Հայաստանի Գիտությունների Ազգային Ակադեմիա
Национальная Академия Наук Армении
National Academy of Sciences of Armenia

Հայաստանի Կենսաբանական Հանդես
Биологический Журнал Армении
Biological Journal of Armenia

•Փորձարարական և տեսական հոդվածներ• Экспериментальные и теоретические статьи• • Experimental and theoretical articles •

Biolog. Journal of Armenia, 3 (68), 2016

ANALYSIS OF SOMATIC MUTATION ENRICHMENT IN GENE EXPRESSION LANDSCAPES FOR CANCER CELL LINES

A.A. ARAKELYAN

Group of Bioinformatics, Institute of Molecular Biology, NAS RA
arakelyanaa@zoho.com

This study is aimed at assessing the association between somatic mutations and gene expression in NCI-60 cancer cell lines using publicly available dataset and bioinformatics approaches. The study results demonstrate that the deregulation in gene expression profiles in cancer only partly can be attributed to the presence of somatic mutations, which implies the importance of assessment of epigenetic mechanisms of regulation and multi-*omics* data integration for cancer systems biology.

NCI-60 cancer cell lines – gene expression landscape – somatic mutations

Ուսումնասիրվել է NCI-60 քաղցկեղային բջջային գծերում լայնածավալ գեների էքսպրեսիայի և սոմատիկ մուտացիաների կապը: Հետազոտության արդյունքները ցույց են տվել, որ դիֆերենցիալ գեների էքսպրեսիան միայն մասամբ են ասոցիացված մուտացիաների առկայությամբ, ինչը առաջ է բերում էպիգենետիկական մեխանիզմների ուսումնասիրության անհրաժեշտությունը և մուլտի*ոմիկայի* տվյալների ինտեգրումը քաղցկեղների համակարգային կենսաբանության ոլորտում:

NCI-60 քաղցկեղային բջջային գծեր – գեների էքսպրեսիա – սոմատիկ մուտացիաներ

Проведена оценка влияния соматических мутаций на экспрессию генов в раковых клеточных линиях NCI-60. Результаты исследования показали, что дифференциальная экспрессия генов только частично зависит от наличия мутаций, что в свою очередь указывает на необходимость исследования эпигенетических механизмов регуляции экспрессии и интеграции мульти-*омных* данных в контексте системной биологии опухолей.

Клеточная линия NCI-60 – экспрессия генов – соматические мутации

Cancer cell lines are important tools for study mechanisms of cancer development and drug discovery [3]. The NCI-60 cancer cell lines set which includes leukemias, lymphomas, and carcinomas of ovarian, renal, breast, prostate, colon, lung, and CNS origin, is probably one of the most comprehensively studied and widely used collection [13]. Those cell lines have fully characterized in terms of somatic mutation typing, genome-wide gene expression, methylation and histone modification profiles [12]. However, in the most cases these data have been analyzed separately [2, 8, 11], which limits the integration of available data into single systems biology context and assessment of the role of genetic and epigenetic alterations on gene expression and path-

way activity deregulations in cancers. In this study we attempted to integrate the somatic mutation and gene expression profiles in cancer cell lines and evaluate the co-existence of differentially expressed genes enriched with somatic mutations.

Materials and methods. NCI-60 gene expression data was downloaded from Gene Expression Omnibus public repository [1]. Somatic mutation data was obtained from Cell lines project of Catalog of somatic Mutations in Cancer (COSMIC) database [4]. Clusters of differentially expressed genes were identified and functionally annotated using oposSOM package [10]. The enrichment with somatic mutations was presented as total number of mutations in the cluster. The significance of the enrichment was calculated using bootstrapping. Enrichment p values < 0.05 were considered as significant.

Results and Discussion. Self organizing map (SOM) algorithm implemented in oposSOM package reduces the gene expression landscape onto two-dimensional grids (maps), where similar gene expression profiles are combined into mini-clusters called metagenes. Each cell line is represented by a single, "personal" SOM portrait. In each portrait, metagenes are into up- and down- regulated clusters called spots (fig.1), populated by co-expressed genes.

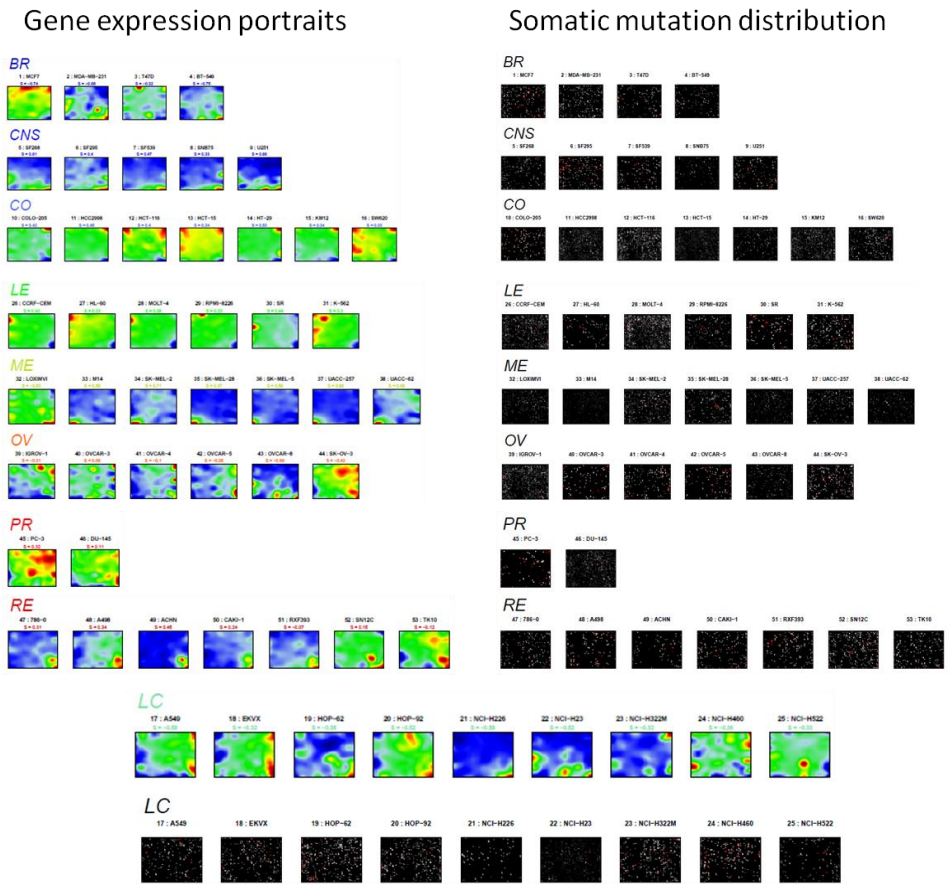
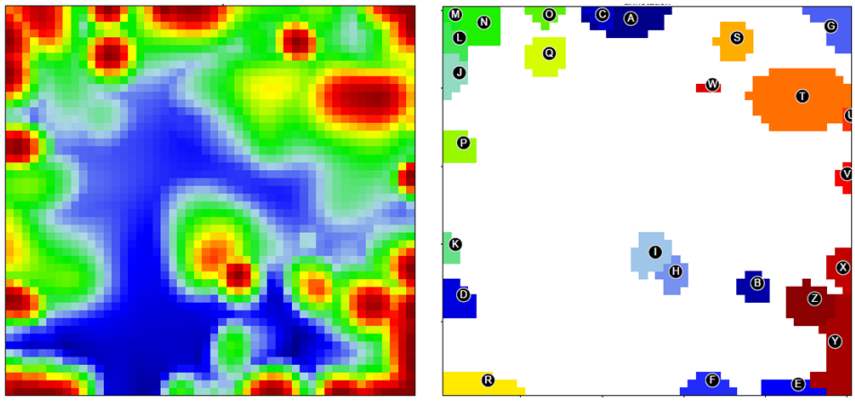


Fig. 1. Gene expression portraits and somatic mutation distribution in NCI-60 cell lines. Each cell line is characterized by its SOM portrait visualizing the gene expression levels. Red and blue spot-like areas contain highly co-regulated gene mini-clusters. Black background with dots represent mutation distribution equivalent to gene expression portraits. BR – breast cancer, CNS – central nervous system, CO – colorectal cancer, ME – melanoma, LE – leukemia, LC – lung cancer, OV – ovarian cancer, PR – prostate cancer, RE – renal cancer.

The results show that gene expression portraits in some cases show considerable similarities between cell lines one tissue groups (CNS, CO, PR, and RE), while in other tissue groups show differentially distributed gene expression spots (BR, ME, LE, LC, OV). To have more global look on differentially genes we integrated spots from all individual cell line portraits and thus it provides an overview of all relevant regions becoming activated in the data set (fig. 2).



A.XENOBIOTIC_METABOLISM
B.G2M_CHECKPOINT
C.ESTROGEN_RESPONSE_LATE
D.HEME_METABOLISM
E.EPITHELIAL_MESENCHYMAL_TRANSITION
F.GLYCOLYSIS
G.ESTROGEN_RESPONSE_LATE
H.PANCREAS_BETA_CELLS
I.MYC_TARGETS_V2
J.ALLOGRAFT_REJECTION
K.E2F_TARGETS
L.MYC_TARGETS_V1
M.MYC_TARGETS_V1
N.MYC_TARGETS_V1
O.ALLOGRAFT_REJECTION
P.INTERFERON_GAMMA_RESPONSE
Q.HEME_METABOLISM
R.EPITHELIAL_MESENCHYMAL_TRANSITION
S.MYOGENESIS
T.APICAL_SURFACE
U.COAGULATION
V.CHOLESTEROL_HOMEOSTASIS
W.CHOLESTEROL_HOMEOSTASIS
X.UV_RESPONSE_DN
Y.EPITHELIAL_MESENCHYMAL_TRANSITION
Z.TNFA_SIGNALING_VIA_NFKB

Fig. 2. Global summary map of gene expression in cancer cell lines and top functional category associated with spot genes.

In total 132 up-regulated and 109 down-regulated spots were detected in cell cancer lines. Next, we mapped somatic mutation counts into the SOM portraits to evaluate whether some clear pattern of mutation-expression dependence in cell-lines (fig.1). The results suggest that the somatic mutations are almost evenly distributed among the entire expression landscape, making it hard to distinguish mutation enriched spots. The somatic mutation enrichment analysis revealed that in all cancer cells contain mutation enriched spots as well as non-enriched spots (fig. 3).

Finally we evaluated if spots are enriched with cancer driver mutations. From 53 studied cell lines 16 had up- and down-regulated spots enriched with cancer driver mutations.

We have analyzed gene expression profiles and somatic mutations in NCI-60 cancer cell lines. Our data suggest that the gene expression profiles in cancer cell lines demonstrate considerable difference even from the same tissue origin, which is consistent with previous studies [9]. Moreover, cell lines of different tissue origin show different expression levels in respect to cancer hallmarks. Furthermore, it seems that in the most cases somatic mutations may affect the protein-protein interactions, signal activity in pathways, but not the gene expression levels [5].

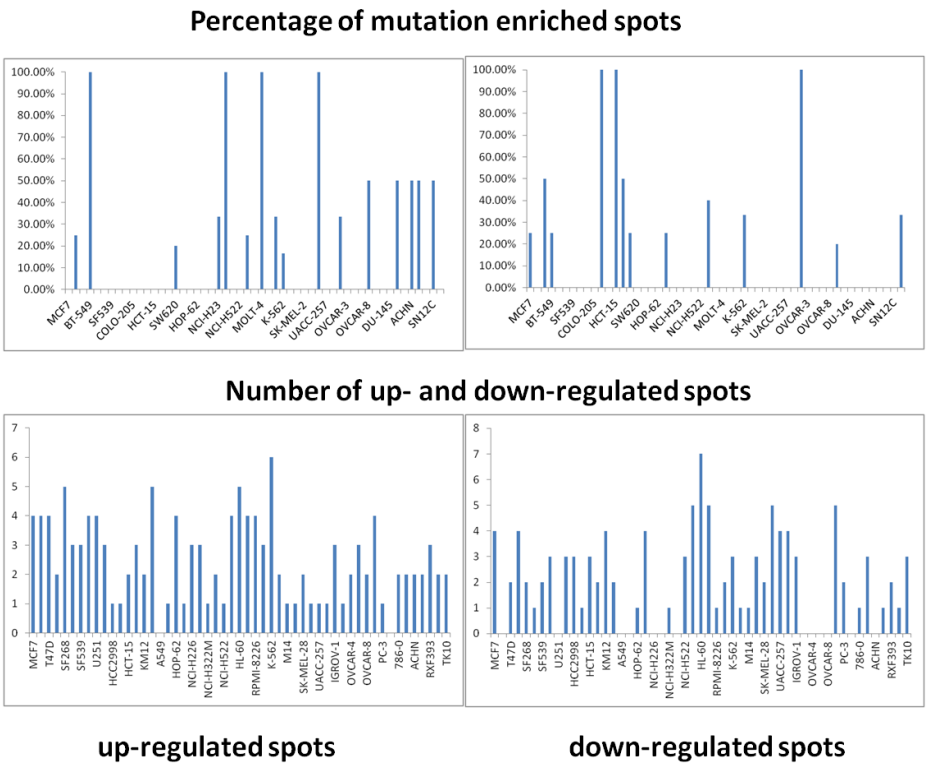


Fig. 3. Distribution of mutation enriched spots across cancer cell lines

Moreover, the recent data suggest that the alterations in epigenetic mechanisms (methylation and chromatin modifications), chromosome rearrangements play a leading role in regulation of gene expression in cancers [6,7]. Thus it is important to evaluate the pathological events in cancers by integration of multi-*omics* data derived from several sources, such as gene expression, DNA methylation and genomic data, into a single analytical framework for simultaneous assessment of their overall impact.

Acknowledgements

This work was supported by the RA MES State Committee of Science, in the framework of the research project 15T-1F150.

REFERENCES

1. *Barrett T., Wilhite S.E., Ledoux P., Evangelista C., Kim I.F., Tomashevsky M., Marshall K.A., Phillippy K.H., Sherman P.M., Holko M., Yefanov A., Lee H., Zhang N., Robertson C.L., Serova N., Davis S., Soboleva A.* NCBI GEO: archive for functional genomics data sets--update. Nucleic Acids Res., 41 (Database issue), p.D991-995, 2013.
2. *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., Sadee W.* MicroRNA expression profiles for the NCI-60 cancer cell panel. Mol. Cancer Ther., *6*, 5, p.1483-1491, 2007.

3. *Deng J., Liang H., Zhang R., Hou Y., Liu Y., Ying G., Pan Y., Hao X.* Clinical and experimental role of ring finger protein 180 on lymph node metastasis and survival in gastric cancer. Br J Surg., *103*, 4, p.407-416, 2016.
4. *Forbes S.A., Bindal N., Bamford S., Cole C., Kok C.Y., Beare D., Jia M., Shepherd R., Leung K., Menzies A., Teague J.W., Campbell P.J., Stratton M.R., Futreal P.A.* COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. Nucleic Acids Res., 39 (Database issue), D945-950, 2011.
5. *Hofree M., Shen J.P., Carter H., Gross A., Ideker T.* Network-based stratification of tumor mutations. Nat Methods, *10*, 11, p. 1108-1115, 2013.
6. *Hopp L., Löffler-Wirth H., Binder H.* Epigenetic Heterogeneity of B-Cell Lymphoma: DNA Methylation, Gene Expression and Chromatin States. Genes (Basel), *6*, 3, p.812-840, 2015.
7. *Hopp L., Nersisyan L., Löffler-Wirth H., Arakelyan A., Binder H.* Epigenetic Heterogeneity of B-Cell Lymphoma: Chromatin Modifiers. Genes (Basel) , *6*, 4, p.1076-1112, 2015.
8. *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., Wooster R.* Mutation analysis of 24 known cancer genes in the NCI-60 cell line set. Mol. Cancer Ther., *5*, 11, p.2606-2612, 2006.
9. *Lee J.K., Havaleshko D.M., Cho H., Weinstein J.N., Kaldjian E.P., Karpovich J., Grimshaw A., Theodorescu D.* A strategy for predicting the chemosensitivity of human cancers and its application to drug discovery. Proc. Natl. Acad. Sci. USA, *104*, 32, p.13086-13091, 2007.
10. *Löffler-Wirth H., Kalcher M., Binder H.* oposSOM: R-package for high-dimensional portraying of genome-wide expression landscapes on bioconductor. Bioinformatics, *31*, 19, p.3225-3227, 2015.
11. *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., Weinstein J.N.* Proteomic profiling of the NCI-60 cancer cell lines using new high-density reverse-phase lysate microarrays. Proc. Natl. Acad. Sci. USA, *100*, 24, p.14229-34, 2003.
12. *Reinhold W.C., Sunshine M., Liu H., Varma S., Kohn K.W., Morris J., Doroshow J., Pommier Y.* CellMiner: a web-based suite of genomic and pharmacologic tools to explore transcript and drug patterns in the NCI-60 cell line set. Cancer Res., *72*, 14, p. 3499-3511, 2012.
13. *Shoemaker R.H.* The NCI60 human tumour cell line anticancer drug screen. Nat Rev Cancer, *6*, 10, p. 813-823, 2006.

Received on 23.06.2015