Regulation of Signaling Networks in B Lymphocytes: Delineation of Cell Cycle Regulatory Modules

Thesis Submitted

To

Jawaharlal Nehru University

New Delhi, India,

for the Award of the Degree

of

Doctor of Philosophy

By

Noor Jailkhani

Immunology Group

INTERNATIONAL CENTRE FOR GENETIC ENGINEERING AND BIOTECHNOLOGY

NEW DELHI
INDIA

JULY 2011

INTERNATIONAL CENTRE FOR GENETIC ENGINEERING AND BIOTECHNOLOGY

ICGEB Campus, P.O. Box : 10504, Aruna Asaf Ali Marg,
New Delhi - 110 067, India
http.//www.icgeb.trieste.it

Tel : 91-11-26741358/61
91-11-26742357/60
91-11-26741007
Fax : 91-11-26742316
E-mail : icgeb@icgeb.res.in

Certificate

This is to certify that the research work embodied in the thesis entitled '**Regulation of Signaling Networks in B Lymphocytes: Delineation of Cell Cycle Regulatory Modules'** has been carried out under my supervision at the Immunology Group, International Centre for Genetic Engineering and Biotechnology, for the degree of Doctor of Philosophy. This work is original and no part of this thesis has been submitted for any other degree or diploma to any other university.

Dr. Kanury V. S. Rao
Supervisor and Head
Immunology Group
ICGEB
(International Centre for Genetic Engineering and Biotechnology)
New Delhi - 110067
India

Dr. V. S. Chauhan
Director
ICGEB
(International Centre for Genetic Engineering and Biotechnology)
New Delhi - 110067
India

Intl. Centre for Genetic Engg. & Biotechnology
DIRECTOR
N. Delhi

Declaration

I hereby declare that the research work embodied in this thesis entitled '**Regulation of Signaling Networks in B lymphocytes: Delineation of Cell Cycle Regulatory Modules'** has been carried out by me under the supervision of Dr. Kanury V. S. Rao at the Immunology Group, International Centre for Genetic Engineering and Biotechnology, New Delhi, India.

Noor Jailkhani

Dated: 14th of JULY, 2011

Noor Jailkhani
Graduate Student
Immunology Group
ICGEB
New Delhi, India

Dedicated to My Parents

Acknowledgement

All our lives, in one way or the other are undirected graphs, made up of nodes and edges. Each node represents a person, and each edge, the interaction or relationship between the two people. I am, by definition, connected to everyone in my own graph or network. Each graph has functional modules or clusters consisting of groups of nodes/people that work together to carry out different functions. My own network, therefore, has modules and nodes, some of whom have played a very central role during my PhD, while others have also contributed in their own ways. My gratitude towards each of the modules and their constituent nodes who have helped me through my PhD is briefly described below.

It has been a rich and gratifying experience to work under the guidance of Dr. Kanury V.S. Rao. Dr. Rao's very deep, almost passionate interest in Science, and his natural calm temperament have always been a great source of encouragement and inspiration. His unique approach inculcates in his students the capacity for creative and independent thinking.

I thank Dr. Pawan Sharma and Dr. V. Manivel for their encouragement and support throughout my PhD. I am thankful to Prof. V.S. Chauhan for his generous support to carry out this study. I am thankful to my doctoral committee members Dr. Sunil K. Mukherjee and Dr. Raj Bhatnagar for their valuable suggestions and constant encouragement. I am also thankful to all the faculty members of ICGEB for their support and help. I would also like to thank Dr.Riitta Lahesmaa, Director at the Turku Centre for Biotechnology for all the encouragement.

The work detailed in this thesis could not have been possible without the help of the core module consisting of my colleagues in the lab, Srikanth and Zaved Bhai, and Collaborators Shubhada Hegde from Dr. Shekhar Mande's Lab in CDFD, Hyderabad. I am indebted to Dr. Shekhar Mande for allowing me to work in his lab and also for his very valuable suggestions which have helped shape this work. They all made invaluable contributions to this project, and just as importantly taught me how collaborations within and outside the lab can make Science so much more rewarding.

I am grateful to my past lab mates Raina, Dhiraj, Manish, Ranjan Sir, Sarwar, Manjari, Viren, Rashmi, Nooshin, Azhar, Lekha, Faisal, Parul and Shilpi for initiating me into the lab. I thank my current lab members for their friendship and pleasant interactions. In particular I would like to thank Dr. Samrat, for his great company and constant encouragement specially towards the later part of my PhD; Swati, Shilpa and Parul for all the good times and conversations, scientific and often not so scientific; Varshneya for the same and also for having proof read and commented on some of my work; Ubaid, Ajay, Kamiya, Neha, Pallavi, Joyita, Shyama, Viplov, Neha A, Neha S, Ekta, Kunal, Shashank, Sumit, Nimmi, Naresh Sir and Sachin Sir for a good working atmosphere.

Acknowledgement

I am grateful for my interaction with the team at Sphaerapharma. A special thanks to Kanchan for all her support and good will. I am also thankful to Dr. Sundeep Dugar and Dr. N. Krishnamoorthy for their encouragement and interest in my work.

I am thankful to Pradeep and Ashok for their constant assistance and timely help with all the lab supplies. My flow cytometry experiments went as well as they did due to the timely help from the team at BD Biosciences, India. I would like to particularly thank Mr. Dinesh and Mr. Nagarjuna for their help. I am thankful to the administrative staff at ICGEB for their help. The financial support from CSIR is duly acknowledged.

I couldn't have made it as enjoyably through my PhD as I did without all my friends who are all bottomless wells of friendship and support. I thank them for listening to my endless stories on pretty much everything in life and for all the great times that we have had. In particular I want to thank Ankita, Ruhel, Adarsh, Vijaya, Shipra, Kaveri, Pratibha, Mini and Neha.

I owe my thanks to my seniors, friends, batch mates and colleagues from other labs - Surendra, Bhavna, Hardeep, Anup, Swarna, Sakshi, Vandana, Shashank, Vaibhao, Deepali, Kuhulika, Sanchari, Nehul, Vahid, Ravi Sir and Khalid for all the pleasant interactions and help. A special thanks to Rachna, Shikha, Prakash and Amjad for their guidance and encouragement as seniors, especially during the early parts of my PhD.

Most of all I thank my family and close friends. They have comforted me during all my lows and celebrated my successes through graduate school. I also thank them for patiently listening to all my stories. Words cannot adequately express my gratitude for the love and unwavering support of my parents. I have been lucky to have them nearby, and also to have inherited their passion for science, learning and critical thinking. They are my inspiration for all good things in life and also my best friends.

Noor Jailkhani

In addition to standard abbreviations for signaling molecules and chemical symbols, the abbreviations listed below have been used in the text:

µg	: Microgram
µl	: Microlitre
µM	: Micromolar
Ab	: Antibody
ALL	: Acute lymphoblastic leukemias
Anti-Ig	: Anti-immunoglobulin
BCR	: B-cell Receptor
BIND	: Biomolecular Interaction Network Database
BIOGRID	: Biological General Repository for Interaction Datasets
BSA	: Bovine Serum Albumin
CDK	: Cyclin Dependent Kinase
CH1	: CH1 Murine Lymphoma cell line
CI	: Combination index
CKIs	: CDK inhibitory proteins
CML	: Chronic myelogenous (or myeloid) leukemia
DMEM	: Dulbecco's Modified Eagle Medium
DMSO	: Dimethyl Sulphoxide
ECL	: Electrochemiluminescence
EDTA	: Ethylene di-amine tetra acetic acid
EGF	: Epidermal Growth Factor
EGFR	: Epidermal Growth Factor Receptor
ERK	: Extracellular signal regulated protein kinase
ESI	: Electrospray ionization
FACS	: Fluorescence-activated cell sorting
FBL	: Feeb back Loops
FCS	: Fetal Calf Serum
FDR	: False discovery rate
FFLs	: Feed Forward Loops
G1 Phase	: Gap 1 phase of the cell cycle
G1/S	: G1/S checkpoint

G1S phase	:	G1S transition phase
G2/M phase	:	Gap2 or Mitotic phase
GNF	:	Genomics Institute of the Novartis Research Foundation
h	:	Hour
HEPES	:	N-[(2-hydroxyl ethyl) piperizine-N'2-ethane sulphonic acid]
HPLC	:	High Pressure Liquid Chromatography
HPRD	:	Human protein reference database
HTS	:	High Throughput Screens
IMP nodes	:	Important nodes
IntAct	:	an open source molecular interaction database
iTRAQ	:	Isobaric tag for relative and absolute quantitation
kb	:	Kilobases
kDa	:	Kilodalton
LC	:	Liquid chromatography
M	:	Molar
mAb	:	Monoclonal Antibody
MAPK	:	Mitogen Activated Protein Kinase
mg	:	milligram
MIARE	:	Minimum Information about an RNAi Experiment
min	:	minute (s).
MINT	:	the Molecular INTeraction database
ml	:	milliliter
mM	:	millimolar
MS	:	Mass spectrometry
MS/MS	:	Tandem Mass spectrometry
MSigDB	:	The Molecular Signatures Database
nM	:	nanomolar
OTEs	:	Off target effects
PAGE	:	Polyacrylamide Gel Electrophoresis
PBS	:	Phosphate Buffered Saline
PBST	:	Phosphate Buffered Saline with 0.1% tween 20.
PDGF	:	Platelet Derived Growth Factor
PDT	:	Population doubling Time Experiments

PH(+)	:	Philadelphia chromosome positive
PI3K	:	Phosphatidylinositol 3-kinases
PKB	:	Protein Kinase B (AKT)
PMSF	:	Phenyl methyl sulphonyl fluoride
PPI	:	Protein -protein interaction
PSK	:	Phospho-serine/Threonine kinases
PSPs	:	Phospho- Ser/Thr Phosphatases
PTK	:	Phospho-tyrosine kinases
PTPs	:	Phospho- tyrosine phosphatases
RB1	:	Retinoblastoma protein 1
RNA	:	Ribonucleic acid
RNAi	:	RNA interferance
RNAse	:	Ribonuclease
RT	:	Room temperature
RTK	:	Receptor Tyrosine Kinase
RTs	:	Residence Times
SD	:	Standard Deviation
SDS	:	Sodium Dodecyl Sulphate
siRNA	:	Short Interfering RNA
SNAVI	:	Signaling Networks Analysis and Visualization tool
S-Phase	:	DNA Synthesis Phase
TEMED	:	N,N,N',N'-tetramethyl-ethane-1,2-diamine
TFs	:	Transcription factors
TiO2	:	Titanium dioxide

CONTENTS

Chapter 1: Introduction **1-4**

Chapter 2: Review of Literature **5-49**

2.1 The Cell-Signaling Machinery: Kinases and Phosphatases 6-8
2.2 The Cell Signaling Network 8-10
2.3 Network Biology 10-21
2.3.1 Applications of Graph Theory in Biological Networks 11-12
2.3.2 Structural Properties of Biological Networks 12-13
2.3.3 Diameter, Shortest Path and Mean Path Length 13-14
2.3.4 Clustering 14-14
2.3.5 Centrality Parameters 14-18
2.3.6 Motifs 18-19
2.3.7 Modules 19-20
2.3.8 Hierarchical Modularity 20-21
2.4 Shortest Path Analysis 21-22
2.5 Robustness and Sensitivity in Complex Systems 22-24
2.6 The Eukaryotic Cell Cycle 24-30
2.7 The 2-Wave theory 30-30
2.8 Dysregulation of the Cell Cycle: Cancer 31-34
2.9 RNAi Screens 34-41
2.9.1 Meta-Analysis of Genome wide RNAi screens 40-41
2.10 Multi-component Therapeutics for Networked Systems 41-47
2.10.1 Synergy: when the whole is greater than the sum of parts 45-46
2.10.2 Applications in Oncology 46-47
2.11 Cell Cycle Deregulation in B cell Lymphomas: Importance of Signaling Networks 47-49

Chapter 3: Rationale for the Study **50-54**

Chapter 4: Materials and Methods **55-79**

4.1 Materials 56-57
4.2 Buffers and Solutions 58-59
4.3 Methods 60-79
4.3.1 Cell Lines and Cell Culture Medium 60-60
4.3.2 Western Blot analysis 60-61
4.3.3 siRNA Library Screening 61-62
4.3.4 The siRNA Screening Procedure 63-64
4.3.5 Propidium Iodide Staining and Sample Acquisition using Flow Cytometry 64-64

4.3.6 Primary and Secondary Screen data analysis..... 64-66
4.3.6.1 Cell Cycle analysis..... 64-65
4.3.6.2 Screen data analysis..... 65-66
4.3.7 Population Doubling Time Experiments (PDTs)..... 66-67
4.3.8 Calculation of Residence Times..... 67-67
4.3.9 Isolation of Human Naive B cells..... 68-68
4.3.10 Pharmacological Inhibitor Assays..... 68-68
4.3.11 Sorting of cells in the G1 and S phases..... 68-69
4.3.12 Mass spectrometry based determination of the Cellular Phosphoproteome..... 69-73
4.3.13 Network analysis methods..... 73-79

Chapter 5: Results..... 80-127

5.1 An RNAi Screen targeting Cellular Kinases and Phosphatases identifies Regulators of the Cell Cycle..... 81-83
5.2 Silencing of Signaling Intermediates induces differential effects on Cell Doubling times..... 84-86
5.3 Functional relevance of the Targets identified by the siRNA Screen... 86-87
5.4 Analysis of the Phosphoproteome of cycling CH1 cells..... 87-92
5.5 Defining the Source and Target relationships for phase-specific regulation of the Cell Cycle..... 92-95
5.6 Extracting the Protein-Protein Interaction Networks that control G1 and S phases of the Cell Cycle..... 95-95
5.7 Identification of important Intermediates using the shortest path analysis..... 95-99
5.8 Delineation of IMP node-dependent Regulatory Modules..... 99-103
5.9 Extracting Cell Cycle Regulatory Modules and the identification of vulnerable Nodes..... 103-108
5.10 Verifying phase-specificity of vulnerable Nodes..... 108-112
5.11 High stress, high betweenness Nodes represent vulnerable constituents of phase-specific Regulatory Modules..... 112-114
5.12 A comparative analysis of independently conducted screens yields a common IMP node network..... 114-116
5.13 IMP node modules define the signaling axes that govern G1 and G1S windows of the Cell Cycle..... 117-118
5.14 IMP node modules capture the network of core pathways that mediate commitment of cells to the division cycle..... 118-122
5.15 The G1 and G1S IMP node modules represent conserved elements of mitogen-activated signaling cascades..... 123-127

Chapter 6: Discussion..... 128-134

Chapter 7: Summary and Conclusion..... 135-139

Chapter 8: Supplementary Information..... 140-165

Supplementary Figures 140-148

Supplementary Tables..... 149-165

Chapter 9: Bibliography..... 166-176

Chapter 10: Publications and Patents..... 177-178

Chapter 1

Introduction

Chapter 1: Introduction

Heterogeneity in both the inductive and maintenance mechanisms involved in cancers has undoubtedly complicated efforts at developing more effective regimens for chemotherapy. The multiplicity of mutational events that characterize a given cancer, which then vary significantly from one cancer to another, have rendered it difficult to identify the most appropriate targets for inhibition (Hartman, Garvik et al. 2001). In addition to such factors, however, an issue that may also require further examination is the rationale that dominates current efforts aimed at anti-cancer drug development. In such exercises, the general strategy is to identify those signaling intermediates where mutations have led to constitutive activity in a given cancer, with the aim of inhibiting them (Evan and Vousden 2001; Vermeulen, Van Bockstaele et al. 2003). The prevailing systems view, however, describes that components of the signaling machinery organize into a complex network, with multiple levels of regulation that include both feed-forward and feedback mechanisms (Barabasi and Oltvai 2004; Zhu, Gerstein et al. 2007). This therefore implies a non-linear response behaviour of the signaling network (Alon 2007), wherein the effects of inhibiting an intermediate need not necessarily be the inverse of that which is obtained upon its constitutive activation.

At least one factor influencing the outcome would be the level of functional redundancy exhibited by this intermediate, in the context of the overall network topology (Tononi, Sporns et al. 1999; Edelman and Gally 2001). That is, although activation of mitogenic pathways may derive from an oncogenic mutation in a given signaling intermediate, inhibition of this intermediate may however have only a minimal effect if its functional role is compensated for through contributions from other intermediates. This possibility is especially exacerbated in cancer cells where

mutations in more than one signaling molecules is often the norm (Vogelstein and Kinzler 2004). In other words the greater degree of plasticity associated with oncogenic pathway activation, relative to its suppression, also indicates that the ideal targets for pathway inhibition need not necessarily coincide with those that are involved in its activation. Such increments to our understanding of the complex properties of biological systems therefore illuminate that drug development efforts will be significantly aided by a better resolution of the signaling circuitry that controls cell cycle, as well as a description of the least redundant nodes (nodes which are functionally least replaceable from the context of network integrity/function) that participate in this process.

Despite the extensive accumulation of information on mitogen-activated signaling cascades, however, a clear picture on how they integrate to modulate the cell cycle program is still lacking (Papin, Hunter et al. 2005). Indeed the current challenge is to develop approaches that can distil the available information, and provide a coherent view of the core pathways that are involved (Papin, Hunter et al. 2005; del Sol, Balling et al. 2010). In the present study, we combined the results of a siRNA screen targeting the signaling machinery with extensive graph theoretical analysis to extract the core modules that processed mitogenic signal in the context of the individual phases of the cell cycle. Subsequent experiments confirmed both the functional significance of these modules and their phase-specific mode of action, whereas an examination of the component links provided novel insights into mechanisms mediating the seamless transition of cells from one phase to the next. Importantly, we could demonstrate that these modules constituted functionally conserved features of both mitogen and oncogene dependent signaling networks, and that the least redundant nodes present in them provided effective targets for chemotherapy. Further, when viewed from a

broader perspective, our results also support that multi-module targeting - wherein the temporal dimensions of a biological process are also taken into account - may provide a useful refinement to current drug development efforts.

Chapter 2

Review of Literature

Chapter 2: Review of Literature

2.1 The Cell-Signaling Machinery: Kinases and Phosphatases

The reversible phosphorylation of proteins plays a critical role in the regulation of cellular processes, with nearly 30% of the cellular proteins being phosphorylated at any given time (Cohen 2000). Coordinated coupling of biochemical reactions involving protein phosphorylation and dephosphorylation represents the hallmark of the intracellular signal transduction machinery. Distinct classes of enzymes known as kinases and phosphatases respectively drive these reactions. While phosphorylation is achieved through the action of kinases, dephosphorylation is mediated through the action of phosphatases. Protein kinases catalyze the following reaction: $Mg\text{-}ATP^{1-}$ + Protein-OH → Protein-OPO_3^{2-} + MgADP + H^+. Protein phosphatases reverse this process by catalyzing the dephosphorylation of the protein according to the following chemical reaction: Protein-OPO_3^{2-} + H2O → Protein-OH + $HOPO_3^{2-}$. The opposing activities of these two classes of enzymes maintain a fine balance of the phosphorylation state of the signaling molecules and this balance maintains cellular functions in the homeostatic state. The cell employs acute modulations of this balance in order to regulate its functions in response to various external cues. Chronic perturbations, on the other hand, fix a more stable imprint on the cell phenotype, which often causes diseases such as cancer (Downward 2003; Hennessy, Smith et al. 2005; Bradham and McClay 2006; Bollen, Gerlich et al. 2009). Alterations in activity of such signaling intermediates, either due to mutations in the corresponding genes or epigenetic modulation of their expression levels, is often the cause of many cancers. Signal transduction is the primary means by which cells coordinate their metabolic,

morphologic, and genetic responses to environmental cues. Typically, these responses are mediated through the engagement of specific cell-surface receptors that are either endowed with kinase activity, or closely associated with kinases (Lemmon and Schlessinger 2010). This then triggers a cascade of phosphorylation events through the downstream intermediates, to eventually influence the gene transcription machinery of the cell. The ability to transmit signals deriving from the external stimulus is due to the fact that phosphorylation generally alters the conformation of a protein, thereby leading to an alteration of its function (Cohen 2000). This effect could either be at the level of its enzymatic activity, or at that of its ability to interact with other proteins. Thus, protein phosphorylation is central to the process of signal transduction as it modulates the formation and/or disruption of regulatory interactions between the components of the signaling machinery (Cohen 2000). The role of kinases during signal transduction has been extensively investigated over the past several decades and the consensus view is that subsets of kinases form distinct cascades of signaling pathways. It is, therefore, not surprising that protein kinases have occupied the centre-place in our thrust to understand the mechanisms governing signal transduction. Kinases are key components of signal transduction pathways that promote cell proliferation (Downward 2003; Hennessy, Smith et al. 2005; Bradham and McClay 2006). Over the years several discrete kinase cascades have been characterized and dysregulated functioning of many of them have been implicated in cancers. Some of the more prominent of such cascades include the PI-3K/AKT, the MAP kinase, and the Wnt signaling pathways (Downward 2003; Hennessy, Smith et al. 2005; Reya and Clevers 2005; Bradham and McClay 2006). The kinase components of these pathways are being actively investigated as possible targets for cancer chemotherapy. Kinases are also the key regulators of the cell cycle and its checkpoints (Kastan and Bartek

2004). Different sets of kinases drive the progression of cells from early cell cycle stages in interphase, to the later stages of mitosis and cytokinesis .They also regulate cell cycle checkpoints, thereby maintaining the fidelity of cell division (reviewed elsewhere (Vermeulen, Van Bockstaele et al. 2003; Malumbres and Barbacid 2007; Malumbres and Barbacid 2009). Dysregulation of these kinases often leading to altered regulation of the cell cycle or its checkpoints has been found in many cancers (Vermeulen, Van Bockstaele et al. 2003; Kastan and Bartek 2004). . Inhibitors of many of these kinases are now being exploited in cancer therapy.

In contrast to this, regulation by cellular phosphatases has generally been considered to occur through isolated interactions between a given phosphatase and its target substrate. Emerging evidence, however, is beginning to suggest that phosphatases also inter-regulate each other, and that such interactions can lead to the formation of discrete phosphatase-specific cascades. The crosstalk between such regulatory axes of phosphatases and kinase cascades provides for complex modes of regulation, with non-linear signal input/output relationships.

2.2 The Cell Signaling Network

Although the signaling machinery was initially conceived to form a set of discrete, linear pathways or cascades, subsequent detailing however revealed an extensive degree of both intra- and inter-pathway crosstalk. Further, the extensive crosstalk that exists between these cascades leads to a complex network configuration for the signaling machinery. Nonetheless, a network configuration confers functional robustness to the signaling machinery, which arises from the many redundancies and feedbacks that are inherent to complex structures. This allows the systems to

dynamically adapt or compensate for losses or environmental perturbations (Barabasi and Oltvai 2004; Papin, Hunter et al. 2005).

The human cellular signaling network includes genes for over 1500 signaling receptors, and more than 500 kinases and 150 phosphatases (Manning, Whyte et al. 2002; Alonso, Sasin et al. 2004; Papin, Hunter et al. 2005; Shi 2009). In addition, the repertoire of signaling molecules is further enhanced through the use of alternate splice variants of several of these enzymes (Papin, Hunter et al. 2005). Protein kinases can be grouped into two families. The large majority of protein kinases are those that phosphorylate proteins on Ser/Thr residues (PSKs), while others phosphorylate the Tyr residues (PTKs) (Manning, Whyte et al. 2002). The number of protein phosphatases is far less and includes Ser/Thr-, Tyr-, and dual specific- (i.e. capable of dephosphorylating at Ser, Thr and Tyr residues) phosphatases. However in contrast to kinases, the tyrosine phosphatases (PTPs) are far greater in number than the Ser/Thr Phosphatases (PSPs) (Shi 2009).

The signaling network has a modular architecture, where individual modules govern distinct functions of the cell. It is now recognized that these interactions in fact represent a complex network wherein small-scale regulatory configurations such as switches, gates, feedback loops, and feed-forward motifs process information related to both duration and strength of the input signal (Barabasi and Oltvai 2004; Papin, Hunter et al. 2005). Motifs are sets of recurring regulation patterns which occur in biological networks such as transcription networks, signaling networks and neuronal networks (Alon 2007). They have been defined by Barabasi *et al.* as the elementary units of cellular networks and consist of small groups of interconnected nodes carrying out a specific information processing function (Barabasi and Oltvai 2004). A large number

of motifs have been identified in signaling networks in bacterial, plant and mammalian systems. The main classes of motifs and their functions are discussed later in the text. Modules or modularity in signaling networks have been defined by Papin *et al.* as an "intuitive grouping of reactions from an intracellular signaling network that have a related functions" and by Barabási *et al.* as a "group of functionally (or physically) linked molecules or nodes that work together to achieve distinct functions" (Barabasi and Oltvai 2004; Papin, Hunter et al. 2005). It is now generally accepted that the signaling network represents the central functional module of the cell that, in turn, is connected to the other modules that are responsible for the various phenotypic functions (Hartwell, Hopfield et al. 1999). This includes the control of processes of cell growth and cell division. It is not surprising, therefore, that the etiology of cancers frequently derives from aberrations in the functioning of the cellular signaling machinery, which arise from mutational and/or epigenetic modulation of its component nodes (Evan and Vousden 2001; Downward 2003; Kastan and Bartek 2004; Hennessy, Smith et al. 2005; Bradham and McClay 2006; Bollen, Gerlich et al. 2009) .

Mapping the complete topology of the signaling network, understanding its behaviour in quantitative terms, and resolving how this network enforces context-specific cellular responses are some of the key questions that are currently being addressed (Shiraishi, Matsuyama et al. 2010; Fitzgerald, Schoeberl et al. 2006).

2.3 Network Biology

Biologists have traditionally taken the reductionist approach to solving biological questions. Though such approaches have revealed a wealth of knowledge and information regarding individual molecules and components, the integration of such information to reveal a systems level understanding of cellular processes still remains a

challenge (Bloom 2001). To understand cellular systems, it is imperative to resolve the networks that control cellular processes, understand their topology and dynamics, and integrate experimental data from different "omics" approaches.

The emergence of high throughput techniques and their applications in cell biology have allowed the rapid and simultaneous detection of various cellular readouts at a given time (Barabasi and Oltvai 2004). The advent of such techniques has resulted in the generation of large sets of "omics" data, and has also greatly advanced the understanding of signaling mechanisms. The proteomics data from such techniques is available from many databases and has facilitated the reconstruction of signaling networks. As described by Papin *et al.*, "A network reconstruction includes a chemically accurate representation of all of the biochemical events that are occurring within a defined signaling network, and incorporates the interconnectivity and functional relationships that are inferred from experimental data. Network reconstructions provide the framework for the application of mathematical methods that can quantitatively describe the properties of signaling networks." (Papin, Hunter et al. 2005).

2.3.1 *Applications of Graph Theory in Biological networks*

In the recent past, the principles of graph theory have emerged as instrumental tools in the field of computational biochemistry and in the understanding of bio-molecular networks (Huber, Carey et al. 2007; Mason and Verwoerd 2007). Graph theory is the study of graphs, in which the pair wise relations between objects in a particular collection is studied, and represented as a collection of nodes and edges which connect them (Huber, Carey et al. 2007; Mason and Verwoerd 2007). The principles of graph theory have applications in diverse fields and have been applied widely in computer

sciences, physic, chemistry, electrical engineering and operations research and sociology among other fields.

The understanding of biological networks is helped by the ever growing field of complex networks. There are main classes of bio-molecular networks which have been well studies – metabolic networks (which consists of the networks of biochemical reactions between metabolic substrates), PPI or protein-protein interaction networks (which depict physical protein-protein interaction) and transcriptional regulatory networks (which describe the transcriptional regulation among genes) (Mason and Verwoerd 2007). In protein networks, the basic network nomenclature includes proteins which constitute the 'nodes' of the network and the interactions between them are the links or 'edges' (Barabasi and Oltvai 2004). In protein-protein interaction networks, the links represent a direct physical interaction between the proteins. In pure mathematical terms, such networks are called graphs. The translation of such protein interaction data into networks or graphs has allowed the applications of the well established principles of graph theory for the evaluation of structural properties of networks in living cells (Mason and Verwoerd 2007). Some concepts of complex networks which also find applications in PPI and other biological networks are explained below.

2.3.2 Structural properties of biological networks

Graphs may be directed or undirected. In undirected graphs there is no direction associated with the edges whereas in directed graphs there is a direction associated with the edges. Transcription regulatory networks, metabolic networks and signaling networks are generally directed where the directions of the edges depict the nature of regulation between the nodes or the direction of information flow (Mason and

Verwoerd 2007). Protein-protein interaction networks on the other hand represent the physical interactions between the set of proteins and are generally modelled as undirected graphs.

Recent advances in high throughput data collection techniques, following the post genomic era, has facilitated the generation of a lot of data for biological networks. It has added to the already existing wealth of information on various cellular processes including signal transduction mechanisms. Although an increasing number of databases provide invaluable information for biological research, the characterization of properties that arise from whole-cell function requires integrated, mathematical descriptions of the relationships between different cellular components (Albert 2005). Motivated by these developments, there has been a significant amount of work done on identifying and interpreting the key structural properties of these networks in recent years. Some basic concepts of graph theory which find applications in biological networks are discussed below.

2.3.3 Diameter, Shortest path and Mean Path Length

In networks, the distance between nodes is measured in terms of path length which is defined as the number of nodes that are passed when information flows between two nodes (Barabasi and Oltvai 2004).

The maximum distance between any two nodes is known as the graph/network diameter (Mason and Verwoerd 2007).The mean path length is the average number of nodes that are passed when travelling between any two random nodes in the network and reflects how readily information is transmitted in the system (Barabasi and Oltvai 2004). There may be many possible paths between two nodes in a network which may or may not have the same path length. The shortest path between two nodes is the path

with the least number of links/nodes between two given nodes (Barabasi and Oltvai 2004; Mason and Verwoerd 2007).

Studies on most biological networks have shown that the mean shortest path lengths and network diameter are small as compared to the size of the network (nodes and links). Such networks are known as small world networks and show small world properties i.e. most nodes are connected by at least one path and though most nodes are not neighbours, they can be reached by crossing a small number of intermediates (Barabasi and Oltvai 2004; Mason and Verwoerd 2007). Such small world properties of biological networks highlight their efficiency in information flow.

2.3.4 Clustering

Small world networks have another characteristic property of forming cliques or groups of nodes which are highly connected to each other. In such highly clustered networks there is a high probability that neighbours of a given node are also connected. The clustering properties of a network can be understood by calculating the clustering co-efficient, which gives the average clustering of nodes in a network and this gives insight into the network structure (Barabasi and Oltvai 2004; Mason and Verwoerd 2007). Clustering coefficient and related functions are independent of network size. Networks are often classified based on mean path length and clustering coefficient (Barabasi and Oltvai 2004; Mason and Verwoerd 2007).

2.3.5 Centrality Parameters

The main aim of studying cellular networks is not only to get a systems level understanding of the biological function or disease but also to identify sets of disease genes and how they are regulated. Centrality parameters are used to determine the

importance of a node in a network or graph. Different centralities address different features of the node and have been frequently used in cellular network analysis for characterizing the functional significance of a node in a given network. Some centrality parameters which are being used in cellular networks are discussed below-

Degree Centrality:

The most elementary characteristic of a node is its degree (or connectivity) ,k, which tells us how many links the node has to other nodes (Barabasi and Oltvai 2004; Mason and Verwoerd 2007). In an undirected network the degree of a node is represented by the total number of direct links it has to other nodes. In the case of directed graphs, a node can have an incoming degree which represents the number of directed edges coming into the node; and an outgoing degree representing the number of directed edges going out from the node. A node with high degree is better connected in the network and therefore may play a more important role in maintaining the network structure (Barabasi and Oltvai 2004; Mason and Verwoerd 2007).

Significance studies by various groups have observed that node degree and essentiality may be related indicating that important nodes are involved in a large number of interactions (Jeong, Mason et al. 2001; He and Zhang 2006). Though this may be true, direct relations between degree and centrality have to be treated with caution since a lot of the studies have been carried out on published data that may be biased towards well studied proteins. Therefore low degree centrality of a protein may result from the fact that not a lot of its interactions have been well studied yet.

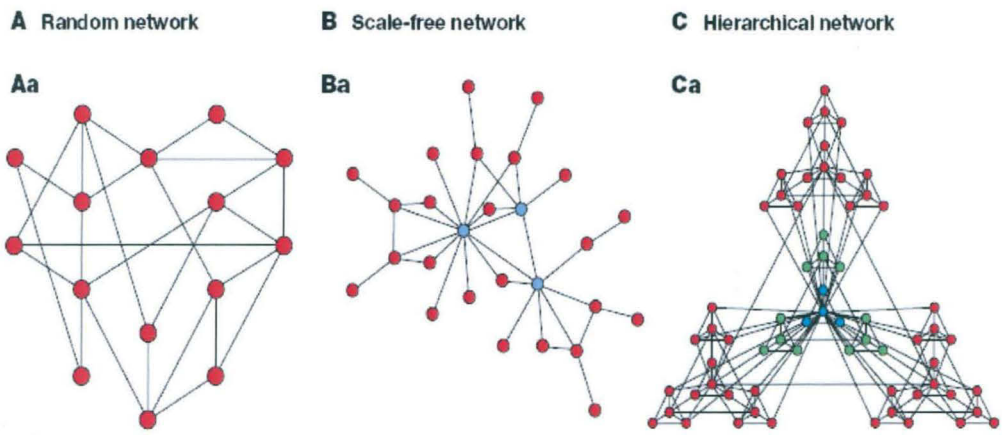


Figure 2.1 Network Models. *There are three models that had a direct impact on our understanding of biological networks. (A) Random Networks: The node degrees follow a Poisson distribution which indicates that most nodes have approximately the same number of links (close to the average degree <k>). (B) Scale-free networks: are characterized by a power-law degree distribution; The probability that a node is highly connected is statistically more significant than in a random graph, the network's properties often being determined by a relatively small number of highly connected nodes that are known as hubs. (C) To account for the coexistence of modularity, local clustering and scale-free topology in many real systems it has to be assumed that clusters combine in an iterative manner, generating a hierarchical network represented above.(Source :Barabasi and Oltvai 2004).*

The *degree distribution* is the probability distribution of the degrees over the entire graph (Barabasi and Oltvai 2004; Mason and Verwoerd 2007). Based on degree distribution, networks can be of different types (Fig. 2.1). For examples, in most biological networks the degree distribution follows the power law and such networks are known as scale free networks (Barabasi and Oltvai 2004; Mason and Verwoerd 2007). As opposed to random networks which follow the Poisson or Gaussian degree distribution where the degree of most nodes is close to the mean node degree, in scale free networks most nodes have a low degree whereas a significant number of nodes have a an unusually high degree (compared to the mean degree) and are called *hubs* (Barabasi and Oltvai 2004; Mason and Verwoerd 2007) . In the case of protein-protein interaction networks, the degree of a protein indicates the number of proteins it

interacts with, and a hub protein is one which interacts with a large number of proteins (Fig 2.2).

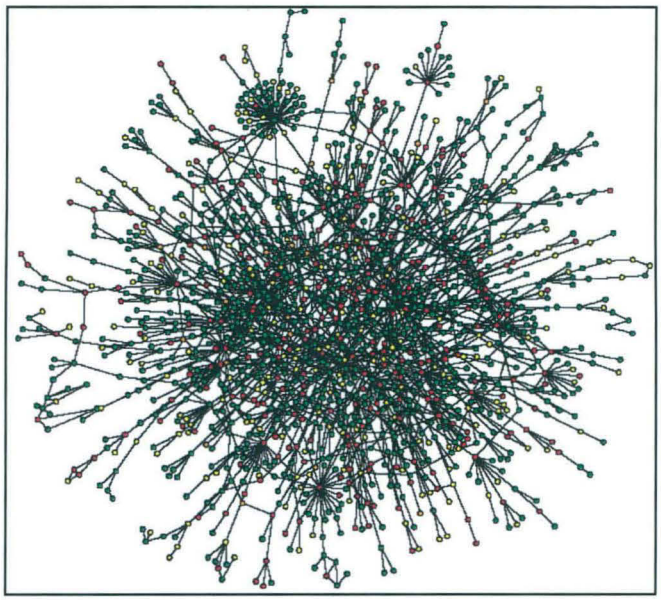


Figure 2.2 . Yeast protein interaction network. *A map of protein–protein interactions in Saccharomyces cerevisiae, which is based on early yeast two-hybrid measurements, illustrates that a few highly connected nodes (which are also known as hubs) hold the network together. The largest cluster, which contains ~78% of all proteins, is shown. The colour of a node indicates the phenotypic effect of removing the corresponding protein (red = lethal, green = non-lethal,orange = slow growth, yellow = unknown). (Source :Barabasi and Oltvai 2004).*

Closeness centrality

The closeness centrality of a node is based on the shortest path of this node to another node (Mason and Verwoerd 2007).

Betweenness centrality and Stress centrality

Betweeness centrality is a measure of the influence that a node has over the spread of information through the network. It follows that an important node will lie between many paths between nodes in a network (Mason and Verwoerd 2007). It is based on the number of shortest paths going through a specific node. It measures the extent to which a node lies in between other nodes, which is calculated as a fraction of the

shortest paths between node pairs that pass through the node of interest (Chen, Wang et al. 2009).

A similar centrality index is the 'stress centrality". Stress centrality defines the ability of a node to hold together communicating nodes, and a "stressed" node is one that is traversed by a high number of shortest paths (Manimaran, Hegde et al. 2009). Consequently, the higher the stress of a node the greater is its involvement in regulatory processes.

Betweenness, on the other hand, is a more elaborate centrality index that describes the capacity of a node to serve as a junction in the network and, thereby, regulate the network in a coherent manner (Joy, Brock et al. 2005; Mason and Verwoerd 2007; Manimaran, Hegde et al. 2009). Thus stress and betweenness represent complementary indices that together describe the functional importance of a node in a regulatory module.

2.3.6 Motifs

Network motifs are believed to be the principle building block of complex networks. They are patterns of interconnections occurring in complex networks at numbers that are significantly higher than those in randomized networks (Barabasi and Oltvai 2004). In biological networks motifs are sets of recurring regulation patterns which occur in biological networks such as transcription networks, signaling networks and neuronal networks and are likely to be associated with an optimized biological function (Barabasi and Oltvai 2004; Alon 2007). They have been defined by Barabasi *et al* as the elementary units of cellular networks and consist of small groups of interconnected nodes carrying out a specific information- processing function (Barabasi and Oltvai 2004). A large number of motifs have been identified in signaling networks in

bacterial, plant and mammalian systems. The main classes of network motifs in biological networks and their biological functions have been reviewed elsewhere (Barabasi and Oltvai 2004; Papin, Hunter et al. 2005; Alon 2007; Shoval and Alon 2010). Motifs can be of different types and believed to perform different signal processing functions within the network. Real networks are characterised by their distinct set of motifs. The high degree of conservation of motifs observed in yeast PPI networks and transcriptional networks of other species indicate that they are indeed of biological significance. In real networks, motifs tend to overlap and cluster together to form "motif clusters" or connected sub graphs (Barabasi and Oltvai 2004).

2.3.7 Modules

Cellular functions cannot be attributed to single molecules and are carried out by groups of molecules in a modular manner (Barabasi and Oltvai 2004). In general, modularity refers to a group of physically or functionally linked molecules (nodes) that work together to achieve a (relatively) distinct function (Hartwell, Hopfield et al. 1999; Barabasi and Oltvai 2004). Modularity is the hallmark of biology where groups of molecules work together to perform specific cellular functions (Hartwell, Hopfield et al. 1999; Barabasi and Oltvai 2004). For example components of a modules can be structural such as proteins which are part of the ribosome while others can be temporal i.e. molecules working together at a specific stage of a temporal process such as signal transduction or cell cycle. The modularity in a network is reflected by its clustering co-efficient and most real networks have a larger clustering co-efficient than random networks. Most protein-protein interaction networks also have a high clustering co-efficient indicating their modular topology.

In the context of cellular networks, these three modules are believed to be related and/or overlapping. It is thought that since the nodes in a topological module are closely related, they may be involved in a specific cellular function as well. And any disruption of the functional module may result in a diseased state; therefore this may also represent a disease module. Therefore it is important to note that disease modules overlap with topological and/or functional modules, each disease may have its own unique module and multiple disease modules may overlap.

Various network based approaches are now being used to identify disease modules to aid the understanding of the molecular basis of the disease and also aid drug development efforts (Barabasi, Gulbahce et al. 2011).

2.4 Shortest path analysis

Finding the shortest path to any target is of interest to all of us. The concept that information flows through the shortest paths has evolved from studies on different kinds of complex networks. The flow of information is intrinsic to any complex system. It appears that in the case of complex networks, finding the shortest path between two given nodes requires knowledge about the networks global topology, following which finding the shortest path would merely be a matter of distributed computation. However studies have shown that this is not true for random scale free networks, where information flows through the shortest path even without global topology knowledge (Boguna and Krioukov 2009). This navigation technique is called 'greedy routing'. The concept of 'greedy routing' is based on the conjecture of hidden metric spaces that underlie real networks (Boguna and Krioukov 2009). In greedy routing nodes pass information to the node that is closest to the target node in the abstract space called matrix space. This form of information flow or navigation has

been well studied in small world social networks. The concepts of shortest paths are also finding vital applications in cellular network. Many groups of researchers have used shortest paths to extract critical information from cellular networks (Martinez-Antonio and Collado-Vides 2003; Pereira, An et al. 2009; Jamal, Ravichandran et al. 2010).

2.5 Robustness and Sensitivity in complex Systems

The phenomenon of robustness versus sensitivity is paradoxical in any complex network. A key feature of many complex systems is their robustness. Robustness is the system's ability to keep functioning despite perturbations. Robustness is coupled with fragility toward non-trivial rearrangements of the connections between the system's internal parts (Jeong, Mason et al. 2001; del Sol, Fujihashi et al. 2006). Many complex systems display a surprising degree of tolerance against errors. For example, relatively simple organisms grow, persist and reproduce despite drastic pharmaceutical or environmental interventions, an error tolerance attributed to the robustness of the underlying metabolic network. Complex communication networks display a surprising degree of robustness: although key components regularly malfunction, local failures rarely lead to the loss of the global information-carrying ability of the network. The stability of these and other complex systems is often attributed to the redundant wiring of the functional web defined by the systems' components (Albert, Jeong et al. 2000).

Such error tolerance is not shared by all redundant systems: it is displayed only by a class of homogeneously wired networks, called scale-free networks including cellular networks. Such networks display an unexpected degree of robustness, the ability of their nodes to communicate being unaffected even by unrealistically high failure rates. Scale-free networks display a surprisingly high degree of tolerance against random

failures, a property not shared by their exponential counterparts (Albert, Jeong et al. 2000). This robustness is probably the basis of the error tolerance of many complex systems, ranging from cells to distributed communication systems.

However, error tolerance comes at a high price in that these networks are extremely vulnerable to attacks (that is, to the selection and removal of a few nodes that play a vital role in maintaining the network's connectivity). The error tolerance comes at the expense of attack survivability: the diameter of these networks increases rapidly and they break into many isolated fragments when the most connected nodes are targeted (Albert, Jeong et al. 2000). Such decreased attack survivability is useful for drug design.

Functional characteristics of some complex systems and their network topologies are better understood in terms of the system's need to respond sensitively to external change by switching from one mode of behaviour to another. This requirement is apparent in the case of biochemical signaling and metabolic networks, whose role is to facilitate the response of cellular systems to external stimuli or to changes in the availability of resources. Responsiveness implies an ability to adjust, perhaps dramatically, even to small environmental changes (Bar-Yam and Epstein 2004). Complementing the importance of effective response to environmental change when appropriate is the need for robustness to many other possible alterations. Robustness entails a lack of sensitivity to environmental variations, retaining the same behaviour even when subject to large stimuli. Both properties are necessary for effective reaction and adaptation to environmental changes (Bar-Yam and Epstein 2004). The response of a system can be understood to be propagated through the network of connections, where the initial stimulus affects one or more nodes. Thus, the topology provides direct

information about the nature of the response. It is natural in this context to characterize the size of a stimulus by the number of nodes that it initially affects and the response by the number that subsequently changes state. Indeed, there is ultimately no reason to expect that the size of the response should be determined solely or even primarily by the size of the stimulus rather than by its specific relationship to the network topology. The complementary importance of robustness and sensitivity suggests that we re-examine previous analyses, which showed, with respect to removal of nodes from the network, that scale-free network topologies are robust to failure and sensitive to attack (Albert, Jeong et al. 2000; Callaway, Newman et al. 2000; Cohen 2000; Jeong, Mason et al. 2001). Assuming, as is often done, that the former is an advantage, whereas the latter is a disadvantage, may be misleading in some contexts, because both characteristics can be advantageous if the sensitivity enables the system to respond effectively to environmental changes.

2.6 The Eukaryotic Cell Cycle

In any multi-cellular organism, the balance between cell division and cell death maintains a constant cell number. Both cell division cycle and cell death are highly regulated events. Whether the cell will proceed through the cycle or not, depends upon whether the conditions required at the checkpoints during the cycle are fulfilled. In higher eukaryotic cells, such as mammalian cells, signals that arrest the cycle usually act at a G1 checkpoint. Cells that pass this restriction point are committed to complete the cycle. Regulation of the G1 phase of the cell cycle is extremely complex and involves many different families of proteins such as retinoblastoma family, cyclin dependent kinases, cyclins, and cyclin kinase inhibitors.

The cell cycle is an ordered set of events leading to the division of a cell into two daughter cells. It consists of interphase which includes the G1 (Gap 1 phase), S phase (DNA synthetic phase) and G2 (Gap 2 phase) and Mitosis or M phase. Progression of the cycle from one phase to the next is controlled by certain key cell cycle regulators such as the cyclins and CDKs which drive the cycle through regulatory 'checkpoints' (Johnson and Walker 1999). The cell cycle checkpoints are control mechanisms that ensure that the cell has accurately completed a particular phase before progressing to the next. Two main checkpoints exist: the G1/S checkpoint and the G2/M checkpoint. G1/S transition is a rate-limiting step in the cell cycle and is also known as restriction point. The G1-S check point controls the passage of cells from the G1 into the S phase and is one of the most tightly regulated process after which replication begins. This stage is particularly sensitive to extra-cellular signals and acts as the signal sensing component of the cell cycle (Donjerkovic and Scott 2000; Sherr and Roberts 2004) (Fig 2.3). Studies have shown that the G1 to S interval or the G1/S checkpoint is the only part of the cell cycle which is dependent on growth factor mediated signaling (Sherr and Roberts 2004). If cells growing in culture are deprived of growth factors before the G1/S checkpoint is crossed, they arrest in the G1 phase and exit cycle without replicating their DNA. If however this occurs after the cells have crossed the G1/S checkpoint, they continue to cycle till they complete G2/M phase and then arrest in the subsequent G1 phase (Matsushime, Quelle et al. 1994; Sherr 1994; Sherr and Roberts 2004).

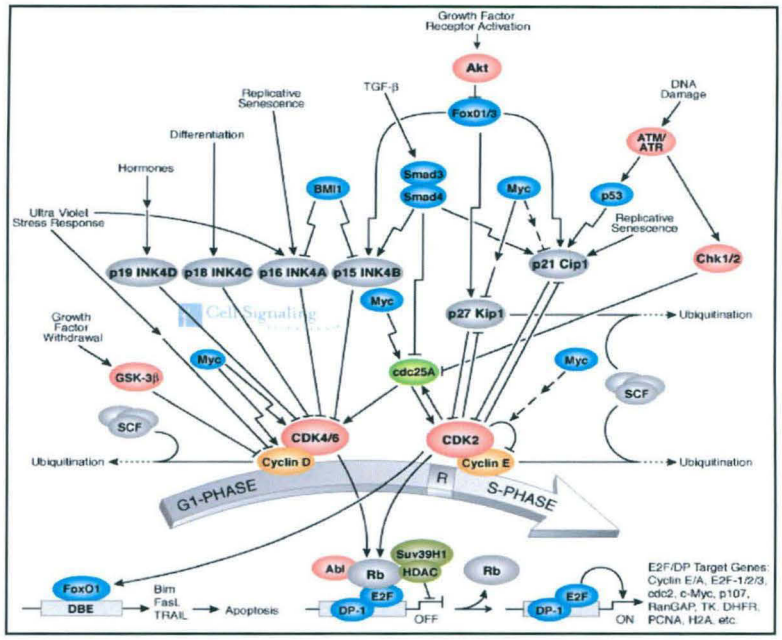


Figure 2.3. The G1/S Checkpoint control: *The figure depicts the regulation of the G1/S checkpoint by multiple mechanisms. (Source :www.cellsignal.com).*

The progression of the G1 phase of the mammalian cell cycle is regulated by the mammalian G1 cyclins and their associated kinases which integrate information flow from outside the cell and in response to mitogens drive the G1 phase and initiate DNA replication (Sherr and Roberts 2004). The G1 cyclins include the D-Type and the E-type cyclins. The D-type cyclins include cyclins D1, D2 and D3 which bind to either Cdk4 or Cdk6 in different combinations and allosterically regulate them. In an analogous fashion, the E-Type cyclins, E1 and E2, regulate the activity of cdk2. Different combinations of D type cyclins are expressed in different cell types, whereas cyclin E-Cdk2 complexes are ubiquitously expressed. These cyclin-Cdk complexes had been thought to be necessary and "rate-limiting" for entry into and progression through the G1 phase of the cell cycle (Sherr and Roberts 2004).

The activation of signal transduction pathways in response to mitogens promotes the activation of cyclin D-Cdk complexes in the early part of G1. Activation of these

complexes happens at many levels including gene transcription , cyclin-D translocation and stability, assembly of cyclin-Cdk partners and import of complexes to the nucleus where they can phosphorylate their substrates (Sherr and Roberts 1999). The mitogenic stimulation of cells arrested in the G0 stage of the cell cycle results in active accumulation of these complexes and progression into the G1 phase. These complexes are not required after the initiation of DNA synthesis in S phase, until the cell enters the next G1 phase (Matsushime, Roussel et al. 1991). Mitogen withdrawal from actively cycling cells negatively affects the transcription, translocation, assembly and import cyclin D-Cdk complexes. If this occurs before a cell crosses the G1/S checkpoint, the cell arrests in the G1 phase (Sherr and Roberts 2004). If however, this occurs when the cells have crossed the G1/S checkpoint; the cell completes mitosis and exits at the next G1. Therefore the D type cyclins act as growth sensors, providing a link between the mitogenic cues and the potentially autonomous cell cycle machinery. The activity of cyclin E-Cdk2 is periodic and maximal at the G1/S transition. Their expression is also partly mitogen dependent (Sherr and Roberts 2004).

Following mitogen dependent activation, CycD-Cdk complexes promote cell cycle progression by inactivating two classes of cell cycle inhibitors -

- Negative regulators of S phase gene expression (including the pocket protein family members Rb, p107 and p130)
- CDK inhibitory proteins (CKIs)

In the early part of G1, cyclin dependent kinases initiate Rb phosphorylation (Dowdy, Hinds et al. 1993; Ewen, Sluss et al. 1993; Sherr and Roberts 1995). Following which, in late G1 phase, cyclin E-Cdk2 complexes preferentially phosphorylate Rb on additional but distinct sites (Reed 1996; Donjerkovic and Scott 2000). Cyclin A and B

dependent cdks activated later during cell division, maintain Rb in a hyperphosphorylated form till cells exit mitosis. Rb returns to its hypophosphorylated or active state in the subsequent G1 phase. In its active state Rb remains bound to E2F family of transcription factors. The hyperphosphorylation of Rb due to its sequential phosphorylation, disrupts its association with E2F family members, histone deacetylases and other chromatic remodelling proteins (Donjerkovic and Scott 2000; Sherr and Roberts 2004). This activates the E2F mediated transcriptional program and allows DNA synthesis to occur. Both cyclin A and E are E2F responsive genes and are required for G1/S transition (Donjerkovic and Scott 2000).

A second, non catalytic role of G1 cyclins is regulating CDK inhibitors or CKIs. There are two main classes of CDK inhibitors based on structure and the CDK target (Sherr and Roberts 1999). These are the INK4A proteins consisting of p16, p15, p18 and p19 and the Cip/Kip family of proteins including p21, p27 and p57. The INK4A family members bind only to CDK4/6 and to no other CDKs or cyclins. The Cip/Kip family of inhibitors affects the activities of Cyclin D, E and A dependent kinases. They act as inhibitors of Cdk2 and positive regulators of cyclin D dependent Kinases, Cdk4 and Cdk6. The INK4A proteins act by antagonizing the assembly of cyclin D dependent kinases by binding to CDK4 (or CDK6) (Sherr and Roberts 1999). On the other hand, the members of Cip/Kip family, namely p21 and p27 are potent inhibitors of Cdk2 and are sequestered by Cyclin D-Cdk4/6 complexes. Their binding to cyclin D-Cdk4 complexes, stabilizes the complex and facilitates its nuclear import without inhibiting cyclin D associated kinase activity (Sherr and Roberts 1999). High amounts of INK4 results in formation of INK4-CDK4/6 complexes and destabilization of cyclin D. This results in the release of Cip/Kip proteins, which bind and inhibit cyclin E (and A)

dependent CDK2 activity (Sherr and Roberts 1999; Donjerkovic and Scott 2000; Sherr and Roberts 2004).

In resting or quiescent cells, the levels of p27 are relatively high whereas p21 levels are low and increase in response to mitogenic signals during G1 phase progression (Sherr and Roberts 1999). Titration of unbound p27 and p21 into higher order complexes with assembling cyclin D –dependent kinases relieves the Cyclin E-Cdk2 complexes from Cip/Kip, thereby facilitating cyclin E-cdk2 activation in late G1 (Sherr and Roberts 1999). Unlike cyclin D- Cdk complexes, cyclin E-Cdk2 does not sequester p27, but phosphorylates it on Thr 187 (Sheaff, Groudine et al. 1997; Vlach, Hennecke et al. 1997). Once sequestered to these complexes, the levels of the CKIs falls below a critical point allowing cyclin E-Cdk2 to facilitate its own activation by phosphorylating p27 and marking it for degradation. Mitogenic withdrawal leads to disassembly of these complexes and mobilizes the latent cyclin D-Cdk bound pool of p27 to bind to cyclin E –Cdk2 and facilitate cell cycle exit. This is known as the titration model (Sherr and Roberts 2004).

Cyclin E and A are E2F responsive genes and their synthesis increases after phosphorylation of Rb by Cyclin dependent kinases. Their activities maintain p27 at low levels in the S phase through phosphorylation triggered proteolysis. This emergence of Cdk2 activity in late G1 requires the inactivation of CKIs and Rb (Sherr and Roberts 1999). On activation, Cdk2 reinforces Cdk4 to complete Rb phosphorylation and also inhibits p27. This changes the program to reduce the cell's dependency on mitogens for completion of the cell cycle, and, in this sense, results in an irreversible commitment of cells to enter S phase.

After the initiation of the S phase, the progression of cell cycle is independent of mitogens. The subsequent phases of the cell cycle including the S-phase, the G2 phase and mitosis are regulated by the action of cyclin A-Cdk1 and Cyclin B-Cdk1 complexes. The G2/M DNA damage checkpoint prevents the cell from entering mitosis (M-phase) if the genome is damaged. The cdc2-cyclin B kinase is pivotal in regulating this transition (Vermeulen, Van Bockstaele et al. 2003).

2.7 The 2-Wave theory

Recently, it has been proposed that growth factor dependent mechanisms of cell cycle progression may be considered to consist of two phases or "two waves" (Jones and Kazlauskas 2000). The first wave consists of the early burst of signaling after growth factor stimulation, is insufficient for cell cycle progression and regulates early G1 or $G1_E$ phase (Jones and Kazlauskas 2000). The second wave or late phase begins nearly 3-7 hours (in hepg2 cells) after addition of growth factors, is required for S phase entry and regulates late G1 or $G1_L$ phase (Jones and Kazlauskas 2001).The findings indicate that the 2 phases may be regulated by different sets of enzymes or signaling pathways (Jones and Kazlauskas 2001). Studies also suggest activation of pathways that drive cell cycle are not dependent on the type of signal and that multiple growth factors/mitogens result in the activation of common signaling cascades which drive the cell cycle progression (Jones and Kazlauskas 2001).

The G1 phase may therefore temporally be separated into the early and late G1 phase. Based on the results of the present study, we believe it is the cells in this late G1 phase which are crossing or preparing to cross the G1/S checkpoint and are therefore in the 'G1S' phase. We speculate that the networks that control or drive the cells through these phases of early and late G1 may differ in terms of structure and topology

2.8 Dysregulation of the cell cycle: Cancer

It is believed that most cancers share an acquired set of capabilities (Fig 2.4). These include evading apoptosis, self sufficiency in growth signals, insensitivity to anti-growth signals and limitless replicative potential, all of which are directly or indirectly associated with the *loss of cell cycle regulation* (Hanahan and Weinberg 2000). Normal cells require growth factor mediated signals to enter the cell cycle. The loss of dependence on growth factor signals is a hallmark for cancer. This self sufficiency in growth signals may occur due to alteration in any one or more of the following-

(i) Alteration of the growth signal (at the receptor level),

(ii) Alteration in the transducers of those signals (Signaling intermediates),or

(iii) Alteration in the circuits that translate signals into action (such as the cell cycle)

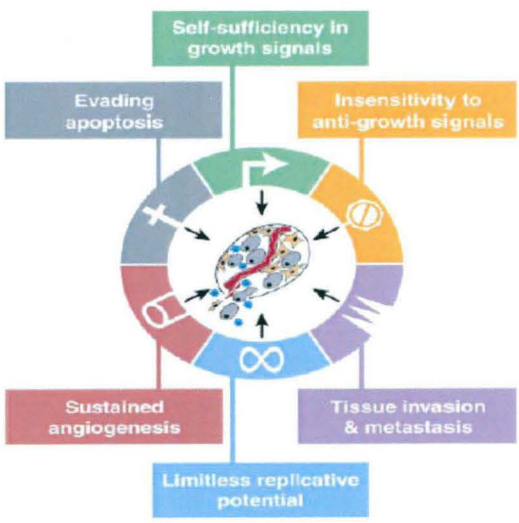


Figure 2.4. The Acquired Capabilities of Cancer. *(Source :Hanahan and Weinberg 2000).*

From among the factors mentioned above, the growth factor independence acquired from the deregulation of cell surface receptors and cell cycle regulators is fairly well understood (Hanahan and Weinberg 2000). Cell surface receptors which transmit extra-cellular signals into the cell are deregulated in various cancer types. Growth

factor receptors, largely belonging to the tyrosine kinase family, are often over-expressed in many forms of cancers. For example, (EGF-R/*erbB*) is upregulated in stomach, brain, and breast tumors (Yarden and Ullrich 1988), while the HER2/*neu* receptor is over-expressed in stomach and mammary carcinomas (Slamon, Clark et al. 1987; Yarden and Ullrich 1988). Similarly, the cell cycle machinery is well studied and the broad principles of cell cycle regulation are well accepted. It is also well known that tumour-associated cell cycle defects are often mediated by alterations in cyclin-dependent kinase (CDK) activity. Mis-regulated CDKs induce unscheduled proliferation as well as genomic and chromosomal instability. However, according to current models, mammalian CDKs are essential for driving each cell cycle phase, so therapeutic strategies that block CDK activity are unlikely to selectively target tumour cells (Malumbres and Barbacid 2009).

The most complex mechanism of acquired growth factor autonomy derive from alteration in components of downstream cytoplasmic circuitry that receive and process the signal from the cell surface growth factor receptors. Many signaling intermediates, mostly Kinases, have been found to be mutated or overexpressed in human cancers. Some of the well studied signaling pathways include MAPK pathway, the PI3K pathway and Wnt signaling pathways among others (Downward 2003; Hennessy, Smith et al. 2005; Reya and Clevers 2005; Bradham and McClay 2006). It is largely believed that signaling pathways are deregulated in most cancers. Though this fact has not been proved, there is a large body of evidence suggesting the same (Hunter 1997). Signaling pathways such as the MAPK and the PI3K pathways activated in response to growth factors promote the activation of G1/S phase regulators at multiple levels including gene transcription, translation, assembly, transport and activation of different cell cycle regulators (Zhang and Liu 2002). Multiple kinases involved in these

pathways such as Raf and AKT are known to be deregulated in different types of cancers. Similarly the ras proteins control signaling pathways that are key regulators of several aspects of normal cell growth and malignant transformation. They are aberrant in most human tumours due to activating mutations in the *RAS* genes themselves or to alterations in upstream or downstream signaling components (Downward 2003).

These regulators may also include immediate early genes like Jun, Fos, Myc etc and delayed early genes such as D type cyclins etc (Roussel 1997). An example of how growth factor mediated signaling pathways regulate cell cycle is the regulation of cyclin D. The G1 phase progression is dependent on the activity of D type cyclins and their CDK partners CDK4/6. Activation by growth factors results is increased association of cyclinD-CDK4 and activation of these complexes to promote G1 progression (Sherr and Roberts 1999; Sherr and Roberts 2004). Withdrawal of growth factors results in arresting Cyclin D synthesis, its degredation and loss of cyclin D activity. In general, growth factors promote cell cycle progression by the phosphorylation of RB and the inactivation of CDK inhibitory CIP/KIP proteins such as p27 (Sherr and Roberts 1999; Sherr and Roberts 2004). RB phosphorylation results in the subsequent release of E2F transcription factors and the initiation of S phase. The critical role of these pathways in regulating G1/S phase of cell cycle makes them attractive targets for therapeutic intervention.

One of the earliest studied examples in this of such a pathway was the MAPK pathway. Gradually, many other pathways and signaling regulators were identified and the emerging cross talk between these pathways, further complicated the molecular understanding of the disease. As a result, though many signaling intermediates have also been identified in the aetiology of many cancers, the exact mechanism with which

they bring about cell transformation is not well understood. In this study we have investigated the role of such upstream regulators specifically the kinases and phosphatases in cell cycle regulation to obtain a systems view of the upstream signaling networks that regulate the cell cycle and its phases.

2.9 RNAi screens

One of the most effective tools of functional genomics is the use of RNAi interference to understand gene functions. In the last few years a large number of genome scale RNAi screens have been performed to understand a plethora of biological processes in drosophila, animals, eukaryotes, humans, nematodes, viruses infection and mammalian cell based systems (Mohr, Bakal et al. 2010). These processes include cell cycle, signal transduction pathways, host pathogen interactions, fat regulation and cancer biology among others (Mohr, Bakal et al. 2010) (Table 2.1). Such screens have not only enabled a systems level understanding of cellular processes, they have also helped deepen the mechanistic insights by identification of novel components in these processes.

RNA interference is essentially a loss-of–function technique, where the effect of gene knockdown can be assessed directly as a functional (biological) readout. It reveals the role of a gene in a specific biological context and depending on the experimental design it may be small scale or high throughput (Mohr, Bakal et al. 2010). Though the use of RNAi at a small scale has also revealed a wealth of information, its most important impact has been the ability to perform high throughput RNAi based screens. Such formats allow the function of a large number of genes to be interrogated concurrently (Fig 2.5).

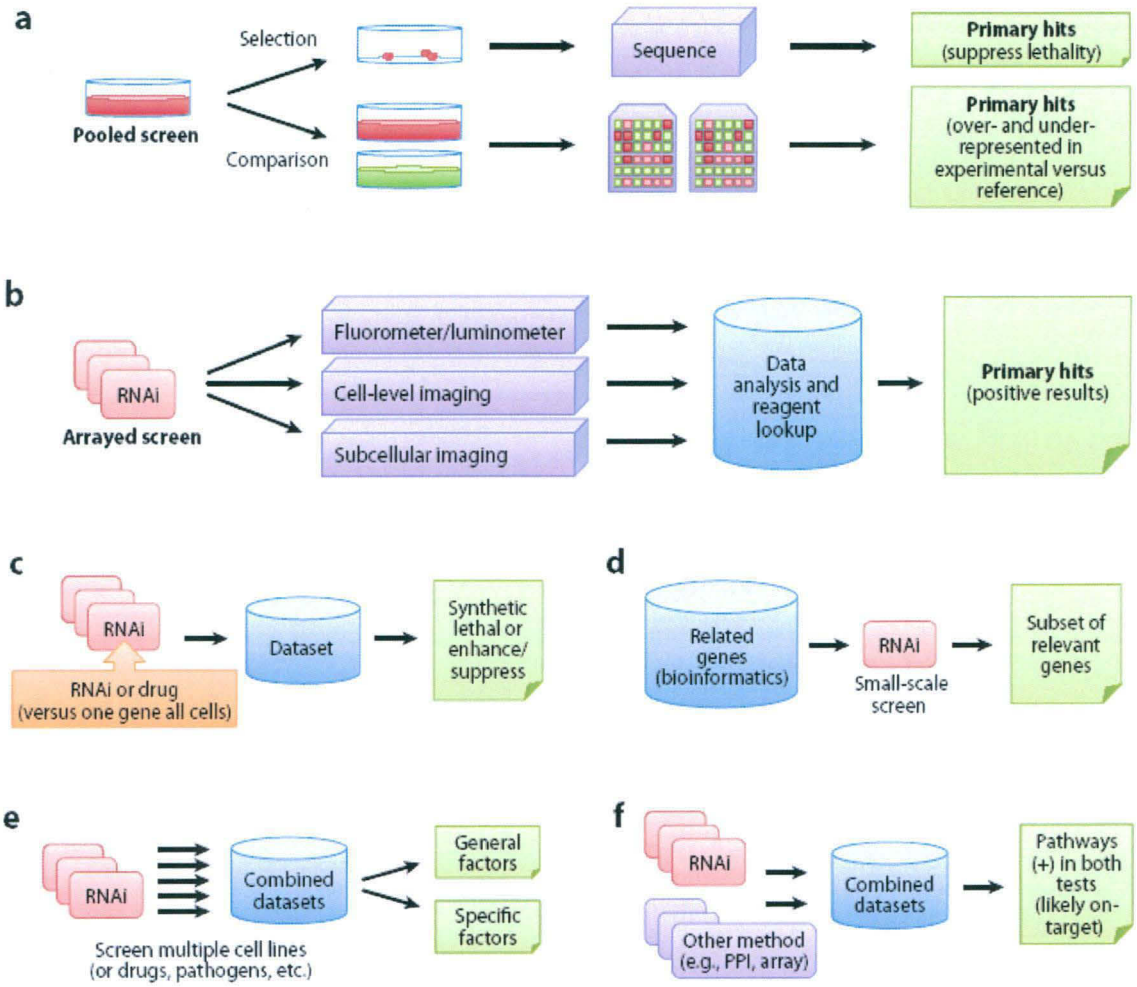


Figure 2.5. Approaches to high-throughput cell-based RNAi screening. *(a) Pooled RNAi high-throughput screen (HTS) approach. (b) Arrayed RNAi HTS approach. (c) Modification of a pooled or arrayed approach via prior addition of a treatment, such as RNAi against a single gene in all cells or treatment with a small molecule. (d) Identification of related genes via informatics-based analysis (e.g., all kinases or genes previously implicated in a specific pathway or complex), followed by screening with reagents directed against the identified subset of genes. (e) A HTS with one assay type using multiple cell lines, the same cell line with multiple pathogens, or a similar multiplexed approach, followed by data integration to identify specific and general factors. (f) Parallel RNAi HTSs and an additional experimental high-throughput genomic or proteomic approach, followed by data integration to identify high-confidence hits. Abbreviations: PPI, protein-protein interaction; + sign, positive result. (Source : Mohr, Bakal et al. 2010).*

A cell based RNAi high throughput screening can be done in two formats: a pooled format, in which the library is introduced at random into cells, or an arrayed format, in which single genes are targeted by reagents in individual wells of a microtiter plate (Iorns, Lord et al. 2007; Lord, Martin et al. 2009; Sharma and Rao 2009) (Fig 2.5). The

array format is more flexible and is often performed in 96 or 384 well plate format, with each well containing a unique RNAi reagent. Depending on the experimental design and cell assay different detection methods are used. The main advantage of the arrayed format is that it allows multiple genes to be interrogated and multiple phenotypes to be detected in a single screen at a given time. This interrogation of multiple genes at a given time helps short listing of a set of genes implicated in a given biological function or process (Mohr, Bakal et al. 2010). The screens have enabled us to get biological data at the genome level which was previously possible at a much smaller scale (Table 2.1).

Important factors to consider in RNAi (High throughput screens) HTSs include (*a*) performing at least one replicate test in the primary screen, (*b*) including an appropriate type and number of "no treatment" and other controls, (*c*) a thoughtful array of the library and controls (e.g., randomized), (*d*) an early assessment of data quality to detect plate or well-level problems such as dispensing errors, (*e*) data normalization, and (*f*) setting appropriate cut-off values for significant results (Boutros, Bras et al. 2006; Stone, Marine et al. 2007; Wiles, Ravi et al. 2008; Birmingham, Selfors et al. 2009; Zhang and Heyse 2009). Despite the recent improvements in addressing all of these factors during RNAi HTSs, subsequent data analysis, and follow-up tests, false discovery remains a significant and difficult problem to address. Statistical and experimental approaches can help to minimize the problem (Echeverri, Beachy et al. 2006; Birmingham, Selfors et al. 2009).

Table 2.1 Results of genome-scale[a] cell-based RNAi high-throughput screens in mammalian or *Drosophila* cells.

Cell type	Screen type	Reagent	Primary hits	Secondary hits[b]	Field of study
Human cells					
HeLa	Plate reader & imaging	esiRNA	275	37	Cell division
U2OS	Imaging	siRNA	1,152	18	Cell cycle
NCI-H1155	Plate reader	siRNA	87	6	Cancer biology
NIH3T3	Pooled	shRNA	15	3	Stress resistance
293T	Plate reader	siRNA	—	295	Host-pathogen interactions
293T, HeLa, MCF-7	Pooled	shRNA	30	8	Cell death
DLD1	Plate reader	siRNA	740	268	Signal transduction
HEK293	Pooled	shRNA	13,140	21	Cell adhesion
HeLa	Imaging	siRNA	305	124	Host-pathogen interactions
HeLa	Plate reader	siRNA	530	23	Signal transduction
HeLa-derived TZM-b1	Plate reader	siRNA	386	273	Host-pathogen interactions
HeLa P4/R5	Plate reader	siRNA	931	232	Host-pathogen interactions
Jurkat	Pooled	shRNA	11	5	Cancer biology
MCF-10A[c]	Pooled	shRNA	201	166	Cancer biology
MNT-1	Plate reader	siRNA	98	35	Pigmentation
RDG3	Imaging	siRNA	—	171	Stress resistance
BJtsLT	Pooled	shRNA	100	37	Cancer biology
DLD-1	Pooled	shRNA	368	83	Cancer biology
Huh7/Rep-Feo	Plate reader	siRNA	236	96	Host-pathogen interactions
Jurkat	Pooled	shRNA	252	7	Host-pathogen interactions
Huh 7.5.1	Imaging	siRNA	521	262	Host-pathogen interactions
Mouse cells					
NIH 3T3	Pooled	shRNA	—	28	Cancer biology
L929	Plate reader	siRNA	666	432	Cell death
B16-F0	Pooled	shRNA	78	22	Cancer biology
Oct4-Gip ESCs	FACS[d] & imaging	esiRNA	296	21	Stem cell biology
Oct4-Gip ESCs	FACS	siRNA	148	104	Stem cell biology
***Drosophila* cells**					
Kc167, S2R+	Imaging	dsRNA	438	—	Viability
S2	Imaging	dsRNA	—	121	Host-pathogen interactions
Clone 8	Plate reader	dsRNA	238	213	Signal transduction
Clone 8	Plate reader	dsRNA	509	96	Signal transduction
Kc167	Plate reader	dsRNA	—	90	Signal transduction
S2	Plate reader	dsRNA	474	121	Signal transduction
S2	Imaging	dsRNA	—	86	Host-pathogen interactions
S2	Imaging	dsRNA	305	~190	Host-pathogen interactions
S2	Imaging	dsRNA	210	112	Host-pathogen interactions
S2	Plate reader	dsRNA	—	14	Host-pathogen interactions
S2	Plate reader	dsRNA	1,133	284	Protein secretion

S2	FACS	dsRNA	488	—	Cell cycle and/or cell size
S2	Imaging	dsRNA	—	—	Signal transduction
S2	FACS	dsRNA	66	23	RNA biology
S2	Plate reader	dsRNA	75	4	Ion transport
S2R+	Plate reader	dsRNA	138	7	RNA biology
S2R+	Plate reader	dsRNA	1,168	331	Signal transduction
S2R+	Imaging	dsRNA	699	—	Signal transduction
S2R+	Imaging	dsRNA	1,500	27	Signal transduction
S2	Imaging	dsRNA	90	24	RNA biology
Kc167	Plate reader	dsRNA	81	47	Cell death
S2	Plate reader	dsRNA	47	1	RNA biology
S2*	Plate reader	dsRNA	18	5	Chromatin regulation
S2R+	Imaging	dsRNA	346	—	Signal transduction
S2	Imaging	dsRNA	162	54	Host-pathogen interactions
DL1	Plate reader	dsRNA	176	110	Host-pathogen interactions
Kc167	Imaging	dsRNA	526	—	Lipids
Kc167	Plate reader	dsRNA	265	120	Transcription and/or translation
Primary neurons	Imaging	dsRNA	336	104	Neural outgrowth
S2	Plate reader	dsRNA	821	152	Mitochondria
S2	Plate reader	dsRNA	—	—	Phagocytosis
S2	Imaging	dsRNA	—	—	Centrioles and/or centrosomes
S2	Imaging	dsRNA	847	227	Lipids
S2	Imaging	dsRNA	292	133	Cancer biology
S2	Imaging	dsRNA	23	—	Transcription and/or translation
S2-derived RZ-14	Plate reader	dsRNA	177	—	RNA biology
S2R+	Imaging	dsRNA	119	39	Centrioles and/or centrosomes
S2R+	Imaging	dsRNA	133	72	RNA biology
S2R+	Imaging	dsRNA	~500	1	Mitochondria
S2R+	Plate reader	dsRNA	303	173	Circadian rhythms
Clone 8	Plate reader	dsRNA	~100	11	Signal transduction
S2	Plate reader	dsRNA	218	116	Host-pathogen interactions
Kc167	Plate reader	dsRNA	996	202	Cell cycle and/or cell size
S2R+	Plate reader	dsRNA	42	33	Cell death
S2R+	Imaging	dsRNA	—	—	Signal transduction
S2R+	Imaging	dsRNA	15	7	Cell cycle and/or cell size

[a] *For this summary, genome-scale was defined with a cut-off of approximately 5000 genes (mammalian cell screens) or at least 70% of the genome (Drosophila cell screens).*

[b] *Here, secondary hits were used to refer to the largest set of primary hits that passed an additional test verifying the result at the reagent level (retest after re-synthesis or with another assay or cell type) or in most cases, at the gene level (retest with another reagent or single reagents from a pool, for example).In some cases, only a subset of primary hits were tested in secondary assays. For most reports, only a small number of genes (typically, one to five) were confirmed with a rigorous test, such as rescue of the RNAi effect with a cDNA, or were confirmed at the level of biological significance with another type of assay or an in vivo analysis.*

[c] *Additional cell types tested with smaller shRNA pools.*

[d] *FACS, fluorescence-activated cell sorter. (Source: Mohr, Bakal et al. 2010)*

Clearly, these new methods of high throughput screening provide tremendous power. In principle, all genes of the genome are being tested, and the effect on the gene product—depletion—is likely to be simple and easy to interpret. Importantly, this

method of depleting gene products is perfectly suited to screening in diploid cells, as both alleles of a gene are simultaneously suppressed. However there are some drawbacks and challenges in siRNA screening (Mohr, Bakal et al. 2010).

There is significant variability observed in high-throughput screen data caused due to variable methods of data analysis. Also, the results from such screens are also very sensitive to experimental design, nature of reagents used, cell based systems and statistical cut-offs used. A major challenge in RNAi screens data analysis is minimising false discovery rates (FDRs) due to the presence of false positives and false negatives (Mohr, Bakal et al. 2010). Experimental noise in the HTS screens and inherent bias in certain screen assays contributes to both false positives and negative. On the other hand false negative may emerge due to low potency of the RNAi reagents and weak knockdown of target genes (Lee, Santat et al. 2009; Mohr, Bakal et al. 2010). Also knockdowns may result in weak phenotypes which lie below the statistically significant cut-offs. False discovery tolerances may also vary among researchers. All these factors contribute to the limited overlap observed among related siRNA screens. One of the inherent drawbacks of RNAi screening is the off- target effects or OTEs observed. False positives are often attributed to the presence of OTEs which may be sequence dependent or sequence independent (Jackson, Bartz et al. 2003; Echeverri, Beachy et al. 2006; Kulkarni, Booker et al. 2006; Ma, Creanga et al. 2006).

Another major drawback of siRNA screens is their inability to deal with redundancy in the cellular networks. The screens are designed in a manner which allows the depletion of a single gene product at a time. Critical regulators may be missed as hits and reported as false negatives if there are redundant pathways involved. Often the systems used for screening are cell lines chosen for experimental convenience. Therefore, genes

identified in such screens need to be tested in more physiologically relevant systems. Lastly, the siRNA screens are likely to identify only those genes whose depletion confers recessive, loss-of-function, or hypomorphic effects (Goff 2008). Unlike a conventional mutagenesis screen, the full range of mutations, such as over expression or dominant mutations, are not likely to be found.

All these factors contribute to the false discovery rates (FDRs). Development of appropriate approaches to minimize false discovery rates remains a challenge as, to a large extent, minimizing false-positive results increases the number of false-negative results, and vice versa (Mohr, Bakal et al. 2010). Efforts at standardization include MIARE (for Minimum Information about an RNAi Experiment, ttp://www.miare.org), and information about reagents and data is being collected at the Probe (http://www.ncbi.nlm.nih.gov/probe) and PubChem (http://pubchem.ncbi.nlm. nih.gov/) databases at NCBI (http://www. ncbi.nlm.nih.gov).

2.9.1 Meta-analysis of genome wide RNAi screens.

Recent work on the meta-analysis of three different siRNA screens conducted to identify host factors involved in HIV survival highlights these weaknesses (Goff 2008; Bushman, Malani et al. 2009). Closer comparisons of the three screens reveal a disconcerting fact: Although all screens had the same goal, they did not recover the same sets of genes. Indeed, there was very little overlap of < 7 %between some pairs of sets (Goff 2008; Bushman, Malani et al. 2009). Clearly, despite their common aim, these screens uncovered wildly different hits. The identities of the hits recovered seem to depend strongly on the conditions and readouts of the screens. Thus, new screens may identify even more sets of host genes. Indeed, it is likely that even these large-scale screens have missed some important players. Therefore, despite the many genes

identified in various siRNA screens, it is evident that these screens produce a parts list of genes involved in specific cellular processes with little or no insight into the mechanistic understanding of the process (Goff 2008; Bushman, Malani et al. 2009). Most of the lists are incomplete due to the inherent weaknesses of siRNA screens.

Integration of orthogonal high-throughput experimental approaches has emerged as an important approach which can help overcome the above mentioned weaknesses of high throughput siRNA screens (Bakal, Linding et al. 2008; Major, Roberts et al. 2008; Mohr, Bakal et al. 2010). Such approaches are based on the combination of high throughput screen data with proteomic, genomic and/or genome level bio-informatics data. Also, experimental data can also be merged or overlaid onto Protein –protein interaction or PPI networks. Such approaches have lead to important insights into areas such as signal transduction, host-pathogen interactions and cancer biology among others.

2.10 Multi-component therapeutics for networked systems

One of the main purposes of high throughput experimental approaches is the identification of druggable targets. However the advent of high-throughput approaches and the generation of high throughput proteomic, genomic and metabolomic data have not satisfactorily yielded to the expectation of identifying individual drug targets (Korcsmáros, Szala et al. 2007). The number of novel drug targets and drugs has fallen significantly behind expectation. Another important fact that has emerged is that the targeting or inhibition of individual targets has often been found to be less effective at treating a disease (Sams-Dodd 2005; Zimmermann, Lehar et al. 2007). Though some of the 'one drug-one target' drugs have revolutionized modern medicine, many patients are unable to benefit from these therapies because of pharmacogenomic effects (Roses

2000; Zimmermann, Lehar et al. 2007). Often the rationale of drug discovery is to restore the healthy state by inhibiting the component which is central to the disease. In such cases, the action of an alternative target might predominate and these patients could benefit from a combination that simultaneously impacts the principal and alternative targets (Zimmermann, Lehar et al. 2007).

The one drug- one target approaches have been challenging for multiple reasons. The tremendous redundancy in cellular systems results in a 'buffering' effect such that other similar components can perform the functions of the targeted components (Hanahan and Weinberg 2000; Hartman, Garvik et al. 2001). Such limitations can be overcome by targeting the disease at multiple fronts. Such multi-target approaches can prove to be more effective and less vulnerable to adaptive resistance because the biological system is less able to compensate for the action of two or more drugs simultaneously (Zimmermann, Lehar et al. 2007).

Multi-target therapies have recently been rediscovered by drug developers. The multi target approach is being explored for the treatment of many complex diseases including cancers and tuberculosis (Cohn, Catlin et al. 1990; Csermely, Agoston et al. 2005; Fitzgerald, Schoeberl et al. 2006). As mentioned earlier, given the complexity and redundancy in cellular systems, the biggest challenge is the identification of multiple druggable targets for combinatorial targeting and the subsequent development of multi-target drugs. Current efforts in this direction, such as that of employing combination screens, are limited by their inability to sample the vast space of high-order combinations that are theoretically possible. In contrast integration of high throughput data with additional ***network based approaches*** which identify a shorter list of key targets can greatly reduce this space and prove to be a more amenable solution

(Csermely, Agoston et al. 2005; Lehar, Stockwell et al. 2008; Pawson and Linding 2008; Barabasi, Gulbahce et al. 2011; Farkas, Korcsmaros et al. 2011). Though most drug discovery efforts have been largely reductionist and evaluate the roles of individual genes/proteins in a disease, however systems perspectives make it clear that most diseases and cellular processes cannot be attributed to a single gene or protein. A better understanding of ***disease networks*** reveal why inhibition of individual targets is not sufficient for restoration of a healthy state.

Cellular networks help us to understand the complexity of the cell. The network approach examines the effects of drugs in the context of cellular networks (Csermely, Agoston et al. 2005). Cellular networks offer a lot of possibilities to point out their key elements as potential drug targets. Network based approaches make it apparent that depending on the network topology and node characteristics, perturbing a given node/target may have local or global effects. Network based approaches also highlight that the redundancy and complexity of cellular networks result in limited efficacy of individual node perturbations of one drug-one target approaches.

Networks have a number of vulnerable points and, therefore, can be attacked in many ways (Korcsmáros, Szala et al. 2007). These vulnerable points can be identified by understanding network topology and node characteristic. As described earlier, cellular networks contain hubs, that is, elements which have a large number of neighbours. These networks can be dissected to overlapping modules that form hierarchical communities such as signaling networks which have interdigitated pathways and multiple layers of cross-talk (Korcsmáros, Szala et al. 2007). Networks can be attacked by attacking nodes with the highest degree, attacking hub-links, targeting bridges or links which connect nodes having high betweeness centralities or attacking bridges

having links with the highest weighted centrality (Fig 2.6) (Korcsmáros, Szala et al. 2007). Such network based approaches can help delineate the critical functional and disease modules and the subsequent identification of critical nodes within these modules which may prove to be ideal drug targets.

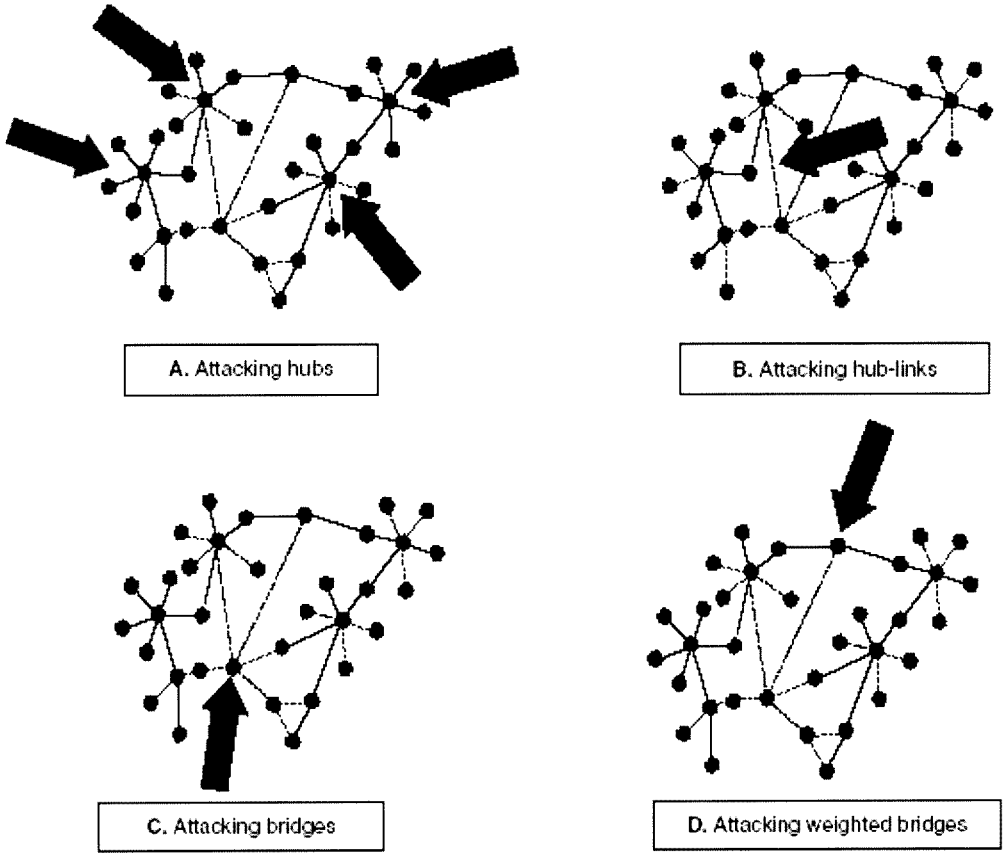


Figure 2.6. Attack scenarios on networks. *The figure summarizes a number of malicious attacks on vulnerable points of networks.* ***A.*** *Attacks on nodes with highest degree (hubs).* ***B.*** *Attacks on 'hub-links' with the highest degree of their end points.* ***C.*** *Attacks on bridging elements with links having a high betweenness centrality.* ***D.*** *Attacks on bridges having links with the highest weighted centrality. (Source :Korcsmáros, Szala et al. 2007).*

Drug-driven network attacks are targeted to find the most efficient way to influence network behaviour (Fig 2.6). Several classes of drugs, such as antibiotics, fungicides, anticancer drugs as well as numerous chemical compounds, such as pesticides, are designed to destroy the normal function of cellular networks (Korcsmáros, Szala et al. 2007).

The sections above discuss the identification of novel targets for synergistic multi-therapeutic interventions using network approaches. Recent studies have attempted to identify functional and disease modules using network approaches (Barabasi, Gulbahce et al. 2011). The identification of such modules seems functionally more relevant since they consist of groups of critical nodes involved in regulating specific cellular processes or disease. Also, since both biological processes and disease progression are dynamic processes, the identification of dynamic temporal modules for these would provide a more realistic understanding of the constituent nodes (Barabasi, Gulbahce et al. 2011). A "multi-modular" therapeutic targeting approach can prove to be the paradigm shift in therapeutic strategies for complex disorders.

2.10.1 Synergy: when the whole is greater than the sum of parts

There is widespread evidence that the combinations of compounds can be more effective than the sum of the effectiveness of the individual agents themselves, a result that can be rationalised using principles of systems and molecular biology (Keith, Borisy et al. 2005). An obvious application of this lies in the treatment of multigenetic disorders like cancers. For example, certain cancers develop due to multiple mutations and their chemotherapeutic regimes require combination therapies (Kinzler and Vogelstein 1996; Keith, Borisy et al. 2005). Similarly, many other multi-component drugs are now being used to treat disorders like asthma, HIV etc.

One of the prime advantages of multi-targeted therapeutics is the production of "synergy" that is, the combinational effect to be greater than the sum of the individual effects, making multi-component therapeutics a systematic approach, rather than the reductionism of an additive effect (Zimmermann, Lehar et al. 2007). So following the identification of synergistic targets, synergy testing for dose determination is also

required. It requires testing specific doses, as compound combinations are frequently found to be synergistic over one range of doses and antagonistic over another. Though there are different additive mathematical models to compute synergy, such as Loewe additivity and Bliss Independence (Chou and Talalay 1981; Zimmermann, Lehar et al. 2007), the best definition for synergy must reflect the requirement that, when used in combination, a benefit is observed that could not be achieved by the individual components (Straube, Aicher et al. 2011). In some cases, adding more of the same drug will incur unacceptable negative consequences. For example, when the dose of the first drug is just below a threshold of toxicity, even a small amount more might be unacceptable, whereas combining it with a drug that possesses non-overlapping (orthogonal) toxicities might provide enormous benefit, even if the combined effect on efficacy is only Loewe additive. For example, an increase in tumour-cell killing from 50% to 75% without additional side effects might be clinically meaningful, but not mathematically synergistic (Keith, Borisy et al. 2005).

Such Multi-Target therapeutics employs multi-component drugs which comprise several biologically active compounds with mutually interdependent activities that are required for an optimal effect. Ideally, when these components interact in a biological system, they yield a significant and desired pharmacological effect. Such drugs can be particularly effective interventions in networked systems (Barabasi, Gulbahce et al. 2011).

2.10.2 Applications in Oncology

The principles of multi-target therapeutics can be especially applicable to oncology (Zimmermann, Lehar et al. 2007). The etiology of many cancers and the process of oncogenesis is known to be multigenic (most cancers have four to seven mutations). A

multi-therapeutic strategy can yield effective and durable responses. Interestingly, even in the case of most tumorigenic viruses, evolution has favoured viral strains that can simultaneously block redundant mechanisms and safeguards that the cell uses to prevent inappropriate cell cycling and –proliferation (Zimmermann, Lehar et al. 2007). These viruses use a multi-target approach to drive their own proliferation. Given that multiple nodes in the system must be modified to induce cancer, many researchers have proposed that multiple interventions will be required to counter this process (Csermely, Agoston et al. 2005; Zimmermann, Lehar et al. 2007; Lehar, Stockwell et al. 2008; Farkas, Korcsmaros et al. 2011). Many such drugs that target multiple kinases are already being tested for their efficacy against different types of cancers and are under clinical trials (Elbauomy Elsheikh, Green et al. 2007; Tomlinson, Baldo et al. 2007; Sathornsumetee and Reardon 2009). It is evident now that such approaches may also prove to be more effective for other multigenic disorders such as Diabetes.

2.11 Cell Cycle Deregulation in B cell Lymphomas: Importance of Signaling Networks.

Although studies have shown that cell cycle subversion is a key step in tumorigenesis, the available information on human tumors has only recently reached the critical threshold of knowledge that allows a reasonably clear understanding of the mechanisms of cell cycle inactivation and their contribution to the genesis and progression of human cancer (Sanchez-Beato, Sanchez-Aguilera et al. 2003). Disruption of the physiologic balance between cell proliferation and death is a universal feature of all cancers. In general terms, human B-cell lymphomas can be subdivided into 2 main groups, low- and high-growth fraction lymphomas, according to the mechanisms through which this imbalance is achieved (Sanchez-Beato, Sanchez-

Aguilera et al. 2003). Most types of low-growth fraction lymphomas are initiated by molecular events resulting in the inhibition of apoptosis, such as translocations affecting *BCL2*, in follicular lymphoma, or *BCL10* and *API2/MLT1*, in mucosa-associated lymphoid tissue (MALT) lymphomas (Willis, Jadayel et al. 1999; Zhang, Siebert et al. 1999; Albinger-Hegyi, Hochreutener et al. 2002). This leads to cell accumulation as a consequence of prolonged cell survival. In contrast, high-growth fraction lymphomas are characterized by an enhanced proliferative activity, as a result of the deregulation of oncogenes with cell cycle regulatory functions, such as *BCL6*, in large B-cell lymphoma, or c-*myc*, in Burkitt lymphoma (Bastard, Deweindt et al. 1994; Offit, Lo Coco et al. 1994).

Both low- and high-growth fraction lymphomas can evolve to highly aggressive tumors; the process takes place through the acquisition of alterations in the major tumor suppressor pathways, most frequently affecting CKI (Sanchez-Beato, Sanchez-Aguilera et al. 2003). Low- and high-growth fraction lymphomas are both able to accumulate other alterations in cell cycle regulation, most frequently involving tumor suppressor genes such as *p16INK4a , p27KIP1 and pathways such as the p53 and Rb pathways*. In fact, inactivation of CKI is a marker of highly aggressive lymphomas, irrespective of their histologic grade (Sanchez-Beato, Sanchez-Aguilera et al. 2003). As a consequence, these tumors behave as highly aggressive lymphomas. The simultaneous inactivation of several of these regulators confers increased aggressivity and proliferative advantage to tumoral cells. These factors indicate that the molecular basis of cancers cannot be attributed to single molecular events, but are in fact a result of accumulated alterations in tumor suppressive pathways. However, despite the important progress that has been made toward the elucidation of the mechanisms of

lymphogenesis, many points still need to be clarified, such as the plethora of molecules and pathways that may be affected by their alterations.

Chapter 3

Rationale for the Study

Chapter 3: Rationale for the study

Human cancers usually evolve through a multistage process that involves the accumulation of mutations and epigenetic abnormalities that affect expression of several genes (Puxeddu, Romagnoli et al. 2011). This multi-factorial basis, which also then varies widely between individual cancers, is primarily responsible for the limited success rate in the development of anti-cancer chemotherapies (Malumbres and Barbacid 2001; Vogelstein and Kinzler 2004). In addition to this, however, we speculated that another mitigating cause could be the fact that present strategies are based on the expectation of linear behaviour from components of the signaling machinery. That is, if the oncogenic activation of a given signaling molecule contributes to cell transformation, then pharmacological inhibition of that molecule should produce the reciprocal effect of inducing death of that cell. Such logic, however, does not take into account the current paradigm that the molecular components of a cell organize into a complex network of interactions, and that biological responses represent emergent features of such networks (Barabasi and Oltvai 2004). This perspective is equally true of mitogen-activated, or, oncogene-activated signaling networks, where the consequent emergent responses can then either be cell cycle regulation, or dysregulation.

As discussed earlier in the review of literature, the hallmark of dynamic networks, including signaling networks, is non-linear response behaviour. This property derives from multiple factors some of which include a high degree of pathway redundancy, and the diversity of regulatory elements that combine feedforward, feedback and other such mechanisms in local domains distributed across the network. Consequently, the principle of reversibility need not necessarily hold true for the behaviour of such

networks. This issue becomes particularly more complex in the case of cancers where mutational/epigenetic activation of multiple molecules converges to induce, and then sustain, the dysregulated cell cycle. We therefore rationalized that a resolution of the regulatory elements involved in the mitogenic signaling network would be necessary, to facilitate a more rational way of approaching the challenge of cancer chemotherapy.

Despite the accumulation of a large body of literature on signaling pathways involved in cell cycle regulation, the manner in which they integrate together and the regulatory features that process input-output relationships, is still poorly understood (Murray 2004). Our first task therefore, was to address this outstanding issue. Here, given that the cell cycle process extends over several hours, we also recognized the need to resolve the temporal modulations in network topology, and the consequent effects on functioning of the signal processing elements. This posed a particularly difficult problem given that satisfactory methods for capturing the temporal dimension of a biological process are presently lacking. However, as discussed in the results, we were indeed successful in developing a novel approach that resolved not only the question of network topology and its regulation, but also its architectural modulations that drive cells through individual stages of the cell cycle response. Central to this success was our adoption of a working hypothesis that derived from current information on the structural features of signaling networks. It is now accepted that the opposing requirements for sensitivity and robustness constraints the overall structure into a bow tie, or hourglass, configuration (Csete and Doyle 2004; Supper, Spangenberg et al. 2009). Functionally, this implies the convergence of diverse and redundant input processes onto a conserved core module, or set of proteins. Such core elements then serve as the key regulators of plasticity in the cellular response, by calibrating a range of output processes (Csete and Doyle 2004; Supper, Spangenberg et al. 2009).

It was probable therefore, that the mitogen-activated signaling network also adopts such a structure, thereby incorporating conserved regulatory elements that can processes input signals, and translate them into an output response that regulates the cell cycle. This possibility has important implications for cancers since oncogenic activation of network components can then be expected to influence by perturbing the input information to these modules, thus resulting in an aberrant output response manifested in terms of cell cycle dysregulation. Consequently, in addition to yielding new mechanistic insights, the identification of such regulatory elements should also facilitate development of more effective modalities for chemotherapy.

A useful approach for probing properties of networks is to perform targeted perturbations of the component nodes, and then monitor for the consequent effects on the output response. We adopted such a strategy by performing a siRNA screen against components of the signaling machinery to identify those targets whose depletion resulted in a perturbation of the cell cycle. At this stage, we took note of the fact that genetic screens suffer from limitations that miss important regulatory molecules, and also frequently yield 'hits' that function in a cell type-specific manner. This aspect was exemplified when we compared our results with that of two other such screens described on unrelated cell lines in the literature. The overlap in targets identified between these three datasets was found to be less that 10%. This divergence could possibly reflect differences in the oncogenic pathway choices of these three distinct cell lines. We, therefore, segregated the 'hits' identified in our screen on the basis of the specific cell cycle stages that they perturbed, and extracted the network of protein-protein interaction pathways that extended from them, to the components of the cell cycle regulatory machinery that influenced that specific phase. In other words, this exercise yielded the network of signaling pathways that regulated each of the

temporally distinct phases of the cell cycle. By then applying the tools of graph theory in an iterative manner, we were also able to eventually identify the core ***modules*** that processed mitogenic signals, in each of these individual phases. Importantly, the innovative criteria employed in our analysis ensured that the modules identified were those that were unique for that corresponding cell cycle phase. The specific and functional relevance of these individual modules could be validated through extensive experimentation.

Chapter 4

Materials and Methods

4.1 MATERIALS

MATERIAL		SOURCE
Cell lines		
Raji, Jurkat, HL60, U937, THP1, Namalwa, A549, HeLa, HepG2, HEK293, U2OS and Huh7.		ATCC, Rockville, MD, USA
Cell line AW8507		Gift from Dr. Shubha Chiplunkar (ACTREC)
Cell line MCF7		Gift from Dr. Ramesh Bamezai (JNU)
siRNAS		
Mouse Kinase siRNA SET V1.0 (758)		Qiagen, GmbH
Mouse Phosphatase siRNA SET V1.0 (294)		Qiagen, GmbH
RNAi Human/Mouse starter kit (with alexa 488 labelled siRNA)		Qiagen, GmbH
Hiperfect transfection reagent		Qiagen, GmbH
siRNAs for Validation phosphatases		Dharmacon,(Chicago, IL),USA
MISSION siRNA Mouse Kinase Panel		Sigma Chemical Company, St. Louis, MO, USA
Reagents for SDS PAGE/Western Blot		
Hybond-N transfer membrane		Amersham Biosciences, Uppsala, Sweden
Protein Marker		Bio-Rad Labs, Hercules, CA, USA
Odyssey Blocking buffer		LI-COR Biosciences, NE,USA
IRDye 800/700 antibody		LI-COR Biosciences, NE,USA
Tissue culture reagents		
RPMI-1640		Invitrogen, Carlsbad, CA, USA
DMEM		Invitrogen, Carlsbad, CA, USA
Fetal Bovine Serum		Invitrogen, Carlsbad, CA, USA
Antibiotic-Penicillin-Streptomycin		Invitrogen, Carlsbad, CA, USA
Vybrant Dyecycle dye		Invitrogen, Carlsbad, CA, USA
B cell negative isolation kit		Invitrogen, Carlsbad, CA, USA

MATERIAL		SOURCE
Flow Cytometry Reagents		
DNA QC particles		BD Biosciences, San Jose, CA
Rainbow beads		BD Biosciences, San Jose, CA
Propidium Iodide		Sigma Chemical Company, St. Louis, MO, USA
RNAse A		Sigma Chemical Company, St. Louis, MO, USA
Softwares		
Flowjo		Tree Star, Inc., USA
MASCOT		Matriz science
BD FACSdiva software		BD Biosciences, San Jose, CA
Inhibitors		
Imatinib Mesylate		Gift from Sphaera pharma
LY294002		Tocris Bioscience, Bristol , UK
TCS 2312		Tocris Bioscience, Bristol , UK
2-APB		Tocris Bioscience, Bristol , UK
PD173074		Tocris Bioscience, Bristol , UK
KT5720		Tocris Bioscience, Bristol , UK
Others		
TiO2 beads		ZirChrom Separations Inc
SCX ICAT cartridges		Applied biosystems, USA
iTRAQ 8-plex/4-plex labelling kit		Applied biosystems, USA

4.2 BUFFERS AND SOLUTIONS

Buffers and solutions for PAGE and western blotting

Lysis Buffer for preparation of cytoplasmic extract (Buffer A)

20mM HEPES, 10mM NaCl, 1.5nM MgCl2, 0.2mM EDTA, 0.5% Triton X-100, 0.5mM DTT, 1 mM sodium orthovanadate, I mM NaF, and a cocktail of protease inhibitors.

Lysis Buffer for preparation of nuclear extract (Buffer B)

20mM HEPES, 20% glycerol, 0.8M NaCl, 1.5mM MgCl2, 0.2mM EDTA and 0.2% Triton X-100, and all the above mentioned protease and phosphatase inhibitors.

2X SDS-PAGE Buffer:

100 mM Tris-HCl, pH 6.8, 4% SDS, 20% glycerol, 4% β-mercaptoethanol, 0.01% bromophenol blue.

SDS –PAGE Running buffer

25 mM Tris, 192 mM glycine, 0.1% SDS.

Transfer buffer:

25mM Tris base, 250mM Glycine, 0.01% SDS (Sodium Dodecyl Sulfate) and 20% Methanol.

1X PBS (Phosphate-Buffered Saline): 13.7mM NaCl, 0.27mM KCl, 10mM di-Sodium hydrogen phosphate (Na2HPO4) and 0.2mM Potassium di-hydrogen phosphate (KH2PO4).

Blocking buffer: 5% (w/v) Bovine Serum Albumin in 1X PBS-T.

Wash buffer: 0.1% (v/v) Tween-20 in 1X PBS .Primary and secondary antibodies are prepared in wash buffer and blots were developed by Enhanced Chemiluminescence (ECL) kit.

Odyssey blocking buffer: Odyssey blocking solution in a 1:1 ratio with PBS.

Odyssey antibody dilution buffer:

For blots detected using the odyssey infrared scanner, the primary and secondary antibodies were prepared in a 1:1 ratio of odyssey blocking solution to PBST.

Low detergent lysis buffer (For MS):

It contains 0.01% triton X-100, 0.2mM HEPES, 10mM NaCl, 1.5mM MgCl2, 0.2mM EDTA), phosphatase inhibitors (10mM NaF, and 1mM NaV) and protease inhibitors (1mM PMSF, 10ug/ml each of leupeptin, aprotinin and pepstatin).

Propidium Iodide Staining Buffer:

0.1% Sodium Citrate, 0.1% Triton X, 100 µg/ml RNAse, 50-100 µg/ml PI.

Buffers for immunoprecipitation

RIPA buffer: 50mM Tris, 100mM NaCl, 1% (v/v) NP-40, 0.25% (w/v) Sodium deoxycholate , 1mM EDTA, 1mM EGTA along with phosphatase and protease inhibitors as applicable.

4.3 METHODS

4.3.1 Cell Lines and Cell Culture Medium

Murine B cell Lymphoma CH1 (TIB-221) cells were cultured in RPMI 1640 (Invitrogen) medium with 2mM L-glutamine adjusted to contain 1.5 g/L Na bicarbonate, 4.5 g/L glucose, 10 mM HEPES, 0.05 mM 2-mercaptoethanol (90%) ,1.0 mM Na pyruvate and 10% fetal calf serum.

Human cell lines Raji, Jurkat, HL60, U937 and THP1 were cultured in RPMI 1640 medium with 2mM L-glutamine adjusted to contain 1.5g/L Na bicarbonate,4.5g/L glucose,10 mM HEPES and 10% fetal calf serum.

Human Burkitts Lymphoma Namalwa was cultured in RPMI 1640 medium with 2mM L-glutamine adjusted to contain 1.5g/L Na bicarbonate,4.5g/L glucose,10 mM HEPES, and 1.0 mM Na pyruvate and 10% fetal calf serum.

Human cell lines A549, HeLa, HepG2, HEK293, U2OS, AW8507, MCF7 and Huh7 and were maintained in DMEM (Dulbecco's Modified Eagle Medium, Invitrogen) supplemented with 10% FCS. All cell lines were cultures at 37 C and 5% CO2.

4.3.2 Polyacrylamide Gel Electrophoresis and Western Blot analysis.

Western blot analysis was done to assess efficiency of siRNA knockdown. At appropriate times after knockdown i.e 24 hrs, 48 hrs, 72 hrs and 96hrs, aliquots of cells were collected, centrifuged, washed with 1xPBS and the cell pellets stored in liquid nitrogen. Just prior to electrophoresis, cells were lysed in lysis buffer (20mM HEPES, 10mM NaCl, 1.5nM MgCl2, 0.2mM EDTA, 0.5% Triton X-100, 0.5mM DTT, 1mM sodium orthovanadate, 1mM NaF, and a cocktail of protease inhibitors). Lysates were incubated on ice for 30 minutes followed by removal of the nuclear material and other debris through centrifugation. The supernatant was retained for western blotting or IP. The lysates were diluted by 2X SDS sample buffer before boiling, the detergent-

soluble proteins were then resolved by SDS–PAGE and transferred to nitrocellulose membranes. Membranes we blocked with Odyssey blocking buffer and cut to enable proteins of different sizes to simultaneously be probed with different antibodies. Following treatment with the suitable primary antibody, membranes were probed with IRDye 800/700 conjugate anti- rabbit/anti-goat/ anti-mouse (depending on the requirement) at a 1:15000 dilution, detected using odyssey infrared scanner and analysed using Odyssey 2.1 software.

4.3.3 siRNA Library Screening

The primary siRNA screens were performed using siRNA libraries targeting mouse Kinases (Mouse Kinase siRNA SET V1.0) and Phosphatases (Mouse Phosphatase siRNA SET V1.0) containing 2 siRNA duplexes for each target gene (Qiagen).The siRNAs were designed using the innovative HiPerformance siRNA design algorithm. A pool of the 2 siRNAs was used and a total of 758 Kinases and 294 phosphatases were screened.

The Validation screen for the shortlisted Kinase hits was performed using pools of 3 individual siRNA duplexes against each gene target from the MISSION siRNA Mouse Kinase Panel (Sigma).The screen was performed similar to the primary screen. For Validation of Phosphatases, siGENOME (Dhramacon) 'smartpool' siRNAs containing 4 siRNAs per target gene were used.

***Standardization of the screening procedure*:**

(i) ***No. of Cells Plated/ Well***

Experiments were done to determine the optimum number of cells to be used per well for transfection and subsequently for cell cycle analysis. TIB-221 cells were seeded at

varying numbers ranging from 10,000/well to 50,000/well and transfected with GFP-specific siRNA. Subsequently, cell viability was determined using trypan blue exclusion at 24, 48, and 72 hours. A seed density of 20,000 cells/well was found to be optimal with a cell viability that was consistently >85%.

(ii) *siRNA Transfection Conditions*

Transfection efficiency: To ensure efficient transfection Alexa 488 labelled siRNA was used to transfect CH1 cells. 20,000 cells were plated in a 96-well plate and transfected using HiPerfect transfection reagent. Transfection was carried out as per the manufacturer's protocol. The efficiency was monitored at up to 72 hours post transfection using confocal microscopy and was found to be >98%.

Optimal siRNA concentration and the kinetics of Knockdown:

To determine optimal knockdown conditions, siRNAs against six representative proteins were chosen. Using the optimised cell numbers (20,000/well) experiments were set up with different concentrations of siRNA ranging from 50 to 400nM using HiPerfect transfection reagent as per manufacturer's protocol. The effect of different siRNA knockdowns on protein level was determined by Western blot analysis in cell lysates obtained at 24, 36, 48, 72 and 96h after transfection. Across experiments the siRNA concentration of 200nM was found to be optimal, giving a knockdown of between 70-80%, which was retained up to 96h post transfection. The western blots for two representative molecules BTK and PLK1, along with those for three validated screen targets are shown in Figure S1A. The relatively higher concentration of siRNA required is likely due to the fact that these cells grow as a suspension.

4.3.4 The siRNA Screening Procedure

The mouse Kinase and phosphatase libraries (Qiagen) containing 2 siRNAs per target gene were obtained in a 96 well format. The siRNAs were diluted to a working concentration of 10μM.

CH1 mouse B cell lymphoma cells were transfected with the pool of 2 siRNA duplexes per target gene, using HiPerfect transfection reagent as per the manufacturer's protocol for transfection of suspension cell lines (Qiagen). Briefly, 20,000 cells of mouse lymphoma cell line CH1 in 25uL of RPMI were seeded into each well of a flat bottom 96 well plate. The cells were incubated at 370C and 5% CO2 for 30 minutes. Meanwhile, 25uL (per well) of transfection mix, containing RPMI medium and a 1:1 volume ratio of siRNA and transfection reagent Hiperfect i.e. 200nM (3ul) siRNA and 3ul HiPerfect were prepared in parallel U-bottom 96 well plates. To each well in the plate containing cells, 25ul of transfection mix was added. The mixture was pipetted a few times to ensure proper mixing of transfection complexes with the cells. The plates were incubated at 370C and 5% CO2. At 6 hours post transfection the volume of each well was made up to 150 μL with complete medium (RPM1 with 10% FCS, 1mM sodium Pyruvate, βme).

At 24 and 48 hours post transfection the medium of the plate was replaced with fresh complete medium. Our transfection conditions ensured that >98% of the cells were transfected with the siRNA (Figure S1B). Further, by using siRNA pools against a representative set of cellular kinases we determined that specific depletion was evident by 36 h after transfection, and that this attained maximal levels by 48h. Depleted protein levels then persisted up to 96h of the culture (Figure S1A).

Since during standardization experiments the target protein levels were found to be the lowest at 48 hours post transfection, the cell cycle analysis was performed at 72 hours

(which is 24 hours after the point at which protein levels are minimum). Each plate had 5 negative control wells which included cells treated with scrambled siRNA (2 wells), GFP siRNA (2 wells) and 1 well contained mock transfected cells (i.e. only HiPerfect). All the screens were performed in duplicates.

4.3.5 Propidium Iodide staining and sample acquisition using flow cytometry

For propidium iodide staining and DNA content analysis the cells were centrifuged in the 96 well plates using a plate rotor and the medium was aspirated. The cells were stained for 30 minutes at 40C with 150 uL propidium iodide staining solution containing 0.1mg/mL Propidium Iodide (Sigma), 3ul/mL TritonX -100 (Sigma), 1 mg/mL sodium citrate (Sigma) and 20ug/mL RNase (Sigma). Plates which were not acquired immediately were kept at 40C. The samples were acquired in a 96-well plate format using an automated BD FACSCalibur HTS or High Throughput Sampler (Beckton Dickinson, FACSCalibur).

4.3.6 Primary and secondary screen data analysis and hit selection

4.3.6.1 Cell cycle analysis

The data obtained was analysed using FlowJo software and the DNA histograms obtained were analysed to quantitate the subG1, G1, S and G2 populations. During data analysis we observed minor shifts in G1 and G2 peak positions and widths across cell cycle histograms. This prevented a reliable quantification and comparison of S and G2 phases of the cell cycle using automated platforms such as those available with FlowJo. The cell cycle histograms were therefore analysed using manual gating of the sub-G1, G1, S and G2 phases. The G1, S and G2 phases were analysed after gating out

the subG1 population such that their quantified values represent percentages of the total live cell population.

For each plate the gates were defined for the negative control samples and the same gates were copied to the other histograms and used for analysis. Although with respect to the control sample there were some well-dependent (instrument based) and sample dependent shifts in peaks of some samples in each plate. To increase precision of analysis small adjustments in the gates of individual peaks of each sample allowed reliable quantification of cell cycle phases. The same procedure was used for the analysis of the data obtained from the second screen.

4.3.6.2 Screen data analysis

To allow the comparison of data sets from different plates the data was normalized by calculating and representing each quantified parameter sub G1, G1, S and G2 populations into Z scores (also called normal scores). Each Z-score is a dimensionless quantity which depicts the deviation of the data point from the mean of the negative controls in units of the standard deviations of the negative control. The Z scores were calculated for each parameter with respect to the mean and standard deviations of the 5 negative controls of its respective plate. Each Z score was calculated as follows:

$$Z\ score=(x-\mu)/\int$$

Where x is the G1/S/G2 population to be standardised, μ is the plate mean/mean of controls and $\int$ is the standard deviation of the plate/plate controls. Both the screens were performed in duplicates, therefore for each siRNA knockdown there were 2 Z scores. The Z scores were obtained for each duplicate and the mean was calculated. For the first screen a z score cut off of $Z > \pm 2.5$ was chosen and 83 kinases and 22 phosphatases were shortlisted as our primary hits.

The molecules shortlisted were screened in the second screen using new siRNAs obtained from an alternative source and having sequences which were different from those used in the primary screen. All siRNAs which resulted in a Z score cut off of ± 2.0 were shortlisted as the final hits.The final list of hits included 38 kinases and 5 phosphatases as the final hits.

An analysis of the distribution of scores obtained across the entire screen yielded a normal distribution profile with a very low 'skewness' value (Figure S2). Further, a comparison of *Z* scores between the replicates revealed that the potency of identified targets remained largely comparable (Figure S3). Also, both intra- and inter-plate variations for the control wells were low with a percent standard deviation of the median value of <10%. In order to assess the overall reproducibility of the screen we calculated the z-factor with PLK1 as positive control. This resulted in a z-factor of 0.56 which is considered to represent an overall high quality for the assay conditions. These collective results, therefore, confirm the robustness, reproducibility, and sensitivity of our primary screen. (Figure S2, S3).

4.3.7 Population Doubling Time Experiments (PDTs).

To assess the effects of siRNA knockdown on cell cycle dynamics and population doubling time, the following experiments were carried out with the validated siRNA hits. The siRNA transfections were set up in duplicates and carried out in a 96-well plate format as described above. At 48 hours post knockdown (point of maximum knockdown) the first population doubling sample was taken and the cells were counted using trypan blue exclusion to give the "day 0" counts. Cells were counted similarly at 72 hours and 96 hours to give the "day 1" and "day 2" counts respectively. Here it should be emphasized that the loss of function due to siRNA knockdown is observed

up to 96 hours (Figure S1A), which is within the recommended time frame for siRNA mediated transient knockdown experiments [4]. Population doubling times (PDTs) were calculated using the formula (log 2/log N) hr, where N is final cell count/initial cell count (Nf/Ni) and hr is time (in hours) between initial and final cell count. The population counts were used to plot population growth curves, and also to calculate the PDTs.

4.3.8 Calculation of Residence Times

Using the population doubling times and the cell cycle histograms, the average residence times (RT) of cells in different cell cycle phases were calculated as described earlier [5]. The residence time in a particular phase G1/S/G2 (hrs) = % cells in G1/S/G1 x PDT (hrs). A table indicating the population doubling times and respective residence times is given in Table S3.

PDTs were determined for both mock-silenced, as well as 'hit'-silenced CH1 cells. The doubling time for mock-silenced CH1 cells was 26 hrs (SD ±3 hours), and the individual phase-specific RTs were calculated to be G1-8.41 (SD ±1.5), S-13.48 (SD ±1), G2-4.11 (SD ±0.5). The results discussed in the text clearly indicate that each siRNA shortlisted as a 'hit' affected one or more cell cycle phase in terms of RT extensions. Since our interest was to understand regulatory networks involved in G1 and S phase control, we grouped the hits based on residence times alone. Many patterns of perturbation in RT were observed in the resulting data. The residence times calculated are a reflection of the Z scores. A normalization procedure was carried out to negate the plate to plate fluctuations of control cell percentages and allow comparison between the different PDTs.

4.3.9 Isolation of Human Naive B cells

Normal B-lymphocytes were purified from the blood of donors by using the B-cell negative selection kit from Dynal and following the protocol recommended by the manufacturer.Purity of these cells were >90% as assessed by CD19 staining. Antibodies were purchased from Cell Signaling Technologies. Purity of these cells were >90% as assessed by CD19 staining. Antibodies were purchased from Cell Signaling Technologies.

4.3.10 Pharmacological Inhibitor Assays

For the inhibitor assays, the chemical inhibitors were reconstituted in DMSO. To test the effect of specific inhibitors on cell cycle, cells were plated in a 24 well plate. Cells were treated with increasing concentrations of inhibitors, in multiple of the IC50s for the specific kinases. The cells were treated with inhibitors of specific kinases (or kinase combinations) and the cell cycle analysis and effect on cell death was assayed using FACS at 24, 48 and 72 hours post treatment.

4.3.11 Sorting of cells in the G1 and S phases by flow cytometry

To obtain separate populations of CH1 cells in the G1 and S phase, a fluorescence activated cell sorting (FACS) approach based on DNA content was used. For this, cells were stained with Vybrant Dye Cycle Orange stain (Invitrogen), a cell permeant nucleic acid stain that stains live cell populations. Cells were sorted on a FACSAria flow cytometer (Becton Dickenson), using 488nm excitation and 564-606nm emission (bandpass filter).

Cytometer verification: The resolution and linearity of the FACSAria instrument was checked using Chicken erythrocyte nuclei (CEN) stained with PI (DNA QC Particle

Kit, BD Biosciences).Calf thymocyte nuclei (DNA QC Particle Kit, BD Biosciences) stained with PI were used to check doublet discrimination on FL2 (pulse width versus pulse area). The fluorescence sensitivity was checked using SPHERO rainbow calibration particles (BD Biosciences).

Data acquisition and sorting: A viable cell sorting based on DNA content was performed in order to separate the G1 phase and S phase cell populations of CH1 cells. A cell suspension of 10 x 106 cells/mL was taken for sorting and a total of nearly 500 x 106 cells were sorted. Data from the FACSAria was acquired using the BD FACSDiva software. Data from FL2 (pulse width and area) was acquired with linear amplification. A two-way cell sorting was performed by a dual lazer BD FACSAria cell sorter (BD) equipped with a 488 nm argon and 633 nm helium neon lazer The 488nm lazer was used to excite the Vybrant dye cycle orange dye and the emissions were detected with 564-606nm band width filter (PE) and 600-620nm (PE Texas Red). The cells were sorted into 15ml polystyrene tubes (falcon) containing 500 µl of complete medium. Water recirculation for cooling of sample collection tube was used to maintain the sorted cells at 4°C. After sorting, the cells were immediately centrifuged at 3000 rpm. The supernatant was discarded and the cell pellets were processed for protein extraction and mass spectrometric analysis. The purity of the sorted populations was >95%.

4.3.12 Mass spectrometry based determination of the cellular phosphoproteome.

2 x 108 cells in each phase (G1 and S) were taken separately and lysed in low detergent lysis buffer (containing 0.01% triton X-100, 0.2mM HEPES, 10mM NaCl, 1.5mM MgCl2, 0.2mM EDTA) with phosphatase inhibitors (10mM NaF, and 1mM NaV) and protease inhibitors (1mM PMSF, 10ug/ml each of leupeptin, aprotinin and

pepstatin) for 30 minutes at 4^0C. The samples were centrifuged for 5 minutes at 3000 rpm and supernatants containing cytoplasmic cell extracts were collected separately. The nuclear pellets obtained were washed once with low detergent lysis buffer. They were lysed with Buffer B (20mM HEPES, 20% glycerol, 0.8M urea, 1.5mM MgCl2, 0.2mM EDTA and 0.2% Triton X-100, and all the above mentioned protease and phosphatase inhibitors) for 1 hour at 40C. The sample was centrifuged for 10 minutes at 10,000 rpm. The supernatants were collected in separate eppendorfs. The nuclear and cytoplasmic protein contents for G1 and S phase were determined using a commercial Bradford assay reagent (Bio Rad). From the G1 and S phase samples, aliquots of 100μg cytoplasmic protein fraction and 100 μg of nuclear protein fraction were taken for iTRAQ labeling. For this analysis we took cytoplasmic extracts of the G1 and S phase obtained from four different sorting experiments (i.e.four replicates), and the corresponding nuclear extracts were taken from two of such experiments (two replicates). Each of these were separately digested with trypsin, and the resulting peptides labeled with the individual isobaric components of either the iTRAQ 8-plex (for the cytoplasmic fractions), or the iTRAQ4-plex (for the nuclear fraction) kits as described below:

1. *Acetone Precipitation*

The acetone and sample eppendorfs were chilled to 4°C. To each sample six volumes of cold acetone was added and the eppendorf was inverted 3 times. The samples were incubated overnight at -20°C to allow protein precipitation. The eppendorfs were centrifuged at 10,000 rpm for 20 min at 4°C and the acetone was removed to get the precipitated pellet.

2. *Reducing the proteins and Blocking the Cysteines.*

To each tube containing 100 µg of precipitated protein pellet, 20 µl of dissolution buffer was added. Then 1 µl of denaturant and 2 µl of reducing agent were added, the tubes were vortexed and incubated at 60°C for 1 hour. To each tube 1 µl of cysteine blocking reagent was added and tubes were incubated at RT for 10 min.

3. *Tryptic digestion*

A fresh vial of trypsin was reconstituted with 25 µl of Milli-Q water. Each sample was digested with 10 µl trypsin (12.5ng/µl) solution overnight (12-16 hours) at 37°C.

4. *Labeling with iTRAQ reagents.*

The vials of iTRAQ reagents were allowed to equilibrate to RT. Each vial was reconstituted by the addition of 70 µl of ethanol. The 8-plex iTRAQ labels were used to label the four replicate samples each of the G1 and S phase cytoplasmic extracts, and the 4-plex iTRAQ labels were used to label the two replicate samples each of the G1 and S phase nuclear extracts. The labelling was done as follows: 8-plex iTRAQ tags 113, 114, 115, 116 for G1 cytoplasmic; 8-plex iTRAQ tags 117,118,119 and 121 for S cytoplasmic; 4-plex iTRAQ tags 113 and 114 for G1 nuclear; 4-plex iTRAQ tags 115 and 116 for S nuclear sample. The contents of each iTRAQ vial were accordingly transferred to the appropriate sample tubes and incubated at RT for 1 hour. The G1 and S cytoplasmic fractions labelled with 8-plex iTRAQ labels were pooled into one tube. The G1 and S nuclear fractions labelled with 4-plex iTRAQ labels were pooled into another tube.

The contents were dried on a SpeedVac. Phosphopeptide enrichment was performed using TiO2 beads (ZirChrom Separations Inc.).150ug of beads were taken in an eppendorf and equilibrated with loading buffer (80%ACN/0.1% Formic acid with 350mg/ml 2, 5-Dihydroxybenzoic acid or DHB) twice. The final wash was given with 2% Formic acid (FA). The samples were acidified with 2% FA and then dissolved in

500ul of loading buffer. Each sample was mixed with beads and incubated for 45 min with rotation. They were centrifuged at 2000 rpm and the supernatant was collected in a separate eppendorf. The beads were washed with loading buffer followed by 80% ACN/0.1%FA without DHB. The bound peptides were eluted first with 25ul of ammonia water, pH 10.5 in 20% ACN and then with ammonia water, pH10.5 in 40% ACN. Eluates were pooled and dried using a SpeedVac. Dried eluate was reconstituted into 20ul of 5mM ammonium formate with 30% ACN, pH3.0. They were fractionated manually on SCX ICAT cartridges (Applied biosystems) into 5fractions using a gradient up to 500mM ammonium formate (5mM, 100mM, 200mM, 300mM and 500mM). All the fractions were then lyophilized.

Each fraction was reconstituted in 3% ACN/0.1% FA and was separated by reverse phase nano-LC (Agilent 1100 series; column: Zorbax 300SB-C18, 3.5um, 150 x 0.075mm). The peptides were eluted using a 90 min gradient with solvent A (H2O/ACN, 97/3 vol/vol) and B (H2O/ACN, 10/90 vol/vol): 10 min from 0% to 10% B, 45 min from 15% to 35% B, 60 min from 35% to 60% B and 65 min from 60% to 90% buffer B. Eluted peptides were directly analyzed through electrospray ionization (ESI) on a hybrid quadrupole time-of-flight mass spectrometer (QSTAR XL, Applied Biosystems). The MS/MS spectra of three most intense peaks with 2 to 5 charges in the MS scan were automatically acquired in information-dependent acquisition with previously selected peaks excluded for 90 seconds. In addition to this, phospho-peptides were also analyzed by ESI-LC-MS/MS on a triple quadrupole (Qq linear ion trap; QTRAP 4000, Applied Biosystems) mass spectrometer, in a neutral loss scanning mode.

Data analysis

The MS/MS spectra obtained were extracted and searched against mouse protein database (Swiss Prot 51.6) using MASCOT (Matrix Science) with trypsin specificity, 2 missed cleavages, Methylthio as fixed modifications and iTRAQ8plex (K), iTRAQ8plex (N-term), iTRAQ8plex (Y), iTRAQ4plex (K), iTRAQ4plex (N-term), iTRAQ4plex (Y),Oxidation (M), Phospho (ST), Phospho (Y) were taken as variable modification. Precursor mass tolerance was set at ±0.6 Da and fragment mass tolerance was set to ±0.4 Da. Only MS/MS spectra that exceeded the MASCOT identity threshold at the 95% significance level were taken for further analysis. These spectra were manually examined for the presence of characteristic peptide fragment ions (i.e. band y-type of fragment ions), and the identification was considered positive when the MASCOT scores were higher than the corresponding identity threshold score, and when sufficient peptide sequence coverage (at least 50%) was observed. Here, an additional criterion imposed for selection was that the phosphopeptide must be present to comparable levels in all the four replicates of the phase-specific cytoplasmic sample, or both replicates of the corresponding nuclear extract. The mean of the label-specific intensities of each subset (G1/S) was taken and were compared to determine phase-specific changes in the level of each phosphopeptide. Table S4 provides the combined list of the phosphopeptides identified on both the mass spectrometers, along with the description of the parent protein from which they were derived.

4.3.13 Network analysis methods

Rationale for assigning cell cycle phase specific source and targets.

We identified thirty eight kinases and five phosphatases as hits from our siRNA screen (Figure 5.1B, 5.1D). Kinases and phosphatases are generally considered as molecules involved in signaling mechanisms. We wanted to classify the siRNA screen hits as

signaling molecular sources affecting cell cycle. Initially we could classify the hits as molecules causing *extension* or *no extension* in cell cycle duration from the population doubling time data (Figure 5.1D). Then we further classified these hits as affecting a particular phase based only on extent of the increase in the residence time of a particular phase, but not reduction in the residence time of a particular phase to simplify further data analysis. These hits must affect the cell cycle machinery either directly or indirectly to induce such drastic changes in the residence times of each cell cycle phases.

Based on the above criteria, we defined 12 molecules as sources for causing G1 extension, 9 for S phase extension, 4 for G2 phase extension on the basis of their residence time (Table S3). We further classified 15 molecules as both G1S phase specific sources since they affected both the phases to an equal extent. Similarly we had 1 molecule in G1G2 and 2 in SG2 sources. By this process we defined sources for cell cycle regulation based on our siRNA screen results and population doubling experiments.

By extensive literature curation on cell cycle we shortlisted molecules involved directly in the regulation of cell cycle (Table S5). The target list consisted of 93 molecules grouped into 3 classes, those which are involved in regulating the G1 phase, the S phase and the G2 phase.

Many cell cycle regulators are known to play a role in G1 to S phase progression or in both G1 and S phase regulation. These have been listed as possible G1 as well as S phase targets. These molecules include core cell cycle regulators such as cyclins and CDKs, transcription factors, molecules involved in DNA replication and kinases/phosphatases which are known to play a role in cell cycle. Some of the molecules shortlisted as targets were also screened in the siRNA libraries and were

shortlisted as the screen and PDT hits. These molecules may therefore appear as sources, targets and possible intermediates of the shortest paths during network analysis.

This set of 93 molecules was considered as molecular targets controlling cell cycle. Since we wished to study the phase specific regulation of cell cycle we categorized these molecules as G1-targets (51), S-targets (53) and G2-targets (29) based on their involvement in regulation of a particular phase. We combined both G1-targets and S-targets molecules list to get G1S-target molecules. From the experimental and literature information we could define sets of source-target relationships for phase specific cell cycle regulation.

***Construction of mammalian protein-protein interaction database*:** We downloaded seven different mammalian protein-protein interaction (PPI) databases (i.e., reported to be present in rat/mouse/human) from the following sources, BIND (Bader, Donaldson et al. 2001), BioGRID (Breitkreutz, Stark et al. 2008) , IntAct (Kerrien, Alam-Faruque et al. 2007), MINT (Chatr-aryamontri, Ceol et al. 2007) , HPRD (Keshava Prasad, Goel et al. 2009), NetworKIN (Bakal, Linding et al. 2008). These databases were consolidated by combining rat/mouse/human gene name to yield a single PPI database with ~12000 nodes and ~50000 edges. Since we wished to incorporate all direct experimental biochemical interactions with high confidence, curated from low throughput experiments, we incorporated filter criteria to prune out those interactions (edges) that were not reported by at least two different experimental methods. We also removed interactions predicted purely by *in silico* and high throughput methods. This yielded us a high confidence PPI network of ~5200 nodes and ~12000 edges

essentially including all known direct experimental interactions in mammalian systems.

Shortest path analysis and ranking of cell cycle phase specific intermediates: We defined source-target relationships involved in regulation of each of the cell cycle phases as *viz.* $G1_{sources}$-$G1_{targets}$ as G1 phase regulators, $S_{sources}$-$S_{targets}$ as S phase regulators and $G1S_{sources}$-$G1S_{targets}$ as G1S phase regulators. We wished to identify key unique regulatory intermediates (IMP nodes) for a given source-target subset for regulating a particular phase of cell cycle.

Algorithm: We defined the 2 path length PPI network to be our core network (N). Given a network N with a subset of nodes as sources (S) and targets (T) where S, $T \subset N$ we had to identify the intermediate nodes and quantify their occurrences for all pairs of S and T.

1. Trace all possible shortest paths between all pairs sources S and targets T.
2. Find the intermediate nodes (I) that occur in all possible shortest paths for a given S and T. Identify intermediate nodes I for all pairs of S and T.
3. Count the number of times each of the intermediate I occurs in shortest paths for all pairs of S and T.
4. Compute a z-score to normalize the occurrences of an intermediate with respect to the number of the source (S) and target (T) nodes for a given source-target subset.

$$IS_{S\text{-}T} = \frac{nI_{S\text{-}T} - \langle I_{S\text{-}T} \rangle}{\sigma I_{S\text{-}T}}$$

Where, $nI_{S\text{-}T}$ is the number of times the intermediate node I occurred in a given source- target subset, $<I_{S\text{-}T}>$ is the median of number of occurrences of all intermediates nodes in the source-target subset, $\sigma I_{S\text{-}T}$ is the standard deviation of the number of occurrences of all intermediates nodes in the source-target subset.

5. This process of identifying the intermediates and scoring was repeated for all source-target subsets.
6. We employed a cut off criteria of intermediate score $IS_{S\text{-}T} \geq 1.28$ (p-value < 0.1) to shortlist only significantly high scoring intermediates for our further analysis.

Comparison and short listing of intermediates with control subsets.

Since we were interested in identifying the unique phase specific IMP nodes for G1, S and G1S phases of cell cycle we included a control criteria. Wherein molecules present in each group was compared with the distinct source-target group (as shown in the table below). The overlapping or common molecules between these groups were filtered to retain only the unique nodes for each phase. we believe this analysis would further yield us highly phase specific IMP nodes for subsequent network analysis.

Phase specific regulatory source-target subset	Distinct control Networks
$G1_{sources}$-$G1_{targets}$	$S_{sources}$-$S_{targets}$ $G2_{sources}$-$G2_{targets}$
$S_{sources}$-$S_{targets}$	$G1_{sources}$-$G1_{targets}$ $G2_{sources}$-$G2_{targets}$
$G1S_{sources}$ -$G1S_{targets}$	$G2_{sources}$ -$G2_{targets}$

***Extraction of phase specific IMP node subnetworks*:** We wanted to extract subnetwork regulating cell cycle phases incorporating the phase specific regulatory intermediates we identified earlier. We considered the phase specific intermediates as the seed list and traced all possible shortest paths of path length 2 between all pairs of nodes in the seed list. We identified the nodes present in the shortest paths of 2 path length between all pairs of seed nodes and merged them to create a connected subnetwork. This process was employed for all the three subsets of phase specific intermediates yielding three such networks of 2 path length.

***Identification of cell cycle phase specific regulatory motifs*:** Once we identified cell cycle phase specific subnetworks we wanted to resolve it further by manual literature curation on all the edges present in the subnetwork. By extensive manual literature curation we could include directions to the edges for those the experimental evidence was present. We compiled a flat file with human Entrez gene name of source protein as first column, target protein as second column, + for activation interaction, - for inhibition and 0 for neutral interaction as third and the PMIDS for the interaction as a separate fourth column. This network was a mixed graph of both directed and undirected edges for each of the IMP node network (Table S7).

We used this network for identification of network motifs involving the phase specific IMP nodes. We used SNAVI, (Ma'ayan, Jenkins et al. 2009) a software tool for network analysis, for motif detection in our literature curated IMP node subnetworks. With the phase specific regulatory intermediate nodes as seed list and the corresponding IMP node subnetwork as background we used SNAVI to identify the regulatory motifs for each phase separately. This yielded cell cycle phase specific

motifs for G1, S and G1S phases, *viz.* scaffolds, bifans, feed-forward loops, feed-back loops (Figure 5.8).

Cell cycle phase specific regulatory modules and regulators: We merged the motifs we identified for each phase to obtain a module for regulation of each phase. Since we wanted to identify the least redundant node in the modules we calculated centrality parameters betweenness and stress. And we identified a set of key regulatory nodes for each cell cycle phase based on its very high stress and betweenness centrality compared to other nodes (Figure 5.9). For this we employed a selection criteria where we shortlisted only those nodes exhibiting stress and betweenness greater then 2.5 times the mean of all nodes in the corresponding phases. This signifies a z score >1.28 corresponding to a p value <0.1. This we consider would identify the statistically most important nodes for further examinations in the phase specific modules.

Network Randomization of IMP nodes sub networks: In order to validate the motifs we identified in the phase specific IMP nodes networks to be functionally significant we merged the IMP node networks since there are overlapping interactions to compile a single network with mixed edges. This network was randomized over 100 times by shuffling the edge properties. Motifs involved by the phase specific IMP nodes were scanned for in the randomized networks using SNAVI. The figure 5.8 shows the results of the randomization exercise validating the motifs identified statistically.

Chapter 5

Results

Chapter 5: Results

5.1 An RNAi screen targeting cellular kinases and phosphatases identifies regulators of the cell cycle.

To identify signaling molecules that regulate cell cycle progression we performed a siRNA screen against all known kinases (758 proteins) and phosphatases (294 proteins) in cycling cells of the murine B lymphoma cell line CH1. (For standardization of screening procedure please refer to Experimental procedures and Supplementary Figure S1). Changes in cell cycle phase distribution at 72h after siRNA transfection were determined by staining cells with propidium iodide, followed by a quantitative analysis for the DNA content by flow cytometry (Experimental procedures). Here, only those effects that perturbed the cell cycle without significantly affecting cell viability were considered. A primary screen followed by a validation exercise (Fig. 5.1 and Experimental Procedures), identified 38 kinases and 5 phosphatases whose silencing significantly affected the cell cycle phase distribution (Supplementary Table S1). Microarray analysis subsequently confirmed that the genes coding for all of these validated target proteins were indeed expressed in CH1 cells (Supplementary Table S2).

A hierarchical clustering analysis of the nature of effect caused by depletion of these proteins revealed that the most common one was that of a simultaneous increase in the pool size of cells resident in the G1 and S phases (Fig. 5.2A). In addition, there were also a significant number of cases where siRNA treatment led to an increase specifically in either the G1 (12), or the S (9) phase populations, whereas only a limited number of siRNAs had any detectable affect on the G2 phase (Fig. 5.2A).

Figure 5.1

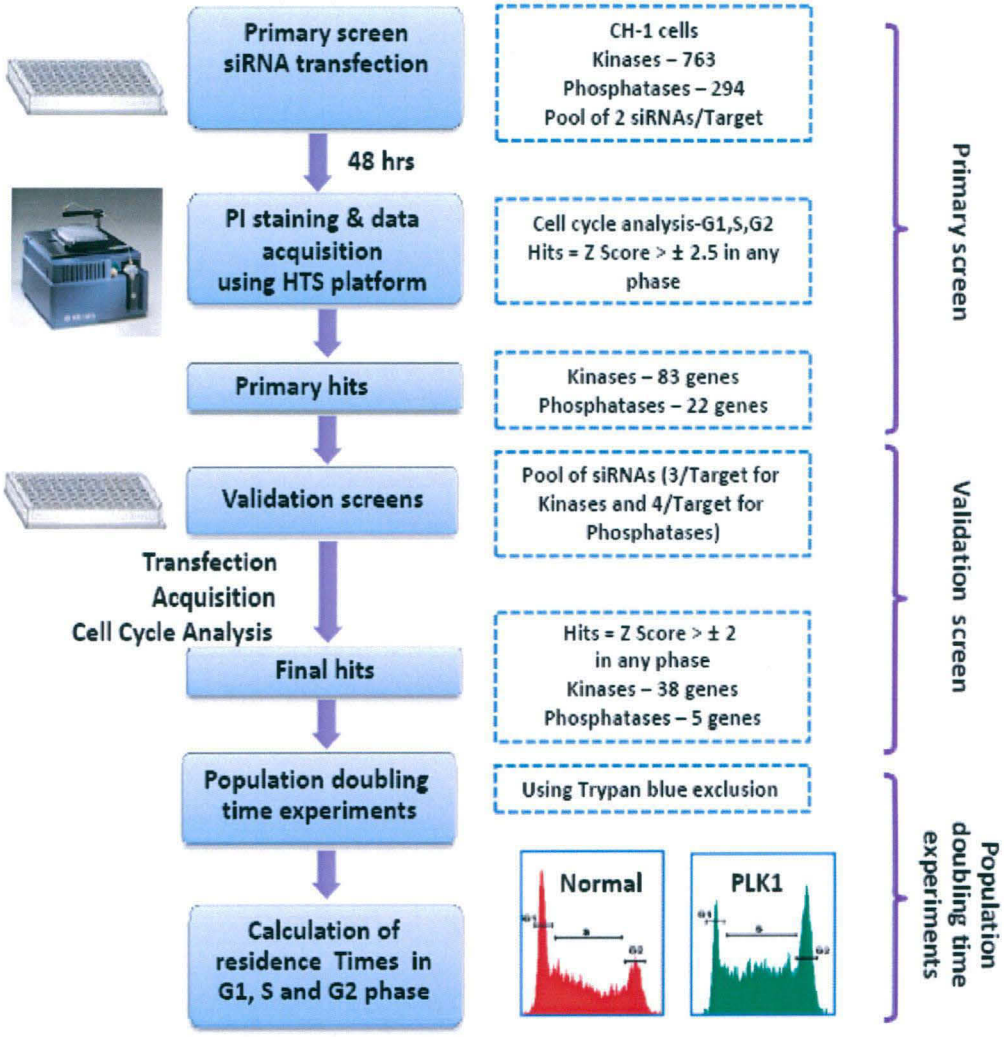


Figure 5.1. Design of the siRNA screen and the cell cycle phase-specific effects of the identified targets.

The Panel provides a schematic of the strategy employed to identify genes involved in cell cycle regulation. The overall z-factor obtained for our screen was 0.56, confirming its high overall quality. Details of the design, assay validation, and target verification are provided in the text, Experimental Procedures and Supplementary Figures S1, S2 and S3.

Figure 5.2

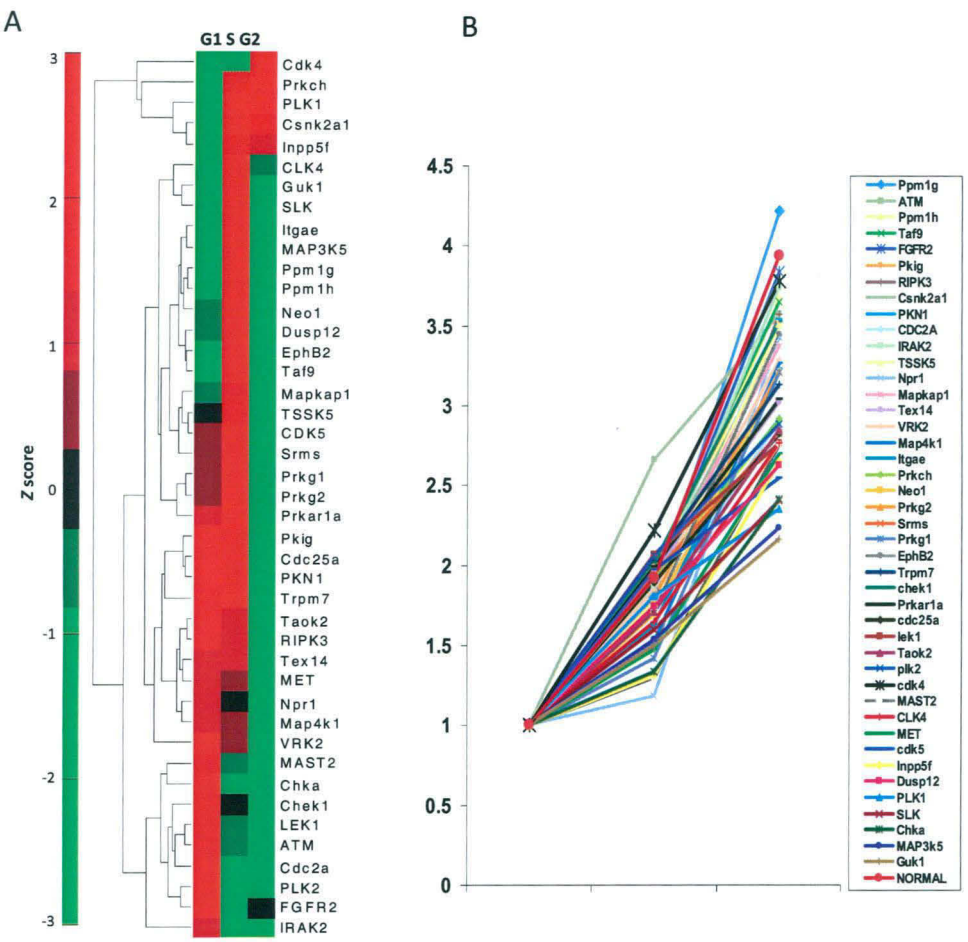


Figure 5.2. The cell cycle phase-specific effects of the identified targets.

Panel A shows the hierarchical clustering of the siRNA hits according to the observed phenotype on cell cycle. The dendrogram represents the various clusters of genes identified by our screen for the observed cell cycle distribution pattern. This pattern is based on the normalized z-scores of each of the siRNA hits in the respective cell cycle phases.

Panel B depicts the population growth curves of CH1 cells obtained in response to transfection with either non-silencing siRNA (Normal), or with siRNA against each of the validated hits. Here, cells were plated at 48 h after siRNA transfection (Day 0), and results are plotted in terms of the fold increase in cell population at the indicated time points. Values are the mean of three separate determinations. The subset highlighted by the bracket indicates those target-specific siRNAs that produced a significant increase in the population doubling times ($p < 0.05$, denoted by the star). Details of calculation of the population doubling times are provided in Experimental Procedures.

5.2 Silencing of signaling intermediates induces differential effects on cell doubling times.

An examination of the effects of target-specific depletion on cell population doubling times (PDTs) yielded a diverse spectrum of effects that ranged from negligible, to a marked reduction in the proliferation rate (Fig. 5.2B). While the PDT of control cells (i.e. transfected with non-silencing siRNA) was 26±3 h, 16 of the 43 siRNAs tested produced a significant extension with PDTs of 30h or more. The individual PDT values were then employed to calculate the residence time (RT) of cells in G1, S, and G2 phases under each of the knockdown conditions (Experimental Procedures and Supplementary Fig. S4), and a comparison of these results, relative to the control values, are shown in Figure 5.3. It is evident that each siRNA pool caused specific perturbations in one or more of the individual phases of the cell cycle, although the effects differed widely between the various knockdowns. For example, the G1 phase RT varied from about 2-fold lower to about 2-fold greater than that in control cells (Fig. 5.3). A faster progression through the G1 phase (i.e. reduced RT) was seen in response to treatment with five target-specific siRNAs, whereas the G1 RT was markedly increased in the case of 22 siRNAs. No significant effect was observed for the remaining 16 siRNAs (Fig. 5.3). A similar degree of variation was also observed for the S phase, whereas the G2 phase was only sensitive to 19 of the siRNAs tested of which 8 caused an increase in its RT (Fig. 5.3).

To simplify our subsequent analysis, we categorized the target-specific siRNAs solely on the basis of the phase-specific extensions in RT that they induced (Fig. 5.3). That is, although in some cases extension in RT of a given phase was compensated for

Figure 5.3

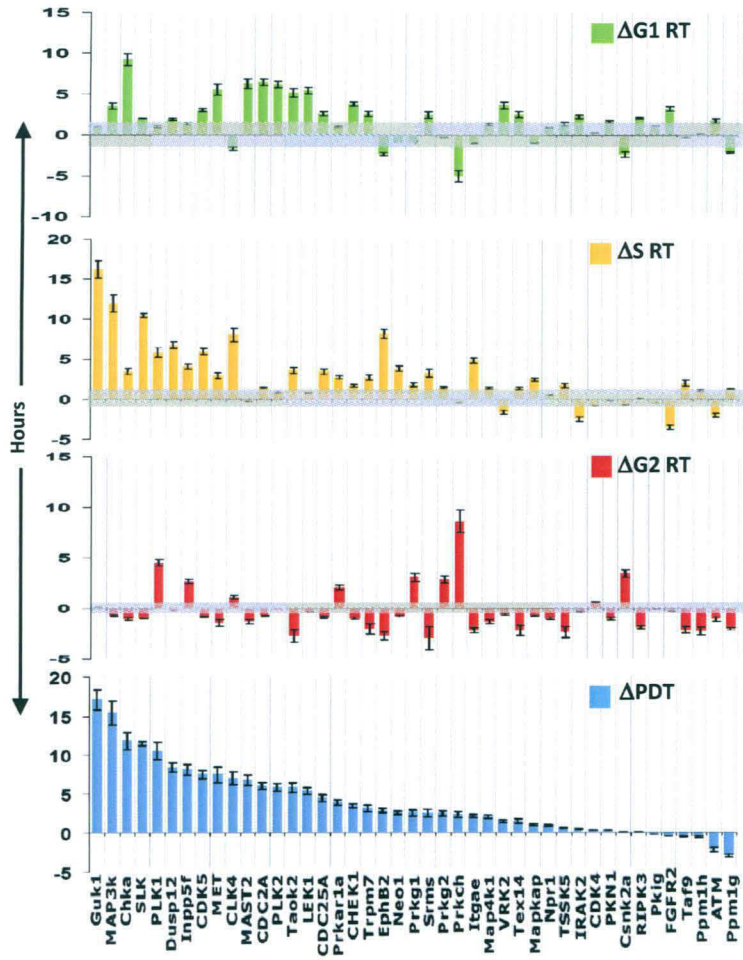


Figure 5.3. The cell cycle phase-specific effects of the identified targets.

Figure 5.3 provides a further analysis wherein the results in panels B and C were combined to determine the effects of siRNA-mediated target silencing on residence time (RT) of the individual cell cycle phases. Changes (Δ), relative to the corresponding values in GFP-silenced cells, in PDT and RT of the individual phases are shown here as the mean (± SD, grey bar) of three experiments. Details of their calculation are provided in Supplementary Figure S4 and Experimental procedures.

by shortening of another phase, the latter effect was not taken into account. Accordingly, the largest group of siRNAs (15) corresponded to those that induced a significant extension of both the G1 and the S phases (Fig. 5.3). Extension in RT of only the G1 phase was observed for an additional group of 12 cases, whereas another set of 9 siRNAs specifically extended the S phase alone. Perturbations prolonging the G2 phase were seen in response to 4 siRNAs and the remaining three target-specific siRNAs affected either the S and G2 (2), or the G1 and G2 phases (Fig. 5.3). These results are consistent with the known biological roles of these target proteins in the literature (Supplementary Table S3).

Since the role of mitogenic signals is to move cells through the G1 phase and then ensure their entry into the S phase (Jones and Kazlauskas 2000; Sears and Nevins 2002), our subsequent analysis concentrated on only those targets where depletion resulted in an extension in RT of either specifically the G1, the S, or both of these phases simultaneously. This latter group of targets was termed as those extending RT of the G1S phase and available literature suggests that it likely represents a window extending from the later stages of G1, to initiation of the S phase (Jones and Kazlauskas 2000).

5.3 Functional relevance of the targets identified by the siRNA screen.

To assess the significance of the hits identified by our screen, we examined the expression profile of these target proteins across different tissue and cancer cell lines. Expression data from GNF BioGPS, a database compiled by the Genomics Institute of Novartis Research Foundation was downloaded and a hierarchical clustering analysis performed to classify our targets according to the tissue type in which they were over-expressed. The majority of hits were highly expressed in tissues of the immune

system, with an additional cluster that was also over-expressed in cancer cells (Fig. 5.4). In addition, smaller clusters were also highly over-expressed in neuronal and reproductive tissue although the significance of this is presently not clear. A further, analysis of the disease relationship of our hits, by using the Novoseek gene-disease relationship scores, revealed a strong association with either one or more forms of cancer for nearly half of the molecules (21 of 43; Fig. 5.5). Of the remaining, there were sporadic reports in the literature implicating several in different forms of cancer. These results, therefore, further support a functional role for the targets identified by our siRNA screen, in the context of cell cycle regulation.

5.4 Analysis of the phosphoproteome of cycling CH1 cells.

In order to get a "coarse grained" view of the signaling pathways that were constitutively active in cycling cells we resorted to a mass spectrometric determination of the cellular phosphoproteome, in CH1 cells sorted for the G1 and S phase sub-populations. An analysis of the phosphoproteins for their functional properties revealed several of the signaling molecules identified to be known regulators of survival pathways (e.g. AKT2, DAPK1, MAP3K3, CSNK2A1, and CSNK1D; Supplementary Table S4). Here, CSNK2A1 was also described by our siRNA screen as a molecule that influenced RT of the G2 phase (Fig. 5.2A). Similarly, whereas CDK5 was short listed as a G1 and S phase regulator by our siRNA screen, our mass spectrometric analysis revealed that the CDK5 activator protein (CDK5R1; Supplementary Table S4) was phosphorylated in both of these phases. The microtubule-associated Ser/Thr kinase MAST2 was also similarly phosphorylated in both the G1 and the S phase, with its isoform – MAST1 – representing a confirmed hit in our screen. TAOK2, a G1 and S phase regulator identified in our screen (Fig. 5.2A),

Figure 5.4

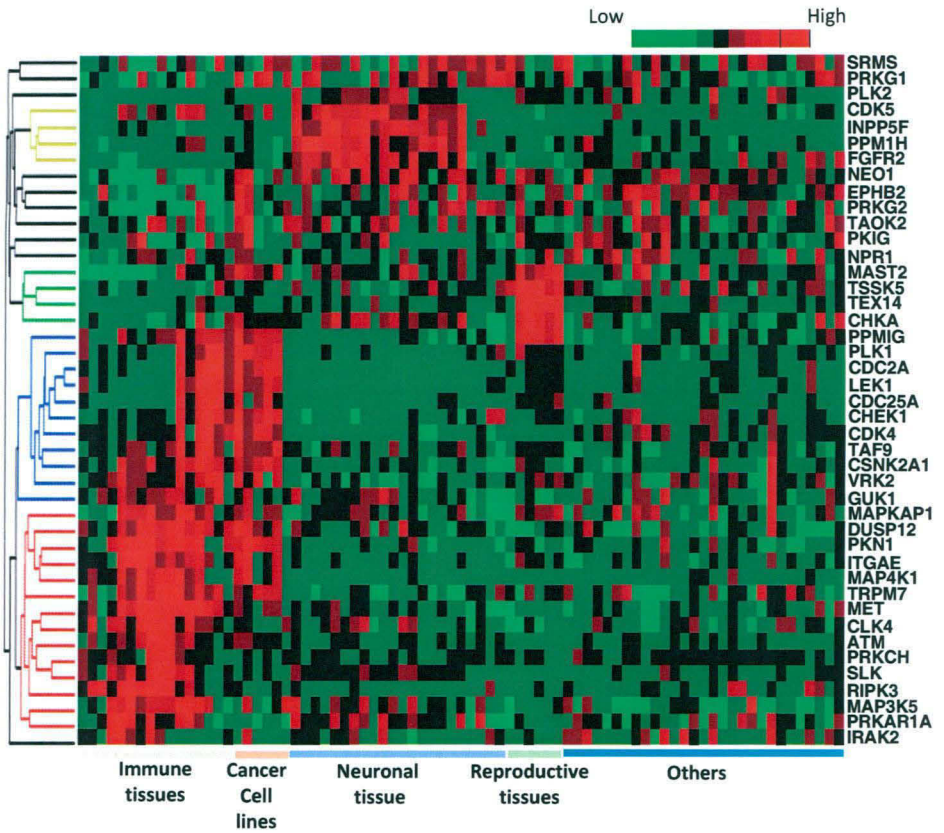


Figure 5.4. Functional relevance of the identified targets.

The heatmap displays the gene expression pattern of the human orthologs of our validated hits, in the indicated range of tissue types. These profiles for the different human tissue types were downloaded from GNF BioGPS and the colours in the dendrograms indicate the normalized levels of expression described in the database. Expression pattern of the target genes, based on their normalized values, across the different tissues are plotted here. The bar above the panel defines the colour code employed to depict the different expression levels.

Figure 5.5

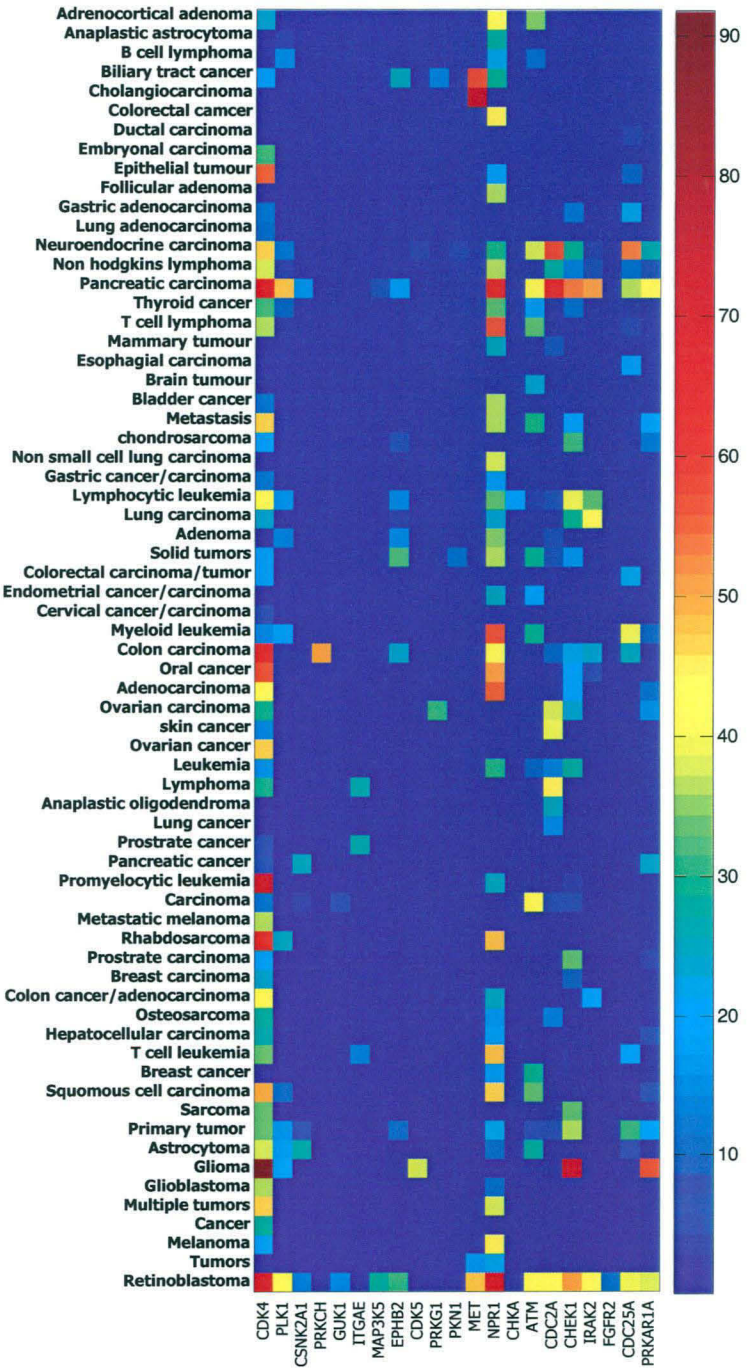


Figure 5.5. Functional relevance of the identified targets.

The figure depicts the gene-disease relationship score (indicated by the colour bar) of our hits with various forms of cancer. These scores were obtained from Novoseek gene-disease database.

was also present – along with TAOK1 its closely related family member – in the phosphoproteome.

Additional phosphoproteins identified included those involved in G protein coupled receptor signaling (GBGT2, GTPB2, GPR139, ADCY6, RAB6A, RGS9), along with regulators of transcription (e.g. RB1CC1, SMAD5, FOXP4 etc), translation (e.g. E2AK3, EIF3C etc), and DNA replication (e.g. ORC3L). Finally, other functional classes of proteins present were ribosomal proteins (RPS10, RPS17, RPLP2 etc), transporters (ABCD4, ABCG5, ABCA1 etc), ion channels (KCNC3, KCNC4 etc), and members of the ubiquitin ligase pathway (e.g. UBE3C). Figure 5.6 shows the ion spectrum of a representative phosphopeptide derived from the voltage-gated potassium channel KCNQ5. Comparable levels of this phosphopeptide were present in both the G1 and S phases.

By employing the multiplexing isobaric tagging approach, we also quantified the relative presence of a given phosphoprotein, between the G1 versus the S phase (Experimental Procedures). A total of 303 parent phosphoproteins were identified to be present in the G1 and the S phases (Supplementary Table S4) of which 61 were differentially phosphorylated between the two phases (Fig. 5.6). While 11 such proteins showed increased phosphorylation in the G1 phase, the remaining were selectively enriched in the S phase (Supplementary Table S4). This phase-specific distinction in composition of the phosphoproteome conforms to the expected shift in topology of the signaling network from the G1 to the S phase. Though these results can be further explored at the proteome level also, the distinct phosphorylation states of proteins observed in different cell cycle phase prompted us to understand the phase

Figure 5.6

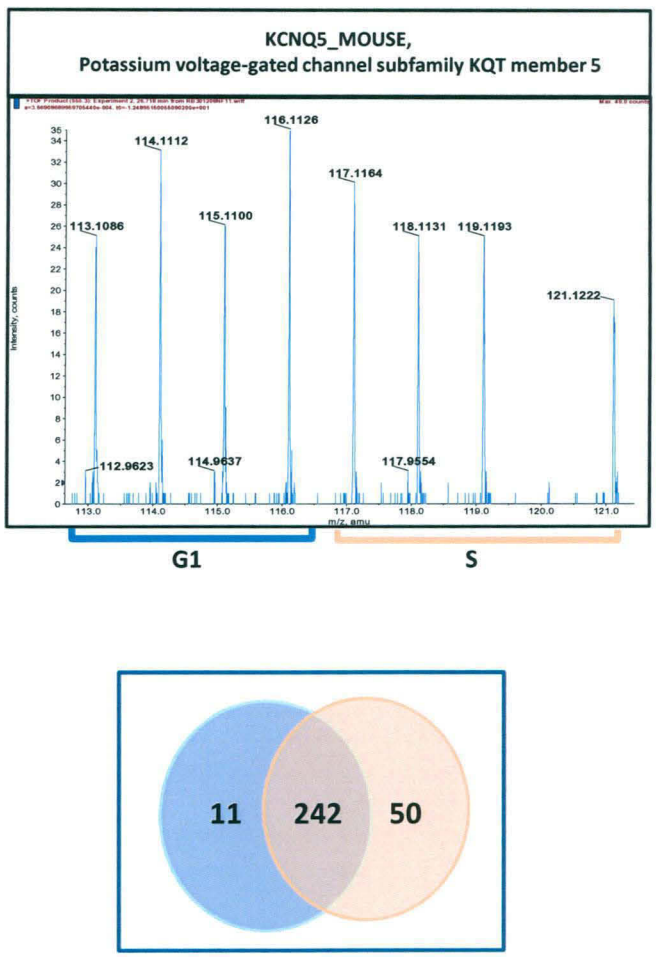


Figure 5.6. Enrichment of the dataset through a description of the cellular phosphoproteome.

The top section of figure 5.6 shows ***the ion-spectra*** *of a representative peptide obtained in our phosphoproteome analysis of the G1 and S population of CH1 cells (see Text). The numbers on top of each peak represent the molecular mass of the iTRAQ label associated with it, in cells in the G1 and the S phase. Here, tryptic peptides derived from four batches of independently sorted cells were separately labeled with each of the iTRAQ labels and then pooled prior to analysis (see Experimental Procedures).*

The Venn diagram in the bottom of this panel represents the distribution of the phosphoproteins obtained between the G1 and S phases of the cell cycle.

specific signaling networks that may regulate the distinct cell cycle phases (Olsen, Vermeulen et al. 2010).

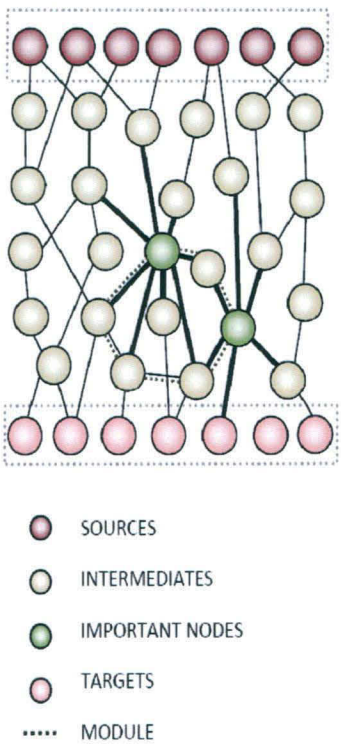
5.5 Defining the source and target relationships for phase-specific regulation of the cell cycle.

When we examined results from two previously reported genome-wide surveys performed in the human cell lines HeLa and U2OS (Mukherji, Bell et al. 2006; Kittler, Pelletier et al. 2007), we found that there was a <5% overlap between them at the level of the signaling molecules (i.e. kinases and phosphatases) that were identified. Relative to this, the overlap between targets identified in this study and the signaling molecules described in either of these two previous reports was between 5-10%. This poor concordance between results from independently conducted screens exemplifies the emerging limitations of RNAi-based screening approaches where differences in cell type and the experimental conditions employed, exert a significant influence on the outcome (Goff 2008; Bushman, Malani et al. 2009). This is in addition to the intrinsic drawbacks of this procedure that lead to a preponderance of false negatives that arise both from the redundant functioning of molecular intermediates, and from insufficient target silencing (Bushman, Malani et al. 2009). It was, therefore, imperative to explore whether a further interrogation of our results would yield more meaningful, and consistent, insights into signal-mediated regulation of the cell cycle.

Having identified the signaling intermediates whose depletion introduced temporal perturbations in the cell cycle, we next aimed to reconstruct the network of pathways through which these effects were enforced. The rationale underlying our approach is illustrated in Figure 5.7A. Here, we defined the targets identified by our screen as 'source nodes' (nodes in dark red) to imply that these signaling proteins provide

Figure 5.7

A



B

List of cell cycle phase specific sources		
G1 Sources	**S Sources**	**G1S Sources**
PKIG	ITGAE	TRPM7
VRK2	EPHB2	NPR1
CHEK1	MAPKAP1	PRKAR1A
FGFR2	NEO1	SRMS
MAST2	PPM1H	TSSK5
RIPK3	PPM1G	MET
ATM	TAF9	CDC2A
IRAK2	GUK1	CDC25A
PLK2	CLK4	TAOK2
LEK1		SLK
PKN1		CDK5
TEX14		MAP3K5
		DUSP12
		MAP4K1

Figure 5.7. A systems approach to the analysis of cell cycle regulation and the identification of phase-specific IMP nodes.

Panel A *illustrates the rationale underlying the examination of source to target sub-networks in order to identify key intermediate effectors of cell cycle perturbation. Here the sources denote the hits identified by the siRNA, while the targets denote the key molecules involved directly in progression of cell cycle and growth.*

Panel B *provides the 'source' characterization of the hits obtained in the siRNA screen, based on the phase-specific effects obtained. The genes described here are the human orthologs of the murine counterparts (see Text).*

sources for inducing perturbations in the cell cycle. Further, since the effects of siRNA-mediated depletion of each of these source proteins was likely to be eventually enforced through modulations in activity of one or more constituents of the core cell cycle regulatory machinery, we treated this latter group of proteins as target nodes (Fig. 5.7A, light red nodes). The objective then was to trace the network of pathways connecting the sources to the targets, and search for any associated topological features (shown as green nodes linked through dotted edges) that may be central to the regulation of cell cycle progression through the G1 and S phases. In addition to providing insights into the underlying regulatory mechanisms, we expected that this strategy would also illuminate any other key regulators that may have been missed by the screen.

The proteins whose depletion resulted in an extension in RT of either the G1, the S, or both of these phases (Fig. 5.3 and Supplementary Table S3), were labelled as source nodes specific for the respective phases (Fig. 5.7B). Here, targets that influenced both the G1 and S phases were denoted as G1S sources (Fig. 5.7B). Following this phase-specific segregation of source nodes, our next goal was to define the putative end-targets for them. That is, constituents of the core cell cycle machinery that directly influence progression of cells through the G1 and S phases (Murray 2004). An extensive literature survey identified seventy-four such proteins, which could then be grouped either as G1 or S phase-specific 'targets' based on their known functional roles. Supplementary Table S5 provides a list of these targets, along with links that describe evidence in support of their categorization. Due to the contiguous nature of cell cycle progression, however, it was not possible to distinguish G1S-specific targets from those involved either in the G1 or the S phases. Therefore the list

of G1 and S phase targets were combined together, and taken as targets for the G1S phase.

5.6 **Extracting the protein-protein interaction networks that control G1 and S phases of the cell cycle.**

We next mapped the network of pathways linking the source and target nodes in each of the relevant phases by employing the approach outlined in Figure 5.8. Combining the BIND, IntAct, HPRD, BioGRID, MINT, and NetworKIN databases generated an undirected 'parent network' that consisted of nearly fifty thousand protein-protein interactions. To filter out all low-confidence interactions, we then selected only those whose existence was supported by at least two independent experimental approaches. Interactions described solely through a combination of *in silico* and high-throughput methods were also excluded. The resulting network - consisting of 5200 nodes and 12,000 interactions (or, edges) – therefore retained only interactions that were based on sound experimental verification.

5.7 Identification of important intermediates using the shortest path analysis.

By taking this high-confidence network, we next traced all possible shortest paths from each of the classified phase-specific sources (G1, S, or G1S, see Fig. 5.7B) to all of the corresponding phase-specific targets described in Supplementary Table S4. Shortest paths were traced from each $G1_{source}$ to all $G1_{target}$,' from each S_{source} to all S_{target} and also from all $G1S_{source}$ to all $G1S_{target}$, (STEP 2, Fig. 5.8). Shortest path lengths from source to target generally varied from 3 to 7, with several such paths often being present between a given source and target combination. To then identify those molecular intermediates (IMP nodes) that were enriched in a cell cycle phase-specific manner, we counted the number of times an intermediate occurred in the

Figure 5.8

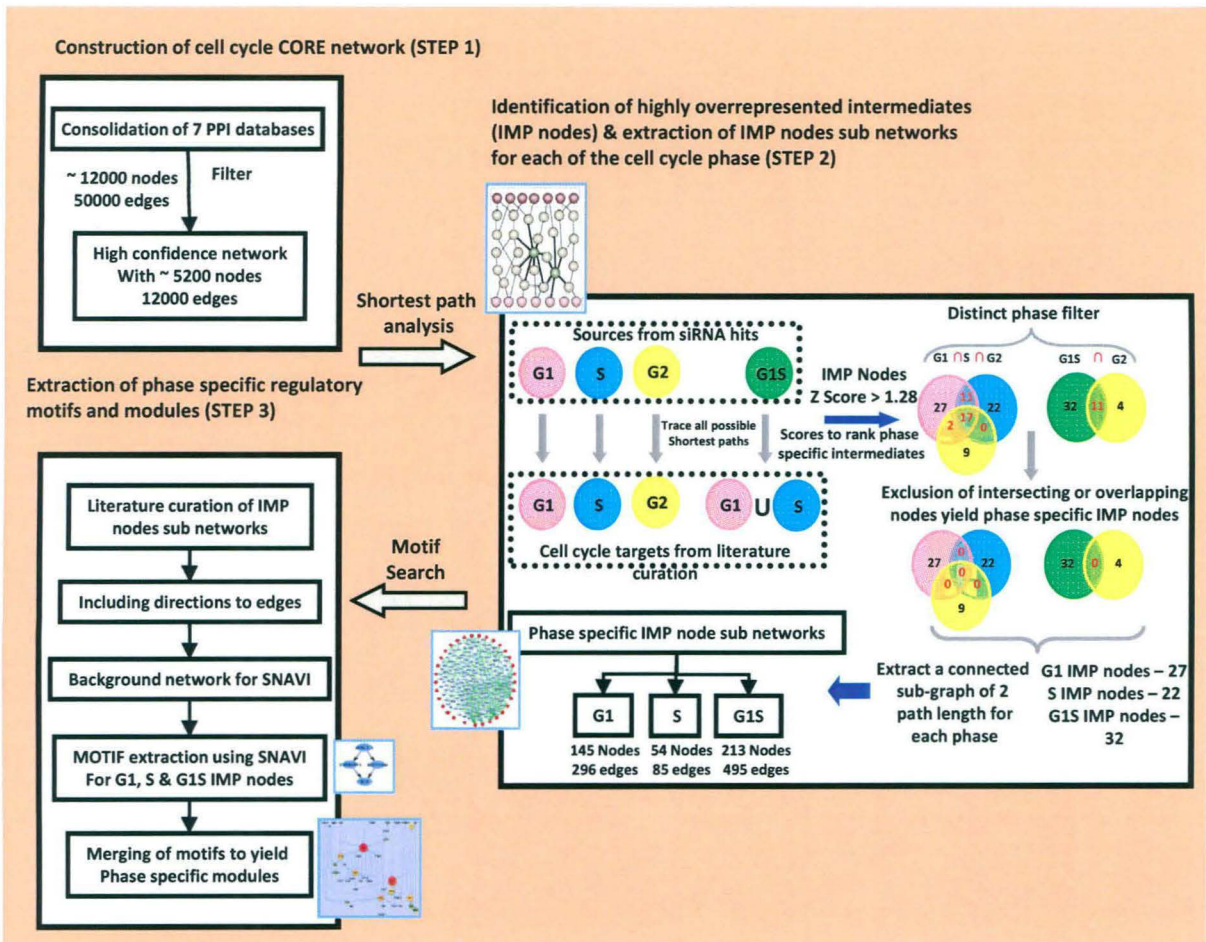


Figure 5.8. A systems approach to the analysis of cell cycle regulation and the identification of phase-specific IMP nodes.

The flow chart provides a step-wise summary of the in silico methodologies employed for eventual identification of the phase-specific regulatory modules. STEP 1 describes the stages involved in delineation of the core network, whereas STEP 2 illustrates the subsequent analysis of this core network to extract the phase-specific IMP node sub-networks. Finally, STEP 3 depicts the stages through which an analysis of the IMP node sub-networks eventually yielded the phase-specific modules. These individual steps are described in detail in Experimental Procedures.

ensemble of shortest paths for each phase (Supplementary Table S6). A ranking of the relative frequency of occurrence of the individual intermediates in each cell cycle phase, but in a manner that facilitated comparison across the three phases, was then achieved by applying Z-scores to normalize for differences in the number of source and target nodes between the respective cell cycle phases (Experimental Procedures, and Supplementary Table S6). Here, to facilitate comparison, a similar $G2_{source}$ to $G2_{target}$ analysis was also performed and the component nodes were similarly ranked. A cut-off Z-score of >1.28 (i.e. $p < 0.1$, Experimental Procedures) was then employed for selecting the most frequently occurring nodes in all possible shortest paths for a each of the cell cycle phases. From each of these individual lists, we next removed all those nodes that were also significantly present (i.e. Z-score > 1.28) in shortest paths describing the other, distinct, phases. That is, any intermediate short-listed in the $G1_{source}$ to $G1_{target}$ network was eliminated as a potential G1-specific IMP node if this node was also over-represented in shortest paths describing either the S, or the G2 phases. S-specific IMP nodes were delineated in a similar manner by comparing against the high-ranking node list from the G1 and G2 phases, whereas the G1S IMP nodes were identified by comparison against nodes from the G2 phase alone (Experimental Procedures). This process identified 27, 22 and 32 IMP nodes from the G1, S and G1S phase-specific shortest paths respectively, and these are listed in Figure 5.9.

Intriguingly, with the exception of CDC25A, none of the source nodes described in Figure 5.7B were present in the list of extracted IMP nodes. This was in spite of the fact that shortest paths had been extracted from an undirected network thereby ensuring the absence of any inherent bias against their selection. A closer examination revealed that several of the nodes from both groups were eliminated as possible IMP

Figure 5.9

List of cell cycle phase specific IMP nodes

G1 Phase		S Phase		G1S Phase	
PTK2B	CSNK2A1	ITGB7	RBL1	ABL1	LYN
TICAM2	WT1	ADORA2B	CCNA2	ANXA1	GADD45G
KIF23	NFKBIA	NTN1	DAB1	GRB2	SKP2
YWHAQ	TICAM1	FOS	ETF1	RIS2	CDKN1B
BCL2	DLGAP4	YWHAG	COIL	SRC	CSNK2A1
IKBKAP	E2F1	PIN1	GRIN2B	RB1	CDC25A
MAPK8	MYC	CSE1L		UBB	ELK1
ESR1	LCK	DOK1		AKT1	JUN
PTEN	TANK	DCC		EP300	TOPBP1
AKT1	XRCC6	RIS2		BRCA1	E2F1
SRC	TRIM27	STAT1		MCM7	MYC
ARHGEF7		E7		ESR1	TUBA4A
ZBTB17		EZR		POLR2A	EPC1
EEF1D		PKM2		STAT1	WT1
GRB2		EED		MAPK3	PTPN2
SYNGAP1		TAF1		CCNA2	CCND1

Figure 5.9. A systems approach to the analysis of cell cycle regulation and the identification of phase-specific IMP nodes.

Figure 5.9 lists the IMP nodes eventually identified through this network analysis for regulation of the G1, S, and G1S phases.

nodes because their frequency of occurrence was below the threshold value of significance. The remaining were then eliminated in the second round of selection since these nodes were also present as frequently occurring intermediates in the shortest paths describing the mutually exclusive cell cycle phases. Thus these nodes were either not enriched enough, or specific enough, to be extracted as cell cycle phase-specific IMP nodes.

5.8 Delineation of IMP node-dependent regulatory modules.

The identification of intermediate nodes that were specifically enriched in pathways mediating a given cell cycle phase also then enabled the search for any topological features unique to the respective phase-specific sub-network. For this we first took the IMP nodes from each cell cycle phase as the seed nodes, and extracted connected sub-networks of two-path length cut-off from the core network (STEP 3 of Fig. 5.8). The resulting G1 IMP node network was composed of 145 nodes and 296 edges, whereas the S IMP node network contained 54 nodes and 85 edges, with 215 nodes and 495 edges describing the G1S IMP node network (Fig. 5.10 A-C). These sub-networks likely encapsulate the specific elements responsible for regulation of the respective phases of the cell cycle.

To therefore delineate the presence of regulatory elements, we adopted the approach outlined in STEP 3 of Figure 5.8. To this end, a survey of the literature first helped to identify the nature of the biochemical reaction represented by each of the edges in the IMP node sub-networks (compiled in Supplementary Table S7). The resulting classification of these edges into functional categories of activation, inhibition (both directed), or neutral (undirected) then facilitated conversion of the individual phase-specific IMP node sub-networks into mixed networks that consisted

Figure 5.10

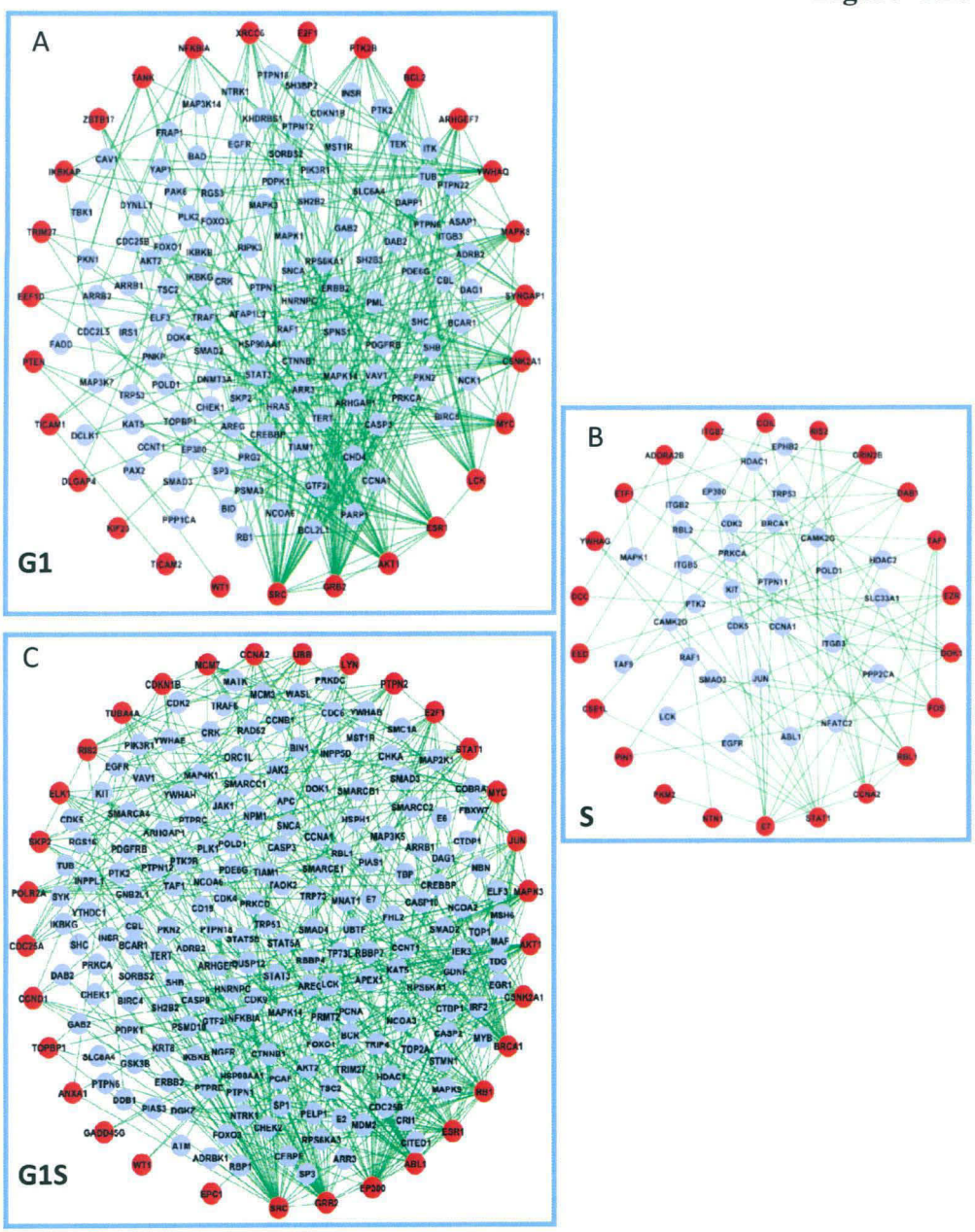


Figure 5.10. Extraction of phase-specific sub-network.

Panels A, B, and C describe the phase specific IMP node sub networks extracted from our cell cycle core network for G1, S and G1S phase regulation. The IMP nodes are defined in red, while the blue nodes denote the intermediate node linking any two or more IMP nodes.

of directed and un-directed edges. Taking each of these, we next probed for any embedded motifs that also included the respective IMP nodes. This was achieved by employing SNAVI - a software tool that implements standard network analysis methods (Ma'ayan, Jenkins et al. 2009). Figure 5.11A lists the various types of motifs that could be identified, their biological significance, and the number of times that such motifs occurred in each of the individual phase-specific IMP node sub-networks. Integrating the phase-specific IMP node networks, and randomizing the edges over a hundred iterations led to a marked reduction in the number of motifs involving directed edges (i.e. excluding the scaffolds; Fig. 5.11B). This confirmed that the phase-specific motifs identified were indeed statistically significant.

The number of scaffolds present in the G1S-specific IMP node network was 2-fold greater than that of G1-, and over 4-fold greater than that in the S phase-specific IMP node sub-network (Fig. 5.11A). Similarly, the numbers of Feed Forward Loops (FFLs), Bifans, and Diamond motifs were also significantly higher than the corresponding numbers present in either the G1 or the S phase sub-networks (Fig. 5.11A). The large number of FFLs identified in the G1S phase sub-network was particularly noteworthy. FFLs regulate both filtering of noise and signal amplification during signal transduction, and previous studies have suggested that they function as important transducers of signal to cell cycle responses (Csikasz-Nagy, Kapuy et al. 2009). Our findings may thus support that FFLs serve as the crucial signal-regulatory elements driving G1 to S phase transition. Further, the redundancy provided by over-representation of bifans and diamond motifs could contribute towards robustness of the system against external perturbations. Finally, our identification that the relative enrichment of signal-dependent regulatory motifs follows the order of G1S > G1 > S is consistent with fact that the G1/S checkpoint represents the step that is most tightly

Figure 5.11

A

Motif	Structure	Biological significance	Occurrences in phase specific network			Example
			G1	S	G1S	
Scaffold		Signal localization and amplification	13	7	23	
Feed forward loop		Noise filter and signal dynamics	11	1	60	
Feed back loop		Signal dampening and dynamics	1	0	7	
Bifan		Signal synchronization	3	0	17	
Diamond		Reflection of redundancy in the system	0	0	19	

B

Motif type	Occurrences of Motifs		Z score
	Real	Random±(S.D)	
Scaffold	31	31±6	NS
Feed forward loop	64	22±7	21
Feed back loop	7	3±2	2
Bifan	18	6±3	6
Diamond	19	6±2	6.5

Figure 5.11. Extraction of phase-specific sub-network and their constituent regulatory motifs.

Figure A provides a compilation of the different functional motifs identified in each of the phases, along with a brief description of their biological implications. In addition, a typical topological representation of each kind of motif is also shown. Here both 3- and 4-node motifs in the FFL categories have been grouped together for the sake of convenience.

Figure B compares the number of the different motifs obtained for the integrated IMP node network (Real) versus that in a hundred randomized networks (Random, values given as mean ± S.D.). In these randomizations both the nodes and the number of links were kept at a constant. The statistical significance of the difference in numbers between the Real and Random cases is indicated by the corresponding Z-scores which were calculated as (<Real> - <Random>)/S.D. Here NS indicates that there was no significant change in the case of scaffolds. This is not surprising given that scaffolds are composed of undirected edges.

regulated by growth factor-mediated signaling (Zetterberg, Larsson et al. 1995; Sears and Nevins 2002).

5.9 Extracting cell cycle regulatory modules and the identification of vulnerable nodes.

The final step towards delineation of the phase-specific regulatory modules involved merging of the motifs identified for each phase in Figure 5.11A to generate the corresponding clusters (see STEP 3 of Fig. 5.8). Thus, all motifs detected in the G1-specific IMP node network were merged together to obtain the G1 phase motif cluster. Similarly, motifs present in the G1S- or the S-specific IMP node networks were also independently merged to generate motif clusters defining the corresponding phases. These phase-specific motif clusters then, likely define regulatory modules that specifically function during the cell cycle phases from which they were derived. A compilation of the nodes present in these modules, along with a classification of the functional groups to which they belong, is provided in Supplementary Table S8.

Having extracted the cell cycle phase-specific regulatory modules, we next sought to identify those constituent nodes that were most critical for their functioning. For this we employed the centrality measures of 'stress' and 'betweenness'. Stress centrality defines the ability of a node to hold together communicating nodes and, therefore, is representative of the degree of its involvement in regulatory processes (Manimaran, Hegde et al. 2009). Betweenness, on the other hand, is a more elaborate centrality index that describes the capacity of a node to serve as a junction in the network and, thereby, regulate the network in a coherent manner (Manimaran, Hegde et al. 2009). Thus stress and betweenness represent complementary indices that together describe the functional importance of a node in a regulatory module. We

calculated the stress and betweenness measures for each node in each phase-specific module, and the resulting plot of these values for the G1 and G1S specific motif clusters is shown in Figure 5.12. Here, as expected, the S-specific module was comprised of only a few nodes (see Supplementary Table S8) with just two giving any measurable value for stress and betweenness. The motif clusters for the G1, G1S, and S phases are shown in Figures 5.13A-C, where the nodes are colour-coded according to their stress and betweenness values. The larger the size of the node the higher is the respective value of stress, whereas the betweenness value is indicated by the shade of the colour for that node.

We next segregated the functionally most significant nodes as those with stress and betweenness values that were both greater than 2.5-fold that of the mean for all nodes in the corresponding module (Experimental Procedures, Fig.5.12). The composition of the resulting shortlist was then compared across the individual phases. Intriguingly, the high stress and high betweenness nodes identified for the G1, G1S, and S windows of the cell cycle did not constitute distinct sets of molecules. Rather, the group of vulnerable nodes identified for a given window overlapped significantly with that for the adjacent window. For example, while PTK2B was exclusively present in the G1 module ESR1, SRC, and GRB2 commonly occurred as high-stress and high-betweenness nodes in both G1 and G1S modules (Fig.5.12). Interestingly, although AKT1 and E2F1 were also present in both modules (Supplementary Table S8), these proteins exhibited properties of high stress and betweenness only in the context of the G1 module (see Fig. 5.12 and 5.13). The nodes exclusively present in the G1S module were ABL1, JUN, BRCA1, EP300, and RB1 (Fig. 5.12). In the S module, only CCNA2 and RBL1 showed functional non-redundancy (i.e. high stress and

Figure 5.12

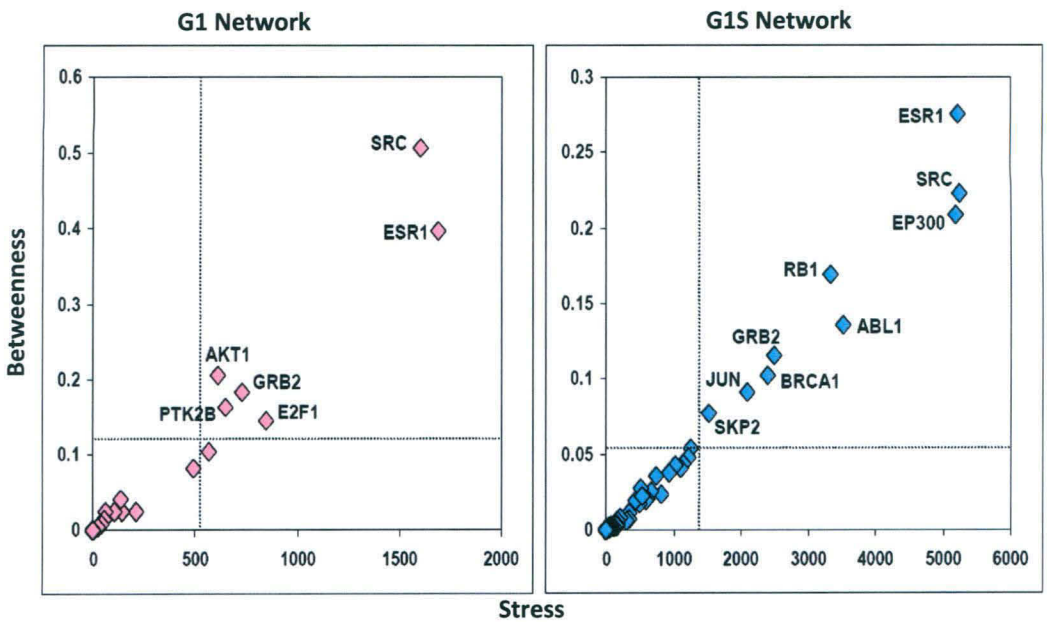


Figure 5.12. Delineating phase-specific regulatory modules and defining the vulnerable nodes.

Figure 5.12 summarizes the results of our analysis of centrality measures of the nodes in each of the phase-specific regulatory modules. This is depicted as a Stress versus Betweenness plot for all the molecules in the G1 and G1S cell cycle modules. Only those nodes that are above the cutoff of 2.5 times that of the mean value (indicated by the dotted line, see text) are identified here.

Figure 5.13

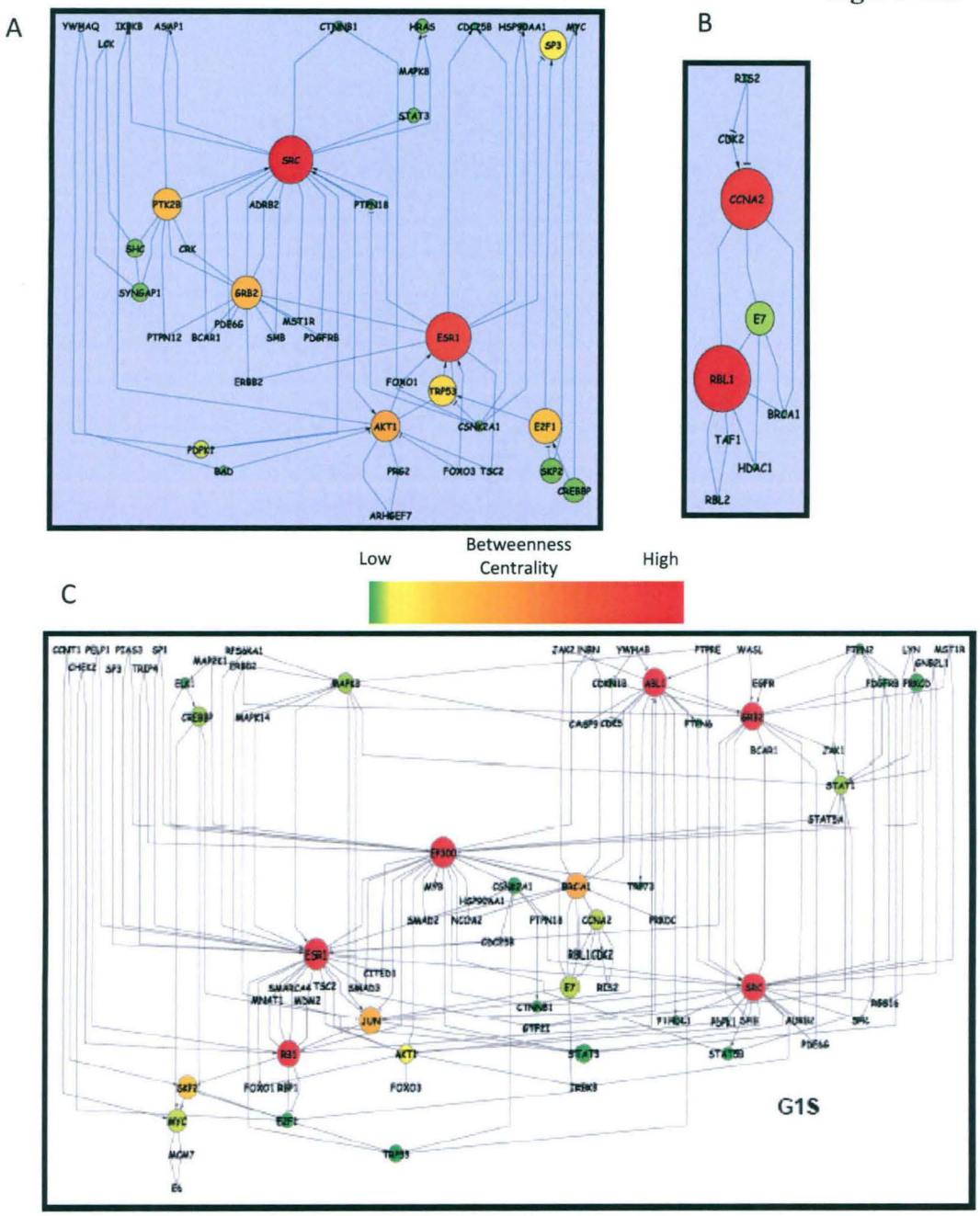


Figure 5.13. Delineating phase-specific regulatory modules and defining the vulnerable nodes.

Panels A,B and C show the IMP node based regulatory modules obtained for each of the individual phases. These represent phase-specific modules extracted by merging the motifs identified for the G1, G1S and S phases respectively. The nodes are colour coded according to their betweenness centrality measure. The size of the nodes represents the stress parameter, where the size of the nodes increases with increasing stress value. The gradation of the betweenness values is defined by the color bar.

betweenness, Fig. 5.13 B). The observed overlap in structure between modules describing the adjacent windows of the cell cycle is consistent with the fact that individual stages of this process do not represent discrete events, but rather reflect windows that capture temporally defined stages of a contiguously evolving biochemical cascade.

Interestingly the majority of the G1-unique nodes represented signaling molecules, whereas nodes specific for the G1S IMP node network were distributed between signaling molecules, transcription regulators, and proteins involved in DNA replication (Supplementary Table S8). The transcription factors identified in the latter were those with an established role in the cell division cycle (e.g. E2F1, TRP53, TRP73, MYC, JUN etc.), in addition to components of the general transcription machinery (GTF21) as well as factors that influenced chromatin remodeling (e.g. EP300, SMARCA4) (Supplementary Table S8). The identification of proteins involved in DNA replication within the G1S network is consistent with the requirement for cells to cross the G1/S checkpoint and these included MCM7, CDC6, TOPB1, and SKP2. The former two molecules are essential for initiation and regulation of eukaryotic genome replication respectively, whereas TOPB1 regulates topoisomerase II activity (Kumagai, Lee et al. 2006). On the other hand, SKP2 - a subunit of the ubiquitin protein complex - regulates stability of the origin recognition complex subunit ORC1L, and the DNA replication factor CDT1. Further, the SKP2 containing ubiquitin protein ligase complex also controls stability of cell cycle regulators such as E2F1, CCNA2, CDKN1B, and the CDK inhibitor p27 (Arias and Walter 2007).

The nodes enriched in S phase network included CDK2 and CCNA2, which form a complex that drives cells through the S phase (Arias and Walter 2007). Other such enriched nodes included here were BRCA1 the protein involved in DNA damage repair, the DNA replication factor RIS2, and RBL1 – a key regulator of entry into cell division (Harper and Elledge 2007). Figure 5.14 summarizes the distinct classes of biochemical activities described by each of the IMP node sub-networks in the form of a heat map. The observed layering of the initial activation of signaling events (G1 and G1S) over the overlapping phase involving the induction of transcriptional processes (G1S), and subsequent recruitment of the DNA replication machinery (G1S and S) accurately recapitulates the broad sequence of events that govern cell cycle progression.

These collective results, therefore, substantiate the information-rich nature of the IMP node modules extracted in Figure 5.13, and also that together these modules indeed capture the dynamic transitions in biochemical activities that enforce the commitment of cells to the division cycle.

5.10 Verifying phase-specificity of vulnerable nodes.

To validate the relevance of the phase-specific modules more rigorously, we wanted to experimentally verify both the selection of high-stress high-betweenness nodes, and the characterization of their phase-specificity. For this, we first compared between the effects obtained upon perturbation either of AKT or ABL. Our IMP node modules had suggested that the functional significance of AKT1 was restricted to the G1-specific module whereas ABL1 was specific to the G1S module. Therefore, we treated CH1 cells with pharmacological agents that either inhibited AKT phosphorylation (LY294002), or, ABL activity (Imatinib mesylate). Here, the concentration of both

Figure 5.14

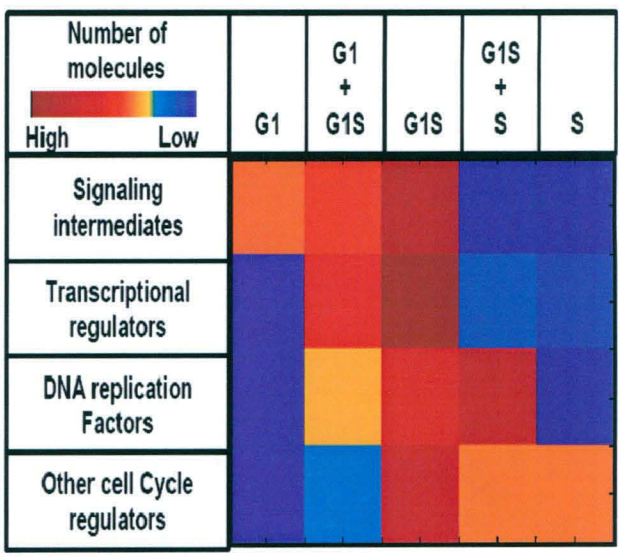


Figure 5.14. Charting phase-dependent modulations in IMP node network architecture.

Figure 5.14 *depicts the stage-dependent modulation in biochemical activities of the core cell cycle regulatory network spanning from early G1, to completion of the S phase. This is presented as a heat map that defines the percent contribution of signaling, transcription regulatory, DNA replicative and cell cycle regulatory molecules, in each of the indicated windows.*

inhibitors was adjusted such that a similar PDT was obtained in both cases, thereby enabling a direct comparison between the RT of the individual phases. Figure 5.15A shows that inhibition of AKT resulted in a specific extension of the G1 phase, with a marginal reduction in RT of the S phase of the cycle. In contrast, the increased PDT obtained upon ABL inhibition involved contributions – in terms of extensions in RT - from both the G1 and S phases (Fig. 5.15A). These results therefore verify the distinctions in phase-specificity that were earlier suggested for these two nodes, on the basis of centrality index estimations.

We also tested the effects of perturbing the remaining high-stress, high-betweenness nodes that were identified (Fig. 5.15B). As earlier discussed, the large majority of high-stress high-betweenness nodes identified were those that were either unique to G1S, or shared between the G1 and G1S modules. Consistent with this characterization, perturbation at any of these nodes resulted in a simultaneous increase in RT of both the G1 and the S phase (Fig. 5.15B). Importantly though, RT of the G2 phase was unaffected. Further, in line with our identification of this protein as a G1-specific regulatory factor, silencing of PTK2B led to a biased extension of the G1 phase (Fig. 5.15B). The only discrepant result obtained was for E2F1 whose silencing perturbed both the G1 and S phases, although this protein was characterized as a node whose vulnerability was restricted to the G1 phase (Fig.5.12). This, however, may not be surprising given that the downstream products of E2F-dependent transcriptional regulation (e.g. cyclins A, B and E) drive both the later stages of G1, as well as the S phase of the cycle.

Finally, with the exception of CDK5, silencing of nodes with low values for stress and betweenness (i.e. between 1-1.5 fold of the mean value) had no significant effect

Figure 5.15

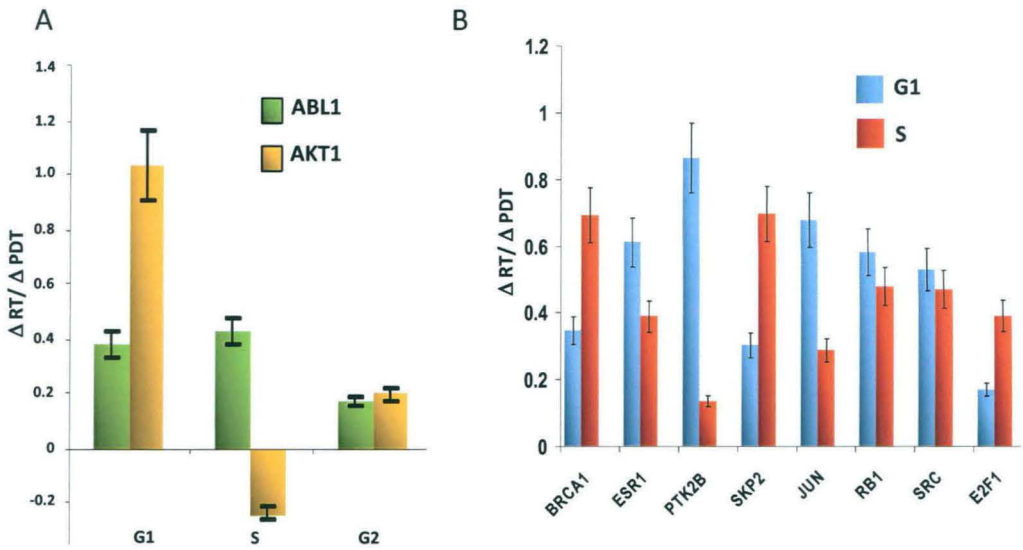


Figure 5.15. Vulnerable nodes provide targets for synergistic disruption of the cell cycle.

Panel A *shows the effects of inhibition of either AKT or ABL on RTs of the individual phases. The increase in RT (DRT) of a given phase, relative to untreated cells, is expressed as a ratio of the corresponding increase- again relative to untreated cells - in PDT (DPDT). Here, LY294002 and Imatinib mesylate were used at concentrations that were half of their respective* IC_{50} *values and the increase in PDT obtained over that of untreated cells was 3h.*

The effects of perturbation of the remaining high-stress high-betweenness nodes identified in Figure 5.9 are similarly shown in ***Panel B****. With the exception of SRC, where PP1a was used for inhibition, perturbation was achieved through siRNA-mediated depletion in all cases. Accordingly, the control cells used for the purposes of comparison were those treated with non-silencing (i.e. GFP-specific) siRNA. The results for GRB2 and EP300-silencing are not included as high levels (>80%) of cell death were obtained in these cases. Values in both panels are the mean (±SD) of three independent experiments. No significant effect on RT of the G2 phase was noted in any of the cases. See also Supplementary Figure S5.*

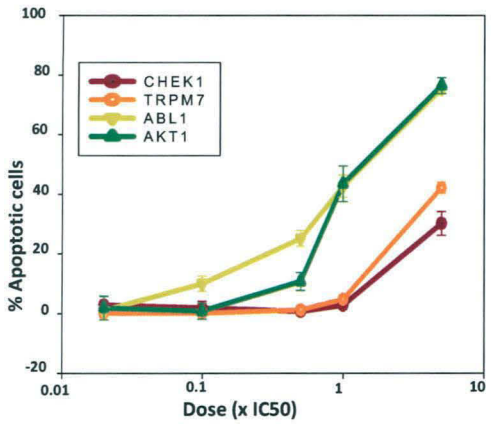
on phase-specific RTs. Although CDK5 is suspected to play a role in cell cycle regulation, its properties and function have been relatively less explored (Dhavan and Tsai 2001). This would then account for the lesser number of associations known for CDK5 and, thereby, the resulting low values for stress and betweenness. Thus, on the one hand, the results in Figure 5.15A and B provide experimental support for the functional description of nodes, on the basis of centrality indices, from the corresponding regulatory modules. On the other, however, they also underscore the possibility that some additional critical nodes may have been missed either due to insufficient information, or noise in the curated PPI databases.

5.11 High stress, high betweenness nodes represent vulnerable constituents of phase-specific regulatory modules.

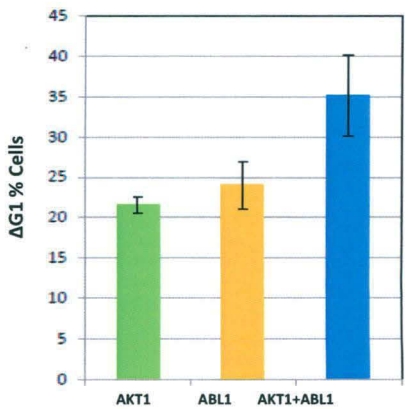
The identification of critical nodes on the basis of centrality measures also implied that these were the functionally least redundant, or most vulnerable, nodes within the respective cell cycle phase-specific modules. Consequently, inactivation of one or more of these nodes should then exert a profound influence on the cell cycle. Consistent with this, treatment of CH1 cells with an inhibitor of either AKT1 activation, or that of ABL1 activity resulted in a dose-dependent increase in the frequency of apoptotic cells (Fig. 5.16 A). Importantly, this apoptotic response was preceded by an accumulation of cells in the G1 phase (Fig. 5.16B), confirming that cell death was indeed the consequence of an arrest in the cycle. Here, for the purposes of comparison, we also examined the effects of inhibition of five representative target proteins that had been identified by our siRNA screen (Fig. 5.1). Of these, two (CHEK1 and FGFR2) were characterized as G1-specific source nodes and the

Figure 5.16

A



B



C

TARGET KINASE	PHASE SPECIFICITY	INHIBITOR USED	IC50
ABL1	G1 and S	Imatinib Mesylate	0.25 μM
AKT1	G1	LY294002	6.6 μM
CHEK1	G1	TCS 2312	60 nM
TRPM7	G1 and S	2-APB	10 μM
FGFR2	G1	PD 173074	3.3 nM
PKA (Prkar1a)	G1 and S	KT 5720	60 nM
MET	G1 and S	PHA-665752	9 nM
SRC	G1 and S	PP1a	200nM

D

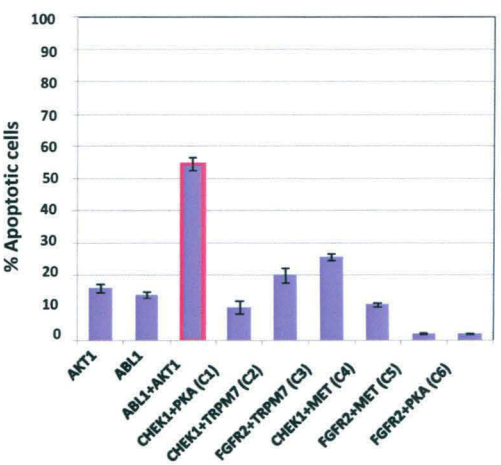


Figure 5.16. Vulnerable nodes provide targets for synergistic disruption of the cell cycle.

Panel A *depicts the dose-response profile for apoptosis in CH1 cells treated with pharmacological inhibitors of the indicated kinases. Doses for each inhibitor used were in multiples of their corresponding IC_{50} values as noted. Values (mean ± S.D. of three experiments) are expressed as the percent of apoptotic cells obtained 72h later, after normalizing for spontaneous apoptosis. (Also see figure S6A)*

The bar graph in ***Panel B*** *shows the corresponding accumulation of CH1 cells in the G1 phase obtained at 18h with the 5 x IC_{50} concentration of inhibitors. Values (mean ± S.D. of three experiments) are expressed as the increase in percent of the G1 population, over that in untreated cells.*

Panel C *lists the target kinases examined, their phase-specificity, the pharmacological inhibitors employed for inhibition of these kinases, and the corresponding IC_{50} values of these inhibitors.*

In Panel D, *CH1 cells were treated with inhibitors against the indicated kinases (or kinase combinations) and the consequent effect on cellular apoptosis was determined at 48h later. The concentration of the relevant inhibitor employed was five-fold greater than its corresponding IC_{50} value in all cases and results are the mean (± S.D.) of three experiments. For the panels A, B and D, cells were treated with a single addition of the inhibitor, or, inhibitor combination.*

remaining three (TRPM7, PRKAR1A and MET) as G1S source nodes (see Fig. 5.7B). Whereas inhibitors of CHEK1 and TRPM7 were significantly less potent than that of AKT1 and ABL1, only marginal effects were obtained upon either PRKAR1A, FGFR2, or MET inhibition (Fig. 5.16A). The inhibitors used in these experiments, along with the IC_{50} values for their respective targets, are listed in Figure 5.16 C.

Our interpretation that the vulnerable character of AKT1 and ABL1 emerged in temporally distinct windows of the cell cycle regulatory module also implied that simultaneous inhibition of both AKT1 and ABL1 should produce a cooperative effect on cell cycle arrest, and the consequent apoptosis. This expectation was indeed borne out in a subsequent experiment (Fig. 5. 16D), and the Loewe Combination Index (CI) obtained was 0.56 ± 0.08. Here, a CI of <1 is indicative of a synergistic response (Chou and Talalay 1981). In contrast, the extent of apoptosis induced by combinations of inhibitors of either CHEK1 or FGFR2 (G1-specific source nodes) with that of TRPM7, PRKAR1A, or MET (G1S-specific source nodes) remained poor (Fig. 5.16D). Thus, the cumulative results in Figure 5.16 confirm that both AKT and ABL exhibit a degree of vulnerability that is significantly greater than at least some of the targets identified in our siRNA. Further, the cooperative effects of combined inhibition observed also supports our inferred distinction in phase-specific involvement (G1 versus G1S) of at least two of the nodes (AKT1 and ABL1) identified on the basis of centrality indices.

5.12 A comparative analysis of independently conducted screens yields a common IMP node network.

Since cell cycle regulation is common to all dividing cells, it would be reasonable to expect that these core modules – and the IMP node network from which they were

extracted – would represent a common feature of dividing cells, independent of the tissue type from which they were derived. To verify our hypothesis, therefore, we also analyzed results of the previously described RNAi-based screens for cell cycle regulators that were performed in HeLa and U2OS cells (Mukherji, Bell et al. 2006; Kittler, Pelletier et al. 2007). As noted earlier, the overlap in signaling molecules identified between these two screens was <5%.

From each of these reports we separately shortlisted the signaling molecules described as 'hits' for either the G1 or S phase (the G1S phase was not characterized in these studies), and categorized them as G1+S source nodes. We next repeated the analysis described in Figure 5.7A wherein shortest paths were traced from each of these source nodes, to each of the G1+S target nodes described earlier. By then subtracting against the corresponding G2-specific source to target network, we extracted the IMP nodes as described in Figure 5.8. A similar exercise was also performed with the results of our present screen, which then allowed us to compare the resulting IMP node list with that obtained from the HeLa and U2OS cell screens. Remarkably, in spite of the nominal degree of concordance at the level of the RNAi-defined hits, there was a >70% overlap in the IMP nodes derived from the three screens (Fig. 5.17A). Further, when IMP node networks were independently generated from each of these three lists, the resulting overlap between them was >92% (Fig. 5.17B). These findings thus establish the significance of our methodology for further interrogating the siRNA screen results. In addition, they also support the likelihood that the IMP nodules identified from our data constitute conserved regulatory elements of the mitogen-dependent signaling network.

Figure 5.17

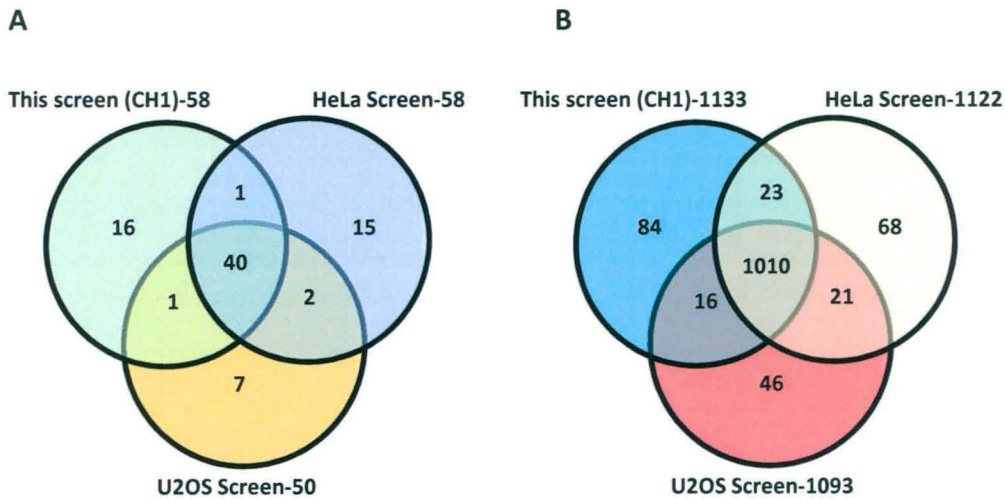


Figure 5.17.

Panel A *shows the venn diagram comparing the overlap of IMP nodes identified from the three different screens.*

Panel B *shows the venn diagram comparing the overlap of nodes from the IMP node networks of the three different screens.*

5.13 IMP node modules define the signaling axes that govern G1 and G1S windows of the cell cycle.

The extensive experimental validation of the functional significance of the IMP node modules supported that a further analysis may shed light onto mechanisms by which mitogen-dependent signaling cascades drive the cell cycle. In the G1 IMP node module, AKT1 and PTK2B were the two kinases that were uniquely present as high-stress high-betweenness nodes, whereas SRC and the adaptor molecule GRB2 represented the least redundant constituents of both G1 and G1S modules (Fig. 5.12). That is, these four molecules likely represented key constituents of the signaling cascade that governs the early G1 phase of the cycle, with AKT1 and PTK2B playing a more specific role in this process. An examination of the known functions of these nodes in the literature supported a central role for AKT1 in mediating the downstream processes required for cells to enter the cycle. Of the remaining nodes, SRC likely serves as a common sensor that bridges AKT1 activation with both RTKs and mitogenic GPCRs. Whereas the SRC/GRB2 complex represents a constituent of the signalosome complex recruited by RTKs (Luttrell and Luttrell 2004), PTK2B mediates SRC activation by GPCRs (Litvak, Tian et al. 2000). This may then explain earlier observations that distinct mitogenic signals activate a common signaling cascade in the early G1 response (Jones and Kazlauskas 2001).

In comparison with the G1 IMP node network, the only unique kinase present as a high-stress high-betweenness node in the G1S IMP node network was ABL1 (Fig. 5.12). While this implication of a specific and key regulatory role for this protein in the G1S phase was intriguing it was, however, experimentally supported by the results in Figure 5.15A. In addition studies from another group have also demonstrated that

depletion in levels of ABL1 caused a marked delay in entry of cells into the S phase (Plattner, Kadlec et al. 1999; Furstoss, Dorey et al. 2002). ABL1 is normally distributed between the cytoplasmic and nuclear compartments of the cell, and its regulatory effects are enforced through activation of its kinase activity. In similarity with AKT1, activation of the cytoplasmic pool of ABL1 is also primarily dependent upon the upstream signaling intermediate SRC, which occurs both through direct and indirect mechanisms (Plattner, Kadlec et al. 1999; Sirvent, Benistant et al. 2008). The latter involves mediation by PLCγ, and by phospho-tyrosine containing SH2 domains generated by active SRC (Sirvent, Benistant et al. 2008). These collective findings, therefore, yield interesting insights into the nature of receptor-proximal signals that govern commitment of cells to the division cycle. Thus, the early G1 is likely regulated through the complement of pathways initiated by the SRC-AKT1 signaling axis. Continuance through late G1, and then entry into the S phase, however appears to be contingent upon a switch in this axis where ABL1 now replaces AKT1 as the non-redundant target of SRC.

5.14 IMP node modules capture the network of core pathways that mediate commitment of cells to the division cycle.

To further probe for information on cell cycle regulatory mechanisms, we merged the G1 and G1S IMP node modules and extracted the motifs incorporating the combined list of high-stress high-betweenness nodes. By then enriching the resulting links with information from the literature, we translated these into a pathway map that extended from the relevant surface receptors, to the effector responses that were influenced. This is shown in Figure 5.18.

Figure 5.18

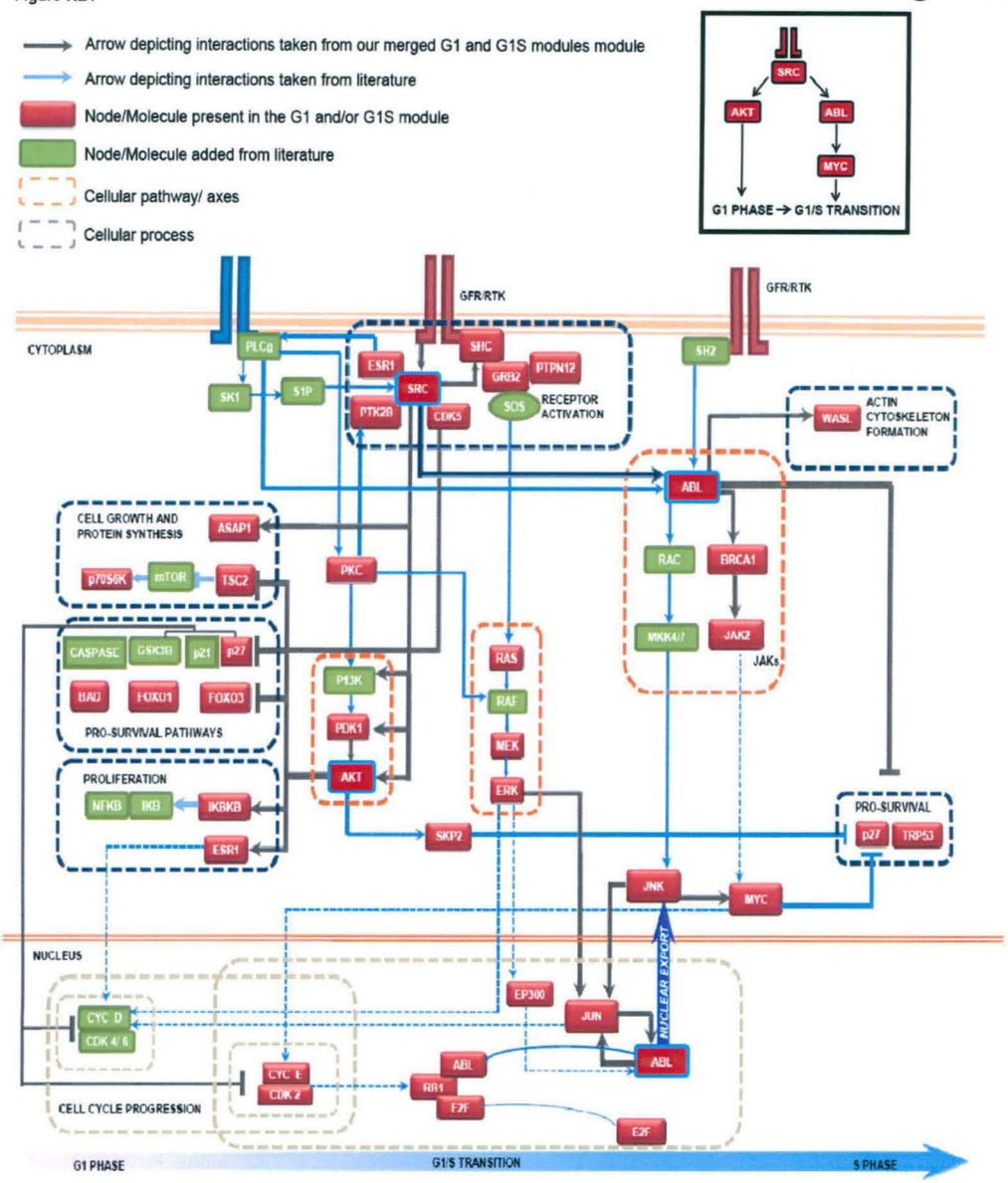


Figure 5.18. Network of core pathways that mediate commitment of cells to the division cycle.

In the pathway model the nodes in red are proteins which are present in the G1 and/or G1S modules. Nodes present in green are derived from literature (they include intermediates or effectors of cellular processes captured by our merged modules). The grey arrows depict interactions taken from the merged G1 and G1S modules. The blue arrows are links added from literature. The dotted orange boxes highlight cellular signaling pathways and the dotted grey boxes group molecules involved in regulating similar cellular processes. The dashed blue arrows indicate a known role of the source node in regulating the target node, derived from literature.AKT and PTK2B are G1 specific high-stress and betweeness nodes. SRC and GRB2 show high stress and betweeness in both modules. ABL and JUN are present as high stress- betweeness nodes only in the G1S modules. A detailed description of the pathways captured is given in the text.

It is evident that this map indeed represents an accurate synopsis of the multiplicity of biochemical events that cumulatively drive commitment of cells to the division cycle. More importantly it also describes the individual modes of functioning of the G1- and G1S-specific signaling cascades, and also the links between them that coordinate their respective activities. Thus, while this analysis confirms that AKT1-dependent pathways indeed initiate cell cycle progression, it also reveals that this activity is supported by a concomitant inhibition of pro-apoptotic pathways, in addition to being coordinated with mechanisms regulating cell growth (Fig. 5.18). Similarly, this figure also rationalizes the key role ascribed to ABL1 during regulation of late G1 and the subsequent G1/S transition. As shown, pathways emanating from the cytoplasmic pool of ABL1 influence both cell cycle progression and cell growth, while also suppressing pro-apoptotic mechanisms (Fig. 5.18). A prominent downstream consequence of cytoplasmic ABL1 activation is the enhancement in cellular c-MYC concentrations (Boureux, Furstoss et al. 2005). MYC is known to function as a central regulator that links external signals to the cell cycle machinery (Obaya, Mateyak et al. 1999). Importantly, the cell cycle regulatory effects of MYC predominantly impinge on the later stages of G1, and facilitate entry of cells into the S phase (Heikkila, Schwab et al. 1987; Santoni-Rugiu, Falck et al. 2000).

A notable feature of our model is that it also highlights a role for the nuclear pool of ABL1 in regulating MYC levels. A significant proportion of nuclear ABL1 is normally maintained in an inactive form through its existence in a complex with RB (Welch and Wang 1995). As cells approach the G1/S boundary however, phosphorylation of RB causes the release of ABL1. At least some fraction of this released ABL1 interacts with JUN, as a result of which ABL is activated. This process then leads to the potent activation of JUN through a feed-forward regulatory

mechanism wherein ABL1 directly phosphorylates JUN, and also its upstream kinase JNK (Barila, Mangano et al. 2000). It is logical then to infer that signaling pathways initiated from both the cytoplasmic and nuclear pools of ABL1 would converge to activate JUN and, thereby, *MYC* upregulation. This cooperative action between the two ABL1 subsets may well explain the long-standing question of how cellular MYC levels are temporally modulated during cell cycle progression. While expression of this protein is induced in the early G1 phase these levels are, however, further enhanced as cells prepare for S phase entry (Jones and Kazlauskas 2001). This increase is likely essential for recruiting – through transcription regulatory mechanisms - the wider spectrum of biochemical activities that are required for driving G1 to S transition. Interestingly, consistent with our proposal, the timing of this second phase of MYC enhancement coincides with that of dissociation of the RB-ABL complex (Welch and Wang 1993). These cumulative results therefore provide a mechanism-based rationalization for our assignment of key, but distinct, regulatory roles to AKT1 and ABL1 in the early and late G1 stages respectively. Of particular note here is that the canonical MAP kinase pathway also plays a supporting role for the effector functions of both AKT1 and ABL1. Significantly, the SRC-mediated activation of this pathway – in a manner that is not strictly dependent upon either AKT1 or ABL1 – allows for its recruitment in both the early and late G1 (Fig. 5.18). Consequently, the MAP kinase pathway likely serves as a common link that facilitates transition between these two phases.

Our model also suggests a plausible mechanism that enables ABL1 to replace AKT1 as the dominant signaling node in the G1S phase. As illustrated, this process hinges upon the release of nuclear ABL1 from its complex with RB. AKTI-dependent initiation of cells into the cycle culminates with the phosphorylation of RB, which is

normally evident by 8h after stimulation of cells with a mitogenic signal (Ezhevsky, Ho et al. 2001). This leads to the release of bound ABL1. As we proposed earlier, activation of this subset synergizes with cytoplasmic ABL signaling to ensure enhancement in MYC levels. In addition, some fraction of this nuclear ABL1 is also likely to partition into the cytoplasmic compartment of the cell. This could be achieved either through its direct export in the non-phosphorylated form, or, following acetylation by the histone acetyltransferase EP300 (di Bari, Ciuffini et al. 2006). The resulting increase in the pool of cytoplasmic ABL1 would then further intensify signals generated from this subset of the kinase.

Our proposal of a marked amplification in ABL1-dependent signaling as cells approach the G1/S boundary is well supported by recent findings demonstrating that cellular levels of phosphorylated ABL1 peak at the G1S phase (Olsen, Vermeulen et al. 2010). However, we further resolve here that this likely occurs through both enrichment of the cytoplasmic pool, as well as through activation of the fraction that remains in the nucleus. Importantly, this amplification would be further reinforced by the fact that signals from both subsets synergize at the level of activating downstream effector pathways such as MYC regulation, and the synthesis of cyclins. These cumulative effects may then explain how ABL1 acquires the characteristics of a vulnerable node in the later stages of G1. Significantly, this switch in the regulatory signaling node from AKT1 to ABL1 occurs in a seamless manner because, in addition to recruiting the additional pathways required for S phase entry, ABL1 also supplants the functions of AKT1 in terms of suppressing apoptotic pathways and regulating cell growth. This then would then also explain the increased functional redundancy of AKT1 at this stage.

5.15 The G1 and G1S IMP node modules represent conserved elements of mitogen-activated signaling cascades.

In addition to the strong experimental validation of their functional significance, our results so far had also provided a mechanistic understanding of how the G1 and G1S IMP node modules regulate their respective phases of the cell cycle in a coordinated manner. Significantly, our comparative analysis yielding a similar IMP node network from the results of siRNA screen results performed in three different cell lines also suggested the likelihood that the IMP node modules thus derived constitute invariant features that are common to all dividing cells. To investigate this we took a panel of fourteen human cell lines derived from a diverse range of tissue types (see Supplementary Figure S6B). These cells were treated with a single dose of either the combined inhibitors of AKT and ABL or, for purposes of comparison, combinations of inhibitors against the siRNA-identified targets (C1, C2, and C3 in Fig. 5.16D). The frequency of apoptotic cells obtained in each case was then determined 48h later and the results are presented in Figure 5.19A. The efficacy of the inhibitor combinations C1, C2, and C3 was highly restricted with the combination C2 inducing significant (i.e. >30%) apoptosis only in four of the cell lines tested. Further, with the exception of Jurkat cells, effects on the remaining cell lines were predominantly due to CHEK1 inhibition with little or no contribution from PKA inhibition (Supplementary Fig. S7B). This was also true of the effects of combination C1 on THP1 and U937 cells, whereas the combination C3 was virtually ineffective (Fig. 5.19 A).

In contrast to these limited effects, simultaneous inhibition of AKT and ABL induced significant levels of apoptosis in all the cell lines tested, which included HeLa and U2OS cells (Fig. 5.19A). This effect was further enhanced by the 72h time point

Figure 5.19

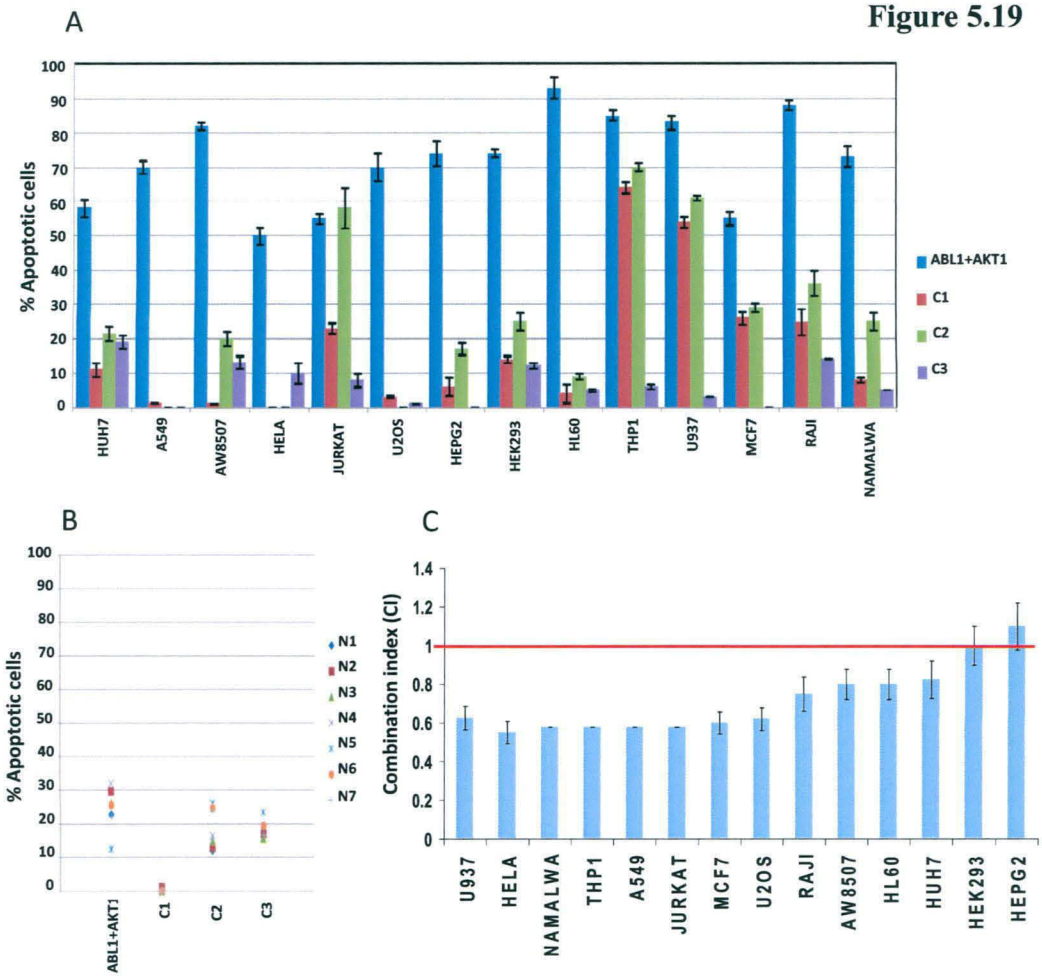


Figure 5.19. Vulnerable nodes constitute regulatory elements that are conserved across a broad range of cell types.

The effects of combined inhibition of the indicated kinase combinations (at concentrations of 5 x IC_{50} values) on fourteen different cancer cell lines, following a protocol similar to that described for Figure 5.13D, are shown in ***Panel A****. (Also refer to Supplementary figure S7).*

Panel B *shows the results of a similar experiment with indicated inhibitor combinations performed on normal B-lymphocytes purified from the peripheral blood of six different individuals.*

Panel C *gives the CI values for the combined inhibition of AKT and ABL obtained for the individual cell lines. CI values were calculated as previously described (Chou and Talalay 1981). For four cell lines CI values could not be estimated since either one (for Namalwa and THP1) or both (for Jurkat and A549) inhibitors showed no effect when added individually (indicated as broken bars). The red line defines the cut-off (CI=1) for interpreting responses as being synergistic and values are the mean (± S.D.) of three determinations.*

where the number of apoptotic cells was >80% in all cases. These collective results, therefore, further support the functional distinction between at least the hits from the siRNA screen tested here, and the high vulnerability nodes identified from a graph theoretical derivation of the phase-specific regulatory modules. That is, in comparison with the former group the latter represents nodes that are involved in a less redundant manner, during cell cycle regulation of a broader range of cancer cell types. Notably, none of the inhibitor combinations examined had any significant affect on viability of normal B-lymphocytes purified from the peripheral blood of six different individuals (Fig. 5.19B). This confirms that the functional relevance of AKT1 and ABL1 is only expressed in actively cycling cells.

To assess whether responses to the combined inhibition of ABL and AKT in Figure 5.19A were synergistic we determined the CI values from dose-response curves either to LY294002, Imatinib mesylate, or the combination of both in each instance. The results are presented in Figure 5.19C where a synergistic effect is clearly evident in twelve of the fourteen cell lines. Interestingly, this also included examples where either one (Imatinib mesylate in Namalwa and THP1 cells), or both inhibitors (Jurkat and A549 cells) had no significant effect on cell survival when administered individually (Fig. 5.20); further emphasizing the extent of cooperativity achieved by the inhibitor combination in these cells. For the remaining two cell lines however (HEK293 and HEPG2), dual inhibition of both ABL and AKT yielded an outcome that largely represented an additive consequence of the individual effects. We reasoned that this absence of synergy for the inhibitor combination in this latter group of cell lines could derive from subtle, cell type-inherent, differences in the architecture of the IMP node modules. This in turn may then influence the degree of vulnerability exhibited by AKT1 and ABL1. Thus, for example, our pathway model in Figure 5.18

Figure 5.20

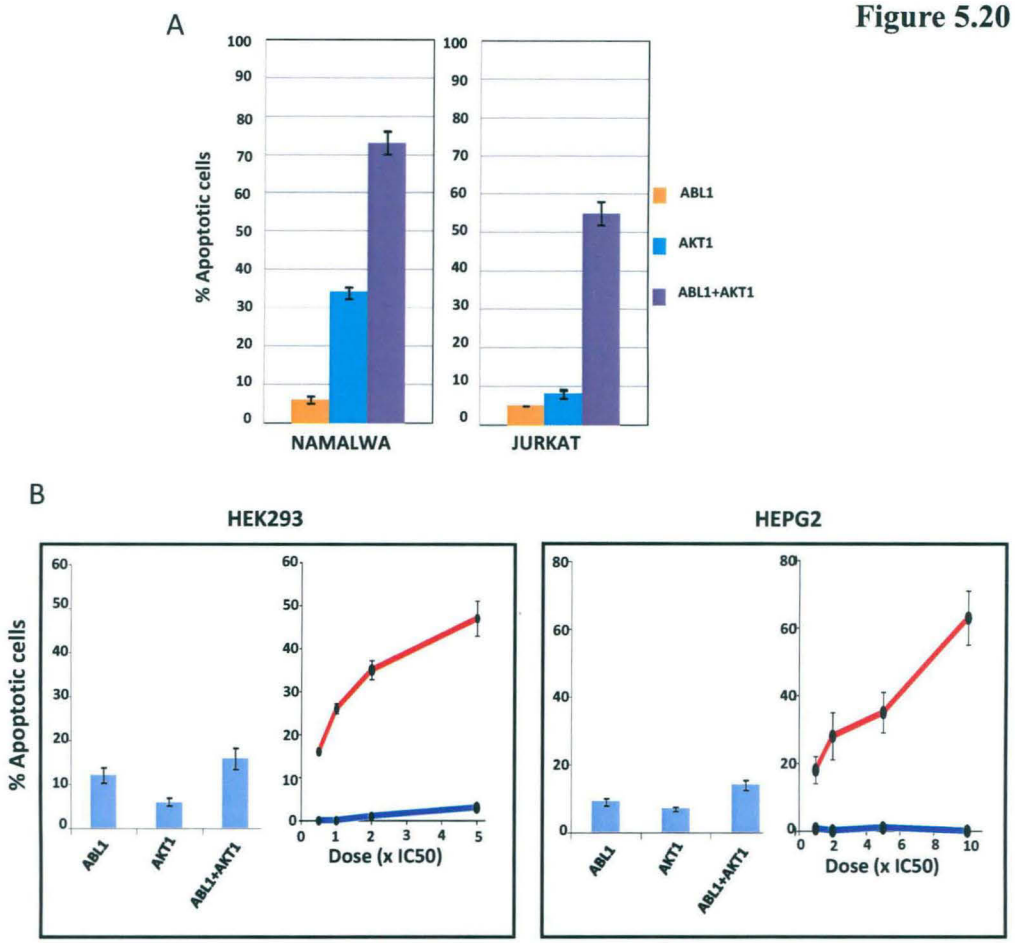


Figure 5.20. Vulnerable nodes constitute regulatory elements that are conserved across a broad range of cell types.

Panel A *shows representative results for the synergy between AKT and ABL inhibition where one (Namalwa) or both (Jurkat) inhibitors were ineffective when added individually.*

Panel B *shows the effects of the inclusion of PP1a, in addition to LY294002 and Imatinib mesylate, on the apoptotic responses of HEK293 and HEPG2 cells. Results are shown in two sub-panels for each cell line where the sub-panel on the left indicates the effects either of LY294002, Imatinib mesylate alone (each at 4 x* IC_{50}*), or a combination of both. The sub-panel on the right gives the results of addition of increasing concentrations of PP1a either alone (blue line), or in the presence of the combined inhibitors for AKT and ABL (each at 4 x* IC_{50}*, red line). Values again are the mean (± S.D.) of three determinations.*

would suggest that the degree of involvement of the MAP kinase pathway could potentially determine the extent to which both AKT1 and ABL1 were required, in order to maintain cells in active cycle. Consequently, we probed this latter group further by also including SRC inhibitor PP1a, along with those for ABL and AKT. As earlier discussed, SRC principally functions by mediating either AKT1 activation in the G1 phase, or ABL1 activation in the G1S phase.

Figure 5.20 B shows that although SRC inhibition alone had no significant effect on either of the cell lines tested, a strong synergistic effect was nonetheless obtained when it was combined with the inhibitors of both AKT and ABL. Thus, by combining SRC inhibition along with that for AKT1 and ABL1, it seem possible to also guard against any cell type-specific variations in topology of the G1 and G1S IMP node modules. Important to note here is that in all cases involving inhibition of vulnerable nodes - or their combinations - in Figure 5.19A, increased apoptosis was always preceded by an accumulation of cells in the G1 phase, supporting that cell death was indeed mediated through an arrest in the cell cycle (Supplementary Fig. S7C). These cumulative results, therefore, confirm that the G1 and G1S IMP node modules described in Figure 5.13 indeed represent signal processing elements that commonly mediate cell cycle progression in at least a wide range of cell types. Further, the vulnerable nodes that they incorporate provide sensitive targets for achieving an efficient disruption of the cell cycle regulatory network.

Chapter 6
Discussion

Chapter 6: Discussion

The trade-off between sensitivity and robustness constrains the architecture of the signaling network into a bow tie or hourglass structure, which represents the convergence of diverse and redundant input processes onto a conserved core module or set of proteins (Csete and Doyle 2004; Supper, Spangenberg et al. 2009). Such core elements function as the key regulators of plasticity in the cellular response, by calibrating a range of output processes (Kitano and Oda 2006; del Sol, Balling et al. 2010). It was therefore likely that the mitogen-activated signaling network also incorporates such conserved regulatory elements that process input signals, and translates them into the output response of cell cycle progression. We approached the question of how the signaling network regulates cell cycle by first attempting a definition of those intermediates that were capable of inducing perturbations in the output response. The latter was achieved through a siRNA screen that targeted components of the signaling machinery, followed by a scoring for effects that were specific to the individual phases of the cell cycle.

Although RNAi-based screens have been employed to identify regulators of cell proliferation and/or survival, limitations inherent to this approach have confounded interpretation of the results. Key caveats here are the preponderance of false negatives because of which many important regulators are missed, and the fact that the majority of hits obtained are those that function in a cell type-restricted manner (Bushman, Malani et al. 2009). Consequently, convergence between results from different laboratories is often too low to permit any meaningful interpretation. This aspect was also underscored by the insignificant overlap obtained between the three datasets that included our present screen results, and that of the two earlier performed screens.

Therefore, we were prompted to explore graph theoretical approaches for unravelling the mechanistic implications of the hits identified. We were encouraged in this endeavor by earlier studies where analysis of network interactions in receptor-specific pathways indeed yielded insights into signaling mechanisms, and also emphasized the potential of this approach for identifying new drug targets (Bushman, Malani et al. 2009; Astsaturov, Ratushny et al. 2010).

The molecules identified in our siRNA screen were categorized as sources for phase-specific perturbation, and subjected to iterative rounds of a network-based analysis. This exercise eventually yielded the core signaling modules that regulated the G1, G1S, and S phases of the cycle, and also delineated the most vulnerable nodes present in them. Central to this analysis was the identification of IMP nodes, which were selected on the basis of their unique over-representation in each of the phase-specific regulatory sub-networks. Their high frequency of occurrence implied correlation with a strong functional role that was specific to the corresponding cell cycle phase. That is, phase-specific IMP nodes were likely those that combined the traits of minimal functional redundancy and increased specificity of action, in the context of regulation of the respective window of the cell cycle. By extension then, the modules defined by the phase-specific group of IMP nodes also likely constitute the key regulatory elements that process mitogenic (and oncogenic?) signals for the respective cell cycle phases. Here, a re-examination of data from the two previously described screens validated both the relevance and significance of our approach. The virtually similar IMP node network that was extracted in all cases also lent preliminary support to our proposition of the existence of a conserved set of core elements that regulate mitogenic signals.

Intriguingly, although the vulnerable nodes identified from the IMP node modules included prominent signaling molecules such as AKT1, ABL1, PTK2B, and SRC, none of these were, however, detected either in our CH1 cells screen, or in the previously described screens in HeLa and U2OS cells (Mukherji, Bell et al. 2006; Kittler, Pelletier et al. 2007). That this discrepancy was due to limitations in the RNAi screen procedure, and not an artifact of our analyses, could be confirmed in experiments where inhibition of AKT1 and ABL1 induced an apoptotic response in all of the three cell lines. Importantly, these and other such similar experiments provided convincing verification of the functional properties ascribed to the vulnerable nodes and, therefore, also served to validate the parent IMP node modules from which they were derived.

Examination of the incorporated links revealed novel details on the mode of functioning of these modules. For instance, the G1 and G1S modules both participate in the critical function of coordinating cell cycle progression with cell growth, in a setting where pro-apoptotic pathways are inhibited. The shared activities between the individual modules maintains contiguity during progression of cells from early G1 up to the S phase, whereas integration between the three modules promotes temporal evolution of the spectrum of biochemical processes that drive commitment of cells to the cycle. Particularly significant here was our discovery that the G1 and G1S phases were controlled by distinct signaling axes that were composed of SRC-AKT and SRC-ABL respectively. Further, we also subsequently elucidated that it was the switch between these two axes that guided transition of cells from early to late G1/G1S. This inference, initially derived from the graph theoretical analysis, could be experimentally verified in CH1 cells by specifically targeting either AKT or ABL with pharmacological inhibitors. In more recent experiments we have found similar effects

of AKT- and ABL-inhibition in all of the cell lines examined in Figure 5.19 (Jailkhani, N., et. al., unpublished results), supporting that this regulatory mechanism is a common property of dividing cells.

In addition to defining the receptor-proximal signals, however, our cumulative results further established that the IMP node modules in fact also capture the key regulatory elements that translate the effector consequences of mitogenic/oncogenic signaling to the respective stages of the cell cycle. In this connection, our experiments demonstrating that these cumulative functional properties of the G1 and G1S modules were conserved across a diverse panel cell lines is particularly significant. That is, these phase-specific modules presumably describe core information-processing units of the mitogen-activated signaling network, which function - in at least a broad range of cell types - to ensure irreversible commitment to the cycle. It can therefore be anticipated that future studies on the systems properties of these modules may provide important clues on mechanisms underlying cell transformation, and also on factors that govern properties related to cellular growth rate such as tumor aggressiveness and metastasis.

Having identified the IMP node modules and characterized their vulnerable constituents, we were then able to exploit this information to develop a potentially more effective strategy for chemotherapy. In view of the pathway redundancies that are inherent to signaling networks, current emphasis is on approaches that involve a concerted pharmacological intervention at multiple key targets (Lehar, Stockwell et al. 2008). Here, the likelihood of generating a synergistic effect constitutes another potential advantage of this strategy. The challenge, however, is to identify the most appropriate combination of nodes to be targeted, and present approaches in this direction are limited by the vast space of high-order combinations that are theoretically

possible (Lehar, Krueger et al. 2008). Here, our present strategy of extracting vulnerable nodes from the core elements that process input signals in a context-specific manner, alone serves to reduce the dimensions of combination space to be explored by several orders of magnitude. Further, as we have demonstrated, the relative functional conservation of these regulatory elements also implies that the benefits of targeting such nodes would be at least less cell type-restricted.

From a broader point of view, the above results also support that targeting nodes derived from modules that regulate temporally distinct stages provides for a more effective means for breaching the fragility barrier of the cell cycle regulatory network. That is, these findings emphasize the potential of exploiting the temporal dimensions of a biological process, an aspect that is presently ignored in ongoing drug development exercises. More specifically though, our results also underscore the potential utility of simultaneous AKT1 and ABL1 (and, SRC?) inhibition, either independently or in conjunction with existing drugs, for anti-cancer treatment. At one level, targeting combinations of nodes whose functional vulnerabilities are conserved can potentially override current limitations faced due to problems of cell- and tissue-type heterogeneity. In addition, the observed synergism between AKT1 and ABL1 inhibition also entails a broader therapeutic window, thus providing an advantage over several of the chemotherapeutic agents currently in use (Kaelin 2005). Our proposal is particularly well supported by emerging evidence that a role for constitutive ABL activity is not restricted to CML alone but, rather, is now implicated in a broader spectrum of cancers (Sirvent, Benistant et al. 2008). This would be consistent with our present description of ABL1 as the key regulatory signaling molecule that drives the G1S phase. Also to be noted in this context is the possibility of exploiting the other

vulnerable nodes that were identified in the individual IMP node modules, through the development of specific inhibitors.

Thus, in summary, our present report describes an integration of results from a siRNA screen with a graph theoretical analysis, to delineate the core signaling modules that regulated the early phases of the cell cycle. The significance of this approach could be established through a comparative analysis with other – independently conducted – screens, and its relevance was underscored by the fact that the modules identified represented regulatory elements that were conserved across a broad range of cell types. A particular highlight of our results was the strong experimental corroboration that could be generated for the inferences gained from the network analysis. This, in turn, allowed for elaboration of novel mechanistic features pertaining to signal-mediated regulation of the cell cycle. Notable among these was the delineation of the distinct core signaling axes that governed early and late G1 phases, and subsequent elucidation of the seamless manner in which the switch between them was integrated. In addition, successful resolution of the least redundant nodes in each of the regulatory modules permitted the development of a strategy for multi-module targeting in order to achieve synergistic disruption of the cell cycle. These latter findings thus pave the way for overcoming the daunting challenge presently faced, in identifying target combinations that can yield synthetic lethal effects in mammalian cells (Kaelin 2005; Kaelin 2009).

Chapter 7

Summary and Conclusion

Chapter 7: Summary and Conclusion

The eukaryotic cell cycle with its vast numbers of regulators and multiple levels of regulation is a very complex cellular process. Despite the redundancy and tissue specific expression of cell cycle regulators, the core cell cycle machinery consisting of the cyclins, CDKs and related molecules is largely conserved across tissues and even organisms.

While a lot is known about the regulation of the cell cycle, not enough is understood about the growth factor dependent upstream signaling networks that regulate the cell cycle. Traditional work in biology has focused on studying individual parts of cell signaling pathways. Though multiple cell surface receptors, ligands and their downstream signaling pathways have been independently characterized, a systems view of their integration into signaling networks regulating temporal cell cycle stages and a resolution of their architecture is lacking. A systems approach will help us to understand the underlying structure of the signaling networks and how the changes in these networks affect transmission and flow of information.

As described in the review of literature, in the recent past, attempts to understand the systems level regulation of the signaling networks regulating the cell cycle have been undertaken. These often employ various high throughput biology techniques including genome level gene expression data and RNAi methodologies among others. While most of these have the limitation of being tissue type dependent, many also end up with only a parts-list of regulators, without shedding any novel insights into the mechanistic regulation of the cell cycle via these signaling networks. While some of these limitations have been recognized, a step in the forward direction has been the use of orthogonal experimental and analytical approaches by various groups in the recent

literature. It has emerged that the analysis of signaling networks requires a combination of experimental and theoretical approaches.

In this study, we have combined the results of an RNAi screen targeting the cellular signaling machinery, with graph theoretical analysis to extract the core modules that process both mitogenic and oncogenic signals to drive cell cycle progression. Since the focus of our study was to understand the growth factor dependent networks, we delineated the G1 phase modules, G1-S phase module and S phase module and their constituent nodes with the help of network analysis. An analysis of the modules revealed that the modules were not distinct, but instead consisted of overlapping nodes which represented the transition events occurring between the contiguous cell cycle stages of the G1, G1-S and S. We found that these modules encapsulate mechanisms for coordinating seamless transition of cells through the individual cell cycle stages and, importantly, are functionally conserved across different cancer cell types.

We also identified and experimentally verified the least redundant nodes in these modules. Pharmacological targeting of the least redundant nodes within these modules led to synergistic disruption of the cell cycle in a tissue-type independent manner. Thus targeting combinations of critical nodes from regulatory modules driving distinct phases of the cell cycle provides for a "*multi-module*" targeting based chemotherapeutic strategy that may be effective against a broad spectrum of cancers.

An analysis of the constituent motifs within the modules helped reveal novel mechanistic insights into the regulation of the early G1 and the G1-S stages of the cell cycle. The IMP node modules identified, defined the signaling axes that govern the G1 and G1S windows of the cell cycle. They capture the network of core pathways that mediate commitment of cells to the division cycle. Based on these finding the

mechanistic model presented in this study suggests that the G1 and G1S IMP node modules represent conserved elements of the mitogen activated signaling cascades.

Therefore, the present study provides a new perspective on the systems level regulation of signaling networks that control the eukaryotic cell cycle and its growth factor dependent stages. It also delineates the signaling events that integrate together to regulate the cell cycle and identifies the regulatory features that process input-output relationships.

Novel aspects of the study

1. The interrogation of siRNA screen data through an iterative graph theoretical analysis represents a novel adaptation of existing analytical approaches that can find more general applicability.

2. We capture, for the first time, the topological modulations in network architecture that drive progression of cell through the individual phases of the cycle. Further, novel mechanisms that mediate this process were also revealed. Indeed the approach described here may be used to delineate the dynamic features of any relevant cellular process.

3. The broad concept that aberrations in cellular behavior may be mapped to the functioning of a few core regulatory modules, which can then be exploited for the development of therapeutic strategies, is also a novel and significant one. This is further exemplified by demonstration of the potential of multi-module targeting approaches.

4. In the present study, we describe a novel signaling axes that capture the network of core pathways that mediate the commitment of cells to the division cycle.

5. Finally, as discussed above, our results confirm that multi-module targeting represents a novel strategy that constitutes a viable addition to existing multi-component therapeutic strategies. In addition, we also describe an approach for rational design of the same.

Chapter 8

Supplementary Information-Figures and Tables

Figure S1

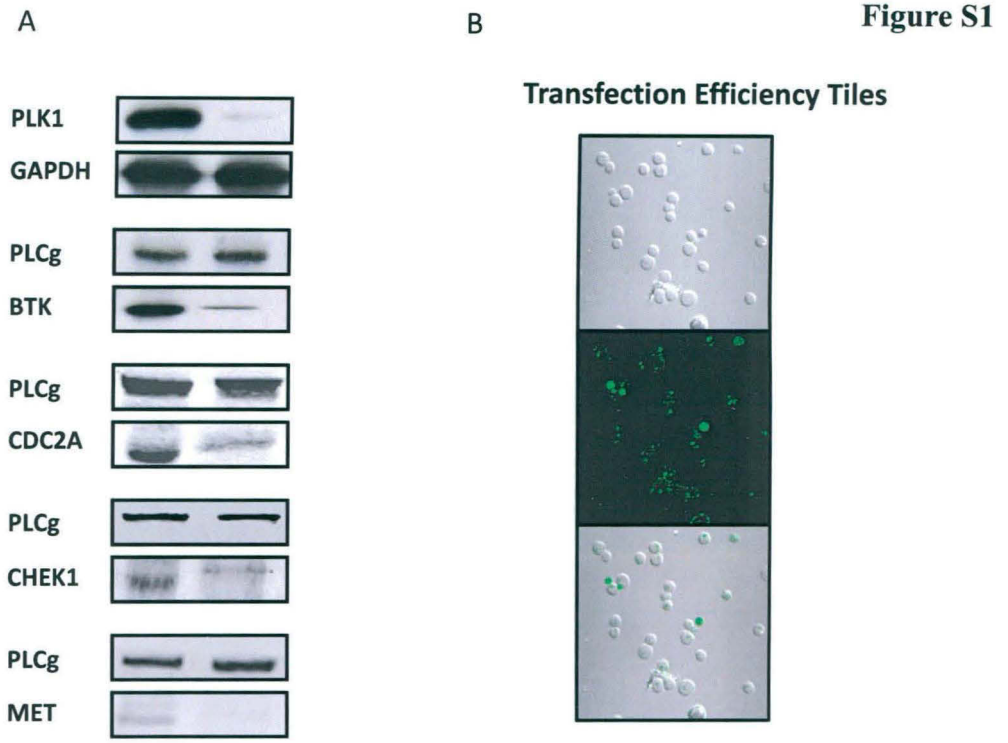


Figure S1. Standardization of the siRNA Screening Procedure: Western blots showing knockdown efficiency.

Panel A *shows the effect of siRNA knockdown of 5 targets at 96 hours post transfection. siRNAs targeting BTK and PLK1 were used for screen standardizations. The reductions in protein levels were monitored at 24, 36, 48, 72 and 96 hours post transfection. At each of these time points cells were harvested, lysed and the levels of the respective proteins in the cytoplasm detected by Western blot analysis. Silencing of select proteins, which includes 3 validated screen hits (CHEK1, MET and CDC2A) at 96 hours after transfection is shown here. Here either GAPDH (for PLK1), or PLCg (for the remaining molecules) were also probed in parallel to serve as the loading control. The antibodies were obtained from Cell Signaling Technology. (Also refer Figure 5.1).*

Panel B *shows the transfection efficiency in CH1 cells to be >98%. Cells were transfected with Alexa-488 labeled siRNAs and the transfection efficiency was monitored up to 72 hours post transfection using confocal microscopy.*

Figure S2

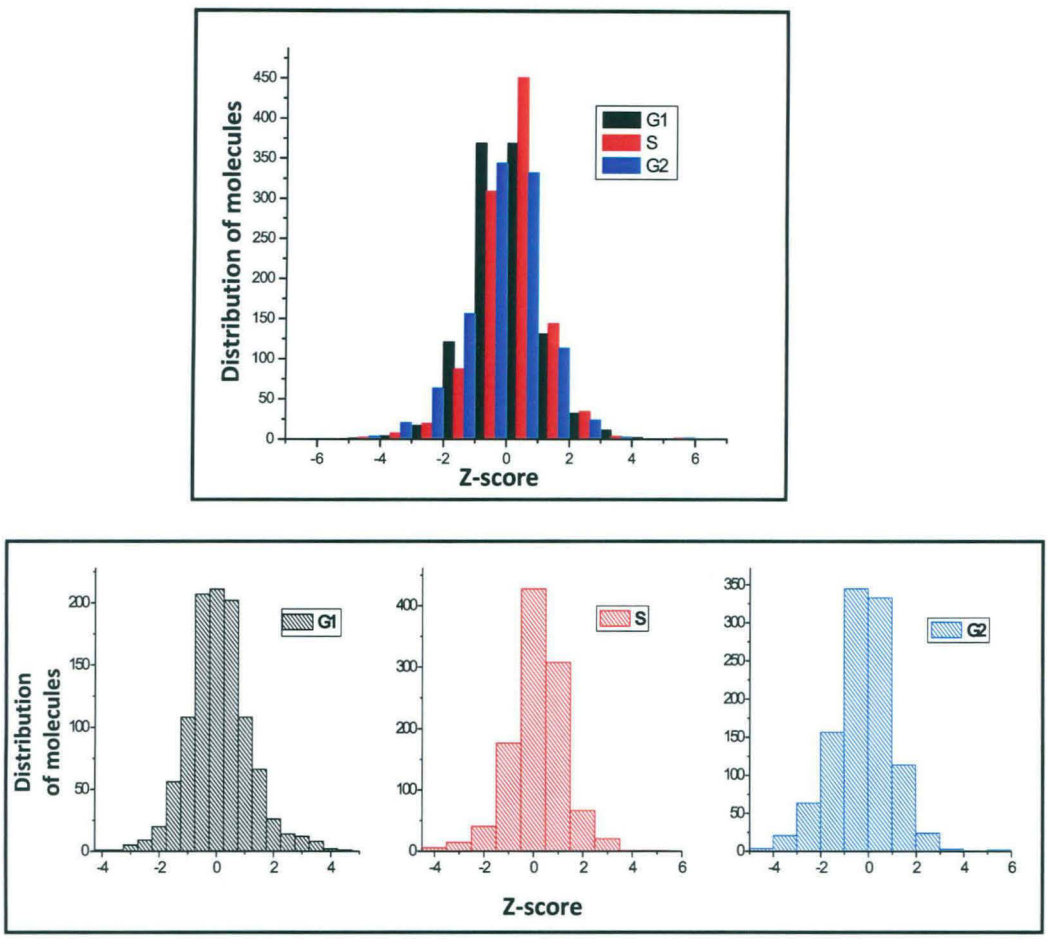


Figure S2: Standardization of the siRNA Screening Procedure: Distribution profiles of complete primary screen.

*Plotted in **figure S2 above** are the mean z-scores of the replicates for the entire screen with siRNAs targeting kinases and phosphatases. The graph shows the distribution of the data points for all the molecules targeted. The normal nature of the distribution curve, with minimal skewness, confirms the robustness of the screen.*

Figure S3

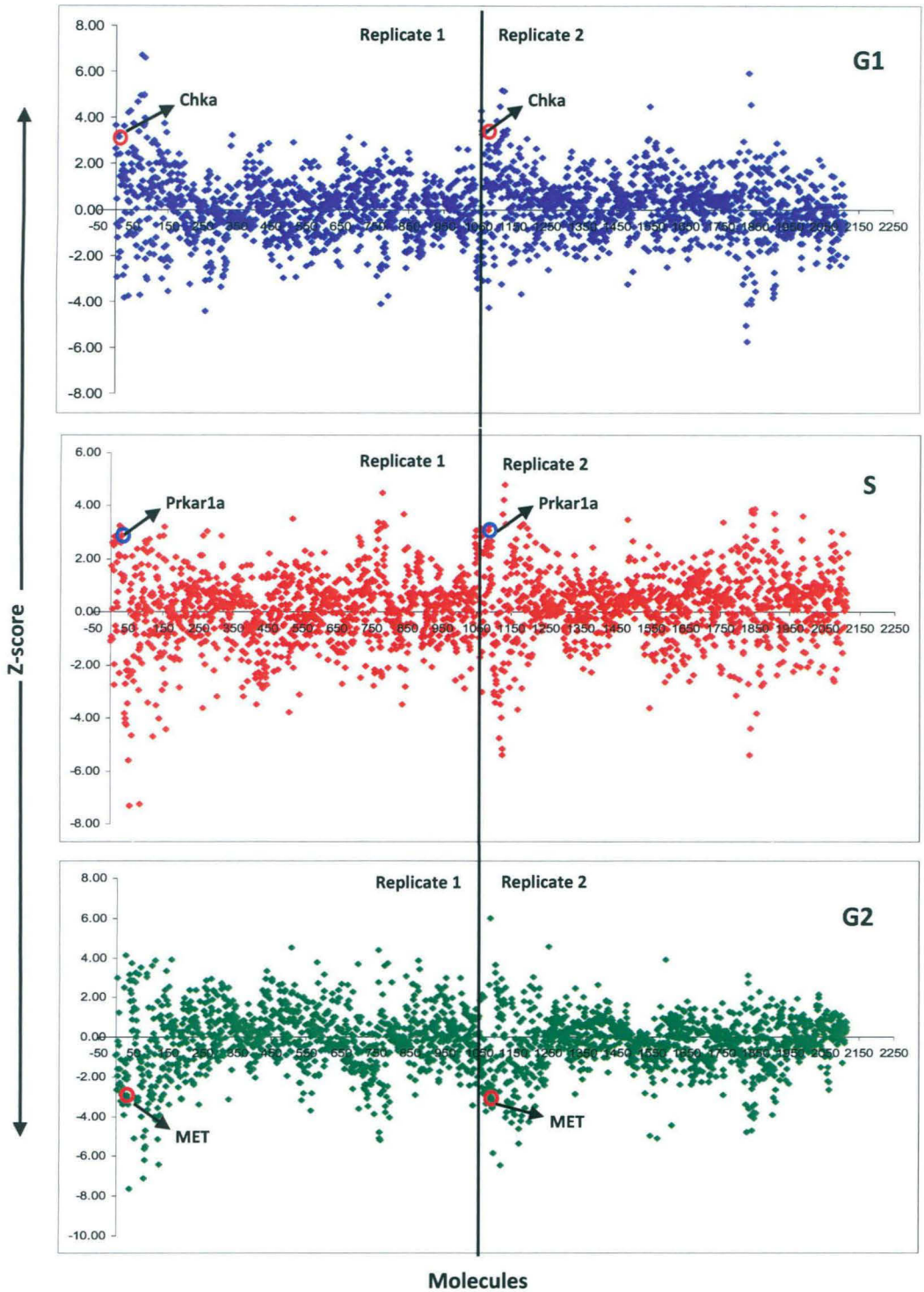


Figure S3: Scatter plot of z-scores to verify reproducibility of the primary siRNA screen

Plotted above are the scatter plots for the two individual replicate z-scores for the three different phases (Refer Figure 5.1).

Figure S4

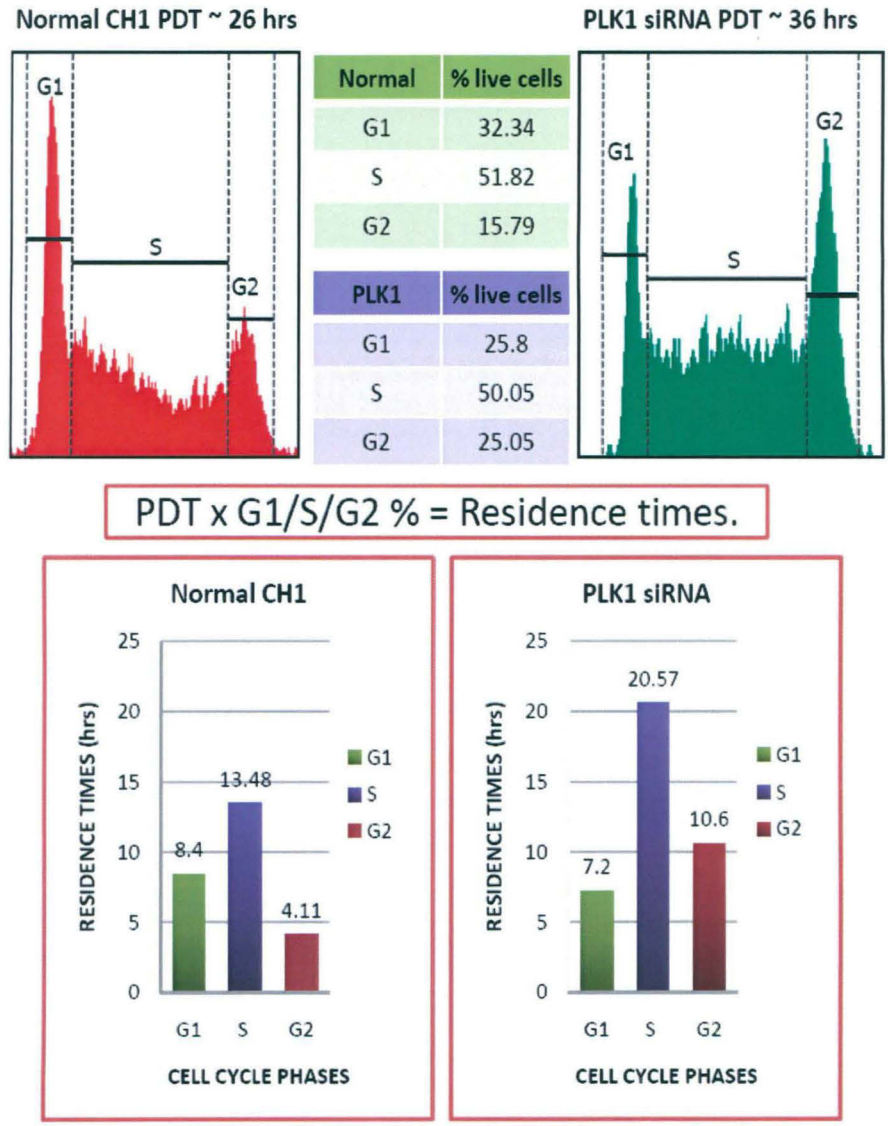
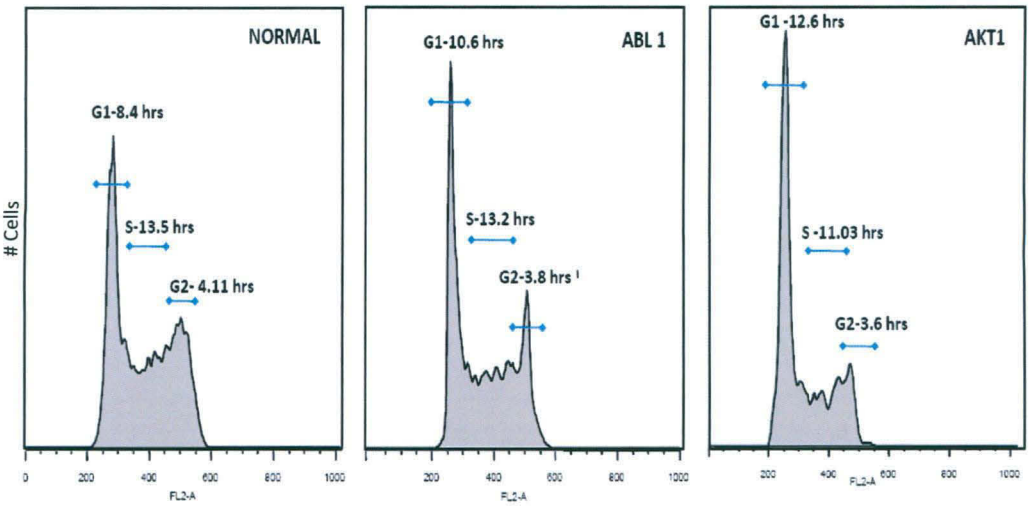


Figure S4: Functional relevance of the identified targets and Calculation of residence time from PDT and cell cycle distribution.

Figure S4 explains the method followed to calculate RT for CH1 cells after targeted siRNA mediated perturbation of screen hits. The approach is illustrated by taking the case of PLK1 as a typical example (Refer Figure 5.2,5.3 and Figure 5.4).

Figure S5

A



B

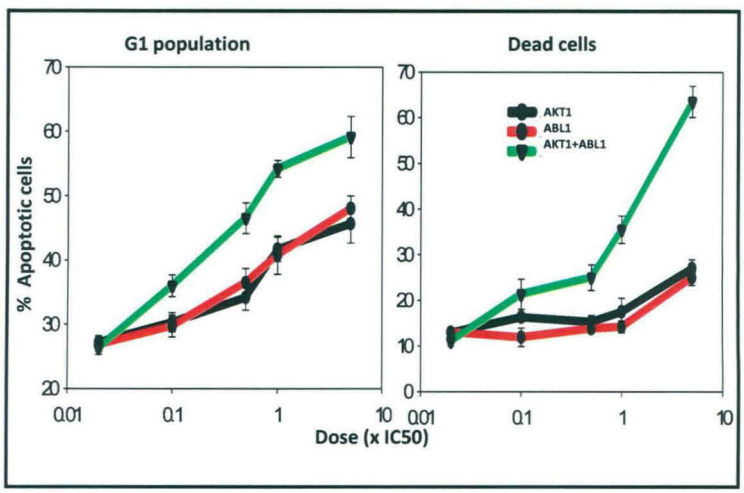
	siRNA sequence 1	siRNA sequence 2	siRNA sequence 3	siRNA sequence 4
BRCA1	GAGCAGCCCUUCACAAAUA	GGAGGAAAGUGGAUAGUUA	GUAGUAAUCUCCAGAAAGA	GAAUGCAGCUUACAAAUGG
ESR1	GGACUUGAAUCUCCAUGAU	UCAAGUCGGUUCCGCAUGA	GAAAGGCGGCAUACGGAAA	GACAAUCGACGCCAGAAUG
PTK2B	GAACCCACCUGGAUUAUCA	GAACAUGGCUGAUCUCAUA	GGACGAAGACUAUUACAAA	GAACAAGCUUGGCCAGUUG
SKP2	CCGGUGCUAUGAUAUAAUA	CGAACUCAGUGAUAAAUGU	CGACUUAAGUGACAGUAUC	GGAAACGAGUCAAGGGCAA
JUN	GAACAGGUGGCACAGCUUA	GAAACGACCUUCUACGACG	CCAAGAACGUGACCGACGA	GCCAAGAACUCGGACCUUC
RB1	GGAGUUUGAUUCCAUUAUA	GCAUAUCUCCGACUAAAUA	UCGAAGCCCUUACAAGUUU	UGCGUUAUCUACUGAAAUA
E2F1	CCACGAGGCCCUUGACUAU	UACAAGCUGUGGAUUCUUC	GGAGGGUGAGGGCAUUAGA	GCAAACAAGGCCCCAUUGA

Figure S5*: **Panel A** shows the histograms obtained from a cell cycle analysis of CH1 cells at 48 hours post treatment with 0.5 x IC50 of ABL1 and AKT1 inhibitor (i.e. Imatinib mesylate and LY294002). The respective RTs of the G1, S and G2 phases are indicated (Figure 5.15A).*

***Panel B** lists the siRNA sequences employed for silencing of the high stress and betweeness targets shown in Figure 5.15B.*

Figure S6

A



B

		LIST OF HUMAN CELL LINES USED
	CELL LINE	DESCRIPTION
1	A549	Human alveolar epithelial cell line
2	AW8507	Human buccal cancer cell lines
3	MCF7	Human breast epithelial adenocarcinoma
4	HEK293	Human embryonic kidney
5	HELA	Human epithelial carcinoma cell line
6	HEPG2	Human hepatocellular liver carcinoma cell line
7	HL-60	Human promyelocytic leukemia cells
8	HUH7	Human hepatoma cell line
9	JURKAT	Human T cell lymphoblast-like cell line
10	NAMALWA	Human burkitts lymphoma
11	RAJI	Human burkitts lymphoma
12	THP1	Human acute monocytic leukemia cell line
13	U20S	Human osteosarcoma expressing wt p53 and Rb
14	U937	Human leukemic monocyte lymphoma cell line

Figure S6: ***Panel A*** *depicts the dose-response profile for increase in G1 population, and increase in apoptotic cells in CH1 cells treated with pharmacological inhibitors of ABL1, AKT1 and the combination of the respective doses. Doses for each inhibitor used were in multiples of their corresponding IC50 values as indicated. Values (mean ± S.D. of three experiments) are expressed as the percent of cells obtained at 48h post treatment (Also refer Figure 5.16A).*

Panel B *lists and describes all the human cell lines used in this study.*

Figure S7

A

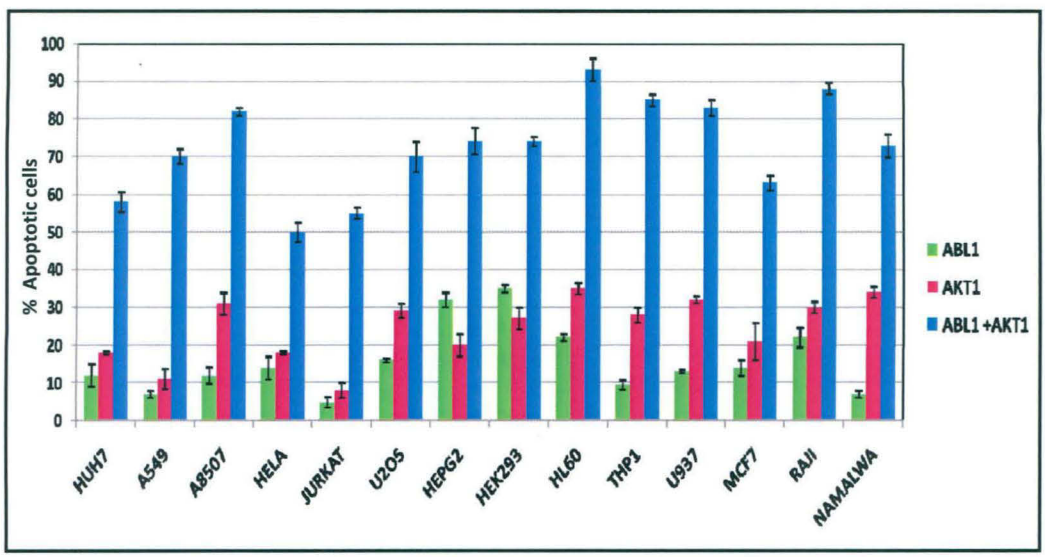


Figure S7: *In **Panel A**, the effect of inhibition of ABL1, AKT1 and their combination (at concentrations corresponding to 5 times their respective IC50 values) on inducing apoptosis of fourteen different human cancer cell lines is shown. Here, the cells were treated with a single addition of the inhibitor or inhibitor combination, and the frequency of apoptotic cells was determined 48h later. In all cases, results are the mean (± S.D.) of three experiments (Also refer figure 5.19).*

Figure S7

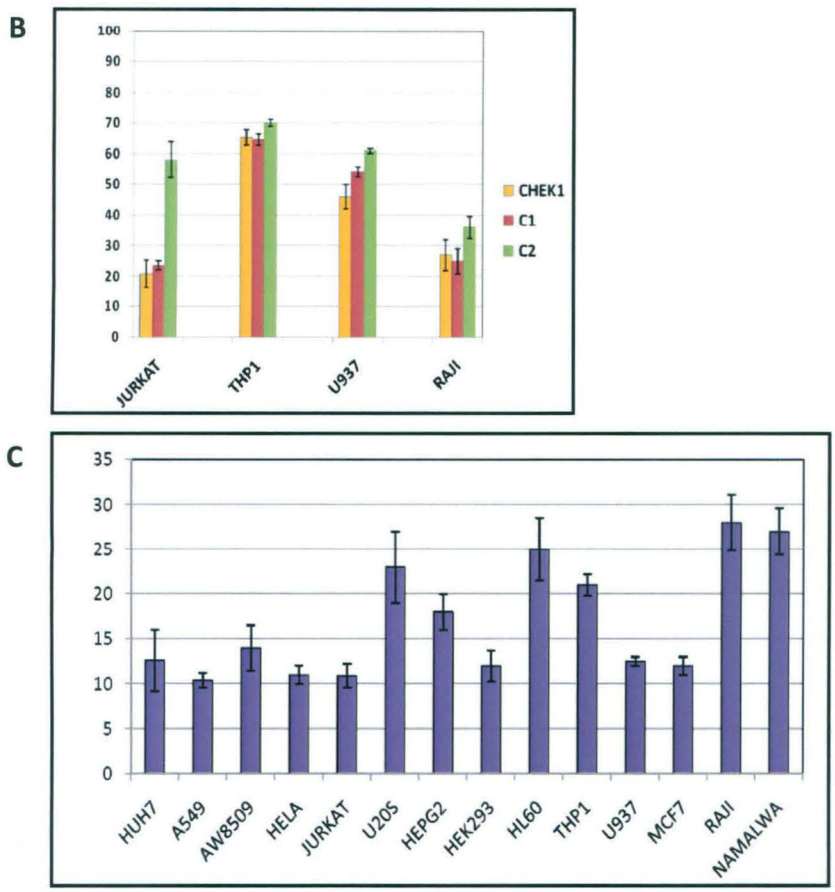


Figure S7: *In **Panel B**, cells of the indicated cancer cell lines were treated with inhibitors against either the indicated kinase, or the kinase combinations and the consequent effect on cellular apoptosis was determined 48 hours later. The concentration of the relevant inhibitor employed was five-fold greater than its corresponding IC50 value in all cases and results are the mean (± S.D.) of three experiments. The protocol followed was similar to panel C. As is evident, with the exception of Jurkat cells, apoptosis resulted from the dominant effects of CHEK1 inhibition in all cases (Refer Figure 5.19A).*

***Panel C** depicts the increase in G1 population observed in the fourteen different human cancer cell lines following treatment with inhibitors of both ABL1 and AKT1 (at concentration corresponding to their IC50 values). Values (mean ± S.D., n=3) are those obtained at 24 h post-treatment and are expressed as increase in the percent of cells in the G1 phase, relative to that in the absence of any inhibitor (i.e. vehicle only).*

Table S1: siRNA Screen Results for Targeted Kinases and Phosphatases

Complete data for primary and sencondary screens
Primary screen hits
Validated targets
siRNA targets inducing cell death
Validated targets inducing cell death

			PRIMARY SCREEN		VALIDATION SCREEN		
				siRNA target			
GENE NAME	Entrez Gene ID	Description	Target 1 (plate A)	Target 2 (plate B)	Target 1	Target 2	Target 3
ATM	11920	ataxia telangiectasia mutated homolog (human)	CACCTGTTTGTTAGTTTATTA	AAGGTCTACGATACTCTTAAA	CAUCUAAUGGUCUAACGUATT	CCAAUUCUUCAGCGUAAUUTT	CUGUAUCCCACACUUAGUATT
Cdc2a	12534	cell division cycle 2 homolog A (S. pombe)	TCCATATGAATTTAAATATAA	TCGGGAAATCTCTCTATTAAA	CUAUGAUCCUGCCAAACGATT	CAAUCAAACUGGCUGAUUUTT	CAUGGAUUCUUCACUCGUUTT
Cdk4	12567	cyclin-dependent kinase 4	ACCCTAGTGTTTGAGCATATA	CCGAACTGATCGGGACATCAA	GAGCGUUGGCUGUAUCUUUTT	CAGUUCGUGAGGUGGCCUUTT	GUUUCUAAGCGGCCUGGAUTT
CDK5	12568	cyclin-dependent kinase 5	CCGGGAGATCTGTCTACTCAA	TGGACTTTATTTAATTTCATA	GCUACAACAUCCUUGGUGATT	GACUAUAAGCCCUACCCAATT	CCAAUGUACCCAGCUACAATT
Chek1	12649	checkpoint kinase 1 homolog (S. pombe)	CACAGGGATATTAAACCAGAA	CAGGAATATTCTGATTGGAAA	CAGGUUACUGUAUCAACAATT	GCUUGUAAACAGUCAGUUATT	GUAUGAAUCAGGUUACUGUTT
Chka	12660	choline kinase alpha	CAGCTTGGTGGCCGAACCCAA	CAGGGCCTACCTGTGGTGTAA	CCUAUGAGAAGUAUCCUUUTT	CACUGAAGAGUUACGUUUATT	GACUCAAAUUCAGCAGGGATT
CLK4	12750	CDC like kinase 4	AAGTATGGTGTTGATAATGAA	CCGGAGATACATTGATGAATA	GUUUGACCUGGUUCGAAGATT	CUGUCAAAGUGGAGACGUUTT	CAGAUCUGCCAGUCUAUAATT
Csnk2a1	12995	casein kinase II, alpha 1 polypeptide	CAGAAGATTTATATGACTATA	ACCGAAGAGCCCTTTAAATAA	GAUAUGUGGAGCUUGGGUUTT	CAGAAGAUUUAUAUGACUATT	CAGAAAGCUUCGGCUAAUATT
EphB2	13844	Eph receptor B2	CACGGACAAGCTACAACACTA	TTCCAGGAACATTAACATGAA	CUUUACCUCUUUCGACGUUTT	GUGUAUGCCGCAACGGCUATT	CUCGAAACAUCCUCGUCAATT
FGFR2	14183	fibroblast growth factor receptor 2	CCGAATGAAGACCACGACCAA	CACATAAACGGCAGTGTTAAA	GGGAUAUCAACAACAUAGATT	GUGAUGAAGAUAGCAGACUTT	CAGAUGACCUUCAAGGACUTT
Guk1	14923	guanylate kinase 1	CGGGACTGACCTCTTAATAAA	CAGGGAGATGATGCAGCGTGA	GGGAGAUGAUGCAGCGUGATT	GUGAUGUCUGUGGUCUAAATT	GAUGAUGCAGCGUGAUAUUTT
IRAK2	108960	interleukin-1 receptor-associated kinase 2	CACCCTCATCCCGACAATAAA	CTCAGTATCCTTGTAAGATAA	CAUUCGAAACCAAGCAUAUTT	GUUCGAAUCAGCGUAUUGATT	GAUUUGGAGUGGUGUACAATT
Itgae	16407	integrin, alpha E, epithelial-associated	TACCATGATGAGGAACTTCTA	CACAAAGGAGGTGACCATGAA	GGUUAGUUUCAAUGGCGA	GAGUGACAUUGGUCGGGAA	UGUGAGGACUCGGAGCGGUU
LEK1	108000	centromere protein F	GAAAGUAGCGCUUGUUGAA	GAGUGAACUUGGAUUGAUA	GAAAUUAGGCUUAAGUUGA	GCGAGUAUCUGACUUAGAA	
MAP3K5	26408	mitogen activated protein kinase kinase kinase 5	AAGGATGAAGATTGAAACTAA	CTGGATCGAATGAGTATCTTA	CAGAUAGUCCACCGGGAUATT	CUGAGUAGCCUUCUGGGUATT	GACAAUAUCUAUUUGGCAUTT
Map4k1	26411	mitogen activated protein kinase kinase kinase kinase 1	CAGGCTGAAACTGTTCGCTTA	CTGGGCCTGCTTGAAAGGAAA	CUGAAACUGUUCGCUUAUATT	CUUUCAUUGGGACACCGUATT	GAUACAAUGAGCUGUGCGATT
Mapkap1	227743	mitogen-activated protein kinase associated protein 1	TAGGGCTTAGCCAGTAGTTTA	AAGGATGGCCTTCTTGGACAA	AGAUAGACCCUGUUACGAA	GAUUAGAACGACUCCGCAA	UCACAAAUUGACAUCGCUA
MAST2	17776	microtubule associated serine/threonine kinase 2	ATGAAGAAGATTAATAAGCAA	AAAGAGGGACATGGAATTAAA	CGAAUACCACGGUCCCAAATT	CGGUUAUCUGUGUCCGAGUTT	CUCUAACACUCCAGAUAGUTT
MET	17295	met proto-oncogene	CTGGACAATGACGGAAAGAAA	AAGAATAATAAATTTGATTTA	GUUAUUCACUGCAUUCGAUTT	CCAAAGUCCUCCUGUCGGATT	CUCAAUAUCAGUAGUAGUUTT
Neo1	18007	neogenin	ACGGATGTGAATGCAGAGATA	CAGGCGGAGCTTACTGTGCAA	GAGGUUAUGCCAUCGGUUA	CAAAUACGCUCUAUGAGUU	GAAACGAGCUGCGUGCAAA
Npr1	18160	natriuretic peptide receptor 1	CAGGGTTGTCCTGAACTTTAA	CCCGGCCAAGATCCAAACAAA	GAUACUGCCUCUUUGGAGATT	CCUAAAUGUGGCUUUGACATT	CCUGUUUGAACUUAAACAUTT
Pkig	18769	protein kinase inhibitor, gamma	AAGCTCAGTTCTGATGATCTA	GAGATGTTATTTATTCAGCTA	GGAAUGCGGUCCCUGACAU	AGGCUGUGAGCGUGAGGAA	CCAGGAAAUAAAUGCCGAU
PKN1	320795	protein kinase N1	TTGAAGAAAGGTGACATTGTA	CGCGATGCAGAAGATGTGAAA	GCAUUGAGGAGCUACGACATT	CACAUGUGGUGCUGCCUGATT	GAGAAAUUGACUGAGUCCATT
PLK1	18817	polo-like kinase 1 (Drosophila)	CCCACTGTATGAATTGTATAA	CAGGAGAGGTCTGGGCATAAA	GGUAUCAGCUGUGUGACAATT	CAACCAAAGUGGAAUAUGATT	CUACGAACGUGGUUCCGCATT
PLK2	20620	polo-like kinase 2 (Drosophila)	CCCGAAGTCCTCAACAAACAA	AAGAGCAGTATTTATTATCAA	CAUUCGAAACCACAAAUCUTT	GUGGUUAAAGUCUGAUAAATT	CCUUUCAGUGGGUCACCAATT
Prkar1a	19084	protein kinase, cAMP dependent regulatory, type I, alpha	CAAGATAAGTCTAAATTTAAA	AAGCTTCTAGATTGTTACTAA	CGGUUAUUCAGCAAGGUGATT	CUAUCAGUGCUGAAGUUUATT	CUUUGAAGUGCGUUAAGUUTT
Prkch	18755	protein kinase C, eta	CAGGATGAGTTTAGAAACTTT	CACGATGAAGTTCAATGGCTA	CGUCUUAACUCCGAUUGAUTT	CACUGUAAACUGGCCGAUUTT	CUAUGAUUAACCAGGAUGATT

Prkg1	19091	protein kinase, cGMP-dependent, type I	CTGGACGATGTTTCTAACAAA	ACCAAGGAGAATAGAGAATAA	CUGAUGAAAUCCUCAGCAATT	CAAGUGAAGACCCAGUCUUTT	GCUAUACUUUACAACUGUATT
Prkg2	19092	protein kinase, cGMP-dependent, type II	CCCGAGGTCATTCTTAACAAA	CCGGGTTTCTTGGGTAGTCAA	GAAGUGUCUCCUUGUUGAATT	CAAUUUGAUUCUCAAGGGATT	GAUUUGAUCCGGAGGCUUUTT
RIPK3	56532	receptor-interacting serine-threonine kinase 3	CCCGACGATGTCTTCTGTCAA	AGAGCTGTTATTTAAAGTCAA	CUCAAGUUCGGCCAAGUAUTT	CAGAGAAGCUGAGUUGGUATT	GAGGUAAAGCAUUAUCUGUTT
SLK	20874	STE20-like kinase (yeast)	ATGGTTGAGATTGACATATTA	GAGGATAAATTTAGCAACAAA	CGCAAUUUGGGAGCUUGAATT	CAGAACAAAUGCAGCGCUATT	GUUAUAACCAGUGAUAGAATT
Srms	20811	src-related kinase lacking C-terminal regulatory tyrosine and N-terminal myristylation sites	CCCATGGGAAATACTGCACAA	CACAAAGTCAGCTTCAATAAA	GACUACAUCUUUGCCCAAATT	CGCUAUAUGUUCCCUCGGUTT	CUGUAUGAGGUCUUCACUUTT
Taf9	108143	TAF9 RNA polymerase II, TATA box binding protein (TBP)-associated factor	TACGTTAACGTGGGAGATCTA	TAGGCAAAGAACTTGCATCAA	ATCAGATGCTGGAGTTTGCCT	ATGCACTGAAACGGAAACGTG	
Taok2	381921	TAO kinase 2TAO kinase 2	TTCGAAGGTGATTATGATTTA	CCCAAATGTTATTGTAAGGAA	GAAGCUACGGCAUCCUAAU	CCUUAUACCACAUUGCACA	CAAGGAUGCUGUACGGGAA
Tex14	83560	testis expressed gene 14	CACGGTGAAGTCAGACATATA	AACCATAACTTTGTTATATAA	GGUUUACGAGCGUAUCGCATT	GGAAUCACUUCUCGCUUCUTT	GAGAUUGUGUAUGUUGAGUTT
Trpm7	58800	transient receptor potential cation channel, subfamily M, member 7	AAGCTGTTAAATAGTGTTTAA	ATGGAAGTATTTAAAGAAGTA	GAGAUUGGGUGAACACUUUTT	GAAUGGAUUGUUAUCGCUUTT	GAAGUAUCAGCGGUAUCAUTT
TSSK5	73542	testis-specific serine kinase 5	CCCAGAGATCCTCATGAGCAA	AAGCACATGAACATAGTTCAA	GAAACUUGAUUAUUUAGGATT	GAUAUUGUCCCAAUGGGAATT	CAGAUUCAUGGUAAUGGAATT
VRK2	69922	vaccinia related kinase 2	TTGATGTATGTTCATAATTTA	CTGGATGTACTGGAATATATA	GAAACUUGAUUAUUUAGGATT	GAUAUUGUCCCAAUGGGAATT	CAGAUUCAUGGUAAUGGAATT
Baiap1	14924	BAI1-associated protein 1	CAGGACGAGCACACTCTCCAA	AAGGACCACCACCAAACCAAA	CCUAUGAAGGAAACUAUUATT	CAAUUAUUGUAAAUGGACATT	GAGCUAAUAACUGUUCACATT
BMPR1A	12166	bone morphogenetic protein receptor, type 1A	TACCAATATCATACTATTTAA	AGCCATCATTTAAATAATTAA	CAGCUAUACACUUACAUCATT	GAUACCUUGCCUUUCUUAATT	GAACUAUUGCCAAACAGAUTT
CASK	12361	calcium/calmodulin-dependent serine protein kinase	TACTGTCATGATAATAACATA	CGGGATCGTTATGCCTACAAA	GACAUCAUCCAGAUUAUUATT	CUUAUUAGGUGCACAUGGUTT	CCAACUAUCACUCCAGGUUTT
CDK10	234854	cyclin-dependent kinase (CDC2-like) 10	AAGGGTTGGTTAACTGGAGTA	CGGGCCTATGGTGTTCCAGTA	GGAAGAUGCCGAAGCGUUA	GGUCAUGGGUUACUGCGAA	GGACACCGAGUGAGAAUAU
CDK2	12566	cyclin-dependent kinase 2	CCGAGAGATCTCTCTCCTTAA	CTGGATGTCATCCACACAGAA	GAGAAGUUGUGGCGCUUAATT	GGAGAAGUUGUGGCGCUUATT	CUCAUCAAGAGCUAUCUGUTT
CDK6	12571	cyclin-dependent kinase 6	ACGGACGGACAGAGAAACCAA	CCCAGCTGAGTTATGATGTAA	CAGAGUAGUGCAUCGUGAUTT	CCACUUACUUGGAUAAAGUTT	GACAGAUAUUGACGAACUATT
CDK8	264064	cyclin-dependent kinase 8	CAGTGACAGTCTGTAGCAGTT	GTAGCAGTTGAAGCTGTGAAA	CUAUCAGCGUUCCAAUCCATT	GUACCUAAGUUUGGAGUUATT	GGUUGUACCUAAGUUUGGATT
CDK9	107951	cyclin-dependent kinase 9 (CDC2-related kinase)	CACAGCCTTGAGGGAAATCAA	CAGTTTGATTGGTTAAATTAA	GACACAAAGUGAGGUGGGATT	GAGAAUGUGGUGAACCUAATT	GACCAAAGCCUCACCGUAUTT
CDKl1	71091	cyclin-dependent kinase-like 1 (CDC2-related kinase)	CACAGCTACAACATTAGATAA	CAGCAAGTATTTAGCATGAAT	GCAAGUAUUUAGCAUGAAUTT	CAGCUCUACCUGAUAAGGATT	CCUACUCUGCACUGGGCUUTT
Ckmt2	76722	creatine kinase, mitochondrial 2	CTCCAGCTTCAAGACACTTAA	TCCCTCCAAGTTTATGGGTAA	GGCAAUAUGAAACGCGUAUTT	GUACGAUAUUUCCAACAUATT	GCUACUACAAGCUGUCUGATT
CSK	12988	c-src tyrosine kinase	CACAGGCAAGCTGCCAGTCAA	CACCATCGTGGCTGTCACCAA	CUGCUACAGACAAUAGGGATT	CACACUGUGUGUGAGCUGUTT	CGCAUCAUGUAUCAUGCGATT
DAPK3	13144	death-associated kinase 3	CACGCAGTTCCTCAAACAAAT	GACGAACATCTCAGCAGTGAA	CGCAUUAAGCUCAUCGACUTT	CCGAGAACAUCAUGUUGCUTT	GUGAGCAUCCUGCGCGAGATT
ErbB2	13866	v-erb-b2 erythroblastic leukemia viral oncogene homolog 2	CCCGCACAGGCTGGACATAAA	CTGGACATTGATGAGACTGAA	GCAAUUUGGAGCUUACCUATT	CCAUCAAGUGGAUGGCAUUTT	GAGAUUCCGGGAGUUGGUATT
ErbB4	13869	v-erb-a erythroblastic leukemia viral oncogene homolog 4 (avian)	CCAGAACAAATTCCTATGTTA	CAGAACAAGAATTGACTCCAA	GUAUAGUGCUGAUCCCACATT	CUCCAAUAGGAAUCAGUUUTT	CCAUUAACCAGAGAAUAGUTT
Frap1	56717	FK506 binding protein 12-rapamycin associated protein 1	CACCAGTGAAATGCTGGTCAA	CACAATCATTCTTCTCATTGA	GGAUCAACCACCAGCGCUATT	CUGACUACGCCUCCCGCAUTT	GGAAGAUCCUGCACAUUGATT
MAP3K6	53608	mitogen-activated protein kinase kinase kinase 6	ACGCTGAGTGATGACATCAAA	AACGACTTCGTCATAAGAATA	GGCUAUUUCCGGGAGACCATT	CUGACAUUGUCAUGAAUCUTT	GUCAGAAGAGUCGAGUAAATT
Map4k2	26412	mitogen activated protein kinase kinase kinase kinase 2	ACCAAGATTCCTGACACCAAA	CTGGCTCTACTGTGTGAACAA	GACAGAUCCUCUCAAUGAATT	CCUUUGACUUCACUAUUGATT	CCAUCAAACGGGCUCCAUUTT
Mapk3	26417	mitogen activated protein kinase 3	GTGCCTGTATCTAATATATAA	CTGCAGATTGTTAGAAAGTGA	CCCAAACAAGCGCAUCACATT	GGCUGAAGGAGUUGAUCUUTT	CACCUUCUCUCACUUUGCUTT
NEK8	140859	NIMA (never in mitosis gene a)-related expressed kinase 8	AAAGCTCTTTATCTTAGTAAA	CAGAACATCCTTCTTGACAAA	CUCAUGAUUGCUAUGGAGUTT	CUUGACAGACCGAGGCAUUTT	GUACUAUGAGAAUUUCCUATT
Nme3	79059	expressed in non-metastatic cells 3	ATGGAGGTTGGCAAGAATGTA	CTTGTTGCCTTTCATAATAAA	GCCGACUGGUUAAAUACAUTT	CUGGUUAAAUACAUGAGUUTT	GCUACUGCGGGAGCAUUAUTT
Pank3	211347	pantothenate kinase 3	CTGGGCAGCATTACTAAATTA	CTCGATGAAACTGTTGGCATA	GUUACUAUCACCAAUAACATT	CAUUGGUUCUGUGGCACGATT	GAUAGGUUGGUCCGAGAUATT
Papss1	23971	3'-phosphoadenosine 5'-phosphosulfate synthase 1	CACCAGCTTTATATCGCCTTA	CTGGACTGTGCTGGTAGAGTA	CACUAAUGAGGAUCCUAAATT	CUUCGAUGAUGAAGCAUAUTT	GGAAAUGAUGACUUUAUGUTT
PBK	52033	PDZ binding kinase	TAGATATATGTAGTTAATATA	AACGTGATTGCTACAATGGAA	GGAAGAACGUGCUUGGCCATT	GUCAUUCCGGCCACGGUUUTT	GGUAUACAAGCUUCUAGCUTT
Pfkm	18642	phosphofructokinase, muscle	CACCACAGTGATGAAGTGTAA	CTGGAGCACATTTCAAGGAAA	AGACUUGUGUGAAGAUCUUTT	GAUGAGAGGCCAGAUCUUATT	GGCCAGAUCUUAAACCUCATT
Pi4k2b	67073	phosphatidylinositol 4-kinase type 2 beta	AACTGGATTGATAATAAACAA	ACCCTATATTTCTGACATGAA	CUAUGUGGAGGUGUUAGAATT	CAUUCGUUUCCAGUUGUUUTT	CCAUCAUUUGUAUUCCGGATT
Pik3c2a	18704	phosphatidylinositol 3-kinase, C2 domain containing, alpha polypeptide	TTGGCAGAAATTATAAACTTA	CCCAAACTTAATATATACTAA	CAAUGAAAUCCGUUGCCUATT	GAUGCUACGGCCAGCAUAATT	CCUACAUUGGCAACCGGCUTT

Pik4cb	107650	phosphatidylinositol 4-kinase, catalytic, beta polypeptide	CCCAAGGGAAATGGAAGGCAA	CACCGGGAAGCAGGAAAGCAA	CCAGUUAGGUUCUUGGGCATT	CAAACACCUUGGCCUUAUATT	GUUGUAAUGACUAACCUAUTT
PINK1	68943	PTEN induced putative kinase 1	CTGGGCTAAATGCATTCTTAA	CAGCGAAGCCATCTTAAGCAA	CACAGUAUCUGAGGGACCUTT	GUCUUUCUCUCGGCGAUCATT	CAACCUUUGCAAAUACUAATT
Pip5k1b	18720	phosphatidylinositol-4-phosphate 5-kinase, type 1 beta	AAGGATCATCCTTACTCCCAA	CAGGATACTACATGAATCTTA	CUGACUUCCUGGUGACGGATT	GAGAUGUGGUCAUUGUGCUTT	GAAUGAAUGUGGCUCGGCUTT
Pkm2	18746	pyruvate kinase, muscle	AACCCTGACTTTATCACAGAA	CTGGGCTTTGCTTCTGTAGAA	CUGACUUCCUGGUGACGGATT	GAGAUGUGGUCAUUGUGCUTT	GAAUGAAUGUGGCUCGGCUTT
Prkce	18754	protein kinase C, epsilon	AAGAAGGACGTTATCCTACAA	CTCGGAAACACCCTTATCTAA	CUGACAACGAGGACGACUUTT	CGUUAUCCUACAAGACGAUTT	CUCCUAUCGGCUACGACGATT
Rage	26448	renal tumor antigen	TCCGTTTAATATATTTGTAAA	CGCGTTCAGGATGAAGAACTA	CAUCAAACCUCACCUGAAATT	GAUGUCAUCGGCACACCUUTT	GAGAAGACACCCAUUAUCATT
RET	19713	ret proto-oncogene	CAGCGCGGTGGTGGAGTTTAA	ATGGACACATTTGATAGCTAA	CAGACAACUGCAGCGAGGATT	GGAACAUCUUGGUGGCUGATT	CAGUGUAUGAUGAGGACGATT
RIOK1	71340	RIO kinase 1 (yeast)	CTGGGATATAGCTATGAATAA	AACGGACTTATTAAGGTATTA	GAGAUACAUGGCUGUAUUATT	GAGUUGUACCUGCAGGUCATT	CAUGCAUUGGAGUUCUUGATT
Rp2h	19889	retinitis pigmentosa 2 homolog (human)	CAGAGTCAATATGATCTCATA	CCCACCAAATTTAGTATGTTA	UGAAACAGUAGGUCGCUUA	GGUAUUAUUUGCCGAUGAU	GGGAGAAGGUUGAUCCAAA
Rps6kb2	58988	ribosomal protein S6 kinase, polypeptide 2	CTCAGTGAGAGTGCCAACCAA	TGCCATGAAGGTCTTAAGGAA	GCCAUGAAGGUCUUAAGGATT	GGAACAUUCUAGAAUCUGUTT	CAGACAGGUGGCAAACUCUTT
JAK2	16452	Janus kinase 2	AAGATTTAATATTTAATGAAA	ATGATTGGCAATGATAAACAA	GUAUUACGCCUGUGUAUCATT	CUUCUACACAGAACAGUUUTT	GAGAUUAUGGUCAACUGCATT
SRPK1	20815	serine/arginine-rich protein specific kinase 1	CAGGCGTTAATGGAACACATA	AAAGAAGTTTGTAGCAATGAA	GAAGUUUGUAGCAAUGAAATT	GAGUGAUUGGAGUCGUUAATT	GUGGUUAUCAUGGGAUAUUTT
Stk23	56504	serine/threonine kinase 23	CTGGAGGTCTTCTATCTCCTA	TGCCTGTTTGTTTCTTAATAA	GUGAAGAUCGGUGAUUUGUTT	CAUACUAAGUGCAAGAUCATT	CCAUUGUUCAACUUAUCGATT
TIE1	21846	tyrosine kinase receptor 1	CTCAAGAAAGTCTTTAATTTA	CCCAGTGAGAATGTGACATTA	GCAACUUUGGCCAAGUGAUTT	CCAUUGAGUCCCUCAACUATT	GUCAUGACCAUGAUGGCGATT
Trim24	21848	tripartite motif protein 24	CAGGTGGAACAGGATATTAAA	CACATTGATGGTGGAGATAAA	CAGGUUACGGCACCUUCUUTT	GGUUUACCUUCCAACCAGUTT	CUUCCAACCAGUUAUCCAATT
TYRO3	22174	TYRO3 protein tyrosine kinase 3	ATGGCTGATGTGGGTAAGAAA	CCCGATCTTTCAATCGAGAAA	CACAGUAGCUGUUAACCCUTT	GCAUUGAGGCCACAUUGGATT	GGAAUUGCAUGCUGGCCGATT
Uck1	22245	uridine-cytidine kinase 1	CGAGGGCATCTTGGTATTCTA	GAGAATCAATAAATAACATAA	CACAAAUCAUGGUGCAACATT	GAUGUUAGAUCCCUUGUUUTT	GAUGUUAGGCUGUCUCGAATT
ULK2	29869	Unc-51 like kinase 2 (C. elegans)	CTACATGAAATTAGTAAGAAA	AAGGCTAGTTGTTTAAAGAAA	CCAGGAUAGCUUCUACCGATT	CACAGAUGCGGAUACCCGUTT	GAGACCUGUUCCAGAUGAATT
Umpk	80914	uridine monophosphate kinase	CCGGATGCCTTTGACAACGAA	CTGGTGATGCCTAATTATGAA	CAAAGACUCUGCGAGUAAUTT	CCGAUAAACAGAGAUUUAUTT	GAGCUUAAUGCCUAGUAUUTT
TTBK1	140810	tau tubulin kinase 1	CAGGAGGTTGATTCCAAAGAA	CACACTGGTCATGACATGTTA	GUGUGCAGAGGGUGCAGGA	CGUACUGUGUGCAGAGGGU	CUCUUUGGCGUCUCAGCUC
Pik3cb	74769	phosphatidylinositol 3-kinase, catalytic, beta polypeptide	CACATGGAAATCTTTAATTTA	CTCGGTTGGAGTGATCTTTAA	CACUAAUUCAGUUCCAGUATT	GUAUGUGUUUGGCGAUCAUTT	CUCUCUAUGCCGUCGUGGATT
Csnk1g2	103236	casein kinase 1, gamma 2	TACGACGTGTACAGAAATGAA	TAGGAAAGAATCTGTATACAA	CUCAUAAAGUCCAGCUUGUTT	CAAACUUCCGAGUCGGCAATT	CGUGUUUGACUACGAGUAUTT
TXK	22165	TXK tyrosine kinase	ATCGTGCCATATGTTATTTAA	CAGAAACGTGGTGACCTTAAA	CAGAAACGUGGUGACCUUATT	CUGCUAUAUUCACAGGGAUTT	CCAUAUACAGUGUAAUGUATT
Rps6ka2	20112	ribosomal protein S6 kinase, polypeptide 2	AAAGGTGTTCTTGGTGAGGAA	CTGGAATAAGCTGTACCGCAA	GGAAUAAGCUGUACCGCAATT	GAGACAUCCUGGCAGAGGUTT	CCCAAUUGUGCAGCAGCUATT
CAMKK2	207565	calcium/calmodulin-dependent protein kinase kinase 2, beta	CAGGAGATTGCTATCCTCAAA	CAGGATCTGATCAAAGGCATA	GACAAUACUUAUUACGCAATT	CAACUGUGAUCCUGGUAAATT	CUGGAAACCUGCUCACGAATT
BCKDK	12041	branched chain ketoacid dehydrogenase kinase	AAGGATGTGGTGACCCTGTTA	GAGGAAAGCTTCAGAATCTAA	CUACUUGCAGCAAGAGUUATT	GCUCUAUUCCGGUCGCUCATT	CCUGUUAGCUGAGGGUCUUTT
BLK	12143	B lymphoid kinase	CACCGGAGTTGTACAATGATA	CCCATGAAACTTCCTCAGAAA	CUCUCAAGUUGGUCCGGAATT	CUAACCAGGACCCAGAUGATT	CGAAGAACCUUUGGGCCCATT
CaMK2a	12322	calcium/calmodulin-dependent protein kinase II alpha	CAGGAAATTATCAAAGTGACA	CGCAAACAGGAAATTATCAAA	GAACGAUGGUGUGAAGGAATT	CUCUUUGUCCGUGGAAUGUTT	CUUAUCUUCGAUCUGGUUATT
BRD4	57261	bromodomain containing 4	CCCATGGATATGGGAACAATA	TCCCTGATTACTATAAGATTA	CCUGAUUACUAUAAGAUUATT	GACAAACCAACUGCAAUAUTT	CUGAUUACUAUAAGAUUAUTT
Bub1b	12236	budding uninhibited by benzimidazoles 1 homolog, beta (S. cerevisiae)	AACGGGCATTTGAATCTGAAA	AGCGAATTTGTTACTCATGAA	CUACUAGAAUUAAGUGCUUTT	GCAAGAUUGAGCCUAGUAUTT	GUACCAAGAUGGCCAUGUUTT
CAMKK1	55984	calcium/calmodulin-dependent protein kinase kinase 1, alpha	TAGGAAGTGAATTACAAACAA	CTCCATAAGAGTGTTTATAAA	GCUAUAUUGCUUUGUCUAUTT	CGCUAUAUUGCUUUGUCUATT	CAUCAAGCCAUCUAACCUATT
TEC	21682	cytoplasmic tyrosine kinase, Dscr28C related (Drosophila)	AAGTTATATATTTATAAGTAA	AACAATAATATCATGATTAAA	GAAAUGGUGUGCCAUCGGATT	CCAUCUACAUCGUUACCGATT	GAGAUAAGUAUGGGUGGUATT
LIMK2	16886	LIM motif-containing protein kinase 2	CACGGAGGACAGGATGGATAA	ATGGGTTTGGTTATAAATACA	GAGAUCAUUGGGCAGGUAUTT	GCAAAGUGAUGGUCAUGAATT	CUGAAAGGCUCUCUGACUATT
Matk	17179	megakaryocyte-associated tyrosine kinase	AAGGTCAGTGACTTTGGCTTA	CTCCATAAAGAGCATTTGGAA	GCUUGUACAUUGUCAUGGATT	GUGACUUUGGCUUAGCCAATT	CUCGUCACCCUGGAGACUATT
PAK4	70584	p21 (CDKN1A)-activated kinase 4	AACCATGTTTGGAAAGAAGAA	CAGGGCCACTTTGAAGAATCA	GGGACUACCGGCACGAGAATT	GCCCAAAUCUUCGAGGGAUTT	GAGACUUCCUGGAUCGGAUTT
PAK6	214230	p21 (CDKN1A)-activated kinase 6	CACAGGGATGTTTCTAAGAAA	CAGGAATAGTAAGGGAATAAA	GAGACUUCCUGGAUCGGAUTT	CAGUCAAGAUGAUGGACCUTT	CAGUUCAAGGCUGCGCUCATT
Pak7	241656	p21 (CDKN1A)-activated kinase 7	CAGAATATCTTTCTTATCTTA	TTGGAGCATGGCTTTAATTAA	CUCAGUAUCUCCUCGAGCATT	CACUGAAUGAGAAUAUUGATT	CAAAUCCCGAGUUGUGUCATT
PLK3	12795	polo-like kinase 3 (Drosophila)	CACGGATAACATGGAACTGAA	TTCGAGGATGCTGACAACATA	CAACUACGUGGCUCCAGAATT	GAAAGUCCCUGGCCCACAUTT	GUCUACCUGGUAUUGGAAATT
Rps6ka4	56613	ribosomal protein S6 kinase, polypeptide 4	AACGTCTTACTTGACTCAGAA	CAGGAGCATACCCGCACCGAA	CUUACCUGGUCCUGGAGUUTT	CCAUGAAGGUGCUACGCAATT	GAUAUCUAGGAGAAUCUUATT

Table S2: Expression status of RNAi hits in CH1 cells

SYMBOL	Control1	Control2	Control3	Log2 C1	Log2 C2	Log2 C3	Average	ACCESSION	PROBE_SEQUENCE
ATM	228.1315833	241.3628667	235.4602167	7.833722382	7.915059927	7.879339513	7.876040607	NM_007499.1	CTCAAGATCCATTCGTAGGATACGTGCTGACTCTTAGGTCATGCTTGTGC
Cdc2a	2171.7975	1763.2885	2618.775583	11.08467388	10.78405282	11.35467672	11.07446781	NM_007659.2	TTTCGGCCTTGCCAGAGCGTTTGGAATACCGATACGAGTGTACACACACG
Cdk4	1969.026583	1538.9585	2147.850667	10.94326687	10.58773861	11.06867798	10.86656115	NM_009870.2	TATAGACCAGGACCTGAGGACATACCTGGACAAAGCACCTCCACCGGGCC
CDK5	102.9948367	104.0833967	101.3625525	6.686428204	6.701596139	6.66338095	6.683801764	NM_007668.2	CTCATGTAACCTGCCACTTCCACCTTCTCGCTGAGCTAGTTAGAACCTCG
Chek1	128.7799167	89.55185333	109.6320992	7.008763812	6.484651385	6.776526457	6.756647218	NM_007691.2	CTGATGGTATGCAGTACTGTCTTCCGAGCCTGTATCTCTCAAGTCATGAT
Chka	68.5927375	63.71801083	66.98071833	6.099983929	5.993629324	6.065673943	6.053095732	NM_013490.1	GTGCAGGTAGATTAGAGCTGTGAAATCCATGTACATTAATACCCAATGGG
CLK4	352.6707917	278.6910333	378.8115	8.462178286	8.122522775	8.565336319	8.383345793	NM_007714.2	TTGACCTGGTTCGAAGAATGTTGGAGTATGACCCAGCGAGAAGGATCACC
Csnk2a1	84.747025	81.50536167	88.42566833	6.405090819	6.348823062	6.466393313	6.406769065	NM_007788.2	CTCTGGTGTGCCACCGAAGAGCCCTTTAAATAAATCGAAGCTTAAAAGAA
EphB2	71.79065417	84.42171333	71.38840833	6.165724139	6.399542204	6.157617931	6.240961424	AK017630	ACACCCAGTTAGGTAGTGTCCTCTTGATCCTGGAGGGGTGAGTAGGATGG
FGFR2	102.9885275	109.9512942	98.38736667	6.686339826	6.780720776	6.620401175	6.695820592	NM_201601.1	CAATGCAGAAGTGCTGGCTCTGTTCAATGTGACGGAGATGGATGCTGGGG
Guk1	2037.240333	2268.919917	1839.263167	10.99240047	11.14778997	10.8449122	10.99503422	NM_008193	TCTGACCACAAGGTGGCTCTGCACCGTGGGCTGACTCACTGGCTGAATGA
IRAK2	119.45645	103.6982742	116.323475	6.900340943	6.696248074	6.861998463	6.81952916	NM_172161.2	CCCACTTCTTAAAAGACCTACAGCCGGTGTCTCCGCAAGTTCTCAAGTTC
Itgae	78.73677083	77.44368917	80.18291083	6.29896564	6.275075775	6.325222887	6.299754767	NM_008399.1	GCAGTGCCTTTAATGAGCACGGTTCCCAGTTTGCCATGCGTTTTGATGAC
LEK1	122.8139083	102.5953667	134.4617333	6.940330141	6.680821769	7.071051842	6.89740125	XM_129658.4	GCAATGGACTCATCTTTTGGTTGCGTGCCAAGCTGAGGGTAAAGAATGGC
MAP3K5	219.502	214.1621833	240.1481417	7.778090275	7.742559942	7.907780835	7.809477017	NM_008580.1	GGCTGAGGATTACACGCTAGTGGATGTTCTTTACTACGTCACACGTGATG
Map4k1	2712.708083	2654.90175	1959.94	11.40551809	11.37444276	10.93659377	11.23885154	NM_008279.1	AAGGCCCACGGATGACCCTACGGCCCCCAGCAACCTCTACATCCAGGAAT
Mapkap1	1106.777933	1156.805967	984.517225	10.11215007	10.17593118	9.943272637	10.07711796	NM_177345.2	GGCAGTCCCAGCAAGACTGTACATAGCCATGCCAGACATCTCCTTGCAAC
MAST2	1469.094333	1741.503917	1650.588583	10.52071132	10.766118	10.68876485	10.65853139	NM_008641.1	GGGGGCCAGCAAGATCATCAGGACTTAGCACTGACATCAGATGAGCTCTT
MET	64.97770667	75.10282167	69.7119625	6.021872921	6.230795207	6.123334337	6.125334155	NM_008591.1	GGATATCCTCAATCTTCTCCACGTTCATTGGGGAACACTACGTCCACGTG
Neo1	68.53786	76.8164475	66.61904917	6.098829241	6.263343341	6.057862858	6.140011813	AK077830	CCTAGTACCACATTGAGCATCACATAACCCTCTTATTTGGGGGCTAGGGT
Npr1	67.74838083	79.973965	76.91179917	6.082114562	6.321458511	6.265133036	6.222902036	NM_008727.4	GGGACTGGAGGGGGACTCCTAAGTTTATAGGGCTGACTGAAATACCCAGT
Pkig	7685.802667	7595.295583	7408.57925	12.90798022	12.8908904	12.85498119	12.88461727	NM_001039391.1	GCAGCTCTAAGTTTCCTCATGGAGCCTCACAGATCCAGTTTTGGGAAAGG
PKN1	115.01665	128.7960333	100.91434	6.845698913	7.008944351	6.656987387	6.837210217	NM_177262.3	TGAATCTCGGTGCTGACTCAGACAGCTCGTCCCAAAAGAGCCCACCAGGG
PLK1	1996.109583	1797.121167	1902.442	10.96297521	10.81147197	10.89363676	10.88936131	NM_011121.2	CCCTGCCCTCGCTCAGCCCACTGTATGAATTGTATAAATGTTTCTATTAA
PLK2	90.06156417	90.80175917	90.80175917	6.492839629	6.504648343	6.504648343	6.500712105	NM_152804.1	GGTTGCTCCCTTGCTGTTTAAAGGCTACAATCAGAGCAGCTTTTGGCTGC
Prkar1a	3075.84925	3001.592167	3238.833417	11.58676908	11.55151225	11.66125855	11.59984663	NM_021880.1	CTGTGTCTTTGTGGACTCTTGGCACCATTGTCTCTGGATGGGTGGGCGAA
Prkch	61.4066425	58.37544417	48.86356417	5.940322819	5.867289717	5.610687194	5.80609991	AK076352	GATATGATTGTCCTTGAGAATCTGCTGCACCACCCCTAGCAGACGCCTCC
Prkg1	70.40987	65.37300417	66.21308	6.137705774	6.030623092	6.049044336	6.072457734	NM_011160	TGATTTGCTTTGTATTTTCCCTGCAGCTGTAATTGCAGAGTGCCTGTTGG
Prkg2	91.37781083	98.46538917	82.974185	6.513771975	6.621544798	6.374590649	6.503302474	NM_008926.2	CTCCAGGGCTCCTTAGACACCAGGGATCTTTACAGCATCCTTAAGTGGGG

RIPK3	117.7401583	108.509055	105.97945	6.879462662	6.761671629	6.727640735	6.789591676	NM_019955.1	AGTTCGGCCAAGTATGACCAAGCACAGTTCGGCAGGGGTAGGGGCTGGCA
SLK	80.33471417	73.8528875	87.97344167	6.327951634	6.206582423	6.458996149	6.331176735	NM_009289.1	AGGGCAGACGAACTAAGACAGATGCTCGGCCATGTTGGTGACGTAGCATC
Srms	80.41175167	70.85510583	75.95346667	6.329334452	6.146799913	6.247043909	6.241059425	NM_011481.2	GAGGAGGAGCCGCAAGGATGTGCTTTGACAGACGAGAGCCAGCAGTTTCC
Taf9	62.018635	60.03240167	80.12658667	5.954629868	5.907669481	6.324209115	6.062169488	NM_027139.3	GGTATCACAGAATACAGCCGAGTCAGCAAATGCACTGAAACGGAAACGTG
Taok2	477.6816833	671.5090083	373.4992083	8.899905748	9.391262943	8.544961375	8.945376689	XM_355941	CAGGGGTCTTGCAGTCAGGGTGCCTGAGTCTCACCATTCTCCTATTGCTG
Tex14	293.675625	240.9205333	275.8449417	8.198079717	7.912413547	8.107713715	8.07273566	NM_031386.1	AGCGACTTGTCAGACTGATTGGAACTGGACCGTGCAAGCATTGTGGCTGT
Trpm7	107.562115	124.6832167	103.8977658	6.749026218	6.96212347	6.699020821	6.80339017	AK009470	GGGGCATCCTTATAATCTCTATATTTGAGAGGCAGGCGGTGGTGCAGGAG
TSSK5	69.6971875	75.88793917	74.06727083	6.123028535	6.245798712	6.210764274	6.193197174	NM_183099.1	GGAACACCAGCCCCATCGCATGCTGAACCTGATTCGTCGAGGTCCCATCT
VRK2	274.7814583	284.7961083	273.9807917	8.102140847	8.153785622	8.097930942	8.11795247	NM_027260.1	CCAGCTATATGAGTTTCCAAGGTTCATATTGCAAGCCCTATCTGGACTGC
Ppm1h	77.28228917	83.50705667	75.570835	6.272065924	6.383826211	6.239757658	6.298549931	NM_176919.2	GCCCTAGAAGAAAGCAGCTCTGGCAGGAAAGTTTCAGCTACTACACCGAG
Cdc25a	1258.136667	1090.016125	1378.143	10.29707293	10.09013376	10.42850988	10.27190552	NM_007658	GCCTGTCACGAGGAGAGAAAAGCCGCCCAAGTGCTGCTGGCCAAATACCA
Ppm1g	1665.153833	1534.250833	1595.684167	10.70143975	10.58331865	10.63995941	10.6415726	NM_008014.2	CCCCAGGAACAGATGATTTCTGCCCTTCCTGACATCAAGGTGCTGACTCT
Inpp5f	341.5078	275.8449417	316.05275	8.41577472	8.107713715	8.304021558	8.275836664	NM_178641.3	CAATGTCTCTGGCATCCTTTAGAGCCTTACTCTGGTTAACTGACTCTCAG
Dusp12	61.30982333	60.50856583	61.39306333	5.938046342	5.919067485	5.940003753	5.932372527	NM_023173	CATATATACAGTTACGTTGCTTGGCAATTCAGATATTTAGTTATGTAAGT
Pik3cb	158.0111167	165.2849333	153.8567583	7.303882251	7.36881141	7.265444007	7.312712556	NM_029094.1	GCTTTCTTCTCCAAAGACATGTGGGAGCCTTCAATCCTCAGTGGCATGGG
Csnk1g2	6984.464167	7513.731833	6004.9845	12.76993372	12.87531391	12.55194481	12.73239748	NM_134002.1	GGCCAGGCCTTGGTCGTGGGAACTGAGTTCTTGCTTTACGACGTGTACAG
TXK	73.29028833	68.7705175	69.027035	6.195550135	6.103718297	6.109089611	6.136119347	NM_013698	TTCAACAGTTTCATGTTTGGGGCTGTATAAATGTTTATAGGAAACATAGC
Rps6ka2	89.17359	110.0961275	90.08039083	6.478544594	6.782619914	6.493141182	6.584768563	XM_123003.1	GCTTTTTAAGTGAGGATCGTCCTGATGGTATCTAATGTCAGCAGTATTCC
CAMKK2	124.1925833	155.569575	140.5037333	6.956435209	7.281416127	7.134464654	7.12410533	NM_145358.1	CTGCACGGGGCTGCTGCACCCTGCGTTTCCATAGCAGCATGTCCTACGGA
BCKDK	444.2751	463.834725	395.19995	8.795309476	8.857467021	8.626438954	8.759738484	NM_009739.2	CACCTAGACACTCCCTACAATGTCCCAGATGTGGTCATCACCATCGCCAA
BLK	433.75515	406.4142083	485.9140583	8.760737077	8.66680703	8.924557362	8.784033823	NM_007549.2	CACACTTAGCTCTTCTTCCAGGTAGGTTGCAGAATCAGGGAAGAATGGTG
BRD2	2968.041583	3409.325417	3355.886167	11.53529559	11.73527059	11.71247806	11.66101475	NM_010238.1	TGGCATAAAGCCCCCTGAGCTGTAGCTGTCCTGTTCACTTCTAAGGGCCC
BRD4	5666.351083	6048.141417	5375.820917	12.46820428	12.56227616	12.39226936	12.47424993	NM_020508.2	CCCCTGGTGAGGTTACATGTGCCATGAGAACCCACTCTACACCATAGTGC
Bub1b	1448.455167	1329.478083	1597.622	10.50029932	10.37664428	10.64171039	10.50621799	NM_009773	GAGCACACTTAGACACATCAGCAGTCCTTCACTAGGGGTCACTCTCAGTA
CAMKK1	174.5630667	144.2470417	151.6561333	7.447604542	7.172397921	7.244660035	7.288220833	NM_018883.1	TCTACATCAACTTATGCCTAGGAGCAGCAGGTAGCCCACTGTCAAAGGGC
Cib1	1837.6835	2179.0365	1700.262083	10.8436726	11.08947465	10.73154143	10.88822956	NM_011870.2	ACTGCTGAGATGTGGCCAAGGTTATGCCTGCGGTGTCAGGGCCAGTGCCG

siRNA hits affecting cell cycle distribution

siRNA targets affecting cell viability

Table S3: Classification of nodes present in modules according to specificity

Sr No	Molecule	Observed Phenotype (Based on increase in RT)	Pubmed IDs for Cell Cycle role known			Role in Cell Proliferation	Residence Time (Hours)			Population Doubling Time (Hours)	Standard Deviation (± hours)			
			G1 and G1/S	S	G2		G1	S	G2	PDT	G1 (SD)	S (SD)	G2 (SD)	PDT(SD)
1	ATM	G1	7923116	11544175:8843193	--	--	10.2	11.5	3	23.9	0.5	0.7	0.2	1.5
2	CHEK1	G1	11544175	11544175	11544175	--	12.2	15.2	3.1	29.5	0.8	1.1	0.3	2.3
3	FGFR2	G1	--	--	--	10572038:15199177	11.7	10	3.8	25.6	1.6	1.7	0.6	3.9
4	IRAK2	G1	--	--	--	--	10.7	11	3.8	26.5	1.1	1.3	0.5	2.9
5	LEK1	G1	14555653	--	--	--	13.9	14.2	3.8	31.4	1	1.2	0.3	2.6
6	MAST2	G1	--	--	--	--	14.7	13.2	2.8	32.8	1.4	1.4	0.4	3.2
7	PKIG	G1	--	--	--	--	9.6	13	4.1	25.8	0.8	1	0.4	2.2
8	PKN1	G1	---	---	---	--	10.1	13.4	3.1	26.3	0.7	1.1	0.4	2.2
9	PLK2	G1	12972611	--	--	12972611	14.6	14.3	3.9	31.8	1	1.3	0.3	2.6
10	RIPK3	G1	--	--	--	10358032	10.5	13.5	2.2	26.1	0.5	0.7	0.2	1.4
11	TEX14	G1	--	--	--	--	11	14.8	1.9	27.5	1.6	2.2	0.5	4.2
12	VRK2	G1	18286197		--	--	12.1	11.9	3.5	27.5	1.3	1.4	0.6	3.3
13	CDK4	G1 and G2	19013283	---	16476733:10318807	---	8.7	12.7	4.8	26.4	0.6	0.9	0.3	1.8
14	CDC25A	G1 and S	14681206:10454565	10454565:10995786	19192479	--	11	16.9	3.2	30.5	1	1.5	0.4	2.9
15	CDC2A	G1 and S	17700700	17635936	19570916:19013283	--	14.9	14.9	3.4	32	0.9	1.2	0.3	2.4
16	CDK5	G1 and S	8302605:1358458	---	---	--	11.5	19.5	3.3	33.6	0.8	1.2	0.3	2.3
17	CHKA	G1 and S	18296102	---	---	--	17.7	17	3.1	37.8	1.4	1.7	0.4	3.5
18	DUSP12	G1 and S	--	--	--	16274472	10.3	20.2	3.9	34.4	0.7	1.3	0.3	2.4
19	MAP3K5	G1 and S	--	---	---	15210709	11.9	25.4	3.4	41.4	1.3	2.3	0.5	4.1
20	MAP4K1	G1 and S	14981547	15634691	--	--	9.8	14.9	2.8	28.1	0.9	1.4	0.4	2.7
21	MET	G1 and S	15064724	---	---	--	14	16.4	2.6	33.5	1.7	2.1	0.6	4.4
22	NPR1	G1 and S	--	--	--	--	9.4	14	3.1	27	0.8	1.3	0.4	2.6
23	PRKAR1A	G1 and S	---	---	---	16569736	9.4	16.2	6.2	29.9	0.8	1.3	0.7	2.9
24	SLK	G1 and S	--	--	16236704	---	10.4	24	3.1	37.5	0.3	0.5	0.1	0.9
25	SRMS	G1 and S	--	--	--	--	10.9	16.7	1.2	28.6	1.8	2.6	0.4	4.9
26	TAOK2	G1 and S	--	--	--	10660600	13.6	17.1	1.4	31.8	1.4	1.7	0.3	3.4

27	TRPM7	G1 and S	19515901	---	---	18590694:19515901	11	16.2	2.1	29.2	1.3	1.8	0.5	3.7
28	TSSK5	G1 and S	--	--	--	--	9.8	15.2	1.8	26.6	1.4	2.1	0.4	3.9
29	CSNK2A1	G2	12396231	---	12396231	---	6.1	12.7	7.6	26.1	0.9	1.4	0.8	3.1
30	PRKCH	G2	16434969	---		8454583:8336936:9310352	3.4	13	12.7	28.3	0.9	1.9	1.6	4.4
31	PRKG1	G2	---	---	---	11073964	7.6	15.3	7.2	28.6	0.9	1.7	1	3.6
32	PRKG2	G2	--	--	--	--	8.1	15	7	28.5	0.9	1.5	0.8	3.1
33	CLK4	S	--	---	---	---	6.8	21.5	5.2	33.1	0.9	2.3	0.7	3.9
34	EPHB2	S	--	--	--	15973414	6.1	21.6	1.4	28.9	0.5	1.6	0.2	2.4
35	GUK1	S	--	---	---	---	9.4	29.7	4.2	43.1	0.8	2	0.4	3.3
36	ITGAE	S	--	---	---	---	7.4	18.3	1.9	28.2	0.7	1.3	0.3	2.3
37	MAPKAP1	S	--	--	--	17012250	7.5	15.9	3.4	27.1	0.5	1.1	0.3	1.9
38	NEO1	S	--	--	--	--	7.6	17.3	3.3	28.6	0.7	1.4	0.4	2.5
39	PPM1G	S	--	17054950	--	---	6.3	14.8	2.1	23.1	0.5	1	0.2	1.7
40	PPM1H	S	--	---	---	18059182	8.6	14.5	1.9	25.5	1.2	2.2	0.4	3.8
41	TAF9	S	--	12837753	--	12837753	8.1	15.5	2	25.6	1.1	1.7	0.3	3.1
42	INPP5F	S and G2	--	---	---	---	9.8	17.5	6.8	34.2	1	1.3	0.5	2.8
43	PLK1	S and G2	--	---	15838519:16557283	---	9.4	19.3	8.6	36.6	1.1	2	0.7	3.8
	Normal						8.4	13.5	4.1	26	1.5	1	0.5	2.5

--- *Role in Cell cycle/Proliferation not well known*

The table provides the classified list of molecules based on their exclusivity for regulation of a particular phase in the identified modules.

Table S4: List of phosphoproteins identified

	Molecules showing increased phosphorylation in G1 phase
	Molecules showing increased phosphorylation in S phase

SR. NO.	PHOSPHOPEPTIDE CONTAINING PROTEIN	NAME IN NETWORK	PROTEIN DESCRIPTION	IDENTIFIED PHOSPHOPEPTIDE
1	ABCG5	ABCG5	ATP-binding cassette sub-family G member 5	YTAMLALCR
2	RRAGA	RRAGA	Ras-related GTP-binding protein A	MPNTAMKKK
3	ADCY6	ADCY6	(Ca(2+)-inhibitable adenylyl cyclase)	LSRNIVR
4	AKT2	AKT2	RAC-beta serine/threonine-protein kinase	THFPQFSYSASIR
5	LEPR	LEPR	Leptin receptor precursor	TLASVVKASVFR
6	RT35	MRPS28	28S ribosomal protein S35	TFSSVASPAAPR
7	RBCC1	RB1CC1	RB1-inducible coiled-coil protein 1	RSTELVLSPDMPR
8	HNRPQ	SYNCRIP	Heterogeneous nuclear ribonucleoprotein Q (hnRNP Q) (hnRNP-Q)	LFVGSIPK
9	TAF5	TAF5	Transcription initiation factor TFIID subunit 5	DSIGSKSKK
10	AFF4	AFF4	AF4/FMR2 family member 4	SPAQSDSTTQRR
11	NELFA	WHSC2	Negative elongation factor A	LLDISELNTVGAGREAK
12	DAPK1	DAPK1	Death-associated protein kinase 1	NRFGNDLHVSNKLFVLDAGASGSK
13	DLDH	DLD	Dihydrolipoyl dehydrogenase	VVHVNGFGKITGK
14	DNAS1	DNAS1	Deoxyribonuclease-1 precursor	DSHLVAVGK
15	SYIC	IARS	cytoplasmic	VQQVPENISFPAEEEK
16	NUP53	NUP153	Nucleoporin NUP53	VSTMRPLATAYK
17	ADDB	ADD97	Beta-adducin (Erythrocyte adducin subunit beta) (Add97)	SPSTESQLMSKGDADTK
18	ALEX	ALEX	Protein ALEX	SPGTSDLK
19	NOCT	CCRN4L	Nocturnin (CCR4 protein homolog)	TVSSMGNGTSR
20	CHST5	CHST5	Carbohydrate sulfotransferase 5	VHVLVLSSWR
21	CLD8	CLDN8	Claudin-8	SPSIYSK
22	CSDC2	CSDC2	Cold shock domain-containing protein C2 (RNA-binding protein PIPPin)	TYSATARASAGPVFK
23	DCTN6	DCTN6	Dynactin subunit 6	GSSTPVKN
24	DHX36	DHX36	Probable ATP-dependent RNA helicase DHX36	GSGQGSSGGGGGGSR
25	E41L2	E41L2	Band 4.1-like protein 2 (Generally expressed protein 4.1)	MTTEVGSASEVKK
26	FA18B	FAM18B	Protein FAM18B	VGSKK
27	GDF7	GDF7	Growth/differentiation factor 7 precursor	AVTASESSPLALR
28	T2EB	GTF2B	Transcription initiation factor IIE subunit beta (TFIIE-beta)	QGISSMQESGPK
29	HES5	HES5	Transcription factor HES-5	SSAKAAAAAVSTSR
30	HXD10	HOXD8	Homeobox protein Hox-D10 (Hox-4.5) (Hox-5.3)	KMNESASGQEPTK
31	KCNC3	KCNC3	Potassium voltage-gated channel subfamily C member 3	SPITPGSR
32	LNP	KIAA1715	Protein lunapark (Protein ulnaless)	TRPQAPR
33	SCF	KITLG	Kit ligand precursor (C-kit ligand) (Stem cell factor)	NDSSSSNR
34	KLF4	KLF4	Krueppel-like factor 4	RTATHTCDYAGCGKTYTK
35	LMBL3	L3MBTL1	Lethal(3)malignant brain tumor-like 3 protein	LSGDTSPPTTPSFPR
36	LDHA	LDHA	L-lactate dehydrogenase A chain	IVSSKDYCVTANSK
37	LRP6	LRP6	Low-density lipoprotein receptor-related protein 6 precursor	MTSVATAK
38	MAP9	MAP9	Microtubule-associated protein 9	VSSASGR
39	RM35	MRPL35	mitochondrial precursor	MTTSFWKR
40	MYO7A	MYO7A	Myosin-VIIa	SPTLSSQFKR
41	MYST1	MYST1	Probable histone acetyltransferase MYST1	TYAEMDPTTAALEK
42	PEX14	PEX14	Peroxisomal membrane protein PEX14	EPLIATAVK
43	PHLB2	PHLDB2	Pleckstrin homology-like domain family B member 2	SPSPMGTSVR
44	ARVC	PKP2	Armadillo repeat protein deleted in velo-cardio-facial syndrome homolog	SPSVDSTRK
45	RAB6A	RAB6A	Ras-related protein Rab-6A	VAAALPGMESTQDR
46	RGL2	RGL2	Ral guanine nucleotide dissociation stimulator-like 2 (RalGDS-like factor)	APSVISRVLKK
47	RNF25	RNF25	E3 ubiquitin-protein ligase RNF25	TSGAGPNQQRPGETQK
48	RS10	RPS10	40S ribosomal protein S10	SRPETGRPRPK
49	SACS	SACS	Sacsin	TTSCLK
50	SC5A3	SLC5A3	Sodium/myo-inositol cotransporter	TTTFWSKK
51	SMAD5	SMAD5	SMAD5 protein	MTSMASLFSFTSPAVK
52	SPEG	SPEG	Striated muscle-specific serine/threonine protein kinase	SPRGSVAER
53	SRGP2	SRGAP3	SLIT-ROBO Rho GTPase-activating protein 3	DINSVAGVLK
54	SYNPO	SYNPO	Synaptopodin	VSAPPSAASTFSR
55	TAC2N	TAC2N	Membrane targeting tandem C2 domain-containing protein 1	IQTQTPKKK
56	TAF1C	TAF1C	TATA box-binding protein-associated factor RNA polymerase I subunit C	MTIVSAK
57	TAOK1	TAOK1	Serine/threonine-protein kinase TAO1	TASLVTR
58	TDRD3	TDRD3	Tudor domain-containing protein 3	TMNSEAFSGLK
59	TM6S1	TM6S1	Transmembrane 6 superfamily member 1	TPRSRILGCSSMLR
60	TTC16	TTC16	Tetratricopeptide repeat protein 16 (TPR repeat protein 16)	MTDPTKRR
61	ZN318	ZNF318	Zinc finger protein 318	TPGVPK
62	ABCA1	ABCA1	ATP-binding cassette sub-family A member 1	VLTGSEK
63	ABCAD	ABCA13	ATP-binding cassette sub-family A member 13	MATDELASK
64	COPD	ARCN1	Coatomer subunit delta (Delta-coat protein)	LMNTGK
65	ARHG3	ARHGEF3	Rho guanine nucleotide exchange factor 3	DLTLEDLQDGEVR
66	CASP7	CASP7	Caspase-7 precursor	DLEIMQILTRVNDR
67	CD5R2	CDK5R1	Cyclin-dependent kinase 5 activator 2 precursor (CDK5 activator 2)	GESRLK
68	CSK21	CSNK2A1	Casein kinase II subunit alpha	GGPNIITLADIV
69	DCAK3	DCLK3	Serine/threonine-protein kinase DCAMKL3	RAPEKESK
70	DPP3	DPP3	Dipeptidyl-peptidase 3	AMSAKFER
71	FYB	FYB	FYN-binding protein	GGKNELSFK
72	GAK	GAK	Cyclin G-associated kinase	MASMK
73	GBGT2	GBGT2	Guanine nucleotide-binding protein	MAQDLSEK
74	GOGA4	GOLGA4	Golgin subfamily A member 4	LHAEELASK
75	IER2	IER2	Immediate early response gene 2 protein	ELYLSAK
76	ISL2	ISL2	Insulin gene enhancer protein ISL-2 (Islet-2)	THVDKQAEK
77	KCNC4	KCNC4	Potassium voltage-gated channel subfamily C member 4	IIINVGGTR
78	KCNQ5	KCNQ5	Potassium voltage-gated channel subfamily KQT member 5	SASANISR
79	KI18A	KIF18A	Kinesin family member 18A	VVHVVDKHILSFDPK
80	MAGI2	MAGI2	WW and PDZ domain-containing protein 2	MSKSLK

81	PXR	NR1I2	Orphan nuclear receptor PXR	FALTLK
82	PPIL1	PPIL1	Peptidyl-prolyl cis-trans isomerase-like 1	GGASIYGK
83	R3HD2	R3HDM2	R3H domain-containing protein 2	SNSNTTQETLEIMK
84	SCRT1	SCRT1	Transcriptional repressor scratch 1	TYATSSNLSR
85	TAOK2	TAOK2	Serine/threonine-protein kinase TAO2	LDETQEAEFQAL
86	TGFB1	TGFB1	Transforming growth factor beta-1 precursor (TGF-beta-1)	NNAIYEKTK
87	THNSL	THNSL1	Threonine synthase-like 1	AEIIGSQR
88	TPST1	TPST1	Protein-tyrosine sulfotransferase 1	SGTTLMRAMLDAHPDIR
89	UGDH	UGDH	UDP-glucose 6-dehydrogenase	TYGMGK
90	1433Z	YWHAZ	14-3-3 protein zeta/delta (Protein kinase C inhibitor protein 1) (KCIP-1) (SEZ-2)	TAFDEAIAELDTLSEESYK
91	ABCD4	ABCD4	ATP-binding cassette sub-family D member 4	TLMGPIVTK
92	ADA33	ADAM 33	ADAM 33 precursor	TPGDTK
93	Add3 (Addl)	ADD3	Gamma-adducin	AEDSSK
94	CENG3	AGAP3	(MR1-interacting protein) (MRIP-1) (CRAM-associated GTPase) (CRAG)	LLVMLLAHGSK
95	Aktip (Fif) (Ft1) (Fts)	AKTIP	AKT-interacting protein	MNPLWSMSAGSVRK
96	ALPK1	ALPK1	Alpha-protein kinase 1	SSQNSSSYSPWWLK
97	Stradb (Als2cr2) (Papk) (Syradb)	ALS2CR2	STE20-related kinase adapter protein beta	MSLLDCFCASR
98	ANKS1	ANKS1A	Ankyrin repeat and SAM domain-containing protein 1A	RSLSTN
99	Arhgef12 (Kiaa0382) (Larg)	ARHGEF12	Rho guanine nucleotide exchange factor 12	MSGTQSTITDR
100	Arhgef2 (Kiaa0651) (Lbc1) (Lfc)	ARHGEF2	Rho guanine nucleotide exchange factor 2	AAVASVTPEK
101	ATRX	ARTX	Transcriptional regulator ATRX	TMSRIGAARK
102	DDEF1	ASAP1	5-biphosphate-dependent ARF1 GTPase-activating protein (PIP2-dependent ARF1 GAP)	KTCASSSAK
103	ATP4A	ATP4A	Potassium-transporting ATPase alpha chain 1	SPECTHESPLETR
104	BAG4	BAG4	BAG family molecular chaperone regulator 4	RSGYGPSDGPSYGR
105	Bai1	BAI1	Brain-specific angiogenesis inhibitor 1	QQTPNKR
106	BMI1	BMI1	Polycomb complex protein BMI-1	TRPLLNIRSDK
107	Brca2 (Fancd1)	BRCA2	Breast cancer type 2 susceptibility protein homolog	DDTAAKTLVLCISDIISPSTK
108	BRD2	BRD2	Bromodomain-containing protein 2	QPMDMGTIK
109	Brd7 (Bp75)	BRD7	Bromodomain-containing protein 7	ERTDACQSGEDSGCWQR
110	BRSK2	BRSK2	BR serine/threonine-protein kinase 2	MTSTGK
111	CB39L	CAB39L	Calcium-binding protein 39-like	GYEAPQIALR
112	CAC1E	CAC1E	Voltage-dependent R-type calcium channel subunit alpha-1E	DSASASRSR
113	Camk2b (Camk2d)	CAMK2B	Calcium/calmodulin-dependent protein kinase type II beta chain	STVASMMHR
114	CAMKV	CAMKV	CaM kinase-like vesicle-associated protein	TTGKLHTCK
115	CCD45	CCDC45	Coiled-coil domain-containing protein 45	RSQVQKTYSR
116	CD2L5	CDC2L5	Cell division cycle 2-like protein kinase 5	SPYSPVLRR
117	SPEC1	CDC42SE1	Sperm antigen with calponin homology and coiled-coil domains 1	SSTSLAFESRLSK
118	CDKL5	CDKL5	Cyclin-dependent kinase-like 5	SPSYR
119	CELR2	CELR2	Cadherin EGF LAG seven-pass G-type receptor 2 precursor	LLETEER
120	CFAB	CFB	Complement factor B precursor	VKDASEVVTPR
121	Cflar (Cash)	CFLAR	CASP8 and FADD-like apoptosis regulator	YNSRLQNGR
122	CH3L1	CH3L1	Chitinase-3-like protein 1 precursor	TRNTNLK
123	CHD1	CHD1	Chromodomain-helicase-DNA-binding protein 1	DSSSGTER
124	CLCN1	CLCN1	skeletal muscle	FVSASSTYGELR
125	KC1D	CSNK1D	Casein kinase I isoform delta	IEYIHSK
126	CP3AB	CYP2R1	Cytochrome P450 3A11	DFGPVGIMSK
127	DAB2P	DAB2	Disabled homolog 2-interacting protein (DAB2-interacting protein)	TRPARESPQERPGSRR
128	Dapp1 (Bam32)	DAPP1	Dual adapter for phosphotyrosine and 3-phosphotyrosine and 3-phosphoinositide	AELLGGNMSTQDPSELWGR
129	DCAK1	DCLK1	Serine/threonine-protein kinase DCAMKL1	SPGPSR
130	Dek	DEK	Protein DEK	NSSGTTR
131	DHX33	DHX33	Putative ATP-dependent RNA helicase DHX33	MSMPIMSSR
132	WAPL	DIF2	Wings apart-like protein homolog (Dioxin-inducible factor 2) (DIF-2)	SPTKAVYNAR
133	DNMBP	DNMBP	Dynamin-binding protein	FSIMDFYSEK
134	DOCK1	DOCK1	Dedicator of cytokinesis protein 1	QTSVDSGIVQ
135	DOCK2	DOCK2	Dedicator of cytokinesis protein 2	VVHVPYRDSK
136	E2AK3	E2AK3	Eukaryotic translation initiation factor 2-alpha kinase 3 precursor	SMSSSGTK
137	Egf	EGF	Pro-epidermal growth factor	SVVGGSDLSGK
138	IF38	EIF3C	Eukaryotic translation initiation factor 3 subunit 8	MDEEFTK
139	ELAV1	ELAVL1	ELAV-like protein 1	ILQVSFK
140	EPN2	EPN2	Epsin-2	RDTVKVPK
141	Eps8	EPS8	Epidermal growth factor receptor kinase substrate 8	AALEDSNGSSELQEIMRR
142	Erc2 (Cast1) (D14Ertd171e) (Kiaa0378)	ERC2	ERC protein 2	TISNLEGSPSR
143	EXOC1	EXOC1	Exocyst complex component 1	STDYGK
144	Ezh2 (Enx1h)	EZH2	Polycomb protein EZH2	EFAAALTAER
145	F126A	FAM126A	Protein FAM126A (Down-regulated by CTNNB1 protein A)	ISRNAVTSMSIR
146	FAM3B	FAM3B	Protein FAM3B precursor	LLSGGGRDK
147	Fbn1 (Fbn-1)	FBN1	Fibrillin-1	GGGGHDALKGPNVCGSR
148	FXL21	FBXL21	F-box/LRR-repeat protein 21	YTPDEVHTEVSK
149	FBX18	FBXO18	F-box only protein 18	YTATTK
150	FGF13	FGF13	Fibroblast growth factor 13	CVSSPSKGK
151	Fhod1 (Fhos1)	FHOD1	FH1/FH2 domain-containing protein 1	RTMTVVLDPK
152	FKBP4	FKBP4	FK506-binding protein 4	VLQLYPSNKAAK
153	FLCN	FLCN	Folliculin	SPTATESRS
154	FMN1B	FMN1	isoform IV	AETGSEGSQGKGR
155	Fmr1 (Fmr-1)	FMR1	Fragile X mental retardation protein 1 homolog	GSRPYRNR
156	FOXP4	FOXP4	Forkhead box protein P4	MTGSPTLVK
157	GBLP	GBLP	Guanine nucleotide-binding protein subunit beta 2-like 1	MTEQMTLR
158	GCKR	GCKR	Glucokinase regulatory protein (Glucokinase regulator)	LGSGGVVKK
159	GDF2	GDF2	Growth/differentiation factor 2 precursor (GDF-2)	SSTPASNIVR
160	Gemin5	GEMIN5	Gem-associated protein 5	HASLQTSQR
161	GLCE	GLCE	D-glucuronyl C5-epimerase	GSRAKHN
162	GBG5	GNAT1	Guanine nucleotide-binding protein	MSGSSSVAAMKK
163	GPC6	GPC6	Glypican-6 precursor	VWSALPYTICK
164	GP139	GPR139	Probable G-protein coupled receptor 139	HGKPIKVSP
165	GASP2	GPRASP2	G-protein coupled receptor-associated sorting protein 2 (GASP-2)	SSSRARPK

166	GSTA3	GSTA3	Glutathione S-transferase A3	LRSDGSLMFQQVPMVEIDGMK
167	GTPB2	GTPBP2	GTP-binding protein 2 (GTP-binding-like protein 2)	VGADITVLR
168	HAX1	HAX1	HS1-associating protein X-1 (HAX-1) (HS1-binding protein)	DSMLK
169	HBS1L	HBS1L	HBS1-like protein	TMNSEAFSGLK
170	Hdac6	HDAC6	Histone deacetylase 6	MTSTGQDSSTR
171	Hras1 (Hras)	HRAS	GTPase HRas	MTEYK
172	WAPL	IER3	WAPL_MOUSE(Dioxin-inducible factor 2) (DIF-2)	SPTKAVYNAR
173	Ilf3	ILF3	Interleukin enhancer-binding factor 3	AEHMTRTLR
174	ITAE	ITAE	Integrin alpha-E precursor (Integrin alpha M290) (CD103 antigen)	TVQYSYLGYSLAVLHK
175	ITPR2	ITPR2	5-trisphosphate receptor type 2	TPVTGNHGVPTMTLSSMMETCQK
176	JHD3C	JMJD2C	JmjC domain-containing histone demethylation protein 3C	SPMTLVK
177	TIP60	KAT5	Histone acetyltransferase HTATIP	MTGSLVSDR
178	KLDC4	KBTBD4	Kelch domain-containing protein 4	TAAKMEKK
179	Kcna3	KCNA3	Potassium voltage-gated channel subfamily A member 3	SNSTLSK
180	KI21B	KIF21B	Kinesin family member 21B	VVHVPYRDSK
181	KIF2B	KIF2B	Kinesin-like protein KIF2B	TSGQTSVNAHSSR
182	LAMC1	LAMC1	Laminin subunit gamma-1 precursor	FHTSRPESFAIYKR
183	ARH	LDLRAP1	Low density lipoprotein receptor adapter protein 1	IVATAKASGKK
184	LIMA1	LIMA1	LIM domain and actin-binding protein 1	SPKTSSPSLR
185	TBCD4	LOC686547	TBC1 domain family member 4 (Akt substrate of 160 kDa)	SPSAMQQQK
186	Lphn1 (Kiaa0821) (Lec2)	LPHN1	Latrophilin-1	AQSVLYQSDLDESESCTAEDGATSRPL SSPPGR
187	Lrp2	LRP2	Low-density lipoprotein receptor-related protein 2	FRCDNSR
188	LRC59	LRRC59	Leucine-rich repeat-containing protein 59	MTKAGSKGGNLR
189	LRRK1	LRRK1	Leucine-rich repeat serine/threonine-protein kinase 1	GSRSVAK
190	LUC7L	LUC7L	Putative RNA-binding protein Luc7-like 1	FSRSRSR
191	MAP1B	MAP1B	Microtubule-associated protein 1B	ATKTTTTPEVK
192	Map3k1 (Mekk) (Mekk1)	MAP3K1	Mitogen-activated protein kinase kinase kinase 1	CGSHSAELAAAR
193	M4K3	MAP3K3	Mitogen-activated protein kinase kinase kinase kinase 3	GTNLSRKEK
194	Map4k1 (Hpk1)	MAP4K1	Mitogen-activated protein kinase kinase kinase kinase 1	TTWLYCINNLLMSLSGK
195	JIP3	MAPK8IP3	C-jun-amino-terminal kinase-interacting protein 3 (JNK-interacting protein 3)	SPTAAGFSQRR
196	MAST1	MAST1	Microtubule-associated serine/threonine-protein kinase 1	KYGFTLRAIR
197	TRIM1	MID2	Midline-2	TYRPSSAMSSER
198	Mknk1 (Mnk1)	MKNK1	MAP kinase-interacting serine/threonine-protein kinase 1	SRDANPCLTPAGL
199	MPPH1	MPP1	M-phase phosphoprotein 1 (MPP1)	VSELSLCDLAGSER
200	Mtch1	MTCH1	Mitochondrial carrier homolog 1	GLSPRLMSNALSTVTR
201	Muc4	MUC4	Mucin-4	LSTLPQSQHTGSK
202	ERG19	MVD or MPD	Diphosphomevalonate decarboxylase	ILILVVSADK
203	MYH6	MYH6	Myosin-6	LSRIITR
204	NACAM	NACA	Nascent polypeptide-associated complex subunit alpha	ESSATSSSKRAPK
205	PPRB	NCOA6	Peroxisome proliferator-activated receptor-binding protein	SSSGSASSGSVSQK
206	Ndufa4l2	NDUFA4L2	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 4-like 2	MAGTSLGTR
207	Nefh (Kiaa0845) (Nfh)	NEFH	Neurofilament heavy polypeptide	TSVSSVSASPSR
208	NELF	NELF	Nasal embryonic luteinizing hormone-releasing hormone factor	SEAMSSVAAKVR
209	NMBR	NMBR	Neuromedin-B receptor (NMB-R)	MTSLKSNTK
210	NDK6	NME6	Nucleoside diphosphate kinase 6	TLMGPTR
211	Nod1 (Card4)	NOD1	Nucleotide-binding oligomerization domain-containing protein 1	LQSLWASGR
212	Notch1 (Motch)	NOTCH1	Neurogenic locus notch homolog protein 1	CSQPSGTCLNGGRCEVASGTEACVAS GSFVGQR
213	Nr3c2 (Mlr)	NR3C2	Mineralocorticoid receptor	CSVSSPSNTNNR
214	TEN1	ODZ1	Teneurin-1	SPFLSEKR
215	ORC3	ORC3L	Origin recognition complex subunit 3	SMESFTSK
216	OVGP1	OVGP1	Oviduct-specific glycoprotein precursor	TMVLEKATVSPR
217	PALM	PALM	Paralemmin	SPGGSTMMK
218	PCLO	PCLO	Protein piccolo	SPSTISLKESK
219	Pdlim5 (Enh)	PDLIM5	PDZ and LIM domain protein 5	DFNMPLTISSLK
220	PER2	PER2	Period circadian protein homolog 2	TSGKIQTK
221	PHLPP	PHLPP	PH domain leucine-rich repeat-containing protein phosphatase	TSALLQPK
222	PI51B	PIP51KB	Phosphatidylinositol-4-phosphate 5-kinase type-1 beta	MHLTYDLKGSTYK
223	Pkn2 (Prkcl2)	PKN2	Serine/threonine-protein kinase N2	ATSVALPGWSPSDNRSSFMSR
224	Polb	POLB	DNA polymerase beta	VTGIGPSAAR
225	RPO3C	POLR3K	DNA-directed RNA polymerase III subunit C	MTMSAVVK
226	RNAS4	POP4	Ribonuclease 4 precursor	GICSTTNILCK
227	PRGC1	PPARGC1A	Peroxisome proliferator-activated receptor gamma coactivator 1-alpha	SPGSRSSSR
228	Ppp1r12a (Mypt1)	PPP1R12A	Protein phosphatase 1 regulatory subunit 12A	TSSSYTRR
229	AKAP4	PRKA4	A-kinase anchor protein 4 precursor (Protein kinase A-anchoring protein 4) (PRKA4)	VSEHILK
230	NETR	PRSS12	Neurotrypsin precursor	RPPRTPR
231	Psmb8 (Lmp7) (Mc13)	PSMB8	Proteasome subunit beta type-8	ATAGSYISSLRMNK
232	PTN18	PTPN12	Tyrosine-protein phosphatase non-receptor type 18	TSSALPATSRPPGGVLR
233	PTN5	PTPN5	Tyrosine-protein phosphatase non-receptor type 5	MCCSER
234	PURG	PURG	Purine-rich element-binding protein gamma	IRQTMMR
235	Rabep1 (Rab5ep) (Rabpt5) (Rabpt5a)	RABEP1	Rab GTPase-binding effector protein 1	GQLESTLR
236	Rad9a (Rad9)	RAD9	Cell cycle checkpoint control protein RAD9A	CCISLSGSHSHLVVQLHCK
237	Ranbp2	RANBP2	E3 SUMO-protein ligase RanBP2	CVSCQNTNPTSNK
238	RAVR1	RAVER1	Ribonucleoprotein PTB-binding 1 (Protein raver-1)	RLFSHPREPTLGAHGPSR
239	REEP4	REEP4	Receptor expression-enhancing protein 4	AMPSDMDS
240	Reps1	REPS1	RalBP1-associated Eps domain-containing protein 1	ELCGATRLGYFGR
241	RGN	RGN	Regucalcin	TTSCCFGGK
242	RGS9	RGS9	Regulator of G-protein signaling 9 (RGS9)	SPIYKEMLAK
243	RIC8A	RIC8A	Synembryn-A	YTGYGNAAGLLAAR
244	RNH2B	RNASEH2B	Ribonuclease H2 subunit B (RNase H2 subunit B) (Ribonuclease HI subunit B) (Deleted in lymphocytic leukemia 8 homolog)	EDYTKFNTK
245	NC6IP	RNMTL1	Putative RNA methyltransferase NCOA6IP	VRSKVEK

246	RLA0	RPLP0	60S acidic ribosomal protein P0 (L10E)	AEAKEESEESDEDMGFGLFD
247	RLA2	RPLP2	60S acidic ribosomal protein P2	KEESEESDDDMGFGLFD
248	RS17	RPS17	40S ribosomal protein S17	TPRGAV
249	SALL1	SALL1	Sal-like protein 1 (Zinc finger protein Spalt-3)	KSKPPNVTAFEAK
250	SCFD1	SCFD1	Sec1 family domain-containing protein 1	VGSKMAATASIR
251	SDK1	SDK1	Protein sidekick-1 precursor	GSTFEPAR
252	Sec23ip	SEC23IP	SEC23-interacting protein	HTPVSSR
253	SF3B3	SF3B3	Splicing factor 3B subunit 3	TSGPESSVK
254	SFRS7	SFRS7	arginine/serine-rich 7	SRSASLR
255	SGCD	SGCD	Delta-sarcoglycan (SG-delta)	FEVKTVSGK
256	Sgol1 (Sgo1)	SGOL1	Shugoshin-like 1	DQISLCSRLINPAK
257	SH24A	SH24A	SH2 domain-containing protein 4A	AQTKPVKEK
258	Sh2b3 (Lnk)	SH2B3	SH2B adapter protein 3	MNEPTVQPSR
259	SI1L2	SIPA1L2	Signal-induced proliferation-associated 1-like protein 2	SPSTLTGK
260	ACATN	SLC33A1	Acetyl-coenzyme A transporter 1	SPTISHK
261	CTL4	SLC44A4	Choline transporter-like protein 4	NEAPTGGKTRK
262	SMG5	SMG5	Protein SMG5 (SMG-5 homolog) (EST1-like protein B)	SPAQEPPQAR
263	SMG7	SMG7	Protein SMG7 (SMG-7 homolog) (EST1-like protein C)	EVNQGRSFPPK
264	SMN	SMN1	Survival motor neuron protein	KNATTPLK
265	Sik1 (Msk) (Sik) (Snf1lk)	SNF1LK	Serine/threonine-protein kinase SNF1-like kinase 1	QVCQSSVRTPR
266	Sox10 (Sox-10) (Sox21)	SOX10	Transcription factor SOX-10	HPEEGSPMSDGNPEHPSGQSHGPPT PPTTPK
267	MINT	SPEN	Msx2-interacting protein (SPEN homolog) (SMART/HDAC1-associated repressor protein)	TPGTYPEDSR
268	SRRM1	SRRM1	Serine/arginine repetitive matrix protein 1	SPSLSSK
269	SRRM2	SRRM2	Serine/arginine repetitive matrix protein 2	QSCSGSSPR
270	STAR7	STARD7	StAR-related lipid transfer protein 7	SPAGGNEKSK
271	STAR8	STARD8	StAR-related lipid transfer protein 8 (StARD8)	SPDWYNK
272	STK36	STK36	Serine/threonine-protein kinase 36	LALTDSASLK
273	Strn3 (Gs2na) (Sg2na)	STRN3	Striatin-3	NQLLSCSADGTIR
274	ST3A1	SULT1A	Amine sulfotransferase	MEVVENIENYEIR
275	SYTL3	SYTL3	Synaptotagmin-like protein 3 (Exophilin-6)	KASTPDILK
276	TAP2	TAP2	Antigen peptide transporter 2 (APT2)	TMLVIAHR
277	TBX6	TBX6	T-box transcription factor TBX6 (T-box protein 6)	EFSAVGTEMIITK
278	TEX2	TEX2	Testis-expressed sequence 2 protein	STSCLLK
279	Tirap (Mal)	TIRAP	Toll/interleukin-1 receptor domain-containing adapter protein	ALLITPGFLR
280	Tnfrsf1b (Tnfr-2) (Tnfr2)	TNFRSF1B	Tumor necrosis factor receptor superfamily member 1B	THSGSCR
281	TNI3K	TNNI3K	Serine/threonine-protein kinase TNNI3K	NSGSFEDGN
282	Tnpo1 (Kpnb2)	TNPO1	Transportin-1	ESQSPDTTIQR
283	Trim29	TRIM29	Tripartite motif-containing protein 29	TSYQPSSPSR
284	TSC2	TSC2	Tuberin	SSSASSQEEK
285	Ttn	TTN	Titin	MGQKYSFR
286	UBE3C	UBE3C	Ubiquitin-protein ligase E3C	VSAPYITEECLRK
287	UHRF2	UHRF2	E3 ubiquitin-protein ligase UHRF2	TDSSEVVK
288	RENT1	UPF1	Regulator of nonsense transcripts 1(Nonsense mRNA reducing factor 1) (NORF1)	TEAANVEKITTK
289	Usp2 (Ubp41) (MNCb-0190)	USP2	Ubiquitin carboxyl-terminal hydrolase 2	SQLSSTLKR
290	Usp47	USP47	Ubiquitin carboxyl-terminal hydrolase 47	STETSDFENIESPLNERGSSTSVDNR
291	Usp8 (Kiaa0055) (Ubpy)	USP8	Ubiquitin carboxyl-terminal hydrolase 8	FYCSHCR
292	Vamp8	VAMP8	Vesicle-associated membrane protein 8	NIMTQNVERILSR
293	CSPG2	VCAN	Versican core protein precursor	EETVGMGGSDDER
294	CSPG2	VCP	Valosin-containing protein p97/p47 complex-interacting protein 1)	LGSGGVVKK
295	VCIP1	VCPIP1	Deubiquitinating protein VCIP135	LGSGGVVKK
296	Vrk1	VRK1	Serine/threonine-protein kinase VRK1	SVESQGAIHGSMSQPAAGCSSSDSSR R
297	WDHD1	WDHD1	WD repeat and HMG-box DNA-binding protein 1	VRSQVEDAEDR
298	WNT5B	WNT5B	Protein Wnt-5b precursor	YDSAAAMRITR
299	FOG2	ZFPM2	Zinc finger protein ZFPM2	TTASPKR
300	ZN326	ZNF326	Zinc finger protein 326	SPDHSKGK
301	ZN592	ZNF592	Zinc finger protein 592	EGSKGSPKMPK
302	ZN638	ZNF638	Zinc finger protein 638 (Nuclear protein 220)	RLPNLPSHSRNK
303	ZNFX1	ZNFX1	NFX1-type zinc finger-containing protein 1	VKGSIAEEVSSIR

Table S4 gives the complete list of phosphoproteins identified by mass spectrometry along with the peptide sequences. Distinction between molecules differentially identified between G1 and S phase have been highlighted in the table.

Table S5: List of molecules classified as cell cycle targets

	G1 TARGETS	PMIDs
1	ATM	11544175:8843193
2	ATR	11544175:18606783
3	CCNC	1833066:7568034:9552396:8730095
4	CCND1	19167337
5	CCND2	19167337
6	CCND3	19167337
7	CCNE1	19167337:7797073:8156587
8	CCNE2	19167337:7797073:8156587
9	CCNG1	12556559
10	CCNG2	11956189:9139721
11	CCNH	7533895:9130709
12	CDC25A	14681206:10454565
13	CDC2A	17700700
14	CDC34	8248134:11514588
15	CDC37	8703009:9228064:14701845
16	CDC7	10846177
17	CDK2	19167337
18	CDK4	19167337
19	CDK5	8302605:1358458
20	CDK6	19167337
21	CDK7	9832506:7929589
22	CDK8	9584184:17612495
23	CDKN1C	7729684
24	CDKN2B	19167337
25	CDKN2C	19167337
26	CDKN2D	19167337
27	CHEK1	11544175
28	CHEK2	11544175
29	E2F1	10619603:19603422
30	E2F2	19457859
31	E2F3	15467444
32	E2F4	19585502:19562678
33	E2F5	15467444
34	E2F6	17457055:19549334
35	FOS	17216194
36	HIPK2	12851404
37	JUN	17216194
38	JUNB	18793152
39	MDM2	15485814:8628312:10995885:9087426
40	MNAT1	10026172:11113200
41	MYC	12631706
42	p27	8730099:10385618:8033213
43	PLK2	12651910:12972611
44	RB1	11018009:7777526:9694794
45	RBL1	1152119
46	RBL2	1152119:9188854:8253384
47	TFDP1	14618416:18687693
48	TFDP2	9704927
49	Trp53	7742528:9843965:11884608:10072388
50	Trp63	17114587:14737098
51	TRP73	10562283

	S TARGETS	PMIDs
1	ATM	11544175:8843193
2	ATR	11544175:18606783
3	CCNA1	11566717:19167337
4	CCNA2	11566717:19167337
5	CCNC	1833066:7568034:9552396:8730095
6	CCNE1	19167337:7797073:8156587
7	CCNE2	19167337:7797073:8156587
8	CCNH	7533895:9130709
9	CDC25A	10454565:10995786
10	CDC2A	17635936
11	CDC45L	19054765
12	CDC6	16439999:18687693
13	CDC7	10846177
14	CDK2	19167337
15	CDK7	9832506:7929589
16	CDK8	9584184:17612495
17	CHEK1	11544175
18	CHEK2	11544175
19	E2F1	10619603:19603422
20	E2F2	19457859
21	E2F3	15467444
22	E2F4	19585502:19562678
23	E2F5	15467444
24	E2F6	17457055:19549334
25	FOS	17216194
26	GMNN	19487297:11125146
27	JUN	17216194
28	JUNB	18793152
29	MCM2	17680271
30	MCM3	19064704:19377303
31	MCM4	16754955
32	MCM5	16754955
33	MCM6	16754955
34	MCM7	16754955: 19647517
35	MDM2	15485814:8628312:10995885:9087426
36	MNAT1	10026172:11113200
37	MYC	19554081
38	ORC1L	19197067
39	ORC2L	17680271
40	ORC3L	17716973
41	ORC4L	18652488:19647517
42	ORC5L	17716973
43	ORC6L	17716973
44	PCNA	19595719
45	RB1	11018009:7777526:9694794
46	RBL1	1152119
47	RBL2	1152119:9188854:8253384
48	RIS2	18006686:11125146
49	SKP1A	18353424
50	SKP2	19477924
51	TFDP1	14618416:18687693
52	TFDP2	9704927
53	Trp53	7742528:9843965:11884608:10072388

	G2 TARGETS	PMIDs
1	ANAPC2	11285280
2	ATM	11544175
3	ATR	11544175
4	AURKB	19342897
5	CCNB1	11285280
6	CCNB2	9926943
7	CCNB3	7474080
8	CCNG1	16322753 :18408012
9	CCNH	7533895:9130709
10	CDC16	12629511
11	CDC20	11285280
12	CDC25A	19192479
13	CDC25C	18604163
14	CDC2A	19570916:19013283
15	CDK10	8208557
16	CENPA	18007590
17	CENPB	16183641
18	CENPC	12490152
19	CENPE	9763420
20	CENPH	16875666
21	CHEK1	11544175
22	CHEK2	11544175
23	MAD2L1	19461085
24	MAD2L2	10366450
25	p27	17954563: 16310807
26	PLK1	19033445
27	PLK2	19001868
28	PLK3	11971976
29	STK6	19342897

List of molecules classified as cell cycle targets.

Table S5 provides the complete details of classification of molecules according to their functional role in cell cycle and their corresponding PMID for the literature evidence.

Table S6: Occurances of nodes in the shortest paths

G1 Subnetwork		
Molecule	N/O	Score
TRP53	612	10.79
RB1	471	8.29
PTK2B	457	8.04
PRKCA	385	6.76
TRAF6	332	5.82
TICAM2	332	5.82
KIF23	316	5.54
YWHAQ	313	5.48
BCL2	266	4.65
CDK2	258	4.51
IKBKAP	251	4.38
MAPK8	234	4.08
ESR1	195	3.39
POLD1	192	3.34
PTEN	191	3.32
CREBBP	184	3.19
AKT1	178	3.09
EP300	169	2.93
MCM7	160	2.77
SRC	153	2.64
AREG	147	2.54
EPC1	146	2.52
UBB	145	2.50
POLR2A	130	2.24
ARHGEF7	124	2.13
ZBTB17	118	2.02
BRCA1	115	1.97
EEF1D	115	1.97
CASP3	114	1.95
GRB2	103	1.76
SYNGAP1	100	1.70
RAF1	95	1.61
WT1	94	1.60
CSNK2A1	94	1.60
TICAM1	92	1.56
NFKBIA	92	1.56
ABL1	90	1.53
E2F1	89	1.51
DLGAP4	89	1.51
MYC	89	1.51
LCK	88	1.49
XRCC6	86	1.45
TANK	86	1.45
TRIM27	82	1.38
MAPK1	80	1.35
PRC1	76	1.28
RACGAP1	76	1.28
SMAD3	76	1.28
YWHAE	71	1.19
LIMA1	71	1.19
CDKN1B	70	1.17
MRGBP	64	1.06
CCNA2	63	1.05
PCNA	62	1.03
MNAT1	61	1.01

S Subnetwork		
Molecule	N/O	Score
TRP53	372	8.19
ITGB7	367	8.08
NTN1	278	6.10
ADORA2B	278	6.10
FOS	229	5.01
ABL1	225	4.92
YWHAG	204	4.45
PIN1	201	4.38
TRAF6	190	4.14
CSE1L	189	4.12
CDK2	177	3.85
DOK1	154	3.34
DCC	151	3.27
AREG	141	3.05
POLD1	138	2.98
RIS2	138	2.98
STAT1	138	2.98
E7	135	2.92
EZR	128	2.76
PKM2	127	2.74
MCM7	125	2.69
BRCA1	106	2.27
PRKCA	103	2.20
SMAD3	99	2.11
RB1	93	1.98
EED	82	1.74
TAF1	75	1.58
POLR2A	75	1.58
EP300	70	1.47
UBB	70	1.47
RBL1	70	1.47
CCNA2	68	1.42
RAF1	67	1.40
DAB1	67	1.40
ETF1	67	1.40
GRIN2B	66	1.38
COIL	66	1.38
MAPK1	65	1.36
CREBBP	60	1.25
ABL2	56	1.16
CASP3	56	1.16
MCM3	55	1.14
EPC1	55	1.14
JUN	54	1.11
CDK5	51	1.05
YWHAE	51	1.05
EGFR	50	1.02
E2F1	49	1.00

G2 Subnetwork		
Molecule	N/O	Score
LASP1	175	8.93
PRKD1	159	8.10
TRP53	137	6.97
PRKCA	126	6.40
CDK7	74	3.72
FZR1	60	2.99
EGFR	59	2.94
POLD1	55	2.74
CDK2	46	2.27
CHEK2	41	2.01
ETS2	40	1.96
PRKG1	34	1.65
AREG	34	1.65
CASP3	31	1.50
NFE2L2	28	1.34
CREBBP	27	1.29
ACTB	26	1.24
HDAC1	25	1.19
CTNNB1	25	1.19
CREB1	24	1.14
UBTF	23	1.08
SYK	22	1.03

G1S Subnetwork		
Molecule	N/O	Score
CDK2	833	11.38
TRP53	648	8.84
ABL1	465	6.32
POLD1	463	6.30
ANXA1	434	5.90
CREB1	424	5.76
PRKG1	391	5.31
GRB2	388	5.26
EGFR	344	4.66
CREBBP	311	4.21
RIS2	294	3.97
PRKCA	289	3.90
SRC	273	3.68
RB1	268	3.62
UBB	260	3.51
AKT1	238	3.20
EP300	236	3.18
BRCA1	230	3.09
MCM7	230	3.09
ESR1	212	2.85
CTNNB1	191	2.56
POLR2A	181	2.42
STAT1	180	2.41
MAPK3	174	2.32
CCNA2	166	2.21
AREG	165	2.20
LYN	162	2.16
GADD45G	152	2.02
SKP2	149	1.98
CDKN1B	148	1.97
CSNK2A1	137	1.81
ELK1	136	1.80
JUN	136	1.80
CDC25A	136	1.80
TOPBP1	127	1.68
E2F1	126	1.66
TUBA4A	119	1.57
MYC	119	1.57
EPC1	114	1.50
CASP3	113	1.48
WT1	105	1.37
PTPN2	104	1.36
CCND1	103	1.35
ARHGEF7	100	1.31
YWHAE	100	1.31
GADD45A	98	1.28
PCNA	96	1.25
MCM5	94	1.22
ZBTB17	93	1.21
PPP2R4	90	1.17
PITPNM1	88	1.14

N/O **Number of Occurances in the shortest paths**

Table S6 gives the complete list of molecules and the number of times each molecule occurred in the shortest pathnetworks along with the normalized Z scores for each molecule. Only molecules with a score of greater than one are shown in the table above.

Table S7: Network file used as SNAVI background

Source	Target	Effect	Interaction
EP300	CITED1	0	Binding
ESR1	CITED1	0	Binding
MAP2K1	MAPK3	+	Phosphorylation
MAPK3	CASP9	-	Phosphorylation
PTPRE	SRC	+	Dephosphorylation
ADRBK1	SRC	0	Binding
MAPK3	ADRBK1	-	Phosphorylation
PELP1	RB1	0	Binding
MAPK3	DUSP12	0	Phosphorylation
RB1	PRMT2	0	Binding
IER3	MAPK3	+	Binding
MAPK3	IER3	+	Phosphorylation
MAPK3	KRT8	0	Phosphorylation
PRKCA	FOS	0	Phosphorylation
FOS	SMAD3	0	Binding
FOS	NFATC2	0	Binding
JUN	FOS	0	Binding
MAPK1	FOS	+	Phosphorylation
TAF1	FOS	0	Binding
MAPK1	ETF1	0	Phosphorylation
PPP2CA	ETF1	0	Binding
GRIN2B	EPHB2	0	Binding
TAF9	TAF1	0	Binding
TAF1	RBL2	0	Binding
TRP53	COIL	0	Binding
TAF9	COIL	0	Binding
CDK2	PTPN2	0	Phosphorylation
CDK2	COIL	0	Binding
CDK2	E7	0	Binding
MAPK3	JUN	+	Phosphorylation
CSE1L	STAT1	0	Binding
CSE1L	TRP53	0	Binding
DCC	PTK2	0	Binding
EED	HDAC2	0	Binding
ADORA2B	DCC	0	Binding
ADORA2B	EZR	0	Binding
ADORA2B	NTN1	0	Binding
DLGAP4	PLK2	0	Binding
CAMK2D	STAT1	+	Phosphorylation
GRIN2B	CAMK2D	0	Binding
BRCA1	NPM1	0	Binding
PIN1	NFATC2	0	Binding
PIN1	RAF1	0	Dephosphorylation
SLC33A1	STAT1	0	Binding
ABL1	YWHAG	0	Binding
CAMK2G	STAT1	+	Phosphorylation
GRIN2B	CAMK2G	0	Binding
PRKCA	GRIN2B	0	Phosphorylation
GRIN2B	PTPN11	0	Binding
PTPN11	STAT1	0	Dephosphorylation
ESR1	CCNT1	0	Binding
YWHAQ	PDPK1	-	Binding
TRIM27	PML	0	Binding
INSR	SYNGAP1	0	Binding
RGS3	YWHAQ	0	Binding
YWHAQ	PAK6	0	Binding
FRAP1	AKT1	+	Phosphorylation
AKT1	TSC2	0	Phosphorylation
AKT1	BAD	-	Phosphorylation
AKT1	YAP1	0	Phosphorylation
PLK2	IKBKAP	0	Binding
MAP3K14	IKBKAP	0	Binding
IKBKAP	MAPK8	0	Binding
MAP3K7	IKBKAP	+	Phosphorylation
BIRC5	BCL2	0	Binding
PKM2	E7	0	Binding
HDAC2	E7	0	Binding
HDAC1	E7	0	Binding
PPP2CA	E7	0	Binding
MAPK3	RPS6KA3	+	Phosphorylation
PTPRE	MAPK3	-	Dephosphorylation
MAPK3	STMN1	0	Phosphorylation
MAPK3	SMAD2	0	Phosphorylation
RPS6KA1	MAPK3	+	Binding
MAPK3	STAT5A	+	Phosphorylation
MAPK3	UBTF	+	Phosphorylation
MAPK3	SP1	0	Phosphorylation
MAPK3	STAT3	+	Phosphorylation
NTRK1	MAPK3	0	Binding
MAPK3	MAPK14	+	Phosphorylation
CDKN1B	YWHAQ	+	Binding
SMARCA4	RB1	0	Binding
BCL2L1	MAPK8	0	Phosphorylation
NTRK1	SYNGAP1	0	Phosphorylation
SLC6A4	SYNGAP1	0	Binding
AFAP1L2	SYNGAP1	0	Binding
AFAP1L2	SRC	0	Binding
YWHAQ	PNKP	0	Binding
MAP3K14	TANK	0	Phosphorylation
YWHAQ	RAF1	0	Binding
YWHAG	RAF1	0	Binding
MAPK3	ARRB1	+	Phosphorylation
DOK4	SYNGAP1	0	Phosphorylation
DOK4	SRC	+	Phosphorylation
HRAS	SRC	-	Phosphorylation
HRAS	MAPK8	+	Phosphorylation
EP300	PIAS3	0	Binding
YAP1	YWHAQ	-	Binding
MAPK8	IRS1	-	Phosphorylation
FRAP1	YWHAQ	0	Binding
FADD	MYC	0	Binding
TANK	FADD	0	Binding
MAPK3	GTF2I	+	Binding
AKT2	ESR1	+	Phosphorylation
YWHAQ	BAD	-	Binding
KHDRBS1	SYNGAP1	0	Binding
CCNA1	XRCC6	0	Binding
SP3	E2F1	+	Binding
MYC	CHD4	0	Binding
TRIM27	CHD4	0	Binding
YWHAQ	MST1R	0	Binding
YWHAQ	SH3BP2	-	Binding
EGFR	SYNGAP1	0	Binding
EGFR	EZR	0	Binding
SYNGAP1	SHC	0	Binding
ERBB2	GRB2	0	Binding
ESR1	ERBB2	0	Binding
GRB2	RPS6KA1	0	Binding
NTRK1	GRB2	0	Binding
KHDRBS1	GRB2	0	Binding
NGFR	GRB2	-	Binding
GRB2	NPM1	0	Binding
MAPK1	ESR1	0	Binding
HSP90AA1	ESR1	-	Binding
RGS3	ESR1	0	Binding
PAK6	ESR1	+	Binding
ESR1	TSC2	0	Binding
ESR1	KAT5	0	Binding
SP3	ESR1	-	Binding
MAPK3	ESR1	+	Phosphorylation
RPS6KA1	ESR1	+	Phosphorylation
MAPK14	ESR1	+	Phosphorylation
ESR1	SMAD3	0	Phosphorylation
ESR1	FOXO3	+	Binding
PKN1	ESR1	0	Binding
ESR1	SMAD4	0	Binding
ESR1	PIAS3	0	Binding
RPS6KA3	ESR1	0	Binding
PRMT2	ESR1	+	Binding
PELP1	ESR1	0	Binding
SMARCA4	ESR1	0	Binding
TRIP4	ESR1	0	Binding
ESR1	SP1	0	Binding
SRC	DAPP1	0	Binding
MAPK1	LCK	+	Phosphorylation
MAPK3	LCK	0	Phosphorylation
SH3BP2	LCK	+	Binding
BIRC5	SYNGAP1	0	Binding
ARHGAP1	SYNGAP1	0	Binding
PDGFRB	SYNGAP1	0	Binding
CDC25B	ESR1	+	Binding
YWHAQ	CDC25B	0	Binding
MAPK8	DCLK1	0	Phosphorylation
PAX2	MAPK8	+	Phosphorylation
MAPK8	PRG2	+	Phosphorylation
MAPK8	SMAD2	+	Phosphorylation
MAPK8	STAT3	-	Phosphorylation
TANK	CAV1	0	Binding
CAV1	YWHAQ	0	Binding
PNKP	MYC	0	Binding
PML	MYC	+	Binding
MAP3K7	MAPK8	0	Phosphorylation
NCK1	SYNGAP1	0	Binding
MAPK8	CDC2L5	0	Phosphorylation
ANXA1	KRT8	0	Binding
EGR1	EP300	+	Binding
MAP2K1	ELK1	+	Phosphorylation
MAPK9	ELK1	+	Phosphorylation
WT1	FHL2	+	Binding
CDC6	RIS2	+	Binding
CDK9	SKP2	0	Binding
E2	TOPBP1	+	Binding
LYN	MATK	+	Binding
TBP	POLR2A	+	Binding
TP73L	WT1	-	Binding
INPPL1	SRC	+	Phosphorylation
SKP2	ORC1L	0	Binding
CASP2	CSNK2A1	+	Phosphorylation
UBB	RIS2	-	Ubiquitination
E6	EP300	+	Binding
GDNF	JUN	+	Binding
JUN	PRKDC	+	Phosphorylation
JUN	UBB	-	Ubiquitination
JUN	MAF	+	Binding
JUN	NCOA3	+	Binding
BRCA1	EP300	+	Binding
ABL1	JUN	+	Phosphorylation
ABL1	SORBS2	-	Ubiquitination
CBL	ABL1	-	Ubiquitination
ABL1	TRP53	+	Binding
ABL1	KIT	+	Binding
ABL1	PTPN18	-	Dephosphorylation
YWHAB	ABL1	+	Binding
YWHAB	CDKN1B	-	Binding
YWHAE	CDKN1B	-	Binding
UBB	CDKN1B	-	Ubiquitination
UBB	TRP53	-	Ubiquitination
EP300	STAT1	+	Binding
EP300	SP1	+	Binding
EP300	IRF2	-	Binding
E2F1	NCOA6	+	Binding
EP300	CTNNB1	+	acetylation
SKP2	MYC	-	Binding
SMARCB1	MYC	+	Binding
EGFR	ANXA1	+	Binding
EPC1	TRIM27	-	Binding
JUN	ABL1	+	Binding
PLK1	TUBA4A	+	Phosphorylation
MAPK14	ELK1	0	Phosphorylation
ELK1	EP300	+	Phosphorylation
CTBP1	EP300	-	Binding
CTBP1	RB1	0	Binding
CTBP1	ELK1	0	Binding
PRKCD	STAT1	+	Phosphorylation
PRKCD	SRC	+	Phosphorylation
DDB1	RIS2	-	Ubiquitination
BIRC4	UBB	0	Binding
MAPK9	RB1	0	Phosphorylation
CRI1	EP300	0	Binding
CRI1	RB1	0	Binding
APC	TUBA4A	+	Binding
CDC6	MYC	-	Binding
CEBPE	RB1	0	Binding
CEBPE	E2F1	+	Binding
MATK	SRC	0	Binding
RB1	MDM2	0	Binding
CHKA	JUN	0	Binding
CASP9	RB1	0	Cleavage

Source	Target	Effect	Interaction
FBXW7	MYC	0	Binding
FBXW7	JUN	0	Ubiquitination
DGKZ	SRC	0	Binding
CDKN1B	CCNB1	0	Binding
CDK9	RB1	0	Phosphorylation
VAV1	TUBA4A	0	Binding
MNAT1	RB1	0	Phosphorylation
PSMD10	RB1	-	Ubiquitination
TAOK2	TUBA4A	-	Binding
PTPN2	STAT3	-	Dephosphorylation
PTPN2	STAT5B	0	Binding
PTPN2	STAT1	-	Dephosphorylation
PTPN2	SHC	-	Dephosphorylation
PTPN2	STAT5A	0	Binding
PTPN2	SRC	0	Dephosphorylation
JAK1	STAT1	+	Phosphorylation
PTPN2	JAK1	-	Dephosphorylation
LYN	SNCA	0	Binding
LYN	POLD1	0	Binding
LYN	MAP4K1	0	Binding
PTPRC	LYN	-	Dephosphorylation
LYN	RGS16	+	Phosphorylation
LYN	PRKCD	+	Phosphorylation
CDK4	LYN	0	Binding
CDK4	RB1	-	Phosphorylation
SMARCC1	POLR2A	0	Binding
POLR2A	SMARCB1	0	Phosphorylation
SMARCC2	POLR2A	0	Binding
CASP10	RB1	0	Binding
GNB2L1	STAT1	+	Binding
GNB2L1	SRC	-	Binding
COBRA1	JUN	0	Binding
SYK	UBB	0	Binding
INPP5D	LYN	0	Binding
EP300	APEX1	+	acetylation
GADD45G	APEX1	0	Binding
GADD45G	PCNA	+	Binding
BCR	SRC	0	Binding
TRAF6	UBB	0	Binding
CASP2	RB1	0	cleavage
DOK1	KIT	0	Binding
DOK1	ITGB7	0	Binding
DOK1	ITGB2	0	Binding
DOK1	ITGB3	0	Binding
DOK1	ITGB5	0	Binding
HDAC1	RBL1	0	Binding
MCM7	MNAT1	0	Binding
SMC1A	MCM7	0	Binding
MCM7	RBL1	0	Binding
CD19	GRB2	0	Binding
MCM3	MCM7	+	Binding
E6	MCM7	0	Binding
PIAS1	STAT1	0	Binding
BIN1	MYC	+	Binding
JUN	STAT3	+	Binding
JUN	TBP	+	Binding
JUN	TOP1	+	Binding
EED	ITGB7	0	Binding
POLD1	PTPN2	0	Phosphorylation
DLGAP4	SYNGAP1	0	Binding
POLD1	EEF1D	0	Binding
RIPK3	EEF1D	0	Phosphorylation
KIF23	YWHAQ	0	Binding
PTK2	STAT1	+	Binding
BRCA1	RBL1	0	Binding
BRCA1	E7	0	Binding
NBN	BRCA1	0	Binding
BRCA1	SMAD4	0	Binding
RBBP4	BRCA1	0	Binding
BRCA1	ELK1	-	Binding
SMAD3	BRCA1	-	Binding
RBBP7	BRCA1	0	Binding
BRCA1	E6	0	Binding
BRCA1	SMARCC1	0	Binding
BRCA1	FHL2	0	Binding
BRCA1	SMAD2	0	Binding
JAK1	BRCA1	0	Binding
BRCA1	SMARCB1	0	Binding
PRKDC	BRCA1	0	Binding
EGFR	CDC25A	0	Binding
JAK2	BRCA1	0	Binding
BRCA1	CCNA1	0	Binding
SMC1A	BRCA1	0	Binding
BRCA1	COBRA1	0	Binding
CASP3	BRCA1	-	Cleavage
BRCA1	NCOA2	0	Binding
CCNB1	BRCA1	0	Binding
CREBBP	BRCA1	0	Binding
BRCA1	SMARCC2	0	Binding
TAF1	CCNA2	-	reg
CCNA2	RBL1	0	Binding
CHKA	CCNA2	0	Binding
RBL1	RBL2	0	Binding
RBL1	TAF1	0	Binding
WASL	ABL1	+	Binding
ABL1	RAD52	+	Phosphorylation
ABL1	CD19	0	Binding
ABL1	JAK2	0	Phosphorylation
ABL1	MAP4K1	+	Binding
PTPN12	ABL1	0	Dephosphorylation
ABL1	DDB1	-	Phosphorylation
ABL1	YWHAH	0	Binding
ABL1	GRB2	+	Binding
ABL1	TRP73	+	Phosphorylation
ABL1	YTHDC1	+	Phosphorylation
ABL1	CDK5	+	Phosphorylation
ABL1	PRKDC	0	Phosphorylation
ABL1	BCR	+	Binding
ABL1	CRK	0	Phosphorylation
ABL1	CDKN1B	0	Phosphorylation
ABL1	TUB	+	Phosphorylation
ABL1	CASP9	+	Phosphorylation
ABL1	RB1	-	Binding
ABL1	BCAR1	0	Phosphorylation
ABL1	PTPN6	+	Binding
ABL1	PIK3R1	+	Binding
CDK5	PTPN2	+	Phosphorylation
CDK5	SRC	0	Phosphorylation
CDK5	DAB1	0	Binding
CDK5	EZR	0	Binding
DAB1	ITGB7	0	Binding
DAB1	ITGB2	0	Binding
DAB1	ITGB3	0	Binding
DAB1	ITGB5	0	Binding
LCK	EZR	0	Phosphorylation
STAT1	STAT5A	0	Binding
STAT1	STAT5B	+	Binding
CCNT1	MYC	+	Binding
CCNT1	SKP2	+	Binding
INSR	SRC	0	Phosphorylation
PTPN2	INSR	-	Dephosphorylation
TICAM1	TICAM2	+	Binding
TICAM1	TRAF1	0	Binding
AKT1	DYNLL1	0	Phosphorylation
AKT1	CDC2L5	0	Binding
AKT1	PTEN	0	Phosphorylation
AKT1	RAF1	-	Phosphorylation
AKT1	IRS1	0	Binding
AKT1	FOXO1	-	Phosphorylation
AKT1	PRG2	0	Phosphorylation
AKT1	CASP9	0	Phosphorylation
AKT1	BIRC4	+	Phosphorylation
AKT1	MAP3K5	-	Phosphorylation
AKT1	GSK3B	-	Phosphorylation
ARHGEF7	RAF1	+	Binding
ARHGEF7	SHC	0	Binding
ARHGEF7	NCK1	0	Binding
ARHGEF7	MAPK8	0	Binding
ARHGEF7	PARP1	0	Binding
ARHGEF7	PKN2	0	Binding
ARHGEF7	RB1	0	Binding
ARHGEF7	PSMA3	0	Binding
ARHGEF7	PRG2	0	Binding
ARHGEF7	CASP3	0	Binding
BCL2	PPP1CA	0	Binding
BCL2	SPNS1	0	Binding
BCL2	MYC	0	Binding
MAPK14	BCL2	0	Phosphorylation
BCL2L1	BCL2	0	Binding
UBB	CDC25A	-	Ubiquitination
BCL2	PRG2	0	Binding
BCL2	BID	-	Binding
PRKCA	BCL2	-	Phosphorylation
BCL2	RAF1	0	Binding
BCL2	PARP1	0	Binding
CASP3	BCL2	0	cleavage
E7	RBL1	0	Binding
E7	EP300	-	Binding
E7	JUN	+	Binding
CASP3	RB1	-	Cleavage
CASP3	LYN	-	cleavage
XRCC6	SMAD3	0	Binding
CDKN1B	GRB2	-	Binding
CDKN1B	YWHAH	+	Binding
CDKN1B	MCM7	-	Binding
AREG	XRCC6	0	Binding
AREG	BRCA1	0	Binding
AREG	JUN	0	Binding
TRIM27	RB1	-	Binding
RB1	RBP1	+	Binding
RB1	RBL1	+	Binding
RBBP4	RB1	0	Binding
RBBP7	RB1	0	Binding
RB1	TOP2A	0	Binding
RB1	SKP2	-	Binding
CSNK2A1	ARRB2	-	Phosphorylation
ARRB2	NFKBIA	0	Binding
CHEK1	XRCC6	+	Binding
SRC	ITGB3	-	Phosphorylation
ITGB3	PTK2B	0	Binding
LCK	ITK	+	Phosphorylation
PRKCA	SRC	0	Phosphorylation
PARP1	XRCC6	0	Binding
CSNK2A1	BID	0	Phosphorylation
ARRB1	CSNK2A1	0	Binding
ARRB1	NFKBIA	+	Binding
TRP53	XRCC6	0	Binding
RIPK3	TANK	0	Binding
RIPK3	TICAM1	+	Phosphorylation
EP300	ZBTB17	0	Phosphorylation
EP300	ESR1	0	Binding
EP300	SMAD2	0	Binding
EP300	SMAD3	0	Binding
EP300	MSH6	+	Binding
NBN	EP300	0	Binding
EP300	RB1	+	acetylation
PCAF	EP300	0	Binding
EP300	STAT3	+	Binding
TRIP4	EP300	0	Binding
EP300	MYB	+	acetylation
EP300	JUN	-	acetylation
EP300	TRP73	+	Binding
EP300	NCOA2	0	Methylation
FHL2	EP300	+	Binding
EP300	TP73L	0	Binding
EP300	TDG	0	acetylation
EP300	STAT5B	+	Binding
EP300	STAT5A	0	Binding
ZBTB17	GTF2I	+	Binding
GTF2I	STAT1	+	Binding
VAV1	XRCC6	0	Binding
TANK	IKBKG	+	Binding
CCNA1	RBL1	0	Binding
DNMT3A	MYC	0	Binding
DNMT3A	ZBTB17	0	Binding
CREBBP	ZBTB17	0	Binding
CREBBP	MYC	+	acetylation
CREBBP	E2F1	+	Binding
CREBBP	JUN	-	sumo
ELF3	MYC	0	Binding
ELF3	EP300	0	Binding
PKN1	TANK	0	Binding
EGFR	GRB2	+	Binding
PTPN2	EGFR	-	Dephosphorylation
EGFR	TUBA4A	+	Binding
EGFR	LYN	+	Binding
EGFR	STAT1	0	Phosphorylation
PTEN	TRP53	0	Binding
PPP1CA	PTEN	0	Binding
ABL1	RB1	-	Phosphorylation

Source	Target	Effect	Interaction
TERT	XRCC6	0	Binding
IKBKB	NFKBIA	+	Phosphorylation
TANK	IKBKB	+	Binding
NCOA6	TUBA4A	0	Binding
SRC	SHC	+	Phosphorylation
GRB2	ITK	0	Binding
GRB2	TEK	0	Binding
GRB2	SH2B3	0	Binding
GRB2	PTPN22	0	Binding
KIT	GRB2	0	Binding
INPP5D	GRB2	+	Binding
GRB2	JAK1	+	Phosphorylation
WASL	GRB2	+	Binding
GRB2	PSMD10	0	Binding
PTPRC	GRB2	0	Binding
GRB2	TAOK2	0	Binding
SYK	GRB2	+	Phosphorylation
LCK	ESR1	0	Phosphorylation
MNAT1	ESR1	0	Binding
GADD45G	ESR1	0	Binding
MDM2	ESR1	0	Binding
SRC	TRAF1	0	Binding
SRC	STAT3	+	Phosphorylation
SRC	TIAM1	+	Phosphorylation
SRC	STAT5B	+	Phosphorylation
SRC	YTHDC1	-	Phosphorylation
LCK	STAT1	0	Binding
LCK	TEK	0	Phosphorylation
LCK	PTK2	+	Binding
LCK	SH2B3	0	Binding
LCK	NFKBIA	-	Phosphorylation
LCK	SPNS1	0	Phosphorylation
LCK	TUB	0	Binding
LCK	SYNGAP1	0	Binding
PRKCA	LCK	0	Phosphorylation
LCK	PTK2B	+	Binding
LCK	VAV1	+	Phosphorylation
LCK	SHC	0	Binding
LCK	PTPN6	0	Phosphorylation
PTPN22	LCK	-	Dephosphorylation
ARHGAP1	SRC	0	Binding
ARHGAP1	TUBA4A	0	Binding
PTPN2	PDGFRB	-	Dephosphorylation
PDGFRB	STAT1	+	Binding
PTK2B	ASAP1	-	Phosphorylation
SRC	ASAP1	+	Phosphorylation
PTK2B	SNCA	0	Phosphorylation
PTK2B	SORBS2	0	Binding
PTK2B	SYNGAP1	0	Binding
PTK2B	SHC	0	Binding
PTPN12	PTK2B	0	Dephosphorylation
DYNLL1	NFKBIA	0	Binding
NFKBIA	TRP53	0	Binding
NFKBIA	PIK3R1	0	Binding
NFKBIA	TBK1	0	Binding
CSNK2A1	CDC25B	+	Phosphorylation
CSNK2A1	DCLK1	0	Phosphorylation
CSNK2A1	PSMA3	0	Phosphorylation
CSNK2A1	IRF2	0	Phosphorylation
CSNK2A1	MYB	+	Phosphorylation
CSNK2A1	HSPH1	+	Phosphorylation
CSNK2A1	GDNF	0	Phosphorylation
CSNK2A1	TOP2A	+	Phosphorylation
CSNK2A1	EGR1	-	Binding
PAX2	WT1	0	Binding
CRK	PTK2B	0	Binding
XRCC6	MAPK8	-	Phosphorylation
MAPK8	MYC	0	Binding
TBK1	TANK	+	Phosphorylation
MYC	SMAD3	-	Binding
MYC	SMAD2	-	Binding
KAT5	MYC	0	Binding
SP1	CCND1	0	Binding
GSK3B	CCND1	0	Binding
CCND1	CTNNB1	0	Binding
PRKCA	CCND1	0	Phosphorylation
TAF1	CCND1	0	Binding
CDKN1B	CCND1	0	Binding
PCAF	CCND1	0	Binding
PRKDC	ABL1	+	Phosphorylation

Source	Target	Effect	Interaction
CDC25A	MAP3K5	-	Binding
CDC25A	YWHAB	0	Binding
CDC25A	YWHAE	0	Binding
CDC25A	CDK2	+	Dephosphorylation
CTDP1	POLR2A	0	Binding
JUN	EGR1	0	Binding
TAOK2	ELK1	+	Binding
PIAS1	ELK1	-	Sumolation
SRC	RGS16	+	Phosphorylation
APC	CSNK2A1	-	Binding
MCM7	CDC6	0	Binding
SMARCE1	SRC	0	Binding
RB1	DGKZ	+	Binding
ATM	E2F1	0	Phosphorylation
ATM	TOPBP1	+	Phosphorylation
TUBA4A	HSPH1	0	Binding
EP300	E2	+	Binding
SKP2	DUSP12	-	Ubiquitination
UBB	DUSP12	-	Ubiquitination
MYC	MSH6	0	Binding
POLR2A	SMARCE1	0	Binding
POLR2A	CASP10	0	Binding
CHEK2	RB1	+	Phosphorylation
CHEK2	E2F1	+	Phosphorylation
RIS2	ORC1L	0	Binding
KIT	STAT1	+	Phosphorylation
LYN	DOK1	+	Phosphorylation
RB1	HDAC1	0	Binding
SKP2	RIS2	-	Binding
LYN	CDK2	0	Binding
CDK2	RIS2	-	Phosphorylation
TRAF6	MCM7	0	Binding
RAD52	MCM7	0	Binding
MYC	MCM7	0	Binding
PLK1	MCM7	0	Binding
MCM3	STAT1	0	Binding
MYC	E6	0	Binding
TDG	JUN	0	Binding
POLD1	RIS2	0	Binding
CCNA2	BRCA1	0	Binding
CCNA2	E7	0	Binding
CCNA2	TRP53	0	Binding
CCNA2	CDK2	+	Binding
CCNA2	RIS2	-	Binding
CCNA2	POLD1	0	Binding
CCNA2	UBTF	+	Binding
SMAD3	RBL1	0	Binding
BRCA1	ABL1	0	Binding
BIN1	ABL1	0	Binding
ABL1	DOK1	+	Phosphorylation
ABL1	YWHAE	0	Binding
INPPL1	ABL1	0	Binding
POLD1	ABL1	0	Phosphorylation
CDKN1B	CDK5	-	Phosphorylation
SRC	PDPK1	+	Phosphorylation
PTPN1	SRC	+	Dephosphorylation
AKT1	EP300	-	Phosphorylation
AKT1	CHEK1	+	Phosphorylation
AKT1	IKBKB	+	Binding
AKT1	ARHGEF7	0	Binding
SRC	AKT1	+	Phosphorylation
AKT1	AREG	-	Phosphorylation
AKT1	SORBS2	0	Phosphorylation
AKT1	SLC6A4	0	Binding
AKT1	GAB2	0	Binding
AKT1	PTPN1	0	Phosphorylation
PDPK1	AKT1	+	Phosphorylation
AKT1	HSP90AA1	0	Binding
AKT1	TOPBP1	+	Phosphorylation
AKT1	TERT	+	Phosphorylation
AKT1	SH2B2	0	Phosphorylation
AKT1	ESR1	0	Phosphorylation
ARHGEF7	CCNA2	0	Binding
DAB2	SRC	-	Binding
DAB2	GRB2	0	Binding
RB1	UBTF	-	Binding
HSP90AA1	SRC	0	Binding
UBB	CHEK1	-	Ubiquitination
ADRB2	SRC	0	Binding
POLD1	CDC25A	0	Binding

Source	Target	Effect	Interaction
ADRB2	GRB2	0	Binding
CBL	SRC	-	Ubiquitinationquitination
CBL	GRB2	-	Ubiquitinationquitination
GRB2	GAB2	+	Binding
PTPN18	SRC	+	Dephosphorylation
EP300	PCNA	+	Binding
EP300	MAF	0	Binding
MAP2K1	EP300	+	Phosphorylation
PRKCD	EP300	-	Phosphorylation
SRC	GTF2I	+	Phosphorylation
IKBKG	SRC	0	Binding
UBB	IKBKG	+	Ubiquitination
CCNA1	E2F1	0	Binding
CCNA1	RIS2	-	Phosphorylation
CREBBP	POLR2A	0	Binding
ELK1	CREBBP	+	Binding
TOPBP1	E2F1	-	Binding
E7	E2F1	+	Binding
RB1	E2F1	-	Binding
E2F1	TRP53	+	Binding
SKP2	E2F1	-	Binding
E2F1	POLD1	0	Binding
RBP1	E2F1	-	Binding
NCOA3	E2F1	+	Binding
BCAR1	SRC	0	Binding
BCAR1	GRB2	0	Binding
MST1R	SRC	0	Binding
PTPN6	SRC	+	Dephosphorylation
PDE6G	SRC	0	Binding
DAG1	GRB2	0	Binding
SRC	DAG1	+	Binding
SRC	IKBKB	+	Phosphorylation
SRC	NCOA6	0	Binding
CSNK2A1	ARR3	0	Phosphorylation
ARR3	SRC	0	Binding
GRB2	PTK2	0	Binding
GRB2	PTK2B	0	Binding
GRB2	HNRNPC	0	Binding
GRB2	MST1R	0	Binding
GRB2	PDE6G	0	Binding
GRB2	SHB	0	Binding
GRB2	VAV1	+	Binding
GRB2	PIK3R1	0	Binding
GRB2	SLC6A4	0	Binding
GRB2	PKN2	0	Binding
GRB2	PTPN12	0	Binding
GRB2	PDGFRB	0	Binding
GRB2	MAPK14	0	Binding
GRB2	SH2B2	0	Binding
GRB2	SRC	0	Binding
RB1	ESR1	0	Binding
ESR1	SRC	+	Binding
ESR1	GRB2	0	Phosphorylation
JUN	ESR1	0	Binding
SRC	CTNNB1	+	Binding
SRC	TUB	0	Binding
SRC	TERT	0	Phosphorylation
LCK	DAPP1	0	Phosphorylation
PDGFRB	SRC	0	Binding
PTK2B	SRC	+	Binding
SRC	NFKBIA	0	Phosphorylation
CSNK2A1	PTPN1	0	Phosphorylation
CSNK2A1	CTNNB1	+	Phosphorylation
CSNK2A1	POLD1	0	Phosphorylation
CSNK2A1	HSP90AA1	+	Phosphorylation
CSNK2A1	HNRNPC	0	Phosphorylation
CSNK2A1	NFKBIA	0	Phosphorylation
CSNK2A1	SNCA	0	Phosphorylation
CSNK2A1	PTPN18	-	Phosphorylation
CSNK2A1	MYC	0	Phosphorylation
CSNK2A1	TRP53	-	Phosphorylation
CSNK2A1	CTDP1	+	Phosphorylation
CSNK2A1	UBTF	+	Phosphorylation
TOP1	CSNK2A1	0	Binding
CSNK2A1	HDAC1	+	Phosphorylation
SRC	PKN2	+	Phosphorylation
CRK	GRB2	0	Binding
MYC	XRCC6	0	Binding
TUBA4A	MYC	0	Binding
SHB	SRC	0	Binding

Table S8: Classification of nodes according to their presence in different phase specific modules.

Functional groups	Molecules unique to G1	Common molecules between G1 and G1S	Molecules Unique to G1S		Common molecules between G1S and S	Molecules unique to S
Signaling intermediates	YWHAQ BAD HRAS CRK LCK PTK2B MAPK8 PTPN12 BCAR1 SYNGAP1 ASAP1 PRG2 ARHGEF7	ADRB2 AKT1 CDC25B PDGFRB GRB2 CSNK2A1 ERBB2 MST1R PDE6G PDPK1 PTPN18 SHB SHC SRC CTNNB1	RGS16 CDC25A CHEK2 WASL GNB2L1 PRKCD MNAT1 MAPK14 CDKN1B YWHAB PELP1 PRKDC MAPK3 MAP2K1 CBL MDM2	ABL1 CDK5 EGFR JAK1 LYN PTPN2 PTPN6 PTPRE RPS6KA1 CASP3 CASP9 PIAS3 DUSP12	CDK2	
Transcription Regulators		CREBBP E2F1 ESR1 FOXO1 FOXO3 IKBKB MYC STAT3 TRP53 TSC2	EGR1 ELK1 EP300 GTF2I JUN MYB STAT1 STAT5B NCOA3	TRIP4 TRP73 UBTF CITED1 E2 SMARCA4 SMAD3 UBB E6	BRCA1 E7	TAF1
DNA replication factors		SKP2 SP3	YTHDC1 MCM7 TOPBP1	CDC6	RIS2	
Other Cell Cycle Regulators		HSP90AA1	CCNT1 YTHDC1	RB1 RBP1	CCNA2 RBL1	RBL2 HDAC1

Chapter 9

Bibliography

Chapter 9: Bibliography

Albert, R. (2005). "Scale-free networks in cell biology." J Cell Sci 118(Pt 21): 4947-57.

Albert, R., H. Jeong and A. L. Barabasi (2000). "Error and attack tolerance of complex networks." Nature 406(6794): 378-82.

Albinger-Hegyi, A., B. Hochreutener, M. T. Abdou, I. Hegyi, M. T. Dours-Zimmermann, M. O. Kurrer, P. U. Heitz and D. R. Zimmermann (2002). "High frequency of t(14;18)-translocation breakpoints outside of major breakpoint and minor cluster regions in follicular lymphomas: improved polymerase chain reaction protocols for their detection." Am J Pathol 160(3): 823-32.

Alon, U. (2007). "Network motifs: theory and experimental approaches." Nat Rev Genet 8(6): 450-61.

Alonso, A., J. Sasin, N. Bottini, I. Friedberg, I. Friedberg, A. Osterman, A. Godzik, T. Hunter, J. Dixon and T. Mustelin (2004). "Protein tyrosine phosphatases in the human genome." Cell 117(6): 699-711.

Arias, E. E. and J. C. Walter (2007). "Strength in numbers: preventing rereplication via multiple mechanisms in eukaryotic cells." Genes Dev 21(5): 497-518.

Astsaturov, I., V. Ratushny, A. Sukhanova, M. B. Einarson, T. Bagnyukova, Y. Zhou, K. Devarajan, J. S. Silverman, N. Tikhmyanova, N. Skobeleva, A. Pecherskaya, R. E. Nasto, C. Sharma, S. A. Jablonski, I. G. Serebriiskii, L. M. Weiner and E. A. Golemis (2010). "Synthetic lethal screen of an EGFR-centered network to improve targeted therapies." Sci Signal 3(140): ra67.

Bader, G. D., I. Donaldson, C. Wolting, B. F. Ouellette, T. Pawson and C. W. Hogue (2001). "BIND--The Biomolecular Interaction Network Database." Nucleic Acids Res 29(1): 242-5.

Bakal, C., R. Linding, F. Llense, E. Heffern, E. Martin-Blanco, T. Pawson and N. Perrimon (2008). "Phosphorylation networks regulating JNK activity in diverse genetic backgrounds." Science 322(5900): 453-6.

Bar-Yam, Y. and I. R. Epstein (2004). "Response of complex networks to stimuli." Proc Natl Acad Sci U S A 101(13): 4341-5.

Barabasi, A. L., N. Gulbahce and J. Loscalzo (2011). "Network medicine: a network-based approach to human disease." Nat Rev Genet 12(1): 56-68.

Barabasi, A. L. and Z. N. Oltvai (2004). "Network biology: understanding the cell's functional organization." Nat Rev Genet 5(2): 101-13.

Barila, D., R. Mangano, S. Gonfloni, J. Kretzschmar, M. Moro, D. Bohmann and G. Superti-Furga (2000). "A nuclear tyrosine phosphorylation circuit: c-Jun as an activator and substrate of c-Abl and JNK." Embo J 19(2): 273-81.

Bastard, C., C. Deweindt, J. P. Kerckaert, B. Lenormand, A. Rossi, F. Pezzella, C. Fruchart, C. Duval, M. Monconduit and H. Tilly (1994). "LAZ3 rearrangements in non-Hodgkin's lymphoma: correlation with histology, immunophenotype, karyotype, and clinical outcome in 217 patients." Blood 83(9): 2423-7.

Birmingham, A., L. M. Selfors, T. Forster, D. Wrobel, C. J. Kennedy, E. Shanks, J. Santoyo-Lopez, D. J. Dunican, A. Long, D. Kelleher, Q. Smith, R. L. Beijersbergen, P. Ghazal and C. E. Shamu (2009). "Statistical methods for analysis of high-throughput RNA interference screens." Nat Methods 6(8): 569-75.

Bloom, F. E. (2001). "What does it all mean to you?" J Neurosci 21(21): 8304-5.

Boguna, M. and D. Krioukov (2009). "Navigating ultrasmall worlds in ultrashort time." Phys Rev Lett 102(5): 058701.

Bollen, M., D. W. Gerlich and B. Lesage (2009). "Mitotic phosphatases: from entry guards to exit guides." Trends Cell Biol 19(10): 531-41.

Boureux, A., O. Furstoss, V. Simon and S. Roche (2005). "Abl tyrosine kinase regulates a Rac/JNK and a Rac/Nox pathway for DNA synthesis and Myc expression induced by growth factors." J Cell Sci 118(Pt 16): 3717-26.

Boutros, M., L. P. Bras and W. Huber (2006). "Analysis of cell-based RNAi screens." Genome Biol 7(7): R66.

Bradham, C. and D. R. McClay (2006). "p38 MAPK in development and cancer." Cell Cycle 5(8): 824-8.

Breitkreutz, B. J., C. Stark, T. Reguly, L. Boucher, A. Breitkreutz, M. Livstone, R. Oughtred, D. H. Lackner, J. Bahler, V. Wood, K. Dolinski and M. Tyers (2008). "The BioGRID Interaction Database: 2008 update." Nucleic Acids Res 36(Database issue): D637-40.

Bushman, F. D., N. Malani, J. Fernandes, I. D'Orso, G. Cagney, T. L. Diamond, H. Zhou, D. J. Hazuda, A. S. Espeseth, R. Konig, S. Bandyopadhyay, T. Ideker, S. P. Goff, N. J. Krogan, A. D. Frankel, J. A. Young and S. K. Chanda (2009). "Host cell factors in HIV replication: meta-analysis of genome-wide studies." PLoS Pathog 5(5): e1000437.

Callaway, D. S., M. E. Newman, S. H. Strogatz and D. J. Watts (2000). "Network robustness and fragility: percolation on random graphs." Phys Rev Lett 85(25): 5468-71.

Chatr-aryamontri, A., A. Ceol, L. M. Palazzi, G. Nardelli, M. V. Schneider, L. Castagnoli and G. Cesareni (2007). "MINT: the Molecular INTeraction database." Nucleic Acids Res 35(Database issue): D572-4.

Chen, L., R. S. Wang and X. S. Zhang (2009). "Biomolecular Networks: Methods and Applications in Systems Biology." Wiley 1st edition.

Chou, T. C. and P. Talalay (1981). "Generalized equations for the analysis of inhibitions of Michaelis-Menten and higher-order kinetic systems with two or more mutually exclusive and nonexclusive inhibitors." Eur J Biochem 115(1): 207-16.

Cohen, P. (2000). "The regulation of protein function by multisite phosphorylation--a 25 year update." Trends Biochem Sci 25(12): 596-601.

Cohn, D. L., B. J. Catlin, K. L. Peterson, F. N. Judson and J. A. Sbarbaro (1990). "A 62-dose, 6-month therapy for pulmonary and extrapulmonary tuberculosis. A twice-weekly, directly observed, and cost-effective regimen." Ann Intern Med 112(6): 407-15.

Csermely, P., V. Agoston and S. Pongor (2005). "The efficiency of multi-target drugs: the network approach might help drug design." Trends Pharmacol Sci 26(4): 178-82.

Csete, M. and J. Doyle (2004). "Bow ties, metabolism and disease." Trends Biotechnol 22(9): 446-50.

Csikasz-Nagy, A., O. Kapuy, A. Toth, C. Pal, L. J. Jensen, F. Uhlmann, J. J. Tyson and B. Novak (2009). "Cell cycle regulation by feed-forward loops coupling transcription and phosphorylation." Mol Syst Biol 5: 236.

del Sol, A., R. Balling, L. Hood and D. Galas (2010). "Diseases as network perturbations." Curr Opin Biotechnol 21(4): 566-71.

del Sol, A., H. Fujihashi, D. Amoros and R. Nussinov (2006). "Residues crucial for maintaining short paths in network communication mediate signaling in proteins." Mol Syst Biol 2: 2006 0019.

Dhavan, R. and L. H. Tsai (2001). "A decade of CDK5." Nat Rev Mol Cell Biol 2(10): 749-59.

di Bari, M. G., L. Ciuffini, M. Mingardi, R. Testi, S. Soddu and D. Barila (2006). "c-Abl acetylation by histone acetyltransferases regulates its nuclear-cytoplasmic localization." EMBO Rep 7(7): 727-33.

Donjerkovic, D. and D. W. Scott (2000). "Regulation of the G1 phase of the mammalian cell cycle." Cell Res 10(1): 1-16.

Dowdy, S. F., P. W. Hinds, K. Louie, S. I. Reed, A. Arnold and R. A. Weinberg (1993). "Physical interaction of the retinoblastoma protein with human D cyclins." Cell 73(3): 499-511.

Downward, J. (2003). "Targeting RAS signalling pathways in cancer therapy." Nat Rev Cancer 3(1): 11-22.

Echeverri, C. J., P. A. Beachy, B. Baum, M. Boutros, F. Buchholz, S. K. Chanda, J. Downward, J. Ellenberg, A. G. Fraser, N. Hacohen, W. C. Hahn, A. L. Jackson, A. Kiger, P. S. Linsley, L. Lum, Y. Ma, B. Mathey-Prevot, D. E. Root, D. M. Sabatini, J. Taipale, N. Perrimon and R. Bernards (2006). "Minimizing the risk of reporting false positives in large-scale RNAi screens." Nat Methods 3(10): 777-9.

Edelman, G. M. and J. A. Gally (2001). "Degeneracy and complexity in biological systems." Proc Natl Acad Sci U S A 98(24): 13763-8.

Elbauomy Elsheikh, S., A. R. Green, M. B. Lambros, N. C. Turner, M. J. Grainge, D. Powe, I. O. Ellis and J. S. Reis-Filho (2007). "FGFR1 amplification in breast

carcinomas: a chromogenic in situ hybridisation analysis." Breast Cancer Res 9(2): R23.

Evan, G. I. and K. H. Vousden (2001). "Proliferation, cell cycle and apoptosis in cancer." Nature 411(6835): 342-8.

Ewen, M. E., H. K. Sluss, C. J. Sherr, H. Matsushime, J. Kato and D. M. Livingston (1993). "Functional interactions of the retinoblastoma protein with mammalian D-type cyclins." Cell 73(3): 487-97.

Ezhevsky, S. A., A. Ho, M. Becker-Hapak, P. K. Davis and S. F. Dowdy (2001). "Differential regulation of retinoblastoma tumor suppressor protein by G(1) cyclin-dependent kinase complexes in vivo." Mol Cell Biol 21(14): 4773-84.

Farkas, I. J., T. Korcsmaros, I. A. Kovacs, A. Mihalik, R. Palotai, G. I. Simko, K. Z. Szalay, M. Szalay-Beko, T. Vellai, S. Wang and P. Csermely (2011). "Network-based tools for the identification of novel drug targets." Sci Signal 4(173): pt3.

Fitzgerald, J. B., B. Schoeberl, U. B. Nielsen and P. K. Sorger (2006). "Systems biology and combination therapy in the quest for clinical efficacy." Nat Chem Biol 2(9): 458-66.

Furstoss, O., K. Dorey, V. Simon, D. Barila, G. Superti-Furga and S. Roche (2002). "c-Abl is an effector of Src for growth factor-induced c-myc expression and DNA synthesis." Embo J 21(4): 514-24.

Goff, S. P. (2008). "Knockdown screens to knockout HIV-1." Cell 135(3): 417-20.

Hanahan, D. and R. A. Weinberg (2000). "Hallmarks of cancer: the next generation." Cell 144(5): 646-74.

Harper, J. W. and S. J. Elledge (2007). "The DNA damage response: ten years after." Mol Cell 28(5): 739-45.

Hartman, J. L. t., B. Garvik and L. Hartwell (2001). "Principles for the buffering of genetic variation." Science 291(5506): 1001-4.

Hartwell, L. H., J. J. Hopfield, S. Leibler and A. W. Murray (1999). "From molecular to modular cell biology." Nature 402(6761 Suppl): C47-52.

He, X. and J. Zhang (2006). "Why do hubs tend to be essential in protein networks?" PLoS Genet 2(6): e88.

Heikkila, R., G. Schwab, E. Wickstrom, S. L. Loke, D. H. Pluznik, R. Watt and L. M. Neckers (1987). "A c-myc antisense oligodeoxynucleotide inhibits entry into S phase but not progress from G0 to G1." Nature 328(6129): 445-9.

Hennessy, B. T., D. L. Smith, P. T. Ram, Y. Lu and G. B. Mills (2005). "Exploiting the PI3K/AKT pathway for cancer drug discovery." Nat Rev Drug Discov 4(12): 988-1004.

Huber, W., V. J. Carey, L. Long, S. Falcon and R. Gentleman (2007). "Graphs in molecular biology." BMC Bioinformatics 8 Suppl 6: S8.

Hunter, T. (1997). "Oncoprotein networks." Cell 88(3): 333-46.

Iorns, E., C. J. Lord, N. Turner and A. Ashworth (2007). "Utilizing RNA interference to enhance cancer drug discovery." Nat Rev Drug Discov 6(7): 556-68.

Jackson, A. L., S. R. Bartz, J. Schelter, S. V. Kobayashi, J. Burchard, M. Mao, B. Li, G. Cavet and P. S. Linsley (2003). "Expression profiling reveals off-target gene regulation by RNAi." Nat Biotechnol 21(6): 635-7.

Jamal, M. S., S. Ravichandran, N. Jailkhani, S. Chatterjee, R. Dua and K. V. Rao (2010). "Defining the antigen receptor-dependent regulatory network that induces arrest of cycling immature B-lymphocytes." BMC Syst Biol 4: 169.

Jeong, H., S. P. Mason, A. L. Barabasi and Z. N. Oltvai (2001). "Lethality and centrality in protein networks." Nature 411(6833): 41-2.

Johnson, D. G. and C. L. Walker (1999). "Cyclins and cell cycle checkpoints." Annu Rev Pharmacol Toxicol 39: 295-312.

Jones, S. M. and A. Kazlauskas (2000). "Connecting signaling and cell cycle progression in growth factor-stimulated cells." Oncogene 19(49): 5558-67.

Jones, S. M. and A. Kazlauskas (2001). "Growth factor-dependent signaling and cell cycle progression." FEBS Lett 490(3): 110-6.

Joy, M. P., A. Brock, D. E. Ingber and S. Huang (2005). "High-betweenness proteins in the yeast protein interaction network." J Biomed Biotechnol 2005(2): 96-103.

Kaelin, W. G., Jr. (2005). "The concept of synthetic lethality in the context of anticancer therapy." Nat Rev Cancer 5(9): 689-98.

Kaelin, W. G., Jr. (2009). "Synthetic lethality: a framework for the development of wiser cancer therapeutics." Genome Med 1(10): 99.

Kastan, M. B. and J. Bartek (2004). "Cell-cycle checkpoints and cancer." Nature 432(7015): 316-23.

Keith, C. T., A. A. Borisy and B. R. Stockwell (2005). "Multicomponent therapeutics for networked systems." Nat Rev Drug Discov 4(1): 71-8.

Kerrien, S., Y. Alam-Faruque, B. Aranda, I. Bancarz, A. Bridge, C. Derow, E. Dimmer, M. Feuermann, A. Friedrichsen, R. Huntley, C. Kohler, J. Khadake, C. Leroy, A. Liban, C. Lieftink, L. Montecchi-Palazzi, S. Orchard, J. Risse, K. Robbe, B. Roechert, D. Thorneycroft, Y. Zhang, R. Apweiler and H. Hermjakob (2007). "IntAct--open source resource for molecular interaction data." Nucleic Acids Res 35(Database issue): D561-5.

Keshava Prasad, T. S., R. Goel, K. Kandasamy, S. Keerthikumar, S. Kumar, S. Mathivanan, D. Telikicherla, R. Raju, B. Shafreen, A. Venugopal, L. Balakrishnan, A. Marimuthu, S. Banerjee, D. S. Somanathan, A. Sebastian, S. Rani, S. Ray, C. J. Harrys Kishore, S. Kanth, M. Ahmed, M. K. Kashyap, R. Mohmood, Y. L. Ramachandra, V. Krishna, B. A. Rahiman, S. Mohan, P. Ranganathan, S. Ramabadran, R. Chaerkady and A. Pandey (2009). "Human Protein Reference Database--2009 update." Nucleic Acids Res 37(Database issue): D767-72.

Kinzler, K. W. and B. Vogelstein (1996). "Lessons from hereditary colorectal cancer." Cell 87(2): 159-70.

Kitano, H. and K. Oda (2006). "Robustness trade-offs and host-microbial symbiosis in the immune system." Mol Syst Biol 2: 2006 0022.

Kittler, R., L. Pelletier, A. K. Heninger, M. Slabicki, M. Theis, L. Miroslaw, I. Poser, S. Lawo, H. Grabner, K. Kozak, J. Wagner, V. Surendranath, C. Richter, W. Bowen, A. L. Jackson, B. Habermann, A. A. Hyman and F. Buchholz (2007). "Genome-scale RNAi profiling of cell division in human tissue culture cells." Nat Cell Biol 9(12): 1401-12.

Korcsmáros, T., M. S. Szala, C. Böde, I. A. Kovác and P. Csermely (2007). "How to design multi-target drugs: Target search options in cellular networks." Expert Opinion on Drug Discovery 2: 1-10.

Kulkarni, M. M., M. Booker, S. J. Silver, A. Friedman, P. Hong, N. Perrimon and B. Mathey-Prevot (2006). "Evidence of off-target effects associated with long dsRNAs in Drosophila melanogaster cell-based assays." Nat Methods 3(10): 833-8.

Kumagai, A., J. Lee, H. Y. Yoo and W. G. Dunphy (2006). "TopBP1 activates the ATR-ATRIP complex." Cell 124(5): 943-55.

Lee, G., L. A. Santat, M. S. Chang and S. Choi (2009). "RNAi methodologies for the functional study of signaling molecules." PLoS One 4(2): e4559.

Lehar, J., A. Krueger, G. Zimmermann and A. Borisy (2008). "High-order combination effects and biological robustness." Mol Syst Biol 4: 215.

Lehar, J., B. R. Stockwell, G. Giaever and C. Nislow (2008). "Combination chemical genetics." Nat Chem Biol 4(11): 674-81.

Lemmon, M. A. and J. Schlessinger (2010). "Cell signaling by receptor tyrosine kinases." Cell 141(7): 1117-34.

Litvak, V., D. Tian, Y. D. Shaul and S. Lev (2000). "Targeting of PYK2 to focal adhesions as a cellular mechanism for convergence between integrins and G protein-coupled receptor signaling cascades." J Biol Chem 275(42): 32736-46.

Lord, C. J., S. A. Martin and A. Ashworth (2009). "RNA interference screening demystified." J Clin Pathol 62(3): 195-200.

Luttrell, D. K. and L. M. Luttrell (2004). "Not so strange bedfellows: G-protein-coupled receptors and Src family kinases." Oncogene 23(48): 7969-78.

Ma'ayan, A., S. L. Jenkins, R. L. Webb, S. I. Berger, S. P. Purushothaman, N. S. Abul-Husn, J. M. Posner, T. Flores and R. Iyengar (2009). "SNAVI: Desktop application for analysis and visualization of large-scale signaling networks." BMC Syst Biol 3: 10.

Ma, Y., A. Creanga, L. Lum and P. A. Beachy (2006). "Prevalence of off-target effects in Drosophila RNA interference screens." Nature 443(7109): 359-63.

Major, M. B., B. S. Roberts, J. D. Berndt, S. Marine, J. Anastas, N. Chung, M. Ferrer, X. Yi, C. L. Stoick-Cooper, P. D. von Haller, L. Kategaya, A. Chien, S. Angers, M. MacCoss, M. A. Cleary, W. T. Arthur and R. T. Moon (2008). "New regulators of

Wnt/beta-catenin signaling revealed by integrative molecular screening." Sci Signal 1(45): ra12.

Malumbres, M. and M. Barbacid (2001). "To cycle or not to cycle: a critical decision in cancer." Nat Rev Cancer 1(3): 222-31.

Malumbres, M. and M. Barbacid (2007). "Cell cycle kinases in cancer." Curr Opin Genet Dev 17(1): 60-5.

Malumbres, M. and M. Barbacid (2009). "Cell cycle, CDKs and cancer: a changing paradigm." Nat Rev Cancer 9(3): 153-66.

Manimaran, P., S. R. Hegde and S. C. Mande (2009). "Prediction of conditional gene essentiality through graph theoretical analysis of genome-wide functional linkages." Mol Biosyst 5(12): 1936-42.

Manning, G., D. B. Whyte, R. Martinez, T. Hunter and S. Sudarsanam (2002). "The protein kinase complement of the human genome." Science 298(5600): 1912-34.

Martinez-Antonio, A. and J. Collado-Vides (2003). "Identifying global regulators in transcriptional regulatory networks in bacteria." Curr Opin Microbiol 6(5): 482-9.

Mason, O. and M. Verwoerd (2007). "Graph theory and networks in Biology." IET Syst Biol 1(2): 89-119.

Matsushime, H., D. E. Quelle, S. A. Shurtleff, M. Shibuya, C. J. Sherr and J. Y. Kato (1994). "D-type cyclin-dependent kinase activity in mammalian cells." Mol Cell Biol 14(3): 2066-76.

Matsushime, H., M. F. Roussel, R. A. Ashmun and C. J. Sherr (1991). "Colony-stimulating factor 1 regulates novel cyclins during the G1 phase of the cell cycle." Cell 65(4): 701-13.

Mohr, S., C. Bakal and N. Perrimon (2010). "Genomic screening with RNAi: results and challenges." Annu Rev Biochem 79: 37-64.

Mukherji, M., R. Bell, L. Supekova, Y. Wang, A. P. Orth, S. Batalov, L. Miraglia, D. Huesken, J. Lange, C. Martin, S. Sahasrabudhe, M. Reinhardt, F. Natt, J. Hall, C. Mickanin, M. Labow, S. K. Chanda, C. Y. Cho and P. G. Schultz (2006). "Genome-wide functional analysis of human cell-cycle regulators." Proc Natl Acad Sci U S A 103(40): 14819-24.

Murray, A. W. (2004). "Recycling the cell cycle: cyclins revisited." Cell 116(2): 221-34.

Obaya, A. J., M. K. Mateyak and J. M. Sedivy (1999). "Mysterious liaisons: the relationship between c-Myc and the cell cycle." Oncogene 18(19): 2934-41.

Offit, K., F. Lo Coco, D. C. Louie, N. Z. Parsa, D. Leung, C. Portlock, B. H. Ye, F. Lista, D. A. Filippa, A. Rosenbaum and et al. (1994). "Rearrangement of the bcl-6 gene as a prognostic marker in diffuse large-cell lymphoma." N Engl J Med 331(2): 74-80.

Olsen, J. V., M. Vermeulen, A. Santamaria, C. Kumar, M. L. Miller, L. J. Jensen, F. Gnad, J. Cox, T. S. Jensen, E. A. Nigg, S. Brunak and M. Mann (2010). "Quantitative

phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis." Sci Signal 3(104): ra3.

Papin, J. A., T. Hunter, B. O. Palsson and S. Subramaniam (2005). "Reconstruction of cellular signalling networks and analysis of their properties." Nat Rev Mol Cell Biol 6(2): 99-111.

Pawson, T. and R. Linding (2008). "Network medicine." FEBS Lett 582(8): 1266-70.

Pereira, J. P., J. An, Y. Xu, Y. Huang and J. G. Cyster (2009). "Cannabinoid receptor 2 mediates the retention of immature B cells in bone marrow sinusoids." Nat Immunol 10(4): 403-11.

Plattner, R., L. Kadlec, K. A. DeMali, A. Kazlauskas and A. M. Pendergast (1999). "c-Abl is activated by growth factors and Src family kinases and has a role in the cellular response to PDGF." Genes Dev 13(18): 2400-11.

Puxeddu, E., S. Romagnoli and M. E. Dottorini (2011). "Targeted therapies for advanced thyroid cancer." Curr Opin Oncol 23(1): 13-21.

Reed, S. I. (1996). "Cyclin E: in mid-cycle." Biochim Biophys Acta 1287(2-3): 151-3.

Reya, T. and H. Clevers (2005). "Wnt signalling in stem cells and cancer." Nature 434(7035): 843-50.

Roses, A. D. (2000). "Pharmacogenetics and the practice of medicine." Nature 405(6788): 857-65.

Roussel, M. F. (1997). "Regulation of cell cycle entry and G1 progression by CSF-1." Mol Reprod Dev 46(1): 11-8.

Sams-Dodd, F. (2005). "Target-based drug discovery: is something wrong?" Drug Discov Today 10(2): 139-47.

Sanchez-Beato, M., A. Sanchez-Aguilera and M. A. Piris (2003). "Cell cycle deregulation in B-cell lymphomas." Blood 101(4): 1220-35.

Santoni-Rugiu, E., J. Falck, N. Mailand, J. Bartek and J. Lukas (2000). "Involvement of Myc activity in a G(1)/S-promoting mechanism parallel to the pRb/E2F pathway." Mol Cell Biol 20(10): 3497-509.

Sathornsumetee, S. and D. A. Reardon (2009). "Targeting multiple kinases in glioblastoma multiforme." Expert Opin Investig Drugs 18(3): 277-92.

Sears, R. C. and J. R. Nevins (2002). "Signaling networks that link cell proliferation and cell fate." J Biol Chem 277(14): 11617-20.

Sharma, S. and A. Rao (2009). "RNAi screening: tips and techniques." Nat Immunol 10(8): 799-804.

Sheaff, R. J., M. Groudine, M. Gordon, J. M. Roberts and B. E. Clurman (1997). "Cyclin E-CDK2 is a regulator of p27Kip1." Genes Dev 11(11): 1464-78.

Sherr, C. J. (1994). "G1 phase progression: cycling on cue." Cell 79(4): 551-5.

Sherr, C. J. and J. M. Roberts (1995). "Inhibitors of mammalian G1 cyclin-dependent kinases." Genes Dev 9(10): 1149-63.

Sherr, C. J. and J. M. Roberts (1999). "CDK inhibitors: positive and negative regulators of G1-phase progression." Genes Dev 13(12): 1501-12.

Sherr, C. J. and J. M. Roberts (2004). "Living with or without cyclins and cyclin-dependent kinases." Genes Dev 18(22): 2699-711.

Shi, Y. (2009). "Serine/threonine phosphatases: mechanism through structure." Cell 139(3): 468-84.

Shiraishi, T., S. Matsuyama and H. Kitano (2010)."Large-scale analysis of network bistability for human cancers." PLoS Comput Biol 6(7): e1000851.

Shoval, O. and U. Alon (2010). "SnapShot: network motifs." Cell 143(2): 326-e1.

Sirvent, A., C. Benistant and S. Roche (2008). "Cytoplasmic signalling by the c-Abl tyrosine kinase in normal and cancer cells." Biol Cell 100(11): 617-31.

Slamon, D. J., G. M. Clark, S. G. Wong, W. J. Levin, A. Ullrich and W. L. McGuire (1987). "Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene." Science 235(4785): 177-82.

Stone, D. J., S. Marine, J. Majercak, W. J. Ray, A. Espeseth, A. Simon and M. Ferrer (2007). "High-throughput screening by RNA interference: control of two distinct types of variance." Cell Cycle 6(8): 898-901.

Straube, A., B. Aicher, B. L. Fiebich and G. Haag (2011). "Combined analgesics in (headache) pain therapy: shotgun approach or precise multi-target therapeutics?" BMC Neurol 11: 43.

Supper, J., L. Spangenberg, H. Planatscher, A. Drager, A. Schroder and A. Zell (2009). "BowTieBuilder: modeling signal transduction pathways." BMC Syst Biol 3: 67.

Tomlinson, D. C., O. Baldo, P. Harnden and M. A. Knowles (2007). "FGFR3 protein expression and its relationship to mutation status and prognostic variables in bladder cancer." J Pathol 213(1): 91-8.

Tononi, G., O. Sporns and G. M. Edelman (1999). "Measures of degeneracy and redundancy in biological networks." Proc Natl Acad Sci U S A 96(6): 3257-62.

Vermeulen, K., D. R. Van Bockstaele and Z. N. Berneman (2003). "The cell cycle: a review of regulation, deregulation and therapeutic targets in cancer." Cell Prolif 36(3): 131-49.

Vlach, J., S. Hennecke and B. Amati (1997). "Phosphorylation-dependent degradation of the cyclin-dependent kinase inhibitor p27." Embo J 16(17): 5334-44.

Vogelstein, B. and K. W. Kinzler (2004). "Cancer genes and the pathways they control." Nat Med 10(8): 789-99.

Welch, P. J. and J. Y. Wang (1993). "A C-terminal protein-binding domain in the retinoblastoma protein regulates nuclear c-Abl tyrosine kinase in the cell cycle." Cell 75(4): 779-90.

Welch, P. J. and J. Y. Wang (1995). "Abrogation of retinoblastoma protein function by c-Abl through tyrosine kinase-dependent and -independent mechanisms." Mol Cell Biol 15(10): 5542-51.

Wiles, A. M., D. Ravi, S. Bhavani and A. J. Bishop (2008). "An analysis of normalization methods for Drosophila RNAi genomic screens and development of a robust validation scheme." J Biomol Screen 13(8): 777-84.

Willis, T. G., D. M. Jadayel, M. Q. Du, H. Peng, A. R. Perry, M. Abdul-Rauf, H. Price, L. Karran, O. Majekodunmi, I. Wlodarska, L. Pan, T. Crook, R. Hamoudi, P. G. Isaacson and M. J. Dyer (1999). "Bcl10 is involved in t(1;14)(p22;q32) of MALT B cell lymphoma and mutated in multiple tumor types." Cell 96(1): 35-45.

Yarden, Y. and A. Ullrich (1988). "Growth factor receptor tyrosine kinases." Annu Rev Biochem 57: 443-78.

Zetterberg, A., O. Larsson and K. G. Wiman (1995). "What is the restriction point?" Curr Opin Cell Biol 7(6): 835-42.

Zhang, Q., R. Siebert, M. Yan, B. Hinzmann, X. Cui, L. Xue, K. M. Rakestraw, C. W. Naeve, G. Beckmann, D. D. Weisenburger, W. G. Sanger, H. Nowotny, M. Vesely, E. Callet-Bauchu, G. Salles, V. M. Dixit, A. Rosenthal, B. Schlegelberger and S. W. Morris (1999). "Inactivating mutations and overexpression of BCL10, a caspase recruitment domain-containing gene, in MALT lymphoma with t(1;14)(p22;q32)." Nat Genet 22(1): 63-8.

Zhang, W. and H. T. Liu (2002). "MAPK signal pathways in the regulation of cell proliferation in mammalian cells." Cell Res 12(1): 9-18.

Zhang, X. D. and J. F. Heyse (2009). "Determination of sample size in genome-scale RNAi screens." Bioinformatics 25(7): 841-4.

Zhu, X., M. Gerstein and M. Snyder (2007). "Getting connected: analysis and principles of biological networks." Genes Dev 21(9): 1010-24.

Zimmermann, G. R., J. Lehar and C. T. Keith (2007). "Multi-target therapeutics: when the whole is greater than the sum of the parts." Drug Discov Today 12(1-2): 34-42.

Chapter 10

Publications

Chapter 10: Publications

Publications

1. **Jailkhani, N.**, S. Ravichandran, S. Hegde, Z. Siddiqui, S. C. Mande and K. V. S. Rao. Delineation of key regulatory elements identifies points of vulnerability in the mitogen-activated signaling network (Communicated).

2. **Jailkhani N**, V. K. Chaudhri and K. V. S. Rao. Regulatory Cascades of Protein Phosphatases: Implications for Cancer Treatment. Anticancer Agents in Medicinal Chemistry. 2011.

3. Jamal MS, Ravichandran S, **Jailkhani N**, Chatterjee S, Dua R, Rao KV. Defining the antigen receptor-dependent regulatory network that induces arrest of cycling immature B-lymphocytes. BMC Systems Biology. 4: 169, 2010.

Provisional Patents Filed

1. Title of Invention- "Method for identification of functional modules and novel protein drug targets from biological networks." Patent Application no. 1092/DEL/2010.

2. Title of Invention- "Cancer combination therapy comprising multi-molecular targeting." Patent Application no. 1326/DEL/2010.