Network-based identification of key proteins involved in apoptosis and cell cycle regulation

L. Wu, N. Zhou, R. Sun, X. D. Chen, S. C. Feng, B. Zhang and J. K. Bao

School of Life Sciences and Key Laboratory of Bio-resources and Eco-environment, Ministry of Education, State Key Laboratory of Biotherapy, Sichuan University, Chengdu, 610064, China

Received 17 February 2014; revision accepted 8 April 2014

Abstract

Objectives: Cancer cells differ from normal body cells in their ability to divide indefinitely and to evade programmed cell death. Crosstalk between apoptosis and cell cycle processes promotes balance between proliferation and death, and limits population growth and survival of cells. However, intricate relationships between them and how they are able to manipulate the fate of cancer cells still remain to be clarified. Identification of key factors involved in both apoptosis and cell cycle regulation may help to address this problem.

Materials and methods: Identification of such key proteins was carried out, using a series of bioinformatics methods, such as network construction and key protein identification.

Results: In this study, we computationally constructed human apoptotic/cell cycle-related protein-protein interactions (PPIs) networks from five experimentally supported protein interaction databases, and further integrated these high-throughput data sets into a Naïve Bayesian model to predict protein functional connections. On the basis of modified apoptotic/cell cycle related PPI networks, we calculated and ranked all protein members involved in apoptosis and cell cycle regulation. Our results not only identified some already known key proteins such as p53, Rb, Myc and Src but also found that the proteasome, Cullin family members, kinases and transcriptional repressors play important roles in regulating apoptosis and the cell cycle.

Correspondence: J. K. Bao, School of Life Sciences and Key Laboratory of Bio-resources, Ministry of Education, Sichuan University, Chengdu 610064, China. Tel./Fax: +86 28 85415171; E-mail: baojinku@scu.edu.cn

L. Wu and N. Zhou contributed equally to this work.

Furthermore, we found that the top 100 proteins ranked by PeC were enriched in some pathways such as those of cancer, the proteasome, the cell cycle and Wnt signalling.

Conclusions: We constructed the global human apoptotic/cell cycle related PPI network based on five online databases, and a Naïve Bayesian model. In addition, we systematically identified apoptotic/cell cycle related key proteins in cancer cells. These findings may uncover intricate relationships between apoptosis and cell cycle processes and thus provide further new clues towards future anti-cancer drug discovery.

Introduction

Cancer is characterized by seemingly uncontrolled cell population growth and spread of abnormal cells. In normal cells, the cell cycle is controlled by a complex series of signalling pathways by which a cell grows, replicates its DNA and divides. This process also includes mechanisms which ensure that errors therein are corrected, and if not, the cells ‘commit suicide’ in a systematic cell process known as programmed cell death, or apoptosis (1). In cancer cells, as a result of genetic mutations, this regulatory process malfunctions, resulting in uncontrolled cell proliferation. Deregulation of the cell cycle underlies aberrant cell proliferation that characterizes cancer, and loss of cell cycle checkpoint control promotes genetic instability. Evasion of apoptosis is one of the hallmarks of human cancers and promotes tumour formation and progression, as well as treatment resistance (2).

Apoptosis and cell cycle processes are highly interconnected at the molecular level. Both are affected by activity of many genes that may act synergistically to contribute to cancer development. Most

notably, cell cycle genes such as *p53*, *RB* and *MYC* have been shown to participate in both the cell cycle and in apoptosis; dysfunction or mutation of these and their products have been shown to play an important role in the progression of various types of malignancy (3–5).

Understanding the molecular basis of evasion of apoptosis and indefinite division of cancer cells are essential in identifying novel targets for prevention and treatment of malignancy. Comprehensive knowledge of apoptotic and cell cycle signalling networks may provide a framework for better understanding their relationships as an integrated system (6). With increasing genome-wide data on genetic, regulatory, functional and physical interactions, exploration of a potential 'key regulator' between apoptosis and the cell cycle is imperative for inferring their complex connections in cancer (7,8).

Of note, protein-protein interactions (PPIs) are involved in biological processes and signal transduction pathways at almost every level of cell function. Dissecting the PPI networks for a particular biological process may provide important clues into molecular mechanisms of the process (9). In addition, PPI networks can help to reveal signal transduction pathways and serve to aid in identification of novel regulatory components at the molecular level (10). Thus, PPI networks have quickly become valuable resources to research protein function and to understand molecular mechanisms underlying specific diseases (11).

Recently machine learning methods have been introduced to predict PPIs by combining genomic and experimental features. Among these machine learning approaches, the Naïve Bayesian model has been widely used in analysing PPI networks in a number of model organisms such as annelids, yeasts, dipterans, mice and humans (12–14). However, to our knowledge, few studies have been reported that have analysed specific disease or biological process-related PPI networks by use of the Naïve Bayesian model. Due to the extreme intricacy of apoptosis and cell cycle signalling networks, prediction of novel apoptotic and cell cycle PPIs will help to clarify them.

In this study, we constructed a Naïve Bayesian model to predict apoptotic and cell cycle PPIs. On the basis of apoptotic/cell cycle related PPI networks, we systematically identified apoptotic/cell cycle related key proteins and their pathways involved in cancer cells. Our findings will, perhaps, not only uncover the complex relationships between apoptosis and cell cycle processes but also provide further potential information towards discovery of novel drugs for anti-cancer therapy.

Materials and methods

Construction of gold standard data sets

Gold standard positive (GSP) refers to known PPIs available in databases. First, we built the global human PPI networks by collecting diverse data sets from five online databases. These were from the Biomolecular Object Network Databank (BOND) (15), IntAct (16), Human Protein Reference Database (HPRD) (17), Database of Interacting Proteins (DIP) (18) and from BioGRID (19). Subsequently, the GSP set was established by extracting protein pairs that had at least one apoptotic/cell cycle protein, by Gene Ontology (GO) consortium, in a pair of proteins. We used the gold standard negative (GSN) data set, previously generated by Rhodes *et al.* (20), which includes all possible pairwise combinations between two sets of proteins, that are assigned subcellular localization of the plasma membrane (6541 proteins) and the nucleus (10 317 proteins), respectively, by the GO consortium. Overall, 923 proteins that were assigned to both components were removed. This method is based on an assumption that pairs of proteins from different cell locations are unlikely to interact.

Biological evidence extraction

Interacting protein pairs share features in sequence, functional annotation, domain and expression profiles; thus they can be used to represent a pair of proteins and predict PPIs. In this study, features of amino acid composition, shared functional annotation, enriched protein domain pairs and gene co-expression profiles were used to define interaction signatures of protein pairs. The interaction signature was defined as the combination of two protein signatures from each partner. Interaction signatures were denoted as positive if observed in the GSP, or negative if observed in the GSN. We calculated the number of times an interaction signature appeared in protein pairs of the GSP or the GSN. Total number of interaction signatures and protein signatures are summarized in Table S1.

The sequential feature

An amino acid sequence contains a great deal of useful information and might be effective for estimation of interacting propensity between two proteins (21). Suitable encoding of protein information is needed for PPI prediction. Here we used amino acid composition as protein chain descriptor to gain insight into protein interactions, using sequence information.

Let $P = \{p1, p2, p3, \dots, p20\}$ represent percentage of 20-amino acid compositions in a protein. For a certain protein pair (A and B), we acquired the interacting feature as follows: first, calculate 20-amino acid compositions of protein A and B respectively. Second, let $Pa = \{a1, a2, a3, \dots, a20\}$ denote amino acid percentages of protein A and $Pb = \{b1, b2, b3, \dots, b20\}$ denote protein B. Third, use Pa and Pb to calculate standard deviation (σ) of protein A (σ_a) and B (σ_b), respectively. The interacting feature F_{ab} of the given protein pair was calculated as: $F_{ab} = \sigma_a / \sigma_b (\sigma_a \sigma_b)$. Subsequently, F_{ab} was used to represent the sequential feature of protein pair 'AB'.

Shared functional annotation

Proteins with the same biological function are more likely to physically interact than those without. In addition, proteins sharing a more specific annotation are more likely to interact than those sharing a more common less specific annotation. To quantify similarity between gene annotations, we identified the smallest shared biological process (SSBP) between a pair of proteins (20). SSBP was calculated by performing three steps: first, find all GO terms shared by each pair of proteins. Secondly, find the number of other proteins also sharing these GO terms. Thirdly, obtain GO term with the smallest protein count. In agreement with expectation, smaller the SSBP count, the more likely proteins are to directly interact.

Enriched protein domain pairs

Protein domains, regions with specific structures and dependent functions, can be used to predict PPIs. As protein interactions involve physical associations between them (20,22), it has been proposed that protein domain pairs enriched among known human PPIs may indicate novel interactions (22). To test this proposal in the context of our GSP and GSN sets, we used a collection of hidden Markov models of protein domains from the Pfam database, version 26.0 (23). We assumed that the combination number of domain pairs of each protein pair represent the domain enrichment. Let n represent the domain number of a protein, according to the Pfam database and N denote the combination number of domain pairs; the protein pair 'AB' can be represented by $N_{ab} = n_a \times n_b$.

Gene co-expression

Genes that exhibit mRNA co-expression tend to show protein interaction, especially those in the same complex

or in the same biochemical reaction. Such correlated genes might be regulated by the same transcriptional factor or a set of transcription factors (20,24). To identify genes that are co-expressed, we used microarray data (NCBI GEO database, accession GSE31432) of human ovarian carcinoma (SCOV3) cells, treated with trastuzumab, pertuzumab or a combination to measure pairwise co-expression level of related genes, in apoptosis and the cell cycle (25). Co-expression level was calculated as the Pearson correlation coefficient (PCC):

$$\text{PCC}_{(X,Y)} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{(n-1)\sigma_X\sigma_Y}$$

where X and Y are expression level data vectors of length n for two genes, $\bar{X}$ and $\bar{Y}$ are means, and σ_X and σ_Y are standard deviations.

Integrating evidence by Naïve Bayesian classifiers

We integrated evidence from different sources by the Naïve Bayesian probabilistic model, described extensively in a number of studies previously (20,26). Briefly, we defined a pair of proteins that interact with each other as 'positive' and those that do not interact as 'negative'. Following the Bayesian theorem, the prior odds (O_{prior}) of finding a PPI pair can be calculated as:

$$O_{\text{prior}} = P_{\text{pos}} / P_{\text{neg}}$$

where P_{pos} is probability that a pair of proteins interacts within all the possible protein pairs while P_{neg} stands for the possibility that a pair of proteins does not interact. When considering given n evidence (E), posterior odds ($O_{\text{posterior}}$) of an interaction can be computed as:

$$O_{\text{posterior}} = \frac{P(\text{pos}|E_1, \dots, E_n)}{P(\text{neg}|E_1, \dots, E_n)}$$

Likelihood ratio (LR) is defined as:

$$\text{LR}_{(E_1, \dots, E_n)} = \frac{P(E_1, \dots, E_n|\text{pos})}{P(E_1, \dots, E_n|\text{neg})}$$

Posterior odds of an interaction can be calculated as product of prior odds and the likelihood ratio ($O_{\text{posterior}} = O_{\text{prior}} \times \text{LR}_{(E_1, \dots, E_n)}$). If we assume that evidence is conditionally independent, composite LR can be calculated simply as following:

$$\text{LR}_{(E_1, \dots, E_n)} = \prod_{i=1}^n \text{LR}(E_i)$$

which is the Naïve Bayesian model. Evidence for this, is the data sets used to infer PPIs between proteins. As prior odds is a constant, composite LR corresponding to

a specific piece of biological evidence, can be used to measure the predictive power or confidence degree for predicting PPIs. Cut-off likelihood ratio is represented as an indicator whether a protein pair interacts (that is, 'yes' if the composite LR is above the LR cut-off, 'no' if not). By filtering in the Naïve Bayesian model, resultant protein pairs with composite LR above the LR cut-off, were identified as true positives and used to construct the apoptotic/cell cycle related network.

Evaluation of the Naïve Bayesian classifier

We computed three standard evaluation criteria (27) to evaluate the performance of the Naïve Bayes classifier as follows:

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}$$

where TP, FP, TN and FN, denote true positive (TP), false positive (FP), true negative (TN), and false negative (FN) respectively. For further evaluation, we employed a common unbiased measure – the receiver operating characteristic (ROC) curve (28), which is a plot between true positive proportion (sensitivity) and false positive proportion ($1 - \text{specificity}$). We also calculated the area beneath the ROC curve (AUC); this has positive correlation with the classifier prediction performance.

Identification of key regulators in apoptotic/cell cycle related PPI networks

We determined key regulators in apoptotic/cell cycle related PPI networks based on the following two standards. First, hub proteins that are well known, tended to play more crucial roles in PPI networks (29). In this study, a method termed PeC (which integrates both PPIs and gene expression data), proposed by Li *et al.*, was applied to identify hub proteins in the apoptotic/cell cycle related PPI networks (30). This method has been proven to be more effective for hub protein identification in yeast PPI networks than many other methods of measurement, such as degree, density of maximum neighbourhood component (DMNC) (31) and Bottle Neck (32) which is only based on PPI network topological features. With this method, we calculated PeC scores

of all protein members in apoptotic/cell cycle related PPI networks. Briefly, probability that a protein co-clustered and co-expressed with all apoptotic/cell cycle members, was evaluated, combining edge clustering coefficient (ECC) and Pearson correlation coefficient. Edge clustering coefficient is used to analyse topology of PPI networks. For an edge (u, v) connecting node u and node v , ECC of an edge (u, v) is defined as:

$$\text{ECC}(u, v) = \frac{|N_u \cap N_v| + 1}{\min\{d_u, d_v\}}$$

where N_u (or N_v) is the set of neighbours of vertex u (or v), and d_u (or d_v) denotes degree of vertex u (or v). $\text{ECC}(u, v)$ is a local variable which characterizes closeness of two proteins u and v . Also, we defined the probability of paired proteins u and v to be in the same cluster as following:

$$p_c(u, v) = \text{ECC}(u, v) \times \text{PCC}(u, v)$$

Pearson's correlation coefficient of a pair of proteins (u and v) as equal to PCC of their corresponding paired genes. For protein v , its $\text{PeC}(v)$ is defined as the sum of probabilities that the protein and its neighbours belong to a same cluster:

$$\text{PeC}(v) = \sum_{u \in N_v} p_c(u, v)$$

Value of PeC was determined not only by how many neighbours the protein has, but also how likely it is to be co-clustered with its neighbours. For protein v , higher the value of PeC, the more important it is in the apoptotic/cell cycle related PPI network. Secondly, we indicated that hub proteins should connect known apoptotic/cell cycle members, thereby making them worthy of particular attention when developing novel drug targets. This method was similar to previous studies that focused on identifying novel cancer-related genes (33). These proteins identified as both hub proteins and connect apoptotic/cell cycle members were extracted as key proteins implicated in apoptotic/cell cycle related PPI networks.

Evaluation of PeC method of hub protein identification

To evaluate performance of the PeC method of hub protein identification, global human PPI networks and above-mentioned gene expression data were used. Human hub proteins were collected from the DEG databases (34). Of all 14 704 proteins in global human PPI networks, 2452 are hub proteins, the rest are non-hub or not-known to be hub or non-hub. Subsequently, we carried out a comparison between PeC and topological features-based measure methods such as degree, DMNC

and Bottle Neck. Proteins were ranked according to their values calculated by each measure method. A certain number of top proteins was selected as candidates for being hub proteins. Then, we counted how many of them were true hub proteins. Number of hub proteins detected by PeC and three other measurement methods were compared.

Results

Performance of the Naïve Bayesian classifier

First, a set of 149 730 non-redundant PPIs (14 704 proteins) were obtained from five online databases: BOND, IntAct, HPRD, DIP and BioGRID. Subsequently, an apoptotic/cell cycle related PPI network was established by extracting protein pairs that had at least one apoptotic or cell cycle protein, by GO consortium, in a pair of proteins. Ultimately, a total number of 78 531 distinct PPIs among 12 841 proteins was queried from the five online databases as our GSP set. Correspondingly, a GSN set of 1 000 060 protein pairs was defined, in which one protein was assigned to the plasma membrane cell component and the other to the nuclear cell component, by the GO consortium.

On the basis of the two Gold sets, we integrated heterogeneous biological evidence from four independent data sets and used likelihood ratios as the reliability of individual data sets to infer human apoptotic/cell cycle related PPIs. Each data set was divided into several bins dependent on their intrinsic characters, and LR for each bin was calculated, as shown in Table S1.

The Naïve Bayesian classifier is based on the assumption that predictive evidence is strongly independent; thus, we assessed statistical dependence between each pair of data sets by calculating the PCC (13,35). These results indicated that data from the sequential feature, SSBP, enriched protein domain pairs and gene co-expression, are conditionally independent which can be used as individual predictors in a Naïve Bayesian model (Table S2).

For training and evaluation, we randomly divided GSP and GSN into five parts respectively. In practice, four of five of protein pairs of GSP and GSN were used as training sets and the rest were used as test sets. The classifier was trained by 863 068 training protein pairs and tested by 215 767 pairs. Then, we counted numbers of TP, TN, FP and FN to calculate sensitivity, specificity and accuracy under different cut-off values (Table S3). ROC, a commonly used unbiased standard, can reflect classifier performance by its AUC. Plotted by pROC (36) library of R language, the ROC of test results has an AUC of 0.815 (Fig. S1). Evaluation indi-

cates that the Naïve Bayesian classifier developed by us performs well in prediction of apoptotic/cell cycle PPIs. To find the optimal threshold we plotted Fig. S2 to show TP/FP ratios at different LR cut-offs, indicating that 160 is the optimal threshold with an accuracy of 92.754%, specificity, higher than 99.99%, and FPR less than 0.004% (Table S3).

Construction of the apoptotic/cell cycle related PPI network

For better understanding of interactions of 1022 apoptotic proteins and 1250 cell cycle proteins, with all other human proteins, we constructed the global human apoptotic/cell cycle related PPI network. First, we downloaded 20 248 reviewed human proteins from UniProt (<http://www.uniprot.org>). On the basis of the Naïve Bayesian model and LR cut-off 160, we achieved 8301 interacting apoptotic/cell cycle protein pairs. After removing duplicate interactions with the known 78 531 apoptotic/cell cycle protein pairs of the GSP, we obtained the global human apoptotic/cell cycle related network containing 85 428 interactions of 12 944 proteins. The procedure of constructing the apoptotic/cell cycle related PPI network is illustrated in Fig. 1.

Of note, apoptosis can be induced by two core pathways, the extrinsic death receptor pathway, and the intrinsic mitochondrial pathway (37). Life or death of the cell is regulated by a sophisticated balance between pro-apoptotic and anti-apoptotic proteins and their interactions (38). Caspases and Bcl-2 family members are pivotal regulators of apoptosis.

Cell cycle regulation is achieved through a family of serine/threonine kinase holoenzyme complexes, consisting of regulatory cyclin subunits that bind to and activate catalytic cyclin-dependent kinases (CDKs) (39). Due to the importance of caspases, Bcl-2 and CDKs family members in apoptosis and cell cycle regulation, we extracted their PPI interactions for further analysis. We found many novel predicted interactions, including that caspase-3 can interact with RAIDD and caspase-8 can interact with Sirt1, Jak and Stat1. In addition, we found that caspase-9 can interact with Chk1, Beclin1 can interact with Sirt1, and Mcl1 can interact with FADD/caspase-8. We also discovered that CDK2 can interact with Sin3a, and CDK6 can interact with PTEN (Fig. 2). In addition, we found proteins such as 14-3-3 protein epsilon (YWHA_E), 78 kDa glucose-regulated protein (GRP78), 14-3-3 protein theta (YWHA_Q), Vimentin (VIM), COP9 signalosome complex subunit 6 (COPS6), Baculoviral IAP repeat-containing protein 5 (BIRC5) and Transcription factor Sp1 (Sp1) can connect both caspases and CDKs family members, indicating

Figure 1. Flowchart for construction the apoptotic/cell cycle related protein-protein interactions (PPI) network.

their important role in regulating apoptosis and cell cycle processes (Fig. 3a). Furthermore, we found that CDK1, CDK2, CDK5 and caspase-8 can connect both p53 and Rb, indicating they may be important modifiers of p53 and Rb (Fig. 3b). Surprisingly, we found poly-ubiquitin-C (*UBC*), with unexpectedly high levels of connectivity to caspases/CDK-related proteins, which might have an unrecognized function of regulation of apoptosis and cell cycle mechanisms (Fig. 3c). Furthermore, we also found JNK can interact with both caspase family members (caspase-3, caspase-8 and caspase-10) and Bcl-2 family members (BAD, Bid and Bcl-2), indicating its central role in apoptosis regulation (Fig. 3d). Interestingly, we also found that VCP can interact with CDK family members (CDK1, CDK2, and CDK4), caspase family members (caspase-7, caspase-9) and Bcl-2 family members (BAD, Bid and bcl-2) (Fig. 3e).

Identification of key regulators involved in both apoptosis and cell cycle processes

To identify key regulators in the apoptotic/cell cycle related PPI networks, we first evaluated performance of the PeC hub protein identification method, by comparing PeC and topological features-based measurement methods degree, DMNC and Bottle Neck. From Fig. 4 we can see that PeC significantly outperformed the other three methods for predicting hub proteins from human PPI networks. These results suggest that our method used for predicting hub proteins might be effective for predicting hub proteins in the apoptotic/cell cycle related PPI network.

Subsequently, we applied PeC to identify hub proteins in the apoptotic/cell cycle related PPI network. These were both hub proteins and connect to known apoptosis/cell cycle related proteins, and were extracted as key proteins implicated in the apoptosis/cell cycle related PPI network. With this method, we calculated and ranked all protein members in the apoptosis/cell cycle related PPI network based on their PeC-scores.

Then, we retrieved the top 100 key regulators scored and ranked by PeC. Also, we retrieved the top 100 key regulators scored and ranked by degree, then compared the regulators predicted by PeC and by degree. There were 78 proteins that were predicted by both PeC and degree of the top 100 key regulators. This information is shown in Table S4. We found that some well-characterized proteins associated with apoptosis and cell cycle progression, such as p53, Rb, Myc and Src, were predicted by both PeC and degree; the remaining 23 proteins predicted by PeC but ignored by degree when predicting the top 100 proteins are shown in Table 1. These were mainly proteasome subunits.

To further illustrate the top 100 ranked proteins identified by PeC, we further enrolled these proteins into the DAVID database for functional annotation and pathway enrichment analysis (40). The top 100 proteins ranked by PeC can be classified into four groups. The first were mainly proteasome units, including proteasome subunits alpha type 1, 2, 3, 4, 5, 6 (*PSMA1*, 2, 3, 4, 5, 6), proteasome subunits beta type-1, 3, 5 (*PSMB1*, 3, 5), 26S protease regulatory subunits 4, 7, 6A, 8, 10B (*PSMC1*, 2, 3, 5, 6), 26S proteasome non-ATPase regulatory subunits 1, 2, 4, 6, 7, 8 (*PSMD1*, 2, 4, 6, 7, 8), *UBC*, *UBE2D1* and transitional endoplasmic reticulum ATPase (*VCP*) with an enrichment score 26.21. The second group was comprised of Cullin family proteins including Cullins 1, 2, 3, 4A, 4B, 5 (*CUL1*, 2, 3, 4A, 4B, 5) with enrichment score 15.0. The third group was comprised of kinases such as CDK2, CDK4, CDK9, casein kinase 1 (CK1) isoform alpha, CK1 isoform alpha epsilon (*CSNK1A1*, *CSNK1E*), casein kinase 2 (CK2) subunit alpha, CK2 subunit alpha' (*CSNK2A1*, *CSNK2A2*), GSK3B, IkbKE, MAPK6, AMPK subunit alpha-1 (*PRKAA1*) and Serine/threonine-protein kinases 4, 24 (*STK4*, 24) with enrichment score 7.50. The fourth group was comprised of transcriptional repressors such as histone deacetylases 1, 3 (*HDAC1*, 3), interleukin enhancer-binding factor 3 (*ILF3*) and nuclear receptor corepressor 1 (*NCOR1*) with enrichment score 6.24.

Figure 2. Apoptosis and cell cycle signalling network. Our predicted interactions by Naïve Bayesian classifier are marked by dotted lines.

Gene functional classification of the top 100 proteins ranked by PeC is summarized in Table 2.

Thus, we saw in detail the functional annotations of the top 100 proteins ranked by PeC. These key regulators were enriched in 14 categories with P -value < 0.05 and $FDR < 0.05$, as shown in Table 3. Major functional annotations for these genes are Pathways in cancer (hsa05200), Proteasome (hsa03050), Cell cycle (hsa04110) and Wnt signalling pathway (hsa04310).

Discussion

Pathways involving the cell cycle and apoptosis have been found to play an important role in tumourigenesis. While there is abundant evidence that pathways involving p53, Rb and many key mediators of cell cycle regulation and apoptosis contribute to various types of cancer, how these critical proteins participate in apoptosis and cell cycle regulating processes has not been clearly understood.

In this study, we constructed the global human apoptotic/cell cycle related PPI network which contains 85 428 interactions of 12 944 proteins, on the basis of five online databases and a Naïve Bayesian model. We found that apoptosis and the cell cycle are highly linked processes and share common molecules. Some proteins such as 14-3-3 protein epsilon, 78 kDa glucose-regulated protein, 14-3-3 protein theta, Vimentin, COP9 signalosome complex subunit 6, Baculoviral IAP repeat-containing protein 5 and Transcription factor Sp1 can connect both caspases and CDK family members, showing that they may be important mediators in regulating apoptosis and the cell cycle by interacting with caspases and CDK family members. In addition, CDK2 has been well known to be an important modifier of p53 and pRb (41); our finding shows CDK1, CDK5 and caspase-8 also are modifiers of p53 and Rb. As p53 and Rb have been widely reported to be key mediators of apoptosis and cell cycle regulation, our findings may help clarify their molecular bases in apoptosis and cell cycle control.

Figure 3. Caspases, Bcl-2 and CDK family members PPI network. (a) Global caspases, Bcl-2 and CDK family members PPI network, magenta nodes are caspase family members, yellow nodes Bcl-2 family members and green nodes CDK family members. (b) Sub-network of p53 and pRb. (c) Sub-network of UBC. (d) Sub-network of JNK. (e) Sub-network of VCP.

Furthermore, we also found VCP, also known as Ubiquitin-dependent molecular chaperone p97 can interact with CDK family members, caspase family members and Bcl-2 family members; these intricate relationships of VCP indicate that it may serve as a bridge linking different core apoptotic and cell cycle pathways. VCP has recently been reported to be involved in the ubiquitin-dependent proteasome degradation pathway of Inhibitor $\kappa B\alpha$ ($I\kappa B\alpha$), an inhibitor of nuclear factor- κB ($NF\kappa B$); inhibition of VCP has been shown to promote cancer cell apoptosis *via* UPR activation and disruption of autophagy (42,43). These findings suggest VCP may be an important therapeutic target in tumours.

With the PeC method, we calculated and ranked all protein members in the apoptotic/cell cycle related PPI network based on their PeC scores. We found that the top 100 key regulators predicted by PeC can be classified into four groups including proteasome, Cullin family members, kinases and transcriptional repressors.

The proteasome is a multicatalytic proteinase complex responsible for degradation of most intracellular proteins, including proteins crucial to cell cycle regulation and programmed cell death (44). The 26S proteasome is a large (1500–2000 kDa) multi-subunit complex that identifies and degrades proteins tagged for destruction by the ubiquitin system. Orderly degradation of cell proteins is critical for normal cell cycling and function, and inhibition of the proteasome pathway results in cell cycle arrest and apoptosis (45). Dysregulation of this enzymatic system may also play a role in tumour progression, drug resistance and altered immune surveillance, making the proteasome an appropriate and novel therapeutic target in cancer (46,47).

In addition, the ubiquitin-proteasome system plays a pivotal role in maintaining balance between normal growth and uncontrolled cell proliferation by controlling abundance of a large variety of cell proteins (48). Alterations in ubiquitination are observed in most cancer cells

Figure 4. Comparison of number of hub proteins detected by PeC and the other three previously proposed hub protein measurement methods.

(49). The Cullin family of ubiquitin ligases, traditionally composed of Cullin 1, 2, 3, 4A, 4B, 5 and 7, represents the largest class of RING-type E3 ligases (CRLs) (48). Without intrinsic catalytic activity, Cullins serve as scaffolds that facilitate assembly of multimeric E3 ligase complexes and transfer ubiquitin from E2-conjugating enzyme to the substrate. Cullin-mediated substrate degradation dictates a wide range of cell processes such as proliferation, differentiation and apoptosis. Once cell regulatory mechanisms of Cullin encounter malfunctions

or perturbations, accumulation of oncoproteins or excessive degradation of tumour suppressors will inevitably occur, which may provoke cells into malignant transformation and tumourigenesis (50). In particular, Cullin1, the most completely characterized member of the Cullin family, has been shown to be closely associated with gastric carcinogenesis, and its overexpression predicts poor prognosis of patients with gastric carcinoma (51). In addition, Cullin3 has been reported to be down-regulated in breast and kidney cancer, suggesting a tumour suppressive role. In addition, Cullin4A has been found to be amplified or over-expressed in primary breast cancer (48). Thus, more investigation is required to uncover intricate relationships between Cullins and cancer.

Protein kinases, enzymes that transfer phosphate groups from high-energy molecules, control a series of cell processes, including metabolism, transcription, differentiation, cell cycle progression and apoptosis (52). Previous studies have described that aberrant activation of CDKs may contribute to tumourigenesis (53) and this study has highlighted that CDK2, 4, 9 play an important role in the apoptosis/cell cycle related PPI network. The serine/threonine-specific CK1 and 2 protein kinase families have been reported to play an important role in development of cancer (54,55). CK1 α is an essential component of the Wnt/ β -catenin signalling pathway. It binds to axin, and phosphorylates β -catenin, ultimately causing degradation of β -catenin (56). CK1 δ/ϵ has been

Table 1. List of 23 proteins predicted by PeC but ignored by degree when predicting the top 100 key regulators

Gene name	Protein name	PeC score	Degree
CIQBP	Complement component 1 Q subcomponent-binding protein, mitochondrial	32.56846	162
FBXO6	F-box only protein 6	33.66942	156
HTT	Huntingtin	39.875	140
PSMA1	Proteasome subunit alpha type-1	31.50527	121
PSMA2	Proteasome subunit alpha type-2	34.3422	112
PSMA4	Proteasome subunit alpha type-4	27.31699	82
PSMA5	Proteasome subunit alpha type-5	28.253	101
PSMA6	Proteasome subunit alpha type-6	34.43836	119
PSMB1	Proteasome subunit beta type-1	27.77776	112
PSMB3	Proteasome subunit beta type-3	28.15656	77
PSMB5	Proteasome subunit beta type-5	27.32591	86
PSMC1	26S protease regulatory subunit 4	28.10321	95
PSMC2	26S protease regulatory subunit 7	33.6666	109
PSMC3	26S protease regulatory subunit 6A	29.00672	102
PSMC5	26S protease regulatory subunit 8	36.97223	138
PSMC6	26S protease regulatory subunit 10B	26.8077	94
PSMD1	26S proteasome non-ATPase regulatory subunit 1	29.09145	115
PSMD2	26S proteasome non-ATPase regulatory subunit 2	33.34872	139
PSMD4	26S proteasome non-ATPase regulatory subunit 4	31.84587	137
PSMD6	26S proteasome non-ATPase regulatory subunit 6	29.21329	90
PSMD7	26S proteasome non-ATPase regulatory subunit 7	27.57661	76
PSMD8	26S proteasome non-ATPase regulatory subunit 8	27.35977	71

Table 2. Gene functional classification of top 100 proteins ranked by PeC

Classification	Count	Gene name	Enrichment score
Group 1	23	PSMA1, PSMA2, PSMA3, PSMA4, PSMA5, PSMA6, PSMB1, PSMB3, PSMB5, PSMC1, PSMC2, PSMC3, PSMC5, PSMC6, PSMD1, PSMD2, PSMD4, PSMD6, PSMD7, PSMD8, UBC, UBE2D1, VCP	26.20859121
Group 2	6	CUL1, CUL2, CUL3, CUL4A, CUL4B, CUL5	14.99829133
Group 3	13	CDK2, CDK4, CDK9, CSNK1A1, CSNK1E, CSNK2A1, CSNK2A2, GSK3B, IKBKE, MAPK6, PRKAA1, STK24, STK4	7.503054387
Group 4	4	HDAC1, HDAC3, ILF3, NCOR1	5.680204261

Table 3. Pathway enrichment analysis of top 100 proteins ranked by PeC

KEGG pathway	Count	Genes	Fold enrichment	P-value	FDR
hsa05200:Pathways in cancer	22	TRAF2, HSP90AA1, TP53, SMAD3, NFKB1, RB1, CDK4, STAT1, STK4, PTEN, CDK2, MAPK1, CUL2, EP300, HDAC1, JUN, GSK3B, IKBKG, TRAF6, ABL1, MYC, APC	4.109241845	2.61E-08	2.79E-05
hsa03050:Proteasome	20	PSMB5, PSMA2, PSMA1, PSMC6, PSMC5, PSMB1, PSMA6, PSMA5, PSMC3, PSMB3, PSMC2, PSMA4, PSMD1, PSMC1, PSMA3, PSMD2, PSMD4, PSMD6, PSMD7, PSMD8	26.0702384	4.63E-23	4.95E-20
hsa04110:Cell cycle	18	YWHAZ, YWHAB, TP53, SMAD3, PRKDC, RB1, CDK4, CDK2, YWHAG, YWHAH, EP300, HDAC1, GSK3B, YWHAQ, PCNA, ABL1, MYC, CUL1	8.822168675	6.18E-12	6.6E-09
hsa04310:Wnt signaling pathway	14	CSNK1A1, TP53, SMAD3, CSNK2A2, EP300, CSNK2A1, CSNK1E, PPP2CA, GSK3B, JUN, RUVBL1, MYC, CUL1, APC	5.680204261	6.34E-07	0.000678
hsa04722:Neurotrophin signaling pathway	12	MAPK1, YWHAZ, YWHAG, YWHAH, JUN, GSK3B, YWHAB, YWHAQ, TP53, NFKB1, TRAF6, ABL1	5.928876798	0.00000375	0.00401
hsa04114:Oocyte meiosis	11	MAPK1, PPP1CA, YWHAZ, YWHAG, YWHAH, PPP2CA, YWHAB, YWHAQ, PPP1CC, CDK2, CUL1	6.126506024	0.00000864	0.009231
hsa05220:Chronic myeloid leukemia	10	MAPK1, HDAC1, IKBKG, TP53, SMAD3, NFKB1, RB1, ABL1, CDK4, MYC	8.168674699	0.0000025	0.002673
hsa05222:Small cell lung cancer	10	TRAF2, IKBKG, TP53, NFKB1, RB1, TRAF6, CDK4, MYC, PTEN, CDK2	7.293459552	0.00000652	0.006965
hsa05215:Prostate cancer	10	MAPK1, HSP90AA1, EP300, GSK3B, IKBKG, TP53, NFKB1, RB1, PTEN, CDK2	6.883714634	0.0000105	0.011265
hsa04120:Ubiquitin mediated proteolysis	9	CUL3, CUL2, CUL5, CUL4A, CUL4B, UBE2D1, TRAF6, BRCA1, CUL1	4.024711987	0.001504	1.594687
hsa05212:Pancreatic cancer	8	MAPK1, IKBKG, TP53, SMAD3, NFKB1, RB1, CDK4, STAT1	6.807228916	0.000139	0.147969
hsa05213:Endometrial cancer	7	MAPK1, GSK3B, ILK, TP53, MYC, PTEN, APC	8.247219648	0.000165	0.175792
hsa05210:Colorectal cancer	7	MAPK1, JUN, GSK3B, TP53, SMAD3, MYC, APC	5.105421687	0.002182	2.307076
hsa05219:Bladder cancer	5	MAPK1, TP53, RB1, CDK4, MYC	7.293459552	0.004414	4.61579

implicated in colon, pancreatic and breast cancer progression *via* the suppression of apoptosis, regulation of mitotic function, and activation of β -catenin signalling (57). In addition, recent studies also suggest that CK1 α promotes phosphorylation of MDM2 acidic domain, and regulates p53 stability; inhibition of CK1 α kinase activ-

ity may be beneficial to cancer treatment by activating p53 (58). CK2, reported to be important for cell survival and resistance to apoptosis, is over-expressed in a number of malignancies. CK2 negatively regulates caspase-8 activity and CK2 inhibitors increase the susceptibility of cancer cells to chemotherapeutic agents or apoptotic

stimuli (59). In addition, one recent study has shown that inhibition of CK2 induces p53-dependent cell cycle arrest and sensitizes glioblastoma cells to tumour necrosis factor (TNF α)-induced apoptosis through SIRT1 inhibition (60). Thus, the CDKs family and CK1 and 2 protein kinase families can be considered as potential therapeutic targets for various human cancers.

Transcriptional repressors such as NCORs and class I HDACs are receiving growing attention as potential therapeutic targets for human malignancies (61). Relevance of NCORs in cancer is still a recent field of investigation, but several HDAC-specific inhibitors are currently being tested in preclinical and clinical trials, including first trials with patients suffering from astrocytic gliomas (62). However, expression and functional role of NCORs and HDACs in cancer cells still remains to be clarified. Thus, more research is needed to understand the influence of these transcriptional repressors on tumour proliferation.

In summary, we computationally constructed the global human apoptosis/cell cycle related PPI network based on five online databases and a Naïve Bayesian model. In addition, we systematically identified apoptosis/cell cycle related key regulators and their relative signalling pathways in cancer cells. Our results not only identified some already known key proteins such as p53, Rb, Myc and Src involved in apoptosis and cell cycle regulation but also found that the proteasome, Cullin family members, kinases and transcriptional repressors played important roles in regulating apoptosis and cell cycle processes. These findings will help uncover intricate crosstalk between apoptosis and cell cycle mechanisms and thus provide further new revelations for possible anticancer drug discovery.

Acknowledgements

We are grateful to Dr. Huai-long Xu, Xin Li and Xiao Li (Sichuan University) for their kind help on this manuscript. This study was supported in part by National Natural Science Foundation of China (nos 81373311, 31300674, 81173093, 30970643 and J1103518), the Special Program for Youth Science and the Technology Innovative Research Group of Sichuan Province, China (No. 2011JTD0026).

Conflicts of interest

The authors have no conflicts of interest to declare.

References

- 1 Rastogi N, Mishra DP (2012) Therapeutic targeting of cancer cell cycle using proteasome inhibitors. *Cell Div.* **7**, 26.

- 2 Kasibhatla S, Tseng B (2003) Why target apoptosis in cancer treatment? *Mol. Cancer Ther.* **2**, 573–580.
- 3 Sherr CJ, McCormick F (2002) The RB and p53 pathways in cancer. *Cancer Cell* **2**, 103–112.
- 4 Senderowicz AM (2004) Targeting cell cycle and apoptosis for the treatment of human malignancies. *Curr. Opin. Cell Biol.* **16**, 670–678.
- 5 Alenzi FQ (2004) Links between apoptosis, proliferation and the cell cycle. *Br. J. Biomed. Sci.* **61**, 99–102.
- 6 Wen X, Lin ZQ, Liu B, Wei YQ (2012) Caspase-mediated programmed cell death pathways as potential therapeutic targets in cancer. *Cell Prolif.* **45**, 217–224.
- 7 Evan GI, Vousden KH (2001) Proliferation, cell cycle and apoptosis in cancer. *Nature* **411**, 342–348.
- 8 Kastan MB, Bartek J (2004) Cell-cycle checkpoints and cancer. *Nature* **432**, 316–323.
- 9 Hartwell LH, Hopfield JJ, Leibler S, Murray AW (1999) From molecular to modular cell biology. *Nature* **402**, C47–C52.
- 10 Liu Y, Zhao H (2004) A computational approach for ordering signal transduction pathway components from genomics and proteomics data. *BMC Bioinformatics* **5**, 158.
- 11 Sharan R, Ulitsky I, Shamir R (2007) Network-based prediction of protein function. *Mol. Syst. Biol.* **3**, 88.
- 12 Xia K, Dong D, Han JD (2006) IntNetDB v1.0: an integrated protein-protein interaction network database generated by a probabilistic model. *BMC Bioinformatics* **7**, 508.
- 13 Li X, Cai H, Xu J, Ying S, Zhang Y (2010) A mouse protein interactome through combined literature mining with multiple sources of interaction evidence. *Amino Acids* **38**, 1237–1252.
- 14 Xu Y, Hu W, Chang Z, Duanmu H, Zhang S, Li Z *et al.* (2011) Prediction of human protein-protein interaction by a mixed Bayesian model and its application to exploring underlying cancer-related pathway crosstalk. *J. R. Soc. Interface* **8**, 555–567.
- 15 Alfaro C, Andrade CE, Anthony K, Bahroos N, Bajec M, Bantoft K *et al.* (2005) The Biomolecular Interaction Network Database and related tools 2005 update. *Nucleic Acids Res.* **33**, D418–D424.
- 16 Kerrien S, Alam-Faruque Y, Aranda B, Bancarz I, Bridge A, Derow C *et al.* (2007) IntAct – open source resource for molecular interaction data. *Nucleic Acids Res.* **35**, D561–D565.
- 17 Keshava Prasad TS, Goel R, Kandasamy K, Keerthikumar S, Kumar S, Mathivanan S *et al.* (2009) Human protein reference database – 2009 update. *Nucleic Acids Res.* **37**, D767–D772.
- 18 Xenarios I, Salwinski L, Duan XJ, Higney P, Kim SM, Eisenberg D (2002) DIP, the database of interacting proteins: a research tool for studying cellular networks of protein interactions. *Nucleic Acids Res.* **30**, 303–305.
- 19 Winter AG, Wildenhain J, Tyers M (2011) BioGRID REST Service, BiogridPlugin2 and BioGRID WebGraph: new tools for access to interaction data at BioGRID. *Bioinformatics* **27**, 1043–1044.
- 20 Rhodes DR, Tomlins SA, Varambally S, Mahavisno V, Barrette T, Kalyana-Sundaram S *et al.* (2005) Probabilistic model of the human protein-protein interaction network. *Nat. Biotechnol.* **23**, 951–959.
- 21 Shen J, Zhang J, Luo X, Zhu W, Yu K, Chen K *et al.* (2007) Predicting protein-protein interactions based only on sequences information. *Proc. Natl. Acad. Sci. USA* **104**, 4337–4341.
- 22 Ng SK, Zhang Z, Tan SH (2003) Integrative approach for computationally inferring protein domain interactions. *Bioinformatics* **19**, 923–929.
- 23 Punta M, Coghill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C *et al.* (2012) The Pfam protein families database. *Nucleic Acids Res.* **40**, D290–D301.

- 24 Ge H, Liu Z, Church GM, Vidal M (2001) Correlation between transcriptome and interactome mapping data from *Saccharomyces cerevisiae*. *Nat. Genet.* **29**, 482–486.
- 25 Sims AH, Zweemer AJ, Nagumo Y, Faratian D, Muir M, Dodds M *et al.* (2012) Defining the molecular response to trastuzumab, pertuzumab and combination therapy in ovarian cancer. *Br. J. Cancer* **106**, 1779–1789.
- 26 Li D, Liu W, Liu Z, Wang J, Liu Q, Zhu Y *et al.* (2008) PRINCESS, a protein interaction confidence evaluation system with multiple data sources. *Mol. Cell. Proteomics* **7**, 1043–1052.
- 27 Zhang YN, Pan XY, Huang Y, Shen HB (2011) Adaptive compressive learning for prediction of protein-protein interactions from primary sequence. *J. Theor. Biol.* **283**, 44–52.
- 28 Swets JA, Pickett RM, Whitehead SF, Getty DJ, Schnur JA, Swets JB *et al.* (1979) Assessment of diagnostic technologies. *Science* **205**, 753–759.
- 29 He X, Zhang J (2006) Why do hubs tend to be essential in protein networks? *PLoS Genet.* **2**, e88.
- 30 Li M, Zhang H, Wang JX, Pan Y (2012) A new essential protein discovery method based on the integration of protein-protein interaction and gene expression data. *BMC Syst. Biol.* **6**, 15.
- 31 Lin CY, Chin CH, Wu HH, Chen SH, Ho CW, Ko MT (2008) Hubba: hub objects analyzer – a framework of interactome hubs identification for network biology. *Nucleic Acids Res.* **36**, W438–W443.
- 32 Przulj N, Wagle DA, Jurisica I (2004) Functional topology in a network of protein interactions. *Bioinformatics* **20**, 340–348.
- 33 Ostlund G, Lindskog M, Sonnhhammer EL (2010) Network-based Identification of novel cancer genes. *Mol. Cell. Proteomics* **9**, 648–655.
- 34 Zhang R, Lin Y (2009) DEG 5.0, a database of essential genes in both prokaryotes and eukaryotes. *Nucleic Acids Res.* **37**, D455–D458.
- 35 Wang N, Xu HL, Zhao X, Wen X, Wang FT, Wang SY *et al.* (2012) Network-based identification of novel connections among apoptotic signaling pathways in cancer. *Appl. Biochem. Biotechnol.* **167**, 621–631.
- 36 Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC *et al.* (2011) pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* **12**, 77.
- 37 Ola MS, Nawaz M, Ahsan H (2011) Role of Bcl-2 family proteins and caspases in the regulation of apoptosis. *Mol. Cell. Biochem.* **351**, 41–58.
- 38 Wyllie AH (2010) “Where, O death, is thy sting?” A brief review of apoptosis biology. *Mol. Neurobiol.* **42**, 4–9.
- 39 Malumbres M, Barbacid M (2009) Cell cycle, CDKs and cancer: a changing paradigm. *Nat. Rev. Cancer* **9**, 153–166.
- 40 da Huang W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57.
- 41 Zhang W, Lee JC, Kumar S, Gowen M (1999) ERK pathway mediates the activation of Cdk2 in IGF-1-induced proliferation of human osteosarcoma MG-63 cells. *J. Bone Miner. Res.* **14**, 528–535.
- 42 Magnaghi P, D’Alessio R, Valsasina B, Avanzi N, Rizzi S, Asa D *et al.* (2013) Covalent and allosteric inhibitors of the ATPase VCP/p97 induce cancer cell death. *Nat. Chem. Biol.* **9**, 548–556.
- 43 Vij N (2008) AAA ATPase p97/VCP: cellular functions, disease and therapeutic potential. *J. Cell Mol. Med.* **12**, 2511–2518.
- 44 Voorhees PM, Dees EC, O’Neil B, Orlowski RZ (2003) The proteasome as a target for cancer therapy. *Clin. Cancer Res.* **9**, 6316–6325.
- 45 Frankland-Searby S, Bhaumik SR (2012) The 26S proteasome complex: an attractive target for cancer therapy. *Biochim. Biophys. Acta* **1825**, 64–76.
- 46 Shah SA, Potter MW, McDade TP, Ricciardi R, Perugini RA, Elliott PJ *et al.* (2001) 26S proteasome inhibition induces apoptosis and limits growth of human pancreatic cancer. *J. Cell. Biochem.* **82**, 110–122.
- 47 Driscoll JJ, Woodle ES (2012) Targeting the ubiquitin+proteasome system in solid tumors. *Semin. Hematol.* **49**, 277–283.
- 48 Lee J, Zhou P (2010) Cullins and cancer. *Genes Cancer* **1**, 690–699.
- 49 Confalonieri S, Quarto M, Goisis G, Nuciforo P, Donzelli M, Jodice G *et al.* (2009) Alterations of ubiquitin ligases in human cancer and their association with the natural history of the tumor. *Oncogene* **28**, 2959–2968.
- 50 Zheng Q, Zhao L-Y, Kong Y, Nan K-J, Yao Y, Liao Z-J (2013) CDK-associated Cullin 1 can promote cell proliferation and inhibit cisplatin-induced apoptosis in the AGS gastric cancer cell line. *World J. Surg. Oncol.* **11**, 1–11.
- 51 Bai J, Zhou Y, Chen G, Zeng J, Ding J, Tan Y *et al.* (2011) Overexpression of Cullin1 is associated with poor prognosis of patients with gastric cancer. *Hum. Pathol.* **42**, 375–383.
- 52 Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S (2002) The protein kinase complement of the human genome. *Science* **298**, 1912–1934.
- 53 Payton M, Chung G, Yakowec P, Wong A, Powers D, Xiong L *et al.* (2006) Discovery and evaluation of dual CDK1 and CDK2 inhibitors. *Cancer Res.* **66**, 4299–4308.
- 54 Knippschild U, Gocht A, Wolff S, Huber N, Löhler J, Stöter M (2005) The casein kinase 1 family: participation in multiple cellular processes in eukaryotes. *Cell. Signal.* **17**, 675–689.
- 55 Ahmad KA, Wang G, Unger G, Slaton J, Ahmed K (2007) Protein kinase CK2 – a key suppressor of apoptosis. *Adv. Enzyme Regul.* **48**, 179–187.
- 56 Verheyen EM, Gottardi CJ (2010) Regulation of Wnt/beta-catenin signaling by protein kinases. *Dev. Dyn.* **239**, 34–44.
- 57 Voronkov A, Krauss S (2013) Wnt/beta-catenin signaling and small molecule inhibitors. *Curr. Pharm. Des.* **19**, 634–664.
- 58 Wu S, Chen L, Becker A, Schonbrunn E, Chen J (2012) Casein kinase 1alpha regulates an MDMX intramolecular interaction to stimulate p53 binding. *Mol. Cell. Biol.* **32**, 4821–4832.
- 59 Hellwig CT, Ludwig-Galezowska AH, Concannon CG, Litchfield DW, Prehn JH, Rehm M (2010) Activity of protein kinase CK2 uncouples Bid cleavage from caspase-8 activation. *J. Cell Sci.* **123**, 1401–1406.
- 60 Dixit D, Sharma V, Ghosh S, Mehta VS, Sen E (2012) Inhibition of Casein kinase-2 induces p53-dependent cell cycle arrest and sensitizes glioblastoma cells to tumor necrosis factor (TNFalpha)-induced apoptosis through SIRT1 inhibition. *Cell Death Dis.* **3**, e271.
- 61 Campos B, Bermejo JL, Han L, Felsberg J, Ahmadi R, Grabe N *et al.* (2011) Expression of nuclear receptor corepressors and class I histone deacetylases in astrocytic gliomas. *Cancer Sci.* **102**, 387–392.
- 62 Galanis E, Jaeckle KA, Maurer MJ, Reid JM, Ames MM, Hardwick JS *et al.* (2009) Phase II trial of vorinostat in recurrent glioblastoma multiforme: a north central cancer treatment group study. *J. Clin. Oncol.* **27**, 2052–2058.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Fig. S1. ROC curve of our Naïve Bayesian classifier.

Fig. S2. TP/FP ratios at different cut-offs. When cut-off is 160, TP/FP has the maximum value, regarded as the optimal threshold.

Table S1. Likelihood ratios for individual data sets to infer human apoptotic/cell cycle PPIs.

Table S2. Pearson correlation coefficients between evidence sources on the GSP and the GSN data set.

Table S3. Sensitivity, specificity and accuracy under different cut-off values.

Table S4. Information of the top 100 proteins predicted by PeC and degree.