ORIGINAL ARTICLE

Bioinformatics Analysis of Differentially Expressed Genes in Liver Cancer for Identification of Key Genes and Pathways

Aisya Fathiya Che Rosli, Siti Razila Abdul Razak, Nurulisa Zulkifle

Cluster for Oncological & Radiological Sciences, Advanced Medical & Dental Institute, Universiti Sains Malaysia, Bertam, 13200 Kepala Batas, Pulau Pinang

ABSTRACT

Introduction: Liver cancer is among the main leading cause of mortality in Malaysia and the world. Therefore, there is an urgent need to understand the complex mechanisms and pathways involved in liver cancer. **Methods:** Microarray datasets GSE84402, GSE60502, GSE29721 and GSE19665 were downloaded from GEO database. The datasets were normalised and differentially expressed genes (DEGs) were calculated using GeneSpring software. GO and KEGG pathway enrichment analyses were then performed using DAVID. Finally, Cytoscape stringApp plugin was utilised to construct a protein-protein interaction (PPI) network. **Results:** A total of 1382, 714, 1038 and 1828 DEGs satisfying *p* value cut-off 0.01 and fold change cut-off 2.0 are identified from each dataset. 412 DEGs appeared in at least three datasets, consisting of 167 up-regulated and 245 down-regulated genes. These genes are most significantly enriched in terms related to cell division and mitotic nuclear division. Construction of PPI network produced a network with 275 nodes and 2157 edges with confidence score 0.7. Topological analysis identified CDK1, TOP2A and NDC80 as key genes. MCODE plugin extracted five modules from the network with mitotic cell cycle process being the most enriched term in module 1. Meanwhile, platelet degranulation, epoxygenase P450 pathway, cellular response to zinc ion and complement and coagulation cascade are the terms enriched in module 2, 3, 4 and 5. **Conclusion:** The key genes and pathways identified from this study provide information on the molecular mechanism underlying liver cancer to increase our understanding regarding liver cancer development and progression at molecular level.

Keywords: Bioinformatics, Liver cancer, Differentially expressed genes, Enrichment analysis, Protein–protein interaction

Corresponding Author:
Nurulisa Zulkifle, PhD
Email: nurulisa@usm.my
Tel: +604-5622504

INTRODUCTION

Liver cancer or hepatocellular carcinoma (HCC) is among the leading cause of global cancer-related death (1). In Asia alone, at least 580,000 new cases are predicted to happen every year (2). This high rate of prevalence is caused by late diagnosis, which resulted in low survival rate of the patients. The most common molecular biomarkers for diagnosing liver cancer include alpha-fetoprotein (AFP), which is widely used since its discovery in the blood of liver cancer patients in 1964 (3). However, AFP sensitivity and specificity are not fully optimal (4). The sensitivity declines to 25% in tumors that are 3 cm or less and in 80% of small hepatocellular carcinomas, the differential expression of AFP are almost undetected (5). The second most commonly used biomarker for liver cancer evaluation is Des-gamma-carboxyprothrombin (DCP). DCP could provide the best sensitivity and specificity that enable physicians to differentiate if the patients have liver cancer or cirrhosis (6). The combination of AFP and DCP has been suggested in order to improve liver cancer detection (7). In term of molecular targeted therapy, sorafenib, a multikinase inhibitor is reported to be the most commonly used drug to treat liver cancer (8). Sorafenib targets a number of molecules involved in signal transduction like PDGFR and RAF and angiogenesis such as VEGFR(1-3) and PDGFR (8). Despite the positive outcome of sorafenib in term of patient's survival, adverse events such as weight loss, fatigue, diarrhea and hand-food skin reaction are recorded among the users (8). Many other biomarkers and target molecules have been identified apart from the examples above. In fact, research in this field is still advancing in order to unravel novel detection and treatment methods that are reliable in term of specificity, sensitivity and safety.

The study of gene expression profiling is very useful in the discovery of potential biomarkers as well as therapeutics targets. This is achieved by comparing and analysing the genome-wide expression changes in health and

disease (9). One of the most powerful tools for analysing gene expression profiles at genome level is microarray technology (10). With this technology, researchers are able to study the function of genes and their products on a genome-wide basis. Microarray data generally could be used to uncover the mechanisms involved in the development of diseases, to categorise and group a disease for example differentiate many types of tumor, to monitor the patient's response to therapy and to identify diagnostic or prognostic biomarkers of cancer (10). Although RNA-Seq approach proved to be better in detecting low abundance transcripts and having a broader dynamic range which allowed for the detection of more differentially expressed genes with higher fold-change, microarray is still the most opted one due to its cheaper cost and less complex data analysis (11).

In this study, four microarray datasets were analysed to screen for differentially expressed genes (DEGs) between liver cancer tissues and their normal counterparts. Functional enrichment and network analysis were employed to identify key genes and pathways with the aim to produce a systematic perspective in order to understand the key players and mechanisms in the development and progression of liver cancer.

MATERIALS AND METHODS

Selection and processing of microarray datasets

In this study, Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/) was used to retrieve the microarray dataset. Selected datasets were normalised and analysed using GeneSpring software (Agilent Technologies) to identify the DEGs. Benjamini and Hochberg false discovery rate method was applied to fix the occurrence of false positive results. Genes that have adjusted p value <0.01 and fold change >2.0 were considered significant and selected for further analysis.

Acquiring intersection of DEGs in all datasets

To identify the overlapping DEGs that consistently appeared in all selected datasets, Venn diagram analysis was conducted using online web tools at http://bioinformatics.psb.ugent.be/webtools /Venn/.

Functional and pathways enrichment analysis

The list of DEGs was submitted to Database for Annotation, Visualization and Integrated Discovery (DAVID) database (https://david.ncifcrf.gov/) to identify the gene ontology (GO) terms covering biological process (BP), cellular component (CC) and molecular function (MF) domains and the pathways involved in liver cancer. Kyoto Encyclopedia of Genes and Genomes (KEGG) database was opted to map the datasets for biological interpretation. A p value <0.05 was considered significant.

Construction of PPI network and analysis of modules

The interaction among DEGs was extracted from Search Tool for the Retrieval of Interacting Genes (STRING) database (12) with confidence score >0.7 and PPI network was constructed in Cytoscape ver3.6 (13). Then, the network was analysed using NetworkAnalyzer to identify the key genes in the network. Subnetworks or modules were extracted using MCODE plugin (14). The criteria for module selection were set at: nodes >5, score >5, degree cut-off=2, node score cut-off=0.2, max depth=100 and k-core=2. The GO and KEGG pathway enrichment analysis of each module were analysed using ClueGO plugin (15). The p value was calculated by right-side hypergeometric test followed by Benjamini-Hochberg multiple test correction. In this study, terms and pathways with adjusted p value <0.05 was considered significant.

RESULTS

Microarray datasets

To search for the suitable datasets, the keyword 'liver cancer' is used as a query and the search is limited to species '*Homo sapiens*' and type 'expression profiling by array' only. A total of 684 items appeared in the search results as of 11th December 2017. After a careful review, only five gene expression profiles: GSE84402 (16), GSE64041 (17), GSE60502 (18), GSE29721 (19) and GSE19665 (20), which altogether contain 112 liver carcinoma and 112 matched non-cancerous liver specimens from tissue biopsy, are selected for downstream analysis.

Differentially expressed genes

Analysis performed by GeneSpring software identified 1382, 714, 1038 and 1828 DEGs satisfying p value cut-off 0.01 and fold change cut-off 2.0 from GSE84402, GSE60502, GSE29721 and GSE19665, respectively. At this point, dataset GSE64041 did not give any significant DEGs and therefore is excluded in further analysis. The top 20 DEGs for each microarray datasets are tabulated in Table I. To screen for genes that have different expression in all datasets, the intersection of the DEGs is obtained by Venn analysis. A total of 412 genes are found to appear consistently in at least three datasets consisting of 167 up-regulated genes and 245 down-regulated genes (Fig. 1).

Enriched terms and pathways

The up-regulated genes are significantly enriched in 96 GO-BP terms, 37 GO-CC terms, 32 GO-MF terms and 9 KEGG pathways (p <0.05). On the other hand, the down-regulated genes are significantly enriched in 79 GO-BP terms, 27 GO-CC terms, 42 GO-MF terms and 22 KEGG pathways (p <0.05). The top five GO terms and top three pathways enriched in both up- and down-regulated genes are tabulated in Table II.

PPI network analysis

The PPI constructed from 412 DEGs consist of 275 nodes and 2157 edges (Fig. 2). The nodes represent

Table I: The top 20 differentially expressed genes in four selected microarray datasets

Up-regulated							
GSE84402		GSE60502		GSE29721		GSE19665	
Gene	log FC	Gene	log FC	Gene	log FC	Gene	log FC
GPC3	3.320	*SPINK1*	4.304	*PEG10*	4.670	*ASPM*	3.747
TOP2A	3.207	*GPC3*	3.458	*GPC3*	4.418	*CRNDE*	3.656
ASPM	2.924	*SPP1*	2.560	*CCL20*	3.722	*FAM83D*	3.532
CCNB1	2.892	*GMNN*	2.417	*SULT1C2*	3.716	*CDKN3*	3.485
CTHRC1	2.820	*ACSL4*	2.400	*ACSL4*	3.462	*TRIM16*	3.457
PEG10	2.817	*ASPM*	2.368	*CDKN3*	3.413	*ANLN*	3.172
GINS1	2.800	*AKR1B10*	2.349	*ASPM*	3.408	*TOP2A*	3.167
CDKN3	2.757	*RRM2*	2.336	*TOP2A*	3.406	*HMMR*	3.110
SULT1C2	2.581	*CD24*	2.178	*CCNB1*	3.402	*CCNB1*	3.095
RRM2	2.473	*TOP2A*	2.134	*ANLN*	3.371	*RBM24*	3.078
Down-regulated							
GSE84402		GSE60502		GSE29721		GSE19665	
Gene	log FC	Gene	log FC	Gene	log FC	Gene	log FC
SLC22A1	-4.138	*HAMP*	-4.114	*THRSP*	-4.593	*HAMP*	-6.77
ABCA8	-3.957	*CYP2E1*	-3.954	*GBA3*	-3.969	*PDGFRA*	-5.819
MT1M	-3.925	*CYP1A2*	-3.702	*TTC36*	-3.946	*DCN*	-5.758
GYS2	-3.871	*CYP3A4*	-3.495	*CNDP1*	-3.927	*LUM*	-5.540
OIT3	-3.853	*ADH1B*	-3.414	*CLEC1B*	-3.717	*CRHBP*	-5.490
FCN3	-3.773	*FCN3*	-3.360	*GLS2*	-3.712	*IGH*	-5.405
CYP1A2	-3.616	*APOF*	-3.196	*APOF*	-3.684	*C7*	-5.367
APOF	-3.604	*MT1M*	-3.031	*CXCL14*	-3.628	*OIT3*	-5.309
GBA3	-3.556	*CYP2A6*	-2.980	*HAO2*	-3.595	*FCN3*	-5.190
ADH1B	-3.504	*SLCO1B3*	-2.972	*NAT2*	-3.587	*SRPX*	-5.179

FC, fold change.

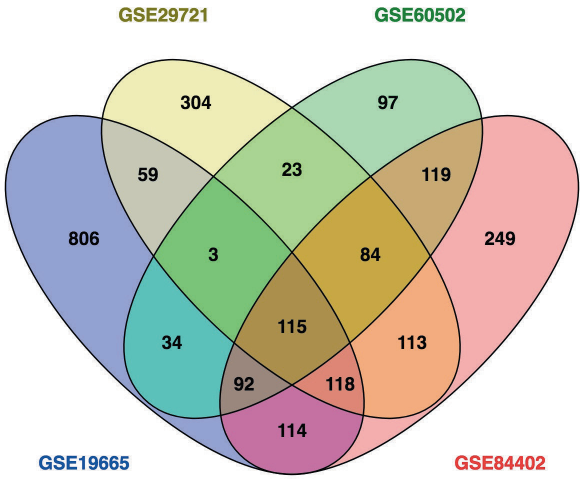

Figure 1: Venn diagram showing the intersection of DEGs from four datasets

each protein and the edges represent interactions between them. There are 137 single nodes (not shown) that do not formed a network and most of them are the down-regulated genes. Nodes degree and betweenness centrality are calculated by NetworkAnalyzer tool integrated in Cytoscape to identify the key genes. Nodes with high degree are regarded as hub genes and play a major role in robustness of the network. Apart from detecting hub nodes, the measurement of betweenness centrality (BC) to identify bottleneck nodes is also important. The BC measures the total number of shortest paths going through a certain node and the nodes with high betweenness; known as bottleneck nodes; act like a bridge in between highly interconnected network clusters (21). In this study, hub nodes with large BC value are considered as the key genes (Table III).

Table II: The top five over-represented GO terms and top three enriched KEGG pathways in up-regulated and down-regulated DEGs

Up-regulated			
Genes	Term	Count	*p* value
GOTERM_BP	GO:0051301~cell division	41	5.39×10^{-32}
	GO:0007067~mitotic nuclear division	30	1.70×10^{-23}
	GO:0007062~sister chromatid cohesion	19	4.17×10^{-18}
	GO:0000082~G1-S transition of mitotic cell cycle	14	1.47×10^{-11}
	GO:0006260~DNA replication	16	2.15×10^{-11}
GOTERM_CC	GO:0005654~nucleoplasm	78	1.96×10^{-21}
	GO:0000777~condensed chromosome kinetochore	17	7.76×10^{-17}
	GO:0000775~chromosome, centromeric region	14	3.48×10^{-15}
	GO:0005829~cytosol	73	6.44×10^{-14}
	GO:0005634~nucleus	94	2.06×10^{-12}
GOTERM_MF	GO:0005524~ATP binding	46	9.72×10^{-13}
	GO:0005515~protein binding	123	1.02×10^{-10}
	GO:0019901~protein kinase binding	20	2.70×10^{-9}
	GO:0003678~DNA helicase activity	5	6.95×10^{-5}
	GO:0003682~chromatin binding	14	8.61×10^{-5}
KEGG PATHWAY	hsa04110:Cell cycle	17	9.67×10^{-14}
	hsa03030:DNA replication	11	1.87×10^{-12}
	hsa04115:p53 signaling pathway	6	6.65×10^{-4}
Down-regulated			
Category	Term	Count	*p* value
GOTERM_BP	GO:0006805~xenobiotic metabolic process	14	4.04×10^{-11}
	GO:0019373~epoxygenase P450 pathway	9	4.12×10^{-11}
	GO:0055114~oxidation-reduction process	29	1.02×10^{-8}
	GO:0006898~receptor-mediated endocytosis	16	3.99×10^{-8}
	GO:0071294~cellular response to zinc ion	7	1.41×10^{-7}
GOTERM_CC	GO:0005576~extracellular region	65	3.78×10^{-16}
	GO:0070062~extracellular exosome	83	3.97×10^{-13}
	GO:0072562~blood microparticle	19	1.61×10^{-12}
	GO:0031090~organelle membrane	14	1.12×10^{-10}
	GO:0005615~extracellular space	47	1.86×10^{-9}
GOTERM_MF	GO:0020037~heme binding	19	3.08×10^{-13}
	GO:0016705~oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen	12	1.99×10^{-10}
	GO:0019825~oxygen binding	11	4.87×10^{-10}
	GO:0005506~iron ion binding	16	2.04×10^{-9}
	GO:0004497~monooxygenase activity	11	4.32×10^{-9}
KEGG PATHWAY	hsa01100:Metabolic pathways	48	2.59×10^{-6}
	hsa04610:Complement and coagulation cascades	10	7.98×10^{-6}
	hsa00830:Retinol metabolism	9	3.99×10^{-5}

GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; BP, biological process; CC, cellular component; MF, molecular function.

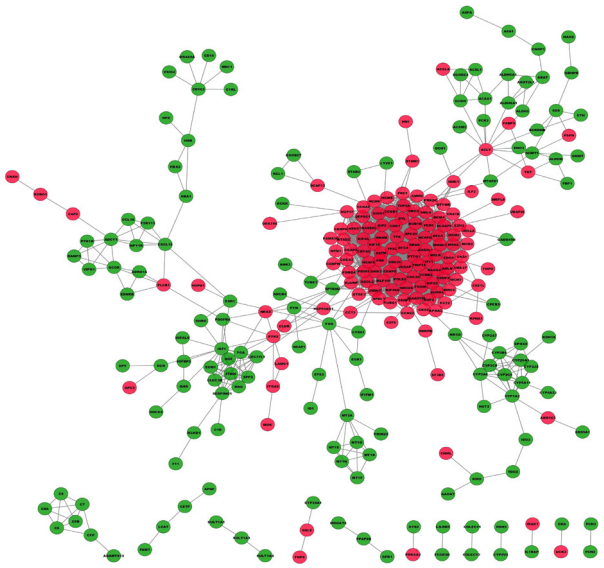

Figure 2: PPI network of the up-regulated (red) and down-regulated (green) DEGs

Table III: Key genes (*) in PPI network identified by degree and betweenness centrality measurements

Node	Degree	Node	Betweenness centrality
CDK1	86	*PPAP2B*	1
CCNA2	75	*SQLE*	1
CCNB1	75	*SULTIA3*	1
MAD2L1	74	*LCAT*	0.67
TOP2A	74	*CETP*	0.67
CCNB2	72	*CFP*	0.33
KIF11	67	*FOS*	0.32
NCAPG	66	**CDK1*	0.29
CDC20	66	**TOP2A*	0.22
TTK	65	*C7*	0.2
AURKA	65	*C6*	0.2
NDC80	64	*ACLY*	0.2
RRM2	63	*ESR1*	0.17
CENPA	63	**NDC80*	0.16
MELK	61	*CXCL12*	0.16
PBK	61	*NR1I2*	0.15
DTL	61	*IGF1*	0.12
BUB1B	61	*HBA1*	0.07
PRC1	60	*MSH2*	0.07
NUSAP1	60	*SPTBN2*	0.07
KIF2C	59	*CYP1A2*	0.06

Module analysis

MCODE algorithm analysis identified five modules that may represent functional molecular complexes in cell (Fig. 3). The modules are analysed using ClueGO plugin to reveal the GO terms and KEGG pathways enriched in each module. The ClueGO visualisation is designed like a PPI network where the nodes are the GO terms or KEGG pathways and the edges are the relationship among the terms and pathways. Nodes having more than one color indicate that more than one term are sharing the same set of genes. In this analysis, module 1 is shown to be involved in several functions and pathways, mainly mitotic cell cycle process, while the terms enriched in module 2-5 are specific to a particular function (Fig. 4).

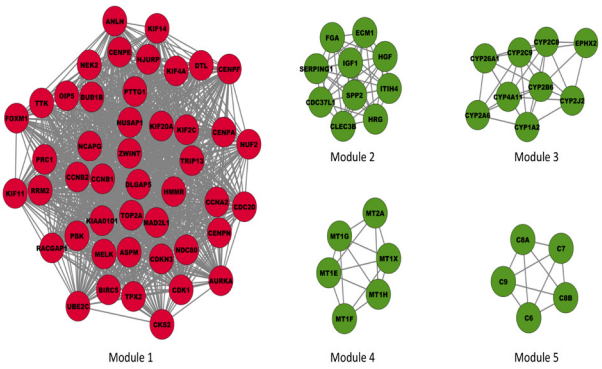

Figure 3: Modules obtained from the MCODE algorithm analysis

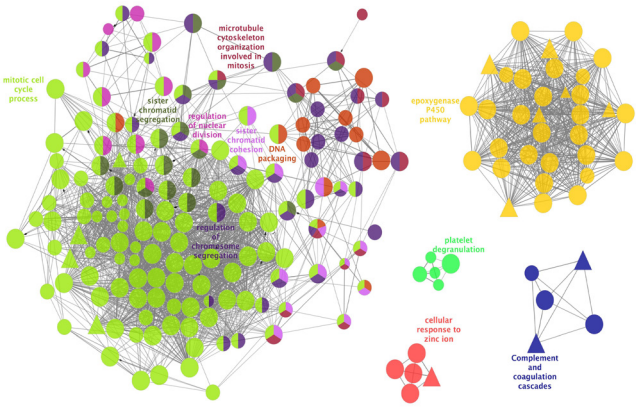

Figure 4: ClueGO network showing the interaction of enriched GO terms (eclipse) and KEGG pathways (triangle) of the five PPI network modules

DISCUSSION

The most widely used liver cancer biomarkers such as AFP and DCP failed to pass the cut-off criterion in this study, which might be caused by the differences in sample used in this analysis. In our selected microarray datasets, the samples used are from tissue biopsy while AFP and DCP are more commonly acted as serum biomarkers. Examples of common tissue-associated biomarkers include GPC3, HSP70 and TAG72 (22). In this study, GPC3 or Glypican-3, a heparan sulfate proteoglycan involved in cell growth regulation is observed to be among the top DEGs in three datasets (GSE84402, GSE60502 and GSE29721), validating its role as a potent liver cancer biomarker. In term of therapeutics target development, antibody-based therapies targeting GPC3 and other glypicans are being investigated in pre-clinical and clinical studies, with the goal of treating solid tumors that do not respond to standard therapies (23). Apart from GPC3, other top DEGs include TOP2A, ASPM, CCNB2, CDNK3, FCN3 and APOF that appeared in at least three datasets in this study, suggesting their potential as liver cancer biomarkers and therapeutics targets.

DAVID functional annotation and pathway analyses revealed that the up-regulated genes are mostly enriched in function related to mitosis and cell cycle pathways, which are very closely related to cancer. The overexpression of cell cycle related genes may be partly responsible in hepatocellular carcinomatosis due to uncontrolled cell division process, which is in agreement with previous findings that showed the perturbation in the cell cycle regulation is one of the main factors that give rise to cancer (24,25). Meanwhile, the down-regulated genes are involved in several terms and pathways at lower *p* value compared to the up-regulated genes. They mainly participated in diverse metabolism-associated signaling pathways for example xenobiotic metabolic process and epoxygenase P450 pathway. In treatment of cancer, expression of drug and xenobiotic metabolising enzymes (DXME) plays a

major role in patients' response and resistance to drug therapy. It does so by controlling the distribution of drug into the cells (26). There is a strong genomic instability that leads to highly variable expression of DXME, hence, understanding the behavior of DXME toward drugs could be beneficial for better precision medicine (27). Apart from metabolism-associated signaling pathways, ClueGO analysis revealed three more pathways enriched in down-regulated modules, which are platelet degranulation, complement and coagulation cascade and cellular response to zinc ion.

Topological analysis of PPI network identified TOP2A, CDK1 and NDC80 as the key genes in the network. TOP2A, or DNA topoisomerase II alpha is a cell cycle related gene that encodes a DNA topoisomerase, an enzyme that controls and alters the topologic states of DNA during transcription. It is reported that TOP2A is overexpressed in tumor tissues compared to adjacent non-tumor tissues (28). The gene is highly regulated at transcriptional and translational level suggesting that TOP2A overexpression may arise from aberration at these levels (29). It is located on chromosome 17 and is related to the onset of malignancy and chemoresistance, often shortening the survival time of liver cancer patients (30). Previous study by immunohistochemical staining method proved that liver cancer patients with lower TOP2A level index had significantly longer disease-free survival and exhibited a longer cumulative survival period than those with higher TOP2A level index (31). Another key gene identified is CDK1, the primary regulator of cell cycle from CDK family. The cell cycle process consisted of G1, S, G2 and M phase and functioned to regulate cell proliferation. CDK1 encodes cyclin-dependent kinase 1, known to have a major role in the control of G2-M transition, G1 progression and G1-S transition in eukaryotic cell (32). Previous study has demonstrated that CDK1 is upregulated in liver cancer and modulated the G2-M checkpoint in liver cancer cell cycle and cell proliferation (33). CDK family such as CDK1 and CDK2 has also been reported to be overly expressed in many types of cancers, such as lung cancer, colon cancer and adenomatous tissue carcinomas (34,35). The third key gene is NDC80, a coiled-coil protein critical for cell mitosis and holds a portfolio in chromosome segregation by interacting with several proteins that modulate the G2-M phase through its coiled-coil domains (36). NDC80 is overexpressed in a variety of human cancers such as gastric and breast cancer (37,38). Taken together, all these data suggests that TOP2A, CDK1 and NDC80 involved in the pathogenesis of liver cancer by affecting mitosis and cell cycle process, which generally agreed with our findings.

TOP2A and CDK1 have already been developed as targeted therapy of cancer. Targeting TOP2A gene using enzyme binders such as etoposide or DNA lesions can result in topoisomerase 2-mediated DNA damage and targeting TOP2A with the combination of etoposide with doxorubicin, a chemotherapy drug can triggered the cell death mechanism in liver cancer cells (39). Meanwhile, knockdown of CDK1 gene has shown to reduce cell proliferation in liver cancer cells (40). Various CDK1 inhibitors have been developed to treat cancer for example flavopiridol, olomoucine, and staurosporine (41). As for NDC80, various studies have showed a significant differential expression of this gene in cancer, hence, it could be nominated as a new target for gene therapy in treating liver cancer. In regards to biomarker development, the key genes identified in this study are also up-regulated in other diseases such as breast, colon and lung cancer hence their potential to be the biomarkers specific for liver cancer may not be ideal.

CONCLUSION

In summary, we have identified TOP2A, CDK1 and NDC80 as key genes in liver cancer using bioinformatics approach. These genes are closely related to cancer and have been studied as biomarkers and targeted therapy in cancer diagnosis and treatment. By performing the functional and pathway enrichment analyses, our study has provides information regarding the molecular mechanism underlying liver cancer in order to increase our understanding regarding liver cancer development and progression at molecular level.

ACKNOWLEDGEMENT

This work is supported by FRGS grant (FRGS/1/2014/SG05/USM/02/1) from Ministry of Higher Education, Malaysia.

REFERENCES

1. World Health Organization. Cancer [Online]. 2017. Available: http://www.who.int/mediacentre/factsheets/fs297/en/.
2. Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. CA Cancer J Clin. 2014;64:9-29.
3. Ius T. Detection of embryo-specific alpha-globulin in the blood serum of a patient with primary liver cancer. Vopr Med Khim. 1964;10:90-91.
4. Ju6rez-Hern6ndez E, Motola-Kuba D, Ch6vez-Tapia NC, Uribe M, Barbero Becerra V. Biomarkers in hepatocellular carcinoma: An overview. Expert Review of Gastroenterology & Hepatology. 2017;11:549-558.
5. Saffroy R, Pham P, Reffas M, Takka M, Lemoine A, Debuire B. New perspectives and strategy research biomarkers for hepatocellular carcinoma. Clin Chem Lab Med. 2007;45:1169-1179.
6. Marrero JA, Su GL, Wei W, Emick D, Conjeevaram HS, Fontana RJ, Lok AS. Des-gamma carboxyprothrombin can differentiate hepatocellular carcinoma from nonmalignant chronic liver disease in American patients. Hepatology. 2003;37: 1114-1121.
7. Bertino G, Neri S, Bruno CM, Ardiri AM, Calvagno

GS, Malaguarnera M, Clementi S, Bertino G, Di carlo I. Diagnostic and prognostic value of alpha-fetoprotein, des-gamma-carboxy prothrombin and squamous cell carcinoma antigen immunoglobulin M complexes in hepatocellular carcinoma. Minerva Med. 2011;102:363-371.
8. Josep M, Llovet MD, Jordi Bruix MD. Molecular targeted therapies in hepatocellular carcinoma. Hepatology. 2008;48(4):1312–1327.
9. Tarca AL, Romero R, Draghici S. Analysis of microarray experiments of gene expression profiling. American Journal of Obstetrics and Gynecology. 2006;195: 373-388.
10. Allison DB, Cui X, Page GP, Sabripour M. Microarray data analysis: From disarray to consolidation and consensus. Nat Rev Genet. 2006;7:55-65.
11. Zhao S, Fung-Leong WP, Bittner A, Ngo K, Liu X. Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells. PLoS ONE. 2014:9(1):e78644.
12. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P, Jensen LJ, von Mering C. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Research. 2017;45:D362-D368.
13. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13:2498-504.
14. Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics. 2003;4:2. doi: 10.1186/1471-2105-4-2.
15. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, Fridman WH, Pagиs F, Trajanoski Z, Galon J. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics. 2009;25(8):1091-1093. doi: 10.1093/bioinformatics/btp101.
16. Wang H, Huo X, Yang XR, He J, Cheng L, Wang N, Deng X, Jin H, Wang N, Wang C, Zhao F, Fang J, Yao M, Fan J, Qin W. STAT3-mediated upregulation of lncRNA HOXD-AS1 as a ceRNA facilitates liver cancer metastasis by regulating SOX4. Mol Cancer. 2017;14:16(1):136.
17. Makowska Z, Boldanova T, Adametz D, Quagliata L, Vogt JE, Dill MT, Matter MS, Roth V, Terracciano L, Heim MH. Gene expression analysis of biopsy samples reveals critical limitations of transcriptome-based molecular classifications of hepatocellular carcinoma. J Pathol Clin Res. 2016;2(2):80-92.
18. Wang YH, Cheng TY, Chen TY, Chang KM, Chuang VP, Kao KJ. Plasmalemmal Vesicle Associated Protein (PLVAP) as a therapeutic target for treatment of hepatocellular carcinoma. BMC Cancer. 2014;14:815.
19. Stefanska B, Huang J, Bhattacharyya B, Suderman M, Hallett M, Han ZG, Szyf M. Definition of the landscape of promoter DNA hypomethylation in liver cancer. Cancer Res. 2011;71(17):5891-5903.
20. Deng YB, Nagae G, Midorikawa Y, Yagi K, Tsutsumi S, Yamamoto S, Hasegawa K, Kokudo N, Aburatani H, Kaneda A. Identification of genes preferentially methylated in hepatitis C virus-related hepatocellular carcinoma. Cancer Sci. 2010;101(6):1501-1510.
21. Yu H, Kim PM, Sprecher E, Trifonov V, Gerstein M. The importance of bottlenecks in protein networks: Correlation with gene essentiality and expression dynamics. PLoS Comput Biol. 2007;3(4):e59. doi: 10.1371/journal.pcbi.0030059
22. Chauhan R, Lahiri N. Tissue- and serum-associated biomarkers of hepatocellular carcinoma. Biomark Cancer. 2016;8(Suppl 1):37–55.
23. Li N, Gao W, Zhang Y-F, Ho M. Glypicans as cancer therapeutic targets. Trends in Cancer. 2018;4(11):741-754.
24. Greenbaum LE. Cell cycle regulation and hepatocarcinogenesis. Cancer Biology & Therapy. 2004;3:1200-1207.
25. Kastan MB, Bartek J. Cell-cycle checkpoints and cancer. Nature. 2004;432(7015):316-323. doi: 10.1038/nature03097.
26. Rochat B. Importance of influx and efflux systems and xenobiotic metabolizing enzymes in intratumoral disposition of anticancer agents. Curr Cancer Drug Targets. 2009;9:652–674.
27. Li Y, Steppi A, Zhou Y, Mao F, Miller PC, He MM, Zhao T, Sun Q, Zhang J. Tumoral expression of drug and xenobiotic metabolizing enzymes in breast cancer patients of different ethnicities with implications to personalized medicine. Scientific Report. 2017;7:4747
28. Panvichian R, Tantiwetrueangdet A, Angkathunyakul N, Leelaudomlipi S. TOP2A Amplification and overexpression in hepatocellular carcinoma tissues. BioMed Research International. Epub 2015:381062. doi: 10.1155/2015/381602.
29. Isaacs RJ, Davies SL, Sandri MI, Redwood C, Wells NJ, Hickson ID. Physiological regulation of eukaryotic topoisomerase II. Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression. 1998;1400:121-137.
30. Wong N, Yeo W, Wong WL, Wong NL, Chan KY, Mo FK, Koh J, Chan SL, Chan AT, Lai PB, Ching AK, Tong JH, Ng HK, Johnson PJ, To KF. TOP2A overexpression in hepatocellular carcinoma correlates with early age onset, shorter patients survival and chemoresistance. Int J Cancer. 2009;124:644-652.
31. Watanuki A, Ohwada S, Fukusato T, Makita F, Yamada T, Kikuchi A, Morishita Y. Prognostic significance of DNA topoisomerase II alpha expression in human hepatocellular carcinoma. Anticancer research. 2002;22:1113-1119.

32. Qiao M, Shapiro P, Fosbrink M, Rus H, Kumar R, Passaniti A. Cell cycle-dependent phosphorylation of the RUNX2 transcription factor by cdc2 regulates endothelial cell proliferation. J Biol Chem. 2006;281:7118-28.
33. Li KK, Ng IO, Fan ST, Albrecht JH, Yamashita K, Poon RY. Activation of cyclin-dependent kinases CDC2 and CDK2 in hepatocellular carcinoma. Liver. 2002;22:259-268.
34. Evan GI, Vousden KH. Proliferation, cell cycle and apoptosis in cancer. Nature. 2001;411:342-348.
35. Geleta B, Makonnen E, Abay SM. Cyclic Dependent Kinase (CDK): Role in cancer pathogenesis and as drug target in cancer therapeutics. J Cancer Sci Ther 2016, 8:6
36. Chen Y, Riley DJ, Zheng L, Chen PL, Lee WH. Phosphorylation of the mitotic regulator protein Hec1 by Nek2 kinase is essential for faithful chromosome segregation. J Biol Chem. 2002;277:49408-49416.
37. Qu Y, Li J, Cai Q, Liu B. Hec1/Ndc80 is overexpressed in human gastric cancer and regulates cell growth. J Gastroenterol. 2014;49:408-418.
38. Glinsky GV, Berezovska O, Glinskii AB. Microarray analysis identifies a death-from-cancer signature predicting therapy failure in patients with multiple types of cancer. J Clin Invest. 2005;115:1503-1521.
39. Wong N, Yeo W, Wong WL, Wong NL, Chan KY, Mo FK, Koh J, Chan SL, Chan AT, Lai PB, Ching AK, Tong JH, Ng HK, Johnson PJ, To KF. TOP2A overexpression in hepatocellular carcinoma correlates with early age onset, shorter patients survival and chemoresistance. Int J Cancer. 2009;124:644-652.
40. Cai J, Li B, Zhu Y, Fang X, Zhu M, Wang M, Liu S, Jiang X, Zheng J, Zhang X, Chen P. Prognostic biomarker identification through integrating the gene signatures of hepatocellular carcinoma properties. EbioMedicine. 2017;19:18-30.
41. Wang Q, Su L, Liu N, Zhang L, Xu W, Fang H. Cyclin dependent kinase 1 inhibitors: a review of recent progress. Curr Med Chem. 2011;18:2025-2043.