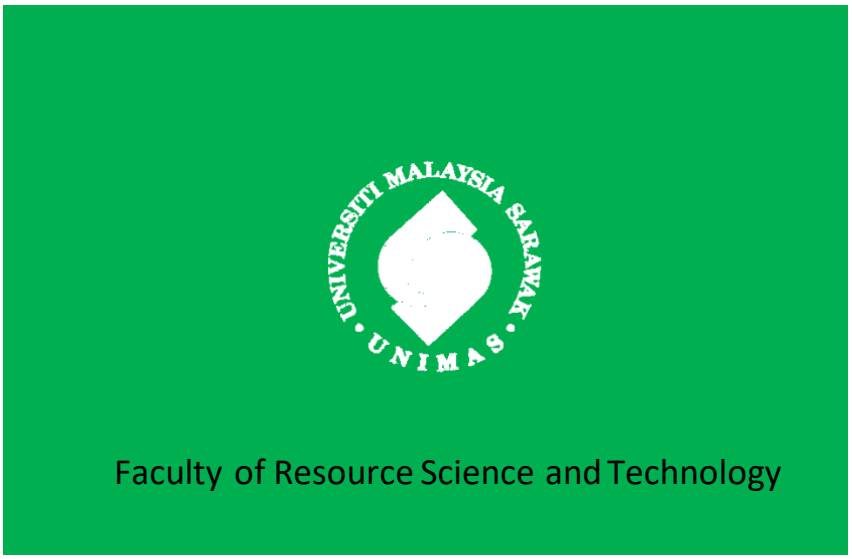

Bibliographic Analysis of p53-gene Family and Gene Expression Analysis of p53 in Nasopharyngeal Cancer

Emmanuella Nina Christina Wakiman (69597)

Bachelor of Science with Honours
(Resource Biotechnology)
2022

Bibliographic Analysis of p53-gene Family and Gene Expression Analysis of p53 in Nasopharyngeal Carcinoma

Emmanuella Nina Christina Wakiman

A thesis submitted in partial fulfilment of the requirement of
The Degree of Bachelor Science with Honours
(Resource Biotechnology)

SUPERVISOR: ASSOCIATE PROFESSOR DR. EDMUND SIM UI HANG

Programme of Resource Biotechnology
Faculty of Resource Science and Technology
UNIVERSITY MALAYSIA SARAWAK
2022

UNIVERSITI MALAYSIA SARAWAK

Grade: _____

Please tick ( )

Final Year Project Report ☒

Masters ☐

PhD

DECLARATION OF ORIGINAL WORK

This declaration is made on the 12 of June 2022.

Student's Declaration:

I Emmanuella Nina Christina, 69597, Faculty of Resource Science and Technology (PLEASE INDICATE STUDENT'S NAME, MATRIC NO. AND FACULTY) hereby declare that the work entitled, Bibliographic Analysis of p53-Gene Family and Gene Expression Analysis of p53 in Nasopharyngeal Carcinoma is my original work. I have not copied from any other students' work or from any other sources except where due reference or acknowledgement is made explicitly in the text, nor has any part been written for me by another person.

15/06/2022 — Date submitted

Emmanuella Nina Christina Wakiman (69597) — Name of the student (Matric No.)

Supervisor's Declaration:

I, PROF DR EDMUND SIM UI HANG (SUPERVISOR'S NAME), hereby certify that the work entitled, as stated in the student's declaration section (TITLE) was prepared by the above named student, and was submitted to the "FACULTY" as a * partial/fullfulfillment for the conferment of BSc. Hons. (Resource Biotechnology) (PLEASE INDICATE THE DEGREE), and the aforementioned work, to the best of my knowledge, is the said student's work

Prof. Dr. Edmund Sim Ui Hang — Received for examination by: _____ (Name of the supervisor)

15 June 2022 — Date: _____

I declare this Project/Thesis is classified as (Please tick (√)):

Act 1972)*	☐	**CONFIDENTIAL** (Contains confidential information under the Official Secret
	☐	**RESTRICTED** (Contains restricted information as specified by the
organisation	☑	where research was done)*
		OPEN ACCESS

Validation of Project/Thesis

I therefore duly affirmed with free consent and willingness declared that this said Project/Thesis shall be placed officially in the Centre for Academic Information Services with the abide interest and rights as follows:

- This Project/Thesis is the sole legal property of Universiti Malaysia Sarawak (UNIMAS).
- The Centre for Academic Information Services has the lawful right to make copies for the purpose of academic and research only and not for other purpose.
- The Centre for Academic Information Services has the lawful right to digitise the content to for the Local Content Database.
- The Centre for Academic Information Services has the lawful right to make copies of the Project/Thesis for academic exchange between Higher Learning Institute.
- No dispute or any claim shall arise from the student himself / herself neither third party on this Project/Thesis once it becomes sole property of UNIMAS.
- This Project/Thesis or any material, data and information related to it shall not be distributed, published or disclosed to any party by the student except with UNIMAS permission.

Student's signature _____ (15/06/2022)

Supervisor's signature: _____ (20 July 2022)

Current Address:

No. 34, Medan Ginseng, Jalan Sibiyu, 97000 Bintulu, Sarawak

Notes: ***** If the Project/Thesis is **CONFIDENTIAL** or **RESTRICTED**, please attach together as annexure a letter from the organisation with the period and reasons of confidentiality and restriction.

[The instrument was duly prepared by The Centre for Academic Information Services]

ACKNOWLEDGEMENT

First and foremost, I would like to thank all individuals, parties, and organizations that have contributed and cooperated throughout this Final Year Project especially the Faculty of Resource Science and Technology, University Malaysia Sarawak for this invaluable opportunity to conduct the research. Special thanks are directed to Associate Professor Dr. Edmund Sim Ui Hang for his undying support, willingness to share his knowledge and experiences, as well the useful advices throughout my whole final year project. Furthermore, I would like to express my appreciation and gratitude toward my beloved family, especially my parents, Mr Wakiman and Mdm Rain Udup, for their undying supports and motivation throughout the research. I would also like to thank the postgraduate students inside Immunology Laboratory, UNIMAS, namely Adrianne Vong Dien Yu, Alyaa Syafiqah binti Shamsuri and Daphne Olivia Jawai anak Sadai for their guidance and invaluable supervision towards the completion of this final year project.

Bibliographic Analysis of p53-gene Family and Gene Expression Analysis of p53 in Nasopharyngeal Carcinoma

Emmanuella Nina Christina Wakiman

Resource Biotechnology Programme
Faculty of Resource Science and Technology
Universiti Malaysia Sarawak

ABSTRACT

Nasopharyngeal carcinoma (NPC) is relatively known as a unique malignancy associated with an oncovirus, Epstein-Barr virus (EBV), displaying genetic, epidemiological and regional distribution properties. Many researches have proved that p53 mutations are involved in NPC, but the precise gene expression of p53 is only partially understood. Identification on the dysregulation expression level of p53 in NPC can provide new insight into the future research and pathology of NPC by the discovery of a new prognostic marker. p53 is a tumor suppressor gene that most frequently mutated found in a large variety of human cancer epidemiology. In this study, bibliographical data was conducted fundamentally by Rstudio to study the publication and locality trend of p53 gene family from 2013 to 2022 across the world. As a whole, the data represents that the highest annual scientific production was on 2013, 7924 authors (87.41%) published 1 document each. With respect to citations, 2019 has the highest average article citations per year and the main country in terms of number of publications is China. The result clearly shows that USA is the most cited countries. On the other hand, for gene expression of p53 in NPC, the differential expression level of p53 between NP69 and HK1 are not able to be concluded since the intensity of GAPDH for NP69 and HK1 in Duplicate 1 are not the same. Besides, there is a band presence in GAPDH negative control.

Keywords: p53, nasopharyngeal carcinoma, gene expression, bibliographic analysis

ABSTRAK

Karsinoma nasofaring agak dikenali sebagai malignan unik yang dikaitkan dengan virus onkogenik, virus Epstein-Barr (EBV), yang memaparkan genetik, epidemiologi dan pengedaran serantau. Banyak penyelidikan telah membuktikan bahawa mutasi p53 terlibat dalam NPC, tetapi ungkapan gen p53 yang tepat hanya difahami sebahagiannya. Pengenalpastian pada tahap ekspresi disregulasi p53 dalam NPC boleh memberikan wawasan baru dalam penyelidikan masa depan dan patologi NPC dengan penemuan penanda prognostik baru. P53 adalah gen penindas tumor yang paling kerap bermutasi yang terdapat dalam pelbagai jenis epidemiologi kanser manusia. Dalam kajian ini, data bibliografi dijalankan secara asasnya oleh Rstudio untuk mengkaji trend penerbitan dan lokaliti gen p53 dari 2013 hingga 2022 di seluruh dunia. Secara keseluruhannya, daa menunjukkan bahawa pengeluaran saintifik tahunan tertinggi adalah pada tahun 2013, sebanyak 7924 penulis (87.41%) menerbitkan 1 dokumen setiap seorang. Berkenaan dengan petikan, 2019 mempunyai petikan artikel purata tertinggi setiap tahun dan negara yang mempunyai bilangan penerbitan tertinggi adalah China. Hasil kajian jelas menunjukkan bahawa petikan terbanyak diambil dari Amerika Syarikat. Sebaliknya, untuk ekspresi gen p53 dalam NPC, tahap ekspresi pembezaan p53 antara NP69 dan HK1 tidak dapat disimpulkan kerana intensiti GAPDH untuk NP69 dan HK1 dalam pendua 1 tidak sama. Selain itu, terdapat kehadiran band dalam kawalan negatif GAPDH.

Kata kunci: *p53, karsinoma nasofaring, ekspresi gen, data bibliografi*

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AGE	Agarose Gel Electrophoresis
BLAST	Basic Local Alignment Search Tool
cDNA	Complementary Deoxyribonucleic Acid
dNTPs	Deoxyribonucleotide Triphosphates
DNA	Deoxyribonucleic Acid
EBV	Epstein-Barr Virus
FBS	Fetal Bovine Serum
GAPDH	Glyceraldehyde-3-Phosphate Dehydrogenase
KSFM	Keratinocyte-serum Free Media
M MLV-RT	Moloney Murine Leukemia Virus Reverse Transcriptase
NCBI	National Centre for Biotechnology Information
NPC	Nasopharyngeal Carcinoma
mRNA	Messenger Ribonucleic Acid
p53	Protein 53
PCR	Polymerase Chain Reaction
PBS	Phosphate Buffer Saline
rRNA	Ribosomal Ribonucleic
RPMI	Roswell-Park Memorial Institute
RT-PCR	Reverse-Transcription Polymerase Chain Reaction

CHAPTER 1

INTRODUCTION

Nasopharyngeal carcinoma (NPC) is relatively known as a unique malignancy associated with an oncovirus, Epstein-Barr virus (EBV), that takes place in epithelium covering the nasopharynx's lateral wall. According to the geographical distribution as mentioned by Le et al., (2019), NPC has demonstrated the highest incidence rate, particularly in Southeast China, Southeast Asia, North Africa, and the Arctic regions and quite rare in the other part of the world. Bibliographic analysis was used in this research to assess the productivity of scientific outputs quantitatively. According to Ellegaard and Wallin (2015), bibliographic analysis is an integral part of research evaluation methodology and is now developed as scientific specialties. Basically, it is essential for models of bibliographic data to consider several types of information, for instance, the citation network, author's research areas and the paper content as articles normally accompanied by metadata such as publication trend and locality.

Gene expression is a process in which the informations stored in DNA are converted into instructions to produce a functional product, such as RNA and proteins, which involves transcription and translation processes (Schwanhausser *et al.*, 2011). Protein 53 or p53 which is located on the short arm of chromosome 17 (Agaoglu *et al.*, 2004), is a tumor suppressor gene that is most frequently mutated gene in human cancer compared with other genes. It is phosphorylated in normal cells as a result of stress signal such as presence of genomic damage. When there is mutation in the cell, the phosphorylated p53 then transactivates genes that function in apoptosis, a programmed cell death, hence allows the repair of damaged cell or elimination of that particular cell permanently in accordance with the composition of the previous normal amino acids due to its protooncogene nature. According to Volgelstein, Lane and Levine (2000), clustering of mutations can take place after the p53 function is

inactivated by interaction with viral proteins, which eventually may result in aggressive cancer growth.

While p53 has been known for decades, two other members of the p53 family, known as p63 and p73, have been discovered more recently. Murray-Zmijewski *et al.,* (2006) stated that the three genes possess a high degree of homology and there is increasing evidence that they have risen from the triplication of a common ancestral gene . All three genes consist of important structural elements including a transactivation domain, an oligomerization domain and a central DNA-binding domain (Muller *et al.,* 2006). p63 and p73 have been shown to induce apoptosis similarly to p53 via activation of several of its downstream target genes (Schilling *et al.,* 2009). Yet, both family members also exhibit functions distinct from p53.

Many researches have proved that p53 mutations are involved in nasopharyngeal carcinogenesis, but the precise gene expression of p53 is only partially understood. In this research, the bibliographic analysis of p53 gene family was evaluated and the expression level of p53 in NPC was analyzed. The expression level of p53 in HK1 nasopharyngeal cancer cell line is to provide insight into the pathology of NPC and can add new biological marker for ongoing p53 studies. Human nasopharyngeal cell lines, NP69 and HK1 cell lines were used in this research for the isolation of p53. The null hypothesis of this study is there is no significant difference in the expression of p53 between NP69 and HK1 cell lines.

The objectives of this study were:

1. To study the p53-gene family publication trend and locality
2. To determine the gene expression level of p53 between a normal nasopharyngeal cell line, NP69, and a nasopharyngeal carcinoma cell line, HK1

CHAPTER 2

LITERATURE REVIEW

2.1 p53 Gene Family

p53-gene family compose a family of transcription factors involved in cell response to stress and developmental process (Pflaum *et al.,* 2014). p53 was discovered in 1979 and known as a tumor suppressor gene that plays vital role inside the nucleus that encodes protein and controls cell division and cell death, and loss of p53 activity is considered to be ubiquitous to all cancers (Sauer *et al.,* 20008). According to National Human Genome Research Institute (2022), p53 mutation may cause the protein to not function properly, thus lead to uncontrolled growth and possibly spreading of cancer in the body, and it also known as anti-oncogene. Recently, a few studies have emphasised new transcriptional targets downstream of p53 promoting genomic integrity in tumor suppression though it was characterized as a regulator for responses to a few of stress signals such as acute damage of DNA, hypoxia and oncogene expression (Boutelle & Attardi, 2021).

p63 and p73 were identified in 1997 and they are rarely mutated in human cancers unlike p53 (Pflaum *et al.,* 2014). Since their sequence has high level of similarity with p53, it allows them to transactivate p53-responsive genes which causes cell cycle arrest and apoptosis to occur. Hence, they can form a family of transcription factor along with p53. The overall protein architecture is highly conserved among the three family members and the highest degree of sequence is observed in the DNA binding core domain (Sauer *et al.,* 2008).

2.2 p53 and NPC

To date, point mutations in the p53 gene are found in NPC but the rate of mutation is remarkably smaller compared to other kind of tumors (Agaoglu *et al.,* 2004). Although also known as tumor protein p53 (TP53) gene, these changes have been observed in LiFraumeni syndrome and vast majority types of cancer. There are numerous studies conducted in conjunction with the p53 and NPC research, and there were reports stated that p53 is rarely mutated in NPC (Van *et al.*,1997). However, there were studies reported a high percentage of cases of NPC, as detected by immunohistochemistry, to have over expression of the p53 protein. According to a study by Agaoglu *et al., (*2004), they confirm that p53 overexpression is present in a subset of patients with NPC although the relationship between NPC and p53 is not transparent. Initially, the p53 was classified as an oncogene ever since it was seen in excessive amounts in malignant cells. It is able to prevent growth of cell caused by oncogenes and also could inhibit the tumorigenic potential of cells in animals.

2.3 Nasopharyngeal Cancer

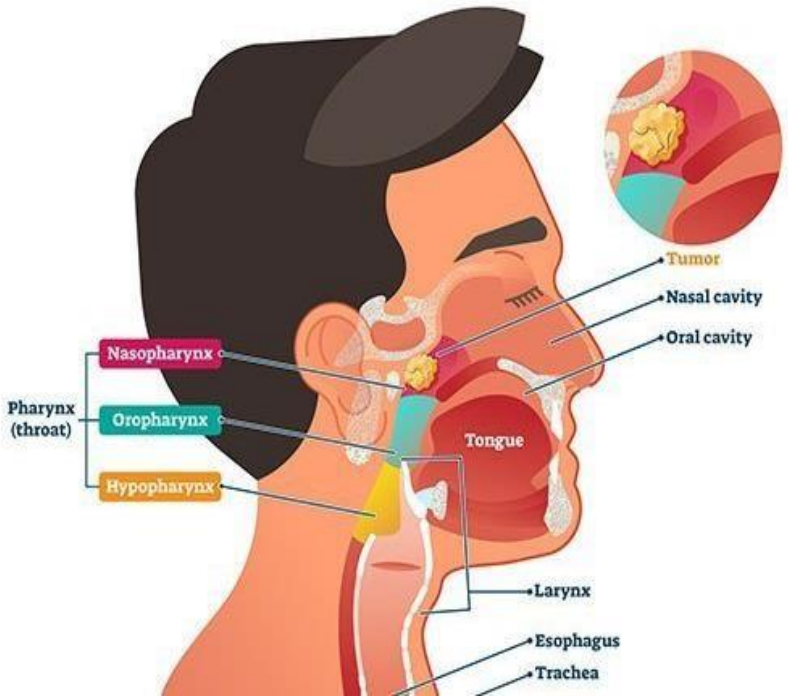

Figure 1: Location of nasopharyngeal tumor

(Adapted from

https://www.emedicinehealth.com/what_is_survival_rate_of_nasopharyngeal _carcinoma/article_em.htm)

According to Chou *et al.,* (2008), nasopharyngeal cancer is a type of head and neck cancer which starts in the nasopharynx, which is located in the upper part of the throat and near skull's base. In Malaysia, NPC is the fourth most common cancer and it is predominant among Chinese, natives of Sabah, native of Sarawak especially Bidayuh and Malay. It mainly consists of three categories, which are squamous cell carcinoma, non-keratinizing carcinoma and undifferentiated carcinoma (Brennan, 2006). Normally, cancer develops when cells begin to develop uncontrollably within almost every body part and are able to spread to other parts and become a tumor. However, not all tumor cells are benign, or in

other word, they are non-cancerous. Some of them are malignant, or cancerous. Infection of EBV and epigenetic changes disrupt normal cell function during growth of the tumor, thus leading to NPC pathogenesis (Tao & Chan, 2007). Even though survival rate has increased due to the presence of advanced technology, there are still a lot of controversies floating over the available treatment methods. Firstly, the cancer sufferers still encounter distant metastasis even after they went through radiotherapy treatments, especially for cases involving those in advanced stages (Alotaibi *et al.,* 2019). Next, most patients did not give some positive response for the treatments mentioned above, especially those in advanced stages too.

2.4 Cancer Cell Lines

2.4.1 NP69 Cell Line

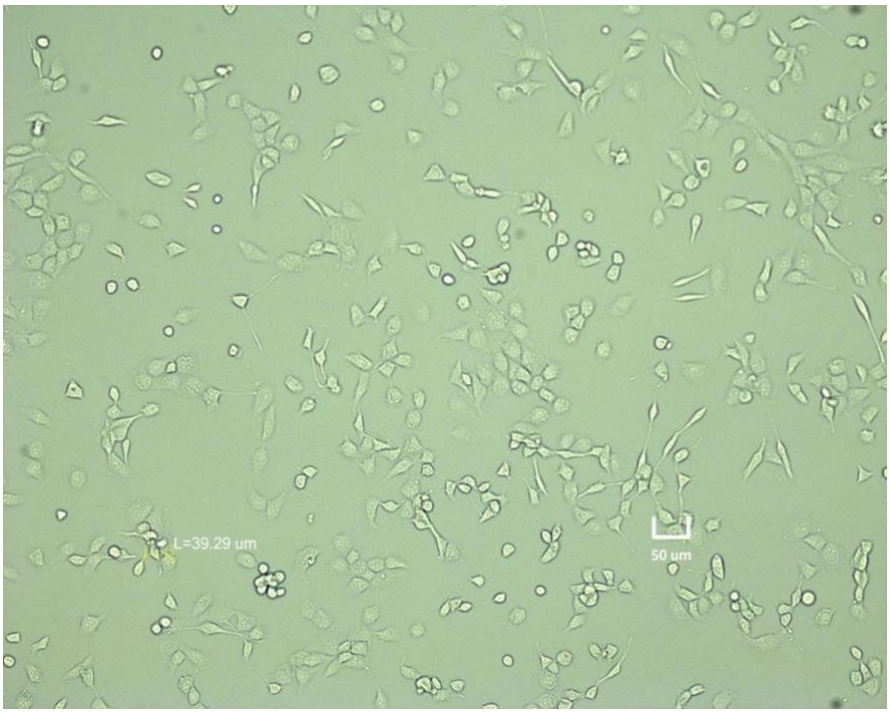

Figure 2: Image of NP69 cell line from Inverted Fluorescent Microscope**.**

NP69 cell line was isolated from passage 1 cultures of biopsy-isolated primary nasopharyngeal epithelial cells transfected with PX8 plasmid containing the simian virus 40 large T antigen. It is non-tumorigenic and possessed anchorage-dependent growth. The keratin profile and response to TGF-beta, while possessing several of genetic signatures observed in NPC retain a lot of the normal nasopharyngeal cells criteria.

2.4.2 HK1 Cell Line

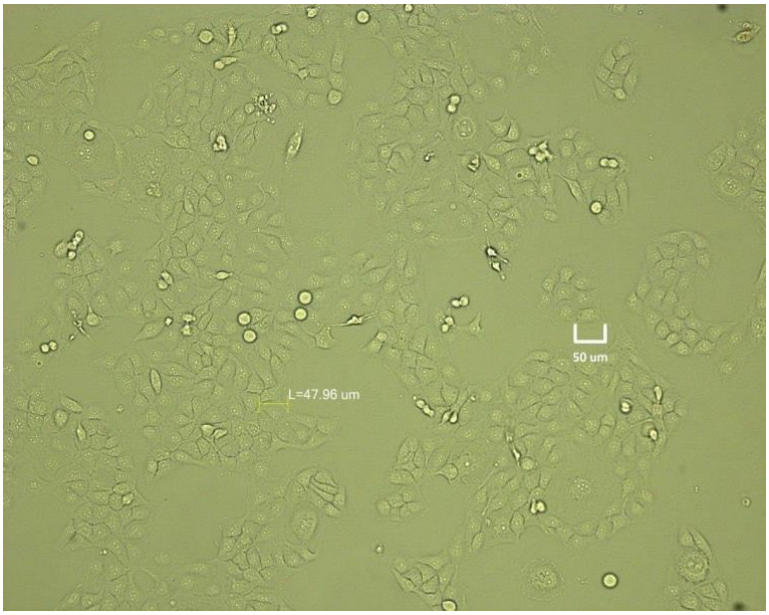

Figure 3: Image of HK1 cell line from Inverted Fluorescent Microscope.

HK1 (accession number: CVCL_7084) cell line was isolated by Professor DP Huang and her colleagues in Hong Kong which was outlined to be one of the EBV-negative NPC cell lines. It was established from the well-differentiated squamous carcinoma of the nasopharynx of a 58 year old male. (Expasy is operated by the SIB Swiss Institute of Bioinformatics | Terms of Use, 2021). Two touch smears from freshly biopsied NPC tissues of HK1 were inconclusive for EBNA expression. It was shown to be tumorigenic cell line in nude mice after transplantation and able to form a well-differentiated tumor at the sites of inoculation eventually, which similar in morphology to the recurrent human tumor from which they originally derived (Huang *et al.*, 1980).

2.5 Housekeeping Genes

Housekeeping genes are known as cellular maintenance genes which regulate basic and ubiquitous cellular functions. Thus, they are valid candidates to act as internal control for the gene expression of p53 in NPC (Turabelidze, Guo & DiPietro, 2010). To date, the most commonly used housekeeping genes are GAPDH, beta actin and ubiquitin. They play a vital role in the final result interpretation. They act as a reference gene that is required to correct for basic simple differences, such as RNA quality, differences in cellular input, and efficiency of reverse transcription process (de Kok *et al.*, 2005)

2.5.1 *GAPDH*

Glyceraldehyde-3-phosphate dehydrogenase (*GAPDH)* has been used extensively by cancer researchers as internal reference controls, due to the reasons that their mRNA expression levels are established to be constant and high in various types of tissues and cells (Zhu *et al,* 2008) and for normalization of gene expression data. As mentioned by Tang *et al.,* (2012), it catalyzes redox reaction in glycolysis by converting glyceraldehyde-3phosphate to 1,3-biphosphoglycerate through the reduction of NAD^+ to NADH. Through mRNA expression analysis of *GAPDH* measured from a panel of 72 human tissues, there was no significant specific effect of gender and age noticed on expression of *GAPDH*, thus providing standard values for levels *GAPDH* mRNA expression in the tissues studied. These data have established comparative levels of expression and may be used to increase value to gene expression data in which *GAPDH* is used as the internal control (Barber *et al.,* 2005).

CHAPTER 3

MATERIAL AND METHODS

3.1 Bibliographic Analysis of p53 gene family

3.1.1 Data Collection

Scopus was used to thoroughly search for the textual term “p53 Gene Family” to retrieve all the document set as it contains bibliographic data. For this research, Scopus database was chosen because it has the largest database in the world and has thousands of research papers. The document set was filtered for limitation of 10 years, ranging from 2013 to 2022.

3.1.2 Selection Criteria of Eligible Studies

Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Checklist was used for validating search queries purpose. PRISMA is an evidence-based minimum set of items for reporting in systematic reviews and meta-analyses and there were 27 checklist items found in PRISMA 2020 Checklist. Since Scopus database was used to figure out publications that had been collected, the duplicated search files found was deleted and displayed again to exclude the unnecessary files. Next, the full-text articles were evaluated for eligibility and those full-text articles that did not fulfil the objective scope was deleted. After that, the total number of studies included in qualitative synthesis was exported into a file.

3.1.3 Data Analysis

All the data previously retrieved from Scopus database was analyzed with Rstudio.

This bibliometrix package was installed with the latest version 4.2.0 in Windows 11. Biblioshiny was chosen because it precisely provides very useful visual representations of bibliographic data, thus helping to comprehend the nature of the literature as well as identify potential gaps in the literature. Second, biblioshiny code was utilised to assess the updated shiny web-interface (version 2.0). Once the R and Rstudio packages were installed, biblioshiny code was keyed in the Rstudio, namely library(bibliometrix) and biblioshiny(). The screen was redirected to the biblioshiny web and bibliographic data was analysed. The downloaded file from Scopus was imported to the biblioshiny web. The meta package in the Rstudio was then used to carry out the publication trend and locality trend of p53 gene family. The results were recorded.

3.2 Gene Expression Analysis of p53 in NPC

3.2.1 Cell Culture

One normal nasopharyngeal epithelial cell line, NP69 and one nasopharyngeal cancer cell line, HK1, were used in this research. The cell lines were provided by the postgraduate students in T25 flasks and they were cultured at 37℃ with 95% of humidity and 5% of carbon dioxide, which were securely grown in a CO^{2+} incubator. The normal cell, NP69 was cultured in Keratinocyte-serum free media (KSFM) supplemented with human recombinant epidermal growth, bovine pituitary extracts and 1% antibiotics, Penicillin-Streptomycin. As for the NPC cell line, HK1, it was cultured and maintained in Roswell-Park Memorial Institute (RPMI), supplemented with 10% Fetal Bovine Serum (FBS) and 1% of antibiotics,

Penicillin-Streptomycin. Once the cells reached 80% confluency, RNA extraction was done.

There were several reasons of choosing the NP69 cell line. The cell line was verified to originate from human and was tested negative for contamination of inter-species from mouse, rat, Golden Syrian hamster, and non-human primate.

The reason for choosing HK1 cell line as it is a well-differentiated squamous carcinoma and viral particles of EBNA has not been demonstrated in the cells from the primary culture or other subcultures tested.

3.2.2 Total RNA Extraction

In this research, the GENEzol method was used to extract total RNA from the NP69 and HK1. The growth media inside the culture flask was removed by using pipette and the cell monolayer was rinsed with 1 ml of ice-cold Phosphate Buffer Saline (PBS). Then, 1 ml of GENEzol reagent was added to the T25 flask, and the cell was lysed for 2 minutes using a cell scraper. The mechanical lysis was carried out by pipetting up and down the cell a few times and the cell will be incubated for 5 minutes at room temperature. Next, the GENEzol mixture was transferred to a 1.5 ml sterile microcentrifuge tube, followed by 0.2 ml of chloroform pipetted in the solution to separate the solution into the aqueous, interphase, and organic phases. Then, the cell solution was vortexed vigorously for 12 seconds and was left to incubate for 5 minutes at room temperature.

Subsequently, the sample was centrifuged at 10,000 rpm for 15 minutes at 4℃. Three layers were shown after centrifuge which indicated aqueous layer (RNA), interphase (DNA), and organic phase (protein). The aqueous layer was transferred carefully, avoiding the pipette tips from touching the interphase, into a new 1.5 ml microcentrifuge tube placed on an ice tray. 0.5 ml of isopropanol was added and mixed well. The aqueous layer was centrifuged again for 10 minutes at 4℃ at 10,000 rpm to obtain a RNA pellet at the bottom