IDENTIFICATION OF INFORMATIVE SUBPATHWAYS AND GENES USING IMPROVED DIFFERENTIAL EXPRESSION ANALYSIS FOR PATHWAYS METHOD

NURUL ATHIRAH BINTI NASARUDIN

UNIVERSITI TEKNOLOGI MALAYSIA

IDENTIFICATION OF INFORMATIVE SUBPATHWAYS AND GENES USING IMPROVED DIFFERENTIAL EXPRESSION ANALYSIS FOR PATHWAYS METHOD

NURUL ATHIRAH BINTI NASARUDIN

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Master of Philosophy

School of Computing
Faculty of Engineering
Universiti Teknologi Malaysia

NOVEMBER 2020

ACKNOWLEDGEMENT

In preparing this thesis, I was in contact with many people, researchers, academicians, and practitioners. They have contributed towards my understanding and thoughts. In particular, I wish to express my sincere appreciation to my main thesis supervisor, Dr. Zalmiyah Zakaria, for encouragement, guidance, critics and friendship. I am also very thankful to my co-supervisors Dr. Chan Weng Howe and Professor Dr. Mohd Saberi Mohamad for their guidance, advices and motivation. Without their continued support and interest, this thesis would not have been a success as presented here.

I am also indebted to Universiti Teknologi Malaysia (UTM) for funding my masters study. The UTM librarians also deserve special appreciation for their assistance in supplying relevant literatures.

My fellow postgraduate friends should also be recognised for their support during this entire journey. My sincere appreciation also extends to all my colleagues and others who have provided assistance at various occasions. Their views and tips are useful indeed. Also, never to forget all my family members who have always been my pillar of strength, courage and perseverance.

ABSTRACT

Pathway-based analysis is introduced to define useful biological knowledge by considering the whole pathway features. However, most of these analyses have several shortcomings, such as less sensitivity towards data that could lead to some important information being missed. Because of the deficiency, pathway-based analysis has been shifted to subpathway-based analysis, which is seen to be more relevant in understanding the biological reactions. This is strengthened by the fact that several studies have found abnormalities in pathways caused by certain regions that respond in the etiology of diseases. In addition, subpathway-based analysis has been found to be more effective and sensitive than the whole pathway. Due to this orientation, many tools have been developed to accomplish the inadequate interpretation in biology system. The Differential Expression Analysis for Pathway (DEAP) is one of the methods in subpathway-based analysis which identifies a local region perturbed by complex diseases in large pathway data. However, the method has shown low performance in identifying informative pathway and subpathway. Hence, this research proposes a modified DEAP method (termed iDEAP) for enhancing the identification of perturbed subpathways in pathway activities and aimed at achieving higher performance in the detection of differential expressed pathways. To this end, firstly, asearch algorithm adapted from DMSP algorithm was implemented to DEAP in search for informative subpathways. Secondly, the relation among subpathways was taken into account by averaging the maximum absolute value (termed DEAP score) to emphasize the reaction among subpathways so that efficient identification of informative pathways can be achieved. Three gene expression data sets were applied in this study (head and neck tumour, colorectal cancer and breast cancer). The results were obtained in terms of the number of differential expressed pathways (head and neck tumor-81 pathways, colorectal cancer-78 pathways, breast cancer-95 pathways) and they suggest that the proposed method yielded better performance as compared to previous work. In fact, when the selected genes from the results were evaluated using 10-fold CV in terms of accuracy, the proposed method showed higher accuracy for Colorectal (90%) and Breast cancer (94%). Finally, a biological validation was conducted on the top five (5) significant pathways and selected genes based on biological literatures.

ABSTRAK

Analisis berasaskan laluan diperkenalkan untuk mentakrif pengetahuan biologi yang bermanfaat dengan mempertimbangkan keseluruhan ciri laluan. Walau bagaimanapun, terdapat beberapa kekangan pada kebanyakan analisis ini seperti kurang kepekaan terhadap data yang boleh membawa kepada keciciran beberapa maklumat penting. Menerusi kelemahan yang dikenal pasti, penggunaan analisis berasaskan laluan telah beralih kepada analisis berasaskan sub-laluan, yang lebih relevan dalam memahami reaksi biologi, kerana beberapa kajian telah menemukan keabnormalan dalam laluan yang disebabkan oleh bahagian tertentu yang bertindak balas dalam etiologi penyakit. Di samping itu, analisis sub-laluan didapati lebih berkesan dan sensitif daripada keseluruhan laluan. Oleh kerana orientasi ini, pelbagai peralatan dibangunkan untukmemenuhi tafsiran yang tidak lengkap dalam sistem biologi. Analisis Ekspresi Berbeza untuk Laluan (DEAP) adalah salah satu kaedah dalam analisis berasaskan sub-laluan yang mengenal pasti kawasan setempat yang dipengaruhi oleh penyakit kompleks dalam data laluan besar. Walau bagaimanapun, kaedah ini menunjukkan prestasi rendah dalam mengenal pasti laluan bermaklumat dan sub-laluan. Oleh itu, penyelidikan ini mengusulkan kaedah DEAP yang telah diubah suai (dinamakan iDEAP) untuk meningkatkan pengesanan sub-laluan yang terganggu dalam aktiviti laluan dan bertujuan untuk meningkatkan keberkesanan dalam mengesan laluan yang dinyatakan di atas. Pertama, algoritma carian yang disesuaikan daripada algoritma DMSP dilaksanakan kepada DEAP untuk mencari sub-laluan bermaklumat. Kedua, penyelidik telah mengambil kira hubungan antara sub-laluan dengan purata nilai mutlak maksimum (disebut sebagai skor DEAP) untuk mengambil kira tindak balas antara sublaluan supaya pengenalan laluan bermaklumat dapat dicapai secara efektif. Terdapat tiga set data ungkapan gen yang digunakan dalam kajian ini (tumor kepala dan leher, kanser kolorektal dan kanser payudara). Keputusan diperolehi dari segi bilangan laluan yang dijumpai dan menunjukkan bahawa kaedah yang dicadangkan menghasilkan prestasi yang lebih baik berbanding penyelidikan terdahulu. Selain itu, gen yang dipilih daripada keputusan dinilai menggunakan CV 10 kali ganda dari segi ketepatan. Akhir sekali, pengesahan biologi dijalankan kepada lima (5) jalur penting dan gen terpilih berdasarkan literatur biologi.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

CV	-	Cross Validation
SVM	-	Support Vector Machine
HNSCC	-	Head and Neck Squamous Cell Carcinoma
IGA	-	Individual Gene Analysis
GSA	-	Gene Set Analysis
GSEA	-	Gene Set Enrichment Analysis
iDEAP	-	Improved Differential Expression Analysis for Pathway
DEAP	-	Differential Expression Analysis for Pathway
KEGG	-	Kyoto Encyclopaedia of Genes and Genomes
GEO	-	Gene Expression Omnibus
TN	-	True Negative
TP	-	True Positive
PANTHER	-	Protein Analysis Through Evolutionary Relationships
eBIOLOGICAL VALIDATION	-	Electronic Biological Validation

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Overview

In the era of life science, emerging high–throughput technologies such as next-generation sequencing, omics technology, and microarray have brought a massive dimension of biological data. The biological data that have been discovered including data in genome, transcriptome, epigenome, proteome, metabolome, molecular imaging, and molecular pathways. As a result, biological data is exponentially increasing the database size due to the information at various levels of biological systems (Penisi, 2011). Microarray technology has been introduced, which is known for its capability of analyzing thousands of data with multiple samples simultaneously. Therefore, many sophisticated analytic methods have been developed to analyze microarray data for intepreting important biological function. Differential expression analysis is an analysis in finding genes that are differentially expressed across biological conditions. It has been commonly used for finding biomarkers, drug target and candidates for understanding the molecular mechanism of complex disease (Walker, 2001). Traditionally, genes expression is analyzed gene-by-gene without considering the interaction and association mechanism. By ignoring the biological interaction and structure, the analysis would become less effective and lead to misleading interpretation.

The earliest approach introduced in gene-by-gene analysis is individual gene analysis or IGA that naturally produces a list of altered genes from a cutoff threshold (Nam and Kim, 2008). Subsequently, systems-level methodologies have pushed forward the transition of IGA to gene set analysis (GSA), since cutoff-based method has deficiency in consideration of many informative genes, which causes low statistical identification efficiency of true positive. GSA methods have also received attention among researchers since it is free from issues of “cutoff-based” methods.

Moreover, this method is able to identify gene sets in a subtle way, coordinated by a single-step process. The biological meaning of gene expression data can be directly inferred by applying a sample or a gene randomization test. Gene Set Enrichment Analysis (GSEA) is one of the popular methods in GSA, which interprets the ranked genes based on the correlation between their expressions from two sample classes (Subramanian et al., 2005). The significance of the informative gene set is analyzed based on maximum running sum, where each gene set is calculated simultaneously throughout the ranked gene list. Then, the significant gene sets are analyzed based on two types of comparison methods which are competitive approaches that compare gene sets relative to another and self-contained approaches that compare individual gene sets across conditions without consideration for other sets. Even though GSA methods give advantage to researchers in characterizing a group of genes, they still have a limitation when being applied to pathway dataset. Most of GSA methods neglect the graph structure of the pathway data, therefore, they might miss important information such as the biological interaction between melocules that leads to inaccurate result.

At this point, pathway topology-based analysis has been introduced to overcome the limitation of GSA methods by considering the pathway structure. This analysis has integrated the benefits of GSA and extended them with information from gene-gene interaction in the pathway database. In addition, there are two hypothesis tests that can be observed in this analysis: first, entire pathways are tested for differential expression; second, an informative path identified represents the entire pathway with massive information to that differential expression. As a result, researchers can identify the associated pathways with a biological condition related to targeted phenotype accurately. Previous studies stated that the pathway structure information is able to provide relevant biological insights and contributes for comprehension of higher-order of biological system functions (Emmert- Streib et al., 2012). One of the popular methods in the topology-based analysis is signaling pathway impact analysis (SPIA), which associates the information from the classical enrichment analysis with pathway information in identifying perturbated pathway under a given condition (Tarca et al., 2009). But as a whole, the topology-based analysis methods test the generic hypothesis of pathways without identifying specific paths (Daraghici et al., 2007).

Recently, topology-based analysis has shifted towards subpathway-based analysis, which provides information of biological phenomena more precisely, and contributes in identifying regions of the pathway that are dysregulated by disease (Bezerianos et al., 2017). This analysis is based on assumption that biological process related to complex diseases can be described through local region topologies in the pathway. In addition, previous studies have proved that deformities in subpathway regions of the pathway might contribute to etiology of disease (Li et al., 2012). From this evolution, several subpathway-based analysis methods have been developed which share the same target in the search pathway portion related to disease modelling, drug targeting and other objectives (Chen et al., 2011; Martini et al., 2013; Judeh et al., 2013; Nam et al., 2014). The earliest subpathway-based analysis methods are TAPPA (Gao et al., 2007), Subpathway-GM (Li et al., 2013), TEAK (Judeh et al., 2013) and many more. These methods identify subpathways through the incorporation of genes information and metabolites pathway data by taking account their topology structures and interactions. The overview of subpathway-based analysis is illustrated in Figure 1.1.

In the present biological studies, identification of perturbed subpathways and genes in cancer-related pathways is crucial to provide insights for better biological interpretation of the biological processes. Comprehensive interpretation of biological processes is important to drugs discovery and targeted treatment design. For the past few years, the development of subpathway-based analysis methods shows an increasing trend to take advantage on the incorporation of pathway activity data in order to enhance the outcome. However, there are also challenges discovered by previous studies. One of the challenges is how to examine the subpathway (Li et al., 2009). Most of subpathway methods independently search the subpathway without implementing any search algorithm. This reduces the tendency to find the perturbed subpathway related to disease. In addition, the pathway structure is complex since it involves the combination of many subpathways and interaction. Due to this problem, an efficient subpathway-based analysis method is functional to identify specific region that is differentially expressed by utilizing every information within a pathway.

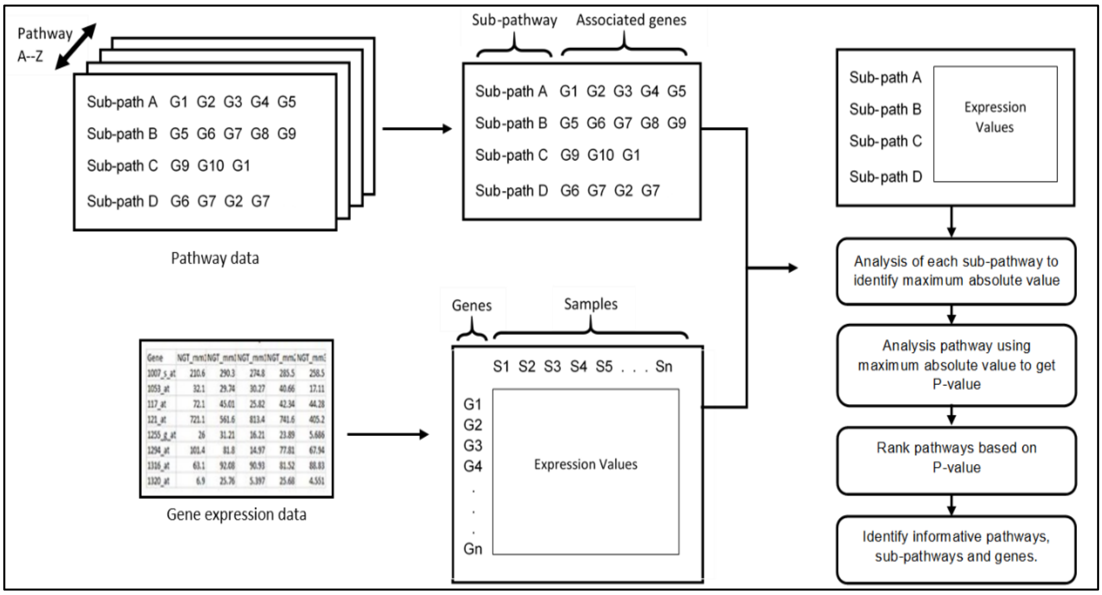
Pathway A–Z
Sub-path A G1 G2 G3 G4 G5
Sub-path B G5 G6 G7 G8 G9
Sub-path C G9 G10 G1
Sub-path D G6 G7 G2 G7
Pathway data
Sub-pathway
Associated genes
Sub-path A G1 G2 G3 G4 G5
Sub-path B G5 G6 G7 G8 G9
Sub-path C G9 G10 G1
Sub-path D G6 G7 G2 G7
Sub-path A
Sub-path B
Sub-path C
Sub-path D
Expression Values
Analysis of each sub-pathway to identify maximum absolute value
Analysis pathway using maximum absolute value to get P-value
Rank pathways based on P-value
Identify informative pathways, sub-pathways and genes.
Gene expression data
Genes
Samples
S1 S2 S3 S4 S5 . . . Sn
G1
G2
G3
G4
Gn
Expression Values

Figure 1.1: Overview illustration of subpathway-based analysis

1.2 Research Background

In the past few years, there is a large gap between data collection in molecular biology and data analysis method to derive accurate functional information. As the data constantly growing, the capability of obtaining an informative list of genes from different phenotypes has become a routine in research nowadays. Even though there are various methods developed to analyse high-throughput data, the ability to interpret biological interaction is as challenging as ever. In fact, living organisms are complex systems with evolving phenotypes that cause thousands of complex interactions taking part in various pathways data. The complexity of pathway data has affected the performance identification of informative pathways where there are uncountable reactions need to consider. Hence, the ability to correctly define perturbed pathways under case study in pathway-based analysis becomes a challenge in order to transform the abundant high-throughput data into biological knowledge (Mitrea et al., 2013). According to (Khatri et al., 2012), the identification of number of significant pathways under case study currently has come into failure due to the weakness and limitation of pathway-based analysis methods. This shows that an effective pathway-based analysis method is essential in order to achieve more promising results.

However, recent methods that have been developed still have limitation and weaknesses. For instance, they can only search informative pathways without knowing the abnormal condition within the pathway and which set of interaction leads to diseases (Khatri et al., 2012). Thus, the generation analysis has emerged to subpathway-based analysis that finds the informative local region known as subpathway by considering all the interactions between subpathways in each pathway. From there, the informative subpathway can be represented as the whole pathway and assist the medical team to discover diseases in short time through complex pathway data. Since few years ago, many researchers have started to develop methods in order to find relevant subpathways related to targeted phenotype. But, most of the methods still have constraints that need to be improved. For example, Subpathway-GM (Li et al., 2013) and Teak (Judeh et al., 2013) have a limitation in defining subpathway in a given pathway. Both methods do not consider the interaction between nodes that can affect the efficiency of identifying significant subpathway under case study. Meanwhile, TEAK method implements two ways of subpathway extraction which are known as linear subpathway and non-linear subpathway. This method has some weaknesses where the nodes inside subpathway could be redundant, hence causing analysis confusion.

With the complexity of the pathway data, the identification of significant subpathway has shown less promising result due to the presence of many interactions within the pathway (Amadoz et al., 2018). As shown in Figure 1.2, Ras signalling pathway comprises of many interactions and biological molecules that have their own roles in biological system. It is impossible to obtain an accurate result with huge number of interactions and molecules within the pathway data. Therefore, current approaches are designed to analyse specific local region in biological system by assuming that each subpathway is independent of each other. The lack of a method that accounts for dependence among subpathways at a time point limits our ability to observe changes at a pathway level in a biological system (Li et al., 2015).

Recent advancements of omics research and the intensive biological researches have made available some well-known online biological pathway databases such as Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa and Goto, 2000), Gene Ontology (GO) (Ashburner et al., 2000), Biocarta (www.biocarta.com) and many more. Generally, many of the biological pathway databases are not specific to certain biological context such as cancer. By implementing subpathway-based analysis, many informative pathways can be identified and directly improve the biological database. In addition, the knowledge of genes within informative subpathway highly related to diseases can be applied for future study such as cancer classification. In the study of complex diseases like cancers, the pathway data might contain irrelevant genes that do not contribute to the development of cancer or involve in cancer-related biological processes. The presence of non-informative genes in the classifier construction might impair the performance of classification (Wang et al., 2008). Therefore, it is crucial to efficiently identify the informative subpathways related to cancer in order to enhance the classification performance.

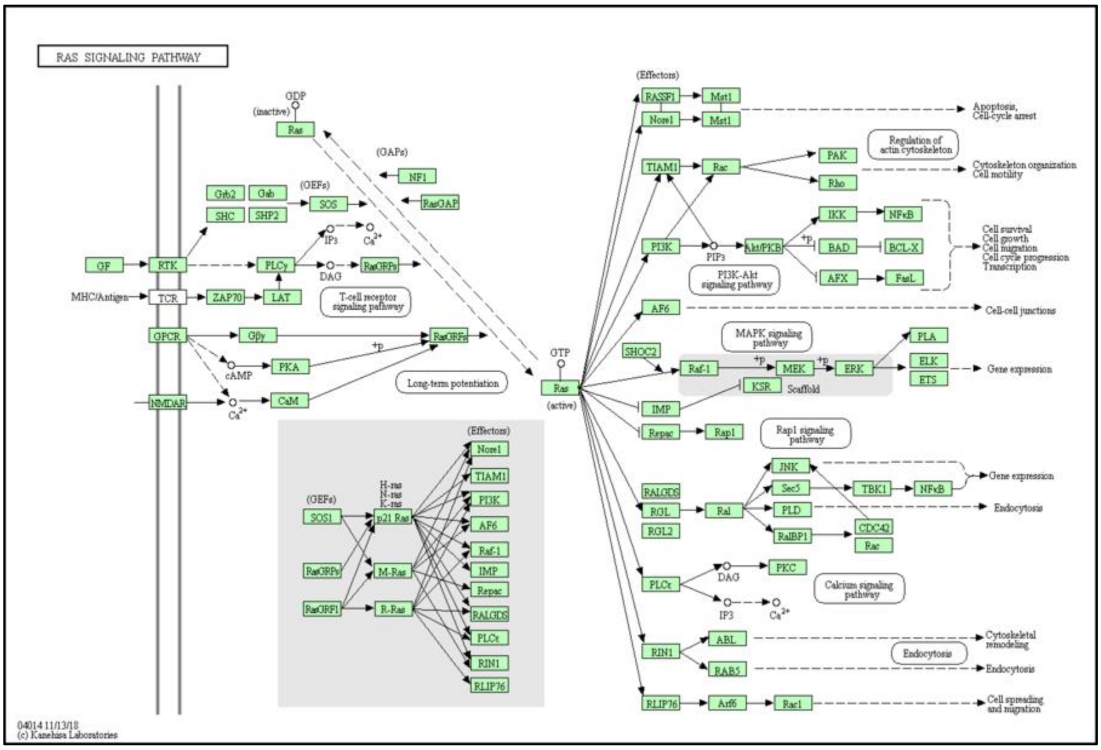

Figure 1.2: Example of complex pathway data that comprises of many interactions and biological molecules. Ras signaling pathway map (https://www.genome.jp/kegg/)

1.3 Problem Statement

The researcher focuses on the problem regarding limitation in identifying the perturbed subpathway that is significantly related to cancer disease. Since the pathway data consists of various biological interactions, the subpathway-based method is required in order to improve the performance of identifying target region or known as subpathway, which interacts with targeted phenotype. Previous studies showed that some information of the genes is enough to identify significant pathway related to targeted phenotype. However, the problem arises when defining the position of the significant genes requiring interactions among each other in order to obtain the perturbed region in each pathway. In addition, the weakness can also be seen when the subpathways are assumed independently by neglecting the interactions between them (Li *et al*., 2015). In previous method (DEAP), a single subpathway with maximum score was selected to represent the corresponding significant pathway. In order to improve the performance of subpathway-based analysis method, an effective and practical approach is needed to address the problems. The method should be able to identify the informative subpathway within the pathway and take into account the interaction between subpathways to improve the performance of such identification.

It can be concluded that the main problem in this research is the weakness of subapathway-based analysis method in identifying the significant subpathway and the inefficiency of analysis when neglecting the interaction between subpathways which affects the performance of pathway identification. Thus, this research intends to address the aforementioned problems based on the following research questions:

How to effectively identify important subpathway related to complex diseases from differential expression data in a given pathway?

How to effectively validate the identified informative subpathways and genes?

1.4 Research Goal

The goal of this research is to propose an improved differential expression analysis for pathway method to efficiently identify the informative subpathways and genes in a pathway.

1.4.1 Research Objectives

Several objectives have been set as the research direction. The objectives are expressed as below:

(a) To propose an improved Differential Expression Analysis for the pathway (iDEAP) with Detect Module from Seed Protein (DMSP) algorithm features for more efficient identification of informative subpathway and genes in better prediction of pathway related to cancer.

To verify and validate the performance and result of improved differential expression analysis for the pathway (iDEAP) with previous research and biological database.

1.5 Research Scope

The scope of the research is bounded under some limitations, as stated below:

(a) Three types of cancer microarray dataset applied in this research are obtained from Gene Expression Omnibus (GEO) database and pathway data with a total of 177 pathways obtained from the Protein Analysis Through Evolutionary Relationships (PANTHER) database.

This research is carried out in Phyton with R programming base with implementation of “Rpy2” Python package index (Gautier, 2008), used as statistical analysis freely available at http://rpy.sourceforge.net

Classification support vector machine (10-fold CV) is used for performance measurement in this research.

Genecards that are available at www.gencards.com are used for biological validation of selected genes in subpathway.

1.6 Significance of Research

This research is conducted to improve the performance of subpathway-based analysis method by modifying the search algorithm and taking account all the interactions between subpathways for identifying the informative subpathway and genes under case study. The significance of this study can be summarized as follows:

(a) Investigate the potential improvement of identification of perturbed subpathway within the pathway.

Provide clear information on the perturbed region related to targeted phenotype in a given pathway by using a computational method and analysis that provide better understanding in biological processes.

The development of subpathway-based analysis method can provide precise information in complex diseases which will eventually help medical team in targeted treatment design.

1.7 Thesis Outline

In this section, general description of each subsequent chapter is stated as below:

(a) Chapter 1 presents the introduction of this research including background of the problem, problem statement, goal, objectives, scope and significance of the study.

Chapter 2 presents the concept and recent trends applied by previous researchers related to the research topic. Besides, the details regarding the techniques and methods applied in the subpathway-based analysis on cancer diseases are explained and presented.

Chapter 3 states the research methodology including the research framework adopted in this study, datasets used, performance measurements and software requirements to achieve the goal and objectives.

Chapter 4 describes the proposed method in detail, an improved differential expression analysis for pathway by modifying the search algorithm and averaging the maximum absolute value (termed DEAP score) of subpathway. Besides, the data preparation and the result are explained and discussed wisely.

Chapter 5 concludes the research study. The contribution, limitations, and future work suggestions for this research are also presented.

REFERENCES

Alinejad, V., Dolati, S., Motallebnezhad, M., & Yousefi, M. (2017). The role of IL17B-IL17RB signaling pathway in breast cancer. Biomedicine & Pharmacotherapy, 88, 795-803.

Al-Mahdi, R., Babteen, N., Thillai, K., Holt, M., Johansen, B., Wetting, H. L., & Wells, C. M. (2015). A novel role for atypical MAPK kinase ERK3 in regulating breast cancer cell morphology and migration. Cell adhesion & migration, 9(6), 483-494.

Ardalan Khales, S., Ebrahimi, E., Jahanzad, E., Ardalan Khales, S., & Forghanifard, M. M. (2018). MAML1 and TWIST1 co-overexpression promote invasion of head and neck squamous cell carcinoma. Asia-Pacific Journal of Clinical Oncology.

Balomenos, P., Dragomir, A., Tsakalidis, A. K., & Bezerianos, A. (2020, July). Identification of differentially expressed subpathways via a bilevel consensus scoring of network topology and gene expression. In 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) (pp. 5316-5319). IEEE.

Beamer, S., Asanović, K., & Patterson, D. (2013). Direction-optimizing breadth-first search. Scientific Programming, 21(3-4), 137-148.

Banerjee, K., & Resat, H. (2016). Constitutive activation of STAT 3 in breast cancer cells: A review. International journal of cancer, 138(11), 2570-2578.

Ben-Shaul, Y., Bergman, H., & Soreq, H. (2005). Identifying subtle interrelated changes in functional gene categories using continuous measures of gene expression. Bioinformatics, 21(7), 1129-1137.

Bezerianos, A., Dragomir, A. and Balomenos, P. (2017). Time-Varying Methods for Pathway and Sub-pathway Analysis. In *Computational Methods for Processing and Analysis of Biological Pathways* (pp. 47-68). Springer, Cham.

Boeckx, C., Weyn, C., Bempt, I. V., Deschoolmeester, V., Wouters, A., Specenier, P., & Peeters, M. (2014). Mutation analysis of genes in the EGFR pathway in Head and Neck cancer patients: implications for anti-EGFR treatment response. BMC research notes, 7(1), 337.

Breslin, T., Edén, P., & Krogh, M. (2004). Comparing functional annotation analyses with Catmap. BMC Bioinformatics, 5(1), 193.

Bravatà, V., Cammarata, F. P., Forte, G. I., & Minafra, L. (2013). "Omics" of HER2-positive breast cancer. Omics: a journal of integrative biology, 17(3), 119-129.

Capaccione, K. M., & Pine, S. R. (2013). The Notch signaling pathway as a mediator of tumor survival. Carcinogenesis, 34(7), 1420-1430.

Cedars, E., Johnson, D. E., & Grandis, J. R. (2018). Jak/STAT Signaling in Head and Neck Cancer. In Molecular Determinants of Head and Neck Cancer (pp. 155-184). Humana Press, Cham.

Chalabi, N., Satih, S., Delort, L., Bignon, Y. J., & Bernard-Gallon, D. J. (2007). Expression profiling by whole-genome microarray hybridization reveals differential gene expression in breast cancer cell lines after lycopene exposure. Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression, 1769(2), 124-130.

Chen X, Xu J, Huang B et al (2011) A sub-pathway-based approach for identifying principal network. Bioinformatics 27:649–654

Choudhary, M. M., France, T. J., Teknos, T. N., & Kumar, P. (2016). Interleukin-6 role in head and neck squamous cell carcinoma progression. World journal of otorhinolaryngology-head and neck surgery, 2(2), 90-97.

Chung, S., Low, S. K., Zembutsu, H., Takahashi, A., Kubo, M., Sasa, M., & Nakamura, Y. (2013). A genome-wide association study of chemotherapy-induced alopecia in breast cancer patients. Breast Cancer Research, 15(5), R81.

Draghici, S., Khatri, P., Tarca, A.L., Amin, K., Done, A., Voichita, C., Georgescu, C. and Romero, R. (2007). A systems biology approach for pathway level analysis. *Genome research*, *17*(10), 000-000.

Davies PA, Pistis M, Hanna MC, Peters JA, Lambert JJ, Hales TG, Kirkness EF. The 5-HT3B subunit is a major determinant of serotonin receptor function. Nature. 1999;397:359–363.

De Carvalho, T. G., De Carvalho, A. C., Maia, D. C. C., Ogawa, J. K., Carvalho, A. L., & Vettore, A. L. (2013). Search for mutations in signaling pathways in head and neck squamous cell carcinoma. Oncology reports, 30(1), 334-340.

Dorman, S. N., Viner, C., & Rogan, P. K. (2014). Splicing mutation analysis reveals previously unrecognized pathways in lymph node-invasive breast cancer. Scientific reports, 4, 7063.

Dou, X. W., Liang, Y. K., Lin, H. Y., Wei, X. L., Zhang, Y. Q., Bai, J. W., ... & Tian, J. (2017). Notch3 Maintains Luminal Phenotype and Suppresses Tumorigenesis and Metastasis of Breast Cancer via Trans-Activating Estrogen Receptor-α. Theranostics, 7(16), 4041.

Eckert, L. B., Repasky, G. A., Ülkü, A. S., McFall, A., Zhou, H., Sartor, C. I., & Der, C. J. (2004). Involvement of Ras activation in human breast cancer cell signaling, invasion, and anoikis. Cancer research, 64(13), 4585-4592.

Elkhadragy, L., Chen, M., Miller, K., Yang, M. H., & Long, W. (2017). A regulatory BMI1/let-7i/ERK3 pathway controls the motility of head and neck cancer cells. Molecular oncology, 11(2), 194-207.

Emmert-Streib, F., Tripathi, S. and de Matos Simoes, R. (2012). Harnessing the complexity of gene expression data from cancer: from single gene to structural pathway methods. *Biology Direct*, *7*(1), 44.

Fearon, E. R. (2011). Molecular genetics of colorectal cancer. Annual Review of Pathology: Mechanisms of Disease, 6, 479-507.

Flanagan, J. M., Healey, S., Young, J., Whitehall, V., Trott, D. A., Newbold, R. F., & Chenevix-Trench, G. (2004). Mapping of a candidate colorectal cancer tumor-suppressor gene to a 900-kilobase region on the short arm of chromosome 8. Genes, Chromosomes, and Cancer, 40(3), 247-260.

Fu, Y. P., Edvardsen, H., Kaushiva, A., Arhancet, J. P., Howe, T. M., Kohaar, I., & Ambs, S. (2010). NOTCH2 in breast cancer: association of SNP rs11249433 with gene expression in ER-positive breast tumors without TP53 mutations. Molecular cancer, 9(1), 113.

Fukusumi, T., Guo, T. W., Sakai, A., Ando, M., Ren, S., Haft, S., & Califano, J. A. (2018). The NOTCH4–HEY1 Pathway Induces Epithelial–Mesenchymal Transition in Head and Neck Squamous Cell Carcinoma. Clinical Cancer Research, 24(3), 619-633

Garcia, R., & Jove, R. (1998). Activation of STAT transcription factors in oncogenic tyrosine kinase signaling. Journal of biomedical science, 5(2), 79-85.

Gershon MD, Tack J. The serotonin signaling system: from basic understanding to drug development for functional GI disorders. Gastroenterology. 2007;132:397–414

Giudice, F. S., & Squarize, C. H. (2013). The determinants of head and neck cancer: Unmasking the PI3K pathway mutations. Journal of carcinogenesis & mutagenesis.

Gooch, J. L., Christy, B., & Yee, D. (2002). STAT6 mediates interleukin-4 growth inhibition in human breast cancer cells. Neoplasia, 4(4), 324-331.

Goodsell, D. S. (1999). The molecular perspective: the ras oncogene. The oncologist, 4(3), 263-264.

Grabowski, P., Schönfelder, J., Ahnert-Hilger, G., Foss, H. D., Heine, B., Schindler, I., & Scherübl, H. (2002). Expression of neuroendocrine markers: a signature of human undifferentiated carcinoma of the colon and rectum. Virchows Archiv, 441(3), 256-263.

Hagerstrand, D., Tong, A., Schumacher, S. E., Ilic, N., Shen, R. R., Cheung, H. W., & Rosenbluh, J. (2013). Systematic interrogation of 3q26 identifies TLOC1 and SKIL as cancer drivers. Cancer discovery, CD-12.

Haynes, W. A., Higdon, R., Stanberry, L., Collins, D., & Kolker, E. (2013). Differential expression analysis for pathways. PLoS Comput Biol, 9(3), e1002967.

Hong, Y., Ho, K. S., Eu, K. W., & Cheah, P. Y. (2007). A susceptibility gene set for early onset colorectal cancer that integrates diverse signaling pathways: implication for tumorigenesis. Clinical Cancer Research, 13(4), 1107-1114.

Houghton, J. A., Harwood, F. G., Gibson, A. A., & Tillman, D. M. (1997). The fas signaling pathway is functional in colon carcinoma cells and induces apoptosis. Clinical cancer research, 3(12), 2205-2209.

Huang, F., Wang, D., Yao, Y., & Wang, M. (2017). PDGF signaling in cancer progression. Int J Clin Exp Med, 10(7), 9918-9929.

Hyakusoku, H., Sano, D., Takahashi, H., Hatano, T., Isono, Y., Shimada, S., & Oridate, N. (2016). JunB promotes cell invasion, migration and distant metastasis of head and neck squamous cell carcinoma. Journal of Experimental & Clinical Cancer Research, 35(1), 6.

Iacopetta, B. (2003). TP53 mutation in colorectal cancer. Human mutation, 21(3), 271-276.

Ihnatova, I., Popovici, V., & Budinska, E. (2018). A critical comparison of topology-based pathway analysis methods. *PloS one*, *13*(1), e0191154.

Javaid, S., Zhang, J., Smolen, G. A., Yu, M., Wittner, B. S., Singh, A., & Schott, B. J. (2015). MAPK7 regulates EMT features and modulates the generation of CTCs. Molecular Cancer Research, molcanres-0604.

Joyce, T., Oikonomou, E., Kosmidou, V., Makrodouli, E., Bantounas, I., Avlonitis, S., & Pintzas, A. (2012). A molecular signature for oncogenic BRAF in human colon cancer cells is revealed by microarray analysis. Current cancer drug targets, 12(7), 873-898.

Jung, K., Kang, H., & Mehra, R. (2018). Targeting phosphoinositide 3-kinase (PI3K) in head and neck squamous cell carcinoma (HNSCC). Cancers of the Head & Neck, 3(1), 3.

Kao, P. Y., Leung, K. H., Chan, L. W., Yip, S. P., & Yap, M. K. (2017). Pathway analysis of complex diseases for GWAS, extending to consider rare variants, multi-omics, and interactions. Biochimica et Biophysica Acta (BBA)-General Subjects, 1861(2), 335-353.

Kelly, P., Moeller, B. J., Juneja, J., Booden, M. A., Der, C. J., Daaka, Y., ... & Casey, P. J. (2006). The G12 family of heterotrimeric G proteins promotes breast cancer invasion and metastasis. Proceedings of the National Academy of Sciences, 103(21), 8173-8178.

Khammanivong, A., Gopalakrishnan, R., & Dickerson, E. B. (2014). SMURF1 silencing diminishes a CD44-high cancer stem cell-like population in head and neck squamous cell carcinoma. Molecular cancer, 13(1), 260.

Khatri, P., Sirota, M., and Butte, A. J. (2012). Ten years of pathway analysis: current approaches and outstanding challenges. PLoS Comput Biol. 8(2): 1-10.

Khodarev, N. N., Minn, A. J., Efimova, E. V., Darga, T. E., Labay, E., Beckett, M., Mauceri, H.J., Roizman, B. and Weichselbaum, R. R. (2007). Signal transducer and activator of transcription 1 regulates both cytotoxic and prosurvival functions in tumor cells. Cancer research, 67(19), 9214-9220.

Kok, K., Nock, G. E., Verrall, E. A., Mitchell, M. P., Hommes, D. W., Peppelenbosch, M. P., & Vanhaesebroeck, B. (2009). Regulation of p110δ PI 3-kinase gene expression. PloS one, 4(4), e5145.

Kontomanolis, E. N., Kalagasidou, S., Pouliliou, S., Anthoulaki, X., Georgiou, N., Papamanolis, V., & Fasoulakis, Z. N. (2018). The Notch Pathway in Breast Cancer Progression. The Scientific World Journal, 2018.

Koromilas, A. E., & Sexl, V. (2013). The tumor suppressor function of STAT1 in breast cancer. Jak-Stat, 2(2), e23353.

Li, X., Shen, L., Shang, X., & Liu, W. (2015). Subpathway analysis based on signaling-pathway impact analysis of signaling pathway. PloS one, 10(7), e0132813.

Li, C., Li, X., Miao, Y., Wang, Q., Jiang, W., Xu, C., Li, J., Han, J., Zhang, F., Gong, B. and Xu, L. (2009). SubpathwayMiner: a software package for flexible identification of pathways. *Nucleic acids research*, *37*(19), e131-e131.

Li, C., Shang, D., Wang, Y., Li, J., Han, J., Wang, S., Yao, Q., Wang, Y., Zhang, Y., Zhang, C. and Xu, Y. (2012). Characterizing the network of drugs and their affected metabolic subpathways. *PLoS One*, *7*(10), e47326.

Lin, Q., Lai, R., Chirieac, L. R., Li, C., Thomazy, V. A., Grammatikakis, I., & Hamilton, S. R. (2005). Constitutive activation of JAK3/STAT3 in colon carcinoma tumors and cell lines: inhibition of JAK3/STAT3 signaling induces apoptosis and cell cycle arrest of colon carcinoma cells. The American journal of pathology, 167(4), 969-980.

Lu, Y., & Han, J. (2003). Cancer classification using gene expression data. Information Systems, 28(4), 243-268.

Lui, V. W., Hedberg, M. L., Li, H., Vangara, B. S., Pendleton, K., Zeng, Y., & Freilino, M. (2013). Frequent mutation of the PI3K pathway in head and neck cancer defines predictive biomarkers. Cancer discovery.

Lui, V. W., Xi, S., Raymond, C. L., Koppikar, P., & Grandis, J. R. (2006). Activation of STAT5 contributes to tumor growth and epithelial-mesenchymal transition in head and neck cancer.

Mangone, F. R., Brentani, M. M., Nonogaki, S., Begnami, M. D. F., Campos, A. H. J., Walder, F., & Federico, M. H. (2005). Overexpression of Fos-related antigen-1 in head and neck squamous cell carcinoma. International journal of experimental pathology, 86(4), 205-212.

Manzat Saplacan, R. M., Balacescu, L., Gherman, C., Chira, R. I., Craiu, A., Mircea, P. A., & Balacescu, O. (2017). The role of PDGFs and PDGFRs in colorectal cancer. Mediators of inflammation, 2017.

Maraziotis, I. A., Dimitrakopoulou, K., & Bezerianos, A. (2007). Growing functional modules from a seed protein via integration of protein interaction and gene expression data. *Bmc Bioinformatics, 8*(1), 408.

Miller, F. R., Soule, H. D., Tait, L., Pauley, R. J., Wolman, S. R., Dawson, P. J., & Heppner, G. H. (1993). Xenograft model of progressive human proliferative breast disease. JNCI: Journal of the National Cancer Institute, 85(21), 1725-1732.

Miller, M. A., & Zachary, J. F. (2017). Mechanisms and morphology of cellular injury, adaptation, and death. In Pathologic Basis of Veterinary Disease (Sixth Edition) (pp. 2-43).

Mitrea, C., Taghavi, Z., Bokanizad, B., Hanoudi, S., Tagett, R., Donato, M., Voichita, C. Draghici, S. (2013). Methods and approaches in the topology-based analysis of biological pathways. *Frontiers in physiology*, *4*, 278.

Mönch, R. (2016). The Growth Factor PDGF and its Signaling Pathways in Colorectal Cancer.

Mukhopadhyay, U. K., Cass, J., Raptis, L., Craig, A. W., Bourdeau, V., Varma, S., & Ferbeyre, G. (2016). Dataset of STAT5A status in breast cancer. Data in brief, 7, 490.

Naghavi, A. O., Ahmed, K. A., Kim, Y., & Caudell, J. J. (2017). Head and Neck Cancer Genes Predictive of Radioresistance and Detriment to Local Control. International Journal of Radiation Oncology• Biology• Physics, 99(2), S122-S123.

Nakamura, Y., Tanaka, F., Yoshikawa, Y., Mimori, K., Inoue, H., Yanaga, K., & Mori, M. (2008). PDGF-BB is a novel prognostic factor in colorectal cancer. Annals of surgical oncology, 15(8), 2129-2136.

Nakanishi, Y., Walter, K., Spoerke, J. M., O'Brien, C., Huw, L. Y., Hampton, G. M., & Lackner, M. R. (2016). Activating mutations in PIK3CB confer resistance to PI3K inhibition and define a novel oncogenic role for p110β. Cancer research, canres-2201.

Nam, D., and Kim, S. Y. (2008). Gene-set approach for expression pattern analysis. Briefings in Bioinformatics. 9(3): 189-197.

Ning, Z., Feng, C., Song, C., Liu, W., Shang, D., Li, M., ... & Yu, X. (2019). Topologically inferring active miRNA-mediated subpathways toward precise cancer classification by directed random walk. Molecular Oncology, 13(10), 2211-2226.

Nunez, A. R. (2016). The role of the interleukin-12/STAT4 axis in breast cancer.

Page, F., & Bishop, W. (1898). Recurrent Carcinoma Of The Female Breast Entirely Disappearing Under The Persistent Use Of Thyroid Extract Continued For Eighteen Months. The Lancet, 151(3900), 1460-1461.

Pang, X., Tang, Y. L., & Liang, X. H. (2018). Transforming growth factor β signaling in head and neck squamous cell carcinoma: Insights into cellular responses. Oncology letters, 16(4), 4799-4806.

Parr, C., Watkins, G., & Jiang, W. G. (2004). The possible correlation of Notch-1 and Notch-2 with clinical outcome and tumour clinicopathological parameters in human breast cancer. International journal of molecular medicine, 14(5), 779-786.

Patel, N., & Patel, K. M. (2015). A survey on: enhancement of minimum spanning tree. J. Eng. Res. Appl, 5(1 Part 3), 06-10.

Peck, A. R., Witkiewicz, A. K., Liu, C., Klimowicz, A. C., Stringer, G. A., Pequignot, E., ... & Girondo, M. A. (2012). Low levels of Stat5a protein in breast cancer are associated with tumor progression and unfavorable clinical outcomes. Breast Cancer Research, 14(5), R130.

Phan, N. N., Wang, C. Y., Chen, C. F., Sun, Z., Lai, M. D., & Lin, Y. C. (2017). Voltage-gated calcium channels: Novel targets for cancer therapy. Oncology letters, 14(2), 2059-2074.

Pennisi, Elizabeth. "Will computers crash genomics?." (2011): 666-668.

Qiu, W., Schönleben, F., Li, X., Ho, D. J., Close, L. G., Manolidis, S., & Su, G. H. (2006). PIK3CA mutations in head and neck squamous cell carcinoma. Clinical Cancer Research, 12(5), 1441-1446.

Rebhan, M., Chalifa-Caspi, V., Prilusky, J., & Lancet, D. (1998). GeneCards: a novel functional genomics compendium with automated data mining and query reformulation support. Bioinformatics (Oxford, England), 14(8), 656-664.

Reyes-Gibby, C. C., Wang, J., Silvas, M. R. T., Yu, R., Yeung, S. C. J., & Shete, S. (2016). MAPK1/ERK2 as novel target genes for pain in head and neck cancer patients. BMC genetics, 17(1), 40.

Riaz, N., Morris, L. G., Lee, W., & Chan, T. A. (2014). Unraveling the molecular genetics of head and neck cancer through genome-wide approaches. Genes & diseases, 1(1), 75-86.

Richardson, C., Zhang, S., Hernandez Borrero, L. J., & El-Deiry, W. S. (2017). Small-molecule CB002 restores p53 pathway signaling and represses colorectal cancer cell growth. Cell Cycle, 16(18), 1719-1725.

Rivetti, S., Lauriola, M., Voltattorni, M., Bianchini, M., Martini, D., Ceccarelli, C., & Rosati, G. (2011). Gene expression profile of human colon cancer cells treated with cross-reacting material 197, a diphtheria toxin non-toxic mutant. International journal of immunopathology and pharmacology, 24(3), 639-649.

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics, 26(1), 139-140.

Sánchez-Muñoz, A., Gallego, E., de Luque, V., Pérez-Rivas, L. G., Vicioso, L., Ribelles, N., & Alba, E. (2010). Lack of evidence for KRAS oncogenic mutations in triple-negative breast cancer. BMC cancer, 10(1), 136.

Sarhan, A. M. (2009). Cancer classification based on microarray gene expression data using DCT and ANN. Journal of Theoretical & Applied Information Technology, 6(2).

Sarnataro, D., Grimaldi, C., Pisanti, S., Gazzerro, P., Laezza, C., Zurzolo, C., & Bifulco, M. (2005). Plasma membrane and lysosomal localization of CB1 cannabinoid receptor are dependent on lipid rafts and regulated by anandamide in human breast cancer cells. FEBS letters, 579(28), 6343-6349.

Savas, S., Hyde, A., Stuckless, S. N., Parfrey, P., Younghusband, H. B., & Green, R. (2012). Serotonin transporter gene (SLC6A4) variations are associated with poor survival in colorectal cancer patients. PLoS One, 7(7), e38953.

Sikdar, S., Datta, S., & Datta, S. (2016). Exploring the importance of cancer pathways by meta-analysis of differential protein expression networks in three different cancers. Biology direct, 11(1), 65.

Slattery, M. L., Lundgreen, A., John, E. M., Torres-Mejia, G., Hines, L., Giuliano, A. R., & Wolff, R. K. (2015). MAPK genes interact with diet and lifestyle factors to alter risk of breast cancer: the Breast Cancer Health Disparities Study. Nutrition and cancer, 67(2), 292-304.

Slattery, M. L., Lundgreen, A., Kadlubar, S. A., Bondurant, K. L., & Wolff, R. K. (2013). JAK/STAT/SOCS-signaling pathway and colon and rectal cancer. Molecular carcinogenesis, 52(2), 155-166.

Slattery, M. L., Lundgreen, A., Kadlubar, S. A., Bondurant, K. L., & Wolff, R. K. (2013). JAK/STAT/SOCS-signaling pathway and colon and rectal cancer. Molecular carcinogenesis, 52(2), 155-166.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S. and Mesirov, J. P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*, *102*(43), 15545-15550.

Suda, T., Yoshihara, M., Nakamura, Y., Sekiguchi, H., Godai, T. I., Sugano, N., & Kameda, Y. (2011). Rare MDM4 gene amplification in colorectal cancer: The principle of a mutually exclusive relationship between MDM alteration and TP53 inactivation is not applicable. Oncology reports, 26(1), 49-54.

Sugano, N., Suda, T., Godai, T. I., Tsuchida, K., Shiozawa, M., Sekiguchi, H., & Kameda, Y. (2010). MDM2 gene amplification in colorectal cancer is associated with disease progression at the primary site, but inversely correlated with distant metastasis. Genes, Chromosomes and Cancer, 49(7), 620-629.

Sun, W., Gaykalova, D. A., Ochs, M. F., Mambo, E., Arnaoutakis, D., Liu, Y., & Ahn, S. (2014). Activation of the NOTCH pathway in head and neck cancer. Cancer research, 74(4), 1091-1104.

Tarca, A.L., Draghici, S., Khatri, P., Hassan, S.S., Mittal, P., Kim, J.S., Kim, C.J., Kusanovic, J.P. and Romero, R. (2008). A novel signaling pathway impact analysis. *Bioinformatics*, *25*(1), 75-82.

Tarca, A. L., Draghici, S., Khatri, P., Hassan, S. S., Mittal, P., Kim, J. S., ... & Romero, R. (2008). A novel signaling pathway impact analysis. *Bioinformatics*, *25*(1), 75-82.

Tarjan, R. (1972). Depth-first search and linear graph algorithms. SIAM journal on computing, 1(2), 146-160.

Thomas, P. D., Kejariwal, A., Campbell, M. J., Mi, H., Diemer, K., Guo, N., Ladunga, I., Ulitsky-Lazareva, B., Muruganujan, A., Rabkin, S. and Vandergriff, J. A. (2003). PANTHER: a browsable database of gene products organized by biological function, using curated protein family and subfamily classification. Nucleic acids research, 31(1), 334-341.

Uchiyama, T., Takahashi, H., Endo, H., Sugiyama, M., Sakai, E., Hosono, K., & Nakajima, A. (2011). Role of the long form leptin receptor and of the STAT3 signaling pathway in colorectal cancer progression. International journal of oncology, 39(4), 935-940.

Walker, M. G. (2001). Drug target discovery by gene expression analysis cell cycle genes. *Current cancer drug targets*, *1*(1), 73-83.

Wang, J. W., Wei, X. L., Dou, X. W., Huang, W. H., Du, C. W., & Zhang, G. J. (2018). The association between Notch4 expression, and clinicopathological characteristics and clinical outcomes in patients with breast cancer. Oncology letters, 15(6), 8749-8755.

Wang, Y., Klijn, J. G., Zhang, Y., Sieuwerts, A. M., Look, M. P., Yang, F., Talantov, D., Timmermans, M., Meijer-van Gelder, M.E., Yu, J. and Jatkoe, T. (2005). Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. The Lancet, 365(9460), 671-679.

Wang, X., Dalkic, E., Wu, M., Chan, C. (2008). Gene Module Level Analysis: Identification to Networks and Dynamics. Current Opinion in Biotechnology. 19(5): 482-491.

White, R. A., Malkoski, S. P., & Wang, X. J. (2010). TGFβ signaling in head and neck squamous cell carcinoma. Oncogene, 29(40), 5437.

Yamada, M., Monden, T., Konaka, S., & Mori, M. (1993). Assignment of human thyrotropin-releasing hormone (TRH) receptor gene to chromosome 8. Somatic cell and molecular genetics, 19(6), 577-580.

Yan, G. R., Xu, S. H., Tan, Z. L., Liu, L., & He, Q. Y. (2011). Global identification of miR-373-regulated genes in breast cancer by quantitative proteomics. Proteomics, 11(5), 912-920.

Yoo, A., Chow, E., Henderson, K., McLendon, W., Hendrickson, B., & Catalyurek, U. (2005, November). A scalable distributed parallel breadth-first search algorithm on BlueGene/L. In Proceedings of the 2005 ACM/IEEE conference on Supercomputing (p. 25). IEEE Computer Society.

Zhang, S., Liu, J., Xu, K., & Li, Z. (2018). Notch signaling via regulation of RB and p AKT but not PIK3CG contributes to MIA PaCa 2 cell growth and migration to affect pancreatic carcinogenesis. Oncology letters, 15(2), 2105-2110.

Zhang, W., Ding, E. X., Wang, Q., Zhu, D. Q., He, J., Li, Y. L., & Wang, Y. H. (2005). Fas ligand expression in colon cancer: a possible mechanism of tumor immune privilege. World journal of gastroenterology: WJG, 11(23), 3632.

Zhao, Y. Y., Yu, G. T., Xiao, T., & Hu, J. (2017). The Notch signaling pathway in head and neck squamous cell carcinoma: A meta-analysis. Advances in clinical and experimental medicine: official organ Wroclaw Medical University, 26(5), 881-887.

Zhao, Y., Fu, D., Xu, C., Yang, J., & Wang, Z. (2017). Identification of genes associated with tongue cancer in patients with a history of tobacco and/or alcohol use. Oncology letters, 13(2), 629-638.

Zheng, Z. Y., Tian, L., Bu, W., Fan, C., Gao, X., Wang, H., & Zwaka, T. P. (2015). Wild-type N-Ras, overexpressed in basal-like breast cancer, promotes tumor formation by inducing IL-8 secretion via JAK2 activation. Cell reports, 12(3), 511-524.

Zheng, Y., Sun, S., Yu, M., & Fu, X. (2019). Identification of potential hub-lncRNAs in ischemic stroke based on Subpathway-LNCE method. Journal of cellular biochemistry, 120(8), 12832-12842.

Zhou, C. Z., Qiu, G. Q., Fang Zhang, L. H., & Peng, Z. H. (2004). Loss of heterozygosity on hromosome 1 in sporadic colorectal carcinoma. World journal of gastroenterology, 10(10), 1431.Lv, Y. and Gao, J. (2011) ‘Condition prediction of chemical complex systems based on Multifractal and Mahalanobis-Taguchi system’, in *ICQR2MSE 2011 - Proceedings of 2011 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering*, pp. 536–539.