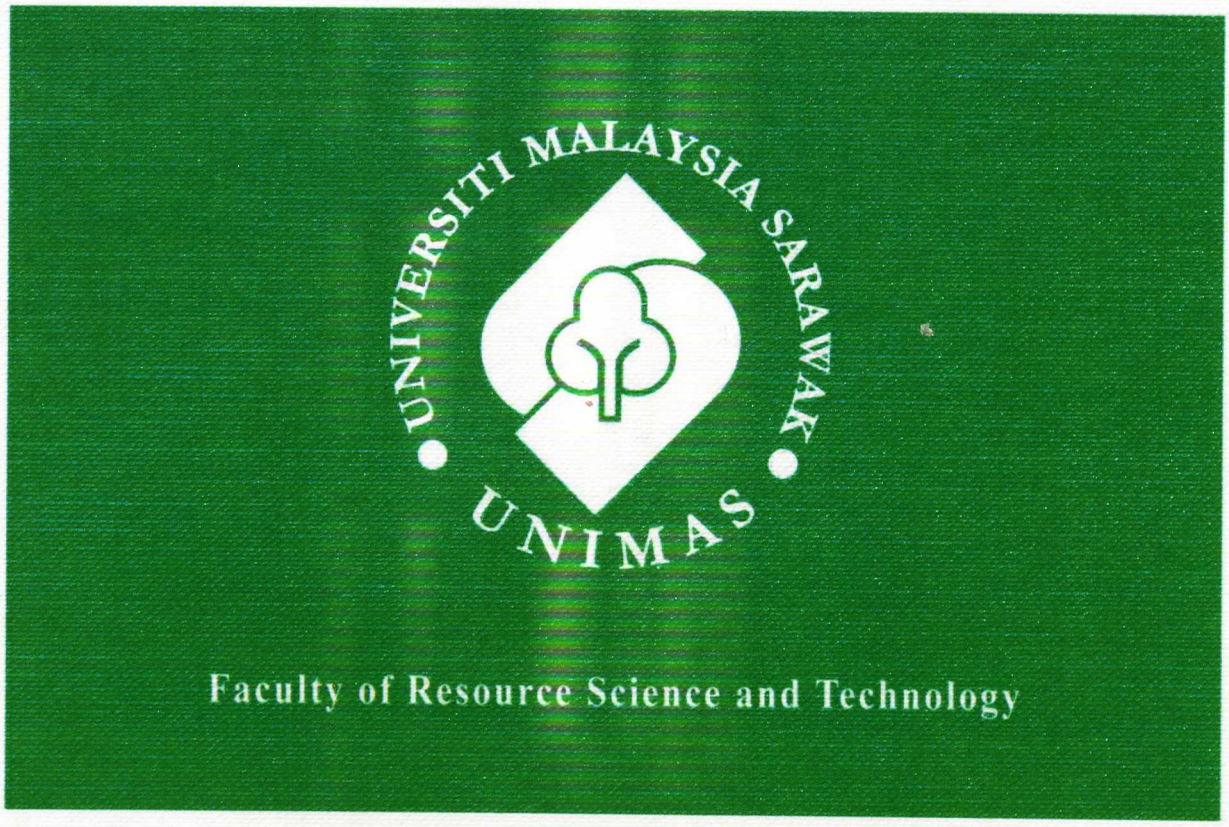

Identification of protein partners of the human ribosomal protein, eS31 using Bioinformatic Analysis

Leong Kar Xin 54693

Bachelor of Science with Honours
Resource Biotechnology
2018

Identification of protein partners of the human ribosomal protein, eS31 using Bioinformatic Analysis

LEONG KAR XIN

A thesis is submitted in partial fulfilment for the requirement for the degree of Bachelor Science with Honours (Resource Biotechnology)

Supervisor:Associate professor Dr Edmund Sim Ui-Hang

Resource Biotechnology

Faculty of Resource Science and Technology
UNIVERSITI MALAYSIA SARAWAK
2018

UNIVERSITI MALAYSIA SARAWAK

Grade: _____

Please tick (√)

Final Year Project Report ☑

Masters ☐

PhD ☐

DECLARATION OF ORIGINAL WORK

This declaration is made on the 9 day of JUNE year 2018

Student's Declaration:

I **LEONG KAR XIN(54693), FACULTY OF RESOURCE SCIENCE AND TECHNOLOGY(FRST)** (PLEASE INDICATE NAME, MATRIC NO. AND FACULTY) hereby declare that the work entitled,

Identification of protein partners of the human ribosomal protein, eS31 using Bioinformatic Analysis

is my original work. I have not copied from any other students' work or from any other sources with the exception where due reference or acknowledgement is made explicitly in the text, nor has any part of the work been written for me by another person.

9 JUNE 2018
Date submitted

LEONG KAR XIN (54693)
Name of the student (Matric No.)

Supervisor's Declaration:

I, ASSOCIATE PROF DR EDMUND SIM UI HANG (SUPERVISOR'S NAME), hereby certify that the work entitled, Identification of protein partners of es31 using Bioinformatic (TITLE) was prepared by the aforementioned or above mentioned student, and was submitted to the "FACULTY" as a * partial/full fulfillment for the conferment of (PLEASE INDICATE THE DEGREE TITLE), and the aforementioned work, to the best of my knowledge, is the said student's work

Received for examination by: _____ Date: 5 June 2018
(Name of the supervisor)

Prof. Madya Dr. Edmund Sim Ui Hang
Jabatan Biologi Molekul
Fakulti Sains dan Teknologi Sumber
UNIVERSITI MALAYSIA SARAWAK

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

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

I declare this Project/Thesis is to be submitted to the Centre for Academic Information Services and uploaded into UNIMAS Institutional Repository (UNIMAS IR) (Please tick (√)):

☐ **YES**
☐ **NO**

Validation of Project/Thesis

I hereby 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:

0 This Project/Thesis is the sole legal property of Universiti Malaysia Sarawak (UNIMAS).
0 The Centre for Academic Information Services has the lawful right to make copies of the Project/Thesis for academic and research purposes only and not for other purposes.
0 The Centre for Academic Information Services has the lawful right to digitize the content to be uploaded into Local Content Database.
0 The Centre for Academic Information Services has the lawful right to make copies of the Project/Thesis if required for use by other parties for academic purposes or by other Higher Learning Institutes.
0 No dispute or any claim shall arise from the student himself / herself neither a third party on this Project/Thesis once it becomes the sole property of UNIMAS.
0 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 himself/herself without first obtaining approval from UNIMAS.

Student's signature _____ (Date) 9/6/2018

Supervisor's signature: _____ (Date) 5/6/18
Prof. Madya Dr. Edmund Sim Ui Hang
Jabatan Biologi Molekul
Fakulti Sains dan Teknologi Sumber
UNIVERSITI MALAYSIA SARAWAK

Current Address:
Universiti Malaysia Sarawak,
Fakulti Sains dan Teknologi Sumber, 94300 Kota Samarahan, Sarawak.

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

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

ACKNOWLEDGEMENT

I would like to thank to God for walking me through the entire semesters and granted me strength and sense of determination to complete my final year project. A profound thanks to my project supervisor, Associate Professor Dr Edmund Sim Ui Hang for his invaluable guidance and advice were given throughout the execution of the project. I am indeed very grateful to learn new knowledge from him as he has guided me patiently whenever I faced any problems. The clarification and guidance given from him allowed me to understand my mistakes so that I would be able to improve from the mistakes that I have made.

Next, I would like to express my sincere gratitude towards all post graduates in Immunological Human Molecular Genetics Laboratory, Faculty of Resources Science and Technology(FRST), especially Yew Keh Li and Sim Chor Chien for their patience in guiding me and pointed out mistakes during the lab training, so that I could learn from my mistakes.

Besides that, my utmost gratitude to my beloved family members, especially to my parents, Leong Wai Chuen and Wee Kam Moi, who have been giving constant moral support and motivation during the completion of this project. I would also like to express my sincere thanks to all of my course mates and lab mates, especially 2 of the Final Year Project students who were working on Bioinformatic together with me, Daphne Olivia Jawai Anak Sadai and Anis Sabrina Binti Jubaidi for providing invaluable guidance and knowledge sharing throughout the completion of the project. This project would be difficult without knowledge sharing from both of them. Last but not least, a virtual thanks to online sources such as Google, Wikipedia, Science Direct, Google Scholar etc for the accessible information in allowing to complete my final year project.

Identification of protein partners of the human ribosomal protein, eS31 using Bioinformatic Analysis

Leong Kar Xin

Resource Biotechnology
Programme Faculty of Science and
Technology Universiti Malaysia
Sarawak

ABSTRACT

Ribosomal protein(RPs) are one of the most important component found in ribosome that play an essential roles in ribosome biogenesis and includes some extraribosomal functions. Apart from these functions, RPs also often associated with cancer and disease development. Previous study indicated that there is low expression of eS31 in NPC cells is unrelated to p53 expression pattern. However, the protein partners and its function resulting in this tumor-suppressing properties and the function of other protein-protein interacting partners still remains unclear. An approach known as Bioinformatic analysis is used to derive the possible protein interacting partners based on computational approach data analysis. This method can be divided into 4 basic steps, which are identification of target and template sequences, three-dimensional protein modelling, identification of structural neighbours and candidate partners and molecular docking analysis. By using this approach, eS31 3D structural model has constructed that have allowed the prediction of 110 structural neighbors. From these, there are 5 of them which are from human origin were selected for further analysis based on logical prospect of binary protein-protein interactions. Further analysis revealed that there are 3 candidate docking partners. Thus, this research has provided theoretical evidence for the interactions between eS31 and other protein partners. Then, it has been predicted that there are 3 potential protein partners that might play distinct roles and functions from its interaction with eS31.

Key words: Ribosomal protein, ribosome biogenesis, eS31, protein modelling, prediction of protein-protein interactions .

ABSTRAK

Protein ribosom (RPs) merupakan salah satu komponent yang penting dalam ribosom dengan memainkan peranan yang penting dalam biogenesis ribosom dan merangkumi sesetenegah fungsi extraribosomal. Selain daripada fungsi-fungsi tersebut, RPs juga sentiasa dikaitkan dengan kanser dan perkembangan penyakit. Kajian terdahulu menyatakan bahawa eS31 mempunyai ungkapan yang rendah adalah tidak berkaitan dengan coral ekspresi p53. Namun demikian, rakan kongsi proteins dan fungsi berkaitan dengan sifat penindasan tumor dan fungsi dalam rakan protein berinteraksi protein yang lain masih kurang jelas. Satu pendekatan dikenali sebagai analisis Bioinformatik digunakan untuk memperolehi rakan protein yang berkemungkinan berinteraksi melalui analisis data pendekatan komputasi. Cara ini boleh dibahagikan kepada 4, iaitu pengenalan urutan dan susunan templat, pemodelan protein tiga dimensi, pengenalpastian jiran struktur dan rakan kongsi calon dan analisis docking molekul. Melalui pendekatan ini, model struktur 3D eS31 telah membina yang megizinkan ramalan 110 jiran struktur. Daripada yang ini, terdapat 5 dari mereka yang berasal dari manusia dipilih untuk analisis selanjutnya berdasarkan prospek logik interaksi protein-protein binari. Analisis lebih lanjut mendedahkan bahawa terdapat 3 calon pasangan dok. Oleh itu, penyelidikan ini telah memberikan keterangan teoretis untuk interaksi antara eS31 dan rakan protein lain. Kemudian telah diramalkan bahawa terdapat 3 rakan protein potensial yang mungkin memainkan peranan dan fungsi yang berbeza dari interaksinya dengan eS31.

***Kata kunci**: Induksi kalus, proliferasi pucuk, eksplan daun.*

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ATF4	Activating transcription factor 4
RPs	Ribosomal proteins
ERAD	Endoplasmic reticulum-associated degradation
NMR	Nuclear magnetic resonance spectroscopy
VAST	Vector Alignment Sequence Tool
MMDB	Molecular Modeling Database
PDB	Protein Data Bank
NCBI	National Center for Biotechnology Information
SSEs	Secondary structure elements
PPI	Protein-protein interaction
RMSD	Root-mean-square deviation
FFT	Fast Fourier Transform
PSI-BLAST	Position-Specific Iterative Basic Local Alignment Search Tool
PSSM	Position-specific scoring matrix
RPG	Ribosomal Protein Gene
QMEAN	Qualitative Model Energy Analysis

RNA	Ribonucleic acid
NF-κB	nuclear-factor kappa-B
SAPK	Stress- activated protein kinase

1.0 INTRODUCTION

Ribosomal proteins are one of the essential components involved in ribosome biosynthesis. In mammalian cells, the biogenesis of cytoplasmic ribosomes requires the assembly of 4 RNA molecules and 79 different proteins. Apart from ribosome biogenesis, ribosomal proteins have distinct functions varied from its types. It plays an essential role during the assembly of ribosome particles and the regulation of ribosome biogenesis. Any steps that perturbed the process of ribosome biogenesis would induce ribosomal stress. Other than ribosome biogenesis, it is also found that ribosomal proteins have other extraribosomal functions such as DNA repair and translation (Wool, 1996). Other than that, previous studies also indicated that human ribosomal proteins also have some implications for the human disease, for example, RPS4 and RPL6 that was predicted to be the candidate genes for Turner (Fisher et al. 1990) and Noonan syndromes.

eS31 or RPS27a is a component of the 40S subunit of the ribosome and belongs to the S27AE family of ribosomal proteins. According to Sim et al(2017), low expression of eS31 in NPC cells is unrelated to p53 expression pattern. From this finding, the expression of eS31 can be associated with its tumor-suppressing properties, but however, the protein partners and its function resulting in this tumor-suppressing properties and the function of other protein-protein interacting partners still remains unclear. Research about the interactions of human ribosomal protein, eS31 with other proteins is less likely available compared to other ribosomal protein genes. In order to gain a pertinent insight into its role in molecular interactions with other molecules, bioinformatics approach is another alternative for structural studies using computational software. The derivation of the functions of a protein and its interactions with other protein partners can therefore be

obtained based on the computational data analysis of the targeted protein sequence, structure and association with other protein partners.

Firstly, the construction of three-dimensional (3D) structural model which will be used as the templates for searching structural neighbors. This structure gives useful insights into the molecular biology as well as in structural biology whereby we can predict the protein function through the constructed protein structure. Bioinformatic algorithm accountable for the structural similarity of the 3D-protein structure through the arrangement of vectors in 3-D space will also be used to identify the significant similarities between substructures in the protein folds. Identification of structural neighbors and candidate partners are carried out after qualitative evaluation of the constructed 3D protein structure. This method is used to determine the function of the protein partners resulted from the interactions between candidate partners with the structural neighbor of the eS31. Molecular docking approach will also be used as the prediction of the function of the interactions of identified protein partners with constructed protein model. This involves the computational generation of structural complexes by fitting two or more reliable 3D protein models by considering the criteria such as surface complementarity and electrostatistics. Herein, this project provides a theoretical insight into the partners that are involved in the interactions with eS31 and its function resulting from this interaction.

2.0 LITERATURE REVIEW

2.1 Ribosomal proteins

Ribosomal proteins (RPs) are proteins that made up of ribosomes which are found in eukaryotic organism. They make up the ribosomal subunits with the association of rRNAs which are synthesized in the nucleolus which then contributed to the stabilization of both large 60S and small 40S ribosomal subunits (Doudna and Rath, 2002). Thus, ribosome consists of four ribosomal RNA (rRNA) species and 79 RPs in a mammal. Unlike rRNAs, which are encoded by up to hundred copies of genes, human RP is only encoded by a single gene. Thus, the sequence functional genes generate a huge amount of processed pseudogenes, which may then restricted the cloning of functional genes. Defective in these RP genes also resulted in post-translational modifications and rare neurological disorders that are not associated with cancer such as Coffin-Lowry syndrome. Since ribosomal proteins are part of the component of ribosomes, thus its function is usually correlated together which is essential for cellular ribosome-mediated protein synthesis for example ribosome biogenesis and protein translation (Maki et al, 2002).

Apart from that, RPs were believed to possess ribosome independent function in tumorigenesis, immune signaling, and development (Zhou et al, 2015). Studies also indicated that interference of any steps or mutation involved in the ribosome biogenesis by extracellular or intracellular stimulations results in ribosomal stress (Zhang & Lu, 2009; Zhou et al, 2012) Ribosomal stress is caused by the accumulation of free-ribosome RPs in the nucleolus and prevent itself from tight regulation to produce ribosomal subunits. These ribosomal stress causes RPs release from the nucleolus and induced many biological changes such as p53 activation and also p53-independent events (Xu et al, 2016). They are

believed to be caused by chemical agents or radiation that disturbed rRNA production or bring about RP degradation, lack of nutrients and gene deregulation.

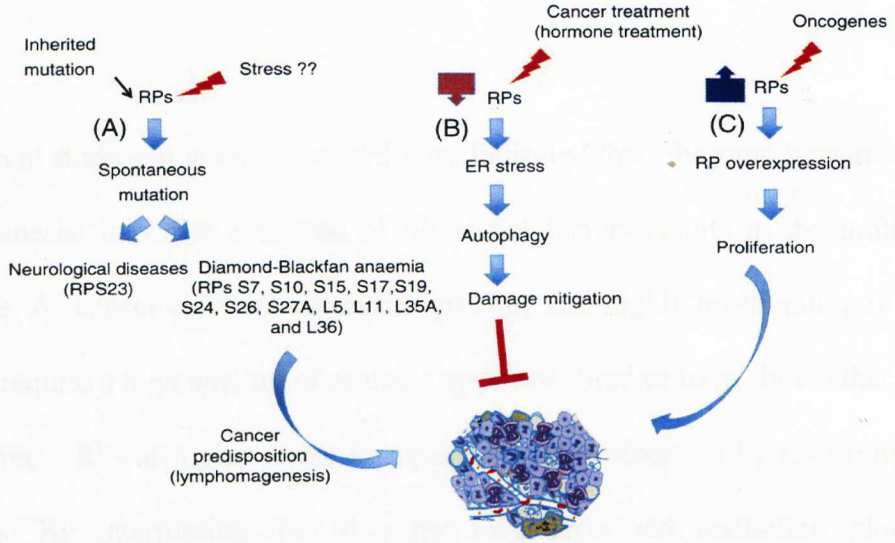

Figure 1: Effects of disruption of RPs by inherited mutations or stress factors. Ribosomopathies(A) are caused by mutations in RPs. Treatments that can lower the level of RPs (B). Regulation of RPs by oncogenes(C) that caused overexpression of RP promotes cancer growth. Source: Lorente, J., Velandia, C., Leal, J. A., Garcia-Mayea, Y., Lyakhovich, A., Kondoh, H., & LLeonart, M. E. (2018). The interplay between autophagy and tumorigenesis: exploiting autophagy as a means of anticancer therapy. Biological Reviews, 93(1), 152-165.

2.2 Ribosomal biogenesis and tumorigenesis

Studies related with ribosome biogenesis have been studied extensively in bacteria such as *Escherichia Coli* (Stokes & Brown, 2015) and yeast *Saccharomyces cerevisiae* (Woolford & Baserga, 2013), however, there are fewer studies conducted in mammals. Ribosome biogenesis is defined as the process in which rRNA transcripts and RPs are produced and associated with functional ribosomes. Ribosome biogenesis begins in the nucleolus of the cell and requires 4 different species of rRNAs, 80RPs and 3 RNA polymerase (Pol I, Pol II and Pol III). This complex and tightly regulated process involved 4 steps which are the coordinated transcription, modifications, and folding of rRNA transcripts; translation, modification, and folding of RPs; binding of RPs to the appropriate

rRNA scaffolds and lastly binding of and release of ribosome assembly factors. Ribosome biogenesis and protein translation are well coordinated rRNAs which they are important for animal development, cellular growth, proliferation, and differentiation (Zhou et al, 2015).

Recent studies in mammalian cell lines indicated that ribosome biogenesis is linked to tumorigenesis, in which mutation of ribosomal factors results in the proliferation of cancer cells. As cancer cells are constantly growing and highly proliferating in the human body, thus require a high amount of proteins synthesis. Studies have shown that some other ribosome-free RPs also play a role in suppressing tumorigenesis by activating the tumor suppressors. By interrupting the ribosome biogenesis and restricting global protein synthesis, several tumor suppressors inhibited the growth and proliferation of tumor cell (Ruggero & Pandolfi, 2003; Silvera et al., 2010). Some examples of RPs involved are RPS15 and RPS 20.

RPs role in tumorigenesis was disturbed when ribosome biogenesis was affected or through extraribosomal functions. They regulate tumorigenesis by disturbing the cells in different manners via p53-dependent and independent pathways shown in Figure 1 and Figure 2. As shown in Figure 1 for the p53-dependent pathway, RPs are released from nucleus during nucleolar stress.They bound to E3 ubiquitin ligase MDM2 and inhibited its ligase activity against p53, which was then caused the accumulation of p53 and suppressed tumorigenesis. One of the example in p53- independent pathway is through the regulation of ATF4, which is a transcription factor that overexpressed in tumor and the major coordinator of tumor cell survival in stress. Reports indicated that RPL41 induced proteasomal degradation of ATF4 by promoting the phosphorylation and translocation of ATF4 from nucleus to cytoplasm which then triggered cell death (Xu et al, 2016).

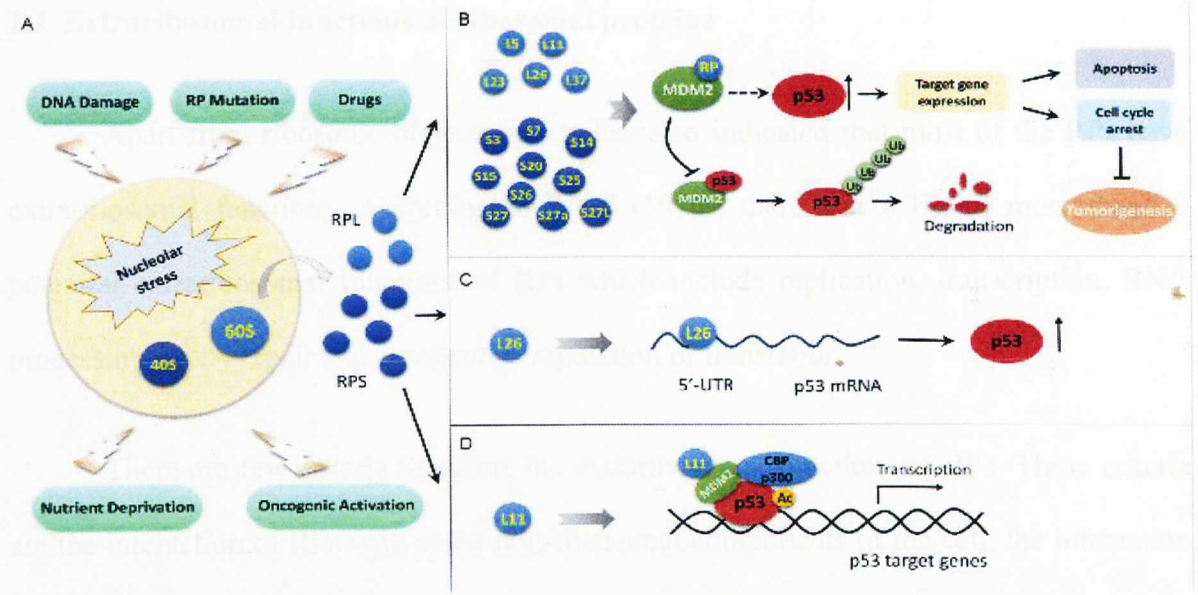

Figure 2: Different manners of p53- dependent carried out by RPs to regulate tumorigenesis. Under nucleolar stress in the cell, p53 expression can be regulated in the cell by binding with RPs. Adapted from Xu, X., Xiong, X., & Sun, Y. (2016). The role of ribosomal proteins in the regulation of cell proliferation, tumorigenesis, and genomic integrity. *Science China Life Sciences*, *59*(7), 656-672.

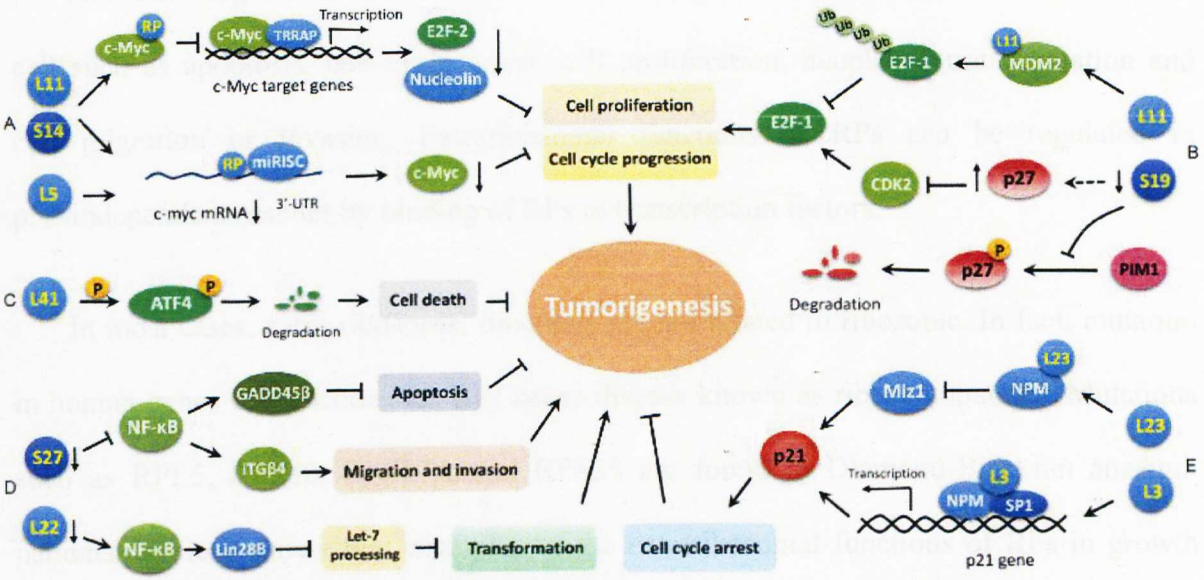

Figure 3: Different manners of p53- independent carried out by RPs to regulate tumorigenesis. RPs regulate tumorigenesis through the activation of various transcription factors. Adapted from Xu, X., Xiong, X., & Sun, Y. (2016). The role of ribosomal proteins in the regulation of cell proliferation, tumorigenesis, and genomic integrity. *Science China Life Sciences*, *59*(7), 656-672.

2.3 Extraribosomal functions of ribosomal proteins

Apart from ribosome biogenesis, studies also indicated that most of the RPs have extraribosomal functions. According to Wool (1996), there was a list of more than 30 potential extraribosomal functions of RPs which include replication, transcription, RNA processing, DNA repair and autogenous regulation of translation.

There are few criteria to define the extraribosomal functions of RPs. These criteria are the interaction of RPs with some non-ribosomal components of the cell, the interaction has physiological effects on cellular functions and these effects occur apart from the ribosome and it will not affect or influence the protein synthesis (Warner & McIntosh, 2009).

Abnormal regulation of the extraribosomal function RPs leads to several effects on the cell such as apoptosis, cell cycle arrest, cell proliferation, neoplastic transformation and cell migration or invasion. Extraribosomal functions of RPs can be regulated in p53-independent manner by binding of RPs to transcription factors.

In most cases, extra-ribosomal functions are not related to ribosome. In fact, mutation in human genes that encodes for RPs cause disease known as ribosomopathies. Mutations such as RPL5, RPS10, RPS27A and RPS28 are found in Diamond-Blackfan anaemia patients. Table 1 shows few examples of the extraribosomal functions of RPs in growth regulation and tumorigenesis (Xu et al, 2016).

Table 1: Some known examples of extraribosomal functions of RPs in growth regulation and tumorigenesis

Extraribosomal functions	RPs
Apoptosis	S3, S3a, S6, S7, L3, L6, L7
Cell proliferation	S6, S9, S14, S24, S27, L11
Cell cycle	S20, S25, S26, S27a, L23, L26
Neoplastic transformation	P1, S3a, S14, L5, L22, L41
Cell migration and invasion	S3, S15a, S24, S24, S27,L15

(Adapted from Xu et al, 2016)

2.4 Ribosomal Proteins in Cancer

The role of ribosomal proteins is often correlated with cancer as the constant growth and proliferating cancer cells required a high amount of protein for protein synthesis purpose. Studies also indicated that some RPs have been found to be up-regulated and down-regulated in human cancer tissues in adjacent to the comparison with the normal tissues. Some examples of the dysregulation of RPs in human cancers are indicated in Table 2.

A typical example of the association of ribosomal proteins with carcinogenesis can be determined through Diamond-Blackfan anemia (DBA). DBA is a cancer susceptibility syndrome in which the mutation of RPS19 which resulted in the loss of RPs function and eventually leads to the development of cancer. Apart from that, study related the association of RPs with carcinogenesis of NPC was also carried out to determine the expression profiles of RPs in NPC-derived cell lines. According to Sim et al.(2017), 4 RP genes, namely uS8, uS4, eS31, and uL14 were reported to be essentially underexpressed in

NPC cell lines in response to nonmalignant nasopharyngeal epithelial cells, which then strengthen the evidence of the association of RP in cancer.

Even though study indicated that ribosome-free RPs can act as oncoproteins, however the differential functions of RPs on coordinating finely with ribosome biogenesis and protein synthesis during normal cell growth and proliferation are still unclear and need further investigation (Zhou et al., 2015).

Table 2: Some known examples of dysregulation of RPs in human cancers

Cancer type	RPs	Alteration	Effects to the disease
Prostate cancer	S2, S14, S15a, S27, L7a, L19, L23a, L31, L37	Upregulation	L19 expression is correlated with Gleason scores of patients with prostate cancer. Increased L19 expression is high;y predictive of shotter patient survival.
Gastric cancer	S13, S27, L6, L13, L15, L23	Upregulation	Increased L6 expression in human gastric tissues is associated with poor prognosis.
Colorectal cancer	S3, S11, S19, S27, S27a, L7, L10a, L13, L19, L29, L36a(L44)	Upregulation	HIgh level of 27L in faeces or cancer tissues is associated with a better prognosis of CRC patients.
	Sa, S8, S12, S18, S24, S27L, L13a, L18, L28, L32, L35a	Downregulation	

Liver cancer	P-S6, S27, S15a, L34	Upregulation	The expression of p-S6 is a negative independent predictor of metastasis-free survival after adjustment for tumor stage.
	L22	Downregulation	
Melanoma	S3	Upregulation	Overexpression of S3 predicts poor prognosis of melanoma patients.
Pancreatic cancer	L15	Downregulation	Low expression of L15 is associated with tumor progression and presitc poor overall survival.

(Adapted from Xu et al, 2016)

2.5 eS31

eS31, also known as RPS27A, is one of the components of the 40S subunit of the human ribosome. It is an 80-amino acid ubiquitin C-terminal extension protein (CEP80). It is a highly conserved protein that has a significant role in aiming cellular proteins for the degradation by 26S proteasome (NCBI, 2017). Ubiquitin is a highly conserved amino acid protein synthesized in the cell by proteolysis of larger proteins containing either polyubiquitin chains or ubiquitin fused to CEPs. In human, ubiquitin-CEP gene that code for ubiquitin fused to eS31, eS31 gene encodes a fusion protein including ubiquitin at N-terminus and at the C-terminus. Pseudogenes originated from eS31 gene are present in the human genome. Similarly, with RPL40 and RPS30, they are all synthesized as a fusion protein with a ubiquitin-like protein(Kirschner & Stratakis, 2000).

eS31 also contains C4-type zinc fingers domain and is located in the cytoplasm. These zinc finger domains are generally a small protein motif that has finger-like structure that has various binding specificities by producing tandem contacts with targeted molecules. Zinc-binding motifs are very stable structures. Example of C4-type zinc fingers is found in the estrogen receptor which is shown in figure 4. In the figure, two estrogen receptors form a dimer by binding with two zinc ions which are represented by orange balls. Most steroid hormone receptors contain such motif (MoBio ,n.d.).

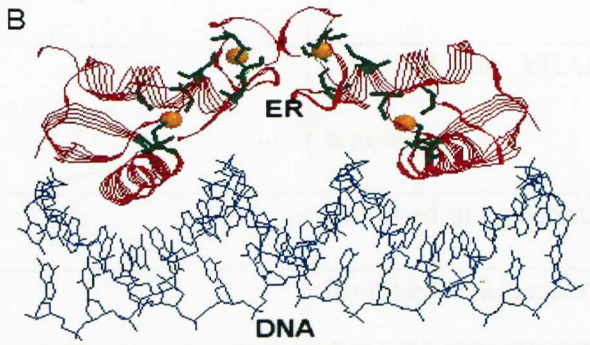

Figure 4: The complex of the estrogen receptor's zinc finger domain and DNA. Source: MoBio (n.d.). Zinc Finger. Retrieved from http://www.web-books.com/MoBio/Free/Ch4F2.htm

Overexpression of eS31 is related to certain cancer cells such as colorectal, breast and renal cancers and leukemia. This occurred during ribosomal stress whereby overexpressed eS31 and the translated regulate the activation of p53 through MDM2 protein binding to inhibit MDM2-mediated p53 ubiquitination. According to Sim et al.(2017), underexpression of eS31 in NPC cells unrelated to p53 expression pattern.

According to UniProt, there were 3 binary interactions identified in previous study, which are the inetractions with Ubiquilin 1(UBQL1),Deleted In Azoospermia Associated Protein 2 (DAZP2) and heterogeneous human epithelial colorectal adenocarcinoma cells (CACO2). When eS31 is covalently bound to another protein, it is temporarily united to the target proteins through isopeptide bond either as a monoubiquitin or polyubiquitin,