5-2010

Initial Analysis o f the 3' Spliced Region of the Receptor Tyrosine Kinase RET

Ananya Sengupta
Montclair State University

ABSTRACT

The RET proto-oncogene is expressed in the developing kidney and enteric nervous system during vertebrate embryogenesis. It has two major splicing variants; RET9 and RET51. The expression pattern of the two variants appears to vary among the developing organisms. RET9 plays a major role in development of kidney and excretory system whereas the role of RET 51 seems to be of maintenance of these systems after development. The protein RET is a transmembrane receptor which expresses itself in the cell surface and initiates the signal transduction pathways for cellular differentiation and proliferation of the cell. The different expression patterns shown by RET 9 and RET51 isoforms and the work of other researchers indicating that different RET splice variants result in different phenotype when expressed independently, are demonstrations of the importance of these splice variations in RET functioning. In order to expand our understanding on the mechanism of RET splicing, the 3′ end of the RET 9 isoform was isolated and a new construct was created that contained the exons and introns involved in the RET9 and RET51 splice variations. This region was placed into a mammalian expression vector in both orientations and the ability to measure changes in splicing with these constructs was measured. As a sequential step, in this study two different constructs of a specific region of RET were designed and analyzed using RT-PCR and Q-PCR to demonstrate the mechanism of splicing. These constructs are planned to be further used as a model for the study of conditions that lead to splicing of Human RET proto-oncogene. While no differences were seen in initial studies, the development of the constructs and the evaluation of these constructs will provide an opportunity to study this important splicing event in more detail.

MONTCLAIR STATE UNIVERSITY

Initial Analysis of the 3' Spliced Region of the Receptor Tyrosine Kinase RET

By

Ananya Sengupta

A Master's Thesis Submitted to the Faculty of

Montclair State University

In Partial Fulfillment of the Requirements

For the Degree of

Master of Science

May 2010

College/School College of Science And Mathematics

Department Biology & Molecular Biology

Certified by:

[redacted]

Dr. Robert Prezant
Dean of College or School

05/11/2010

Thesis Committee:

[redacted]

Dr. Quinn Vega
Thesis Sponsor

[redacted]

Dr. Carlos Molina
Committee Member

[redacted]

Dr. Kirsten Monsen
Committee Member

[redacted]

Dr. Quinn Vega
Department Chair

INITIAL ANALYSIS OF THE 3' SPLICED REGION OF THE RECEPTOR TYROSINE KINASE RET

A THESIS

Submitted in partial fulfillment of the requirements

For the degree of Master of Science

by

ANANYA SENGUPTA

Montclair State University

Montclair, NJ

2010

ACKNOWLEDGEMENTS

I would like to express the deepest appreciation to my committee chair, Dr. Quinn Vega, whose encouragement, guidance and support from the initial to the final level enabled me to develop an understanding of the subject. Without his persistent help this thesis would not have been possible.

I would like to thank my committee members, Dr. Carlos Molina and Dr. Kirsten Monsen for their guidance and invaluable help on multiple occasions that helped me to complete the thesis.

Lastly, I am heartily thankful to all of my friends, family & peers who supported me in any respect during the completion of the thesis.

Ananya Sengupta

TABLE OF CONTENTS

INTRODUCTION

RET, an abbreviation for "Rearranged during Transfection", was first identified by Takahashi *et al.* (1985). The transfection of NIH3T3 cells with human lymphoma DNA resulted in the activation of a novel transforming gene. This biologically active gene was generated by the recombination of two fragments of distantly located normal cell DNA[1]. Since then several research works have been undertaken involving the RET gene. These findings have shown that the protein RET is a cell surface receptor and is expressed in cell lineages derived from the neural crest. It plays an important role in regulation of cell proliferation, migration, differentiation and survival during embryogenesis. This protein consists of three extracellular and an intracellular domain. The extracellular region consists of four cadherin-like repeats, a calcium-binding site, and a cysteine-rich domain while the intracellular portion contains a typical bipartite tyrosine kinase domain.[2]

The receptor RET is activated by one of the 4 ligands of GDNF family; GDNF, neutrin, artemin and persephin. The activation of RET is mediated by one of the four coreceptors named as GDNF family coreceptors (GFRα1-4). The interaction of RET with GDNF and GFRα1 occurs through its cadherin domain (CLD4 and CRD).[9] RET activation occurs by initial binding of one of the GFL (GDNF family ligands) with one coreceptor from GFRα1-4 family. This is followed by binding of one of the RET molecules to the above formed GFL-GFRα complex and subsequent joining of the other GFRα & RET molecules. Henceforth, resulting in the activation of RET. Absence of the GFL-GFR complex inhibits the dimerization of the two RET molecules and its downstream signaling. It has been found that wild type RET phosphorylation occurs only after being stimulated by its ligands and there is

an increase in the phosphorylation event with the presence of GFR coreceptors. This event of dimerization and ligand receptor interaction results in strong control of the cells response to growth factor.[10] Murakami et al (2007) have proposed that GDNF plays a important role in RET downstream signaling. In their paper they established the fact that GDNF regulates the downstream signal transduction pathway of RET along with SPHK1 gene.[32] The coreceptor Gfra4 from the GFRα family of receptors has also been reported to play an important role in activating RET along with GDNF during endocrine cell development. [33]

The expression of RET is essential for embryonic development, formation of enteric nervous system and uteric bud development. RET protooncogene has been primarily associated with Hirschsprung disease (congenital agangliosis of the gastrointestinal tract), thyroid carcinoma, and Multiple Endocrine Neoplasia 2 (MEN2A, MEN2B) syndromes.[7,8] Because of its importance in various biological processes, scientists have had an increased interest in the study of the RET proto-oncogene in various research field like tumor biology, neurogenesis, neuronal disorders, nephrogenesis and reproductive medicine and spermatogenesis.[4]

RET- Tyrosine kinase domain:

The binding of RET with GDNF and GFRα1 results in autophosphorylation of RET. The phosphorylation of RET occurs in its intercellular tyrosine kinase domain. The autophosphorylated RET forms the docking site for various adaptor proteins which finally initiates the signaling cascade for the following signaling pathways: RAS/extracellular signal-regulated kinase (ERK), phosphatidylinositol 3-kinase (PI3K)/AKT, p38 mitogen-activated protein kinase (MAPK) and c-Jun N terminal kinase (JNK) pathways[4]. After the activation of RET and its initiation of signaling cascade, RET is ubiqutinated and degraded. The degradation of RET is mediated by proteosomes[5]. The two isoforms share 1,063 amino

acids but differ at their carboxy-terminal tails. RET9 ends with a 9-amino-acid tail and is 1072 amino acids long, whereas RET51 has 51 different amino acids at its carboxy-terminal tail with a length of 1114 amino acids. There are total 18 tyrosine residues in the intracellular domain of RET. Tyr905, Tyr1015, Tyr1062, and Tyr1096 have been found to be of special significance in RET signaling events. The longer isoform of RET (RET51) contains all the 18 tyrosine residues while the shorter isoform of RET (RET9) lacks the two tyrosines (Tyr 1090 and Tyr1096)[11,12]. The Tyr1062 of RET forms a docking site for several adaptor proteins like DOKs, Enigma, Shc, FRS2 and hence forms an important site in GDNF- RET signaling pathway. It forms the multidocking site for both the isoforms but shows some difference in the sequence context. The Tyr1062 has also been seen to play an essential role for GDNF- mediated neuronal differentiation and signaling. Among the other tyrosine's of RET, Tyr 905, Tyr1015 and Tyr1062 undergoes robust phosphorylation with increasing concentration of GDNF during the embryonic development in DRG (dorsal root ganglion) neurons.[11,12] The Tyr1062 has been found to play an important role in differentiating the activity of the two isoforms of RET (RET9 and RET51). The structural and functional difference between RET9 and RET51 results from the alternative modes of binding of adaptor proteins to the Tyr1062 residue. Mutation of Tyr1062 in RET9 isoform results in the reduction of its ability to develop the enteric nervous system but doesn't show any effect on RET51 activity. Y1096 is essential for the binding of the Grb2 adaptor protein. Jain et al (2006) have reported that presence of Y1096 in RET51 helps it to perform its signaling process even in the absence of Y1062. The adaptor protein, Grb2 helps in coordinating cell proliferation and differentiation mediated by RET during embryonic development[41]. Hence, Wong et al. in 2005 reported that among the two isoforms of the gene RET, RET 9 plays a

larger role in the development of kidney and enteric nervous system[13]. Previous research on this has also indicated that expression of RET9 occurs much earlier during the embryonic development than RET51. In the case of Hirschsprung disease (HSCR) it has been found that the people affected by the disease possess undeveloped kidneys. Also the primary cause for the HSCR disease has been recognized as the missense mutation of RET51 isoform but not RET9[14]. This report further establishes the fact that the two isoforms RET9 and RET51 vary in their physiological function and both are required for kidney development, but in different stages.

RET and diseases:

The gene RET is expressed in the developing kidney, parasympathetic nervous system and central nervous system nuclei during vertebrate embryogenesis. In the adult tissues RET expression is seen in the peripheral enteric, sympathetic and sensory neurons plus in neuronal cells of the brain[3]. Mutation of the RET gene is responsible for Medullary Endocrine Neoplasia (MEN2) syndromes, Papillary carcinomas of the thyroid gland (PTC), and Medullary thyroid carcinoma (MTC).

Mutations in the extracellular domain of the RET gene results in MEN2A while intracellular domain mutations causes MEN2B. MEN2B mutation results from the change of Methionine in tyrosine kinase domain of RET. Wild type RET activates with the help of ligands but in case of MEN2B the replacement of methionine with threonine results in a mutant RET tyrosine kinase domain which can activate itself without ligand binding[15]. MEN2A causes substitution of extracellular cysteine with several other residues (Ex. Arginine). These cysteines form intramolecular disulfide bonds in the wild type receptor. Due to the

replacement of cysteine with other residues in the mutant receptor, the cysteine start forming intermolecular disulfide bonds with another RET-MEN2A monomer resulting in the activation of RET without any ligand binding.[16]

A fusion of the RET tyrosine kinase domain with the 5' terminal region of the heterologous genes results in chimeric oncogenes designated as RET/PTC. To date 12 rearranged forms of RET have been found. In each case the intercellular domain of RET is fused to different activating genes namely H4, PCM1, ELKS, RFP. RET, H4, RFG all of which map on the long arm of chr.10 and their fusion is generated by inversions. In case of RET/PTC2, it occurs by reciprocal balanced translocation between chromosome 10 and 17 between RET and RI alpha.[17]

Hereditary medullary thyroid cancer (MTC) is inherited as an autosomal dominant trait, which means each child of an affected parent has a 50/50 probability of inheriting the mutant RET proto-oncogene from the affected parent. DNA analysis makes it possible to identify children who carry the mutant gene; surgical removal of the thyroid in children who carry the mutant gene is curative if the entire thyroid gland is removed at an early age, before there is spread of the tumor.[18,19]

Hirschsprung's disease occurs during the fetal development, resulting from the failure in the development of nerve cells in the enteric nervous system. Absence of these nerve fibers results in intestinal obstruction, enlargement of the colon and ultimately leads to bowel obstruction due to the absence of normal enteric nerves of that section. This disease is named after Harald Hirschsprung who first described the disease in 1886. The disease usually occurs in infancy. It is multifactorial, involves several genes some of which are still unknown and

has high sex bias. It has been found that it gets transferred from the male parent to his son in most of the cases and the reason for that is still unknown.[21]

RNA splicing:

Splicing refers to a process in which a mature functional mRNA sequence is formed by cutting and removal of the introns and ligation of the exons[24]. It has been reported that of the ~25,000 genes encoded by the human genome more than 70% produce transcripts that are alternatively spliced.[51] According to Wang et al. (2008) splicing is essential for removing introns from nearly 90% of the human gene sequences. They have also reported splicing defects are responsible for most of the mutations associated with human hereditary diseases[49]. It has been also recorded that ~16% of the 31,250 point mutations listed in the data base of human disease alleles are located within splice sites[53]. The first chemical step of splicing includes cleavage of the 5' exon/intron boundary through nucleophilic attack by the 2'-hydroxyl of a specific branch point adenosine located within the intron. This step results in the generation of a 5'-exon fragment and a lariat intermediate that contains intron, the 3'-exon sequences, and the branched adenosine. The second chemical step is followed by the cleavage of the 3'-exon/intron boundary through the nucleophilic attack of the 3'-hydroxyl of the 5'-exon at the 3'-splice site. This step results in the ligation of the exons and the release of the intron in the form of a lariat[50]. In the pre-mRNA that undergoes splicing the 5'-splice site is defined by a single element of 8 nucleotides and the 3'-splice site is defined by three sequence elements usually contained within 40 nucleotides upstream of the exon/3'-intron junction[52]. Human genome mostly contains exons flanked by long introns. According to Hertel (2007) experimental and computational analysis of the human genome has shown that

the length of the upstream intron is more important in inducing alternative splicing than the length of the downstream intron[53].

The process of splicing is mediated by the spliceosome machinery, a complex of more than 100 polypeptides and 5 uridine rich small ribonucleoproteins that assembles on the pre-mRNA. U1snRNP is the first factor to interact with the pre-mRNA and recognizes the 5' splice site. It form complex with the pre- mRNA and commits the pre-mRNA to the splicing pathway. Human snRNP is composed of U1 snRNA, SM proteins and U1 specific polypeptides. Forch et al (2002) reported that apart from U1snRNP, the protein TIA also plays an important role in splicing machinery and binds with uridine-rich sequences downstream from 5' single strand and promote U1snRNP binding.[47]

In many cases, splicing can produce a range of various mRNA by varying the exons of the same pre-mRNA; this is known as alternative splicing. These mRNAs then undergo translation and result in the production of different isoforms that differ in their peptide sequences and thus have varying chemical and biological activity[25, 26]. Alternative splicing generates protein isoforms that differ in their binding properties to ligands. It causes 2 to 10 fold changes in binding affinity and it further determines the ligand specificity of a receptor.[27] The isoforms generated by alternative splicing in turn lead to the generation of modified final protein product. Black et al. (2003) have reported that the process of alternative splicing is attributed mainly to the activities of splicing enhancers and repressors that further allows transient interactions with splicing regulators.[26] Splicing enhancers and regulators are the cis-acting auxiliary sequences that occur within both exonic and intronic regions and are responsible for exon inclusion and exon skipping by recruitment or disruption of the spliceosome machinery.[53] Presence of exon splicing enhancers such as SR

proteins results in the stimulation of spliceosome components such as U2AF. The presence of splicing silencers such as hnRNP A1 protein results in shifting of splice sites in pre-mRNA. Similar effects are seen with the presence of intron silencer & enhancer elements like hexanucleotide UGCAUG and GGGGG. The presence of the above mentioned opposing regulatory elements in splicing mechanism has resulted in the elaboration of a yang yang model of alternative splicing. According to this model splicing occurs when positively acting sites (splicing enhancers) outnumber the negative sites (splicing silencers) [54]. The invariable elements like splice site sequences and exon/intron architecture have also been reported to possess potential for mediating differential splicing[53]. Splicing removes many protein interaction domains and effects the signaling pathways. Romero et al (2006) reported in their article that to date 46% of the studied proteins which have alternative splice variants are involved in signal transduction, gene expression or regulation[43].

RET and alternative splicing:

The gene RET has 21 exons and it has been reported to splice alternatively resulting in the formation of three splicing variants (isoforms); RET9, RET51, and RET43. All these three isoforms are similar untill the 19th exon and they diverge thereafter[6]. The isoforms of RET are well conserved among different species like human, mouse, rat, chicken, and pufferfish. When the C-termini of RET was compared among different isoforms it was seen that the amino acids of both the isoforms RET9 and RET51 are similar among all the species examined. Further comparison of the amino acids both upstream and downstream of the Y1062 and Y1096 region of the RET protein showed 80% homology among all species[22]. Recently, similarity in the expression of RET isoforms has been seen within the neuronal cells of human and Zebrafish[23].

The loss of one of the isoforms of RET generated by alternative splicing has shown to cause malformation of kidney and enteric nervous system (ENS) in humans while the loss of the other isoform of RET did not have any effect in development of both[28]. The occurrence of alternative splicing has also been reported in papillary thyroid carcinoma. Fluge et al (2001) reported the presence of multiple alternative splicing variants of RET-PTC in the genome of the patients suffering from the disease. They reported the presence of two types of alternative splicing events in the tyrosine kinase domain of the RET protein, one with exclusion of exon 19 and the other with the retention of intron19, apart from the usual RET9 and RET51 splicing variants.[29] Hir et al. (2000) have reported that RET51 protein isoform has growth advantage in MEN2 tumors than RET9 protein isoform. The mutation in RET gene is responsible for MEN2 syndromes where mutated RET receptor causes alteration of the downstream signaling and changes the pattern of the target gene expression. Hickey et al (2009) have described in their article that there is a significant difference in the expression level of six genes (PIK3CG, RAB3A, INDO, DDX21, CWF19L1, and DSG1) associated with MEN2 syndrome when induced by either RET9 or RET51. Some of these genes are either up-regulated or down-regulated by the two isoforms of RET. Gene INDO is of special importance which is expressed in most human tumor that is induced by RET9 but not by RET51. Hence, splicing of RET plays a significant role in RET's regulation and signaling pattern.[20] According to Romero et al (2006) alternative splicing occurs mostly in those polypeptide segments which are intrinsically disordered. Alternative splicing thus enables functional and regulatory diversity while avoiding structural complications. They analyzed 46 differentially spliced genes of human proteins and have shown that 81% of them contain partially or fully disordered protein region.[43]

There has been extensive research going on so far with the Isoforms of RET and their expression. It has been shown that RET gene displays its multiple functions by expressing various forms of its 3'splicing variants and the aberrant expression of various transcripts may lead to developmental defects and tumor formation. As demonstrations of the importance of these splice variations in RET function, RET 9 and RET51 have shown different expression patterns.[6] Previous research in our laboratory has indicated that ligands like glial cell line-derived neurotrophic factor (GDNF) plays an important role in the rate of expression of the isoforms.

In this study, 3' end of the RET 9 isoform and a specific region of the RET gene containing intron19-exon20 has been isolated. The region, exon19 to exon20 has special importance in the gene RET as this is the region which undergoes differential splicing. There have been several other research works as well with focus on this region. For example, Jijiwa et al. in 2004 created a construct of genomic RET surrounding the above mentioned regions for their studies on neuronal cells.[30]

This study also includes the designing of two genomic DNA constructs of RET by isolating the 3′ genomic DNA sequence of human RET gene with an aim to measure the different conditions of splice variation and also identify conditions which lead to splicing of human RET proto-oncogene. The constructs created, also aim to measure the expression of RET 9 and RET51 isoforms using a quantitative QPCR technique. These constructs may play an important role in further studies related to the event of RET splicing and can be useful to find the regulator which is responsible for the production of the two isoforms of RET.

MATERIALS and METHODS

1. Genomic DNA extraction:

Human genomic DNA was isolated from MCF breast cancer cells using TRIZOL reagent (TRIzol® Reagent, Invitrogen Life Technologies #15596 USA). 5-10 X 106 MCF cells were lysed with in 1 ml of Trizol reagent by repetitive pipetting. Then phase separation of the cells was performed by incubating the homogenized samples for 5 minutes at 15˚-30˚c followed by addition of 0.2 ml of Chloroform per 1ml of Trizol reagent. Centrifugation of the sample at 14,000 rpm for 15 minutes at 2 to 8 ˚c resulted in the separation of mixture into 3 phases with the presence of DNA in between the interphase and phenol phase. The DNA was further precipitated with 300µl of 100% ethanol for 30 minutes. This was followed by centrifugation at 5981 rpm for 5 minutes and a second wash with 75% ethanol. The DNA was ultimately redissolved with 8mM NaOH after the ethanol was air dried. The pH of the sample was adjusted as 7.5 with 159 µl 0.1M HEPES and 10 µl of 1mM EDTA. The isolated DNA concentration was measured using NanoDrop™ spectrophotometer (Thermo scientific).

2. Polymerase Chain Reaction:

The linear genomic DNA sequence of human RET gene was taken from NCBI and EMBL (NCBI ID No. : AJ243297, EMBL ID No. : ENST00000340058). Exon 19 to exon 20 including introns 19-20 of the human RET gene was targeted for PCR amplification based on sequence data collected from the above mentioned reference sources. Polymerase chain reaction was performed with the two sets of designed primers. The first set of primers include one Forward primer at Intron18 (5'CACTGTCTCGCTTGGATGAA3') and two Reverse primers; one at Intron19 (5'AGGAGGCACTCGTCTGTTGT3') and the other at Exon20

(5'GCACACACCACAAGGCTCTA3'). These primers were designed to amplify a 2000 bp and 4000bp regions of the RET genomic DNA respectively. This was followed by the designing of a new set of primers to amplify a more precise region consisting of a 1910bp region with the Forward primer at Intron18 (5'tctcttggagaggtcaggag 3') and reverse primer at Intron 19 (5' CACTCGTCTGTTGTCTGAGC 3') around the splicing site. A reverse primer at Exon 20 (5'TCCCCTTGTGAGTCCATTAC 3') was also designed along with above mentioned forward primer of the second set to amplify a 900 bp region consisting of the 3' end of the RET9 isoform.

Five different dilutions of the primers were made with concentration of forward and reverse primers in each 25µl reaction mixture as 0.1µM-0.5µM with 1 µl of DNA and 12.5 µl of HotStarTaq® PCR master mix (QIAGEN sample and assay Tech.). PCR was performed using HotStar master mix protocol with 15 minutes at 95°c for activation, 1 minute at 94°c for denaturation, 1minute at 53°c for annealing and 3&10 minutes at 72°c as extension for 40 cycles in thermocyler. The amplified product was visualized by running it in 1% agarose gel.

Agarose gel: The amplified DNA was run on 1% agarose gels made with 0.5 g of agarose per 50 ml of 1X TAE. The gels were run in 1X TAE with 5 µL of Ethidium bromide (EtBr) at 100 Volts for 35 minutes.

3. Ligation of RET DNA into pGEM-T cloning vector:

Ligation: A 10 µl reaction mixture was prepared by adding 1µl of the PCR amplified DNA, 1 µl of PGEM-T easy vector (pGEM®-T Easy vector system from PROMEGA Corp. Cat. No. A1360), 5 µl of 2X buffer and 1 µl of T4 DNA ligase (INVITROGEN life technologies Cat. No.15224-041). The reaction mixtures were kept at room temperature for 1 hour.

Transformation: The ligation reaction was centrifuged briefly followed by thawing of the frozen MAX Efficiency® DH5α™ Competent Cells (INVITROGEN Life technologies Cat. No.18258-012). 2 µl of the ligation reaction was added to the 50 µl of the cells. Three such reaction tubes were made one experimental, positive & negative control and kept in ice for 20 minutes. The tubes were then heat shocked for 45-50 seconds at 42°c water bath and plated in the LB/Amp plates. The plates were incubated overnight 16 hrs at 37 °c.

Plasmid isolation: Six colonies from the bacterial plates were picked and each was placed in a test tube containing 10 ml of LB along with 10 µl of 50 mg/ml amp. The bacterial Miniprep procedure was followed according to the QIAprep® Miniprep protocol (QIAGEN, USA). The test tube was incubated for 12-16 hrs at 37°c with vigorous shaking (220 rpm). 1.5 ml of the overnight Bacterial culture from each test tube was taken in a microcentrifuge tube and spun for 1 minute at 13,000rpm. The pellets were resuspended using 250 µl Buffer P1 and vortexed. 250 µl Buffer P2 were added and mixed well. 350 µl of Buffer N3 was added and the tubes were again inverted four to six times. The samples were centrifuged for 10 minutes at 14,000 rpm and the supernatant transferred to a spin column placed in a 2 ml collection tube. The spin column and collection tube were centrifuged at 14,000 rpm for 30 to 60 seconds and the flow through was discarded. 750µL Buffer PE was added to the spin column and centrifuged for thirty to sixty seconds and the flow through discarded, centrifuged for one minute and the collection tube with the flow through discarded. The spin column was placed in a new 1.5 ml microcentrifuge tube and 50 µl of sterile water was placed in the center of the spin column. After one minute the column was centrifuged for one minute and the flow through containing the plasmid DNA was collected.

Restriction enzyme digestion: The Plasmid DNA was digested with ECOR1. A master mix was being prepared for 18 reactions containing 20 µl of 10 X buffer, 3 µl of enzyme and 157 µl of sterile water. Each reaction was made with 8 µl of master mix and 2 µl of sterile water. The 1.5 ml microcentrifuge tubes containing the reactions were kept in the water bath at 37°c for 4-5 hrs.

DNA sequencing: To further confirm that the amplified region doesn't contain any mutated product, the pGEMT-RET insert was sent out for sequencing to an outside lab using SP6 and T7 universal primers. (GENEWIZ).

4. Insertion of the approx. 1900 bp amplified region of the RET gene ligated with pGEMT into mammalian expression vector pFlagCMV-1:

Large scale precipitation of pFLAG-CMV-1: pFLAG-CMV-1 was first obtained in large quantity using QIAGEN (QIAprep Spin Midi/Maxiprep Kit) large scale precipitation protocol. The pFLAG-CMV-1 DNA was first transformed into MAX Efficiency® DH5α™ Competent Cells (INVITROGEN Life technologies Cat. No.18258-012) and a colony was picked up for the bacterial culture. The 10 ml LB culture containing 10 µl of amp was incubated at 37 °c for 6-8 hrs. The culture was then picked up and transferred into a 500ml LB broth with 500 µl of amp for 12-16 hrs. The culture was spun for 10 minutes at 3000rpm and the pellet was redissolved with 10ml of buffer P1. The solution was transferred to a new tube and 10 ml of P2 was added and mixed vigorously. 10ml of buffer P3 was subsequently added and centrifuged for 30minutes at 12,000 rpm. The supernatant was then transferred to a column by passing through two layers of cheese cloth. The column was further washed 2 times with 30 ml of buffer QC and 15 ml of buffer QF was added. The eluted DNA was further collected in a collecting tube and isopropanol was added and the mixture was

centrifuged at 12,000 rpm at 30 minutes. The supernatant was poured out and 5ml of 70% Ethanol was added to the pellet. The pellet was further washed with 200 µl of EB buffer and was transferred to an eppendrof tube.

Restriction enzyme digestion: The plasmid p-FLAG-CMV1, pGEMT-RET (the amplified region of RET ligated with pGEMT) were digested with ECOR1, SAL1 and NOT1 to isolate the RET construct from the cloning vector pGEMT and further insert it into mammalian expression vector pFLAG.

DNA purification using column: A large scale prep of the pGEMT-RET plasmid was performed and the DNA concentration was also determined using (250µg/ml). This time the two large scale prep plasmid samples were digested using another restriction enzyme NOT1 for 2 hrs and the insert containing plasmid pGEMT was further digested with SAL1 for 3 hrs to prevent its ligation to itself. The plasmids were re purified using a CENTRI-SEP COLUMNS protocol (PRINCETON Separations Inc. Cat. No. CS-900). The column was first hydrated by adding 800 µl of water, shaken, vortexed and kept for 30 minutes at room temperature. After the gel inside the column has settled the air bubble were removed by tapping the column and the stopper of the column was separated. The excess column fluid was allowed to drain into a waste tube. The column was spun at 3000 rpm to remove the interstitial fluid. Approximately 300 µl of the fluid was removed. 20 µl of Dye-Deoxy terminator reaction mixture was added to the top of the gel followed by careful dispersion of the sample directly to the center of the gel. The column was placed into a sample collecting tube and the spun at 3000rpm for 2 minutes. The purified sample was collected in the tube.

Ligation: The purified product was then ligated in different vector: insert ratio (1:3, 1:6, 1:10, and 1:15) using T4 DNA ligase (PROMEGA Cat. No. M1801) for 1hr.

Transformation of the ligated products: All the ligated products of the above mentioned ligation reactions were transformed to 50 µl of one shot MAX Efficiency® DH5α™ Competent Cells (INVITROGEN Life technologies Cat. No.18258-012). The bacterial transformation reaction were placed on pre-warmed LB/amp plates and grown overnight at 37°c for 16 hours.

Plasmid isolation: 8 colonies from the 2 plates (1:3 and 1:6 ligation reactions) were further cultured for 16 hrs and the plasmid DNA was isolated by following the QIAprep® Miniprep protocol.

Restriction digestion: Further digestion with the restriction enzyme NOT1 was performed for an hour and a 1% agarose gel run.

Orientation confirmation: The two colonies containing the ligated pFLAG vector plus insert were further digested with the restriction enzyme ECORV.

Confirmation of the ligated product: Both the ligated plasmids were sent for sequencing to further confirm their orientation. The sequencing was done using CMV forward primer from GENEWIZ.

Large scale prep of the plasmid -A large scale prep of the new pFLAG-RET was performed using QIAprep® Maxiprep protocol.

5. Mammalian Cell Transfection

30 µl of lipofectamine was added to 1 ml of serum free media in a microcentrifuge tube. 5 µg of Plasmid DNA was added to 1ml of serum free media. The two tubes are mixed and incubated for 10 minutes at room temperature. The Vero cells (cell line derived from kidney epithelia cells of green monkey) were washed 2 times with 2.5ml of serum free media and 2.4 ml of serum free media was added to the tube. The tube was now added to the plate. The

plate was incubated for 4-6 hrs at 37°c. The media was removed and the cells are washed 2 times with 2.5 ml media plus serum or 1X DPBS. 5 ml of media plus serum was further added and the cells were incubated for 48-72 hrs.

6. Isolation of total RNA

After transfection the RNA was extracted from the Vero cells using Absolute RNA Miniprep kit (STRATEGENE). 4.2µl of β-ME was added to 1000 µl of lysis buffer. The tissue culture dish was aspirated and lysis buffer β-ME mixture was added to the dish. The cell lysate was mixed by repeated pipetting and the lysate was transferred to a microcentrifuge tube. 700 µl of the homogenate was added to the spin column and spun at 14,000 rpm for 5 minutes. 70% Ethanol was added to the filtrate and the mixture was filled to rRNA binding spin cup. The mixture was spun at maximum speed for 30-60 seconds. 600 µl of 1X wash buffer was add and spun for 30-60 seconds. The wash step was repeated twice and 300 µl of 1X low-salt wash buffer was added. The cup was spun for 2 minutes to dry the fibre matrix and then transferred to a 1.5 ml microcentrifuge tube. 100 µl of warmed elution buffer was added and the mixture was incubated for 2 minutes at room temperature. The RNA was furthered stored at -80°c. The concentration of the RNA was measured with 1XTBS in a spectrophotometer.

7. mRNA isolation

The mRNA was isolated from the total RNA obtained from neuronal cells (SKNSH human neuroblastoma cell line) using Oligotex mRNA spin-column protocol. The 100 µl of total RNA (1mg/ml) was pipette into a RNase free microcentrifuge tube and 150 µl of RNase free water, 250 µl of buffer OBB, 15 µl of Oligotex suspension buffer was added. The sample was incubated for 3 minutes at 70°c in a water bath. The sample was removed from the water

bath and kept for 10 minutes at room temperature and centrifuged at 14,000 rpm for 2 minutes. The mRNA pellet was resuspended by adding 400 µl of Buffer OW2 by vortexing and the mixture was transferred to a spin column. The column was centrifuged and 400 µl of Buffer OW2 was added to the column and further centrifuged for a minute. 20-100 µl of hot (70˚c) Buffer OEB was added onto the column and the resin was resuspended. The tube was centrifuged for 1 minute at maximum speed and 100 µl of (70˚c) elution buffer was added for further yield.

8. RT PCR

cDNA synthesis- The SuperScript™II First Strand Synthesis System for RT-PCR (Invitrogen Life Technologies Inc.) was used for the reverse transcription reaction following manufacturer's protocol. The primer pair sets designed for amplifying the RET construct RNA are as follows: The Forward Primer 5'ATGCTATGAGGCTGGCCCGT 3' was designed in exon19 and the reverse primer5'GGCAGTGTAGGGCAACAGCA3' pair was designed in intron 19 spanning 175 bp region for amplification of RET9 isoform. The other set was designed to amplify 150bp region in between exon19 and exon20 for RET51. It included forward primer in the exon19 5'CCCGTGTGCACCCTCGATTT3' and the reverse primer in the beginning of exon 20 5'GGCAGTGTAGGGCAACAGCA3'.

Intron-spanning primers Intron19 forward 5' ATGCTATGAGGCTGGCCCGT3' and Intron19 reverse 5' GGCAGTGTAGGGCAACAGCA 3' were designed to amplify 160bp region of intron19 in human RET.

Two sets of specific exon splice junction primers with the first set containing Forward Primer 5' CTCTATGGCATGTCAGACC 3' in between 7 bases in exon 19 & 12bases in exon 20, reverse Primer 5' CCCTTGTGAGTCCATTAC 3' with all bases in exon 20 and the second

set with the Forward Primer 5' AAACTCTATGGCATGTCAG 3' within 10 bases in exon 19 & 9 bases in exon 20 and reverse Primer 5' CCCTTGTGAGTCCATTAC 3' same as above with all bases in exon 20 were designed.

10µl of the sample containing 1mg/ml total RNA, 10mM dNTP mix, 2µM gene specific primer and RNase free water were heated at 65˚C for 5 minutes. The samples were then chilled on ice for I minute. The samples were further mixed with 10 µl of DNA synthesis Mix (4µl of 5X RT Buffer, 2 µl 50mM MgCl2, 2 µl of 0.1M DTT, 1 µl of RNAseOUT ™, 1µl Superscript ™IIIRT. They were incubated at 50˚C for 50 minutes followed by a final incubation at 85˚C for 5 minutes.

The cDNA prepared from the previous step was amplified using the above mentioned gene specific primers with HotStarTaq® PCR master mix protocol (QIAGEN sample and assay Tech.).

9. Quantitative Real time Polymerase chain reaction

For quantitative real-time polymerase chain reaction (qPCR) experiments the SYBER green Expression Assays (STRATEGENE) were used. The qPCR analyses was performed by using an MAXPRO 3000 sequence detection system (STRATEGENE).The qPCR reaction mixture consisted of DNA diluted 10 fold from the reverse transcriptase reaction (1µl), forward and reverse primers (1 µl 0.5 µM) with 2X SYBER Green Master M ix (1 2.5 µl) in a final volume of 25 µl. The qRT-PCR was performed in triplicate for each cDNA sample.

RESULTS

1. Isolation of 3' end of the RET gene

In order to isolate the 3' end of the RET gene for future splicing studies, genomic DNA was isolated from MCF-7 cells, a human breast cancer cell line. Using 5x 106 cells, 213ng/µl of genomic DNA was isolated. Polymerase chain reaction was performed with the isolated DNA & designed primers, to amplify 3' end of the RET gene. The failure of the amplification was attributed to the low concentration of the genomic DNA extracted from human MCF cells. Thus 100µg of Human Genomic DNA was obtained from Promega and was used to amplify 1910bp in between intron18-exon20 & 1289bp in between intron18-intron19. The first pair of primers was used to amplify the region in between the two exons; exon19 and exon20 where splicing of the RET gene takes place and the two isoforms RET9 and RET51 are formed. 1289bp region amplified with the second pair of primers included the end portion of the 3' carboxy tail region of the RET9 isoform which contains the last exon (exon19) and few bases of intron19. The polymerase chain reaction was performed with 5 different concentrations (0.1µM-0.5 µM) of the two set of primers. The first five lanes represent (1900 bp) amplified region with 5 different concentration of the first set of primers. The 6^{th}-8^{th} lane represent (1200bp) amplified region with 5 different concentrations of the second set of primers.

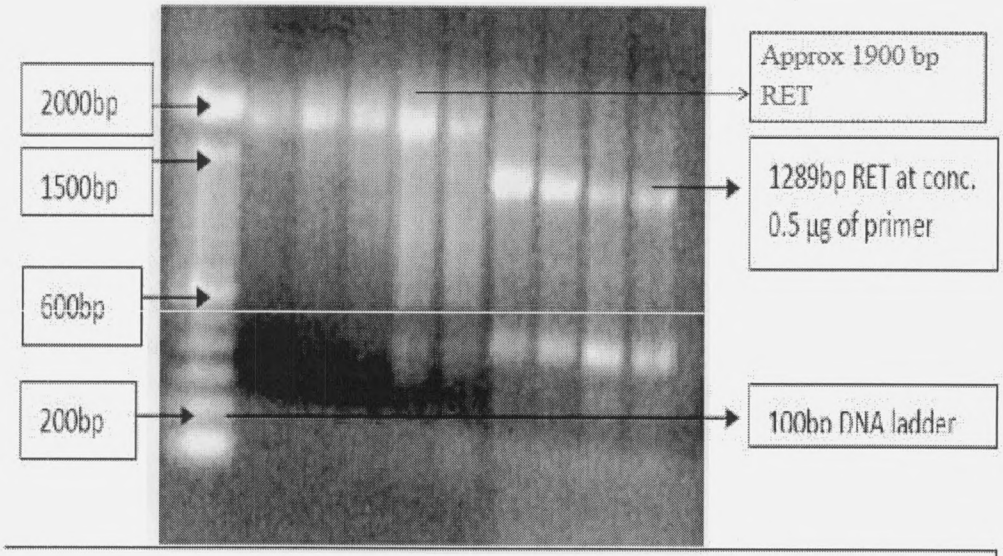


Figure 1: Gel showing both 1910bp and 1289bp region of RET DNA amplified with different concentration gradient of primers

The 1910 bp region was chosen for further experiments as it includes the specific region exon19 and exon 20 where splicing event of RET proto-oncogene occurs.

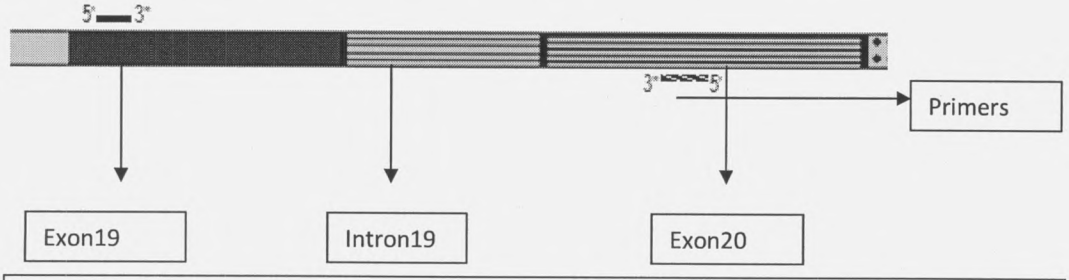


Figure 2: A schematic diagram of the splicing region of the RET transcript. Arrows denote primers used to amplify the region.

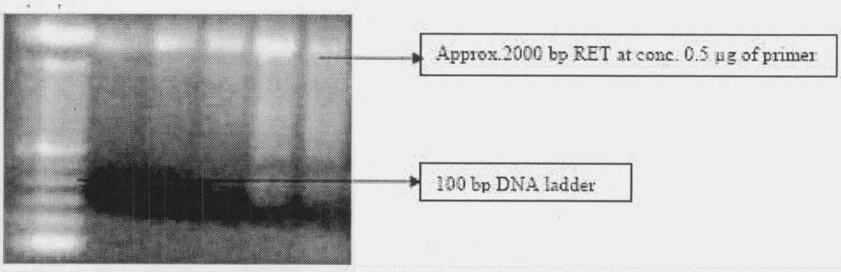


Figure 3: Gel showing 2000bp region of RET DNA amplified with different concentration gradient of primers

2. Cloning of RET insert into pGEM-T vector

The RET (1910 bp) DNA from the previous PCR reaction was ligated to the pGEM-T cloning vector. This vector was used for its cloning efficiency and its characteristic 3' thymidine-overhang which easily ligate (without the use of ligase) with the deoxyadenosine overhang that Taq polymerase leaves in PCR products. A positive and a negative control were also made by adding a control insert DNA in case of positive control and only water in negative control. After ligation, the vector containing 1910bp insert of human RET DNA was transformed into 50 µl of competent DH5- α bacterial cells. The transformed cells were plated onto pre-warmed LB-Ampicillin plates and grown overnight. 16 colonies grew, 8 of which were later cultured in LB/amp broth overnight for miniprep. The excised DNA samples of the miniprep were then restriction digested with ECOR1 and showed the presence of approx. 2kb insert with the 3kb pGEMT plasmid vector when run on a 1% agarose gel. Thus a new plasmid containing the foreign DNA (as an insert) was assumed to have been obtained.

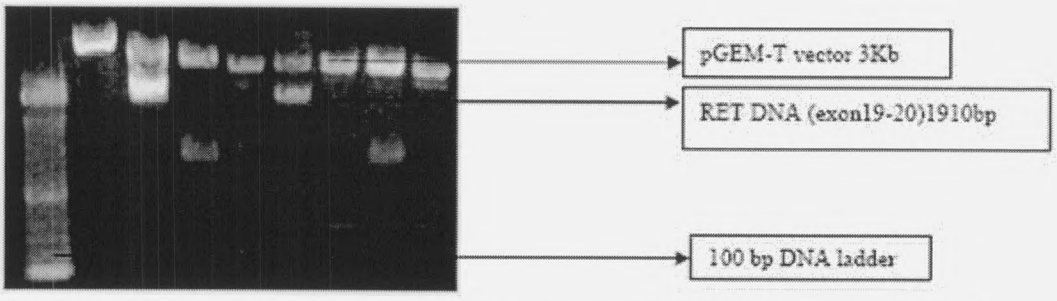


Figure 4: Restriction enzyme ECOR1digestion of the plasmid (pGEMT)

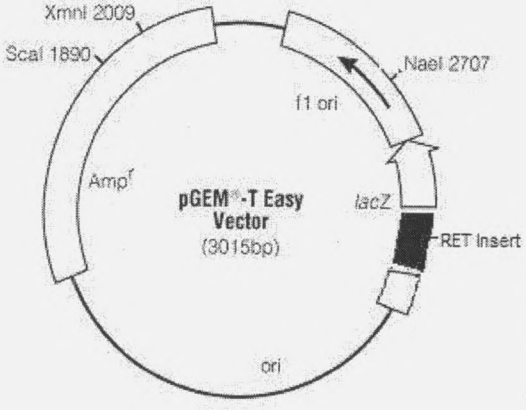


Figure 5: The map showing the exact location of the insert in PGEM-T vector.

The automated sequencing data thus obtained had 99% similarity with human RET genomic sequence (exon19-20) when a blast search was done using NCBI database.

NCBI Blast result of pGEMT- RET sequenced with Forward SP6 primer-

```
>ref|NG_007489.1| Homo sapiens ret proto-oncogene (RET) on chromosome 10
Length=60283

 Score = 1729 bits (936),  Expect = 0.0
 Identities = 941/943 (99%), Gaps = 2/943 (0%)
 Strand=Plus/Plus

Query  8      AGCC-GAGCCTCTGCCCTGAGGA-GGCTTGTTGTATACTGAGTTGTATCTAGTTGTGGCA  65
              ||||  |||||||||||||||||| ||||||||||||||||||||||||||||||||||
Sbjct  54411  AGCCAGAGCCTCTGCCCTGAGGATGGCTTGTTGTATACTGAGTTGTATCTAGTTGTGGCA  54470

Query  66     CATGGCTTGGAGTGACCGGCCATCTCTGTCTTCCAGGACTACTTGGACCTTGCGGCGTCC  125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  54471  CATGGCTTGGAGTGACCGGCCATCTCTGTCTTCCAGGACTACTTGGACCTTGCGGCGTCC  54530

Query  126    ACTCCATCTGACTCCCTGATTTATGACGACGGCCTCTCAGAGGAGGAGACACCGCTGGTG  185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  54531  ACTCCATCTGACTCCCTGATTTATGACGACGGCCTCTCAGAGGAGGAGACACCGCTGGTG  54590

Query  186    GACTGTAATAATGCCCCCCTCCCTCGAGCCCTCCCTTCCACATGGATTGAAAACAAACTC  245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  54591  GACTGTAATAATGCCCCCCTCCCTCGAGCCCTCCCTTCCACATGGATTGAAAACAAACTC  54650

Query  246    TATGGTAGAATTTCCCATGCATTTACTAGATTCTAGCACCGCTGTCCCCTCTGCACTATC  305
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  54651  TATGGTAGAATTTCCCATGCATTTACTAGATTCTAGCACCGCTGTCCCCTCTGCACTATC  54710

Query  306    CTTCCTCTCTGTGATGCTTTTTAAAAATGTTTCTGGTCTGAACAAAACCAAAGTCTGCTC  365
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  54711  CTTCCTCTCTGTGATGCTTTTTAAAAATGTTTCTGGTCTGAACAAAACCAAAGTCTGCTC  54770

Query  366    TGAACCTTTTTATTTGTAAATGTCTGACTTTGCATCCAGTTTACATTTAGGCATTATTGC  425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  54771  TGAACCTTTTTATTTGTAAATGTCTGACTTTGCATCCAGTTTACATTTAGGCATTATTGC  54830

Query  426    AACTATGTTTTTCTAAAAGGAAGTGAAAATAAGTGTAATTACCACATTGCCCAGCAACTT  485
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  54831  AACTATGTTTTTCTAAAAGGAAGTGAAAATAAGTGTAATTACCACATTGCCCAGCAACTT  54890

Query  486    AGGATGGTAGAGGAAAAAACAGATCAGGGCGGAACTCTCAGGGGAGACCAAGAACAGGTT  545
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  54891  AGGATGGTAGAGGAAAAAACAGATCAGGGCGGAACTCTCAGGGGAGACCAAGAACAGGTT  54950

Query  546    GAATAAGGCGCTTCTGGGGTGGGAATCAAGTCATAGTACTTCTACTTTAACTAAGTGGAT  605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  54951  GAATAAGGCGCTTCTGGGGTGGGAATCAAGTCATAGTACTTCTACTTTAACTAAGTGGAT  55010

Query  606    AAATATACAAATCTGGGGAGGTATTCAGTTGAGAAAGGAGCCACCAGCACCACTCAGCCT  665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  55011  AAATATACAAATCTGGGGAGGTATTCAGTTGAGAAAGGAGCCACCAGCACCACTCAGCCT  55070

Query  666    GCACTGGGAGCACAGCCAGGTTCCCCCAGACCCCTCCTGGGCAGGCAGGTGCCTCTCAGA  725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  55071  GCACTGGGAGCACAGCCAGGTTCCCCCAGACCCCTCCTGGGCAGGCAGGTGCCTCTCAGA  55130

Query  726    GGCCACCCGGCACTGGCGAGCAGCCACTGGCCAAGCCTCAGCCCCAGTCCCAGCCACATG  785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  55131  GGCCACCCGGCACTGGCGAGCAGCCACTGGCCAAGCCTCAGCCCCAGTCCCAGCCACATG  55190

Query  786    TCCTCCATCAGGGGTAGCGAGGTTGCAGGAGCTGGCTGGCCCTGGGAGGACGCACCCCCA  845
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  55191  TCCTCCATCAGGGGTAGCGAGGTTGCAGGAGCTGGCTGGCCCTGGGAGGACGCACCCCCA  55250

Query  846    CTGCTGTTTTCACATCCTTTCCCTTACCCACCTTCAGGACGGTTGTCACTTATGAAGTCA  905
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  55251  CTGCTGTTTTCACATCCTTTCCCTTACCCACCTTCAGGACGGTTGTCACTTATGAAGTCA  55310
```

NCBI Blast result of pGEMT- RET sequenced with Reverse T7 primer

```
>ref|NG_007489.1| Homo sapiens ret proto-oncogene (RET) on chromosome 10
Length=60283

 Score = 1707 bits (924),  Expect = 0.0
 Identities = 944/952 (99%), Gaps = 7/952 (0%)
 Strand=Plus/Minus

Query  4      ACTATC-AACGTGGTCCATTAATTTTGCCGCTGAGGGTGAAAGCATCCAGTTAGCATATA  62
              ||||||| ||||| ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  56198  ACTATCAAACGT-GTCCATTAATTTTGCCGCTGAGGGTGAAAGCATCCAGTTAGCATATA  56140

Query  63     CACTATCATTTGGATATCTTGGAAACCCAGTGTTAGTGCCATCAGCTCTCGTGAGTGGTA  122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  56139  CACTATCATTTGGATATCTTGGAAACCCAGTGTTAGTGCCATCAGCTCTCGTGAGTGGTA  56080

Query  123    CAGGACTCTCTCCAGGCCAGTTCGGGTCTGACATGCCTAAAAATGAAAACAGATCATTTG  182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  56079  CAGGACTCTCTCCAGGCCAGTTCGGGTCTGACATGCCTAAAAATGAAAACAGATCATTTG  56020

Query  183    TTCTGCACTGAAGAACCAAAACTTCAAGTGCAGACAGTAAGGCCTTGGCAAAACTCCCTT  242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  56019  TTCTGCACTGAAGAACCAAAACTTCAAGTGCAGACAGTAAGGCCTTGGCAAAACTCCCTT  55960

Query  243    TGATGTTCAAACCAAACTCGTGGTTTCTGTGCCAATTATATTAATGGTTTTGTGTTCCAG  302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  55959  TGATGTTCAAACCAAACTCGTGGTTTCTGTGCCAATTATATTAATGGTTTTGTGTTCCAG  55900

Query  303    GTTTCCAAAAACACTGTTATAGTTAGGGGAAATATCATGTGCTCTCTTGGTAAATGTAAC  362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  55899  GTTTCCAAAAACACTGTTATAGTTAGGGGAAATATCATGTGCTCTCTTGGTAAATGTAAC  55840

Query  363    TTAACAGACACTACCAATTTTGACTTTTATGAACTTGAGACTCTCTAATCTGATGTATAT  422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  55839  TTAACAGACACTACCAATTTTGACTTTTATGAACTTGAGACTCTCTAATCTGATGTATAT  55780

Query  423    TTGTCGAACTGCTTAACTCTTTTTCTAAGGCAGTGTAGGGCAACAGCACACTGCTGCACT  482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  55779  TTGTCGAACTGCTTAACTCTTTTTCTAAGGCAGTGTAGGGCAACAGCACACTGCTGCACT  55720

Query  483    CAATGGCTTACTGCAGCGGAAACGTTCTATAAAACATTCTCTAACGTAAAGGAGGCACTC  542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  55719  CAATGGCTTACTGCAGCGGAAACGTTCTATAAAACATTCTCTAACGTAAAGGAGGCACTC  55660

Query  543    GTCTGTTGTCTGAGCCTACAACGGAGAAGAGGACCTTCCAAATCGAGGGTGCACACGGGC  602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  55659  GTCTGTTGTCTGAGCCTACAACGGAGAAGAGGACCTTCCAAATCGAGGGTGCACACGGGC  55600

Query  603    CAGCCTCATAGCATCGTCAGCACAAATGGAGTGTGGCACTAGACACAAAGCCAGTGAGGA  662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  55599  CAGCCTCATAGCATCGTCAGCACAAATGGAGTGTGGCACTAGACACAAAGCCAGTGAGGA  55540

Query  663    GTCAGTGAGTCAGGAGGCATGCAAAAGGAACCGTCTCAGCTTGaaaaaaaTGCTTGAAAC  722
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  55539  GTCAGTGAGTCAGGAGGCATGCAAAAGGAACCGTCTCAGCTTGAAAAAAATGCTTGAAAC  55480

Query  723    TCCAAGATTTCAAACAATCAATGATGAAGTTAGGAAACAGAAAAAATTTCTAGTATTTAA  782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  55479  TCCAAGATTTCAAACAATCAATGATGAAGTTAGGAAACAGAAAAAATTTCTAGTATTTAA  55420

Query  783    TCGAAGACAAAAGAAGGAGTTTTGGCAAGAACCATATTTACTGTCCATTGTAAGACCACA  842
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  55419  TCAAAGACAAAAGAAGGAGTTTTGGCAAGAACCATATTTACTGTCCATTGTAAGACCACA  55360

Query  843    GCACCACAGACCATGTTCTTTCAAAAAGCAACTGCTCCAGCTTTAGCACTGACTTCATAA  902
```

Sequences producing significant alignments:

Accession	Description	Max score	Total score	Query coverage	E value	Max ident	Links
Transcripts							
NM_020630.4	Homo sapiens ret proto-oncogene (RET), transcript variant	1565	1565	89%	0.0	100%	GM
NM_020975.4	Homo sapiens ret proto-oncogene (RET), transcript variant	276	276	15%	2e-71	100%	GM
Genomic sequences [show first]							
NW_001837940.1	Homo sapiens chromosome 10 genomic contig, alternate a	1729	1729	99%	0.0	99%	
NT_033985.7	Homo sapiens chromosome 10 genomic contig, GRCh37 re	1718	1718	99%	0.0	99%	

A large scale prep of the cloning vector pGEMT containing the insert was performed to generate it in large amounts and make it more purified. Its DNA concentration (250μg/ml) and A260/A280 ratio of 0.099/0.062=1.6 were also determined.

3. Creation of RET-pFLAG-CMV-1 constructs-

To express the amplified product into mammalian cells it is necessary to insert it into a mammalian expression vector. pFLAG-CMV-2 was chosen for the purpose as it is a shuttle vector for E. coli and mammalian cells of 4.7 kb in size, used for expression and secretion of N-terminal FLAG® fusion proteins in the mammalian cells. The promoter-regulatory region of the human cytomegalovirus2 drives transcription of FLAG-fusion constructs. It also results in most efficient replication in Vero cells. A large scale prep of the pFlag plasmid was performed and its concentration (375μg/ml) was measured by a spectrophotometer using A260/A280 ratio as 0.15/.082= 1.8.

Three common restriction digestion sites (ECOR1, ECORV, NOT1 and SAL1) between the two vectors, pFLAG and pGEMT were determined in such a way that they are not common with the restriction sites present within the insert using NEB cutters from Biolabs. The aim was to cut the insert from pGEMT vector and ligate it with pFLAG-CMV-1. Different

methods such as In-gel ligation, Ethanol precipitation, Addition of ECOR1 site to the insert DNA, purification using column was performed. In case of in gel ligation the ECOR1 & SAL1 digested plasmids were precipitated using ethanol and sodium acetate and further ligated using T4DNA ligase for an hour using 1:3 vector to insert ratio. The insert was again amplified from the cloning vector pGEMT by adding ECOR1 sites. The PCR amplified insert with added ECOR1 restriction site was run in a 0.8% agarose gel and it was extracted. Then it was ligated with ECOR1 digested pFLAG vector. And finally success was achieved using column purification and NOT1 and SAL1 restriction enzymes. The plasmids were both double-digested with NOT1& SAL1 and single digested with SAL1 for an hour. The digested regions of the pFLAG plasmid and the 1910bp insert were then visualized by running on a gel. The digested products were further purified using purification columns from Sigma. The purified product was then ligated in different vector: insert ratio (1:3, 1:6, 1:10, and 1:15) using T4 DNA ligase for 1hr. The ligated products were plated into LB/amp plates and 50, 50, 100 colonies were obtained after an overnight incubation at 37°c. The earlier ligation reactions had not shown any colonies except the last one which formed colonies at different vector to insert ratio ligation ranging from approx.50-500.

8 colonies were picked from the plates and Miniprep was performed. The plasmid Miniprep samples were further digested with Not1 and Sal1 restriction enzymes. After running on a 1% agarose gel two samples were assumed to have the ligated vector plus insert.

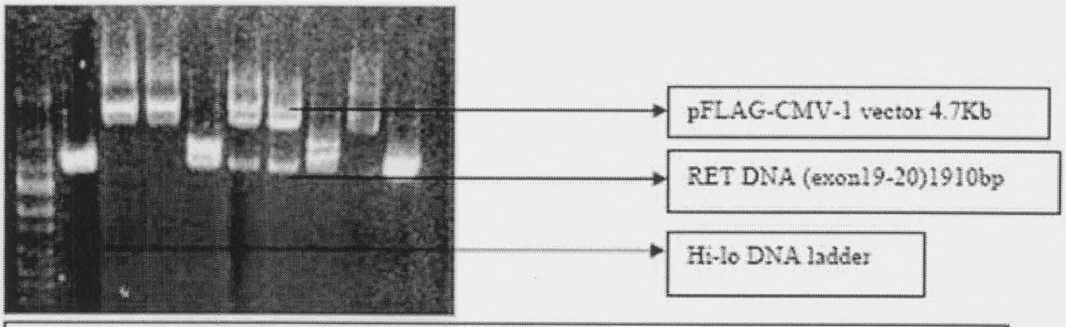


Figure6: Restriction enzyme NOT1&SAL1digestion of the plasmid (pFLAG-CMV-1)

These two samples were further digested with ECORV; which cuts the insert at its extreme end and the vector just before its multicloning site to confirm their orientations. One with almost 2kb insert and approx. 5 kb vector proving its anticipated forward orientation and the other with approx 6 kb fragment and approx.200 bp fragment indicating its inverted orientation were obtained.

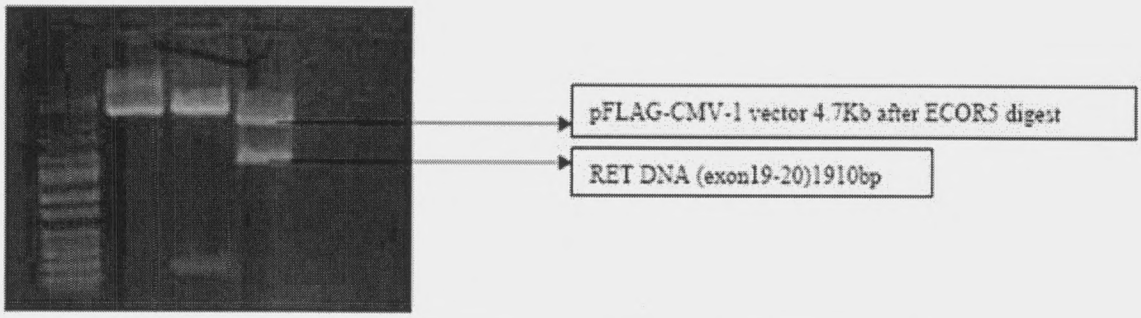
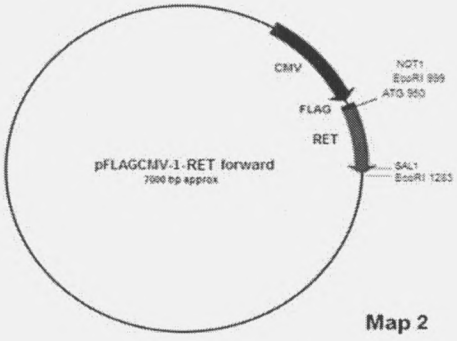
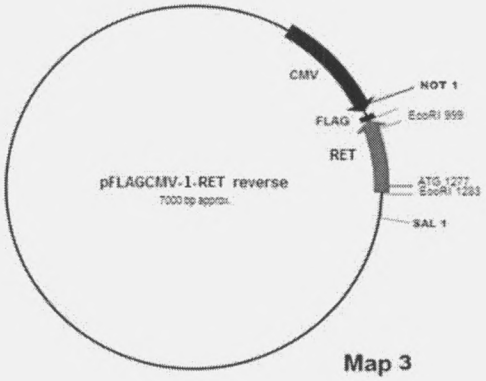


Figure7: Restriction enzyme ECOR5 digestion of the plasmid (pFLAG-CMV1)



Map 2

Figure 8: The map showing the Forward orientation of the amplified product inside pFlagCMV-1 gene.



Map 3

Figure 9: The map showing the reverse orientation of the amplified product inside pFlagCMV-1 gene.

The samples were sent for sequencing at GENEWIZ to further confirm their orientation and presence of insert. The sequencing result confirmed the presence of the insert (RET) ligated with pFLAG. A large scale prep of the new pFLAG-RET was performed to purify it so that it can be transfected into mammalian cells.

NCBI Blast showing RET construct with forward orientation within pFLAG-CMV vector

```
>ref|NG_007489.1| D Homo sapiens ret proto-oncogene (RET) on chromosome 10
Length=60283

 Score = 1594 bits (863),  Expect = 0.0
 Identities = 884/902 (98%), Gaps = 2/902 (0%)
 Strand=Plus/Plus

Query  148    TCTCTTGGAGAGGTCAGGAGATTGGTGCGAGGTGGAGACACAGCCAGAGCCTCTGCCCTG  207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  54370  TCTCTTGGAGAGGTCAGGAGATTGGTGCGAGGTGGAGACACAGCCAGAGCCTCTGCCCTG  54429

Query  208    AGGATGGCTTGTTGTATACTGAGTTGTATCTAGTTGTGGCACATGGCTTGGAGTGACCGG  267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  54430  AGGATGGCTTGTTGTATACTGAGTTGTATCTAGTTGTGGCACATGGCTTGGAGTGACCGG  54489

Query  268    CCATCTCTGTCTTCCAGGACTACTTGGACCTTGCGGCGTCCACTCCATCTGACTCCCTGA  327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  54490  CCATCTCTGTCTTCCAGGACTACTTGGACCTTGCGGCGTCCACTCCATCTGACTCCCTGA  54549

Query  328    TTTATGACGACGGCCTCTCAGAGGAGGAGACACCGCTGGTGGACTGTAATAATGCCCCCC  387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  54550  TTTATGACGACGGCCTCTCAGAGGAGGAGACACCGCTGGTGGACTGTAATAATGCCCCCC  54609

Query  388    TCCCTCGAGCCCTCCCTTCCACATGGATTGAAAACAAACTCTATGGTAGAATTTCCCATG  447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  54610  TCCCTCGAGCCCTCCCTTCCACATGGATTGAAAACAAACTCTATGGTAGAATTTCCCATG  54669

Query  448    CATTTACTAGATTCTAGCACCGCTGTCCCCTCTGCACTATCCTTCCTCTCTGTGATGCTT  507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  54670  CATTTACTAGATTCTAGCACCGCTGTCCCCTCTGCACTATCCTTCCTCTCTGTGATGCTT  54729

Query  508    TTTAAAAATGTTTCTGGTCTGAACAAAACCAAAGTCTGCTCTGAACCTTTTTATTTGTAA  567
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  54730  TTTAAAAATGTTTCTGGTCTGAACAAAACCAAAGTCTGCTCTGAACCTTTTTATTTGTAA  54789

Query  568    ATGTCTGACTTTGCATCCAGTTTACATTTAGGCATTATTGCAACTATGTTTTTCTAAAAG  627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  54790  ATGTCTGACTTTGCATCCAGTTTACATTTAGGCATTATTGCAACTATGTTTTTCTAAAAG  54849

Query  628    GAAGTGAAAATAAGTGTAATTACCACATTGCCCAGCAACTTAGGATGGTAGAGGAAAAAA  687
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  54850  GAAGTGAAAATAAGTGTAATTACCACATTGCCCAGCAACTTAGGATGGTAGAGGAAAAAA  54909

Query  688    CAGATCAGGGCGGAACTCTCANGGGAGACCAAGAACAGGTTGAATAAGGCGCTTCTGGGG  747
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Sbjct  54910  CAGATCAGGGCGGAACTCTCAGGGGAGACCAAGAACAGGTTGAATAAGGCGCTTCTGGGG  54969

Query  748    TGGGAATCAAGTCATAGTACTTCTACTTTAACTAAGTGGATAAATATACAAATCTGGGGA  807
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  54970  TGGGAATCAAGTCATAGTACTTCTACTTTAACTAAGTGGATAAATATACAAATCTGGGGA  55029

Query  808    GGTATTCAGTTGAGAAAGGANCCACCAGCACCACTCAGCCTGCACTGGGANCACAGCCAG  867
              |||||||||||||||||||| |||||||||||||||||||||||||| ||||||||||||
Sbjct  55030  GGTATTCAGTTGAGAAAGGAGCCACCAGCACCACTCAGCCTGCACTGGGAGCACAGCCAG  55089

Query  868    GTTCCCCCAGACCCCTCCTGGGCAGGCAGGTGCCTCTCANAGGNCACCCGGCACTGGCGA  927
              |||||||||||||||||||||||||||||||||||||||| ||| ||||||||||||||||
Sbjct  55090  GTTCCCCCAGACCCCTCCTGGGCAGGCAGGTGCCTCTCAGAGGCCACCCGGCACTGGCGA  55149

Query  928    GCAGCCACTGGCCAAGCCTCAGCCCCAGTCCCAGCCACNTGTCCTCCATCNNGG-TANCG  986
              |||||||||||||||||||||||||||||||||||||| |||||||||| || || ||||
Sbjct  55150  GCAGCCACTGGCCAAGCCTCAGCCCCAGTCCCAGCCACATGTCCTCCATCAGGGGTAGCG  55209

Query  987    AGGTTGCNGNANCTGGCTGGNCCTGGGAGGANNCACCCCCACTGCTGTTT-CACNTCCTT  1045
              ||||||| | | |||||||| |||||||||  |||||||||||||||| ||| |||||
Sbjct  55210  AGGTTGCAGGAGCTGGCTGGCCCTGGGAGGACGCACCCCCACTGCTGTTTTCACATCCTT  55269

Query  1046   TC  1047
              ||
Sbjct  55270  TC  55271
```

NCBI Blast data showing the RET construct with reverse orientation within pFlag-CMV vector

```
>ref|NG_007489.1| D Homo sapiens ret proto-oncogene (RET) on chromosome 10
Length=60283

 Score = 1640 bits (888),  Expect = 0.0
 Identities = 934/967 (96%), Gaps = 7/967 (0%)
 Strand=Plus/Minus

Query  145    TTCCCCTTGTGAGTCCATTACCTTTCACAAAGAAATGTTAACTATCAAACGTGTCCATTA  204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  56238  TTCCCCTTGTGAGTCCATTACCTTTCACAAAGAAATGTTAACTATCAAACGTGTCCATTA  56179

Query  205    ATTTTGCCGCTGAGGGTGAAAGCATCCAGTTAGCATATACACTATCATTTGGATATCTTG  264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  56178  ATTTTGCCGCTGAGGGTGAAAGCATCCAGTTAGCATATACACTATCATTTGGATATCTTG  56119

Query  265    GAAACCCAGTGTTAGTGCCATCAGCTCTCGTGAGTGGTACAGGACTCTCTCCAGGCCAGT  324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  56118  GAAACCCAGTGTTAGTGCCATCAGCTCTCGTGAGTGGTACAGGACTCTCTCCAGGCCAGT  56059

Query  325    TCGGGTCTGACATGCCTAAAAATGAAAACAGATCATTTGTTCTGCACTGAAGAACCAAAA  384
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  56058  TCGGGTCTGACATGCCTAAAAATGAAAACAGATCATTTGTTCTGCACTGAAGAACCAAAA  55999

Query  385    CTTCAAGTGCAGACAGTAAGGCCTTGGCAAAACTCCCTTTGATGTTCAAACCAAACTCGT  444
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  55998  CTTCAAGTGCAGACAGTAAGGCCTTGGCAAAACTCCCTTTGATGTTCAAACCAAACTCGT  55939

Query  445    GGTTTCTGTGCCAATTATATTAATGGTTTTGTGTTCCAGGTTTCCAAAAACACTGTTATA  504
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  55938  GGTTTCTGTGCCAATTATATTAATGGTTTTGTGTTCCAGGTTTCCAAAAACACTGTTATA  55879

Query  505    GTTAGGGGAAATATCATGTGCTCTCTTGGTAAATGTAACTTAACAGACACTACCAATTTT  564
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  55878  GTTAGGGGAAATATCATGTGCTCTCTTGGTAAATGTAACTTAACAGACACTACCAATTTT  55819

Query  565    GACTTTTATGAACTTGAGACTCTCTAATCTGATGTATATTTGTCGAACTGCTTAACTCTT  624
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  55818  GACTTTTATGAACTTGAGACTCTCTAATCTGATGTATATTTGTCGAACTGCTTAACTCTT  55759

Query  625    TTTCTAAGGCAGTGTAGGGCAACAGCACACTGCTGCACTCAATGGCTTACTGCAGCGGAA  684
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  55758  TTTCTAAGGCAGTGTAGGGCAACAGCACACTGCTGCACTCAATGGCTTACTGCAGCGGAA  55699

Query  685    ACGTTCTATAAAACATTCTCTAACGTAAAGGAGGCACTCGTCTGTTGTCTGAGCCTACAA  744
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  55698  ACGTTCTATAAAACATTCTCTAACGTAAAGGAGGCACTCGTCTGTTGTCTGAGCCTACAA  55639

Query  745    CGGAGAAGAGGACCTTCCAAATCGAGGGTGCACACGGGCCAGCCTCATAGCATCGTCAGC  804
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  55638  CGGAGAAGAGGACCTTCCAAATCGAGGGTGCACACGGGCCAGCCTCATAGCATCGTCAGC  55579

Query  805    ACAAATGGAGTGTGGCACTAGACACAAAGCCAGTGAGGAGTCAGTGAGTCAGGAGGCATG  864
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  55578  ACAAATGGAGTGTGGCACTAGACACAAAGCCAGTGAGGAGTCAGTGAGTCAGGAGGCATG  55519

Query  865    CAAAAGGAACCGTCTCAGCTTGAAAAAA-TGCTTGAAACTCCAAGATTTCNAACA-TCAA  922
              |||||||||||||||||||||||||||| |||||||||||||||||||| |||| ||||
Sbjct  55518  CAAAAGGAACCGTCTCAGCTTGAAAAAAATGCTTGAAACTCCAAGATTTCAAACAATCAA  55459

Query  923    TGATGAAGTTAGGAAACAGAAAAANTTTCTAGTATTTAATCNAANACAAAAGAAGNAGTT  982
              |||||||||||||||||||||||| |||||||||||||| || |||||||||| ||||
Sbjct  55458  TGATGAAGTTAGGAAACAGAAAAAATTTCTAGTATTTAATCAAAGACAAAAGAAGGAGTT  55399
```

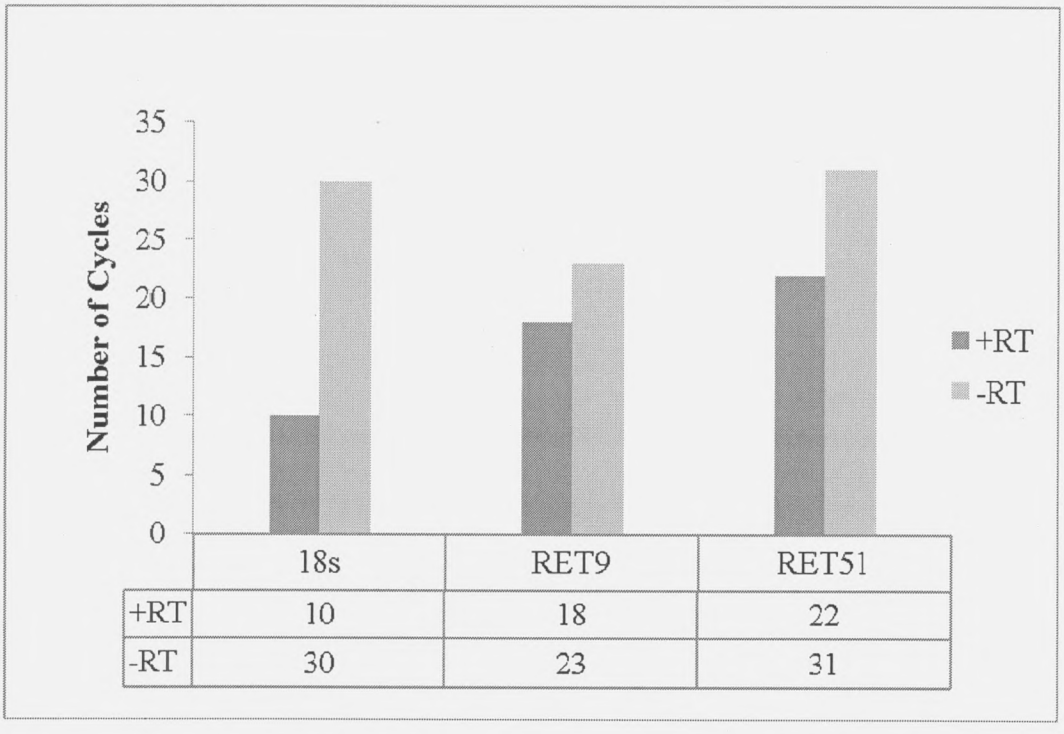
Sequences producing significant alignments:

Accession	Description	Max score	Total score	Query coverage	E value	Max ident	Links
Transcripts							
NM_020630.4	Homo sapiens ret proto-oncogene (RET), transcript variant	1343	1343	63%	0.0	97%	GM
NM_020975.4	Homo sapiens ret proto-oncogene (RET), transcript variant	276	276	12%	2e-71	100%	GM
Genomic sequences [show first]							
NW_001837940.1	Homo sapiens chromosome 10 genomic contig, alternate a	1596	1596	74%	0.0	98%	
NT_033985.7	Homo sapiens chromosome 10 genomic contig, GRCh37 re	1585	1585	74%	0.0	97%	

4. RNA analysis-

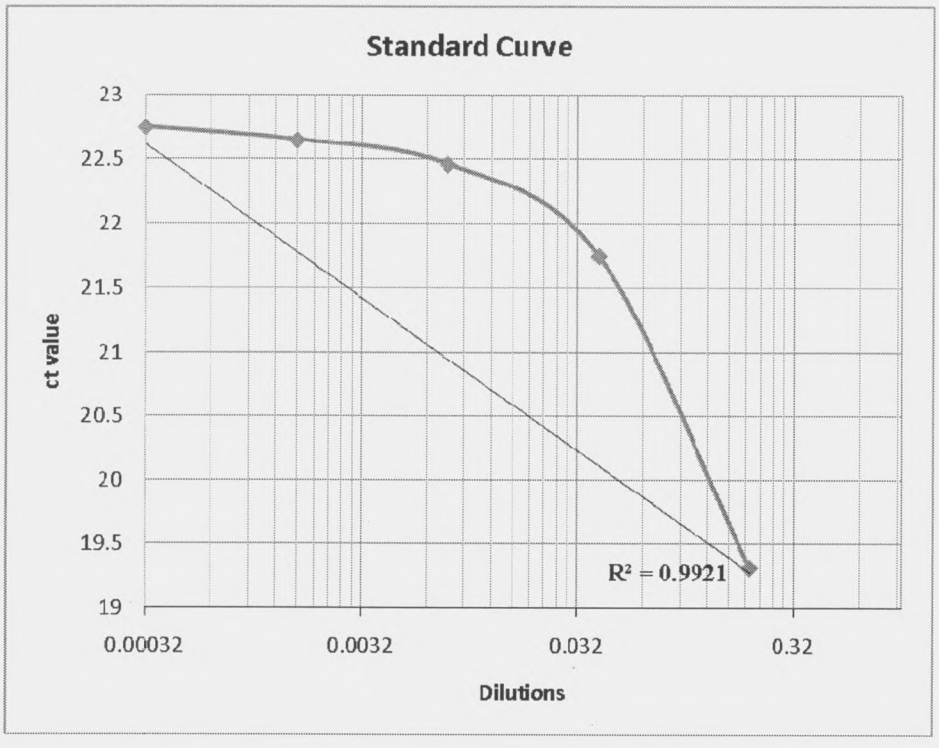
The purified pFLAG-RET constructs both forward and reverse orientation of 797µg/ml (reverse) and 1087 µg/ml(forward) were transfected into Vero cells. The total RNA was extracted and the primers were designed to amplify RET9 and RET51 RNA.

To be used as a positive control, RNA (RAGD72 1mg/ml) from neuronal cells were used. The first set of primer pairs designed to amplify RET9 and RET51 were tested with this RNA. The graph below represents the quantitative expression of RET isoforms in neuronal RNA with Invitrogen superscript reverse transcriptase - cDNA and SYBR real time pcr reagent. The following graph shows the expression of RET9 and RET51 with and without reverse transcriptase. The dark bars represent the ct value obtained from cDNA without reverse transcriptase and the light bars are for ct values with reverse transcriptase. The first pair of bars represent the control ct value obtained with 18s universal primers. In real time PCR data, lower the number of cycle higher is the expression level of RNA. Thus, it shows that in case of neuronal RNA, the RET9 isoform is expressed in higher level than RET51.



	18s	RET9	RET51
+RT	10	18	22
-RT	30	23	31

Serial dilutions of the positive control RNA were performed to obtain a standard curve for the quantitative PCR, using RET51 primers and neuronal RNA. The graph below represents the expression of the RET51 RNA obtained from 5 fold serial dilution of the cDNA obtained with Invitrogen superscriptII cDNA synthesis kit.



The table below represents the concentration of the dilution series. Dilution was made using a 1:5 ratio of the template cDNA and DI water.

Dilutions	ct Value
0.2	19.32
0.04	21.75
0.008	22.46
0.0016	22.65
0.00032	22.75

The RNA specific primers for meant to amplify RET51 mRNA was further were further used to amplify the construct RNA (total RNA obtained by inserting the constructs into the Vero cells). The amplified product from the RT PCR reaction was sent for sequencing.

Sequencing data from GENEWIZ of RET mRNA amplified with RET51 forward primers

```
>ref|NM_020630.4| GM Homo sapiens ret proto-oncogene (RET), transcript variant 4,
mRNA
Length=4174

 GENE ID: 5979 RET | ret proto-oncogene [Homo sapiens] (Over 100 PubMed links)

                                                          Sort alignments for this sub
                                                            E value  Score  Percent id
                                                            Query start position  Subj

 Score =  560 bits (303),  Expect = 4e-157
 Identities = 516/652 (79%), Gaps = 19/652 (2%)
 Strand=Plus/Plus

Query  150   TTCTTCTNTGAGATGC-TTTT-GAAATGNNATCTGGACTGATCAAGATC-GNGTCTGCTC  206
             ||| ||| || ||||| |||| ||||| |||||  |||| ||| | |  |||||||||
Sbjct  3435  TTCCTCTCTGTGATGCTTTTTAAAAATG-TTTCTGGTCTGAACAAAACCAAAGTCTGCTC  3493

Query  207   TGANCGTTTTTATTTG-AAAGTGTCGGACTTTGCANNCAG-TNACATTTAGGCATTATTA  264
             ||| | |||||||||| ||| |||| |||||||||  ||| | ||||||||||||||||
Sbjct  3494  TGAACCTTTTTATTTGTAAA-TGTCTGACTTTGCATCCAGTTTACATTTAGGCATTATTG  3552

Query  265   CACCTA-GCCNTTTCTAAAAGAAAGTGAAAATAANTGCAATTACCATNTTGCTGAGGATC  323
             || ||| |  ||||||||||| |||||||||||| || |||||||| |||| || | |
Sbjct  3553  CAACTATG-TTTTTCTAAAAGGAAGTGAAAATAAGTGTAATTACCACATTGCCCAGCAAC  3611

Query  324   TTACGATGGTCCAGTAANAAACANATCCCGCCGGAACTCTCAGGCGAGACCAAAAACAGG  383
             ||| ||||||  || || ||||| |||  | ||||||||||||| |||||||| ||||||
Sbjct  3612  TTAGGATGGTAGAGGAAAAAACAGATCAGGGCGGAACTCTCAGGGGAGACCAAGAACAGG  3671

Query  384   TTGAANNANGCGCTTCTGGGGTGGGAAACTANTCATANTACTTCNNCTTTAANC-ANNTG  442
             |||||  | |||||||||||||||||| | | ||||| ||||||  ||||||  |  ||
Sbjct  3672  TTGAATAAGGCGCTTCTGGGGTGGGAATCAAGTCATAGTACTTCTACTTTAA-CTAAGTG  3730

Query  443   NANNAATATACAAAANTGGGGATGTATT-ANNTGAAGAAAGGAACCACCAGCACCACTCA  501
              |  ||||||||||  |||||| ||||| |  || |||||||| ||||||||||||||||
Sbjct  3731  GATAAATATACAAATCTGGGGAGGTATTCAGTTG-AGAAAGGAGCCACCAGCACCACTCA  3789

Query  502   ACCTGCCA-AGNAAACACAACCTGGNTCCCCCNCACCCCTCCNGGGCAGGCNGGTGCCTC  560
               |||| ||  |  | |||| || || ||||||  |||||||| |||||||| ||||||||

Sbjct  3790  GCCTG-CACTGGGAGCACAGCCAGGTTCCCCCAGACCCCTCCTGGGCAGGCAGGTGCCTC  3848

Query  561   TCNNANGCCACCCGGGAGTGGTGAGCANNCNCNGGTCNAA-NCTCANNNNNNA-NCCCAG  618
             || | |||||||||  ||| ||||| | | || | ||  |||        |  ||||||
Sbjct  3849  TCAGAGGCCACCCGGCACTGGCGAGCAGCCACTGG-CCAAGCCTCA-GCCCCAGTCCCAG  3906

Query  619   CCACATGTCCTCCATCNNGNNTAACNAGGTTACNNGAGCTGGCTGGACCTGNGANGACGN  678
             |||||||||||||||  |  || | ||||| |  |||||||||| |||| || ||||
Sbjct  3907  CCACATGTCCTCCATCAGGGGTAGCGAGGTTGCAGGAGCTGGCTGGCCCTGGGAGGACGC  3966

Query  679   NNCNNCACTGNTGTTTTCACATCCTTTCCCTTACCCACCTTCANGACGGTTGTCACTTAT  738
               |  ||||| ||||||||||||||||||||||||||||||||| ||||||||||||||||
Sbjct  3967  ACCCCCACTGCTGTTTTCACATCCTTTCCCTTACCCACCTTCAGGACGGTTGTCACTTAT  4026

Query  739   GAAGTCAGTGCTAAAGCTGGANCANNTGCTTTTNNAAAGAACATGNTCTGTG  790
             |||||||||||||||||| ||  |||||||  |||||||||| ||||||
Sbjct  4027  GAAGTCAGTGCTAAAGCTGGAGCAGTTGCTTTTTGAAAGAACATGGTCTGTG  4078
```

<u>Sequencing data from GENEWIZ of RET mRNA amplified with RET51 reverse primers</u>

```
>ref|NW_001837940.1| D Homo sapiens chromosome 10 genomic contig, alternate assembly
(based on HuRef), whole genome shotgun sequence
Length=3646701

                                                  Sort alignments for this subject s
                                                    E value  Score  Percent identit
                                                    Query start position  Subject s

 Features in this part of subject sequence:
   ret proto-oncogene isoform a

 Score =  246 bits (272),  Expect = 3e-62
 Identities = 422/622 (67%), Gaps = 18/622 (2%)
 Strand=Plus/Minus

Query  172      TATCACGTGCTCTCCTGATGAATGCAAATTA-CANAAATCTACGAATTTTGACTTTCCTG  230
                ||||| |||||||| || | |||| || ||| || | | |||| ||||||||||||  ||
Sbjct  1051198  TATCATGTGCTCTCTTGGTAAATGTAACTTAACAGACA-CTACCAATTTTGACTTTTATG  1051140

Query  231      AACTTGGTNCTCGNCNNTCTGATGAATATTGANNNAACTGCTTANGTCTNTTTCTNNNTT  290
                ||||||   |||     ||||||| |||||     |||||||||  ||| |||||
Sbjct  1051139  AACTTGAGACTCTCTAATCTGATGTATATTTGTCGAACTGCTTAACTCTTTTTCTAAGGC  1051080

Query  291      AGGGAAGGAAAA-AGCTACAGTGTTGAATTGATAGTTTTAATGCANCTNNCATGGTCGAT  349
                || | |||  || ||| ||| || || | | |  |  ||| |||| |    | | || ||
Sbjct  1051079  AGTGTAGGGCAACAGC-ACACTGCTGCACTCAATGGCTTACTGCAGCGGAAACGTTCTAT  1051021

Query  350      ACATGATNNACTAGCGCATAGGAGGTCATCGACTGA-GNNTGAG--TACGGAAAGGTCNN  406
                | |  ||   ||| || | ||||||   ||| |||  |  ||||  |||   || |
Sbjct  1051020  AAAACATTCTCTAACGTAAAGGAGGCACTCGTCTGTTGTCTGAGCCTAC---AACGGAGA  1050964

Query  407      NGANCACCTTG-AANTCGAGGGTGN-CTTGGACCCNNCNCATACCATCNTCANCACTAAT  464
                 ||  |||||  || |||||||||  |  || ||   | |||| |||| ||| ||| |||
Sbjct  1050963  AGAGGACCTTCCAAATCGAGGGTGCACACGGGCCAGCCTCATAGCATCGTCAGCACAAAT  1050904

Query  465      GAAATGTGGCACTAGANACAAANCCNGTGAGGAGCCAATGACTCCGGAGCCATGCNTGAN  524
                | | |||||||||||| ||||| || ||||||||| || || || || ||||| |||||   | 
Sbjct  1050903  GGAGTGTGGCACTAGACACAAAGCCAGTGAGGAGTCAGTGAGTCAGGAGGCATGCAAAAG  1050844

Query  525      GAACCCTCTNAGCTCGAANATAATGNTTGAANCTCCNNNANTTCCNCCAATCTANGATGA  584
                ||||| ||| |||| ||| | |||| ||||| ||||   |  |||   ||||| | |||||
Sbjct  1050843  GAACCGTCTCAGCTTGAAAAAAATGCTTGAAACTCCAAGATTTCAAACAATCAATGATGA  1050784

Query  585      ANTTNGGACACAGANNTTATTTCNAGTATTTAATCAGNNNAC-NGNGANNGANTCTTGGC  643
                | || ||| |||||    ||||| ||||||||||| |   ||    ||  || | |||||
Sbjct  1050783  AGTTAGGAAACAGAAAAAATTTCTAGTATTTAATC-GAAGACAAAAGAAGGAGTTTTGGC  1050725

Query  644      TAGAACCATATATAGNGTCCNNTGANANANCACAGCACNNCGNACNGTGCTCTTTCaaaa  703
                 |||||||||| || ||||  ||  | | ||||||||  |  ||  || ||||||||||
Sbjct  1050724  AAGAACCATATTTACTGTCCATTGTAAGACCACAGCACCACAGACCATGTTCTTTCAAAA  1050665

Query  704      nnnCATCTGCTCCAACTTATANCTNCTGACNTGNTNNGTGACAATCNNCCTGAACGTNGT  763
                   || |||||||| ||| || |  ||||| |  |  ||||||| |  ||||||| || |
Sbjct  1050664  -AGCAACTGCTCCAGCTT-TAGC-ACTGACTTCATAAGTGACAACCGTCCTGAAGGTGGG  1050608

Query  764      AATGNGANNNNGATGTGANAAC  785
                 | | ||    ||||||| |||
Sbjct  1050607  TAAGGGA-AAGGATGTGAAAAC  1050587
```

Sequences producing significant alignments:

Accession	Description	Max score	Total score	Query coverage	E value	Max ident	Links
Transcripts							
NM_020630.4	Homo sapiens ret proto-oncogene (RET), transcript variant	560	703	81%	4e-157	97%	GM
NM_020975.4	Homo sapiens ret proto-oncogene (RET), transcript variant	200	200	12%	1e-48	98%	GM
Genomic sequences [show first]							
NW_001837940.1	Homo sapiens chromosome 10 genomic contig, alternate a	560	766	85%	4e-157	100%	
NT_033985.7	Homo sapiens chromosome 10 genomic contig, GRCh37 re	555	760	85%	2e-155	100%	

The sequencing result showed the presence of introns in the amplified region and the amplified region was 1500bp long containing exon19 and intron19. The same reaction was performed with the neuronal RNA and a similar result was obtained. A new primer set was designed to verify the presence of unspliced RNA product in the neuronal cells RNA and the mRNA was purified from the total RNA. A new transfection with the purified pFLAG-RET constructs both forward and reverse orientation of 797µg/ml (reverse) and 1087 µg/ml(forward) was further transfected into Vero cells.

The amplification of the construct RNA, neuronal RNA obtained from SK-N-SH cells, mRNA (extracted from neuronal RNA) with all the different sets of primers showed the presence of two bands in 1% agarose gel. The size of the bands was approx 1500 bp and 150 bp respectively.

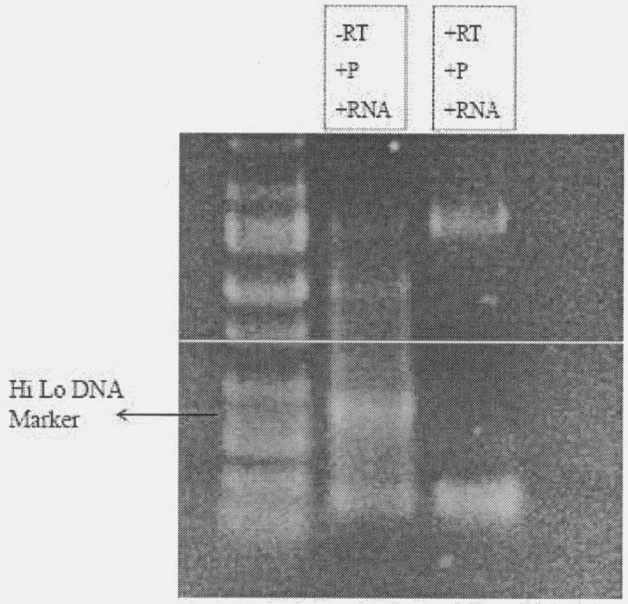
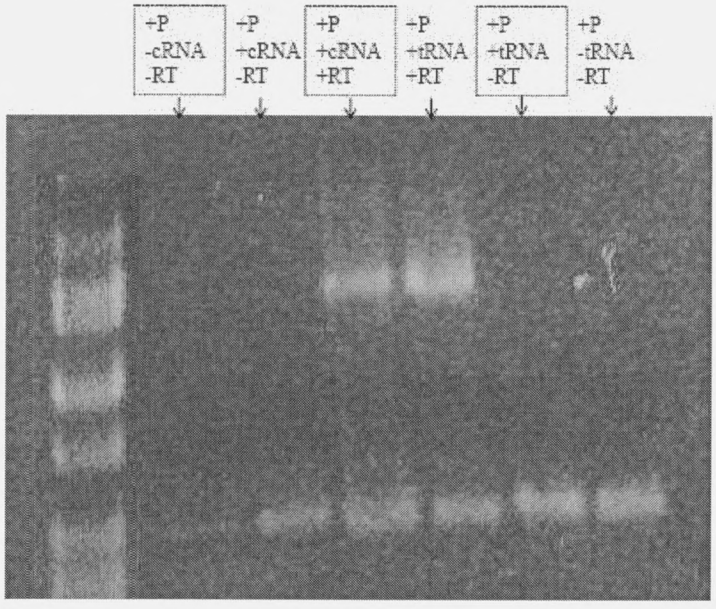


Figure 12: mRNA amplification of the construct using RT

Amplification of neuronal RNA and RET RNA with RNA primers



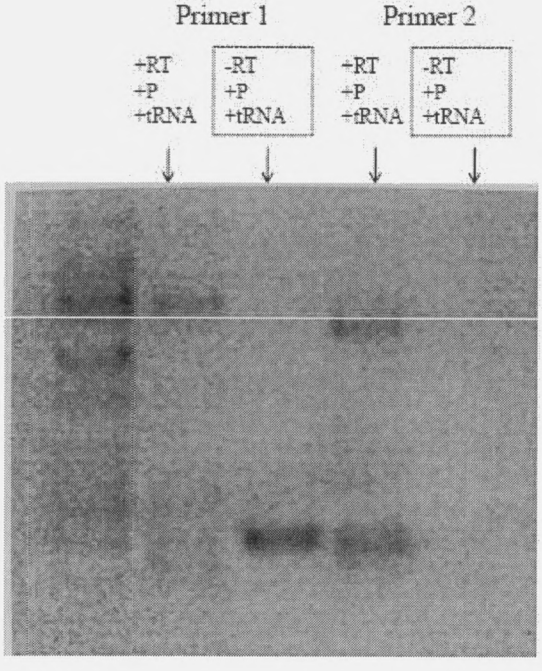


Figure 13: Amplification of mRNA using two sets of exon splice junction primers and construct RET

DISCUSSION

Alternative splicing in the gene RET has resulted in the formation of its two isoforms; RET9 and RET51. These two isoforms are expressed differentially in mammalian cells and they perform important roles in the development of kidney and enteric nervous system. They differ in their carboxy- terminal tails with respect to the number of amino acids. This difference has in turn resulted in the difference of signaling properties in between the two isoforms. [13] The ligand GDNF also plays an important role in RET signaling along with the coreceptor GFRα1. The aim of this research is to investigate the alternative splicing mechanism in RET protooncogene. Alternative splicing has been associated with various diseases like cystic fibrosis, muscular atrophy, Parkinson's disease including various forms of cancer. There are many examples with transmembrane proteins where among the splice variants generated by alternative splicing, one form is up- regulated in cancer than the other ones[39]. The protein RET is a transmembrane receptor protein and is essential for transmitting growth signals from the extracellular matrix. The diseases like thyroid cancer, MEN2 syndromes, Hirschsprung's disease has been associated with the RET protooncogene. Recent studies have indicated that the two isoforms of RET have their individual physiological function in signal transduction pathways mediated by the gene RET. It has also been indicated that there may be certain connections between the expression of these two isoforms and MEN2 syndromes.[40] To investigated further on the splicing mechanism of the gene RET, it thus becomes necessary to design a construct of the RET proto-oncogene which contains the specific region where splicing event takes place in a RET gene.

In this study a construct of RET proto-oncogene consisting of the region specific for splicing event of RET proto-oncogene was constructed. The created construct has approx. 2000 bp of

RET gene starting from the last 132bp of intron18 and ends in the open reading frame (amino acid encoding region) of exon20. This region constitutes of exactly 1910 base pairs. The splicing event of RET occurs in this region for its two isoforms. The RET9 mRNA contains the 152 base pair long exon19 and about 780bp of intron19 (NCBI ref. No. NM_020630). It has an open reading frame until 32bp of intron19. RET 51 mRNA contains both exon19 and exon20 (NCBI ref. No. NM_020975) and has an open reading frame until – 258bp in exon20. The primers were designed to amplify these regions for the creation of the construct. The amplified region had both exon19 and intron19 which forms the end portion of the 9 amino acid tail of RET9 isoform. Human exon19 is 152 bp long and has Y1062 which is the last common tyrosine residue among all the 3 isoforms of RET (RET9, RET51 and RET43). Y1062 is also located in the end part of this exon. Carter et al (2000) have reported that Y1062 lies directly upstream of the splice donor site of the exon19.[22] The amplified region also contained the open reading frame of exon20 which also included the two tyrosines Y1090 and Y1096.

The 1910 base pairs amplified region of RET was further inserted in the mammalian expression vector pFLAG to ease its mRNA expression in mammalian cells. During the creation of pFLAG RET construct, two plasmids (one with forward and other with reverse orientation of RET) were seen in the agarose gel and confirmed by sequencing. These two forward and reverse orientation constructs were further transfected into Vero cells. It was expected that the spliceosome machinery would not act on the construct with reverse orientation and there would be no splicing event. Pasman et al (1996) has shown that the spliceosome machinery works as a three dimensional diffusion model where functional interaction between complexes (assembled independently at splice sites) occurs irrespective

of the presence of introns between them.[36] Even if the splicing event occurs, the correct order of the exons and introns will not be maintained in the two isoforms of RET generated by the reverse construct. Vanacova et al. (2005) had used reverse orientation of the insert as negative control in their research involving the role of introns in splicing mechanism.[35]

The Vero cell line used was initiated from the kidney of the African green monkey and was expected to express the insert RET efficiently after transfection. The real time quantitative PCR (Q-PCR) was chosen for validating the mechanism of splicing and formation of two isoforms RET9 and RET51 in the construct. Real time PCR refers to a process through which the amplification of products are directly monitored in each amplification cycle.[37]

The quantification of mRNA can be efficiently performed by using inexpensive DNA binding dye such as SYBR Green 1 which can be used in real time PCR assays without individual probe design and optimization. The dye emits fluorescence when bound to the dsDNA in a sequence independent manner. Thus representing fluorescence intensities proportional to DNA concentrations.[37] To perform the Q-PCR and validate splicing event, primers were designed in the intron-exon boundaries of RET construct. A primer set spanning intron19 was also designed to avoid the potential signals from genomic DNA contamination and determining unspliced RET gene. The splice site specific primers were expected to allow differentiation of the alternatively spliced regions of the RET construct.

To design a positive control probe for the constructs, initially the real time PCR was performed with the cDNA from neuronal cells mRNA that expresses RET. The primers designed for both RET9 and RET51 were tested on this cDNA and the amplification plots were noted. The experiment was repeated multiple times using different reverse transcriptase

kits for cDNA preparation. The same SYBR green kit was used every time to validate the similar amplification plots in Strategene MaxPro 3000 machine.

A standard curve with cDNA from neuronal cells and the designed primers was made. The standard curve was generated from (5 fold) serial dilutions of the neuronal cDNA. The relative concentrations were expressed in arbitrary units and the logarithms of concentrations were plotted against cross points. The least square fit was calculated and used as standard curve.[38] A standard curve is required for accurate quantification of gene expression levels. The data obtained from standard curve can be used to compare the abundance of the known RET9 and RET51 isoforms in the neuronal cells with the unknown quantity of the isoforms from the mRNA of the generated construct.

The primers were then used to amplify the RNA obtained from transfection of the constructs into Vero cells. The reverse transcriptase PCR of the total RNA extracted from Vero cells was sent for sequencing. The sequencing result showed that the RNA contained approx. 1000 bp of RET RNA which included the exon19, intron19 and the beginning of exon20; instead of the 150bp region of spliced exon19 and exon20. To further investigate the unexpected result; again three exon splice junction specific primers were designed. Futcher et al have reported that the study of splicing event generated by genes which posses isoforms that retain introns are complicated because RT-PCR performed to amplify mRNA with retained introns, can amplify genomic DNA instead. They also reported that if a gene has an antisense transcript, that will never be spliced and will yield the same product as an unspliced sense transcript.[41] Until now there has been no report of the presence of antisense transcript in RET; hence 1500bp region amplified by the RT PCR may have occurred due to the presence of genomic DNA in the extracted RNA of the constructs. mRNA was further purified from

the extracted total RNA of the construct but due to low concentration of the obtained mRNA, the result was obscure. Hence, further purification of the RNA is required to confirm occurrence of the splicing event in the created construct.

Earlier studies with splicing of different genes have indicated that a gene's promoter region (especially the region that regulates gene expression) can also regulate splicing. Hence splicing can depend on the transcription factor that is bound to the promoter. Futcher et al in their review article has reported that even though splicing can be mediated by signals inside the transcript which undergoes splicing, there can be outer signals too which induces splicing mechanism in certain genes.[42] The created constructs can thus be used to study whether the ligand GDNF plays an important role in the splicing mechanism of RET. Apart from different regulators of the splicing machinery; the extracellular signaling from factors such as ligands has been reported to play important role in the splicing events of genes.[26] Takahashi in 2001 also reported the importance of GDNF in RET signaling and their prevalence in diseases like MEN2A/ FMTC and HSCR.[19]

The construct can further be used to reveal the factor/ factors which direct the formation of the RET9 and RET51 isoforms in the spliceosome machinery. Similar work has been performed with other genes and one of them is by Bonnal et al. with Fas receptors. Bonnal et al. in 2008 identified the presence of a factor RBM5 which act as a regulator of alternative splicing event in apoptosis inducing Fas receptors that have 8 known splicing variants. The presence of splicing signals has also been reported in the introns themselves such as 5'GT rich signals. These type of spliciosomal snRNP specific splicing signals has been invariably reported in eukaryotes and are highly conserved among deep-branching lineages.[35]

Park et al. (2004) has performed RNAi screening to identify the splice regulatory proteins in 3 alternatively spliced genes of *Drosophila.* According to them the typical process of identifying splicing regulators of an alternatively spliced gene starts with construction of a minigene that accurately undergoes splicing in vitro. The second step requires identification of RNA sequences required for regulation and finally leading to the identification of proteins that bind to these elements.[44] Hence the constructs of RET gene designed in this study will serve as the first step in the identification process of splice regulators of the RET proto-oncogene.

The real time PCR data with neuronal RNA has revealed that RET9 isoform is expressed in higher amounts than RET51. The constructs can also be used to evaluate the increase in production of RET9 over RET51 mRNA and further during the formation of RET protein any reversible changes in the amounts of the two isoforms can also be evaluated. The construct created can thus be used to quantify the fold of increase in the amounts of the two isoforms.

Ohno et al (2007) developed a transgenic alternatively splicing reporter system to identify the switch that regulates the expression of two isoform of let-2 gene that encodes collagen in *C. elegans*. This gene has two isoforms and they vary in their expression level during the developmental stages of the nematodes. One isoform is expressed in higher amounts in larval stage while the expression of another one is higher in the adult stage of the nematode. The expression pattern of this gene is similar to RET proto-oncogene in which the RET9 is expressed in higher amount in earlier stages of kidney development while RET51 is required in later stages and shows higher expression in protein level than RET 9 protein. The isoforms of the let-2 also differs in between its two exons- exon9 and exon10 similar to exon19 and

exon 20 in case of RET proto-oncogene. By using GFP based reporter system they have identified a novel member of the highly conserved STAR (signal transduction activators of RNA) family RNA-binding proteins, ASD-2 (for Alternative Splicing Defective-2), as a regulator of the *let-2* alternative splicing.[48]

The sequencing data of the construct can be used to locate the transcriptional factor binding site, exon/intron border consensus sequences, exon splicing enhancers (ESE) sequences which are binding sites for specific serine/Arginine-rich (SR) proteins involved in the splicing machinery. Cebrain et al. 2005 have used PuPa SNP finder, a web based tool to find SNP in RET proto-oncogene. The tool also locates the above mentioned regions which are essential to understand the splicing event of a gene.[45] According to Ni et al. (2007) identification of SR proteins and hnRNPs are essential to understand the alternative splicing event in any gene.[46]

In conclusion this study has been able to take the first few steps towards the goal of finding the regulator which is responsible for the splicing of the gene RET. Our immediate aim is to use the construct developed in this study to transfect in vitro along with the ligand GDNF and GFRα1 and quantify the expression of the two isoforms of RET; RET9 and RET51. The construct is also planned to be used along with reporter genes to find the switch responsible for the splicing of the gene RET.

BIBLIOGRAPHY:

1. Takahashi et al. (September 1985) "Activation of a Novel transforming gene, ret, by DNA rearrangement." Cell; 42:581-588
2. Santoro et al. (2004) "Minireview: RET: Normal and Abnormal Functions." Endocrinology;145:5448-5451
3. Arighi et al. (2005) "Ret tyrosine kinase in development and cancer. Cytokine and growth factor reviews."Journal Of Endocrinology;16:441-467
4. Ichihara et al. (2004) "RET and neuroendocrine tumors." Cancer Lett.;204:197-211
5. Piechala et al. (2006) "Glial cell line-derived neurotrophic factor-dependent recruitment of Ret into the lipid rafts enhances signaling by partitioning Ret from proteosome-dependent degradation." The Journal of Neuroscience;26:2777-2787
6. King-Yiu Lee et al. (2003) "3' Splicing variants of ret receptor tyrosine kinase is differentially expressed in mouse embryos and in adult mice" Elsevier;1627:26-38
7. Airaksinen et al. (2002) "The GDNF family: Signaling, biological functions and therapeutic value." Nature Review Neuroscience;3:383-394
8. Blume et al. (2001) "Oncogenic Kinase signaling." Nature; 411:355-365
9. Amoresano et al. (2004) "Direct interactions among Ret, GDNF and GFRα1 molecules reveal new insights into the assembly of a functional three-protein complex" Cell signaling; 17 :717-727
10. Sandra et al. (November 2006) "Quantitative analysis of the activation mechanism of the multicomponent growth-factor receptor Ret." Nature Chemical Biology;2:11
11. Coulpier et al. (November 16, 2001) "Coordinated activation of autophosphorylation sites in the RET receptor tyrosine kinase: Importance of tyrosine1062 for GDNF-mediated neuronal differentiation and survival." JBC;6;789
12. Elena et al. (2005) "RET tyrosine kinase signaling in development and cancer. cytokines and growth factor" Nature reviews;16:441-465
13. Wong et al. (November 2005) "Phosphorylation of 1062 is critical for the In Vivo activity of the Ret9 receptor tyrosine kinase isoform. JBC;25(21):9661-9673
14. Attie et al. (1995) "Diversity of RET proto-oncogene mutations in familial and sporadic Hirschsprung disease." Human Molecular Genetics; 4: 1381-1386

15. Gujral et al.(November2006) "Molecular Mechanisms of RET Receptor- Mediated Oncogenesis in Multiple Endocrine Neoplasia 2B" Cancer Res. 66:22-28
16. Machens et al.(2006) "Multiple endocrine Neoplasia type2 and the RET protooncogene: From Bedside to bench side" Molecular and cellular endocrinology; 247:34-40
17. Ciampi et al. (2007) "Mini review: RET/ PTC Rearrangements and Braf Mutations in Thyroid Tumerogenesis." Endocrinology;148(3):936-941
18. Menacho et al. (September 2009) "Current concepts in RET-related genetics, signaling and therapeutics." Trends in genetics;35:7-21
19. Takahashi (2001) "The GDNF/ RET signaling pathway and human disease." Cytokine and growth factor reviews; 12:361-373
20. Hickey et al. (2009) "RET-Mediated Gene Expression Pattern is affected by Isoform but Not Oncogenic Mutation." Genes Chromosomes and Cancer; 48:429–440
21. Carlomango et al. (1996) "Molecular heterogeneity of RET loss of function in Hirschsprung's disease." The EMBO Journal;15(11):2717-2725
22. Carter et al. (2001) "Conservation of RET proto-oncogene splicing variants and implications of RET isoform function" Cytogenetics & Cell genetics; 95:169-176
23. Heanue et al. (2008) "Ret isoform function and marker gene expression in the enteric nervous system is conserved across diverse vertebrate species" Mechanism of Development;12: 5687-699
24. Lodish et al. "Molecular Cell Biology" Sixth edition
25. David Brett et al. (December, 2001) "Alternative splicing and genome complexity."Nature Genetics 30: 29–30
26. Black (February,2003) "Mechanism of Alternative pre-messenger RNA splicing" Annual rev. Biochemistry; 72: 291-336
27. Stamm et al. (2005) "Function of alternative splicing" Gene;344: 1-20
28. De Graff et al. (2001) "Differential activities of the RET tyrosine kinase receptor isoforms during mammalian embryogenesis." Genes Dev. 15:2443-2444
29. Fluge et al. (2001) "Expression and alternative splicing of c-ret RNA in papillary thyroid carcinomas" Oncogene;20:885 ± 892

30. Jijiwa et al. (September, 2004) "A Targeting Mutation of Tyrosine 1062 in Ret Causes a Marked Decrease of Enteric Neurons and Renal Hypoplasia" Molecular and cellular biology; 8026–8036
31. Schedl et al. (2000) "A cross talk in kindey development" Curr. Opin. Genet. Dev.10 543-549
32. Murakami et al (2007) "RET signaling-induced SPHK1 gene expression plays a role in both GDNF-induced differentiation and MEN2-type oncogenesis." Journal of neurochemistry;102: 1585-1594
33. Lindahl et al. (2000) "Expression and alternative splicing of Mouse Gfra4 suggest roles in endocrine cell development."Molecular and cellular neuroscience;15:522-533
34. Bonnal et al. (October, 2008) "RBM5/Luca-15/ H37 Regulates Fas Alternative Splice site pairing after exon definition." Molecular cell; 32: 81-95
35. Stepanka et al. (March, 2005) "Spliciosomal introns in the deep-branching eukaryotes Trichomonas vaginalis" PNAS; (12)102
36. Pasman et al. (1996) "The 5' and 3' splice sites come together via a three dimensional diffusion mechanism." Nucleic acid research;24(9)
37. Yaming et al. (2006) "Nuclear and cytoplasmic mRNA quantification by SYBR green based real-time PCR." Elsevier; 39(4):356-362
38. Min Li et al. (2007) "Detection and quantification of HTLV-1 and HTLV-2 mRNA species by real time PCR" Journal of Virological methods; 142: 159-168
39. Venables P. Julian (2004) "Aberant and Alternative splicing in cancer" Cancer research; 7647-7654.
40. Hir et al. (2000) "Relative Expression of the RET9 and RET51 Isoforms in Human Pheochromocytomas." Oncology;58(4)
41. Jain et al. (2006) "Critical and distinct roles for key RET tyrosine docking sites in renal development" Genes & Dev. 20: 321-333
42. Futcher et al. (October 2008) "Bound to splice" Nature; 455
43. Romero et al (April, 2006) "Alternative splicing in concert with protein intrinsic disorder enables increased functional diversity in multicellular organisms" PNAS;103(22):8390-8395

44. Park et al (November, 2004) "Identification of alternative splicing regulators by RNA interference in Drosophila" PNAS; 101(45):15974–15979
45. Cebrain et al. (2005) "Polymorphisms in the Initiators of RET (Rearranged during Transfection) Signaling Pathway and Susceptibility to Sporadic Medullary Thyroid Carcinoma" The Journal of Clinical Endocrinology & Metabolism: 90(11) 6268-6274
46. Ni et al. (2007) "Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay."Genes and Development;21:708–718
47. Forch et al. (2002) "The splicing regulator TIA-1 interacts with U1-C to promote U1 snRNP recruitment to 5' splice sites" The EMBO Journal; 21(24): 6882±6892
48. Ohno (2008) " STAR family RNA-binding protein ASD-2 regulates developmental switching of mutually exclusive alternative splicing in vivo" Genes and Development; 22:360–374
49. Wang et al (2007) "Splicing in disease: disruption of the splicing code and the decoding machinery" Nat. Rev. Genet. 8:749-761
50. Fedor J. Martha (2007) "Combinational control facilitates splicing regulation of gene expression and enhances genome diversity." JBC;283:1209-1210
51. International human genome sequencing consortium (2004) Nature 431,931-945
52. Reed R. (1996) "Initial splice-site recognition and pairing during pre-mRNA splicing" Curr. Opin. Genet. Dev. 6, 215-220
53. Hertal J. K. (2007) "Combinatorial control of Exon recognition." JBC;283(3):1211-1215
54. Zhang et al. (2009) "Splicing designer exons reveals unexpected complexity in pre-mRNA splicing." RNA;15:367-376