Faculty of Resource Science and Technology

Characterization and Expression Studies of Grb10b Genes in Zebrafish.

Nelly Wong Pei Pei (19244)

Bachelor of Science with Honours

(Resource Biotechnology Program)

2010

Characterization and Expression Studies of Grb10b Genes in Zebrafish.

Nelly Wong Pei Pei

A thesis submitted in partial fulfilment of the requirements for the degree of Bachelor of Science with Honours

Resource Biotechnology

Faculty of Resource Science and Technology
UNIVERSITI MALAYSIA SARAWAK
2010

ACKNOWLEDGEMENTS

Table of Contents

List of Table

List of Figures

List of Abbreviations

PH	Pleckstrin Homology
BPS	Between PH and SH2
SH2	Src homology 2
IGF-IR	Insulin-like Growth Factor Receptor
EGF	Epidermal Growth Factor
PDGF	Platelet-Derived Growth Factor
EphB2/ELK	Ephrin receptor
IR	Insulin Receptor
SRS	Silver-Russell syndrome
IRS	Insulin Receptor Substrate
IRK	Tyrosine Kinase Domain of the IR
PI3-K	Phosphatidylinositol 3-kinase
GH	Growth Hormone
DDC	Duplication-degeneration-complementation
GIGYF	Grb10 interacting GYF domain
RT-PCR	Reverse-transcription PCR

Characterization and Expression Studies of Grb10b Gene in Zebrafish

Nelly Wong Pei Pei

Resource Biotechnology
Faculty of Resource Science and Technology
University Malaysia Sarawak

ABSTRACT

Grb10 gene in mammals encodes for the growth factor receptor-bound protein 10 which belongs to a family of adapter protein that was found to interact with a number of receptor tyrosine kinases and signaling molecules. Two copies of Grb10 genes were identified in Zebrafish (*Danio rerio*), which include Grb10a and Grb10b. Characterization and expression studies for Grb10b gene location from the few major organs (eyes, brain, gills, stomach, intestine, liver, heart, skin, muscle, bone) in Zebrafish were carried out by construction of cDNA from its RNA and performing RT-PCR using a pair of primer consisting of forward primer (5'-gACCTgTgCCAgTTTCTCgT-3') and reverse primer (5'-ATTCATACgCTgCTggCTTT-3'). RT-PCR amplification showed three organs with the correct product size, which is approximately 800bp. Thus, the expression study was partially successful and showed that Grb10b is a duplicated gene.

Key words: Grb10, Grb10b, expression, gene duplication.

ABSTRAK

Gen Grb10 tergolong dalam kumpulan protein penyesuai di mana ia mengekodkan "growth factor receptor-bound protein 10" yang dijumpai berinteraksi dengan beberapa reseptor tirosin kinase dan molekul signaling. Dua salinan gen Grb10 telah pun dikenalpasti dalam Zebrafish (Danio rerio), iaitu Grb10a dan Grb10b. Kajian karakterasasi dan lokasi ekspresi gen Grb10b telah pun dijalani pada beberapa sampel-sampel organ Zebrafish melalui pembinaan cDNA daripada RNAnya dan amplifikasi RT-PCR dengen menggunakan sepasang primer yang mengandungi forward primer (5'-gACCTgTgCCAgTTTCTCgT-3') dan reverse primer (5'-ATTCATACgCTgCTggCTTT-3'). Amplifikasi RT-PCR berjaya menunjukkan tiga sampel organ yang mempunyai saiz produk yang tepat, iaitu kira-kira 800bp. Oleh demikian, kajian ekspresi berjaya dengan sebahagian dan gen Grb10b menunjukkan pola-pola duplikasi.

Kata kunci: Grb10, Grb10b, ekspresi, duplikasi gen.

CHAPTER I

Introduction

1.1 Introduction

Growth factor receptor bound protein 10 (Grb10) belongs to adapter protein family, which consists of Grb7, Grb14 and a protein of *Caenorhabditis elegans* called Mig-10. They are categorized together based on the reported high similarity in their sequence homology and molecular structure. The similarity of structure can be observed with a few highly conserved domains in their sequences, which are the src-homology 2 domain (Margolis *et al*., 1992; Liu & Roth, 1995; Ooi *et al*., 1995, cited from Lim *et al*., 2004), the pleckstrin homology (PH) domain (Haslam *et al*., 1993, cited from Lim *et al*., 2004), the proline-rich region(Kay *et al*., 2000), and the recently discovered, region between SH2 and PH region called BPS region (He *et al*., 1997).

In human and mouse, Grb10 protein were identified as isoforms which produced from the alternative splicing of single gene. These isoforms produced multiple mRNAs that transcribed into different Grb10 proteins. Despite the single copy of Grb10 found in mammals, two copies of Grb10 genes were found in zebrafish, denoted as Grb10a and Grb10b (Lee, 2008). Characterization and expression of the gene in zebrafish was studied on the tissues of *Danio rerio* in different stages (Lee, 2008) and the existence of two copies of gene was believed to be resulted from gene duplication.

In human and mouse, maternal uniparental disomy imprints of Grb10 has been associated with Silver-Russell syndrome, when patients were found to have maternal duplication on chromosome 7, where the gene that regulates growth and development is located, supporting that Grb10 is a strong candidate in causing SRS (Monk *et al*., 2000).

However, there are still reviews stating evidence against the relationship (McCann *et al*., 2001). Thus, the exact relation between Grb10 gene and SRS still remains a controversial.

1.2 Objective

The main objective of this project includes:

1. RNA extraction from organ tissues of *Danio rerio*.

2. Construction of cDNA for each tissue's RNA

3. Screening for expression of Grb10b DNA in target tissues by using forward-reverse primers and cDNA constructed.

4. Amplification of Grb10b gene by using Reverse Transcriptase PCR.

5. Expression studies of Grb10b gene in *Danio rerio*

2.1 Introduction of Grb10 gene

Grb10 gene encodes for the growth factor receptor-binding bound protein 10, which belongs to the Grb7/10/14 superfamily of adaptor protein and was characterized by the N-terminal proline-rich region, a central pleckstrin homology (PH) domain, a domain between the PH and SH2 domain, known as PBS region and a C-terminal SH2 domain (Morgolis *et al*., 1992; Liu and Roth, 1995). The encoded adaptor protein has been reported to mediate the coupling of a number of cell surface receptor tyrosine kinase and signaling molecules with specific signaling pathway (Arnaud *et al*., 2003), especially the association of Grb10 with the insulin (Liu and Roth, 1995; Hansen *et al*., 1996) and the insulin-like growth factor-I (IGF-I) (Dey *et al*., 1996; Morrione et al..1996 and Dong *et al*., 1997a,b). Apart from the tyrosine-kinases, Grb10 was also found to interact with non-tyrosine kinases such as Tec and Bcr-Abl. Initially, Grb10 was found as a binding partner in the epidermal growth factor (EGF) receptor mechanism (Ooi *et al*., 1995, cited from Lim *et al*., 2004). However, the functional role of Grb10 in EGF mechanism still remains a controversial.

2.2 Description of *Danio rerio*

Danio rerio or commonly known as Zebrafish is a type of tropical freshwater fish. This fish is commonly seen captivated in aquarium for ornament.

The Zebrafish belongs to the minnow family (Talwar and Jhingran, 1991, cited from Fishbase, 2004), also known as *Cyprinidae*, which consists of the carps, also known

as the true minnows (Talwar and Jhingran, 1991, cited from FishBase, 2004). The fish is native to the streams of the southeastern Himalayan region (Mayden *et al.*, 2007) and was introduced into aquarium in Japan, Canada, United States and Australia.

Apart from being as ornament, Zebrafish was established as a model organism in researches by George Streisinger from the University of Oregon. Its importance is consolidated by the large scale forward genetic screens, commonly known as Boston Screens. An online database of zebrafish genetic, genomic and developmental information – the Zebrafish Information Network (ZFIN) was established.

As a model organism to study vertebrate development and gene function (Mayden *et al.*, 2007), *D. rerio* allows advancement in the fields of developmental biology, oncology (Xiang J. *et al.*, 2008), toxicology (Hill *et al.*, 2005), reproductive studies, teratology, genetics, neurology, environmental sciences, stem cell and regenerative medicine (Major and Poss, 2007) and evolutionary theory (Parichy, 2006).

According to Dahm (2006), *D. rerio* possesses a few advantages to be established as model organism. The greatest advantages are the sequenced genetic code, well understood, easily observable and testable developmental behaviours. The availability of well-characterized mutants and the rapid development of embryos make manipulation and observation works easier. Besides, zebrafish has sufficient genetic properties that it can be the substitution to the higher vertebrate models, such as mice, rat and frog in genetic studies as it has high similarity genetic properties with humans, therefore gene and cell (Eisen and Weston, 1993) manipulation can be done on zebrafish embryos as the development can easily be observed throughout the embryonic developmental stage. Unlike higher vertebrate models, zebrafish embryos develop outside of the mother, which make the subsequent development events easily assessable for studies (Eisen and Weston,

1993). Furthermore, the transparency of the embryo and the progeny allows direct observation on the embryonic cell in living embryos (Eisen and Weston, 1993). The rapid development of the embryo and the ability of the cell to be labeled in living embryos makes it possible to follow the entire development of a same cell (even single cell) to be tracked and observed for characterization, from birth through the expression of its differentiated phenotype (Eisen and Weston, 1993).

Figure 2.1 Zebrafish (Danio rerio)

(Retrieved from http://www.topnews.in/)

2.3 Structure and Function of Grb10 Domains

2.3.1 Proline-rich region

The proline-rich region of the Grb10 superfamily is a highly-conserved region. It is found in all isoforms or splice variants of the family. Furthermore, the second motif presents in human and mouse Grb10 isoforms were found to have 63.6% identity. Due to this reason, there is also similarity in the function of the gene in both mammals. According to Kay *et al*. (2000), the proline-rich region has the property as ligands for the protein-interaction domains. The known interaction are the src-homology domain 3 (SH3) and the WW domains (small globular domain rich in tryptophan). Besides that, *in vitro* interaction between the SH3 domain of c-Abl and Grb10 proline-rich region was found, however, *in vivo* interaction has yet identified. Thus, there was a conflict report which opposed that instead of proline-rich region, SH2 domain is the one interacted with Bcr-Abl. This result was strongly supported by occurrence in both *in vitro* and *in vivo* (Bai *et al*., 1998).

Two novel proteins were later identified to have interaction with the proline-rich region of Grb10 gene are GIGYF1 and GIGYF2 (Giovannone *et al*., 2003). Giovannone *et al*. reported that deletion analysis showed at least two of three proline-rich motif of mGrb10 are required for the binding of both novel proteins and the interaction may be promoted by the IGF-I receptor in mGrb10 with the increase of binding affinity of GIGYF1 to mGrb10δ. In addition, mGrb10 and GIGYFs are suggested to be able to cooperate in regulating receptor signaling because the overexpression of mGrb10 binding region of GYGYF1 have shown to enhance the IGF-I receptor-stimulated tyrosine autophosphorylation (Giovannone *et al*., 2003). So far, only mouse Grb10 have been put to research, therefore human Grb10 interaction with the GIGYFs remain to be investigated. Apart from the protein interaction mechanism, Grb10 proline-rich motif could play a role

in the insulin-mediated metabolic reaction and also in insulin- and IGF-I-induced mitogenic reaction.

2.3.2 PH Domain

The pleckstrin homology (PH) domain was originally identified in the cytoskeletal protein pleckstrin (Haslam *et al*., 1993; Mayer *et al*., 1993) and is various proteins involving in cellular signaling, cytoskeletal organization and regulation of membrane trafficking (Lemmon and Ferguson, 1998; Rebecchi and Scarlata, 1998). The main function of the 3 domain is in the binding of phosphoinositides which are produced when cell surface receptors are activated when PH domain proteins are largely recruited to the plasma membrane (Maffucci and Falasca, 2001). Though the mechanism is found in Grb7, the roles of PH domain in Grb10 remain unidentified. However, hGrb10β, which is lack of PH domain, remains its ability in insulin-stimulated translocation. PH domain is therefore considered as unnecessary for the ability of Grb10 to translocate (Dong *et al*., 1997). Later in another report by Dong *et al*. (1997), the intact of PH domain, mediated with BPS region and SH2 domain, enhances the association of Grb10 with the insulin and IGF-I receptors and this claim is further supported in another review of Dong *et al*. (1998) explaining the presence of PH domain intact produces more favourable confirmation for Grb10 to carry out the interactions. Though PH domain is found in mediating intramolecular interactions, its molecular details remain to be seen.

2.3.3 SH2 Domain

The first identified src-homology (SH) 2 domain was found in oncoprotein Src and Fps. The length of the molecule is about 100 amino acid residues long and it was found regulating the intracellular signaling cascades by interacting with high affinity to phosphotyrosine-containing target peptides in a high sequence-specific and strict phosphorylation-depedent manner (Sadowski *et al*., 1986).

In general, the function of SH2 domains has been defined to specifically recognize the phosphorylated state of tyrosine residues (3-6 amino acid residues) and this is important to allow the SH2 domain-containing proteins to localize to tyrosine phosphorylated sites (Songyang *et al*., 1993; Pawson, 1995). This function is the fundamental event in forming phosphorylated tyrosine which leads to the activation of a cascade of protein-protein interaction, where SH2 domain-containing proteins are recruited to tyrosine-phosphorylated site (Kashishian *et al*., 1992 & Fantl *et al*., 1992)

Stein *et al*. (2003) reported that SH2 domain of hGrb10 is in dimeric crystal structure in solution. The sequence alignment also reveals that the Grb10 SH2 dimer interface are highly-conserved region among the superfamily except for Phe-496. The finding also provides information on Grb10's SH2 domain binding preferences, where binding of dimeric, turn-containing phosphotyrosine sequences, such as the activation loop of the two beta subunits of insulin and IGF-I receptor, are in favour.

SH2 domain of Grb10 has also been reported to take part in various signaling pathway in association with the mediators, such as the receptor tyrosine kinase for insulin (Liu and Roth, 1995), IGF-I (O'Neill *et al*., 1996), EGF (Margolis *et al*., 1992) and PDGF (Wang *et al*., 1999).

2.3.4 BPS Domain

BPS region was termed when mapping studies using the yeast-2-hybrid system and in vitro protein interaction assays revealed the presence of a small domain, which is located between the PH and SH2 domain (He *et al*., 1998). Full length Grb10 was found to interact more efficiently with the insulin receptor in vitro. Rather than SH2 domain alone that contributed to the interaction, there is another addition domain that is also involved in the interaction (Frantz *et al*, 1997). BPS domain was identified as one of the highly-conserved region and it's especially unique to the Grb superfamily (He *et al*., 1998). So far, the known role of the Grb10 BPS region is the interaction with IGF-I receptor and insulin receptor (Wang *et al*., 1999). Other than that, mGrb10 BPS domain was found to interact with Nedd4 protein, which is a ubiquitin protein ligase composed of three WW domains and a hect domain bearing homology to a ubiquitin ligase (Kumar *et al*., 1992 & 1997). In the yeast-2-hybrid system, although BPS domain is eligible to interact independently with Nedd4, but presence of SH2 domain boosted a stronger interaction (Morrione *et al*., 1999). In the insulin receptor mechanism, tyrosine residues are reported to be essential for the interaction with both BPS and SH2 domains (He *et al*., 1998). However, whether the interaction is direct or autophosphorylation of tyrosines occurs and causes conformational change within the domains to expose the binding site of BPS domain remain to be investigated.

2.4 Grb10 Tissue Distribution and Imprinting

2.4.1 Mouse Grb10

In mouse, Grb10 mRNA was found during gestation in placenta and in most tissues of 13 days old embryos (Miyoshi *et al*., 1998) and the protein were highly expressed in muscle tissues, including the face and trunk, intercostals muscles, diaphragm and cardiac muscles and the limbs (Charalambous *et al*., 2003), liver, bronchioles and cartilage of the atlas, ribs and long bones, adrenal gland, pancreatic bud and in the brain of embryos.

Whereas in adult mouse, northern blotting analysis unveiled the presence of Grb10 in heart, kidney, brain, lung, skeletal muscle, and testis (Ooi *et al*., 1995, cited from Lim *et al*., 2004), but not in spleen and liver. Isoforms of Grb10 protein were detected in skeletal muscle, testis and brain, also medium level of expression in the adipose tissues, by using the polyclonal antibody raised against the N-terminus of mGrb10 (Ramos *et al*., unpublished data). The tissue expression levels of the isoforms appear to be tissue specific and the distribution Grb10 in proteins in insulin-responsive tissues, such as skeletal muscle and adipose tissue justify the functions of Grb10 in regulating insulin signaling.

2.4.2 Human Grb10

Northern analysis uncovered the human Grb10 expressing in the highest level in skeletal muscle and pancreas (Liu and Roth, 1995; O'Neill *et al*., 1996). Apart from that, high levels of Grb10 transcript were found expressing in brain and cardiac muscle, while intermediate levels were detected in placenta, lung, liver, kidney, spleen, prostate, testis, ovary, small intestine and colon.

Expression of Grb10 is also found to be expressed in seven different human breast cancer cell lines and multiple bands of endogenous Grb10 are also observed in human cancer cell lines derived from cervix (HeLaS3, HeLa 229), liver (HepG2, Huh7), and in non-cancerous human skeletal muscle cells. (Dong *et al*., 1997). However, beyond these observations, the reasons of the existence of Grb10 in the cancer cell lines remain unexplained.

2.5 Grb10 Imprinting and Silver-Russell Syndrome (SRS)

Genomic imprinting is a phenomenon whereby a gene copy produced from the mRNA transcription, derives in parental-sex-specific manner and it plays a vital role in regulating development, growth and behavior (McGrath and Sotter, 1984; Cattanach and Kirk,1985). Imprinting is believed to be regulated by DNA methylation at the CpG islands through gene silencing event took place in germ line (Surani, 1998). Like any other mechanism, disruption bears consequences. Expression of originally silent allele may lead to double gene dosage, whereas suppression of normally active allele may results in deficiency (Polychronakos and Kukuvitis, 2002).

In mouse, Grb10 gene is reported to be maternally imprinted. The gene is located on proximal chromosome 11, where the gene the resposibles for pre- and post natal growth retardation is reported to be located (Miyoshi *et al*., 1998; Cattanach *et al*., 1998). Overgrowth of embryo and placenta was reported Charalambous *et al*. (2003) when disruption occurs on the maternal Grb10 allele in mice, by resulting in the overgrown size of mutant mice. From these observations, Grb10 is suggested to be a strong candidate in causing the SRS in human.

SRS in human is a genetic disorder which is characterized by pre- and postnatal growth failure and some other dysmorphic features. Some researches claimed that duplication of chromosome 7 imprint contributed to the syndrome as maternal uniparental disomy of the chromosome 7 was demonstrated in approximately 7% of the cases reported (Monk *et al*., 2000). Furthermore, the Fluorescent *In situ* Hybridization (FISH) carried out by Monk *et al*. (2000) confirmed the presence of a tandem duplication at the region around the gene for growth factor receptor-binding protein 10 (Grb10) and insulin-like growth factor-binding protein 1 and 3 (IGFBP1-3), where Grb10 is known as growth repressor. However, there is also evidence against Grb10 as contributor to SRS, based on the absence of imprinting in growth plate cartilage (McCann *et al*., 2001). However, the exact function of Grb10 remain to be investigated.

Figure 2.2 Overgrown size mice with mGrb10 disrupted (left) compared with normal littermate (right) (Charalambous, 2003).

2.6 Gene Duplication

Gene duplication is the copying of a region of DNA that contains gene (Zhang, 2003). This duplication is suggested to the functional divergent in two genes which have

similar molecular structures but carry different regulatory functions and this might also lead to permanent preservation of the two genes. Two hypotheses are proposed in explaining the functional retention in duplicated genes. One of the hypothesis is the neofunctionalization hypothesis, which argues that one of the daughter gene from duplication retains the ancestral function, while the other acquired for new functions (Ohno, 1970, cited from Wikipedia.). Duplication-degenerative-complementary (DDC) is the second hypothesis which proposed that the functions of the ancestral gene are partitioned between the duplicated genes, in such a way that the genes perform the subfunctions of the ancestral gene by complementing each other in functioning (Force *et al*., 1998). In the latest review by He and Zhang (2005), they proposed a new hypothesis which is the rapid subfunctionalization followed by prolonged and substantial neofunctionalization, and termed the new model as subneofunctionalization (SNF). Thus, gene duplication plays a vital role in the evolution of new function (Zhang, 2003).

<div align="center">

Chapter III

Methodologies

</div>

3.1 <u>Primer Design and Sequences</u>

The widely used free web based primer design tool, the Primer3 program, was used for picking out the primers from the sequence input. The program can be retrieved from http://frodo.wi.mit.edu/primer3/. The cDNA sequences were input into the input column and the product size ranges was set from 500-700bp. The rest of the default settings were left unaltered. Then, click the "Pick Primer" button. A few pairs of primer sequence will be suggested from the sequence that had been input, which will be suitable to use as primer. After generating the possible primer sequences, NetPrimer or Primer Premier, which can be retrieved from http://www.premierbiosoft.com/netprimer/index.html, was used to analyze the rating of the sequences. The primer sequences was pasted into the sequence input column and clicked on the "Analyze" button.

3.2 <u>Fish Dissection</u>

Adult Zebrafish were being used for the dissection. For the expression work, 10 target organs were being isolated which include eyes, brain, gills, heart, stomach, intestine, liver, muscle, bone and skin. Due to the small and fragile size of the fishes, modification of tools such as the scissors is necessary to ensure smooth processing.
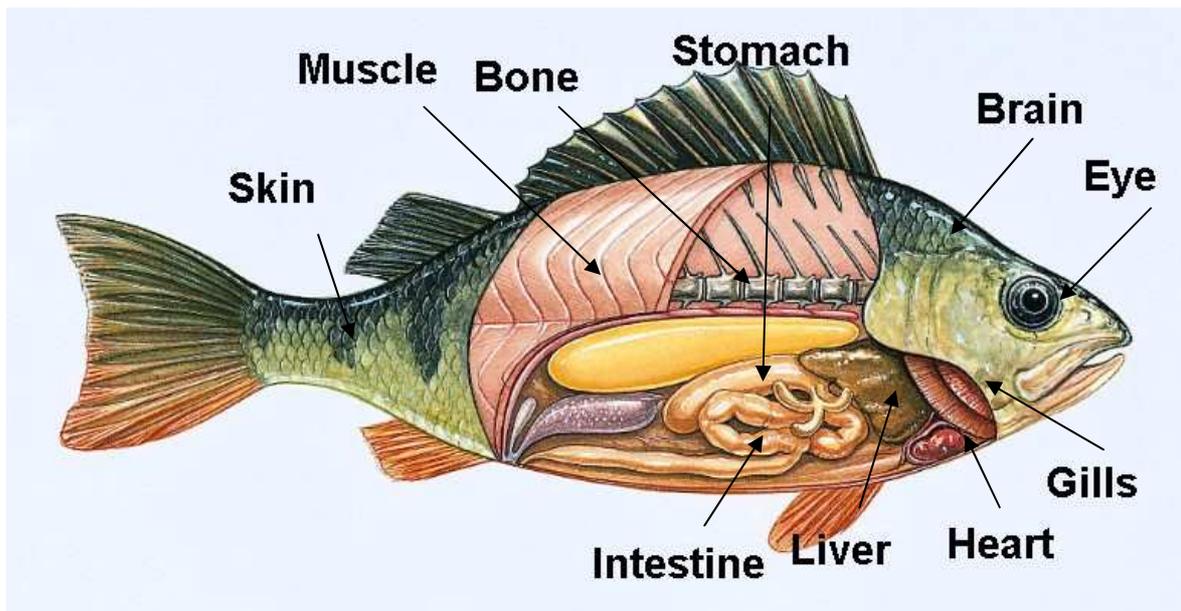
Figure 3.1 Anatomy of bony fish
Retrieved from http://www.dkimages.com/discover/previews/1171/10686362.JPG

3.3 Total RNA Extraction using Tri-Reagent

Targeted organs were minced into fine pieces before approximately 100ml of each organ was placed into a 1.5ml Eppendorf tube. Then, 1ml of Tri-Reagent was added into the tube. Homogenization was done using a few syringes at different sizes. After homogenization steps, tubes were centrifuged at 12,000 rpm for 10 minutes with the temperature of 4°C. After spinning, 200µl of chloroform was added into each tube. The tubes were then vortexed manually for 15 seconds. After that, they were left incubated at room temperature for 5 minutes before proceeding. After incubation, tubes were spun down at 12,000 rpm for 15 minutes. After centrifugation, the upper layer among the three layers was pipetted out and transferred into a new tube. The new transfers were then added with 500µl of cold isopropanol and then incubated again at room temperature for 10 minutes. After incubation, the tubes were centrifuged at 12,000 rpm for 10 minutes to precipitate out the RNA. The supernatant was discarded and RNA pellet was then washed with 1ml of 75% ethanol. The tubes were centrifuged at 7,500 rpm for another 5 minutes and supernatant were discarded.