11-2016

The Biological Function of Hammerhead RANA Ribozyme in Human Genome

Ibtehal M. A. Awad

United Arab Emirates University

College of Science

Department of Biology

THE BIOLOGICAL FUNCTION OF HAMMERHEAD RNA RIBOZYME IN HUMAN GENOME

Ibtehal M. A. Awad

This thesis is submitted in partial fulfilment of the requirements for the degree of Master of Science in Molecular Biology and Biotechnology

Under the Supervision of Dr. Khaled Amiri

November 2016

Declaration of Original Work

I, Ibtehal M. A. Awad, the undersigned, a graduate student at the United Arab Emirates University (UAEU), and the author of this thesis entitled "*The Biological Function of Hammerhead Rna Ribozyme in Human Genome*", hereby, solemnly declare that this thesis is my own original research work that has been done and prepared by me under the supervision of Dr. Khaled Amiri, in the College of Science at UAEU. This work has not previously been presented or published, or formed the basis for the award of any academic degree, diploma or a similar title at this or any other university. Any materials borrowed from other sources (whether published or unpublished) and relied upon or included in my thesis have been properly cited and acknowledged in accordance with appropriate academic conventions. I further declare that there is no potential conflict of interest with respect to the research, data collection, authorship, presentation and/or publication of this thesis.

Student's Signature: Ibtel Date: 14/12/2016

Advisory Committee

1) Advisor: Dr. Khaled Amiri

Title: Chairman of the Department

Department of Biology

College of Science

2) Co-advisor: Dr. Asma Al Menhali

Title: Assistant Professor

Department of Biology

College of Science

3) Member: Dr. Ranjit Vijayan

Title: Assistant Professor

Department of Biology

College of Science

Approval of the Master Thesis

This Master Thesis is approved by the following Examining Committee Members:

1) Advisor (Committee Chair): Dr. Khaled Amiri
   Title: Chairman of the Department
   Department of Biology
   College of Science
   Signature _____ Date 21/11/2016

2) Member: Dr. Rabah Iratni
   Title: Associate professor
   Department of Biology
   College of Science
   Signature _____ Date 21.11.16

3) Member (External Examiner): Dr. Ghada Al-Kafaji
   Title: Assistant Professor
   Department of Molecular Medicine
   Institution: College of Medicine and Medical Sciences, Arabian Gulf University. Bahrain
   Signature _____ Date 21/11/16

This Master Thesis is accepted by:

Dean of the College of Science: Professor Ahmed Murad

Signature _____ Date 14/12/2016

Dean of the College of the Graduate Studies: Professor Nagi T. Wakim

Signature _____ Date 14/12/2016

Copy 10 of 10

Abstract

Hammerhead RNA [HHR] sequence originally found in viroids is also detected in many species. HHR are often found in introns of different genes ranging from prokaryotes to eukaryotes. HHR can form specific tertiary structure that make specific region susceptible to cleavage. However, the role of the HHR is not clear in other life forms. This study aims to investigate the role of HHR in the intron 6 of (RECK) gene. Our study shows that when blocking the HHR sequence in the intron 6, the production of RECK protein declines which may be due to the effect of HHR in splicing and RNA stability which was seen by a decrease in the RNA levels of exons near to intron 6. Previous experiments designed to identify the significance of HHR sequence were addressing the self-cleavage of hammerhead RNA *in vitro* and comprised of partially constituted DNA construct expressed in cell culture. However, hammerhead RNA is a robust molecule and self-cleavage can take place even in biologically irrelevant context and therefore the current studies do not address its biological significance directly.

The aim of this study is to address the biological role of this sequence ex *vivo.* The study aims at blocking the endogenous mRNA with intronic hammerhead sequence in *ex vivo.* This study in principle should elucidate the biogenesis of mRNA whose HHR sequence is blocked. The biogenesis of mRNA will be assessed with RT-PCR. We will utilize bridge antisense (BNAs) RNA to block the HHR sequence in the mRNA. Furthermore, this study will investigate DNA pattern around the loci harboring HHR sequence. In this instance, efficient code will be written to search through the human genome data to identify locations that harbor patterns of the HHR sequence. Next, several thousands of bases upstream and downstream of this sequence will be extracted and analyzed to locate interesting patterns that may include gene, promoter regions and intronic regions.

Keywords: Human genome, RECK gene, Intron 6, Consensus sequence, Fetal bovine serum.

Title and Abstract (in Arabic)

وظيفة بيولوجية HammerHead RNA الحمض النووي الريبوزي في الجينوم البشري

الملخص

تم الكشف عن الحمض الريبوزي HHR في أنواع كثيرة من الكائنات الحية. و غالبا ما توجد HHR في الانترون (Intron) في جينات مختلفة تتراوح من كائنات بدائية النوى الى كائنات حقيقية النوى . و مع ذلك فان الوظيفة البيولوجية ل HHR لم تعرف او غير المتواجد في Intron 6 من الجين HHR واضحة في معظم الكائنات الحية. و تهدف هذه الدراسه الى التعرف على الوظيفة البيولوجية يؤدي الى انخفاض معدل البروتين Intron 6 في HHR. أظهرت دراستنا انه عند منع تسلسل (RECK) في الجينوم البشري .

الهدف من هذه الدراسه هو التعرف على الوظيفه البيولوجيه ل في الانفسام الذاتي في خلايا الكائنات الحية في المختبر و ذلك عن طريق HHR التجارب السابقة تهدف الى التعرف على اهمية تسلسل دراسة زراعة الخلايا التي تحتوي على HHR فقد تبن ان HHR هو جزي قوي الانقسام يمكن ان يحدث حتى في سياق ذي صله بيولوجيا و بالتالي فان الدراسات الحالية لا تعالج اهميته الحيويه مباشرة. و الهدف من هذه الدراسة هو تناول الدور البيولوجي لهذا التسلسل خارج الكائن الحي . فان هذه الدراسه من حيث المبدا يجب توضيح نشوء حيوي مرنا الذي يمنع HHR سيتم تقييم نشوء حيوي مرن مع RT-PCR وسوف تستخدم BANs لمنع تسلسل HHR و علاوة على ذلك فان هذه الدراسة نهدف الى التعرف على الوظيفة البيولوجه ل HHR في الجينوم البشري.

مفاهيم البحث الرئيسية: الجينوم البشري، RECK الجينات، إنترون 6، توافق تسلسل، المصل البقري الجنين

Acknowledgements

All the praises, thank and prayers are to Allah. I am deeply indebted to every person in my family especially my parents, sisters, brothers and brothers in law for their assistance and regular support. I would like to express the deepest appreciation to my supervisor Dr. Khaled Amiri for supporting and encouraging me through my thesis. Moreover, I would like to thank him for putting their faith and belief in my effort and without their supervision and constant help this dissertation would not have been possible. Also I would to express my special thanks to Dr. Rabah Iratni Dr. Ghada Al-Kafaji and Noushad Karuvantevida for their cooperation and help. I would like to express my special thanks and appreciation to Ms. Hidaya M. Abdul Kader for supporting and encouraging me through my thesis. For that, I would like to thank the Molecular biology and biotechnology master program Coordinator: Dr. Synan Abu-Qamar. finally, I would like to express my special appreciation and thanks to United Arab Emirates University.

Dedication

I lovingly dedicate my thesis to my beloved parents (Mrs. Majid Awad and MS. Seham Abu Hashem), sisters, brothers and brothers in law who have never failed to give me a moral support and for giving me all I need and encouraging me to continue my higher education.

Moreover, I dedicate this thesis to my supervisor Dr. Khaled Amiri, who came up with the idea of my thesis, without him it would be impossible to complete my higher education.

Also, I dedicate this thesis to Ms. Hidaya Mohammed Abdul-Kader who has really been there through the hard times.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

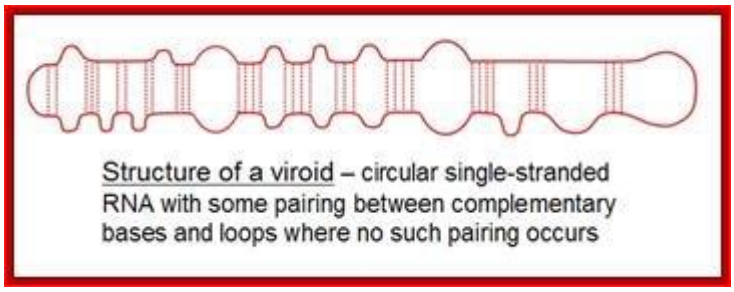
ALU	Arthrobacter Luteus
AS-RNA	Antisense RNA
BNA	Bridge Nucleic Acid
CS	Consensus Sequence
FBS	Fetal Bovine Serum
HHR	Hammerhead Ribozyme
RECK	Reversion Inducing Cysteine Rich Protein with Kazal Motifs
RIPA Buffer	Radioimmunoprecipitation Assay Buffer
RPMI 1640 medium	Roswell Park Memorial Institute 1640 Medium

Chapter 1: Introduction

1.1 Literature Review

Hammerhead Ribozyme (HHR) is a small, self-cleaving, and naturally occurring single stranded RNA. It was originally discovered in sub viral particles (viroid)and plant satellite viruses (Perrault et al, 2011) where they process RNA transcripts containing multimeric genomes to yield individual genome (Perrault et al, 2011). Sub viral HHR RNA domain is embedded into viroid genome. Viroid genome is small, slightly larger than prion. Its genome consists of short single stranded RNA (between 200 to 500 bases) that do not code for any protein (Perrault et al, 2011). Viroid are generally plant pathogens whose role include post-replicative (rolling circle mechanism) processing of multimeric genome into single genome units (Perrault et al, 2011).

A



B

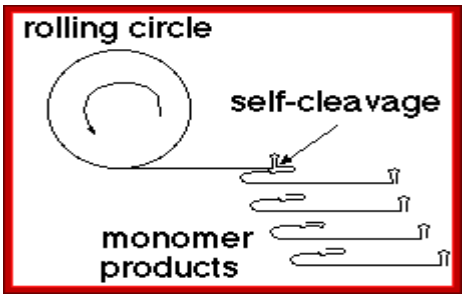


Figure 1: (A) Structure of a viroid RNA (B) Representation of replication pathway of viroid RNA showing post replicative processing of multimeric genome into single genome

1.2 Ribozyme

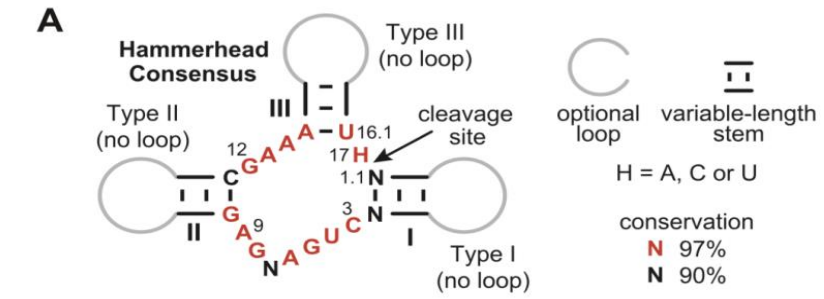
Ribozyme like the HHR, hairpin, hepatitis delta virus (HDV) and Varkud Satellite (VS) were found to catalyze sequence-specific intramolecular cleavage of RNA (Ferré-D'Amaré1 et al, 2010). They range between 50 and 150 nucleotides in length (Ferré-D'Amaré1 et al, 2010), and are known as the "small self-cleaving ribozymes." Crystal structures have revealed how RNA active sites can bind preferentially to the transition state of a reaction, whereas mechanistic studies have shown that nucleobases can efficiently perform general acid–base and electrostatic catalysis (Ferré-D'Amaré1 et al, 2010).

1.2.1 Hammerhead RNA Ribozyme

Specifically, HHR RNA ribozyme is composed of a catalytic core of conserved nucleotides flanked by three helices (Marcos et al, 2010), two of which form essential tertiary interactions for fast self-scission under physiological conditions (Marcos et al, 2010). The cleavage reaction proceeds by an attack of a 2' hydroxyl oxygen of a catalytic site cytosine on the phosphorus atom attached to the 3' carbon of the same residue, breaking the sugar phosphote backbone and producing a 2', 3' cyclic phosphate (Marcos et al, 2010).

The 5'-product, as a result of this cleavage reaction mechanism, possesses a 2', 3'-cyclic phosphate terminus, and the 3'-product possesses a 5'-OH terminus (Marcos et al ,2010).

A



B

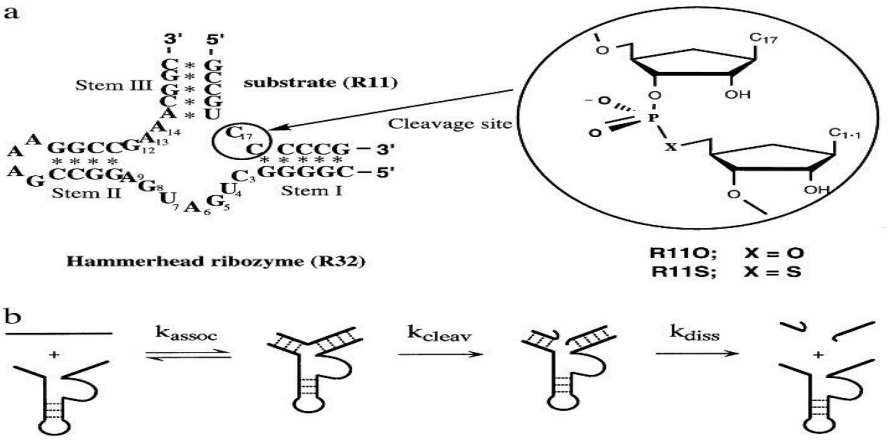


Figure 2: (A) Consensus secondary structure model for HHR ribozymes showing three stems I, II and III (B) (A) HHR secondary structure demonstrating site of cleavage (C_{17}) and (b) intermolecular reaction pathway for HHR (Adopted from De-Min-Zhou et al; 1997;vol94)

An active HHR core can be formed intra or inter-molecularly (figure 1 and 2B) and was clearly demonstrated by designing active HHR that consists of two oligonucleotides, namely, the substrate and catalytic strand (Jonathan P et al.,2011). Furthermore, it is observed in biological systems (viroid and plant satellite RNA viruses) that HHR invariant core sequence can be distally located and that a proper folding of RNA is sufficient to constitute an active HHR ribozyme. A catalytic turnover rate of minimal HHR ribozymes of ~ 1/min (a range of 0.1/min to 10/min) is commonly observed, depending upon the flanking non-conserved sequences, the lengths of the three helical stems, and under standard reaction conditions of high Mg2+ (~1,0mM) (Jonathan P et al.,2011),

pH 7.5 and 25°C (Amiri et al, 1996; Khvorova et. al. 2003; Martick and Scott, 2006). Most experiments on intermolecular HH utilized the minimal structure. These observations prompted the use of the HHR as a therapeutic tool as it can specifically cleave substrate mRNA of complementary sequence and therefore down regulated specific messages. Hence, HHR is viewed as a tool in Biotechnology (Jonathan P et al.,2011).

Recently, extensive and reiterative genomic *in silico* data search identified thousands of sequence homology across the tree of life. Each of these RNA sequence has a unique overall architecture and active site organization. The catalytic versatility of ribozymes may explain its abundance in contemporary organisms and also supports a role for catalytic RNAs early in evolution (Ferré-D'Amaré1 et al, 2010).

The HHR is found to be distributed across all of phlogeny (Figure 3). Initial studies showed presence of HHRs in metagenomic and bacterial sequences. Later they were also detected in metagenomic data from very diverse origins showing their presence in viruses, prokaryotes, archea and their bacteriophages. Studies also showed presence of HHRs in eukaryotic genomes, although they were mostly seen in intergenic and non coding regions(Ferré-D'Amaré1 et al , 2010).

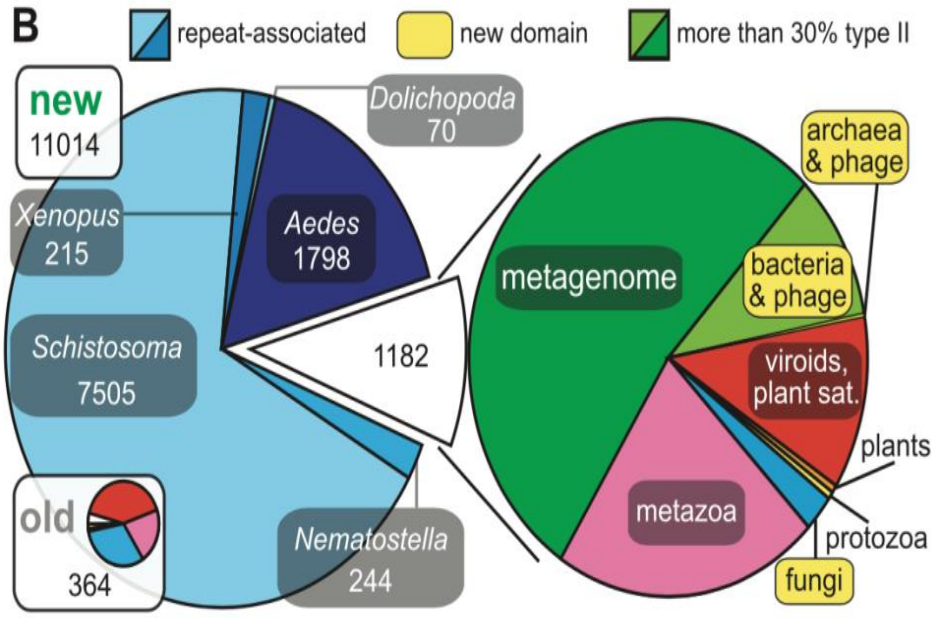


Figure 3: The expanded phylogenetic distribution of self-cleaving ribozyme class (Hammann et al., 2012)

The HHR sequences, in the majority of cases, reside in repetitive DNA sequence (retro elements), 3' UTR, and in the introns of several genes (Ferbeyre et al., 1998). Of note, HHR sequence is found in the intron of Reversion Inducing Cysteine Rich Protein with Kazal Motifs gene (RECK). RECK gene is located on chromosome 9 in human. The HHR lies in the middle of the intron 6 (Ferbeyre et al., 1998). Chromosome 9 is made up of about 141 million DNA building blocks (base pairs) and represents approximately 4.5 percent of the total DNA in cells. Chromosome 9 contains approximately 800 to 900 genes that provide instructions for making proteins. These proteins perform a variety of different roles for cellular functions (Ferbeyre et al., 1998).

1.3 RECK Gene

Reck gene is named after its structure and activity. The human RECK gene is 87 kb in length and has 21 exons and 20 introns with 13 SNPs (Figure 4). The RECK gene is a relatively newly discovered gene with important implications for cancer research. It is thought to be a metastasis suppressor. Up regulated RECK is linked to significantly prolonged survival rates in patients with severe forms of malignancies (Lindgren, 2014). RECK is normally expressed in all cells of the body to various degrees and has an important role in the balance between destructive and constructive features of the extracellular matrix (ECM) (Lindgren, 2014). The RECK protein is a membrane-bound glycoprotein that inhibit matrix metalloproteinases (MMPs) which has the function of breaking down the ECM. There is a significant correlation between RECK gene expression and the formation of new vessels, presumably via the mediation of VEGF (Vascular Endothelial Growth Factor) which is an important and powerful inducer of angiogenesis (Lindgren, 2014).

Research has shown that the down regulation of RECK is caused by the Ras oncogene, which otherwise also is a common cause of tumor development in the early stage (Engström et al, 2014). Reck gene is a potential inhibitor and a prognostic marker available at early clinical stages.

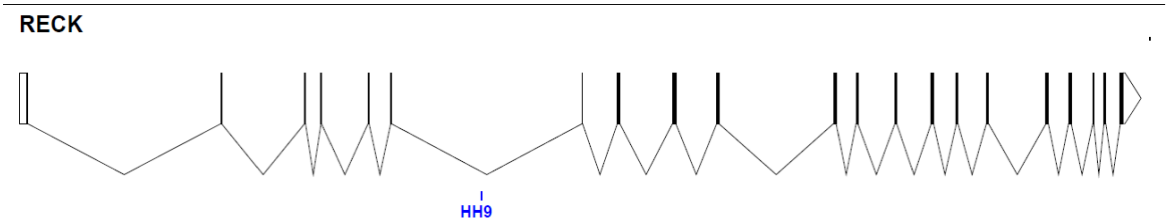


Figure 4: Schematic representation of RECK gene showing exons as vertical lines and introns are shown as lines joining exons. HH9 shows the position of HHR sequence on intron 6

RECK protein is made up of 971 amino acids (Takahashi et al, 1998) the protein is cysteine rich (9%). Although RECK gene contains HHR sequence spanning nucleotides 36254-36330 (NCBI GRCh38.p7 Primary Assembly) in the intron 6, the function is yet unclear. Few studies were attempted to study the biological activity of HHR *in vitro* (Amiri et al, 1996). It is clear from previous studies *in vitro* that any molecule that can form basic secondary structure of HHR is bound to cleave in the presence of divalent ion (Amiri et al, 1996, Khvorova et al, 2003).

Previous experiments designed to identify the significance of HHR sequence in eukaryotes addressed self-cleavage of HHR *in vitro* (Chek, 1983). In principle these experiments reconstitute HHR molecules *in vitro*. Briefly, intermolecular reconstitution of HHR *in vitro* involve addition of two strands (the substrate strand and enzyme strand) to form the secondary structure (Fig: 5). Subsequently, divalent ion (Mg^{2+}) is added to initiate self-cleavage (Chek, 1983). It is very likely that in vitro self-cleavage study does not portray the biological context in which HHR is found. Self-cleaving RNA is not unique to viroids but in earlier studies demonstrated their presence in unicellular and lower life and was termed as group I and group II introns (Chek, 1983). It should be noted that HHR is a robust molecule and self-cleavage can take place even in biologically irrelevant context as long as a stable structure is formed (Chek, 1983).

Therefore, the current studies do not address its biological significance directly. It is imperative to develop a system that allows investigation on the role of HHR in biologically relevant context (Chek, 1983).

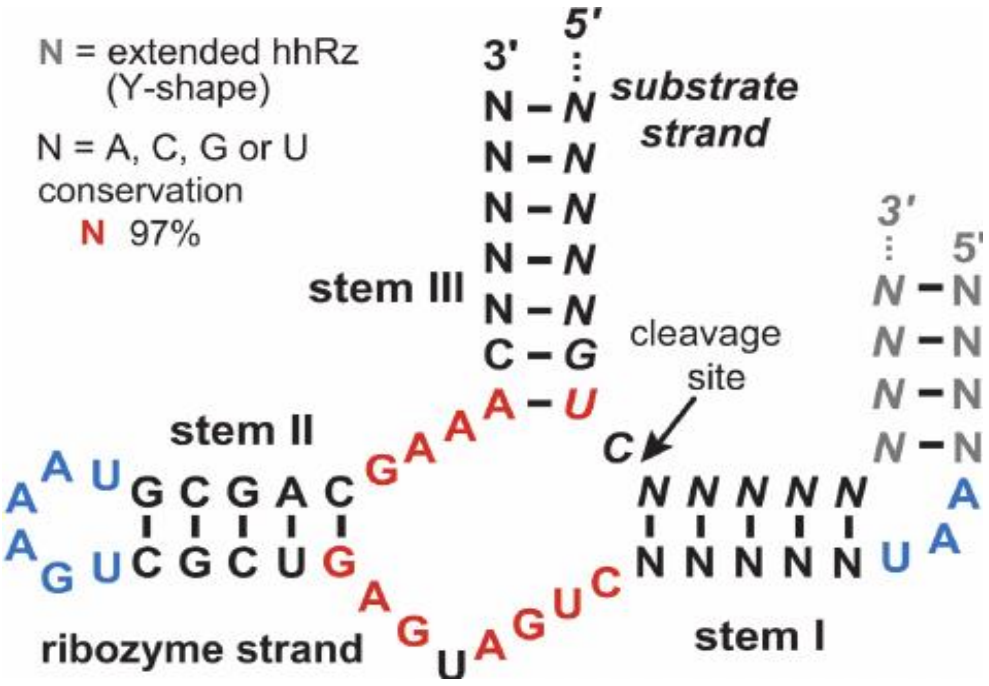


Figure 5: Model of trans-acting hammerhead ribozyme used by RiboSoft. Nucleotides in red correspond to the conserved catalytic core, other nucleotides may vary. Nucleotides in blue are important for the tertiary interaction between stems I and II. The sequence of stem II pictured here corresponds to the optional extended HHRz model (39) used by RiboSoft and includes the extra nucleotides in gray as well. - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/283481517_fig1_Figure-1-Model-of-trans-acting-hammerhead-ribozyme-used-by-RiboSoft-Nucleotides-in-red [accessed Oct 26, 2016]

1.4 Objective

In this study, we are addressing the fundamental question (biological role of HHR) in the biological context, namely cellular level. The study aims at blocking the endogenous mRNA with intronic HHR sequence. This study in principle should elucidate the biogenesis of mRNA whose HHR sequence is blocked. Furthermore, this study will investigate DNA pattern around the loci harboring HHR sequence.

In this instance, efficient code will be written to search through the human genome data to identify locations that harbor patterns of the HHR sequence. Next, several

thousands of bases upstream and downstream of this sequence will be extracted and analyzed to locate interesting patterns that may include gene and promoter regions and intronic regions.

Chapter 2: Materials and Methods

Materials

2.1 Choice of Cell Line

The Lung carcinoma adherent cell line A549 from ATCC was used for the experiments. This cell line serves as a good transfection host and has a doubling time of 22hours. This is a hypotriploid human cell line with the modal chromosome number of 66, occurring in 24% of cells. The cells are well-attached with an epithelial-like morphology and will form a monolayer of tightly compacted small cells. Previous studies revealed that A549 cells could synthesize lecithin with a high percentage of desaturated fatty acids utilizing the cytidine diphosphocholine pathway (M. Lieber et al, 1976). The cell line was grown and maintained at 37 ^{0}C (air, 95%; carbon dioxide (CO_2), 5%) in RPMI media supplemented with penicillin/streptomycin antibiotics and with fetal bovine serum to a final concentration of 10%.

2.2 Designing the Antisense

The HHR sequence was divided into three regions, namely P-box, U-box and control region based on nomenclature used in (Marcos and Robles, 2010). Short Antisense oligonucleotides (Bridged Nucleotides) of about 20-25 base pairs were designed to complement these regions including consensus HHR sequence. The oligonucleotides were provided by biosynthesis (Texas, USA).

Bridged oligonucleotides were used for the experiments. Bridged nucleic acids (BNAs) are modified nucleotides. They are sometimes also referred to as constrained or inaccessible nucleotide molecules. The bridge is synthetically incorporated at the 2', 4'-position of the ribose to afford a 2', 4'-BNA monomer. BNAs are structurally rigid oligo-nucleotides with increased binding affinities and stability.

Antisense oligonucleotides are synthetic polymers. The monomers are chemically-modified ribonucleotides like those in RNA. There are usually only 15–20 of them, hence "oligo". Their sequence (3′ → 5′) is antisense; that is, complementary to the sense sequence of a molecule of mRNA. Antisense oligonucleotides are synthesized in the hope that they can be used to block a gene by altering the synthesis of a particular protein. This would be achieved by the binding of the antisense oligonucleotide to the mRNA from which that protein is normally synthesized. Binding of the two may physically block the ability of ribosomes to move along the messenger RNA preventing synthesis of the protein or hasten the rate at which the mRNA is degraded within the cytosol or prevent splicing errors that would otherwise produce a defective protein. In this study we used BNA to block the HHR site on mRNA and BNA is used to increase the stability of mRNA. This will provide sufficient time for mRNA splicing before degradation.

2.3 Transfection

Chemical based transfection using the cationic liposome lipofectamine2000 (Invitrogen) was used successfully to transfect bridge antisense RNAs (BNAs) into the cell line A549 for RNA interference studies. Lipofectamine2000 has a high

transfection efficiency and high level of transgene expression. The standard protocol for transfection experiments are available at the Invitrogen from Life Technologies.

2.4 RNA and Protein Studies

Total RNA and protein were extracted from the transfected cell line for RT-PCR and western blotting respectively. PCR primers for various exons and intron 6(which harbor the HHR sequence) of RECK gene was designed using the NCBI primer designing tool.

2.5 Bioinformatics

Genome databases were used for the analysis of data and to identify sequence similarities. They were: NCBI-BLAST2 Nucleotide tool at the European Bioinformatics Institute (www.ebi.ac.uk), ENSEMBL (www.ensembl.org), University of California Santa Cruz Genome Bioinformatics(www.genome.ucsc.edu) and the National Center for Biotechnology Information (www.ncbi.nlm.nih.gov). A graphical representation of the human RECK gene mRNA was created using the Gene Structure Display Server (http://gsds.cbi.pku.edu.cn).

Methods

2.6 Gene Knockdown Experiment (Transfection)

All cell culture experiments were carefully aseptically conducted in a class II safety hood. Cell culture was initiated by reviving frozen A549 cell line into RPMI media added with streptomycin/penicillin antibiotics and supplemented with 20% FBS.

2.6.1 Cell Counting and Seeding in 6 Well Plates

Cells maintained in culture in T75 culture flasks at 70-80% confluency were trypsinsed. The cells were counted for subsequent experiments by taking 10 µL of harvested cells and the average from four grids of a hemocytometer were used to calculate cell density for seeding (Fig: 6). The counted cells were seeded at a cell density of 100,000 cells per well in six well culture plates and maintained at 37˚C until 70-80% confluency.

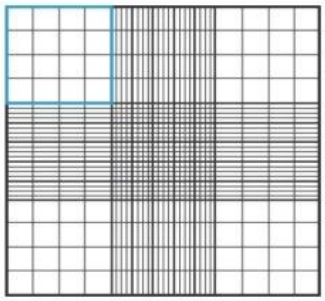


Figure 6: Grids seen in a hemocytometer. The corner square in blue represents one of the four squares to be counted

2.6.2 Antisense Experiment

Once the cells reached the required confluency, transfection experiment was conducted according to manufacturer protocol (Invitrogen by Life Technologies Lipofectamine 2000 Reagent protocol 2013). Firstly, Lipofectamine 2000 reagent was diluted in the medium used for culturing the cells. Then diluted the required amount of the designed antisense oligonucleotides in the medium used for culturing cells. The diluted oligonucleotide was then added to the diluted Lipofectamine 2000 reagent and incubated at room temperature for five minutes for AS-RNA-lipid complex to form and this complex was then added to the cells. The total volume of

media in each well was made up to 3 ml. The transfected cells were then visualized and harvested for further analysis after 24hrs of treatment. The design of our experiment is summarized in table no1.

Table 1: Transfection experiment plan

	Control	**Lipo Control**	**Oligo Control**	**AS-U-25**	**AS-U-20**	**AS-P-21**
Volume of Media	300 µL	282 µL	300 µL	282 µL	282 µL	282 µL
Volume of Lipofectamine	-	18 µL	-	18 µL	18 µL	18 µL
Volume of Media	300 µL	300 µL	298.8 µL	298.8 µL	298.8 µL	298.8 µL
Volume of Oligonucleotide	-	-	1.2 µL	1.2 µL	1.2 µL	1.2 µL

2.7 Molecular Analysis of the Transfected Cell Lines

The transfected cells were analyzed at the RNA and protein levels.

2.7.1 RNA Analysis

RNA extraction: RNA was extracted from the transfected cells using Qiagen RNeasy Mini kit. All procedures were carried out at room temperature. For extraction, the transfection media was removed from the 6 well culture plates and cells washed with ice cold PBS. The cells were scraped from the culture dish after adding 350µL lysis

buffer and transferred to 1.5 mL tubes and centrifuged for 3 minutes at maximum speed. The supernatant was carefully removed and equal volume of ice cold 70% ethanol was added and mixed by pipetting. This solution (total of 700 µL) was then added to spin column and centrifuged at 10,000rpm for 1 minute. The flow through was discarded and 700 µL of wash buffer added to spin column and centrifuged at 10,000rpm for 1 minute. The flow through was again discarded and 500 µL of buffer RPE was added to the spin column and centrifuged at 10,000rpm for 1 minute. The flow through was discarded and this step was repeated and the spin column was centrifuged at 10,000rpm for 2 minutes. The flow through was discarded and the spin column was re-spun at maximum speed for 1 minute to dry the column. The spin column was then placed in elution tubes and 50 µL of RNase free water was added and incubated at room temperature for 1 minute and centrifuged at 10,000rpm for 1 minute to elute the RNA. Finally, the extracted RNA was quantified using the nanodrop spectrophotometer.

RT-PCR for RNA: RT- PCR was done in the study to convert the various regions of the RECK gene into cDNA using the following primer sets.

Table 2: PCR primers for detection of RECK gene

S.No.	Primer Name	Primer Sequence	Product Size (bp)	Position on Gene	Detection of
1	EX_4	**Forward** 5'-AAA TTA ATT GTC ATA TTT AGG AAC TGG-3' **Reverse** 5'-GGT TCC AGG ACA ATG ACA CC-3'	186	23622-23807	Exon 4

2	EX_5	**Forward** 5’-TTG AAA CAT CTG GCA TGC TAT C-3’ **Reverse** 5’-GCA CAG CTA AAA CCA AAC TGC-3’	217	27291-27507	Exon 5
3	EX_6	**Forward** 5’-TTG CCC TGT GCT GAA TAG C-3’ **Reverse** 5’-TTT CTG TAA GGT AAA AGG GAA GTT G-3’	225	29087-29311	Exon 6
4	EX_7	**Forward** 5’-TGT TGG AAG ATG GTT TCC C-3’ **Reverse** 5’-GTG CAC ATT GCT TCA GAT GG-3’	288	43988-44275	Exon 7
5	EX_8	**Forward** 5’-TCA TCG TCT GTG GCA TGG-3’ **Reverse** 5’-TGC CAA ATC TAA ATA CCA AAG G-3’	462	46807-47268	Exon 8
6	Int-6-1	**Forward** 5’-TTT TTG TGC CAG GTT TGC CC-3’ **Reverse** 5’-AAG GAA AAC GGC ACA GTT GC-3’	1320	40042-41361	Intron 6
7	Int-6-3	**Forward** 5’-TGT GAT ACT TTG CCC CAG CA-3’ **Reverse** 5’-TTT GCT GAT GTT CAC GCT GC-3’	1482	36088-37569	Intron 6
8	Int-6-4	**Forward** 5’-TCC ATA GCC TGT CCT CCG TT-3’ **Reverse** 5’-AGC TAG AGG GGT TTG TGG GA-3’	1443	30330-31772	Intron 6

9	Int-6-5	**Forward** 5'-TTC CCA GAG ACT GGC AAT GG-3' **Reverse** 5'-CCT AGG CTG TGA TCT GCC AC-3'	1123	34906-36028	Intron 6
10	Int-6-6	**Forward** 5'-TCT TGA CCC ATC CAG GTG GT-3' **Reverse** 5'-TGC AGA ACC CTT CCT GCA TT-3'	1160	29528-30687	Intron 6
11	HHR-1	**Forward** 5'-TGG ATC TGT ACT GCA GTG G-3' **Reverse** 5'-TGA GAA GAA TGA ACT GGA AAA GG-3'	487	35962-36448	Hammerhead HH-9

A 25µL reaction mix was prepared as follows:

The PCR was performed using T100^{tm}ThermoCycler from Bio Rad. The following parameters was set on the machine for the RT_PCR.

Table 3: Components for one-step RT-PCR used for all RNA samples multiplexed with house-keeping gene GAPDH

S.No	Component	Volume/Reaction	Final Concentration
1	RNase-free Water	13 µL	-
2	5x QIAGEN OneStep RT-PCR Buffer	5 µL	1x
3	dNTP Mix (containing 10Mm of each dNTP)	1 µL	400 µM of each dNTP
4	Forward Primer	0.5 µL	0.6 µM
5	Reverse Primer	0.5 µL	0.6 µM
6	GAPDH Forward Primer	0.5 µL	0.6 µM
7	GAPDH Reverse Primer	0.5 µL	0.6 µM
8	QIAGEN OneStep RT_PCR Enzyme Mix	1 µL	-
9	Template RNA (10ng/µL)	3 µL	30ng
	Total Volume	**25 µL**	

Table 4: RT-PCR thermal cycler parameters

Cycle	Stage	Temperature	Duration
Lid 105 ˚C			
Once		50˚C	30 min
Once	Hot-start	95 ˚C	15 min
35	Denaturation	94 ˚C	1 min
	Annealing	58 ˚C	30 seconds
	Extension	72 ˚C	1 min
Once	Post-Extension	72 ˚C	10 min
Once	Hold	12 ˚C	∞

The resulting PCR products were subjected to gel electrophoresis for quantifying the band intensities of the various regions to confirm the effect of gene knockdown.

2.8 Protein Analysis

2.8.1 Protein Extraction

Total protein was extracted from the transfected cells using the RIPA buffer. All procedures were carried out on ice. For extracting proteins, the transfection media was removed from the 6 well culture plates and the cells washed with ice cold PBS. The cells were scraped after adding 60µL of RIPA buffer (added with 10µL/mL of protease inhibitor to buffer) and transferred to 1.5 ml tubes. The tubes were incubated and vortexed at intervals for 30 minutes. They were then centrifuged at

15,000rpm for 15 minutes at 4˚C and the supernatant was collected. Finally, the protein was quantified using Bradford Assay and stored at -80 ˚C until used for western blotting.

2.8.2 Western Blotting for Protein

The quantified proteins were reduced and denatured by addition of SDS sample buffer and heated at 95˚C for 5 minutes. About 100ng of whole protein was loaded in an SDS polyacrylamide gel for performing an SDS-PAGE. The gel was then transferred to a nitrocellulose membrane using the wet transfer method. And the membrane was blocked using 5% milk for 1-2 hours. Then the membrane was incubated in primary antibody (RECK MaxPad rabbit polyclonal antibody (D01)-catalog number: H0008434-D01) for 1-2 hours diluted in PBST followed by incubation of the membrane with enzyme conjugated secondary antibody (goat – anti – rabbit IgG-HRP) for 1-3 hours diluted in PBST. The bands on membrane was then detected using ECL detection reagent to visualize the bands and the membrane was scanned to analyze the bands.

2.9 Bioinformatics: A Database Search for Homologies and Similarities

The RECK gene intron 6 which harbors HH9 sequence was blasted to the human Refseq RNA to find if there are similar sequence present in mRNA of other genes. Two regions namely one near 5’ end and one near 3’ end of the intron 6 showed similarities to mRNA of many other genes. We named them 5’ and 3’ consensus sequence (CS). 5’ CS and 3’ CS was further analyzed in ten selected genes using NCBI gene bank to find the position of these sequence on the gene and the feature of the gene on which it is contained. The 5’ CS and 3’ CS were blasted against all the

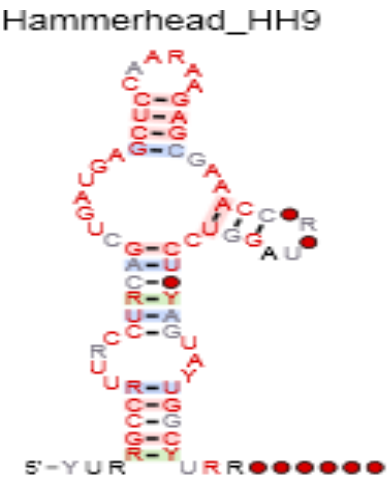
introns of RECK gene using BLAST align two sequences tool. 5’CS and 3’CS were also aligned to each other and with ALU sequence to find if they belong to retrotransposons. The intron 6 was also blasted to RefSeq Genome to find if they are seen in introns of other genes.

Chapter 3: Results and Discussion

3.1 Hammerhead Ribozyme

HHR was originally found in viroid and other satellite viruses. However, later studies revealed its presence in many other forms of life, including eukaryotes. HHR ribozyme is found in intron 6 of RECK gene on chromosome 9 in homo sapiens. The HHR in RECK gene comprise of naturally occurring helix II motifs flanked by conserved U and P boxes (figure 7A). In addition to many novel motifs, they identified various examples of RNAs that conform to the well - established consensus sequence for HHR - cleaving ribozymes in prokaryotes. Although the precise biological functions of these HHRs remain unclear, the fact that nearly all carry conserved sequence and structural features, previously proven to promote RNA cleavage by internal phosphodiester transfer suggests that they also promote high-speed RNA cleavage. Almost without exception, the catalytic core of each representative matches the consensus HHR sequence. Also, the three base-paired stems enclosing the catalytic core typically show variability in sequence and length, with stem II commonly formed by as few as two base pairs. (Figure 7B).

A



B

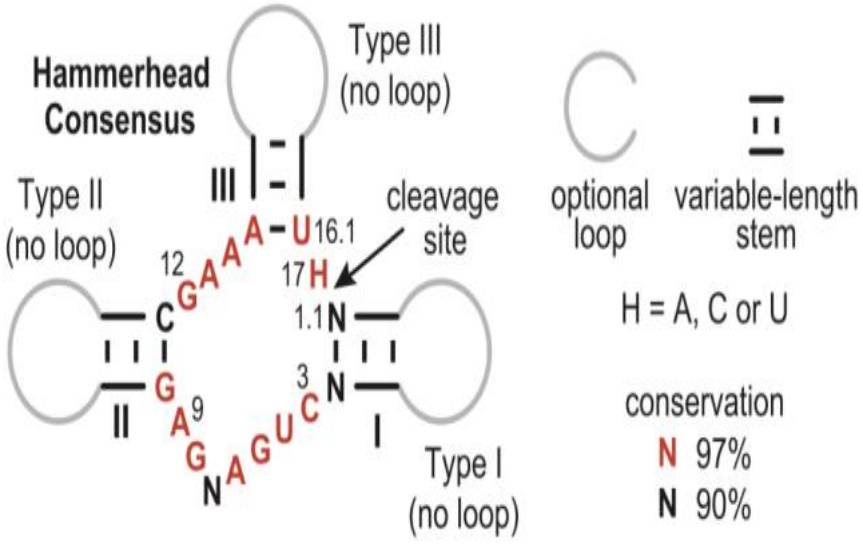


Figure 7: (A) Predicted secondary structure of HHR in intron 6 of RECK gene and (B) the consensus secondary structure model for HHR ribozymes

Past studies demonstrated that similar HHRs were embedded in genomes of several eukaryotic organisms. While many HHR associated with repeated sequences, which are found in multiple copies in eukaryotes genomes, were previously found in various species such as: *Schistosoma mansion, Schistosoma japonica, Opisthorchis viverrini, Ambystomamexicanum, Xenopustropicalis, Petromyzon marinus* (Marcos,

Robles, 2010). It was noted that these motifs were seen in intronic domains of their genes and had close features with reverse transcriptase like genes (Fig: 8).

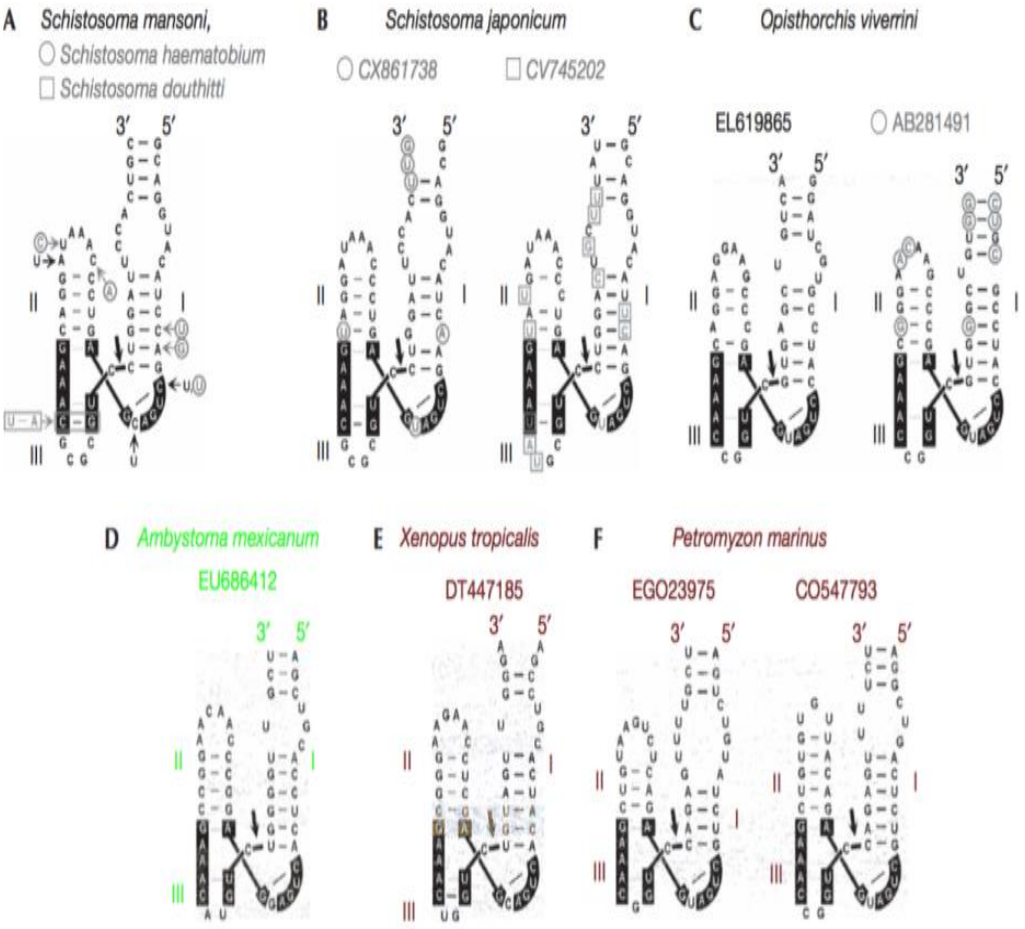


Figure 8: Examples of HHR motifs detected in eukaryotic species. (Marcos, Robles, 2010)

Human HHR located on the middle of intron 6 (14 kb), the HHR fold was observed to be ultra-conserved (90% identity for 60 nt) among the introns of RECK in the genomes of reptiles, birds, and mammals including humans. Furthermore, sequence heterogeneity was mostly restrained to the Helix III- loop 3 region and was always compatible with a catalytically active HHR (figure9) (Marcos, Robles, 2010).

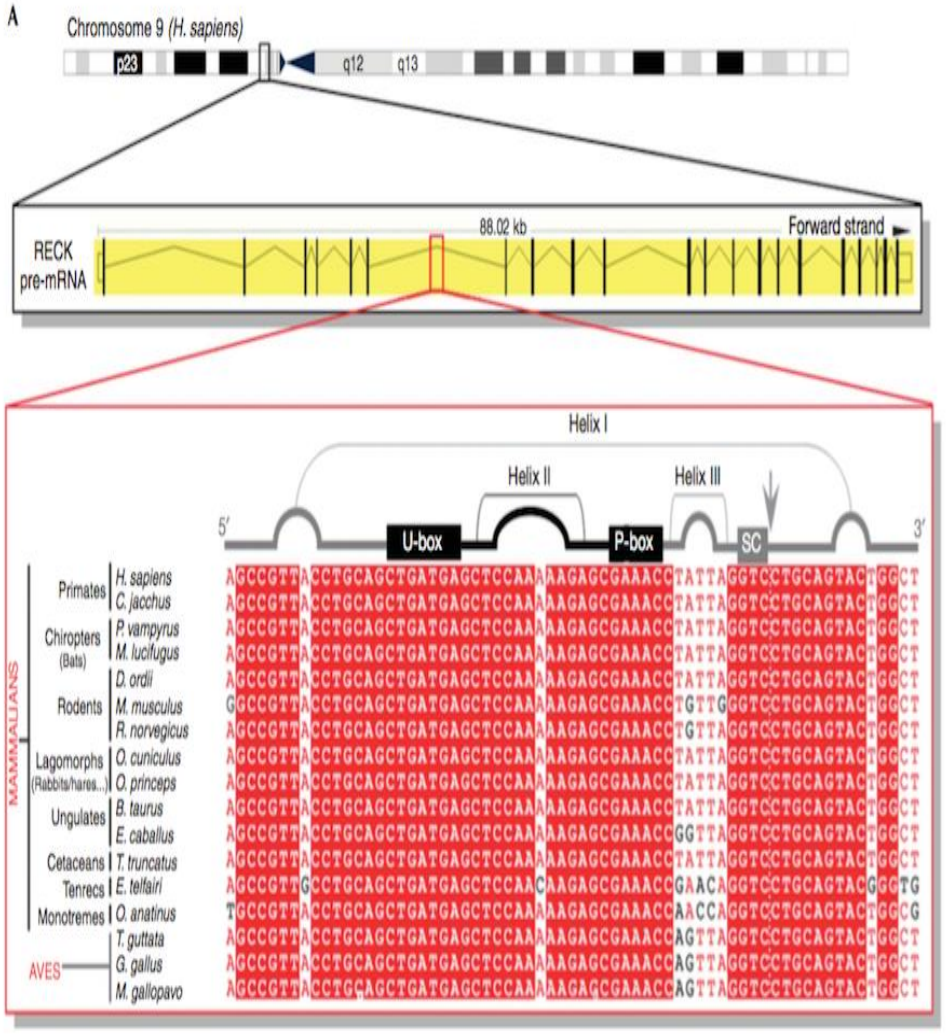


Figure 9 : An intronic HHR is ultra-conserved in the genome of different eukaryotic species, also the approximated location of the RECK gene within the human chromosome 9 (Marcos, Robles, 2010)

3.2 Effect of HHR in Biological Context

The main objective of this study was to find out the possible biological function of the HHR sequence harbored in the intron 6 of RECK gene. The protein analysis of cells in which the HHR sequence was knocked down using antisense oligonucleotides, showed a great decrease in the level of RECK protein compared to the normal cells (Fig 10).

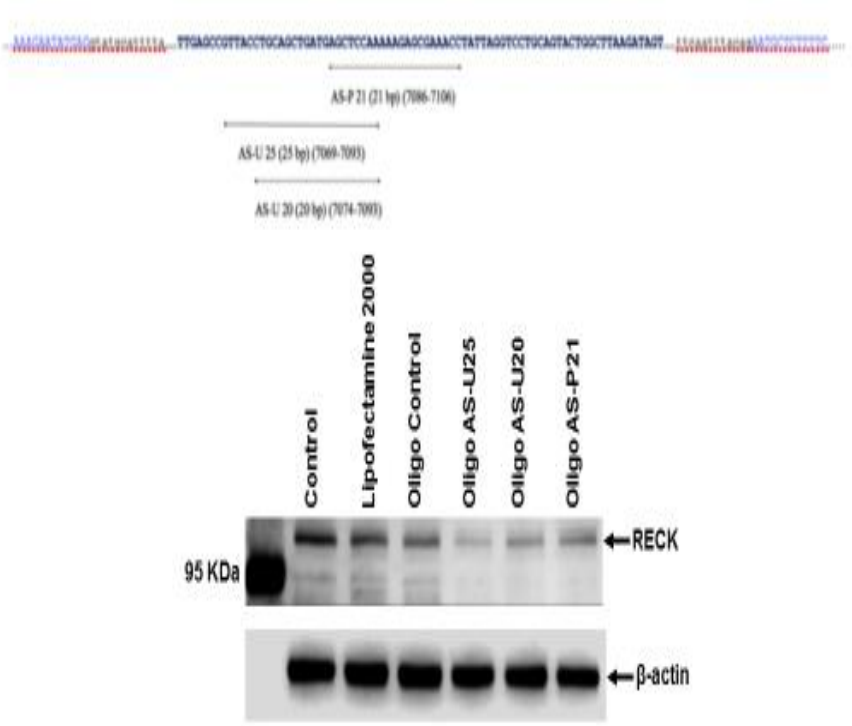


Figure 10: Western blot showing a decrease in the RECK protein levels in knockdown cells compared to the normal cells

The lower expression of RECK protein in knockdown cells suggest the decrease in the mRNA levels. This was confirmed by RT-PCR of RNA from these cells. As expected, the gel electrophoresis of the PCR products showed lower band intensities in knockdown cells compared to normal cells (Fig 11). This decrease was more prominent in the exons closer to intron 6 namely, exon 4, 6 and 7. While the exon further away namely exon 8, showed a much lower change in the band intensities compared with the normal cells. The discrepancy in the quantity of exons 4,6 and 7 compared to exon 8 may result from aberrant splicing. Collectively, the results suggest that the biological function of RECK HHR might be either affecting the splicing (processing of the mRNA) or possibly its stability.

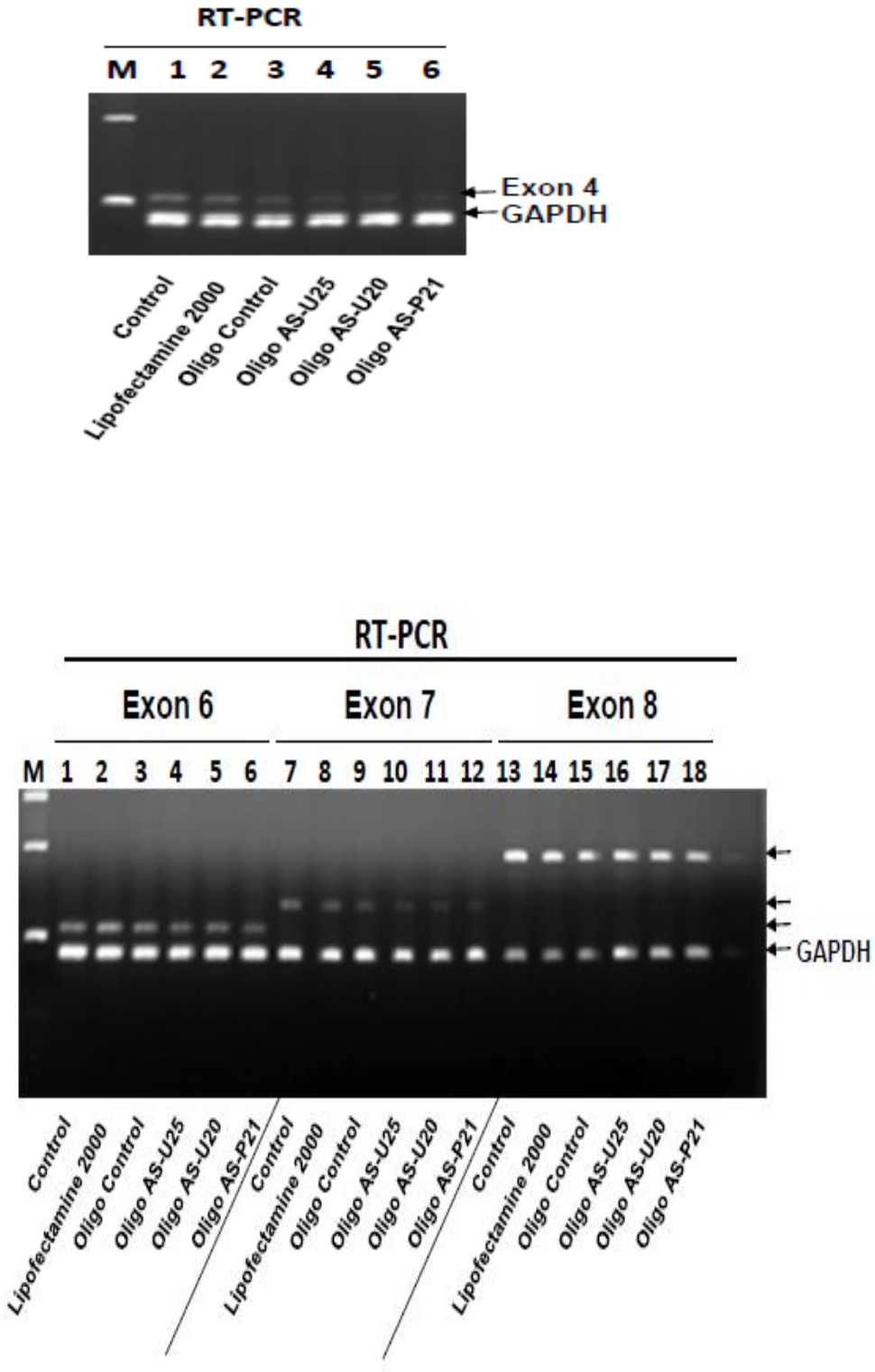


Figure 11: Electrophoresis gel images showing a decrease in RNA levels of exons 4,5,7 and 8 of RECK gene

3.3 Bioinformatics: Homologies and Similarities

In our study, all queries trying to align the HHR sequence to various genes gave a negative result except for the RECK gene (87551 bp in length) which returned a full hit showing the HHR sequence aligned from nucleotides 35782-35858 which is contained in intron 6 of the RECK gene (Fig:12). However, a RefSeq RNA blast query of intron 6 (14980 bp) of RECK gene showed a striking sequence similarity of about 200 base pairs to many other genes at the 5’ and 3’ ends. These sequences were referred to as 5’CS and 3’CS for our study (Fig: 13)

BLAST® » blastn suite-2sequences » RID-ZTCJV002114

BLAST Results

Blast 2 sequences

gi|568815589:36036905-36124455 Homo sapiens...

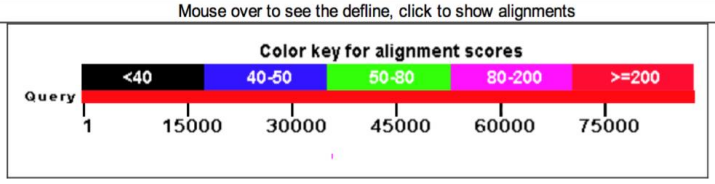
RID ZTCJV002114 (Expires on 10-12 17:35 pm)

Query ID	lcl|Query_175845	**Subject ID**	lcl|Query_175847
Description	gi|568815589:36036905-36124455 Homo sapiens chromosome 9, GRCh38.p7 Primary Assembly	**Description**	None See details
Molecule type	nucleic acid	**Molecule type**	nucleic acid
Query Length	87551	**Subject Length**	77
		Program	BLASTN 2.5.1+

Graphic Summary

Distribution of 1 Blast Hits on the Query Sequence

Mouse over to see the defline, click to show alignments



Descriptions

Sequences producing significant alignments:

Description	Max score	Total score	Query cover	E value	Ident	Accession
None provided	143	143	0%	4e-37	100%	Query_175847

Alignments

Sequence ID: Query_175847 Length: 77 Number of Matches: 1
Range 1: 1 to 77

Score	Expect	Identities	Gaps	Strand	Frame
143 bits(77)	4e-37()	77/77(100%)	0/77(0%)	Plus/Plus	

Features:

```
Query  35782  TTGAGCCGTTACCTGCAGCTGATGAGCTCCAAAAAGAGCGAAACCTATTAGGTCCTGCAG  35841
Sbjct  1      TTGAGCCGTTACCTGCAGCTGATGAGCTCCAAAAAGAGCGAAACCTATTAGGTCCTGCAG  60

Query  35842  TACTGGCTTAAGATAGT  35858
Sbjct  61     TACTGGCTTAAGATAGT  77
```

Figure 12: Shows the position of HHR on RECK gene and its sequence

A

BLAST ® » blastn suite » RID-GVSP3AKG015

BLAST Results

RECK intron 6

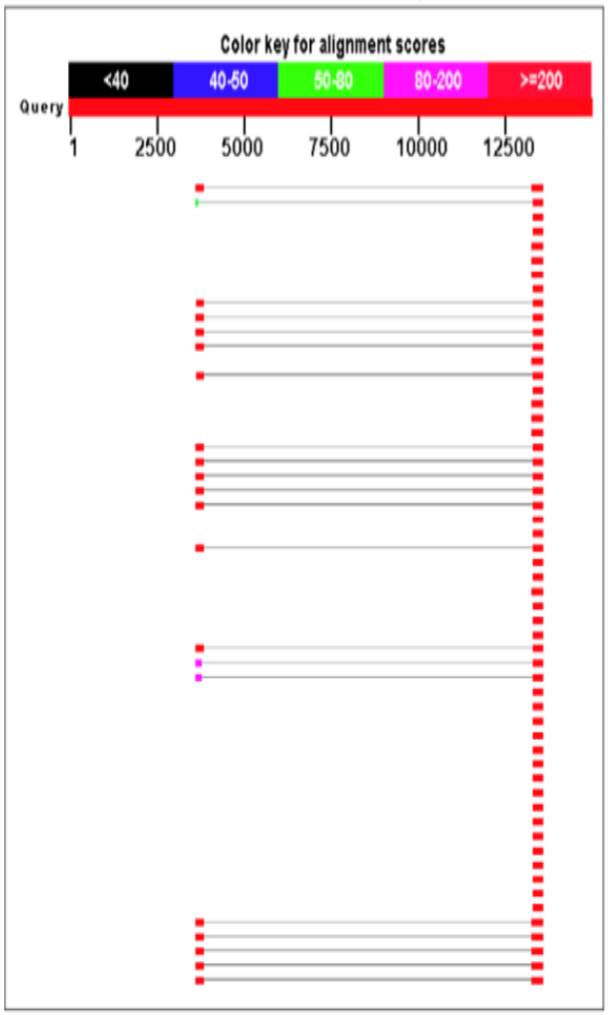
RID GVSP3AKG015 (Expires on 04-14 14:10 pm)

Query ID lcl|Query_90343
Description None
Molecule type nucleic acid
Query Length 14980

Database Name refseq_rna
Description NCBI Transcript Reference Sequences
Program BLASTN 2.3.1+

Graphic Summary

Distribution of 180 Blast Hits on the Query Sequence



B

Descriptions

Sequences producing significant alignments:

Description	Max score	Total score	Query cover	E value	Ident	Accession
Homo sapiens NUAK family kinase 2 (NUAK2), mRNA	385	631	3%	3e-104	88%	NM_030952.2
Homo sapiens KCNQ1 opposite strand/antisense transcript 1 (non-protein coding) (KCNQ1OT1), long non-coding RNA	377	524	2%	4e-102	90%	NR_002728.3
Homo sapiens oxoglutarate dehydrogenase (OGDH), transcript variant 2, mRNA	375	375	2%	2e-101	89%	NM_001003941.2
Homo sapiens long intergenic non-protein coding RNA 1521 (LINC01521), long non-coding RNA	374	769	2%	6e-101	89%	NR_120386.1
Homo sapiens X-linked inhibitor of apoptosis, E3 ubiquitin protein ligase (XIAP), transcript variant 3, non-coding RNA	374	374	2%	6e-101	88%	NR_037916.1
Homo sapiens X-linked inhibitor of apoptosis, E3 ubiquitin protein ligase (XIAP), transcript variant 2, mRNA	374	374	2%	6e-101	88%	NM_001204401.1
Homo sapiens X-linked inhibitor of apoptosis, E3 ubiquitin protein ligase (XIAP), transcript variant 1, mRNA	374	374	2%	6e-101	88%	NM_001167.3
Homo sapiens chromosome 11 open reading frame 58 (C11orf58), mRNA	374	374	1%	6e-101	90%	NM_014267.5
Homo sapiens TLC domain containing 2 (TLCD2), mRNA	372	712	3%	2e-100	89%	NM_001164407.1
Homo sapiens aldehyde dehydrogenase 6 family member A1 (ALDH6A1), transcript variant 3, mRNA	368	1232	3%	3e-99	89%	NM_001278594.1
Homo sapiens aldehyde dehydrogenase 6 family member A1 (ALDH6A1), transcript variant 2, mRNA	368	1232	3%	3e-99	89%	NM_001278593.1
Homo sapiens aldehyde dehydrogenase 6 family member A1 (ALDH6A1), transcript variant 1, mRNA	368	1232	3%	3e-99	89%	NM_005589.3
Homo sapiens ribosomal protein S6 kinase A6 (RPS6KA6), mRNA	368	368	2%	3e-99	88%	NM_014496.4
Homo sapiens ADP ribosylation factor like GTPase 16 (ARL16), mRNA	368	624	3%	3e-99	88%	NM_001040025.1
Homo sapiens P2RX5-TAX1BP3 readthrough (NMD candidate) (P2RX5-TAX1BP3), long non-coding RNA	366	633	1%	1e-98	90%	NR_037928.1
Homo sapiens LIM homeobox transcription factor 1 beta (LMX1B), transcript variant 2, mRNA	366	366	2%	1e-98	86%	NM_001174147.1
Homo sapiens LIM homeobox transcription factor 1 beta (LMX1B), transcript variant 3, mRNA	366	366	2%	1e-98	86%	NM_001174146.1
Homo sapiens LIM homeobox transcription factor 1 beta (LMX1B), transcript variant 1, mRNA	366	366	2%	1e-98	86%	NM_002316.3
Homo sapiens zinc fingers and homeoboxes 3 (ZHX3), mRNA	366	615	3%	1e-98	89%	NM_015035.3
Homo sapiens WD repeat domain 31 (WDR31), transcript variant 4, non-coding RNA	364	571	3%	3e-98	89%	NR_134677.1
Homo sapiens WD repeat domain 31 (WDR31), transcript variant 2, mRNA	364	571	3%	3e-98	89%	NM_001006615.2
Homo sapiens WD repeat domain 31 (WDR31), transcript variant 1, mRNA	364	571	3%	3e-98	89%	NM_001012361.3
Homo sapiens WD repeat domain 31 (WDR31), transcript variant 3, mRNA	364	571	3%	3e-98	89%	NM_145241.4
Homo sapiens transmembrane p24 trafficking protein 4 (TMED4), transcript variant 2, mRNA	363	363	1%	1e-97	89%	NM_001303058.1
Homo sapiens transmembrane p24 trafficking protein 4 (TMED4), transcript variant 4, mRNA	363	363	1%	1e-97	89%	NM_001303060.1
Homo sapiens zinc finger protein 665 (ZNF665), mRNA	363	654	3%	1e-97	88%	NM_024733.3
Homo sapiens coiled-coil domain containing 157 (CCDC157), transcript variant 1, mRNA	361	361	2%	4e-97	88%	NM_001318334.1
Homo sapiens coiled-coil domain containing 157 (CCDC157), transcript variant 2, mRNA	361	361	2%	4e-97	88%	NM_001017437.3
Homo sapiens uncharacterized LOC101929450 (LOC101929450), transcript variant 3, long non-coding RNA	359	359	2%	2e-96	86%	NR_125816.1
Homo sapiens family with sequence similarity 217 member B (FAM217B), transcript variant 3, mRNA	359	359	2%	2e-96	88%	NM_001190827.1
Homo sapiens family with sequence similarity 217 member B (FAM217B), transcript variant 2, mRNA	359	359	2%	2e-96	88%	NM_001190826.1
Homo sapiens family with sequence similarity 217 member B (FAM217B), transcript variant 1, mRNA	359	359	2%	2e-96	88%	NM_022106.2
Homo sapiens DENN domain containing 6A (DENND6A), mRNA	359	886	3%	2e-96	89%	NM_152678.2
Homo sapiens chromosome 1 open reading frame 43 (C1orf43), transcript variant 7, mRNA	357	701	3%	6e-96	89%	NM_001297721.1

:://blast.ncbi.nlm.nih.gov/Blast.cgi

Figure 13: (A) RefSeq RNA BLAST of intron 6 of RECK gene showing the 5'CS and 3'CS. (B) Table of hits returned from (A)

The 5’CS and 3’CS was then extracted and aligned with all the 15 introns of the RECK gene. Interestingly, introns within the vicinity of intron 6 contains 5’CS and 3’CS. The results obtained were summarized into table (table 4). It was observed that a similarity of sequences was seen in the introns closer to intron 6 while other introns further ahead show no similarity. The widespread presence of this sequence in the gene might indicate that this is a retrotransposon.

Table 5: Table of sequence alignments of all introns of RECK gene with the 5’ and 3’ consensus sequence showing similarities and non-similarities

Introns	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
5’ CS	-	-	-	-	-	*	*	*	-	-	-	-	*	-	-	-	*	*	-	-
3’ CS	*	-	-	*	-	*	-	-	-	*	-	*	*	-	*	-	-	*	-	-

Legend: (*) similarity; (-) no similarity.

To further confirm this, we aligned the 5’CS and 3’CS with the ALU sequence and surprisingly, a striking hit was obtained for 3’CS (Fig: 14 A) while no alignment was seen for 5’CS. Moreover, we tried to align the 5’CS along with 3’CS. Partial alignment at the termini was obtained showing that these two sequences are partially complementary to each other and may be involved in circularization of the RNA. It is probable that these regions were acquired from circular viral-like molecule. However further studies are required in this area (Fig: 14 B).

A

BLAST ® » blastn suite-2sequences » RID-FF80B56P114

BLAST Results

Blast 2 sequences

CS2 on ALU

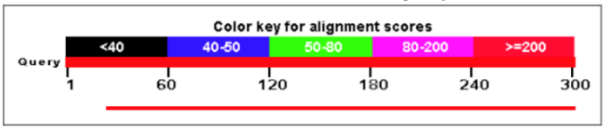
RID FF80B56P114 (Expires on 03-28 16:55 pm)

Query ID lcl|Query_4259
Description None
Molecule type nucleic acid
Query Length 302

Subject ID lcl|Query_4261
Description None
See details
Molecule type nucleic acid
Subject Length 282
Program BLASTN 2.3.1+

Graphic Summary

Distribution of 1 Blast Hits on the Query Sequence



Dot Matrix View

Descriptions

Sequences producing significant alignments:

Description	Max score	Total score	Query cover	E value	Ident	Accession
No definition line found	265	265	92%	1e-75	84%	Query_4261

Alignments

No definition line found

Sequence ID: lcl|Query_4261 Length: 282 Number of Matches: 1
Range 1: 5 to 282

Score	Expect	Identities	Gaps	Strand	Frame
265 bits(143)	1e-75()	234/279(84%)	2/279(0%)	Plus/Minus	

Features:

```
Query  25   GAGACAGAGTTTCGCTCTTGTTTCCCAGGCTGGAGTGTAGTGGCATGATCTCGGCTCACT  84
Sbjct  282  GAGACAGGGTCTCGCTCT-GTCGCCCAGGCTGGAGTGCAGTGGCGCGATCATAGCTCACT  224
Query  85   GCAACCTCCGCCTCCCAGGTTCAAGTGATTCTCTTGCCTCAGCCTCCTGAGTAGCTAGGA  144
Sbjct  223  GCAGCCTCGAACTCCTGGGCTCAAGCGATCCTCCTGCCTCAGCCTCCCGAGTAGCTGGGA  164
Query  145  TTACAGGCATGCACCACCACGCCCGGCTAATTTT-GTATTTTCAGTAGAGACAGGGTTTC  203
Sbjct  163  CTACAGGCGCGCGCCACCACGCCCGGCTAATTTTTGTATTTTTTGTAGAGACGGGGTCTC  104
Query  204  TCCATGTTTTTCAGGCTGGTCTCGAACTCCCGACCTCAGGTGATCCTCCCACCTCGGCCT  263
Sbjct  103  GCTATGTTGCCCAGGCTGGTCTCGAACTCCTGGGCTCAAGCGATCCTCCCGCCTCGGCCT  44
Query  264  CCCAAAGTGCTGGGATTACAGGTGTGAGCCACCTCGCCC  302
Sbjct  43   CCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCC  5
```

B

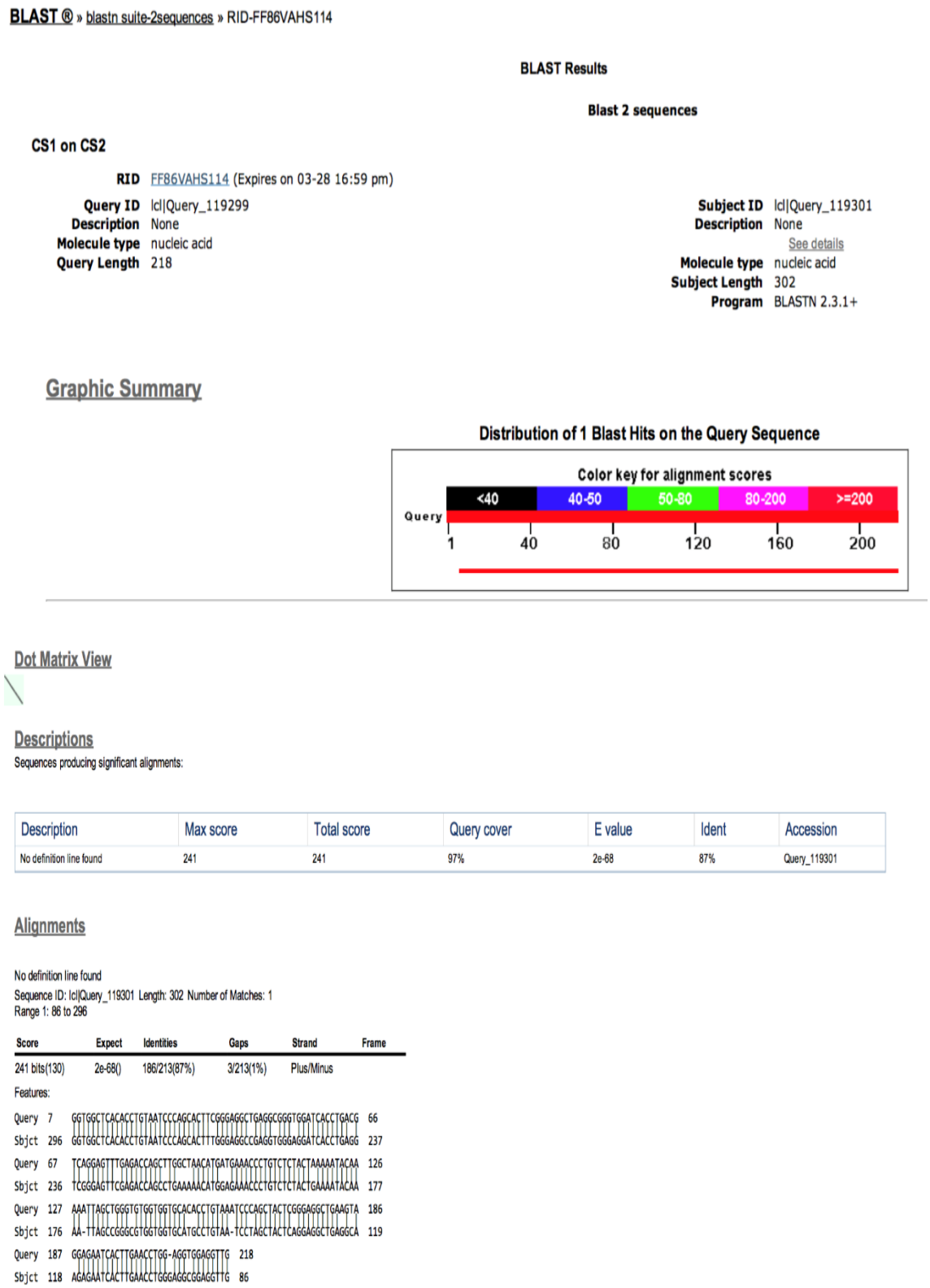


Figure 14: (A) Blast query result on alignment of 3'CS to ALU sequence (B) Blast query result on alignment of 5'CS to 3'CS showing a perfect alignment

Also, a RefSeq Genome blast of intron 6 was performed to check if 5'CS and 3'CS are present in introns of other genes. A 200 blast hit was returned showing that 3'CS was present in different genes located in different chromosomes while 5' CS did not produce any results (Figure:15). In addition, 3'CS is similar to ALU and is widespread over the whole genome. Hence, we believe that 3'CS belongs to retrotransposons. Furthermore, the majority of 3'CS falls in the intronic regions.

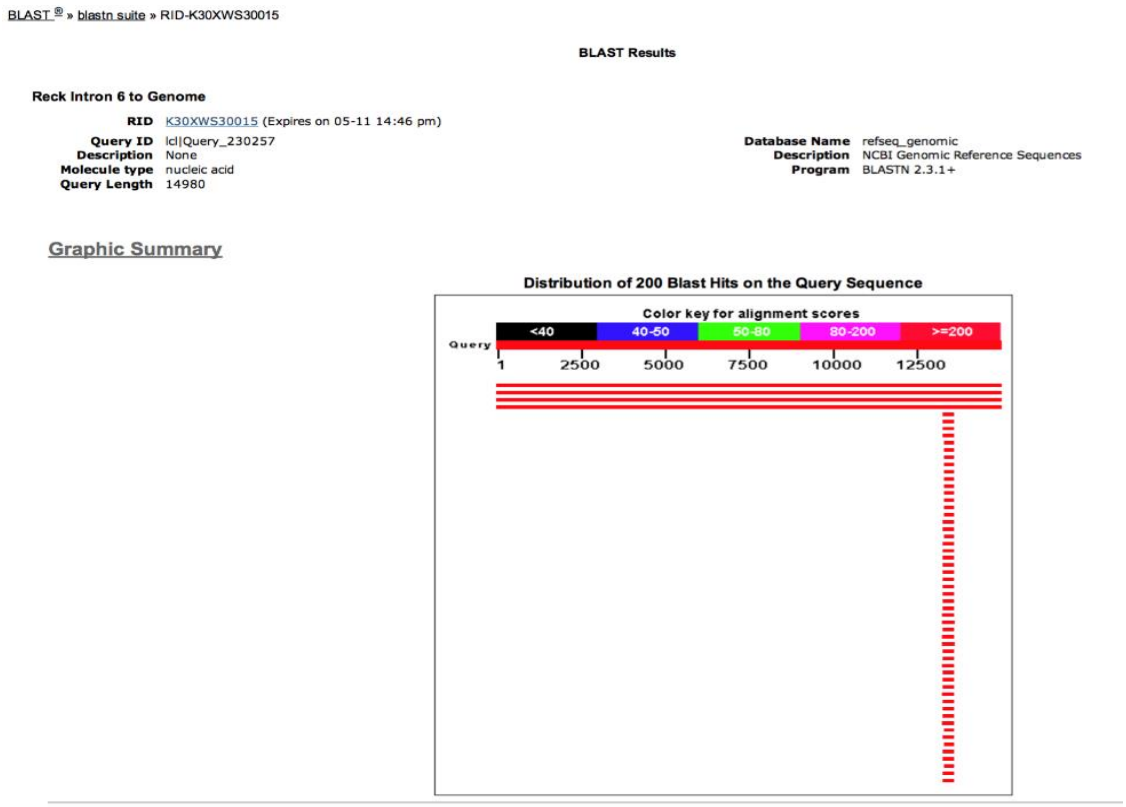


Figure 15: RefSeq Genomic blast query result of intron 6 showing the similarity of 3'CS to many other genes

For further analysis, 10 genes from similarity hits results (Fig: 12) were selected as presented in table 5. It is clear that 5'CS and 3'CS are embedded in exonic regions of various category of genes. The two main region 5'CS and 3'CS show no phylogenic clustering.

Table 6: Summary of the position of 5'CS and 3'CS in 10 selected genes and the element of the gene in which it is contained

S.No.	Name of Gene	Length of mRNA (bp)	Position of 5' CS on intron 6	Position of 5' CS on gene	Position of 3' CS on intron 6	Position of 3' CS on gene	Region of similarity on gene
1	Homo sapiens NUAK family kinase 2 (NUAK2), mRNA	3466	3654-3871	2670-2886	13281-13582	2586-2886	STS
2	Homo sapiens TLC domain containing 2 (TLCD2), mRNA	5901	3664-3873	2432-2640	13289-13585	2419-2716	STS
3	Homo sapiens ribosomal protein S6 kinase A6 (RPS6KA6), mRNA	8223	-	-	13269-13566	7134-7440	STS
4	Homo sapiens ADP ribosylation factor like GTPase 16 (ARL16), mRNA	1207	3660-3873	911-1123	13278-13576	911-1207	Exon,Poly A site,regulatory
5	Homo sapiens zinc fingers and homeoboxes 3 (ZHX3), mRNA	9991	3650-3873	8511-8734	13289-13585	8512-8808	STS

6	Homo sapiens zinc finger protein 665 (ZNF665), mRNA	3291	3651-3873	2154-2376	13285-13582	2157-2455	Exon, STS`
7	Homo sapiens DENN domain containing 6A (DENND6A), mRNA	4679	3650-3873	3462-3684	13293-13576 13279-13585	3391-3674 2670-2976	Exon
8	Homo sapiens calcium modulating ligand (CAMLG), mRNA	2255			13293-13575	1769-2052	Exon
9	Homo sapiens 2-oxoglutarate and iron dependent oxygenase domain containing 1 (OGFOD1), mRNA	3020			13278-13587	2373-2682	Exon
10	Homo sapiens transmembrane protein 170B (TMEM170B), mRNA	8413			13278-13585	4890-5200	Exon

Chapter 4: Conclusion

In this study we examined the significance of HHR experimentally and computationally. We hypothesize that HHR is involved in mRNA biogenesis. We have selected HHR sequence embedded in intron 6 of RECK gene which is 77bp. It comprises of naturally occurring helix II motifs flanked by conserved U and P boxes.

Our analysis clearly shows that HHR in intron 6 is involved in RECK mRNA biogenesis as there is decreased RECK protein level in HHR knockdown cells. This observation is strengthened with the concurrent decline in mRNA level. The discrepancy in the level of exon 4,6,7 and 8 suggest that knockdown of HHR sequence with antisense results in aberrant mRNA splicing. To sum up our observations we hypothesize that HHR might be involved in proper RNA processing and stability

In the database search it was seen that the intron 6 has two consensus sequences namely 3'CS and 5'CS which were also present in many other genes. These sequences were mainly present in the regulatory regions of mRNA. On further analysis of the 3'CS and 5'CS we found that these sequences were partially complementary to each other and might be involved in the circularization of the mRNA. However, further studies are required to establish this. Furthermore, the 5'CS and 3'CS were not limited only to intron 6, but it appeared in other introns of RECK gene which supports the mRNA circularization step. Although 3'CS shows similarity of ALU motif, the 5'CS show no similarity to ALU. It was also noted that 3'CS was more widespread in the gene compared to 5'CS. All these observations

suggest that 3’ CS is a retrotransposon. It should be noted that our future study is aiming to decipher the role of the HHR in conjunction with 5’CS and 3’CS.

This study is important as it contributes to other studies around the world aiming towards defining the exact biological function of HHR *in-vivo*. The mechanism of interference of HHR with the RNA processing or stability opens a new field for further future studies.

References

Adrian RFA and William GS (2010): Small Self –Cleaving Ribozymes, Cold spring Harbor Perspective in Biology, 10: 1-9.

Amiri KMA and Hagerman PJ (1996): The global conformation of an active hammerhead RNA during the process of self-cleavage. J Mol Biol., 16:125-34.

Arnold G , Eliza W, Brigit B , Ricarda , J , Volker AE, and Jens E. (2003): Comparison of different antisense strategies in mammalian cells using locked nucleic acids,2¢-O-methyl RNA, phosphorothioates and small interfering RNA. Nucleic Acid Research, 31: 3185 – 3190.

Christian H , Andrej L, and Jonathan, P (2012): The ubiquitous hammerhead ribozyme. RNA 18:871-883.

Egholm M., Buchardt O, Christensen L, Behrens C., Freier SM, Driver, DA, Berg RH, Kim SK, Norden B, and Nielsen PE (1993): PNA hybridizes to complementary oligonucleotides obeying the Watson-Crick hydrogen-bonding rules. Nature, 365: 566 –568.

Ferbeyre G, Bourdeau V, Pageau M, Miramontes P, and Cedergren R (2000).Distribution of hammerhead and hammerhead-like RNA motifs through the GenBank. Genome Res 10: 1011–1019.

Ferbeyre G, Smith JM, and Cedergren R (1998): Schistosome satellite DNA encodes active hammerhead ribozymes. Mol Cell Biol 18: 3880–3888.

García-Robles, Inmaculada, Jesús Sánchez-Navarro, and Marcos de la Peña. 2012. "Intronic hammerhead ribozymes in mRNA biogenesis." Biological Chemistry 393, no. 11: 1317-1326.

Hartmann TB, Thiel D, Dummer R, Schadendorf D, Eichmüller S (2004): SEREX identification of new tumour-associated antigens in cutaneous T-cell lymphoma. Br J Dermatol 150: 252–258.

Hoem G, Raske CR, Garcia-Arocena D, Tassone F, Sanchez E, Ludwig AL, Iwahashi CK, Kumar M, Yang JE, Hagerman PJ (2011): CGG-repeat length threshold for FMR1 RNA pathogenesis in a cellular model for FXTAS. Hum. Mol. Genet. 20:2161–2170.

Jonathan P, Zasha W, Adam R, Olivia P, Pascal C, and Roland, RB (2011): Identification of HammearheadRibozumes in All Domains of Life Reveals Novel Structure Variations , PLoS computational biology, 7: 1-10.

Khvorova, A., Lescoute, A., Westhof E. &Jayasena, S. D. (2003): Sequence elements outside the hammerhead ribozyme catalytic core enable intracellular activity. Nature Struct. Biol. 10, 708–712.

La Pen˜a M and Garcia-Robles I (2010): Ubiquitous presence of the hammerhead ribozyme motif along the tree of life. RNA 16: 1943–1950

Marcos DLP, and Inmaculada, GR (2010): Intronic hammerhead ribozymes are ultraconserved in the human genome, EMBO reports, 11: 711-716.

Marcos DLP, and Inmaculada GR (2010). Intronic hammerhead ribozymes are ultraconserved in the human genome, EMBO, 11: 711-716.

Prody M, Horan LH, Noller HF, and Scott WG (1): A discontinuous hammerhead ribozyme embedded in a mammalian messenger RNA. Nature 454: 899–902.

Martick, M. and Scott, W. G. (2006): Tertiary contacts distant from the active site prime a ribozyme for catalysis. Cell 126, 309–320 (2006).

Perreault J, Weinberg Z, Roth A, Popescu O, Chartrand P, Ferbeyre G, and Breaker RR (2011): Identification of hammerhead ribozymes in all domains of life reveals novel structural variations. PLoSComputBiol 7

Lieber M, et al. A continuous tumor-cell line from a human lung carcinoma with properties of type II alveolar epithelial cells. Int. J. Cancer 17: 62-70, 1976. PubMed: 175022

Bergers, Gabriele, Rolf Brekken, Gerald McMahon, Thiennu H. Vu, Takeshi Itoh, Kazuhiko Tamaki, Kazuhiko Tanzawa, et al. “Matrix Metalloproteinase-9 Triggers the Angiogenic Switch during Carcinogenesis.” Nat Cell Biol 2, no. 10 (Oktober 2000): 737–744. doi:10.1038/35036374.