Sains Malaysiana 47(12)(2018): 3095–3105
http://dx.doi.org/10.17576/jsm-2018-4712-20

Prediction of Colorectal Cancer Driver Genes from Patients' Genome Data

(Penentuan Gen Pemandu Kanser Kolorektum daripada Data Genomik Pesakit)

MUHAMMAD-IQMAL ABDULLAH & NOR AZLAN NOR MUHAMMAD*

ABSTRACT

Colorectal cancer refers to the cancer that occurs in the colon and rectum. It has been established as the third most common cancer and the forth one in causing worldwide mortality. Cancer caused by the mutation of several genes that usually involved in the regulation of cell proliferation, growth and cell death. The mutation that leads to abnormal function of genes, either in enabling the genes to gain or loss of function was termed as driver mutation and the genes with driver mutation ability was termed as driver genes. The identification of driver genes provides insight on mechanistic process of cancer development where this information can be used to further understand their mode of action for causing dysregulation in signaling pathways. In this study, two bioinformatic tools, i.e. CGI and iCAGES were used to predict potential driver genes from the genome of eight colorectal cancer patients with annotated variants datasets. 44 unique driver genes and 21 pathways have been identified; such as p53 signaling, PI3K-AKT, Endocrine resistance, MAPK and cell cycle pathways. The identification of these pathways can lead to the identification of potential drugs targeting these pathways.

Keywords: Cancer driver genes; colorectal cancer; pathway analysis; precision medicine

ABSTRAK

Kanser kolorektum adalah kanser yang berlaku pada kolon dan juga rektum. Ia merupakan kanser ketiga yang paling kerap dilaporkan serta yang keempat menjadi punca kematian tertinggi di seluruh dunia. Kanser berlaku disebabkan oleh mutasi daripada gen tertentu yang terlibat di dalam pengawalaturan proliferasi, pertumbuhan dan kematian sel. Mutasi yang mencetuskan perubahan pada fungsi gen sama ada meningkatkan atau menghilangkan fungsinya dikenali sebagai mutasi pemandu, manakala gen yang mempunyai mutasi pemandu dikenali sebagai gen pemandu kanser. Pengenalpastian gen pemandu membolehkan mekanisme pembentukan kanser dapat dikenal pasti melalui peranannya dalam pengawalaturan tapak jalan. Dalam kajian ini, dua perisian bioinformatik iaitu CGI dan iCAGES telah digunakan untuk mengenal pasti calon gen pemandu daripada set data varian yang telah dianotasi daripada genom lapan pesakit kanser kolorektum. Sebanyak 44 gen pemandu dan 21 tapak jalan terlibat telah dikenal pasti; antaranya adalah tapak jalan pengisyaratan p53, PI3K-AKT, rintangan endokrin dan MAPK. Kesemua tapak jalan ini berpotensi untuk dijadikan sasaran terapi.

Kata kunci: Analisis tapak jalan; gen pemandu kanser; kanser kolorektum; perubatan jitu

INTRODUCTION

Colorectal cancer (CRC) is the third most common cancer involving 2.4 million people per year worldwide. According to the age-standardized rate, the incidence of colorectal cancer in Malaysia showed higher occurrence in men by 20.9 cases per 100,000 compared to in women at 16.8 per 100,000 populations (Chye et al. 2008). Majority of the cases occurred in Chinese ethnic followed by the Malays and Indians (Radzi et al. 2016). Dietary habit has been listed as a main factor that affects intestinal macro and microenvironment. Consumption of food that affects the intestinal lining gave rise to the development of polyps. It was due to the harmful chemicals that are able to penetrate the cells and disturb genome stability by a process known as mutation (Kanner 2007; O'Keefe 2016; Ryan-Harshman & Aldoori 2007). The examples of bad dietary habit include consumption of carcinogenic foods, lack of dietary fiber in the diet that lead to the intestinal bleeding as well as a disturbance to the microbiome environment and probiotic metabolisms (Louis et al. 2014).

Consumption of food high in fiber enables gut microbiota to produce short-chain fatty acid such as butyrate during fermentation. Butyrate has an anti-inflammatory and antineoplastic properties; and it can also provide the best energy source for colonocytes. Therefore, the risk of intestinal bleeding will be reduced as inflammation is suppressed by the anti-inflammatory substances produced by gut microbiota. Imbalance diet such as high consumption of protein and fat will produce substances with carcinogenic effects such as branched fatty acids that may induce inflammation and hydrogen sulfide that could cause DNA damage and lead to mutation if the repair mechanism is disrupted (O'Keefe 2016). These mutations may eventually affect the genes that are

necessary to regulate cellular growth and to maintain the balance of cell proliferation.

Cancer driver genes are defined as genes that provide selective advantage for the cell to grow more rapidly as the proliferation and death-escaping characteristics of the cell are enhanced. These properties enable cells to grow more rapidly than their neighbouring cell in a specific condition and environment (Tokheim et al. 2016). Cancer driver genes such as APC (adenomatous polyposis coli) gene plays a critical role in regulating Wnt signaling pathway that responsible for cell proliferation. APC forms a destruction complex with several other proteins to inhibit ß-catenin, a protein that activates Wnt signaling and preventing the cells to proliferate continuously. Mutation in APC may affect its function to perform destruction complex and unable to control cell proliferation (Kwong & Dove 2009; Zhan et al. 2017). Driver genes can be mutated at both protein coding and non-coding regions but more commonly at the protein coding region. This mutation causes abnormal function of the genes and is known as driver mutation (Foo et al. 2015). The term 'driver' is introduced to distinguish it from a passenger gene that does not contribute to any abnormal cell growth. Passenger genes could be any genes or recorded as driver genes in TCGA database, however does not carry any driver mutation that affect their structure and function (Marx 2014).

Identification of cancer driver genes is one of the important aspects in whole genome sequencing projects. Two prominent consortiums working on the collection of cancer genes data are Cancer Genome Atlas (TCGA) and International Cancer Genome Consortium (ICGC). The top driver genes that currently reported contribute to colorectal carcinogenesis are *APC, TP53* and *KRAS*. *APC* acts as tumor suppressor that form 'destruction complex' with Axin and GSK-3β to degrade β-Catenin in Wingless (Wnt) signaling pathway. Mutation which could lead to loss function of APC allows β-catenin keep acting as signal to promote cell proliferation through Wnt signaling pathway (Alberts 2008; Kwong & Dove 2009). TP53 is activated by cellular stresses including DNA damage and acts as tumor suppressor that bind to regulatory regions of downstream target genes to induce cell cycle arrest, DNA repair and apoptosis (Iacopetta et al. 2006). Inactivation of TP53 is associated with later stage in cancer as the role of TP53 as an important upstream key regulator to control the number of cancer cells through apoptosis become malfunction (Rivlin et al. 2011). Proto-oncogene is a gene that normally has a role in activating or continuing cell proliferation, growth and differentiation activity. Mutation in this type of genes could lead to gain of function and transform proto-oncogenes to new role in driving cancer known as oncogenes. Oncogene such as *KRAS* could drive the cancer through mutation at specific binding sites. This could prevent *KRAS* to degrade and enable the pathways continuously run to induce cell growth and proliferation.

Precision medicine in cancer requires the information of cancer driver genes in order to perform specific therapy to target altered function of component in signaling pathways. Genes like KRAS has been targeted through biologic agents like cetuximab and panitumumab in clinical trials (Tan & Du 2012).

Various bioinformatics tools for driver genes prediction have been developed to provide analysis tools for tumor sequencing project to distinguish driver and passenger genes; amongst other are SIFT (Sorting Intolerants from Tolerant) and Polyphen 2 (Polymorphism Phenotyping v2) (Adzhubei et al. 2015; Sim et al. 2012), HotNet2, MUFFINN (Mutations For Functional Impact on Network Neighbors) (Cho et al. 2016; Leiserson et al. 2014) and MuSIC (Mutational Significance in Cancer) (Dees et al. 2012). Here, two computational methods were used, i.e. Integrated Cancer Genome Score (iCAGES) and Cancer Genome interpreter (CGI). Both tools allowed for a single patient-based data analysis in determining high confidence prediction on driver genes via a combination of different tools that use different concept of systematic approaches.

iCAGES analyses the impact of variants using the scores obtained from eleven prediction tools that predict the impact of variants on the protein structure and function. Identification of cancer driver genes uses radial support vector machine (radial svm), a machine learning algorithm that combines all scores from the prediction tools. These tools were SIFT, PolyPhen 2, GERP++, Mutation Assessor, SiPhy, LRT, PhyloP, FATHMM, Mutation Tester, CADD and VEST. Driver gene prioritization unified radial SVM score with other bioinformatic tool known as Phenolyzer that predicts driver genes based on the oncogenicity of point coding mutation. Linear regression (LR) model is built to determine driver gene score with 0.11 as binary cut off for suggested cancer driver genes.

CGI determine the characteristics of cancer driver genes through the classification of genes based on their tumorigenic mode (Oncogene or Tumor Suppressor). Functional impact according to types of mutation, and occurrence among human population. CGI search and match query variants with TCGA database of cancer genes. Variants with no matches in TCGA library will undergo protein affecting mutation (PAM) assessment by algorithm known as Oncodrive-MUT. Oncodrive-MUT was built based on the random forest machine learning and classification algorithm that separate driver and passenger genes mutation by 0.78 accuracy based on Matthews Correlation Coefficient.

Materials and Methods

Data source

Variant annotated datasets (VCF) were obtained from eight patients (patient identification: C187, C194, C273, C373, C404, C414, C449, C474) involved in the CRC patient genome sequencing project led by UKM Medical Molecular

Medicine Institute (UMBI). Illumina Hiseq (30× coverage) was used for the sequencing and Gene Analysis Toolkit (GATK) was used in variant calling.

PREDICTION OF CRC DRIVER GENE USING BIOINFORMATICS TOOLS

Data were filtered by selecting all variants that occur at the exonic region as they are required as input in both tools. In iCAGES, the analysis was performed *via* http://icages.wglab.org/ using hg19 as a reference genome. The required input format consists of the information on the chromosome, coordinates, references and alternate nucleotides. The output for the driver genes prediction consists of a list of genes with radial SVM score, funseq, cnv, phenolyzer and iCAGES gene score. As this analysis was focused on the nonsynonymous variants that occur at the coding region, the funseq and cnv scores were set at 0 and not included in the driver genes prediction. Predicted driver genes with gene score more than 0.11 were selected for the comparison with CGI results.

CGI analysis was performed on https://www.cancergenomeinterpreter.org and colon adenocarcinoma was selected as a reference for the cancer type. The input files must include all information on the chromosomes, coordinates and nucleotide changes. The genes that are stated as a driver should consist of tumor suppressor genes where loss of function is detected and oncogenes which gain of function is likely to be occurred in certain variants. In this analysis, consensus results from both tools were chosen and considered as highly confidence predicted driver genes (HCDG) (Figure 1).

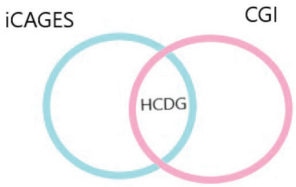

FIGURE 1. Selection of high confidence driver genes (HCDG) was based on the consensus result by iCAGES and CGI

PATHWAY ANALYSIS

KEGG MAPPER in combination with the information obtained from the bibliomic search were applied in identifying all pathways related to the predicted driver genes (Figure 2). Pathway mapping was performed using relevant KEGG pathway as a template in mapping the genes and to visualize the pathways that are likely to be affected by the driver genes was done using Cytoscape ver 3.6.0. Gene symbol of unique driver genes were set as an input in KEGG Mapper and the results were downloaded in KGML format for visualization.

RESULTS AND DISCUSSION

CANCER DRIVER GENES PREDICTION

The consensus predicted genes obtained from iCAGES and CGI were named as high confidence driver genes (HCDG). Forty-four HCDG were identified using our approach (Table 1). With this approach too, we have identified some genes that previously have not been listed in TCGA COADREAD and CGC Colorectal cancer genes dataset and have not been previously reported to be associated with colorectal cancer. Those genes were *NCOR2*, *STAG2*, *PTPN12*, *TSC2*, *MYH9*, *NCOR1*, *PTCH1*, *ATF2* and *ERBB3*. These genes are potential candidates for newly identified colorectal cancer driver genes.

PATHWAYS ASSOCIATED WITH PREDICTED DRIVER GENES

Nonsynonymous variants that occurred at the exonic region were selected as an input for iCAGES and CGI. Predicted driver genes from both tools were named as HCDG (high confidence driver genes) and their respective pathways were identified from KEGG database and literature (Table 2).

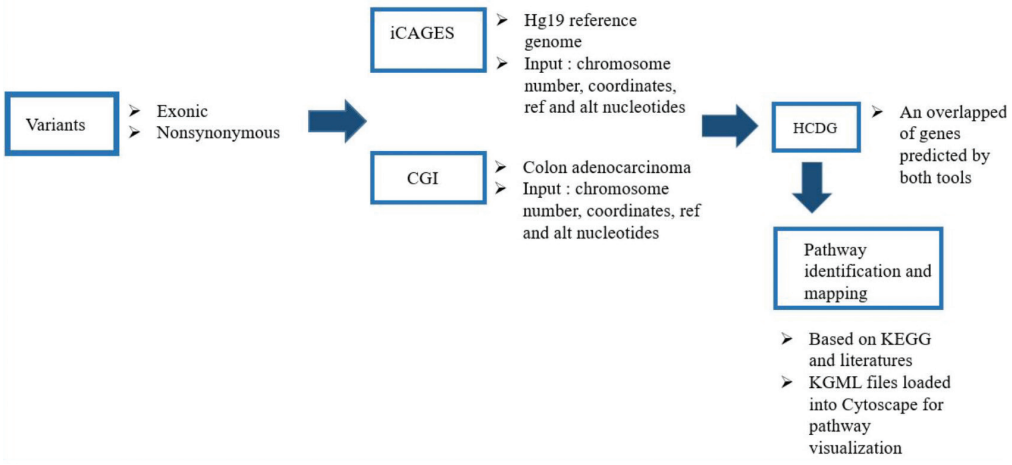

FIGURE 2. A flowchart for the prediction of colorectal cancer driver genes

TABLE 1. Consensus predicted genes obtained from iCAGES and CGI methods are considered as highly confident driver genes (HCDG). Genes with bold characters are CRC driver genes listed in TCGA – COADREAD. Genes with * sign are listed in CGC as CRC driver genes

	Patient ID	C187	C194	C273	C373	C404	C414	C449	C474
	Gene name								
1	*TP53**	■	■		■	■	■		
2	*FBXW7**		■						
3	*POLR2B**	■							
4	*SRGAP3**	■							
5	*SMC1A*	■							
6	*KRAS**			■		■	■		
7	*ITSN1*			■					
8	*TCF7L2**				■	■			
9	*CUL1*				■	■		■	
10	*ARID1A**						■		
11	*BRAF**							■	■
12	*STAG2*								■
13	*BMPR2*								■
14	*MYH10*								■
15	*ERCC2**			■					
16	*CDH1**		■		■				
17	*CHEK2**		■						
18	*NOTCH2**		■		■				
19	*SIN3A*				■				
20	*HDAC3*				■				
21	*CLTC**				■				
22	*TCF4*				■				
23	*RAD54L*				■				
24	*PHOX2B**						■		
25	*HLA-A*					■			
26	*NTRK1*						■		
27	*ERBB3*						■		
28	*ATF2*						■		
29	*PTCH1*						■		
30	*NCOR1*	■	■			■	■		■
31	*MYH9*		■						
32	*TSC2*							■	
33	*ERBB4*							■	
34	*PTPN12*							■	
35	*KLF5*	■							
36	*TSC1*								■
37	*SPTAN1*								■
38	*HGF*								■
39	*HSP90AB1*								■
40	*WNT5A*								■
41	*MAP3K1*			■					
42	*RAX*								■
43	*CDC27*			■					
44	*NCOR2*								■

TABLE 2. List of predicted potential driver genes, their involvement in cellular pathways and previous reports on their occurrence in various cancers. Genes with bold characters are CRC driver genes listed in TCGA – COADREAD. Genes with * sign are listed in CGC as CRC driver genes

Driver Genes	Pathways	Cancer types	References
*TP53**	p53 signaling	Colorectal cancer Melanoma Breast Cancer	(Petitjean et al. 2007) (Ragnarsson-Olding et al. 2002)
*FBXW7**	-	Colorectal cancer	(Iwatsuki et al. 2010)
*SRGAP3**	G-protein signaling	Colorectal cancer Gastric cancer	(Huang et al. 2015)
*KRAS**	Ras, PI3K-AKT, MAPK signaling	Colorectal cancer Bladder cancer Non-small cell lung cancer Breast Cancer	(Phipps et al. 2013) (Tian et al. 2013)
*TCF7L2**	Wnt signaling	Colorectal cancer	(Slattery et al. 2009) (Folsom et al. 2008)
*ARID1A**	Chromatin organization pathway	Colorectal cancer Pancreatic cancer	(Mathur et al. 2017) (Waddell et al. 2015)
*BRAF**	MAPK	Colorectal cancer Melanoma Glioma	(Davies et al. 2002) (Fang & Richardson 2005) (Chen et al. 2014)
ITSN1	RET Signaling	Lung cancer	(R. Wang et al. 2015)
CUL1	Cell cycle pathway	Colorectal cancer	(W. Wang et al. 2015)
BMPR2	Tgf-Beta signaling	Colorectal cancer Gastric cancer	(Kodach et al. 2008) (Park et al. 2010)
MYH10	-	Colorectal cancer	(Park et al. 2010)
*ERCC2**	Nucleotide excision repair	Breast Cancer Ovary Cancer	(Rump et al. 2016)
*CDH1**	Cell cycle pathway	Gastric cancer	(Rump et al. 2016)
*CHEK2**	p53 signaling	Breast cancer	(Apostolou & Papasotiriou 2017)
*NOTCH2**	Notch Signaling	Oesophageal cancer Leukemia	(C. Wang et al.016) (Lobry et al. 2011)
SIN3A	Thyroid hormone signaling	Breast cancer	(Lewis et al. 2016)
HDAC3	Viral carcinogenesis	Colorectal cancer Liver Cancer	(Bhaskara et al. 2011) (Barneda-Zahonero & Parra 2012)
*CLTC**	Lysosome	Breast cancer Renal cancer	(Yao et al. 2015) (Argani et al. 2003)
TCF4	Wnt signaling	Colorectal cancer	(Kim et al. 2003)
RAD54L	Homologous recombination	Colorectal cancer Ovary Cancer prostate cancer	(Pelttari et al. 2012)
*PHOX2B**	Neural Crest Differentiation	Neuroblastoma	(Bachetti et al. 2010)
HLA-A	-	Colorectal cancer	(Menon et al. 2002)
NTRK1	PI3K-Akt signaling pathway	Colorectal cancer	(Créancier et al. 2015) (Iacopetta et al. 2006)
ERBB3	MAPK Signaling	Breast cancer Ovarian cancer prostate cancer	(Ma et al. 2014) (Sithanandam & Anderson 2008)
ATF2	PI3K-Akt signaling	Renal cell carcinoma skin cancer	(Wu et al. 2016) (Bhoumik et al. 2008)

continue

Continued TABLE 2

Driver Genes	Pathways	Cancer types	References
PTCH1	Hedgehog signaling	gastric cancer	(X. De Wang et al. 2013)
NCOR1	Endocrine resistance signaling	Thyroid cancer	(Fozzatti et al. 2013)
MYH9	regulation of cytoskeleton	Lung cancer	(Katono et al. 2015)
TSC2	p53 signaling pathway	Breast cancer	(Mehta et al. 2011)
ERBB4	MAPK signaling	Colon cancer	(Williams et al. 2015)
PTPN12	-	Liver cancer	(Luo & Lam 2013)
KLF5	-	Breast Cancer	(Jia et al. 2016)
TSC1	mTOR signaling pathway	Bladder cancer	(Metha et al. 2011)
SPTAN1	Tgf-Beta Signaling	Oral Cancer	(Prasad et al. 2016)
HGF	MAPK signaling pathway	Esophageal cancer	(Ren et al. 2005)
HSP90AB1	PI3K-Akt signaling	Gastrointestinal cancer	(Moser et al. 2009)
WNT5A	Transcription AR nuclear signaling	Breast Cancer	(Prasad et al. 2016)
MAP3K1	MAPK signaling	Prostate cancer Gastric cancer	(Pham et al. 2013)
RAX	-	-	-
CDC27	Cell cycle pathway	Colorectal cancer Breast cancer	(Qiu et al. 2016) (Qiu et al. 2016)
NCOR2	Notch signaling	Breast cancer	(Van Agthoven et al. 2009)
STAG2	Cell cycle pathway	Bladder cancer	(Black 2014)

A total of 21 pathways have been identified related to the predicted CRC driver genes (Table 3). The information on the pathways where these genes were involved were obtained from KEGG database and literature searching. The result has suggested the involvement of driver genes in dysregulating the multiple pathways causing to the colorectal cancer carcinogenesis. The highest number of pathways might be dysregulated by driver genes were found in patient C474, followed by patient C414 and patient C373.

P53 signaling pathway shows positive signal of perturbation that affects 7 out of 8 patients in this study. Genes *TP53*, *CHEK2* and *TSC2* are predicted to drive colorectal carcinogenesis and related with p53 signaling pathway. Gene *TP53* activated by *CHEK2*, *CHEK1* and *ATM* to regulate various cell activities, meanwhile, *TSC2* regulated by *TP53* to inhibit mTOR signaling pathway (Figure 3). Activation of mTOR signaling pathway will lead to increase in cell proliferation (Harris & Levine 2005; Li et al. 2015; Shao et al. 2001).

Driver gene *NCOR1* was found in the endocrine resistance pathway were predicted to be dysregulated in six patients (C187, C194, C373, C404, C414, and C474). In normal condition, *NCOR1* forms a complex with *ESR1* to inhibit the genes that involved in cell proliferation (Figure 4).

Driver genes such as *KRAS*, *NTRK1*, *ERBB3*, *ATF2*, *ERBB4*, and HSP90AB1 were found in patient C273, C404, C414, C449, and C474 and they were involved in Phosphoinositide 3-kinase-AKT (PI3K-AKT) pathway (Figure 5). PI3K-AKT signaling pathway is involved in mechanisms of growth stimulation, increase cellular proliferation and reduce apoptosis. Crucial component in PI3K-AKT is *AKT3* is linked with driver gene, *TSC2* regulator of mammalian target of rapamycin (mTOR) signaling pathway. mTOR signaling pathway has role in cell growth and proliferation.

MAPK signaling pathway was predicted as one of the highly dysregulated pathways in patient C273, C404, C414, C449, and C474. Three genes were involved in this pathway, i.e. *BRAF*, *ERBB3*, and *MAP3K1*. *ERBB3* encodes a member of RTK (Receptor Tyrosine Kinase) and RTK activation initiates MAPK signaling pathway which then leads to the activation of another key signaling molecules i.e. Ras, which was encoded by *KRAS* gene. *BRAF* encodes for B-Raf protein; one of regulator sin MAPK-ERK signaling pathway that involved in cell proliferation and differentiation whilst *MAP3K1* encodes for MEKK1 protein which is a key regulator in the MAPK-JNK signaling pathway. MEKK1 regulates proliferation, differentiation and inflammation (Fang & Richardson 2005; Ma et al. 2014).

CONCLUSION

From this study, 44 unique driver genes have been predicted in eight Malaysian colorectal cancer patients through extensive customized bioinformatics analysis along with the identification of 21 unique pathways that were found to be related with those driver genes. Amongst the identified pathways, that potentially perturbed driven by mutation of regulators are p53 signaling, PI3K-AKT, MAPK and endocrine resistance were predicted to be potentially perturbed driven by the mutation of regulators. Prediction of cancer driver genes will give detailed insight on the mechanisms of cancer development and a crucial step for targeted therapy process.

TABLE 3. Type of pathways that were associated with driver genes in each patient

	C187	C194	C273	C373	C404	C414	C449	C474
p53 signaling pathway	■	■	■	■	■	■	■	
PI3K-AKT signaling pathway			■		■	■	■	■
Wnt Signaling pathway				■	■			
tgf-β signaling pathway								■
G-protein signaling	■							
MAPK signaling pathway			■		■	■	■	■
RET signaling pathway			■					
Chromatin Organization						■		
Nucleotide excision repair			■					
Cell cycle pathway		■	■	■	■			■
Notch Signaling		■						■
Thyroid hormone signaling				■				
Viral carcinogenesis				■				
Homologous recombination				■				
Neural Crest Differentiation						■		
Hedgehog Signaling Pathway						■		
Endocrine resistance	■	■		■	■	■		■
Tight junction		■						
Regulation of cytoskeleton		■						
mTOR signaling pathway								■
Transcription AR nuclear signaling								■

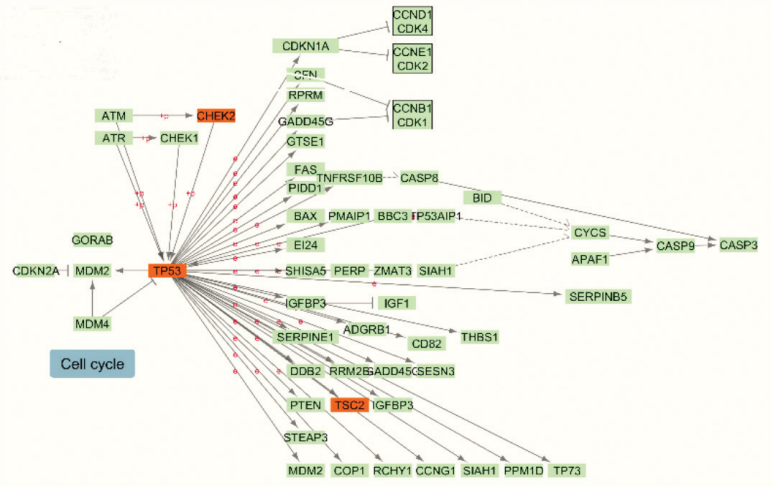

FIGURE 3. The involvement of the predicted drivers *TP53*, *CHEK2* and *TSC2* in p53 signaling pathway

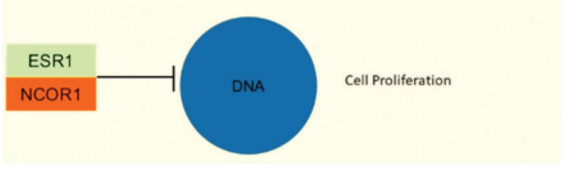

FIGURE 4. NCOR1 in endocrine resistance pathways

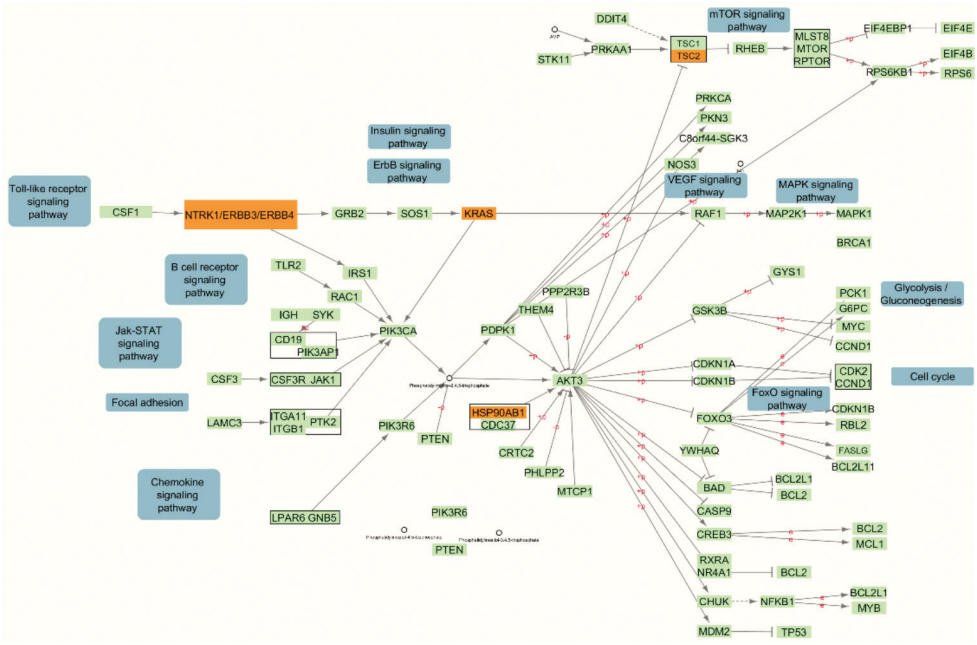

FIGURE 5. Location of *NTRK1*, *ERBB3*, *ERBB4*, *KRAS*, *TSC2* and *HSP90AB1* in PI3K-AKT signaling pathways

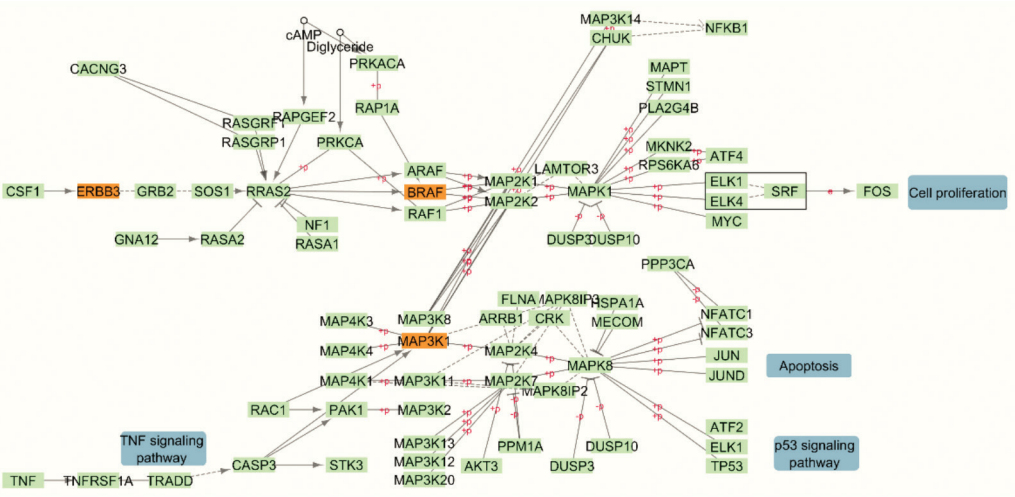

FIGURE 6. *ERBB3*, *BRAF* and *MAP3K1* in MAPK signaling pathways

ACKNOWLEDGEMENTS

This work was supported by LRGS/2014/UKM-UKM/K/04 awarded to Assoc. Prof. Dr. Zeti Azura Mohamed Hussein by the Ministry of Higher Education, Malaysia.

REFERENCES

Adzhubei, I., Jordan, D.M. & Sunyaev, S.R. 2015. Predicting functional effect of human missense mutations using PolyPhen-2. *Current Protocols in Human Genetics* Unit 7.20.

Alberts, B. 2008. Tumor suppressor genes and oncogenes: Genes that prevent and cause cancer. In *Molecular Biology of the Cell*, 5th ed., edited by Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. & Walter, P. Garland Science. pp. 1230-1256.

Apostolou, P. & Papasotiriou, I. 2017. Current perspectives on CHEK2 mutations in breast cancer. *Breast Cancer: Targets and Therapy* 9: 331-335.

Argani, P., Lui, M.Y., Couturier, J., Bouvier, R., Fournet, J.C. & Ladanyi, M. 2003. A novel CLTC-TFE3 gene fusion in pediatric renal adenocarcinoma with t(X;17)(p11.2;q23). *Oncogene* 22(34): 5374-5378.

Bachetti, T., Di Paolo, D., Di Lascio, S., Mirisola, V., Brignole, C., Bellotti, M. & Caffa, I. 2010. PHOX2B-mediated regulation of ALK expression: *In vitro* identification of a functional relationship between two genes involved in neuroblastoma. *PLoS ONE* 5(10): e13108.

Barneda-Zahonero, B. & Parra, M. 2012. Histone deacetylases and cancer. *Molecular Oncology* 6(6): 579-589.

Bhaskara, S., Knutson, S.K., Jiang, G., Chandrasekharan, M.B., Wilson, A.J., Zheng, S. & Yenamandra, A. 2011. Role for

histone deacetylase 3 in maintenance of genome stability *Genome Stability* 18(5): 436-447.

Bhoumik, A., Fichtman, B., DeRossi, C., Breitwieser, W., Kluger, H.M., Davis, S. & Subtil, A. 2008. Suppressor role of activating transcription factor 2 (ATF2) in skin cancer. *Proceedings of the National Academy of Sciences* 105(5): 1674-1679.

Black, P. 2014. Frequent truncating mutations of STAG2 in bladder cancer. *Urology* 83(4): 691-692.

Chen, Y., McGee, J., Chen, X., Doman, T.N., Gong, X., Zhang, Y. & Hamm, N. 2014. Identification of druggable cancer driver genes amplified across TCGA datasets. *PLoS ONE* 9(5): e98293.

Cho, A., Shim, J.E., Kim, E., Supek, F., Lehner, B. & Lee, I. 2016. MUFFINN: Cancer gene discovery via network analysis of somatic mutation data. *Genome Biology* 17(1): 129.

Chye, G.L.C., Rampal, S. & Yahaya, H. 2008. Cancer incidence in Peninsular Malaysia 2003-2005. *National Cancer Registry*. pp. 53-57.

Créancier, L., Vandenberghe, I., Gomes, B., Dejean, C., Blanchet, J.C., Meilleroux, J. & Guimbaud, R. 2015. Chromosomal rearrangements involving the NTRK1 gene in colorectal carcinoma. *Cancer Letters* 365(1): 107-111.

Davies, H., Bignell, G.R., Cox, C., Stephens, P., Edkins, S., Clegg, S. & Teague, J. 2002. Mutations of the BRAF gene in human cancer. *Nature* 417(6892): 949-954.

Dees, N.D., Zhang, Q., Kandoth, C., Wendl, M.C., Schierding, W., Koboldt, D.C. & Mooney, T. B. 2012. MuSiC: Identifying mutational significance in cancer genomes. *Genome Research* 22(8): 1589-1598.

Fang, J.Y. & Richardson, B.C. 2005. The MAPK signalling pathways and colorectal cancer. *Lancet Oncology* 6(5): 322-327.

Folsom, A.R., Pankow, J.S., Peacock, J.M., Bielinski, S.J., Heiss, G. & Boerwinkle, E. 2008. Variation in TCF7L2 and increased risk of colon cancer: The atherosclerosis risk in communities (ARIC) study. *Diabetes Care* 31(5): 905-909.

Foo, J., Liu, L.L., Leder, K., Riester, M., Iwasa, Y., Lengauer, C. & Michor, F. 2015. An evolutionary approach for identifying driver mutations in colorectal cancer. *PLoS Computational Biology* 11(9): 1-19.

Fozzatti, L., Park, J.W., Zhao, L., Willingham, M.C. & Cheng, S.Y. 2013. Oncogenic actions of the nuclear receptor corepressor (NCOR1) in a mouse model of thyroid cancer. *PLoS ONE* 8(6): 1-10.

Harris, S.L. & Levine, A.J. 2005. The p53 pathway: Positive and negative feedback loops. *Oncogene* 24(17): 2899-2908.

Huang, T., Kang, W., Cheng, A.S.L., Yu, J. & To, K.F. 2015. The emerging role of slit-robo pathway in gastric and other gastro intestinal cancers. *BMC Cancer* 15(1): 1-9.

Iacopetta, B., Russo, A., Bazan, V., Dardanoni, G., Gebbia, N., Soussi, T. & Kerr, D. 2006. Functional categories of TP53 mutation in colorectal cancer: Results of an international collaborative study. *Annals of Oncology* 17(5): 842-847.

Iwatsuki, M., Mimori, K., Lshii, H., Yokobori, T., Takatsuno, Y., Sato, T. & Toh, H. 2010. Loss of FBXW7, a cell cycle regulating gene, in colorectal cancer: Clinical significance. *International Journal of Cancer* 126(8): 1828-1837.

Jia, L., Zhou, Z., Liang, H., Wu, J., Shi, P., Li, F. & Wang, Z. 2016. KLF5 promotes breast cancer proliferation, migration and invasion in part by upregulating the transcription of TNFAIP2. *Oncogene* 35(16): 2040-2051.

Kanner, J. 2007. Dietary advanced lipid oxidation endproducts are risk factors to human health. *Molecular Nutrition and Food Research* 51(9): 1094-1101.

Katono, K., Sato, Y., Jiang, S.X., Kobayashi, M., Nagashio, R., Ryuge, S. & Fukuda, E. 2015. Prognostic significance of MYH9 expression in resected non-small cell lung cancer. *Plos One* 10(3): e0121460.

Kim, P.J., Plescia, J., Clevers, H., Fearon, E.R. & Altieri, D.C. 2003. Surviving and molecular pathogenesis of colorectal cancer. *Lancet* 362(9379): 205-209.

Kodach, L.L., Wiercinska, E., de Miranda, N.F.C.C., Bleuming, S.A., Musler, A.R., Peppelenbosch, M.P. & Dekker, E. 2008. The bone morphogenetic protein pathway is inactivated in the majority of sporadic colorectal cancers. *Gastroenterology* 134(5): 1332-1341.

Kwong, L.N. & Dove, W.F. 2009. APC and its modifiers in colon cancer. *Advances in Experimental Medicine and Biology* 656: 85-106.

Leiserson, M.D.M., Vandin, F., Wu, H.T., Dobson, J.R., Eldridge, J.V., Thomas, J.L. & Papoutsaki, A. 2014. Pan-cancer network analysis identifies combinations of rare somatic mutations across pathways and protein complexes. *Nature Genetics* 47(2): 106-114.

Lewis, M.J., Liu, J., Falk Libby, E., Lee, M., Crawford, N.P.S. & Hurst, D.R. 2016. SIN3A and SIN3B differentially regulate breast cancer metastasis. *Oncotarget* 7(48): 78713-78725.

Li, X.L., Zhou, J., Chen, Z.R. & Chng, W.J. 2015. *p53* mutations in colorectal cancer: Molecular pathogenesis and pharmacological reactivation. *World Journal of Gastroenterology* 21(1): 84-93.

Lobry, C., Oh, P. & Aifantis, I. 2011. Oncogenic and tumor suppressor functions of Notch in cancer: it's NOTCH what you think. *The Journal of Experimental Medicine* 208(10): 1931-1935.

Louis, P., Hold, G.L. & Flint, H.J. 2014. The gut microbiota, bacterial metabolites and colorectal cancer. *Nature Reviews Microbiology* 12(10): 661-672.

Luo, S.Y. & Lam, D.C. 2013. Oncogenic driver mutations in lung cancer. *Translational Respiratory Medicine* 1(1): 6.

Ma, J., Lyu, H., Huang, J. & Liu, B. 2014. Targeting of erbB3 receptor to overcome resistance in cancer treatment. *Molecular Cancer* 13(1): 1-9.

Marx, V. 2014. Cancer genomes: Discerning drivers from passengers. *Nature Methods* 11(4): 375-379.

Mathur, R., Alver, B.H., San Roman, A.K., Wilson, B.G., Wang, X., Agoston, A.T. & Park, P.J. 2017. ARID1A loss impairs enhancer-mediated gene regulation and drives colon cancer in mice. *Nature Genetics* 49(2): 296-302.

Mehta, M.S., Vazquez, A., Kulkarni, D.A., Kerrigan, J.E., Atwal, G., Metsugi, S. & Toppmeyer, D.L. 2011. Polymorphic variants in TSC1 and TSC2 and their association with breast cancer phenotypes. *Breast Cancer Research and Treatment* 125(3): 861-868.

Menon, A.G., Morreau, H., Tollenaar, R.A.E.M., Alphenaar, E., Van Puijenbroek, M., Putter, H. & Janssen-van Rhijn, C.M. 2002. Downregulation of HLA-A expression correlates with a better prognosis in colorectal cancer patients. *Laboratory Investigation* 82(12): 1725-1733.

Moser, C., Lang, S.A. & Stoeltzing, O. 2009. Heat-shock protein 90 (Hsp90) as a molecular target for therapy of gastrointestinal cancer. *Anticancer Research* 29(6): 2031-2042.

O'Keefe, S.J.D. 2016. Diet, microorganisms and their metabolites and colon cancer. *Nature Reviews Gastroenterology & Hepatology* 13(12): 691-706.

Park, S.W., Hur, S.Y., Yoo, N.J. & Lee, S.H. 2010. Somatic frameshift mutations of bone morphogenic protein receptor 2 gene in gastric and colorectal cancers with microsatellite instability. *Apmis* 118(11): 824-829.

Pelttari, L.M., Kiiski, J., Nurminen, R., Kallioniemi, A., Schleutker, J., Gylfe, A. & Aaltonen, L.A. 2012. A finnish founder mutation in RAD51D: Analysis in breast, ovarian, prostate, and colorectal cancer. *Journal of Medical Genetics* 49(7): 429-432.

Petitjean, A., Achatz, M.I.W., Borresen-Dale, A.L., Hainaut, P. & Olivier, M. 2007. TP53 mutations in human cancers: Functional selection and impact on cancer prognosis and outcomes. *Oncogene* 26(15): 2157-2165.

Pham, T.T., Angus, S.P. & Johnson, G.L. 2013. MAP3K1: Genomic alterations in cancer and function in promoting cell survival or apoptosis. *Genes and Cancer* 4(11-12): 419-426.

Phipps, A.I., Buchanan, D.D., Makar, K.W., Win, A.K., Baron, J.A., Lindor, N.M. & Potter, J.D. 2013. KRAS-mutation status in relation to colorectal cancer survival: The joint impact of correlated tumour markers. *British Journal of Cancer* 108(8): 1757-1764.

Prasad, C.P., Chaurasiya, S.K., Guilmain, W. & Andersson, T. 2016. WNT5A signaling impairs breast cancer cell migration and invasion via mechanisms independent of the epithelial-mesenchymal transition. *Journal of Experimental and Clinical Cancer Research* 35(1): 1-15.

Qiu, L., Wu, J., Pan, C., Tan, X., Lin, J., Liu, R. & Chen, S. 2016. Downregulation of CDC27 inhibits the proliferation of colorectal cancer cells via the accumulation of p21Cip1/Waf1. *Cell Death & Disease* 7(651): e2074.

Radzi, M., Hassan, A., Ismail, I., Azri, M., Suan, M., Ahmad, F. & Khamizar, W. 2016. Incidence and mortality rates of colorectal cancer in Malaysia. *Epidemiol. Health* 38: 6-10.

Ragnarsson-Olding, B.K., Karsberg, S., Platz, A. & Ringborg, U.K. 2002. Mutations in the TP53 gene in human malignant melanomas derived from sun-exposed skin and unexposed mucosal membranes. *Melanoma Research* 12(5): 453-463.

Ren, Y., Cao, B., Law, S., Xie, Y., Lee, P.Y., Cheung, L. & Chen, Y. 2005. Hepatocyte growth factor promotes cancer cell migration and angiogenic factors expression: A prognostic marker of human esophageal squamous cell carcinomas. *Clinical Cancer Research* 11(17): 6190-6197.

Rivlin, N., Brosh, R., Oren, M. & Rotter, V. 2011. Mutations in the p53 tumor suppressor gene: Important milestones at the various steps of tumorigenesis. *Genes and Cancer* 2(4): 466-474.

Rump, A., Benet-Pages, A., Schubert, S., Kuhlmann, J.D., Janavičius, R., Macháčková, E. & Foretová, L. 2016. Identification and functional testing of ERCC2 mutations in a multi-national cohort of patients with familial breast- and ovarian cancer. *PLoS Genetics* 12(8): 1-18.

Ryan-Harshman, M. & Aldoori, W. 2007. Diet and colorectal cancer: Review of the evidence. *Canadian Family Physician Médecin de Famille Canadien* 53(11): 1913-1920.

Shao, L., Lai, M. & Huang, Q. 2001. Mutagen sensitivity and p53 expression in colorectal cancer in China. *Postgraduate Medical Journal* 77: 713-716.

Sim, N.L., Kumar, P., Hu, J., Henikoff, S., Schneider, G. & Ng, P.C. 2012. SIFT web server: Predicting effects of amino acid substitutions on proteins. *Nucleic Acids Research* 40(W1): W452-W457.

Sithanandam, G. & Anderson, L.M. 2008. The ERBB3 receptor in cancer and cancer gene therapy. *Cancer Gene Therapy* 15(7): 413-448.

Slattery, M.L., Folsom, A.R., Wolff, R., Herrick, J., Caan, B.J. & Potter, J.D. 2009. Tumor markers and rectal cancer: Support for an inflammation-related pathway. *International Journal of Cancer* 17(4): 978-982.

Tan, C. & Du, X. 2012. KRAS mutation testing in metastatic colorectal cancer. *World Journal of Gastroenterology* 18(37): 5171-5180.

Tian, S., Simon, I., Moreno, V., Roepman, P., Tabernero, J., Snel, M. & Veer, L. 2013. A combined oncogenic pathway signature of BRAF, KRAS and PI3KCA mutation improves colorectal cancer classification and cetuximab treatment prediction. *Gut* 62(4): 540-549.

Tokheim, C., Papadopoulis, N., Kinzler, K.W., Vogelstein, B. & Karchin, R. 2016. Evaluating the evaluation of cancer driver genes. *bioRxiv* 113(50): 060426.

Van Agthoven, T., Sieuwerts, A.M., Veldscholte, J., Meijer-Van Gelder, M.E., Smid, M., Brinkman, A. & Den Dekker, A.T. 2009. CITED2 and NCOR2 in anti-oestrogen resistance and progression of breast cancer. *British Journal of Cancer* 101(11): 1824-1832.

Waddell, N., Pajic, M., Patch, A., Chang, D.K., Kassahn, K.S., Bailey, P. & Johns, A.L. 2015. Whole genomes redefine the mutational landscape of pancreatic cancer. *Nature* 518(7540): 495-501.

Wang, C., Li, Q., Liu, F., Chen, X., Liu, B., Nesa, E.U., Guan, S., Han, L., Tan, B., Wang, N., Wang, X., Song, Q., Jia, Y., Wang, J., Lu, M. & Chen, Y. 2016. Notch2 as a promising prognostic biomarker for oesophageal squamous cell carcinoma. *Scientific Reports* 6 (25722): 1-10.

Wang, X.D., Inzunza, H., Chang, H., Qi, Z., Hu, B., Malone, D. & Cogswell, J. 2013. Mutations in the hedgehog pathway genes SMO and PTCH1 in human gastric tumors. *PLoS ONE* 8(1): 1-8.

Wang, R., Zhang, Y., Pan, Y., Li, Y., Hu, H. & Cai, D. 2015. Comprehensive investigation of oncogenic driver mutations in Chinese non-small cell lung cancer patients. *Oncotarget* 6(33): 1-9.

Wang, W., Chen, Y., Deng, J., Zhou, J., Gu, X., Tang, Y. & Zhang, G. 2015. Cullin1 is a novel prognostic marker and regulates the cell proliferation and metastasis in colorectal cancer. *Journal of Cancer Research and Clinical Oncology* 141(9): 1603-1612.

Williams, C.S., Bernard, J.K., Beckler, M.D., Almohazey, D., Washington, M.K., Smith, J.J. & Frey, M.R. 2015. ERBB4 is over-expressed in human colon cancer and enhances cellular transformation. *Carcinogenesis* 36(7): 710-718.

Wu, D.S., Chen, C., Wu, Z.J., Liu, B., Gao, L., Yang, Q. & Chen, W. 2016. ATF2 predicts poor prognosis and promotes malignant phenotypes in renal cell carcinoma. *Journal of Experimental and Clinical Cancer Research* 35(1): 1-11.

Yao, F., Zhang, C., Du, W., Liu, C. & Xu, Y. 2015. Identification of gene-expression signatures and protein markers for breast cancer grading and staging. *PLoS ONE* 10(9): 1-17.

Zhan, T., Rindtorff, N. & Boutros, M. 2017. Wnt signaling in cancer. *Oncogene* 36(11): 1461-1473.

Centre for Bioinformatics Research
Institute of Systems Biology (INBIOSIS)
Universiti Kebangsaan Malaysia
43600 UKM Bangi, Selangor Darul Ehsan
Malaysia

*Corresponding author; email: norazlannm@ukm.edu.my

Received: 30 May 2018
Accepted: 14 September 2018