ORIGINAL ARTICLE

Computational Prediction of miR130a Target Genes in Cancer

Nur Ainina Abdollah, Nurulisa Zulkifle, Siti Razila Abdul Razak

Oncological and Radiological Sciences Cluster, Advanced Medical & Dental Institute, Universiti Sains Malaysia, 13200 Kepala Batas, Penang, Malaysia.

ABSTRACT

Introduction: Cancer is one of the main causes of mortality globally and the incidence has been rising over the years. Studies have shown that miRNAs have the potential as cancer biomarkers. The miR-130a has been reported to be upregulated in several types of cancer, which indicate the important roles of miR-130a in cancer development and metastasis. The aim of this study is to identify potential target genes and to predict the regulatory function of miR-130a-3p and 5p in cancer. **Methods:** Three bioinformatics platforms namely miRWalk, the Database for annotations, visualization and integrated discovery (DAVID) Gene Functional Classification Tool and miRanda-miRSVR analysis tools were used to identify possible interaction between miR-130a and its target. Protein-protein interaction (PPI) network for the predicted target genes was then constructed. **Results:** The analyses have identified nine predicted target genes for miR-130a-3p (*RAPGEF4, SOS2, NRP1, RPS6KB1, MET, IL15, ACVR1, RYR2* and *ITPR1*), and ten for miR-130a-5p (*BCL11A, SPOPL, NLK, PPARGC1A, POU4F2, CPEB4, ST18, RSBN1L, ELF5* and *ARID4B*), that might play an important role in the development of cancer. Findings from this report suggest that miR-130a may involves in controlling cancer related genes; MET, ACVR1 and BCL11A. miR-130a-3p may regulates MET which involves in apoptosis and metastasis, and ACVR1 which involves in metastasis and angiogenesis. miR-130a-5p may regulates BCL11A which involves in apoptosis, proliferation and tumorigenesis. **Conclusion:** This study has highlighted the molecular interaction of miR-130a with associated genes and pathways, suggesting therapeutic potential of miR-130a as personalised targeted therapy for cancer.

Keywords: miR-130a, Cancer, Computational methods, miRNA target prediction

Corresponding Author:
Siti Razila Abdul Razak, PhD
Email: sitirazila@usm.my
Tel: +604-5622571

INTRODUCTION

Cancer is a main global public health problem. One in every six deaths worldwide is caused by cancer, making it the second highest cause of death in the world (1). Standard approach for management of cancer relies on the use of surgery, radiotherapy and chemotherapy. However, the adjuvant therapies including radiotherapy and chemotherapy are associated with numerous side effects which limiting compliance and outcome of treatment. Thus, there is a need for new treatment which use the knowledge about specific molecular targets and the pathogenesis of cancer. This approach need to address how the tumour cell is regulated by oncogenes and tumour suppressor genes which causing uncontrollable proliferation and metastasis via disruption of important cell-cycle regulators and cascades of signal transduction (2).

Current trends in the treatment of cancer have been shifted to the use of small molecule such as tubulin inhibitors which target at the microtubule dynamics, and interfere with critical cellular functions including cell signaling, angiogenesis, intracellular trafficking and mitosis (3). In addition, non-coding RNAs (ncRNAs) such as microRNA (miRNA) are one of the most studies regulatory ncRNAs to date. MiRNAs are short stranded noncoding RNA of ~22 nucleotides that participates in various human biological activities such as regulation of apoptosis, cell proliferation, cellular differentiation, development and immune responses (4). Studies showed that miRNAs contribute in the regulation of almost all cellular processes and that the changes in their expression observed has been shown to be involved in multi-step carcinogenesis in cancer (5). As an evidence, bioinformatics predictions suggested that miRNA may potentially regulate thousands of genes as 30% of all protein coding genes may be affected by miRNAs (6). miRNA has been shown to be involved whether as oncogene miRNA (oncomiR) or tumour suppressor gene in different pathways in cancer.

Several studies reported that miR-130a is an oncomiR as it was highly expressed in cancer cells including in gastric and cervical cancer (7-10). In gastric cancer cells, miR-130a high expression suppressed the expression of dual-luciferase reporter assays, collapsing

response mediator protein 4 (*CRMP4*) gene and runt-related transcription factor 3 (*RUNX3*) which enhanced cell proliferation and apoptosis, increases cell colony formation, invasion, migration, and adhesion (7, 8). In cervical cancer, upregulation of miR-130a modulates nuclear factor-kB (*NF-kB*) and tumour necrosis factor-alpha (*TNF-α*) which then target phosphatase and tensin homolog on chromosome 10 (*PTEN*) and leading to increase in cell proliferation in carcinogenesis of cancer cells (9, 10). On the other hand, miR-130a acts as tumour suppressor in lung cancer, breast cancer and nasopharyngeal cancer (11-15). In lung cancer, a study reported that the suppression of miR-130a regulates macrophage polarization in non-small cell lung cancer (NSCLC) (11). The result showed that miR-130a acted as a molecular switch as it expresses at higher level in M1 compared to M2 macrophages (11). Thus, this study concluded that downregulation of miR-130a in NSCLC resulted in poor prognosis, increased tumour stage and metastasis (11). The expression of miR-130a was shown to be a tumour suppressor in breast cancer cell line (12). As the expression of miR-130a was shown to be downregulated, the overexpression of miR-130a inhibits cell proliferation, invasion and migration of cancer cells by targeting at ras-related protein Rab-5A (*RAB5A*) and fos-related antigen 1 (*FOSL1*) (12, 13). Low expression of miR-130a-3p inhibits the expression of transcription regulator gene, *BACH2* which suppressing the viability, proliferation, invasion, cell cycle and promotes apoptosis in nasopharyngeal carcinoma cells (15).

Association between miRNAs and target genes can be predicted using several miRNA prediction databases. The platforms may be used together to complement each other in order to achieve the objective of predicting the target genes of miR130a. Hence, in this study, three bioinformatics platforms namely miRWalk, the Database for annotations, visualization and integrated discovery (DAVID) Gene Functional Classification Tool and miRanda-miRSVR analysis tools were used to identify possible interaction between miR130a and its target, which eventually will be used to understand the mechanisms of interactions in cancer. The use of this specific molecular targeted therapies are important to increase specificity of treatment as well as to increase the survival rate of the cancer patients as been shown by the miRNA-targeted therapeutics that have already reached phase I clinical trials for the treatment of mesothelioma and liver cancer (16).

MATERIALS AND METHODS

Combinatorial analysis

The target genes of miR-130a-5p and miR-130a-3p were predicted using combinatorial analysis by eight prediction programs namely DIANAmT, miRanda, miRDB, miRWalk, PICTAR4, PICTAR5, PITA and TargetScan in miRWalk database version 2, and was performed in June 2017 (17). Target genes that identified in not less than four databases were regarded as the predicted target genes of miR-130a-3p and miR-130a-5p.

Functional group analysis

Target genes with score of four and above obtained from comparative analysis in miRWalk were analysed in DAVID Gene Functional Classification Tool for gene ontology (GO) and pathway analysis as well as Kyoto Encyclopaedia of Genes and Genomes (KEGG). Both GO and KEGG analysis use two-sided Fisher's exact test and false discovery rate (FDR) that was calculated to the correct *P* value for the classifications. FDR < 0.05 was considered as significant.

miRanda – mirSVR analysis

Target genes with score of four and above obtained from comparative analysis in miRWalk were analysed in microrna.org for miRanda-miRSVR analysis. In the analysis, the efficiency of miRanda-predicted microRNA target sites uses the support vector machine (SVM) and were ranked by mirSVR score. The mirSVR score is a great tool for making predictions as it uses wide range of miRanda prediction rules into a single integrated model, without defining seed subclasses (18). In this study, the cut-off point for mirSVR score was < -1 as it represents the top 7% of miRSVR scores, where it has more than 35% probability of having the expected probability of observing a log expression change of at least -1 (18).

Combinatorial analysis of gene ontology, KEGG and miRanda – miRSVR analysis

Predicted target genes from gene ontology and KEGG analysis were compared with predicted target genes from miRanda-miRSVR analysis for both miR-130a-5p and miR-130a-3p. The overlapping target genes from all the analysis were then listed as the predicted target genes for miR-130a-5p and miR-130a-3p.

Protein-protein interaction network for predicted miR-130a-3p and miR-130a-5p target genes

Protein partners of each predicted miR-130a-3p and miR-130a-5p target genes were extracted from the International Molecular Exchange (IMEx) database (https://www.imexconsortium.org/). IMEx consortium is an international collaboration between major public interaction data providers which currently consists of DIP, IntAct, MatrixDB, MINT, MBInfo, I2D, InnateDB, UniProt, Molecular Connections, SIB, UCL-BHF and HPIDB as full members, and BioGRID and PrimesDB as observer members (19). Protein-protein interaction (PPI) network were then constructed and analysed in Cytoscape, an open source software platform for visualizing molecular interaction network (20). List of cancer related gene was obtained from Cancer Gene Census by the Catalogue of Somatic Mutations in Cancer (COSMIC) at https://cancer.sanger.ac.uk/cosmic (21).

RESULTS

Target genes prediction of miR-130a-3p and 5p by miRWalk combinatorial programs

In this study, eight established miRNA-target prediction programs namely DIANAmT (22), miRanda (23), miRDB (24), miRWalk, PICTAR4 (25), PICTAR5 (25), PITA (26), and TargetScan (6) were chosen based on popularity for combinatorial analysis. Total number of target genes identified were 19,058, where 2525 and 214 target genes which are redundant in four databases and more (score of > 4) were predicted for miR-130a-3p and miR-130a-5p, respectively.

Gene Ontology analysis

The functions of the predicted target gene of miR-130a-3p and miR-130a-5p were analysed separately using GO annotation with David Tool Analysis with FDR<0.05 as the cut-off standard. The GO annotation from David Tool Analysis had predicted 31, 13 and 15 terms for three ontologies consisting of biological processes (BP), cellular components (CC) and molecular functions (MF) respectively for miR-130a-5p. However, the only significant term was nucleus from CC with FDR of 9.1 x 10^3 (Table I). There were no significant terms for BP and MF. On the other hand, David Tool Analysis predicted that there were 99 and 114 terms respectively for two ontologies: CC and MF for miR-130a-3p. Seven most significant terms for CC are perinuclear region of cytoplasm, membrane, Golgi apparatus, cell surface, SNARE complex, Golgi apparatus and intracellular. Six most significant terms for MF were protein binding, metal ion binding, transcription factor activity sequence-specific DNA binding, transcriptional activity RNA polymerase II core promoter proximal region sequence specific binding, ATP binding and SNAP receptor activity (Table I). These terms suggest possible regulation of miR-130a-3p in the listed cellular components and molecular functions. The predicted target genes from BP were unable to be processed by the software due to the complexity of the data.

Table I: The enriched GO categories of predicted target genes for miR-130a-5p and miR130a-3p

Term	Count	Fold enrichment	FDR
miR-130a-5p Gene Ontology: Cellular Components			
GOTERM_CC_DIRECT~nucleus	52	1.90	7.50E-03
miR-130a-3p Gene Ontology: Cellular Components			
GO:0048471~perinuclear region of cytoplasm	133	1.85	2.58E-09
GO:0016020~membrane	331	1.30	2.02E-04
GO:0005794~Golgi apparatus	149	1.49	6.82E-04
GO:0009986~cell surface	100	1.59	3.74 E-03
GO:0031201~SNARE complex	20	3.26	5.56 E-03
GO:0000139~Golgi membrane	106	1.55	6.85 E-03
GO:0005622~intracellular	205	1.33	0.02
GO:0048471~perinuclear region of cytoplasm	133	1.85	2.58E-09
miR-130a-3p Gene Ontology: Molecular Function			
GO:0005515~protein binding	1214	1.15	1.95E-11
GO:0046872~metal ion binding	323	1.30	2.02E-04
GO:0003700~transcription factor activity, sequence-specific DNA binding	169	1.47	2.86E-04
GO:0001077~transcriptional activator activity, RNA polymerase II core promoter proximal region sequence-specific binding	55	1.94	2.87 E-03
GO:0005524~ATP binding	234	1.31	0.01
GO:0005484~SNAP receptor activity	16	3.42	0.04

(P < 0.05, FDR < 0.05). Count refers to number of genes

KEGG analysis

Target enrichment in KEGG was carried out to provide better understanding on the functions and regulatory networks of the predicted target genes of miR-130a-5p and miR-130a-3p. KEGG analysis from David Analysis Tool had predicted 400 target genes for miR130a-3p which involved in nine pathways as shown in Table II A with a cut-off standard of FDR < 0.05. The highest number of genes were recorded to be involved in pathways in cancer. There were also some important signalling pathways predicted and that include F*oxO, TGF-beta*, calcium, *cGMP-PKG* and *cAMP* signalling pathway. Predicted targets was also correlated with HTLV-I infection and melanogenesis. This suggest that miR-130-3p may involves in the tumorigenesis through participation in numbers of metabolic pathways in order to regulate the expression of target genes. However, no significant pathway was predicted as target genes for miR-130a-5p (Table IIA).

miRanda-miRSVR analysis

Target genes with of score four and above obtained from comparative analysis in miRWalk were analysed in microrna.org for miRanda-mirSVR analysis. The total number of targeted genes with miRSVR score < -1 for miR-130a-5p was 1,207 from total 8,302 and miR-130a-3p was 599 from total 7,753. It is important to note that alternative isoforms were also included in the targeted genes in miRanda-mirSVR analysis.

Comparing targeted genes from GO analysis with mirSVR score

List of targeted genes from the most significant terms (FDR < 0.05) from GO analysis were compared with list of targeted genes from miRanda-mirSVR analysis with score < -1. For miR-130a-5p, only nucleus from CC of GO analysis was significant. The targeted genes were then compared to the list of targeted genes from miRanda-mirSVR analysis. Result shows that there were ten targeted genes that gives an overlapping positive result namely *BCL11A, SPOPL, NLK, PPARGC1A, POU4F2, CPEB4, ST18, RSBN1L, ELF5* and *ARID4B* (Table III). For miR-130a-3p, predicted target genes for CC were compared to the list of targeted genes from miRanda-mirSVR analysis with score of < -1. Result

Table II: (A) The enriched pathways of predicted target genes for miR-130a-3p (P < 0.05 and FDR < 0.05). (B) The enriched predicted target genes for miR130a-3p KEGG pathway analysis (P <0.05 and FDR < 0.05) and miRanda-mirSVR analysis (mirSVR score < -1) for miR130a-3p

(A) miR-130a-3p KEGG pathways		
Term	Count	FDR
hsa04068:FoxO signaling pathway	42	1.82E-05
hsa05200:Pathways in cancer	83	8.38E-04
hsa05215:Prostate cancer	29	1.53E-03
hsa05166:HTLV-I infection	58	5.99E-03
hsa04350:TGF-beta signaling pathway	27	6.55E-03
hsa04916:Melanogenesis	30	8.06E-03
hsa04020:Calcium signaling pathway	44	0.01
hsa04022:cGMP-PKG signaling pathway	41	0.02
hsa04024:cAMP signaling pathway	46	0.04

(B) KEGG pathway and miRanda-mirSCR: miR-130a-3p	
Targeted genes	**mirSVR score 3p**
ATG12	-1.04
PRKAG2	-1.31
SOS2	-3.17
TGFBR2	-1.47
WNT1	-1.14
CASP8	-1.16
PLCB1	-1.27
PTGER2	-2.22
MET	-1.12
APPL1	-1.18
NRP1	-1.66
IL15	-1.01
MYBL1	-3.42
RPS6KB1	-1.15
ZFYVE9	-2.03
INHBB	-1.03
ACVR1	-2.7
BMP6	-1.14
CALM2	-1.2
STIM2	-1.3
ITPR1	-1.21
RYR2	-1.93
RAPGEF4	-1.22
TSHR	-1.27

shows that there were 85 target genes that gives a positive result. Result for molecular functions shows that there were 198 target genes that gives an overlapping positive result. It is interesting to note that some of the genes were also overlapped within different terms in cellular components and molecular functions.

Table III: The enriched predicted target genes for miR-130a-5p from GO analysis (P <0.05 and FDR < 0.05) and miRanda-mirSVR analysis (mirSVR score < -1)

miR-130a-5p GO: Cellular components (nucleus)	
Targeted genes	**mirSVR score 5p**
BCL11A	-2.77
SPOPL	-2.26
NLK	-2.41
PPARGC1A	-2.01
POU4F2	-1.96
CPEB4	-1.23
ST18	-1.2
RSBN1L	-1.2
ELF5	-1.19
ARID4B	-1.18

Comparing targeted genes from KEGG analysis with mirSVR score

miR-130a-3p KEGG pathways have a significant standard for FDR values, thus mirSVR scores were analysed for each target genes involved in nine pathways for miR-130a-3p. Results shows that there were 24 target genes out of 400 predicted target genes from KEGG pathway analysis which pass the cut-off standard (Table II B). From that number, eight target genes were overlapping with target genes from other pathways in this analysis namely *SOS2, TGFBR2, WNT1, PLCB1, ITPR1, PTGER2, RYR2* and *CALM2*.

Comparing targeted genes from GO with KEGG pathways analysis

In this study, predicted target genes between GO ontology, KEGG pathway and miRanda-miRSVR analysis were then overlapped to improve the predictive accuracy. Predicted target genes of CC and MF from GO analysis and KEGG analysis with mirSVRscore <-1 were compared. There were ten predicted target genes for miR-130a-5p which are *BCL11A, SPOPL, NLK, PPARGC1A, POU4F2, CPEB4, ST18, RSBN1L, ELF5* and *ARID4B*. The enriched predicted target genes were from the analysis which gives FDR < 0.05 for GO analysis and mirSVR score < -1 for miRanda-mirSVR analysis. There were nine predicted target genes for miR130a-3p which are *RAPGEF4, SOS2, NRP1, RPS6KB1, MET, IL15, ACVR1, RYR2* and *ITPR1*. The enriched predicted target genes were from the overlapping predicted target genes analysis which gives FDR < 0.05 for GO analysis (CC and MF), KEGG analysis and mirSVR score < -1 for miRanda-mirSVR analysis. Similar analysis was carried out for molecular functions and results shows 21 overlapping predicted target genes from both molecular functions of GO and KEGG pathways analysis namely *TGFBR2, RPS6KB1, MET, PRKAG2, MYBL1, NRP1, ZFYVE9, ACVR1, RYR2, INHBB, APPL1, IL15, STIM2,*

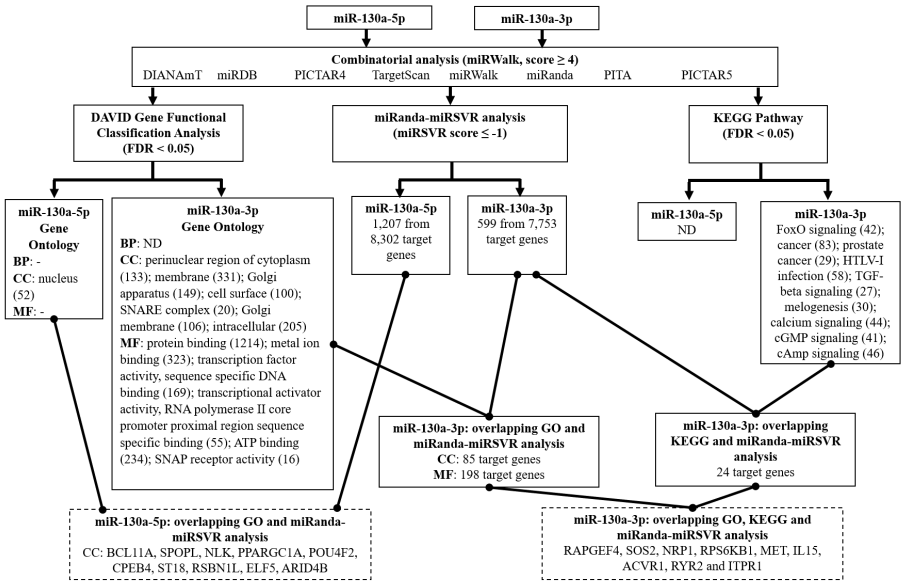

Figure 1: Summary of result from bioinformatics analysis using target gene predictions, combinatorial analysis of gene ontology (GO), miRanda-miRSVR and KEGG analysis for miR-130a-3p and 5p (biological process; BP, cellular components; CC, molecular function; MF).

CASP8, ATG12, SOS2, TSHR, RAPGEF4, ITPR1, CALM2 and *PLCB1*. Summary of result from bioinformatics analysis using target gene predictions, combinatorial analysis of gene ontology (GO), miRanda-miRSVR and KEGG analysis for miR-130a-3p and 5p is presented in Figure 1.

miR-130a-regulated protein-protein interaction networks

Following bioinformatics analysis to predict miR-130a target genes, protein-protein interaction (PPI) network for the predicted target genes was constructed in order to analyse the connection among the predicted target genes and to observe the topological characteristics of miR-130a-regulated PPI networks. Figure 2 (A) and (B) shows the PPI network for predicted miR-130a-3p and miR-130a-5p target genes with red nodes representing target genes and blue nodes are the protein partners for each target genes. In both networks, none of the predicted target genes are connected directly to each other. The closest relationship between the target genes and miR-130a can be observed in pairs in which they are connected by one node (protein) as a linkage. The *RPS6KB1-ITPR1, MET-SOS* and *MET-NRP* pairs are observed as target genes of miR-130a-3p while the *POU4F2-NLK, BCL11A-PPARGC1A, BCL11A-ARID4B* and *ARID4B-RSBN1L* pairs are observed as target genes of miR-130a-5p. The number of protein partners, or connectivity, for each target genes are varying with *MET* having the highest of 87 partners, followed by *POU4F2* with 79 partners, NLK with 41 partners, *RPS6KB1* with 23 partners, and 17 partners for both *ARID4B* and *PPARGC1A*. Meanwhile, *ACVR1* and *ST18* that interact with only one protein are the target genes with the lowest connectivity.

Cancer related gene was also identified from the networks and are represented by the V shape nodes. Figure 2 (A)

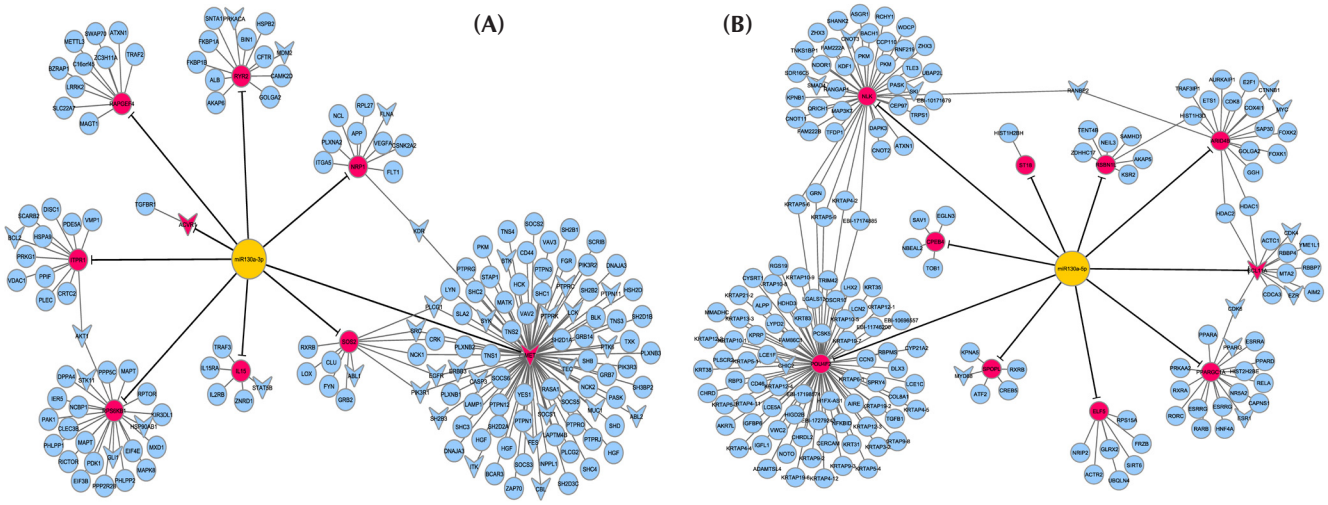

Figure 2: (A) PPI network of predicted target genes for miR-130a-3p: MET, RPS6KB1, RAPGEF4, SOS2, NRP1, IL15, ACVR1, RYR2 and ITPR1 (red nodes). The cancer related proteins are indicated by V shape nodes. **(B)** PPI network of predicted target genes for miR-130a-5p: POU4F2, NLK, PPARGC1A, ARID4B, BCL11A, RSBN1L, ELF5, SPOPL, CPEB4 and ST18 (red nodes). The cancer related proteins are indicated by V shape nodes.

and (B) shows that three of the predicted target genes themselves (*MET, ACVR1* and *BCL11A*) are the cancer related genes. In miR-130a-3p regulated PPI network, all the target genes are interacting with at least one cancer related protein partner, except for *RAPGEF4*. Meanwhile for miR-130a-5p regulated PPI network, P*OU4F2, NLK, SPOPL, PPARGC1A, BCL11A* and *ARID4B* are shown to be associated with cancer related protein while *CPEB4, ELF5, RSBN1L* and *ST18* are not.

DISCUSSION

The miRWalk algorithm is a computational program based on the Perl programming language where it recognises the multiple consecutive Watson-Crick complementary sequences between miRNA and gene sequences (17). The miRWalk algorithm walks on the complete sequence of a gene starting with at least heptamer (seven nucleotides) and extends it length until mismatch arises (17). A study has reported that miRWalk is the only database during that timepoint which offers the possible miRNA-binding sites within the complete sequence (promoter, CDS, 5'- and 3'UTR) of three genomes (human, mouse and rat) (17). Comparative studies of the earlier miRNA target prediction programs proved that there was no constant program that is better than all others (27, 28). The platforms may be used together to complement each other in order to achieve the objective of predicting the target genes of miR-130a. He *et. al* (2015) have argued that bioinformatics prediction programs main weakness is that it is not possible to incorporate different variants that may affect miRNA targeting and prediction outcomes in a single model (29). Therefore, a series of bioinformatics approaches were used to examine the underlying mechanism between miR-130a-3p and miR-130a-5p and target genes especially in this study.

miRWalk database provides a comparative analysis from miRWalk algoritm and clusters of different prediction programs namely DIANA-mT, miRanda, miRDB, RNAhybrid, PICTAR4, PICTAR5, PITA, RNA22 and TargetScan. miRWalk database then combines and analyse all the predicted target sites from different combinatorial databases (17). Every predicting programs are based on different matching criteria including base pairing, target accessibility and evolutionary conservation of target sites (30). As different predicting program uses different approaches, criteria and calculating algorithms, combining eight different predicting programs may improve the prediction precision for this study. It is reported that although there is low total overlaps between miRNA-target prediction programs, there is significant similarity between at least five databases (17). It is a common practice for researchers to look at the intersection of identified miRNA binding site produced by a numbers of different miRNA-target prediction programs (31, 32).

The GO ontology analysis indicated that the predicted target genes for miR-130a-3p were significantly enriched in perinuclear region of cytoplasm of CC and protein binding in MF. In addition, it also revealed that that the highest enrichment of CC for miR-130-5p is in nucleus (FDR < 0.05). The enrichment of these terms suggest that they have potential to be an important target gene for future study. In KEGG pathway analysis for miR-130-3p, the highest number of genes were enriched in cancer pathways. This suggest that miR-130-3p may involves in the tumorigenesis through participation in numbers of metabolic pathways in order to regulate the expression of target genes. However, no significant pathways were predicted for miR-130a-5p. miRanda-miRSVR analysis reported that there were 599 and 1,209 target genes for miR-130a-3p and miR-130a-5p (miRSVR score < -1).

Based on the predicted genes from bioinformatics analysis, the protein-protein interaction (PPI) of the predicted target genes was constructed using data from IMEx consortium database and analysed in Cytoscape platform. Since protein fulfil its functions through interaction and association with other protein, PPI is regarded as the functional components of living cell (33). On the other hand, miRNA is a well-known regulator at post-transcriptional level with copious evidence regarding its role in cancer regulation. Theoretically, miRNA influences the stability of PPI network by controlling protein abundance, either like an on-off switch to completely suppress protein expression or by adjusting the protein expression level according to cell types or conditions. In PPI network, protein with high connectivity is usually been targeted by many miRNAs. This is proved by previous study that shows genes with high number of miRNA target sites would have high connectivity at their protein level (34). This suggests that for proteins with more interacting partners, they may involve in various functions and require a tighter control hence their genes tend to be regulated by many miRNAs. Our result shows that *MET, POU4F2, NLK* and *RPS6KB1* are among the target genes that are highly connected (>20), hence we predicted that beside miR-130a-3p and 5p, these target genes could also be regulated by other types of miRNAs, most probably due to their functional complexity. As these target genes may have multiple miRNA target sites, it is also possible that some of their interacting proteins may be regulated by another miRNA apart from miR-130a-3p and 5p. Meanwhile, target genes such as *ACVR1* and *ST18* which have the lowest connectivity are predicted to be specific to miR-130a-3p and 5p regulation and have more specific functions.

MET, ACVR1, BCL11A are the predicted target genes listed in COSMIC as cancer related gene. This suggest that miR-130a-3p and miR-130a-5p regulate carcinogenesis via *MET*-, *ACVR1*- and *BCL11A*-mediated signaling pathways by controlling the expression of these genes. MET is a transmembrane tyrosine kinase that are

frequently overexpressed in various types of cancers. It is regulated by more than 30 types of miRNA, including miR-130a (35, 36). *MET* has been shown previously to play an important role in NSCLC (43) and colorectal cancer (44) through EGFR signaling pathway. In NSCLC, miR130a targets *MET* and reduces apoptosis-inducing ligand TRAIL resistance by downregulating miR-221 and miR-222. The downregulation of miR-221 and miR-222 also reduced NSCLC migratory capacity (36). Thus, the regulation of miR-130a-3p on MET has been shown to be involved in apoptosis and metastasis in NSCLC.

On the other hand, studies on miRNA regulation of *ACVR1* and *BCL11A* are not as extensive as *MET*. *ACVR1*, or Activin receptor 1, is shown to be a direct target of miR-384 and is involved in the inhibitory effects of miR-384 on breast cancer progression via Wnt/β-catenin signaling pathway (37). Furthermore, recent study by Xin *et al* has identified *ACVR1* as one of the target genes that are regulated by exosomal miR-455-5p and miR-1255a in breast cancer, which further highlight the significant of *ACVR1* in cancer regulation through miRNA (38). Recently, a report has highlighted that ACVR1 and TGFBR1 as important transmembrane receptors for bone morphogenetic proteins (BMP) which promotes metastasis, angiogenesis and lymph angiogenesis in gastric cancer (39). This report agrees with our findings in PPI network for miR-130a-3p which links TGFBR1 with cancer related protein, ACVR1 [Figure 3(A)]. These finding rule out our prediction regarding *ACVR1* could be specifically regulated by miR-130a-3p. The discrepancy could be due to incomplete PPI data, as research in human interactome is still currently expanding.

BCL11A involves in the regulation of C2H2 type zinc-finger protein which plays important role during the fetal to adult erythropoiesis transition (40). *BCL11A* expression is found to be elevated in *NSCLC* and regulated by miR-30a (41). *BCL11A* has been shown to play a critical regulator in lymphoid malignancies (42) and in triple-negative breast cancer (43). *BCL11A* plays primary role for lymphopoiesis and negatively regulates p53 activities as it regulates the expression of *BCL2, BCL2-XL* and *MDM2* in lymphoid development of B-cell chronic lymphocytic leukemia (CLL) (44). In triple-negative breast cancer (TNBC), *BCL11A* plays vital role for mammary stem and progenitor cells (43). Thus, this suggest that miR-130a-5p through *BCL11A* may directly regulates tumorigenesis and the stemness of cancer cells in TNBC and indirectly regulates apoptosis in CLL. Following this, further bioinformatics studies could be proposed to see if the 3′ UTR of *ACVR1* and *BCL11A* contained regions that matched the seed sequence of miR-130a to confirm that these genes are indeed targeted by miR-130a.

CONCLUSION

Protein-protein interaction analysis for miR-130a-3p shows that all of the predicted target genes were linked with one another. Therefore, downstream analysis can be carried out of the relationship between miR-130a-3p and predicted target genes of the other genes. On the other hand, protein-protein interaction analysis for miR-130a-5p shows that only four of the predicted target gene of miR-130a-5p associated with one another namely *NLK, PPARGC1A, BCL11A* and *ARID4B*. Thus, downstream analysis should focus on these targeted genes. Future analysis for experimental validation between miR-130a-3p, miR-130a-5p and its associated target genes are important to further understand their involvement in the regulation of carcinogenesis. In addition, all of these target genes were reported to regulate different types of cancer, hence suggesting that the computational analyses that were employs in this study could be used as a guideline in identifying target gene, and potentially be explored to be used as therapeutic target for cancer treatment.

ACKNOWLEDGEMENTS

The study was supported by Universiti Sains Malaysia through Short Term Grant (304.CIPPT.6313203).

REFERENCES

1. Naghavi M, Abajobir AA, Abbafati C, Abbas KM, Abd-Allah F, Abera SF, Aboyans V, Adetokunboh O, Afshin A, Agrawal A, Ahmadi A, Ahmed MB, Aichour AN, Aichour MTE, Aichour I, Aiyar S, Alahdab F, Al-Aly Z, Alam K, Alam N, Alam T, Alene KA, Al-Eyadhy A, Ali SD, Alizadeh-Navaei R. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the Global Burden of Disease Study 2016. The Lancet. 2017;390(10100):1151-210.
2. Fong KM, Sekido Y, Minna JD. Molecular pathogenesis of lung cancer. The Journal of thoracic and cardiovascular surgery. 1999;118(6):1136-52.
3. Arnst KE, Banerjee S, Chen H, Deng S, Hwang DJ, Li W, Miller DD. Current advances of tubulin inhibitors as dual acting small molecules for cancer therapy. Medicinal Research Reviews. 2019;39(4).
4. He L, Hannon GJ. MicroRNAs: small RNAs with a big role in gene regulation. Nature Reviews Genetics. 2004;5:522-31.
5. Filipowicz W, Bhattacharyya SN, Sonenberg N. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? . Nature Reviews Genetics. 2008;9:102-14.
6. Lewis BP, Burge CB, Bartel DP. Conserved seed

pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell. 2005;120(1):15-20.
7. Lee SH, Jung YD, Choi YS, Lee YM. Targeting of RUNX3 by miR-130a and miR-495 cooperatively increases cell proliferation and tumor angiogenesis in gastric cancer cells. Oncotarget. 2015;6(32):33269.
8. Zhou Y, Li R, Yu H, Wang R, Shen Z. microRNA-130a is an oncomir suppressing the expression of CRMP4 in gastric cancer. OncoTargets and therapy. 2017;10:3893.
9. Feng Y, Zhou S, Li G, Hu C, Zou W, Zhang H, Sun L. Nuclear factor-κB–dependent microRNA-130a upregulation promotes cervical cancer cell growth by targeting phosphatase and tensin homolog. Archives of biochemistry and biophysics. 2016;598:57-65.
10. Zhang J, Wu H, Li P, Zhao Y, Liu M, Tang H. NF-κB-modulated miR-130a targets TNF-α in cervical cancer cells. Journal of translational medicine. 2014;12(1):155.
11. Lin L, Lin H, Wang L, Wang B, Hao X, Shi Y. miR-130a regulates macrophage polarization and is associated with non-small cell lung cancer. Oncology reports. 2015;34(6):3088-96.
12. Pan Y, Wang R, Zhang F, Chen Y, Lv Q, Long G, Yang K. MicroRNA-130a inhibits cell proliferation, invasion and migration in human breast cancer by targeting the RAB5A. International journal of clinical and experimental pathology. 2015;8(1):384.
13. Kong X, Zhang J, Li J, Shao J, Fang L. MiR-130a-3p inhibits migration and invasion by regulating RAB5B in human breast cancer stem cell-like cells. Biochemical and biophysical research communications. 2018;501(2):486-93.
14. Chen X, Zhao M, Huang J, Li Y, Wang S, Harrington CA, Qian DZ, Sun XX, Dai MS. microRNA-130a suppresses breast cancer cell migration and invasion by targeting FOSL1 and upregulating ZO-1. Journal of cellular biochemistry. 2018;119(6):4945-56.
15. Chen X, Yue B, Zhang C, Qi M, Qiu J, Wang Y, Chen J. MiR-130a-3p inhibits the viability, proliferation, invasion, and cell cycle and promotes apoptosis of nasopharyngeal carcinoma cells by suppressing BACH2 expression. Bioscience reports. 2017:BSR20160576.
16. Slaby O, Laga R, Sedlacek O. Therapeutic targeting of non-coding RNAs in cancer. Biochemical Journal. 2017;474(24):4219-51.
17. Dweep H, Sticht C, Pandey P, Gretz N. miRWalk – Database: Prediction of possible miRNA binding sites by "walking" the genes of three genomes. Journal of Biomedical Informatics. 2011;44(5):839-47.
18. Betel D, Koppal A, Agius P, Sander C, Leslie C. Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites. Genome Biology. 2010;11(8).
19. Orchard S, Kerrien S, Abbani S, Aranda B, Bhate J, Bidwell S, Bridge A, Briganti L, Brinkman FSL, Cesareni G, Chatr-aryamontri A, Chautard E, Chen C, Dumousseau M, Goll J, Hancock REW, Hannick LI, Jurisica I, Khadake J, Lynn DJ, Mahadevan U, Perfetto L, Raghunath A, Ricard-Blum S, Roechert B, Salwinski L, Stьmpflen V, Tyers M, Uetz P, Xenarios I, Hermjakob H. Protein Interaction Data Curation - The International Molecular Exchange Consortium (IMEx). Nature Methods. 2012;9(4):345-50.
20. 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 Research. 2003;13(11):2498-504.
21. Sondka Z, Bamford S, Cole CG, Ward SA, Dunham I, Forbes SA. The COSMIC Cancer Gene Census: describing genetic dysfunction across all human cancers. Nature Reviews Cancer. 2018;18(11):696-705.
22. Kiriakidou M, Nelson PT, Kouranov A, Fitziev P, Bouyioukos C, Mourelatos Z, Hatzigeorgiou A. A combined computational-experimental approach predicts human microRNA targets. Genes & Development. 2004;18(10):1165-78.
23. Enright AJ, John B, Gaul U, Tuschl T, Sander C, Marks DS. MicroRNA targets in Drosophila. Genome Biology. 2004;5(1):R1.
24. Wang X. miRDB: a microRNA target prediction and functional annotation database with a wiki interface. RNA. 2008;14(6):1012-7.
25. Krek A, Grьn D, Poy MN, Wolf R, Rosenberg L, Epstein EJ, MacMenamin P, Piedade Id, Gunsalus KC, Stoffel M, Rajewsky N. Combinatorial microRNA target predictions. Nature Genetics. 2005;37:495-500.
26. Kertesz M, Iovino N, Unnerstall U, Gaul U, Segal E. The role of site accessibility in microRNA target recognition. Nature Genetics. 2007;39:1278-84.
27. Rajewsky N. microRNA target predictions in animals. Nature Genetics. 2006;38(S8-S13).
28. Sethupathy P, Megraw M, Hatzigeorgiou AG. A guide through present computational approaches for the identification of mammalian microRNA targets. Nature Methods. 2006;3(881-886).
29. He S, Zeng S, Zhou Z-W, He Z-X, Zhou S-F. Hsa-microRNA-181a is a regulator of a number of cancer genes and a biomarker for endometrial carcinoma in patients: a bioinformatic and clinical study and the therapeutic implication. Drug Design, Development and Therapy. 2015;9:1103-75.
30. Liu B, Li J, Cairns MJ. Identifying miRNAs, targets and functions. Briefings in Bioinformatics. 2014;15(1):1-19.
31. Megraw M, Sethupathy P, Corda B, Hatzigeorgiou AG. miRGen: a database for the study of animal microRNA genomic organization and function.

Nucleic Acids Research. 2007;35(Database issue).
32. Pandey P, Brors B, Srivastava PK, Bott A, Boehn SN, Groene H-J, Gretz N. Microarray-based approach identifies microRNAs and their target functional patterns in polycystic kidney disease. BMC Genomics. 2008;9(624).
33. Hartwell LH, Hopfield JJ, Leibler S, Murray AW. From molecular to modular cell biology. Nature. 1999;402:C47-C52.
34. Liang H, Li W-H. MicroRNA regulation of human protein–protein interaction network. RNA. 2007;13(9):1402-8.
35. Zhang J, Babic A. Regulation of the MET oncogene: molecular mechanisms. Carcinogenesis. 2016;37(4):345-55.
36. Acunzo M, Visone R, Romano G, Veronese A, Lovat F, Palmieri D, Bottoni A, Garofalo M, Gasparini P, Condorelli G, Chiariello M, Croce CM. miR-130a targets MET and induces TRAIL-sensitivity in NSCLC by downregulating miR-221 and 222. Oncogene. 2012;31(5):634-42.
37. Wang Y, Zhang Z, Wang J. MicroRNA-384 inhibits the progression of breast cancer by targeting ACVR1. Oncology Reports. 2018;39(6):2563-74.
38. Xin Y, Wang X, Meng K, Ni C, Lv Z and Guan D. Identification of exosomal miR-455-5p and miR-1255a as therapeutic targets for breast cancer. Bioscience Reports. 2020; 40 (1) BSR20190303.
39. Sun Z, Liu C, Jiang WG and Ye L. Deregulated bone morphogenetic proteins and their receptors are associated with disease progression of gastric cancer. Computational and Structural Biotechnology Journal. 2020; 18: 177-188.
40. Smith EC, Luc S, Croney DM, Woodworth MB, Greig LC, Fujiwara Y, Nguyen M, Sher F, Macklis JD, Bauer DE, Orkin SH. Strict in vivo specificity of the Bcl11a erythroid enhancer. Blood. 2016; 128(19)2338-2342.
41. Jiang B-y, Zhang X-c, Su J, Meng W, Yang X-n, Yang J-j, Zhou Q, Chen Z-y, Chen Z-h, Xie Z, Chen S-l, Wu Y-l. BCL11A overexpression predicts survival and relapse in nonsmall cell lung cancer and is modulated by microRNA-30a and gene amplification. Molecular Cancer. 2013;12(61).
42. Satterwhite E, Sonoki T, Willis TG, Harder L, Nowak R, Arriola EL, Liu H, Price HP, Gesk S, Steinemann D and Schlegelberger B. The BCL11 gene family: involvement of BCL11A in lymphoid malignancies. Blood, The Journal of the American Society of Hematology. 2001; 98(12) pp.3413-3420.
43. Khaled WT, Lee SC, Stingl J, Chen X, Ali HR, Rueda OM, Hadi F, Wang J, Yu Y, Chin SF and Stratton M. BCL11A is a triple-negative breast cancer gene with critical functions in stem and progenitor cells. Nature communications. 2015; 6(1) 1-10.
44. Yu Y, Wang J, Khaled W, Burke S, Li P, Chen X, Yang W, Jenkins NA, Copeland NG, Zhang S and Liu P. Bcl11a is essential for lymphoid development and negatively regulates p53. Journal of Experimental Medicine. 2012; 209(13) 2467-2483.