International Journal of Medical, Medicine and Health Sciences
ISSN: 2517-9969
Vol:10, No:7, 2016

Metabolomics Profile Recognition for Cancer Diagnostics

Valentina L. Kouznetsova, Jonathan W. Wang, Igor F. Tsigelny

***Abstract*—**Metabolomics has become a rising field of research for various diseases, particularly cancer. Increases or decreases in metabolite concentrations in the human body are indicative of various cancers. Further elucidation of metabolic pathways and their significance in cancer research may greatly spur medicinal discovery. We analyzed the metabolomics profiles of lung cancer. Thirty-three metabolites were selected as significant. These metabolites are involved in 37 metabolic pathways delivered by MetaboAnalyst software. The top pathways are glyoxylate and dicarboxylate pathway (its hubs are formic acid and glyoxylic acid) along with Citrate cycle pathway followed by Taurine and hypotaurine pathway (the hubs in the latter are taurine and sulfoacetaldehyde) and Glycine, serine, and threonine pathway (the hubs are glycine and L-serine). We studied interactions of the metabolites with the proteins involved in cancer-related signaling networks, and developed an approach to metabolomics biomarker use in cancer diagnostics. Our analysis showed that a significant part of lung-cancer-related metabolites interacts with main cancer-related signaling pathways present in this network: PI3K–mTOR–AKT pathway, RAS–RAF–ERK1/2 pathway, and NFKB pathway. These results can be employed for use of metabolomics profiles in elucidation of the related cancer proteins signaling networks.

***Keywords*—**Cancer, metabolites, metabolic pathway, signaling pathway.

I. Introduction

METABOLITES exist in every living organism, their profiling can provide in-depth biological insights that advance research across a variety of areas: Oncology, diabetes, and other disorders, as well as in neurology, immunology, gerontology, naming a few. Metabolomics research lately focused on cancer biomarker discovery. Cancer is the second leading cause of death in the United States, and responsible for 582,623 deaths in 2015, and lung cancer is a sad winner of them—157,444 deaths [1]. It has become ever more imperative that researchers focus their attention in this field. In metabolomics, intermediates of various metabolic pathways, metabolites are products of bodily processes that may indicate the presence of cancers. In general, metabolites are good biomarkers for cancer diagnostics because they can be easily measured non-invasive methods. We studied the metabolites in lung cancer in attempt to select the best sets for using as biomarkers for diagnostics and monitoring of cancer progress. A number of studies appeared lately describing the sets of metabolites, which levels have aberrations in different cancers [2]–[5], nevertheless often the results are contradictory and not repeated in various publications; no consensus is reached that can make metabolite profiling a reliable tool for diagnostics.

TABLE I
METABOLITES USED AS METABOANALYST INPUT

Metabolite	HMDB ID	KEGG ID	Media	Action	Ref
1-Methylnicotinamide	HMDB00699	C02918	Urine	decrease	[6]
3-Hydroxyisovaleric acid	HMDB00754	-	Urine	decrease	[6], [7]
Acetic acid	HMDB00042	C00033	Urine	increase	[6]
Acetone	HMDB01659	C00207	Urine	increase	[6], [7]
Adipic acid	HMDB00448	C06104	Urine	decrease	[6], [7]
Alpha-hydroxyisobutyric acid	HMDB00729	-	Urine	increase	[7]
Betaine	HMDB00043	C00719	Urine	increase	[6], [7]
Choline	HMDB00097	C00114	Urine	decrease	[7]
cis-Aconitic acid	HMDB00072	C00417	Urine	increase	[7]
Citric acid	HMDB00094	C00158	Urine	increase	[6], [7]
Creatine	HMDB00064	C00300	Urine	increase	[6], [7]
D-xylose	HMDB00098	C00181	Urine	decrease	[7]
Ethanol	HMDB00108	C00469	Urine	increase	[7]
Formic acid	HMDB00142	C00058	Urine	increase	[6], [7]
Fumaric acid	HMDB00134	C00122	Urine	increase	[7]
Glycine	HMDB00123	C00037	Urine	decrease	[6]
Glycolic acid	HMDB00115	C00160	Urine	decrease	[6], [7]
Hippurate	HMDB00714	C01586	Urine	increase	[6], [7]
Hypoxanthine	HMDB00157	C00262	Urine	increase	[6], [7]
L-alanine	HMDB00161	C000541	Urine	increase	[6], [7]
L-Carnitine	HMDB00062	C15025	Urine	increase	[6], [7]
L-Valine	HMDB00883	C00183	Urine	increase	[6]
Levoglucosan	HMDB00640	-	Urine	decrease	[6]
Mannitol	HMDB00765	C00392	Urine	increase	[7]
p-Hydroxyphenylacetic acid	HMDB00020	C00642	Urine	increase	[7]
Propylene glycol	HMDB01881	C00583	Urine	decrease	[7]
Succinic acid	HMDB00254	C00042	Urine	increase	[6], [7]
Sucrose	HMDB00258	C00089	Urine	increase	[6]
Tartaric acid	HMDB00956	C00898	Urine	decrease	[6]
Taurine	HMDB00251	C00245	Urine	increase	[6]
trans-Aconitic acid	HMDB00958	C02341	Urine	increase	[6]
Trigonelline	HMDB00875	C01004	Urine	decrease	[6]
Trimethylamine N-oxide	HMDB00925	C01104	Urine	increase	[6]

V. L. Kouznetsova is with the Moores Cancer Center and San Diego Supercomputer Center, University of California, San Diego, La Jolla, CA 92093-0505 USA (corresponding author to provide phone: 858-822-2118; fax: 858-534-5117; e-mail: vkouznetsova@ ucsd.edu).

J. W. Wang is with REHS Program of San Diego Supercomputer Center, University of California, San Diego, La Jolla, CA 92093-0505 USA (e-mail: jonny9986@ gmail.com).

I. F. Tsigelny is with the Department of Neurosciences, the Moores Cancer Center, and San Diego Supercomputer Center, University of California, San Diego, La Jolla, CA 92093-0505 USA (e-mail: itsigeln@ ucsd.edu).

II. Materials and Methods

A. Compilation of Endogenous Metabolite Data for the Lung Cancers

In a process of studies various metabolites sets were found as representing different types of cancer (http://www.hmdb.ca/). Each set of metabolites may represent a biomarker for diagnosis a specific cancer. Below we describe an example of lung cancer metabolomic profiling [6], [7].

B. Metabolite Database (MDB) Analysis

Thirty-three human metabolites, which level is changed in lung cancer were selected from literature and are shown in Table I. Twenty-two metabolites increased their concentrations in urine in lung cancer patients while concentrations of eleven metabolites decreased.

C. Metabolic Pathway Analysis

For metabolic pathway analysis we used web-based software MetaboAnalyst, version 3.0 [8]. Its Pathway Analysis Module supports integrating enrichment analysis and pathway topology analysis and visualization for 21 model organisms, including Human, Mouse, Rat, Cow, Chicken, Zebrafish, Arabidopsis thaliana, Rice, Drosophila, Malaria, *S. cerevisae*, *E. coli* etc., with a total of approximately 1600 metabolic pathways [9]. The Enrichment Analysis module identifies compounds from the input lists that are over-represented and have significant aberrations to their concentrations are included in certain metabolic pathway. The topological analysis, called Pathway Analysis, module measures the centrality of a metabolite in a metabolic network or a metabolic pathway. Central or highly important metabolites with higher connectivity are "hubs," located in the center of a metabolic pathway or process. Hubs are involved in syntheses of many other metabolites; therefore, we attached special importance to them. To measure the statistical significance of having drawn a sample consisting of a specific number of **k** successes (out of **n** total draws) from a population of size **N** containing **K** successes, we used an over-representation hypergeometric test [10]. For pathway topology analysis we used relative betweenness centrality, which is an indicator of a node's centrality in a network [11].

D. Signaling Pathway Analysis

We analyzed signaling pathways using Ingenuity® Pathway Analysis (IPA®)—the web-based software for the analysis, integration, and interpretation of data derived from 'omics experiments including metabolomics [12], [13].

III. Results

Topological Pathway Analysis yields 37 metabolic pathways shown in Fig. 1 and partially described in Table II.

The Glyoxylate and dicarboxylate metabolic pathway along with Citrate cycle have best *p*-values. The third place is taken by Taurine and hypotaurine metabolic pathway. Figs. 2-4 display these pathways.

Glyoxylate and dicarboxylate metabolic pathway (Fig. 2) is responsible for the synthesis of macromolecules from dicarboxylates compounds such as ethanol and acetate [14]. It involves six metabolites; concentrations of five of them were increased in cancer patients' urine: *cis*-aconitic acid, citric acid, formic acid, glycolic acid, and succinic acid and concentration one—tartaric acid—was decreases.

TABLE II
EXTRACT FROM RESULT OF PATHWAY ANALYSIS

Pathway	Total	Expected	Hits	Raw *p*
Glyoxylate and dicarboxylate metabolism	50	0.62	6	2.40E-05
Citrate cycle (TCA cycle)	20	0.25	4	8.28E-05
Taurine and hypotaurine metabolism	20	0.25	3	1.73E-03
Phenylalanine metabolism	45	0.56	4	2.05E-03
Glycine, serine and threonine metabolism	48	0.60	4	2.61E-03
Alanine, aspartate and glutamate metabolism	24	0.30	3	2.97E-03
Pyruvate metabolism	32	0.40	3	6.79E-03
Methane metabolism	34	0.42	3	8.06E-03
Propanoate metabolism	35	0.44	3	8.74E-03
Nitrogen metabolism	39	0.49	3	1.18E-02
Nicotinate and nicotinamide metabolism	44	0.55	3	1.64E-02
Selenoamino acid metabolism	22	0.27	2	2.97E-02
Glycolysis or Gluconeogenesis	31	0.39	2	5.58E-02
Aminoacyl-tRNA biosynthesis	75	0.93	3	6.46E-02
Tyrosine metabolism	76	0.95	3	6.67E-02
Synthesis and degradation of ketone bodies	6	0.07	1	7.26E-02
Butanoate metabolism	40	0.50	2	8.75E-02
Primary bile acid biosynthesis	47	0.59	2	1.15E-01
Cyanoamino acid metabolism	16	0.20	1	1.82E-01
Arginine and proline metabolism	77	0.96	2	2.49E-01
Thiamine metabolism	24	0.30	1	2.61E-01
Purine metabolism	92	1.15	2	3.19E-01
Glycerophospholipid metabolism	39	0.49	1	3.89E-01
Fructose and mannose metabolism	48	0.60	1	4.56E-01
Porphyrin and chlorophyll metabolism	104	1.3	1	7.36E-01

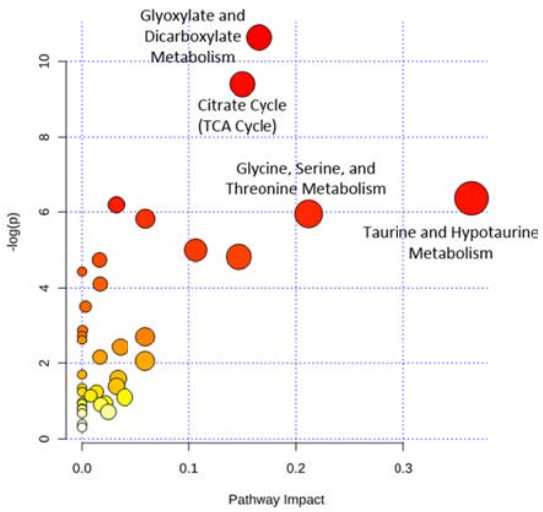

Fig. 1 Mapping lung cancer 37 collective pathways

Citrate or citric acid cycle, also known as tricarboxylic acid (TCA) cycle or the Krebs cycle is a series of chemical reactions used by all aerobic organisms to generate energy

through the oxidation of acetyl-CoA derived from carbohydrates, fats and proteins [14].

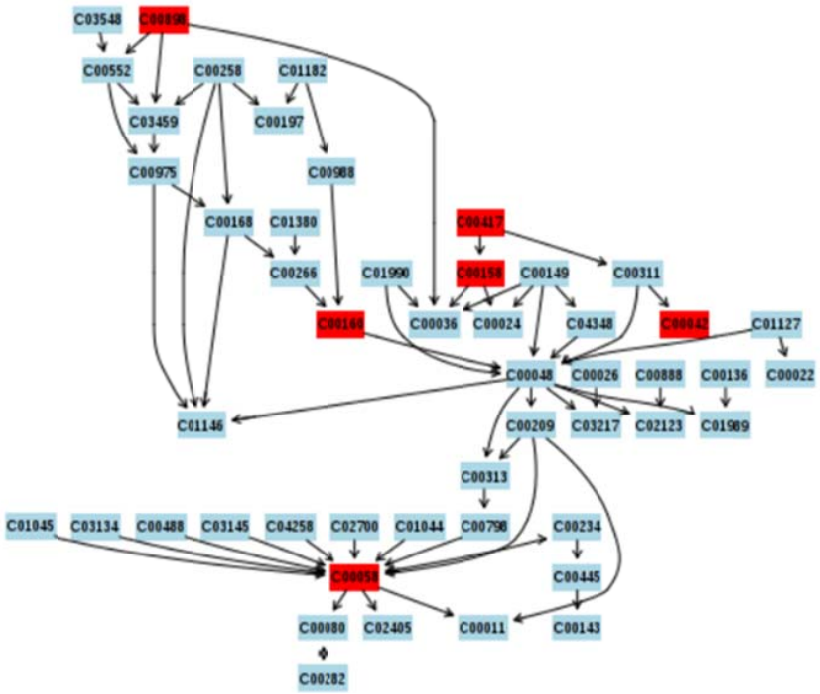

Fig. 2 Glyoxylate and dicarboxylate metabolic pathway includes six metabolites involved in lung cancer: C00898—tartaric acid, C00417—*cis*-aconitic acid, C00158—citric acid, C00160—glycolic acid, C00042—succinic acid, and C00058—formic acid. The formic acid is a hub. Another hub is glyoxylic acid (C00048)

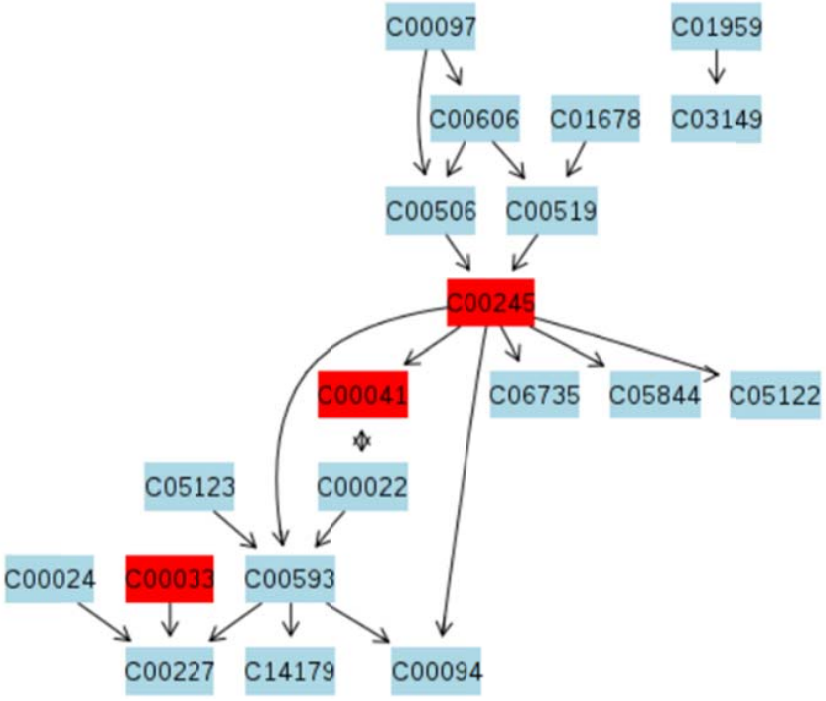

Fig. 3 Taurine and hypotaurine metabolic pathway includes three metabolites involved in lung cancer: C00033—acetic acid, C00041—L-Alanine, and C00245—taurine. The taurine is a hub. Another hub is sulfoacetaldehyde (C00593)

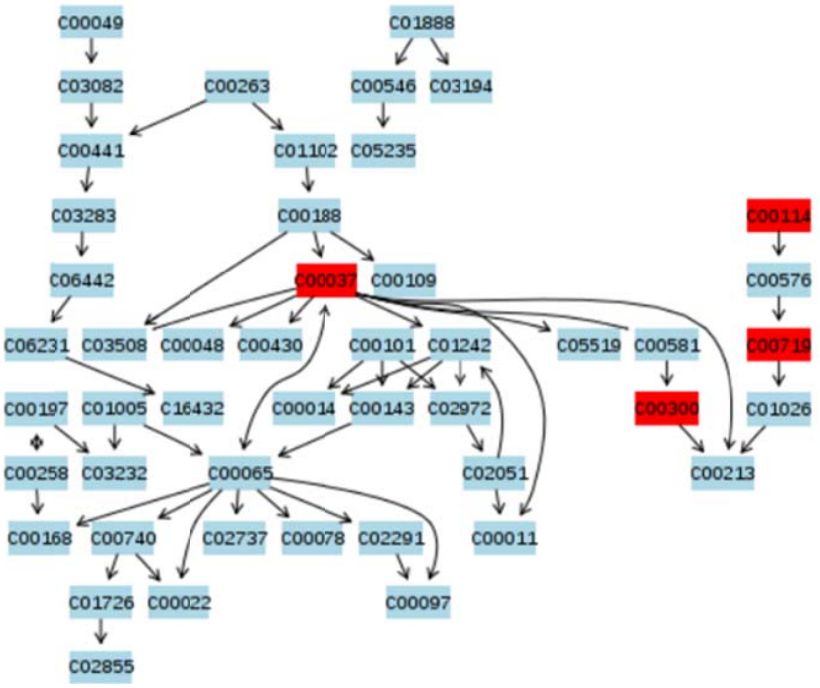

Fig. 4 Glycine, serine and threonine metabolic pathway includes four metabolites involved in lung cancer: C00114—choline, C00037—glycine, C00719—betaine, and C00300—creatine. The glycine is a hub. Another hub is L-serine (C00065)

TABLE III
METABOLITES AFFECTING GENES RELATED TO CANCER

Metabolite	Protein	Action	Cancer relation
Citric acid	ERK1/2	inhibiting	Anticancer
	caspase	activating	Anticancer
	LDH	inhibiting	Anticancer
	AKT	activating	Tumorigenesis
Glycine	ERK1/2	activating	Tumorigenesis
	LDH	inhibiting	Anticancer
Ethanol	LDH	activating	Tumorigenesis
	ERK1/2	activating	Tumorigenesis
	ERK1/2	inhibiting	Anticancer
	TGF-beta	activating	Tumorigenesis
	MAPK p38	activating	Tumorigenesis
		inhibiting	Anticancer
	NFKB	inhibiting	Anticancer
		activating	Tumorigenesis
	AKT	inhibiting	Anticancer
		activating	Tumorigenesis
	AMPK	inhibiting	Anticancer
		activating	Tumorigenesis
Taurine	TGF-beta	inhibiting	Anticancer
	ERK1/2	activating	Tumorigenesis
	MAPK p38	inhibiting	Anticancer
	AKT	activation	Tumorigenesis
Mannitol	FAK	inhibiting	Anticancer
	MAPK p38	activating	Tumorigenesis
	AKT	inhibiting	Anticancer
Creatine	AMPK	activating	Tumorigenesis
	AKT	activation	Tumorigenesis
Choline	AKT	inhibiting	Anticancer
	AMPK	inhibiting	Anticancer
	NFKB	inhibiting	Anticancer
Betaine	AKT	inhibiting	Anticancer
	caspase	activating	Tumorigenesis

We not only explored the possible biomarker functions of the selected metabolites but also elucidated their possible impact on cancer development processes. IPA® [13] analyses gave the unique opportunity to elucidate the cancer-related proteins that are affected by the metabolites (Table III). For example, important participant in both metabolic pathways is a citric acid, which inhibits ERK1/2 and LDH proteins and activates caspase. All these actions are anticancer, and simultaneously activates AKT protein promoting tumorigenesis. More studies are needed to get the "general" impact of metabolites to cancer.

Taurine is a metabolite, which participates in two important metabolic pathways: Taurine and hypotaurine pathway shown in Fig. 3 and Nitrogen pathway. Taurine is intermediate product of cysteine metabolism [14]. It acts as an anticancer agent inhibiting TGF-beta and MAPKp38 and in the same time as a tumor promoter activating ERK1/2 and AKT proteins.

Besides taurine, taurine and hypotaurine pathway includes two affecting cancer development metabolites—acetic acid and L-Alanine. Next important mammalian metabolite, which concentration is changed in lung cancer is Glycine. It is a member of at least nine metabolic pathways found by MetaboAnalyst: Glycine, serine and threonine, Methane, Purine, Nitrogen, Cyanoamino acid, Thiamine, and Porphyrin and chlorophyll metabolic pathway and in two biosynthesis: Primary bile acid and Aminoacyl-tRNA. Glycine is a central participant of Glycine, serine and threonine metabolic pathway (see Fig. 4), which synthesizes a cysteine [14]. Other affecting cancer development metabolites are Choline, Betaine, and Creatine. Among them Creatine is procancer,

Choline—anticancer, and Glycine and Betaine are both promoting and inhibiting lung cancer.

The original list of proteins with their concentrations was analyzed by the Ingenuity pathway analysis web-based software. As a result it creates a network containing 13 lung-cancer metabolites of 33.

Using a set of these metabolites the IPA program created a network in which it added proteins that almost exclusively are the parts of the main cancerogenic signaling pathways. Such a fact makes possible to use metabolites sets for prediction of the cancer pathways involved that related to these metabolites and actually can help in the analysis of cancer in patients. Note that at least three of the main cancer-related signaling pathways present in this network: PI3K-mTOR-AKT-pathway, RAS-RAF-ERK1/2 pathway, and NFKB pathway.

We want to point that the network of genes and metabolites (Fig. 5) has been obtained by using as an input data exclusively the 33 metabolites—a set of lung cancer-related metabolites.

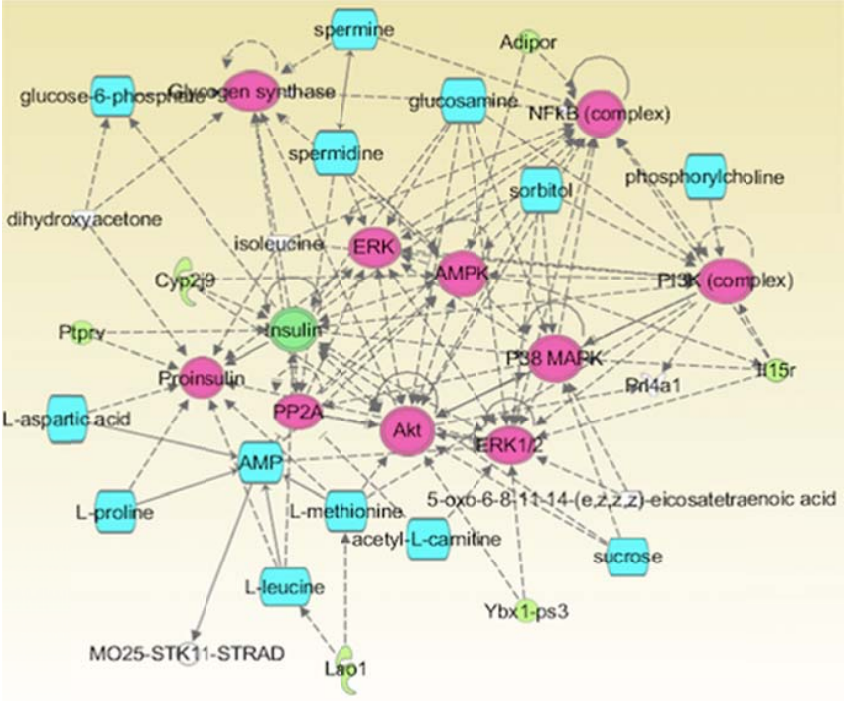

Fig. 5 Top network created by IPA® containing metabolites and proteins: Cyan—metabolites from the original list, magenta—important cancer-related proteins, green—other proteins

IV. Discussion

A number of publications devoted to study of metabolite sets related to various cancers are published lately. Scientists are encouraged by the non-invasive character of metabolite profiling that might became a tool for cancer diagnostics and monitoring. In the same time there are a number of points that are important for these studies At first, when comparing the metabolomic profiles of the cancer cells and an organism having the same cancer, we see that in the cell cultures we can define significantly greater number of metabolites (above the level of error) than in organism studies. The second—profiles of metabolites in urine can reflect the efflux transporters profiles of specific organs. It may actually be an advantage helping to define the organ of metabolites production. The third—development of standards of metabolites extraction and measurements. It is possible to create some sort of the standard markers that are always present in urine and blood to use them as points of comparison for all other metabolites concentrations.

Another important point is a possibility to use the metabolites interaction with the cancer-signaling-pathway proteins (Fig. 5). This approach may pave the way to starting an incredibly powerful method of cancer diagnostics and monitoring elucidating the specific proteins that may play a leading role in the specified cancer.

V. Conclusions

Studies of metabolomic profiles of various cancers can became a standard tool in cancer diagnostics and cancer development monitoring. The use of metabolites profiles may be used for elucidation of the gene-signaling pathways involved in the specific cancer.

References

[1] Lung Cancer Allience. 2015 Lung Cancer Facts. http://www.lungcanceralliance.org/2015%20General%20Lung%20Cancer%20Fact%20Sheet%20%20.pdf (Last visited 2/9/2016).

[2] Q. Liang Q, C. Wang, B. and Li, "Metabolomic analysis using liquid chromatography/mass spectrometry for gastric cancer," *Appl. Biochem. Biotechnol.*, vol. 176, no. 8, pp. 2170–2184, Aug. 2015.
[3] B. M. Wittmann, S. M. Stirdivant, M. W. Mitchell, J. E. Wulff, J. E. McDunn, Z. Li, *et al.*, "Bladder cancer biomarker discovery using global metabolomics profiling of urine," *PLoS One*, vol. 9, no. 12, p. e115870, Dec. 2014.
[4] A. K. Kaushik, S. K. Vareed, S. Basu, V. Putluri, N. Putluri, K. Panzitt, *et al.*, "Metabolomic profiling identifies biochemical pathways associated with castration-resistant prostate cancer," *J. Proteome Res.*, vol. 13, no. 2, pp. 1088–1100, Feb. 2014.
[5] K. Itoh, S. Aida, S. Ishiwata, T. Yamaguchi, N. Ishida, and M. Mizugaki, "Immunochemical detection of urinary 5-methyl-2'-deoxycytidine as a potential biologic marker for leukemia," *Clin. Chim. Acta*, vol. 234, no. 1–2, pp. 37–45, Jan. 1995.
[6] C. Stretch, T. Eastman, R. Mandal, R. Eisner, D. S. Wishart, M. Mourtzakis, *et al.*, "Prediction of skeletal muscle and fat mass in patients with advanced cancer using a metabolomic approach," *J. Nutr*., vol. 142, no. 1, pp. 14–21, Jan. 2012.
[7] D. S. Wishart, C. Knox, A. C. Guo, R. Eisner, N. Young, B. Gautam, *et al.*, "HMDB: a knowledgebase for the human metabolome," *Nucleic Acids Res.*, vol. 37, Database issue, pp. D303–D610, Jan. 2009.
[8] J. Xia, I. Sinelnikov, B. Han, and D. S. Wishart, "MetaboAnalyst 3.0--making metabolomics more meaningful," *Nucleic Acids Res*., vol. 43, no. W1, pp. 251–257, Jul. 2015.
[9] "MetaboAnalyst 3.0 – a comprehensive tool for metabolomic data analysis," http://www.metaboanalyst.ca/faces/home.xhtml (Last visited 2/10/2016).
[10] J. A. Rice, *Mathematical Statistics and Data Analysis* (Third ed.). Duxbury Press, 2007. p. 42.
[11] L. Freeman, "A set of measures of centrality based on betweenness," *Sociometry*, vol. 40, no. 1, pp. 35–41, Jan. 1977.
[12] "IPA®,"QIAGEN, Redwood City, www.ingenuity.com (Last visited 2/9/2016).
[13] "Ingenuity IPA—Integrate and Understand Complex 'omics Data," Ingenuity, Web, 8 Apr. 2015, http://www.ingenuity.com/products/ipa#/?tab=features (Last visited 2/10/2016).
[14] J. G. Salway, *Metabolism at a Glance* (Third ed.), Blackwell, 2004.