BMB
Reports

Somatic mutation patterns and compound response in cancers

*Ningning He, Nayoung Kim & Sukjoon Yoon**
Department of Biological Sciences, Sookmyung Women's University, Seoul 140-742, Korea

The use of various cancer cell lines can recapitulate known tumor-associated mutations and genetically define cancer subsets. This approach also enables comparative surveys of associations between cancer mutations and drug responses. Here, we analyzed the effects of ~40,000 compounds on cancer cell lines that showed diverse mutation-dependent sensitivity profiles. Over 1,000 compounds exhibited unique sensitivity on cell lines with specific mutational genotypes, and these compounds were clustered into six different classes of mutation-oriented sensitivity. The present analysis provides new insights into the relationship between somatic mutations and selectivity response of chemicals, and these results should have applications related to predicting and optimizing therapeutic windows for anti-cancer agents. [BMB Reports 2013; 46(2): 97-102]

INTRODUCTION

Identifying the effects and mechanisms of known drugs provides perspectives for developing new cancer therapies. Approaches from systems biology and bioinformatics have been widely applied to discover new drug candidates with specific cellular activities and mechanisms, and these approaches have mainly focused on the lineage-based classification of cancer cell lines (1). Somatic mutations are important contributors to cancer progression and drug responses (2). A genotype-oriented analysis of compound response should thus be carried out using a wide variety of cancer cell lines. Accordingly, we used a new statistical method, termed Cell Line Enrichment Analysis (CLEA), to quantitatively analyze associations between genotype and drug sensitivity in cancer cell lines (2). Furthermore, this approach enabled us to measure the correlation between differentially expressed genes and mutational genotypes.

Anticancer compound screening of 60 cell lines by the National Cancer Institute (USA) (NCI60) was initiated in the late 1980s as a way to discover new drugs for leukemia. The NCI60 representing nine distinct tumor types (3): leukemia, colon, lung, CNS, renal, melanoma, ovarian, breast and prostate. The response data quantified the GI_{50} values for more than 40,000 chemical compounds and the results have been made available in a public database (DTP, http://dtp.cancer.gov/). The GI_{50} represents the compound concentration required to inhibit the growth of exposed cells to 50% of that of untreated control cells. The NCI60 panel provides many opportunities for identifying the pathways and mechanisms related to cancer at both the molecular and genetic levels (4,5). Specifically, the NCI60 panel of human tumor cell lines has been characterized at the molecular level. The analysis of RNA expression (DNA microarray data) provides unique transcriptional features for each cell line (6), and single nucleotide polymorphism data have provided estimates of DNA copy number variation at ~120,000 sites (7). Additional types of molecular characterization of these cell lines include microRNA expression (8), DNA mutations (9), protein analysis (10), DNA methylation (11), functional target analysis (12) and the reverse phase protein array (RPPA) analysis (13). These data have been used to discover valuable relationships between compound structure, mechanism of action, cell lineage and tumor mutations, among others.

CLEA is a valuable tool, particularly for the identification of genotype-dependent compound sensitivity. Using this analytical tool, we calculated the enrichment of each compound for each genotypic category of NCI60 cell line and then attempted to generate CLEA maps to select compounds with significant sensitivity againist one of the particular genotypes. Through hierarchical clustering analysis of GI_{50} data against various mutational genotypes, we attempted to confirm the existence of clusters of compounds that were restricted in terms of specific genotypes. The aim of this study was to systematically identify all potential compounds exhibiting specific sensitivity to a tumor genotype, and these findings should have applications for identifying potential compounds for genotype-oriented cancer therapies.

RESULTS AND DISCUSSION

The frequency of mutational genotypes in NCI60 cell lines

The NCI60 cell lines have been extensively characterized at the molecular level, and mutation information for the NCI60 cell lines is publicly available through the DTP website. A total

*Corresponding author. +82-2-710-9415; Fax: +82-2-2077-7322; E-mail: yoonsj@sookmyung.ac.kr.
http://dx.doi.org/10.5483/BMBRep.2013.46.2.226

Received 5 November 2012, Revised 13 November 2012, Accepted 14 November 2012

Keywords: Cancer mutation, Cell line enrichment analysis, Compound response, GI_{50}, NCI60

ISSN: 1976-6696 (print edition) / 1976-670X (electronic edition)
Copyright © 2012 by the The Korean Society for Biochemistry and Molecular Biology
This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

30 different mutations are annotated for these 60 cell lines, and we calculated the frequency of individual mutations in each cell line (Table 1). In summary, TP53 showed the highest mutation frequency, i.e., 44 of 60 cell lines (74.33%) harbored the TP53 mutation. CDKN2A also showed a relative high mutation frequency, as it was detected in 35 of 60 cell lines (58.33%). However, genes such as EGFR, BRCA2 and NF2 showed a low frequency of mutation (5%) in the NCI60 cell lines, and NOTCH1, HRAS, MSH6 and VHL only showed a mutation frequency of 3.33%. Mutations in FBXW7, FLT3, PDGFRA, MAP2K4, and BRCA1 and gene amplification in KRAS, AKT2, and EGFR were just observed in one cell line (1.67%). To ascertain statistical significance in the CLEA analysis (see Methods section for greater detail), genes having mutations or amplifications in more than 3 cell lines (TP53, CDKN2A, PTEN, KRAS, RAF, PIK3CA, APC, c-MYC-amp, STK11, CTNNB1, SMAD4, RB1, MLH1, NRAS, and TN_stromal) were selected for evaluating the association with compound response.

Hierarchical clustering of genotype-specific compounds

To identify patterns of genotype-specific compound responses, we selected subsets of compounds using the GI_{50} profile pattern on the CLEA map. First, the $-\log GI_{50}$ value of 5 (i.e., $GI_{50}=10$ μM) was adopted as the bipartite cutoff to determine whether a compound was sensitive (>5) or insensitive (<5) to any of the cell lines in the NCI60 panel (14). Second, the enrichment score (AUC value) in the CLEA analysis was used to select genotype-specific compounds (see Methods section for greater detail). An AUC value of 0.85 and a P value of 0.01 were used as cutoff values to ensure that compounds had a significant sensitivity for a particular genotype. We only included compounds that demonstrated strong potency (i.e., $-\log GI_{50}$ value of >5) against at least one cell line in the NCI60 panel. As a result, a total of 1,161 non-redundant compounds were compiled, satisfying the above-mentioned filters. Hierarchical clustering was carried out for these selected compounds against 15 genotypic categories (Fig. 1). We identified six major groups of compounds that showed unique sensitivity based on mutational genotype; these compound clusters were sensitive to genetic mutations in the KRAS, STK11, MLH1, CTNNB1 and BRAF genes, or sensitive to TN-stromal genotype. This result provides direct clues for understanding common mechanisms of action (MOA) between diverse compounds with similar applications.

Genotype-dependent sensitivity of compounds

From each of the six clusters shown in Fig. 1, we selected one representative compound and further analyzed its genotype-specific cellular response. Neuroblastoma RAS viral oncogene homolog (NRAS) is a member of the Ras gene family and encodes 21-kDa proteins that are members of the super family of small GTP-binding proteins. NRAS has diverse intracellular signaling functions that include control of cellular proliferation, growth, and apoptosis (15). Somatic activating mutations in RAS are present in up to 30% of all human cancers (16). We found that NSC639187 (Landomycin A) belonged to the cluster of NRAS sensitivity (Fig. 2A). Three cell lines harboring NRAS mutations demonstrated superior compound responses (i.e., GI_{50}) as compared to wild-type cell lines. Landomycin A, a natural antibiotic, is known to induce the inhibition of DNA synthesis, interference with cellular processes critical for DNA synthesis and inhibit cell cycle progression from G1/S phase to S phase (17), and we found that the cytotoxicity of this compound was, on average, >10-fold higher in NRAS-mutant cell lines.

STK11 (LKB1) encodes a serine-threonine kinase that directly phosphorylates, and activates AMPK, a central metabol-

Table 1. Frequency of diverse mutations in the NCI60 cell lines. In this study, 15 mutational genotypes with an occurrence in >3 cell lines (left column) were selected for CLEA analysis to associate mutations with compound response (GI_{50}). c-MYC-AMP, KRAS-Amp, AKT2-Amp and EGFR-Amp represent gene amplifications

Mutation type	Number of cell lines	Frequency (%)	Mutation type	Number of cell lines	Frequency (%)
TP53	44	73.33	EGFR	3	5.00
CDKN2A	35	58.33	BRCA 2	3	5.00
PTEN	17	28.33	NF2	3	5.00
KRAS	14	23.33	NOTCH1	2	3.33
RAF	11	18.33	HRAS	2	3.33
PIK3CA	10	16.67	MSH6	2	3.33
APC	7	11.67	VHL	2	3.33
c-MYC-Amp	6	10.00	FBXW7	1	1.67
STK11	6	10.00	FLT3	1	1.67
CTNNB1	5	8.33	PDGFRA	1	1.67
SMAD4	5	8.33	MAP2K4	1	1.67
RB1	5	8.33	KRAS-Amp	1	1.67
MLH1	4	6.67	AKT2-Amp	1	1.67
NRAS	4	6.67	BRCA1	1	1.67
TN_stromal	4	6.67	EGFR-Amp	1	1.67

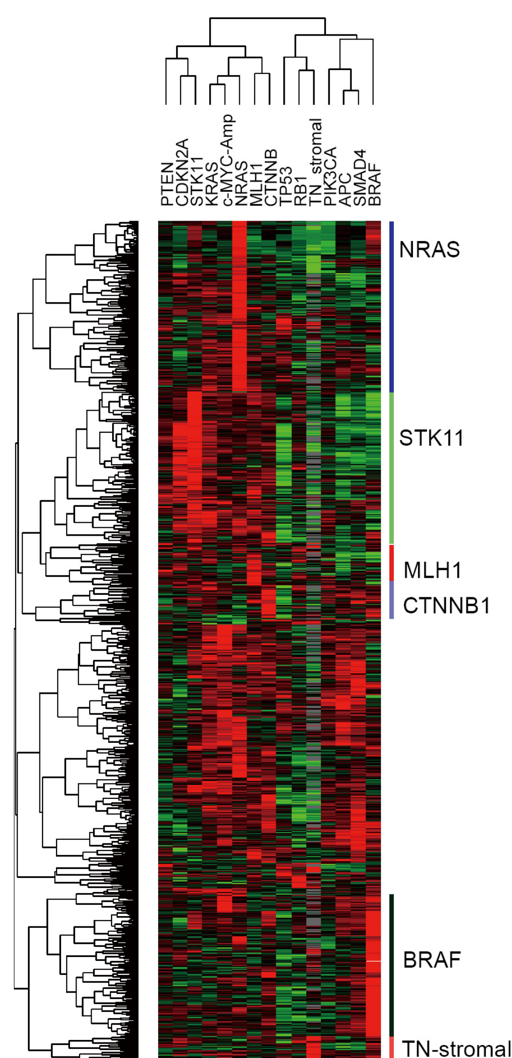

Fig. 1. Hierarchical clustering of the 1,161 compounds. The significance level (P value) of the enrichment score (AUC value) was used against the 15 genotypic categories for clustering. Six major groups of compounds with unique genotype-specific cellular responses, are indicated by the corresponding mutated genes shown on the left. Red color represents sensitive responses (low GI_{50}, AUC > 50) to the genotype, while the green represents resistance responses (high GI_{50}, AUC < 50).

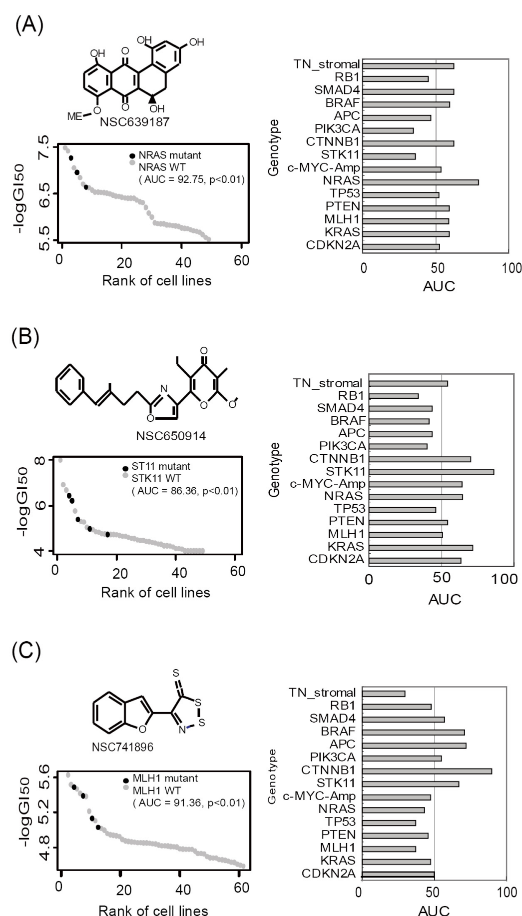

Fig. 2. Chemical structure and genotype-specific cellular response of various compounds. (A) An NRAS mutation-specific compound and its cellular response. (B) A STK11 mutation-specific compound and its cellular response. (C) A MLH1 mutation-specific compound and its cellular response. The enrichment of the mutant cell lines over the wild-type cell lines are displayed in a $-\log GI_{50}$ waterfall plot. Compound specificity for the 15 mutational classes is displayed in a bar graph.

ic sensor. AMPK regulates lipid, cholesterol and glucose metabolism in specialized metabolic tissues, such as the liver, muscle and adipose tissues (18). STK11 protein is involved in two biologically important pathways that lead to cancer. First, STK11 helps to maintain a polarized epithelium, and second, STK11 activates the AMP-dependent kinase (AMPK), which controls the cellular energy balance (19). These insights into STK11 function suggest that it may represent a target of novel therapeutic strategies via its regulation of AMPK activity. Furthermore, Metformin, a widely prescribed oral hypoglycemic for diabetes, is known to activate AMPK. In the present study, NSC650914 (Phenoxan) (Fig. 2B) was identified from cluster associated with STK11 mutations, as shown in Fig. 1, and five cell lines harboring STK11 mutations showed relatively high sensitivity to Phenoxan. This compound is known to affect the mitochondrial respiratory chain in human ovarian carcinoma cell lines treated with tumor necrosis factor-alpha (TNF-α) (20), and TNF-α is known to induce insulin resistance through the AMPK pathway (21).

MutL homolog 1 (MLH1) is a gene commonly associated with hereditary nonpolyposis colorectal cancers (22). NSC741896 (4-(1-benzofuran-2-yl)-5H-1,2,3-dithiazole-5-thione) was defined as having MLH1 mutation-dependent activity (Fig. 2C). Four cell

lines harboring MLH1 mutations showed >10-fold sensitivity to this compound in comparison to MLH1 wild-type cell lines. The study by Konstantinova *et al.* provided the first evidence supporting the *in vitro* antiproliferative activity of 1,2,3-dithiazoles on human breast cancer cell lines (23). Here, we found that the MOA of this compound was more related to MLH1 genotype than cell type (i.e., breast cancer cell line).

CTNNB1 is a regulator of cell adhesion and a key downstream effector in the Wnt signaling pathway. CTNNB1 has also been implicated in tumorigenesis through the phosphorylation and destabilization induced by CK1 and GSK-3beta (24). We found that NSC731431 (Amarbellisine) possessed CTNNB1-dependent sensitivity against the NCI60 cell lines (Fig. 3A). Amarbellisine was previously reported as a strictly growth inhibitory and antiproliferative molecule in a general

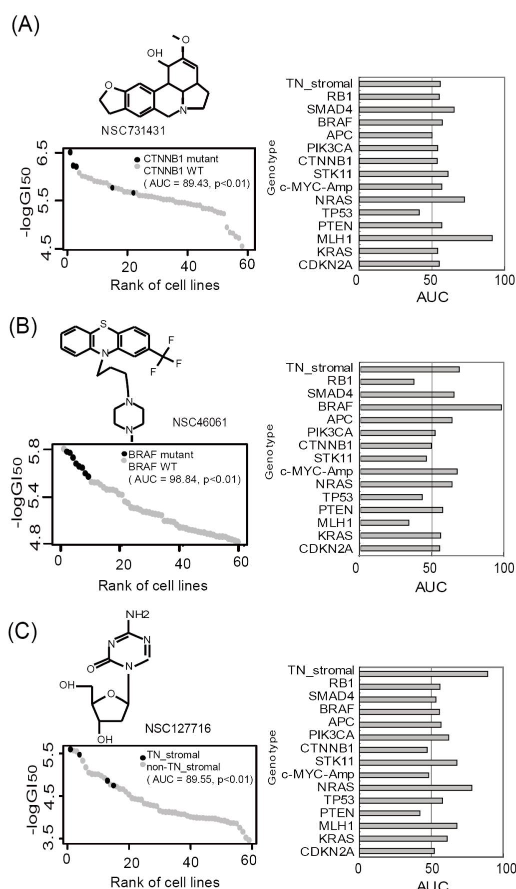

Fig. 3. Chemical structure and genotype-specific cellular response of various compounds. (A) A CTNNB1 mutation-specific compound and its cellular response. (B) A BRAF mutation-specific compound and its cellular response. (C) A TN stromal-specific compound and its cellular response. The enrichment of the mutant cell lines over wild-type cell lines are displayed in a $-\log GI_{50}$ waterfall plot. Compound specificity for the 15 mutational classes is displayed in a bar graph.

cancer drug discovery study (25), and its strong association with CTNNB1 mutations should help to further understanding this drug's MOA and anticancer applications.

BRAF is an oncogene that encodes the B-Raf protein, which is involved in intracellular signaling and cell growth. BRAF was shown to be frequently mutated in human cancers (26), and the V600E mutation of the BRAF gene has been associated with hairy cell leukemia in numerous studies (27). BRAF mutations yielded the most statistically significant associations with compound activity. NSC46061 (butanedioic acid compound with 10-[3-(4-methyl-1-piperazinyl) propyl]-2-(trifluoromethyl)-10H-phenothiazine (2 : 1)) was selected as a representative of the BRAF-dependent compound cluster (Fig. 3B), as it has been reported to represent a signature compound associated with mutations in the RAS-BRAF pathway (28). Additional compounds in the same cluster provide a promising resource for the development of new BRAF-specific cancer therapies.

Tenascin (TN) was previously shown to be highly expressed in cancer cells (29), and we identified NSC127716 (2-deoxy-50-azacytidine; Decitabine, Dacogen) as displaying a TN-stromal-dependent cellular response (Fig. 3C). Decitabine reactivates unmethylated p21 and, in some cases, this effect is independent of wild-type p53. Decitabine was also shown to restore the expression of Apaf1 in primaryAML cells and increase the susceptibility of bladder TCC to cisplatin. Other reports have also demonstrated that this compound potentiates p53 inducibility of NOXA, activates reexpression of p73 in AML cells and mediates cell cycle arrest in the G2/M phase via the p38 MAP kinase pathway (30, 31). Thus, we believe that TN-stromal can be used as a unique marker to predict the sensitivity of cancer cells to decitabine.

NCI60 chemical screening data have been used to identify new anticancer agents and understand the MOA of anti-cancer compounds (32). Although many studies have demonstrated correlations between chemical structure and MOA or cancer lineage, the importance of cancer genotype in compound responses has not been appropriately addressed. By taking advantage of publicly available genotype data regarding the NCI60 cell lines and CLEA technology, we identified a subset of compounds exhibiting significant associations with cancer genotype in terms of their cellular response. In the CLEA analysis, hierarchical clustering of selected potent compounds revealed a total of six clusters of compounds that were sensitive to mutations in KRAS, STK11, MLH1, CTNNB1, and BRAF genes and TN-stromal genotype. Many non-NCI60 compounds in previous studies were also retrospectively validated for their genotype-specific activity (33, 34). In the present study, we identified a subset of compounds with specific activity against the major cancer genotypes present in the NCI60 cell line collection. These results provide a unique resource for optimizing anticancer therapies and new drug discovery.

MATERIALS AND METHODS

NCI60 response database

The NCI60 response database contains more than 40,000 compounds with negative log-transformed GI_{50} values ($-\log GI_{50}$), which can be used to characterize sensitivity across the 60 cancer cell lines. Briefly, the cell lines were grown in 96-well plates and exposed to the test compound for 48 hours. Growth inhibition was expressed in terms of the GI_{50}, i.e., the concentration required to inhibit cell growth by 50% in comparison to that in untreated controls (3). We filtered out 9,075 compounds with missing data ($-\log GI_{50}$ values available for less than 45 cell lines) or those that possessed a minimal level of variance across the 60 cell lines (a standard deviation less than 0.1 across the available lines). The $-\log GI_{50}$ values of the remaining 34,921 compounds were used for further analysis.

Statistical analysis

Cell Line Enrichment Analysis (CLEA) was designed as a statistical analysis method to associate experimental data (compound response, gene/protein expression and protein phosphorylation) with cancer genotypes of gene mutations. We previously reported the use of CLEA for associating chemical activity with other cellular parameters (2). Briefly, the prioritization of cell lines of particular genotypes for a specific compound was analyzed on a Receiver Operating Characteristic curve (ROC) plot. The Area Under the Curve (AUC) of the ROC plot was used as a measure of "sensitivity" or "resistance". The AUC score will around 50 if random enrichment and near 100 if perfect enrichment. The statistical significance (P value) for the AUC score was assigned through 1,000 permutation tests.

Software support

The 2D structures and annotations of the NCI60 compounds were displayed using MarvinSchetch developed by ChemAxon (http://www.chemaxon.com/). Hierarchical clustering was carried out using Cluster3.0 (developed by Human Genome Center, June 2002, ((35). Tree Viewer (developed by Eisen's laboratory, (36)) was used to visualize the clustered data.

Acknowledgements

This research was supported by Sookmyung Women's University Research Grant 1-1003-0116.

REFERENCES

1. Blagosklonny, M. V. (2004) Analysis of FDA approved anticancer drugs reveals the future of cancer therapy. *Cell Cycle* **3**, 1035-1042.
2. Kim, N., He, N., Kim, C., Zhang, F., Lu, Y., Yu, Q., Stemke-Hale, K., Greshock, J., Wooster, R., Yoon, S. and Mills, G. B. (2012) Systematic analysis of genotype-specific drug responses in cancer. *Int. J. Cancer* **131**, 2456-2464.
3. 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.
4. Kwei, K. A., Baker, J. B. and Pelham, R. J. (2012) Modulators of Sensitivity and Resistance to Inhibition of PI3K Identified in a Pharmacogenomic Screen of the NCI-60 Human Tumor Cell Line Collection. *PLoS One* **7**, e46518.
5. 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.
6. Sausville, E. A. and Holbeck, S. L. (2004) Transcription profiling of gene expression in drug discovery and development: the NCI experience. *Eur. J. Cancer* **40**, 2544-2549.
7. Garraway, L. A., Widlund, H. R., Rubin, M. A., Getz, G., Berger, A. J., Ramaswamy, S., Beroukhim, R., Milner, D. A., Granter, S. R., Du, J., Lee, C., Wagner, S. N., Li, C., Golub, T. R., Rimm, D. L., Meyerson, M. L., Fisher, D. E. and Sellers, W. R. (2005) Integrative genomic analyses identify MITF as a lineage survival oncogene amplified in malignant melanoma. *Nature* **436**, 117-122.
8. 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.
9. Ikediobi, 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.
10. 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.
11. 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.
12. 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.
13. 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 integromic microarray study. *Mol. Cancer Ther.* **6**, 820-832.

14. Lee, A. C., Shedden, K., Rosania, G. R. and Crippen, G. M. (2008) Data mining the NCI60 to predict generalized cytotoxicity. *J. Chem. Inf. Model.* **48**, 1379-1388.
15. Clavel, J. (2007) Progress in the epidemiological understanding of gene-environment interactions in major diseases: cancer. *C. R. Biol.* **330**, 306-317.
16. Malumbres, M. (2007) Cyclins and related kinases in cancer cells. *J. BUON* **12(Suppl 1)**, S45-52.
17. Crow, R. T., Rosenbaum, B., Smith, R. 3rd, Guo, Y., Ramos, K. S. and Sulikowski, G. A. (1999) Landomycin A inhibits DNA synthesis and G1/S cell cycle progression. *Bioorg. Med. Chem. Lett.* **9**, 1663-1666.
18. Ahn, Y. J., Kim, H., Lim, H., Lee, M., Kang, Y., Moon, S., Kim, H. S. and Kim, H. H. (2012) AMP-activated protein kinase: implications on ischemic diseases. *BMB Rep.* **45**, 489-495.
19. Keshavarz, P., Inoue, H., Nakamura, N., Yoshikawa, T., Tanahashi, T. and Itakura, M. (2008) Single nucleotide polymorphisms in genes encoding LKB1 (STK11), TORC2 (CRTC2) and AMPK alpha2-subunit (PRKAA2) and risk of type 2 diabetes. *Mol. Genet. Metab.* **93**, 200-209.
20. Uslu, R. and Bonavida, B. (1996) Involvement of the mitochondrion respiratory chain in the synergy achieved by treatment of human ovarian carcinoma cell lines with both tumor necrosis factor-alpha and cis-diamminedichloroplatinum. *Cancer* **77**, 725-732.
21. Zhang, Z., Zhao, M., Li, Q., Zhao, H., Wang, J. and Li, Y. (2009) Acetyl-l-carnitine inhibits TNF-alpha-induced insulin resistance via AMPK pathway in rat skeletal muscle cells. *Febs. Lett.* **583**, 470-474.
22. Bonadona, V., Bonaiti, B., Olschwang, S., Grandjouan, S., Huiart, L., Longy, M., Guimbaud, R., Buecher, B., Bignon, Y. J., Caron, O., Colas, C., Nogues, C., Lejeune-Dumoulin, S., Olivier-Faivre, L., Polycarpe-Osaer, F., Nguyen, T. D., Desseigne, F., Saurin, J. C., Berthet, P., Leroux, D., Duffour, J., Manouvrier, S., Frebourg, T., Sobol, H., Lasset, C. and Bonaiti-Pellie, C. (2011) Cancer risks associated with germline mutations in MLH1, MSH2 and MSH6 genes in Lynch syndrome. *JAMA* **305**, 2304-2310.
23. Konstantinova, L. S., Bol'shakov, O. I., Obruchnikova, N. V., Laborie, H., Tanga, A., Sopena, V., Lanneluc, I., Picot, L., Sable, S., Thiery, V. and Rakitin, O. A. (2009) One-pot synthesis of 5-phenylimino, 5-thieno or 5-oxo-1,2,3-dithiazoles and evaluation of their antimicrobial and antitumor activity. *Bioorg. Med. Chem. Lett.* **19**, 136-141.
24. Lopez-Knowles, E., Zardawi, S. J., McNeil, C. M., Millar, E. K., Crea, P., Musgrove, E. A., Sutherland, R. L. and O'Toole, S. A. (2010) Cytoplasmic localization of beta-catenin is a marker of poor outcome in breast cancer patients. *Cancer Epidemiol. Biomarkers Prev.* **19**, 301-309.
25. Evidente, A., Kireev, A. S., Jenkins, A. R., Romero, A. E., Steelant, W. F., Van Slambrouck, S. and Kornienko, A. (2009) Biological evaluation of structurally diverse amaryllidaceae alkaloids and their synthetic derivatives: discovery of novel leads for anticancer drug design. *Planta Med.* **75**, 501-507.
26. Davies, H., Bignell, G. R., Cox, C., Stephens, P., Edkins, S., Clegg, S., Teague, J., Woffendin, H., Garnett, M. J., Bottomley, W., Davis, N., Dicks, E., Ewing, R., Floyd, Y., Gray, K., Hall, S., Hawes, R., Hughes, J., Kosmidou, V., Menzies, A., Mould, C., Parker, A., Stevens, C., Watt, S., Hooper, S., Wilson, R., Jayatilake, H., Gusterson, B. A., Cooper, C., Shipley, J., Hargrave, D., Pritchard-Jones, K., Maitland, N., Chenevix-Trench, G., Riggins, G. J., Bigner, D. D., Palmieri, G., Cossu, A., Flanagan, A., Nicholson, A., Ho, J. W., Leung, S. Y., Yuen, S. T., Weber, B. L., Seigler, H. F., Darrow, T. L., Paterson, H., Marais, R., Marshall, C. J., Wooster, R., Stratton, M. R. and Futreal, P. A. (2002) Mutations of the BRAF gene in human cancer. *Nature* **417**, 949-954.
27. Alonso, C. M., Such, E., Gomez-Segui, I., Cervera, J., Martinez-Cuadron, D., Luna, I., Ibanez, M., Lopez-Pavia, M., Vera, B., Navarro, I., Senent, L. and Sanz Alonso, M. A. (2012) BRAF V600E mutation in adult acute lymphoblastic leukemia. *Leuk. Lymphoma* [Epub ahead of print].
28. Ikediobi, O. N., Reimers, M., Durinck, S., Blower, P. E., Futreal, A. P., Stratton, M. R. and Weinstein, J. N. (2008) In vitro differential sensitivity of melanomas to phenothiazines is based on the presence of codon 600 BRAF mutation. *Mol. Cancer Ther.* **7**, 1337-1346.
29. Ishihara, A., Yoshida, T., Tamaki, H. and Sakakura, T. (1995) Tenascin expression in cancer cells and stroma of human breast cancer and its prognostic significance. *Clin. Cancer Res.* **1**, 1035-1041.
30. Scott, S. A., Dong, W. F., Ichinohasama, R., Hirsch, C., Sheridan, D., Sanche, S. E., Geyer, C. R. and Decoteau, J. F. (2006) 5-Aza-2'-deoxycytidine (decitabine) can relieve p21WAF1 repression in human acute myeloid leukemia by a mechanism involving release of histone deacetylase 1 (HDAC1) without requiring p21WAF1 promoter demethylation. *Leuk. Res.* **30**, 69-76.
31. Amatori, S., Papalini, F., Lazzarini, R., Donati, B., Bagaloni, I., Rippo, M. R., Procopio, A., Pelicci, P. G., Catalano, A. and Fanelli, M. (2009) Decitabine, differently from DNMT1 silencing, exerts its antiproliferative activity through p21 upregulation in malignant pleural mesothelioma (MPM) cells. *Lung Cancer* **66**, 184-190.
32. Cheng, T., Wang, Y. and Bryant, S. H. (2010) Investigating the correlations among the chemical structures, bioactivity profiles and molecular targets of small molecules. *Bioinformatics* **26**, 2881-2888.
33. Kirchheiner, J., Fuhr, U. and Brockmoller, J. (2005) Pharmacogenetics-based therapeutic recommendations- ready for clinical practice? *Nat. Rev. Drug Discov.* **4**, 639-647.
34. Altmann, A., Daumer, M., Beerenwinkel, N., Peres, Y., Schulter, E., Buch, J., Rhee, S. Y., Sonnerborg, A., Fessel, W. J., Shafer, R. W., Zazzi, M., Kaiser, R. and Lengauer, T. (2009) Predicting the response to combination antiretroviral therapy: retrospective validation of geno2pheno-THEO on a large clinical database. *J. Infect. Dis.* **199**, 999-1006.
35. de Hoon, M. J., Imoto, S., Nolan, J. and Miyano, S. (2004) Open source clustering software. *Bioinformatics* **20**, 1453-1454.
36. Eisen, M. B., Spellman, P. T., Brown, P. O. and Botstein, D. (1998) Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 14863-14868.