Neural network based prediction of less side effect causing cancer drug targets in the network of MAPK pathways

V.K. Md Aksam

GrayMatter Software Services Pvt. Ltd.,
Bangalore, 560103, India
Email: aksam.vk@graymatter.co.in

V.M. Chandrasekaran

School of Advanced Sciences,
VIT University,
Vellore, 632014, India
Email: vmcsn@vit.ac.in

Sundaramurthy Pandurangan*

Department of Basic Sciences,
Alliance College of Engineering and Design,
Alliance University,
Bangalore, 562106, India
Email: sundaramurthy.pandurangan@alliance.edu.in
*Corresponding author

Abstract: Computational side-effect prediction tools assist in rational drug design to decrease the late-stage failure of the drugs. Irrational selection of cancer drug targets in the deregulated MAPK pathways causes side effects. Network centralities and biological features – Degree, Radiality, Eccentricity, Closeness, Bridging, Stress, Pagerank centralities, essentiality, pathway-specific proteins, disease-causing proteins, protein domains are exploited quantitatively. We train an artificial neural network (ANN) with 15 selected features for the binary classification of side effects causing and less side-effect causing drug targets among the non-targeted proteins. Top ranked proteins among the Degree, Eccentricity, betweenness centralities, possessing GO-based molecular function, involved in more than one Biocarta pathways, domain content are prone to cause a number of side effects than other centralities and functional features. We predicted the following 15 less side effect causing cancer drug targets – Shc, Rap 1a, Mos, Tpl-2, PAC1, 4EBP1, GAB1, LAD, MEF2, ZAK, GADD45, TAB2, TAB1, ELK1 and SRF.

Keywords: cancer drug targets identification; network of MAPK pathways; side effects; essential proteins; graph theory.

Reference to this paper should be made as follows: Md Aksam, V.K., Chandrasekaran, V.M. and Pandurangan, S. (2021) 'Neural network based prediction of less side effect causing cancer drug targets in the network of MAPK pathways', *Int. J. Bioinformatics Research and Applications*, Vol. 17, No. 1, pp.69–79.

Biographical notes: V.K. Md Aksam is a Data Scientist, GrayMatter Software Services Pvt. Ltd., Bangalore His specialisation includes network based cancer drug target identification and data analytics. He did his PhD and MSc in VIT University and Pondicherry University, respectively.

V.M. Chandrasekaran is working as a Professor of Mathematics in the School of Advanced Sciences, Vellore Institute of Technology, Vellore. He was awarded his PhD degree by the University of Madras in 1997. He has 25 years of teaching and research experience. His current research interests include Algebra, Graph Theory and Queueing Theory. He has published more than 100 technical papers in refereed journals and conference proceedings. He is serving as an editorial board member of three international journals. He has also reviewed papers for many international journals. He has guided 10 PhD scholars.

Sundaramurthy Pandurangan is a Professor of Mathematics at the Depatment of Basic Sciences, Alliance College of Engineering and Design, Alliance University, Bangalore. His research specialisation includes network biology, systems biology, structural bioinformatics and computational genomics. He has contributed to the development of a bioinformatics web server and a software product for pharma R&D. He graduated from the Indian Institute of Technology Roorkee and mentored a PhD scholar.

This paper is a revised and expanded version of a paper entitled 'Neural network based prediction of less side effect causing cancer drug targets in the network of MAPK pathways' presented in *Inbix-2017*, BISR, Jaipur, India, 7 November, 2017.

1 Introduction

Side effects prediction is of higher importance to both industries and academics based on the identification of new drugs and new drug targets. The significant cost of R&D to identify new targets and late-stage failure of the drug due to several side effects is a growing concern (Siramshetty et al., 2016). Computational tools act successfully in the prediction of side effects at an early stage and reduce the time and cost spent on the drug development. Early side effects detection based on targets found in literature are structure-activity relationship analysis (Hammann et al., 2010), chemical structures (Bender et al., 2007), integration of chemical and biological space (Yamanishi et al., 2012) and quantitative prediction (Kotlyar et al., 2012). Post-genomic era shifts the paradigm of studies from individual gene structure to the network-based effectiveness of drugs (Kotlyar et al., 2012; Tang and Aittokallio, 2014). Acceptance of the fact of interacting proteins carrying cellular functions and inhibitive effects of the individual proteins causing effects were studied (Perez-Lopez et al., 2015). This increases various databases of side effects by considering biological networks (Duran-Frigola and Aloy, 2013). Specifically, on disease like cancer, therapeutics approach of targeting single gene

retain the connection through feedback, compensatory and redundant loops. Understanding of biological signal machinery gives effective intervention schemes (Dorel et al., 2015; Peng and Schork, 2014).

In the literature, several centrality measure of the networks are individually link side effects like Degree (Wang et al., 2013), Betweeness (Wang et al., 2013), Radiality (Tosadori et al., 2014; Isik et al., 2015), Eccentricity (Tosadori et al., 2014; Lan et al., 2010; Feng et al., 2017), Closeness (Tosadori et al., 2014; Lan et al., 2010), Bridging (Hwang et al., 2008), Stress (Tosadori et al., 2014; Isik et al., 2015) and Pagerank (Grolmusz, 2015). Drug targets involved in more number of functional features also individually linked to side effects like Biocarta pathway (Dy and Adjei, 2013; Van Wijk et al., 2015), biological process (Huang et al., 2011), cellular component (Kotlyar et al., 2012), Molecular functions (Huang et al., 2011), Interpro domain (Patil et al., 2010), OMIM diseases and Essentiality. The collection of the features individually acclaimed for the prediction of side effects and their combination can be used to predict more accurately.

In this work, we consider 15 features like centrality measures, disease, and a pathway involving activities, essentiality and biological functional roles analysed from the nodes in the network of MAPK pathways. MAPK pathways are evolutionarily conserved and their role in cancer is well known (Wagner and Nebreda, 2009). Machine learning tools like artificial neural network (ANN) and support vector machine based side-effect prediction are gaining popularity in the recent years (Dimitri and Lió, 2017). We train the ANN on the 46 known targeted proteins and predict the targets in 36 non-targeted proteins can be involved in side effects.

2 Methods

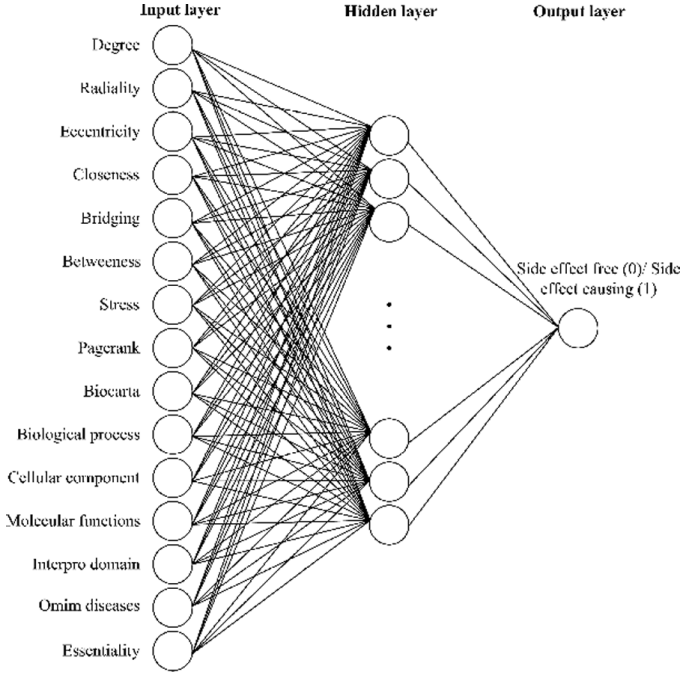
MAPK pathways are extracted from the Science STKE database (http://stke.sciencemag.org/cm/stkecm). Network centralities Betweenness, Radiality, Eccentricity, Bridging, and Stress were measured using Cytoscape Centiscape plugin Cytoscape (Shannon et al., 2003). Gephi (Bastian et al., 2009) is used to find Closeness, Degree and Pagerank centrality. Functional features collected from DAVID database of level 5 (Dennis et al., 2003). Proteins involved in more than one pathway are identified by using Biocarta. Domain feature and disease properties of the proteins are extracted from Interpro (Apweiler et al., 2001) and OMIM database (Feramisco et al., 2009) respectively. Essentiality and functional characteristics of proteins are extracted from DEG10 (Luo et al., 2013) and GO database (Gene Ontology Consortium et al., 2004) respectively. DR.PRODIS (Zhou et al., 2015) database is used on proteins in the network of MAPK pathways to isolate chronic side effects (neuro and psychiatry, cardiovascular, respiratory and diabetes) causing target. ANN analysis of topological and functional features is carried out using MATLAB 2014 with nntool (Figure 1).

We have used backpropagation algorithm along with gradient descent method to train the model (ie) to determine the optimal parameters (weights). The mathematical formula is given below and the same can be found in the method section of the manuscript. We used backpropagation algorithm for weight updating function. Each variable is adjusted according to gradient descent with momentum,

$$dX = mc * dXprev + lr * mc * dperf / dX$$

where *dXprev* is the previous change to the weight or bias.

Figure 1 ANN model to predict the less side effects causing drug targets



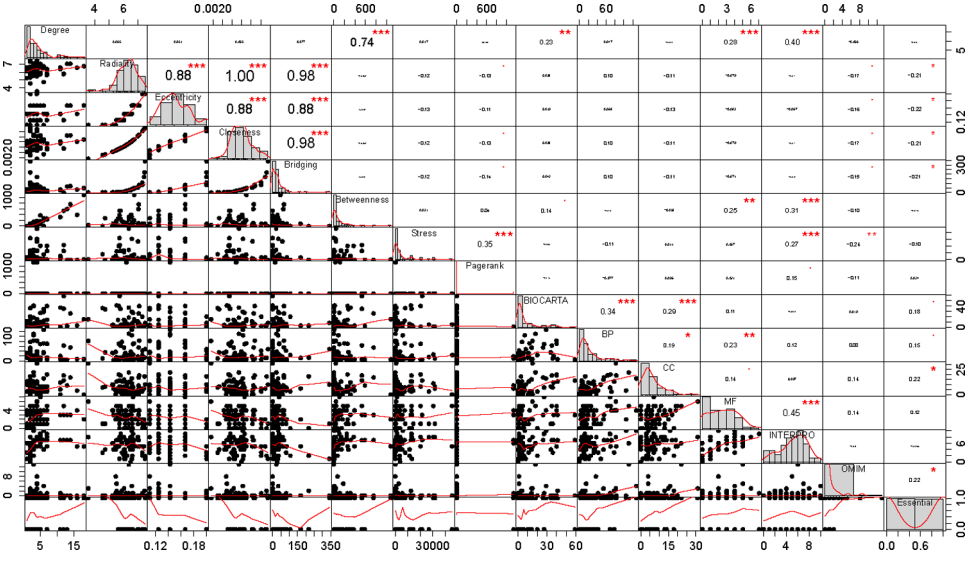
3 Results

3.1 Network of MAPK pathways analysis with the centralities, pathways, GO annotations, protein domains, OMIM diseases and essentiality

Network centralities Degree, Betweenness, Radiality, Eccentricity, Closeness, Bridging, Stress, and Pagerank using pairwise Kendall's correlation reveals three clusters of correlations. A correlation of 0.74 observed between Degree and Betweenness, followed by all pairwise combination of correlation between Radiality, Eccentricity, Closeness and Bridging as above 0.88 and correlation of 0.34 between Stress and Pagerank (Figure 2). Similar observation of pairwise Kendall's correlation were observed between eight different centralities carried out in gene regulatory network (Koschützki and Schreiber, 2008). The strong relationship between rank of centralities in the network of MAPK pathways between Degree and Betweenness shows the proteins with both properties of connectivity and functionality. Furthermore, topranked proteins in Degree and Betweenness overlap as eight proteins in both the centrality with different arrangements (bold proteins in Table 1). Xiujuan et al. (2013) found high Degree and Betweenness centrality containing target proteins causing more number of side effects. In our network, we observe 12 out of 15 top rank degree proteins, and 11 out of 15 top rank betweenness proteins found to cause chronic side effects. Takeshi et al. (2009) also reported the

oncogenes found to be the top rank degree nodes and they are less likely to be drug target due to unwanted side effects. From these pairwise centralities of both Degree and Pagerank with Biocarta pathway, Biological process, Cellular component, Molecular functions, Interpro domain, OMIM diseases and Essentiality shows very low correlations (0.4) like degree and domain content. While looking second clusters, centralities found to correlate among Radiality, Eccentricity, Closeness, and Bridging (Figure 1). Scardoni and Laudanna (2012) interpreted the three high valued centralities Radiality, Eccentricity, and Closeness to centrally influencing the regulations of other proteins in the protein-signaling network. We found all the nodes in the top rank Radiality, Bridging and Closeness centralities with different positions and few top rank Radiality nodes overlap with those three centralities. Side effects among this top rank Radiality, Bridging and Closeness centralities reveals 8 out of 15 proteins causing chronic side effects. Only 6 proteins out of 15 top Eccentricity proteins overlap with top rank Radiality, Bridging and Closeness centralities. Top Eccentricity individually found to be capable of drug target with some hub proteins of passing signal transduction (Feng et al., 2017). In our network, top rank Eccentricity proteins overlapping with Radiality, Bridging and Closeness centralities. CHOP10, HSP27 and SRF found to be side-effect free and MAPKAP-K3, MAPKAP-K2, and ELK1 found to cause one or more side effects documented in the DR.PRODIS database.

Figure 2 Kendall's correlation between the network centralities and biological features (see online version for colours)



Stress and Pagerank centrality correlates with 0.35 between them. Top rank stress centrality proteins claimed to be heavily involved in cellular processes (Scardoni and Laudanna, 2012) and top rank Pagerank claimed the mixed role of network function and drug targets (Li et al., 2016). A top-ranked Stress and Pagerank centrality node contains 9 out of 15 nodes in different position. Side effects among top Stress and Pagerank centrality proteins are 9 and 8 respectively out of 15 nodes. Side effects observed on top rank (15) Stress and Pagerank nodes are 9 and 8 respectively.

Pairwise analysis among Biocarta pathway, Biological process, Cellular component, Molecular functions, Interpro domain, OMIM diseases, and Essentiality reveals no positive correlations on a pairwise basis. Drug target creates perturbations due to the hindrance of significant overlapping pathways (Bezerianos et al., 2017) and crosstalk in the signalling network (Van Wijk et al., 2015). We observed 10 out of 15 top-ranked Biocarta pathways based involved proteins causing chronic side effects. Gene Ontology (GO) annotates hierarchical terms on the biological process, cellular component, molecular functions of the proteins (Gene Ontology Consortium et al., 2004). Top-ranked (15) biological process, cellular component, molecular functions protein's targeted effects involved in 7, 7 and 10 side effects respectively. There are studies on the targeted removal of nodes with different biological process (Lee et al., 2011), molecular functions (Vogt et al., 2014) and cellular component (Kotlyar et al., 2012; Huang et al., 2011). Multi-domain proteins as a target for drugs found to draw number side effects. In our network, we found 10 out of 15 highest domain containing proteins target causes side effects. A large frequency of multidomain proteins found in eukaryotes (Ekman et al., 2005) led to analyse domain content in the network of MAPK pathways. A correlation of 0.45 is the maximum value attained between Molecular functions and Interpro domain. The correlation is due to a maximum number of domain contents of the proteins carry out various molecular functions in the network.

Table 1 Three clusters of centralities (see online version for colours)

Degree	*Betweenness*	*Radiality*	*Eccentricity*	*Closeness*	*Bridging*	*Stress*	*Pagerank*
MKK4	MKK4	SRF	CHOP10	SRF	SRF	ASK1	MEKK2
MKK7	P38alpha	HSP27	MAPKAP-K3	HSP27	HSP27	MKP3	MKK5
P38alpha	P38beta	MAPKAP-K2	ELK1	MAPKAP-K2	MAPKAP-K2	MKP4	M3/6
P38beta	ERK1	CHOP10	MAPKAP-K2	CHOP10	ELK1	MLK3	ASK1
ERK1	ERK2	MAPKAP-K3	HSP27	MAPKAP-K3	MAPKAP-K3	MKP1	MEKK3
ERK2	MKK7	ELK1	SRF	ELK1	ATF2	ERK1	MKP3
JNK1	MKP2	c-Jun	MLK3	c-Jun	CHOP10	MKK1	MKP4
JNK3	JNK1	ATF2	MEKK4	ATF2	c-Jun	MKP2	MKK1
JNK2	JNK3	P38alpha	GCK	P38alpha	P38alpha	M3/6	MKP2
Mkk2	JNK2	P38delta	TAB2	P38delta	P38beta	p90-RSK	ERK2
MKK1	Shc	P38beta	TAB1	P38beta	P38delta	4EBP1	MSK2
MSK2	Mkk2	P38gamma	LZK	P38gamma	P38gamma	ERK2	MSK1
MSK1	MKK1	MKP1	MLTKa/b	MKP1	MKP1	MSK2	B-Raf
MKP5	MSK2	MKK3	MEKK1	MKK3	MKK3	MSK1	Rap 1a

OMIM database containing diseases involving proteins shows 7 out of 15 top diseased targeted effect causes one or more chronic side effects. Top disease involving proteins is P53 involved in Adrenal cortical carcinoma, Breast cancer, Choroid plexus papilloma, Colorectal cancer, Hepatocellular carcinoma, Histiocytoma, Li-Fraumeni syndrome, Li-Fraumeni-like syndrome, Multiplemalignancy syndrome, Nasopharyngeal carcinoma, Osteosarcoma, Pancreatic cancer and Thyroid Carcinoma. Essential targets rather than just target determines the side effects (Xiujuan et al., 2013) and essential proteins found

to cause side effects on drug targets (Liu and Altman, 2015). We found the mixed observation of essential and non-essential proteins causing side effects. Out of 41 essential proteins, only 17 found to cause one or more chronic side effects.

Figure 3 Error plot on training sets (47) with small value mean squared error (MSE)

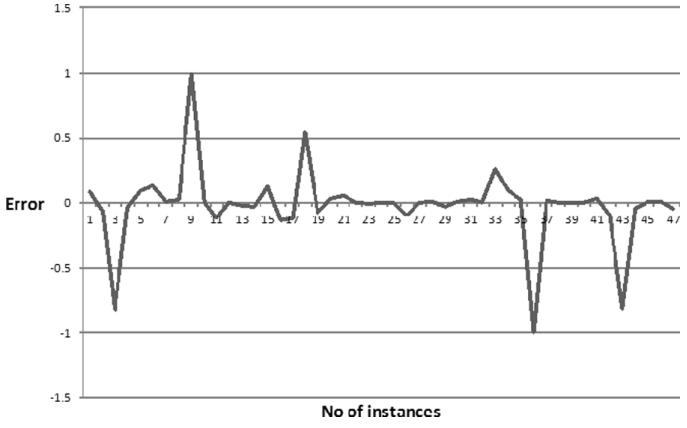
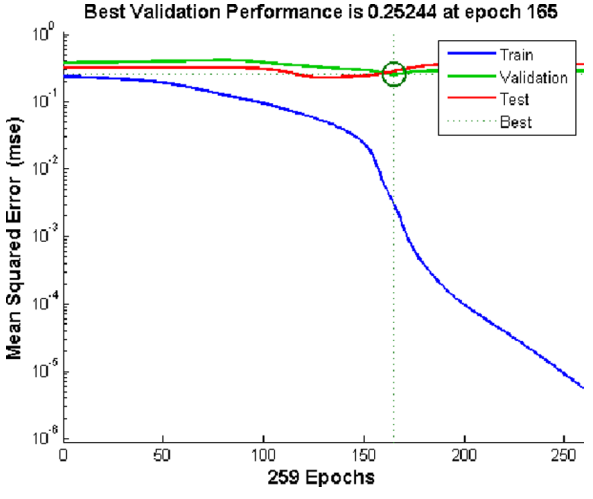


Figure 4 Test and validation data are good data points to overfitting (see online version for colours)



3.2 ANN based prediction of less side effect causing drug targets

ANN model is constructed to predict the side effects by considering network topologies – Degree, Betweenness, Radiality, Eccentricity, Closeness, Bridging, Stress and Pagerank and Functional properties – Biocarta pathway, Biological process, Cellular component, Molecular functions, Interpro domain, OMIM diseases, and Essentiality. We consider 15 collections of attributes in the input layers to learn the neural network in the MATLAB nntool. A multi-layer feed-forward loop is constructed by 10 hidden layers with different input parameters to predict the binary values of side effects (1 – side effects, 0 – no side effects). Network training function trainingDX is used which updates weights and bias

values based on gradient descent momentum along with adaptive learning rate. Repeated training is done and performance of the network to predict side effects with small value mean squared error (MSE) (Figure 3). We simulate by using non-targeted nodes to predict its side effects. To train the network, input and target vectors are randomly divided into three sets of data: 70%, 15% and 15% for training, validation, and test respectively (Figure 4). We normalised the data of small variation in the MSE values shows the performance of the network in the training set of 47 proteins (Figure 3). Most of the proteins contain the value zero indicating to reveal actual values. The predictive simulation carried out in 36 proteins and 15 side effect free non-targeted proteins Shc, Rap 1a, Mos, Tpl-2, PAC1, 4EBP1, GAB1, LAD, MEF2, ZAK, GADD45, TAB2, TAB1, ELK1, SRF as the potential cancer drug targets.

4 Conclusions

We revealed the inter-relationship among the topological and functional characteristics in the network of MAPK pathways. We use Kendall's coefficient of correlation among the 15 features and observed a very few positive correlation between them. Three clusters of relations observed are

1 Correlation of 0.74 between Degree and Betweenness

2 Correlation of nearby 0.88 and above between every pair of Radiality, Eccentricity, Closeness, and Bridging

3 Correlation of 0.34 between Stress and Pagerank.

Furthermore, among three clusters the top centrality nodes overlap within clusters. Top overlapping nodes in cluster 1 and 3 carry out dual topological roles and top overlapping nodes in cluster 2 carries four topological roles (Radiality, Eccentricity, Closeness, and Bridging). Topological analysis of the top overlapping nodes in the clusters contributes towards the complexity of the network. Each of the top centrality nodes (15) targeted effect causes more numbers of side effects (Degree – 12, Betweenness – 11, Radiality, Bridging and Closeness – 8, Eccentricity – 6, Stress – 9 and Pagerank – 8). Very low correlations were observed in the topological and functional features in our network. Functional features based pairwise analysis also reveals no positive correlations among Biocarta pathway, Biological process, Cellular component, Molecular functions, Interpro domain, OMIM diseases, and Essentiality. Top Biocarta overlapping pathway nodes (15) are causing chronic side effects in 10 protein targets. A biological process, Cellular component, Molecular functions protein's targeted effects lead to the 7, 7 and 10 chronic side effects respectively. OMIM based top disease involving seven nodes found to cause chronic side effects. Less number of essential proteins is found to involve in side effects as (41 essential proteins only 17 found to cause one or more chronic side effects. We consider both topological and functional 15 features to train the neural network to the binary prediction of side effects causing and side-effect free protein targets. Non-targeted proteins (36) are used to predict side-effect free proteins. We found side effect free non-targeted proteins (13) Shc, Rap 1a, Mos, Tpl-2, PAC1, 4EBP1, GAB1, LAD, MEF2, GADD45, TAB2, ELK1, SRF as the potential cancer drug targets.

References

Apweiler, R., Attwood, T.K., Bairoch, A., Bateman, A., Birney, E., Biswas, M., Bucher, P., Cerutti, L., Corpet, F., Croning, M.D. and Durbin, R. (2001) 'The interpro database, an integrated documentation resource for protein families, domains and functional sites', *Nucleic Acids Research*, Vol. 29, No. 1, pp.37–40.

Bastian, M., Heymann, S. and Jacomy, M. (2009) 'Gephi: an open source software for exploring and manipulating networks', *Icwsm*, Vol. 8, pp.361–362.

Bender, A., Scheiber, J., Glick, M., Davies, J.W., Azzaoui, K., Hamon, J., Urban, L., Whitebread, S. and Jenkins, J.L. (2007) 'Analysis of pharmacology data and the prediction of adverse drug reactions and off-target effects from chemical structure', *ChemMedChem*, Vol. 2, No. 6, pp.861–873.

Bezerianos, A., Dragomir, A. and Balomenos, P. (2017) 'Networks and pathways in systems pharmacology', *Computational Methods for Processing and Analysis of Biological Pathways*, Springer, pp.11–46.

Dennis, G., Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C. and Lempicki, R.A. (2003) 'David: database for annotation, visualization, and integrated discovery', *Genome Biology*, Vol. 4, No. 9, p.R60.

Dimitri, G.M. and Lió, P. (2017) 'Drugclust: a machine learning approach for drugs side effects prediction', *Computational Biology and Chemistry*, Vol. 68, pp.204–210.

Dorel, M., Barillot, E., Zinovyev, A. and Kuperstein, I. (2015) 'Network based approaches for drug response prediction and targeted therapy development in cancer', *Biochemical and Biophysical Research Communications*, Vol. 464, No. 2, pp.386–391.

Duran-Frigola, M. and Aloy, P. (2013) 'Analysis of chemical and biological features yields mechanistic insights into drug side effects', *Chemistry & Biology*, Vol. 20, No. 4, pp.594–603.

Dy, G.K. and Adjei, A.A. (2013) 'Understanding, recognizing, and managing toxicities of targeted anticancer therapies', *CA: A Cancer Journal for Clinicians*, Vol. 63, No. 4, pp.249–279.

Ekman, D., Björklund, Å.K., Frey-Skött, J. and Elofsson, A. (2005) 'Multidomain proteins in the three kingdoms of life: orphan domains and other unassigned regions', *Journal of Molecular Biology*, Vol. 348, No. 1, pp.231–243.

Feng, Y., Wang, Q. and Wang, T. (2017) 'Drug target protein-protein interaction networks: a systematic perspective', *BioMed Research International*.

Feramisco, J.D., Sadreyev, R.I., Murray, M.L., Grishin, N.V. and Tsao, H. (2009) 'Phenotypic and genotypic analyses of genetic skin disease through the online mendelian inheritance in man (omim) database', *Journal of Investigative Dermatology*, Vol. 129, No. 11, pp.2628–2636.

Gene Ontology Consortium *et al*. (2004) 'The gene ontology (go) database and informatics resource', *Nucleic Acids Research*, Vol. 32, Suppl. 1, pp.D258–D261.

Grolmusz, V.I. (2015) 'Identifying diabetes-related important protein targets with few interacting partners with the pagerank algorithm', *Royal Society Open Science*, Vol. 2, No. 4, p.140252.

Hammann, F., Gutmann, H., Vogt, N., Helma, C. and Drewe, J. (2010) 'Prediction of adverse drug reactions using decision tree modeling', *Clinical Pharmacology & Therapeutics*, Vol. 88, No. 1, pp.52–59.

Huang, L-C., Wu, X. and Chen, J.Y (2011) 'Predicting adverse side effects of drugs', *BMC Genomics*, Vol. 12, No. 5, p.S11.

Hwang, W., Kim, T., Ramanathan, M. and Zhang, A. (2008) 'Bridging centrality: graph mining from element level to group level', *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, pp.336–344.

Isik, Z., Baldow, C., Cannistraci, C.V. and Schroeder, M. (2015) 'Drug target prioritization by perturbed gene expression and network information', *Scientific Reports*, Vol. 5, p.17417.

Koschützki, D. and Schreiber, F. (2008) 'Centrality analysis methods for biological networks and their application to gene regulatory networks', *Gene Regulation and Systems Biology*, Vol. 2, p.193.

Kotlyar, M., Fortney, K. and Jurisica, I. (2012) 'Network-based characterization of drug-regulated genes, drug targets, and toxicity', *Methods*, Vol. 57, No. 4, pp.499–507.

Lan, M-Y., Chen, C-L., Lin, K-T., Lee, S-A., Yang, W-L.R., Hsu, C-N., Wu, J-C., Ho, C-Y., Lin, J-C. and Huang, C-Y.F. (2010) 'From npc therapeutic target identification to potential treatment strategy', *Molecular Cancer Therapeutics*, Vol. 9, No. 9, pp.2511–2523.

Lee, S., Lee, K.H., Song, M. and Lee, D. (2011) 'Building the process-drug–side effect network to discover the relationship between biological processes and side effects', *BMC Bioinformatics*, Vol. 12, No. 2, p.S2.

Li, X-Q., Xing, T. and Du, D. (2016) 'Identification of top-ranked proteins within a directional protein interaction network using the pagerank algorithm: applications in humans and plants', *Curr. Issues Mol. Biol*., Vol. 20, No. 1.

Liu, T. and Altman, R.B. (2015) 'Relating essential proteins to drug side-effects using canonical component analysis: a structure-based approach', *Journal of Chemical Information and Modeling*, Vol. 55, No. 7, pp.1483–1494.

Luo, H., Lin, Y., Gao, F., Zhang, C-T. and Zhang, R. (2013) 'Deg 10, an update of the database of essential genes that includes both protein-coding genes and noncoding genomic elements', *Nucleic Acids Research*, Vol. 42, No. D1, pp.D574–D580.

Patil, A., Kinoshita, K. and Nakamura, H. (2010) 'Domain distribution and intrinsic disorder in hubs in the human protein–protein interaction network', *Protein Science*, Vol. 19, No. 8, pp.1461–1468.

Peng, Q. and Schork, N.J. (2014) 'Utility of network integrity methods in therapeutic target identification', *Frontiers in Genetics*, Vol. 5.

Perez-Lopez, Á.R., Szalay, K.Z., Türei, D., Módos, D., Lenti, K., Korcsmáros, T. and Csermely, P. (2015) 'Targets of drugs are generally, and targets of drugs having side effects are specifically good spreaders of human interactome perturbations', *Scientific Reports*, Vol. 5, p.10182.

Scardoni, G. and Laudanna, C. (2012) 'Centralities based analysis of complex networks', *New Frontiers in Graph Theory*, InTech Open.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B. and Ideker, T. (2003) 'Cytoscape: a software environment for integrated models of biomolecular interaction networks', *Genome Research*, Vol. 13, No. 11, pp.2498–2504.

Siramshetty, V.B., Nickel, J., Omieczynski, C., Gohlke, B-O., Drwal, M.N. and Preissner, R. (2016) 'Withdrawn – a resource for withdrawn and discontinued drugs', *Nucleic Acids Research*, Vol. 44, No. D1, pp.D1080–D1086.

Takeshi, H., Tanaka, H., Suzuki, Y., Nakagawa, S. and Kitano, H. (2009) 'Structure of protein interaction networks and their implications on drug design', *PLoS Computational Biology*, Vol. 5, No. 10, p.e1000550.

Tang, J. and Aittokallio, T. (2014) 'Network pharmacology strategies toward multi-target anticancer therapies: from computational models to experimental design principles', *Current Pharmaceutical Design*, Vol. 20, No. 1, pp.23–36.

Tosadori, G., Spoto, F. and Laudanna, C. (2014) *Biological Network Analysis: From Topological Indexes to Biological Applications towards Personalised Medicine*, Universit`a degli Studi di Verona, Corso di Dottorato di Ricerca in Informatica, ciclo 29/2014.

Van Wijk, R., Tans, S.J., Ten Wolde, P.R. and Mashaghi, A. (2015) 'Non-monotonic dynamics and crosstalk in signaling pathways and their implications for pharmacology', *Scientific Reports*, Vol. 5, p.11376.

Vogt, I., Prinz, J., Worf, K. and Campillos, M. (2014) 'Systematic analysis of gene properties influencing organ system phenotypes in mammalian perturbations', *Bioinformatics*, Vol. 30, No. 21, pp.3093–3100.

Wagner, E.F. and Nebreda, Á.R. (2009) 'Signal integration by jnk and p38 mapk pathways in cancer development', *Nature Reviews Cancer*, Vol. 9, No. 8, pp.537–549.

Wang, X., Thijssen, B. and Yu, H. (2013) 'Target essentiality and centrality characterize drug side effects', *PLoS Computational Biology*, Vol. 9, No. 7, p.e1003119.

Xiujuan, W., Thijssen, B. and Yu, H. (2013) 'Target essentiality and centrality characterize drug side effects', *PLoS Computational Biology*, Vol. 9, No. 7, p.e1003119.

Yamanishi, Y., Pauwels, E. and Kotera, M. (2012) 'Drug side-effect prediction based on the integration of chemical and biological spaces', *Journal of Chemical Information and Modeling*, Vol. 52, No. 12, pp.3284–3292.

Zhou, H., Gao, M. and Skolnick, J. (2015) 'Comprehensive prediction of drugprotein interactions and side effects for the human proteome', *Scientific Reports*, Vol. 5.

Website

http://stke.sciencemag.org/cm/stkecm; CMP10705.