Prediction of Repurposed Drug Compound For Systems Treatments of Colorectal Cancer

Feng-Hsiang Chung 鍾豐翔[1] and HC Lee 李弘謙[1,2]
[1]Institute of Systems Biology and Bioinformatics, National Central University, Zhongli, Taiwan
[2]Department of Physics, Chung Yuan Christian University, Zhongli, Taiwan

Introduction

Notwithstanding advances made in the treatments in some types of cancers, progress achieved in the last 40 years in reducing the overall cancer mortality rate has been disappointing. At the same time, many previously thought successful drugs have been withdrawn, mostly due to side-effect issues. Cancer is now recognized as is a disease caused by a breakdown of a large part of the biological system in the tumor, not just of the failure of one or two of its biological functions. Here, we used "gene set enrichment analysis" to identify repurposed candidate drug compound intended for cancer treatment with minimal side effects. The intended task of the drug compound is to reprogram the tumor from a more advanced state of (a specific) cancer to a less advanced state.

Materials & Methods

Samples were two different types of frozen colonic biopsies, from prospectively collected adenomas and from normal mucosa of 32 individuals. Microarray data were downloaded from GEO database (GEO accn. GSE08671) and handled using HG-U133 Plus 2.0 platform (Affymetrix, Santa Clara). The protein-protein interaction experimental data derived from Human Protein Reference Database (HPRD) [1] and Gene Ontology database [2] were used for network analysis and functional gene sets construction. The Connectivity Map (Cmap) [3], a collection of genome-wide transcriptional expression data of bioactive drugs and small molecules on cultured human cells, was used for drug search.

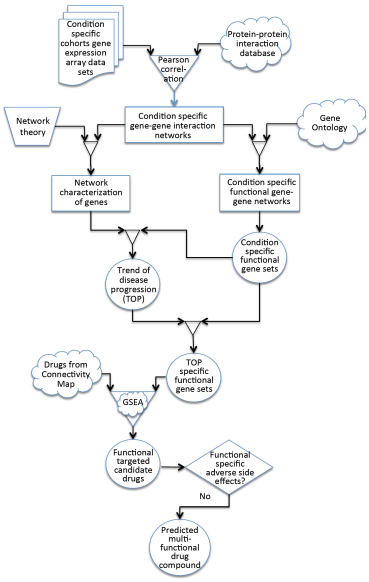

Figure 1. Flowchart of methodology

For a given disease type, a gene-gene interaction network was constructed by integration of data-set wide Pearson's correlation coefficient of gene-pairs protein-protein interaction database. Functional sub-networks were then identified based on Gene Ontology term association determined by conditional hypergeometric test. A trend of progression (TOP) procedure [4] was devised for selecting hub genes. Functional gene-sets defined by the sub-net-works were used in queries on repo-drugs in the Cmap database using GSEA [5]. Individual drugs are identified as function-specific effective, neutral, or adverse. The predicted drug compound is a set of effective and non-adverse repo-drug covering all the functional modules.

Results

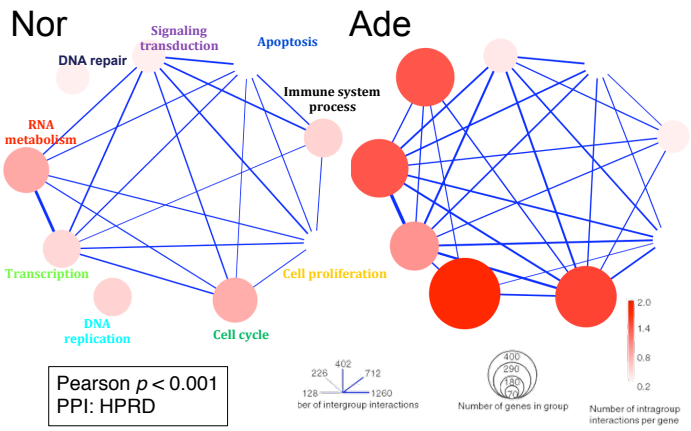

Figure 2. Functional sub-network in colonic adenoma and parental healthy mucosa. A gene-gene interaction network was constructed by cohort expression array data and protein-protein interaction DB. The GGI network is then reduced to function-function network using classification by GO terms. Nodes represent GO functions, size of node indicate number of interacting genes included, shade of node indicate number of intra-node interactions, thickness of link indicates number of internode interactions.

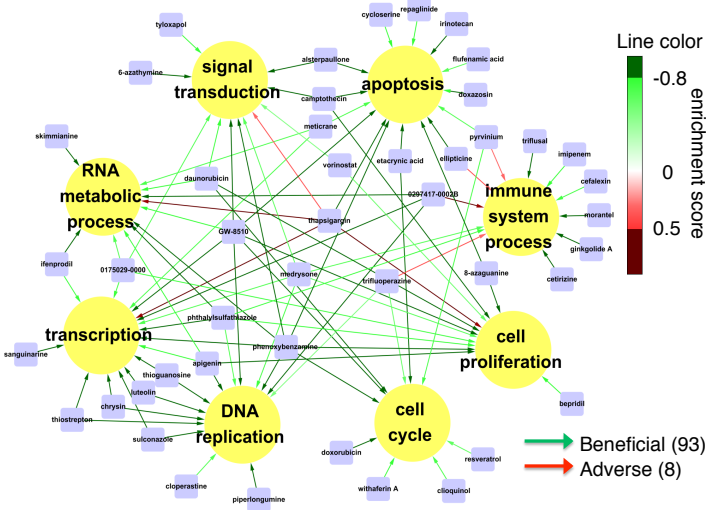

Figure 3. Drug-functional association network. Beneficial links have p-value < 0.001 (by randomization) and enrichment score < -0.7; adverse links have ES > 0.35.

Table 1. Predicted functional specific drugs and carcinogen/mutagen or antitumor agents

Drug/molecule	degree	Functional module (ES)	Drug function	Carcinogen/ mutagen	Anticancer agents
phenoxybenzamine	7	cell cycle (-0.987), DNA replication(-0.983), apoptosis (-0.977), cell proliferation (-0.962), transcription (-0.905), signal transduction (-0.886), RNA metabolic process (-0.81)	an α-adrenergic-antagonist	[1, 2]	
GW-8510	7	cell proliferation (-0.972), signal transduction (-0.936), DNA replication (-0.882), apoptosis (-0.867), cell cycle (-0.834), transcription (-0.822), RNA metabolic process (-0.791)	a CDK2 inhibitor that protects hair-loss in chemotherapy		[3]
thapsigargin	5	Apoptosis (-0.918), signal transduction (0.4), transcription (0.521), cell proliferation (0.528), RNA metabolic process (0.887)	a nonselective inhibitor of endoplasmic reticulum Ca^{2+} ATPase	[4-7]	[8-11]
daunorubicin	4	cell cycle (-0.867), cell proliferation (-0.844), RNA metabolic process (-0.8), signal transduction (-0.786)	a chemotherapeutic antibiotic		[12-14]
apigenin	4	DNA replication (-0.896), cell proliferation (-0.837), transcription (-0.796), RNA metabolic process (-0.784),	a flavone that have the chemopreventive action in vegetables		[15-18]
pyrvinium	3	cell cycle (-0.75), apoptosis (-0.694), immune system process (0.314)	Anthelmintic	[19, 20]	[21, 22]
camptothecin	3	Apoptosis (-0.953), cell proliferation (-0.935), signal transduction (-0.878)	a cytotoxic quinoline alkaloid which inhibits the DNA enzyme topoisomerase I (TOP1)		[26-28]
ellipticine	2	Apoptosis (-0.827), immune system process (0.422)	an antineoplastic agent, and TOP2 inhibitor	[29-31]	[29-31]
8-azaguanine	2	Apoptosis (-0.87), cell proliferation (-0.83)	a purine analog, an antineoplastic agent		[32, 33]
etacrynic acid	2	apoptosis(-0.891), cell cycle(-0.875)	GST Inhibitor-2, diuretics		[34]
alsterpaullone	2	signal transduction (-0.874), apoptosis (-0.866)	CDK inhibitor		[35]
thioguanosine	2	DNA replication (-0.935), transcription (-0.811)	antineoplastic agent		[39, 40]
chrysin	2	transcription (-0.934), DNA replication (-0.913)	a natural flavone, antineoplastic agent		[41, 42]
thiostrepton	2	DNA replication (-0.837), transcription (-0.816)	a natural cyclic oligopeptide antibiotic		[43-45]
luteolin	2	Transcription (-0.856), DNA replication (-0.811)	a flavonoid, and an antineoplastic agent		[46-48]
ifenprodil	2	RNA metabolic process (-0.839), transcription (-0.779)	vasodilator		[49]
triflusal	1	immune system process (-0.891)	a platelet aggregation inhibitor		[50]

Summary

1. The present program of combining functional gene sets determined by two-cohort gene expression data with Cmap allows us to find repurposed drug compounds for treating colorectal cancer with predicted strong beneficial effects on all eight biological functions and no adverse effect on any.
2. Two of the many predicted two-drug compounds are: (phenoxybenzamine, trifusal), and (GW-8510, morantel); the two three-drug compounds are (phthalylsulfathiazole, etacrynic acid, 6-azathymine or tyloxapol).
3. Program is expected to be useful for other systems diseases.

References

[1] Keshava Prasad, T.S., et al., *Human Protein Reference Database--2009 update*. Nucleic Acids Res, 2009. **37**(Database issue): p. D767-72.
[2] Ashburner, M., et al., *Gene ontology: tool for the unification of biology. The Gene Ontology Consortium*. Nature genetics, 2000. **25**(1): p. 25-9.
[3] Lamb, J., et al., *The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease*. Science, 2006. **313**(5795): p. 1929-35.
[4] Chung, F.H., et al., *Integration of network theory with gene expression data on disease progression significantly improves prediction for human colorectal cancer biomarkers (submitted)*. 2011.
[5]] A. Subramanian, A., et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. PNAS, 2005.102: 15545-50.

Work supported by National Science Council (ROC) grants 100-2911-I-008-001 and 99-2911-I-008-100.