrieved. Gene probes with a standard deviation of > 2 across all 59 cell lines and a fold-change of > 4 over the median value on at least one cell line were selected for the further analysis.

Constructing network and data visualization

Pearson's correlation coefficient (PCC) and its p-value between compounds and genes were calculated using drug response and gene expression data on 59 NCI60 cell lines. The criteria of $PCC > 0.7$ and $p\text{-value} < 0.01$ were applied to select compound-gene relationship for further analysis. Finally, the network was constructed for links between compounds and genes using Cytoscape (<http://www.cytoscape.org/>) (Merico *et al.*, 2009). The force-layout algorithm was applied to optimize the topology of network (Garcia *et al.*, 2007).

The 2D structures and annotation of compounds were displayed by MarvinSketch developed by ChemAxon (<http://www.chemaxon.com/>). The structural similarity among compounds was calculated by Pipeline Pilot (Chu *et al.*, 2009). Hierarchical clustering of compound and gene expression data was carried out using complete linkage algorithm (Williams *et al.*, 1976) (<http://bonsai.hgc.jp/~mdehoon/software/cluster/>). Tree Viewer was used to visualize the clustered data (<http://www.treeview.net>). All figures were edited using Adobe Illustrator CS3 (<http://www.adobe.com>).

Results and Discussion

Clustering analysis of compound response data

Among 48,151 compounds, a total of 6,652 compounds satisfied the filtration criteria (see methods section for details). They showed a significant response ($-\log_{10}GI50 < 5$) on at least one cell lines and demonstrated varied response on 60 cell lines with missing data on less than 5 cell lines. We carried out a clustering analysis of GI50 data for these 6,652 compounds over 60 cell lines (Fig. 1A). Three major patterns were observed in the com-

pound classification: sensitive or resistant responses on most cell lines, and some compounds exhibiting both sensitive and resistant responses in a cell line-dependent manner. Many of them showed a difference of over >100-fold in GI₅₀ values among cell lines, which implied a cell line-specific mechanism of the drug action. For example, compounds in the bottom area of the heatmap in Fig. 1A showed resistant response ($-\log\text{GI}_{50} \sim -4$) on most cell lines (green colors), while a few cell lines gave $-\log\text{GI}_{50} > 5$ (red colors).

A total of 9 different lineage groups are represented by NCI60 cell lines. Leukemia cell lines tended to have relatively sensitive response to the majority of cell lines, although they were clustered in two separate groups. They were clustered together with some lung cell lines. All of CNS cell lines were clustered together in a large group, and they were resistant to more numbers of

compounds than other lineage groups. Melanoma groups were shown to distribute in separate clusters from CNS cell lines. Overall, cell lines did not show lineage-specific clustering patterns of drug responses; however, we could find some unique clustering patterns in lineage groups, such as Leukemia, CNS and Melanoma cells. Among 6,652 compounds with varied responses on NCI60 cell lines, we further selected 801 compounds that exhibited a significant correlation with gene probes in the DNA microarray data on NCI60 cell lines (Fig. 1B). Their GI₅₀ profile showed positive or negative correlations with the expression profile of at least one gene probe in the DNA microarray data. We focused on these compounds for the network-based analysis of compound and gene probe pairs.

Fig. 1. $-\log(\text{GI}_{50})$ profile of diverse compounds on NCI60 cell lines. (A) A total of 6,652 compounds were selected from 48,151 NCI60 compounds and used for $-\log(\text{GI}_{50})$ profiling across 59 cell lines. (B) A total of 801 compounds were further selected from 6,652 compounds based on the significant correlation with gene expression signatures in Fig. 2. See Methods section for details of compound filtration. Different cancer types are displayed in different colors.

Fig. 2. Gene expression profile on NCI60 cell lines. (A) The expression ($\log(\text{fold change})$) profile of a total of 5,653 gene probes were selected from 44,928 probes and clustered against 59 NCI60 cell lines. (B) The expression profile of 341 gene probes exhibiting a significant correlation with compound response data on the same NCI60 cell lines. Nine tissue types of 60 cancer cell lines were represented in different colors.

Clustering analysis of gene expression data

In the analysis of gene expression profile from DNA microarray data, lineage-oriented clustering patterns were more obvious than in the compound response profile (Fig. 2). We used a total of 5,635 filtered gene probes for this analysis (See Methods section for details of filtration). Among the 9 different lineages of NCI60 cell lines, Leukemia, CNS and colon cell lines were completely clustered together on a lineage basis (Fig. 2A). Melanoma and Renal groups were also well-clustered with only one exception. Most lung cell lines were included in a single cluster. This observation was roughly consistent with clustering patterns in the compound response. Relatively homogeneous characteristics in cell lines within a lineage group (leukemia, CNS and melanoma) were observed through gene expression analysis, which explains the similar compound responses among cell lines of common lineages. However, cell lines of breast and ovary did not show a lineage-oriented clustering pattern, indicating that they are more heterogeneous than those in other lineage groups. Thus, the analysis of compound responses on these lineages require further sub-clustering of cell lines to find improved correlations between compound response and gene expression profiles.

We could also identify unique gene signatures for several lineage groups. A large number of up-regulated gene signatures were observed in melanoma cell lines. In the cases of leukemia and colon cell lines, significant gene signatures were found with both up- and down-regulated patterns. CNS also exhibited a small lineage-specific gene signature, while renal cell lines did not show any signature groups with a significant expression pattern. Among these 5,635 gene probes, we further selected 341 gene probes with a significant correlation with the GI50 profile of a compound in Fig. 1 (Fig. 2B). They were used for the following network-based analysis.

In conclusion, gene expression profiles reveal lineage-specific clustering patterns, indicating that gene signatures effectively represent the distinct biological features in the lineage-based analysis of cell lines. In the other hand, the compound response data shows relatively complicated patterns in the lineage clustering, indicating that the compound response is the outcome of complicated mechanisms beyond lineage-specific features. Thus, combining the GI50 with gene expression data on NCI60 panel might provide a new tool for understanding lineage-specific versus non-lineage-specific mechanisms in drug action.

Network analysis of the relationship between compound response and gene expression

A large-scale correlation study between compound response and gene expression was carried out using NCI60 cell lines. The relationship was analyzed on a network which represents similarity of GI50s of a compound and fold change of a gene probe over 59 cell lines (Fig. 3). Totals of 801 compounds and 341 gene probes had at least one counterpart with a significantly high correlation ($PCC > 0.7$ & $p\text{-value} < 0.01$). Most of the selected gene probes had more than one connected compound that mediated inter-connection among gene probes, resulting in a globular shape in the overall network topology. However, a small number of compounds and gene probes showed an isolated network containing only one or a few number of compound-gene probe pairs. The data provide a useful resource for studying and predicting the mechanism of drug actions on diverse cancer cells.

To further explore this network, we picked up two sub-networks to examine the correlations between the involved components. Firstly, as shown in the Fig. 4, compound of NSC710203 (Obtusol) shows a close relationship with 207463_x_at, 204818_at, and 3421_x_at. Obtusol is known to have an anti-bacterial activity (Vai-

Fig. 3. Correlation network of compounds and gene probes. The correlation between compound response and gene expression was calculated using $-\log(\text{GI50})$ and $\log(\text{fold change})$ values on NCI60 cell lines. Totals of 801 compounds and 341 genes were selected with a significant correlation ($p < 0.01$) displayed with connecting edges.

Fig. 4. Subnetwork of Obtusol (NSC710203) and correlated gene probes. Three probes for PRSS3 and HSD17B2 genes showed similar pattern of gene expression with the pattern of GI50 of Obtusol on NCI60 cell lines. (A) The structure of subnetwork is displayed for Obtusol and gene probes. (B) NCI60 profile of gene expression and compound response for the selected subnetwork.

rappan *et al.*, 2001; Vairappan, 2003). Among three related gene probes, both 207463_x_at, and 3421_x_at represent the PRSS3 gene, which can promote tumor growth and metastasis of human pancreatic cancer (Jiang *et al.*, 2010). Another gene probe 204818_at is HSD17B2, which can regulate intracellular levels of granulosa cells (GCs) (Svechnikova *et al.*, 2007). These two genes, PRSS3 and HSD17B2 are predicted as Obtusol related genes. Interestingly, Obtusol showed a relatively strong sensitivity on only a small number of colon cancer cell lines, while it had no significant sensitivity on all of other cell lines (Fig. 4B). The related genes, PRSS3 and HSD17B2, showed a unique up-regulation on most of colon cell lines. It seems that unique functions of these genes in colon cancer cell lines might be related to the mechanism of action of Obtusol. However, in order to explain the specificity of Obtusol on a narrow range of colon cell lines, it requires further sub-classification of gene signatures within the cell lines of colon cancers.

Another subnetwork includes a gene probe, 209921_at, and a total of 39 related compounds (Fig. 5). 209921_at presents SLC7A11 (solute carrier family 7 (anionic amino acid transporter light chain, xc- system), member 11), which has been known as a direct target of miR-26b. Its expression is remarkably increased in both breast cancer cell lines and clinical samples (Liu *et al.*, 2011). MiR-26b is a down-regulated gene in human breast cancer cell lines which can impair viability and triggers apoptosis of human MCF7 cells (Davoren *et al.*, 2008). Based on 2D chemical similarity, 27 out of 39

Fig. 5. Subnetwork of SLC7A11 gene and its related compounds. The responses ($-\log(\text{GI50})$) of 39 compounds on NCI60 cell lines showed a significant correlation with the expression pattern of SLC7A11 (solute carrier family 7 (anionic amino acid transporter light chain, xc- system), member 11).

compounds were clustered into 6 groups. The chemical names of cluster centers, their biological (or chemical) information and the number of compounds in each cluster were summarized in Table 2. NSC249956 (Chaplatrin) is isolated from *Liatris gracilis*, which has many anti-tumor constituents (Gravuer *et al.*, 2003). NSC34821 (2-Propen-1-one, 2-[(dimethylamino)methyl]-1-(2-nitrophenyl)-hydrochloride) is a similar chemical of alkylamino-propionophenone which is known to have a potent anti-neoplastic activity (Huang and Hall, 1996). NSC708044 (2-butenic acid, 4,4-dicyano-4-[(phenylmethylene)amino]-methyl ester, (Z,E)-(9CI)) has an effect on chemoresistance, mediated by glutathione transport system (Dai *et al.*, 2007). NSC654206 (Naphthasaritone) and its related compounds were used in the treatment and control of insulin resistance and hyperglycemia (Cote and Goodman, 1973). NSC138333 (Cyclohexanone, 2,6-bis(4-(1,3-dihydro-2H-pyrrolo[3,4-b] quinolin-2-ylcarbonyl)-2-oxotetrahydro-3-furanyl)methyl acetate), one of pyrrolo-quinoline derivative, is a potent inhibitor for PI3-kinase related kinases (Peng *et al.*, 2002). NSC38533 (Cyclohexanone, 2,6-bis(4-morpholinylmethyl)-, dihydrochloride) can be as a potential drug for the treatment of the ER-negative breast cancer MDA-MB-231 and SKBr3 human breast cancer cells (Somers-Edgar *et al.*, 2011). It is implied that SLC7A11 plays an essential role in the action of these 6 compound classes. We have demonstrated that the network-based analysis of compound response and expression patterns is a new promising approach to categorize compounds and predict the mechanism of action.

Table 2. Classification of SLC7A11 related compounds

Class representative compound	No. of compounds	Name	Annotation
NSC249956	8	Chapliatrin	Chapliatrin is isolated from <i>Liatris gracilis</i> which contains many anti-tumor constituents
NSC34821	5	2-Propen-1-one, 2-[(dimethylamino)methyl]-1-(2-nitro phenyl)-,hydrochloride	Similar structure with alkylaminopropiophenone which showed a potent antineoplastic activity
NSC708044	3	2-Butenoic acid, 4,4-dicyano-4-[(phenylmethylene)amino]-,methyl ester,(Z,E)-(9CI)	Compounds Susceptible to chemoresistance mediated by glutathione transport system
NSC654206	4	Naphthsaritone	Naphthsaritone and related compounds was used in the treatment and control of insulin resistance and hyperglycemia,
NSC138333	3	(4-(1,3-Dihydro-2H-pyrrolo[3,4-b]quinolin-2-ylcarbonyl)-2-oxotetrahydro-3-furanyl) methyl acetate	Pyrrolo-quinoline derivatives can be as potent inhibitors for PI3-kinase related kinases
NSC38535 HCl	4	Cyclohexanone, 2,6-bis(4-morpholinylmethyl)-, dihydrochloride	Potential drugs for the treatment of ER-negative breast cancer MDA-MB-231 and SKBr3 human breast cancer cells

Compounds in Fig. 5 were classified into 6 clusters based on chemical similarity (Tanimoto index, >0.8). A total of 27 out of 39 compounds belong to 6 clusters.

In conclusion, we attempted to integrate compound and gene expression data on diverse cancer cell lines. It enabled us to find many unexpected relationships between genes and compounds. Previously, most of analysis have focused on finding gene signatures (Ballif *et al.*, 2011) or compounds clusters (Takigawa *et al.*, 2011), and there are few reports on the cross analysis of compound and gene expression data on diverse cell types. In the present study, we built a network of compound-gene relationships on 59 diverse cell lines retrieved from NCI60 datasets. Patterns in subnetworks provided new clues on the possible mechanisms of compounds.

Acknowledgements

This research was supported by Sookmyung Women's

University Research Grants 1-1003-0120.

References

- Ballif, B.C., Rosenfeld, J.A., Traylor, R., Theisen, A., Bader, P.I., Ladda, R.L., Sell, S.L., Steinraths, M., Surti, U., McGuire, M., Williams, S., Farrell, S.A., Filiano, J., Schnur, R.E., Coffey, L.B., Tervo, R.C., Stroud, T., Marble, M., Netzloff, M., Hanson, K., Aylsworth, A.S., Bamforth, J.S., Babu, D., Niyazov, D.M., Ravnani, J.B., Schultz, R.A., Lamb, A.N., Torchia, B.S., Bejjani, B.A., and Shaffer, L.G. (2011). High-resolution array CGH defines critical regions and candidate genes for microcephaly, abnormalities of the corpus callosum, and seizure phenotypes in patients with microdeletions of 1q43q44. *Hum. Genet.* [Epub ahead of print].
- Blower, P.E., Verducci, J.S., Lin, S., Zhou, J., Chung, J.H.,

- Dai, Z., Liu, C.G., Reinhold, W., Lorenzi, P.L., Kaldjian, E.P., Croce, C.M., Weinstein, J.N., and Sadee, W. (2007). MicroRNA expression profiles for the NCI-60 cancer cell panel. *Mol. Cancer Ther.* 6, 1483-1491.
- Chu, C.W., Holliday, J.D., and Willett, P. (2009). Effect of data standardization on chemical clustering and similarity searching. *J. Chem. Inf. Model* 49, 155-161.
- Cote, P.N. and Goodman, L. (1973). Glucopyranosides derived from 2-hydroxy-1,4-naphthoquinones. *Carbohydr. Res.* 26, 247-251.
- Dai, Z., Huang, Y., Sadee, W., and Blower, P. (2007). Chemoinformatics analysis identifies cytotoxic compounds susceptible to chemoresistance mediated by glutathione and cystine/glutamate transport system xc. *J. Med. Chem.* 50, 1896-1906.
- Davoren, P.A., McNeill, R.E., Lowery, A.J., Kerin, M.J., and Miller, N. (2008). Identification of suitable endogenous control genes for microRNA gene expression analysis in human breast cancer. *BMC Mol. Biol.* 9, 76.
- Deeken, J.F., Robey, R.W., Shukla, S., Steadman, K., Chakraborty, A.R., Poonkuzhali, B., Schuetz, E.G., Holbeck, S., Ambudkar, S.V., and Bates, S.E. (2009). Identification of compounds that correlate with ABCG2 transporter function in the National Cancer Institute Anticancer Drug Screen. *Mol. Pharmacol.* 76, 946-956.
- Ehrich, M., Turner, J., Gibbs, P., Lipton, L., Giovanneti, M., Cantor, C., and van den Boom, D. (2008). Cytosine methylation profiling of cancer cell lines. *Proc. Natl. Acad. Sci. U.S.A.* 105, 4844-4849.
- Garcia, O., Saveanu, C., Cline, M., Fromont-Racine, M., Jacquier, A., Schwikowski, B., and Aittokallio, T. (2007). GOLORize: a Cytoscape plug-in for network visualization with Gene Ontology-based layout and coloring. *Bioinformatics* 23, 394-396.
- Gravuer, K., von Wettberg, E.J., and Schmitt, J. (2003). Dispersal biology of *Liatris scariosa* var. *novae-angliae* (Asteraceae), a rare New England grassland perennial. *Am. J. Bot.* 90, 1159-1167.
- Huang, Y. and Hall, I.H. (1996). Antineoplastic activities of alpha-methyl-beta-alkylaminopropiophenone derivatives in tissue culture cells. *Anticancer Res.* 16, 3589-3595.
- Ikedobi, O.N., Davies, H., Bignell, G., Edkins, S., Stevens, C., O'Meara, S., Santarius, T., Avis, T., Barthorpe, S., Brackenbury, L., Buck, G., Butler, A., Clements, J., Cole, J., Dicks, E., Forbes, S., Gray, K., Halliday, K., Harrison, R., Hills, K., Hinton, J., Hunter, C., Jenkinson, A., Jones, D., Kosmidou, V., Lugg, R., Menzies, A., Mironenko, T., Parker, A., Perry, J., Raine, K., Richardson, D., Shepherd, R., Small, A., Smith, R., Solomon, H., Stephens, P., Teague, J., Tofts, C., Varian, J., Webb, T., West, S., Widaa, S., Yates, A., Reinhold, W., Weinstein, J.N., Stratton, M.R., Futreal, P.A., and Wooster, R. (2006). Mutation analysis of 24 known cancer genes in the NCI-60 cell line set. *Mol. Cancer Ther.* 5, 2606-2612.
- Jiang, G., Cao, F., Ren, G., Gao, D., Bhakta, V., Zhang, Y., Cao, H., Dong, Z., Zang, W., Zhang, S., Wong, H.H., Hiley, C., Crnogorac-Jurcovic, T., Lemoine, N.R., and Wang, Y. (2010). PRSS3 promotes tumour growth and metastasis of human pancreatic cancer. *Gut* 59, 1535-1544.
- Jobson, A.G., Cardellina, J.H.2nd, Scudiero, D., Kondapaka, S., Zhang, H., Kim, H., Shoemaker, R., and Pommier, Y. (2007). Identification of a Bis-guanyldihydrazone [4,4'-Diacetyldiphenylurea-bis(guanyldihydrazone); NSC 109555] as a novel chemotype for inhibition of Chk2 kinase. *Mol. Pharmacol.* 72, 876-884.
- Lee, J.S., Paull, K., Alvarez, M., Hose, C., Monks, A., Grever, M., Fojo, A.T., and Bates, S.E. (1994). Rhodamine efflux patterns predict P-glycoprotein substrates in the National Cancer Institute drug screen. *Mol. Pharmacol.* 46, 627-638.
- Liu, W., Wu, X., Zhang, W., Montenegro, R.C., Fackenthal, D.L., Spitz, J.A., Huff, L.M., Innocenti, F., Das, S., Cook, E.H.Jr., Cox, N.J., Bates, S.E., and Ratain, M.J. (2007). Relationship of EGFR mutations, expression, amplification, and polymorphisms to epidermal growth factor receptor inhibitors in the NCI60 cell lines. *Clin. Cancer Res.* 13, 6788-6795.
- Liu, X.X., Li, X.J., Zhang, B., Liang, Y.J., Zhou, C.X., Cao, D.X., He, M., Chen, G.Q., He, J.R., and Zhao, Q. (2011). MicroRNA-26b is underexpressed in human breast cancer and induces cell apoptosis by targeting SLC7A11. *FEBS Lett.* 585, 1363-1367.
- Merico, D., Gfeller, D., and Bader, G.D. (2009). How to visually interpret biological data using networks. *Nat. Biotechnol.* 27, 921-924.
- Nishizuka, S., Charboneau, L., Young, L., Major, S., Reinhold, W.C., Waltham, M., Kouros-Mehr, H., Bussey, K.J., Lee, J.K., Espina, V., Munson, P.J., Petricoin, E.3rd., Liotta, L.A., and Weinstein, J.N. (2003). Proteomic profiling of the NCI-60 cancer cell lines using new high-density reverse-phase lysate microarrays. *Proc. Natl. Acad. Sci. U.S.A.* 100, 14229-14234.
- Pe'er, D. and Hacohen, N. (2011). Principles and strategies for developing network models in cancer. *Cell.* 144, 864-873.
- Peng, H., Kim, D.I., Sarkaria, J.N., Cho, Y.S., Abraham, R.T., and Zalkow, L.H. (2002). Novel pyrrolo-quinoline derivatives as potent inhibitors for PI3-kinase related kinases. *Bioorg. Med. Chem.* 10, 167-174.
- Scudiero, D.A., Shoemaker, R.H., Paull, K.D., Monks, A., Tierney, S., Nofziger, T.H., Currens, M.J., Seniff, D., and Boyd, M.R. (1988). Evaluation of a soluble tetrazolium/formazan assay for cell growth and drug sensitivity in culture using human and other tumor cell lines. *Cancer Res.* 48, 4827-4833.
- Shankavaram, U.T., Reinhold, W.C., Nishizuka, S., Major, S., Morita, D., Chary, K.K., Reimers, M.A., Scherf, U., Kahn, A., Dolginow, D., Cossman, J., Kaldjian, E.P., Scudiero, D.A., Petricoin, E., Liotta, L., Lee, J.K., and Weinstein, J.N. (2007). Transcript and protein expression profiles of the NCI-60 cancer cell panel: an integrative microarray study. *Mol. Cancer Ther.* 6, 820-832.
- Shoemaker, R.H. (2006). The NCI60 human tumour cell line anticancer drug screen. *Nat. Rev. Cancer* 6, 813-823.
- Somers-Edgar, T.J., Taurin, S., Larsen, L., Chandramouli, A., Nelson, M.A., and Rosengren, R.J. (2011). Mechanisms for the activity of heterocyclic cyclohexanone cur-

- cumin derivatives in estrogen receptor negative human breast cancer cell lines. *Invest. New. Drugs* 29, 87-97.
- Svechnikova, I., Svechnikov, K., and Söder, O. (2007). The influence of di-(2-ethylhexyl) phthalate on steroidogenesis by the ovarian granulosa cells of immature female rats. *J. Endocrinol.* 194, 603-609.
- Takigawa, I., Tsuda, K., and Mamitsuka, H. (2011). Mining significant substructure pairs for interpreting polypharmacology in drug-target network. *PLoS One* 6, e16999.
- Vairappan, C.S. (2003). Potent antibacterial activity of halogenated metabolites from Malaysian red algae, *Laurencia majuscula* (Rhodomelaceae, Ceramiales). *Biomol. Eng.* 20, 255-259.
- Vairappan, C.S., Daitoh, M., Suzuki, M., Abe, T., and Masuda, M. (2001). Antibacterial halogenated metabolites from the Malaysian *Laurencia* species. *Phytochemistry* 58, 291-297.
- Weinstein, J.N. (2006). Spotlight on molecular profiling: Integromic analysis of the NCI-60 cancer cell lines. *Mol. Cancer Ther.* 5, 2601-2605.
- Williams, G.W., Barton, G.M., White, A.A., and Won, H. (1976). Cluster analysis applied to symptom ratings of psychiatric patients: an evaluation of its predictive ability. *Br. J. Psychiatry* 129, 178-185.
- Xu, K. and Cote, T.R. (2011). Database identifies FDA-approved drugs with potential to be repurposed for treatment of orphan diseases. *Brief. Bioinform* 12, 341-345.
- Zhou, L. and Rocke, D.M. (2005). An expression index for Affymetrix GeneChips based on the generalized logarithm. *Bioinformatics* 21, 3983-3989.