European Review for Medical and Pharmacological Sciences 2016; 20: 2532-2540

Investigating key genes associated with ovarian cancer by integrating affinity propagation clustering and mutual information network analysis

J. WANG[1], C. CHEN[2], H.-F. LI[2], X.-L. JIANG[1], L. ZHANG[1]

[1]Department of Gynaecology and Obstetrics, The First Affiliated Hospital of ZheJiang Chinese Medical University, Zhejiang Provincial Hospital of TCM, Hangzhou, Zhejiang Province, China
[2]Department of Radiotherapy, The First Affiliated Hospital of ZheJiang Chinese Medical University, Zhejiang Provincial Hospital of TCM, Hangzhou, Zhejiang Province, China

Abstract. – OBJECTIVE: The objective of the present work was to investigate key genes in ovarian cancer based on mAP-KL method which comprised the maxT multiple hypothesis (m), Krzanowski and Lai (KL) cluster quality index, and affinity propagation (AP) clustering algorithm, and mutual information network (MIN) constructed by the context likelihood of relatedness (CLR) algorithm.

MATERIALS AND METHODS: MAP-KL method was employed to identify exemplars in ovarian cancer, of which the maxT function ranked the genes of train set and test set and obtained top 200 genes; KL cluster index was utilized to determine the quantity of clusters; and then AP clustering algorithm was conducted to identify the clusters and their exemplars. Also, we assessed the classification performance of mAP-KL by support vector machines (SVM) model. Subsequently, the MIN for exemplars and cluster genes was constructed according to CLR algorithm. Finally, topological centrality properties of exemplars in MIN were assessed to investigate key genes for ovarian cancer.

RESULTS: SVM model validated that the classification between normal controls and ovarian cancer patients by mAP-KL had a good performance. A total of 22 clusters and exemplars were detected by performing the mAP-KL method. Based on the topological centrality analyses for exemplars in MIN, we considered the C9orf16, COX5B and ACTB to be key genes in the progress of ovarian cancer.

CONCLUSIONS: We have obtained three key genes (C9orf16, COX5B and ACTB) for ovarian cancer on the basis of mAP-KL method and MIN analysis. These genes might be potential biomarkers for treatment of ovarian cancer, and give insight for revealing the underlying mechanism of this tumor.

Key Words: Ovarian cancer, Cluster, Affinity propagation, Mutual information network, Gene.

Introduction

With roughly 14,000 deaths and 22,000 new cases estimated annually in women around the world, ovarian cancer has the highest mortality-to-new case ratio among all gynecologic malignancies[1]. When it invaded or spread to other parts of the body, symptoms may be vague or not apparent, but they become more noticeable as the cancer progresses[2]. Despite standard treatments have been performed on primary ovarian cancer patients, 70%-90% of them would develop into recurrent tumor[3]. Hence, effective treatments for ovarian cancer are on urgent need.

Recently, due to the accumulation of genetic variants in genes may be involved in the process of ovarian carcinogenesis[4-6], target gene has been proposed and developed into an emerging mean for tumor treatment. The knowledge of target gene for ovarian cancer attract more and more attentions from researchers, and several potential biomarkers are identified, such as CXCR2 and GADD45A[7,8]. However, the quantity of them is still far from the demand. What is more, understanding molecular mechanism even investigating potential biomarkers of ovarian cancer may provide further prognostic and therapeutic insights.

High-throughput technologies have brought unprecedented opportunities for the large-scale analysis of the disease-related genes to ascertain the key molecular mechanisms[9]. And most of all, Sakellariou et al proposed a method by combining affinity propagation (AP) to control the quality of the partition of a known number of clusters, Krzanowski and Lai (KL)[10] index to evaluate the optimum number of clusters, and maxT function (m) to rank the genes in microarray data together, to

Corresponding Author: Chao Chen, MD; e-mail: chenchao201601@sina.com

determine a so-called exemplar which is most representative sample for this cluster[11]. This combination offers a more meaningful way to investigate exemplars or informative genes for disease and the relative target treatment.

However, from the biology point of view, genes interact with each other in complex disease rather than independent entities, which gain the difficulty to explore significant genes as biomarkers. Network-based approach is capable of extracting informative and significant genes dependent on bio-molecular networks, for instance, protein-protein interaction (PPI) network, co-expression network and mutual information network (MIN), rather than individual genes[12,13].

Therefore, in this paper, we combined the mAP-KL method with MIN to select key genes in ovarian cancer. Firstly, the mAP-KL method was implemented to investigate clusters and exemplars, and then the support vector machines (SVM) model was designed to assess classification performance of it. Subsequently, the MIN for cluster genes were constructed based on the context likelihood of relatedness (CLR) algorithm. Finally, key genes of ovarian cancer were explored from exemplars by topological analysis (degree, closeness, betweenness and transitivity) of MIN, which might be potential biomarkers for treatment and diagnoses of ovarian cancer.

Materials and Methods

Gene Expression Data

In the present study, the gene expression data for ovarian cancer with accessing number E-GEOD-26712 was recruited from ArrayExpress database. E-GEOD-26712, which presented on A-AFFY-33 – Affymetrix GeneChip Human Genome HG-U133A [HG-U133A] Platform, was consisted of 185 ovarian tumor samples and 10 normal samples. We primarily divided total samples into two parts including of train set and test set according to the proportion of 3:2. In other words, we kept 117 samples (111 tumors samples and 6 normal samples) to build a train set, and 78 samples (74 tumors samples and 4 normal samples) to construct a test set, respectively. The train set was utilized to perform the balance, and the test set was used to identify the classification models.

In order to control the quality of data and eliminate batch effects caused by experimental parameters and other factors, we compared four kinds of normalized pre-treatment measures, mean-centering[14], z-score[15], quantile[16] and cyclic loess[17], and then carried out log2 transformation for the normalized data. Among them, we selected one with the optimal preprocessing outcome as the pre-processed method which was utilized for further analysis.

MAP-KL

The mAP-KL method was a data-driven and classifier-independent hybrid feature selection method to select a small yet informative subset of genes[18]. The method combined maxT multiple hypothesis testing[19], KL cluster quality index[10] as well as AP clustering algorithm[11]. Its theory showed that the clusters of genes that shared similar biological functions related to the investigated disease. MaxT function, which computes permutation adjusted p-values for step-down multiple testing procedures[20], was employed to rank the genes of the training set and then we reserved the top N genes for further exploitation[21].

In the second step, the KL quality index as included in the ClusterSim package aimed to determine the number of clusters which in essence would be the number of representative genes solely on the ovarian cancer samples of the train set[22]. The KL was calculated as following formula:

Where k was the number of clusters, W_k represented the within-cluster sum of squared errors. The clusters met to the thresholding of gene numbers < 50 were regarded as KL clusters.

$$KL(k) = \left| \frac{\emptyset(k)}{\emptyset(k+1)} \right|$$

of which

$$\emptyset(k) = (k-1)^{2/P} W_{k-1} - k^{2/P} W_k$$

In the next, the AP clustering method was engaged to detect clusters and provide a list of the most informative genes of each cluster, the so-called exemplars. Here, the AP clustering method involved in AP Cluster package regards each data point as a node in a network, and recursively transmits real-valued messages along edges of the network until a good set of exemplars and corresponding clusters emerges[23]. It explored n (n = k, the KL index) clusters among the top N genes according to the pre-defined number, and obtained a list of n exemplars[24]. These were expected to form a classifier that shall discriminate between the normal and fibroid classes in a test set which was formulated through retaining only those n genes to proceed with the classification.

Classification and Evaluation

SVM, which has become popular because of its effective learning properties[25,26], is supervised learning models with associated learning algorithms that analyze data and recognize patterns, used for classification and regression analysis[27]. For the purpose of evaluating the classification performance of the mAP-KL method, we employed SVM with linear kernel, in which 5-fold cross-validation was applied on the train set to analyze the potential classification strength of the models' and then estimated its prediction power on the separate test set. We determined whether the classification results were reasonable according to the following parameter values: the area under the receiver operating characteristics curve (AUC) which is a better measure for assessing the predictive ability of machine learners than accuracy[28]; the Matthews coefficient correlation classification measure (MCC) was a measure of the quality of binary classification and considered the true and false positive and negatives[29]; true negative rate (TNR) or specificity represents the ratio of correctly classified negatives to the actual number of negatives; as well as true positive rate (TPR) or sensitivity is defined to be the ratio of positives correctly classified to the actual number of positives[30].

MIN Construction and Topological Analysis

MIN construction

MIN is a subcategory of network inference methods, which can infer a link between a couple of nodes if it has a high score based on mutual information (MI)[31]. Hence, in this research, we constructed the MIN for the cluster related top N genes. Firstly, we calculated the mutual information matrix (MIM), a square matrix whose i, j-th element is the MI between the genes X_i and X_j, q was a probability measure.

$$MIM_{ij} = \sum_{i,j} q(x_i, y_j) log \frac{q(x_i,y_j)}{q(x_i)q(y_j)}$$

Secondly, we applied the CLR algorithm which is an extension of the relevance network approach to compute the network boundary value[32]. This algorithm computes the MI for each pair of genes and derives a score related to the empirical distribution of the MI values[33]. In particular, instead of considering the information $I(X_i; X_j)$ between genes X_i and X_j, it took into account the edge score:

$$z_{ij} = \sqrt{z_i^2 + z_j^2}$$

of wich

$$z_i = \max\left(0, \frac{I(X_i;X_j-\mu_i)}{\sigma_i}\right)$$

Where μ_i and σ_i represented respectively the sample mean and standard deviation of the empirical distribution of the values $I(X_i; X_j)$.

Topological Centrality Analysis

To in-depth understand the biological functions of the genes in the MIN, we calculated the indicators including of degree[34], closeness[35] and betweenness[36] by topological analysis. Degree quantifies the local topology of each gene by summing up the number of its adjacent genes, and the top 4% degree distributed genes were defined as hub genes. Closeness centrality is a measure of the average length of the shortest paths to access all other genes in the network. Betweenness centrality is a shortest path enumeration-based metric in graphs for determining how the neighbors of a node were interconnected, and was considered the ratio of the node in the shortest path between two other nodes.

Results

Data

In this paper, we compared four pre-treatment methods (mean-centering, z-score, quantile and cyclic loess), and the results were illustrated in Figure 1. After log2 transformation, quantile normalization had a better performance than the others and, hence, we selected it for further exploitation. A total of 22283 genes were obtained in the pre-processed gene expression profile.

Clusters and Exemplars

Following with the maxT multiple hypothesis testing, we reserved the top 200 genes (N = 200), and listed the top 100 in Table I. SMPD3, PFN1, CALR, C9orf16 and WT1 were the top five genes. By performing AP clustering algorithm in conjunction with KL cluster quality for top 200 genes, the quantity of clusters with gene amount < 50

was 22. The gene compositions of different clusters were various. Among them, Cluster 6 possessed the most number of genes with 36, while Cluster 4 only included 9 genes (Table II). Furthermore, for each cluster, an exemplar was detected on the basis of AP clustering method, which might play key roles in the progress of ovarian cancer. The exemplars were displayed in Table II, for example, exemplar for Cluster 1 and Cluster 2 was PFN1 and C9orf16, respectively.

Evaluation by SVM Model

In this step, to find an optimal hyper plane that separates the test samples by a maximal margin, with all positive samples lying on one side and all negative samples lying on the other side, we assessed the classification performance for exemplars obtained from mAP-KL method with the assistance of two classifiers (linear, 5-CV). Based on the SVM model, the classified results had the highest scores (AUC = 1.00, MCC = 1.00, TNR = 1.00 and TPR = 1.00), and thus we concluded that the classification performance was almost ideal during the SVM evaluation, and inferred that the mAP-KL methodology which combined ranking-filtering and cluster analysis was feasible and suitable for identifying clusters of ovarian cancer.

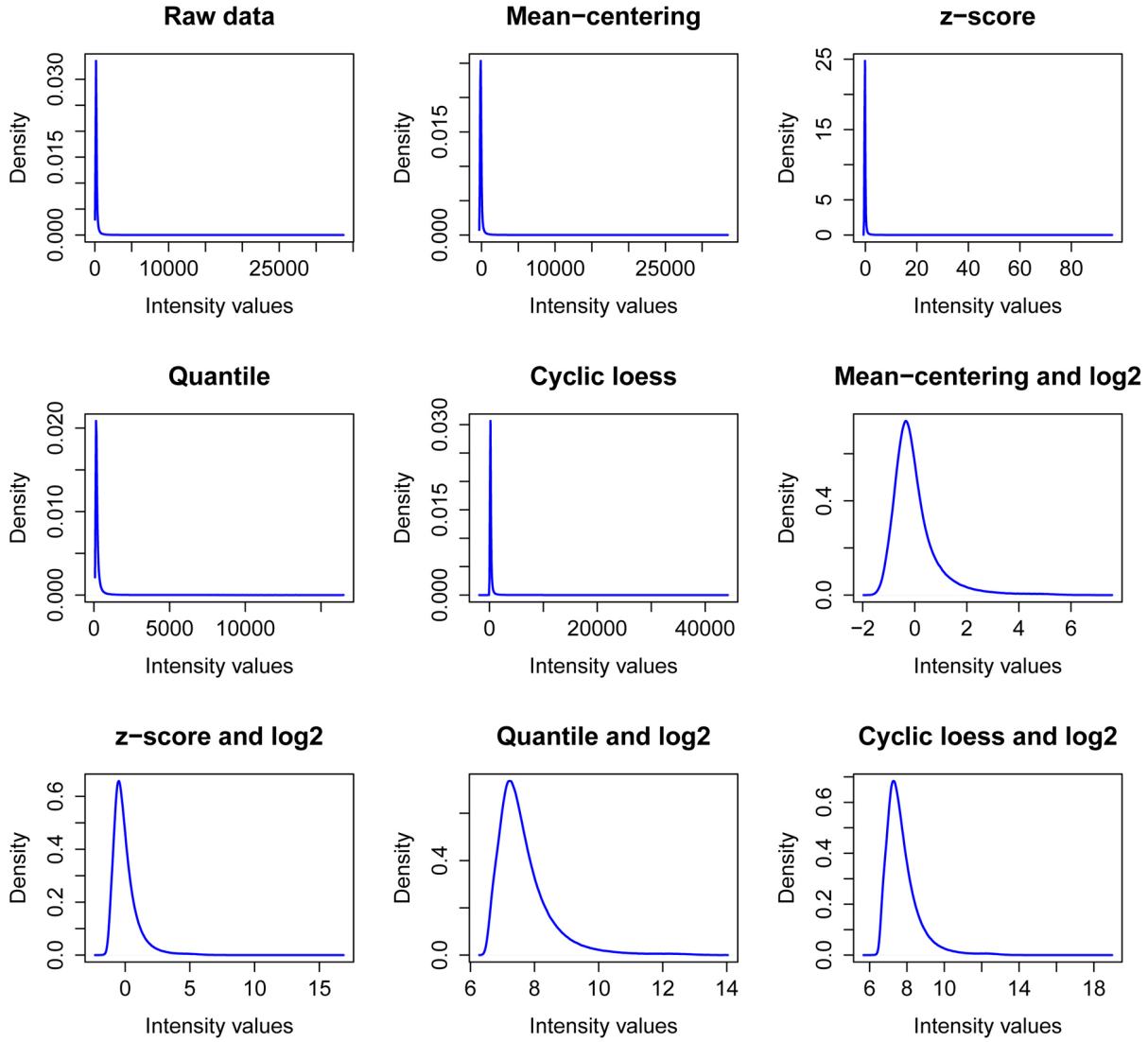

Figure 1. Preprocessing for microarray data by mean-centering, z-score, quantile and cyclic loess methods.

Table I. The top 100 genes based on maxT multiple hypothesis testing.

Number	Gene	Number	Gene	Number	Gene	Number	Gene
1	SMPD3	26	PKM	51	SLC52A2	76	HNRNPA3
2	PFN1	27	COX5B	52	DCUN1D4	77	TRAK2
3	CALR	28	NA	53	SNRPF	78	KRR1
4	C9orf16	29	VPS54	54	PRMT1	79	MTO1
5	WT1	30	MED14	55	UBE2S	80	TNFRSF10B
6	RPS7	31	PREPL	56	BRD2	81	TIMM17B
7	RBM25	32	ARPC4	57	ATP5H	82	ACTN4
8	TCEAL2	33	SMARCB1	58	CALR	83	SNX7
9	SEC61B	34	ARL8B	59	RPL22	84	GPBP1L1
10	PPDPF	35	APPBP2	60	ATXN1	85	IL6ST
11	CORO1B	36	NAP1L2	61	N4BP2L2	86	SMARCA2
12	C9orf16	37	S100A1	62	SEH1L	87	TCEB2
13	MET	38	UQCRQ	63	PTK7	88	ZDHHC17
14	CRABP2	39	TVP23B	64	ALDH3A2	89	AGPAT2
15	ITM2B	40	KLK8	65	NUCB1	90	FAM13B
16	S100PBP	41	MIOS	66	DKFZP586I1420	91	JMJD4
17	IK	42	HNRNPL	67	ATP2A2	92	MAZ
18	PPP2R5C	43	CLPTM1	68	AKAP17A	93	ARPC4
19	GANAB	44	SEC23B	69	GRINA	94	TSC22D4
20	CAST	45	PEA15	70	DFNA5	95	CCSER2
21	KIAA0368	46	FEZ2	71	PGRMC2	96	SYPL1
22	C9orf16	47	ZKSCAN7	72	HECA	97	MRPL4
23	GSTP1	48	FKBP8	73	RYR2	98	PLSCR4
24	PEA15	49	ATP8A1	74	TACC1	99	MDFIC
25	FAM120A	50	ALDH9A1	75	AMPD3	100	CAPNS1

MIN Construction and Topological Analysis

The ideal evaluation of classification for mAP-KL method gave more confidence to the significance of 22 exemplars in ovarian cancer, whereas those also were expected to form a classifier or module that shall distinct between normal controls and ovarian cancer patients. Therefore, we constructed a MIN for top 200 genes (Figure 2). It had 200 nodes and 1902 interactions, of which the yellow ones stood for hub genes. They were C9orf16, ACTB, COX5B, IL6ST, TYP23B, FAM120A, CLPTM1 and RAPGEF2. When extracting cluster genes related network, that was to say, cluster genes were mapped to MIN preferentially, we obtained Figure 3, there were 177 nodes and 1002 edges, and their topological properties in the MIN were showed in Table III. Results for exemplars were inconsistent due to different methods, except for C9orf16, ACTB and COX5B. Interestingly, the three genes were right the common genes between hub genes and exemplars, and thus we defined them as the key genes for ovarian cancer. The degree, closeness and betweenness for C9orf16 were 129, 7.16 and 8495, respectively.

Table II. Clusters identified by mAP-KL method for ovarian cancer.

Cluster	Amount	Exemplar
1	24	PFN1
2	28	C9orf16
3	15	HNRNPA1
4	9	HNRNPA1P10
5	14	CRABP2
6	36	COX5B
7	21	VPS54
8	25	HIST2H2AA3
9	21	HIST2H2AA4
10	29	MIOS
11	15	PEA15
12	10	FEZ2
13	13	TCEB2
14	21	RAD23A
15	19	ZNF45
16	18	PPP4C
17	24	ACTB
18	27	ISG15
19	11	FOLR1
20	16	FAM69A
21	30	SNORD21
22	17	RPL5

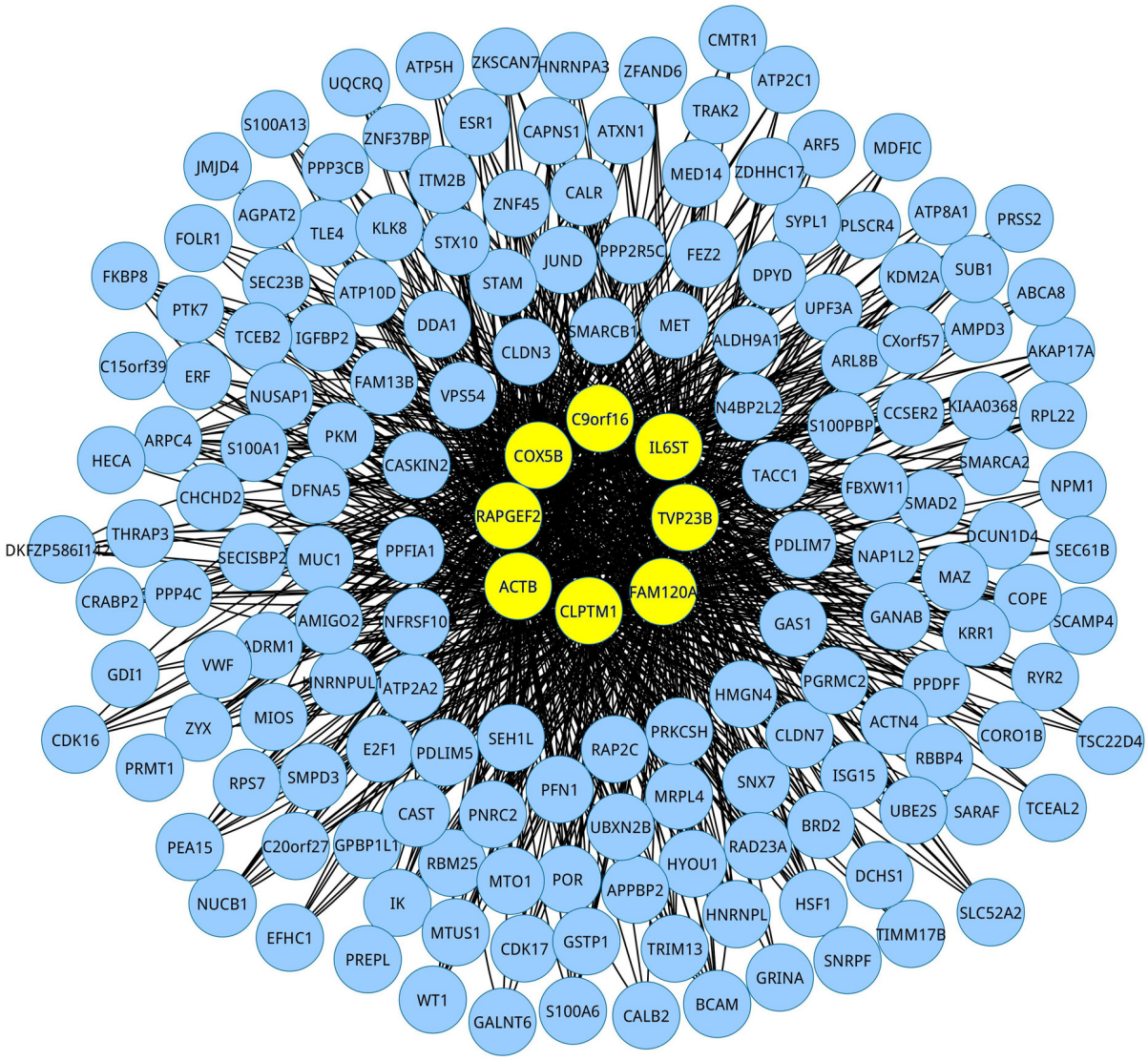

Figure 2. Mutual information network (MIN) for top 200 ranked genes in the microarray data. There were 200 nodes and 1902 interactions, where nodes represented genes, and edges were the interactions between two genes. The yellow nodes were hub genes with top 4% degree distribution of the MIN.

Discussion

In order to investigate the key genes of ovarian cancer, we combined the mAP-KL algorithm and MIN related analysis. The results showed that a total of 22 clusters and exemplars were obtained based on mAP-KL. Additionally, we validated that the classified performance of the mAP-KL method was reasonable and feasible based on SVM parameters. Subsequently, the MIN for top 200 genes and cluster genes were constructed by CLR algorithm, and topological centrality analyses were conducted on them. Among the hub genes and exemplars, we discovered 3 common genes (C9orf16, COX5B and ACTB), and termed key genes. Besides, key genes in cluster MIN had a higher degree, closeness and betweenness centrality than the others, which indicated that they were more significant to ovarian cancer progression.

Taking COX5B and ACTB for examples, COX5B (cytochrome c oxidase subunit 5B) is a peripheral subunit of the cytochrome c oxidase complex (COX). COX is a multi-subunit enzyme complex that couples the transfer of electrons from cytochrome c to molecular oxygen and contributes to a proton electrochemical gradient across the inner mitochondrial membrane to drive ATP synthesis[37]. A previous study[38] demonstrated that COX5A (COX subunit 5A) and COX5B in-

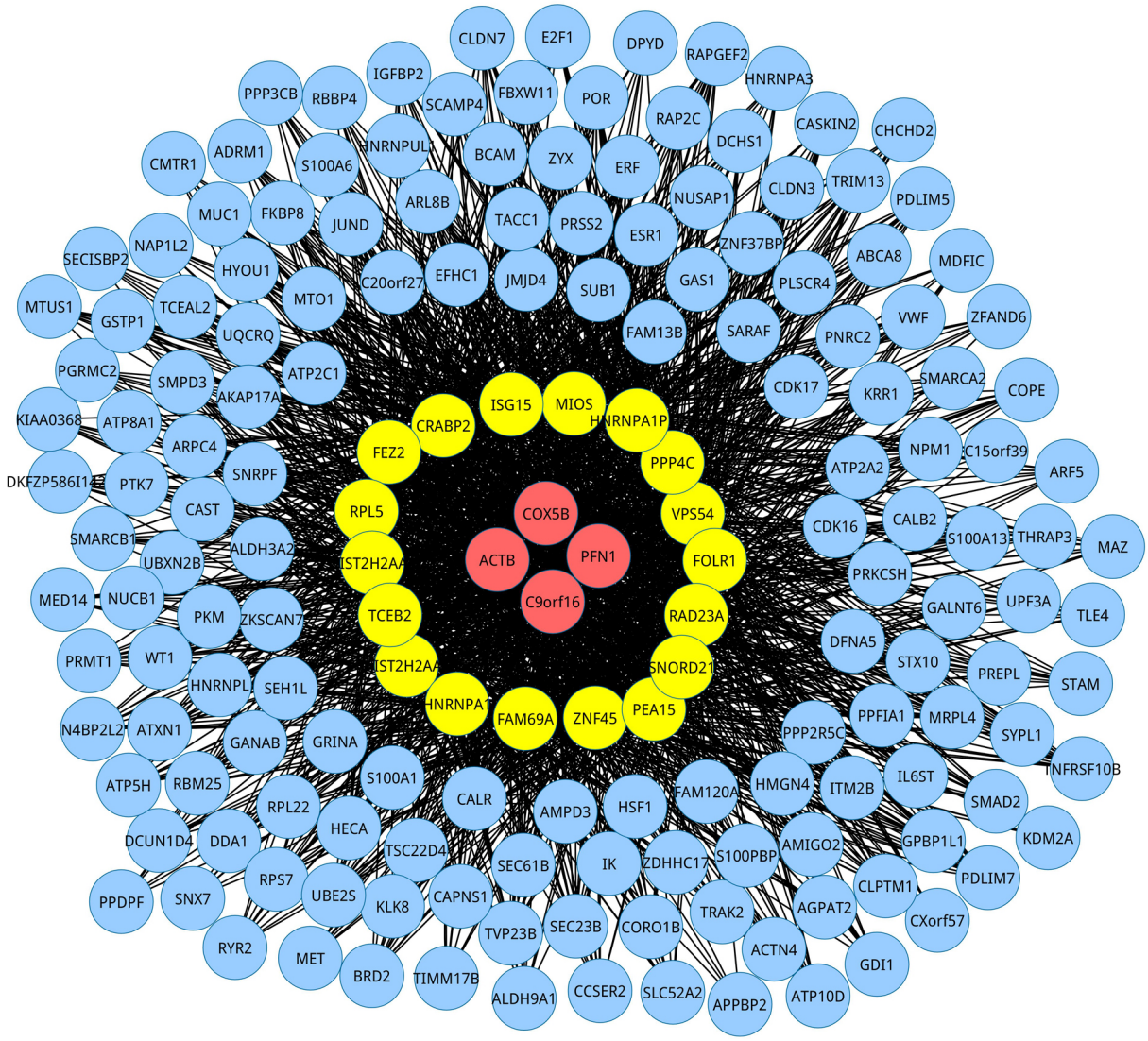

Figure 3. Mutual information network (MIN) for clusters related genes of ovarian cancer. There were 177 nodes and edges, where nodes represented genes, and edges were the interactions between two genes. The yellow and red nodes both stood for 22 exemplars identified by mAP-KL method, what was more, the red nodes also indicated the key genes.

volved in the regulation of cancer cell metabolism. For instance, down-regulation of COX5B in breast cancer cell lines could suppress cell proliferation and induced cell senescence[39]. What was more, Taylor et al[40] had revealed that COX5B was associated with ovarian cancer that might be translated into diagnostic and prognostic biomarkers of the disease. Therefore, COX5B was correlated to ovarian cancer closely.

ACTB (actin, beta), one of six different actin isoforms, has been widely used as a reference gene in quantifying expression levels in tumors[41]. Actions are highly conserved proteins that are involved in cell motility, structure and integrity[42]. Accumulating evidence indicated that ACTB was closely associated with a variety of cancers[43-46], such as breast cancer and lung cancer. Furthermore, it had also been reported that ACTB played significant roles in ovarian cancer[47]. In this work, ACTB was dug out from numerous genes dependent on mAP-KL method and MIN related analysis, which suggested its potential association with ovarian cancer.

Conclusions

We have identified three key genes (C9orf16, COX5B and ACTB) related to ovarian cancer ba-

Table III. The top 100 genes based on maxT multiple hypothesis testing.

Exemplar	Degree	Closeness	Betweenness
C9orf16	129	7.16	8495
ACTB	120	7.05	5323
COX5B	113	6.88	4018
VPS54	109	3.94	15
FEZ2	109	3.61	234
PFN1	107	4.76	604
CRABP2	106	5.69	2158
PEA15	106	3.46	87
SNORD21	106	5.34	63
MIOS	105	6.35	753
RAD23A	102	4.02	202
FOLR1	102	5.32	198
HIST2H2AA3	100	6.29	207
HNRNPA1P10	99	4.45	160
HIST2H2AA4	98	6.04	123
TCEB2	98	5.68	164
HNRNPA1	97	5.30	558
FAM69A	96	5.46	659
PPP4C	94	7.05	3043
ISG15	93	5.83	1131
ZNF45	89	5.11	890
RPL5	85	4.99	784

sed on mAP-KL algorithm and MIN. These genes might be potential biomarkers for early detection and therapy of ovarian cancer.

Conflicts of interest

The authors declare no conflicts of interest.

References

1) Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. CA: Cancer J Clin 2014; 64: 9-29.
2) Qian X, Qin J, Pan S, Li X, Pan Y, Ma S. Maintenance therapy in ovarian cancer with targeted agents improves PFS and OS: A Systematic Review and Meta-Analysis. PLoS One 2015; 10: e0139026.
3) Armstrong DK. Relapsed ovarian cancer: challenges and management strategies for a chronic disease. Oncologist 2002; 7: 20-28.
4) Risch HA. Hormonal etiology of epithelial ovarian cancer, with a hypothesis concerning the role of androgens and progesterone. J Natl Cancer Inst 1998; 90: 1774-1786.
5) Zhang X, Zhang H. Diminished miR-613 expression as a novel prognostic biomarker for human ovarian cancer. Eur Rev Med Pharmacol Sci 2016; 20: 837-841.
6) Cong J, Liu R, Wang X, Wang J, Wang H, Hou J. Low miR-498 expression levels are associated with poor prognosis in ovarian cancer. Eur Rev Med Pharmacol Sci 2015; 19: 4762-4765.
7) Devapatla B, Sharma A, Woo S. CXCR2 Inhibition combined with sorafenib improved antitumor and antiangiogenic response in preclinical models of ovarian cancer. PLoS One 2015; 10: e0139237.
8) Yuan C, Liu X, Liu X, Yang N, Liu Z, Yan S, Shen K, Kong B. The GADD45A (1506T> C) polymorphism is associated with ovarian cancer susceptibility and prognosis. PLoS One 2015; 10: e0138692.
9) Abu-Jamous B, Fa R, Nandi AK. High-throughput Technologies, in Integrative Cluster Analysis in Bioinformatics. John Wiley & Sons, Ltd., 2015; pp. 53-66.
10) Krzanowski WJ, Lai Y. A criterion for determining the number of groups in a data set using sum-of-squares clustering. Biometrics 1988; 44: 23-34.
11) Frey BJ, Dueck D. Clustering by passing messages between data points. Science 2007; 315: 972-976.
12) Liu ZP, Wang Y, Zhang XS, Chen L. Network-based analysis of complex diseases. IET Syst Biol 2012; 6: 22-33.
13) Chen L, Wang RS, Zhang XS. Reconstruction of Gene Regulatory Networks, in Biomolecular Networks. John Wiley & Sons, Inc., 2009; pp. 47-87.
14) Wylie D, Shelton J, Choudhary A, Adai AT. A novel mean-centering method for normalizing microRNA expression from high-throughput RT-qPCR data. BMC Res Notes 2011; 4: 555.
15) Cheadle C, Vawter MP, Freed WJ, Becker KG. Analysis of microarray data using Z score transformation. J Mol Diagn 2003; 5: 73-81.
16) Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 2003; 19: 185-193.
17) Ballman KV, Grill DE, Oberg AL, Therneau TM. Faster cyclic loess: normalizing RNA arrays via linear models. Bioinformatics 2004; 20: 2778-2786.
18) Sakellariou A, Sanoudou D, Spyrou G. Combining multiple hypothesis testing and affinity propagation clustering leads to accurate, robust and sample size independent classification on gene expression data. BMC Bioinformatics 2012; 13: 270.
19) Pollard K, Dudoit S, van der Laan M. Multiple Testing Procedures: R multtest Package and Applications to Genomics UC Berkeley Division of Biostatistics Working Paper Series. Working Paper 164. http://www. bepress. com/ucbbiostat/paper164, 2008.
20) Westfall PH, Young SS. Resampling-based multiple testing: Examples and methods for p-value adjustment. John Wiley & Sons., 1993; pp. 279.
21) Sakellariou A, Sanoudou D, Spyrou G. Investigating the minimum required number of genes for the classification of neuromuscular disease microarray data. IEEE T Inf Technol B 2011; 15: 349-355.
22) Walesiak M. Cluster analysis with ClusterSim computer program and R environment. Acta Universitatis Lodziensis Folia Oeconomica 2008; 216.

23) Bodenhofer U, Kothmeier A, Hochreiter S. APCluster: an R package for affinity propagation clustering. Bioinformatics 2011; 27: 2463-2464.

24) Team RC. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2012. 2014.

25) Chou K-C. Progress in protein structural class prediction and its impact to bioinformatics and proteomics. Curr Protein Pept Sc 2005; 6: 423-436.

26) Chen C, Chen LX, Zou XY, Cai PX. Predicting protein structural class based on multi-features fusion. J Theor Biol 2008; 253: 388-392.

27) Cortes C, Vapnik V. Support-vector networks. Mach Learn 1995; 20: 273-297.

28) Huang J, Ling CX. Using AUC and accuracy in evaluating learning algorithms. IEEE T Knowl Data En 2005; 17: 299-310.

29) Kantardjieff KA, Rupp B. Matthews coefficient probabilities: improved estimates for unit cell contents of proteins, DNA, and protein–nucleic acid complex crystals. Protein Sci 2003; 12: 1865-1871.

30) Baldi P, Brunak S, Chauvin Y, Andersen CA, Nielsen H. Assessing the accuracy of prediction algorithms for classification: an overview. Bioinformatics 2000; 16: 412-424.

31) Meyer PE, Kontos K, Lafitte F, Bontempi G. Information-theoretic inference of large transcriptional regulatory networks. EURASIP J Bioinform Syst Biol 2007; 1: 1-9.

32) Faith JJ, Hayete B, Thaden JT, Mogno I, Wierzbowski J, Cottarel G, Kasif S, Collins JJ, Gardner TS. Large-scale mapping and validation of Escherichia coli transcriptional regulation from a compendium of expression profiles. PLoS Biol 2007; 5: e8.

33) Meyer PE, Lafitte F, Bontempi G. minet: AR/Bioconductor package for inferring large transcriptional networks using mutual information. BMC Bioinformatics 2008; 9: 461.

34) Haythornthwaite C. Social network analysis: An approach and technique for the study of information exchange. Lib Inform Sci Res 1996; 18: 323-342.

35) Wasserman S, Faust K. Social network analysis: Methods and applications. 1994; 8.

36) Barthelemy M. Betweenness centrality in large complex networks. Eur Phys J B 2004; 38: 163-168.

37) Surhone LM, Tennoe MT, Henssonow SF. Macromolecule, Acid N. COX5B. 2011.

38) Chen Z, Pervaiz S. Involvement of cytochrome c oxidase subunits Va and Vb in the regulation of cancer cell metabolism by Bcl-2. Cell Death Differ 2010; 17: 408-420.

39) Gao SP, Sun HF, Jiang HL, Li LD, Hu X, Xu XE, Jin W. Loss of COX5B inhibits proliferation and promotes senescence via mitochondrial dysfunction in breast cancer. Oncotarget 2015; 6: 43363-43374.

40) Taylor S, Sanchez C, Adrees A, Ozbek HN, Sahni H, Said NA. Global profiling of SPARC-regulated metabolic pathways in ovarian cancer. Cancer Res 2015; 75: 1146-1146.

41) Guo C, Liu S, Wang J, Sun M-Z, Greenaway FT. ACTB in cancer. Clin Chim Acta 2013; 417: 39-44.

42) Gunning PW, Ghoshdastider U, Whitaker S, Popp D, Robinson RC. The evolution of compositionally and functionally distinct actin filaments. J Cell Sci 2015; 128: 2009-2019.

43) Majidzadeh-A K, Esmaeili R, Abdoli N. TFRC and ACTB as the best reference genes to quantify Urokinase Plasminogen Activator in breast cancer. BMC Res Notes 2011; 4: 215.

44) Hauser S, Zahalka T, Ellinger J, Fechner G, Heukamp LC, Von Ruecker A, Mueller SC, Bastian PJ. Cell-free circulating DNA: diagnostic value in patients with renal cell cancer. Anticancer Res 2010; 30: 2785-2789.

45) Mollerup S, Berge G, Bæra R, Skaug V, Hewer A, Phillips DH, Stangeland L, Haugen A. Sex differences in risk of lung cancer: expression of genes in the PAH bioactivation pathway in relation to smoking and bulky DNA adducts. Int J Cancer 2006; 119: 741-744.

46) Warren JD, Xiong W, Bunker AM, Vaughn CP, Furtado LV, Roberts WL, Fang JC, Samowitz WS, Heichman KA. Septin 9 methylated DNA is a sensitive and specific blood test for colorectal cancer. BMC Med 2011; 9: 133.

47) Li Y-L, Ye F, Hu Y, Lu W-G, Xie X. Identification of suitable reference genes for gene expression studies of human serous ovarian cancer by real-time polymerase chain reaction. Anal Biochem 2009; 394: 110-116.