THE PROGRESS OF PROSTATE CANCER IN PATHWAY LEVEL EXPLORED BY PROTEIN NETWORK WITH GENE EXPRESSION

Fei-Hung Hung, Hung-Wen Chiu
Graduate Institution of Biomedical Informatics, Taipei Medical University
250 Wu-Hsing Street, Taipei/Taiwan
hwchiu@tmu.edu.tw

ABSTRACT
Biological pathways are the crucial biological mechanisms in living cells. The huge volume of genomics and proteomics data requires computational methods for predicting or reconstructing pathways. Thus, the application of protein-protein interaction (PPI) or gene expression methods is insufficient to discover meaningful pathways. The integration of PPIs and gene profiles is a better approach to uncover the regulation of pathway and must be utilized well. Previous studies on this topic only focus on the gene level or some limited local groups. This study presents an approach to finding potential fragments of active pathways around known pathways between the various stages of diseases. The proposed method used a maximum score-based function that integrates genomics and proteomics information. This method quantified the strength of gene expression change and the degree of protein-protein interactions to illustrate global status as pathway maps. In this study, we use prostate cancer data as an example to explain which potential fragments of pathway co-constructed a pathway map of prostate cancer at different disease statuses. The resulting map shows a possible correspondence between known pathway and cancer-related genes that are not on the known pathway. Comparing distinct status pathway map reveals a global change of different disease states pathway level. The pathway map of different disease statuses can provide more insight in the progress of cancer.

KEY WORDS
Biological Pathway, Cancer-Related Gene, Gene Expression, Maximum Score-based Function

1. Introduction

Bioinformatics has benefitted greatly from advances in computer science and biology laboratory techniques, an era of rapid accumulation of genomic and proteomic information. For example, the Gene Expression Omnibus (GEO) is one of several public genomic data repositories [1]. The GEO includes 12,211 platforms, 1,024,125 samples, 42,673 series and 3,413 datasets. Computational biology methods can help researchers obtain a better understanding of complex systems (e.g., protein-protein interaction network, regulatory pathways or cancer mechanisms). A signal transduction pathway is a main respondence for extracellular excitement. When signal pathways are involved in activating apoptosis, cell cycle, or proliferation, they have a comprehensive effect on upstream/downstream relationships between interacting proteins/genes. The widely used pathway database is the Kyoto Encyclopedia of Genes and Genomes (KEGG). The KEGG is a database that integrates genomic, chemical, and systemic functional information [2]. The KEGG currently includes 275,060 pathways. Researchers can access these online resources easily through their web-based interface.

Early pathway prediction methods, such as PathFinder, were based only on PPIs. PathFinder is a tool for finding potential pathways [3] that maps GO annotations onto the PPI network and applies the association rule method to identify pathways with high confidence. The recall rate is 78% and precision rate is 40%. When researchers investigate the importance of gene regulation, they often used PPIs and gene expression data to reconstruct some simple signaling networks [4-6]. One method, NetSearch, tried to integrate PPI and gene expression [4]. This approach used gene expression data to cluster proteins and scored protein by clustering, and was capable of reconstructing MAPK signal pathways. The recall rate for this approach is 44%, with a precision rate of 24%. Ruth *et al*. built PathwayOracle Toolkit. This toolkit applies the STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) method [7] to score PPI data, and then adopts Eppstein's k-shortest algorithm is used for pathway prediction [8]. Some of the methods mentioned above only use PPI, which does not sufficiently represent the entire pathway, and some are limited to reconstructing specific species. Even approaches that consider gene-level data only apply that data for clustering, and fail to exhibit true gene expression values.

Researchers have recently identified many disease markers by analyzing genome-wide and proteomic-wide information. However, investigators have shown that many well-known risk factors may be partial emphases rather than global mechanisms of disease. To identify a marker for more complete performance of disease is a challenging. A sub-network marker is more reproducible than individual marker genes selected without network information [9]. Most previous methods cannot identify molecular changes and relationships on the environmental side.

DOI: 10.2316/J.2014.216.818-0036

In the past researches of prostate cancer, Yu *et al*. tested a comprehensive gene expression analysis on 152 human samples and compared with normal neighbouring prostate tissues to confirm an alteration of gene expression in prostate cancer [10]. Chandran *et al*. analyzed Affymetrix oligonucleotide arrays and their results shows that 415 genes are up-regulated and 364 genes are down-regulated in metastatic prostate tumor [11].

Some studies report genes that are not included in published pathways as cancer-related genes [12-15]. Thus, researchers must locate the crux of pathways and their environment and apply gene-wide and protein-wide data to find the relationship between those genes and published pathways. Some activated pathways cut across the published pathways from those cancer-related genes that are not emphasized on the published pathways. Therefore, a significant change in gene level is needed, and the character of cancer can match this requirement. Cancer is strongly associated with defects in signal transduction pathways. In cancer tissue, the function of pathways is uncontrolled and inappropriate. When a gene shows a significant change, an activated pathway across this gene allows researchers to infer a pathway from here, even if they do not know which pathway is activated.

2. Materials and Methods

This study involves the collection of three kinds of data. Protein-protein interaction data is used for network construction. Protein location information prevents an impossible interaction. Finally, gene expression profile reveals the strength of change. All of the proteins/genes used in this study were normalized to a specific symbol using data downloaded from the Uniprot [16].

To construct a PPI network, protein-protein interaction information was collected from the Interologous Interaction Database (I2D) [17]. The data is combined from 6 commonly used PPI databases (BIND, BioGrid, HPRD, INNATEDB, IntAct, and MINT). We filtered out specific PPIs (e.g., experimental or predicted data). The remaining PPIs are non-redundant PPIs and the number exceeds 70,000.

To avoid interactions that do not naturally exist, this study follows the basic protein targeting pathways to remove them. It means that all reactions in the results can really happen in cell. The real reactions happen between cytoplasm and nucleus, cytoplasm and mitochondria, cytoplasm and endoplasmic reticulum, cytoplasm and chloroplast, cytoplasm and peroxisome, endoplasmic reticulum and golgi apparatus, lysosome and golgi apparatus, secretory vesicles and golgi apparatus, plasma membrane and golgi apparatus, plasma membrane and secretory vesicles, plasma membrane and endosome, and lysosome and endosome.

A change on gene expression is applied to locate pathways on the PPI network. More severe changes are needed, and we collected gene expression data of cancer. Many tumors' gene data samples are available in the GEO data set. This study uses prostate cancer (GDS2545), which is a metastatic prostate tumors and primary prostate tumors (Affymetrix Human Genome U95 Version 2 Array) that includes 12,625 identifiers/genes. That study is that normal tissue adjacent to the tumor and normal donor tissue also examined. Specifically, metastasis reflects the most adverse clinical outcome and provides insight into the molecular mechanisms underlying the metastatic process. This dataset is from 18 donor and 64 primary prostate tumor samples. The stages are divided into four stages. The four stages are "Stage 1: normal prostate tissue", "Stage 2: normal prostate adjacent to tumor", "Stage 3: primary prostate tumor" and "Stage 4: metastatic prostate tumor". According the header description of GDS2545, these samples are divided into three groups: "normal prostate adjacent to tumor versus normal prostate tissue", "primary prostate tumor versus normal prostate adjacent to tumor" and "metastatic prostate tumor versus primary prostate tumor". Next, we used a self-developed tool to analyze these three groups based on a function of R package. Stage 1 includes 18 GSMs. Stage 2 includes 63 GSMs. Stage 3 includes 65 GSMs. Stage 4 includes 25 GSMs. Table 1 shows the list of all stages.

Table 1
Classified GSMs from GDS2545

GDS2545: Metastatic prostate cancer (HG-U95A)	
Stage	List of GSMs
Stage 1	GSM152804,GSM152805,GSM152806,GSM152807,GSM152808, GSM152809,GSM152810,GSM152811,GSM152812,GSM152813, GSM152814,GSM152815,GSM152816,GSM152817,GSM152818, GSM152819,GSM152820,GSM152821
Stage 2	GSM153115,GSM153116,GSM153117,GSM153118,GSM153119, GSM153120,GSM153121,GSM153122,GSM153123,GSM153124, GSM153125,GSM153126,GSM153127,GSM153128,GSM153129, GSM153130,GSM153131,GSM153132,GSM153133,GSM153134, GSM153135,GSM153136,GSM153137,GSM153138,GSM153139, GSM153140,GSM153141,GSM153142,GSM153143,GSM153144, GSM153145,GSM153146,GSM153147,GSM153148,GSM153149, GSM153150,GSM153151,GSM153152,GSM153153,GSM153154, GSM153155,GSM153156,GSM153157,GSM153158,GSM153159, GSM153160,GSM153161,GSM153162,GSM153163,GSM153164, GSM153165,GSM153166,GSM153167,GSM153168,GSM153169, GSM153170,GSM153171,GSM153172,GSM153173,GSM153174, GSM153175,GSM153176,GSM153177
Stage 3	GSM152931,GSM152932,GSM152933,GSM152934,GSM152935, GSM152936,GSM152937,GSM152938,GSM152939,GSM152940, GSM152941,GSM152942,GSM152943,GSM152944,GSM152945, GSM152946,GSM152947,GSM152948,GSM152949,GSM152950, GSM152951,GSM152952,GSM152953,GSM152954,GSM152955, GSM152956,GSM152957,GSM152958,GSM152959,GSM152960, GSM152961,GSM152962,GSM152963,GSM152964,GSM152965, GSM152966,GSM152967,GSM152968,GSM152969,GSM152970, GSM152971,GSM152972,GSM152973,GSM152974,GSM152975, GSM152976,GSM152977,GSM152978,GSM152979,GSM152980, GSM152981,GSM152982,GSM152983,GSM152984,GSM152985, GSM152986,GSM152987,GSM152988,GSM152989,GSM152990, GSM152991,GSM187524,GSM187525,GSM187526,GSM187527
Stage 4	GSM152856,GSM152857,GSM152858,GSM152859,GSM152860, GSM152861,GSM152862,GSM152863,GSM152864,GSM152865, GSM152866,GSM152867,GSM152868,GSM152869,GSM152870, GSM152871,GSM152872,GSM152873,GSM152874,GSM152875, GSM152876,GSM152877,GSM152878,GSM152879,GSM152880

The Wilcoxon rank-sum test (also called the Mann–Whitney U test) is applied to identify the expression change of a gene between different statuses. When the *p*-value is

less than 0.05, the expression change of a gene between two statuses, original status and developed status, is considered to be significant. A gene with a significant change on gene expression is named as a locating gene/point. This is because one or more activated pathway crosses that point in the developed status. Therefore, we should find fixed-searching-depth fragments of pathway from a locating point. We suggest that an activated pathway could exhibit more severe change than inactivated pathways at the same situation. Thus, the strength of change should be calculated. This study uses a parameter GCS (Gene Expression Change Score) to measure the strength of gene expression change. The GCS equation is defined as follows (1):

$$GCS_i = (1 - p_i)\ 100/n\ , if\ n = {}_i 0, GCS = {}_i 0$$

here n_i is the total interaction number of gene i on the PPI network (non-loops). To decide the strength of edge between gene i and gene k, the score equation is defined as *ECS* (Edge Change Score) (2):

$$ECS_{ik} = (GCS_i + GCS_k)/2$$

An *ECS* is the average of two *GCS*s in an interaction. When we determine how to measure the strength of edge's change, the strength of pathway fragment's change in a fixed searching depth could be calculated as *PCS* (Pathway Change Score) (3):

$$PCS = \frac{\sum ECS}{Searching\ Depth}$$

A *PCS* is the average of all *ECS*s in a merged-fragment subnetwork. A searching-depth x of fragment includes x+1 nodes/genes. To rank the *PCS*s, the highest scores are considered as potential pathway fragments. In the study, we adopt searching-depth 2 because a fragment including at most 5 nodes/genes (length 5) can be created from these results. Results show that every locating point leads to thousands of fragments that include some loops and two-way fragments. All top 5 fragments (non-loops and non-redundancies) were merged into a subnetwork. Merging these subnetworks produces the final results, which is a pathway map. The proposed process involves several scoring and searching steps, as illustrated in Figure 1.

3. Results

This study uses prostate cancer data (KEGG pathway map hsa05215 and GEO GDS2545) to develop and test the proposed method. According to the available data, three groups (A: Stage 1 versus Stage 2, B: Stage 2 versus Stage 3 and C: Stage 3 versus Stage 4) were produced from GDS2545.

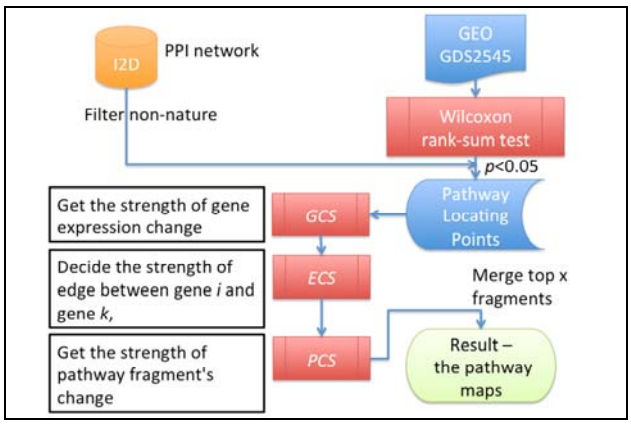

Figure 1. Schematic overview of this study.

For the Wilcoxon rank-sum test, if the *p*-value of the change of gene expression between original status and developed status is less than 0.05, it is considered as significance. The group A contains 1,485 genes with a significant change in gene expression. The group B contains 2,619 genes with a significant change in gene expression. The group C contains 4,380 genes with a significant change in gene expression. For group A, 17 significant protein families can be mapped to the KEGG prostate cancer map. For group B, 24 significant protein families can be mapped to the KEGG prostate cancer map. For group C, 48 significant protein families can be mapped to the KEGG prostate cancer map. Table 2 shows the locating points on the KEGG prostate cancer map at these three groups. The next step calculates the gene expression change score *GCS* of each point.

To understand which pathway is active, the edge expression change score *ECS* is defined to score the intensity of change in the link/edge between normal and other situations. The number of pathway fragments that need to be scored depends on the degree of interaction. From the distributions of fragment score, rare fragments that pass the locating points have strong reaction.

For merging top-x score fragments, three pathway maps of these groups for prostate cancer are shown in Figure 2. (Pathway map A), Figure 3. (Pathway map B) and Figure 4. (Pathway map C). The top score fragments are calculated based on the pathway change score *PCS*. Apart from finding the potential pathways of complicity in cancer, this approach also shows the progress of cancer in pathway level. The resulting pathway map shows some interesting results in the hubs of nodes that belong to the KEGG prostate cancer map. Specifically, 7 genes appear as hubs in the pathway map A (AURKA, EPRS, HSPA9, MAP3K7, MLST8, NR2C2, and RAF1). 4 hubs appear in the environment pathway B (CTNNB1, EGFR, FGF2, and PLCG1). Furthermore, 17 hubs appear in the environment pathway C (AKT1, CALCOCO1, CREB3L4, CTNNB1, CTSD, IDE, MLST8, NR2C2, PIK3CG, PPP4C, PRKCD, RAF1, RB1, SFN, SUMO3, TANK and ZBTB17).

4. Conclusion

We proposed a heuristic method to measure the change of pathway expression. We pointed out that only PPIs or gene expressions are not enough for pathway inference. This method could solve the problem to integration of protein-protein interaction and gene expression. A series of computational test was conducted to show that our algorithm could draw pathways maps with stage status and showed the progress of cancer.

As for group A, pathway map in Figure 2 shows the progress from normal to tumor. AURKA, EPRS, HSPA9, MAP3K7, MLST8, NR2C2, and RAF1 appear on the first pathway map. Among these, AURKA, EPRS, MAPK3K7 and RAF1 could be found on the related page of GeneCards. In the pathway map for group B (see Figure 3), CTNNB1 and EGFR could be found on the related page of GeneCards. In the pathway map for group C (see Figure 4), AKT1, CREB3L4, CTNNB1, PIK3CG, RAF1, RB1 and SFN) could be found on the related page of GeneCards. It proves these genes are associated with prostate cancer to a certain degree. As for other genes that are not show the association to prostate cancer on the GeneCards, they have to be proved by further works.

5. Discussion

Via three result pathway maps, different genes play important roles at different disease stages respectively. The number of genes with "significant gene expression change" grows. The direction is from membrane to nuclear. The proposed method integrates gene expression data and protein-protein interactions for pathway research. This approach uses quantitative identification to find the fragments of activated pathways and construct the neighbourhood around known pathways. This study reveals the role and importance of the neighbourhood around cancer pathways. In the global pathway maps the results show the potential relationships of cancer-related genes that do not appear on the known pathway map. These relationships provide a possible approach to find potential and unknown cancer-related genes.

Acknowledgement

This study was financially supported by NSC 102-2221-E-038-003 of Taiwan.

References

[1] T. Barrett, D. Troup, S. Wilhite, P. Ledoux, C. Evangelista, I. Kim, M. Tomashevsky, K. Marshall, K. Phillippy, P. Sherman, R. Muertter, M. Holko, O. Ayanbule, A. Yefanov & A. Soboleva, NCBI GEO: archive for functional genomics data sets—10 years on, *Nucleic Acids Res., 39*(Database issue), 2011, D1005–D1010.

[2] M. Kanehisa, S. Goto, S. Furumichi, M. Tanabe & M. Hirakawa, KEGG for representation and analysis of molecular networks involving diseases and drugs, *Nucleic Acids Res., 38*(Database issue), 2010, D355–D360.

[3] G. Bebek & J. Yang, PathFinder: mining signal transduction pathway segments from protein-protein interaction networks, *BMC Bioinformatics, 8*, 2007, 335–350.

[4] M. Steffen, A. Petti, J. Aach, P. D'haeseleer & G. Church, Automated modelling of signal transduction networks, *BMC Bioinformatics, 3*, 2002, 34–44.

[5] Y. Liu & H. Zhao, A computational approach for ordering signal transduction pathway components from genomics and proteomics Data, *BMC Bioinformatics, 5*, 2007, 158.

[6] J. Scott, T. Ideker, R. Karpa & R. Sharan, Efficient algorithms for detecting signaling pathways in protein interaction networks, *J Comput Biol., 13*(2), 2006, 133–144.

[7] D. Szklarczyk, A. Franceschini, M. Kuhn, M. Simonovic, A. Roth & P. Minguez, The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored, *Nucleic Acids Res., 39*(Database issue), 2010, D561–D568.

[8] D. Ruths, J. Tseng, L. Nakhleh & P. Ram, De Novo Signaling Pathway Predictions Based on Protein-Protein Interaction, Targeted Therapy and Protein Microarray Analysis, *LNCS, 4532*, 2007, 108–118.

[9] H. Chuang, E. Lee, Y. Liu, D. Lee & T. Ideker, Network-based classification of breast cancer metastasis, *Mol. Syst. Biol., 3*, 2007, 140.

[10] YP. Yu, D. Landsittel, L. Jing, J. Nelson, B. Ren, L. Liu, C. McDonald, R. Thomas, R. Dhir, S. Finkelstein, G. Michalopoulos, M. Becich & JH. Luo, Gene expression alterations in prostate cancer predicting tumor aggression and preceding development of malignancy, *J Clin. Oncol., 22*(14), 2004, 2790-2779.

[11] UR. Chandran, C. Ma, R. Dhir, M. Bisceglia, M. Lyons-Weiler, W. Liang, G. Michalopoulos, M. Becich & FA. Monzon, Gene expression profiles of prostate cancer reveal involvement of multiple molecular pathways in the metastatic process, *BMC Cancer, 7*(64), 2007.

[12] M. Thomasson, H. Hedman, T. Junttila, K. Elenius, B. Ljungberg & R. Henriksson, ErbB4 is downregulated in renal cell carcinoma–a quantitative RT-PCR and immunohistochemical analysis of the epidermal growth factor receptor family, *Acta Oncol., 43*(5), 2004, 453–459.

[13] T. Behbahani, C. Thierse, C. Baumann, D. Holl, P. Bastian, A. von Ruecker, S. M¨uller, J. Ellinger & S. Hauser, Tyrosine kinase expression profile in clear cell renal cell carcinoma, *World J Urol.,* [Epub ahead of print].

[14] T. Shi, L. Liou, P. Sadhukhan, Z. Duan, A. Novick, J. Hissong, A. Almasan & J. DiDonato, Effects of resveratrol on gene expression in renal cell carcinoma, *Cancer Biol. Ther., 3*(9), 2004, 882–888.

[15] J. Copland, B. Luxon, L. Ajani, T. Maity, E. Campagnaro, H. Guo, S. LeGrand, P. Tamboli & C. Wood, Genomic profiling identifies alterations in TGFbeta signaling through loss of TGFbeta receptor expression in human renal cell carcinogenesis and progression, *Oncogene, 22*(39), 2003, 8053–8062.

[16] M. Magrane & U. Consortium, UniProt Knowledgebase: a hub of integrated protein data, *Database (Oxford), 2011*, 2011, bar009.

[17] K. Brown & I. Jurisica, Unequal evolutionary conservation of human protein interactions in interologous networks, *Genome Biol., 8*(5), 2007, R95.

[18] F. Belinky, I. Bahir, G. Stelzer, S. Zimmerman, N. Rosen, N. Nativ, I. Dalah, T. Iny Stein, N. Rappaport, T. Mituyama, M. Safran & D. Lancet, Non-redundant compendium of human ncRNA genes in GeneCards, *Bioinformatics, 29*(2), 2013 255-61.

Table 2
The test result of locating points at three groups of GDS2545

Family	Member or Alias Name	Group A		Group B		Group C	
		Significant	p-value	Significant	p-value	Significant	p-value
GSTP1	DFN7, FAEES3, GST3, GSTP, GSTP1 , PI			GSTP1	0.000003	GSTP1	0.000000003076
NKX3-1	BAPX2; NKX3; NKX3.1; NKX3A						
PTEN	10q23del; BZS; CWS1; DEC; GLM2; MHAM; MMAC1; PTEN1; TEP1					TEP1	0.000001723
GF	EGF, PDGFA, PDGFB, INS, PDGFC_D, IGF1, TGFA,	PDGFA	0.00493	IGF1	0.006689	INS	0.00003993
						IGF1	0.00002488
GFR	EGFR, ERBB1, FGFR1, PDGFRA, ERBB2, HER2, INSRR, IGF1R, PDGFRB, FGFR2	FGFR2	0.002864	FGFR1	0.001295	EGFR	0.006172
						ERBB2	0.000000005907
				FGFR2	0.0000217	IGF1R	0.0009311
						FGFR2	0.0000009558
PI3K	PIK3CA, PIK3CB, PIK3CD, PIK3CG, PIK3R1, PIK3R2, PIK3R3, PIK3R5, PIK3C2A, PIK3C2B, PIK3C2G			PIK3CG	0.007083	PIK3CA	0.00005876
						PIK3R1	0.000006806
				PIK3R2	0.0474	PIK3C2A	0.002192
						PIK3C2B	0.0436
PDK1	PDPK1; PDK1; PDPK2; PRO0461	PDPK1	0.02111			PDK1	0.003776
PKB/Akt	AKT1, AKT2, AKT3			AKT2	0.0142	AKT2	0.0335
						AKT3	0.001146
SRD5A2	SRD5A2, MGC138457						
Grb2	GRB2, ASH; EGFRBP-GRB2; Grb3-3; MST084; MSTP084; NCKAP2	GRB2	0.009916	GRB2	0.01819	GRB2	0.000000000001861
SOS	SOS1, SOS2	SOS1	0.02208	SOS2	0.02002		
Ras	HRAS; C-BAS/HAS; C-H-RAS; C-HA-RAS1; CTLO; H-RASIDX; HAMSV; HRAS1; KRAS; NRAS; RASH1	HRAS	0.01225	KRAS	0.009854	HRAS	0.00003995
						KRAS	0.02608
Raf	BRAF, RAF1, ARAF, ARAF1, CRAF, PKS2; RAFA1					ARAF	0.0000000009557
MEK1	MAP2K1, MEK1, MAPKK1 , MKK1; PRKMK1,	MAP2K1	0.02523				
MEK2	MAP2K2, MEK2, MAPKK2, MKK2; PRKMK2	MAP2K2	0.001385	MAP2K2	0.005339	MAP2K2	0.00005548
ERK	MAPK1, MAPK2, MAPK3, ERK-1; ERK1; ERT2, PRKM1; PRKM2, PRKM3					MAPK1	0.01324
AR	AR, AIS; DHTR; HUMARA; HYSP1; KD; NR3C4; SBMA; SMAX1; TFM			AR	0.0032	AR	0.000000000004022
HSP	htpG, HSP90A, HSP90B, HSP90B1, TRA1, ECGP; GP96; GRP94					HSP90B1	0.00000000003789
Casp9	CASP9, APAF-3; APAF3; CASPASE-9c; ICE-LAP6; MCH6; PPP1R56	CASP9	0.04144				
BAD	BAD, BBC2, BBC6, BCL2L8					BAD	0.03744
FKHR	FOXO1, FKH1; FKHR; FOXO1A			FOXO1	0.006002	FOXO1	0.00007959
p21	CDKN1A, P21, CIP1, CAP20; CDKN1; MDA-6; SDI1; WAF1; p21CIP1			CDKN1A	0.03011		
p27	CDKN1B, P27, KIP1, CDKN4, MEN1B; MEN4; P27KIP1						
MDM2	MDM2, ACTFS; HDMX; hdm2			MDM2	0.0007657		
GSK3	GSK3A, GSK3B			GSK3A	0.000428	GSK3B	0.000000000541
				GSK3B	0.006835		
IKKA	IKBKA, IKKA, CHUK	CHUK	0.00003134			CHUK	0.003317
IKKB	IKBKB, IKKB	IKBKB	0.03559			IKBKB	0.03504
IKKG	IKBKG, IKKG, NEMO			IKBKG	0.00163		
mTOR	MTOR, FRAP, FRAP1; FRAP2; RAFT1; RAPT1					MTOR	0.000002106
CREB1	CREB1					CREB1	0.00007958
CREB2	ATF4, CREB2	ATF4	0.00002567				
CREB3	CREB3, LUMAN; LZIP					CREB3	0.000002061
CREB5	CREB5; CREBPA					CREB5	0.03238
CREB3L1	CREB3L1			CREB3L1	0.0002449	CREB3L1	0.002783
CREB3L2	CREB3L2	CREB3L2	0.003839			CREB3L2	0.003317
CREB3L3	CREB3L3						
CREB3L4	CREB3L4						
β- Catenin	CTNNB1, CTNNB; MRD19; armadillo						
IkB	NFKBIA, IKBA; MAD-3; NFKBI	NFKBIA	0.01786				
NFkB	NFKB1, RELA, NFKB2, RELB, REL	RELB	0.006975			NFKB1	0.00001596
						RELA	0.03504
						REL	0.02375
CDK2	CDK2, p33					CDK2	0.00003416
cyclin E	CCNE, CCNE1, CCNE2, CYCE2					CCNE2	0.00003416
Rb	RB1, RB; pRb; OSRC; pp110; p105-Rb						
E2F	E2F1, E2F2, E2F3			E2F2	0.01107	E2F2	0.03911
						E2F3	0.000000166
CBP	EP300, CREBBP, KAT3, CBP, RSTS					CREBBP	0.02731
p53	TP53, P53, BCC7; LFS1; TRP53			TP53	0.03197	TP53	0.001042
TCF/LEF	TCF7, TCF7L1, TCF7L2, LEF1,			TCF7L2	0.0002974	TCF7	0.005841
				LEF1	0.005033	TCF7L2	0.003514
						LEF1	0.03955
cyclin D1	CCND1, BCL1; D11S287E; PRAD1; U21B31	CCND1	0.000001195			CCND1	0.001473
BCL2	BCL2, PPP1R50						
PSA	KLK3, APS; KLK2A1; PSA; HK3	KLK3	0.03711	KLK3	0.04634	HK3	0.04221

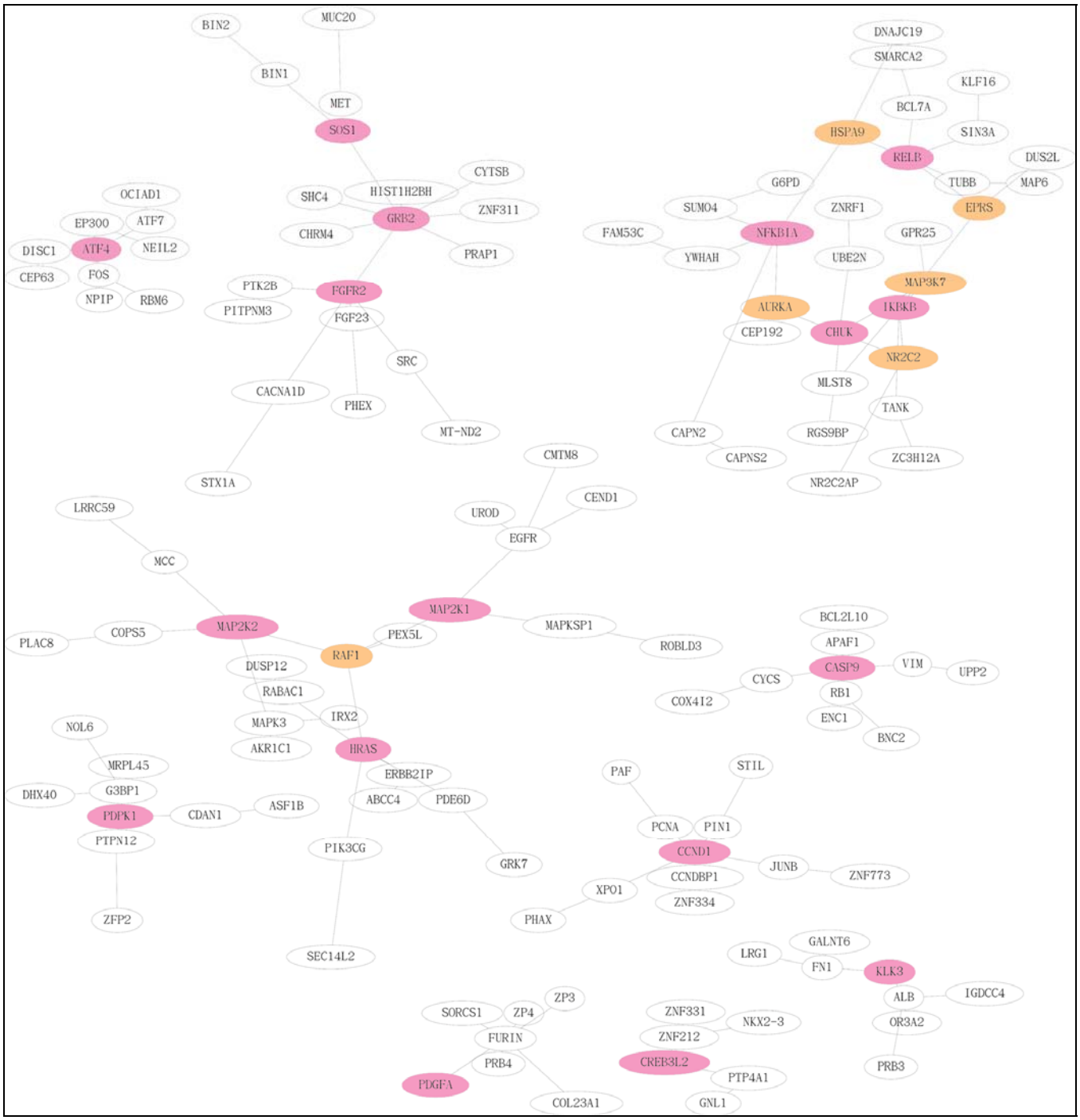

Figure 2. The pathway map A for group A. This map shows the situation from normal to tumor in prostate cancer. The pink ellipses show a node in KEGG prostate cancer pathway map with significant change in gene expression at group A. The orange ellipses show a node in KEGG prostate cancer pathway map without significant change in gene expression at group A, but a hub for pink ellipses in this map.

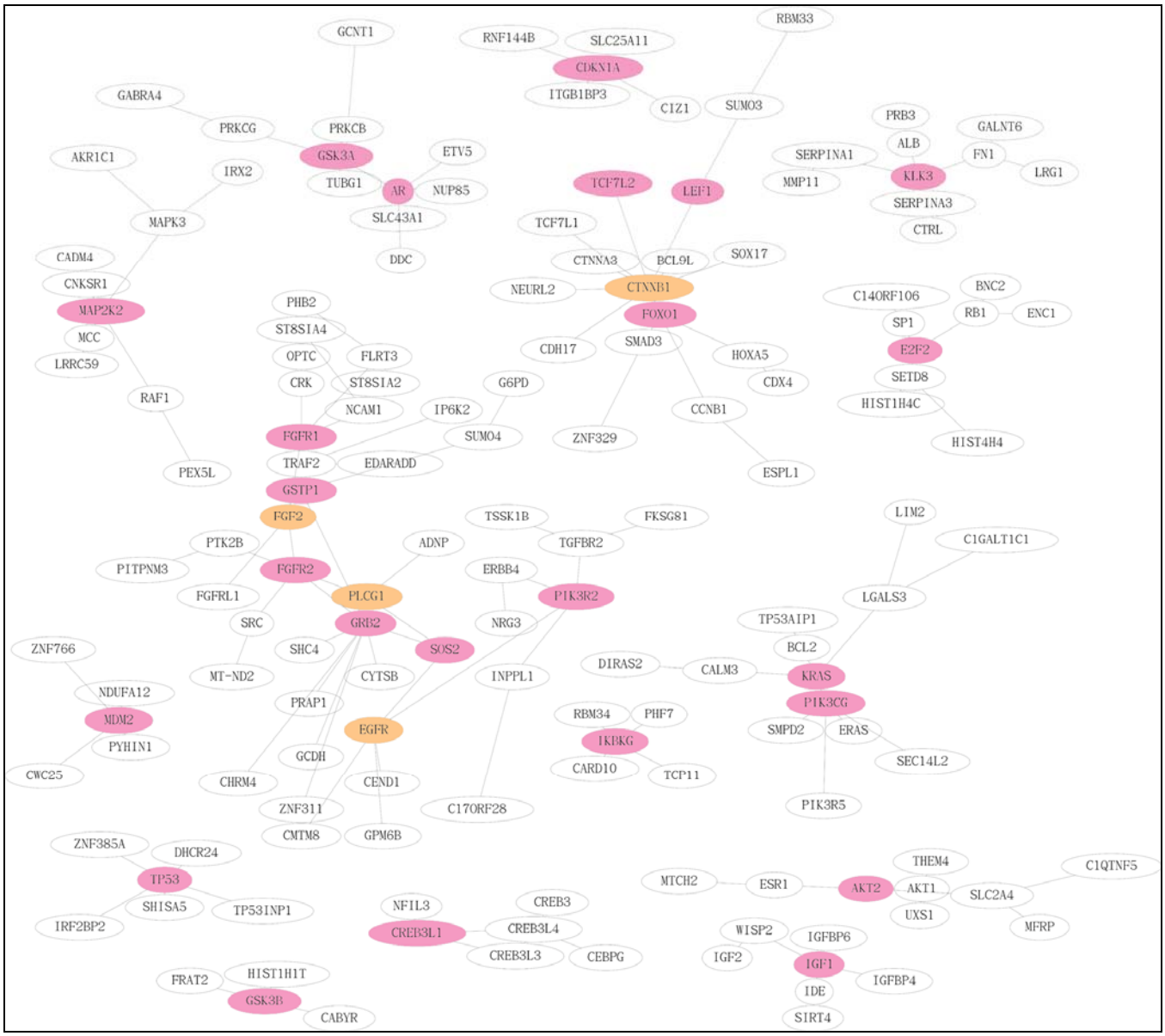

Figure 3. The pathway map B for group B. This map shows the situation from early prostate tumor to primary prostate tumor. The pink ellipses show a node in KEGG prostate cancer pathway map with significant change in gene expression at group B. The orange ellipses show a node in KEGG prostate cancer pathway map without significant change in gene expression at group B, but a hub for pink ellipses in this map.

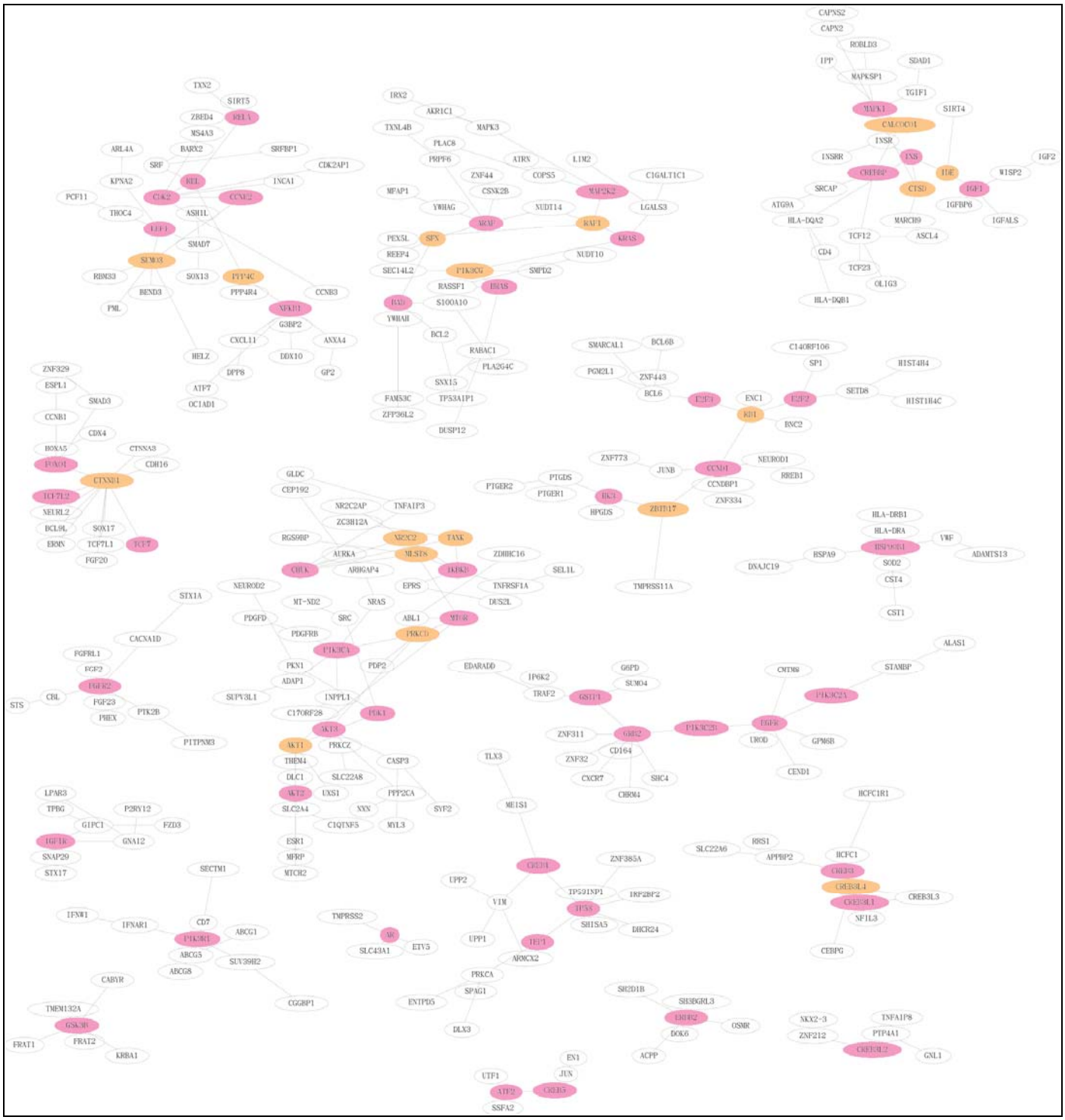

Figure 4. The pathway map C for group C. This map shows the situation from primary prostate tumor to metastatic prostate tumor. The pink ellipses show a node in KEGG prostate cancer pathway map with significant change in gene expression at group C. The orange ellipses show a node in KEGG prostate cancer pathway map without significant change in gene expression at group C, but a hub for pink ellipses in this map.