2015

Tel Aviv University

SHNAPS and PERRY

July 27th, 2015

Dear PSB Editor,

Please find enclosed our manuscript entitled **'INFERENCE OF PERSONALIZED DRUG TARGETS VIA NETWORK PROPAGATION'** by Shnaps and Perry et al., submitted for publication in the session of 'Innovative Approaches to Combining Genotype, Phenotype, Epigenetic, and Exposure Data for Precision Diagnostics'.
The submitted paper contains original, unpublished results, and is not currently under consideration elsewhere. All co-authors concur with the contents of the paper.

We present a computational strategy to simulate drug treatment in a personalized setting. The method is based on integrating patient mutation and differential expression data with a protein-protein interaction network, and testing the impact of in-silico deletions of different proteins on the flow of information in the network. We apply our method to AML data from TCGA and validate the predicted drug targets using known targets from DrugBank and COSMIC. We show the superiority of our patient-specific approach by comparing the personalized setting predictions to those of the conventional setting. Our predicted targets were highly enriched with known drug targets (DrugBank targets: $p < 10^{-5}$, COSMIC targets: $p < 10^{-10}$). We utilized our prediction score to rank patient sensitivity to targets, reproducing the findings of in-vitro experiments studying the largest sub-group of AML patients.

Sincerely yours,

Roded Sharan, Professor
School of Computer Science,
Tel-Aviv University, Tel-Aviv, Israel 69978
Email: roded@post.tau.ac.il
Phone: +972-3-640-7139

INFERENCE OF PERSONALIZED DRUG TARGETS VIA NETWORK PROPAGATION

ORTAL SHNAPS☯[1]

School of Computer Science, Tel Aviv University
Tel Aviv 69978, Israel
Email: ortalits@gmail.com

EYAL PERRY☯

School of Computer Science, Tel Aviv University
Tel Aviv 69978, Israel
Email: eyal.perry88@gmail.com

DANA SILVERBUSH

School of Computer Science, Tel Aviv University
Tel Aviv 69978, Israel
Email: dana.silverbush@gmail.com

RODED SHARAN

School of Computer Science, Tel Aviv University
Tel Aviv 69978, Israel
Email: roded@post.tau.ac.il

We present a computational strategy to simulate drug treatment in a personalized setting. The method is based on integrating patient mutation and differential expression data with a protein-protein interaction network, and testing the impact of in-silico deletions of different proteins on the flow of information in the network. We apply our method to AML data from TCGA and validate the predicted drug targets using known targets from DrugBank and COSMIC. We show the superiority of our patient-specific approach by comparing the personalized setting predictions to those of the conventional setting. Our predicted targets were highly enriched with known drug targets (DrugBank targets: $p < 10^{-5}$, COSMIC targets: $p < 10^{-10}$). We utilized our prediction score to rank patient sensitivity to targets, reproducing the findings of in-vitro experiments studying the largest sub-group of AML patients.

☯These authors contributed equally to this work.

1. Introduction

Precision medicine, an approach where medical treatment is tailored for a specific group of patients, is an arising paradigm in medical research and practice. Indeed, it is well known that some drugs affect only a specific sub-group of patients, while even harming other patients with the same disease [1], [2]. In recent years, computational tools emerged to stratify diseases into informative subtypes [3] and to predict sensitivity of subtype in order to optimally couple patients with existing medical treatments [4].

In spite of these advances, the development of new treatments in the context of precision medicine is still scarce. Development of new drugs is an expensive and time consuming process; it takes about 15 years and up to 800 million dollars to convert a promising new compound into a marketed drug [5]. Consequently, there is increasing interest in computational prediction of drug targets. Chiang et al., Gottlieb et al., Hu et al. and Lamb et al. [6]–[9] used similarity among diseases to employ drugs designed for one disease to medicate another, as well as to prioritize new compounds as potential drugs. Lamb et al. [7] created a database containing ranked drug response gene expression profiles, allowing to query the database with a disease-specific genetic signature to identify drug response profiles that correlate to it. GBA [9] predicts novel associations between drugs and diseases by assuming that if two diseases are treated by the same drug, alternative drugs treating only one of them might treat also the other. Finally, Gottlieb et. al. [6] predicts novel associations between drugs and diseases by utilizing multiple drug–drug and disease–disease similarity measures for the prediction task. Some of the methods, such as [6], [7] could be extended for personalized prediction of drugs, yet to this date efforts for personalized design of drugs had focused on experimental work [10] or small scale networks tailored for specific condition [11], [12].

As drugs often act by inhibiting a target in a manner resembling a knockout, attempts were also made to predict candidates for drug targets by predicting the effect of gene knockouts. These attempts focused on metabolic drugs and used metabolic network models, testing the impact of in-silico deletion of genes on the network's fluxes. For example, Fatumo et.al. [13] simulated knockouts by deleting reactions from a metabolic network to identify enzymes essential for the malaria parasite Plasmodium falciparum. Papp et al. [14] used a metabolic flux model to predict the knockout fitness effect of nonessential genes in Saccharomyces cerevisiae. In their review of current paradigms for predicting inhibitory effects, Csermely et al. [15] conclude with the need for approaches allowing the examination of multi-targets inhibition, as our new approach allows.

In this work we present a novel approach to tackle the drug target inference problem from a personalized perspective using in-silico knockouts based on propagation methods in a protein-protein interaction (PPI) network. Figure 1 provides an overview of the method: we start from a general PPI network and personal disease-related data. We rely on the framework described by Vanunu et al. [16] to prioritize casual genes by network propagation. We perform multiple network propagations in order to simulate the current patient state, the patient state after gene knockouts (by removing the

gene's node from the network) and an estimated "healthy" state. We use these different states in order to rank the gene knockouts and retrieve a candidates list for potential novel drug targets.

The framework we presented is general and can be applied to any personalized disease-related data, with cancer being a pronounced candidate for application. Cancer is wildly heterogeneous, in that gene combinations can vary greatly between patients suffering from the same type of cancer. This is especially true in acute myeloid leukemia (AML), which has striking heterogeneity in gene mutations and expression aberrations across samples [17]–[19]. We therefore evaluated our performance by applying it on patients suffering from AML, based upon data generated by the TCGA Research Network: http://cancergenome.nih.gov/. of mutated and differentially expressed genes [19].

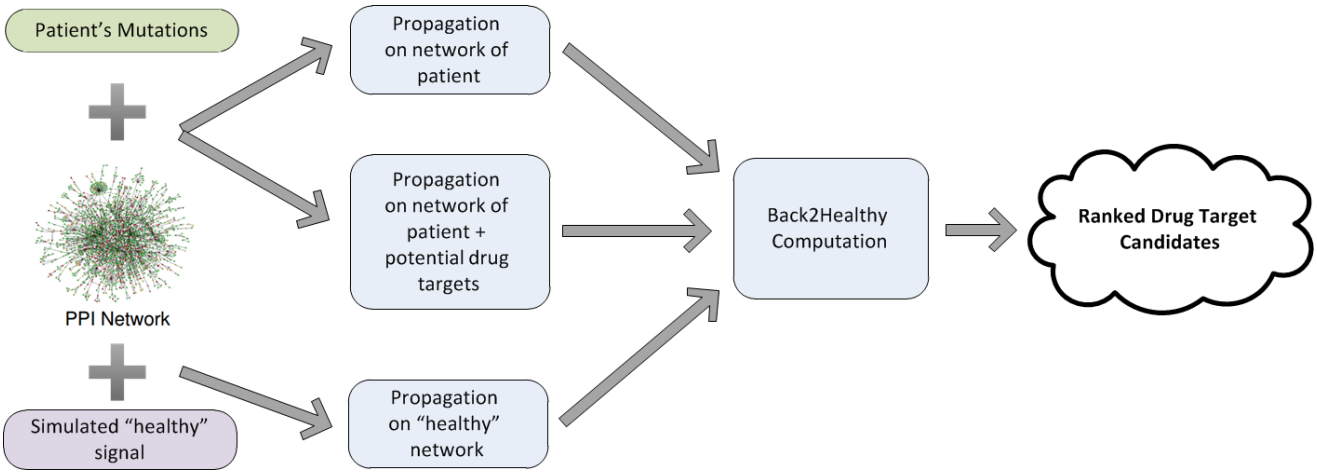

Fig. 1. An overview of the algorithmic pipeline.

2. Results

We present a novel approach to tackle the drug target inference problem from a personalized perspective using in-silico knockouts in a PPI network. As described in Figure 1, we start from a general PPI network and individual-specific disease-related data. We rely on the framework described by Vanunu et al. [16] to infer gene prioritization by network propagation. We perform multiple network propagations in order to simulate the current patient state, the patient state after gene knockouts (by removing the gene's node from the network) and an estimated "healthy" state. We use these different states in order to rank the gene knockouts and retrieve a candidates list for potential novel drug targets.

To evaluate our performance we applied our method to TCGA gene expression and mutation data of patients suffering from acute myeloid leukemia (AML). We gathered this dataset from the

COSMIC cancer gene census [20]. First, we show that we can identify common AML causal genes and components in consensus pathway by synthesizing the individual mutations set propagations and ranking according to propagation scores. Second, we show that by integrating results from a personalized knockout process we can infer potential drug targets and rank their efficacy in a patient or a sub-group of patients.

We executed the algorithm using different settings for its *alpha* parameter (0.5, 0.75 and 0.9, see Methods) and different *prior knowledge gene set*s (labeled P, see Methods). To evaluate the results, we used three sets of known causal genes, varying in confidence and size: 10 known AML causal genes from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (http://www.genome.jp/kegg/) [21], [22], 72 known AML causal genes from COSMIC and 363 known cancer causal genes from COSMIC. The application of the method to each patient resulted in a propagation score for each gene. We aggregated the rank of each gene over all patients to yield a gene-based score, retaining the top 10% affected genes in the network. We then computed the hypergeometric enrichment of this set of genes with the different sets of known causal genes. All choices of α resulted in significant and similar p-values ($p < 10^{-5}$), as shown in Figure 2A. Resulting from the high similarity between all choices of α, we decided to base our choice on additional tests where a subset of patients was selected, in which $\alpha = 0.9$ proved to be the best one~~, though the best one out of the three is 0.9, as shown in Figure 2A~~. This value was used in the sequel.

Next, we wish to assess different individual-specific data sources for the assembly of prior knowledge gene sets (P) used to initiate the propagation (see Methods). Each individual holds information about each gene - whether it is mutated and/or differentially expressed. Therefore, we examined four settings – defining P based on (i) mutated genes; (ii) differentially expressed genes; (iii) both, but running them separately and averaging the propagation scores obtained; and (iv) same as (iii) but taking the maximum scores rather than the average. We apply the method to each patient starting from the respected prior knowledge variant $P_{patient}$ and aggregate the results to retain the top 10% affected genes according to the current variant. We again compute the hypergeometric enrichment of this gene set with the different sets of known causal genes. All prior knowledge variants resulted in significant p-values ($p < 10^{-5}$). The best variant was the first – setting P to be the set of mutated genes in each patient (Figure 2B), a choice we use in the sequel.

The causal genes trigger malignant behavior by perturbing signaling pathways that regulate three core cellular processes: cell fate, cell survival, and genome maintenance [23]. In AML cell survival and proliferation is enhanced through an aberrant signal pathway [24] represented in the KEGG database [21]. We computed the hypergeometric enrichment of the top 10% affected genes within this particular KEGG pathway and found that the affected genes comprise 15 out of 21 pathway components (Figure 3) resulting in a significant p-value ($p < 10^{-11}$).

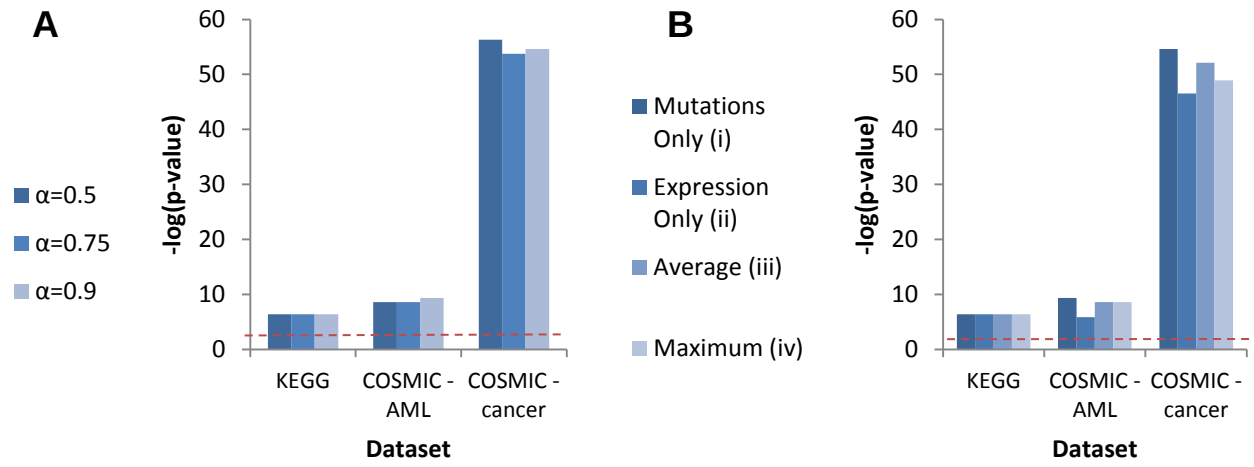

Fig. 2. Performance evaluation under different parameter (A) and prior knowledge set (B) choices. The red line stands for a p-value of 0.01.

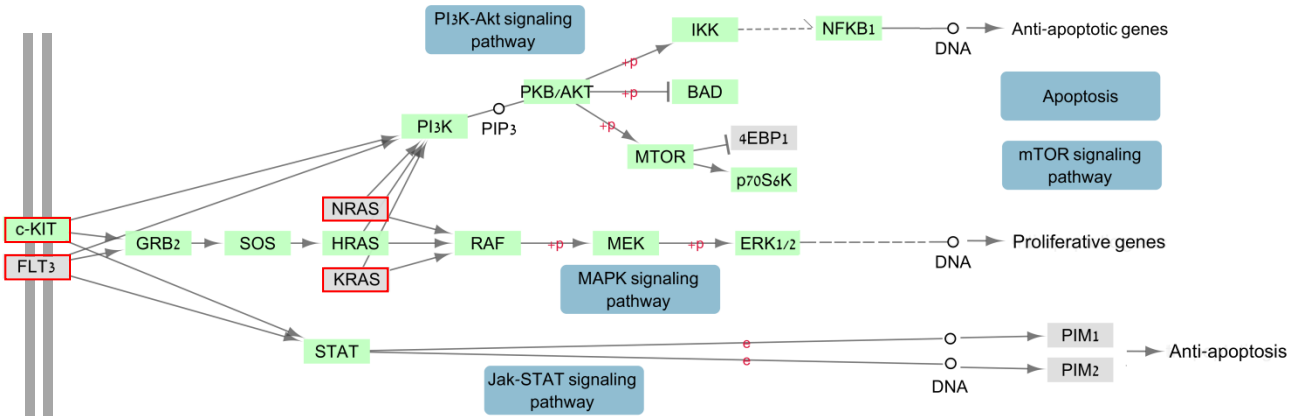

Fig. 3. AML KEGG pathway marking top 10% affected genes (green) and mutations appearing in at least 2 patients (red box).

The previous results imply that our propagation based scores are able to infer disease-related genes and agree with observations made by Rufallo et al. [25]. We hypothesized that good drug targets for the disease could be genes whose knockout is predicted to reverse the disease-related effects [7]. To identify such genes in-silico, we rerun the propagation based scoring while removing each gene in turn from the network, and assessing the similarity between the obtained scores and those that characterize a "healthy" state. To this end, we use a Back2Healthy distance score (B2H; See Methods), taking the top scoring genes as our candidates for potential personalized drug targets. We focus on non-trivial targets by excluding the patient's mutated genes from our ranking.

The process above infers drug targets for each patient. As information about personalized drug targets is very scarce and hard to validate, we aggregated the results over all patients, evaluating the results using known AML drug targets derived from the DrugBank database [26]–[29] and COSMIC [20]. The top 10% scoring genes were highly enriched with known drug targets from both sources (Figure 4A, DrugBank: $p < 10^{-5}$, COSMIC: $p < 10^{-10}$). To assess the personalized approach we took, we generated a "consensus patient", using consensus (appearing in at least 5 patients) mutated and differentially expressed genes derived by aggregating all AML patients. The results were insignificant (Figure 4B), underscoring the utility of a personalized approach.

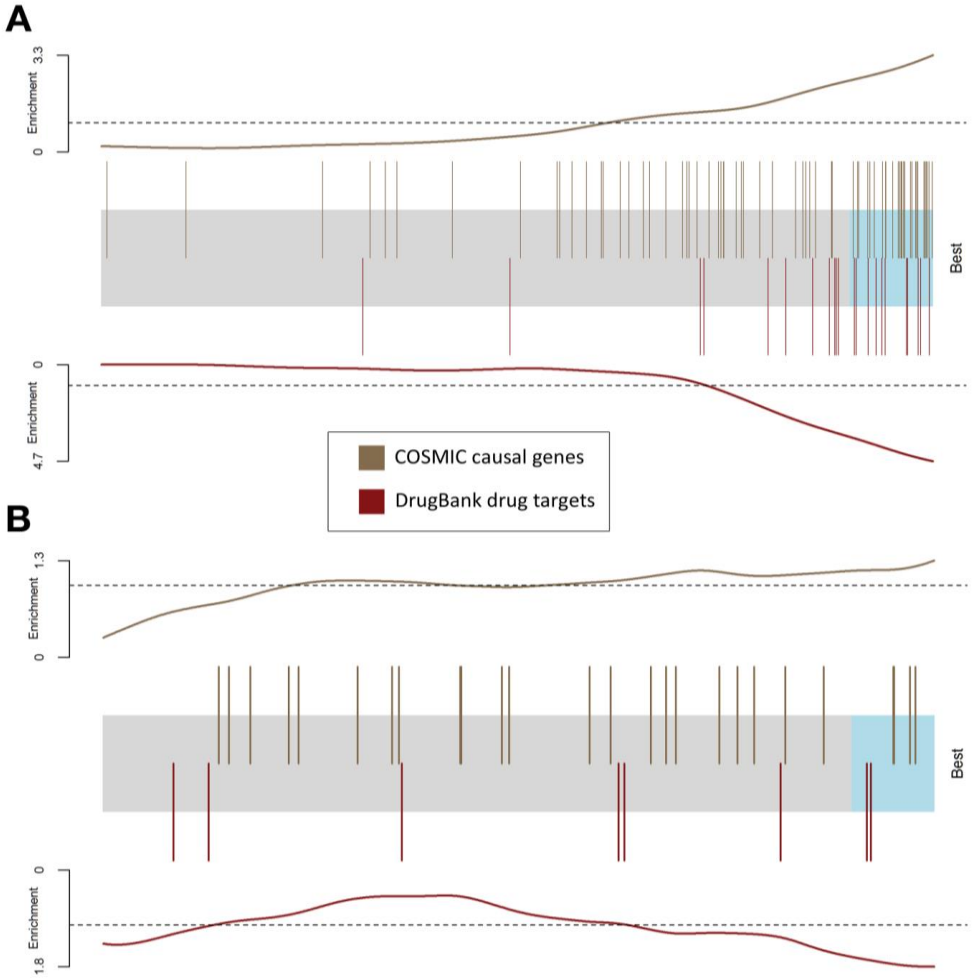

Fig. 4. Performance of drug target prediction. The candidate genes are represented by a shaded rectangle, where the top 10% are shaded cyan. Every overlaid bar stands for a single gene in a collection of known or potential drug targets. Traces above/below the bar represent relative enrichment. (A) The barcode plot was generated by running our method on each AML patient independently and aggregating the results. (B) The barcode plot was generated by running a similar single pipeline on the consensus patient.

We used the method to predict the sensitivity of the largest sub-group of AML patients – carriers of the FLT3 mutation – to known inhibitors. Even though the original goal of the framework is to predict novel drug targets, we hypothesized that we can use same method to prioritize known drug targets in respect to the sensitivity in a specific group of patients. The following inhibitors were experimentally examined as potential drug targets and their influence on FLT3 mutated cell lines was carefully documented. Jin et al. [30] tested PI3K inhibitor and found FLT3 mutated cell lines to be poorly responsive to it. Nishioka et. al [31] showed that the MEK inhibitor caused those cell lines to respond moderately by leading to decreased abnormal proliferation, nearly resembling a healthy cell phenotype, yet showing unchanged abnormal levels of apoptosis. Keeton et al. [32] demonstrated how PIM inhibitor caused FLT3 mutated cell lines to respond with high sensitivity, which led to development of the PIM inhibiting drug AZD1208. Our method shows in-silico sensitivity to PIM knockout, intermediate sensitivity to MEK knockout, and lesser sensitivity to PI3K knockout (Figure 5). These results corroborate the findings of [30]–[32].

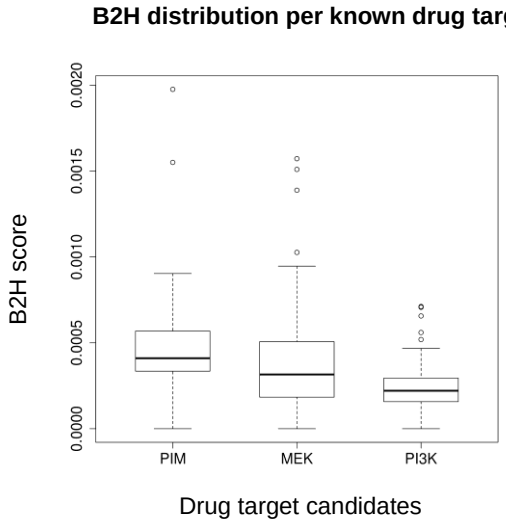

Fig 5. Sensitivity of FLT3 mutated cell lines, as predicted by B2H scores, corroborating the findings of [30]–[32] in-vitro experiments

3. Methods

3.1. *Computing propagation score*

We use the network propagation method described in Vanunu et al. [16]. The input consists of a network $G = (V, E, w)$ with V as the set of proteins, E as the set of their interactions, $w(u, v)$ represents the reliability of the interaction between u and v, and a prior knowledge protein set P. Our goal is to prioritize the proteins in V with respect to P. We do so by defining a function $F: V \to \mathcal{R}$ that is both smooth over the network and accounts for the prior knowledge about each node.

Regarding our framework, the set of protein-protein interactions and their reliability are taken from the HIPPIE network [33]. The prior knowledge set P is derived from the patient's mutation data.

As described by Vanunu et al. we use Laplacian normalization to produce the normalized network edge weight w'. Briefly we construct a sparse $|V| \times |V|$ matrix W from the edge weights w, and construct a diagonal matrix D with $D[i, i] = \Sigma_j W[i, j]$. The normalized edge weight matrix is computed as $W' = D^{-1/2} W D^{-1/2}$.

We define a prior knowledge function $Y: V \to \{0,1\}$ such that:

$$\forall v \in V : \; Y(v) = \begin{cases} 1 & v \in P \\ 0 & v \notin P \end{cases}$$

With the normalized weight matrix W' and the prior knowledge function Y, we use the iterative procedure described by Law et al. [27] to compute F. Namely, starting with $F^{(0)} = Y$, we update F at iteration t as follows:

$$F^{(t)} = \alpha W' F^{(t-1)} + (1 - \alpha) Y$$

The procedure is repeated iteratively until convergence; namely we stop the iterations when the following condition occurs:

$$\left\| F^{(t)} - F^{(t-1)} \right\|_2 < 10^{-4}$$

The propagation score for each gene is its rank among all genes after propagating the network, where lower ranks means higher $F(v)$ value. In case of ties the ranks of the genes is averaged.

3.2. *Back2Healthy distance score*

Let S_{before} , S_{after} be vectors of propagation scores for a chosen gene set (here, the set of differentially expressed genes of some patient) A, where S_{before} was generated by propagating on the original PPI network, while S_{after} was generated by propagating on a "knockout" network, where one of the genes was removed. We define the Back2Healthy (B2H) distance between S_{before} and S_{after} as follows:

Let k be the size of the prior gene set of the patient (the patient's set of mutated genes). For $1 \leq i \leq n$ $(n = 1000)$, we generate a score vector S_i for A by propagating the original PPI network

and setting the prior knowledge set P to be k random nodes (disjoint from A) in order to simulate a "healthy" distribution of propagation scores for A.

Next, for $a \in A$, define

$$Q_{before_a} = \frac{\left|\{\, 1 \le i \le n | S_i[a] < S_{before}[a]\}\right|}{n}$$

$$Q_{after_a} = \frac{\left|\{\, 1 \le i \le n | S_i[a] < S_{after}[a]\}\right|}{n}$$

Hence, Q_{before_a} represents the quantile of $S_{before[a]}$ in our simulated distribution, and similarly for Q_{after_a}. Finally, $B2H(S_{before}, S_{after})$ is defined as:

$$B2H(S_{before}, S_{after}) = \frac{\Sigma_{a \in A} |Q_{before_a} - Q_{after_{la}}|}{|A|}$$

3.3. Data Sets

3.3.1. *Patients data*

The TCGA data portal [7] analyzed 200 clinically annotated adult cases of AML (updated to 29/04/2015). The analysis includes whole-genome sequencing of the primary tumor and matched normal skin samples from 50 patients and exome capture and sequencing for another 150 paired samples of AML tumor and skin [19]. For 174 of the patients both mutations and expression were collected. The change in gene expression is calculated by Cosmic methodology [20], in which a Z-score is computed based on the sequencing platform (in order of preference: IlluminaHiSeq RNASeqV2, IlluminaGA RNASeqV2, IlluminaHiSeq RNASeq, IlluminaGA RNASeq), and a gene is considered significantly differentially expressed if its Z-score exceeds a threshold of 2 STDEV.

To construct individual specific network for propagation we project the matching set of mutations and differentially expressed genes of an individual on a human PPI network taken from HIPPIE [33], and comprising 186,217 interactions among 15,029 proteins.

The projected individual networks have in average 7.6 mutations and 340 differentially expressed genes.

3.3.2. *Drug targets*

The DrugBank database [27] combines detailed drug data with comprehensive drug target information. We screened for AML drugs, and retained their targets composing a collection of 22 drug targets.

3.3.3. *Known causal genes*

We obtained 3 sets on known causal genes, varying in confidence and size, and matched them to nodes in the PPI network. 10 known AML causal genes were collected from the Kyoto Encyclopedia

of Genes and Genomes (KEGG) database (http://www.genome.jp/kegg/) [21], [22], 94 known AML causal genes from COSMIC (72 of which are in our PPI network) and 533 known cancer causal genes from COSMIC (363 in the network).

References

[1] P. I. Poulikakos, C. Zhang, G. Bollag, K. M. Shokat, and N. Rosen, "RAF inhibitors transactivate RAF dimers and ERK signalling in cells with wild-type BRAF," *Nature*, vol. 464, no. 7287, pp. 427–430, Mar. 2010.

[2] P. M. Rothwell, "Can overall results of clinical trials be applied to all patients?," *The Lancet*, vol. 345, no. 8965, pp. 1616–1619, Jun. 1995.

[3] M. Hofree, J. P. Shen, H. Carter, A. Gross, and T. Ideker, "Network-based stratification of tumor mutations," *Nat. Methods*, vol. 10, no. 11, pp. 1108–1115, Nov. 2013.

[4] M. Niepel, M. Hafner, E. A. Pace, M. Chung, D. H. Chai, L. Zhou, B. Schoeberl, and P. K. Sorger, "Profiles of Basal and Stimulated Receptor Signaling Networks Predict Drug Response in Breast Cancer Lines," *Sci. Signal.*, vol. 6, no. 294, Sep. 2013.

[5] J. A. DiMasi, R. W. Hansen, and H. G. Grabowski, "The price of innovation: new estimates of drug development costs," *J. Health Econ.*, vol. 22, no. 2, pp. 151–185, Mar. 2003.

[6] A. Gottlieb, G. Y. Stein, E. Ruppin, and R. Sharan, "PREDICT: a method for inferring novel drug indications with application to personalized medicine," *Mol. Syst. Biol.*, vol. 7, p. 496, Jun. 2011.

[7] J. Lamb, E. D. Crawford, D. Peck, J. W. Modell, I. C. Blat, M. J. Wrobel, J. Lerner, J.-P. Brunet, A. Subramanian, K. N. Ross, M. Reich, H. Hieronymus, G. Wei, S. A. Armstrong, S. J. Haggarty, P. A. Clemons, R. Wei, S. A. Carr, E. S. Lander, and T. R. Golub, "The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease," *Science*, vol. 313, no. 5795, pp. 1929–1935, Sep. 2006.

[8] G. Hu and P. Agarwal, "Human Disease-Drug Network Based on Genomic Expression Profiles," *PLoS ONE*, vol. 4, no. 8, Aug. 2009.

[9] A. P. Chiang and A. J. Butte, "SYSTEMATIC EVALUATION OF DRUG-DISEASE RELATIONSHIPS TO IDENTIFY LEADS FOR NOVEL DRUG USES," *Clin. Pharmacol. Ther.*, vol. 86, no. 5, pp. 507–510, Nov. 2009.

[10] P. P. Zarrinkar, R. N. Gunawardane, M. D. Cramer, M. F. Gardner, D. Brigham, B. Belli, M. W. Karaman, K. W. Pratz, G. Pallares, Q. Chao, K. G. Sprankle, H. K. Patel, M. Levis, R. C.

Armstrong, J. James, and S. S. Bhagwat, “AC220 is a uniquely potent and selective inhibitor of FLT3 for the treatment of acute myeloid leukemia (AML),” *Blood*, vol. 114, no. 14, pp. 2984–2992, Oct. 2009.

[11] R. Chuang, B. A. Hall, D. Benque, B. Cook, S. Ishtiaq, N. Piterman, A. Taylor, M. Vardi, S. Koschmieder, B. Gottgens, and J. Fisher, “Drug Target Optimization in Chronic Myeloid Leukemia Using Innovative Computational Platform,” *Sci. Rep.*, vol. 5, Feb. 2015.

[12] M. F. Ciaccio, V. C. Chen, R. B. Jones, and N. Bagheri, “The DIONESUS algorithm provides scalable and accurate reconstruction of dynamic phosphoproteomic networks to reveal new drug targets,” *Integr. Biol. Quant. Biosci. Nano Macro*, vol. 7, no. 7, pp. 776–791, Jul. 2015.

[13] S. Fatumo, K. Plaimas, J.-P. Mallm, G. Schramm, E. Adebiyi, M. Oswald, R. Eils, and R. König, “Estimating novel potential drug targets of Plasmodium falciparum by analysing the metabolic network of knock-out strains in silico,” *Infect. Genet. Evol. J. Mol. Epidemiol. Evol. Genet. Infect. Dis.*, vol. 9, no. 3, pp. 351–358, May 2009.

[14] B. Papp, C. Pál, and L. D. Hurst, “Metabolic network analysis of the causes and evolution of enzyme dispensability in yeast,” *Nature*, vol. 429, no. 6992, pp. 661–664, Jun. 2004.

[15] P. Csermely, V. Ágoston, and S. Pongor, “The efficiency of multi-target drugs: the network approach might help drug design,” *Trends Pharmacol. Sci.*, vol. 26, no. 4, pp. 178–182, Apr. 2005.

[16] O. Vanunu, O. Magger, E. Ruppin, T. Shlomi, and R. Sharan, “Associating genes and protein complexes with disease via network propagation,” *PLoS Comput. Biol.*, vol. 6, no. 1, p. e1000641, Jan. 2010.

[17] G. Marcucci, T. Haferlach, and H. Döhner, “Molecular Genetics of Adult Acute Myeloid Leukemia: Prognostic and Therapeutic Implications,” *J. Clin. Oncol.*, p. JCO.2010.30.2554, Jan. 2011.

[18] H. Wang, H. Hu, Q. Zhang, Y. Yang, Y. Li, Y. Hu, X. Ruan, Y. Yang, Z. Zhang, C. Shu, J. Yan, E. K. Wakeland, Q. Li, S. Hu, and X. Fang, “Dynamic transcriptomes of human myeloid leukemia cells,” *Genomics*, vol. 102, no. 4, pp. 250–256, Oct. 2013.

[19] “Genomic and Epigenomic Landscapes of Adult De Novo Acute Myeloid Leukemia,” *N. Engl. J. Med.*, vol. 368, no. 22, pp. 2059–2074, May 2013.

[20] S. A. Forbes, N. Bindal, S. Bamford, C. Cole, C. Y. Kok, D. Beare, M. Jia, R. Shepherd, K. Leung, A. Menzies, J. W. Teague, P. J. Campbell, M. R. Stratton, and P. A. Futreal, “COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer,” *Nucleic*

Acids Res., vol. 39, no. Database issue, pp. D945–D950, Jan. 2011.

[21] M. Kanehisa and S. Goto, "KEGG: Kyoto Encyclopedia of Genes and Genomes," *Nucleic Acids Res.*, vol. 28, no. 1, pp. 27–30, Jan. 2000.

[22] M. Kanehisa, S. Goto, Y. Sato, M. Kawashima, M. Furumichi, and M. Tanabe, "Data, information, knowledge and principle: back to metabolism in KEGG," *Nucleic Acids Res.*, vol. 42, no. Database issue, pp. D199–205, Jan. 2014.

[23] B. Vogelstein, N. Papadopoulos, V. E. Velculescu, S. Zhou, L. A. Diaz, and K. W. Kinzler, "Cancer Genome Landscapes," *Science*, vol. 339, no. 6127, pp. 1546–1558, Mar. 2013.

[24] C. Scholl, D. G. Gilliland, and S. Fröhling, "Deregulation of signaling pathways in acute myeloid leukemia," *Semin. Oncol.*, vol. 35, no. 4, pp. 336–345, Aug. 2008.

[25] M Ruffalo, M Koyuturk, and R Sharan, "Network-Based Integration of Disparate Omic Data To Identify 'Silent Players' in Cancer," *submitted*, 2015.

[26] C. Knox, V. Law, T. Jewison, P. Liu, S. Ly, A. Frolkis, A. Pon, K. Banco, C. Mak, V. Neveu, Y. Djoumbou, R. Eisner, A. C. Guo, and D. S. Wishart, "DrugBank 3.0: a comprehensive resource for 'omics' research on drugs," *Nucleic Acids Res.*, vol. 39, no. Database issue, pp. D1035–1041, Jan. 2011.

[27] V. Law, C. Knox, Y. Djoumbou, T. Jewison, A. C. Guo, Y. Liu, A. Maciejewski, D. Arndt, M. Wilson, V. Neveu, A. Tang, G. Gabriel, C. Ly, S. Adamjee, Z. T. Dame, B. Han, Y. Zhou, and D. S. Wishart, "DrugBank 4.0: shedding new light on drug metabolism," *Nucleic Acids Res.*, p. gkt1068, Nov. 2013.

[28] D. S. Wishart, C. Knox, A. C. Guo, S. Shrivastava, M. Hassanali, P. Stothard, Z. Chang, and J. Woolsey, "DrugBank: a comprehensive resource for in silico drug discovery and exploration," *Nucleic Acids Res.*, vol. 34, no. Database issue, pp. D668–672, Jan. 2006.

[29] D. S. Wishart, C. Knox, A. C. Guo, D. Cheng, S. Shrivastava, D. Tzur, B. Gautam, and M. Hassanali, "DrugBank: a knowledgebase for drugs, drug actions and drug targets," *Nucleic Acids Res.*, vol. 36, no. Database issue, pp. D901–906, Jan. 2008.

[30] L. Jin, Y. Tabe, H. Lu, G. Borthakur, T. Miida, H. Kantarjian, M. Andreeff, and M. Konopleva, "Mechanisms of apoptosis induction by simultaneous inhibition of PI3K and FLT3-ITD in AML cells in the hypoxic bone marrow microenvironment," *Cancer Lett.*, vol. 329, no. 1, pp. 45–58, Feb. 2013.

[31] C. Nishioka, T. Ikezoe, J. Yang, and A. Yokoyama, “Inhibition of MEK/ERK signaling induces apoptosis of acute myelogenous leukemia cells via inhibition of eukaryotic initiation factor 4E-binding protein 1 and down-regulation of Mcl-1,” *Apoptosis Int. J. Program. Cell Death*, vol. 15, no. 7, pp. 795–804, Jul. 2010.

[32] E. K. Keeton, K. McEachern, K. S. Dillman, S. Palakurthi, Y. Cao, M. R. Grondine, S. Kaur, S. Wang, Y. Chen, A. Wu, M. Shen, F. D. Gibbons, M. L. Lamb, X. Zheng, R. M. Stone, D. J. Deangelo, L. C. Platanias, L. A. Dakin, H. Chen, P. D. Lyne, and D. Huszar, “AZD1208, a potent and selective pan-Pim kinase inhibitor, demonstrates efficacy in preclinical models of acute myeloid leukemia,” *Blood*, vol. 123, no. 6, pp. 905–913, Feb. 2014.

[33] M. H. Schaefer, J.-F. Fontaine, A. Vinayagam, P. Porras, E. E. Wanker, and M. A. Andrade-Navarro, “HIPPIE: Integrating Protein Interaction Networks with Experiment Based Quality Scores,” *PLoS ONE*, vol. 7, no. 2, p. e31826, Feb. 2012.