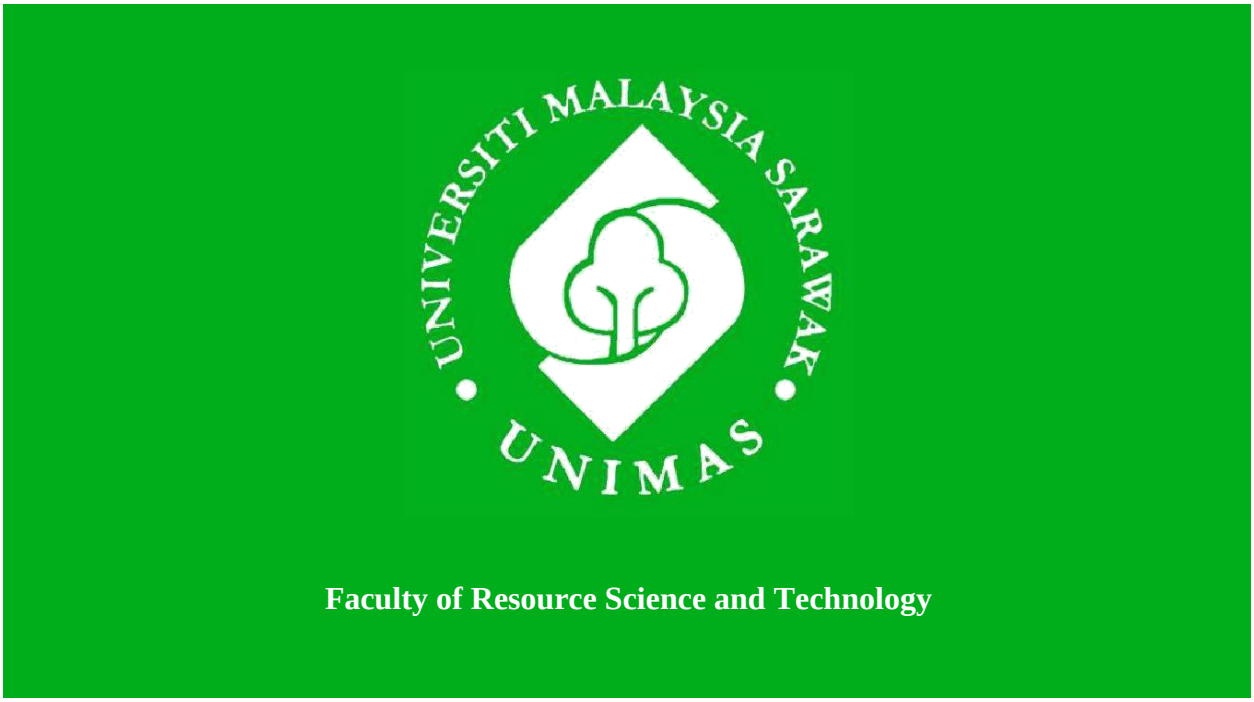

Sequence Analysis of Ribosomal Protein Gene, L14 in Human Colorectal Carcinoma Cell Line.

Najian Binti Ibrahim (37423)

Bachelor of Science with Honours

(Resource Biotechnology)

2015

Sequence Analysis of Ribosomal Protein Gene, L14 in Human Colorectal Carcinoma Cell Line.

Najian Binti Ibrahim (37423)

A Thesis Submitted In Partial Fulfillment of the Requirement of

The Degree of Bachelor of Science with Honours (Resource Biotechnology)

Supervisor: Associate Professor Dr.Edmund Sim Ui Hang

Resource Biotechnology
Department of Molecular Biology

Faculty of Resource Science and Technology
University Malaysia Sarawak
2015

Acknowledgement

Firstly, I humbly thanked God who granted me with physical and emotional strength to complete this project to a success. Next, I would like to express my highest appreciation to my supervisor, Associate Professor Dr. Edmund Sim Ui Hang who has been guiding me by sharing his brilliant ideas and thoughts that of course helped me a lot in executing this research.

Big thanks goes to the post graduate students of the Immunological Human Molecular Genetics Laboratory, UNIMAS. Particularly, Ms. Stella Chan Li Li, Ms. Kherlee Ng, and Ms. Shruti Talwar who have been consistently guiding me by sharing their invaluable experiences and generously helped and discussed my project even with their packed schedule in order to complete my final year project successfully.

In addition, I would also like to give credits to my lab mates, namely Athma Hafeeza, Suariati Pabilah, Yew Keh Li, Cassandra Chee, Lisha Vasudevan, Jaiyogesh Ramesh Patel, and Nazatul Syahira Roslan for their endless support and thoughts that I have put into use in my project.

Last but certainly not the least, I would like to thanked both of my parents, Prof. Dr. Ibrahim Mamat and Mrs. Rohayati Ujud for always being there in terms of providing me with moral and financial support. It is also because of their endless prayers that have led me to finally finishing this project despite all the challenges that were faced.

UNIVERSITI MALAYSIA SARAWAK

Grade:

_____

Please tick (√)

Final Year Project Report ☐

Masters ☐

PhD ☐

DECLARATION OF ORIGINAL WORK

This declaration is made on the ……..day of……..2012.

Student's Declaration:

I ---------- ------- (PLEASE INDICATE STUDENT'S NAME, MATRIC NO. AND FACULTY) hereby declare that the work entitled ---------- -------- 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.

_____ _____

Date submitted Name of the student (Matric No.)

Supervisor's Declaration:

I---------- (SUPERVISOR'S NAME) hereby certifies that the work entitled ----------(TITLE) was prepared by the above named student, and was submitted to the "FACULTY" as a * partial/full fulfillment for the conferment of ---------- (PLEASE INDICATE THE DEGREE), and the aforementioned work, to the best of my knowledge, is the said student's work.

Received for examination by: _____

Date:_____

(Name of the supervisor)

I declare that 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**

Validation of Project/Thesis

I therefore duly affirm with free consent and willingly declare that this said Project/Thesis shall be placed officially in the Centre for Academic Information Services with the abiding 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 digitalise the content 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 itself neither third party on this Project/Thesis once it becomes the 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 signature _____ Supervisor signature:
_____

(Date) (Date)

Current Address:
_____
_____
_____

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 is duly prepared by The Centre for Academic Information Services]

Table of Contents

List of Abbreviations

AGE	=	Agarose Gel Electrophoresis
bZIP	=	basic region-leucine zipper
CRC	=	Colorectal Carcinoma
CSC	=	Cancer stem cells
EtBr	=	Ethidium Bromide
GAP DH	=	Glyceraldehyde-3-Phosphate Dehydrogenase
L, 60S	=	Large subunits
M-MLV	=	Moloney Murine Leukemia Virus Reverse Transcriptase
MEM	=	Minimum Essential Medium
NMT	=	N-myristoyltransferase
PAR	=	Proteins Associated with Ribosome
PKC	=	Protein Kinase C
RPs	=	Ribosomal Proteins
rRNA	=	Ribosomal RNA
RPL	=	Ribosomal Protein Large subunit
RT PCR	=	Reverse Transcriptase Polymerase Chain Reaction
S, 40S	=	Small Subunits
SFM	=	Serum Free Media

List of Tables

List of Figures

Sequence Analysis of Ribosomal Protein Gene, L14 in Human Colorectal Carcinoma Cell Line.

Najian Binti Ibrahim (37423)

Resource Biotechnology
Faculty of Resource Science and Technology
Universiti Malaysia Sarawak

ABSTRACT

Structural information of ribosomal proteins (RPs) has increased the understanding of the structure, function and evolution of the ribosome. RPL 14 is one of the most conserved ribosomal proteins and have a major role in the ribonucleoprotein complex. Previous investigations have revealed that RPs gene expression in colorectal carcinomas (CRC) differs from that was found in normal colorectal cells; some are upregulated and some are down regulated. It has been inferred that the unregulated expression of RPs are associated with the differentiation, progression or metastasis of CRC. This project analyses one of the RP genes, RPL 14 which is taken from colorectal cancer cells. Analysis shown that RPL 14 gene extracted has not undergo any mutation as the nucleotide sequence reflects 99% identity from the sequence in Genbank. Besides that, translation result of RPL 14 illustrated that RPL 14 gene encodes a few biologically important sites such as phosphorylations sites for many protein kinases, myristoylation site, and leucine zipper pattern which makes RPL 14 one of the common housekeeping genes in human.

Keywords: Ribosomal proteins, RPL 14, colorectal cancer, housekeeping genes.

ABSTRAK

Maklumat struktur tentang protein ribosom telah memberikan kefahaman yang lebih mendalam berkenaan tentang struktur, fungsi dan evolusi organel ribosom. RPL 14 adalah salah satu protein ribosom yang kekal dan terlibat secara signifikan dalam penghasilan komplex ribonucleoprotein. Kajian terdahulu telah berjaya mengupas bahawa ada perbezaan di dalam tahap pengeksperesan gen protein ribosom di dalam kanser kolorektal. Oleh sebab ini, inferens telah dicadangkan bahawa ketidakstabilan ekspresi gen protein ribosom ada perkaitan dengan pembezaan, perkembangan atau metastasis kanser kolorektal itu sendiri. Projek ini telah menganalisa salah satu gen protein ribosom iaitu RPL 14 yang diekstrak dari sel kanser kolorektal. Analisa menunjukkan bahawa RPL 14 tidak mengalami sebarang mutasi di mana urutan nukleotida menunjukkan 99% identiti. Selain itu, keputusan translasi RPL 14 menggambarkan gen RPL 14 mengekod untuk beberapa lokasi proses biologi yang penting seperti tapak pengfosforilan untuk protein kinase, tapak mirisotoylasi, dan corak zip leusin di mana sekaligus membuktikan RPL 14 sebagai salah satu gen penjaga dalam manusia.

Kata kunci: *Protein ribosom, RPL 14, kanser kolorektal, gen penjaga.*

1.0 INTRODUCTION

Ribosomes, which serve to synthesize proteins are made of several elements which are ribosomal RNAs (rRNAs), ribosomal proteins (RPs) and accessory factors called protein associated with ribosomes (PAR). Besides this function to combine with rRNA in protein biosynthesis, RPs also fill various roles that are called extraribosomal functions. These functions though are vast, they are still unclear. However, it is widely interpreted that some of these functions include DNA replication, transcription and repair, RNA splicing and modification, cell growth and proliferation, regulation of apoptosis and development, and cellular transformation (Mao-De & Jing, 2007).

It is generally accepted that the rRNA is the principal catalytic component of the ribonucleoprotein complex. Its activity depends upon a precise three-dimensional organization, ribosomal proteins appear to hold on to major role of promoting the folding, and consequently, stabilizes the complex RNA tertiary structure. Thus, this common role of proteins exhibits how specific amino acids mutation can greatly affect the ribosome's functions. Protein mutation that disrupt their interaction with RNA can modulate the local RNA structure and hence its activity (Davies, White & Ramakrishnan, 1995). Recent studies confirmed that RP gene mutations or disturbance in expression levels were found in many inherited genetic diseases such as Diamond-Blackfan anaemia syndrome, Turner syndrome, and Noonan syndrome. Consequently, similar results appeared to be in carcinoma of breast, oesophagus, liver and also colon. Jass (2006) also reported that the development and progression of colorectal cancer (CRC) does involves complex process of multiple genetic changes whereby a series of genetic alterations happens which includes mutations that activate oncogenes and inactivate tumour suppressor genes. To add, the expression of RPs in CRC are commonly seen as up regulated although there are a few studies reporting the otherwise.

A preliminary study done by Sim et al. (2006) found 33 RPs genes to be differentially expressed in CRC cases of Malaysia. Consequently, when these 33 RPs were coordinated with the Cancer Gene Expression Database (CGEP), 22 of them were confirmed to be persistently expressed differentially in CRC. Attention should be given to RPL 14 as it has been listed to be one of the 22 RP genes that are differentially expressed whereby its level of expression was seen to be up-regulated in CRC compared to the normal colorectal sample. Therefore, this project aimed to analyse the nucleotide sequence of RPL 14 gene in CRC as this gene might be a potential diagnostic biomarker of this disease. The two main objectives of this study were:

(a) To analyse the sequence of RPL 14 gene in CRC cell line, HCT 116 and thus compare it with RPL 14 *homo sapiens* sequence that are readily available in Genbank database.

(b) To predict the amino acids compositions as well as the motifs that RPL 14 protein encodes for, so that the importance of RPL 14 protein can be deduced.

RT-PCR method was performed on RNAs of colorectal carcinoma cells to obtain the amplified DNA sequences of RPL14. The DNA sample of RPL 14 was then sent to 1st Base Laboratory Malaysia for DNA sequencing. The RPL 14 DNA sequence was analyse using a few of bioinformatics tools that can be retrieved online such as *BLAST* (Altschul, Gish, Miller, Myers & Lipman, 1990) and *CLUSTAL W* (Thompson, Higgins & Gibson, 1994).

2.0 LITERATURE REVIEW

2.1 Colorectal Carcinoma (CRC)

Colorectal carcinoma (CRC) develops in the colon or the rectum, also known as the large intestine (Figure 2.1.1). According to Cappell (2005), CRC arises from mucosal colonic polyps. The two most common histologic types are hyperplastic and adenomatous.

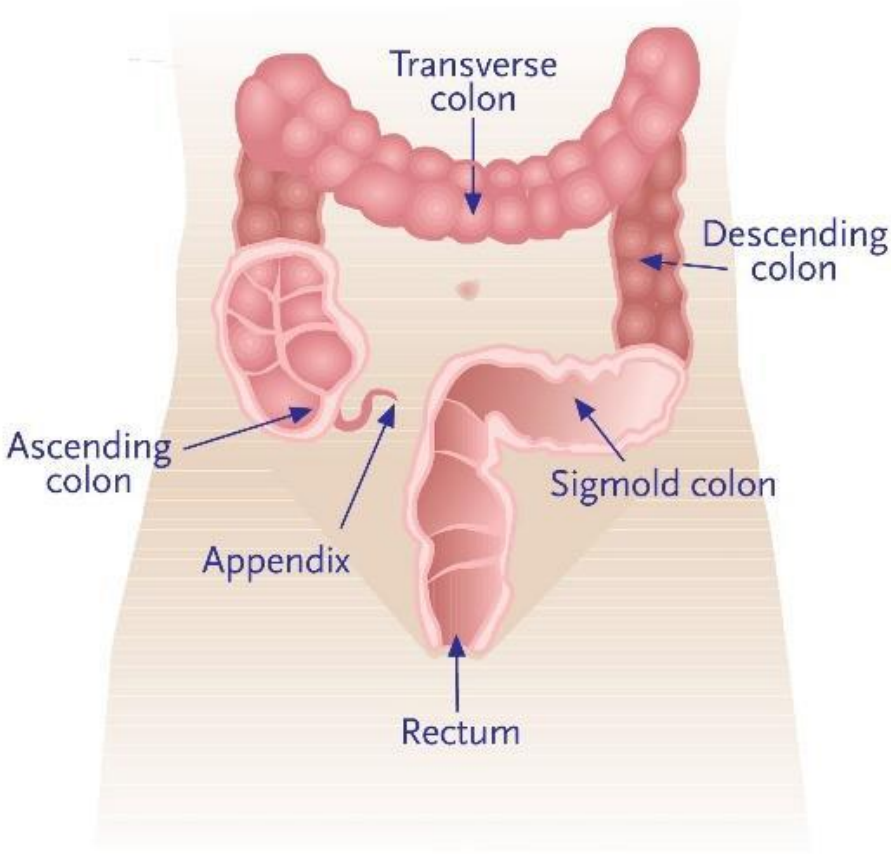

Figure 2.1.1: Human colorectal system.

Retrieved from http://www.newhealthguide.org/Rectum-Function.html

Histologically, hyperplastic polyps contain growing amount of glandular cells with decreased cytoplasmic mucus, but lack nuclear hyperchromatism. On the other hand, adenomatous nuclei are commonly hyperchromatic, enlarged, cigar-shaped, and crowded together in a palisade pattern (Tsai & Lu, 1995). Adenomas can be classified as tubular or villous. However in most cases, colorectal carcinoma generally originates from adenomas

according to screening tests, clinical, and pathologic findings (Cappell, 2005). Once carcinoma forms in the inner lining of the large intestine, it can grow into the wall of the colon or rectum. Carcinoma that has grown into the wall can also invade the blood or lymph vessels resulting to probable metastasizing growth of carcinoma.

Statistically, CRC cases has increased in Malaysia from 8.1% to 12.0% in 1995 (Ministry of Health, 1995). Globally, CRC is the third most fatal cancer and its occurrence has been increasing in Asia, including Malaysia. National Cancer Registry reported 2,246 CRC cases has been registered nationwide, representing 12.3% of all registered cases in 2007. Cappell (2005) suggested that CRC is caused by a series of genetic mutations causing to progressively disordered local DNA replication and enhanced colonocyte replication. The progressive combination of multiple genetic mutations results in the evolution of normal mucosa to benign adenoma to carcinoma. To add, environmental factors could also be an important factor to the carcinoma. Reports by Mao-De and Jing (2007) proposed that disruption in expression level in ribosomal proteins could associate with tumour progression and metastasis. The idea came after RPs were isolated and screened from cDNA libraries of CRC. The mRNAs are seen to be overexpressed and appeared to be more abundant in CRC sample compared to the normal cell lines hence suggesting its role in proliferation and neoplasia of CRC.

2.2 Ribosomal Proteins and RPL 14 gene

Initially, ribosomal proteins (RPs) are recognised for their role of biosynthesizing proteins for any living organism. RPs, together with ribosomal RNAs (rRNAs) will join and bind with each other to make up the ribosome organelle in all cells. As for human ribosomes, it constitutes of 4 different rRNAs species and 80 distinct RPs. There are two main subunits of ribosomes which are the large subunit (L, 60S) and small subunit (S, 40S). Protein associated with ribosomes called PAR is also responsible in ribosomes synthesis and recycling when PAR is combined with the two ribosomal units (Mao-De & Jing, 2006).

In 2001, 80 different RP genes were successfully mapped on the cytogenetic map of human genome (Mao-De & Jing, 2006). While it is agreed that most of RPs involves solely on protein synthesis (RPs stabilize specific rRNA structure in mature ribosomal subunits by supporting accurate folding of rRNAs during ribosomal assembly), attention has been growing to understand the probable extraribosomal functions of RPs. Extraribosomal function of RPs simply means that the RPs are not in any way related to help synthesize protein, instead, they may be involve in other functions such as DNA replication, transcription and repair, RNA splicing and modification, cell growth and proliferation, regulation of apoptosis and development, and cellular transformation. Interestingly, RP genes were constantly identified in of screening the differentiated expressed genes of human diseases especially in genetic diseases and cancers. According to Mao-De and Jing (2006), the disturbances RPs structures or expression levels were associated with various diseases such as Turner syndrome, Noonan syndrome, and Diamond-Blackfan anaemia syndrome. The same results was observed in several kinds of carcinoma including breasts, liver, oesophagus and colorectal itself. RPs can directly regulate the effect of oncogene and tumor suppressors on DNA replication, transcription and translation.

For example, overexpression of RPS3a gene has been shown to promote tumorigenesis in nude mice via suppression of apoptosis by inducing synthesis of anti-apoptosis proteins. These studies indicate that extraribosomal functions of distinct ribosomal proteins may also regulate tumorigenesis.

Consequently, alterations in the synthesis or functioning of ribosomal proteins can lead to various hematologic disorders whereby ribosomal proteins is thought to have some genetic interaction with p53 gene which could explain the association between ribosomal proteins and cancer. Previous report had conveyed that upon receiving cellular stress, RP will bind to MDM2 and block MDM2-mediated p53 ubiquitination and degradation, thus resulting in p53- dependent cell cycle arrest (Narla & Ebert, 2010). Therefore, via this mechanism, we could note that ribosomal proteins plays a big part in connecting deregulated cell growth with inhibition of cell division. Figure shows how the Loss of ribosomal proteins activates p53. In a normal cell, MDM2 binds to p53 in the nucleoplasm, leading to p53 ubiquination and degradation. However, defective ribosome biogenesis leads free ribosomal proteins moving out of the nucleolus and binding MDM2, thereby preventing the degradation of p53 which then causes G1 cell cycle arrest (Narla & Ebert, 2010).

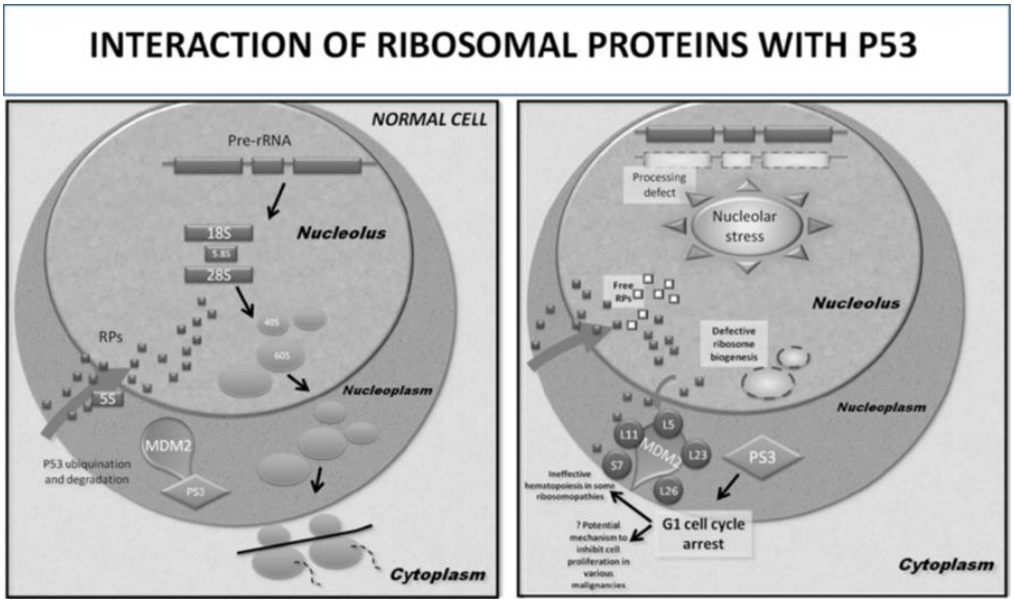

Figure 2.2.1: Interaction of ribosomal proteins with p53

A number of studies have been conducted in investigating the pattern of various RP gene expression in malignant diseases. The mRNA of RPL31 cDNA was selected from a normal colon cDNA library and was found to be overexpressed in all 23 CRC by slot-blot hybridization. In situ hybridization results showed that RPL31 was more abundant in CRC than normal colonic epithelium or stromal tissues thus signifying its role in proliferation and neoplasia (Chester et al., 1989).

The same results were shown by Pogue-Guile et al. (1991) revealing that RPS3, S6, S8, S12, L5, and P0 increased in CRC and adenomatous polyps except for RPL26 and RPL35. However, it cannot be concluded that abundance of the RPs is due to the presence of a higher percentage of dividing cells in tumour as the increment in CRC was not significantly different from the of normal colonic mucosa for every case. Nevertheless, it can still be an important starting indicator to understanding roles of ribosomal proteins in colon neoplasia.

Ribosomal protein L14, or widely addressed by RPL14 gene belongs to the L14E family of ribosomal proteins. RPL 14 encodes for ribosomal proteins from subunit 60S. It contains a basic region-leucine zipper (bZIP)-like domain (Wool, Chun & Gluck, 1996). Figure 2.2.1 shows the 3D morphology structure of RPL 14.

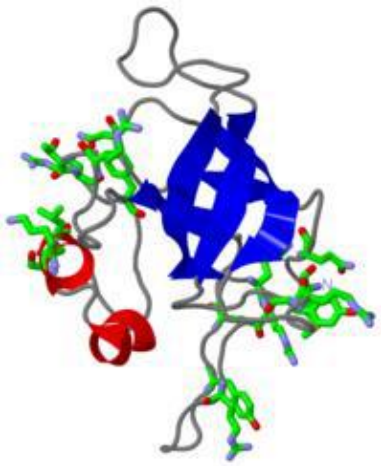

Figure 2.2.2: 3D structure of RPL 14.

Retrieved from http://proteopedia.org/wiki/index.php/Ribosomal_protein_L14

Besides that, Wool et al. (1996) reported that RPL 14 contains a trinucleotide (GCT) repeat tract whose length is highly polymorphic whereby these triplet repeats result in a stretch of alanine residues in the encoded protein. Transcript variants utilizing alternative polyA signals and alternative 5'-terminal exons exist but all encode the same protein. There are multiple processed pseudogenes of this gene dispersed through the genome. Human ribosomal protein gene (RPL14) is located on chromosome 3p21.

2.3 Human Colorectal Cell Lines (HCT 116)

Cancer stem cells (CSCs) are defined as the subpopulation of cells within a tumour that can self-renew, differentiate into multiple lineages, and drive tumour growth (Yeung *et al.*, 2010). This project chooses HCT 116 to study the expression of RPL 14 in human colorectal cell line. HCT 116 is cell line of an adult, male colon ascendens infected with colorectal carcinoma. Figure 2.3.1 shows the morphology of HCT 116 cells.

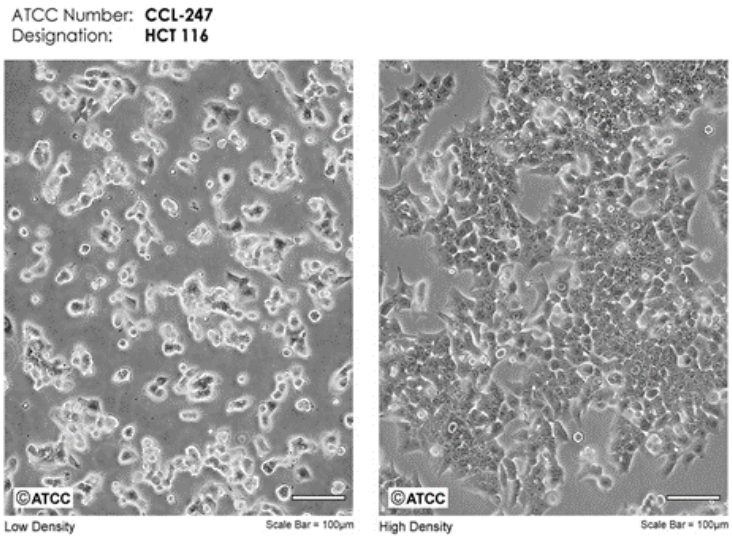

Figure 2.3.1: HCT 116 cells morphology, scale bar 100µm.
Taken from www.atcc.org/Products/Cells_and_Microorganisms/Cell_Lines.aspx

HCT 116 is a highly aggressive cell line that has lost the capacity to differentiate and does not express CDX1 resulting in an aggressive but poorly differentiated tumour. As suggested by Wong et al. (2004), reduced expression of CDX1 in colorectal cancers (by methylation) may have possible association with reduced capacity of the CSCs to differentiate thus contributing to an increased growth rate of undifferentiated tumour. HCT 116 does not significantly separate into different types of colony-forming cells, nor does it contain other subpopulations of cells with higher tumour-forming capacity. This is proven in a study by Yeung *et al.* (2010) whereby HCT 116 is seen to be able to form tumours in NOD/SCID mice that were treated with CD44+CD24+ and CD44−CD24− cells from the HCT 116 cell line. The results showed that the selection does not affect the tumour forming capacity of

HCT116 cells, as there was no significant difference in the tumour sizes between the CD44+CD24+ and CD44−CD24− subpopulations. Thus, HCT 116 is inferred to contain a high proportion of CSCs that have lost its capability to differentiate (Choi *et al.*, 2000).

In summary, colorectal cancer cell lines contain subpopulations of CSCs which are categorized by their cell surface markers and colony morphology. They possess self-renewal capability and could also differentiate into multiple lineages (Ahmed *et al.*, 2013). Figure 2.3.2 shows colorectal carcinoma cell lines which vary in growth rate and morphology. The individual cell cultures are depicted by phase-contrast micrographs 24 hours after trypsinization and seeding. Fast-growing cancer cell lines are indicated with a yellow dot and slower-growing cell lines are indicated by a red dot (Ahmed *et al.*, 2013).

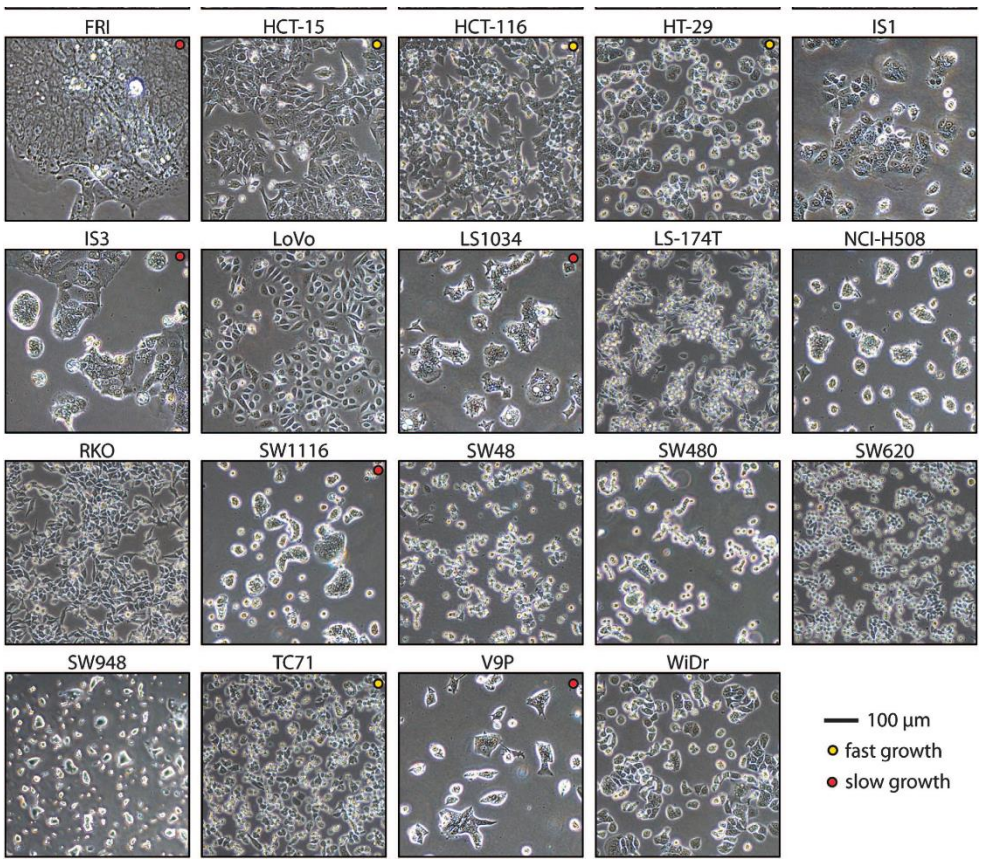

Figure 2.3.2: Different types of CRC cell lines, scale 100mm.
Taken from "Epigenetic and genetic features of 24 colon cancer cell lines" by Ahmed *et al.*, 2013, O*ncogenesis*, 2, p. 3.

2.4 Sequence Analysis

When analysing a gene sequence, normally we would start by comparing nucleic acid sequences. All the four nucleotides, A, T, C, and G, are found in the database with approximately the same frequencies and have roughly the same probability of mutating one into another. Therefore, as a result, DNA-DNA comparisons are largely based on straightforward text matching, which makes them rather slow and less sensitive. There are many methods to increase efficiency in analysing a sequence. One of it is by doing amino acid sequence comparisons as it has several distinct advantages over nucleotide sequence comparisons which makes it a more sensitive method of analysing DNA sequences. First, it is because there are 20 amino acids but only four bases, an amino acid match carries with it >4 bits of information as opposed to only two bits for a nucleotide match (Altschul, Gish, Miller, Myers & Lipman, 1990). Also, the likelihoods of amino acids substitutions occurring during evolution are substantially different, and thus by taking this into account greatly improves the sequence analysis performance. There are two basic alignment method which are global (full-length) and local alignment, which includes only parts of the analyzed sequences (subsequences) (Thompson, Higgins & Gibson, 1994). Of course global alignment is best to describe relationships between sequences across species but however in this study local alignments was used. This study utilized *BLAST* (Altschul, Gish, Miller, Myers & Lipman, 1990) and *CLUSTAL W* (Thompson, Higgins & Gibson, 1994) to make sequences alignments.

3.0 MATERIALS AND METHODS

Table 3.0.1, 3.0.2, and 3.0.3 display the list of materials and apparatus as well as online tools that have been used in this study.

Table 3.0.1: Materials and reagents

Materials and Reagents	Brand	Country of Manufacturer
Agarose Gel	Vivantis	USA
Chloroform-Isoamyl Alcohol 24:1	Spectrum®	USA
DEPC-treated water	Ambion®	USA
Ethidium Bromide	R&M Chemicals	UK
HCT 116 Colorectal carcinoma cell line	ATCC	USA
Ladder mix DNA (0.5µg/µl, 50 µg)	Fermentas	USA
Loading dye	Fermentas	USA
RNase free water	Promega	USA
TRIzol reagent	Invitrogen™	USA
Moloney Murine Leukemia Virus Reverse Transcriptase (M-MLV)	Promega	USA
Pfu DNA Polymerase	Thermo Scientific	USA
Forward and Reverse RPL 14 Primers	Integrated DNA Technologies (IDT)	USA

Table 3.0.2: Tools and apparatus

Tools and Apparatus	Brand	Country of Manufacturer
Express Cool Turbo Freezer	LG	Korea
Incubator	Bibby STUART Scientific S101D	UK

Microwave Oven	National	Japan
Mikro 22R Centrifuge	Hettich Zentrifugen	Germany
pH meter	Orion 2 Star pH Benchtop, Thermo Scientific	USA
Power Pack™ Power Supply 3000	Major Science	USA
RNA Wood Hood	Köttermann	Germany
UV spectrophotometer	Ultrospec 1100 pro	UK
UV transluminator	Wise UV. WUV-M10	UK
Weighing Machine	Setra	USA
T100™ Thermal Cycler	Bio-Rad	USA

Table 3.0.3: Online tools

Online Tools	**Host/Inventor/Website**
NCBI primer designing tools	National Centre for Biotechnology Information (NCBI)/ (http://www.ncbi.nlm.nih.gov/tools/primer-blast/)
Oligoanalyzer 3.1	Integrated DNA Technologies, Inc./ (https://sg.idtdna.com/calc/analyzer)
Oligocalc	Northwestern University/ (Kibbe, 2007)/ (http://www.basic.northwestern.edu/biotools/oligocalc.html)
Basic Local Allignment Search Tool (BLAST)	National Centre for Biotechnology Information (NCBI)/ (Altschul, Gish, Miller, Myers & LIpman, 1990)/ (http://blast.ncbi.nlm.nih.gov/Blast)
CLUSTAL W	The European Bioinformatics Institute (EMBL-EBI)/ (Thompson, Higgins & Gibson, 1994)/ (http://www.ebi.ac.uk/Tools/msa/clustalw2/)
Jalview	The Barton Group/ (Waterhouse, Procter, Martin, Clamp & Barton, 2009)/ (http://www.jalview.org/)
Prosite	Swiss Institute of Bioinformatics Biozentrum, University of Basel (2009)/ (http://prosite.expasy.org/)