Informatics in Medicine Unlocked 9 (2017) 86–92

Contents lists available at ScienceDirect

Informatics in Medicine Unlocked

journal homepage: www.elsevier.com/locate/imu

Identification of cluster of proteins in the network of MAPK pathways as cancer drug targets

V.K. MD Aksam [a], V.M. Chandrasekaran [a], Sundaramurthy Pandurangan [b,*]

[a] School of Advanced Sciences, VIT University, Vellore 632014, India
[b] PointCross Life Sciences, Inc., Bangalore 560045, India

ARTICLE INFO

Keywords:
Network of MAPK pathways
Clusters
Drug targets identification
Drug resistance mechanism

ABSTRACT

The quest to develop computational drug target identification methods in complex diseases like cancer is growing in recent years. Feedback, feed-forward loops and cross-talks observed among the MAPK pathways led to the definition of a network of MAPK pathways and considered for single or multiple therapeutic interventions. We developed a computational method to identify clusters of drug targets by analysing the directed network's topological properties and the individual node's functional roles. We aim to identify the primary drug target nodes possessing more cancerous properties and less number of cellular functional roles. For every primary drug targets, we collect the alternate substrate activating nodes for local resistance analysis. Alternate substrate activation free nodes identified as single drug target are SOS, ATF1, BAD, GAB1, LAD, NFAT4, ATF2, MEF2, eEF2K, 4EBP1 and HSP27. Among the remaining identified nodes and their corresponding alternate substrate activating nodes with their cancer retaining and side effects causing properties studied as three different classes-single, multiple and dangerous targets. C-Raf1 and MAPKAP-K observed as a single efficient target due to the absence of resistance mechanism. Due to the resistance mechanism observed among the targeted M3/6, GADD45, and MKK6 multiple target intervention of their corresponding alternate nodes might prove to be the efficient targets. Targeted effect on MLK3, ZAK, DLK and MLTKa/b will impair the network due to intertwined and proximity nature among themselves.

1. Introduction

Effective therapeutic target strategy for the complex disease like cancer challenged in recent years due to late-stage failure in clinical trials [1]. Cancer is well known as “signaling disease”, and therapeutic inhibition of signal transduction network in human malignancies is gaining remarkable success. Intervention with the multiple drug targets is found to be more efficient than single target strategy [2–4]. Aiming to elucidate single/multiple targets in addicted signal transduction by a mutation in disease network is more complicated. Side effects are caused due to loss of functional properties of the targeted proteins. Computational approaches have attempted a systematic search of the pharmacological inhibitors playing vital role in controlling intracellular signaling event [5]. One such an approach is the development of an algorithm for multiple target optimal intervention (MTOI) in an arachidonic acid metabolic network using its structure and dynamics [2]. In this work, we develop a computational method to identify clusters of drug targets using topological and functional properties of the complex directed signaling network. Here, we confined to the network of MAPK pathways whose deregulation is the cause of cancer [6].

Understanding of cancer mechanism and the search for drug targets in MAPK pathways are dates back to 3 decades [7–9]. Disrupting the signal transduction that abnormally regulates cell growth and programmed cell death (apoptosis) are the therapeutic strategy. MAPK pathways uniquely form complex network due to the cross-talks among them [10]. Upstream and downstream of the pathways are involved in making cross-talks with each other, and there is no cross-talk observed at the MAPK level. Furthermore, drug resistance attained in MAPK pathways is due to the synergistic activation of them through cross-talks [11,12]. Tackling the problem of drug resistance in MAPK pathways [13–15] considered in this work.

We build a “Tailored” drug target identification method for the network of MAPK pathways by exploring topological and functional properties. The topological structure of the network is analysed to choose the best centrality. For instance, efficient and destructive free nodes are isolated in a topological sense. Functional and pathological properties of

* Corresponding author
E-mail addresses: mdaksam.vk@vit.ac.in (V.K. MD Aksam), vmcsn@vit.ac.in (V.M. Chandrasekaran), sundaramurthy@pointcross.com (S. Pandurangan).

http://dx.doi.org/10.1016/j.imu.2017.07.001
Received 30 May 2017; Received in revised form 3 July 2017; Accepted 5 July 2017
Available online 8 July 2017

the nodes in a network of MAPK pathways were derived from the Gene Ontology domain Biological Process (GO: BP) to cluster the proteins. Aiming to collect the nodes which are having more number of cancerous properties and less number of cellular functional roles. Causes of resistance mechanism prevailed in the network of MAPK pathways are identified and analysed to overcome them. The nodes in the clusters are analysed to determine the drug resistance mechanism acquired through the alternative activations of their substrates. Furthermore functional and cancerous properties of the identified drug target nodes and their corresponding alternate nodes are used to study the causes of retaining cancer and side effects.

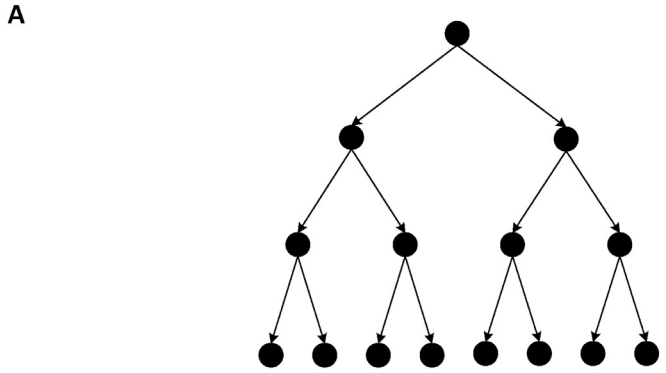
2. Results

2.1. Topology of the network of MAPK pathways

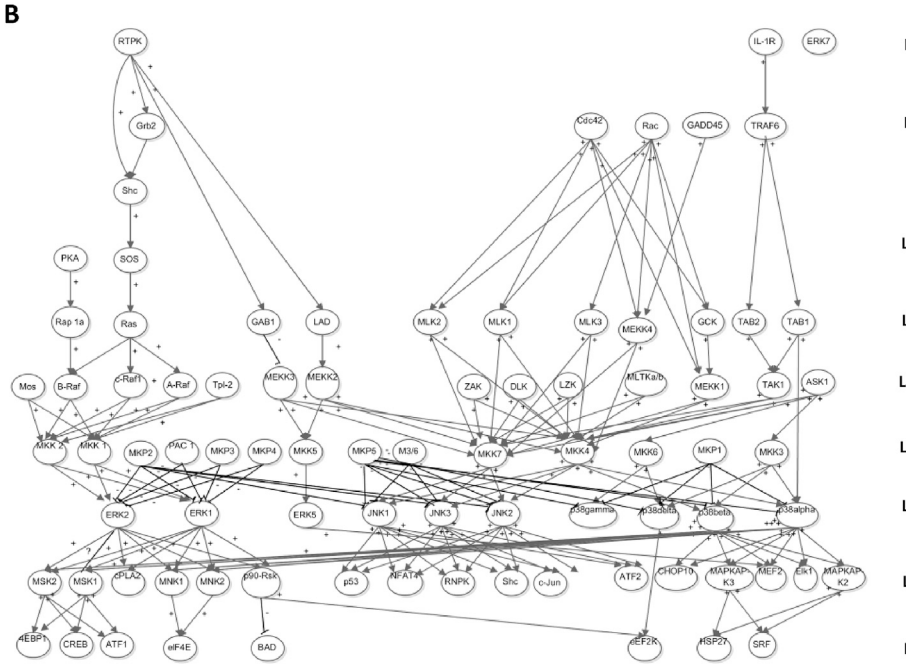
Topological properties based analysis carried out to identify the efficient and destruction free centrality nodes as drug targets. The network of MAPK pathways is directed network due to a chain of activation from receptor to transcription factors (Fig. 1). We assume the network of MAPK pathways as directed ordered network on activation time of the proteins. Directed ordered networks was first introduced by Pavel et al. [15] in the food web network by ordered nodes based on animal's body size. In this work, layer wise (9 layers) analysis carried out by considering the directed order (Fig. 1). The

Table 1
Some of the nodes observed as skipping the immediate substrate layer to activate the more further layers.

Signal events starts from node(layer) substrate node(layer)
RTPK (layer 1) substrate GAB1, LAD (layer 4)
Cdc42, Rac, GADD45, TRAF6 (layer 2) substrate MLK2, MLK1, MLK3, MEKK4, GCK, TAB2, TAB1(layer 4)
Cdc42, Rac (layer 2) substrate MEKK1(layer 5)
MLK2, MLK1, MLK3, MEKK4 (layer 4) substrate MKK7, MKK4 (layer 6)
TAB1(layer 4) substrate P38alpha(layer 7)
P38delta (layer 7) substrate eEF2K (layer 9)



Layer	Number of nodes	Average in degree	Average out degree
1	1	0	2
2	2	1	2
3	4	1	2
4	8	1	0



Layer	Number of nodes	Average in degree	Average out degree
1	2	0	2.5
2	6	1.166667	2.833333
3	2	0.5	1
4	11	1.5	1.6
5	14	0.785714	2.428571
6	14	3	3.071429
7	10	4.4	5
8	16	2.875	0.875
9	8	1.875	0

Fig. 1. A) Model hierarchal network with increasing number of nodes with 2^{n-1} order contains increasing number of nodes in the layers. Average in and out degrees are 1 and 2 respectively. B) Network of MAPK pathways found to be hierarchical up to 6 layers, converge at a 7th layer and diverges further. Average in and out degree at a 7th layer is 4.4 and 5 respectively. Network properties reveal to be small world having network diameter 9 with average number of neighbours 4.410.

Table 2
Some of the signals observed to start from intracellular environment, but are not from layer 1 (receptor).

Layer – nodes starting from layers
layer 2 - Cdc42, Rac and GADD45.
layer 3 - PKA
layer 5 - Mos, Tpl-2, ZAK, DLK, LZK, MLTKa/b and ASK1
layer 6 - MKP2, PAC1, MKP3, MKP4, MKP1, MKP5 and M3/6

Table 3
Layerwise average number of cancerous and functional properties of the nodes.

Layer	No of nodes	Average no. of cancerous property	Average no. of functional property	Sum
1	2	2.5	25	27.5
2	6	2.166667	25.5	27.66667
3	2	1	11	12
4	11	1.272727	11.36364	12.63636
5	14	0.785714	10.5	11.28571
6	14	1.142857	10.42857	11.57143
7	10	1.4	13.9	15.3
8	16	1.75	15.0625	16.8125
9	8	2	13.375	15.375

topological structure of the network reveals two properties. 1- Most of the signaling events start from the intracellular environment (Table 1), and very few events start from receptors (RTPK and IL-1R). 2- Most of the nodes skip activation of substrates in the next layer and activates the nodes in subsequent layers(Table 2). Due to the above two properties, betweenness centrality [16] may not isolate the efficient nodes. We use in and out-degree centrality in understanding the topology of the network layerwise (Fig. 1A). The number of nodes, average in and out degree found to be increasing and decreasing simultaneously. Furthermore, we also observed the network as a hierarchical tree up to 6th layer, converges at the 7th layer and diverges further(Fig. 1B). Highest average in and out degrees of the nodes found in a 7th layer with 4.4 and 5 respectively. Targeted effect on this layer will be destructive to the network topology due to their high degree nature [17]. After thorough analysis, we set the criteria of less than or equal to degree 3 (both in and out degree) for the nodes to define a cluster of destruction free nodes.

We checked our network's directed core/periphery structures like a bow-tie, which claimed as the cause of attaining network robustness [18]. Our network is far from forming the core/periphery network structures due to the presence of very few complete subgraphs (e.g., *3*-cliques).

2.2. Classification of cancer and other functions of the nodes

Gene Ontology: biological process (GO: BP) based protein functional annotation is used to evaluate the relationship among the nodes in the

Table 4
Top ten degree proteins and their associated GO terms.

Uniprot ID	Signaling Proteins	Out Degree	In Degree	Cytohubba degree	Biological Process
P27361	ERK1	6	6	12	cell cycle, macromolecule metabolic process, cellular metabolic process, primary metabolic process, interspecies interaction between organisms, regulation of biological process, regulation of cellular process.
P28482	ERK2	6	6	12	activation of immune response, regulation of immune system process, positive regulation of immune system process, system process, nitrogen compound metabolic process, transport, response to stress, cell cycle, cell communication, multicellular organismal development, behavior, response to external stimulus, response to biotic stimulus, anatomical structure morphogenesis, response to endogenous stimulus, positive regulation of metabolic process, regulation of metabolic process, regulation of localization, macromolecule localization, regulation of locomotion, positive regulation of locomotion, response to chemical stimulus, taxis, macromolecule metabolic process, cellular metabolic process, primary metabolic process, interspecies interaction between organisms, establishment of protein localization, positive regulation of biological process, negative regulation of biological process, positive regulation of cellular process, negative regulation of cellular process, regulation of response to stimulus, positive regulation of response to stimulus, anatomical structure development, regulation of biological process, regulation of developmental process, regulation of cellular process, negative regulation of developmental process, establishment of localization, cellular localization, establishment of localization in cell, response to other organism, cellular response to stimulus.
P45983	JNK1	6	5	11	ossification, cell motion, response to stress, multicellular organismal development, cell death, response to abiotic stimulus, response to chemical stimulus, macromolecule metabolic process, cellular metabolic process, primary metabolic process, positive regulation of biological process, negative regulation of biological process, positive regulation of cellular process, negative regulation of cellular process, anatomical structure development, regulation of biological process, regulation of cellular process, cellular response to stimulus.
P53779	JNK3	6	5	11	response to stress, macromolecule metabolic process, cellular metabolic process, primary metabolic process, regulation of biological process, regulation of cellular process, cellular response to stimulus.
O14733	MKK7	3	13	16	response to stress, macromolecule metabolic process, cellular metabolic process, primary metabolic process, regulation of biological process, regulation of cellular process, cellular response to stimulus.
P45984	JNK2	6	5	11	response to stress, positive regulation of metabolic process, regulation of metabolic process, response to chemical stimulus,macromolecule metabolic process, cellular metabolic process, primary metabolic process, positive regulation of biological process, positive regulation of cellular process, regulation of biological process, regulation of developmental process, regulation of cellular process, positive regulation of developmental process, cellular response to stimulus.
P45985	MKK4	5	13	18	response to stress, macromolecule metabolic process, cellular metabolic process, primary metabolic process, regulation of biological process, regulation of cellular process, cellular response to stimulus.
Q15759	P38beta	9	4	13	response to stress, macromolecule metabolic process, cellular metabolic process, primary metabolic process, regulation of biological process, regulation of cellular process.
Q16539	P38alpha	9	5	14	regulation of immune system process, alcohol metabolic process, cell motion, response to stress, multicellular organismal development,behavior, response to external stimulus, response to biotic stimulus, anatomical structure morphogenesis, positive regulation of metabolic process, regulation of metabolic process, regulation of homeostatic process, response to chemical stimulus, taxis,macromolecule metabolic process, cellular metabolic process, primary metabolic process, positive regulation of biological process, positive regulation of cellular process, anatomical structure formation involved in morphogenesis, anatomical structure development, cellular developmental process, regulation of biological process, regulation of developmental process, regulation of cellular process, positive regulation of developmental process, regulation of multicellular organismal process, response to other organism, cellular response to stimulus.

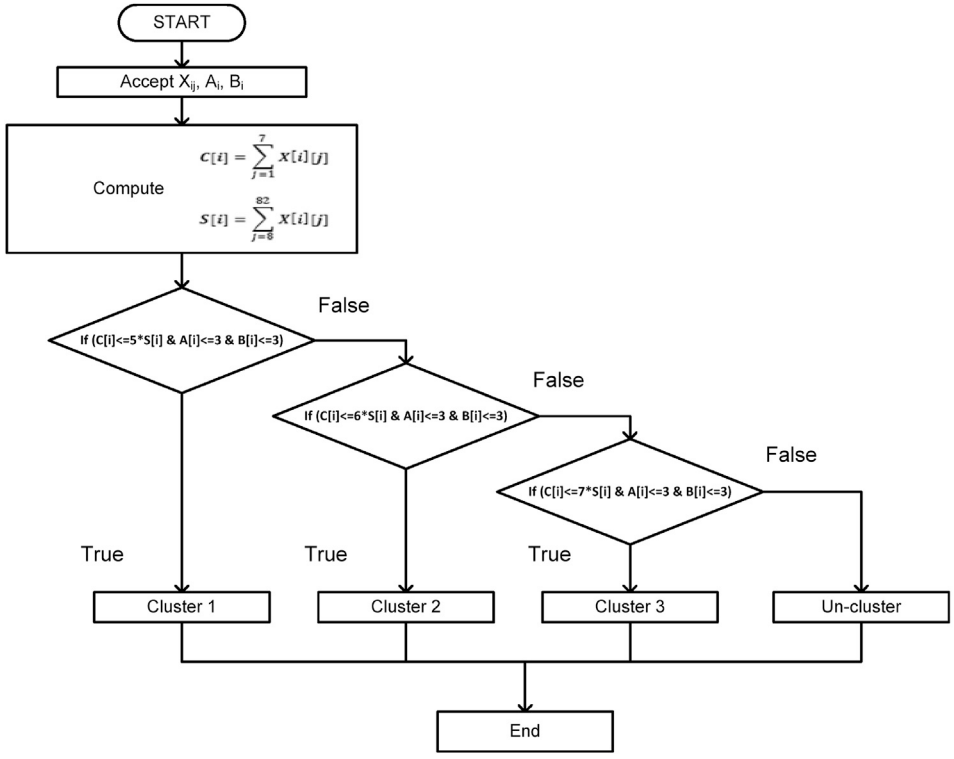
network of MAPK pathways. GO: BP annotations assigned to each of the nodes. In general, GO: BP representing protein functional annotation is used to evaluate the relationship between the sets of proteins [19]. we classify cancer-related and other functional GO: BP separately by using DAVID database level-2 for GO: BP annotations. The annotations like response to stress, cell proliferation, and response to chemical stimulus, regulation of growth, cell death, cell division and regulation of anti-apoptosis (0–6) are the key GO: BP contributing to cancerous processes. The other 75 processes (7–81) include different biological processes listed in Supp. Table - 1. The 83 nodes in the network assigned with 82 distinct GO: BP annotations and an adjacency matrix[Aij] constructed with elements 0 or 1.

$$[\mathbf{Aij}] = \begin{cases} \mathbf{1} & \textbf{\textit{for the nodes which has}} \textbf{ a cancer and functional attributes specific GO:BP} \\ \mathbf{0} & \textbf{\textit{for the nodes which doesn't}} \textbf{ have a cancer and functional attributes specific GO:BP} \end{cases}$$

Cluster-based drug target identification strategy aims to find the nodes with fewer cellular functional properties and more cancerous properties and by the ratio between them. Nine layers of the network are analysed to find the distribution of the proportion of GO: BP in each layer to determine the strategy for clustering. Average cellular functional and cancerous processes are calculated for nodes in the each layer (Table 3). In Layer 1, receptors RTPK and IL-1R possess various cellular functions and cancerous properties 25 and 2.5 respectively on an average. Molecularly targeting the intracellular elements than targeting receptor would cause fewer side effects [36]. Layer 2 with six nodes possess 2 average cancerous properties and 25 cellular functional processes. Average cancerous and functional properties of the nodes in layers 3, 4, 5 (around 1 cancerous and 11 functional properties). Nodes in the layers 7, 8 and 9 possess 13, 15 and 13 average cellular functions, and 1.4, 1.7 and 2 average cancerous properties respectively(Table 3). We set three ratios 1:5, 1:6 and 1:7 to obtain three clusters with cancerous and other functional attributes. Inhibition of nodes in a cluster possessing more cellular functional process would cause various side effects. Trade off between cancerous and other functional properties while identifying drug target is significant to eliminate or reduce side effects. In Table 3, Layerwise average number of cancerous and functional properties of the nodes reveals minimum value at level 5 as 0.785 and 10.5 respectively. So, we set nearest cut-off value as 1:5, 1:6 and 1:7. For instance in the cluster with the ratio 1:7, inhibiting node with one cancerous activity can target 7 cellular functional properties. The loss of those functional properties would contribute to the side effects which can be taken care by the alternative drugs. The nodes in the cluster with 1:5 ratios of cancer and other functional properties would be the preferable drug targets. Furthermore, Top ten degree proteins and their associated GO terms enlisted to understand the distribution (Table 4).

2.3. Topological and functional attributes based cluster identification

Integration of topological and functional attributes of the nodes in the network used for cluster identification. We set a cut-off in & out degree - 3, cancerous and other functional attributes ratio- 1: 5/1: 6/1:7. Cut-off values set by analysing 9 layers of nodes, the minimum average ratio observed between cancer and other functional attributes as 1:7. Cluster 1, 2, 3 contains the potential drug targets with the ratio of one cancerous attributes to 5, 6, 7 biological processes respectively. The topological parameter fixed with in & out degree and the functional parameter describing the three clusters can be involved fewer side effects. Flow chart, algorithm and the obtained cluster of nodes enlisted (Fig. 2) and supp. Table 2. The adjacency matrix [Aij] entries (0 or 1) summed over



Cluster	Protein name
cluster 1	SOS, c-Raf1, Eef2k, GAB1, LAD, MEF2, MLK3, MLTKa/b, MAPKAP-K3, HSP27
cluster 2	ATF1,NFAT4,GADD45,MKK6,ATF2
cluster 3	4EBP1,BAD,M3/6,ZAK,DLK

Fig. 2. Flow chart of the clustering procedure. The sum over cancerous attributes given in GO: BP up to i = 1 to 7 is stored in C[i] for the node i and other cellular functions summed as S[i]. In and out degree cut-off fixed as 3 due to simultaneous degree change in each layer of the network. Along with functional attributes ratio of 1:5, 1:6 and 1:7 are the criteria to cluster 1, 2 and 3 respectively. 83 nodes in the network of MAPK pathways clustered with 10, 5 and 5 nodes in cluster 1, 2 and 3 respectively.

Table 5
Single and multiple targets identified in the network of MAPK pathways.

Protein	Alternative protein	Single/multi-drug target
c-Raf1	Mos	single
	A-Raf	single
	B-Raf	single
	Tpl-2	single
M3/6	MKP5	Multi
	MKP2	Multi
GADD45	Cdc42	Multi
	Rac	Multi
MLK3	MEKK1	–
	MEKK2	–
	MEKK3	–
	MEKK4	–
	ASK1	–
	MLK1	–
	MLK2	–
	LZK	–
	DLK	–
	ZAK	–
	MLTKa/b	–
	TAK1	–
MLTKa/b	MEKK1	–
	MEKK2	–
	MEKK3	–
	MEKK4	–
	ASK1	–
	MLK1	–
	MLK2	–
	LZK	–
	DLK	–
	ZAK	–
	MLK3	–
	TAK1	–
ZAK	MEKK1	–
	MEKK2	–
	MEKK3	–
	MEKK4	–
	ASK1	–
	MLK1	–
	MLK2	–
	LZK	–
	DLK	–
	MLTKa/b	–
	MLK3	–
	TAK1	–
DLK	MEKK1	–
	MEKK2	–
	MEKK3	–
	MEKK4	–
	ASK1	–
	MLK1	–
	MLK2	–
	LZK	–
	MLTKa/b	–
	ZAK	–
	MLK3	–
	TAK1	–
MKK6	MKK3	Multi
MAPKAP-K3	MAPKAP-K2	single

cancerous attributes as C[i] where i = 0 to 6 and for the other cellular functions summed as S[i] where i = 7 to 81. Out of 83 nodes in the network, 10, 5 and 5 nodes are clustered in the clusters 1, 2 and 3 respectively, and other nodes remain unclustered (Fig. 2).

2.4. Local resistance analysis

The nodes in the clusters are analysed to study the drug resistance mechanism acquired through the alternate activation of them by other proteins. The nodes free from the alternate activation mechanism proposed to target by single target approach, and the other alternatively activated nodes can be targeted by combination therapy to avoid the resistance factors. Computational and experimental studies on the versatile MAPK pathways studied due to complex feedback/feed-forward and cross-talk regulations on multiple timescales [20]. Multiple concomitant activations and the overlapping cross-talks between the pathways are the key mechanisms to attain resistance [21]. Furthermore, the resistance acquired due to epigenetic and genetic aberrations can be re-enabled by inhibiting the nodes in the MAPK pathways [22,23]. Switching mechanism from B-RAF to C-RAF are the key observed mechanism [24], which leads to the local analysis of the alternate switching proteins. Nodes in the upstream and downstream of the pathways are involved in making cross-talks, and there is no cross-talk observed at the MAPK level [25].

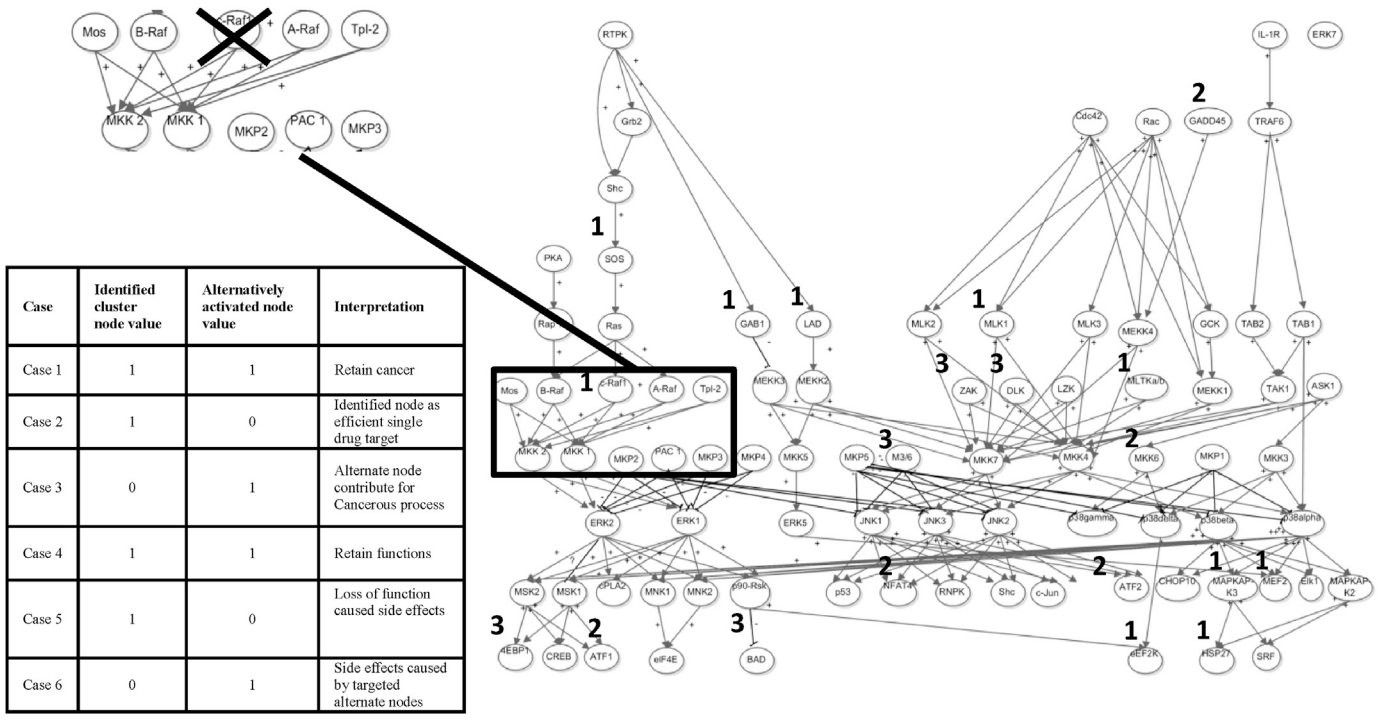
Twenty nodes in the cluster 1, 2 and 3 looked for their alternate activating substrate nodes. SOS, ATF1, BAD, GAB1, LAD, NFAT4, ATF2, MEF2, eEF2K, 4EBP1 and HSP27 are the nodes without any alternate activating nodes and can treat as a single target. c-Raf1, M3/6, GADD45, MLK3, ZAK, DLK, MLTKa/b, MKK6 and MAPKAP-K3 identify as the nine nodes with their alternate activating proteins(Table 5). Identified node's substrate activating other proteins elucidated to observe the procession of similar GO: BP annotations, which could retain the cancerous mechanism or cause side effects if targeted (Supp. Table-1). We classified six sub-cases to consider them as single/multiple targets due to their substrate activating roles(Fig. 3). Case 1- if targeting the nodes in the identified clusters, alternate activating nodes would retain the cancerous mechanism. Case 2- we can treat the cluster nodes which lack cancer causing alternate activating nodes as drug targets. Case 3- targeting only the alternate activating nodes which contribute more to the cancer mechanism than the nodes in the clusters. Case 4 – Alternate activating nodes would retain the functional properties while targeting the identified cluster nodes. Case5 - targeted cluster nodes which cause loss of functions. Case 6 – targeted alternate activating nodes which cause loss of functions. The cases 4, 5 and 6 can be used to predict the side effects caused due to the inhibition of functional process. We develop an algorithm to count the attributes falling in each of the cases to classify the nodes either as single or multiple targets (supp. algorithm 1).

2.5. Efficiency of single or multiple target solution

Eleven single target nodes SOS, ATF1, BAD, GAB1, LAD, NFAT4, ATF2, MEF2, eEF2K, 4EBP1 and HSP27 side effects are checked in Dr.PRODIS Database [28]. ATF1 is found to have chronic side effects like - heart disease (C0018799), Psychiatric Disorders (C0004936), Nervous System Disorders (C0027765) and all the other nodes are free from such chronic side effects.

While looking for multi-target strategy, alternate activating nodes also to be targeted along with the primary targets (cluster nodes) as they help to retain cancer mechanism (Fig. 3). Nine nodes c-Raf1, M3/6, GADD45, MLK3, ZAK, DLK, MLTKa/b, MKK6 and MAPKAP-K3 which poses alternatively activated pairs further analysed to considered as either single or multiple targets (Table 5). The GO: BP attributes fall under the above mentioned 6 cases counted to interpret whether the targeted nodes retain cancerous properties or causing side effects(Supp. Table 2). While targeting the B-RAF, the signaling flow from B-RAF is taken care by its alternate node C-RAF [3]. The above study leads to the local resistance analysis among the alternative switching proteins. For instance, c-Raf1 considered as the primary target, and its alternate activating nodes are Mos, A-Raf, B-Raf and Tpl-2. As per case 1 and case 3 based local resistance analysis, cancerous attributes of c-Raf1 are not found in its alternate activation nodes Mos, A-Raf, B-Raf and Tpl-2. Based on the case 2 analysis, the cancerous attributes 1-cell proliferation and 4-cell death are found in primary target c-raf1, and no such attributes possessed by alternate nodes. Inhibition of c-raf1 will be an effective drug target due to case 2 and loss of functional attributes like 36-regulation of biological process, 37-regulation of cellular process, 53-organelle organisation are collected in case5. Identified single target c-raf1 inhibition proved to be efficient in melanomas and targeting their overlapping feedback mechanism also perceived [29].

Deregulation of phosphatases results in progression and also phosphatases act as tumour suppressor [30,31]. Identified target node M3/6



Case	Identified cluster node value	Alternatively activated node value	Interpretation
Case 1	1	1	Retain cancer
Case 2	1	0	Identified node as efficient single drug target
Case 3	0	1	Alternate node contribute for Cancerous process
Case 4	1	1	Retain functions
Case 5	1	0	Loss of function caused side effects
Case 6	0	1	Side effects caused by targeted alternate nodes

Fig. 3. Local resistance analysis in the network of MAPK pathways. Primary target c-Raf1 with its alternate activating nodes Mos, A-Raf, B-Raf and Tpl-2 highlighted for local resistance analysis. Six cases of cancerous and other functional properties of primary and alternate target shown in Table.

is a dual specific phosphatase for JNK and p38 controlling the phosphorylation mechanism. MKP2 and MKP5 are the other two phosphatases alternately act along with M3/6. Local analysis on a pair of M3/6 and MKP5 shows 0- response to stress and 2- response to chemical stimulus in case 2,3 respectively. Multiple targets can be a better strategy due to inhibition of the pair M3/6 and MKP5. The other pair M3/6 and MKP2 showed 2- response to chemical stimulus in case 1 (retaining cancer). Multiple target strategy elucidated by this approach and cross complex regulatory role played by phosphatases are observed [32].

GADD45 is a node controlling p38 pathway, and their role has been found in cancer [33,34]. The alternate activating nodes of GADD45 are Cdc42 and Rac. GADD45 paired with Cdc42 would retain cancerous properties 0- response to stress and 4- cell death. Targeting them implies to 10 functional attributes of GO: BP collected in case 4. However, targeting Cdc42 alone leads to the gain of 28 attributes collected in case-6 that may cause serious side effects. Targeting the other alternate node Rac leads to same properties 0- response to stress and 4- cell death as in case-2. GADD45 paired with Cdc42 and Rac as multiple targets would cause more side effects. Both the Rac and Cdc42 are found to play a role in cancer mechanism, and their inhibition strategy discussed in the literature [35–37].

MLK3, MLTK a/b, ZAK and DLK are the clustered drug targets in MAPKKK level with proximity among themselves. All the nodes have similar 12 alternate activation nodes and among themselves in combination (Table 5). All the nodes along with their alternate nodes considered as multiple targets and their proximity made them wrong targets.

MKK3 and MKK6 are the nodes at the MAPKK level confined to p38 pathway observed to be multiple targets by local analysis. Multiple targets MKK3 and MKK6 are proved to be efficient in lung, head and neck cancers [38,39]. Further MAPKAP-K3 is found to be a single efficient target with alternate activating node MAPKAP-K2. Mammalian MAPKAP's are found to be regulating cell cycle and targeting them in cancer has been perceived [26].

3. Conclusions

We have developed a computational framework for a primary drug target identification strategy by exploring topological and functional features of the network of MAPK pathways. Furthermore, local resistance analysis is done to elucidate alternate cancer retaining nodes to be inhibited by multiple target approach. Those single or multiple targets are aimed to restore the deregulated network functions to normal condition. Mechanism-based system level analysis was carried out to overcome the resistance mechanism acquired due to the alternate substrate activating nodes. More proximal and side effects causing targets avoided by employing this method. Similar work can be done to analyse the complete network of signaling pathways regulating cancer mechanism. GO-based cellular component and molecular functional properties of the nodes used for further analysis. Also, available open source data can be used to study specific cancer types.

4. Methods

A network of MAPK pathways constructed using individual ERK1/2, ERK5, JNK, and p38 pathways curated by science STKE [40]. We considered a network of MAPK pathways as the *directed ordered network* on activation time. Definition of *directed ordered networks* was introduced by Pavel et al. [15] in the food web by ordering nodes among each other on animal's body size. Which defines ordered network as a graph having nodes and edges constituted through ordered set. For instance, a distinct pair of nodes i, j either ordered as $i < j$ or $j < i$ and follow transitive relation. The orientation of ordered nodes in the network is defined as *directed ordered network*.

Network centrality measure used in this work based on *in-degree* and *out-degree* concept. Individual node roles defined by representing the number of receiving signals and outgoing signals as in and out degree respectively. In and out-degree centrality measure are used by others in various networks to identify the most influential nodes [27].

Gene Ontology: biological process (GO: BP) [6] annotations assigned to all the nodes in the network of MAPK pathways.

Furthermore, we classify cancer-related GO: BP and other functional GO: BP separately and used DAVID database level 2 annotations [41]. Level 2 keeps up great coverage and additionally giving significant term specificity. We obtain 82 GO:BP annotations for 83 nodes (Matrix order[Aij] - 83×82).

An adjacency matrix[Aij] is formed with 0 or 1.

$$[\mathbf{Aij}] = \begin{cases} 1 & \textit{for the nodes which has} \text{ a cancer and functional attributes specific GO:BP} \\ 0 & \textit{for the nodes which doesn't} \text{ have a cancer and functional attributes specific GO:BP} \end{cases}$$

The algorithm for primary target identification is implemented using Python code. Also, local resistance analysis is coded separately as a module. Python code is added to the github (https://github.com/mdaksamvk/drug-target-identification-clustering-and-local-resistance-analysis).

Funding declaration

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest

MD Aksam VK, declares that he has no conflict of interest. V.M. Chandrasekaran declares that he has no conflict of interest. Sundaramurthy Pandurangan declares that he has no conflict of interest.

Ethical approval

This article does not contain any studies with animals performed by any of the authors.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.imu.2017.07.001.

References

[1] Jardim Denis L, Groves Eric S, Breitfeld Philip P, Kurzrock Razelle. Factors associated with failure of oncology drugs in late-stage clinical development: a systematic review. Cancer Treat Rev 2017;52:12–21.

[2] Yang Kun, Hongjun Bai, Qi Ouyang, Luhua Lai, Chao Tang. Finding multiple target optimal intervention in disease-related molecular network. Mol Syst Biol 2008;4(1):228.

[3] Lehár Joseph, Andrew Krueger S, William Avery, Adrian Heilbut M, Lisa Johansen M, Price Roydon E. Synergistic drug combinations tend to improve therapeutically relevant selectivity. Nat Biotechnol 2009;27(7):659–66.

[4] Iadevaia Sergio, Yiling Lu, Fabiana Morales C, Gordon Mills B, Prahlad Ram T. Identification of optimal drug combinations targeting cellular networks: integrating phospho-proteomics and computational network analysis. Cancer Res 2010;70(17):6704–14.

[5] Meng Hu, Liu Ying, Lai Luhua. Diverse ways of perturbing the human arachidonic acid metabolic network to control inflammation. Accounts Chem Res 2015;48(8):2242–50.

[6] Dhillon, Singh Amardeep, Suzanne Hagan, Rath O, Kolch W. MAP kinase signalling pathways in cancer. Oncogene 2007;26(22):3279–90.

[7] Avruch Joseph. MAP kinase pathways: the first twenty years. Biochimica Biophysica Acta (BBA)-Molecular Cell Res 2007;1773(8):1150–60.

[8] Kyriakis John M, Avruch Joseph. Mammalian MAPK signal transduction pathways activated by stress and inflammation: a 10-year update. Physiol Rev 2012;92(2):689–737.

[9] Margutti Simona, Laufer Stefan A. Are MAP kinases drug targets? Yes, but difficult ones. ChemMedChem 2007;2.8:1116–40.

[10] Antonia L. Pritchard, Nicholas Hayward K. Molecular pathways: mitogen-activated protein kinase pathway mutations and drug resistance. Clinic Cancer Res 2013;19(9):2301–9.

[11] Roesch A. Tumor heterogeneity and plasticity as elusive drivers for resistance to MAPK pathway inhibition in melanoma. Oncogene 2015;34(23):2951–7.

[12] Ferrarelli Leslie K. Tackling kinase inhibitor resistance. Sci Signal 2016;9.425.ec97–ec97.

[13] Fey Dirk, David Matallanas, Jens Rauch, Oleksii Rukhlenko S, Boris Kholodenko N. The complexities and versatility of the RAS-to-ERK signalling system in normal and cancer cells. In: Seminars in Cell & Developmental Biologyvol. 58. Academic Press; 2016. p. 96–107.

[14] McCubrey James A, Stephen LAbrams, Timothy Fitzgerald L, Lucio Cocco, Alberto Martelli M, Giuseppe Montalto, Melchiorre Cervello. Roles of signaling pathways in drug resistance, cancer initiating cells and cancer progression and metastasis. Adv Biol Regul 2015;57:75–101.

[15] Paulau Pavel V, Feenders Christoph, Blasius Bernd. Motif analysis in directed ordered networks and applications to food webs. Sci Rep 2015;5.

[16] Freeman LC. A set of measures of centrality based on betweenness. Sociometry 1977:35–41.

[17] Kitano Hiroaki. Biological robustness. Nat Rev Genet 2004;5(11):826–37.

[18] Csermely Peter, András London, Ling-Yun Wu, Brian Uzzi. Structure and dynamics of core/periphery networks. J Complex Netw 2013;1(2):93–123.

[19] Chagoyen Monica, Pazos Florencio. Quantifying the biological significance of gene ontology biological processes—implications for the analysis of systems-wide data. Bioinformatics 2010;26(3):378–84.

[20] Rauch Nora, Oleksii Rukhlenko S, Walter Kolch, Boris Kholodenko N. MAPK kinase signalling dynamics regulate cell fate decisions and drug resistance. Curr Opin Struct Biol 2016;41:151–8.

[21] Tortora Giampaolo, Roberto Bianco, Gennaro Daniele, Fortunato Ciardiello, James McCubrey A, Maria Rosaria Ricciardi. Overcoming resistance to molecularly targeted anticancer therapies: rational drug combinations based on EGFR and MAPK inhibition for solid tumours and haematologic malignancies. Drug Resist Updat 2007;10(3):81–100.

[22] Wellbrock Claudia. MAPK pathway inhibition in melanoma: resistance three ways. Biochem Soc Trans 2014;42(4):727–32.

[23] Margutti Simona, Laufer Stefan A. Are MAP kinases drug targets? Yes, but difficult ones. ChemMedChem 2007;2(8):1116–40.

[24] Inamdar Gajanan S, Madhunapantula SubbaRao V, Robertson Gavin P. Targeting the MAPK pathway in melanoma: why some approaches succeed and other fail. Biochem Pharmacol 2010;80(5):624–37.

[25] Fey Dirk, David Croucher R, Walter Kolch, Boris Kholodenko N. Crosstalk and signaling switches in mitogen-activated protein kinase cascades. Front Physiology 2012;3:355.

[26] Kostenko Sergiy, Alexey Shiryaev, Nancy Gerits, Ugo Moens. The roles of mammalian mitogen-activated protein kinases-activating protein kinases (MAPKAPKs) in cell cycle control. Prog Cell Cycle Control Res 2008:295–320.

[27] Yazdani Azam, Yazdani Akram, Boerwinkle Eric. A causal network analysis of the fatty acid metabolome in African-Americans reveals a critical role for palmitoleate and margarate. OMICS A J Integr Biol 2016;20(8):480–4.

[28] Zhou H, Gao M, Skolnick J. Comprehensive prediction of drug-protein interactions and side effects for the human proteome. Sci Rep 2015;5:11090. http://dx.doi.org/10.1038/srep11090. PubMed ID: 26057345; PMC4603786.

[29] Holderfield M, Nagel TE, Stuart DD. Mechanism and consequences of RAF kinase activation by small-molecule inhibitors. Br J cancer 2014;111(4):640–5.

[30] Östman Arne, Hellberg Carina, Böhmer Frank D. Protein-tyrosine phosphatases and cancer. Nat Rev Cancer 2006;6.4:307–20.

[31] Stebbing J, Lit LC, Zhang H, Darrington RS, Melaiu O, Rudraraju B. The regulatory roles of phosphatases in cancer. Oncogene 2014;33(8):939–53.

[32] Caunt Christopher J, Keyse Stephen M. Dual-specificity MAP kinase phosphatases (MKPs). Febs J 2013;280(2):489–504.

[33] Salvador Jesús M, Brown-Clay Joshua D, Fornace Jr Albert J. Gadd45 in stress signaling, cell cycle control, and apoptosis." Gadd45 Stress Sensor Genes. New York: Springer; 2013. p. 1–19.

[34] E Tamura R, de Vasconcellos JF, Devanand Sarkar, Libermann TA, Fisher PB, Zerbini LF. GADD45 proteins: central players in tumorigenesis. Curr Mol Med 2012;12(5):634–51.

[35] Hong Lin, Ray Kenney S, Phillips Genevieve K, Simpson Denise, Schroeder Chad E, Nöth Julica. Characterization of a Cdc42 protein inhibitor and its use as a molecular probe. J Biol Chem 2013;288(12):8531–43.

[36] Zins Karin, Trevor Lucas, Patrick Reichl, Dietmar Abraham, Seyedhossein Aharinejad. A Rac1/Cdc42 GTPase-specific small molecule inhibitor suppresses growth of primary human prostate cancer xenografts and prolongs survival in mice. PLoS One 2013;8(9):e74924.

[37] Prudnikova Tatiana Y, Rawat Sonali J, Chernoff Jonathan. Molecular pathways: targeting the kinase effectors of RHO-family GTPases. Clin Cancer Res 2015;21(1):24–9.

[38] Galan-Moya Eva M, Miguel A, Maria Llanos-Valero, Juan Callejas-Valera L, Pedro Melgar-Rojas. Balance between MKK6 and MKK3 mediates p38 MAPK associated resistance to cisplatin in NSCLC. PloS One 2011;6(12):e28406.

[39] Warr Nick, Pam Siggers, Gwenn-Aël Carré, Sara Wells, Andy Greenfield. Genetic analyses reveal functions for MAP2K3 and MAP2K6 in mouse testis determination. Biol Reproduction 2016;94(5):103.

[40] http://stke.sciencemag.org/cm/stkecm;CMP_10705.

[41] Huang Da Wei, Sherman Brad T, Lempicki Richard A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009;4(1):44–57.