Transcriptional regulation of the Human VIPR2 gene

by

Ganaesh Kumaar Ramanathan

A thesis presented in fulfilment of the requirements for the degree of

Master of Philosophy Molecular Cell Biology

2012

Strathclyde Institute of Pharmacy and Biomedical Sciences

University of Strathclyde

Abstract

Human vasoactive intestinal peptide receptor -2 (VIP-R-2) mediates the wide range of physiological processes such as organogenesis, circadian rhythm, immune-modulation and energy homeostasis through its ligands VIP and PACAP. The knowledge concerning the promoter region, transcription factor binding sites are vital for the drug development. The lack of knowledge concerning the organisation of the promoter region slowed the progress in the development of novel therapeutics. The work done on this thesis focused on the exhaustive characterization of the organisation of the promoter and other regulatory regions of the human *VIPR2* gene.

The organisation of the promoter region of the human *VIPR2* gene was characterised using bioinformatics tools and the evolutionary conservation analysis indicated that the organisation of the promoter region and the cis-elements are evolutionarily resilient. The length of the minimal promoter region found to be 408bp and spans 240 bp upstream and 168 bp downstream in relation to the translation start site. The minimal promoter found to have no TATA box, CAAT box or initiator element, but there are multiple GC boxes. Two GC box identified in the region upstream to the start codon and other two GC box are found in downstream to the translation start codon. Several evolutionary conserved elements responsible for the tissue specific receptor expression in adipocytes, neurons, activated lymphocytes, pituitary cells, lung epithelial cells and myocytes. The novel elements in the Intron-1 responsible for developmentally regulated expression of the human VIP-R-2 receptor were also identified. Cloning the Intron-1 region with reporter vector for functional studies was also attempted. This study has identified several key development related and tissue specific regulatory elements in the intron-1 and in the 6kb 5' flanking region of the

human *VIPR2* gene and the findings suggests the role of VIP-R-2 receptor in organogenesis and embryogenesis.

Dedicated to Mom-Dad and Sis

Acknowledgements

I would like to express my heartfelt gratitude to my supervisors Dr Elizabeth Ellis and Dr Eve Lutz, for their patience, guidance and support during the entire course of the research study. I am specifically indebted to my second supervisor for her valuable advice, guidance, and patience during the write-up and thankful to my first supervisor for her insightful comments during the project, which I hope I would not lose throughout my life. I gained much knowledge, experience and self-confidence throughout this study. Without their support I don't believe that I would improve myself that much.

I am also thankful to all student and staff members in SIPBS for all the support I got from them. Last but not least, I would like to thank all my friends Srinivas, Biju, Arul, Abhiramavalli, Preethi, Thiagarajan and Vishnu for their friendship and support.

Ganaesh.R

Copyright statement

Signed: Date:

List of Abbreviations

ADCYAP1R1	Adenylate cyclase activating polypeptide 1 (pituitary) receptor type I
a-FTP	Alpha-fetoprotein
AML	Acute Myeloid Leukemia-1Protein
AP1	Activating protein 1
APRT	Human Adenine phosphoribosyl transferase
Aprt	Mouse Adenine phosphoribosyl transferase
ATP	Adenosine tri phosphate
BDGP	Berkeley Drosophila Genome Project
bp	Base pair
BRN	Brain Class III POU transcription factor
C/EBP	CCAAT/Enhancer binding protein
cAMP	cyclic Adenosine Mono Phosphate
CD-97	Clusters of Differentiation-97
cdc2	Cell division cycle 2
cDNA	Complementary Deoxyribo Nucleic Acid
CDPCR	cut like homoedomain protein-3
CDXA	Caudal type homeobox transcription factor-Alpha
CEBP	CCAAT Enhancer binding protein transcription factor
CEBPgamma	CCAAT Enhancer Binding Protein gamma
CIRL	Calcium independent receptor for Latrotoxin
CL1	CIRL/Latrophilin 1
COUP	Chicken Ovalbumin upstream promoter transcription factor

CpG	Cytosine nucleotide-phosphate-Guanine nucleotide
CREB	cAMP response element binding protein
CREBATF	cAMP response element binding protein/Activating transcription factor
DECR	2,4-dienoyl-CoA reductase
dNTPs	Deoxyribonucleotide triphosphate
DTT	Dithiothreitol
E2A	Immunoglobulin Enhancer binding factors E12/E47
EBOX	Enhancer Box
EDTA	Ethylenediaminetetraacetic acid
EGF LAG	Epidermal growth factor like repeats Laminin A G type repeats
EGF-TM7	TM7 containing Epidermal Growth Factor like domains
EMR1	EGF-like module-containing mucin-like hormone receptor-like 1
EMR2	EGF-like module-containing mucin-like hormone receptor-like 2
EMR3	EGF-like module-containing mucin-like hormone receptor-like 3
endA1	Endonuclease-deficient mutant endA1
ETF	Epidermal Growth Factor Receptor-specific transcription factor
FRA	fos-related antigen
FREAC	Forkhead related activator transcription factor
FXR	Farnesoid X receptor
GPCR	Guanine nucleotide binding protein coupled receptor
Gs	Heterotrimeric G-protein that stimulates adenylate cyclase
GTE	Glucose Tris EDTA
gyrA96	Mutant form of DNA gyrase

HCl	Hydrochloric acid
HNF alpha	Hepatocyte Nuclear Factor alpha
HNF	Hepatocyte Nuclear factor
hsdR17	Restriction endonuclease deficient mutant hsdR17
HTF	Hepatocarcinogensis related transcription factor
IC50	Half maximal inhibitory concentration
IK2	Ikaros-2 transcription factor
IL-4	Interleukin-4
IL-6	Interleukin-6
kb	Kilo base
KID	Kidney transcription factor
KoAc	Potassium Acetate
LB broth	Luria Bertani broth
Ldh-A	Lactate dehydrogenase A
LEF1	Lymphoid Enhancer Binding Factor-1
LNB-TM7	TM7 containing Long N-termini structurally similar to Group B GPCRs
GPCRs	G-Protein Coupled Receptors
LN-TM7	TM7 containing Long N-termini
LRF	Liver regeneration factor
LyF	Lymphoid transcription factor-1
M	molar
MAZ	Myc associated Zinc finger protein
MgCl2	Magnesium Chloride

Mgl	Macrophage galactose/*N*-acetylgalactosamine-specific C-type lectin
mRNA	Messanger Ribo Nucleic Acid
Myc	Myelocytomatosis
NaOH	Sodium Hydroxide
NFAT1	Nuclear factor of activated T cells-1
NFKappaB	Nuclear factor Kappa light chain enhancer of activated B cells
NFY	Nuclear transcription Factor Y
NKx	Cardiac homeobox transcription factor
nm	Nanometre
NRSF	Neuron restrictive silencing factor
Oct-1	Octomer binding protein-1
OG2	Osteocalcin gene 2 transcription factor
PAC1	Pituitary adenylate cyclase-activating polypeptide receptor 1
PAX	Paired Box
PBX	Pre-B cell Leukemia transcription factor
PCNA	Proliferating Cell Nuclear Antigen
PCR	Polymerase Chain Reaction
PEG	Poly ethylene glycol
pH	per Hydrogen concentration
PHI	Peptide Histidine Isoleucine
PHM	Peptide Histidine Methionine
PHV	Peptide Histidine Valine
Pit-1A	Pituitary specific positive transcription factor-1A

POU	Pituitary specific Pit-1A/Oct-1, Oct-2/Unc-86 Neural transcription factor
POU3F	POU domain class 3 transcription factor
PPARA	Peroxisome proliferator-activated receptor alpha
recA1	Recombinase-deficient mutant recA1
relA1 lac	relaxed phenotype permits RNA synthesis without protein synthesis
RFX	Regulatory Factor X
Rps19	Ribosomal protein S 19
SDS	Sodium Dodecyl Sulphate
SOC	Super Optimal broth with catabolite repression
Sp1	Specificity protein 1
SREBP	Sterol regulatory element binding protein
STAT	Signal Transducers and Activators of Transcription protein
supE44	Amber (UAG) codon suppressor
TAE	Tris base/Acetic acid/EDTA
TBP	TATA Binding Protein
TBX	T-Box
TCF	T-cell specific transcription factor
TCF	T-cell specific transcription factor 7
TE	Tris/EDTA
TFE	Transcription factor E
TFIIA	Transcription factor II A
TFIIB	Transcription factor II B
TFIID	Transcription factor II D

TFIIE	Transcription factor II E
TFIIF	Transcription factor II F
TFIIH	Transcription factor II H
thi-1	requires thiamine
TITF	Thyroid transcription factor
TK	Thymidine Kinase
TM7	Seven span transmembrane molecules
TMD	Transmembrane domain
T.S.S	Transcription start site
VDR	1,25-dihydroxyvitamin D3 receptor
VMYB	myb related transcription factor
w/v	Weight/Volume
WT1	Wilms Tumour protein-1
YY1	Yin Yang-1
ZF5	Zinc finger protein 5

APPENDIX

List of Figures

List of Tables

Chapter 1
Introduction

1.1 General Introduction

Human vasoactive intestinal peptide receptor 2 (human VIP-R-2; VIP2R; $VPAC_2$) is a member of the G-protein coupled receptor family. Its role is to act as a receptor for neuropeptides, in particular the human vasoactive intestinal peptide (VIP) and the human pituitary adenylate cyclase-activating polypeptide (PACAP). These neuropeptides have a variety of effects on those cells that express the VIP-R-2 receptor, including controlling water secretion in the gut, and allowing signals to move between neurons. Factors that control the expression of the *VIPR2* gene that encodes the human VIP-R-2 receptor are of vital importance as they influence whether cells can respond to VIP and PACAP.

In this Chapter, a brief introduction to PACAP and VIP will be given, including their site of production and biological function. A detailed description of the human VIP-R-2 receptor is also provided, including the organization of the *VIPR2* gene in the genome, its tissue distribution and a brief introduction to the transcriptional regulation of the *VIPR2* gene including the factors known to influence its expression will be provided.

1.2 Peptides that bind to the VIP-R-2 Receptor

1.2.1 Vaso-active-intestinal Peptide (VIP)

VIP was first identified in 1970 in porcine small intestine as a 28 amino acid peptide that dilated the canine femoral artery (Said and Mutt, 1970, 1972). The gene

encoding human VIP was subsequently cloned in 1985, and was mapped to chromosome 6. Translation of the human VIP mRNA sequence yields a 170 amino acid precursor named pre-pro-VIP. Post-translational processing of the precursor yields the 28 amino acid peptide VIP in addition to a 27 amino acid peptide called peptide-histidine-methionine (PHM) in humans or peptide- histidine-isoleucine (PHL) in rodents (Itoh *et al*., 1983). Two alternatively processed forms of the latter peptide have been identified, namely peptide-histidine-valine and -glycine. The peptides VIP and PHM/PHI are located in adjacent exons, and share 48% of their amino acid sequence identity.

Human VIP is expressed in a range of tissues including the liver, resipiratory system, endocrine system, nervous, immune system and gastrointestinal system of the mammals. VIP-containing nerve fibers are can be identified in most lymphoid organs of the respiratory and gastrointestinal tract (Pearse *et al*., 1977; Bellinger *et al*., 1990; Kulkarni-Narla *et al*., 1999; Dey *et al*., 1981).

VIP has been associated with functions such as neurotransmitter release (Duckles and Said, 1982), neuro-protection (Brenneman and Eiden, 1986), glycogen metabolism (Sorg and Magistretti, 1992), prolactin secretion from the pituitary (Reichlin, 1988), adrenal medullary secretion of catecholamine (Malhotra *et al*., 1988), regulation of the T cell (Ottaway, 1987), electrolyte secretion in ileal mucosal cells, as a smooth muscle relaxant and protection against oxidative injury (Gozes and Brenneman, 1989; Laburthe *et al*., 1993; Schwartz *et al*., 1974; Said, 1991, 1996). Other studies identified several important functions of VIP in pain perception [Dickinson *et al*., 1999), suppression of inflammation (Said, 2000) and immunomodulation (Delgado and Ganea, 2001). VIP has been associated with

watery diarrhea syndrome, impotence, asthma, lung injury, a variety of tumors and neurodegenerative diseases [Said, 2000; Gozes *et al*., 1999). VIP appears to regulate the immune system by decreasing and delaying the interleukin-2 production and inhibiting the proliferative response of mitogen-activated T cells (Ottaway, 1987). Another study identifies VIP as a growth regulator for foetus and early brain development (Gressens *et al*., 1993; Waschek *et al*., 1995; Gozes *et al*., 1999).

1.2.2 Pituitary Adenylate Cyclase-Activating Polypeptide (PACAP)

PACAP exist into two forms, one as PACAP-38 which is a 38 amino acid peptide and the other as the 27 amino acid peptide PACAP-27 which has 68% sequence similarity with VIP; both forms were isolated from the ovine hypothalamus (Miyata *et al*., 1989; Miyata *et al*., 1990). As PACAP-27 shows a 68% identity to VIP, this identifies PACAP as a member of the VIP-glucagon-GRF-secretin superfamily of structurally-related peptides.

PACAP is present in the gastrointestinal tract, testis, and adrenal gland (Arimura and Shioda, 1995; Ghatei *et al.,* 1993), is abundant in hypothalamus, and to lesser level in other regions of the central nervous system (Ghatei *et al*., 1993). PACAP is produced in retinal afferents, and is known to function as a biological clock regulator. In the periphery it is thought to function as a non-cholinergic neurotransmitter, and as an adrenal medullary catecholamine stimulator (Przywara *et al*., 1996) and regulator of the pancreatic exocrine and endocrine system (Yada *et al*., 1994).

Two distinct receptors for PACAP have been identified in ligand binding studies, Type I and type II. Where Type II bind both VIP and PACAP with high affinity, Type I has much greater affinity for PACAP than VIP (Ohtaki *et al*., 1993).

1.3 The VIP-R-2 receptor

1.3.1 Introduction to the VIP-R-2 receptor

The VIP-R-2 receptor, also known as the $VPAC_2$ receptor, belongs to the Class II subfamily of the seven-transmembrane G-protein-coupled receptor (GPCR) superfamily (Figure: 1). GPCRs are more than 1000 in number, and are considered as one of the largest families of proteins in the mammalian genome (Lee *et al.,* 2004; Fredriksson *et al*., 2003; Lander *et al*., 2001; Venter *et al*., 2001). GPCRs are categorized into six families: Class I: the rhodopsin family (A), Class II: the secretin-receptor family (B), Class III: the metabotropic- glutamate receptor family (C), Class IV: fungal pheromone P- and alpha factor receptors (D), Class V: fungal pheromone A- and M-factor receptors (E), and Class VI: cyclic AMP receptors (F).

The *VIPR2* gene encoding the rat VIP-R-2 receptor was cloned from a rat olfactory bulb cDNA library (Lutz *et al*. 1993). When expressed in cell lines, the recombinant rat and human VIP-R-2 receptors were able to bind to VIP (IC_{50}, 3 to 4 nm), PHI and PHV (IC_{50}, 10 to 30 nm), PACAP-27 (IC_{50}, 10 nm) and PACAP-38 (IC_{50}, 2nm), and could also bind GRF and secretin with a very low affinity (IC_{50}, 5000 to 30000 nm).

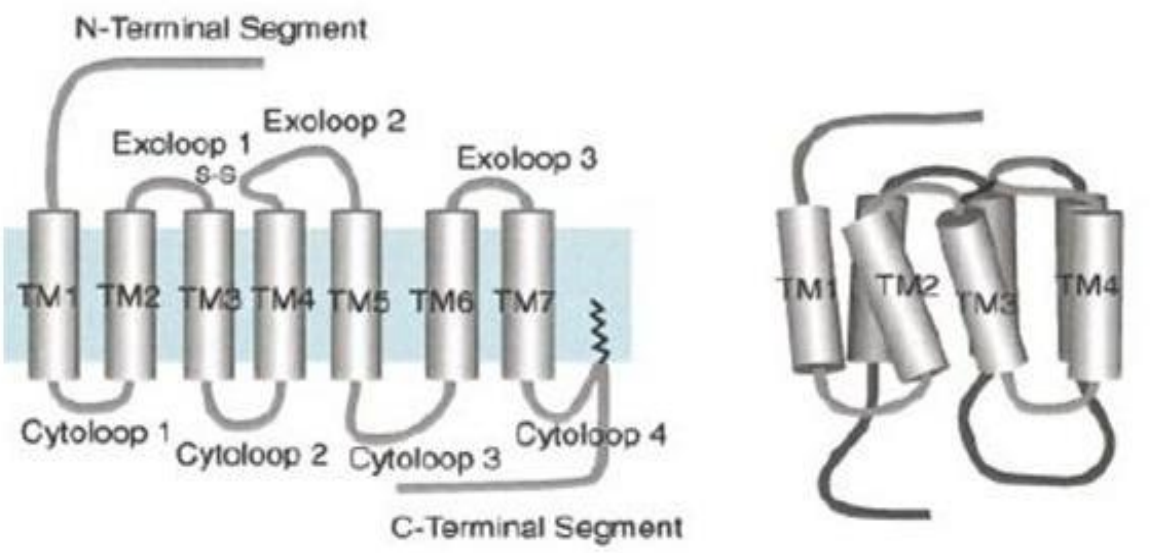


Figure 1.1 Schematic representation of the seven transmembrane GPCR.

This illustration shows the relative position of the three extracellular loops, four intracellular loops, extracellular N-terminal and intracellular C-terminal domain.

(Taken from http://nlp.postech.ac.kr/Research/POSBIOTM/content/struc.html Accessed on 04/07/2009)

1.3.2 Receptors related to VIP-R-2

VIP-R-2 receptors belong to the B1 subfamily of Class II GPCR which are known to regulate intracellular cAMP by coupling to adenylate cyclase through the stimulatory G protein (G_s), though a few other members of this group stimulate G_q and/or G_i/G_o and are involved in other signaling pathways such as phospholipase C (Laburthe *et al*., 2002 and Vaudry *et al*., 2000).

VIP-R-2 is a secretin-related receptor, belong to the group of receptors for large peptide hormones and neuropeptides (Fredriksson *et al*., 2003) (Figure: 2); this family often acts in a paracrine and autocrine fashion. A number of studies have revealed that, although the VIP-R-2 receptor like other Class II receptors has low sequence similarity with the other classes of GPCRs, it shares many characteristic features with its receptor family, such as the presence of a large N-terminal extracellular domain with 10 conserved amino acids including six cysteines, an N-terminal leader sequence, several N-glycosylation sites (Laburthe *et al.*, 1996 and Laburthe *et al.*, 2002) (Figure: 1).

1.3.3 Tissue distribution and biological functions of the VIP-R-2 receptor

The VIP-R-2 receptor is expressed in the thymus, pancreas, macrophages, lymphocytes, heart, kidney, testis, adipose tissue, blood vessels, skin, gastrointestinal tract, prostate, smooth muscle, uterus, cerebellum and several other peripheral tissues (Basille *et al*., 2006; Harmar *et al*., 1998; Gourlet *et al*., 1997). In the central nervous system, VIP-R-2 receptors are found in the thalamus, superchiasmatic nucleus, hippocampus, brain-stem, spinal cord and dorsal root ganglia. (Harmar *et al.*, 1998)

(Table-1). Expression of the VIP-R-2 receptor in CNS, several endocrine tissues (Usdin *et al*., 1994) suggests that the VIP-R-2 receptor plays an important role in neuroendocrine system (Usdin *et al*., 1994).

The expression of the VIP-R-2 receptor in several pituitary derived clonal cell lines such as rat somatomammotroph GH3 and GH4C1 cells, mouse adrenocorticotroph AtT20, gonadotroph T3-1 cells and mouse embryonic stem cells (Cazillis *et al*., 2004; Hirose *et al*., 2005) indicates the role of the VIP-R-2 receptor in the development of pituitary cell lineages, in embryogenesis and organogenesis (Rawlings and Hezareh, 1996; Hezareh *et al*., 1996; Mackenzie *et al*., 2001).

The VIP-R-2 receptor mediates the anti-inflammatory role of VIP (Jurranz *et al*., 2008). Recent research indicates that it also plays an important role in the biological clock (Harmar *et al*, 2002; Cutler *et al*., 2003; Huges *et al*., 2004), growth, male reproduction and basal energy expenditure (Asnicar *et al*., 2002). It is involved in cutaneous active vasodilation (Kellog *et al*., 2010) immune hypersensitivity (Dickson and Finlayson, 2009) and immune regulation in humans (Miller *et al*., 2006). Recent clinical studies implicate the abnormal expression of the VIP-R-2 receptor in several disease processes such as schizhoprenia (Levinson *et al*, 2011), polyps and gall bladder stone formation (Zhang *et al*., 2006).

1.4 Structure of the human *VIPR2* Gene

The VIP-R-2 receptor is encoded by the *VIPR2* gene. The rat *VIPR2* gene was first cloned by Lutz *et al*., 1993 from a rat pituitary cDNA library, followed by the human *VIPR2* gene from SUP-T1 cells by Svoboda *et al*., 1994 and the mouse receptor from

MIN-6 insulin secreting beta cells by Inagaki *et al*., 1994. The structure of the human *VIPR2* gene was characterized by Lutz *et al*., 1999, and the gene was located near the telomere of human chromosome region 7q, the exact location being 7q36.3 with the 5' end of the gene oriented towards the telomere.

The human *VIPR2* gene was cloned from the SUPT-1 lymphoblastic cell line cDNA library and found to consist of 13 exons. The first five exons encode the amino-terminus of the receptor; exons 5-12 encode the transmembrance domain of the receptor and the exon 13 enode the carboxyl tail (Lutz *et al.,* 1999).

The *VIPR2* gene is interupted by varying length of intronic sequences, the largest intronic sequence being intron-4 with the size of 45 kb, and the smallest intronic sequence being intron-11 with the size of 68 bp. It is the one of the largest genes in the Class II GPCR family with the total size of 117 kb. The initiator codon of the 438 amino acid open reading frame is located in Exon-1, and the termination signal and the poly-adenylation signal sequences are located in exon-13. In the human *VIPR2* gene, exon 1 is located within the CpG Island, similar to the mouse *Vipr2* gene. The human *VIPR2* exon which contains the translation start site and the genomic sequence flanking the exon are extremely GC-rich. The mouse *Vipr2* gene which spans 68.6 kb has 12 introns and 13 exons is smaller when compared to human *VIPR2* gene. The intron with the size of 2.1 kb identified between exon-10 and exon-11 found conform to the GT-AG rule of exon/intron splice junctions (Lutz *et al*., 1999). The genomic organization of the human *VIPR2* gene is illustrated in Figure 1.2.

1.4.1 Characterization of *VIPR2* transcripts

The sizes of the *VIPR2* transcripts were determined by northern blot analysis as two hybrizing bands at 4.6 and 2.3 kb (Lutz *et al.,* 1999). The coding sequence contains U-rich regions, AUUUA pentamer sequences, 3' UTR (un-translated regions) suggests the possibility of post-transcriptional regulation of the mRNA, as these regions are found to be important regulatory sites in the mRNA of the other related receptors such as calcitonin receptor mRNA, Beta-adrenergic receptor mRNA and 5HT receptor mRNA. Unlike other genes of the class II GPCR, there are no multiple exons encoded in the 5'UTR in *VIPR2* gene. The translation start site and the 5'UTR are encoded within the same exon.

1.4.2 VIP-R-2 Receptor Splice Variants

Following transcription, introns are removed by RNA splicing to generate a mature RNA product. Existence of the splicing mechanism leading to the generation of slightly different transcripts and hence the availability of the several varied gene products. Recent studies have reported a number of VIP-R-2 receptor splice variants. Grinninger *et al*., 2004 reported a mouse VIP-R-2 receptor splice variant in immune cells that lack a part of the last transmembrane domain. The researchers did perform a stimulation study in the human jurkat T cell line, which expressed low levels of wild type of VIP-R-2 receptor and a deletion isoform. Both receptors bound the same amount of radioactive iodine-labeled VIP with similar affinity. However, unlike the wildtype VIP-R-2 receptor, the deletion isoform did not transduce a VIP-elicited increase in intracellular concentration of cyclic AMP. Miller *et al*., 2006 identified

one short-deletion VIP-R-2 receptor splice variant in mouse lymphocyte and another long-deletion VIP-R-2 receptor splice variant in human lymphocyte.

The mouse splice variant binds the VIP normally but showed reduced VIP dependent signalling and consequent modulation in immune response, whereas human splice variant exhibits reduced binding with VIP and exhibit functional differences. Another research group recently identified and characterized the five-transmembrane isoform of human VIP-R-2 receptor of human Sup-TI cell line (Bokaei *et al*., 2006).

These studies suggest that the VIP-R-2 receptor deletion splice variants are commonly linked to the loss of the third intracellular loop, or the loss of the last two transmembrane segments by skipping of exon 10 and 11 during transcription and deletion of TM5 (transmembrane-5), IC3 (Intracellular loop-3), TM6 (transmembrane-6), EC3 (extracellular loop-3), part of TM7 (transmembrane-7) during the expression, resulting in complete abolition or alteration of signaling abilities.

1.5 Regulation of the *VIPR2* Gene

The VIP-R-2 receptor is not expressed in all tissues or under all conditions, and many factors appear to influence its expression. In addition to differences in tissue-specific expression, a study conducted by Sun *et al*., 2006 found altered expression of the VIP-R-2 receptor in T lymphocytes, and revealed aberrant Th1 immunity in Multiple sclerosis. The investigators suggested that decreased expression of the *VIPR2* gene is a causative factor in Multiple Sclerosis.

Table 1.1 Relative abundance of VIPR2 receptor mRNAs in the rat and human*tissues
[Taken from Vaudry *et al* 2000; Basille et al., 2006*; Wei and Moisov, 1996**]

Rat tissues	**VIP-R-2**
Smooth muscle*	+
Cerebullum*	+
Anterior pituitary	++
Adipose tissue**	+
Intermediate lobe of the pituitary	+
Cortex	++
Thymus	+
Pancreas	+
Pancreatic beta islets	++
Liver	+
Testis	++
Early germ cells	++
Seminiferous tubules	+
Spleen	++
Kidney	+
Lung	+
Tracheo-bronchial wall	+
Stomach	+
Lymphocytes	+
Skeletal muscle*	+

Legend
The following symbols provide a semi-quantitative evaluation of the density of VIPR2 mRNAs. (+++) High density; (++) Moderate density; (+) Low density; (-) No hybridization

Recently, Vacic *et al.*, 2011 conducted a large two-stage genome-wide scan for rare copy number variants and reported the discovery of the micro-duplication of a region of chromosome 7 (7q36.3) located around 89kb upstream of the TSS of the human *VIPR2* gene in schizophrenia patients. The investigators suggested that tandem duplication of regulatory sequences was a leading cause for the altered regulation and consequent overexpression of the *VIPR2* gene in schizophrenia, and indicated that this might be causal in the disease. Taken together, these studies associated the altered regulation of the *VIPR2* gene with disease conditions, yet the disease-causing altered regulatory mechanism has not been characterized in detail. Elucidating the regulatory mechanism will provide insights into the processes involved in disease development which may help in future in developing the molecular diagnosis and treatment for disease conditions Schizophrenia and Multiple Sclerosis.

As transcription is one of the most important ways of regulating gene expression, characterizing the elements involved in the regulation of genes can help understand the factors involved in influencing the expression of the *VIPR2* gene, and can explain tissue specificity, disease processes and organ development (Wray *et al.*, 2003). These aspects are discussed below.

1.5.1 Transcriptional Regulation

Transcription is the process in which the genetic information is transcribed from DNA into RNA. Regulation of transcription can alter gene expression by changing the rate of transcription. This is achieved through specific DNA sequences (promoter elements and enhancer elements) located near the 5' end of the regulated gene, which

are bound by general and specific transcription factors, including activators and repressors) (Figure: 1.3).

1.5.1 Promoters

Promoters are the DNA sequences that are responsible for the initiation and regulation of the transcription of a particular gene. These include the core promoter region which contains the transcription start site, the proximal promoter region which is located up to -250 bp upstream of the transcription start site (TSS); and the distal promoter regions which are located further away from the TSS and which contain additional regulatory elements such as enhancers (see Figure: 1.3). These promoter regions are not precisely defined and can be located several kb away from the TSS.

1.5.2 The core promoter

The core or basal promoter is the minimal portion of the promoter, located near to the start of the gene which plays an important role in initiating transcription and contains the transcription start site (Lewin B, 2008). The core promoter contains elements such as the TATA box and the initiator (Inr), sites that are crucial for initiation of transcription. General transcription factors bind to these sites and form a pre-initiation complex together with RNA polymerase II (Buratowski *et al*., 1989).

1.5.3.1 TATA-Box

In mammalian protein-encoding genes, the core promoter often contains a consensus DNA sequence 5'-TATAAA-3' known as the TATA box, to which TATA binding protein (TBP) binds. This is located approximately 35 nucleotides upstream of the TSS (Kornberg 2007 and Hampsey, 1998).

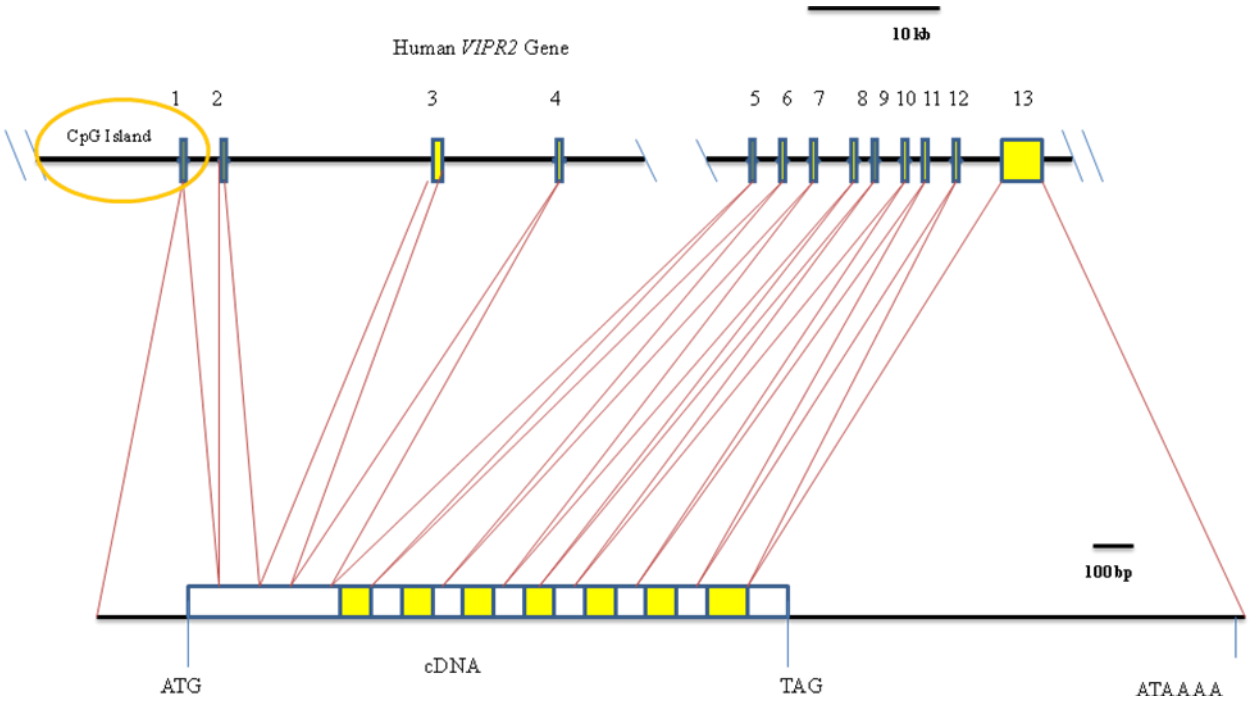


Figure 1.2 Schematic representations of 3974 bp cDNA and corresponding human VIPR2 gene. The exons are shown by vertical lines and numbered from 1 to 13, inbetween the vertical lines (exons) are introns. The human VIPR2 gene is related to its corresponding cDNA by descending lines (adapted from Lutz et al., 1999).

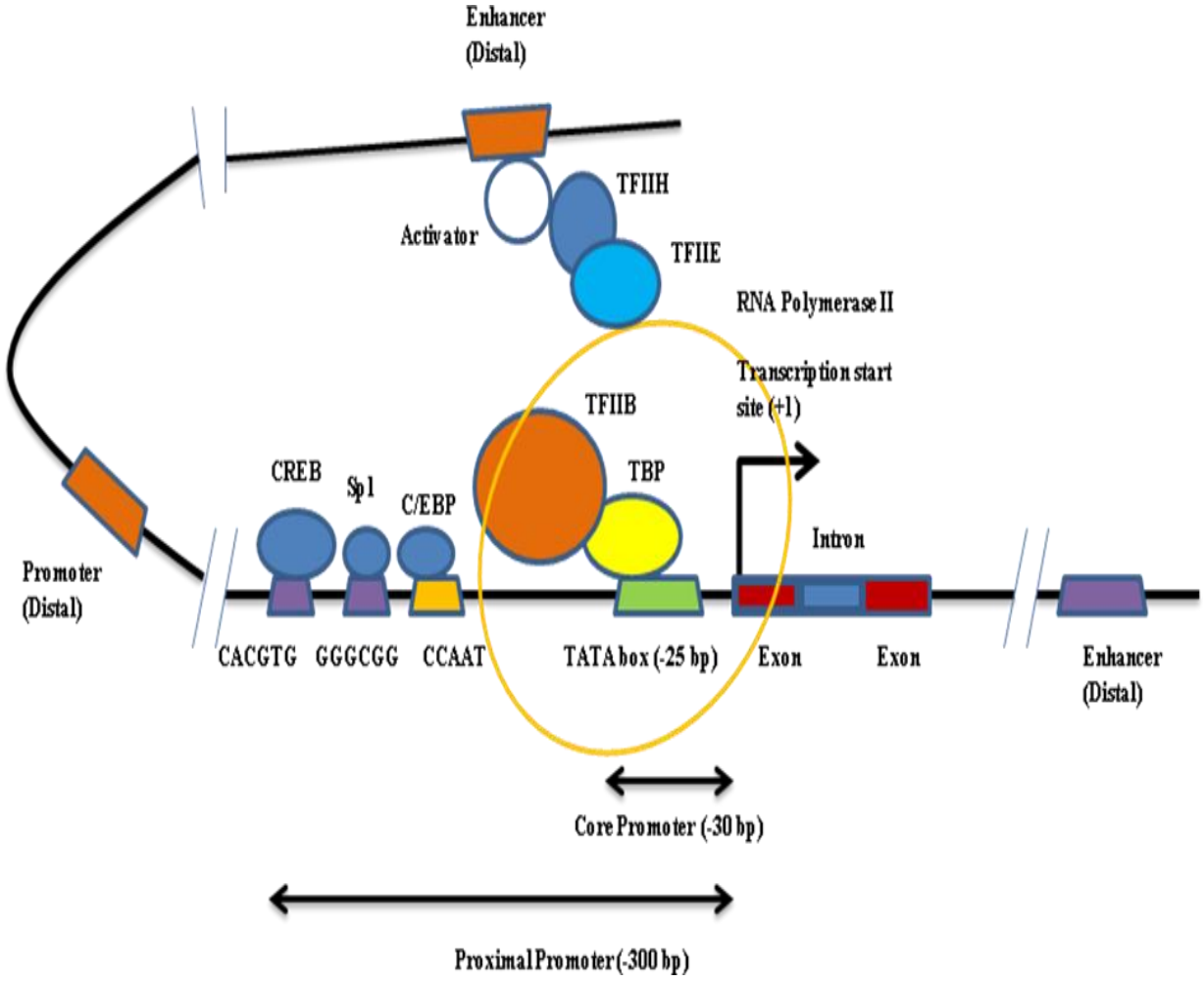


Figure 1.3 Elements found in a typical eukaryotic promoter

1.5.3.2 Initiator Element and Downstream Promoter Element

The core promoter of many eukaryotic genes lacks a TATA box. These are said to be "TATA-less" promoters and they tend to use an alternate means of transcriptional initiation. These promoters instead contain of an Initiator element (Inr) that encompasses the TSS and a short sequence known as the downstream promoter element or DPE.

1.5.3 Proximal Promoter region

The proximal promoter region (also known as the Upstream Control Region or Upstream Promoter Element) is located around 250bp upstream of the transcription start site, and contains binding sites for several specific transcription factors. This region will contain either a CAAT box or a GC box or both. The CAAT box contains a sequence with the consensus 5'- GGCCAATCT-3'. This is located around 75-100 bp upstream of the transcription start site, and is bound by general transcription factors. It is important in transcription initiation and its presence leads to sufficient transcription of a particular gene. The GC box is a short GC-rich sequence with the consensus 5'GGGCGG-3' and is also important in raising the level of basal transcription and in some cases controlling transcription initiation in TATA-less promoters (Blake *et al*., 1990). The GC box is bound by a transcription factor Sp1.

1.5.4 Distal Promoter Regions

Distal promoter regions are located much further upstream or downstream of the transcription start site and include a variety of elements known as enhancers. These

are extremely important in the transcriptional regulation of the particular gene and are often referred to as regulatory elements, controlling tissue-specific expression as well as responses to physiological stimuli and adaptive responses to stress. Enhancers are bound by specific transcription factors that act as activators (or repressors). These factors can induce a loop in the DNA, allowing them to interact with the factors bound to the core promoter to enhance the transcriptional level of the gene.

1.5.5 CpG Islands

Some promoter regions contain stretches of DNA that are rich in cytosine and guanine nucleotides and are known as CpG islands. These regions are associated with genes that are actively expressed. A true CpG island must be >500 bp long with GC content >55% and observed CpG/expected CpG ratio of 0.65 (Takai and Jones 2002). CpG island promoters often possess multiple TSS, unlike TATA Box promoters which have well defined single TSS (Sandelin *et al*., 2007). In humans, CpG islands represent 1% of the genome and are associated with promoters for 70% of human genes, which includes mostly house-keeping genes, and approximately half of genes with tissue restricted gene expression pattern (Antequera and Bird, 1993, 1999; Antequera, 2003). CpG islands are also associated with the majority of genes that are expressed very early in development (Ponger *et al.*, 2001).

CpG regions found outside of the CpG islands upstream of TSS are heavily methylated, while CpGs within the island are non-methylated. CpG islands are usually free of nucleosomes and free of proteins that bind the methylated CpG such as MBD2, MBD3 and MBD4 (Figure: 6) (Hendrich and Bird, 1998). Recent studies

point out that those promoters that are transcriptionally active in the early period of development escape methylation because of the presence of transcription factors attached to the islands and which protect the region from methylation. The resulting absence of methylation in these actively transcribed genes is then imprinted as the cells divide at the totipotency stage, and is transmitted to all subsequent somatic cell lineages. Genes that fall into this category include human alpha-globulin gene, myotonin protein kinase gene, rat enkephalin gene, neurofilament gene and pro-opiomelanocortin gene. (Macleod *et al.*, 1998; Daniels *et al*., 1997, 1995; Gardiner-Garden and Frommer, 1994; Yoshikawa *et al*., 1988; Rachdi *et al*., 2003).

1.5.6 Role of Intronic and Exonic regions in transcriptional regulation

It is generally thought that the 5' sequence upstream of the TSS is the predominant site in which to find regulatory elements, but recent evidence suggests that sequences within introns and exons can also be home to several transcriptional factor binding sites, and may play a significant role in the regulation of gene expression by participating cooperatively with promoter regions (Bai *et al*., 1993; Howard and Davidson, 2004).

Several studies have attributed several regulatory roles for sequences that arise within introns. Recent evidence revealed an enhancer role for sequences contained within an intron, as it can mediate the enhancement of gene expression in several species such as human (Jonsson *et al*., 1990), *C. elegans* (Ho *et al*., 2001), mice (Palmiter., 1991), rice (Jeon *et al*., 2000) and *Arabidopsis* (Ross *et al*., 2003). Another study reported the involvement of intron 1 in the complex pattern of regulation of rat renin gene. In this study, investigators reported the presence of five

negative regulatory elements, and two positive regulatory elements within intron 1; these seven regulatory elements regulated the Upstream Promoter Element depending on the cell in which the gene is located (Voigtlander *et al.*, 1999).

Several other studies also reported many intronic sequences performing several complex regulatory roles as stimulatory introns, for example in human EDN (Eosinophil-derived neurotoxin) gene, human ECP (Eosinophil-cationic protein) gene, human utropin gene, human aldolase B gene (Intron-1), human CFTR (cystic fibrosis transmembrane conductance regulator) gene (Intron-1) (Handen and Rosenberg, 1997; Jenuwein *et al.,* 1997; Galvagni *et al*., 2002; Sabourin *et al*., 1996; Ott *et al*., 2009), and as inhibitory introns in the case of human keratin-18 gene (Intron-1) and human CD4 gene (Intron-1) expression (Umezawa *et al*., 1997; Swada *et al*., 1994).

Previous studies have also attributed a promoter/enhancer role within exon 1 in several genes: PAX6 gene (Zheng *et al*., 2001), elastin gene (Pierce *et al.,* 2006), human prostate tissue specific GPCR gene (Weng *et al.,* 2005) and human insulin like growth factor-1 gene (McLellan *et al*., 2006).

1.5.7 Features of the *VIPR2* Promoter

The human *VIPR2* gene was cloned and analyzed by Lutz *et al.,* 1999. (Figure: 3). A putative TSS of the *VIPR2* gene was identified at 187 bp upstream of the translation start site, and a polyadenylation signal was located at 2416 bp downstream of the stop codon (Lutz *et al*,. 1999).

The human *VIPR2* gene promoter region has similar features to those of house keeping gene promoters, such as the absence of the classical proximal promoter sequences (no TATA box and CAAT box), and presence of several consensus-binding sites for the transcription factors Sp1 (GC-box), and a CpG island.

In addition, there are several putative binding sites including sites for Pit-1A, CAC binding protein and Lyf, located in the 5' region upstream to the TSS. Further upstream to the 5' part of the exon 1 is a region which contains multiple repeats and GATA sites (Lutz *et al.,* 1999). However it is not known whether the putative regulatory elements shown in the 5' flanking region contain all the necessary sequences for the transcription and little is known regarding the contribution of exon and intron sequences, or those tissue specific elements in regulating *VIPR2* gene expression.

The nucleotide sequences of 5' flanking region of most other Class II GPCR genes have now been characterized and validated by several research groups (Minagawa *et al.,* 2000; Steel and Lutz., 2007; Kundu and Rao 1999; Chew *et al.,* 2001; Donohoe and Blomberg 1997; Tsai- morris *et al.,* 1996, Ho *et al.,* 1999; Petersenn *et al.,* 1998; Bettoun *et al.,* 1998; McCuaig *et al.,* 1994; Buggy *et al.,* 1995; Geiger *et al.,* 2000; Antonini *et al*., 2004 and Zolnierowicz *et al*., 1994). These have reported that the 5' promoter regions of these Class II GPCR genes (including human *VIPR2* gene) also exhibit features of CpG Island promoters such as multiple TSS, no TATA box, and are GC-rich with several consensus binding sites for Sp1. Most of the promoters of Class II GPCR genes have features similar to that of the promoters of phosphoribosyltransferase gene, human CDC2 gene, human thymidine kinase (TK)

gene and human PCNA gene with promoter elements largely confined between the 5' end of the CpG Island and the TSS.

These data indicate that the human *VIPR2* promoter has features typical of CpG island promoters, and suggests it is likely to be expressed during embyrogenesis and may therefore play a key role during development.

A study led by McEwen and Ornitz, 1998 investigated the murine Fibroblast growth factor receptor-3 gene that shares key features of human *VIPR2* gene promoter such as a TATA-less promoter, CpG islands and Sp1 binding sites around TSS. Intron 1 of this gene has several cis-regulatory sequences including the Sp1 binding sites. The investigators suggested that the Sp1 binding site located in the proximal promoter, and intron-1 work together to enhance and synergistically regulate gene transcription by taking the form of the syncretic module – enhancer-promoter or "Prohancer". The investigators also suggested alternate regulatory mechanisms such as intronic enhancer/promoter interaction and Sp1-Sp1 protein interaction leading to looping of the intervening cis elements. This Prohancer or syncretic gene regulation model is best seen in rat and human manganese superoxide dismutase genes, where the enhancer element acts as a prohancer by involving cytokine-inducible transcription, in the absence of a classical promoter (No TATA box or CAAT box) (St Clair *et al*., 2002; Porntadavity *et al*., 2001). Taken together, these studies identified the intron and exon sequences as prohancer, transcriptional regulator and tissue specific enhancer. It is possible that the *VIPR2* gene promoter also has these elements.

1.6 Aims and Hypothesis

Although the protein-coding region of the *VIPR2* gene has been characterised, very little is known about the regulatory elements flanking the gene. The aim of this project is to characterise the functional regulatory elements that regulate *VIPR2* gene expression, through identifying the basal promoter and the other cis-elements which contribute to tissue-specific regulation. The research aims to address the following hypothesis based on several assumptions:

The first assumption is that the flanking region 5' to exon 1 of the human *VIPR2* gene contains a basal promoter which is located within the GC rich region. The second assumption is that transcription of the *VIPR2* gene is initiated without a TATA box and that there is an alternative transcriptional initiation mechanism for the *VIPR2* gene.

The hypothesis to be addressed is that the tissue-specific regulation of VIPR2 is dependent upon the presence of consensus binding sites for several transcription factors in the 5' flanking region upstream of the putative TSS of the *VIPR2* gene.

1.7 Experimental approach

The first part of this work is to identify putative regulatory regions and conserved consensus sequences of transcription factor binding sites in the upstream region of the *VIPR2* gene in several species. This is carried out using multiple nucleotide sequence alignment software and other bioinformatic tools. The second stage of the work involves making reporter constructs containing the identified regulatory elements. The third part of the work is to transfect the reporter-constructs in the cell

lines and carries out functional studies to confirm the functionality of the identified elements.

Chapter 2
Materials and Methods

2.1 Materials

Standard laboratory chemicals were of analytical grade and were purchased from BDH Chemicals Ltd., UK, or Sigma Aldrich, UK. Kits for plasmid purification and gel extraction are from Qiagen (Crawley, West Sussex, United Kingdom) and Promega (Southampton, Hampshire, United Kingdom). KOD Hot start DNA Polymerase was obtained from Merck Bioscience. The 1Kb DNA ladder was obtained from Web Scientific or from New England Biolabs. Restriction enzymes and buffers were obtained from New England Biolabs and Promega.

2.2 Methods

2.2.1 Miniprep of plasmid DNA by alkaline lysis

A bacterial colony was picked and grown overnight in LB broth supplemented with ampicillin (0.1 mg/ml) as a selective agent. The culture was then centrifuged at 12,000 rpm for 2 mins to pellet the bacteria. The pellet was resuspended in ice cold 50 mM sterile Glucose/ 25 mM Tris-HCl/ 10mM EDTA/100 ug/ml RNAse A (GTE) and left at room temperature for 5 mins, followed by the addition of freshly prepared 0.2 M NaOH/ 1% SDS solution and mixed by inversion 2 to 3 times and kept at room temperature for 5 minutes.

To this mixture, 3M Potassium acetate (pH 5.5) was added and then mixed by inversion, then spun at room temperature at 14000 rpm for 7 min. The supernatant was then removed and transferred to a fresh tube for ethanol precipitation, by adding

2 volumes of 100% ethanol and mixed by inversion and spun for 5 minutes at 14000 rpm. Then the supernatants were discarded and the pellet washed with 70% ethanol (150 µL). The pellet was air-dried for 10-15 min and then resuspended in 50 µl TE (pH 8) and stored at -20°C freezer.

2.2.2 Wizard plus SV Minipreps DNA purification system

This commercial kit was used to isolate plasmid miniprep DNA, as per the manufacturer's instruction.

2.2.3 Plasmid maxiprep

A single colony was inoculated into 2 ml of L-broth supplemented with 100 µg/ml ampicillin and incubated at 37°C in an orbital shaker overnight. The next day, the 2 ml was used to inoculate 500 ml of L-broth (supplemented with ampicillin) and incubated at 37°C overnight with gentle shaking. The following day, the 500 ml culture was centrifuged at 3,000 rpm for 10 mins at 4°C. The medium was discarded and the cells resuspended with 9 ml of GTE solution, followed by addition of 1 ml fresh lysozyme solution (10 mg/ml in 10mM Tris-Hcl, pH 8.0). To the resuspended cell pellet, 20 ml of 0.2M NaOH/1% SDS solution was added and the contents thoroughly mixed by inversion then stored at room temperature for 10 min. After cell lysis, 10 ml of ice cold potassium acetate solution was added to the mixture, the bottle capped and mixed thoroughly by inversion, then stored on ice for 10 minutes. This was centrifuged at 5,000 rpm for 20 minutes at 4°C. The supernatant was thoroughly filtered through 4 layers of cheese cloth into a fresh centrifuge bottle and 0.6X volume of isopropanol was added. This was incubated at room temperature for

10 minutes followed by centrifugation at 5000 rpm for 20 minutes at 20°C. The supernatant was carefully discarded; the pellet and walls of the bottle were rinsed with 3ml of 70% ethanol and centrifuged again at 5000 rpm. The ethanol was removed using a Pasteur pipette, the plasmid pellet air dried, and then dissolved in 3 ml of TE.

2.2.4 PCR amplification of DNA

All PCR reactions were set up in a laboratory bench-top. Standard lab precautions were taken to prevent contamination of the PCR reactions, including the use of separate pipettes, and aerosol resistant pipette tips. All reagents (PCR grade water, 10x PCR buffer, dNTPs mix (2 mM each), $MgSO_4$ (25mM), DMSO (10%) and polymerase were supplied by the manufacturer (Merck Biosciences), master mixes were used wherever possible (to reduce the number of pipetting steps) and all reagents dedicated for PCR were stored per manufacturer's instructions. Reactions contained 0.3µM each primer, 0.2 mM each dNTP and 1mM $MgSO_4$. Thermocycling was carried out using Hybaid and Perkin Elmer thermal cyclers. Different types of PCR were used:

Touchdown PCR was done using the following protocol mentioned in Table 4, using KOD DNA Polymerase enzyme, Cosmid 66e9 (obtained from Dr Eve Lutz's Lab, primarily isolated from Lawrence Livermore National Lab chromosome-7 specific library) as template along with the primers.

2.2.5 DNA Agarose gel electrophoresis

0.8% (w/v) agarose gels were prepared from agarose (Invitrogen Life Technologies) and 1X TAE (40 mM Tris-Acetate/ 1 mM EDTA pH 8.3 at 25°C) buffer. The TAE buffer was also used as a running buffer. The agarose was dissolved in the electrophoresis buffer, by heating the mixture in the microwave at power-7 for 2 min. The agarose solution (30 ml) was allowed to cool down to 60°C, and ethidium bromide (10 mg/ml stock solution) was added to a final concentration of 1µg/ml. The electrophoresis was carried out at room temperature and ran at 100 volts/cm for 1.2 Hr. Gels were then examined on a UV transilluminator, and photographed.

2.2.6 Restriction enzyme digest

The restriction enzyme digest for the manipulation of plasmid DNA or PCR products were generally set up in a total reaction of 20-30 µl reaction mix containing the appropriate buffer solution, 0.5-1.5 µg of DNA, 1-2 units of restriction enzyme and double distilled water. The reaction components were mixed by vortexing, then spun briefly in a microcentrifuge, and then incubated at 37°C for 2-3 hrs or at room temperature overnight before analysis.

2.2.7 DNA excision and extraction

Following the electrophoresis separation of the restriction digest or PCR products, a sterile razor blade was used to excise the gel which had a fragment of interest and were transferred to separate 1.5 ml Eppendorf tubes. DNA extraction from the excised gel was carried out using a QIAEX II gel extraction kit (Qiagen, UK). The

resin was used according to the manufacturer's instructions, and the DNA was eluted by the addition of 20-30 µl of TE or double distilled water.

2.2.8 Extension

Extension was carried out to make blunt ends of the cohesive ends of EcoRI in order to clone into the Sma1 site of the pGL3 basic vector (Promega, UK). The extension of EcoRI digested 1-1 Intron was carried-out for 15 minutes at 68°C with 10 mM dNTP (Promega), 0.5µL of 3u/µL Pfu polymerase in Pfu buffer.

2.2.9 DNA extraction

To remove unnecessary restriction enzymes and interfering nucleotides for subsequent downstream steps, MicroClean (Web scientific, UK) was used according to the manufacturer's instructions.

2.2.10 Ligation reaction/Insert DNA

Ligation reactions were set up in a total volume of 10 µL containing 100-200 ng of linearized vector DNA (pGL3 basic), insert DNA (to give 2:1 molar ratio of insert: vector DNA), 1x rapid ligation buffer [60mM Tris-HCl (pH 7.8), 20mM $MgCl_2$, 20mM DTT, 2mM ATP and 10% PEG)] (Promega), 3 Weiss units of T4 DNA ligase (Promega) and double distilled water to 10 µL. The ligation reaction was carried out at room temperature for 15 minutes. 4 µL of the ligation reaction was used to transform 200 µl of XL-1 Blue competent cells (Genotype: *rec*A1 *end*A1 *gyr*A96 *thi*-1 *hsd*R17 *sup*E44 *rel*A1 *la*c) (Stratagene).

2.2.11 TOPO Cloning

TOPO cloning was done to clone the PCR amplified product of Intron1-1 as per manufacturer's instruction.

2.2.12 Transformation with one shot Top 10 competent cells

2 µl of TOPO cloning reaction was added to a vial of one shot chemically competent E.Coli (Invitrogen) and mixed gently then incubated on ice for 30 min. The cells were incubated for 30 seconds at 42°C, and then transferred to ice. 250 µl of S.O.C medium were added to the cells and the tube tightly caped and shaken (200 rpm) at 37ºC for 1 hr. Approximately 10-50 µL from each transformation was spread on to a pre-warmed ampicillin plate and incubated overnight at 37°C.

2.2.13 Phenol-chloroform extraction protocol for plasmid DNA

700 µl of sample (a portion of 1-1 maxiprep) was put into a 1.5 mL micro-centrifuge tube and an equal volume of phenol: chloroform: isoamyl alcohol (25:24:1) was added to the tube. This was vortexed vigorously to mix the phases, then centrifuged at top speed (12000-14000 rpm) in a micro-centrifuge for 1-2 minutes to separate the phases. The aqueous phase was transferred to a new tube and the procedure repeated once more with the aqueous phase.

2.2.14 Ethanol precipitation for plasmid DNA

Sodium acetate (3M) (pH 5.2) was added $1/10^{th}$ volume of the DNA solution. Then 95% ethanol was added 3.0 times the DNA solution (after addition of Sodium acetate) and the final mixture was incubated on ice (-20ºC) for 20 minutes then

centrifuged at 12000 rpm for 30 minutes. Carefully decant the supernatant. DNA pellet was rinsed with 70% ethanol, and then centrifuged again for 15 min at 12000 rpm. The supernatant was again discarded, air dried briefly and the pellet dissolved in the desired buffer.

2.2.15 Bioinformatics tools and softwares

ClustalW was used to align multiple sequences from different species. This software revealed similarities, differences and identities in the sequences analyzed (Thompson *et al*., 1994).

EMBOSS CpG plot/CpG report/Isochore was used to identify and plot CpG islands in a nucleotide sequence. The ratio of observed/expected GC nucleotides was calculated over a window and moved along the nucleotide sequence. The ratio is depicted graphically along with the region which reflects the software's definition for CpG Island. The above mentioned ratio is calculated on the user defined window. The window is slided along the loaded sequence and the ratio is recalculated until the end of the sequence was reached. CpG plot identifies a region as CpG island only when the percentage G + percentage C content is over 50% and the calculated ratio is over 0.6 on an average of 10 windows and each window not less than 200 bases. The observed CpG in a window is the number of times a 'C' is found in the sequence immediately followed by a 'G'.The expected CpG in a window is the number of CpG, we expected to see in a particular window size, based on the frequency of C's and G's in that window.So, the expected frequency of CpG's in a window is the number of 'C' multiplied by the number of 'G' in the window, divided by the window length (Larsen *et al*., 1992 and Bernardi, 2000).

GC profile is a web based bioinformatic tool which was used to analyse GC content qualitatively, quantitatively and visualise variation of GC content over a sequence of interest. The results obtained through this tool was used to understand the relationship between the GC content and the other genomic features such as CpG islands, distribution of functional genes etc (Gao and Zhang., 2006).

MATCH/TRANSFAC 2009 is a web based tool which was used to identify transcription factor binding sites over the sequence interest using positional weight matrices. This analytical tool is integrated with TRANSFAC database which is a library of experimentally validated transcription factor binding sites (Wingender, 2008 and Kel *et al.*, 2003).

The algorithm used by the MATCH analytical software rates the identified transcription factor binding site sequences by giving out scores or values for putative hits. For each sequence it matches, it gives out two different scores, the matrix similarity score and the core similarity score. The matrix similarity score is the rating of the quality of match between the 'test sequence' and 'the matrix' and the core similarity score is the rating of the quality of match between the 'test sequence' and the 'five most conserved consecutive positions' or 'core sequence' of a matrix.

The core similarity score and the matrix similarity score, both ranges from 0 to 1, where 0 denotes no match and 1 denotes exact match. First the core similarity helps in the pre-selection of possible matches, then matrix similarity score is calculated over those matches whose core similarity score exceeds certain cut-off. The MATCH

analytical software then combines these two scores and provides output as a single score.

Transcription factor binding sites identified were of high quality and those factors with a lower score are filtered and not shown in the analysis. The transcription factor binding sites, which were identified by MATCH/Transfac 2009, were given with the quality value. The quality value usualy range from 1-6 and each value represents the experimental reliability of certain protein-DNA interaction. The Quality-1 (Q1) represents particular transcription factor binding site identified is functionally confirmed. The Quality-2 (Q2) represents the binding site identified were validated by binding studies using purified or recombinant proteins. The Quality-3 (Q3) represents the identified binding site were immunologically validated. The Quality-4 (Q4) represents the identified binding site were validated by binding activity through known binding sequence, the Quality-5 (Q5) represents the identified binding site were validated by binding activity of uncharacterized protein on the experimentally validated binding site and the Quality-6 (Q6) represents no quality assigned. The matrix usually come with identifier like V$, I$, P$, F$, N$, B$ which means Vertebrate, Insect, Plants, Fungi, Nematode and Bacteria respectively and come with number 01, 02 refers to different types of matrices for the same factor and if they end with an alphabet C means consensus description which is constructed with an aid of Consindex [(Tsunoda and Takagi., 1999) and (Frech *et al*., 1993)].

The analyses were focused on the tissues concerned to five systems (Immune system, Liver, lung, muscle and nervous system) with a special focus on pituitary.

The transcription factors with score 0.99 and 1.0 were considered for analysis, any transcription factor binding sequence below the above scores were not considered for

analyses, because transcription factors with these scores represents 'near-exact' and 'exact' matches and if any transcription factor binding sites less than the above scores may still be considered, only if its known to be relevant to the VIPR2 gene regulation.

Netstart 1.0 is a web based neural network prediction tool which was used to identify probable translation start sites in the nucleotide sequence of interest (Pedersen, 1997)

Berkley Drosophilla Genome project neural network promoter prediction software is a web based neural network promoter prediction tool which was used to identify probable transcription start sites in the nucleotide sequence of interest (Reese, 2001).

Multi-zPicture is a web-based pair-wise multi-species alignment and visualization software, which uses BLATZ local alignment programme. It displays the gapless blocks of evolutionarily conserved regions or evolutionarily resilient regions (ECRs/ERRs) as smooth trace conservation plots. BlastZ does local alignments for sequences of any length based on belief that the input sequences are related, conserved and separated by regions that lack homology. The displayed regions are collinear only to the reference sequence (Ovcharenko *et al*., 2004).

Evoprinter High definition is the second generation multi-genomic comparative tool which automatically superimposes higher resolution alignments

(eBLATs) and display evolutionarily conserved regions or evolutionarily resilient regions (ECRs) that are shared among evolutionarily distant species. This tool requires only a single curated DNA sequence to do a rapid comparative analysis to identify short conserved sequence blocks and does this using modified BLAT (Enhanced BLAT) algorithm. Each eBLAT output represents three superimposed BLAT alignments of the same genomic sequence that were created using different search and alignment parameters. When these eBLAT alignments of evolutionarily distant genomes were compared; 75% more functional and conserved bases were detected compared to original BLAT alignments. The Evoprinter HD can currently align user's DNA of interest with 25 vertebrate, 5 nematode, 12 drosophila, 3 mosquito and 59 bacterial genomes.

Cis-Decoder is the alignment suite that scans and identifies already known and novel cis-elements from the EvoprinterHD identified evolutionarily resilient sequences or multispecies conserved sequences (MCS) (Brody *et al*., 2007).

To summarise, EvoPrinterHD and Cis-decoder identifies significant number of both novel and already defined conserved transcription factor binding sites that are shared among three or more orthologous DNAs. EvoPrinterHD display evolutionarily conserved regions or evolutionarily resilient regions within the user's DNA of interest by superimposing multiple high resolution enhanced BLATs alignments which were generated from sequences of evolutionarily distance species. By superimposing the different species evolutionary histories, and therefore the combined 'in -silico' mutagenic force, evolutionarily resilient regulatory regions that plays important role in gene expression is revealed (Odenwald *et al*., 2005 and Yavatkar *et al*., 2008).

Chapter 3
Computational analysis of regulatory elements in the VIPR2 Promoter

3.1 Introduction

Although many studies have been carried out to understand transcriptional regulation mechanisms, and although some progress has been made in this area, knowledge about transcription factor binding sites or the regulatory elements in some promoters remains rudimentary. To bridge the knowledge gaps in this area, many computational tools have been developed. A variety of tools are available to identify already known elements or discover new motifs. Understanding how regulatory elements are encoded in genomes remain limited and there is still a vast area to be explored (Loots, 2008).

A variety of approaches have been taken to identify these elements: pair-wise alignment strategy, pattern recognition, database search which is fed with experimentally determined patterns of transcription factor binding sites; alignment tools powered by hidden markov models (Eddy, 1995); dynamic programming and various other algorithms. There are also different types of alignment strategies: global alignment, local alignment and combinatorial or synergistic. Global alignment relies on the overall co-linearity of DNA sequences while local alignment depends on short matches between sequences independent of their location and orientation (Mount, 2004). Combinatorial or synergistic alignment is the one utilized in this study that is based on the combined power of functional evolutionarily conserved motif search using the TRANSFAC (Knuppel *et al*., 1994) database, position weight matrices, pattern recognition, and phylogenetic footprinting (Loots, 2008) The strategy used in this study is described through the flowchart (Figure 3.1).

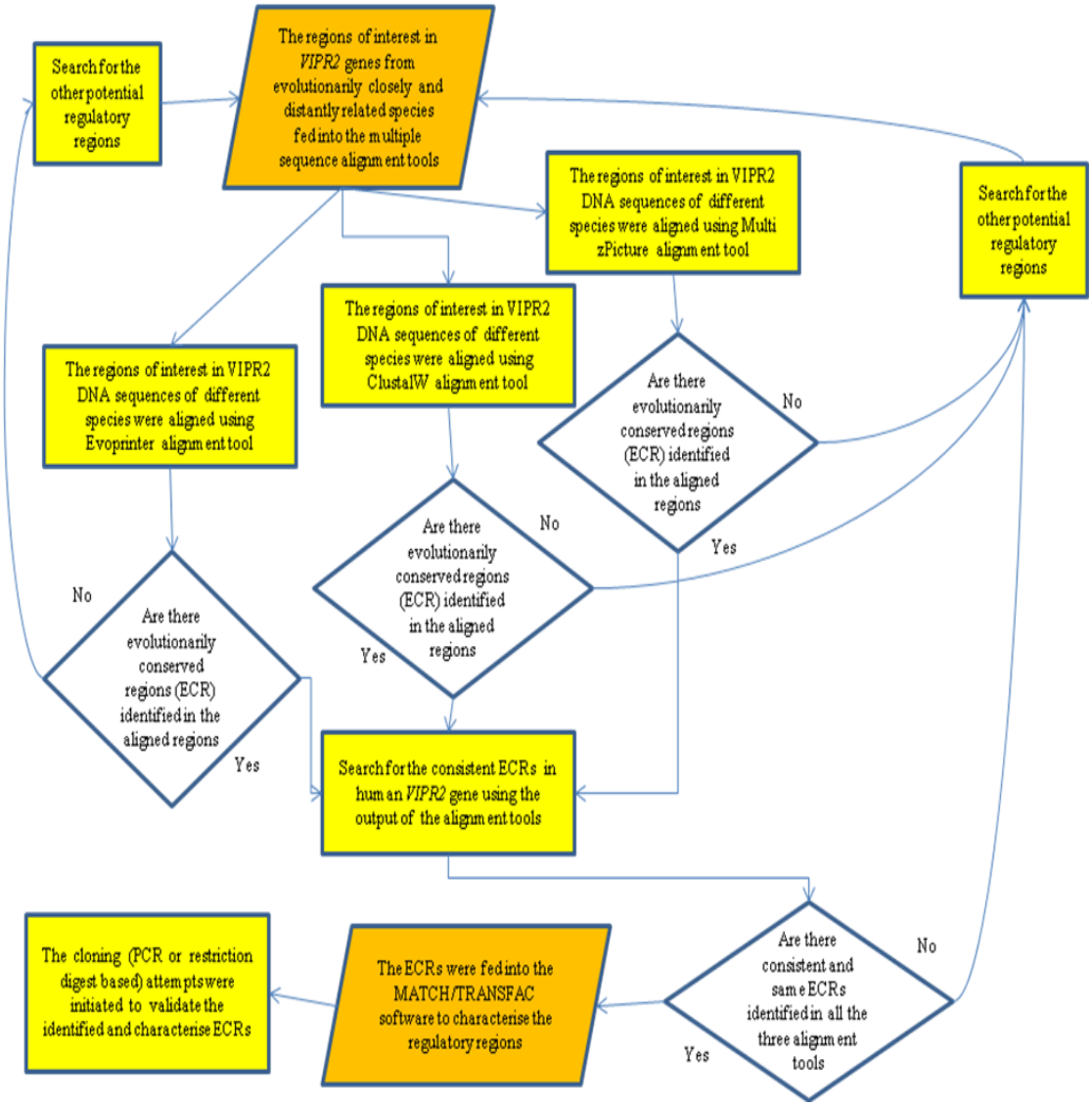


Figure 3.1 Flow chart to demonstrate the strategy followed to identify functional motifs

3.1.1 TRANSFAC 2009.03

TRANSFAC 2009.03 is a popular and powerful database that houses the most comprehensive collection of transcription factor binding sites in the form of position weight matrices. It contains data on transcription factors, their experimentally-validated binding sites and the corresponding derived position weight matrices. It is a probabilistic model that characterizes the DNA binding preference of a transcription factor. The Position Weight Matrix (PWM) is derived from the collection of aligned DNA binding sites that are likely to be bound by a common transcription factor. Data from experimentally evaluated PWM alone cannot find the transcription factor binding site, but needs data about the evolutionary history of a binding element or binding site. The TRANSFAC database is searched using MATCHTM, a weight matrix-based tool (Mayor et al., 2000 and Kel et al., 2003).

3.1.2 Phylogenetic tools

Evolutionary data is obtained from inter-species or cross species evolutionary conserved regulatory elements by phylogenetic footprinting. The software used in this type of analysis are ClustalW (Thompson *et al*., 1994), Multi-zPicture (Ovcharenko *et al*., 2004), EvoprinterHD (Odenwald *et al*., 2005 and Yavatkar et al., 2008) and Cis-Decoder (Brody *et al*., 2007).

Phylogenetic footprinting is the most powerful approach used currently by many molecular biologists and cell biologist. It is a strategy to identify highly conserved DNA motifs presents in inter-species multiple sequence alignment. Phylogenetic footprinting is performed by globally aligning multiple orthologous sequences (from different species) and identifying regions of high conservation in the alignment. The

main assumption here is that sequences which perform important functions are more frequently conserved between evolutionarily distant species, which distinguishes them from non-functional surrounding sequences (Mayor *et al*., 2000).

In this Chapter, a range of bioinformatic tools are used to evaluate the promoter region of the *VIPR2* gene with the aim of identifying both basal and regulatory promoter elements. In addition, the promoter sequences of the *VIPR2* gene from several different species are compared.

3.2 Results

3.2.1 Prediction of Translation start site(s) in the *VIPR2* Gene

It is well known fact that multiple protein variants can exist for a single gene. The existence of multiple variants for a single gene has been explained through mechanisms such as multiple promoters or transcription start sites, and alternative splicing. The other lesser known mechanism involves multiple translation start sites and alternative open reading frames called alternative translation (Kochetov, 2008).

Although a translation start site has already been identified in the human *VIPR2* gene previously, the recent discovery of a variant in the human VIP-R-2 receptor (VPAC2de325-438) that lacks 114 amino acids in human lymphocytes (Miller *et al*., 2006), suggests the possibility of multiple translation start sites in the related *VIPR2* gene.

To identify putative alternate translation start sites in the *VIPR2* gene, the translation start site prediction tool Netstart 1.0 was used and analysis was carried out on ~2.3kb DNA sequence which includes both non-coding and coding regions of the *VIPR2*

gene. Netstart 1.0 uses a neural network algorithm. A 2.3 kb sequence that covered exon-1, intron-1 and exon-2 of the *VIPR2* gene was analyzed. A 1.0 kb (1,086 bp) of sequence upstream of the Exon1/intron 1 boundary was also analyzed. The results are shown in the Table: 2.

The Netstart 1.0 scores range from 0.0 to 1.0. A score of >0.5 represents a probable Translation SS. Two probable Translation SS are identified (-79 to -77 and +349 to +351). None of these predicted Translation SS are the Translation SS previously identified by Lutz *et al.,* 1999. However, it should be noted that none of these predicted sites have not been tested experimentally and are only predictions. In addition, Netstart has not been tested on genome data where introns may be close to the start codon. Therefore these results need to be treated with caution, but are useful for narrowing down possible alternate Translation SS in the *VIPR2* gene. Further work will need to be carried out to establish the true translation SS of the gene. This may include determining the protein sequence of the N-terminus of the receptor protein.

Start	End	Score	Prediction
-888	-886	0.473	Not a probable Translation SS
-782	-780	0.302	Not a probable Translation SS
-720	-718	0.317	Not a probable Translation SS
-79	-77	0.570	**Most probable** Translation SS
+1	**+3**	**0.446**	**Not a probable** Translation SS
+349	+351	0.519	**Most probable** Translation SS
+445	+447	0.333	Not a probable Translation SS
+602	+604	0.238	Not a probable Translation SS
+961	+963	0.034	Not a probable Translation SS
+981	+983	0.105	Not a probable Translation SS

Table 3.1 Bioinformatics result output from the translation start site prediction software Netstart 1.0

3.2.2 CpG Analysis of the *VIPR2* Promoter

To identify GC rich regions, and to reveal the compositional features of the DNA sequence in order to understand the structure, function and evolution of the gene, CpG analyses were performed using two bioinformatic tools, GC profile and EMBOSS CpGPlot/CpGReport/Isochore.

GC Profile is a web based visualizing tool to analyze the variation of GC content in the genome, and is based on quadratic divergence. GC profile identified the average GC content as 66.7%.

EMBOSS identified three CpG islands A, B, C (Figures 3.2 and 3.3). The length of CpG island A is 229 bases, CpG Island B is 1240 bases and CpG island C is 273 bases. The observed/expected ratio: >0.60, percentage C + percentage G: >50. The isochore results show a high GC content with consistent homogeneity (Figure: 3.2).

3.2.3 Identification of the Basal and Regulatory Promoter elements using MATCH

In this study, the ~6kb regulatory region located 5' upstream, and the ~2.5 kb Intron-1 region downstream of the original start codon of the human *VIPR2* gene, was computationally characterised, using MATCH to interrogate the TRANSFAC 2009 database (Mayor *et al*., 2000 and Kel *et al*., 2003).

MATCH identified the minimal promoter region with a size of 408 bp which spans the 240 bp upstream and 168 bp downstream of the start codon (Figure 3.4). Upstream of the basal promoter is the upstream promoter region, which contains several tissue specific elements and further upstream to this region is the distal promoter. There are four GC boxes in the upstream promoter region (+68 to +77,

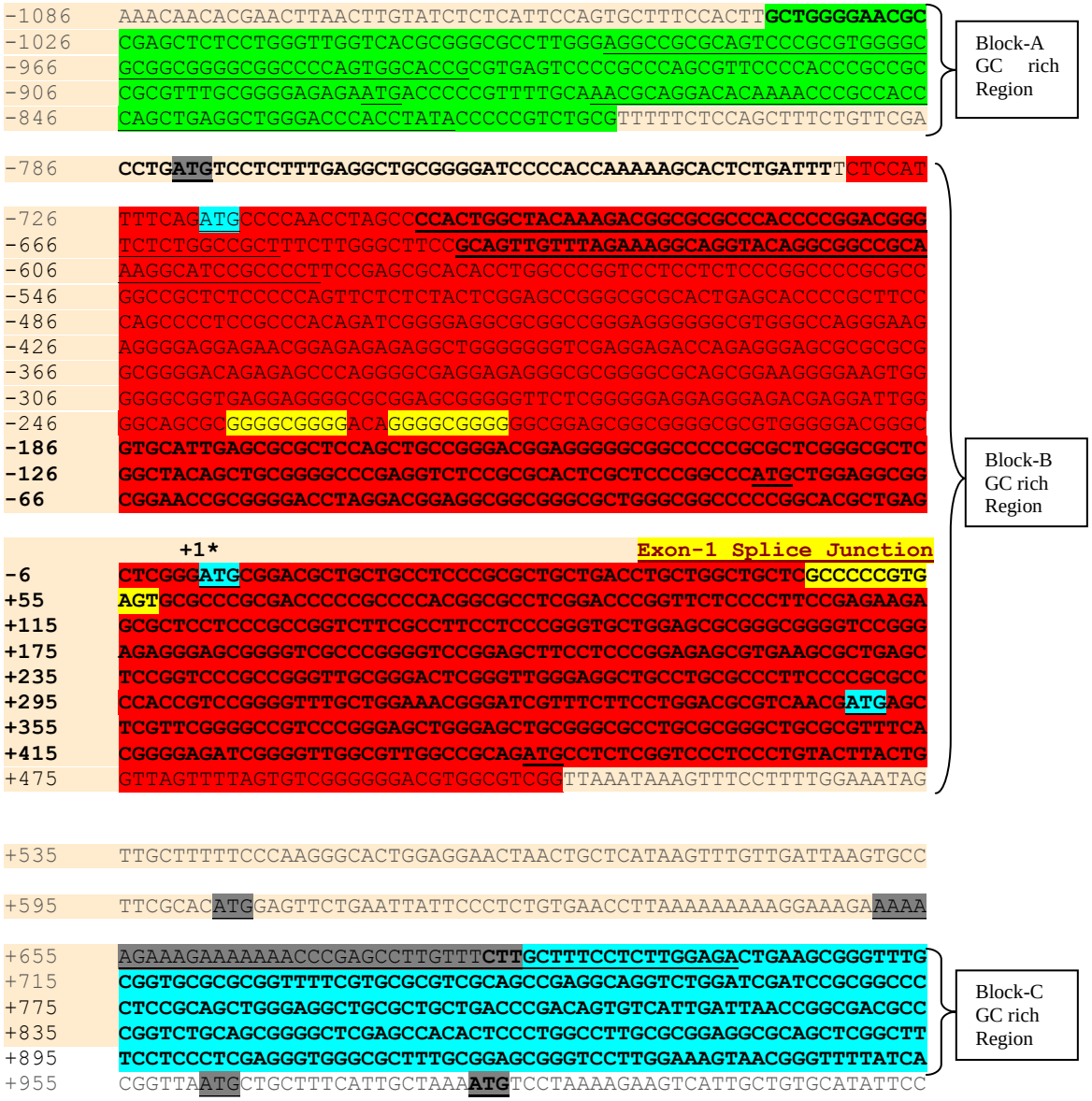
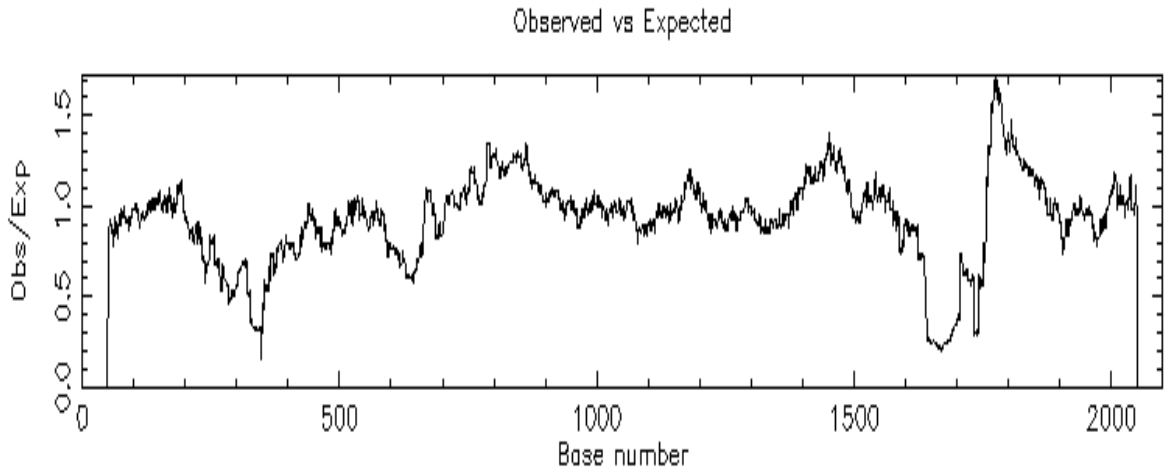
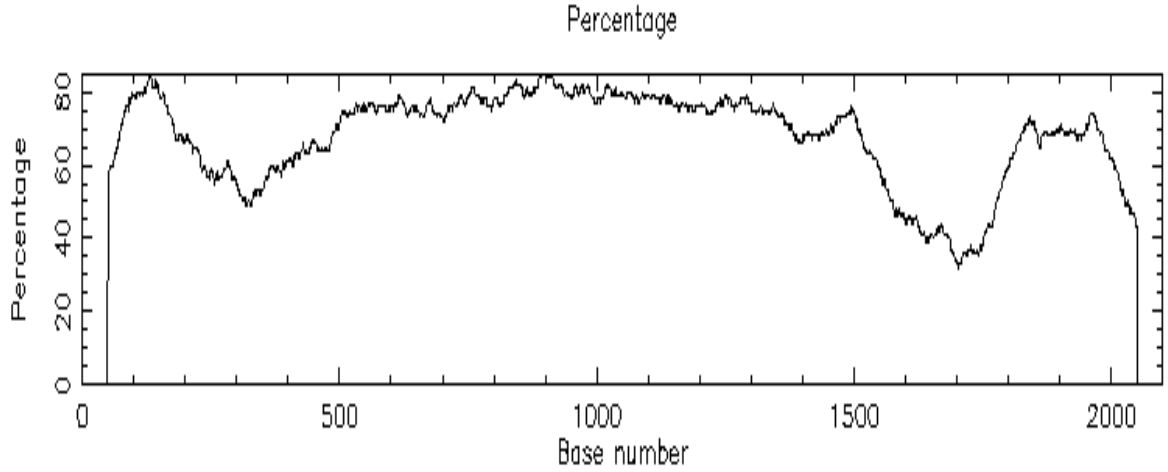


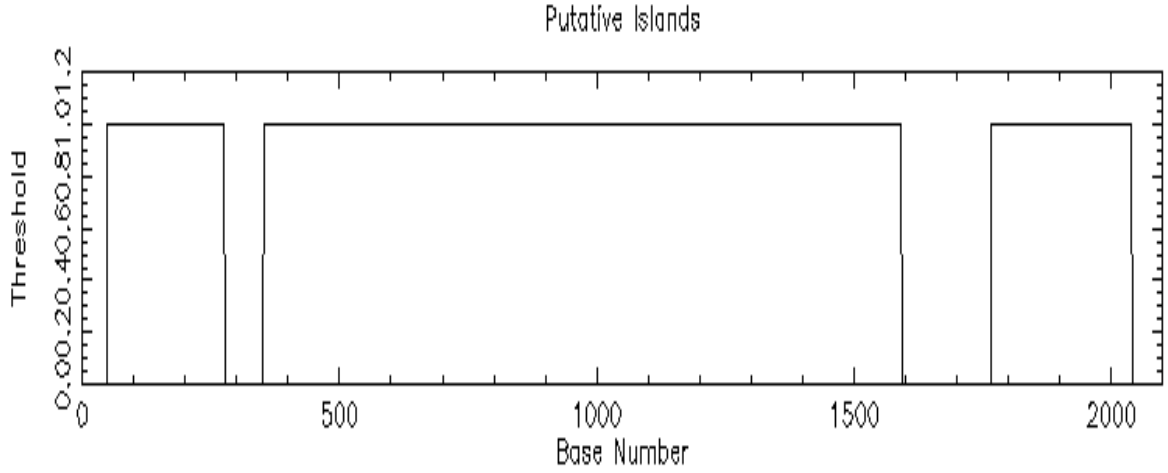
Figure 3.2 Identification of the GC rich regions, putative translation initiation regions and putative promoter regions (greyed and underlined) in the coding and non-coding part of the human VIPR2 gene.

Bases relative to the ATG*(experimentally validated) translation initiation codon (In boldface, designated +1) are numbered on the left. Three blocks of GC-rich region are indicated and highlighted.

A


Observed vs Expected
Obs/Exp
0.0
0.5
1.0
1.5
0
500
1000
1500
2000
Base number


Percentage
Percentage
0
20
40
60
80
0
500
1000
1500
2000
Base number


Putative Islands
Threshold
0.0
0.2
0.4
0.6
0.8
1.0
1.2
0
500
1000
1500
2000
Base Number

/cont

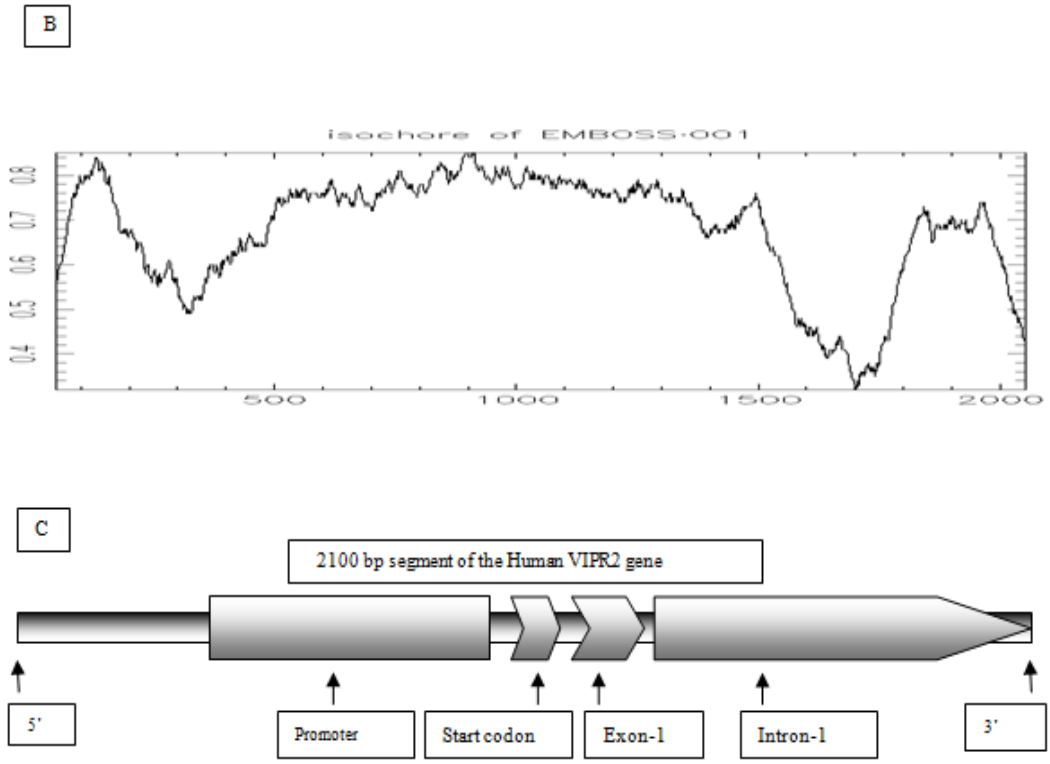


Figure 3.3 CpG analysis of the human VIPR2 gene

(A) Identification of CpG Islands surrounding the start codon. The analysis revealed Observed/Expected ratio of the GC dinucleotides >0.60, Percent C + Percent G >50.0, when the threshold set at 0.6 and a minimal length of 200. The CpG analysis identified three blocks of CpG islands which spans both the coding and non-coding regions. (B) Diagram represents the Isochore analysis of coding and non-coding regions of the gene. (C) Schematic diagram shows the presence of of the promoter, start codon, exon and intron of the human *VIPR2* gene in the CpG islands. All the panels are spatially aligned with each other.

+160 to +168, -228 to -220 and -240 to -232) which potentially bind the Sp1 transcription factor. (Figure 3.4).

MATCH also identified several immune system specific transcription factor binding sites (Figure 3.6) such as YY1, HTF, NFKappa B, LEF1, STAT, IK2, AML; several adipocyte specific cis-elements such as SREBP, GR and PPAR (Figure: 3.5); several Liver specific transcription factor binding sites such as LXR, FXR, HNF1, HNF3 and HNF4 (Figure 3.7); lung specific transcription factor binding sites such as TTF-1, HNF, FREAC (Figure 3.8); and muscle specific transcription factor binding sites such as MyoD, Myogenin, MEF-2, HAND1 (Figure: 3.9). The software also identified several other tissue specific transcription factor binding sites for example nervous system specific transcription binding sites such as Pax3, Pax6, Pax8, Fox, BRN-2, CREB, and pituitary specific transcription factor binding sites such as Pit-1a, and VDR. These putative tissue-specific binding sites are identified in the next section (Figures 3.5-3.9)

3.2.4 Evoluationary Conservation Analysis of VIPR2 promoter

To evaluate the significance of the regulatory elements identified using MATCH, evolutionary conservation analysis was carried on the ~6kb regulatory region located upstream, and the ~2.5 kb Intron-1 region downstream to the start codon of the human *VIPR2* gene, using Multi-zPicture (Ovcharenko et al., 2004) (figures: 3.5-3.10), EvoprinterHD (Odenwald *et al*., 2005 and Yavatkar *et al*., 2008), Cis-Decoder (Brody et al, 2007) and CLUSTALW (Thompson et al., 1994), (figures: 3.11-3.17). The 4 species examined initially were chimpanzee, gorilla, human and marmoset which are expected to be closely related as they are all primates.

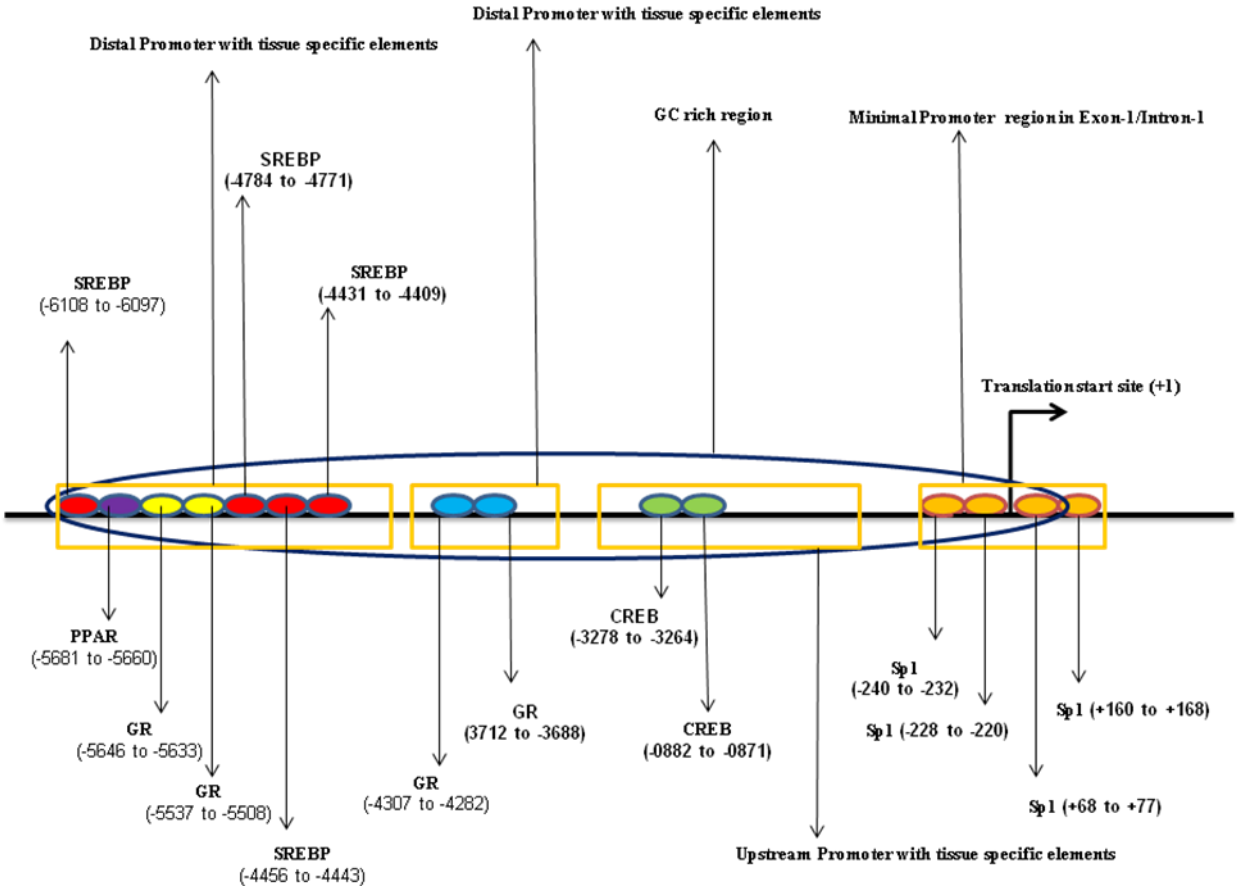
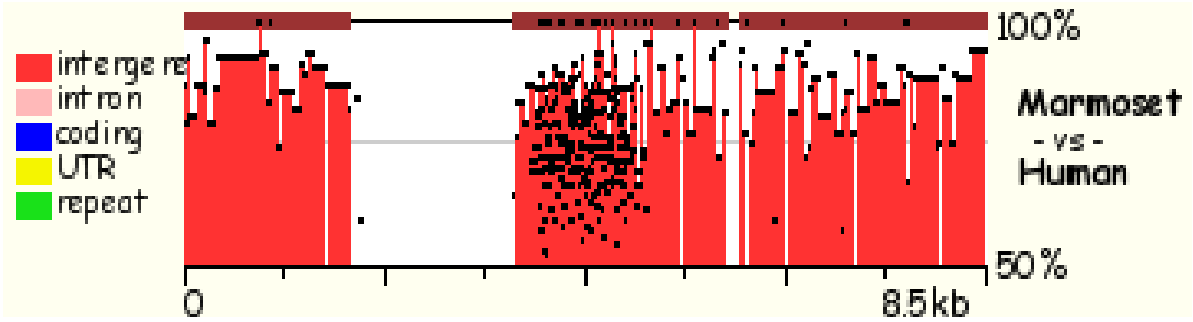


Figure 3.4 Putative transcription factor binding sites in the promoter region of the human VIPR2 gene

Multi-zPicture alignment ID: 11160626423295

Parameter settings:
ECR length: atleast 100 bases
ECR similarity: atleast 70%
Bottom cut-off: 50%



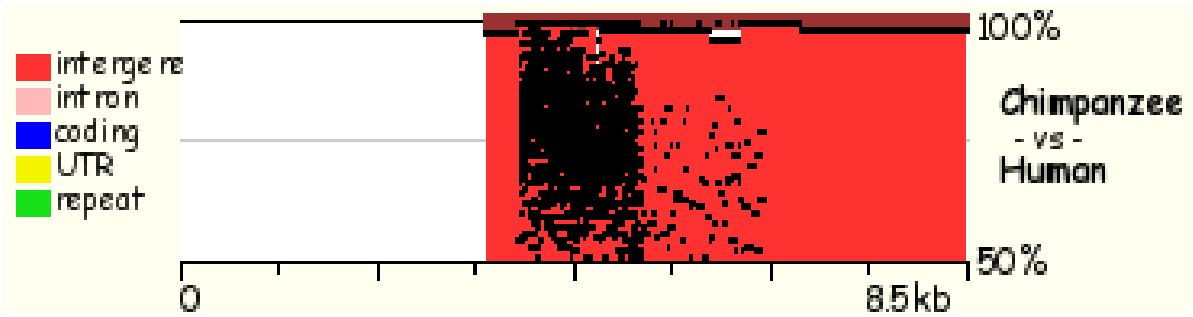
List of evolutionary conserved regions (ECR):
There were 4 ECR(s) detected using a threshold of at least 70% identity over 100 bps

Position in Multi-zPicture alignment	Length	Percent Identity	Position in ClustalW alignment	Region
1-1756	1756bp	88%	-6124 to -4393	Upstream regulatory
3493-5747	2255bp	81%	-2648 to -440	Upstream regulatory
5900-6387	488bp	82%	-282 to +203	Promoter/Exon-1/Intron-1
6415-8520	2106bp	85%	+334 to +2217	Intron-1

Figure 3.5 Multi-zPicture alignment of the regions consisting of 5' upstream to start codon, exon-1, intron-1 and exon-2 of human *VIPR2* gene and marmoset *VIPR2* gene and the identified ECRs

Multi-zPicture alignment ID: 11160621270388

Parameter settings:
ECR length: atleast 100 bases
ECR similarity: atleast 70%
Bottom cut-off: 50%



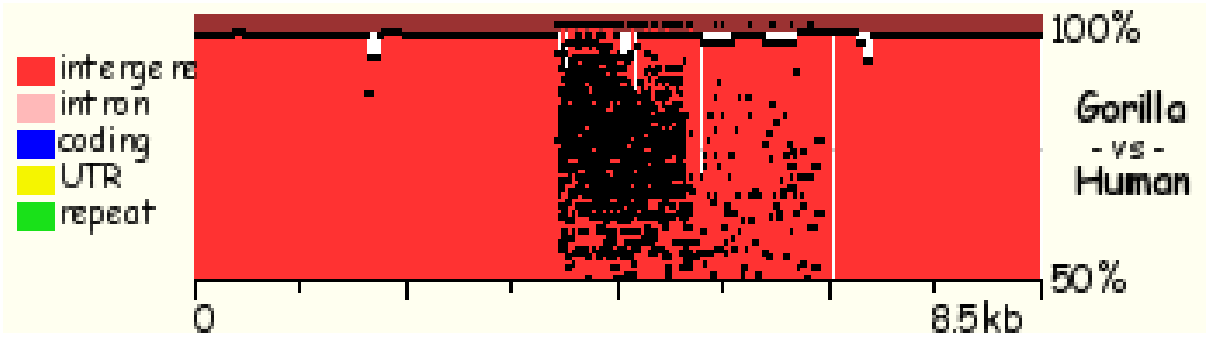
List of evolutionary conserved region(s) (ECR):
There were 1 ECR(s) detected using a threshold of at least 70% identity over 100 bps

Position in Multi-zPicture alignment	Length	Percent Identity	Position in ClustalW alignment	Region
3286-8520	5235bp	98%	-2886 to +2217	Upstream regulatory/Exon-1/Intron-1/Exon-2

Figure 3.6 Multi-zPicture alignment of the regions consisting of 5' upstream to start codon, exon-1, intron-1 and exon-2 of human *VIPR2* gene and chimpanzee *VIPR2* gene and the identified ECRs

Multi-zPicture alignment ID: 11160617023708

Parameter settings:
ECR length: atleast 100 bases
ECR similarity: atleast 70%
Bottom cut-off: 50%



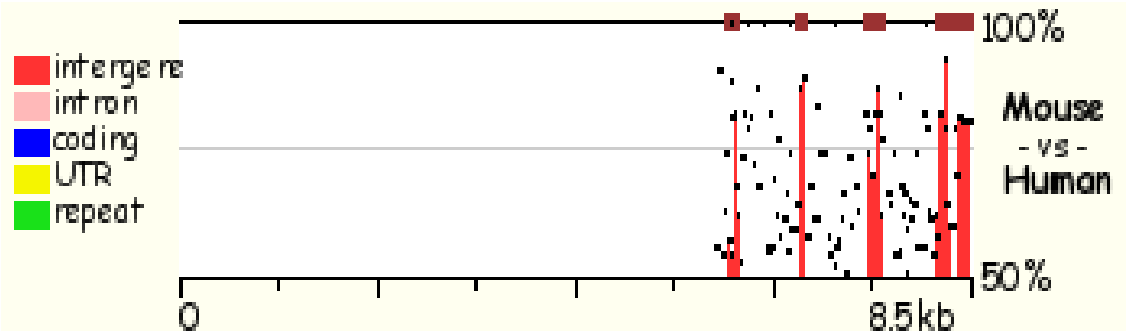
List of evolutionary conserved regions (ECR):
There were 1 ECR(s) detected using a threshold of at least 70% identity over 100 bps

Position in Multi-zPicture alignment	Length	Percent Identity	Position in ClustalW alignment	Region
1-8514	8514bp	97%	-6124 to +2217	Upstream regulatory/Exon-1/Intron-1/Exon-2

Figure 3.7 Multi-zPicture alignment of the regions consisting of 5' upstream to start codon, exon-1, intron-1 and exon-2 of human *VIPR2* gene and gorilla *VIPR2* gene and the identified ECRs

Multi-zPicture alignment ID: 11160517767700

Parameter settings:
ECR length: atleast 100 bases
ECR similarity: atleast 70%
Bottom cut-off: 50%



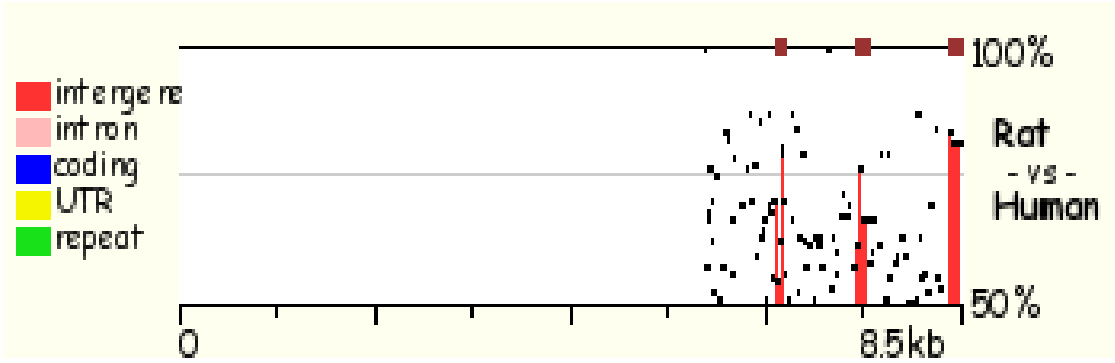
List of evolutionary conserved regions (ECR):
There were 5 ECR(s) detected using a threshold of at least 70% identity over 100 bps

Position in Multi-zPicture alignment	Length	Percent Identity	Position in ClustalW alignment	Region
5872-6014	143bp	69%	-307 to -165	Upstream regulatory
6636-6735	100bp	70%	+441 to +566	Intron-1
7365-7573	209bp	69%	+1154 to +1382	Intron-1
8121-8276	156bp	65%	+1931 to +2119	Intron-1
8320-8517	198bp	75%	+2177 to +2217	Exon-2

Figure 3.8 Multi-zPicture alignment of the regions consisting of 5' upstream to start codon, exon-1, intron-1 and exon-2 of human *VIPR2* gene and mouse *Vipr2* gene and the identified ECRs

Multi-zPicture alignment ID: 11160624845456

Parameter settings:
ECR length: atleast 100 bases
ECR similarity: atleast 70%
Bottom cut-off: 50%



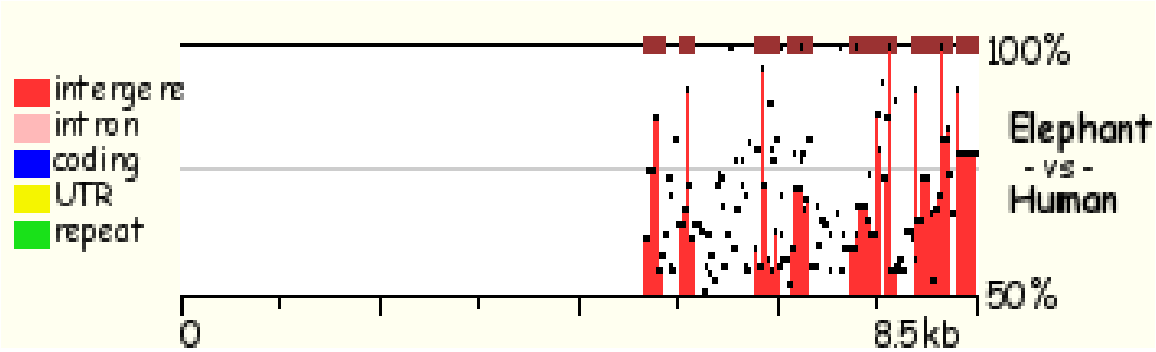
List of evolutionary conserved regions (ECR):
There were 3 ECR(s) detected using a threshold of at least 70% identity over 100 bps

Position in Multi-zPicture alignment	Length	Percent Identity	Position in ClustalW alignment	Region
6484-6583	100bp	70%	+295 to +438	Intron-1
7363-7503	141bp	68%	+1213 to + 1343	Intron-1
8368-8517	150bp	82%	+2002 to +2091	Intron-1/Exon-2

Figure 3.9 Multi-zPicture alignment of the regions consisting of 5' upstream to start codon, exon-1, intron-1 and exon-2 of human *VIPR2* gene and rat *Vipr2* gene and the identified ECRs

Multi-zPicture alignment ID: 11160617634458

Parameter settings:
ECR length: atleast 100 bases
ECR similarity: atleast 70%
Bottom cut-off: 50%



List of evolutionary conserved regions (ECR):
There were 8 ECR(s) detected using a threshold of at least 70% identity over 100 bps

Position in Multi-zPicture alignment	Length	Percent Identity	Position in ClustalW alignment	Region
4956-5161	206bp	70%	-1235 to -1020	Upstream regulatory
5333-5499	167bp	69%	-846 to -692	Upstream regulatory
6139-6395	257bp	69%	-49 to +215	Promoter/Exon-1/Intron-1
6501-6735	235bp	71%	+311 to +571	Intron-1
7152-7339	188bp	69%	+965 to +1154	Intron-1
7351-7647	297bp	71%	+1162 to +1479	Intron-1
7837-8036	200bp	70%	+1647 to +1853	Intron-1
8046-8230	185bp	76%	+1855 to +2054	Intron-1

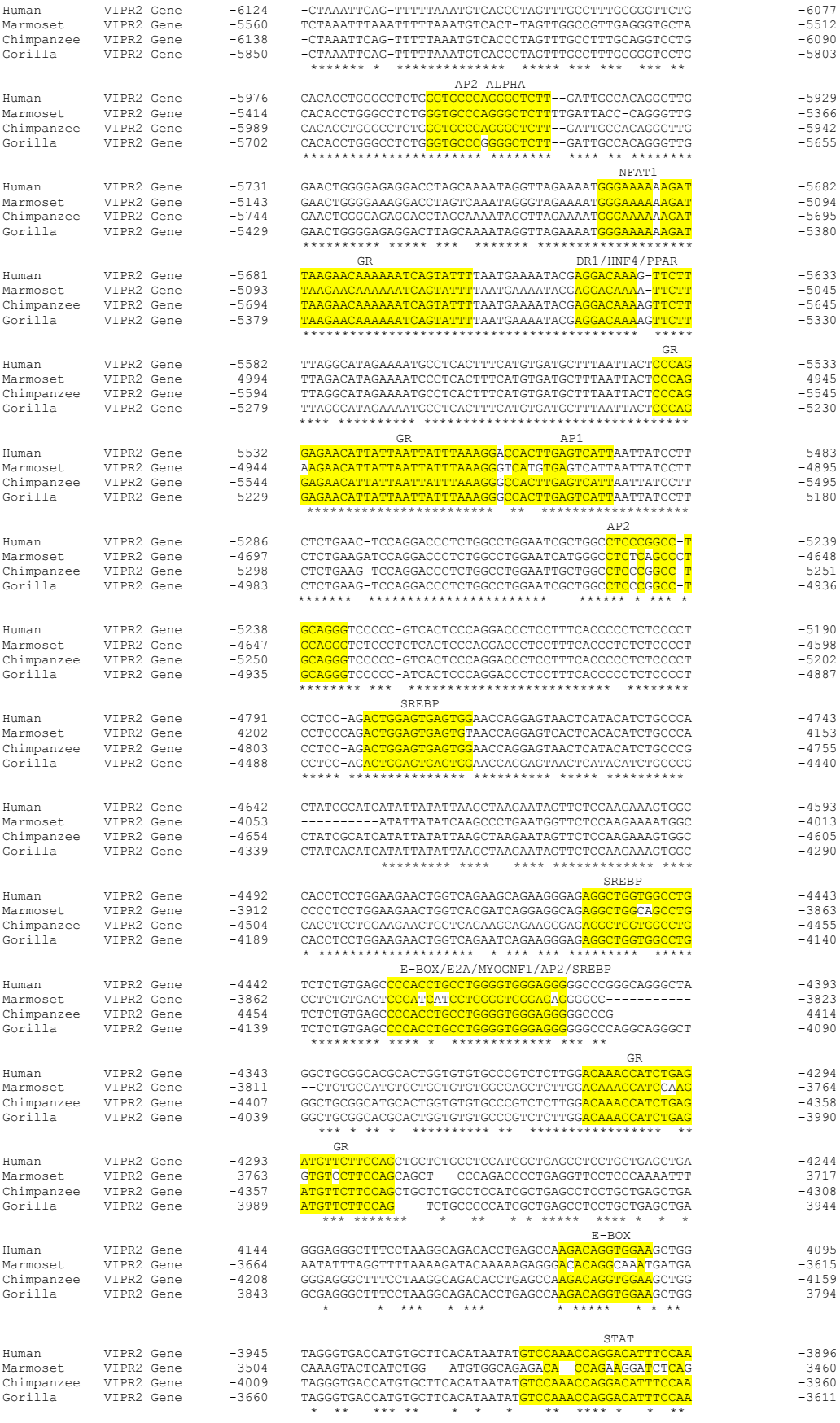
Figure 3.10 Multi-zPicture alignment of the regions consisting of 5' upstream to start codon, exon-1, intron-1 and exon-2 of human *VIPR2* gene and elephant *VIPR2* gene and the identified ECRs

The results show that many of the elements identified are conserved across species. However, some transcription factor binding sites are not conserved, indicating that either there has been a divergence of regulation or that these binding sites are less likely to be functional. The % relatedness is above 70% overall, using a 100 bp window. The results of Multi-zPicture show that many of the elements identified are conserved across species (figures: 3.5-3.10). The Multi-zPicture results are consistent with the CLUSTALW alignments in which the MATCH/TRANSFAC motifs are highlighted. The putative regulatory regions of the human *VIPR2* gene were also compared with rat and mouse *Vipr2* genes. The evolutionary conservation analysis with human, mouse and rat *VIPR2* genes identified no significant evolutionary conserved regions (figures: 3.8 and 3.9). The analysis may suggest that the regulatory mechanism involved in transcription of mouse and rat *VIPR2* gene were different from primates and human.

When evoluationarily distantly related species were examined (rhesus monkey, cow, dog and cat, chicken, elephant, fugu), only certain elements were conserved (figure: 3.10). These included sox, mash, neurogenin and math binding sites (figure 3.16). These are neuron specific elements located in intron 1, suggesting that over evolutionary time the function may have been retained.

3.3 Summary and Conclusions

In this chapter, the strategies used to identify the regulatory elements are explained. The analytical software MATCH was used to search the powerful TRANSFAC database which houses up-to-date experimentally validated transcription factor


Human VIPR2 Gene -6124 -CTAAATTCAG-TTTTTAAATGTCACCCTAGTTTGCCTTTGCGGGTTCTG -6077
Marmoset VIPR2 Gene -5560 TCTAAATTTAAATTTTTAAATGTCACT-TAGTTGGCCGTTGAGGGTGCTA -5512
Chimpanzee VIPR2 Gene -6138 -CTAAATTCAG-TTTTTAAATGTCACCCTAGTTTGCCTTTGCAGGTCCTG -6090
Gorilla VIPR2 Gene -5850 -CTAAATTCAG-TTTTTAAATGTCACCCTAGTTTGCCTTTGCGGGTCCTG -5803
AP2 ALPHA
Human VIPR2 Gene -5976 CACACCTGGGCCTCTGGGTGCCCAGGGCTCTT--GATTGCCACAGGGTTG -5929
Marmoset VIPR2 Gene -5414 CACACCTGGGCCTCTGGGTGCCCAGGGCTCTTTTGATTACC-CAGGGTTG -5366
Chimpanzee VIPR2 Gene -5989 CACACCTGGGCCTCTGGGTGCCCAGGGCTCTT--GATTGCCACAGGGTTG -5942
Gorilla VIPR2 Gene -5702 CACACCTGGGCCTCTGGGTGCCCGGGGCTCTT--GATTGCCACAGGGTTG -5655
NFAT1
Human VIPR2 Gene -5731 GAACTGGGGAGAGGACCTAGCAAAATAGGTTAGAAAATGGGAAAAAAGAT -5682
Marmoset VIPR2 Gene -5143 GAACTGGGGAAAGGACCTAGTCAAATAGGGTAGAAAATGGGAAAAAAGAT -5094
Chimpanzee VIPR2 Gene -5744 GAACTGGGGAGAGGACCTAGCAAAATAGGTTAGAAAATGGGAAAAAAGAT -5695
Gorilla VIPR2 Gene -5429 GAACTGGGGAGAGGACTTAGCAAAATAGGTTAGAAAATGGGAAAAAAGAT -5380
GR
DR1/HNF4/PPAR
Human VIPR2 Gene -5681 TAAGAACAAAAAATCAGTATTTTAATGAAAATACGAGGACAAAG-TTCTT -5633
Marmoset VIPR2 Gene -5093 TAAGAACAAAAAATCAGTATTTTAATGAAAATACGAGGACAAAA-TTCTT -5045
Chimpanzee VIPR2 Gene -5694 TAAGAACAAAAAATCAGTATTTTAATGAAAATACGAGGACAAAAGTTCTT -5645
Gorilla VIPR2 Gene -5379 TAAGAACAAAAAATCAGTATTTTAATGAAAATACGAGGACAAAAGTTCTT -5330
GR
Human VIPR2 Gene -5582 TTAGGCATAGAAAATGCCTCACTTTCATGTGATGCTTTAATTACTCCCAG -5533
Marmoset VIPR2 Gene -4994 TTAGACATAGAAAATCCCTCACTTTCATGTGATGCTTTAATTACTCCCAG -4945
Chimpanzee VIPR2 Gene -5594 TTAGGCATAGAAAATGCCTCACTTTCATGTGATGCTTTAATTACTCCCAG -5545
Gorilla VIPR2 Gene -5279 TTAGGCATAGAAAATGCCTCACTTTCATGTGATGCTTTAATTACTCCCAG -5230
GR
AP1
Human VIPR2 Gene -5532 GAGAACATTATTAATTATTTAAAGGACCACTTGAGTCATTAATTATCCTT -5483
Marmoset VIPR2 Gene -4944 AAGAACATTATTAATTATTTAAAGGGTCATGTGAGTCATTAATTATCCTT -4895
Chimpanzee VIPR2 Gene -5544 GAGAACATTATTAATTATTTAAAGGGCCACTTGAGTCATTAATTATCCTT -5495
Gorilla VIPR2 Gene -5229 GAGAACATTATTAATTATTTAAAGGGCCACTTGAGTCATTAATTATCCTT -5180
AP2
Human VIPR2 Gene -5286 CTCTGAAC-TCCAGGACCCTCTGGCCTGGAATCGCTGGCCTCCCGGCC-T -5239
Marmoset VIPR2 Gene -4697 CTCTGAAGATCCAGGACCCTCTGGCCTGGAATCATGGGCCTCTCAGCCCT -4648
Chimpanzee VIPR2 Gene -5298 CTCTGAAG-TCCAGGACCCTCTGGCCTGGAATTGCTGGCCTCCCGGCC-T -5251
Gorilla VIPR2 Gene -4983 CTCTGAAG-TCCAGGACCCTCTGGCCTGGAATCGCTGGCCTCCCGGCC-T -4936
Human VIPR2 Gene -5238 GCAGGGTCCCCC-GTCACTCCCAGGACCCTCCTTTCACCCCCTCTCCCCT -5190
Marmoset VIPR2 Gene -4647 GCAGGGTCTCCCTGTCACTCCCAGGACCCTCCTTTCACCCTGTCTCCCCT -4598
Chimpanzee VIPR2 Gene -5250 GCAGGGTCCCCC-GTCACTCCCAGGACCCTCCTTTCACCCCCTCTCCCCT -5202
Gorilla VIPR2 Gene -4935 GCAGGGTCCCCC-ATCACTCCCAGGACCCTCCTTTCACCCCCTCTCCCCT -4887
SREBP
Human VIPR2 Gene -4791 CCTCC-AGACTGGAGTGAGTGGAACCAGGAGTAACTCATACATCTGCCCA -4743
Marmoset VIPR2 Gene -4202 CCTCCCAGACTGGAGTGAGTGTAACCAGGAGTCACTCACACATCTGCCCA -4153
Chimpanzee VIPR2 Gene -4803 CCTCC-AGACTGGAGTGAGTGGAACCAGGAGTAACTCATACATCTGCCCG -4755
Gorilla VIPR2 Gene -4488 CCTCC-AGACTGGAGTGAGTGGAACCAGGAGTAACTCATACATCTGCCCG -4440
Human VIPR2 Gene -4642 CTATCGCATCATATTATATTAAGCTAAGAATAGTTCTCCAAGAAAGTGGC -4593
Marmoset VIPR2 Gene -4053 ----------ATATTATATCAAGCCCTGAATGGTTCTCCAAGAAAATGGC -4013
Chimpanzee VIPR2 Gene -4654 CTATCGCATCATATTATATTAAGCTAAGAATAGTTCTCCAAGAAAGTGGC -4605
Gorilla VIPR2 Gene -4339 CTATCACATCATATTATATTAAGCTAAGAATAGTTCTCCAAGAAAGTGGC -4290
SREBP
Human VIPR2 Gene -4492 CACCTCCTGGAAGAACTGGTCAGAAGCAGAAGGGAGAGGCTGGTGGCCTG -4443
Marmoset VIPR2 Gene -3912 CCCCTCCTGGAAGAACTGGTCACGATCAGGAGGCAGAGGCTGGCAGCCTG -3863
Chimpanzee VIPR2 Gene -4504 CACCTCCTGGAAGAACTGGTCAGAAGCAGAAGGGAGAGGCTGGTGGCCTG -4455
Gorilla VIPR2 Gene -4189 CACCTCCTGGAAGAACTGGTCAGAATCAGAAGGGAGAGGCTGGTGGCCTG -4140
E-BOX/E2A/MYOGNF1/AP2/SREBP
Human VIPR2 Gene -4442 TCTCTGTGAGCCCCACCTGCCTGGGGTGGGAGGGGGCCCGGGCAGGGCTA -4393
Marmoset VIPR2 Gene -3862 CCTCTGTGAGTCCCATCATCCTGGGGTGGGAGAGGGGCC----------- -3823
Chimpanzee VIPR2 Gene -4454 TCTCTGTGAGCCCCACCTGCCTGGGGTGGGAGGGGCCCG---------- -4414
Gorilla VIPR2 Gene -4139 TCTCTGTGAGCCCCACCTGCCTGGGGTGGGAGGGGGCCCAGGCAGGGCT -4090
GR
Human VIPR2 Gene -4343 GGCTGCGGCACGCACTGGTGTGTGCCCGTCTCTTGGACAAACCATCTGAG -4294
Marmoset VIPR2 Gene -3811 --CTGTGCCATGTGCTGGTGTGTGGCCAGCTCTTGGACAAACCATCCAAG -3764
Chimpanzee VIPR2 Gene -4407 GGCTGCGGCATGCACTGGTGTGTGCCCGTCTCTTGGACAAACCATCTGAG -4358
Gorilla VIPR2 Gene -4039 GGCTGCGGCACGCACTGGTGTGTGCCCGTCTCTTGGACAAACCATCTGAG -3990
GR
Human VIPR2 Gene -4293 ATGTTCTTCCAGCTGCTCTGCCTCCATCGCTGAGCCTCCTGCTGAGCTGA -4244
Marmoset VIPR2 Gene -3763 GTGTCCTTCCAGCAGCT---CCCAGACCCCTGAGGTTCCTCCCAAAATTT -3717
Chimpanzee VIPR2 Gene -4357 ATGTTCTTCCAGCTGCTCTGCCTCCATCGCTGAGCCTCCTGCTGAGCTGA -4308
Gorilla VIPR2 Gene -3989 ATGTTCTTCCAG----TCTGCCCCCATCGCTGAGCCTCCTGCTGAGCTGA -3944
E-BOX
Human VIPR2 Gene -4144 GGGAGGGCTTTCCTAAGGCAGACACCTGAGCCAAGACAGGTGGAAGCTGG -4095
Marmoset VIPR2 Gene -3664 AATATTTAGGTTTTAAAAGATACAAAAAGAGGGACACAGGCAAATGATGA -3615
Chimpanzee VIPR2 Gene -4208 GGGAGGGCTTTCCTAAGGCAGACACCTGAGCCAAGACAGGTGGAAGCTGG -4159
Gorilla VIPR2 Gene -3843 GCGAGGGCTTTCCTAAGGCAGACACCTGAGCCAAGACAGGTGGAAGCTGG -3794
STAT
Human VIPR2 Gene -3945 TAGGGTGACCATGTGCTTCACATAATATGTCCAAACCAGGACATTTCCAA -3896
Marmoset VIPR2 Gene -3504 CAAAGTACTCATCTGG---ATGTGGCAGAGACA--CCAGAAGGATCTCAG -3460
Chimpanzee VIPR2 Gene -4009 TAGGGTGACCATGTGCTTCACATAATATGTCCAAACCAGGACATTTCCAA -3960
Gorilla VIPR2 Gene -3660 TAGGGTGACCATGTGCTTCACATAATATGTCCAAACCAGGACATTTCCAA -3611

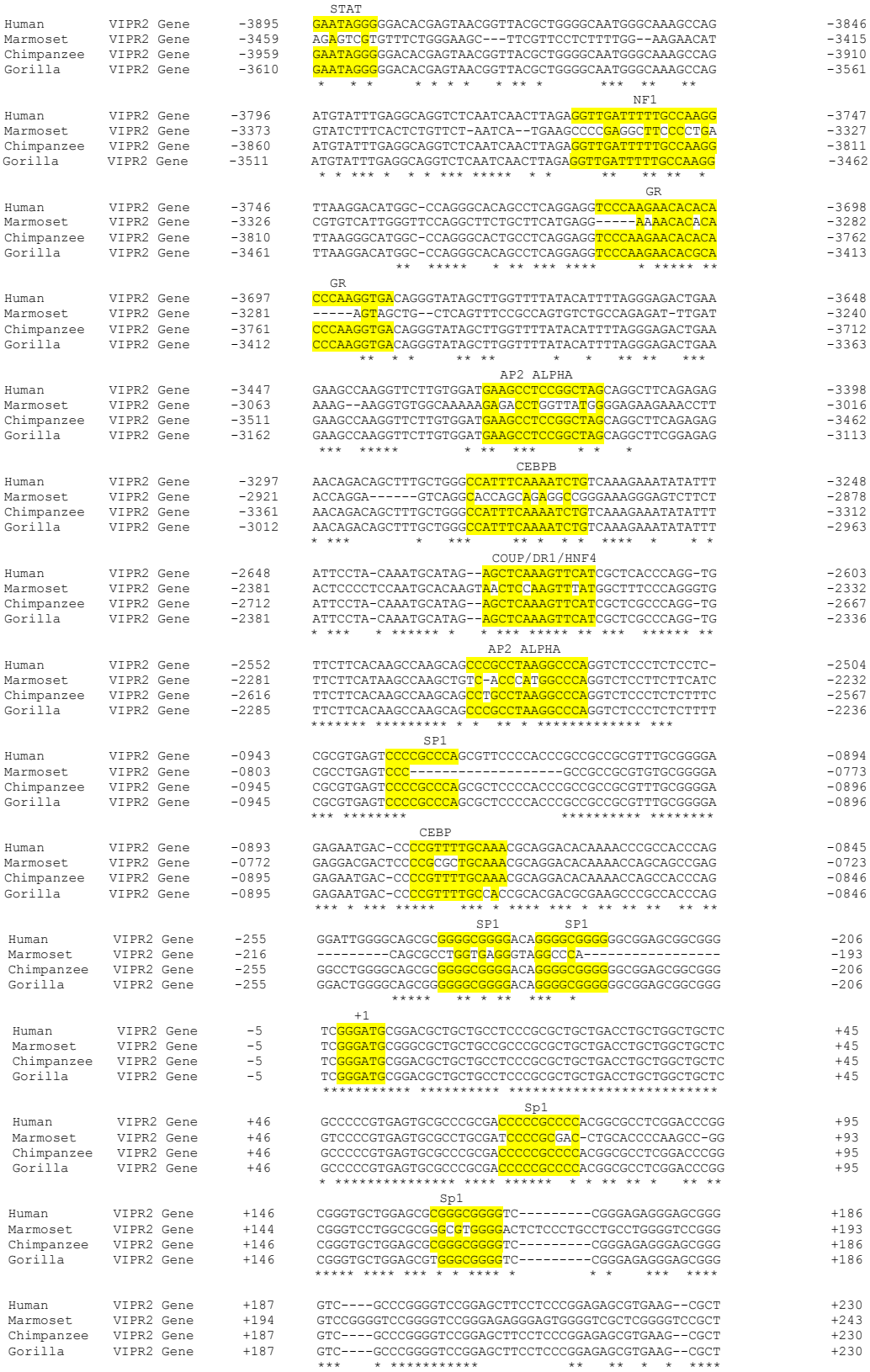
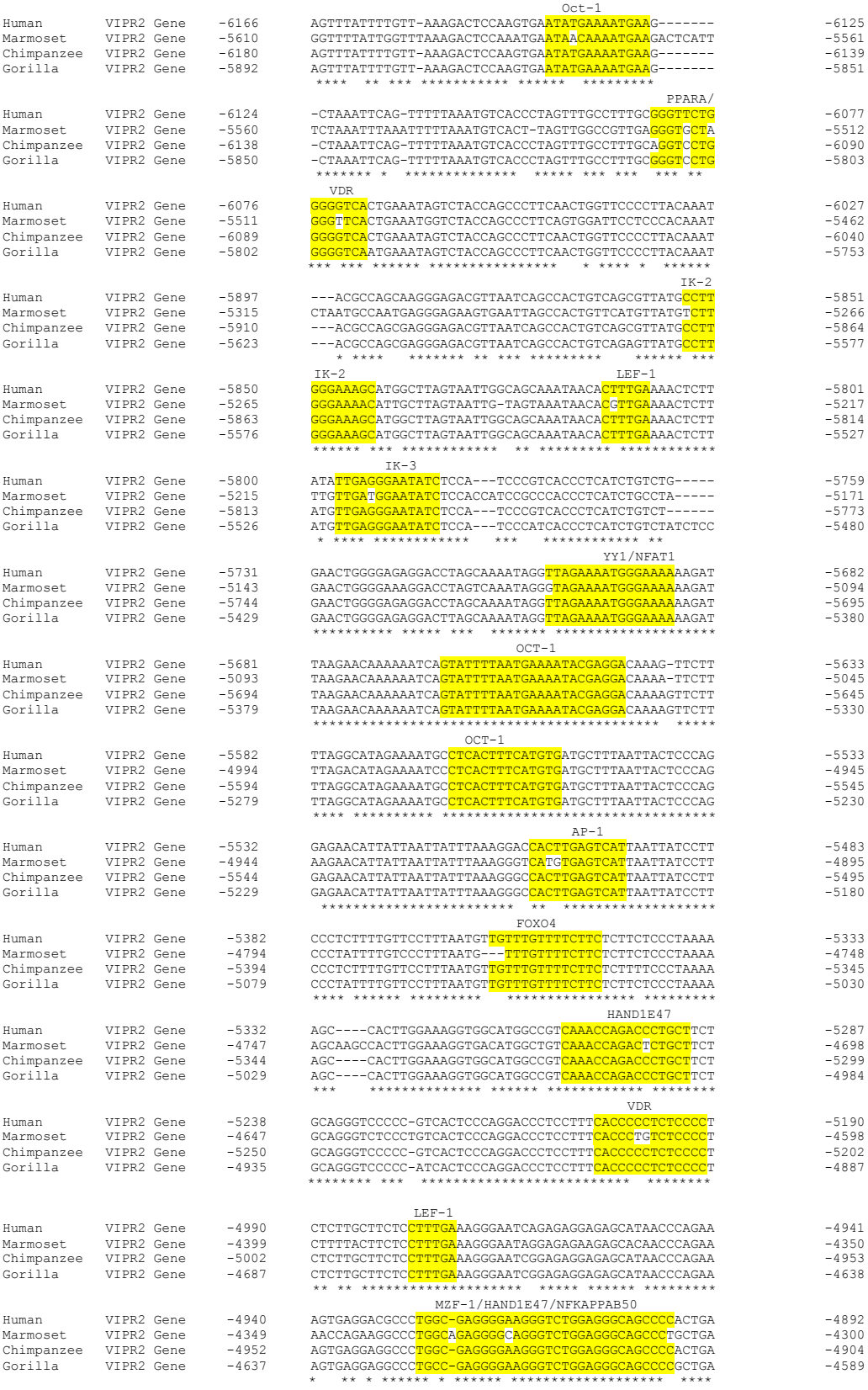
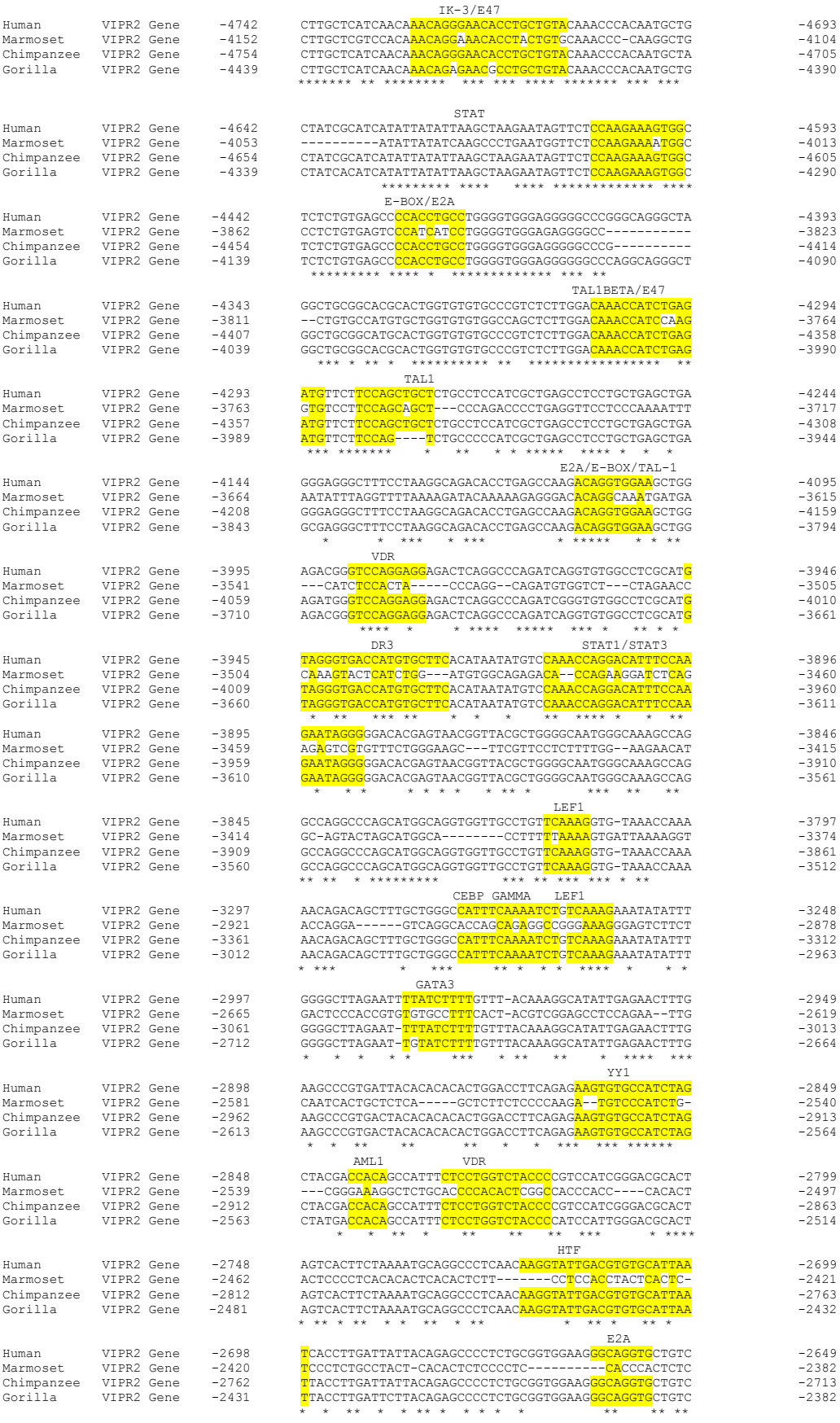
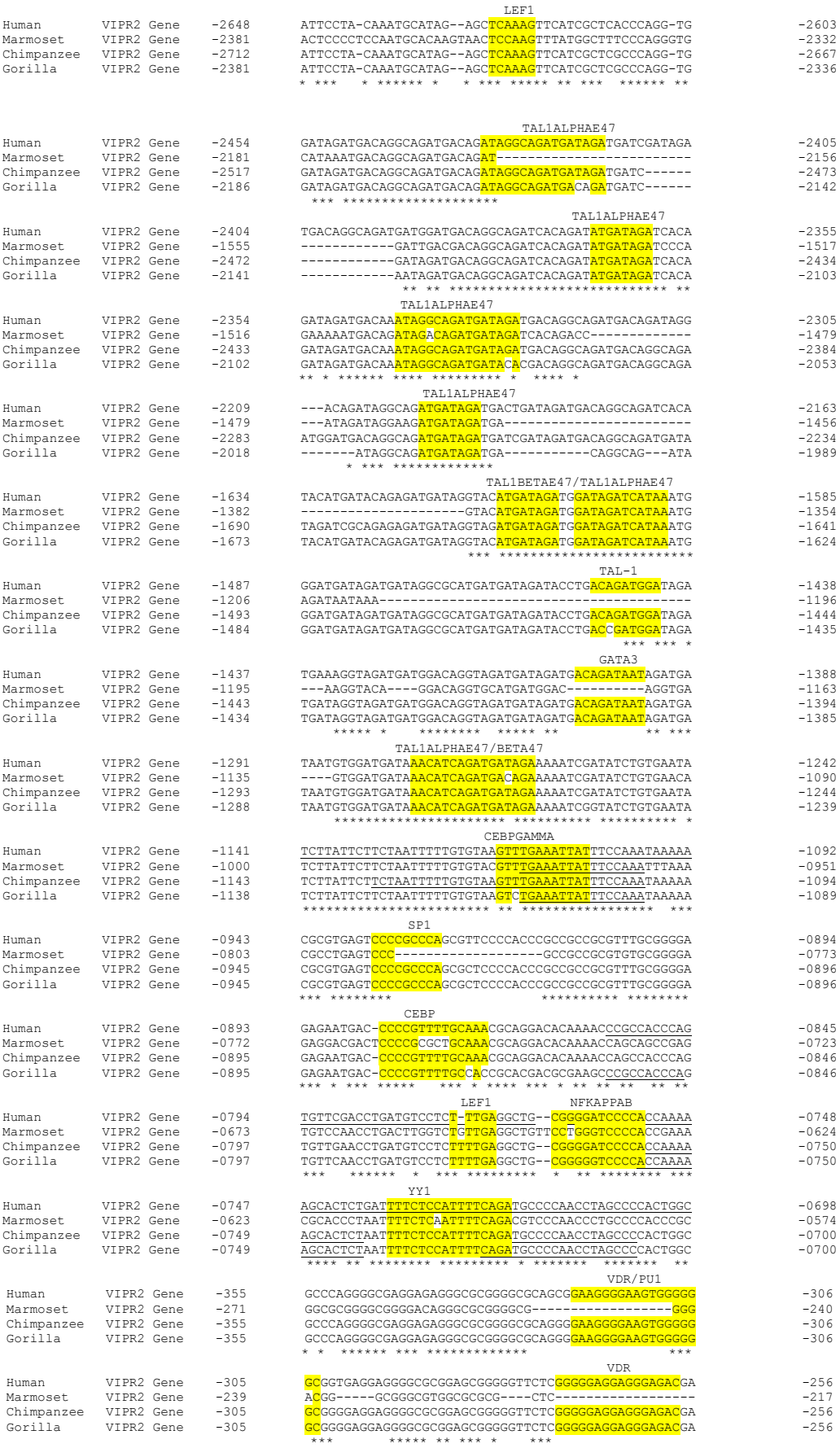


Figure 3.11 Evolutionarily conserved elements important in adipocytes


Oct-1
PPARA/
VDR
IK-2
IK-2
LEF-1
IK-3
YY1/NFAT1
OCT-1
OCT-1
AP-1
FOXO4
HAND1E47
VDR
LEF-1
MZF-1/HAND1E47/NFKAPPAB50
Human VIPR2 Gene -6166 AGTTTATTTTGTT-AAAGACTCCAAGTGAATATGAAAATGAAG------- -6125
Marmoset VIPR2 Gene -5610 GGTTTTATTGGTTTAAAGACTCCAAATGAATAACAAAATGAAGACTCATT -5561
Chimpanzee VIPR2 Gene -6180 AGTTTATTTTGTT-AAAGACTCCAAGTGAATATGAAAATGAAG------- -6139
Gorilla VIPR2 Gene -5892 AGTTTATTTTGTT-AAAGACTCCAAGTGAATATGAAAATGAAG------- -5851
Human VIPR2 Gene -6124 -CTAAATTCAG-TTTTTAAATGTCACCCTAGTTTGCCTTTGCGGGTTCTG -6077
Marmoset VIPR2 Gene -5560 TCTAAATTTAAATTTTTAAATGTCACT-TAGTTGGCCGTTGAGGGTGCTA -5512
Chimpanzee VIPR2 Gene -6138 -CTAAATTCAG-TTTTTAAATGTCACCCTAGTTTGCCTTTGCAGGTCCTG -6090
Gorilla VIPR2 Gene -5850 -CTAAATTCAG-TTTTTAAATGTCACCCTAGTTTGCCTTTGCGGGTCCTG -5803
Human VIPR2 Gene -6076 GGGGTCACTGAAATAGTCTACCAGCCCTTCAACTGGTTCCCCTTACAAAT -6027
Marmoset VIPR2 Gene -5511 GGGTTCACTGAAATGGTCTACCAGCCCTTCAGTGGATTCCTCCCACAAAT -5462
Chimpanzee VIPR2 Gene -6089 GGGGTCACTGAAATAGTCTACCAGCCCTTCAACTGGTTCCCCTTACAAAT -6040
Gorilla VIPR2 Gene -5802 GGGGTCAATGAAATAGTCTACCAGCCCTTCAACTGGTTCCCCTTACAAAT -5753
Human VIPR2 Gene -5897 ---ACGCCAGCAAGGGAGACGTTAATCAGCCACTGTCAGCGTTATGCCTT -5851
Marmoset VIPR2 Gene -5315 CTAATGCCAATGAGGGAGAAGTGAATTAGCCACTGTTCATGTTATGTCTT -5266
Chimpanzee VIPR2 Gene -5910 ---ACGCCAGCGAGGGAGACGTTAATCAGCCACTGTCAGCGTTATGCCTT -5864
Gorilla VIPR2 Gene -5623 ---ACGCCAGCGAGGGAGACGTTAATCAGCCACTGTCAGAGTTATGCCTT -5577
Human VIPR2 Gene -5850 GGGAAAGCATGGCTTAGTAATTGGCAGCAAATAACACTTTGAAAACTCTT -5801
Marmoset VIPR2 Gene -5265 GGGAAAACATTGCTTAGTAATTG-TAGTAAATAACACGTTGAAAACTCTT -5217
Chimpanzee VIPR2 Gene -5863 GGGAAAGCATGGCTTAGTAATTGGCAGCAAATAACACTTTGAAAACTCTT -5814
Gorilla VIPR2 Gene -5576 GGGAAAGCATGGCTTAGTAATTGGCAGCAAATAACACTTTGAAAACTCTT -5527
Human VIPR2 Gene -5800 ATATTGAGGGAATATCTCCA---TCCCGTCACCCTCATCTGTCTG----- -5759
Marmoset VIPR2 Gene -5215 TTGTTGATGGAATATCTCCACCATCCGCCCACCCTCATCTGCCTA----- -5171
Chimpanzee VIPR2 Gene -5813 ATGTTGAGGGAATATCTCCA---TCCCGTCACCCTCATCTGTCT------ -5773
Gorilla VIPR2 Gene -5526 ATGTTGAGGGAATATCTCCA---TCCCATCACCCTCATCTGTCTATCTCC -5480
Human VIPR2 Gene -5731 GAACTGGGGAGAGGACCTAGCAAAATAGGTTAGAAAATGGGAAAAAAGAT -5682
Marmoset VIPR2 Gene -5143 GAACTGGGGAAAGGACCTAGTCAAATAGGGTAGAAAATGGGAAAAAAGAT -5094
Chimpanzee VIPR2 Gene -5744 GAACTGGGGAGAGGACCTAGCAAAATAGGTTAGAAAATGGGAAAAAAGAT -5695
Gorilla VIPR2 Gene -5429 GAACTGGGGAGAGGACTTAGCAAAATAGGTTAGAAAATGGGAAAAAAGAT -5380
Human VIPR2 Gene -5681 TAAGAACAAAAAATCAGTATTTTAATGAAAATACGAGGACAAAG-TTCTT -5633
Marmoset VIPR2 Gene -5093 TAAGAACAAAAAATCAGTATTTTAATGAAAATACGAGGACAAAA-TTCTT -5045
Chimpanzee VIPR2 Gene -5694 TAAGAACAAAAAATCAGTATTTTAATGAAAATACGAGGACAAAAGTTCTT -5645
Gorilla VIPR2 Gene -5379 TAAGAACAAAAAATCAGTATTTTAATGAAAATACGAGGACAAAAGTTCTT -5330
Human VIPR2 Gene -5582 TTAGGCATAGAAAATGCCTCACTTTCATGTGATGCTTTAATTACTCCCAG -5533
Marmoset VIPR2 Gene -4994 TTAGACATAGAAAATCCCTCACTTTCATGTGATGCTTTAATTACTCCCAG -4945
Chimpanzee VIPR2 Gene -5594 TTAGGCATAGAAAATGCCTCACTTTCATGTGATGCTTTAATTACTCCCAG -5545
Gorilla VIPR2 Gene -5279 TTAGGCATAGAAAATGCCTCACTTTCATGTGATGCTTTAATTACTCCCAG -5230
Human VIPR2 Gene -5532 GAGAACATTATTAATTATTTAAAGGACCACTTGAGTCATTAATTATCCTT -5483
Marmoset VIPR2 Gene -4944 AAGAACATTATTAATTATTTAAAGGGTCATGTGAGTCATTAATTATCCTT -4895
Chimpanzee VIPR2 Gene -5544 GAGAACATTATTAATTATTTAAAGGGCCACTTGAGTCATTAATTATCCTT -5495
Gorilla VIPR2 Gene -5229 GAGAACATTATTAATTATTTAAAGGGCCACTTGAGTCATTAATTATCCTT -5180
Human VIPR2 Gene -5382 CCCTCTTTTGTTCCTTTAATGTTGTTTGTTTTCTTCTCTTCTCCCTAAAA -5333
Marmoset VIPR2 Gene -4794 CCCTATTTTGTCCCTTTAATG---TTTGTTTTCTTCTCTTCTCCCTAAAA -4748
Chimpanzee VIPR2 Gene -5394 CCCTCTTTTGTTCCTTTAATGTTGTTTGTTTTCTTCTCTTTTCCCTAAAA -5345
Gorilla VIPR2 Gene -5079 CCCTATTTTGTTCCTTTAATGTTGTTTGTTTTCTTCTCTTCTCCCTAAAA -5030
Human VIPR2 Gene -5332 AGC----CACTTGGAAAGGTGGCATGGCCGTCAAACCAGACCCTGCTTCT -5287
Marmoset VIPR2 Gene -4747 AGCAAGCCACTTGGAAAGGTGACATGGCTGTCAAACCAGACTCTGCTTCT -4698
Chimpanzee VIPR2 Gene -5344 AGC----CACTTGGAAAGGTGGCATGGCCGTCAAACCAGACCCTGCTTCT -5299
Gorilla VIPR2 Gene -5029 AGC----CACTTGGAAAGGTGGCATGGCCGTCAAACCAGACCCTGCTTCT -4984
Human VIPR2 Gene -5238 GCAGGGTCCCCC-GTCACTCCCAGGACCCTCCTTTCACCCCCTCTCCCCT -5190
Marmoset VIPR2 Gene -4647 GCAGGGTCTCCCTGTCACTCCCAGGACCCTCCTTTCACCCTGTCTCCCCT -4598
Chimpanzee VIPR2 Gene -5250 GCAGGGTCCCCC-GTCACTCCCAGGACCCTCCTTTCACCCCCTCTCCCCT -5202
Gorilla VIPR2 Gene -4935 GCAGGGTCCCCC-ATCACTCCCAGGACCCTCCTTTCACCCCCTCTCCCCT -4887
Human VIPR2 Gene -4990 CTCTTGCTTCTCCTTTGAAGGGAATCAGAGAGGAGAGCATAACCCAGAA -4941
Marmoset VIPR2 Gene -4399 CTTTTACTTCTCCTTTGAAAGGGAATAGGAGAGAAGAGCACAACCCAGAA -4350
Chimpanzee VIPR2 Gene -5002 CTCTTGCTTCTCCTTTGAAAGGGAATCGGAGAGGAGAGCATAACCCAGAA -4953
Gorilla VIPR2 Gene -4687 CTCTTGCTTCTCCTTTGAAGGGAATCGGAGAGGAGAGCATAACCCAGAA -4638
Human VIPR2 Gene -4940 AGTGAGGACGCCCTGGC-GAGGGGAAGGGTCTGGAGGGCAGCCCCACTGA -4892
Marmoset VIPR2 Gene -4349 AACCAGAAGGCCCTGGCAGAGGGGCAGGGTCTGGAGGGCAGCCCTGCTGA -4300
Chimpanzee VIPR2 Gene -4952 AGTGAGGAGGCCCTGGC-GAGGGGAAGGGTCTGGAGGGCAGCCCCACTGA -4904
Gorilla VIPR2 Gene -4637 AGTGAGGAGGCCCTGCC-GAGGGGAAGGGTCTGGAGGGCAGCCCCGCTGA -4589


IK-3/E47
Human VIPR2 Gene -4742 CTTGCTCATCAACAAACAGGGAACACCTGCTGTACAAACCCACAATGCTG -4693
Marmoset VIPR2 Gene -4152 CTTGCTCGTCCACAAACAGGAAACACCTACTGTGCAAACCC-CAAGGCTG -4104
Chimpanzee VIPR2 Gene -4754 CTTGCTCATCAACAAACAGGGAACACCTGCTGTACAAACCCACAATGCTA -4705
Gorilla VIPR2 Gene -4439 CTTGCTCATCAACAAACAGAACGCCTGCTGTACAAACCCACAATGCTG -4390
STAT
Human VIPR2 Gene -4642 CTATCGCATCATATTATATTAAGCTAAGAATAGTTCTCCAAGAAAGTGGC -4593
Marmoset VIPR2 Gene -4053 ----------ATATTATATCAAGCCCTGAATGGTTCTCCAAGAAAATGGC -4013
Chimpanzee VIPR2 Gene -4654 CTATCGCATCATATTATATTAAGCTAAGAATAGTTCTCCAAGAAAGTGGC -4605
Gorilla VIPR2 Gene -4339 CTATCACATCATATTATATTAAGCTAAGAATAGTTCTCCAAGAAAGTGGC -4290
E-BOX/E2A
Human VIPR2 Gene -4442 TCTCTGTGAGCCCCACCTGCCTGGGGTGGGAGGGGGCCCGGGCAGGGCTA -4393
Marmoset VIPR2 Gene -3862 CCTCTGTGAGTCCCATCATCCTGGGGTGGGAGAGGGGCC---------- -3823
Chimpanzee VIPR2 Gene -4454 TCTCTGTGAGCCCCACCTGCCTGGGGTGGGAGGGGGCCCG---------- -4414
Gorilla VIPR2 Gene -4139 TCTCTGTGAGCCCCACCTGCCTGGGGTGGGAGGGGGGCCCAGGCAGGGCT -4090
TAL1BETA/E47
Human VIPR2 Gene -4343 GGCTGCGGCACGCACTGGTGTGTGCCCGTCTCTTGGACAAACCATCTGAG -4294
Marmoset VIPR2 Gene -3811 --CTGTGCCATGTGCTGGTGTGTGGCCAGCTCTTGGACAAACCATCCAAG -3764
Chimpanzee VIPR2 Gene -4407 GGCTGCGGCATGCACTGGTGTGTGCCCGTCTCTTGGACAAACCATCTGAG -4358
Gorilla VIPR2 Gene -4039 GGCTGCGGCACGCACTGGTGTGTGCCCGTCTCTTGGACAAACCATCTGAG -3990
TAL1
Human VIPR2 Gene -4293 ATGTTCTTCCAGCTGCTCTGCCTCCATCGCTGAGCCTCCTGCTGAGCTGA -4244
Marmoset VIPR2 Gene -3763 GTGTCCTTCCAGCAGCT---CCCAGACCCCTGAGGTTCCTCCCAAAATTT -3717
Chimpanzee VIPR2 Gene -4357 ATGTTCTTCCAGCTGCTCTGCCTCCATCGCTGAGCCTCCTGCTGAGCTGA -4308
Gorilla VIPR2 Gene -3989 ATGTTCTTCCAG----TCTGCCCCCATCGCTGAGCCTCCTGCTGAGCTGA -3944
E2A/E-BOX/TAL-1
Human VIPR2 Gene -4144 GGGAGGGCTTTCCTAAGGCAGACACCTGAGCCAAGACAGGTGGAAGCTGG -4095
Marmoset VIPR2 Gene -3664 AATATTTAGGTTTTAAAAGATACAAAAAGAGGGACACAGGCAAATGATGA -3615
Chimpanzee VIPR2 Gene -4208 GGGAGGGCTTTCCTAAGGCAGACACCTGAGCCAAGACAGGTGGAAGCTGG -4159
Gorilla VIPR2 Gene -3843 GCGAGGGCTTTCCTAAGGCAGACACCTGAGCCAAGACAGGTGGAAGCTGG -3794
VDR
Human VIPR2 Gene -3995 AGACGGGTCCAGGAGGAGACTCAGGCCCAGATCAGGTGTGGCCTCGCATG -3946
Marmoset VIPR2 Gene -3541 ---CATCTCCACTA-----CCCAGG--CAGATGTGGTCT---CTAGAACC -3505
Chimpanzee VIPR2 Gene -4059 AGATGGGTCCAGGAGGAGACTCAGGCCCAGATCGGGTGTGGCCTCGCATG -4010
Gorilla VIPR2 Gene -3710 AGACGGGTCCAGGAGGAGACTCAGGCCCAGATCAGGTGTGGCCTCGCATG -3661
DR3 STAT1/STAT3
Human VIPR2 Gene -3945 TAGGGTGACCATGTGCTTCACATAATATGTCCAAACCAGGACATTTCCAA -3896
Marmoset VIPR2 Gene -3504 CAAAGTACTCATCTGG---ATGTGGCAGAGACA--CCAGAAGGATCTCAG -3460
Chimpanzee VIPR2 Gene -4009 TAGGGTGACCATGTGCTTCACATAATATGTCCAAACCAGGACATTTCCAA -3960
Gorilla VIPR2 Gene -3660 TAGGGTGACCATGTGCTTCACATAATATGTCCAAACCAGGACATTTCCAA -3611
Human VIPR2 Gene -3895 GAATAGGGGGACACGAGTAACGGTTACGCTGGGGCAATGGGCAAAGCCAG -3846
Marmoset VIPR2 Gene -3459 AGAGTCGTGTTTCTGGGAAGC---TTCGTTCCTCTTTTGG--AAGAACAT -3415
Chimpanzee VIPR2 Gene -3959 GAATAGGGGGACACGAGTAACGGTTACGCTGGGGCAATGGGCAAAGCCAG -3910
Gorilla VIPR2 Gene -3610 GAATAGGGGGACACGAGTAACGGTTACGCTGGGGCAATGGGCAAAGCCAG -3561
LEF1
Human VIPR2 Gene -3845 GCCAGGCCCAGCATGGCAGGTGGTTGCCTGTTCAAAGGTG-TAAACCAAA -3797
Marmoset VIPR2 Gene -3414 GC-AGTACTAGCATGGCA---------CCTTTTTAAAAGTGATTAAAAGGT -3374
Chimpanzee VIPR2 Gene -3909 GCCAGGCCCAGCATGGCAGGTGGTTGCCTGTTCAAAGGTG-TAAACCAAA -3861
Gorilla VIPR2 Gene -3560 GCCAGGCCCAGCATGGCAGGTGGTTGCCTGTTCAAAGGTG-TAAACCAAA -3512
CEBP GAMMA LEF1
Human VIPR2 Gene -3297 AACAGACAGCTTTGCTGGGCCATTTCAAAATCTGTCAAAGAAATATATTT -3248
Marmoset VIPR2 Gene -2921 ACCAGGA------GTCAGGCACCAGCAGAGGCCGGGAAAGGGAGTCTTCT -2878
Chimpanzee VIPR2 Gene -3361 AACAGACAGCTTTGCTGGGCCATTTCAAAATCTGTCAAAGAAATATATTT -3312
Gorilla VIPR2 Gene -3012 AACAGACAGCTTTGCTGGGCCATTTCAAAATCTGTCAAAGAAATATATTT -2963
GATA3
Human VIPR2 Gene -2997 GGGGCTTAGAATTTTATCTTTTGTTT-ACAAAGGCATATTGAGAACTTTG -2949
Marmoset VIPR2 Gene -2665 GACTCCCACCGTGTGTGCCTTTCACT-ACGTCGGAGCCTCCAGAA--TTG -2619
Chimpanzee VIPR2 Gene -3061 GGGGCTTAGAAT-TTTATCTTTTGTTTACAAAGGCATATTGAGAACTTTG -3013
Gorilla VIPR2 Gene -2712 GGGGCTTAGAAT-TGTATCTTTTGTTTACAAAGGCATATTGAGAACTTTG -2664
YY1
Human VIPR2 Gene -2898 AAGCCCGTGATTACACACACACTGGACCTTCAGAGAAGTGTGCCATCTAG -2849
Marmoset VIPR2 Gene -2581 CAATCACTGCTCTCA-----GCTCTTCTCCCCAAGA--TGTCCCATCTG- -2540
Chimpanzee VIPR2 Gene -2962 AAGCCCGTGACTACACACACACTGGACCTTCAGAGAAGTGTGCCATCTAG -2913
Gorilla VIPR2 Gene -2613 AAGCCCGTGACTACACACACACTGGACCTTCAGAGAAGTGTGCCATCTAG -2564
AML1 VDR
Human VIPR2 Gene -2848 CTACGACCACAGCCATTTCTCCTGGTCTACCCCGTCCATCGGGACGCACT -2799
Marmoset VIPR2 Gene -2539 ---CGGGAAAGGCTCTGCACCCCACACTCGGCCACCCACC----CACACT -2497
Chimpanzee VIPR2 Gene -2912 CTACGACCACAGCCATTTCTCCTGGTCTACCCCGTCCATCGGGACGCACT -2863
Gorilla VIPR2 Gene -2563 CTATGACCACAGCCATTTCTCCTGGTCTACCCCATCCATTGGGACGCACT -2514
HTF
Human VIPR2 Gene -2748 AGTCACTTCTAAAATGCAGGCCCTCAACAAGGTATTGACGTGTGCATTAA -2699
Marmoset VIPR2 Gene -2462 ACTCCCCTCACACACTCACACTCTT-------CCTCCACCTACTCACTC- -2421
Chimpanzee VIPR2 Gene -2812 AGTCACTTCTAAAATGCAGGCCCTCAACAAGGTATTGACGTGTGCATTAA -2763
Gorilla VIPR2 Gene -2481 AGTCACTTCTAAAATGCAGGCCCTCAACAAGGTATTGACGTGTGCATTAA -2432
E2A
Human VIPR2 Gene -2698 TCACCTTGATTATTACAGAGCCCCTCTGCGGTGGAAGGGCAGGTGCTGTC -2649
Marmoset VIPR2 Gene -2420 TCCCTCTGCCTACT-CACACTCTCCCCTC----------CACCCACTCTC -2382
Chimpanzee VIPR2 Gene -2762 TTACCTTGATTATTACAGAGCCCCTCTGCGGTGGAAGGGCAGGTGCTGTC -2713
Gorilla VIPR2 Gene -2431 TTACCTTGATTCTTACAGAGCCCCTCTGCGGTGGAAGGGCAGGTGCTGTC -2382


LEF1
Human VIPR2 Gene -2648 ATTCCTA-CAAATGCATAG--AGCTCAAAGTTCATCGCTCACCCAGG-TG -2603
Marmoset VIPR2 Gene -2381 ACTCCCCTCCAATGCACAAGTAACTCCAAGTTTATGGCTTTCCCAGGGTG -2332
Chimpanzee VIPR2 Gene -2712 ATTCCTA-CAAATGCATAG--AGCTCAAAGTTCATCGCTCGCCCAGG-TG -2667
Gorilla VIPR2 Gene -2381 ATTCCTA-CAAATGCATAG--AGCTCAAAGTTCATCGCTCGCCCAGG-TG -2336
TAL1ALPHAE47
Human VIPR2 Gene -2454 GATAGATGACAGGCAGATGACAGATAGGCAGATGATAGATGATCGATAGA -2405
Marmoset VIPR2 Gene -2181 CATAAATGACAGGCAGATGACAGAT------------------------ -2156
Chimpanzee VIPR2 Gene -2517 GATAGATGACAGGCAGATGACAGATAGGCAGATGATAGATGATC------ -2473
Gorilla VIPR2 Gene -2186 GATAGATGACAGGCAGATGACAGATAGGCAGATGACAGATGATC------ -2142
TAL1ALPHAE47
Human VIPR2 Gene -2404 TGACAGGCAGATGATGGATGACAGGCAGATCACAGATATGATAGATCACA -2355
Marmoset VIPR2 Gene -1555 ------------GATTGACGACAGGCAGATCACAGATATGATAGATCCCA -1517
Chimpanzee VIPR2 Gene -2472 ------------GATAGATGACAGGCAGATCACAGATATGATAGATCACA -2434
Gorilla VIPR2 Gene -2141 ------------AATAGATGACAGGCAGATCACAGATATGATAGATCACA -2103
TAL1ALPHAE47
Human VIPR2 Gene -2354 GATAGATGACAAATAGGCAGATGATAGATGACAGGCAGATGACAGATAGG -2305
Marmoset VIPR2 Gene -1516 GAAAAATGACAGATAGACAGATGATAGATCACAGACC------------- -1479
Chimpanzee VIPR2 Gene -2433 GATAGATGACAAATAGGCAGATGATAGATGACAGGCAGATGACAGGCAGA -2384
Gorilla VIPR2 Gene -2102 GATAGATGACAAATAGGCAGATGATACACGACAGGCAGATGACAGGCAGA -2053
TAL1ALPHAE47
Human VIPR2 Gene -2209 ---ACAGATAGGCAGATGATAGATGACTGATAGATGACAGGCAGATCACA -2163
Marmoset VIPR2 Gene -1479 ---ATAGATAGGAAGATGATAGATGA------------------------ -1456
Chimpanzee VIPR2 Gene -2283 ATGGATGACAGGCAGATGATAGATGATCGATAGATGACAGGCAGATGATA -2234
Gorilla VIPR2 Gene -2018 -------ATAGGCAGATGATAGATGA----------CAGGCAG---ATA -1989
TAL1BETAE47/TAL1ALPHAE47
Human VIPR2 Gene -1634 TACATGATACAGAGATGATAGGTACATGATAGATGGATAGATCATAAATG -1585
Marmoset VIPR2 Gene -1382 ----------GTACATGATAGATGGATAGATCATAAATG -1354
Chimpanzee VIPR2 Gene -1690 TAGATCGCAGAGAGATGATAGGTAGATGATAGATGGATAGATCATAAATG -1641
Gorilla VIPR2 Gene -1673 TACATGATACAGAGATGATAGGTACATGATAGATGGATAGATCATAAATG -1624
TAL-1
Human VIPR2 Gene -1487 GGATGATAGATGATAGGCGCATGATGATAGATACCTGACAGATGGATAGA -1438
Marmoset VIPR2 Gene -1206 AGATAATAAA---------- -1196
Chimpanzee VIPR2 Gene -1493 GGATGATAGATGATAGGCGCATGATGATAGATACCTGACAGATGGATAGA -1444
Gorilla VIPR2 Gene -1484 GGATGATAGATGATAGGCGCATGATGATAGATACCTGACCGATGGATAGA -1435
GATA3
Human VIPR2 Gene -1437 TGAAAGGTAGATGATGGACAGGTAGATGATAGATGACAGATAATAGATGA -1388
Marmoset VIPR2 Gene -1195 ---AAGGTACA----GGACAGGTGCATGATGGAC----------AGGTGA -1163
Chimpanzee VIPR2 Gene -1443 TGATAGGTAGATGATGGACAGGTAGATGATAGATGACAGATAATAGATGA -1394
Gorilla VIPR2 Gene -1434 TGATAGGTAGATGATGGACAGGTAGATGATAGATGACAGATAATAGATGA -1385
TAL1ALPHAE47/BETA47
Human VIPR2 Gene -1291 TAATGTGGATGATAAACATCAGATGATAGAAAAATCGATATCTGTGAATA -1242
Marmoset VIPR2 Gene -1135 ----GTGGATGATAAACATCAGATGACAGAAAAATCGATATCTGTGAACA -1090
Chimpanzee VIPR2 Gene -1293 TAATGTGGATGATAAACATCAGATGATAGAAAAATCGATATCTGTGAATA -1244
Gorilla VIPR2 Gene -1288 TAATGTGGATGATAAACATCAGATGATAGAAAATCGGTATCTGTGAATA -1239
CEBPGAMMA
Human VIPR2 Gene -1141 TCTTATTCTTCTAATTTTTGTGTAAGTTTGAAATTATTTCCAAATAAAAA -1092
Marmoset VIPR2 Gene -1000 TCTTATTCTTCTAATTTTTGTGTACGTTTGAAATTATTTCCAAATTTAAA -0951
Chimpanzee VIPR2 Gene -1143 TCTTATTCTTCTAATTTTTGTGTAAGTTTGAAATTATTTCCAAATAAAAA -1094
Gorilla VIPR2 Gene -1138 TCTTATTCTTCTAATTTTTGTGTAAGTCTGAAATTATTTCCAAATAAAAA -1089
SP1
Human VIPR2 Gene -0943 CGCGTGAGTCCCCGCCCAGCGTTCCCCACCCGCCGCCGCGTTTGCGGGGA -0894
Marmoset VIPR2 Gene -0803 CGCCTGAGTCCC-------------------GCCGCCGCGTGTGCGGGGA -0773
Chimpanzee VIPR2 Gene -0945 CGCGTGAGTCCCCGCCCAGCGCTCCCCACCCGCCGCCGCGTTTGCGGGGA -0896
Gorilla VIPR2 Gene -0945 CGCGTGAGTCCCCGCCCAGCGCTCCCCACCCGCCGCCGCGTTTGCGGGGA -0896
CEBP
Human VIPR2 Gene -0893 GAGAATGAC-CCCCGTTTTGCAAACGCAGGACACAAAACCCGCCCACCCAG -0845
Marmoset VIPR2 Gene -0772 GAGGACGACTCCCCGCGCTGCAAACGCAGGACACAAAACCAGCAGCCGAG -0723
Chimpanzee VIPR2 Gene -0895 GAGAATGAC-CCCCGTTTTGCAAACGCAGGACACAAAACCAGCCACCCAG -0846
Gorilla VIPR2 Gene -0895 GAGAATGAC-CCCCGTTTTGCCACCGCACGACGCGAAGCCCGCCACCCAG -0846
LEF1 NFKAPPAB
Human VIPR2 Gene -0794 TGTTCGACCTGATGTCCTCT-TTGAGGCTG--CGGGGATCCCCACCAAAA -0748
Marmoset VIPR2 Gene -0673 TGTCCAACCTGACTTGGTCTGTTGAGGCTGTTCCTGGGTCCCCACCGAAA -0624
Chimpanzee VIPR2 Gene -0797 TGTTGAACCTGATGTCCTCTTTTGAGGCTG--CGGGGATCCCCACCAAAA -0750
Gorilla VIPR2 Gene -0797 TGTTCAACCTGATGTCCTCTTTTGAGGCTG--CGGGGGTCCCCACCAAAA -0750
YY1
Human VIPR2 Gene -0747 AGCACTCTGATTTTCTCCATTTTCAGATGCCCCAACCTAGCCCCACTGGC -0698
Marmoset VIPR2 Gene -0623 CGCACCCTAATTTTCTCAATTTTCAGACGTCCCAACCCTGCCCCACCCGC -0574
Chimpanzee VIPR2 Gene -0749 AGCACTCTAATTTTCTCCATTTTCAGATGCCCCAACCTAGCCCCACTGGC -0700
Gorilla VIPR2 Gene -0749 AGCACTCTAATTTTCTCCATTTTCAGATGCCCCAACCTAGCCCCACTGGC -0700
VDR/PU1
Human VIPR2 Gene -355 GCCCAGGGGCGAGGAGAGGGCGCGGGGCGCAGCGGAAGGGGAAGTGGGGG -306
Marmoset VIPR2 Gene -271 GGCGCGGGGCGGGGACAGGGCGCGGGGCG----------GGG -240
Chimpanzee VIPR2 Gene -355 GCCCAGGGGCGAGGAGAGGGCGCGGGGCGCAGGGGAAGGGGAAGTGGGGG -306
Gorilla VIPR2 Gene -355 GCCCAGGGGCGAGGAGAGGGCGCGGGGCGCAGGGGAAGGGGAAGTGGGGG -306
VDR
Human VIPR2 Gene -305 GCGGTGAGGAGGGGCGCGGAGCGGGGGTTCTCGGGGGAGGAGGGAGACGA -256
Marmoset VIPR2 Gene -239 ACGG-----GCGGGCGTGGCGCGCG----CTC---------- -217
Chimpanzee VIPR2 Gene -305 GCGGGGAGGAGGGGCGCGGAGCGGGGGTTCTCGGGGGAGGAGGGAGACGA -256
Gorilla VIPR2 Gene -305 GCGGGGAGGAGGGGCGCGGAGCGGGGGTTCTCGGGGGAGGAGGGAGACGA -256

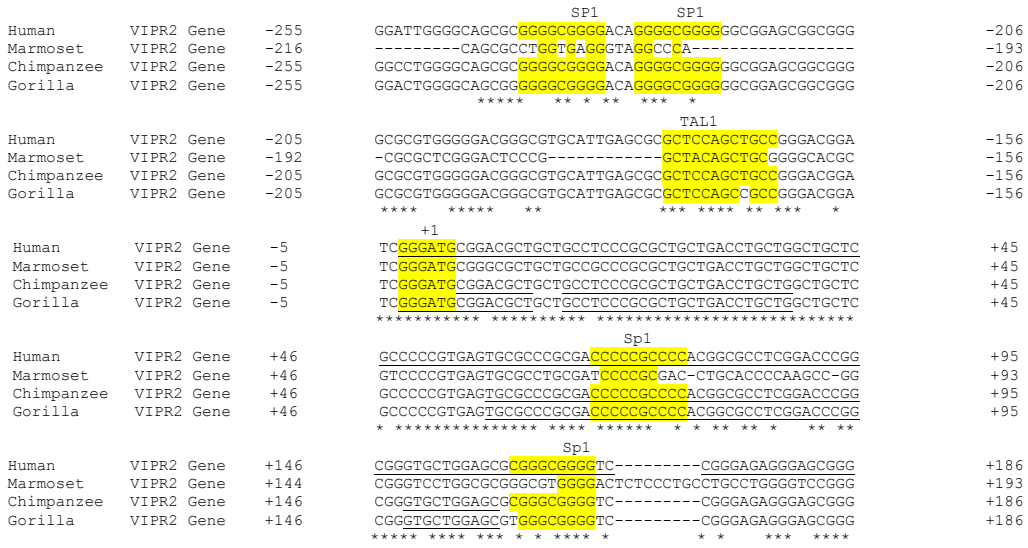
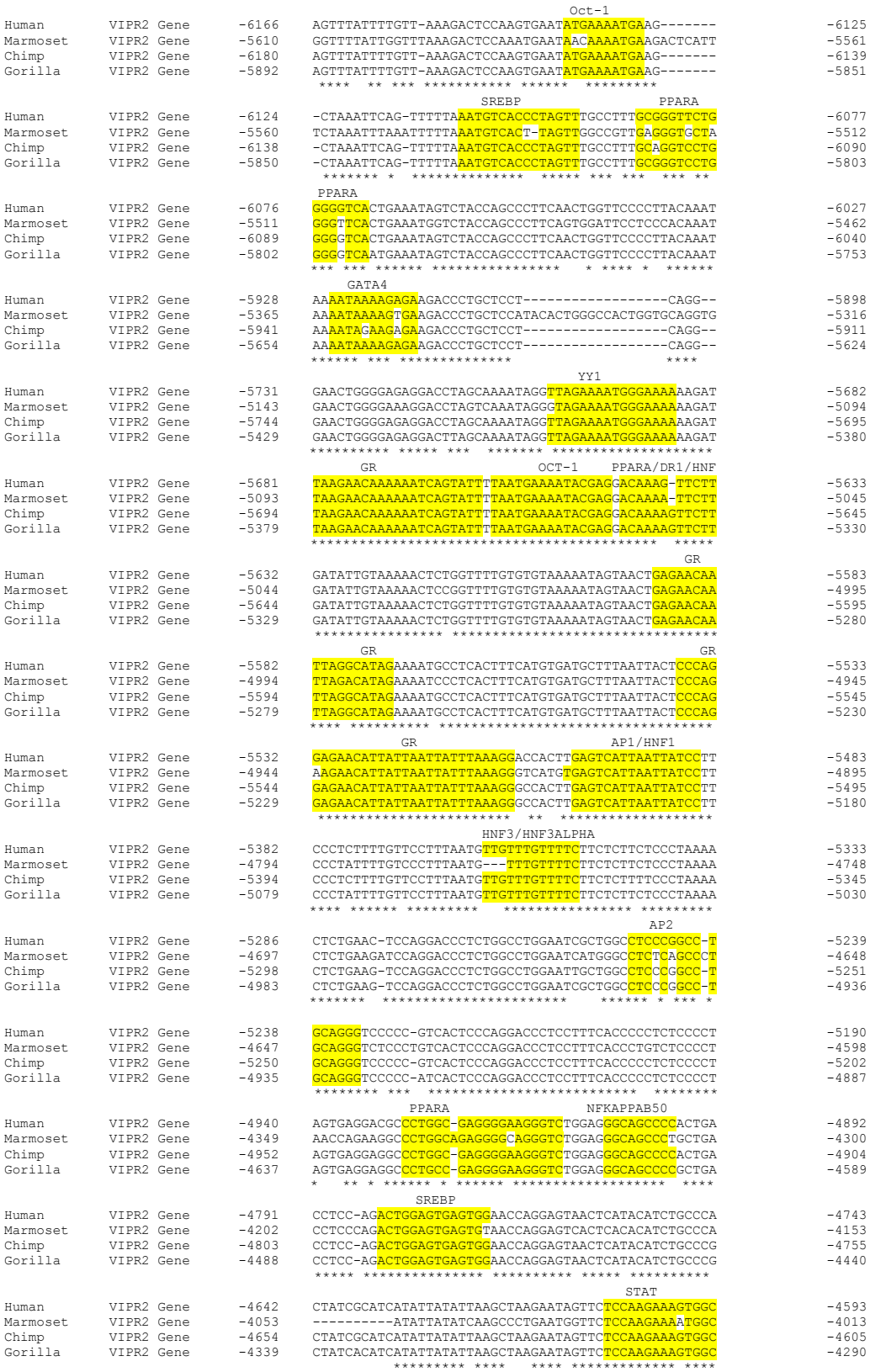
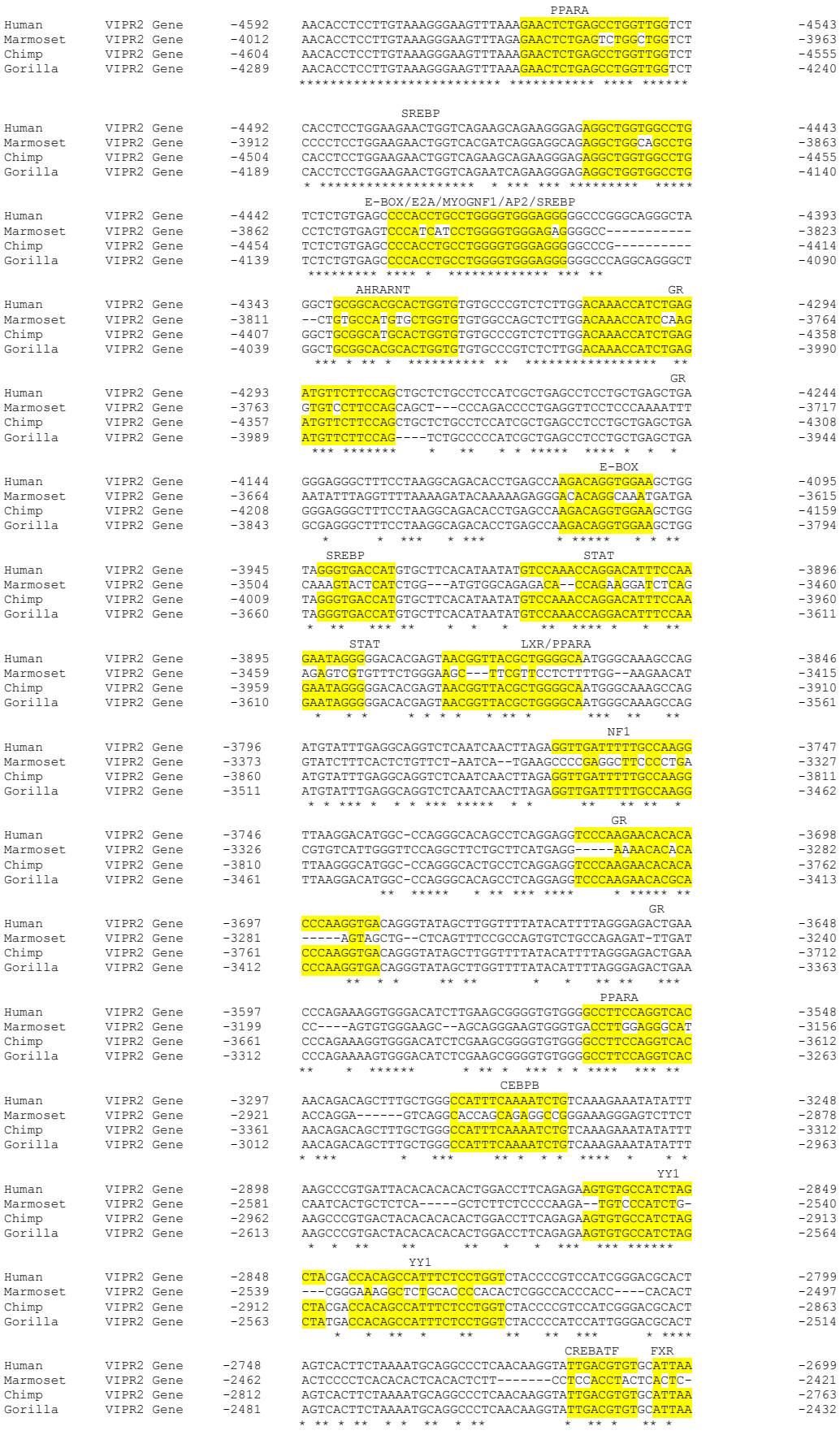
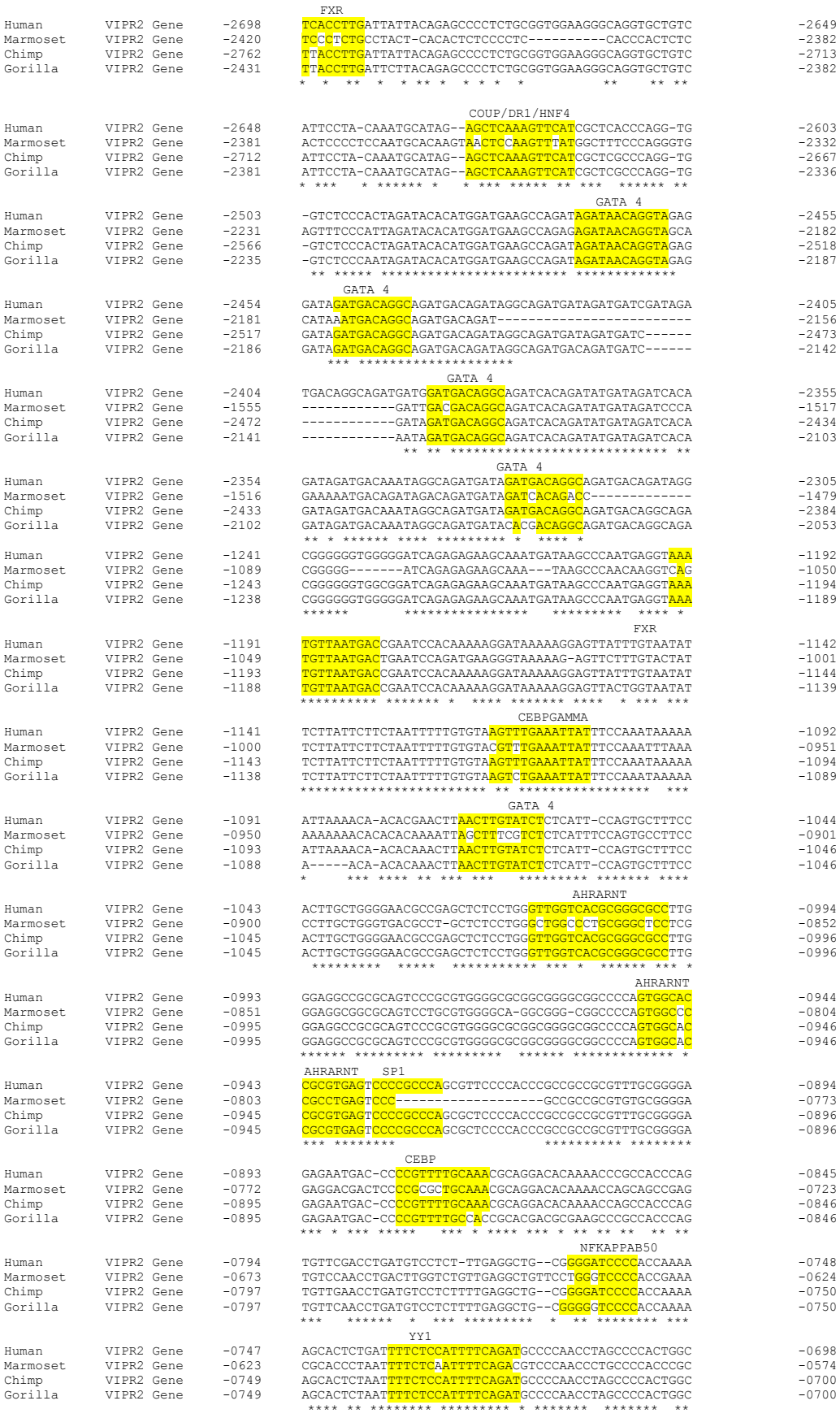


Figure 3.12 Evolutionarily conserved elements important in activated T cells.

The elements which were underscored were found to be conserved with elephant VIPR2 gene.


Oct-1
Human VIPR2 Gene -6166 AGTTTATTTTGTT-AAAGACTCCAAGTGAATATGAAAATGAAG------- -6125
Marmoset VIPR2 Gene -5610 GGTTTTATTGGTTTAAAGACTCCAAATGAATAACAAAATGAAGACTCATT -5561
Chimp VIPR2 Gene -6180 AGTTTATTTTGTT-AAAGACTCCAAGTGAATATGAAAATGAAG------- -6139
Gorilla VIPR2 Gene -5892 AGTTTATTTTGTT-AAAGACTCCAAGTGAATATGAAAATGAAG------- -5851
SREBP PPARA
Human VIPR2 Gene -6124 -CTAAATTCAG-TTTTTAAATGTCACCCTAGTTTGCCTTTGCGGGTTCTG -6077
Marmoset VIPR2 Gene -5560 TCTAAATTTAAATTTTTAAATGTCACT-TAGTTGGCCGTTGAGGGTGCTA -5512
Chimp VIPR2 Gene -6138 -CTAAATTCAG-TTTTTAAATGTCACCCTAGTTTGCCTTTGCAGGTCCTG -6090
Gorilla VIPR2 Gene -5850 -CTAAATTCAG-TTTTTAAATGTCACCCTAGTTTGCCTTTGCGGGTCCTG -5803
PPARA
Human VIPR2 Gene -6076 GGGGTCACTGAAATAGTCTACCAGCCCTTCAACTGGTTCCCCTTACAAAT -6027
Marmoset VIPR2 Gene -5511 GGGTTCACTGAAATGGTCTACCAGCCCTTCAGTGGATTCCTCCCACAAAT -5462
Chimp VIPR2 Gene -6089 GGGGTCACTGAAATAGTCTACCAGCCCTTCAACTGGTTCCCCTTACAAAT -6040
Gorilla VIPR2 Gene -5802 GGGGTCAATGAAATAGTCTACCAGCCCTTCAACTGGTTCCCCTTACAAAT -5753
GATA4
Human VIPR2 Gene -5928 AAAATAAAAGAGAAGACCCTGCTCCT------------------CAGG-- -5898
Marmoset VIPR2 Gene -5365 AAAATAAAAGTGAAGACCCTGCTCCATACACTGGGCCACTGGTGCAGGTG -5316
Chimp VIPR2 Gene -5941 AAAATAGAAGAGAAGACCCTGCTCCT------------------CAGG-- -5911
Gorilla VIPR2 Gene -5654 AAAATAAAAGAGAAGACCCTGCTCCT------------------CAGG-- -5624
YY1
Human VIPR2 Gene -5731 GAACTGGGGAGAGGACCTAGCAAAATAGGTTAGAAAATGGGAAAAAAGAT -5682
Marmoset VIPR2 Gene -5143 GAACTGGGGAAAGGACCTAGTCAAATAGGGTAGAAAATGGGAAAAAAGAT -5094
Chimp VIPR2 Gene -5744 GAACTGGGGAGAGGACCTAGCAAAATAGGTTAGAAAATGGGAAAAAAGAT -5695
Gorilla VIPR2 Gene -5429 GAACTGGGGAGAGGACTTAGCAAAATAGGTTAGAAAATGGGAAAAAAGAT -5380
GR OCT-1 PPARA/DR1/HNF
Human VIPR2 Gene -5681 TAAGAACAAAAAATCAGTATTTTAATGAAAATACGAGGACAAAG-TTCTT -5633
Marmoset VIPR2 Gene -5093 TAAGAACAAAAAATCAGTATTTTAATGAAAATACGAGGACAAAA-TTCTT -5045
Chimp VIPR2 Gene -5694 TAAGAACAAAAAATCAGTATTTTAATGAAAATACGAGGACAAAAGTTCTT -5645
Gorilla VIPR2 Gene -5379 TAAGAACAAAAAATCAGTATTTTAATGAAAATACGAGGACAAAAGTTCTT -5330
GR
Human VIPR2 Gene -5632 GATATTGTAAAAACTCTGGTTTTGTGTGTAAAAATAGTAACTGAGAACAA -5583
Marmoset VIPR2 Gene -5044 GATATTGTAAAAACTCCGGTTTTGTGTGTAAAAATAGTAACTGAGAACAA -4995
Chimp VIPR2 Gene -5644 GATATTGTAAAAACTCTGGTTTTGTGTGTAAAAATAGTAACTGAGAACAA -5595
Gorilla VIPR2 Gene -5329 GATATTGTAAAAACTCTGGTTTTGTGTGTAAAAATAGTAACTGAGAACAA -5280
GR GR
Human VIPR2 Gene -5582 TTAGGCATAGAAAATGCCTCACTTTCATGTGATGCTTTAATTACTCCCAG -5533
Marmoset VIPR2 Gene -4994 TTAGACATAGAAAATCCCTCACTTTCATGTGATGCTTTAATTACTCCCAG -4945
Chimp VIPR2 Gene -5594 TTAGGCATAGAAAATGCCTCACTTTCATGTGATGCTTTAATTACTCCCAG -5545
Gorilla VIPR2 Gene -5279 TTAGGCATAGAAAATGCCTCACTTTCATGTGATGCTTTAATTACTCCCAG -5230
GR AP1/HNF1
Human VIPR2 Gene -5532 GAGAACATTATTAATTATTTAAAGGACCACTTGAGTCATTAATTATCCTT -5483
Marmoset VIPR2 Gene -4944 AAGAACATTATTAATTATTTAAAGGGTCATGTGAGTCATTAATTATCCTT -4895
Chimp VIPR2 Gene -5544 GAGAACATTATTAATTATTTAAAGGGCCACTTGAGTCATTAATTATCCTT -5495
Gorilla VIPR2 Gene -5229 GAGAACATTATTAATTATTTAAAGGGCCACTTGAGTCATTAATTATCCTT -5180
HNF3/HNF3ALPHA
Human VIPR2 Gene -5382 CCCTCTTTTGTTCCTTTAATGTTGTTTGTTTTCTTCTCTTCTCCCTAAAA -5333
Marmoset VIPR2 Gene -4794 CCCTATTTTGTCCCTTTAATG---TTTGTTTTCTTCTCTTCTCCCTAAAA -4748
Chimp VIPR2 Gene -5394 CCCTCTTTTGTTCCTTTAATGTTGTTTGTTTTCTTCTCTTTTCCCTAAAA -5345
Gorilla VIPR2 Gene -5079 CCCTATTTTGTTCCTTTAATGTTGTTTGTTTTCTTCTCTTCTCCCTAAAA -5030
AP2
Human VIPR2 Gene -5286 CTCTGAAC-TCCAGGACCCTCTGGCCTGGAATCGCTGGCCTCCCGGCC-T -5239
Marmoset VIPR2 Gene -4697 CTCTGAAGATCCAGGACCCTCTGGCCTGGAATCATGGGCCTCTCAGCCCT -4648
Chimp VIPR2 Gene -5298 CTCTGAAG-TCCAGGACCCTCTGGCCTGGAATTGCTGGCCTCCCGGCC-T -5251
Gorilla VIPR2 Gene -4983 CTCTGAAG-TCCAGGACCCTCTGGCCTGGAATCGCTGGCCTCCCGGCC-T -4936
Human VIPR2 Gene -5238 GCAGGGTCCCCC-GTCACTCCCAGGACCCTCCTTTCACCCCCTCTCCCCT -5190
Marmoset VIPR2 Gene -4647 GCAGGGTCTCCCTGTCACTCCCAGGACCCTCCTTTCACCCTGTCTCCCCT -4598
Chimp VIPR2 Gene -5250 GCAGGGTCCCCC-GTCACTCCCAGGACCCTCCTTTCACCCCCTCTCCCCT -5202
Gorilla VIPR2 Gene -4935 GCAGGGTCCCCC-ATCACTCCCAGGACCCTCCTTTCACCCCCTCTCCCCT -4887
PPARA NFKAPPAB50
Human VIPR2 Gene -4940 AGTGAGGACGCCCTGGC-GAGGGGAAGGGTCTGGAGGGCAGCCCCACTGA -4892
Marmoset VIPR2 Gene -4349 AACCAGAAGGCCCTGGCAGAGGGGCAGGGTCTGGAGGGCAGCCCTGCTGA -4300
Chimp VIPR2 Gene -4952 AGTGAGGAGGCCCTGGC-GAGGGGAAGGGTCTGGAGGGCAGCCCCACTGA -4904
Gorilla VIPR2 Gene -4637 AGTGAGGAGGCCCTGCC-GAGGGGAAGGGTCTGGAGGGCAGCCCCGCTGA -4589
SREBP
Human VIPR2 Gene -4791 CCTCC-AGACTGGAGTGAGTGGAACCAGGAGTAACTCATACATCTGCCCA -4743
Marmoset VIPR2 Gene -4202 CCTCCCAGACTGGAGTGAGTGTAACCAGGAGTCACTCACACATCTGCCCA -4153
Chimp VIPR2 Gene -4803 CCTCC-AGACTGGAGTGAGTGGAACCAGGAGTAACTCATACATCTGCCCG -4755
Gorilla VIPR2 Gene -4488 CCTCC-AGACTGGAGTGAGTGGAACCAGGAGTAACTCATACATCTGCCCG -4440
STAT
Human VIPR2 Gene -4642 CTATCGCATCATATTATATTAAGCTAAGAATAGTTCTCCAAGAAAGTGGC -4593
Marmoset VIPR2 Gene -4053 ----------ATATTATATCAAGCCCTGAATGGTTCTCCAAGAAAATGGC -4013
Chimp VIPR2 Gene -4654 CTATCGCATCATATTATATTAAGCTAAGAATAGTTCTCCAAGAAAGTGGC -4605
Gorilla VIPR2 Gene -4339 CTATCACATCATATTATATTAAGCTAAGAATAGTTCTCCAAGAAAGTGGC -4290


PPARA
Human VIPR2 Gene -4592 AACACCTCCTTGTAAAGGGAAGTTTAAAGAACTCTGAGCCTGGTTGGTCT -4543
Marmoset VIPR2 Gene -4012 AACACCTCCTTGTAAAGGGAAGTTTAGAGAACTCTGAGTCTGGCTGGTCT -3963
Chimp VIPR2 Gene -4604 AACACCTCCTTGTAAAGGGAAGTTTAAAGAACTCTGAGCCTGGTTGGTCT -4555
Gorilla VIPR2 Gene -4289 AACACCTCCTTGTAAAGGGAAGTTTAAAGAACTCTGAGCCTGGTTGGTCT -4240
SREBP
Human VIPR2 Gene -4492 CACCTCCTGGAAGAACTGGTCAGAAGCAGAAGGGAGAGGCTGGTGGCCTG -4443
Marmoset VIPR2 Gene -3912 CCCCTCCTGGAAGAACTGGTCACGATCAGGAGGCAGAGGCTGGCAGCCTG -3863
Chimp VIPR2 Gene -4504 CACCTCCTGGAAGAACTGGTCAGAAGCAGAAGGGAGAGGCTGGTGGCCTG -4455
Gorilla VIPR2 Gene -4189 CACCTCCTGGAAGAACTGGTCAGAATCAGAAGGGAGAGGCTGGTGGCCTG -4140
E-BOX/E2A/MYOGNF1/AP2/SREBP
Human VIPR2 Gene -4442 TCTCTGTGAGCCCCACCTGCCTGGGGTGGGAGGGGGCCCGGGCAGGGCTA -4393
Marmoset VIPR2 Gene -3862 CCTCTGTGAGTCCCATCATCCTGGGGTGGGAGAGGGGCC---------- -3823
Chimp VIPR2 Gene -4454 TCTCTGTGAGCCCCACCTGCCTGGGGTGGGAGGGGGCCCG---------- -4414
Gorilla VIPR2 Gene -4139 TCTCTGTGAGCCCCACCTGCCTGGGGTGGGAGGGGGCCCAGGCAGGGCT -4090
AHRARNT
GR
Human VIPR2 Gene -4343 GGCTGCGGCACGCACTGGTGTGTGCCCGTCTCTTGGACAAACCATCTGAG -4294
Marmoset VIPR2 Gene -3811 --CTGTGCCATGTGCTGGTGTGTGGCCAGCTCTTGGACAAACCATCCAAG -3764
Chimp VIPR2 Gene -4407 GGCTGCGGCATGCACTGGTGTGTGCCCGTCTCTTGGACAAACCATCTGAG -4358
Gorilla VIPR2 Gene -4039 GGCTGCGGCACGCACTGGTGTGTGCCCGTCTCTTGGACAAACCATCTGAG -3990
GR
Human VIPR2 Gene -4293 ATGTTCTTCCAGCTGCTCTGCCTCCATCGCTGAGCCTCCTGCTGAGCTGA -4244
Marmoset VIPR2 Gene -3763 GTGTCCTTCCAGCAGCT---CCCAGACCCCTGAGGTTCCTCCCAAAATTT -3717
Chimp VIPR2 Gene -4357 ATGTTCTTCCAGCTGCTCTGCCTCCATCGCTGAGCCTCCTGCTGAGCTGA -4308
Gorilla VIPR2 Gene -3989 ATGTTCTTCCAG----TCTGCCCCCATCGCTGAGCCTCCTGCTGAGCTGA -3944
E-BOX
Human VIPR2 Gene -4144 GGGAGGGCTTTCCTAAGGCAGACACCTGAGCCAAGACAGGTGGAAGCTGG -4095
Marmoset VIPR2 Gene -3664 AATATTTAGGTTTTAAAAGATACAAAAAGAGGGACACAGGCAAATGATGA -3615
Chimp VIPR2 Gene -4208 GGGAGGGCTTTCCTAAGGCAGACACCTGAGCCAAGACAGGTGGAAGCTGG -4159
Gorilla VIPR2 Gene -3843 GCGAGGGCTTTCCTAAGGCAGACACCTGAGCCAAGACAGGTGGAAGCTGG -3794
SREBP
STAT
Human VIPR2 Gene -3945 TAGGGTGACCATGTGCTTCACATAATATGTCCAAACCAGGACATTTCCAA -3896
Marmoset VIPR2 Gene -3504 CAAAGTACTCATCTGG---ATGTGGCAGAGACA--CCAGAAGGATCTCAG -3460
Chimp VIPR2 Gene -4009 TAGGGTGACCATGTGCTTCACATAATATGTCCAAACCAGGACATTTCCAA -3960
Gorilla VIPR2 Gene -3660 TAGGGTGACCATGTGCTTCACATAATATGTCCAAACCAGGACATTTCCAA -3611
STAT
LXR/PPARA
Human VIPR2 Gene -3895 GAATAGGGGGACACGAGTAACGGTTACGCTGGGGCAATGGGCAAAGCCAG -3846
Marmoset VIPR2 Gene -3459 AGAGTCGTGTTTCTGGGAAGC---TTCGTTCCTCTTTTGG--AAGAACAT -3415
Chimp VIPR2 Gene -3959 GAATAGGGGGACACGAGTAACGGTTACGCTGGGGCAATGGGCAAAGCCAG -3910
Gorilla VIPR2 Gene -3610 GAATAGGGGGACACGAGTAACGGTTACGCTGGGGCAATGGGCAAAGCCAG -3561
NF1
Human VIPR2 Gene -3796 ATGTATTTGAGGCAGGTCTCAATCAACTTAGAGGTTGATTTTTGCCAAGG -3747
Marmoset VIPR2 Gene -3373 GTATCTTTCACTCTGTTCT-AATCA--TGAAGCCCCGAGGCTTCCCCTGA -3327
Chimp VIPR2 Gene -3860 ATGTATTTGAGGCAGGTCTCAATCAACTTAGAGGTTGATTTTTGCCAAGG -3811
Gorilla VIPR2 Gene -3511 ATGTATTTGAGGCAGGTCTCAATCAACTTAGAGGTTGATTTTTGCCAAGG -3462
GR
Human VIPR2 Gene -3746 TTAAGGACATGGC-CCAGGGCACAGCCTCAGGAGGTCCCAAGAACACACA -3698
Marmoset VIPR2 Gene -3326 CGTGTCATTGGGTTCCAGGCTTCTGCTTCATGAGG-----AAAACACACA -3282
Chimp VIPR2 Gene -3810 TTAAGGGCATGGC-CCAGGGCACTGCCTCAGGAGGTCCCAAGAACACACA -3762
Gorilla VIPR2 Gene -3461 TTAAGGACATGGC-CCAGGGCACAGCCTCAGGAGGTCCCAAGAACACGCA -3413
GR
Human VIPR2 Gene -3697 CCCAAGGTGACAGGGTATAGCTTGGTTTTATACATTTTAGGGAGACTGAA -3648
Marmoset VIPR2 Gene -3281 -----AGTAGCTG--CTCAGTTTCCGCCAGTGTCTGCCAGAGAT-TTGAT -3240
Chimp VIPR2 Gene -3761 CCCAAGGTGACAGGGTATAGCTTGGTTTTATACATTTTAGGGAGACTGAA -3712
Gorilla VIPR2 Gene -3412 CCCAAGGTGACAGGGTATAGCTTGGTTTTATACATTTTAGGGAGACTGAA -3363
PPARA
Human VIPR2 Gene -3597 CCCAGAAAGGTGGGACATCTTGAAGCGGGGTGTGGGGCCTTCCAGGTCAC -3548
Marmoset VIPR2 Gene -3199 CC----AGTGTGGGAAGC--AGCAGGGAAGTGGGTGACCTTGGAGGGCAT -3156
Chimp VIPR2 Gene -3661 CCCAGAAAGGTGGGACATCTCGAAGCGGGGTGTGGGGCCTTCCAGGTCAC -3612
Gorilla VIPR2 Gene -3312 CCCAGAAAAGTGGGACATCTCGAAGCGGGGTGTGGGGCCTTCCAGGTCAC -3263
CEBPB
Human VIPR2 Gene -3297 AACAGACAGCTTTGCTGGGCCATTTCAAAATCTGTCAAAGAAATATATTT -3248
Marmoset VIPR2 Gene -2921 ACCAGGA------GTCAGGCACCAGCAGAGGCCGGGAAAGGGAGTCTTCT -2878
Chimp VIPR2 Gene -3361 AACAGACAGCTTTGCTGGGCCATTTCAAAATCTGTCAAAGAAATATATTT -3312
Gorilla VIPR2 Gene -3012 AACAGACAGCTTTGCTGGGCCATTTCAAAATCTGTCAAAGAAATATATTT -2963
YY1
Human VIPR2 Gene -2898 AAGCCCGTGATTACACACACACTGGACCTTCAGAGAAGTGTGCCATCTAG -2849
Marmoset VIPR2 Gene -2581 CAATCACTGCTCTCA-----GCTCTTCTCCCCAAGA--TGTCCCATCTG- -2540
Chimp VIPR2 Gene -2962 AAGCCCGTGACTACACACACACTGGACCTTCAGAGAAGTGTGCCATCTAG -2913
Gorilla VIPR2 Gene -2613 AAGCCCGTGACTACACACACACTGGACCTTCAGAGAAGTGTGCCATCTAG -2564
YY1
Human VIPR2 Gene -2848 CTACGACCACAGCCATTTCTCCTGGTCTACCCCGTCCATCGGGACGCACT -2799
Marmoset VIPR2 Gene -2539 ---CGGGAAAGGCTCTGCACCCCACACTCGGCCACCCACC----CACACT -2497
Chimp VIPR2 Gene -2912 CTACGACCACAGCCATTTCTCCTGGTCTACCCCGTCCATCGGGACGCACT -2863
Gorilla VIPR2 Gene -2563 CTATGACCACAGCCATTTCTCCTGGTCTACCCCATCCATTGGGACGCACT -2514
CREBATF
FXR
Human VIPR2 Gene -2748 AGTCACTTCTAAAATGCAGGCCCTCAACAAGGTATTGACGTGTGCATTAA -2699
Marmoset VIPR2 Gene -2462 ACTCCCCTCACACACTCACACTCTT-------CCTCCACCTACTCACTC- -2421
Chimp VIPR2 Gene -2812 AGTCACTTCTAAAATGCAGGCCCTCAACAAGGTATTGACGTGTGCATTAA -2763
Gorilla VIPR2 Gene -2481 AGTCACTTCTAAAATGCAGGCCCTCAACAAGGTATTGACGTGTGCATTAA -2432


FXR
Human VIPR2 Gene -2698 TCACCTTGATTATTACAGAGCCCCTCTGCGGTGGAAGGGCAGGTGCTGTC -2649
Marmoset VIPR2 Gene -2420 TCCCTCTGCCTACT-CACACTCTCCCCTC---------CACCCACTCTC -2382
Chimp VIPR2 Gene -2762 TTACCTTGATTATTACAGAGCCCCTCTGCGGTGGAAGGGCAGGTGCTGTC -2713
Gorilla VIPR2 Gene -2431 TTACCTTGATTCTTACAGAGCCCCTCTGCGGTGGAAGGGCAGGTGCTGTC -2382
COUP/DR1/HNF4
Human VIPR2 Gene -2648 ATTCCTA-CAAATGCATAG--AGCTCAAAGTTCATCGCTCACCCAGG-TG -2603
Marmoset VIPR2 Gene -2381 ACTCCCCTCCAATGCACAAGTAACTCCAAGTTTATGGCTTTCCCAGGGTG -2332
Chimp VIPR2 Gene -2712 ATTCCTA-CAAATGCATAG--AGCTCAAAGTTCATCGCTCGCCCAGG-TG -2667
Gorilla VIPR2 Gene -2381 ATTCCTA-CAAATGCATAG--AGCTCAAAGTTCATCGCTCGCCCAGG-TG -2336
GATA 4
Human VIPR2 Gene -2503 -GTCTCCCACTAGATACACATGGATGAAGCCAGATAGATAACAGGTAGAG -2455
Marmoset VIPR2 Gene -2231 AGTTTCCCATTAGATACACATGGATGAAGCCAGAGAGATAACAGGTAGCA -2182
Chimp VIPR2 Gene -2566 -GTCTCCCACTAGATACACATGGATGAAGCCAGATAGATAACAGGTAGAG -2518
Gorilla VIPR2 Gene -2235 -GTCTCCCAATAGATACACATGGATGAAGCCAGATAGATAACAGGTAGAG -2187
GATA 4
Human VIPR2 Gene -2454 GATAGATGACAGGCAGATGACAGATAGGCAGATGATAGATGATCGATAGA -2405
Marmoset VIPR2 Gene -2181 CATAAATGACAGGCAGATGACAGAT---------- -2156
Chimp VIPR2 Gene -2517 GATAGATGACAGGCAGATGACAGATAGGCAGATGATAGATGATC------ -2473
Gorilla VIPR2 Gene -2186 GATAGATGACAGGCAGATGACAGATAGGCAGATGACAGATGATC------ -2142
GATA 4
Human VIPR2 Gene -2404 TGACAGGCAGATGATGGATGACAGGCAGATCACAGATATGATAGATCACA -2355
Marmoset VIPR2 Gene -1555 ------------GATTGACGACAGGCAGATCACAGATATGATAGATCCCA -1517
Chimp VIPR2 Gene -2472 ------------GATAGATGACAGGCAGATCACAGATATGATAGATCACA -2434
Gorilla VIPR2 Gene -2141 ------------AATAGATGACAGGCAGATCACAGATATGATAGATCACA -2103
GATA 4
Human VIPR2 Gene -2354 GATAGATGACAAATAGGCAGATGATAGATGACAGGCAGATGACAGATAGG -2305
Marmoset VIPR2 Gene -1516 GAAAAATGACAGATAGACAGATGATAGATCACAGACC------------- -1479
Chimp VIPR2 Gene -2433 GATAGATGACAAATAGGCAGATGATAGATGACAGGCAGATGACAGGCAGA -2384
Gorilla VIPR2 Gene -2102 GATAGATGACAAATAGGCAGATGATACACGACAGGCAGATGACAGGCAGA -2053
Human VIPR2 Gene -1241 CGGGGGGTGGGGGATCAGAGAGAAGCAAATGATAAGCCCAATGAGGTAAA -1192
Marmoset VIPR2 Gene -1089 CGGGGG-------ATCAGAGAGAAGCAAA---TAAGCCCAACAAGGTCAG -1050
Chimp VIPR2 Gene -1243 CGGGGGGTGGCGGATCAGAGAGAAGCAAATGATAAGCCCAATGAGGTAAA -1194
Gorilla VIPR2 Gene -1238 CGGGGGGTGGGGGATCAGAGAGAAGCAAATGATAAGCCCAATGAGGTAAA -1189
FXR
Human VIPR2 Gene -1191 TGTTAATGACCGAATCCACAAAAAGGATAAAAAGGAGTTATTTGTAATAT -1142
Marmoset VIPR2 Gene -1049 TGTTAATGACTGAATCCAGATGAAGGGTAAAAAG-AGTTCTTTGTACTAT -1001
Chimp VIPR2 Gene -1193 TGTTAATGACCGAATCCACAAAAAGGATAAAAAGGAGTTATTTGTAATAT -1144
Gorilla VIPR2 Gene -1188 TGTTAATGACCGAATCCACAAAAAGGATAAAAAGGAGTTACTGGTAATAT -1139
CEBPGAMMA
Human VIPR2 Gene -1141 TCTTATTCTTCTAATTTTTGTGTAAGTTTGAAATTATTTCCAAATAAAAA -1092
Marmoset VIPR2 Gene -1000 TCTTATTCTTCTAATTTTTGTGTACGTTTGAAATTATTTCCAAATTTAAA -0951
Chimp VIPR2 Gene -1143 TCTTATTCTTCTAATTTTTGTGTAAGTTTGAAATTATTTCCAAATAAAAA -1094
Gorilla VIPR2 Gene -1138 TCTTATTCTTCTAATTTTTGTGTAAGTCTGAAATTATTTCCAAATAAAAA -1089
GATA 4
Human VIPR2 Gene -1091 ATTAAAACA-ACACGAACTTAACTTGTATCTCTCATT-CCAGTGCTTTCC -1044
Marmoset VIPR2 Gene -0950 AAAAAAACACACACAAAATTAGCTTTCGTCTCTCATTTCCAGTGCCTTCC -0901
Chimp VIPR2 Gene -1093 ATTAAAACA-ACACAAACTTAACTTGTATCTCTCATT-CCAGTGCTTTCC -1046
Gorilla VIPR2 Gene -1088 A-----ACA-ACACAAACTTAACTTGTATCTCTCATT-CCAGTGCTTTCC -1046
AHRARNT
Human VIPR2 Gene -1043 ACTTGCTGGGGAACGCCGAGCTCTCCTGGGTTGGTCACGCGGGCGCCTTG -0994
Marmoset VIPR2 Gene -0900 CCTTGCTGGGTGACGCCT-GCTCTCCTGGGCTGGCCCTGCGGGGCTCCTCG -0852
Chimp VIPR2 Gene -1045 ACTTGCTGGGGAACGCCGAGCTCTCCTGGGTTGGTCACGCGGGCGCCTTG -0996
Gorilla VIPR2 Gene -1045 ACTTGCTGGGGAACGCCGAGCTCTCCTGGGTTGGTCACGCGGGCGCCTTG -0996
AHRARNT
Human VIPR2 Gene -0993 GGAGGCCGCGCAGTCCCGCGTGGGGCGCGGCGGGGCGGCCCCAGTGGCAC -0944
Marmoset VIPR2 Gene -0851 GGAGGCGGCGCAGTCCTGCGTGGGGCA-GGCGGG-CGGCCCCAGTGGCCC -0804
Chimp VIPR2 Gene -0995 GGAGGCCGCGCAGTCCCGCGTGGGGCGCGGCGGGGCGGCCCCAGTGGCAC -0946
Gorilla VIPR2 Gene -0995 GGAGGCCGCGCAGTCCCGCGTGGGGCGCGGCGGGGCGGCCCCAGTGGCAC -0946
AHRARNT SP1
Human VIPR2 Gene -0943 CGCGTGAGTCCCCGCCCAGCGTTCCCCACCCGCCGCCGCGTTTGCGGGGA -0894
Marmoset VIPR2 Gene -0803 CGCCTGAGTCCC--------------------GCCGCCGCGTGTGCGGGGA -0773
Chimp VIPR2 Gene -0945 CGCGTGAGTCCCCGCCCAGCGCTCCCCACCCGCCGCCGCGTTTGCGGGGA -0896
Gorilla VIPR2 Gene -0945 CGCGTGAGTCCCCGCCCAGCGCTCCCCACCCGCCGCCGCGTTTGCGGGGA -0896
CEBP
Human VIPR2 Gene -0893 GAGAATGAC-CCCCGTTTTGCAAACGCAGGACACAAAACCCGCCACCCAG -0845
Marmoset VIPR2 Gene -0772 GAGGACGACTCCCCGCGCTGCAAACGCAGGACACAAAACCAGCAGCCGAG -0723
Chimp VIPR2 Gene -0895 GAGAATGAC-CCCCGTTTTGCAAACGCAGGACACAAAACCAGCCACCCAG -0846
Gorilla VIPR2 Gene -0895 GAGAATGAC-CCCCGTTTTGCCACCGCACGACGCGAAGCCCGCCACCCAG -0846
NFKAPPAB50
Human VIPR2 Gene -0794 TGTTCGACCTGATGTCCTCT-TTGAGGCTG--CGGGGATCCCCACCAAAA -0748
Marmoset VIPR2 Gene -0673 TGTCCAACCTGACTTGGTCTGTTGAGGCTGTTCCTGGGTCCCCACCGAAA -0624
Chimp VIPR2 Gene -0797 TGTTGAACCTGATGTCCTCTTTTGAGGCTG--CGGGGATCCCCACCAAAA -0750
Gorilla VIPR2 Gene -0797 TGTTCAACCTGATGTCCTCTTTTGAGGCTG--CGGGGGTCCCCACCAAAA -0750
YY1
Human VIPR2 Gene -0747 AGCACTCTGATTTTCTCCATTTTCAGATGCCCCAACCTAGCCCCACTGGC -0698
Marmoset VIPR2 Gene -0623 CGCACCCTAATTTTCTCAATTTTCAGACGTCCCAACCCTGCCCCCACCCGC -0574
Chimp VIPR2 Gene -0749 AGCACTCTAATTTTCTCCATTTTCAGATGCCCCAACCTAGCCCCACTGGC -0700
Gorilla VIPR2 Gene -0749 AGCACTCTAATTTTCTCCATTTTCAGATGCCCCAACCTAGCCCCACTGGC -0700

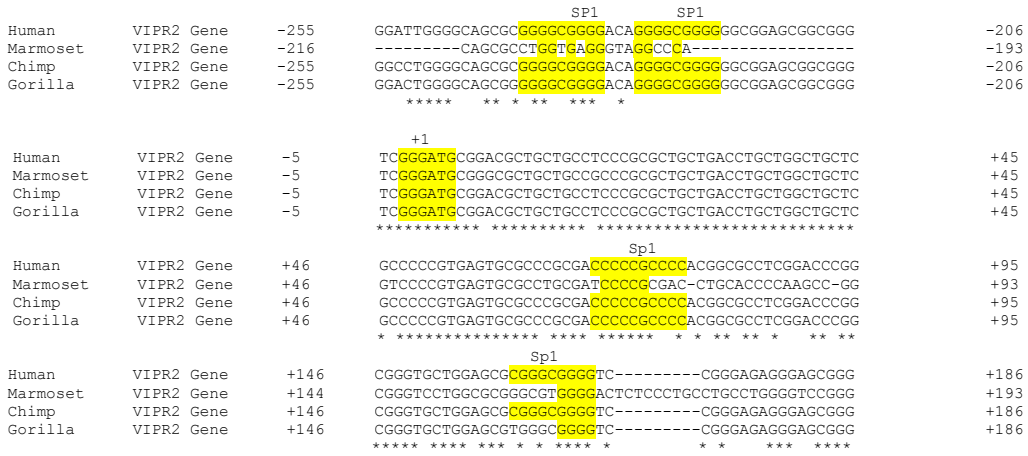
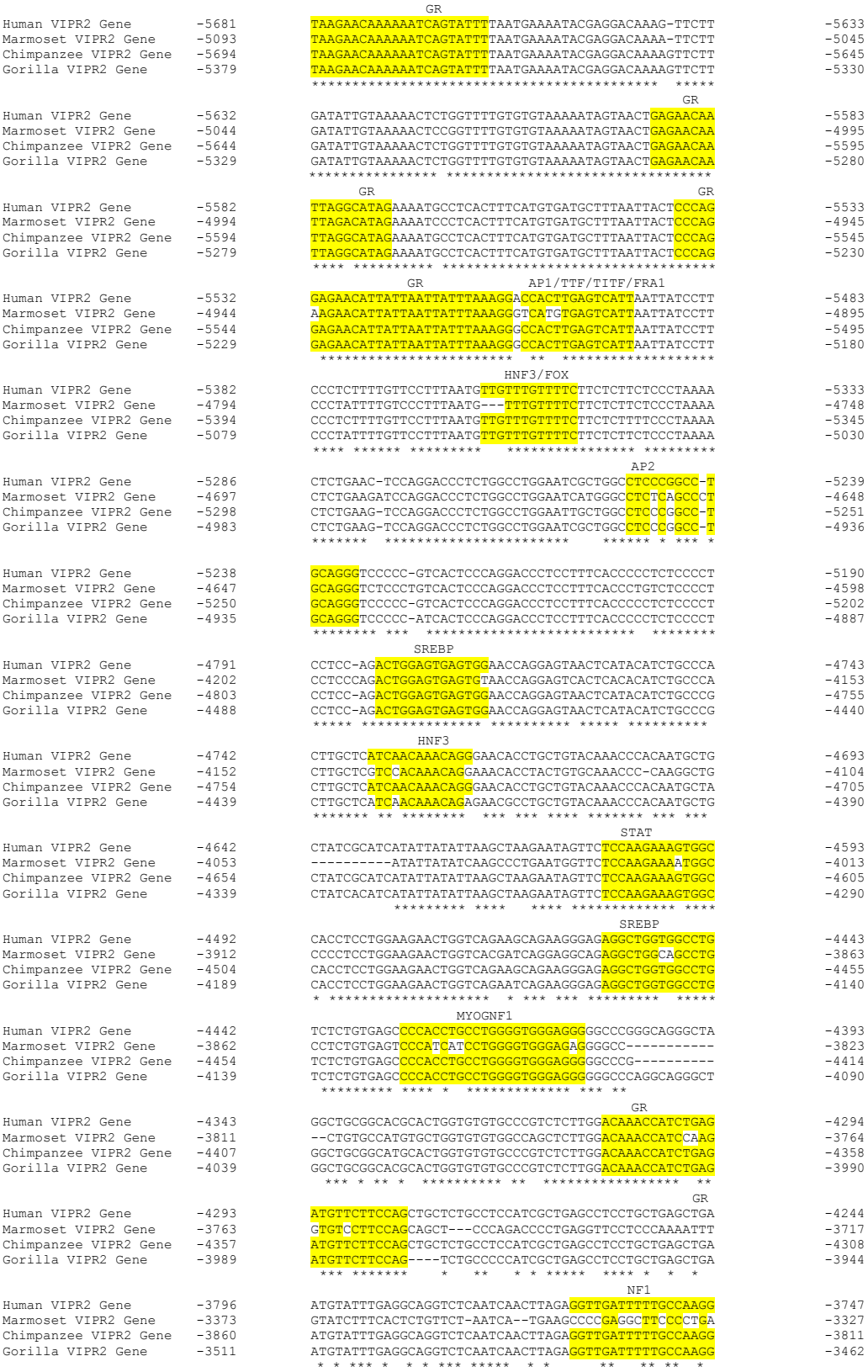
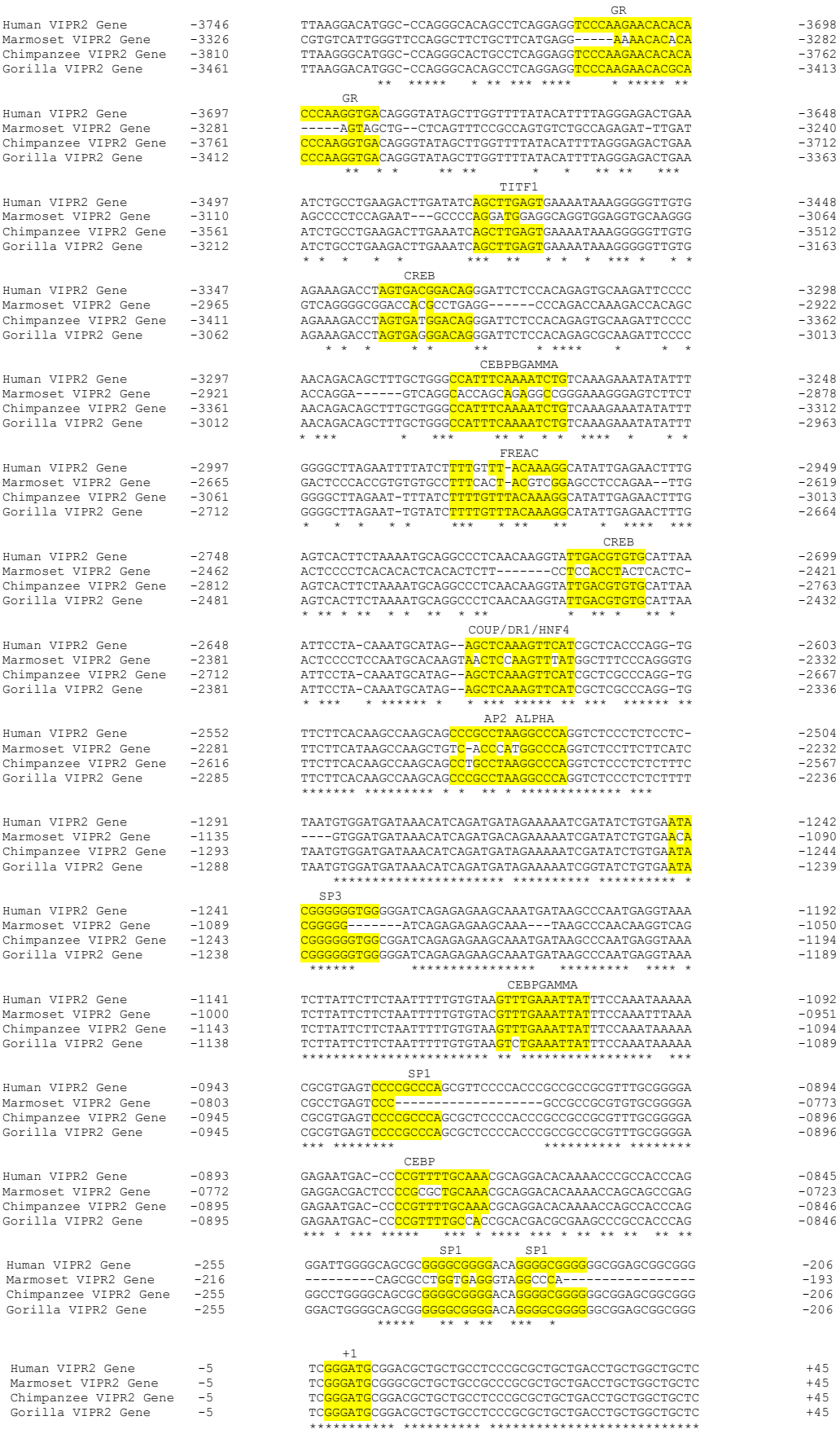


Figure 3.13 Evolutionarily conserved elements important in Hepatocytes.


GR
Human VIPR2 Gene -5681 TAAGAACAAAAAATCAGTATTTTAATGAAAATACGAGGACAAAG-TTCTT -5633
Marmoset VIPR2 Gene -5093 TAAGAACAAAAAATCAGTATTTTAATGAAAATACGAGGACAAAA-TTCTT -5045
Chimpanzee VIPR2 Gene -5694 TAAGAACAAAAAATCAGTATTTTAATGAAAATACGAGGACAAAAGTTCTT -5645
Gorilla VIPR2 Gene -5379 TAAGAACAAAAAATCAGTATTTTAATGAAAATACGAGGACAAAAGTTCTT -5330
GR
Human VIPR2 Gene -5632 GATATTGTAAAAACTCTGGTTTTGTGTGTAAAAATAGTAACTGAGAACAA -5583
Marmoset VIPR2 Gene -5044 GATATTGTAAAAACTCCGGTTTTGTGTGTAAAAATAGTAACTGAGAACAA -4995
Chimpanzee VIPR2 Gene -5644 GATATTGTAAAAACTCTGGTTTTGTGTGTAAAAATAGTAACTGAGAACAA -5595
Gorilla VIPR2 Gene -5329 GATATTGTAAAAACTCTGGTTTTGTGTGTAAAAATAGTAACTGAGAACAA -5280
GR GR
Human VIPR2 Gene -5582 TTAGGCATAGAAAATGCCTCACTTTCATGTGATGCTTTAATTACTCCCAG -5533
Marmoset VIPR2 Gene -4994 TTAGACATAGAAAATCCCTCACTTTCATGTGATGCTTTAATTACTCCCAG -4945
Chimpanzee VIPR2 Gene -5594 TTAGGCATAGAAAATGCCTCACTTTCATGTGATGCTTTAATTACTCCCAG -5545
Gorilla VIPR2 Gene -5279 TTAGGCATAGAAAATGCCTCACTTTCATGTGATGCTTTAATTACTCCCAG -5230
GR AP1/TTF/TITF/FRA1
Human VIPR2 Gene -5532 GAGAACATTATTAATTATTTAAAGGACCACTTGAGTCATTAATTATCCTT -5483
Marmoset VIPR2 Gene -4944 AAGAACATTATTAATTATTTAAAGGGTCATGTGAGTCATTAATTATCCTT -4895
Chimpanzee VIPR2 Gene -5544 GAGAACATTATTAATTATTTAAAGGCCACTTGAGTCATTAATTATCCTT -5495
Gorilla VIPR2 Gene -5229 GAGAACATTATTAATTATTTAAAGGGCCACTTGAGTCATTAATTATCCTT -5180
HNF3/FOX
Human VIPR2 Gene -5382 CCCTCTTTTGTTCCTTTAATGTTGTTTGTTTTCTTCTCTTCTCCCTAAAA -5333
Marmoset VIPR2 Gene -4794 CCCTATTTTGTCCCTTTAATG---TTTGTTTTCTTCTCTTCTCCCTAAAA -4748
Chimpanzee VIPR2 Gene -5394 CCCTCTTTTGTTCCTTTAATGTTGTTTGTTTTCTTCTCTTTCCCTAAAA -5345
Gorilla VIPR2 Gene -5079 CCCTATTTTGTTCCTTTAATGTTGTTTGTTTTCTTCTCTTCTCCCTAAAA -5030
AP2
Human VIPR2 Gene -5286 CTCTGAAC-TCCAGGACCCTCTGGCCTGGAATCGCTGGCCTCCCGGCC-T -5239
Marmoset VIPR2 Gene -4697 CTCTGAAGATCCAGGACCCTCTGGCCTGGAATCATGGGCCTCTCAGCCCT -4648
Chimpanzee VIPR2 Gene -5298 CTCTGAAG-TCCAGGACCCTCTGGCCTGGAATTGCTGGCCTCCCGGCC-T -5251
Gorilla VIPR2 Gene -4983 CTCTGAAG-TCCAGGACCCTCTGGCCTGGAATCGCTGGCCTCCCGGCC-T -4936
Human VIPR2 Gene -5238 GCAGGGTCCCCC-GTCACTCCCAGGACCCTCCTTTCACCCCCTCTCCCCT -5190
Marmoset VIPR2 Gene -4647 GCAGGGTCTCCCTGTCACTCCCAGGACCCTCCTTTCACCCTGTCTCCCCT -4598
Chimpanzee VIPR2 Gene -5250 GCAGGGTCCCCC-GTCACTCCCAGGACCCTCCTTTCACCCCCTCTCCCCT -5202
Gorilla VIPR2 Gene -4935 GCAGGGTCCCCC-ATCACTCCCAGGACCCTCCTTTCACCCCCTCTCCCCT -4887
SREBP
Human VIPR2 Gene -4791 CCTCC-AGACTGGAGTGAGTGGAACCAGGAGTAACTCATACATCTGCCCA -4743
Marmoset VIPR2 Gene -4202 CCTCCCAGACTGGAGTGAGTGTAACCAGGAGTCACTCACACATCTGCCCA -4153
Chimpanzee VIPR2 Gene -4803 CCTCC-AGACTGGAGTGAGTGGAACCAGGAGTAACTCATACATCTGCCCG -4755
Gorilla VIPR2 Gene -4488 CCTCC-AGACTGGAGTGAGTGGAACCAGGAGTAACTCATACATCTGCCCG -4440
HNF3
Human VIPR2 Gene -4742 CTTGCTCATCAACAAACAGGGAACACCTGCTGTACAAACCCACAATGCTG -4693
Marmoset VIPR2 Gene -4152 CTTGCTCGTCCACAAACAGGAAACACCTACTGTGCAAACCC-CAAGGCTG -4104
Chimpanzee VIPR2 Gene -4754 CTTGCTCATCAACAAACAGGGAACACCTGCTGTACAAACCCACAATGCTA -4705
Gorilla VIPR2 Gene -4439 CTTGCTCATCAACAAACAGAGAACGCCTGCTGTACAAACCCACAATGCTG -4390
STAT
Human VIPR2 Gene -4642 CTATCGCATCATATTATATTAAGCTAAGAATAGTTCTCCAAGAAAGTGGC -4593
Marmoset VIPR2 Gene -4053 ----------ATATTATATCAAGCCCTGAATGGTTCTCCAAGAAAATGGC -4013
Chimpanzee VIPR2 Gene -4654 CTATCGCATCATATTATATTAAGCTAAGAATAGTTCTCCAAGAAAGTGGC -4605
Gorilla VIPR2 Gene -4339 CTATCACATCATATTATATTAAGCTAAGAATAGTTCTCCAAGAAAGTGGC -4290
SREBP
Human VIPR2 Gene -4492 CACCTCCTGGAAGAACTGGTCAGAAGCAGAAGGGAGAGGCTGGTGGCCTG -4443
Marmoset VIPR2 Gene -3912 CCCCTCCTGGAAGAACTGGTCACGATCAGGAGGCAGAGGCTGGCAGCCTG -3863
Chimpanzee VIPR2 Gene -4504 CACCTCCTGGAAGAACTGGTCAGAAGCAGAAGGGAGAGGCTGGTGGCCTG -4455
Gorilla VIPR2 Gene -4189 CACCTCCTGGAAGAACTGGTCAGAATCAGAAGGGAGAGGCTGGTGGCCTG -4140
MYOGNF1
Human VIPR2 Gene -4442 TCTCTGTGAGCCCCACCTGCCTGGGGTGGGAGGGGGCCCGGGCAGGGCTA -4393
Marmoset VIPR2 Gene -3862 CCTCTGTGAGTCCCATCATCCTGGGGTGGGAGAGGGGCC----------- -3823
Chimpanzee VIPR2 Gene -4454 TCTCTGTGAGCCCCACCTGCCTGGGGTGGGAGGGGGCCCG---------- -4414
Gorilla VIPR2 Gene -4139 TCTCTGTGAGCCCCACCTGCCTGGGGTGGGAGGGGGCCCAGGCAGGGCT -4090
GR
Human VIPR2 Gene -4343 GGCTGCGGCACGCACTGGTGTGTGCCCGTCTCTTGGACAAACCATCTGAG -4294
Marmoset VIPR2 Gene -3811 --CTGTGCCATGTGCTGGTGTGTGGCCAGCTCTTGGACAAACCATCCAAG -3764
Chimpanzee VIPR2 Gene -4407 GGCTGCGGCATGCACTGGTGTGTGCCCGTCTCTTGGACAAACCATCTGAG -4358
Gorilla VIPR2 Gene -4039 GGCTGCGGCACGCACTGGTGTGTGCCCGTCTCTTGGACAAACCATCTGAG -3990
GR
Human VIPR2 Gene -4293 ATGTTCTTCCAGCTGCTCTGCCTCCATCGCTGAGCCTCCTGCTGAGCTGA -4244
Marmoset VIPR2 Gene -3763 GTGTCCTTCCAGCAGCT---CCCAGACCCCTGAGGTTCCTCCCAAAATTT -3717
Chimpanzee VIPR2 Gene -4357 ATGTTCTTCCAGCTGCTCTGCCTCCATCGCTGAGCCTCCTGCTGAGCTGA -4308
Gorilla VIPR2 Gene -3989 ATGTTCTTCCAG----TCTGCCCCCATCGCTGAGCCTCCTGCTGAGCTGA -3944
NF1
Human VIPR2 Gene -3796 ATGTATTTGAGGCAGGTCTCAATCAACTTAGAGGTTGATTTTTGCCAAGG -3747
Marmoset VIPR2 Gene -3373 GTATCTTTCACTCTGTTCT-AATCA--TGAAGCCCCGAGGCTTCCCCTGA -3327
Chimpanzee VIPR2 Gene -3860 ATGTATTTGAGGCAGGTCTCAATCAACTTAGAGGTTGATTTTTGCCAAGG -3811
Gorilla VIPR2 Gene -3511 ATGTATTTGAGGCAGGTCTCAATCAACTTAGAGGTTGATTTTTGCCAAGG -3462


GR
Human VIPR2 Gene -3746 TTAAGGACATGGC-CCAGGGCACAGCCTCAGGAGGTCCCAAGAACACACA -3698
Marmoset VIPR2 Gene -3326 CGTGTCATTGGGTTCCAGGCTTCTGCTTCATGAGG-----AAAACACA -3282
Chimpanzee VIPR2 Gene -3810 TTAAGGGCATGGC-CCAGGGCACTGCCTCAGGAGGTCCCAAGAACACACA -3762
Gorilla VIPR2 Gene -3461 TTAAGGACATGGC-CCAGGGCACAGCCTCAGGAGGTCCCAAGAACACGCA -3413
GR
Human VIPR2 Gene -3697 CCCAAGGTGACAGGGTATAGCTTGGTTTTATACATTTTAGGGAGACTGAA -3648
Marmoset VIPR2 Gene -3281 -----AGTAGCTG--CTCAGTTTCCGCCAGTGTCTGCCAGAGAT-TTGAT -3240
Chimpanzee VIPR2 Gene -3761 CCCAAGGTGACAGGGTATAGCTTGGTTTTATACATTTTAGGGAGACTGAA -3712
Gorilla VIPR2 Gene -3412 CCCAAGGTGACAGGGTATAGCTTGGTTTTATACATTTTAGGGAGACTGAA -3363
TITF1
Human VIPR2 Gene -3497 ATCTGCCTGAAGACTTGATATCAGCTTGAGTGAAAATAAAGGGGGTTGTG -3448
Marmoset VIPR2 Gene -3110 AGCCCCTCCAGAAT---GCCCCAGGATGGAGGCAGGTGGAGGTGCAAGGG -3064
Chimpanzee VIPR2 Gene -3561 ATCTGCCTGAAGACTTGAAATCAGCTTGAGTGAAAATAAAGGGGGTTGTG -3512
Gorilla VIPR2 Gene -3212 ATCTGCCTGAAGACTTGAAATCAGCTTGAGTGAAAATAAAGGGGGTTGTG -3163
CREB
Human VIPR2 Gene -3347 AGAAAGACCTAGTGACGGACAGGGATTCTCCACAGAGTGCAAGATTCCCC -3298
Marmoset VIPR2 Gene -2965 GTCAGGGGCGGACCACGCCTGAGG------CCCAGACCAAAGACCACAGC -2922
Chimpanzee VIPR2 Gene -3411 AGAAAGACCTAGTGATGGACAGGGATTCTCCACAGAGTGCAAGATTCCCC -3362
Gorilla VIPR2 Gene -3062 AGAAAGACCTAGTGAGGGACAGGGATTCTCCACAGAGCGCAAGATTCCCC -3013
CEBPBGAMMA
Human VIPR2 Gene -3297 AACAGACAGCTTTGCTGGGCCATTTCAAAATCTGTCAAAGAAATATATTT -3248
Marmoset VIPR2 Gene -2921 ACCAGGA------GTCAGGCACCAGCAGAGGCCGGGAAAGGGAGTCTTCT -2878
Chimpanzee VIPR2 Gene -3361 AACAGACAGCTTTGCTGGGCCATTTCAAAATCTGTCAAAGAAATATATTT -3312
Gorilla VIPR2 Gene -3012 AACAGACAGCTTTGCTGGGCCATTTCAAAATCTGTCAAAGAAATATATTT -2963
FREAC
Human VIPR2 Gene -2997 GGGGCTTAGAATTTTATCTTTGTTT-ACAAAGGCATATTGAGAACTTTG -2949
Marmoset VIPR2 Gene -2665 GACTCCCACCGTGTGTGCCTTTCACT-ACGTCGGAGCCTCCAGAA--TTG -2619
Chimpanzee VIPR2 Gene -3061 GGGGCTTAGAAT-TTTATCTTTTGTTTACAAAGGCATATTGAGAACTTTG -3013
Gorilla VIPR2 Gene -2712 GGGGCTTAGAAT-TGTATCTTTTGTTTACAAAGGCATATTGAGAACTTTG -2664
CREB
Human VIPR2 Gene -2748 AGTCACTTCTAAAATGCAGGCCCTCAACAAGGTATTGACGTGTGCATTAA -2699
Marmoset VIPR2 Gene -2462 ACTCCCCTCACACACTCACACTCTT-------CCTCCACCTACTCACTC- -2421
Chimpanzee VIPR2 Gene -2812 AGTCACTTCTAAAATGCAGGCCCTCAACAAGGTATTGACGTGTGCATTAA -2763
Gorilla VIPR2 Gene -2481 AGTCACTTCTAAAATGCAGGCCCTCAACAAGGTATTGACGTGTGCATTAA -2432
COUP/DR1/HNF4
Human VIPR2 Gene -2648 ATTCCTA-CAAATGCATAG--AGCTCAAAGTTCATCGCTCACCCAGG-TG -2603
Marmoset VIPR2 Gene -2381 ACTCCCCTCCAATGCACAAGTAACTCCAAGTTTATGGCTTTCCCAGGGTG -2332
Chimpanzee VIPR2 Gene -2712 ATTCCTA-CAAATGCATAG--AGCTCAAAGTTCATCGCTCGCCCAGG-TG -2667
Gorilla VIPR2 Gene -2381 ATTCCTA-CAAATGCATAG--AGCTCAAAGTTCATCGCTCGCCCAGG-TG -2336
AP2 ALPHA
Human VIPR2 Gene -2552 TTCTTCACAAGCCAAGCAGCCCGCCTAAGGCCCAGGTCTCCCTCTCCTC- -2504
Marmoset VIPR2 Gene -2281 TTCTTCATAAGCCAAGCTGTC-ACCCATGGCCCAGGTCTCCTTCTTCATC -2232
Chimpanzee VIPR2 Gene -2616 TTCTTCACAAGCCAAGCAGCCTGCCTAAGGCCCAGGTCTCCCTCTCTTTC -2567
Gorilla VIPR2 Gene -2285 TTCTTCACAAGCCAAGCAGCCCGCCTAAGGCCCAGGTCTCCCTCTCTTTT -2236
Human VIPR2 Gene -1291 TAATGTGGATGATAAACATCAGATGATAGAAAAATCGATATCTGTGAATA -1242
Marmoset VIPR2 Gene -1135 ----GTGGATGATAAACATCAGATGACAGAAAAATCGATATCTGTGAACA -1090
Chimpanzee VIPR2 Gene -1293 TAATGTGGATGATAAACATCAGATGATAGAAAAATCGATATCTGTGAATA -1244
Gorilla VIPR2 Gene -1288 TAATGTGGATGATAAACATCAGATGATAGAAAAATCGGTATCTGTGAATA -1239
SP3
Human VIPR2 Gene -1241 CGGGGGGTGGGGGATCAGAGAGAAGCAAATGATAAGCCCAATGAGGTAAA -1192
Marmoset VIPR2 Gene -1089 CGGGGG-------ATCAGAGAGAAGCAAA---TAAGCCCAACAAGGTCAG -1050
Chimpanzee VIPR2 Gene -1243 CGGGGGGTGGCGGATCAGAGAGAAGCAAATGATAAGCCCAATGAGGTAAA -1194
Gorilla VIPR2 Gene -1238 CGGGGGGTGGGGGATCAGAGAGAAGCAAATGATAAGCCCAATGAGGTAAA -1189
CEBPGAMMA
Human VIPR2 Gene -1141 TCTTATTCTTCTAATTTTTGTGTAAGTTTGAAATTATTTCCAAATAAAAA -1092
Marmoset VIPR2 Gene -1000 TCTTATTCTTCTAATTTTTGTGTACGTTTGAAATTATTTCCAAATTTAAA -0951
Chimpanzee VIPR2 Gene -1143 TCTTATTCTTCTAATTTTTGTGTAAGTTTGAAATTATTTCCAAATAAAAA -1094
Gorilla VIPR2 Gene -1138 TCTTATTCTTCTAATTTTTGTGTAAGTCTGAAATTATTTCCAAATAAAAA -1089
SP1
Human VIPR2 Gene -0943 CGCGTGAGTCCCCGCCCAGCGTTCCCCACCCGCCGCCGCGTTTGCGGGGA -0894
Marmoset VIPR2 Gene -0803 CGCCTGAGTCCC-------------------GCCGCCGCGTGTGCGGGGA -0773
Chimpanzee VIPR2 Gene -0945 CGCGTGAGTCCCCGCCCAGCGCTCCCCACCCGCCGCCGCGTTTGCGGGGA -0896
Gorilla VIPR2 Gene -0945 CGCGTGAGTCCCCGCCCAGCGCTCCCCACCCGCCGCCGCGTTTGCGGGGA -0896
CEBP
Human VIPR2 Gene -0893 GAGAATGAC-CCCCGTTTTGCAAACGCAGGACACAAAACCCGCCACCCAG -0845
Marmoset VIPR2 Gene -0772 GAGGACGACTCCCCGCGCTGCAAACGCAGGACACAAAACCAGCAGCCGAG -0723
Chimpanzee VIPR2 Gene -0895 GAGAATGAC-CCCCGTTTTGCAAACGCAGGACACAAAACCAGCCACCCAG -0846
Gorilla VIPR2 Gene -0895 GAGAATGAC-CCCCGTTTTGCCACCGCACGACGCGAAGCCCGCCACCCAG -0846
SP1
SP1
Human VIPR2 Gene -255 GGATTGGGGCAGCGCGGGGCGGGGACAGGGGCGGGGGGCGGAGCGGCGGG -206
Marmoset VIPR2 Gene -216 ---------CAGCGCCTGGTGAGGGTAGGCCCA---------- -193
Chimpanzee VIPR2 Gene -255 GGCCTGGGGCAGCGCGGGGCGGGGACAGGGGCGGGGGGCGGAGCGGCGGG -206
Gorilla VIPR2 Gene -255 GGACTGGGGCAGCGGGGGCGGGGACAGGGGCGGGGGGCGGAGCGGCGGG -206
+1
Human VIPR2 Gene -5 TCGGGATGCGGACGCTGCTGCCTCCCGCGCTGCTGACCTGCTGGCTGCTC +45
Marmoset VIPR2 Gene -5 TCGGGATGCGGGCGCTGCTGCCGCCCGCGCTGCTGACCTGCTGGCTGCTC +45
Chimpanzee VIPR2 Gene -5 TCGGGATGCGGACGCTGCTGCCTCCCGCGCTGCTGACCTGCTGGCTGCTC +45
Gorilla VIPR2 Gene -5 TCGGGATGCGGACGCTGCTGCCTCCCGCGCTGCTGACCTGCTGGCTGCTC +45

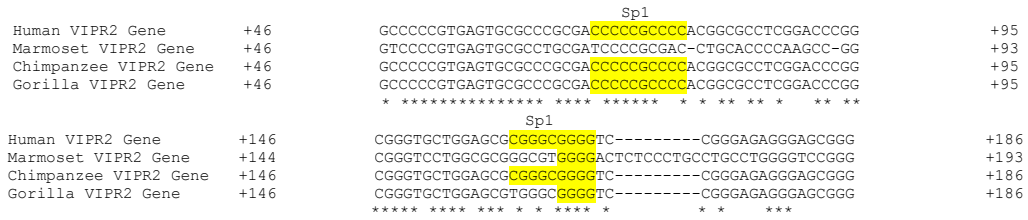
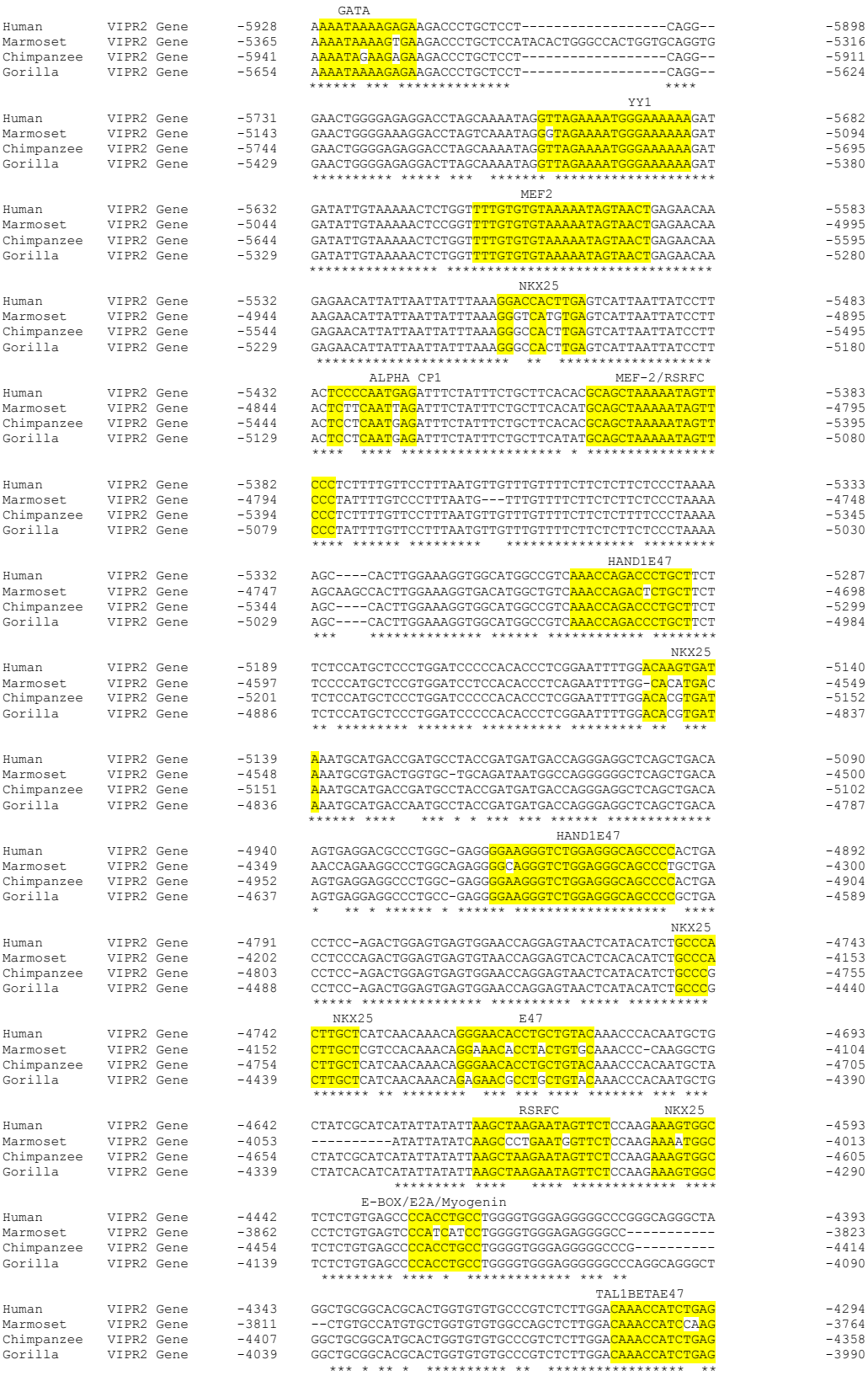
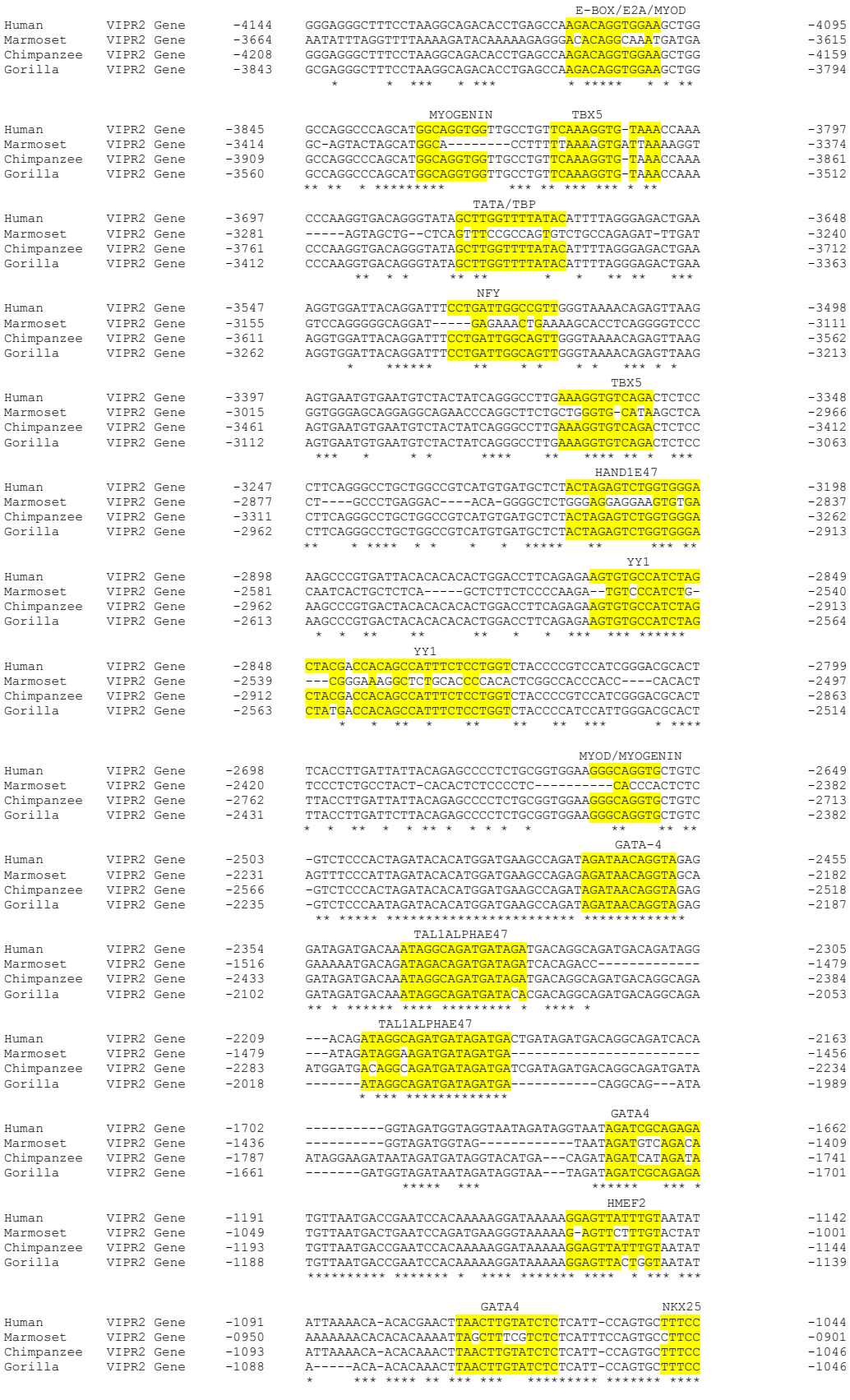


Figure 3.14 Evolutionarily conserved elements important in lung epithelial cells.


GATA
Human VIPR2 Gene -5928 AAAATAAAAGAGAAGACCCTGCTCCT----------CAGG-- -5898
Marmoset VIPR2 Gene -5365 AAAATAAAAGTGAAGACCCTGCTCCATACACTGGGCCACTGGTGCAGGTG -5316
Chimpanzee VIPR2 Gene -5941 AAAATAGAAGAGAAGACCCTGCTCCT----------CAGG-- -5911
Gorilla VIPR2 Gene -5654 AAAATAAAAGAGAAGACCCTGCTCCT----------CAGG-- -5624
YY1
Human VIPR2 Gene -5731 GAACTGGGGAGAGGACCTAGCAAAATAGGTTAGAAAATGGGAAAAAAGAT -5682
Marmoset VIPR2 Gene -5143 GAACTGGGGAAAGGACCTAGTCAAATAGGGTAGAAAATGGGAAAAAAGAT -5094
Chimpanzee VIPR2 Gene -5744 GAACTGGGGAGAGGACCTAGCAAAATAGGTTAGAAAATGGGAAAAAAGAT -5695
Gorilla VIPR2 Gene -5429 GAACTGGGGAGAGGACTTAGCAAAATAGGTTAGAAAATGGGAAAAAAGAT -5380
MEF2
Human VIPR2 Gene -5632 GATATTGTAAAAACTCTGGTTTTGTGTGTAAAAATAGTAACTGAGAACAA -5583
Marmoset VIPR2 Gene -5044 GATATTGTAAAAACTCCGGTTTTGTGTGTAAAAATAGTAACTGAGAACAA -4995
Chimpanzee VIPR2 Gene -5644 GATATTGTAAAAACTCTGGTTTTGTGTGTAAAAATAGTAACTGAGAACAA -5595
Gorilla VIPR2 Gene -5329 GATATTGTAAAAACTCTGGTTTTGTGTGTAAAAATAGTAACTGAGAACAA -5280
NKX25
Human VIPR2 Gene -5532 GAGAACATTATTAATTATTTAAAGGACCACTTGAGTCATTAATTATCCTT -5483
Marmoset VIPR2 Gene -4944 AAGAACATTATTAATTATTTAAAGGGTCATGTGAGTCATTAATTATCCTT -4895
Chimpanzee VIPR2 Gene -5544 GAGAACATTATTAATTATTTAAAGGGCCACTTGAGTCATTAATTATCCTT -5495
Gorilla VIPR2 Gene -5229 GAGAACATTATTAATTATTTAAAGGGCCACTTGAGTCATTAATTATCCTT -5180
ALPHA CP1 MEF-2/RSRFC
Human VIPR2 Gene -5432 ACTCCCCAATGAGATTTCTATTTCTGCTTCACACGCAGCTAAAAATAGTT -5383
Marmoset VIPR2 Gene -4844 ACTCTTCAATTAGATTTCTATTTCTGCTTCACATGCAGCTAAAAATAGTT -4795
Chimpanzee VIPR2 Gene -5444 ACTCCTCAATGAGATTTCTATTTCTGCTTCACACGCAGCTAAAAATAGTT -5395
Gorilla VIPR2 Gene -5129 ACTCCTCAATGAGATTTCTATTTCTGCTTCATATGCAGCTAAAAATAGTT -5080
Human VIPR2 Gene -5382 CCCTCTTTTGTTCCTTTAATGTTGTTTGTTTTCTTCTCTTCTCCCTAAAA -5333
Marmoset VIPR2 Gene -4794 CCCTATTTTGTCCCTTTAATG---TTTGTTTTCTTCTCTTCTCCCTAAAA -4748
Chimpanzee VIPR2 Gene -5394 CCCTCTTTTGTTCCTTTAATGTTGTTTGTTTTCTTCTCTTTTCCCTAAAA -5345
Gorilla VIPR2 Gene -5079 CCCTATTTTGTTCCTTTAATGTTGTTTGTTTTCTTCTCTTCTCCCTAAAA -5030
HAND1E47
Human VIPR2 Gene -5332 AGC----CACTTGGAAAGGTGGCATGGCCGTCAAACCAGACCCTGCTTCT -5287
Marmoset VIPR2 Gene -4747 AGCAAGCCACTTGGAAAGGTGACATGGCTGTCAAACCAGACTCTGCTTCT -4698
Chimpanzee VIPR2 Gene -5344 AGC----CACTTGGAAAGGTGGCATGGCCGTCAAACCAGACCCTGCTTCT -5299
Gorilla VIPR2 Gene -5029 AGC----CACTTGGAAAGGTGGCATGGCCGTCAAACCAGACCCTGCTTCT -4984
NKX25
Human VIPR2 Gene -5189 TCTCCATGCTCCCTGGATCCCCCACACCCTCGGAATTTTGGACAAGTGAT -5140
Marmoset VIPR2 Gene -4597 TCCCCATGCTCCGTGGATCCTCCACACCCTCAGAATTTTGG-CACATGAC -4549
Chimpanzee VIPR2 Gene -5201 TCTCCATGCTCCCTGGATCCCCCACACCCTCGGAATTTTGGACACGTGAT -5152
Gorilla VIPR2 Gene -4886 TCTCCATGCTCCCTGGATCCCCCACACCCTCGGAATTTTGGACACGTGAT -4837
Human VIPR2 Gene -5139 AAATGCATGACCGATGCCTACCGATGATGACCAGGGAGGCTCAGCTGACA -5090
Marmoset VIPR2 Gene -4548 AAATGCGTGACTGGTGC-TGCAGATAATGGCCAGGGGGGCTCAGCTGACA -4500
Chimpanzee VIPR2 Gene -5151 AAATGCATGACCGATGCCTACCGATGATGACCAGGGAGGCTCAGCTGACA -5102
Gorilla VIPR2 Gene -4836 AAATGCATGACCAATGCCTACCGATGATGACCAGGGAGGCTCAGCTGACA -4787
HAND1E47
Human VIPR2 Gene -4940 AGTGAGGACGCCCTGGC-GAGGGGAAGGGTCTGGAGGGCAGCCCCACTGA -4892
Marmoset VIPR2 Gene -4349 AACCAGAAGGCCCTGGCAGAGGGGCAGGGTCTGGAGGGCAGCCCTGCTGA -4300
Chimpanzee VIPR2 Gene -4952 AGTGAGGAGGCCCTGGC-GAGGGGAAGGGTCTGGAGGGCAGCCCCACTGA -4904
Gorilla VIPR2 Gene -4637 AGTGAGGAGGCCCTGCC-GAGGGGAAGGGTCTGGAGGGCAGCCCCGCTGA -4589
NKX25
Human VIPR2 Gene -4791 CCTCC-AGACTGGAGTGAGTGGAACCAGGAGTAACTCATACATCTGCCCA -4743
Marmoset VIPR2 Gene -4202 CCTCCCAGACTGGAGTGAGTGTAACCAGGAGTCACTCACACATCTGCCCA -4153
Chimpanzee VIPR2 Gene -4803 CCTCC-AGACTGGAGTGAGTGGAACCAGGAGTAACTCATACATCTGCCCG -4755
Gorilla VIPR2 Gene -4488 CCTCC-AGACTGGAGTGAGTGGAACCAGGAGTAACTCATACATCTGCCCG -4440
NKX25 E47
Human VIPR2 Gene -4742 CTTGCTCATCAACAAACAGGGAACACCTGCTGTACAAACCCACAATGCTG -4693
Marmoset VIPR2 Gene -4152 CTTGCTCGTCCACAAACAGGAAACACCTACTGTGCAAACCC-CAAGGCTG -4104
Chimpanzee VIPR2 Gene -4754 CTTGCTCATCAACAAACAGGGAACACCTGCTGTACAAACCCACAATGCTA -4705
Gorilla VIPR2 Gene -4439 CTTGCTCATCAACAAACAGAGAACGCCTGCTGTACAAACCCACAATGCTG -4390
RSRFC NKX25
Human VIPR2 Gene -4642 CTATCGCATCATATTATATTAAGCTAAGAATAGTTCTCCAAGAAAGTGGC -4593
Marmoset VIPR2 Gene -4053 ----------ATATTATATCAAGCCCTGAATGGTTCTCCAAGAAAATGGC -4013
Chimpanzee VIPR2 Gene -4654 CTATCGCATCATATTATATTAAGCTAAGAATAGTTCTCCAAGAAAGTGGC -4605
Gorilla VIPR2 Gene -4339 CTATCACATCATATTATATTAAGCTAAGAATAGTTCTCCAAGAAAGTGGC -4290
E-BOX/E2A/Myogenin
Human VIPR2 Gene -4442 TCTCTGTGAGCCCCACCTGCCTGGGGTGGGAGGGGGCCCGGGCAGGGCTA -4393
Marmoset VIPR2 Gene -3862 CCTCTGTGAGTCCCATCATCCTGGGGTGGGAGAGGGGCC---------- -3823
Chimpanzee VIPR2 Gene -4454 TCTCTGTGAGCCCCACCTGCCTGGGGTGGGAGGGGGCCCG---------- -4414
Gorilla VIPR2 Gene -4139 TCTCTGTGAGCCCCACCTGCCTGGGGTGGGAGGGGGCCCCAGGCAGGGCT -4090
TAL1BETAE47
Human VIPR2 Gene -4343 GGCTGCGGCACGCACTGGTGTGTGCCCGTCTCTTGGACAAACCATCTGAG -4294
Marmoset VIPR2 Gene -3811 --CTGTGCCATGTGCTGGTGTGTGGCCAGCTCTTGGACAAACCATCCAAG -3764
Chimpanzee VIPR2 Gene -4407 GGCTGCGGCATGCACTGGTGTGTGCCCGTCTCTTGGACAAACCATCTGAG -4358
Gorilla VIPR2 Gene -4039 GGCTGCGGCACGCACTGGTGTGTGCCCGTCTCTTGGACAAACCATCTGAG -3990


E-BOX/E2A/MYOD
Human VIPR2 Gene -4144 GGGAGGGCTTTCCTAAGGCAGACACCTGAGCCAAGACAGGTGGAAGCTGG -4095
Marmoset VIPR2 Gene -3664 AATATTTAGGTTTTAAAAGATACAAAAAGAGGGACACAGGCAAATGATGA -3615
Chimpanzee VIPR2 Gene -4208 GGGAGGGCTTTCCTAAGGCAGACACCTGAGCCAAGACAGGTGGAAGCTGG -4159
Gorilla VIPR2 Gene -3843 GCGAGGGCTTTCCTAAGGCAGACACCTGAGCCAAGACAGGTGGAAGCTGG -3794
MYOGENIN
TBX5
Human VIPR2 Gene -3845 GCCAGGCCCAGCATGGCAGGTGGTTGCCTGTTCAAAGGTG-TAAACCAAA -3797
Marmoset VIPR2 Gene -3414 GC-AGTACTAGCATGGCA---------CCTTTTTAAAAGTGATTAAAAGGT -3374
Chimpanzee VIPR2 Gene -3909 GCCAGGCCCAGCATGGCAGGTGGTTGCCTGTTCAAAGGTG-TAAACCAAA -3861
Gorilla VIPR2 Gene -3560 GCCAGGCCCAGCATGGCAGGTGGTTGCCTGTTCAAAGGTG-TAAACCAAA -3512
TATA/TBP
Human VIPR2 Gene -3697 CCCAAGGTGACAGGGTATAGCTTGGTTTTATACATTTTAGGGAGACTGAA -3648
Marmoset VIPR2 Gene -3281 -----AGTAGCTG--CTCAGTTTCCGCCAGTGTCTGCCAGAGAT-TTGAT -3240
Chimpanzee VIPR2 Gene -3761 CCCAAGGTGACAGGGTATAGCTTGGTTTTATACATTTTAGGGAGACTGAA -3712
Gorilla VIPR2 Gene -3412 CCCAAGGTGACAGGGTATAGCTTGGTTTTATACATTTTAGGGAGACTGAA -3363
NFY
Human VIPR2 Gene -3547 AGGTGGATTACAGGATTTCCTGATTGGCCGTTGGGTAAAACAGAGTTAAG -3498
Marmoset VIPR2 Gene -3155 GTCCAGGGGGCAGGAT-----GAGAAACTGAAAAGCACCTCAGGGGTCCC -3111
Chimpanzee VIPR2 Gene -3611 AGGTGGATTACAGGATTTCCTGATTGGCAGTTGGGTAAAACAGAGTTAAG -3562
Gorilla VIPR2 Gene -3262 AGGTGGATTACAGGATTTCCTGATTGGCAGTTGGGTAAAACAGAGTTAAG -3213
TBX5
Human VIPR2 Gene -3397 AGTGAATGTGAATGTCTACTATCAGGGCCTTGAAAGGTGTCAGACTCTCC -3348
Marmoset VIPR2 Gene -3015 GGTGGGAGCAGGAGGCAGAACCCAGGCTTCTGCTGGGTG-CATAAGCTCA -2966
Chimpanzee VIPR2 Gene -3461 AGTGAATGTGAATGTCTACTATCAGGGCCTTGAAAGGTGTCAGACTCTCC -3412
Gorilla VIPR2 Gene -3112 AGTGAATGTGAATGTCTACTATCAGGGCCTTGAAAGGTGTCAGACTCTCC -3063
HAND1E47
Human VIPR2 Gene -3247 CTTCAGGGCCTGCTGGCCGTCATGTGATGCTCTACTAGAGTCTGGTGGGA -3198
Marmoset VIPR2 Gene -2877 CT----GCCCTGAGGAC----ACA-GGGGCTCTGGGAGGAGGAAGTGTGA -2837
Chimpanzee VIPR2 Gene -3311 CTTCAGGGCCTGCTGGCCGTCATGTGATGCTCTACTAGAGTCTGGTGGGA -3262
Gorilla VIPR2 Gene -2962 CTTCAGGGCCTGCTGGCCGTCATGTGATGCTCTACTAGAGTCTGGTGGGA -2913
YY1
Human VIPR2 Gene -2898 AAGCCCGTGATTACACACACACTGGACCTTCAGAGAAGTGTGCCATCTAG -2849
Marmoset VIPR2 Gene -2581 CAATCACTGCTCTCA-----GCTCTTCTCCCCAAGA--TGTCCCATCTG- -2540
Chimpanzee VIPR2 Gene -2962 AAGCCCGTGACTACACACACACTGGACCTTCAGAGAAGTGTGCCATCTAG -2913
Gorilla VIPR2 Gene -2613 AAGCCCGTGACTACACACACACTGGACCTTCAGAGAAGTGTGCCATCTAG -2564
YY1
Human VIPR2 Gene -2848 CTACGACCACAGCCATTTCTCCTGGTCTACCCCGTCCATCGGGACGCACT -2799
Marmoset VIPR2 Gene -2539 ---CGGGAAAGGCTCTGCACCCCACACTCGGCCACCCACC----CACACT -2497
Chimpanzee VIPR2 Gene -2912 CTACGACCACAGCCATTTCTCCTGGTCTACCCCGTCCATCGGGACGCACT -2863
Gorilla VIPR2 Gene -2563 CTATGACCACAGCCATTTCTCCTGGTCTACCCCATCCATTGGGACGCACT -2514
MYOD/MYOGENIN
Human VIPR2 Gene -2698 TCACCTTGATTATTACAGAGCCCCTCTGCGGTGGAAGGGCAGGTGCTGTC -2649
Marmoset VIPR2 Gene -2420 TCCCTCTGCCTACT-CACACTCTCCCCTC----------CACCCACTCTC -2382
Chimpanzee VIPR2 Gene -2762 TTACCTTGATTATTACAGAGCCCCTCTGCGGTGGAAGGGCAGGTGCTGTC -2713
Gorilla VIPR2 Gene -2431 TTACCTTGATTCTTACAGAGCCCCTCTGCGGTGGAAGGGCAGGTGCTGTC -2382
GATA-4
Human VIPR2 Gene -2503 -GTCTCCCACTAGATACACATGGATGAAGCCAGATAGATAACAGGTAGAG -2455
Marmoset VIPR2 Gene -2231 AGTTTCCCATTAGATACACATGGATGAAGCCAGAGAGATAACAGGTAGCA -2182
Chimpanzee VIPR2 Gene -2566 -GTCTCCCACTAGATACACATGGATGAAGCCAGATAGATAACAGGTAGAG -2518
Gorilla VIPR2 Gene -2235 -GTCTCCCAATAGATACACATGGATGAAGCCAGATAGATAACAGGTAGAG -2187
TAL1ALPHAE47
Human VIPR2 Gene -2354 GATAGATGACAAATAGGCAGATGATAGATGACAGGCAGATGACAGATAGG -2305
Marmoset VIPR2 Gene -1516 GAAAAATGACAGATAGACAGATGATAGATCACAGACC-------------- -1479
Chimpanzee VIPR2 Gene -2433 GATAGATGACAAATAGGCAGATGATAGATGACAGGCAGATGACAGGCAGA -2384
Gorilla VIPR2 Gene -2102 GATAGATGACAAATAGGCAGATGATACACGACAGGCAGATGACAGGCAGA -2053
TAL1ALPHAE47
Human VIPR2 Gene -2209 ---ACAGATAGGCAGATGATAGATGACTGATAGATGACAGGCAGATCACA -2163
Marmoset VIPR2 Gene -1479 ---ATAGATAGGAAGATGATAGATGA------------------------ -1456
Chimpanzee VIPR2 Gene -2283 ATGGATGACAGGCAGATGATAGATGATCGATAGATGACAGGCAGATGATA -2234
Gorilla VIPR2 Gene -2018 -------ATAGGCAGATGATAGATGA----------CAGGCAG---ATA -1989
GATA4
Human VIPR2 Gene -1702 ----------GGTAGATGGTAGGTAATAGATAGGTAATAGATCGCAGAGA -1662
Marmoset VIPR2 Gene -1436 ----------GGTAGATGGTAG----------TAATAGATGTCAGACA -1409
Chimpanzee VIPR2 Gene -1787 ATAGGAAGATAATAGATGATAGGTACATGA---CAGATAGATCATAGATA -1741
Gorilla VIPR2 Gene -1661 -------GATGGTAGATAATAGATAGGTAA---TAGATAGATCGCAGAGA -1701
HMEF2
Human VIPR2 Gene -1191 TGTTAATGACCGAATCCACAAAAAGGATAAAAAGGAGTTATTTGTAATAT -1142
Marmoset VIPR2 Gene -1049 TGTTAATGACTGAATCCAGATGAAGGGTAAAAAG-AGTTCTTTGTACTAT -1001
Chimpanzee VIPR2 Gene -1193 TGTTAATGACCGAATCCACAAAAAGGATAAAAAGGAGTTATTTGTAATAT -1144
Gorilla VIPR2 Gene -1188 TGTTAATGACCGAATCCACAAAAAGGATAAAAAGGAGTTACTGGTAATAT -1139
GATA4
NKX25
Human VIPR2 Gene -1091 ATTAAAACA-ACACGAACTTAACTTGTATCTCTCATT-CCAGTGCTTTCC -1044
Marmoset VIPR2 Gene -0950 AAAAAAACACACACAAAATTAGCTTTCGTCTCTCATTTCCAGTGCCTTCC -0901
Chimpanzee VIPR2 Gene -1093 ATTAAAACA-ACACAAACTTAACTTGTATCTCTCATT-CCAGTGCTTTCC -1046
Gorilla VIPR2 Gene -1088 A-----ACA-ACACAAACTTAACTTGTATCTCTCATT-CCAGTGCTTTCC -1046

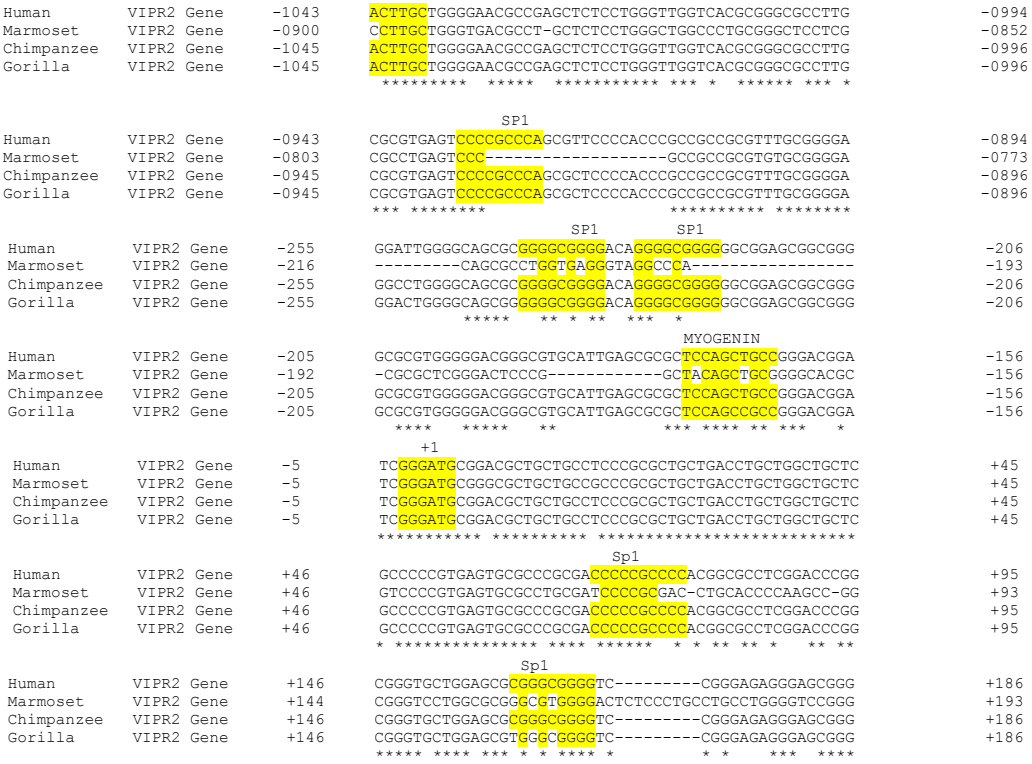
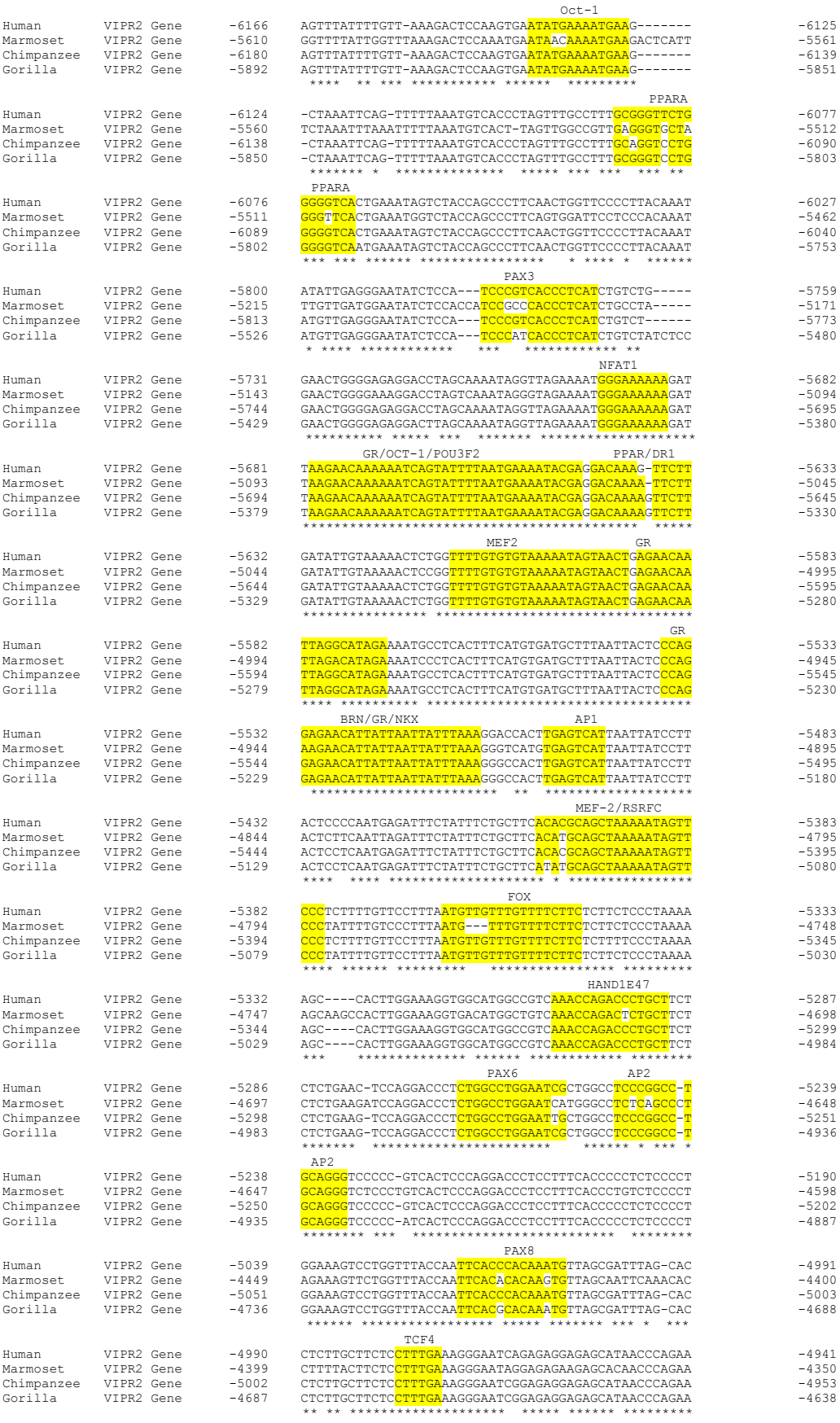
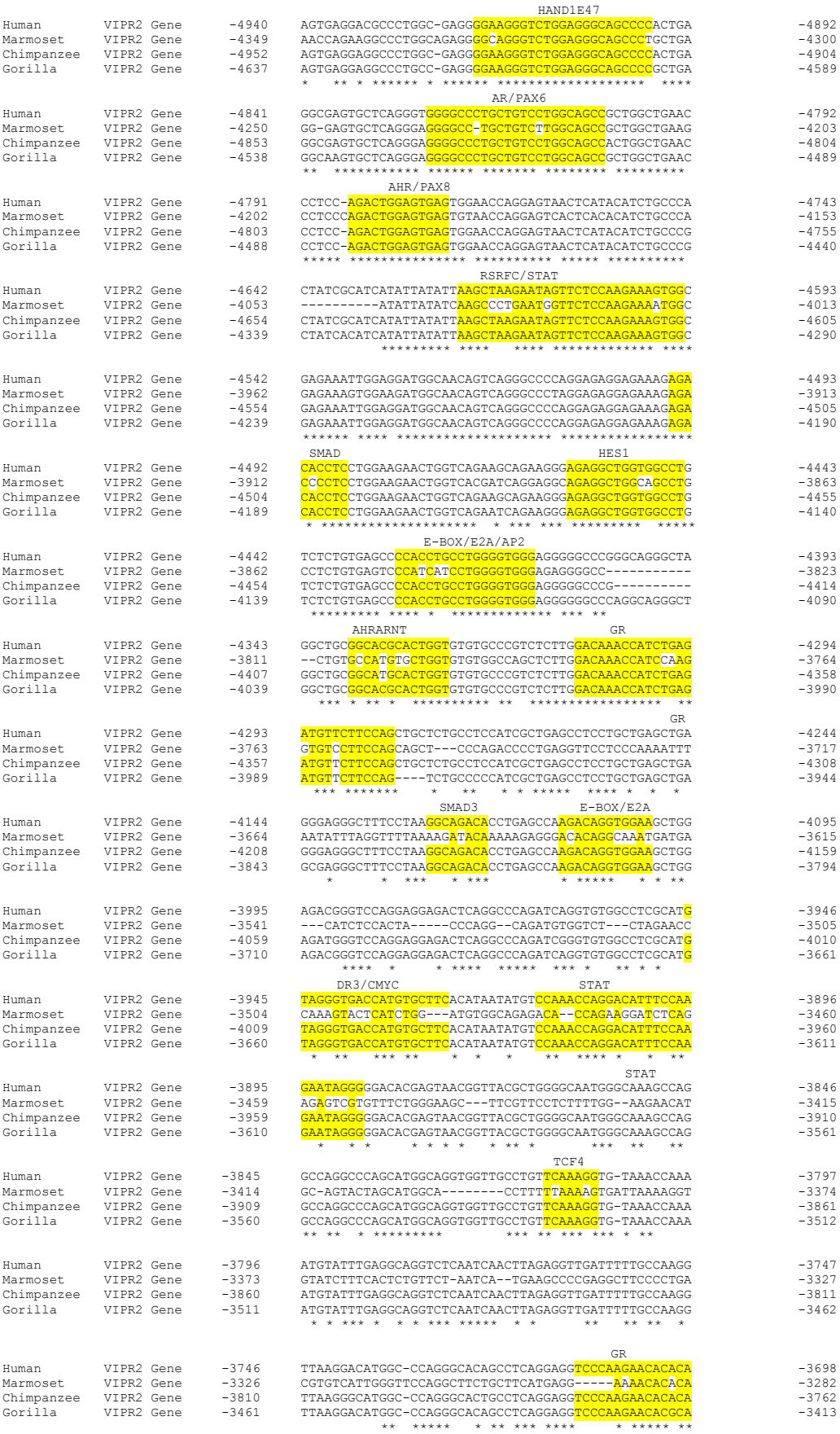
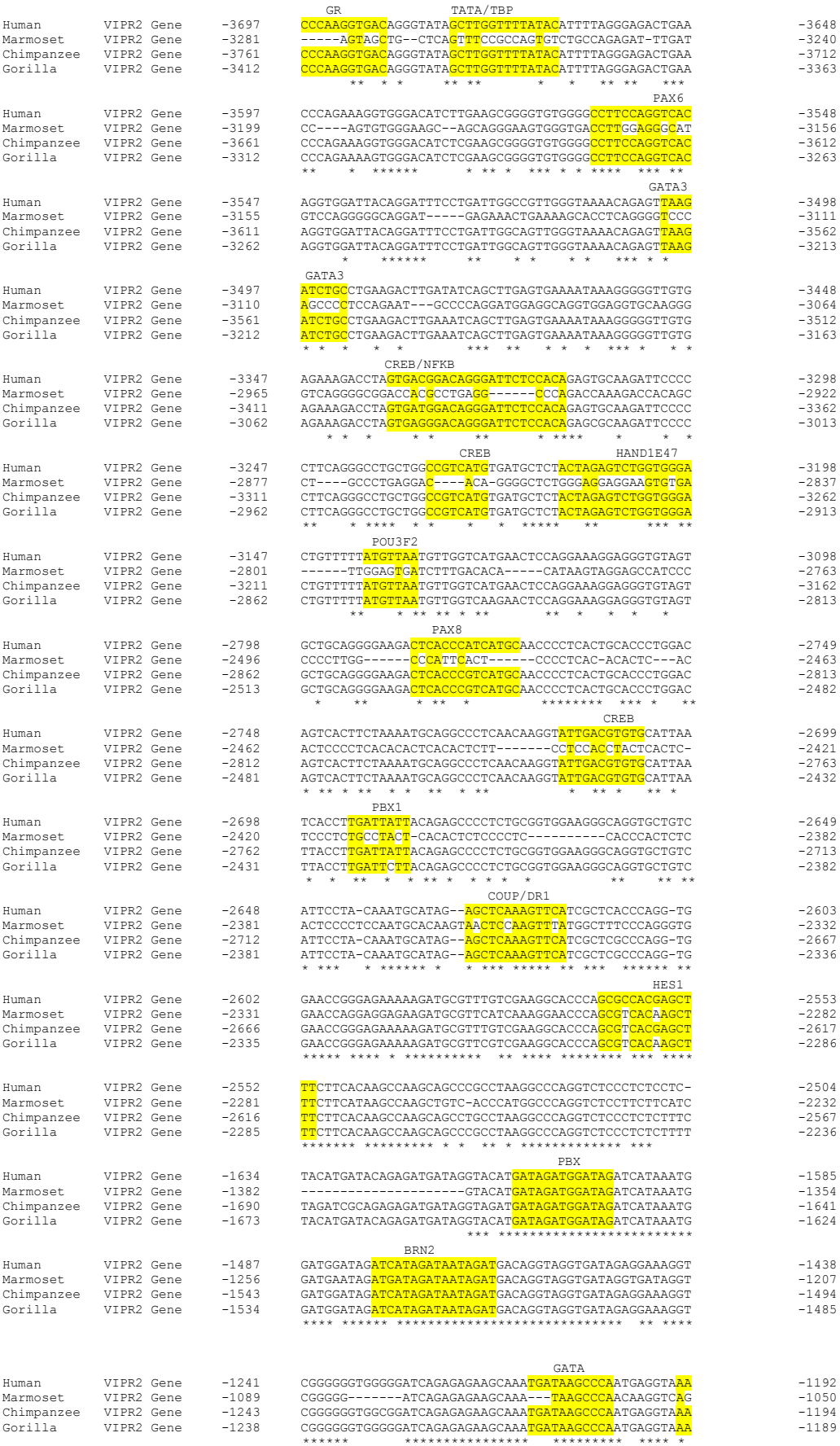


Figure 3.15 Evolutionarily conserved elements important in Myocytes.


Oct-1
PPARA
PPARA
PAX3
NFAT1
GR/OCT-1/POU3F2
PPAR/DR1
MEF2
GR
GR
BRN/GR/NKX
AP1
MEF-2/RSRFC
FOX
HAND1E47
PAX6
AP2
AP2
PAX8
TCF4


HAND1E47
AR/PAX6
AHR/PAX8
RSRFC/STAT
SMAD
HES1
E-BOX/E2A/AP2
AHRARNT
GR
SMAD3
E-BOX/E2A
DR3/CMYC
STAT
TCF4
Human VIPR2 Gene -4940 AGTGAGGACGCCCTGGC-GAGGGGAAGGGTCTGGAGGGCAGCCCCACTGA -4892
Marmoset VIPR2 Gene -4349 AACCAGAAGGCCCTGGCAGAGGGGCAGGGTCTGGAGGGCAGCCCTGCTGA -4300
Chimpanzee VIPR2 Gene -4952 AGTGAGGAGGCCCTGGC-GAGGGGAAGGGTCTGGAGGGCAGCCCCACTGA -4904
Gorilla VIPR2 Gene -4637 AGTGAGGAGGCCCTGCC-GAGGGGAAGGGTCTGGAGGGCAGCCCCGCTGA -4589
Human VIPR2 Gene -4841 GGCGAGTGCTCAGGGTGGGGCCCTGCTGTCCTGGCAGCCGCTGGCTGAAC -4792
Marmoset VIPR2 Gene -4250 GG-GAGTGCTCAGGGAGGGGCC-TGCTGTCTTGGCAGCCGCTGGCTGAAG -4203
Chimpanzee VIPR2 Gene -4853 GGCGAGTGCTCAGGGAGGGGCCCTGCTGTCCTGGCAGCCACTGGCTGAAC -4804
Gorilla VIPR2 Gene -4538 GGCAAGTGCTCAGGGAGGGGCCCTGCTGTCCTGGCAGCCGCTGGCTGAAC -4489
Human VIPR2 Gene -4791 CCTCC-AGACTGGAGTGAGTGGAACCAGGAGTAACTCATACATCTGCCCA -4743
Marmoset VIPR2 Gene -4202 CCTCCCAGACTGGAGTGAGTGTAACCAGGAGTCACTCACACATCTGCCCA -4153
Chimpanzee VIPR2 Gene -4803 CCTCC-AGACTGGAGTGAGTGGAACCAGGAGTAACTCATACATCTGCCCG -4755
Gorilla VIPR2 Gene -4488 CCTCC-AGACTGGAGTGAGTGGAACCAGGAGTAACTCATACATCTGCCCG -4440
Human VIPR2 Gene -4642 CTATCGCATCATATTATATTAAGCTAAGAATAGTTCTCCAAGAAAGTGGC -4593
Marmoset VIPR2 Gene -4053 ----------ATATTATATCAAGCCCTGAATGGTTCTCCAAGAAAATGGC -4013
Chimpanzee VIPR2 Gene -4654 CTATCGCATCATATTATATTAAGCTAAGAATAGTTCTCCAAGAAAGTGGC -4605
Gorilla VIPR2 Gene -4339 CTATCACATCATATTATATTAAGCTAAGAATAGTTCTCCAAGAAAGTGGC -4290
Human VIPR2 Gene -4542 GAGAAATTGGAGGATGGCAACAGTCAGGGCCCCAGGAGAGGAGAAAGAGA -4493
Marmoset VIPR2 Gene -3962 GAGAAAGTGGAAGATGGCAACAGTCAGGGCCCTAGGAGAGGAGAAAGAGA -3913
Chimpanzee VIPR2 Gene -4554 GAGAAATTGGAGGATGGCAACAGTCAGGGCCCCAGGAGAGGAGAAAGAGA -4505
Gorilla VIPR2 Gene -4239 GAGAAATTGGAGGATGGCAACAGTCAGGGCCCCAGGAGAGGAGAAAGAGA -4190
Human VIPR2 Gene -4492 CACCTCCTGGAAGAACTGGTCAGAAGCAGAAGGGAGAGGCTGGTGGCCTG -4443
Marmoset VIPR2 Gene -3912 CCCCTCCTGGAAGAACTGGTCACGATCAGGAGGCAGAGGCTGGCAGCCTG -3863
Chimpanzee VIPR2 Gene -4504 CACCTCCTGGAAGAACTGGTCAGAAGCAGAAGGGAGAGGCTGGTGGCCTG -4455
Gorilla VIPR2 Gene -4189 CACCTCCTGGAAGAACTGGTCAGAATCAGAAGGGAGAGGCTGGTGGCCTG -4140
Human VIPR2 Gene -4442 TCTCTGTGAGCCCCACCTGCCTGGGGTGGGAGGGGGCCCGGGCAGGGCTA -4393
Marmoset VIPR2 Gene -3862 CCTCTGTGAGTCCCATCATCCTGGGGTGGGAGAGGGGCC---------- -3823
Chimpanzee VIPR2 Gene -4454 TCTCTGTGAGCCCCACCTGCCTGGGGTGGGAGGGGGCCCG---------- -4414
Gorilla VIPR2 Gene -4139 TCTCTGTGAGCCCCACCTGCCTGGGGTGGGAGGGGGGCCCAGGCAGGGCT -4090
Human VIPR2 Gene -4343 GGCTGCGGCACGCACTGGTGTGTGCCCGTCTCTTGGACAAACCATCTGAG -4294
Marmoset VIPR2 Gene -3811 --CTGTGCCATGTGCTGGTGTGTGGCCAGCTCTTGGACAAACCATCCAAG -3764
Chimpanzee VIPR2 Gene -4407 GGCTGCGGCATGCACTGGTGTGTGCCCGTCTCTTGGACAAACCATCTGAG -4358
Gorilla VIPR2 Gene -4039 GGCTGCGGCACGCACTGGTGTGTGCCCGTCTCTTGGACAAACCATCTGAG -3990
Human VIPR2 Gene -4293 ATGTTCTTCCAGCTGCTCTGCCTCCATCGCTGAGCCTCCTGCTGAGCTGA -4244
Marmoset VIPR2 Gene -3763 GTGTCCTTCCAGCAGCT---CCCAGACCCCTGAGGTTCCTCCCAAAATTT -3717
Chimpanzee VIPR2 Gene -4357 ATGTTCTTCCAGCTGCTCTGCCTCCATCGCTGAGCCTCCTGCTGAGCTGA -4308
Gorilla VIPR2 Gene -3989 ATGTTCTTCCAG----TCTGCCCCCATCGCTGAGCCTCCTGCTGAGCTGA -3944
Human VIPR2 Gene -4144 GGGAGGGCTTTCCTAAGGCAGACACCTGAGCCAAGACAGGTGGAAGCTGG -4095
Marmoset VIPR2 Gene -3664 AATATTTAGGTTTTAAAAGATACAAAAGAGGGACACAGGCAAATGATGA -3615
Chimpanzee VIPR2 Gene -4208 GGGAGGGCTTTCCTAAGGCAGACACCTGAGCCAAGACAGGTGGAAGCTGG -4159
Gorilla VIPR2 Gene -3843 GCGAGGGCTTTCCTAAGGCAGACACCTGAGCCAAGACAGGTGGAAGCTGG -3794
Human VIPR2 Gene -3995 AGACGGGTCCAGGAGGAGACTCAGGCCCAGATCAGGTGTGGCCTCGCATG -3946
Marmoset VIPR2 Gene -3541 ---CATCTCCACTA-----CCCAGG--CAGATGTGGTCT---CTAGAACC -3505
Chimpanzee VIPR2 Gene -4059 AGATGGGTCCAGGAGGAGACTCAGGCCCAGATCGGGTGTGGCCTCGCATG -4010
Gorilla VIPR2 Gene -3710 AGACGGGTCCAGGAGGAGACTCAGGCCCAGATCAGGTGTGGCCTCGCATG -3661
Human VIPR2 Gene -3945 TAGGGTGACCATGTGCTTCACATAATATGTCCAAACCAGGACATTTCCAA -3896
Marmoset VIPR2 Gene -3504 CAAAGTACTCATCTGG---ATGTGGCAGAGACA--CCAGAAGGATCTCAG -3460
Chimpanzee VIPR2 Gene -4009 TAGGGTGACCATGTGCTTCACATAATATGTCCAAACCAGGACATTTCCAA -3960
Gorilla VIPR2 Gene -3660 TAGGGTGACCATGTGCTTCACATAATATGTCCAAACCAGGACATTTCCAA -3611
Human VIPR2 Gene -3895 GAATAGGGGGACACGAGTAACGGTTACGCTGGGGCAATGGGCAAAGCCAG -3846
Marmoset VIPR2 Gene -3459 AGAGTCGTGTTTCTGGGAAGC---TTCGTTCCTCTTTTGG--AAGAACAT -3415
Chimpanzee VIPR2 Gene -3959 GAATAGGGGGACACGAGTAACGGTTACGCTGGGGCAATGGGCAAAGCCAG -3910
Gorilla VIPR2 Gene -3610 GAATAGGGGGACACGAGTAACGGTTACGCTGGGGCAATGGGCAAAGCCAG -3561
Human VIPR2 Gene -3845 GCCAGGCCCAGCATGGCAGGTGGTTGCCTGTTCAAAGGTG-TAAACCAAA -3797
Marmoset VIPR2 Gene -3414 GC-AGTACTAGCATGGCA---------CCTTTTTAAAAGTGATTAAAAGGT -3374
Chimpanzee VIPR2 Gene -3909 GCCAGGCCCAGCATGGCAGGTGGTTGCCTGTTCAAAGGTG-TAAACCAAA -3861
Gorilla VIPR2 Gene -3560 GCCAGGCCCAGCATGGCAGGTGGTTGCCTGTTCAAAGGTG-TAAACCAAA -3512
Human VIPR2 Gene -3796 ATGTATTTGAGGCAGGTCTCAATCAACTTAGAGGTTGATTTTTGCCAAGG -3747
Marmoset VIPR2 Gene -3373 GTATCTTTCACTCTGTTCT-AATCA--TGAAGCCCCGAGGCTTCCCCTGA -3327
Chimpanzee VIPR2 Gene -3860 ATGTATTTGAGGCAGGTCTCAATCAACTTAGAGGTTGATTTTTGCCAAGG -3811
Gorilla VIPR2 Gene -3511 ATGTATTTGAGGCAGGTCTCAATCAACTTAGAGGTTGATTTTTGCCAAGG -3462
Human VIPR2 Gene -3746 TTAAGGACATGGC-CCAGGGCACAGCCTCAGGAGGTCCCAAGAACACACA -3698
Marmoset VIPR2 Gene -3326 CGTGTCATTGGGTTCCAGGCTTCTGCTTCATGAGG-----AAAACACACA -3282
Chimpanzee VIPR2 Gene -3810 TTAAGGGCATGGC-CCAGGGCACTGCCTCAGGAGGTCCCAAGAACACACA -3762
Gorilla VIPR2 Gene -3461 TTAAGGACATGGC-CCAGGGCACAGCCTCAGGAGGTCCCAAGAACACGCA -3413


GR
TATA/TBP
Human VIPR2 Gene -3697 CCCAAGGTGACAGGGTATAGCTTGGTTTTATACATTTTAGGGAGACTGAA -3648
Marmoset VIPR2 Gene -3281 -----AGTAGCTG--CTCAGTTTCCGCCAGTGTCTGCCAGAGAT-TTGAT -3240
Chimpanzee VIPR2 Gene -3761 CCCAAGGTGACAGGGTATAGCTTGGTTTTATACATTTTAGGGAGACTGAA -3712
Gorilla VIPR2 Gene -3412 CCCAAGGTGACAGGGTATAGCTTGGTTTTATACATTTTAGGGAGACTGAA -3363
PAX6
Human VIPR2 Gene -3597 CCCAGAAAGGTGGGACATCTTGAAGCGGGGTGTGGGGCCTTCCAGGTCAC -3548
Marmoset VIPR2 Gene -3199 CC----AGTGTGGGAAGC--AGCAGGGAAGTGGGTGACCTTGGAGGGCAT -3156
Chimpanzee VIPR2 Gene -3661 CCCAGAAAGGTGGGACATCTCGAAGCGGGGTGTGGGGCCTTCCAGGTCAC -3612
Gorilla VIPR2 Gene -3312 CCCAGAAAAGTGGGACATCTCGAAGCGGGGTGTGGGGCCTTCCAGGTCAC -3263
GATA3
Human VIPR2 Gene -3547 AGGTGGATTACAGGATTTCCTGATTGGCCGTTGGGTAAAACAGAGTTAAG -3498
Marmoset VIPR2 Gene -3155 GTCCAGGGGGCAGGAT-----GAGAAACTGAAAAGCACCTCAGGGGTCCC -3111
Chimpanzee VIPR2 Gene -3611 AGGTGGATTACAGGATTTCCTGATTGGCAGTTGGGTAAAACAGAGTTAAG -3562
Gorilla VIPR2 Gene -3262 AGGTGGATTACAGGATTTCCTGATTGGCAGTTGGGTAAAACAGAGTTAAG -3213
GATA3
Human VIPR2 Gene -3497 ATCTGCCTGAAGACTTGATATCAGCTTGAGTGAAAATAAAGGGGGTTGTG -3448
Marmoset VIPR2 Gene -3110 AGCCCCTCCAGAAT---GCCCCAGGATGGAGGCAGGTGGAGGTGCAAGGG -3064
Chimpanzee VIPR2 Gene -3561 ATCTGCCTGAAGACTTGAAATCAGCTTGAGTGAAAATAAAGGGGGTTGTG -3512
Gorilla VIPR2 Gene -3212 ATCTGCCTGAAGACTTGAAATCAGCTTGAGTGAAAATAAAGGGGGTTGTG -3163
CREB/NFKB
Human VIPR2 Gene -3347 AGAAAGACCTAGTGACGGACAGGGATTCTCCACAGAGTGCAAGATTCCCC -3298
Marmoset VIPR2 Gene -2965 GTCAGGGGCGGACCACGCCTGAGG------CCCAGACCAAAGACCACAGC -2922
Chimpanzee VIPR2 Gene -3411 AGAAAGACCTAGTGATGGACAGGGATTCTCCACAGAGTGCAAGATTCCCC -3362
Gorilla VIPR2 Gene -3062 AGAAAGACCTAGTGAGGGACAGGGATTCTCCACAGAGCGCAAGATTCCCC -3013
CREB
HAND1E47
Human VIPR2 Gene -3247 CTTCAGGGCCTGCTGGCCGTCATGTGATGCTCTACTAGAGTCTGGTGGGA -3198
Marmoset VIPR2 Gene -2877 CT----GCCCTGAGGAC----ACA-GGGGCTCTGGGAGGAGGAAGTGTGA -2837
Chimpanzee VIPR2 Gene -3311 CTTCAGGGCCTGCTGGCCGTCATGTGATGCTCTACTAGAGTCTGGTGGGA -3262
Gorilla VIPR2 Gene -2962 CTTCAGGGCCTGCTGGCCGTCATGTGATGCTCTACTAGAGTCTGGTGGGA -2913
POU3F2
Human VIPR2 Gene -3147 CTGTTTTTATGTTAATGTTGGTCATGAACTCCAGGAAAGGAGGGTGTAGT -3098
Marmoset VIPR2 Gene -2801 ------TTGGAGTGATCTTTGACACA-----CATAAGTAGGAGCCATCCC -2763
Chimpanzee VIPR2 Gene -3211 CTGTTTTTATGTTAATGTTGGTCATGAACTCCAGGAAAGGAGGGTGTAGT -3162
Gorilla VIPR2 Gene -2862 CTGTTTTTATGTTAATGTTGGTCAAGAACTCCAGGAAAGGAGGGTGTAGT -2813
PAX8
Human VIPR2 Gene -2798 GCTGCAGGGGAAGACTCACCCATCATGCAACCCCTCACTGCACCCTGGAC -2749
Marmoset VIPR2 Gene -2496 CCCCTTGG------CCCATTCACT------CCCCTCAC-ACACTC---AC -2463
Chimpanzee VIPR2 Gene -2862 GCTGCAGGGGAAGACTCACCCGTCATGCAACCCCTCACTGCACCCTGGAC -2813
Gorilla VIPR2 Gene -2513 GCTGCAGGGGAAGACTCACCCGTCATGCAACCCCTCACTGCACCCTGGAC -2482
CREB
Human VIPR2 Gene -2748 AGTCACTTCTAAAATGCAGGCCCTCAACAAGGTATTGACGTGTGCATTAA -2699
Marmoset VIPR2 Gene -2462 ACTCCCCTCACACACTCACACTCTT-------CCTCCACCTACTCACTC- -2421
Chimpanzee VIPR2 Gene -2812 AGTCACTTCTAAAATGCAGGCCCTCAACAAGGTATTGACGTGTGCATTAA -2763
Gorilla VIPR2 Gene -2481 AGTCACTTCTAAAATGCAGGCCCTCAACAAGGTATTGACGTGTGCATTAA -2432
PBX1
Human VIPR2 Gene -2698 TCACCTTGATTATTACAGAGCCCCTCTGCGGTGGAAGGGCAGGTGCTGTC -2649
Marmoset VIPR2 Gene -2420 TCCCTCTGCCTACT-CACACTCTCCCCTC----------CACCCACTCTC -2382
Chimpanzee VIPR2 Gene -2762 TTACCTTGATTATTACAGAGCCCCTCTGCGGTGGAAGGGCAGGTGCTGTC -2713
Gorilla VIPR2 Gene -2431 TTACCTTGATTCTTACAGAGCCCCTCTGCGGTGGAAGGGCAGGTGCTGTC -2382
COUP/DR1
Human VIPR2 Gene -2648 ATTCCTA-CAAATGCATAG--AGCTCAAAGTTCATCGCTCACCCAGG-TG -2603
Marmoset VIPR2 Gene -2381 ACTCCCCTCCAATGCACAAGTAACTCCAAGTTTATGGCTTTCCCAGGGTG -2332
Chimpanzee VIPR2 Gene -2712 ATTCCTA-CAAATGCATAG--AGCTCAAAGTTCATCGCTCGCCCAGG-TG -2667
Gorilla VIPR2 Gene -2381 ATTCCTA-CAAATGCATAG--AGCTCAAAGTTCATCGCTCGCCCAGG-TG -2336
HES1
Human VIPR2 Gene -2602 GAACCGGGAGAAAAAGATGCGTTTGTCGAAGGCACCCAGCGCCACGAGCT -2553
Marmoset VIPR2 Gene -2331 GAACCAGGAGGAGAAGATGCGTTCATCAAAGGAACCCAGCGTCACAAGCT -2282
Chimpanzee VIPR2 Gene -2666 GAACCGGGAGAAAAAGATGCGTTTGTCGAAGGCACCCAGCGTCACGAGCT -2617
Gorilla VIPR2 Gene -2335 GAACCGGGAGAAAAAGATGCGTTCGTCGAAGGCACCCAGCGTCACAAGCT -2286
Human VIPR2 Gene -2552 TTCTTCACAAGCCAAGCAGCCCGCCTAAGGCCCAGGTCTCCCTCTCCTC- -2504
Marmoset VIPR2 Gene -2281 TTCTTCATAAGCCAAGCTGTC-ACCCATGGCCCAGGTCTCCTTCTTCATC -2232
Chimpanzee VIPR2 Gene -2616 TTCTTCACAAGCCAAGCAGCCTGCCTAAGGCCCAGGTCTCCCTCTCTTTC -2567
Gorilla VIPR2 Gene -2285 TTCTTCACAAGCCAAGCAGCCCGCCTAAGGCCCAGGTCTCCCTCTCTTTT -2236
PBX
Human VIPR2 Gene -1634 TACATGATACAGAGATGATAGGTACATGATAGATGGATAGATCATAAATG -1585
Marmoset VIPR2 Gene -1382 ----------GTACATGATAGATGGATAGATCATAAATG -1354
Chimpanzee VIPR2 Gene -1690 TAGATCGCAGAGAGATGATAGGTAGATGATAGATGGATAGATCATAAATG -1641
Gorilla VIPR2 Gene -1673 TACATGATACAGAGATGATAGGTACATGATAGATGGATAGATCATAAATG -1624
BRN2
Human VIPR2 Gene -1487 GATGGATAGATCATAGATAATAGATGACAGGTAGGTGATAGAGGAAAGGT -1438
Marmoset VIPR2 Gene -1256 GATGAATAGATGATAGATAATAGATGACAGGTAGGTGATAGGTGATAGGT -1207
Chimpanzee VIPR2 Gene -1543 GATGGATAGATCATAGATAATAGATGACAGGTAGGTGATAGAGGAAAGGT -1494
Gorilla VIPR2 Gene -1534 GATGGATAGATCATAGATAATAGATGACAGGTAGGTGATAGAGGAAAGGT -1485
GATA
Human VIPR2 Gene -1241 CGGGGGGTGGGGGATCAGAGAGAAGCAAATGATAAGCCCAATGAGGTAAA -1192
Marmoset VIPR2 Gene -1089 CGGGGG-------ATCAGAGAGAAGCAAA---TAAGCCCAACAAGGTCAG -1050
Chimpanzee VIPR2 Gene -1243 CGGGGGGTGGCGGATCAGAGAGAAGCAAATGATAAGCCCAATGAGGTAAA -1194
Gorilla VIPR2 Gene -1238 CGGGGGGTGGGGGATCAGAGAGAAGCAAATGATAAGCCCAATGAGGTAAA -1189

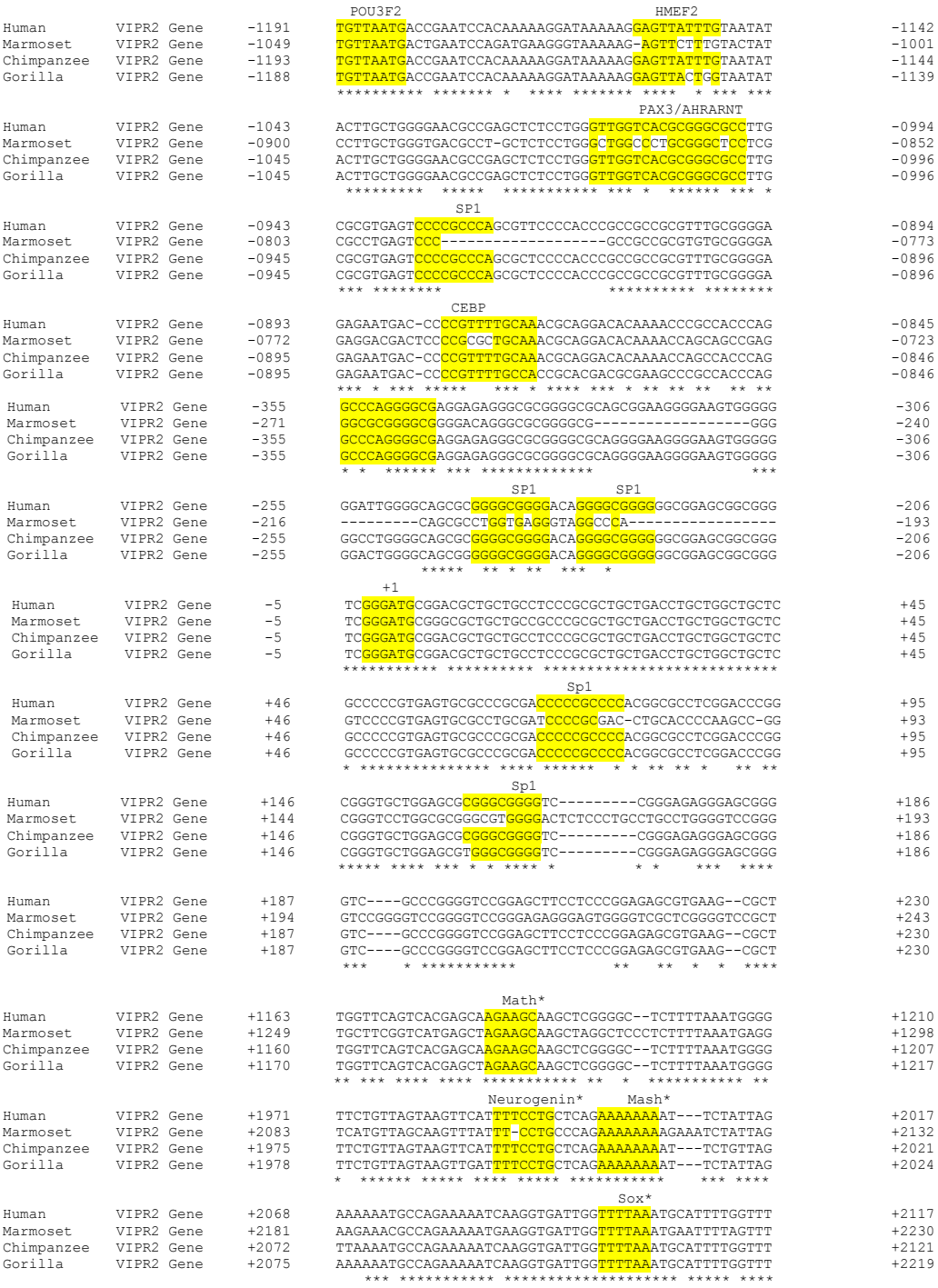
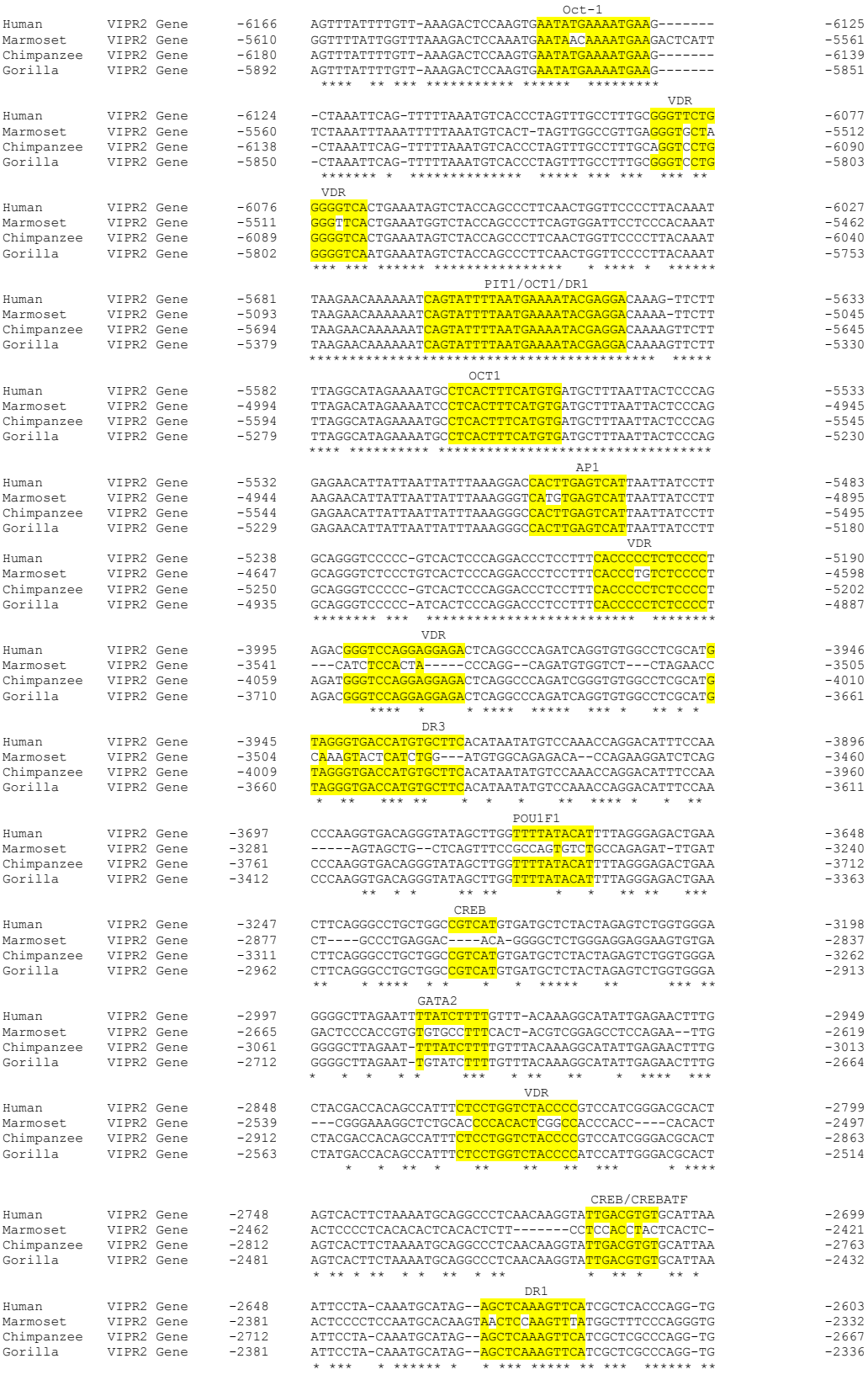


Figure 3.16 Evolutionarily conserved elements important in neurons.

*Elements found to be conserved among Rhesus monkey, cow, dog and cat *VIPR2* genes, identified by Evoprinter.


Oct-1
Human VIPR2 Gene -6166 AGTTTATTTTGTT-AAAGACTCCAAGTGAATATGAAAATGAAG------- -6125
Marmoset VIPR2 Gene -5610 GGTTTTATTGGTTTAAAGACTCCAAATGAATAACAAAATGAAGACTCATT -5561
Chimpanzee VIPR2 Gene -6180 AGTTTATTTTGTT-AAAGACTCCAAGTGAATATGAAAATGAAG------- -6139
Gorilla VIPR2 Gene -5892 AGTTTATTTTGTT-AAAGACTCCAAGTGAATATGAAAATGAAG------- -5851
VDR
Human VIPR2 Gene -6124 -CTAAATTCAG-TTTTTAAATGTCACCCTAGTTTGCCTTTGCGGGTTCTG -6077
Marmoset VIPR2 Gene -5560 TCTAAATTTAAATTTTTAAATGTCACT-TAGTTGGCCGTTGAGGGTGCTA -5512
Chimpanzee VIPR2 Gene -6138 -CTAAATTCAG-TTTTTAAATGTCACCCTAGTTTGCCTTTGCAGGTCCTG -6090
Gorilla VIPR2 Gene -5850 -CTAAATTCAG-TTTTTAAATGTCACCCTAGTTTGCCTTTGCGGGTCCTG -5803
VDR
Human VIPR2 Gene -6076 GGGGTCACTGAAATAGTCTACCAGCCCTTCAACTGGTTCCCCTTACAAAT -6027
Marmoset VIPR2 Gene -5511 GGGTTCACTGAAATGGTCTACCAGCCCTTCAGTGGATTCCTCCCACAAAT -5462
Chimpanzee VIPR2 Gene -6089 GGGGTCACTGAAATAGTCTACCAGCCCTTCAACTGGTTCCCCTTACAAAT -6040
Gorilla VIPR2 Gene -5802 GGGGTCAATGAAATAGTCTACCAGCCCTTCAACTGGTTCCCCTTACAAAT -5753
PIT1/OCT1/DR1
Human VIPR2 Gene -5681 TAAGAACAAAAAATCAGTATTTTAATGAAAATACGAGGACAAAG-TTCTT -5633
Marmoset VIPR2 Gene -5093 TAAGAACAAAAAATCAGTATTTTAATGAAAATACGAGGACAAAA-TTCTT -5045
Chimpanzee VIPR2 Gene -5694 TAAGAACAAAAAATCAGTATTTTAATGAAAATACGAGGACAAAAGTTCTT -5645
Gorilla VIPR2 Gene -5379 TAAGAACAAAAAATCAGTATTTTAATGAAAATACGAGGACAAAAGTTCTT -5330
OCT1
Human VIPR2 Gene -5582 TTAGGCATAGAAAATGCCTCACTTTCATGTGATGCTTTAATTACTCCCAG -5533
Marmoset VIPR2 Gene -4994 TTAGACATAGAAAATCCCTCACTTTCATGTGATGCTTTAATTACTCCCAG -4945
Chimpanzee VIPR2 Gene -5594 TTAGGCATAGAAAATGCCTCACTTTCATGTGATGCTTTAATTACTCCCAG -5545
Gorilla VIPR2 Gene -5279 TTAGGCATAGAAAATGCCTCACTTTCATGTGATGCTTTAATTACTCCCAG -5230
AP1
Human VIPR2 Gene -5532 GAGAACATTATTAATTATTTAAAGGACCACTTGAGTCATTAATTATCCTT -5483
Marmoset VIPR2 Gene -4944 AAGAACATTATTAATTATTTAAAGGGTCATGTGAGTCATTAATTATCCTT -4895
Chimpanzee VIPR2 Gene -5544 GAGAACATTATTAATTATTTAAAGGGCCACTTGAGTCATTAATTATCCTT -5495
Gorilla VIPR2 Gene -5229 GAGAACATTATTAATTATTTAAAGGGCCACTTGAGTCATTAATTATCCTT -5180
VDR
Human VIPR2 Gene -5238 GCAGGGTCCCCC-GTCACTCCCAGGACCCTCCTTTCACCCCCTCTCCCCT -5190
Marmoset VIPR2 Gene -4647 GCAGGGTCTCCCTGTCACTCCCAGGACCCTCCTTTCACCCTGTCTCCCCT -4598
Chimpanzee VIPR2 Gene -5250 GCAGGGTCCCCC-GTCACTCCCAGGACCCTCCTTTCACCCCCTCTCCCCT -5202
Gorilla VIPR2 Gene -4935 GCAGGGTCCCCC-ATCACTCCCAGGACCCTCCTTTCACCCCCTCTCCCCT -4887
VDR
Human VIPR2 Gene -3995 AGACGGGTCCAGGAGGAGACTCAGGCCCAGATCAGGTGTGGCCTCGCATG -3946
Marmoset VIPR2 Gene -3541 ---CATCTCCACTA-----CCCAGG--CAGATGTGGTCT---CTAGAACC -3505
Chimpanzee VIPR2 Gene -4059 AGATGGGTCCAGGAGGAGACTCAGGCCCAGATCGGGTGTGGCCTCGCATG -4010
Gorilla VIPR2 Gene -3710 AGACGGGTCCAGGAGGAGACTCAGGCCCAGATCAGGTGTGGCCTCGCATG -3661
DR3
Human VIPR2 Gene -3945 TAGGGTGACCATGTGCTTCACATAATATGTCCAAACCAGGACATTTCCAA -3896
Marmoset VIPR2 Gene -3504 CAAAGTACTCATCTGG---ATGTGGCAGAGACA--CCAGAAGGATCTCAG -3460
Chimpanzee VIPR2 Gene -4009 TAGGGTGACCATGTGCTTCACATAATATGTCCAAACCAGGACATTTCCAA -3960
Gorilla VIPR2 Gene -3660 TAGGGTGACCATGTGCTTCACATAATATGTCCAAACCAGGACATTTCCAA -3611
POU1F1
Human VIPR2 Gene -3697 CCCAAGGTGACAGGGTATAGCTTGGTTTTATACATTTTAGGGAGACTGAA -3648
Marmoset VIPR2 Gene -3281 -----AGTAGCTG--CTCAGTTTCCGCCAGTGTCTGCCAGAGAT-TTGAT -3240
Chimpanzee VIPR2 Gene -3761 CCCAAGGTGACAGGGTATAGCTTGGTTTTATACATTTTAGGGAGACTGAA -3712
Gorilla VIPR2 Gene -3412 CCCAAGGTGACAGGGTATAGCTTGGTTTTATACATTTTAGGGAGACTGAA -3363
CREB
Human VIPR2 Gene -3247 CTTCAGGGCCTGCTGGCCGTCATGTGATGCTCTACTAGAGTCTGGTGGGA -3198
Marmoset VIPR2 Gene -2877 CT----GCCCTGAGGAC----ACA-GGGGCTCTGGGAGGAGGAAGTGTGA -2837
Chimpanzee VIPR2 Gene -3311 CTTCAGGGCCTGCTGGCCGTCATGTGATGCTCTACTAGAGTCTGGTGGGA -3262
Gorilla VIPR2 Gene -2962 CTTCAGGGCCTGCTGGCCGTCATGTGATGCTCTACTAGAGTCTGGTGGGA -2913
GATA2
Human VIPR2 Gene -2997 GGGGCTTAGAATTTTATCTTTTGTTT-ACAAAGGCATATTGAGAACTTTG -2949
Marmoset VIPR2 Gene -2665 GACTCCCACCGTGTGTGCCTTTCACT-ACGTCGGAGCCTCCAGAA--TTG -2619
Chimpanzee VIPR2 Gene -3061 GGGGCTTAGAAT-TTTATCTTTTGTTTACAAAGGCATATTGAGAACTTTG -3013
Gorilla VIPR2 Gene -2712 GGGGCTTAGAAT-TGTATCTTTTGTTTACAAAGGCATATTGAGAACTTTG -2664
VDR
Human VIPR2 Gene -2848 CTACGACCACAGCCATTTCTCCTGGTCTACCCCGTCCATCGGGACGCACT -2799
Marmoset VIPR2 Gene -2539 ---CGGGAAAGGCTCTGCACCCCACACTCGGCCACCCACC----CACACT -2497
Chimpanzee VIPR2 Gene -2912 CTACGACCACAGCCATTTCTCCTGGTCTACCCCGTCCATCGGGACGCACT -2863
Gorilla VIPR2 Gene -2563 CTATGACCACAGCCATTTCTCCTGGTCTACCCCATCCATTGGGACGCACT -2514
CREB/CREBATF
Human VIPR2 Gene -2748 AGTCACTTCTAAAATGCAGGCCCTCAACAAGGTATTGACGTGTGCATTAA -2699
Marmoset VIPR2 Gene -2462 ACTCCCCTCACACACTCACACTCTT--------CCTCCACCTACTCACTC- -2421
Chimpanzee VIPR2 Gene -2812 AGTCACTTCTAAAATGCAGGCCCTCAACAAGGTATTGACGTGTGCATTAA -2763
Gorilla VIPR2 Gene -2481 AGTCACTTCTAAAATGCAGGCCCTCAACAAGGTATTGACGTGTGCATTAA -2432
DR1
Human VIPR2 Gene -2648 ATTCCTA-CAAATGCATAG--AGCTCAAAGTTCATCGCTCACCCAGG-TG -2603
Marmoset VIPR2 Gene -2381 ACTCCCCTCCAATGCACAAGTAACTCCAAGTTTATGGCTTTCCCAGGGTG -2332
Chimpanzee VIPR2 Gene -2712 ATTCCTA-CAAATGCATAG--AGCTCAAAGTTCATCGCTCGCCCAGG-TG -2667
Gorilla VIPR2 Gene -2381 ATTCCTA-CAAATGCATAG--AGCTCAAAGTTCATCGCTCGCCCAGG-TG -2336

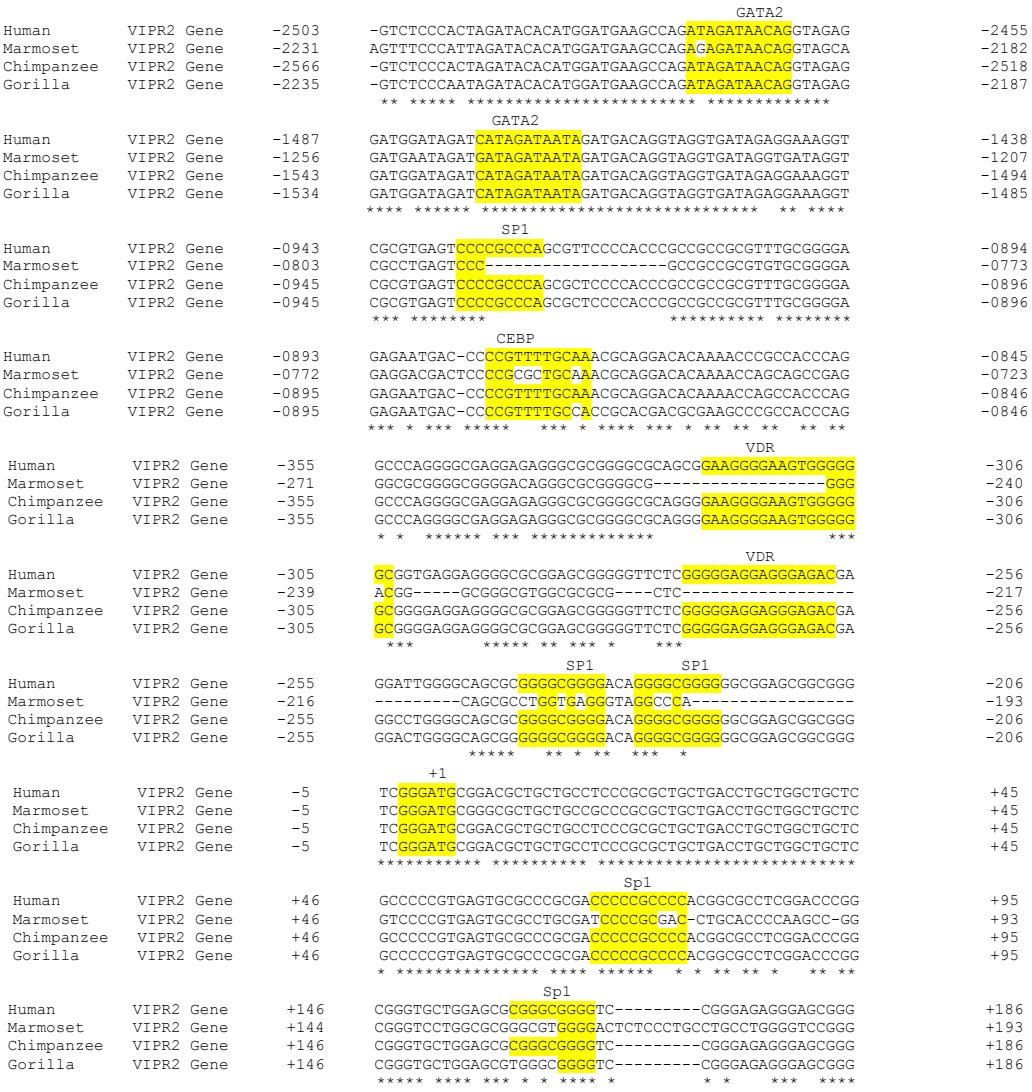


Figure 3.17 Evolutionarily conserved elements important in pituitary cells.

binding sites. This allowed us to identify putative and functional tissue specific transcription factor binding sites and minimal promoter elements. The results of Multi-zPicture show that many of the elements identified in the regions conserved evolutionarily across the species (Figure 3.5-3.10). The results are consistent with the CLUSTALW alignments in which the MATCH/TRANSFAC motifs are highlighted**.**

Using MATCH prevents inherent bias towards short recognition sequences that occur usually in other transcription factor binding site softwares, avoids subjective interpretation such as personal discretion, length and degeneracy issues.We were also able to demonstrate that some of the identified elements were more likely to be functional by using phylogenetic analysis. This allowed us to compare the identified elements in closely related species, bearing in mind 'that functional elements known to be conserved across the species, undergo slower sequence change through time, and exhibit greater constraints than non-functional elements' (Ganley and Kobayashi, 2007). However, computer identification of these elements requires extensive experimental work in a range of cell types to test which of them are truly functional.

Chapter-4

Making *VIPR2* Promoter Constructs

4.1 Introduction

4.1.1 Reporter Genes and their use

A reporter gene is a gene for which a phenotype can be readily measured and easily distinguishable from endogenous background proteins (Alam and Cook 1990). The criteria on which reporter genes are selected are based on sensitivity, convenience and reliability. In a reporter gene assay, cis-regulatory elements control the reporter gene; the cis-regulatory element usually is responsive to changes in gene regulation and expression in host cells. External factors like hormones and growth factors have been shown to activate target cells by activating the secondary messenger pathways which in turn activate the nuclear factors and consequently cis-regulatory elements and thereby alter gene transcription of the reporter gene.

To study different pathways, their interactions, their effects and functional significance of response element on gene expression, specific response elements are integrated into upstream regions controlling expression of the reporter genes. Activation of the response element by appropriate pathways, followed by alteration of the reporter gene expression allows monitoring the effects of pathways on gene expression.

Selection of the right reporter depends on the cell line used, type of experiment, and adaptability of assay to the appropriate detection method [(Wood, 1995) and (Suto and Ignar, 1997)] (Table: 4.1).

Reporter gene	Advantages	Disadvantages
Chloramphenicol acetyltransferase (CAT)	No endogenous activity. Automated ELISA available.	Narrow linear range; use of radioisotopes; stable.
b-Galactosidase (bacterial)	Well characterised; stable; simple colorimetric readouts; sensitive bio- or chemi-luminescent assays available.	Endogenous activity (mammalian cells
Luciferase (firefly)	High specific activity; no endogenous activity; broad dynamic range; convenient assays.	Requires substrate (luciferin) and presence of O2 and ATP.
Luciferase (bacterial)	Good for measuring/analysing prokaryotic gene transcription.	Less sensitive than firefly; not suitable for mammalian cells.
Alkaline phosphatise (human placenta)	Secreted protein; inexpensive colorimetric and highly sensitive luminescent assays available.	Endogenous activity in some cells; interference with compounds being screened.
Green fluorescent protein (GFP)	Auto-fluorescent (no substrate needed); no endogenous activity; mutants with altered spectral qualities available.	Requires post-translational modification; low sensitivity (no signal amplification).

Table 4.1 Reporter genes (Adapted from Naylor 1999)

Chloramphenicol acetyltransferase (CAT) gene was the first reporter to be used for functional studies (Bronstein *et al.,* 1994); it is the bacterial enzyme that can detoxify a protein synthesis inhibitor known as chloramphenicol. CAT is the stable enzyme and the endogenous expression in mammalian cells is nil. The linearity limit and sensitivity of the assay is narrow, when compared to other reporters (Bronstein *et al*., 1994 and Pazzagli *et al*., 1992).

The next most commonly used assay system is the beta-galactosidase based reporter assay system, because it uses a well characterised bacterial enzyme Beta-galactosidase. This assay system is used to monitor efficiency of tranfection and considered the favourite reporter assay system because of its user-friendliness and non-involvement of radioisotopes. The only disadvantage is that this enzyme has endogenous activity in mammalian cells, though its endogenous activity can be reduced by heat inactivation of cell extracts or increasing the pH of the reaction buffer (Young *et al*.,1993).

The other popular reporter is Luciferase which refers to group of enzymes that catalyse the oxidation of various substrates like luciferin and coelenterazine, resulting in emission of light. The widely used reporter assays are the ones which uses the enzyme such as firefly luciferase, heat labile bacterial luciferase and very recently Renilla luciferase. The linearity limit of bacterial luciferase based reporter assay system is low (only 3 orders of magnitude) (Pazzagli et al. 1992 and Manen *et al*., 1997) but the firefly luciferase based assay system is highly sensitive with broad linear range and therefore, the most commonly used reporter assay system [(Joyeux *et al*., 1997) and (Welsh and Kay 1997) (Table 4.1).

The regular firefly luciferase based assay was modified into user-friendly assay by removing the cell disruption step and by introducing membrane permeable, photolysable firefly luciferin esters. Further modification is done to the assay by incorporating the newly developed glow reagents, these modifications, increased the duration and stability of the flash response that can be detected in a scintillation counter and making it popular assay system in high throughput screening. Renillae luciferase based assay system is very appropriate for intact living system, because it catalyses the oxidation of coelentrazine, which is membrane permeable and there is no endogenous activity in mammalian cells like other luciferase based reporter assay systems.

The next reporter gene system is SEAP (Secretable form of alkaline phosphatase) which is mutated form of the alkaline phosphatase. SEAP reporter system has the advantage over luciferase systems, as the protein is secreted from the cell and can be detected by sampling the culture medium. The cells remain intact and viable for downstream steps [(Suto and Ignar, 1997) and (Jones *et al*., 1991)].

The reporter gene systems mentioned above can be used in combination. SEAP with luciferase and beta-galactosidase to normalise the transfection efficiency, two different luciferases such as firefly and renilla luciferases used in combination for taking multiple readouts from a single well and for dual detection of gene transcription. The choosing of reporter assay depends on the type of experiement planned, such as firefly and renilla luciferases based reporter assay systems suitable for kinetic and transfection efficiency experiments as its half lives are short compared to CAT in mammalian cells and also ideal for experiments which involve

analysis of cis-acting elements. Therefore, in this study we used firefly and renilla luciferases based reporter assay system for analysing the activity of cis-acting elements such as promoters and enhancers in the upstream regions of genes (Pazzagli *et al*., 1992).

In this chapter, the *VIPR2* promoter constructs are described. The plan was to transfect into two different cell lines: AtT 20 (a clonal pituitary cell line), and T98g (a partially transformed human glioblastoma) for functional analysis, but this was not achieved in this project.

4.2 Results

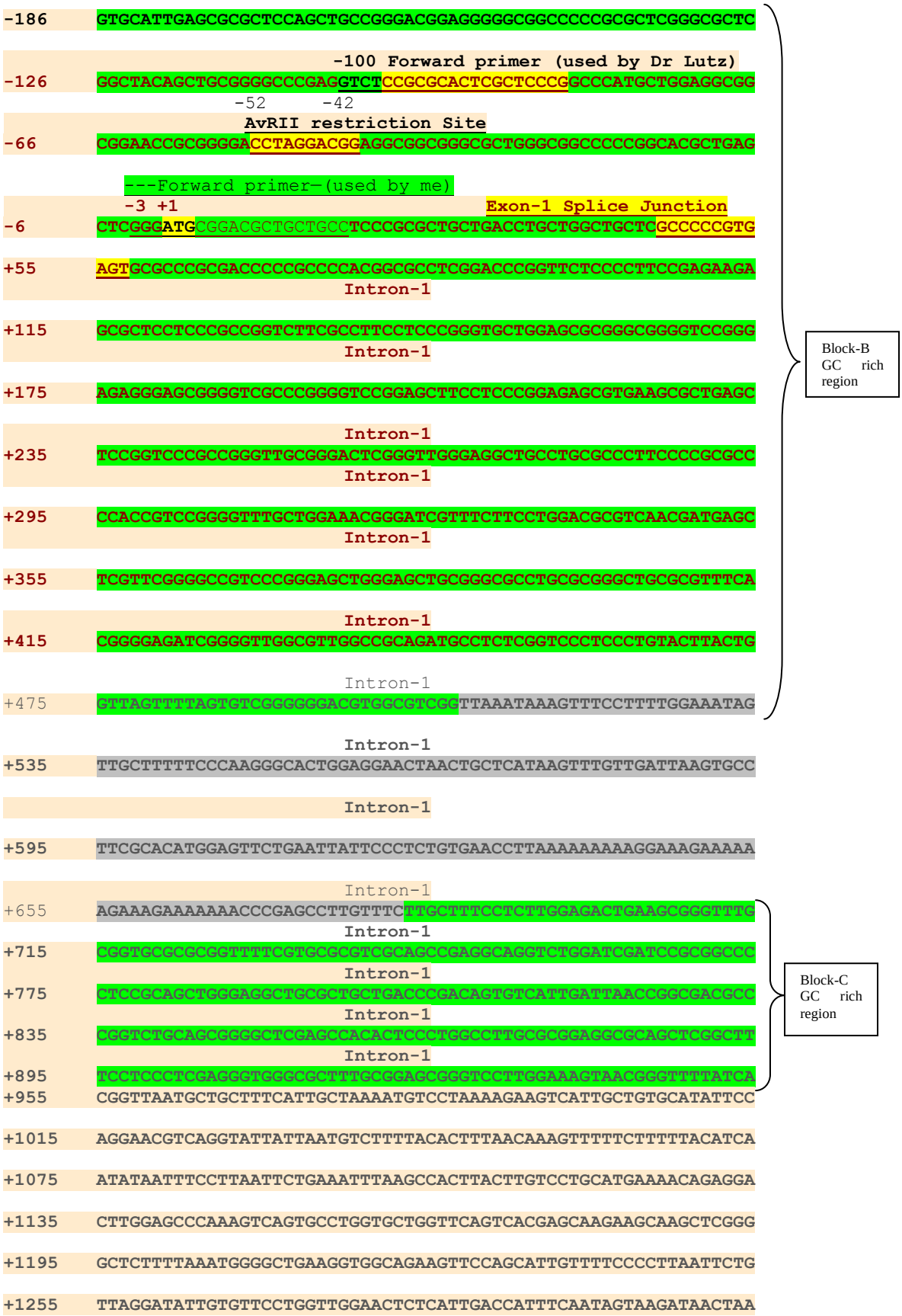
4.2.1 Amplification of Intron 1 of the VIPR2 gene from Cosmid 66e9 by PCR

4.2.1.1 Routine PCR

In order to make a pGL3 construct containing the exon 1- intron 1 sequence of the VIPR2 gene from -3 to +2234 relative to the previously identified ATG translation start site, the polymerase chain reaction (PCR) was used to amplify the 2.3 kb Intron -1 sequence from cosmid 66e9 using KOD polymerase from Novagen (Figure: 4.1).

The forward primer (185) used is 5’ GGG ATG CGG ACG CTG CTG CC 3’ and the reverse primer (2439) used is 5’ GGG TGA GAT CTG TTC ACC TGT TCA 3’. The annealing temperature used was 60ºC and the reaction was run for 25 cycles.

The results in Figure 4.2 show that the expected 2.3 kb PCR amplification product was not obtained, and several non-specific bands were present. The PCR was repeated using with a higher annealing temperature of 65ºC for 120 sec with 25 cycles. Again the PCR did not yield the 2.3 kb band but generated several non-specific bands (Figure 4.3).


-186
GTGCATTGAGCGCGCTCCAGCTGCCGGGACGGAGGGGGCGGCCCCCGCGCTCGGGCGCTC
-100 Forward primer (used by Dr Lutz)
-126
GGCTACAGCTGCGGGGCCCGAGGTCTCCGCGCACTCGCTCCCGGCCCATGCTGGAGGCGG
-52 -42
AvRII restriction Site
-66
CGGAACCGCGGGGACCTAGGACGGAGGCGGCGGGCGCTGGGCGGCCCCCGGCACGCTGAG
---Forward primer--(used by me)
-3 +1
Exon-1 Splice Junction
-6
CTCGGGATGCGGACGCTGCTGCCTCCCGCGCTGCTGACCTGCTGGCTGCTCGCCCCCGTG
+55
AGTGCGCCCGCGACCCCCGCCCCACGGCGCCTCGGACCCGGTTCTCCCCTTCCGAGAAGA
Intron-1
+115
GCGCTCCTCCCGCCGGTCTTCGCCTTCCTCCCGGGTGCTGGAGCGCGGGCGGGGTCCGGG
Intron-1
+175
AGAGGGAGCGGGGTCGCCCGGGGTCCGGAGCTTCCTCCCGGAGAGCGTGAAGCGCTGAGC
Intron-1
+235
TCCGGTCCCGCCGGGTTGCGGGACTCGGGTTGGGAGGCTGCCTGCGCCCTTCCCCGCGCC
Intron-1
+295
CCACCGTCCGGGGTTTGCTGGAAACGGGATCGTTTCTTCCTGGACGCGTCAACGATGAGC
Intron-1
+355
TCGTTCGGGGCCGTCCCGGGAGCTGGGAGCTGCGGGCGCCTGCGCGGGCTGCGCGTTTCA
Intron-1
+415
CGGGGAGATCGGGGTTGGCGTTGGCCGCAGATGCCTCTCGGTCCCTCCCTGTACTTACTG
Intron-1
+475
GTTAGTTTTAGTGTCGGGGGGACGTGGCGTCGGTTAAATAAAGTTTCCTTTTGGAAATAG
Intron-1
+535
TTGCTTTTTCCCAAGGGCACTGGAGGAACTAACTGCTCATAAGTTTGTTGATTAAGTGCC
Intron-1
+595
TTCGCACATGGAGTTCTGAATTATTCCCTCTGTGAACCTTAAAAAAAAAGGAAAGAAAAA
Intron-1
+655
AGAAAGAAAAAAACCCGAGCCTTGTTTCTTGCTTTCCTCTTGGAGACTGAAGCGGGTTTG
Intron-1
+715
CGGTGCGCGCGGTTTTCGTGCGCGTCGCAGCCGAGGCAGGTCTGGATCGATCCGCGGCCC
Intron-1
+775
CTCCGCAGCTGGGAGGCTGCGCTGCTGACCCGACAGTGTCATTGATTAACCGGCGACGCC
Intron-1
+835
CGGTCTGCAGCGGGGCTCGAGCCACACTCCCTGGCCTTGCGCGGAGGCGCAGCTCGGCTT
Intron-1
+895
TCCTCCCTCGAGGGTGGGCGCTTTGCGGAGCGGGTCCTTGGAAAGTAACGGGTTTTATCA
+955
CGGTTAATGCTGCTTTCATTGCTAAAATGTCCTAAAAGAAGTCATTGCTGTGCATATTCC
+1015
AGGAACGTCAGGTATTATTAATGTCTTTTACACTTTAACAAAGTTTTTCTTTTTACATCA
+1075
ATATAATTTCCTTAATTCTGAAATTTAAGCCACTTACTTGTCCTGCATGAAAACAGAGGA
+1135
CTTGGAGCCCAAAGTCAGTGCCTGGTGCTGGTTCAGTCACGAGCAAGAAGCAAGCTCGGG
+1195
GCTCTTTTAAATGGGGCTGAAGGTGGCAGAAGTTCCAGCATTGTTTTCCCCTTAATTCTG
+1255
TTAGGATATTGTGTTCCTGGTTGGAACTCTCATTGACCATTTCAATAGTAAGATAACTAA
Block-B GC rich region
Block-C GC rich region

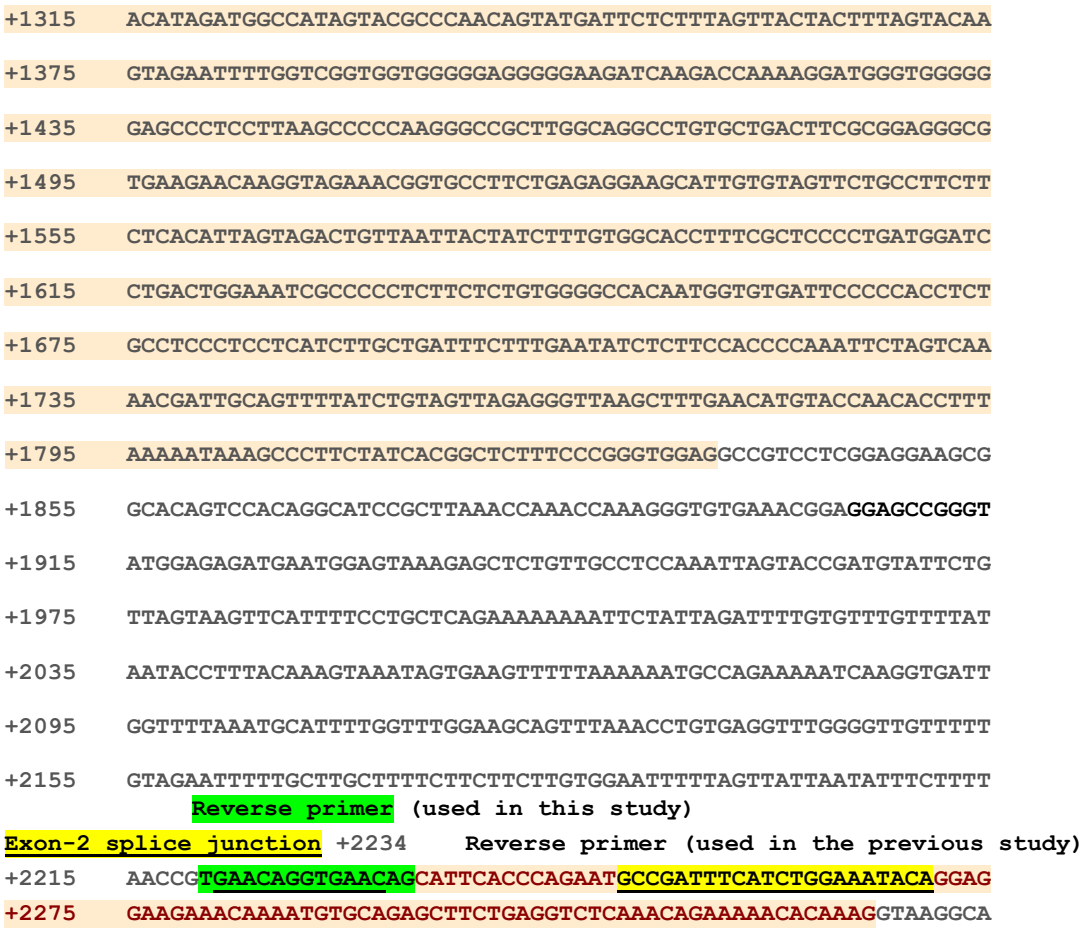


Figure 4.1 The nucleotide sequence of the 5' end of the human VIPR2 gene and primers used for amplification.

Bases relative to the ATG translation initiation codon (In boldface, designated +1) are numbered on the left. AvRII restriction site, exon-1 splice junction, Exon-2 splice junction, Intron-1, Forward primer, Reverse primer, CpG islands are indicated and underlined.

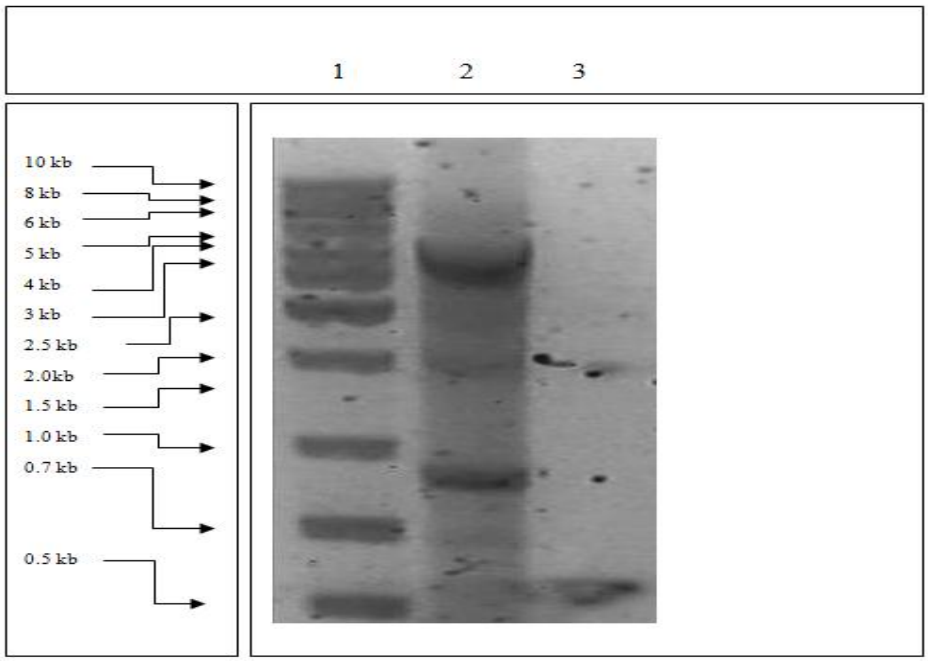


Figure 4.2 Gel electrophoretic analysis following the PCR amplification of 2.3 kb Intron-1 sequence from the VIPR2 Cosmid 66e9.

Analytical Conditions: Number of cycles: 25, Annealing temperature: 60°C. The expected 2.3 kb band was not obtained.

Legend

Lane-1: 1 kb Ladder (Axygen)
Lane-2: Test lane
Lane-3: Negative control

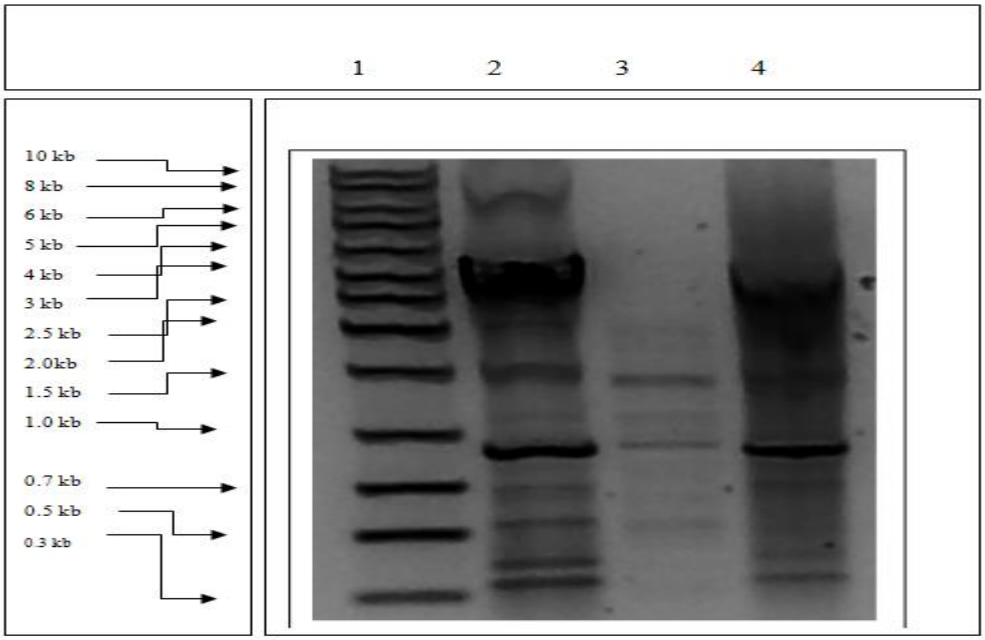


Figure 4.3 Gel electrophoretic analysis following the attempted PCR amplification of 2.3 kb Intron-1 sequence from the VIPR2 cosmid 66e9.

Analytical conditions: Number of cycles: 25, Annealing temperature: 65°C. The expected 2.3 kb band was not obtained

Legend

Lane-1: 1 kb Ladder (Axygen)
Lane-2: Test lane (Amplicon of Cosmid 66e9)
Lane-3: Negative control
Lane-4: Test lane (Amplification of Cosmid 66e9 from different aliquot)

Despite none of the strong bands being the correct size the PCR product was separated by electrophoresis and the region around 2.3 kb, was excised and extracted, and cloned into a TOPO vector, before being transformed into XL-1 blue competent cells. The resulting colonies were grown up and plasmid DNA was prepared using the miniprep protocol as described in materials and methods. The plasmid DNA was then digested with EcoRI. Electrophoresis was carried out on the digested DNA, but no bands of the correct 2.3 kb size were observed, indicating that the correct amplicon had not been achieved.

4.2.1.2 Touch Down PCR

Alternate strategies were tried. Touch down PCR was used to amplify Intron-1 from cosmid 66e9 without amplifying non-specific sequences. The annealing temperature used was 68.2^{o}C with a decrease in temperature (0.5 per cycle) to over 33 cycles (Table: 4.2, Figure 4.4). Non-specific bands were seen again, even after the addition of DMSO with new polymerase to the master mix.

4.2.1.3 GC-rich PCR system to amplify the Intron-1 from Cosmid 66e9

As touch down PCR amplification showed no evidence of giving the expected band size (2.3 kb), we strongly suspected the template was not appropriate. Exon-1 of the human *VIPR2* gene sits within a CpG island (GC content 63.9%) that extends to 897 bp into intron-1. Therefore the GC-rich PCR system (Roche) was tried to amplify the Intron-1 from Cosmid 66e9 (Table 4.3). The amplified product that was obtained was around 3 kb and not the expected 2.3 kb (Figure: 4.5).

A.Master Mix (Final Volume: 20µl); D*= Serially Diluted

Component	**undiluted (µl)D* Cosmid DNA**	**1/10 D*(µl) Cosmid DNA**	**1/100 D* (µl) Cosmid DNA**	**1/1000D D*(µl) Cosmid DNA**	**1/10000D D*(µl) Cosmid DNA**	**1/100000 D* (µl) Cosmid DNA**	**Positive control (µl)**	**Negative control (µl)**
PCR grade water	50	50	50	50	50	50	50	50
dNTP (0.2mM) each	5	5	5	5	5	5	5	5
5’ Primer	1	1	1	1	1	1	1	1
3’ Primer	1	1	1	1	1	1	1	1
10 X PCR buffer	5	5	5	5	5	5	5	5
MgSo4 (1mM)	2	2	2	2	2	2	2	2
Template DNA	5 (150ng/µl)	5	5	5	5	5	5	5
KOD DNA polymerase 0.02 U/ul	1	1	1	1	1	1	1	1

B. Touch down PCR protocol

Forward Primer used: 5’ GGG ATG CGG ACG CTG CTG CC 3’		
Reverse Primer used: 5’ GGG TGA GAT CTG TTC ACC TGT TCA 3’		
Stage: 1		
Step: 1	95ºC	2 min
Stage: 2		
Step: 2	95ºC	30 Sec
Step: 3	68.2 ºC (decrease 0.5 ºC per cycle)	30 Sec
Step: 4	72 ºC	3.6 Min
Repeat steps 2-4 (14 cycles)		
Stage: 3		
Step: 5	95ºC	30 Sec
Step: 6	61.2	30 Sec
Step: 7	72 ºC	220 Sec
Repeat steps 5-7 (19 cycles)		
Step: 8	72ºC	5 Min
Step: 9	4 ºC	Forever

Table 4. 2 Touch-down PCR protocol

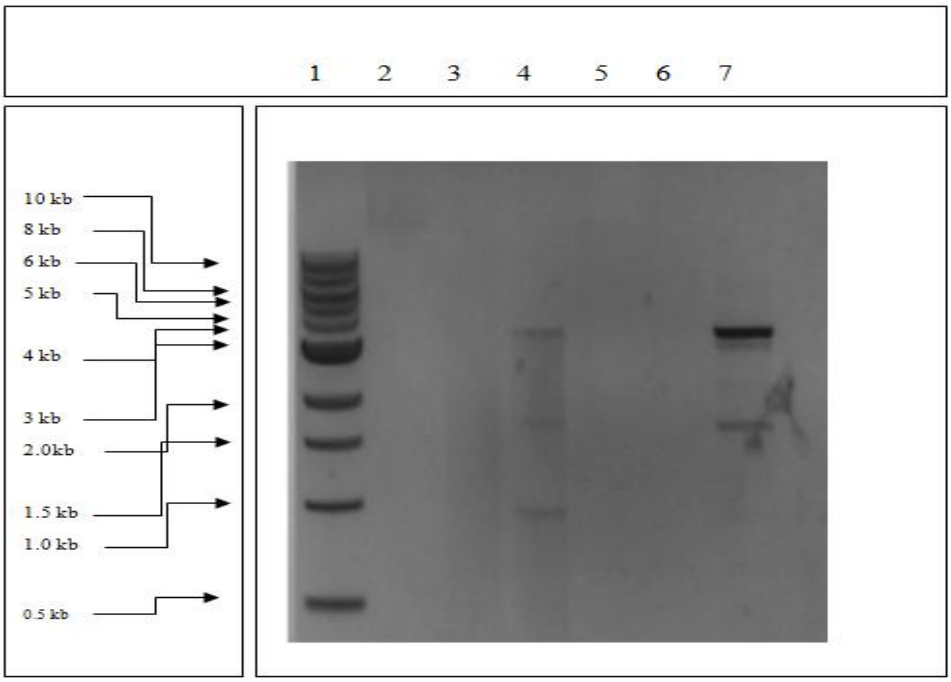


Figure 4.4 Gel electrophoretic analysis of the amplified products generated from serially diluted DNA templates by touchdown PCR

The amplicon was the result of the Touch down PCR amplification of 2.3 kb Intron-1 from Cosmid 66e9. The annealing temperature: 68.2C with a decrease in temperatures (0.5 per cycle to over 33 cycles)

Legend

Lane-1: 1 kb ladder (New England Biolabs)
Lane-2: Undiluted DNA template
Lane-3: 1/10 diluted DNA template
Lane-4: 1/100 diluted DNA template
Lane-5: 1/1000 diluted DNA template
Lane-6: 1/10000 diluted DNA template
Lane-7: 1/100000 diluted DNA template

Master mix-1 (Final Volume= 35 µl).

Master Mix	Resolution buffer (µl)	Water (µl)	dNTP (µl) (10 mM/µl each)	Primer: 185 (µl)	Primer: 2439 (µl)	Clone 1-1 (µl) (Conc:100ng/µl)
1	0.0	29.0	4.0	1.0	1.0	0.5
2	5.0	24.0	4.0	1.0	1.0	0.5
3	10.0	19.0	4.0	1.0	1.0	0.5
4	15.0	14.0	4.0	1.0	1.0	0.5
5	20.0	9.0	4.0	1.0	1.0	0.5
6	25.0	4.0	4.0	1.0	1.0	0.5

Master mix-2 (Final Volume= 15 µl).

Component	Volume	Final concentration
Water PCR grade	4µl	
5X GC-rich reaction buffer with DMSO	10 µl	1 X (1.5 mM MgCl2) 2% DMSO
GC rich PCR enzyme mix	1µl	2 U/50 µl

35 µl of Master mix-1 combined with 15 µl of Master mix-2 in thin walled PCR tube (final volume = 50 µl) and all the six samples were loaded onto the thermal block cycler and PCR was done following the protocol mentioned below:

Polymerase chain reaction protocol

Forward Primer used: 5’ GGG ATG CGG ACG CTG CTG CC 3’		
Reverse Primer used: 5’ GGG TGA GAT CTG TTC ACC TGT TCA 3’		
Stage: 1		
Step: 1	95ºC	3 min
Stage: 2		
Step: 2	94ºC	30 Sec
Step: 3	61ºC	30 Sec
Step: 4	68 ºC	4 Min
Repeat steps 2-4 (20 cycles)		
Stage: 3		
Step: 5	68ºC	15 Min
Step: 6	4 ºC	Forever

Table 4.3 GC rich PCR

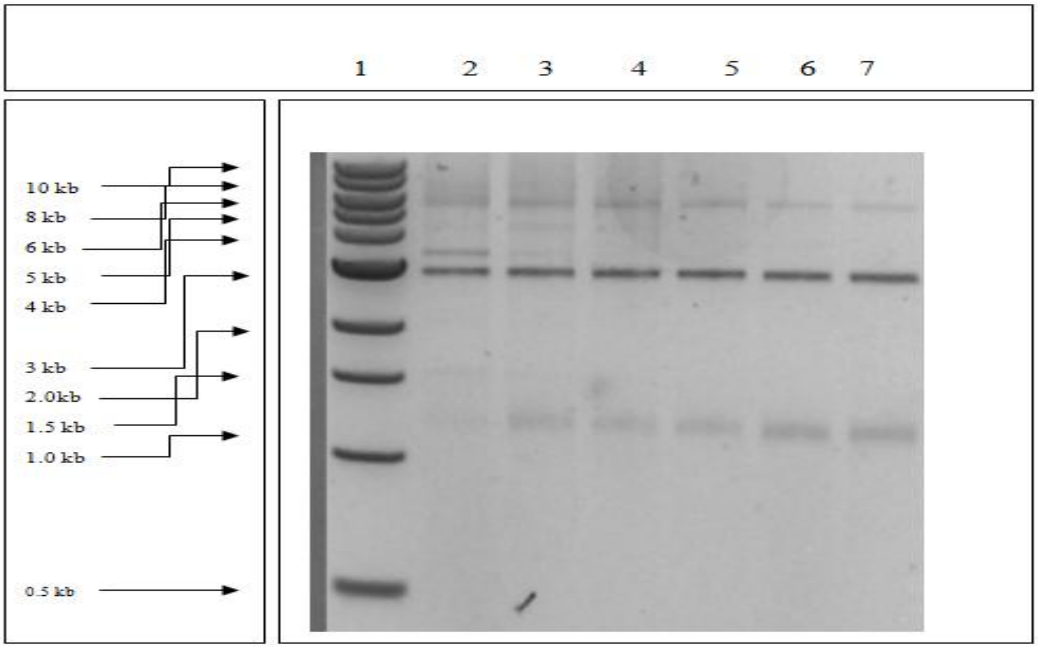


Figure 4.5 Gel electrophoretic analysis following the attempted PCR amplification of 2.3 kb Intron-1 sequence from the VIPR2 Cosmid 66e9 by GC-rich PCR system. The expected 2.3 kb amplicon was not present.

Legend

Lane-1: 1 kb Ladder (New England Biolabs)
Lane-2: GC-6
Lane-3: GC-5
Lane-4: GC-4
Lane-5: GC-3
Lane-6: GC-2
Lane-7: GC-1

4.2.1.4 Expand PCR system to amplify the intron-1 from cosmid 66e9

As the GC-rich PCR system failed to amplify the 2.3 kb of human *VIPR2* intron-1 from cosmid 66e9, the Expand-long-template PCR system was tried. This enzyme system was used to generate a longer DNA region containing a part of exon-1, intron 1 and a part of exon 2 previously in the laboratory (Dr Lutz, unpublished). This did not give expected band size, so it was strongly suspected that the primers may be the reason. Therefore a new reverse primer was designed, which is slightly longer than the previous one. However when this was used in place of the previous primer, we were not able to get the expected amplified product (Table: 4.4 and Figure: 4.6).

Master Mix (Final volume= 50 ul).

Component	Volume (ul)	P1 (ul)	Genomic DNA (ul)	Cosmid 66e9 (ul)	1-1 Intron (ul)	Positive control (ul)	Negative control (ul)	Final Concentration
PCR grade water	50							
dNTP (10 mM)	7							350 mM
Forward Primer	1							300 nM
Reverse Primer	1							300 nM
10 X PCR buffer -1 with MgCl2	5							MgCl2=1.75 mM
Template DNA		1 (200 ng/ul)	1 (200 ng/ul)	1 (150 ng/ul)	1 (100 ng/ul)	1 (100 ng/ul)	1 (sterile H2O)	
Expand enzyme mix	0.75							3.75U

All the six samples were loaded onto the thermal block cycler and PCR was done following the protocol mentioned below:

Polymerase chain reaction protocol

Forward Primer	5’ GGG ATG CGG ACG CTG CTG CC 3’	
Reverse Primer	5’ GGG TGA GAT CTG TTC ACC TGT TCA 3’	
Stage: 1		
Step: 1	94ºC	2 Min
Stage: 2		
Step: 2	94ºC	10 Secs
Step: 3	60ºC	30 Secs
Step: 4	68 ºC	15 Min
Stage: 3		
Step: 5	68ºC	30 Min
Step: 6	4 ºC	Forever
Details		
P1 clone: A human genomic library constructed in the bacteriophage P1 vector, which has the whole human VIPR2 gene. P1 Clone can hold insert size up to 100 kb) Positive control: Human TPA (tissue plasminogen activator) primers		

Table 4. 4 Expand PCR Protocol

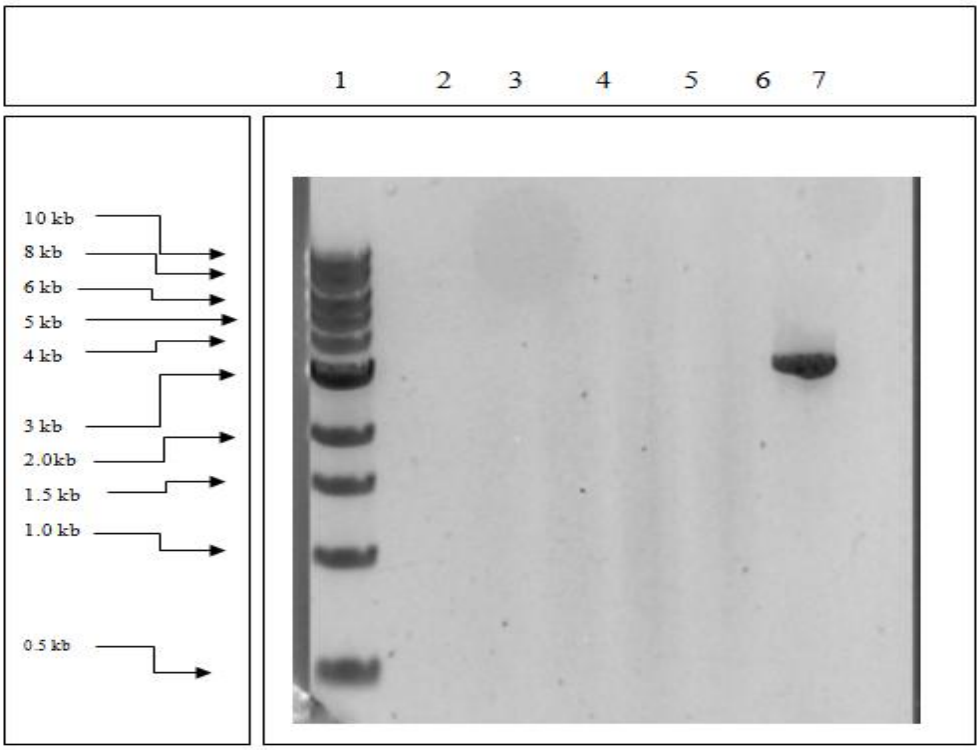


Figure 4.6 Gel electrophoretic analysis following the amplification of 2.3 kb Intron-1 sequence from the VIPR2 gene by Expand PCR system.

The expected 2.3 kb amplicon was not obtained.

Legend

Lane-1: 1 kb Ladder (New England Biolabs)
Lane-2: Negative control
Lane-3: P1 clone (Sternberg et al., 1990 and 1994)
Lane-4: Positive control
Lane-5: Genomic DNA
Lane-6: Cosmid 66e9
Lane-7: 1-1 Intron (which was cloned in pGEMT-easy vector) as template

4.2.2 Cloning strategies and Human VIPR2 Intron-1/pGL3 construct

In the previous sections, regular and touch down polymerase chain reactions (PCR) were tried to amplify the human *VIPR2* Intron-1 sequence from cosmid 66e9. Since the PCR amplification of Intron 1 did not work, it was decided to try to subclone the appropriate region from another cDNA construct made previously in the lab that contained exon 1, intron 1 and exon 2 regions subcloned into pGEM-T easy.

4.2.2.1 Sub-cloning the human VIPR2 Intron-1 into pGL3 vector using EcoRI-AvrII

This cloning strategy is outlined in Figure 4.8. In this, the human *VIPR2* Intron-1 sequence is excised from the pGEM-T easy vector (plasmid pGEMT-VIPR2, created in our lab by Dr Lutz) using the restriction enzymes by EcoRI digestion, followed by blunting the sticky ends of the insert (Intron-1) by extension using pfu polymerase. Then the 5' end of the insert is digested with AvrII and thereby converting the 5' blunt end of the insert into an AvrII sticky end.

pGL3 was digested with NheI and SmaI, and the prepared insert ligated in. The 5' AvrII end of the insert was ligated into the NheI digested end of the pGL3, and the 3' blunt end of the insert with the SmaI digested end of the pGL3 vector.

Figure 4.9 shows the successful release of the 2.3 kb Intron-1 sequence of human VIPR2 from the pGEM-T easy vector. Figure 4.10 shows the gel-purified EcoRI fragment following Pfu treatment. After ligating the fragment into the vector, and transformation in E. coli, 6 colonies were isolated and plasmid miniprep was

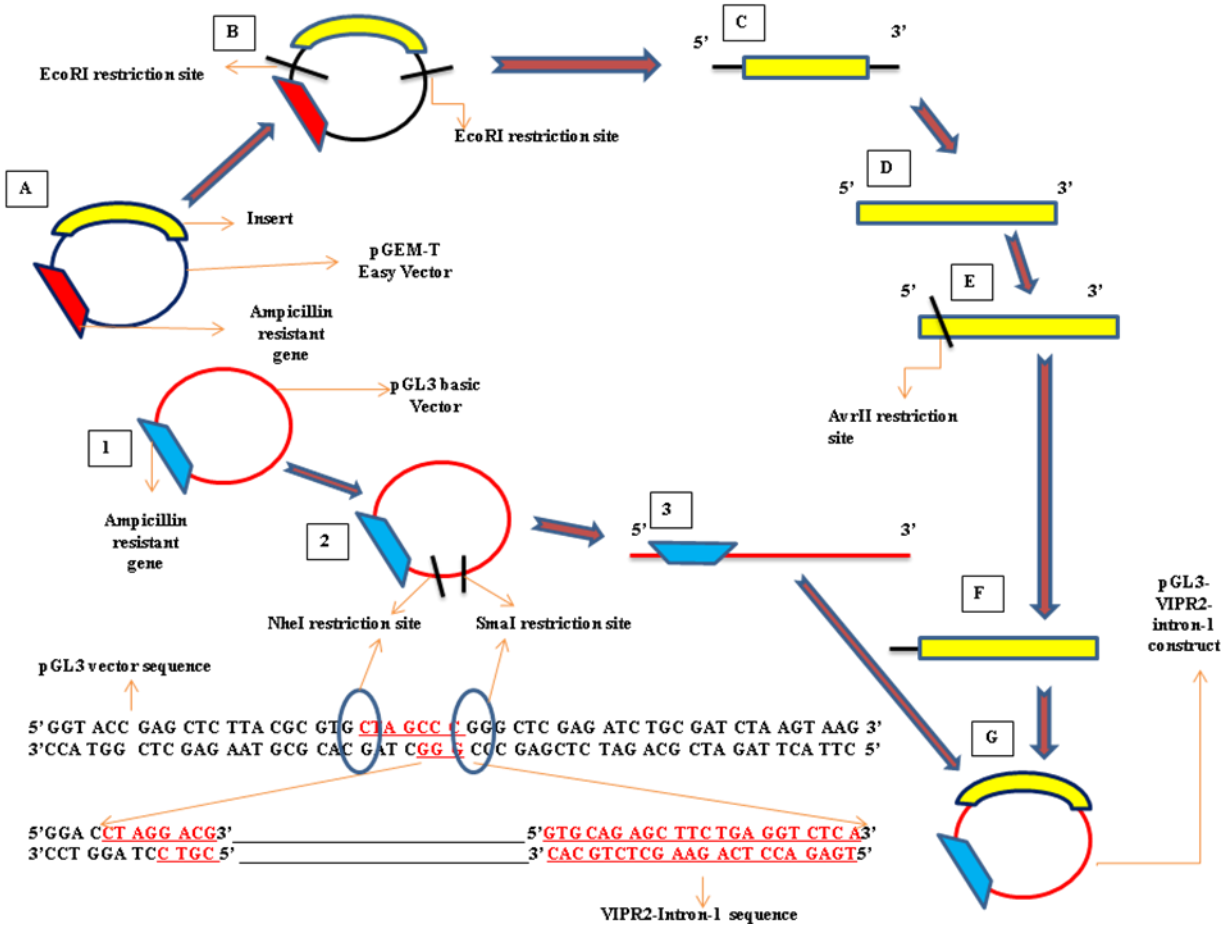


Figure 4.7 Cloning strategy: Subcloning the human VIPR2 Intron-1 insert into pGL3 basic vector using EcoRI/AvRI/NheI/SmaI

A. pGEMT-Easy vector with the human *VIPR2* Intron-1 as its insert (construct)
B. pGEMT-Easy/intron-1 construct with EcoRI restriction site
C. pGEMT-Easy/intron-1 construct linearized after EcoRI digestion producing a sticky end in both 5' and 3' regions
D. Both the sticky ends are converted into blunt end pGEMT-Easy/intron-1 by extension using pfu polymerase
E. Followed by AvrII digestion in the linearised EcoRI digested Insert Intron1
F. AvrII digest leaves a sticky end in 5' region.

1. pGL3 basic vector with ampicillin as selection marker
2. pGL3 basic vector is double digested with NheI and SmaI
3. pGL3 basic vector linearised after digestion leaving sticky end (NheI) in 5' and blunt end (SmaI) in 3' end.

G. Followed by ligation 5'sticky end of AvrII digested insert ligated with 5' sticky end of(NheI digested) pGL3 vector and 3' blunt end of (following extension) insert ligated with 3' blunt end f (SmaI digested) pGL3 vector.

Legend: Blue filled trapezoid (in pGL3 vector), Red filled trapezoid (in pGEM-T Easy vector) are selection markers (Ampicillin), yellow filled block arc is human VIPR2 Intron-1 insert, and black semi-vertical lines ones are restriction enzyme sites.

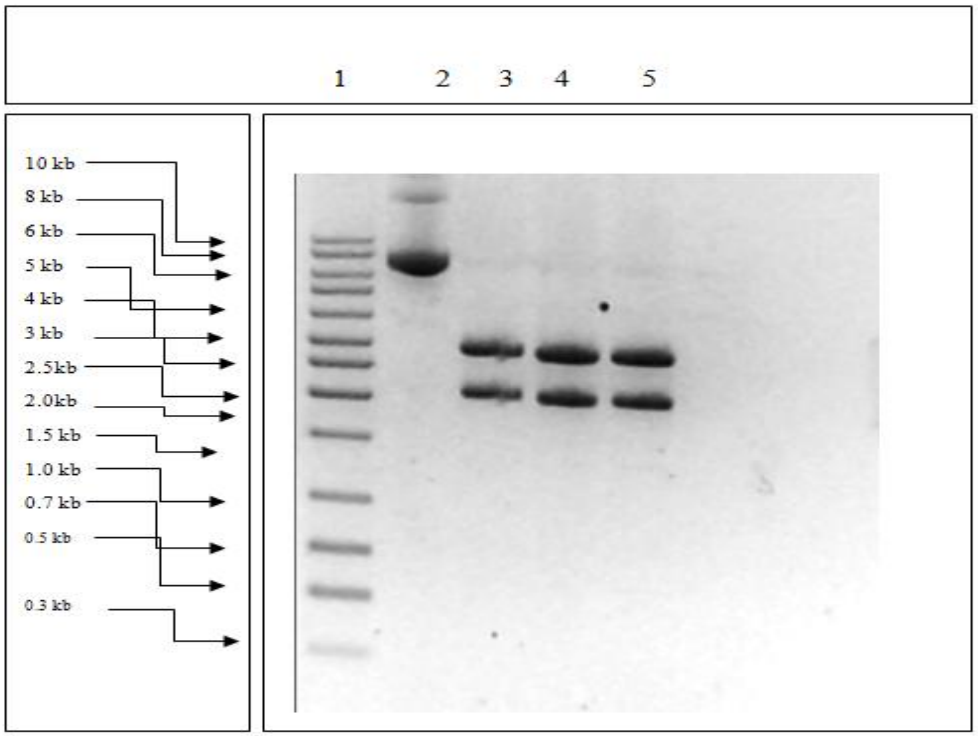


Figure 4.8 Gel electrophoretic analysis of EcoRI digestion pGEMT-VIPR2.

The 2.3kb intron 1 sequence of human VIPR2 is released from the plasmid vector

Legend

Lane-1: 1 kb Ladder (Axygen)
Lane-2: Undigested pGEMT-VIPR2
Lane-3: EcoRI digested pGEMT-VIPR2 (prep 1)
Lane-4: EcoRI digested pGEMT-VIPR2 (prep 2)
Lane-5: EcoRI digested pGEMT-VIPR2 (prep 3)

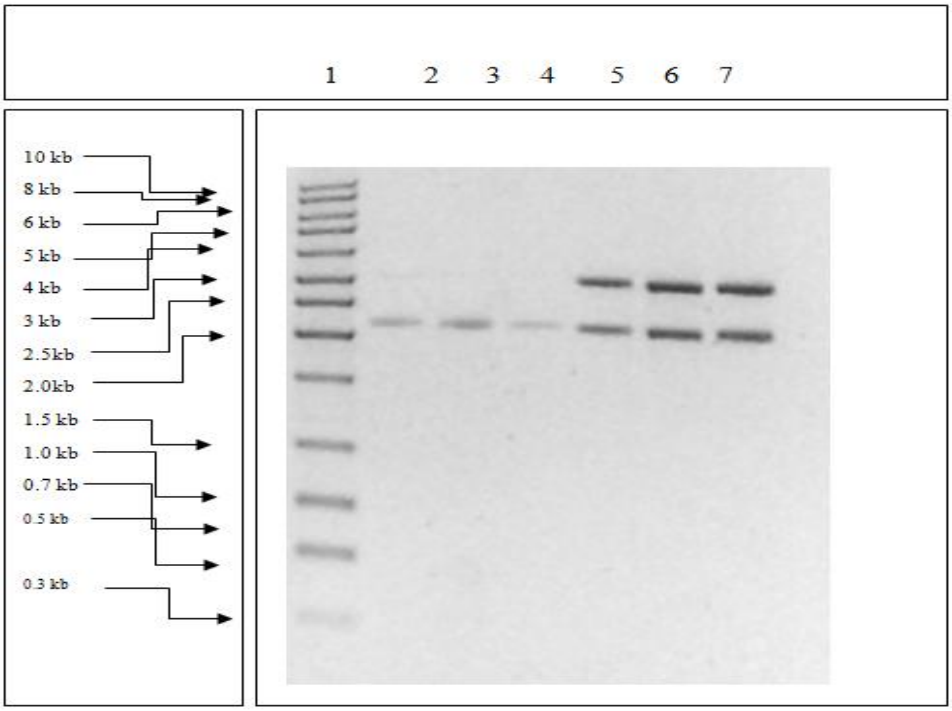


Figure 4.9 Gel electrophoretic analysis of Pfu extended EcoRI digested 2.3 kb Intron-1 Insert of the human VIPR2 gene.

Legend

Lane-1: 1 kb Ladder (Axygen)
Lane-2: 2.3 kb EcoRI fragment, Pfu treated (sample 1)
Lane-3: 2.3 kb EcoRI fragment, Pfu treated (sample 2)
Lane-4: 2.3 kb EcoRI fragment, Pfu treated (sample 3)
Lane-5: EcoRI digested pGEMT-VIPR2 (prep 1)
Lane-6: EcoRI digested pGEMT-VIPR2 (prep 2)
Lane-7: EcoRI digested pGEMT-VIPR2 (prep 3)

prepared. These minipreps were digested with HindIII to check for the insertion of the AvrII/EcoRI fragment into pGL3Basic. The resuls in Figure 4.12 show that the the 2.3kb has not been cloned into the pGL3-Basic vector. It is not clear what the problem is as the yield of miniprep plasmid is low, and the digest with HindIII did not show any band at the expected size.

To verify this result, a further set of digests was carried out using the restriction enzymes HindIII and BamHI. The results in Figure 4.13 show that there is no plasmid DNA remaining after digesting indicating that the yield of plasmid was very low or that the DNA has been degraded by a nuclease.

To prepare more fragment DNAs for ligation, a second digestion of pGEMT-VIPR2 with EcoRI was carried out (Figure 4.14). The digested DNA was pooled and then concentrated using ethanol precipitation before running on a gel and purifying (Figure 4.15). The 2.3kb fragment was then treated in the same way with pfu, and then digested with AvrII, before attempting to clone into pGL3-Basic. After transformation, colonies were isolated and DNA prepared before digestion with BamHI to check for inserts (Figure 4.16). However, no inserts were observed in any of the plasmid clones prepared.

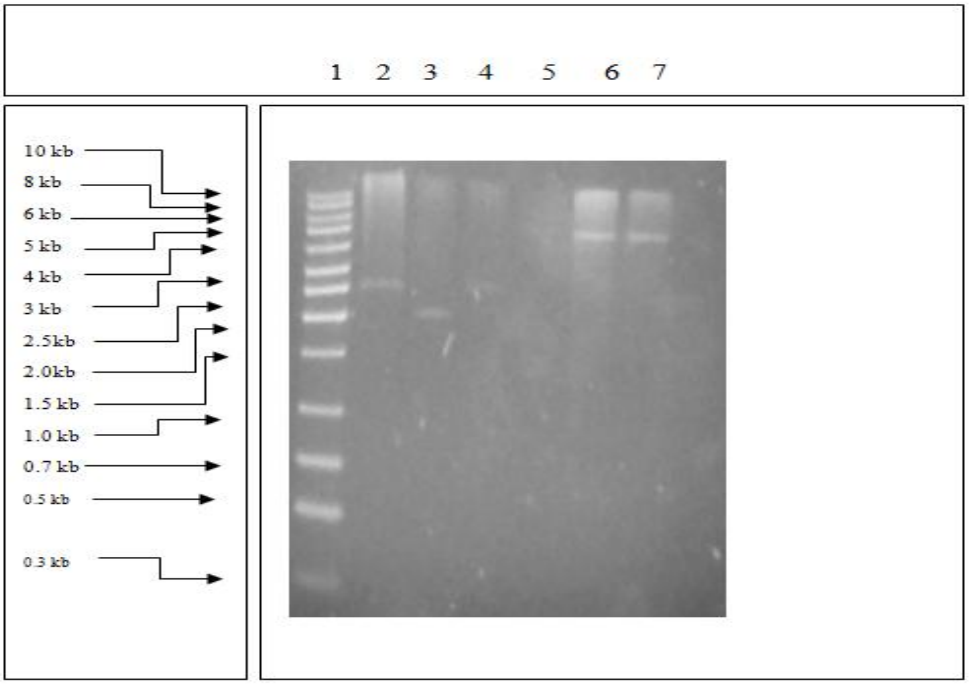


Figure 4.10 Gel electrophoresis analysis of HindIII digests of 6 clones to check whether the cloning of the AvRII/EcoRI 2.3 kb Intron-1 fragment with pGL3 basic vector was successful.

Legend

Lane-1: 1 kb Ladder (Axygen)
Lane-2: Clone-1 digested with HindIII
Lane-3: Clone-2 digested with HindIII
Lane-4: Clone-3 digested with HindIII
Lane-5: Clone-4 digested with HindIII
Lane-6: Clone-5 digested with HindIII
Lane-7: Clone-6 digested with HindIII

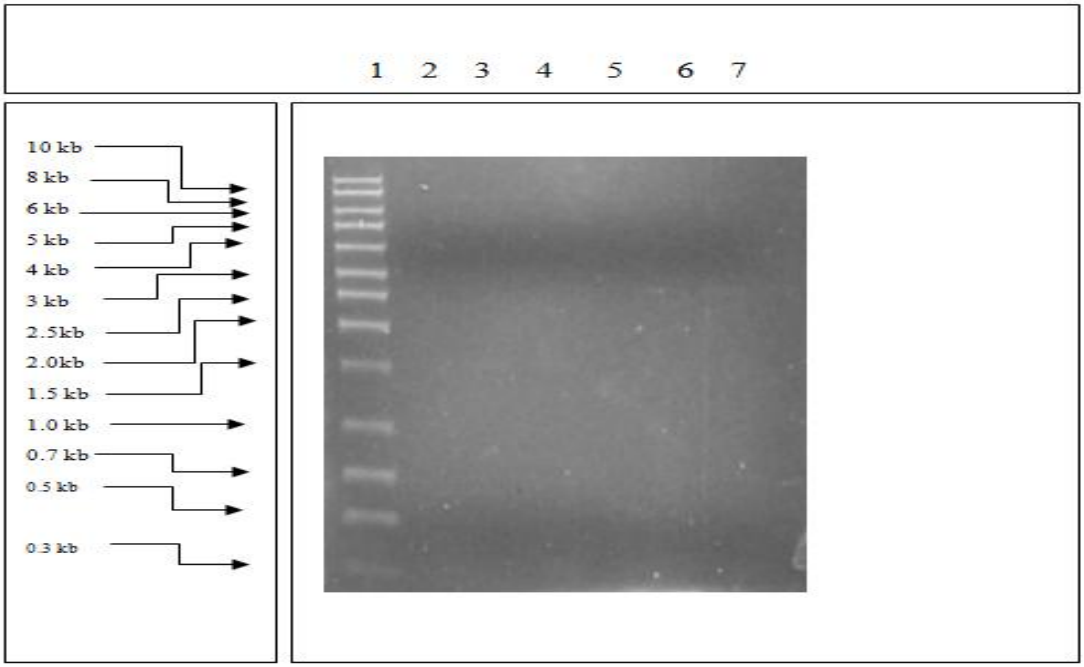


Figure 4.11 Gel electrophoretic analysis of HindIII and BamHI digests of 3 clones to check whether the cloning of the AvRII/EcoRI 2.3 kb Intron-1 fragment with the pGL3 basic vector was successful.

Legend

Lane-1: 1 kb Ladder (Axygen)
Lane-2: Clone-1 (HindIII)
Lane-3: Clone-2 (HindIII)
Lane-4: Clone-3 (HindIII)
Lane-5: Clone-1 (BamHI)
Lane-6: Clone-2 (BamHI)
Lane-7: Clone-3 (BamHI)

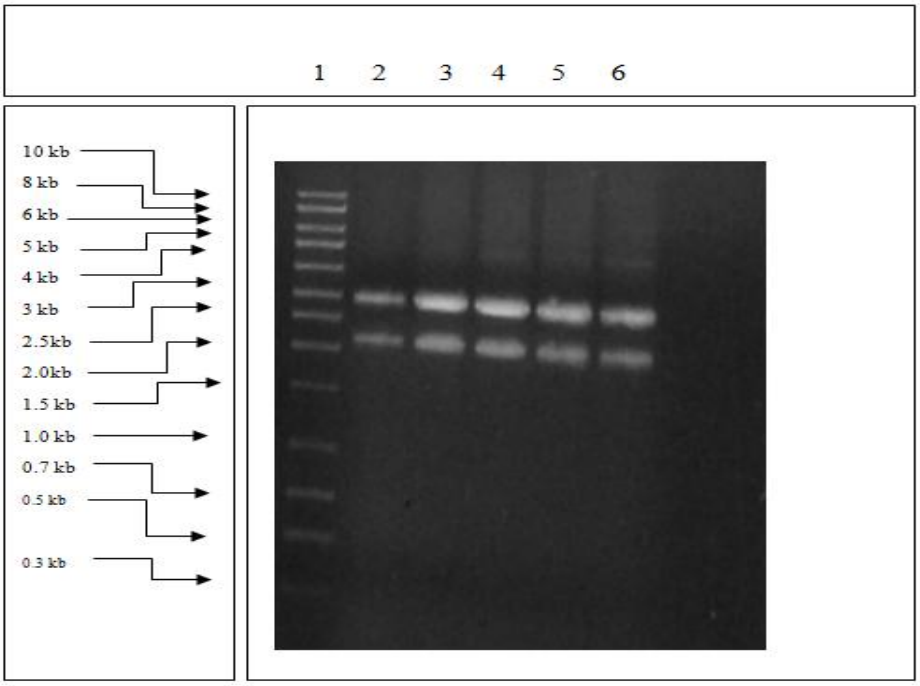


Figure 4.12 (Second cloning attempt) Gel electrophoretic analysis of EcoRI digestion of pGEMT-VIPR2

Legend

Lane-1: 1 kb Ladder (Axygen)
Lane-2: EcoRI-digest of pGEMT-VIPR2 (Prep 1)
Lane-3: EcoRI-digest of pGEMT-VIPR2 (Prep 2)
Lane-4: EcoRI-digest of pGEMT-VIPR2 (Prep 3)
Lane-5: EcoRI-digest of pGEMT-VIPR2 (Prep 4)
Lane-6: EcoRI-digest of pGEMT-VIPR2 (Prep 5)

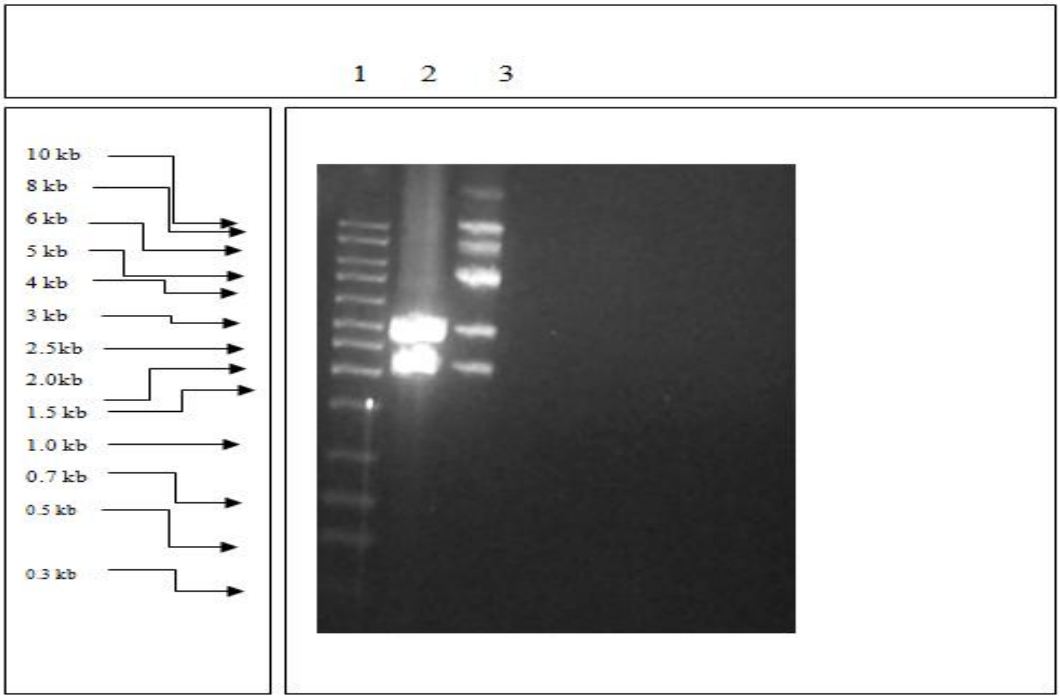


Figure 4.13 (Second Cloning attempt) Gel electrophoretic analysis of the ethanol precipitated EcoRI digest of pGEMT-VIPR2

Legend

Lane-1: 1 kb Ladder (Axygen)
Lane-2: EcoRI digested pGEMT-VIPR2 pool after ethanol precipitation

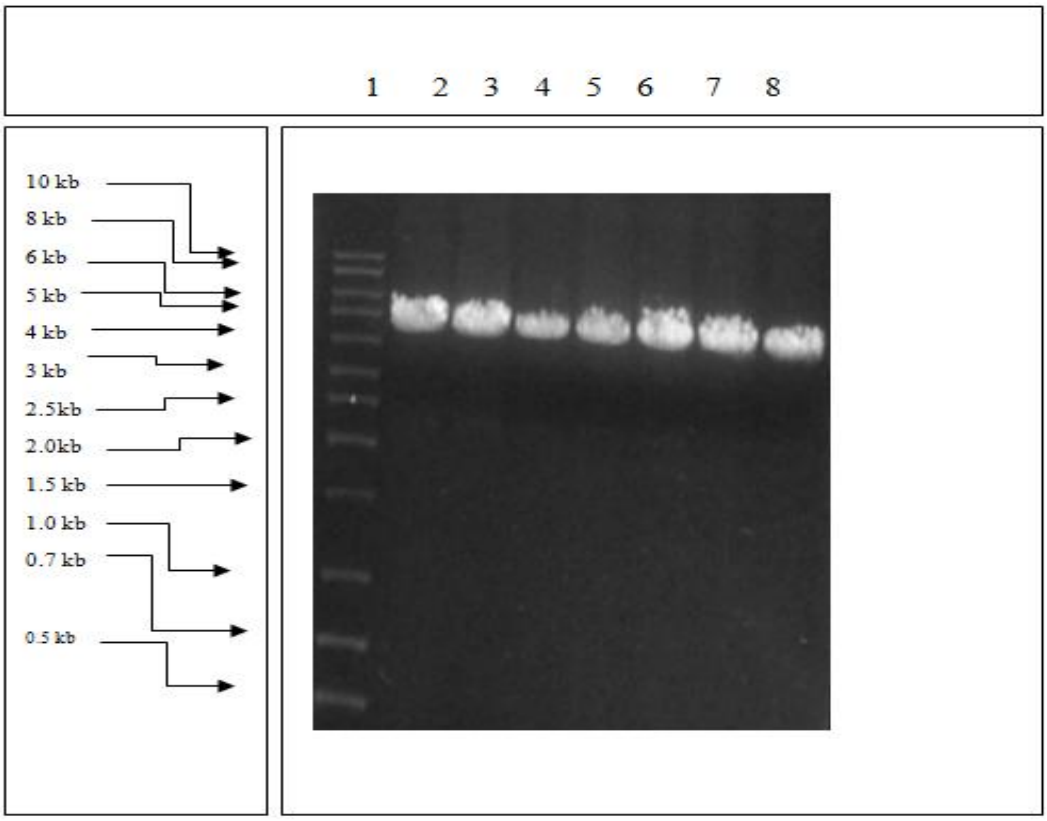


Figure 4.14 (Second cloning attempt) Gel electrophoretic analysis of BamHI digests of minipreped 7 clones to check whether the cloning of the AvRII/EcoRI 2.3 kb Intron-1 fragment with the pGL3 Basic vector was successful.

Legend

Lane-1: 1 kb Ladder (Axygen)
Lane-2: BamHI digest of clone 7
Lane-3: BamHI digest of clone 8
Lane-4: BamHI digest of clone 9
Lane-5: BamHI digest of clone 10
Lane-6: BamHI digest of clone 11
Lane-7: BamHI digest of clone 12
Lane-8: BamHI digest of clone 13

4.2.2.2 Sub-cloning the human *VIPR2* Intron-1 into pGL3 using BsgI-AvrII

As the original cloning strategy did not give any plasmids with the correct insert, an alternative strategy was followed. This is outlined in Figure 4.17. In this, the human VIPR2 Intron-1 sequence is excised from pGEMT-VIPR2 using the restriction BsgI (in the 3' end of the insert). The sticky end of BsgI digested part of the insert is made blunt by extension protocol using pfu polymerase, using PCR machine for 15 minutes at 68oC, followed by cleaning the product using microclean from Web Scientific Ltd. The enzyme AvrII is used to digest the 5' end of the insert. This fragment can then be cloned into the NheI/SmaI digested ends of pGL3 vector as in the previous strategy.

The results of the digestion of pGEMT-VIPR2 with AvrII are shown in Figure 4.19. The fragment after extraction is shown in Figure 4.20. An attempt was made to clone this fragment into pGL3-basic but this was not successful.

4.3 Summary and conclusions

Although we tried to clone the human *VIPR2* intron1 sequence, using various methods such as touch down PCR, GC-rich PCR, Expand PCR and different cloning strategies, we were still not able to obtain the expected clones.

In the PCR based strategies, a major difficulty was obtaining the correct sized DNA fragment for cloning. This may have been a problem with the template DNA or an anomaly with the actual sequence chosen for amplification.

The difficulty encountered during the sub-cloning procedures may have been due to an inability to get adequate concentration of insert DNA. Although some DNA can

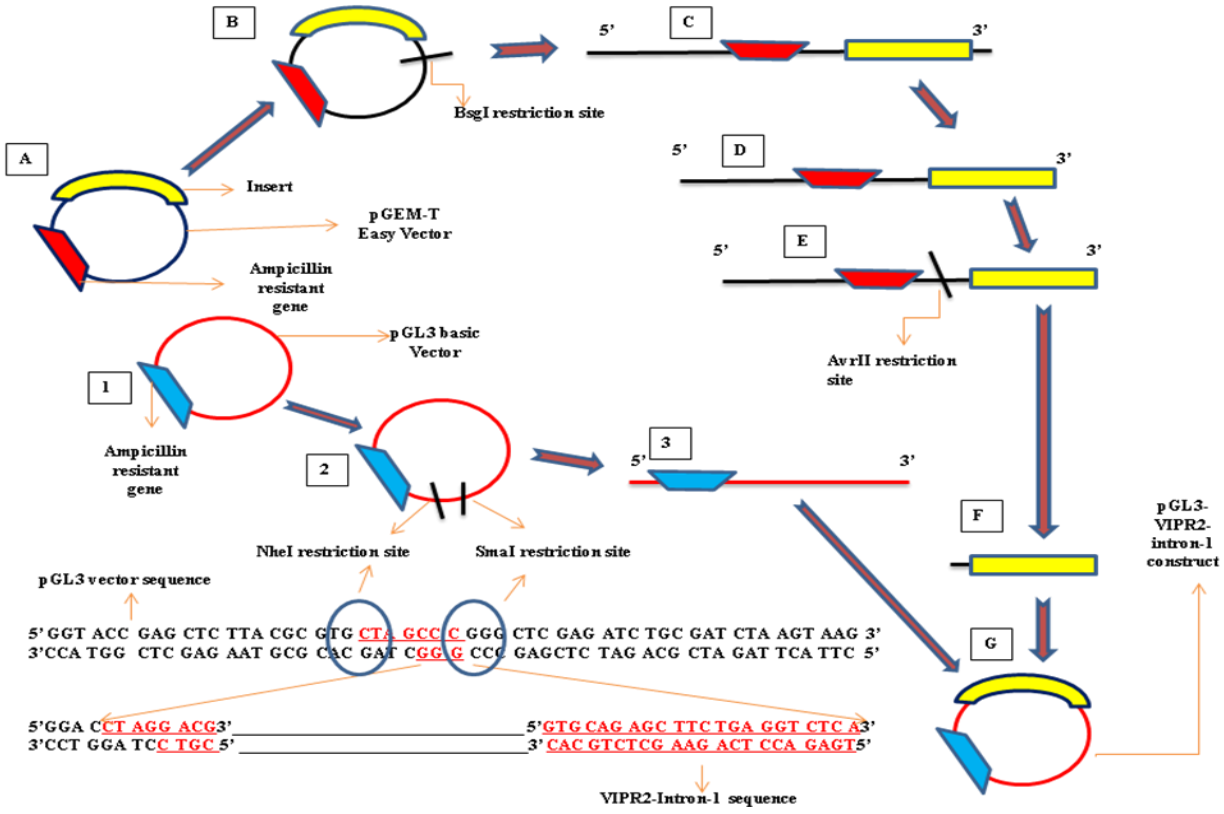


Figure 4.15 Cloning strategy: subcloning the human VIPR2 Intron-1 insert into pGL3 basic vector using AvrII/BsgI/NheI/SmaI

A. pGEMT-Easy vector with the human *VIPR2* Intron-1 as its insert (construct)
B. pGEMT-Easy/intron-1 construct with BsgI restriction site
C. pGEMT-Easy/intron-1 construct linearized after BsgI digestion producing a sticky end in the 3' region
D. The sticky end in the 3' region (BsgI site) is converted into blunt end pGEMT-Easy/intron-1 by extension using pfu polymerase
E. Followed by AvrII digestion in the linearised BsgI digested Insert Intron1
F. AvrII digest leaves a sticky end

1. pGL3 basic vector with ampicillin as selection marker
2. pGL3 basic vector is double digested with NheI and SmaI
3. pGL3 basic vector linearised after digestion leaving sticky end (NheI) in 5' and blunt end (SmaI) in 3' end.

G. Followed by ligation 5'sticky end of AvrII digested insert ligated with 5' sticky end of(NheI digested) pGL3 vector and 3' blunt end of (following extension) of insert ligated with 3' blunt end of (SmaI digested) pGL3 vector.

Legend: Blue filled trapezoid (in pGL3 vector), Red filled trapezoid (in pGEM-T Easy vector) are selection markers (Ampicillin), yellow filled block arc is human VIPR2 Intron-1 insert, and black semi-vertical lines ones are restriction enzyme sites.

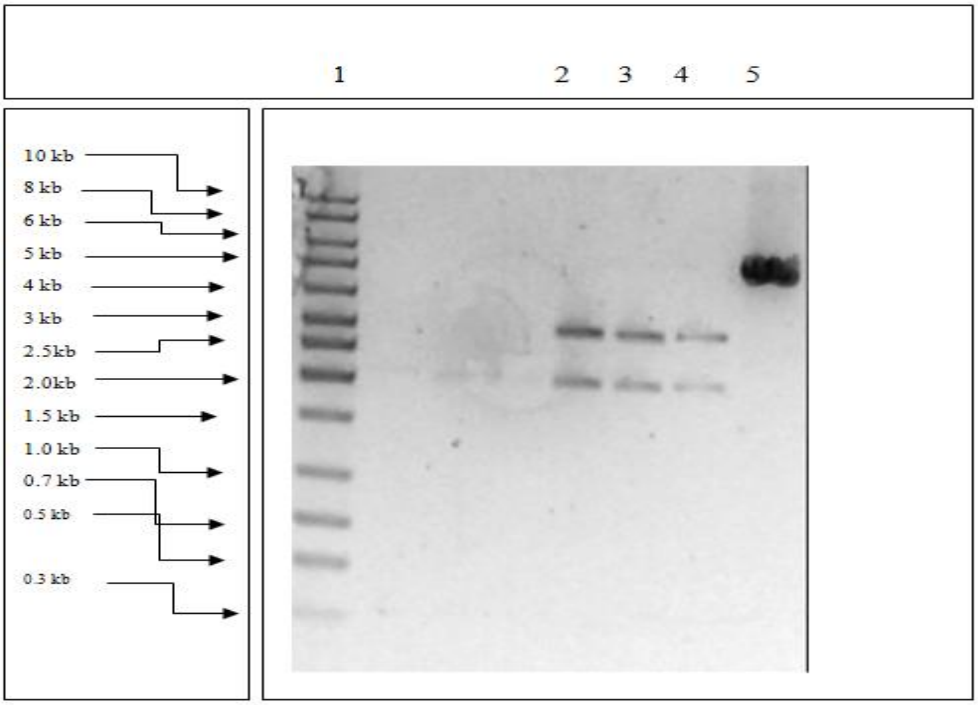


Figure 4.16 Gel electrophoretic analysis of AvRII digested 2.3 kb Intron-1 Insert of pGEMT-VIPR2.

Legend (After AvrII Digest)

Lane-1: 1 kb ladder digest (Axygen)
Lane-2: pGEMT-VIPR2 digested with AvrII (sample 3)
Lane-3: pGEMT-VIPR2 digested with AvrII (sample 2)
Lane-4: pGEMT-VIPR2 digested with AvrII (sample 1)
Lane-5: Uncut pGEMT-VIPR2

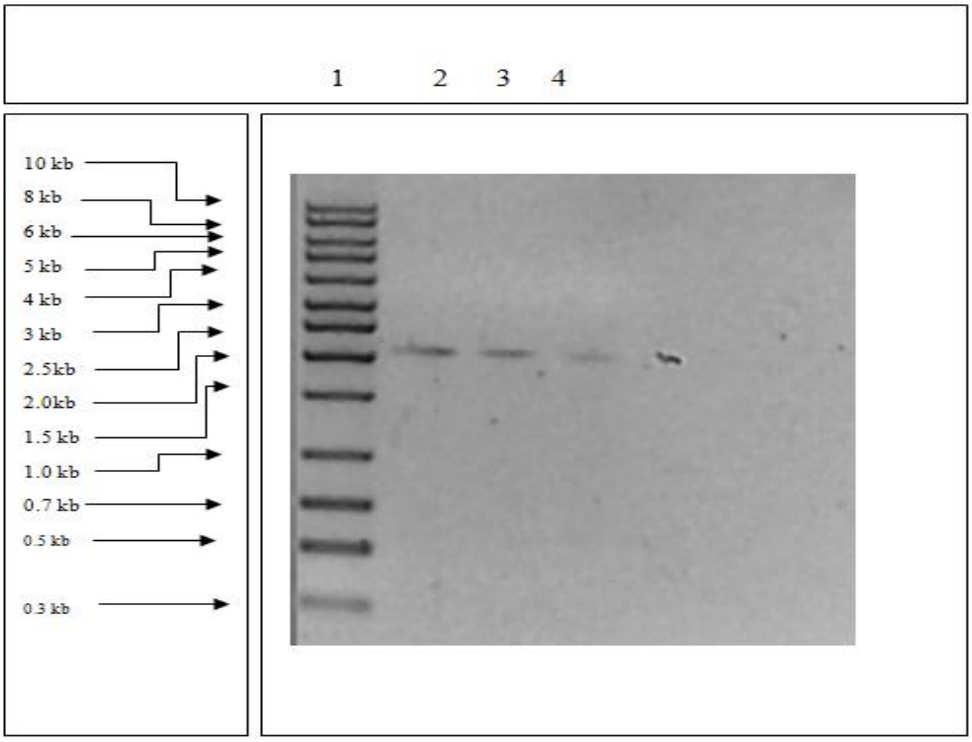


Figure 4.17 Gel electrophoretic analysis of the extracted fragment AvrII digested 2.3 kb Intron-1 Insert from pGEMT-VIPR2.

Legend (After extraction of AvrII digested fragment)

Lane-1: 1 Kb ladder (Axygen)
Lane-2: AvrII sample 3
Lane-3: AvrII sample 2
Lane-4: AvrII sample 1

be seen on the gels, this appeared to be insufficient to clone into the vector. This appears to be one critical impediment we encountered. If we were able to successfully prepare an adequate amount of DNA, the subsequent procedures should be relatively straightforward. In future, an optimized DNA extraction method will be used.

The fact that several clones were obtained that appeared to lack the correct insert or had no insert at all indicates that either there were contaminants on the plates or that the miniprep procedure was not optimized.

In conclusion, it was not possible to make the constructs required to carry out the promoter analysis. Given more time, these approaches would be repeated.

Chapter-5

Discussion, Conclusion and Future directions

This section deals with the discussion of the computer analysis of the regulatory region of the human *VIPR2* gene and the cloning strategies pursued to create various human *VIPR2*/Intron-1 contructs.

The findings of the computer analysis of the conserved ~6 kb 5' flanking region upstream and ~2.5 kb Intron-1 region downstream to the translation start site of the human *VIPR2* gene are elaborated and discussed in detail such as characteristic feature of the minimal and the core promoter involved in the basal expression of the human *VIPR2* gene and the tissue specific elements involved in the expression of the human *VIPR2* gene in activated T cells and Macrophages (Immune cells), Suprachiasmatic Nucleus (SCN) and Arcuate Nucleus (ARC) neurons, Pituitary cells, Adipocytes and Lung epithelial cells.

5.1 Core promoter of the Human *VIPR2* gene

In this study, the ~6kb regulatory region located 5' upstream and ~2.5 kb Intron-1 region downstream to the start codon of the human *VIPR2* gene, was computationally characterised using MATCH (Mayor *et al.*, 2000 and Kel *et al*., 2003), Multi-zPicture (Ovcharenko *et al.*, 2004), EvoprinterHD (Odenwald *et al.*, 2005 and Yavatkar *et al.*, 2008), Cis-Decoder (Brody *et al*., 2007) and CLUSTALW (Thompson *et al.*, 1994).

A number of studies have reported that the core promoter is the minimal portion which has general transcription factor binding sites, TSS, TATA box or several Sp1

binding sites in the case of TATA less promoters. Other studies have reported that the minimal promoter is the few nucleotides found in the core promoter that are absolutely required for the transcription of the gene **(**Smale and Kadonaga 2003). Though elements which make up the core promoter or transcription start site of the human *VIPR2* gene have been identified previously in our laboratory, the bioinformatics tools and the strategies used in the present study, identified minimal promoter and the elements found in it. The present study is more biologically sensible, functional and in-detail.

The strategy used in the present study, identified the elements which make-up the minimal promoter using MATCH analytical software which uses newly updated TRANSFAC 2009 database. The identified cis-elements in the minimal promoter were subjected to the phylogenetic analysis. The results revealed that the elements in the minimal promoter were indeed conserved evolutionarily among related and distant species and therefore considered to be functional.

The region which spans 240 bp upstream and 168bp downstream of the translation start site is identified as minimal promoter region. This 408 bp region is highly GC rich (figure: 3.4, 3.11-3.17) and contains multiple Sp1 binding sites (+68 to +77, +160 to +168, -228 to -220 and -240 to -232) along with core promoter which contains transcription start site around 187 bp upstream to the translation start site.

Macaya et al., 1976 defined the term isochores, as the long stretch of genomic domains (>3kb), with high homogenous GC content. The genomes of the warm-blooded vertebrates considered to be the mosaic of isochores with alternating low and high GC contents when compared to the cold blooded vertebrate genomes that lack GC rich isochores or compositional heterogeneity. Isochores are of two types:

L1 and L2 (<40% GC content) isochores are the GC-poor long genomic domains, represents flanking regions of the genes and H1 (47%), H2 and H3 (>53%) isochores represents the GC-rich long genomic domains with open chromatin structure and high concentration of transcriptionally active genes, regions or functional elements responsible for the gene transcription such as transcription factor binding sites, minimal promoter region, proximal and distal promoter regions. The sharp positive and negative spikes of GC content were also the typical features of the transcription start and stop sites, respectively (Zhang *et al.*, 2004). The identification of H2, H3 isochores, positive and negative GC spikes in the region which spans 240 bp upstream and 168bp downstream of the translation start site confirms the possible presence of minimal and proximal promoters in the 408 bp region (Figure 3.4) (Bernardi *et al.*, 1985 and Bernardi, 2000). The presence of H3 isochores in the Intron-1 (Majewski and Ott, 2002) and identification of several tissue specific regulatory elements suggests that it may play an active role in the tissue specific regulation of the *VIPR2* gene expression.

The absence of TATA or CAAT box or intiator element proximal to the transcription start site, suggest that the core promoter can be classified as a null promoter. But the typical feature of the Null promoter is usually the presence of multiple transcription start sites. Although the presence of multiple translation start sites is revealed by the bioinformatics analysis of the human *VIPR2* gene, the functionality of these sites is questionable *in vivo*. The absence of the typical core promoter element, functional multiple translation start site and tissue specific expression are the unique features of the human *VIPR2* gene (Novina and Roy, 1996).

Like other Class II GPCRs genes such as human secretin receptor gene, growth hormone releasing hormone receptor gene, rat VIP-R-1 receptor gene (Pei and Melmed, 1995) PTH receptor gene (Williams *et al.,* 2000), and mouse glucagon receptor gene (Geiger *et al.,* 2001), the ~6kb 5' flanking human VIPR2 gene has several evolutionary conserved Sp1 binding sites is without canonical TATA box, initiator element or CCAAT box proximal to its translation start site.

In a study conducted by Smale and Baltimore, 1989, to understand the need of Sp1 sites for the transcription of TATA-less promoter genes, G6I gene was chosen for the analysis. In a transcription system without Sp1, transcription initiation was < 0.1% when compared to the observed level in the TATA containing promoter. When Sp1 was added to the transcription system, the initiation was >100 fold. This finding suggests that multiple Sp1 sites are an essential requirement for TATA less promoters. The genes with TATA-less promoters are with multiple Sp1 sites, usually GC rich and the transcription initiated in the same way as the promoter with TATA box, both types of promoter initiate transcription with the help of TFIID but there exists a difference between them in a way TFIID is recruited. In TATA containing promoters, TFIID is recruited directly to the DNA through sequence specific interactions between TATA box and TBP, but in the case of TATA-less promoter, there is no TATA box and therefore absence of the intrinsic specificity for TFIID, the recruitment of TFIID to the DNA is done by the tethering-factor, which must be available physically along with the TBP. Since, the human *VIPR2* gene is a TATA-less gene like G6I and other class II GPCR genes, region surrounding the translation start site containing these multiple Sp1 sites may function as minimal promoter.

5.2 Tissue specific elements of the Human *VIPR2* gene

The computer analysis identified several tissue specific elements which confer tissue specificity to the *VIPR2* gene (figure: 3.11-3.17).

The Immune cell specific elements identified by the bioinformatics tools are NF Kappa B, Lymphoid enhancer binding factor-1, Ikaros-2 and Ikaros-3, TAL-1, MZF-1 and STAT (figure: 3.12). The Neuron specific elements identified are BRN-2, PAX-3, PAX-6, PAX-8, FOX, Math, Mash, Neurogenin and SOX (figure: 3.16). The muscle specific elements identified are MEF-2, Myogenin, MyoD and HAND1 (figure: 3.15). The Adipocyte specific elements identified are SREBP, PPAR and GR. The Lung specific elements identified are TTF-1and FREAC (3.14). The liver specific elements identified are LXR, HNF-1, HNF-3 and HNF-4 (figure: 3.13).

All the tissue specific elements identified are conserved and shared among evolutionarily distant species such as elephant, cow, dog and cat as well as evolutionarily related species such as Chimpanzee, Marmoset and Gorilla (figure: 3.5-3.10)

5.2.1 Transcription regulation of human *VIPR2* gene in Immune cells

Recent studies reveal that VIP plays an important role in regulating Immune system by maintaining Th1 and Th2 cells homeostasis. The VIP exerts its immunomodulary functions through its receptors such as VIP-R-1 and VIP-R-2 receptors. The VIP-R-1 receptor is expressed constitutively in resting T cells, and Macrophages, whereas VIP-R-2 receptor is expressed in the activated CD4+ T cells only upon activation, and not expressed constitutively in the resting T cells (Lara-Marquez *et al.*, 2000)

The expression of the *VIPR2* gene is induced by an antigen stimulated T cell receptor –CD3 complex in activated CD4+ T cells and LPS stimulated Toll like receptors in activated macrophages. The presence of the 5' flanking sequence upstream to the coding region which is conserved among the human, gorilla, chimpanzee, marmoset and elephant *VIPR2* genes, suggests that this proximal promoter sequence may be responsible for tissue specific regulation of *VIPR2* gene expression in activated CD4+ T cells and Macrophages.

5.2.1.1 NFKappaB

NFKappa B is the transcription factor and master regulator of the immune response and activated by more than 150 extracellular signals and various stimuli such as chemicals, heat and biological compounds. Several studies have noted that the activated NFKappa B induced the expression of various cytokine, chemokine, immune receptor genes and thereby plays an important role in cell differentiation and proliferation. The NFKappaB proteins, in unstimulated cells, exist as heterodimers or homodimers and present in the cytoplasm as bound inactive form with its inhibitor molecule IkB.

Under the influence of various stimuli, several cell surface receptors like Toll-like receptors (TLR1, 2, 4, 5, 6), T cell, IL-1 and TNF receptors were activated, resulting in the initiation of early signalling events like activation of various tyrosine kinases in the supramolecular activation complex (SMAC), formation of immunological synapse between immune cells and phosphorylation of several substrate proteins. These phosphorylated proteins form the multiprotein signalling complexes and followed by the activation of multiple downstream signalling pathways such as

Phospholipase -C γ1, mitogen activated protein kinase pathways and activation of exchange factor vav1. All these signalling pathways converge to the level of IKK complex which in turn phosphorylates inhibitory IkB and leading to its degradation by proteosome. The degradation of IkB is followed by the NFKappaB activation. The liberated NFKappaB in-turn trans-located to the nucleus from cytoplasm and binds to cis-element located in the regulatory region of the target genes and thereby induces their expression (Karin and Ben-Neriah, 2000).

Like the inducible role played by the NFKappa B binding site in the genes of Nitric Oxide synthase (Helmer *et al.,* 2002 and Hughes *et al.,* 2008), Intercellular adhesion molecule (ICAM) (Chini *et al.,* 1998), CD48 (Klaman *et al.,* 1995), OX40 (Tone *et al.,* 2007), Oxytocin receptor (Terzidou *et al.,* 2006), Human Opioid receptor (Kraus *et al*., 2003), kinin B1 receptor (Merino *et al.,* 2005 and Ni *et al.,* 1998), a conserved NFKappa B binding site, located in the 5' flanking region (-765 to –764) of the human *VIPR2* gene may be responsible for its tissue specific, inducible expression in the activated T cells and Macrophages.

5.2.1.2 CCAAT/ enhancer binding protein (C/EBP) and CEBP gamma

CCAAT/ enhancer binding proteins are the master regulator of several cellular processes such as cell proliferation, differentiation, Immune response, metabolism, inflammation and acute phase response. They belong to the six isomer family that binds to the dyad symmetrical repeat RTTGCGYAAY, (where R is A or G, and Y is C or T) located in the promoters of several genes. All the members of this family contain highly conserved basic leucin-zipper domain at their C-terminus which is involved in dimerization and DNA binding. They are found in major organs and in

various locations such as fat cells, liver cells, kidney, spleen, brain. Except C/EBP gamma, all other members of the C/EBP activate the transcription of the target genes through their activation domain.

Like the synergistic role played by C/EBP and nuclear factor KappaB binding sites in the human oxytocin receptor gene (Terzidou *et al.*, 2006), the rat serum amyloid A gene (Li and Liao, 1991) and the human intercellular adhesion molecule gene (Chini et al., 1998), the C/EBP binding site located in the evolutionarily conserved 5' flanking region (-884 to -871) of the human *VIPR2* gene, may play a co-activator role with NFKappaB in activating the human *VIPR2* gene expression.

CEBP gamma is the ubiquitous transcription factor generally known to inhibit other C/EBP trans-activational factors, but in certain situation known to play synergistic role with its family members in activating various cytokine promoters and also tissue specific. Like in the cytokine IL-6 and IL-8 promoters, the conserved CEBP gamma element located in the 5' flanking region of the *VIPR2* gene, may act as enhancer in enhancing the VIP-R-2 receptor expression in activated T lymphocytes.

5.2.1.3 Lymphoid enhancer binding factor -1 (LEF-1)

Lymphoid enhancer binding factor-1 is the lymphocyte specific architectural nuclear factor expressed in pre-B and T cells. It binds to the enhancer region of the T cell receptor alpha gene and thereby enhancing its basal gene expression (Travis *et al.*, 1991). The presence of the conserved LEF-1 sequence (-775 to -771) within the proximal promoter region of the human *VIPR2* gene signifies the potential role of *VIPR2* gene in T cell development from thymocytes.

Thymocytes are the hematopoietic stem cells, which develop into physiologically active T cell lymphocytes through negative, positive and beta selection in the thymus. The expression of the VIP-R-2 receptor and the LEF-1 transcription factor in the thymocytes and the LEF-1 transcription factor involvement in the lymphocyte development through Wnt/beta-catenin signalling pathway are well documented (Lara-Marquez *et al.*, 2000, 2001 and Dorfman *et al.*, 2003). When Wnt factors binds to the frizzle receptor (Class F GPCR) of the thymocytes, membrane bound dishevelled protein which is part of Wnt receptor complex become activated. This dishevelled protein inhibits the axin/GSK-3/Protein complex, leading to the entry of beta-catenin in the nucleus, followed by its interaction with TCF/LEF transcription complex and consequently the activation of target genes responsible for lymphocyte development (Staal *et al.*, 2001) Based on existing facts, I can strongly suggest that *VIPR2* gene may be one among those genes involved in T cell development.

To summarise, the VIP-R-2 receptor may be involved in normal T cell development, based on the following facts:

- The presence of the LEF-1 binding site in the conserved 5' flanking region of the Human *VIPR2* gene.
- LEF-1 involvement in the activation of genes involved in T cell development.
- TCF/LEF-1 mediated Wnt signalling pathway requirement for thymocyte development.
- VIP-R-2 receptor and LEF-1 transcription factor expression in thymocytes.

5.2.1.4 Ikaros-2 and -3, TAL-1, and MZF-1.

Ikaros-2 and Ikaros-3

Ikaros family of transcription factors consists of 8 isoforms, and which were generated from a single Ikaros gene by alternate splicing. They are the important regulator of events during lymphopoiesis and are found to be expressed only within fetal and adult hemo-lymphopoetic system. Of the 8 isoforms, only Ikaros-2 and Ikaros-3 binding sites were found within the conserved regions of the 5' upstream flanking region of the human *VIPR2* gene, a single Ik-3 site (-5797 to -5785) and two Ik-2 sites (-4728 to -4709 and -5854 to -5843) (figure: 3.12). Ikaros-2 is found to be expressed abundantly in the developing and matured lymphocytes whereas Ikaros-3 is produced in lesser amount. The trans-activation role played by the Ik elements in VIP-R-1 receptor gene and in other T cell specific genes such as CD2, CD3 and NFkB, implies that Ik binding sites identified in the conserved 5' flanking region of the human *VIPR2* gene, would be playing a important regulatory role in the *VIPR2* gene. This finding also implies the significance of the VIP-R-2 receptor and its expression in the regulation of the events of lymphocyte development (Sun et al., 1996 and Dorsam et al., 2002).

TAL-1

TAL-1 is the immune system specific helix-loop-helix transcription factor considered to be important factor required for all hematopoietic lineages which include erythropoiesis and its total absence, causes defect in erythropoiesis. This TAL-1 is found to be expressed in the vascular endothelial cells, progenitor endothelial cells,

sites for neovascularisation and known to play an important role in vasculogenesis and lymphangiogenesis (Tang *et al.,* 2006).

The presence of several TAL-1 elements in the conserved 5' flanking region (-4110 to -4100, -4306 to -4291, -4286 to -4277, -2197 to 2190, -2341 to -2327, -2367 to -2360, -2331 to -2316, -1599 to -1588, -1609 to -1602, -1450 to -1442, -1267 to 1262 and -175 to -164) of the human *VIPR2* gene, implies the role of the human VIP-R-2 receptor in the regulation of normal lymphopoiesis, erythropoiesis and lymphangiogenesis (Palamarchuk *et al.,* 2006).

MZF-1

MZF-1 is the kruppel family of transcription factor known to be expressed in totipotent hemopoietic cell and progenitor myeloid cells. The presence of the conserved binding site (-4927 to -4897) for this transcription factor in the 5' flanking region of the human VIP-R-2 receptor gene, reinforces the hypothesis that it may play a vital role in blood cell development.

Other miscellaneous binding sites

STAT (-3914 to -3896) is the Signal Transducers and Activator of Transcription plays an important role in cytokine mediated role in several tissues such as T cells, macrophage and mammary glands. The conserved element for STAT identified in the human VIP-R-2 receptor regulatory region, implies this receptor role in immune modulation (Takeda and Akira, 2000).

GATA3 is the T cell specific trans-activating factor belongs to GATA family of transcription factor. It is located in the conserved 5' flanking region (-2976 to -2984) of the *VIPR2* gene. It is the important regulator of T cell development and known to regulate *VIPR2* gene expression in CD4+ T cell (Sun *et al.,* 2006).

FOXO4 (Tzu Ling Lang *et al.,* 2003), **YY1** (Weill *et al.,* 2003) and **CEBP gamma** (Zhou *et al.,* 2009) elements found in the conserved 5' flanking region of the human *VIPR2* gene, may act as repressors and regulate the VIP-R-2 receptor gene expression.

5.2.2 Transcription regulation of the human *VIPR2* gene in the Pituitary cell.

The 5' flanking region upstream to the translation start site of the VIP-R-2 receptor gene found to consists of several conserved cis-elements known to play vital role in tissue specific expression of the human VIP-R-2 receptor gene. The minimal promoter of the human *VIPR2* gene active in Pituitary cell, do not have putative TATA box or CAAT box, proximal to the translation start site but instead, there are several Sp1 binding site surrounding the translation start site. Further upstream to the minimal promoter, there exist several cis-elements which might confer tissue specificity to the human *VIPR2* gene.

5.2.2.1 Pit-1A

Pit-1A is the pituitary specific factor, which is a member of POU family of transcription factors that regulate pituitary and mammalian development. It plays vital role in pituitary development, cell differentiation during organogenesis of anterior pituitary. This element is present in the regulatory region of the growth hormone releasing hormone receptor, growth hormone, prolactin, thyrotropin releasing hormone and thyroid stimulating hormone. The presence of two conserved binding site (-5667 to -5643 and -3672 to -3663) for the Pit-1 A in the 5' flanking

region of the human VIP-R-2 receptor gene, confers this receptor, the pituitary tissue specific expression and also implies that they may play a vital role in normal pituitary development.(Park *et al.,* 1999).

5.2.2.2 Vitamin D receptor (VDR)

Vitamin D receptor is the transcription factor belongs to nuclear receptor family. Upon activation by the ligand Vitamin-D, the VDR then enters the nucleus, bind to the vitamin D receptor element and thereby regulates the expression of the downstream target gene. Like in the case of CYP24 promoter, several VDR elements (-2830 to -2816, -3946 to -3927, -3991 to -3976, -5204 to -5191, and -6084 to -6070) in the conserved 5' flanking region of the human VIP-R-2 receptor gene, may act as enhancer by enhancing the transcription of this gene synergistically and amplify the expression of this receptor to many fold.

5.2.3Transcription regulation of the human *VIPR2* gene in neurons

The 5' flanking region upstream of the translation start site of the *VIPR2* gene was found to consist of several conserved cis-elements known to play vital role in tissue specific expression of the human VIP-R-2 receptor gene. The minimal promoter of the human *VIPR2* gene active in neurons of SCN and ARC, does not have putative TATA box or CAAT box, proximal to the translation start site but instead, there are several Sp1 binding site surrounding the translation start site. Further upstream to the minimal promoter, there exist several cis-elements which confer tissue specificity to the human *VIPR2* gene (figure: 3.16).

5.2.3.1 CREB and E-Box

Cyclic AMP response element binding protein is a neuron specific transcription factor that binds to Cyclic AMP response element, and activates or represses the downstream neuron specific genes. CREB element regulates the expression of the genes of several neuro-peptides or brain related molecules such as neurotrophin BDNF (Brain derived neurotrophic factor), Corticotropin releasing hormone, somatostatin, enkephalin and VGF nerve growth factor inducible (Bozdagi *et al*., 2008). The CREB is known for the involvement of long term memory, cognition, brain development and circadian rhythm.

Suprachiasmatic nucleus (SCN) and Arcuate nucleus (ARC) are the group of neurons known to control both circadian and metabolic rhythms through several molecules, the prominent one is the VIP-R-2 receptor (Bechtold et al., 2008), Histamine 3 receptor (Barrett *et al.,* 2005), CLOCK, PER, CRY, VGF (Wisor *et al.,* 1997 and Barrett *et al.,* 2005);

SCN is the size of the grain situated above the optic chiasm. It is one of the four nuclei (the others were lateral geniculate nucleus, the superior colliculus, and the pretectum) which directly receive nerve signal directly from the retina. It expresses several different molecules such as vasopressin, vaso-active intestinal peptide, transcription factors such as CLK and CYC in drosophila, CRY in mammals and several neurotransmitters.

SCN neuron expresses several genes related to biological clock such as *VIPR2* gene, CLOCK, CRY, and PER genes. Like CREB known to regulate one of the biological clock genes per1 (Tischkau *et al.,* 2003) and E box element known to regulate per1

and CRY in SCN. The three CREB conserved binding sites (-3231 to -3224, -3336 to -3314 and -2715 to -2705), and E-Box (-4030 to -4013) in the 5' flanking region of the human VIP-R-2 receptor, may be regulating the SCN specific VIPR2 gene. The presence of the CREB elements in the *VIPR2* gene also implies that this receptor may be playing an important role in brain development (Carlezon *et al.*, 2005).

5.2.3.2 BRN-2

BRN-2 is transcription factor known to play important role in early neuro-genesis, Disruption of its gene expression, will have tremendous impact on the normal development of endocrine hypothalamus. There is active expression of BRN from the initial stage in the hypothalamus neurons and it is needed for the survival of these neurons and for normal secretion of neuro-peptides. The presence of BRN conserved binding sites (-1478 to -1463 and -5536 to -5510) in the 5' flanking region of the *VIPR2* gene, implies that this receptor would be playing an important role in early events of neural development. (Ryan and Rosenfeld, 1997).

5.2.3.3 PAX3, PAX6 and PAX8

PAX3 is an important transcription factor that plays an important regulatory role in during neural tube development and neural crest migration. It is also vital regulator in emigration of neural crest cells and expressed in dermomyotome of the developing somite and limp buds. The presence of the conserved PAX3 binding elements in the 5' flanking region (-1014 to -997 and -5780 to -5766) of the VIPR2 gene, implies that this receptor plays an important role in neural tube development and organogenesis of brain. (Wehr and Gruss, 1996)

PAX6 is another vital transcription factor plays a key role in optic sulcus, optic stalk and optic vesicle, subsequently in the retina. It is vital for the normal lens and nasal development. Also documented is that this factor would have played important role in the dorsol-ventral polarity of the spinal cord. The presence of the conserved PAX6 binding elements in the 5' regulatory region (-3560 to -3548 and -4825 to -4803) of the *VIPR2* gene did implies that this receptor would have played important role in organogenesis. (Wehr and Gruss, 1996)

PAX8 is the transcriptional factor known to be vital for spinal cord and hindbrain development. The presence of this binding element in the conserved region (-4786 to -4774 and -5019 to -5006) 5' upstream to translation start site of the VIPR2 gene did implies that this receptor would have been key regulator in early brain development. (Wher and Gruss, 1996)

5.2.3.4 Fox

Fox are the group of transcriptional factors known to share common DNA binding element known as Forkhead motif in the target genes. They play important role in normal central nervous system or brain development, and they are abundantly expressed in developing brains. The mutation in the genes of these transcription factors, leads to serious human developmental abnormalities, speech and language impairment. The members of this family are Foxp1, foxp2, foxp3 and foxp4 were involved in normal brain development. The presence of these elements in the

conserved region (-5364 to -5347) of *VIPR2* gene, implies that this receptor would be playing a key role in brain development.

5.2.3.5 Math, Mash, Neurogenin and Sox

Math is the helix loop helix transcription factor belongs to NeuroD family known to play an important role in neuronal differentiation and sustenance of this differentiated state. The binding site for this transcription is located in the Intron-1 region (+1180 to +1185) of the Human *VIPR2* gene which is conserved among several species such as orangutan, rhesus monkey, marmoset, chimpanzee, cow, dog and cat. The presence of this element in the Intron-1 of this gene suggests that the receptor may be playing an important role in neural development and embryogenesis.

Mash is the mammalian achaete-scute homologue 1 transcription factor, now known to be important for serotonergic enteric neurons differentiation. The binding site for this transcription is located in the Intron-1 region (+2001 to +2007) of the Human *VIPR2* gene which is conserved among several species such as orangutan, rhesus monkey, marmoset, chimpanzee, cow, dog and cat. The presence of this element in the Intron-1 of this gene suggests that the receptor may be playing an important role in neural differentiation and neural development.

Neurogenin is the basic helix loop helix transcription factor involved in specifying neuronal differentiation. It is required for the formation of dorsal root ganglia. The binding site for this transcription is located in the Intron-1 region (+1989 to +1995) of the Human *VIPR2* gene which is conserved among several species such as orangutan, rhesus monkey, marmoset, chimpanzee, cow, dog and cat. The presence

of this element in the Intron-1 of this gene suggests that the receptor may be playing an important role in neural development.

Sox is the transcription factor that binds to minor groove of the DNA. It play**s** an important role in neuronal development. The binding site for this transcription is located in the Intron-1 region (+2098 to +2103) of the Human *VIPR2* gene which is conserved among several species such as orangutan, rhesus monkey, marmoset, chimpanzee, cow, dog and cat.

5.2.4 Transcription regulation of Human *VIPR2* gene in Adipocytes

Adipocyte or fat cells are associated with energy homeostasis, lipid metabolism, glucose metabolism, pathological conditions such as obesity, diabetes mellitus and cardiovascular disorders. They secrete several cytokines (TNF alpha, IL-6), complement factors (adipsin, adipocyte complement related protein, acrp30), peptides (angiotensinogen and Plasminogen activator inhibitor type-1) and hormone (leptin) which plays vital role in various physiological processes such as metabolic processes, stress responses, immune response, energy homeostasis, vascular hemodynamics, remodelling, wound healings, satiation, fertility, reproduction and hematopoiesis respectively.

The adipocyte expresses GLUT-1 transporter which play an important role in glucose metabolism. Alteration in the secretion of leptin, cytokines from adipocyte causes various pathologies such as obesity, diabetes mellitus and cardiovascular disorders. So, knowledge gained through various studies dealing with events related to transcription regulation of genes involved in adipogenesis and lipolysis, will help us

develop various therapeutic intervention for several adipocyte associated pathologies.(Morisson *et al.,* 2000).

In adipocytes, VIP and PACAP play an important role in lipolysis, by the cyclic AMP activation through the VIP-R-2 receptor (Akesson *et al.,* 2005). Accumulation of cAMP, followed by the activation of protein kinase A, and subsequent activation of lipase enzyme leads to hydrolysis of triglycerides to free fatty acids and glycerol. Understanding the transcriptional regulation of human *VIPR2* gene in adipocyte, will help us to find potential therapeutic interventions to cure several disorders associated with adipocytes.

The 5' flanking region upstream to the translation start site of the human VIP-R-2 receptor gene found to consists of several conserved cis-elements known to play vital role in tissue specific expression of the human VIP-R-2 receptor gene. The minimal promoter of the human *VIPR2* gene active in adipocytes, do not have putative TATA box or CAAT box, proximal to the translation start site but instead, there are several Sp1 binding site surrounding the translation start site. Further upstream to the minimal promoter, there exist several cis-elements which confer tissue specificity to the human *VIPR2* gene (figure: 3.11).

5.2.4.1 Sterol regulatory element binding protein (SREBP)

Sterol regulatory element binding proteins are helix alpha helix lucine zipper family of transcription factors. There are two SREBPs (SREBP-1 and SREBP-2), which are encoded by two different genes. There are three isoforms (SREBP-1a and SREBP-1c) produced from SREBP-1 gene. Out of these three isoforms, SREBP-1a is expressed more abundantly in the adipocytes. As this transcription factor belongs to helix alpha helix lucine zipper family, it binds with E-box (Wang *et al.,* 1993 and

Yokoyama *et al*., 1993). The key issue about the SREBPs is that they alone cannot able to activate efficiently the genes involved in energy homeostasis, so they require co-activators such as Sp1, Nuclear factor Y, CREB. The requirement for the co-regulators, number of SREBP sites in the promoter region, may vary gene to gene (Forman *et al*., 1995).

SREBPs known to activate the genes (low density lipoprotein receptor gene, fatty acid synthase, lipoprotein lipase) involved in fatty acid metabolism (Kim *et al.,* 1996), de novo lipogenesis, and cholesterol homeostasis (7-dehydrocholesterol reductase; farnesyl diphosphate, geranyl pyrophosphate synthase, lanosterol 14α-demethylase, glucose-6-phosphate dehydrogenase, 6-phosphogluconate dehydrogenase, glycerol-3-phosphate acyltransferase. (Horton *et al.,* 2002).

Like the SREBPs located in the regulatory regions of the genes responsible for the lipogenesis, fatty acid metabolism and cholesterol homeostasis, it is also present in several copies in the conserved regulatory regions of the human VIP-R-2 receptor gene. The presence of the conserved multiple SREBP elements (-6108 to -6097, -4784 to -4771, -4456 to -4443, -4431 to -4409) in the regulatory region of the Human *VIPR 2* gene in the adipocytes, indicates that the human VIP-R-2 receptor would have been playing an important role in adipogenesis, lipogenesis, fatty acid metabolism and cholesterol homeostasis.

5.2.4.2 Peroxisome proliferator activated receptors (PPAR)

Peroxisome proliferator activated receptors were members of the family of nuclear receptors, identified as component of adipocyte regulatory factor which binds to the

well characterised fat cell specific enhancer region of the aP2 gene. There are three known members of this PPAR family; they are PPAR α, γ, δ. The PPAR gamma type give rise to two isoforms, they were PPAR γ1 and PPAR γ2. When PPAR γ combines with the retinoid X factor becomes a potent regulator of adipogenesis and adipocyte differentiation. The PPAR γ is the key adipogenic transcriptional regulator and regulates various genes (CCAAT enhancer binding protein alpha, adipocyte fatty acid binding protein aP2 and Lipoprotein lipase) involved in adipogenesis and lipid metabolism. (Zhang *et al.*, 2008). The presence of conserved PPAR binding site (-5646 to -5633) in the 5' flanking region of the human VIP-R-2 receptor implies that this receptor may be playing a vital role in the adipogenesis and adipocyte differentiation (Morrison and Farmer, 2000).

5.2.4.3 CCAAT/Enhancer binding protein and Glucocorticoids

CCAAT/Enhancer binding protein (CEBP) is member of leucine zipper family of transcription factors. CEBP α, CEBP β and CEBP δ are known to play an important role in adipogenesis.The non-progenitor fibroblasts differentiated into adipocytes when treated with CEBP. Gene ablation study which target CEBP known to show little inclination for adipogenesis. These findings, implies that the CEBP known to regulate genes that are involved in the regulation of adipogenesis. The presence of the conserved binding sites (-0882 to -0871 and -3278 to -3264) in the 5' upstream to the translation start site of the human *VIPR2* gene, indicates that this receptor may be playing an important role in adipogenesis (Morrison and Farmer, 2000).

Glucocorticoids are known key players in the adipocyte development and fat metabolism. Studies have identified Lipin-1 as the important molecule, involved in

adipocyte differentiation, expression of the genes involved in lipid metabolism, insulin sensitivity, and energy expenditure. The key regulatory element known to regulate the Lipin-1 gene is Glucocorticoids through its glucocorticoids response elements (GRE). So, the presence of the several GRE in the conserved regions (-5681 to -5660, -5537 to -5508, -4307 to -4282 and -3712 to -3688) of the 5' flanking regions of the human VIP-R-2 receptor, implies that this receptor does plays an important role in adipocyte differentiation, lipid metabolism and adipogenesis (Zhang *et al.*, 2008).

5.2.5 Transcriptional regulation of the human VIPR2 gene in the lung epithelial cells

VIP-R-2 receptor mRNA was identified in the human respiratory tract. The receptor was found expressed in the lung epithelium and lung macrophages (Groneberg *et al.*, 2001). The regions responsible for the basal and lung specific expression of the human VIP-R-2 receptor is been characterised (figure: 3.14). This regulatory region is highly conserved and contains number of tissue specific cis-elements and minimal promoter need for the VIP-R-2 receptor expression. The minimal promoter is TATA-less, without initiator element and without CCAAT box, but it has number of Sp1 binding sites.

5.2.5.1 Thyroid transcription factor (TTF-1)

Thyroid transcription factor-1 is the homeodomain containing nuclear transcription factor which plays a critical role in the lung epithelial gene expression. It regulates the expression of the several lung epithelium specific genes such as surfactant

proteins –A, -B, -C and Clara cell specific proteins (Bruno et al., 1995; Bohinski et al., 1994). Removal of this TTF in the transgenic mice, causes severe abnormalities in lung morphology, this provides a vital clue that this transcription factor plays an important role in lung morphogenesis and lung epithelium specific gene expression (Oguchi and Kimura, 1995).

The presence of the thyroid transcription factor in the lung epithelium of the developing airway, confirms that the thyroid transcription factor-1 plays an important role in lung development. The presence of the conserved binding site (-5505 to -5493, and -3475 to -3467) for TTF in the ~6 kb flanking region of the human *VIPR2* gene implies that this receptor must be playing a key role in the lung development.

5.2.5.2 Forkhead RElated ACtivator-1 and -2 (FREAC)

Forkhead related activator is a member of the family of transcription factors found in metazoans and saccharomyces. This binding motif is the 100-amino acid binding domain, known to bind the consensus DNA sequence as a monomer and they control the group of genes usually expressed in terminally differentiated cells. The best example for this type of transcription factor which belong to FREAC family is HNF-3 found to regulate genes that are expressed in liver, the other important factors were FREAC-1 and FREAC-2 which are only expressed in lung (Hellqvist et al., 1996)

FREAC binding sites were mostly found in the promoters of the lung specific proteins such as pulmonary surfactant proteins A, B, C and Clara cell 10 kDa protein (CC10). Both FREAC-1 and FREAC-2 known to trans-activate the promoter of the surfactant proteins, but FREAC-1 known to activate the promoter of the CC10 only. Like the FREAC binding sites (-2978 to -2965) contributes to the lung specific

expression of the surfactant proteins A, B, C and CC10, the presence of FREAC sites in the multi-species conserved sequence in the 5' upstream region flanking the human VIP-R-2 receptor gene, may contributes to its lung specific expression (Margana and Boggaram, 1996).

5.2.5.3 Other miscellaneous binding site required for lung specific expression

To identify the transcription factor binding sites essential for the lung specific expression of lung specific proteins, regulatory regions of the SP-B gene was analyzed. At the end of the analysis HNF-3 alpha, Sp3, TTF-1 and Sp1 sites were identified. Mutational analysis were done on these sites and found to be critical for the lung specific expression of the SP-B gene. Those elements which were found in the regulatory region of the SP-B gene were found in the conserved sequences (HNF-3: -5361 to -5350, -4735 to -4723; Sp3: -1241 to -1232) 5' upstream to the flanking region of the human *VIPR2* gene. The elements identified in the *VIPR2* gene may contribute to the lung specific expression of the human VIP-R-2 receptor (Margana and Boggaram, 1997).

5.2.6 Transcriptional regulation of the Human VIPR2 gene in Myocytes

The 5' flanking region upstream to the translation start site of the VIP-R-2 receptor gene found to consists of several conserved cis-elements known to play vital role in tissue specific expression of the human VIP-R-2 receptor gene. The minimal promoter of the human *VIPR2* gene active in smooth muscle, does not have a putative TATA box or CAAT box, proximal to the translation start site but instead, there are several Sp1 binding site surrounding the translation start site. Further

upstream to the minimal promoter, there exist several cis-elements which confer tissue specificity to the human *VIPR2* gene (figure: 3.15).

5.2.6.1 Human myocyte specific enhancer factor MEF-2

Human myocyte specific enhancer factor MEF-2 is the member of MADS gene family, which includes other homeogenes from yeast, plants and animals and all of them share the consensus sequence. It is expressed only in the muscles such as skeletal, cardiac and smooth muscle lineage. The MEF-2 binding site is found in cardiac and skeletal muscle specific genes such as muscle creatine kinase, myogenin, myoD, MRF4, myosine light chain enhancers, myosine light chain-2, cardiac troponin T and muscle specific phosphoglycerate mutase. Like the MEF-2 found in the skeletal, cardiac and smooth muscle specific genes, and activate their muscle specific expression, MEF-2 binds to DNA sequence, located in the conserved 5' upstream (-5612 to -5591 and -5398 to -5383) flanking region to the human *VIPR2* gene and activate its tissue specific expression. The presence of this MEF-2 binding site also implies that the human VIP-R-2 receptor would have played a vital role in commitment events leading to myogenesis (Black and Olson, 1998).

5.2.6.2 Myogenin and MyoD

Myogenin is the muscle specific transcription factor known to be responsible for the normal development and repair of the skeletal muscle and MyoD is the key protein involved in the muscle differentiation. MyoD is the member of the myogenic transcription factors and the important earliest marker of myogenic commitment. Both Myogenin and MyoD are nuclear proteins, considered to be the myogenic

regulatory factors (MRFs). Both the MRFs, expressed exclusively in the skeletal muscle, regulate several muscle specific genes such as creatine kinase gene, myosin light chain gene, nicotinic acetylcholine receptor gene (Eftimie *et al.*, 1991) and considered to be the master regulator for the skeletal muscle development. MyoD is the primary MRF known to play key role in the myoblasts determination, commitment and Myogenin is the secondary MRF considered to be important for differentiation. So presence of the cis-elements for both the factors (Myogenin: -4430 to -4422, -3831 to -3823, -2662 to -2654, -173 to -164 and MyoD: -4111 to -4100) in the conserved 5' flanking regions of the human *VIPR2* gene, suggest that these elements may be responsible for the expression of the human VIP-R-2 receptor in the muscle and therefore the expression of this receptor important for the skeletal muscle development.

5.2.6.3 Heart and Neural crest expressed protein-1 (HAND1)

HAND1 transcription factor is expressed throughout heart tube during embryogenesis. This transcription factor plays an important role in the cardiac cell differentiation and organogenesis of the heart. HAND1 usually heterodimerize with cofactors such as E47 and tend to bind specific binding site and regulate the target gene. Recent finding suggests that HAND1 and MEF-2 synergistically activate the transcription of the Atrial Natriuretic factor and ET-1 genes (Morin *et al.*, 2005). The presence of the binding sites (-5304 to -5290, -4919 to -4897 and -3214 to -3198) for the HAND1/E47 transcription factor in the conserved 5' flanking region of the human *VIPR2* gene suggest that, the VIP-R-2 receptor expression may play a vital role in heart development.

5.2.7 Transcriptional regulation of the Human *VIPR2* gene in Hepatocytes

The 5' flanking region upstream to the translation start site of the VIP-R-2 receptor gene found to consist of several conserved cis-elements known to play vital role in tissue specific expression of the human VIP-R-2 receptor gene. The minimal promoter of the human *VIPR2* gene active in liver, do not have putative TATA box or CAAT box, proximal to the translation start site but instead, there are several Sp1 binding site surrounding the translation start site. Further upstream to the minimal promoter, there exist several cis-elements which confer tissue specificity to the human *VIPR2* gene (figure: 3.13).

5.2.7.1 Liver X receptor (LXR)

Liver X receptor is the nuclear hormone receptor and a ligand activated transcription factor. It plays vital role in cholesterol homeostasis, bile acid metabolism and considered to be the master regulator of lipogenesis. LXR plays a vital role as lipid and sterol sensors and regulate genes (ABCA1, ABCG1, apoE, CYP7A1, ABCG5, and ABCG8) responsible for cholesterol mobilization. LXR regulates fatty acid biosynthesis by hetero-dimerization with FXR and through binding of this heterodimer complex on SREBP-1c target gene. SREBP-1c gene expresses SREBP-1c protein which in turn transactivates several genes involved in fatty acid synthesis. LXR also plays a vital role in carbohydrate metabolism through carbohydrate response element binding protein (ChREBP). LXR together with FXR bind the regulatory region and activate the expression of the ChREBP gene. Once expressed

ChREB protein activates genes responsible for carbohydrate metabolism.The presence of the binding sites for LXR and FXR in the conserved 5' flanking regions (LXR: -3877 to -3860, and FXR: -2703 to -2691, -1194 to -1182) of the human *VIPR2* gene, suggest that VIP-R-2 receptor may be playing an important role in Fat and carbohydrate metabolism (Cha and Repa, 2007)

5.2.7.2 HNF1, HNF3 and HNF4

Hepatocyte nuclear factor 1, 3 and 4 are liver specific transcription factors. HNF1s (HNF1 alpha and HNF1 Beta) which are related to homeobox proteins, HNF3s (HNF3 alpha, HNF3 beta and HNF gamma) which belong to forkhead family and HNF4 a member of Nuclear steroid-thyroid receptor superfamily take part in regulating genes responsible for hepatocyte differentiation and the liver development. The decrease in the mRNA levels of HNF1 alpha, HNF3 and HNF4 and increase in the mRNA levels of the HNF1 beta and HNF3 beta in the late period of embryonic liver development, suggest their presence and their important role in hepatocyte differentiation and therefore liver development (Nagy *et al.*, 1994).

The presence of the conserved binding sites (HNF1: -5643 to -5633; HNF1: -5500 to -5485; HNF3 alpha: -5361 to -5350; HNF4: -2630 to -2617) for these transcription factors in the 5' flanking region of the human *VIPR2* gene suggests that the human VIP-R-2 receptor would be playing several important role in liver development. This conclusion can be substantiated by the fact that the VIP-R-2 receptor is expressed in hepatic progenitor cell (Cassiman *et al.*, 2007).

5.2.7.3 Sterol regulatory element binding protein (SREBP)

SREBP belongs to the family of transcription factors that regulate genes which are responsible for fat metabolism. When the membrane cholesterol is high, SREBPs found to exist in (nucleus envelope and endoplasmic reticulum) membranes bound form. Once the cholesterol concentration becomes low, the bound form started to disintegrate or cleaved by two distinct proteases (site-1 protease and site-2 protease), leading to the generation of active form of SREBP. This mature and active form begin to migrate and enters the nucleus and binds the respective element in the regulatory regions of the genes which were involved in cholesterol uptake or fat metabolism such as low density lipoprotein receptor gene, cytoplasmic hydroxymethylglutaryl-CoA synthase or hydroxymethylglutaryl-CoA reductase gene. This protein is abundantly expressed in liver and adipose tissue. Another study revealed that this transcription factor is needed for the glucose -induced expression of the FAS, S14, Acetyl co-A carboxylase (ACC), and L-type Pyruvate Kinase (L-PK) genes involved in fatty acid metabolism (Foretz et al., 1999). Like SREB proteins, in those genes, their presence and activation of these binding sites (-6108 to -6094, -4784 to -4771, -4456 to -4443, -4431 to -4409 and -3943 to -3934) in the conserved region of the 5' flanking region of the human VIP-R-2 receptor may regulate the liver specific expression of the *VIPR2* gene. Their presence in the regulatory region also indicates that this receptor may play an important role in fat and carbohydrate metabolism in Liver.

5.2.7.4 Other miscellaneous factors

CREBP is the cAMP responsive element binding protein binding site (-2714 to -2706) located in the conserved 5' flanking region of the human *VIPR2* gene, may function as the co-activator of the transcription of the *VIPR2* gene in the hepatocytes by forming trans-activation complex with ChREBP and HNF4 alpha as in the case of glucose mediated L-PK gene expression (Burke *et al.,* 2009).

NF Kappa B is the master regulator of the several genes involved in inflammation and acute phase response. Like in the role of NFKappa B in the ethanol or LPS mediated iNO and COX-2 gene expression, the conserved NFkappaB binding sites (-763 to -755 and -4905 to -4897) found in the 5' flanking region of the human VIPR2 gene, may take part in the trans-activation of the VIP-R-2 receptor expression in liver (Spitzer *et al.,* 2002).

Glucocorticoid receptor (GR) is also known as NR3C1, with which ligand like cortisol binds. It is expressed in all the cells and controls genes which are responsible for metabolism, immunity and development. In the absence of the ligand, GR resides in the cytosol and once activated by ligands, GR translocated into the nucleus and binds to the respective binding element or binding site and activate or represses the target gene. The presence of these GR elements (-5590 to -5573, -5537 to -5508, -4307 to -4282 and -3712 to -3688) in the conserved regions of the 5' flanking region of the human VIP-R-2 recptor, may regulate its gene expression like GR element that activates tyrosine amino transferase gene in rat liver (Schweizer-Groyer et al., 1997).

YY1 is the ubiquitous transcription factor known to repress or activate the promoter of the target gene. YY1, like its role as repressor in the NF Kappa mediated rat serum amyloidal A1 gene expression, YY1 may interfere with the binding of NF Kappa B

with its binding site, which is located in the 5' flanking region of the *VIPR2* gene and thereby repress the VIP-R-2 receptor gene expression or it may bind with its YY1 binding sites (-5702 to -5687, -2862 to -2846, -2842 to -2823 and -736 to -720) located in the 5' regulatory region of the receptor gene and activates its gene expression (Lu et al., 1994).

Peroxisome Proliferator Activated Receptor alpha (PPAR α) is the ligand inducible transcription factor which regulates the transcription of the several genes (Cholesterol 7 7-α- hydroxylase gene, sterol 27-hydroxylase gene, aceyl coenzyme A: cholesterol aceyl transferase, HMG coenzyme A reductase, and Farensyl pyrophosphate synthase), involved in lipid and lipoprotein metabolism through binding with its Peroxisome proliferator response element found in their 5' flanking regions. PPAR-α, as in those genes, may regulate the expression of VIP-R-2 receptor in the hepatocytes. The presence of the PPRE in the conserved regions (-6086 to -6070, -5643 to -5633, -4929 to -4911, -4564 to 4546 and -3561 to -3548) of the 5' flanking region of the human *VIPR2* gene, suggests that, this receptor may play a vital role in fat metabolism (Jossic-corcos et al., 2004).

5.3 Cloning and building an Intron-1 pGL3 construct

The Intron-1 of the human *VIPR2* gene has features which are most seen in the typical regulatory region. The features such as high GC rich content, High GC Isochore, suggest that the Intron-1 may be responsible for transcription regulation of the human *VIPR2* gene. Even if there are no functional regulatory elements in the Intron-1, still it can be used as negative control for future functional studies. So, the task of cloning the Intron-1 was carried out. Though different types of polymerase chain reactions were tried, the cloning the Intron-1 has not materialised because of

high GC content and the secondary structure formation. Sequence with G/C repeats produce lots of inter-strand and intra-strand folding with the neighbouring guanine because of elevated hydrogen bonding. In PCR, this process is confirmed by the presence of the non-specific bands in the electrophoresis. The presence of non-specific bands is purely because of mispriming and misannealing between the template and complementary strands due to high melting Tm. So, cloning by PCR is replaced by cloning restriction digest and ligation. The problem, we encountered is the poor harvest of the Intron-1 insert DNA concentration. If the concentration of the insert DNA is increased, more chance of making Intron-1/pGL3 construct is possible. If the cloning by PCR is still preferred, the betaine and DMSO can be tried out in combination with lengthy primers or longer oligonucleotides or RT-PCR can done using high fidelity taq polymerase using gradient cycler.

5.4 Conclusion and future directions

The outcome of the analysis presented in the thesis is the detailed characterisation of the regulatory regions responsible for basal and tissue specific expression of the *VIPR2* gene in different cell types such as activated T cells, Macrophages, lung epithelial cells, Neurons, myocyte, hepatocytes and adipocytes. The novel elements which are responsible for inducible nature of the human VIP-R-2 receptor were also identified. This is a high resolution analysis, the regulatory factor motif identified, were categorized based on the quality of the source. Now, the minimal promoter is characterised and this characterisation, extended the boundary of minimal promoter to Intron-1. This high resolution study also identified novel neuron specific elements responsible for the brain development, organogenesis and embryogenesis in the

Intron-1. The analysis also identified cis-elements responsible for the normal transcription regulation of the human VIP-R-2 receptor in adipocytes and Lung epithelial cells. These identified elements may pave the way for development of the novel therapeutic interventional strategies by interfereing with the signalling pathways such as wnt signalling in the case of transcription factors which plays vital role in cancer and Immunity.

The next step is to carry out functional studies using AtT20 and T98g cell lines which are elaborated in this thesis. The strategy used to characterise these elements, based on evolutionary conservation, so the elements identified only the resilient elements which are shared among evolutionarily related or distant species.

The experience gained from the cloning, definitely will pave the way for the development of novel techniques to amplify and clone the GC rich fragments. The reason for not getting the clone and not able to make the construct is purely due to lack of time. Further studies such as deletional analysis, transfection studies to be done to experimentaly validate the elements identified by this study.

References

Akesson L, Ahre´n B, Edgren G, and Degerman E (2005). VPAC2-R Mediates the Lipolytic Effects of PituitaryAdenylate Cyclase-Activating Polypeptide/Vasoactive Intestinal Polypeptide in Primary Rat Adipocytes. Endocrinology, 146 (2), 744–750.

Alam J and Cook JL (1990). Reporter genes, Application to the study of mammalian gene transcription. Analytical Biochemistry, 188**,** 245–254.

Antequera F and Bird A (1993). Number of CpG islands and genes in human and mouse. Proceedings of National Academy of Sciences United States of America, 90, 11995-11999.

Antequera F and Bird A (1999). CpG islands as genomic footprints of promoters that are associated with replication origins. Current Biology, 9, R661–R667.

Antequera F (2003). Structure, function and evolution of CpG island promoters. Cellular and Molecular Life Sciences, 60, 1647–1658.

Antonini SR, N'Diaye N, Baldacchino V, Hamet P, Tremblay J, Lacroix A (2004). Analysis of the putative regulatory region of the gastric inhibitory polypeptide receptor gene in food-dependent Cushing's syndrome. Journal of Steroid Biochemistry & Molecular Biology, 91, 171–177.

Arimura A and Shioda S, (1995). Pituitary adenylate cyclase activating polypeptide (PACAP) and its receptors, neuroendocrine and endocrine interaction. Frontiers in Neuroendocrinology 16, 53–88.

Asnicar MA, Köster A, Heiman ML, Tinsley F, Smith DP, Galbreath E, Fox N, Ma YL, Blum WF, Hsiung HM (2002). "Vasoactive intestinal polypeptide/pituitary adenylate cyclase-activating peptide receptor 2 deficiency in mice results in growth retardation and increased basal metabolic rate". Endocrinology 143 (10), 3994–4006

Bai G, Elizabeth W, Stuebing E W, Parker H R, Harlow P and Nemer M (1993). Combinatorial Regulation by Promoter and Intron 1 Regions of the Metallothionein SpMTA Gene in the Sea Urchin Embryo. Molecular and Cellular biology, 13 (2), 993-1001.

Barrett P, Ross A W, Balik A, Littlewood P A, Mercer J G, Moar K M, Sallmen T, Kaslin J, Panula P, Schuhler S, Ebling F J, Ubeaud C and Morgan P J (2005). Photoperiodic regulation of histamine H3 receptor and VGF messenger ribonucleic ccid in the Arcuate Nucleus of the Siberian Hamster. Endocrinology, 146 (4), 1930-

1939.

Basille M, Cartier D, Vaudry D, Lihrmann I, Fournier A, Freger P, Gallo-Payet N, Vaudry H, Gonzalez B (2006). Localization and characterization of pituitary adenylate cyclase-activating polypeptide receptors in the human cerebellum during development. Journal of Comparative Neurology, 496 (4), 468-478.

Bechtold D A, Brown T M, Luckman S M, and Piggins H D (2008). Metabolic rhythm abnormalities in mice lacking VIP-VPAC$_2$ signaling. American Journal of Physiology, Regulatory Integrative Comporative Physiology, 294, R344-R351

Bellinger DL, Lorton D, Romano T (1990). Neuropeptide innervation of lymphoid organs. Annals of the New York Academy of Sciences, 594, 17-33.

Bernardi, G., Olofsson, B., Filipski, J., Zerial, M., Salinas, J., Cuny, G., Meunier-Rotival, M., and Rodier, F. (1985). The mosaic genome of warm-blooded vertebrates. Science 228, 953–958.

Bernardi G (2000). Isochores and the evolutionary genomics of vertebrates. Gene 4; 241 (1), 3-17.

Bettoun D, Minagawa M, Hendy G N, Alpert L C, Goodyer C G, Goltzman D, and White J H (1998). Developmental upregulation of human parathyroid hormone (PTH)/PTH-related peptide receptor gene expression from conserved and human-specific promoters. Journal of Clinical Investigation. 102 (5), 958–967.

Bird A.P. (1986). CpG-rich islands and the function of DNA methylation. Nature 321, 209–213

Black B L, Olson E N (1998). Transcriptional control of muscle development by myocyte enhancer factor-2 (MEF2) proteins. Annual Review of Cell and Developmental Biology, 14, 167-196.

Bokaei PB, Ma XZ, Byczynski B, Keller J, Sakac D, Fahim S, Branch DR (2006). Identification and characterization of five-transmembrane isoforms of human vasoactive intestinal peptide and pituitary adenylate cyclase-activating polypeptide receptors. Genomics, 88 (6), 791-800.

Boutillier AL, Monnier D, Lorang D, Lundblad JR, Roberts JL, Loeffler JP (1995). Corticotropin-releasing hormone stimulates proopiomelanocortin transcription by cFos-dependent and -independent pathways, characterization of an AP1 site in exon. Molecular Endocrinology. 9 (6), 745-55.

Boylan M.O, Jepeal L.I and Wolfe M.M (2006). Sp1/Sp3 binding is associated with cell-specific expression of the glucose-dependent insulinotropic polypeptide receptor gene. American Journal of Physiology, Endocrinology and Metabolism, 290, E1287–E1295

Bozdagi O, Rich E, Tronel S, Sadahiro M, Patterson K, Shapiro ML, Alberini CM, Huntley GW, Salton SR (2008). The neurotrophin-inducible gene VGF regulates hippocampal function and behaviour through a brain-derived neurotrophic factor-dependent mechanism. The Journal of Neuroscience, 28 (39), 9857–9869.

Brenneman DE and Eiden LE (1986). Vasoactive intestinal peptide and electrical activity influence neuronal survival**.** Proceedings of National Academy of Science United States of America, 73, 1159–1162.

Brody, T., Rasband, W., Baler, K., Kuzin, A., Kundu, M. and Odenwald, W. F. (2007). *cis*-Decoder discovers constellations of conserved DNA sequences shared among tissue-specific enhancers. Genome Biology, 8 (5), R75.1- R75.25

Bronstein I, Fortin J, Stanley PE, Stewart GSAB and Kricka LJ (1994). Chemiluminescent and bioluminescent reporter gene assays. Analytical Biochemistry, 219, 169–181.

Buggy J, Hull J and Yoo-Warren H (1995). Isolation and structural analysis of the 5' flanking region of the gene encoding the human glucagon receptor. Biochemical and Biophysical research communications, 208, 339-344.

Buratowski,S., Hahn, S., Guarente, L., Sharp, P.P (1989). Five intermediate complexes in transcription initiation by RNA polymerase II. Cell, 56, 549–561.

Burke S J, Collier J J and Scott D K (2009). cAMP opposes the glucose-mediated induction of the L-PK gene by preventing the recruitment of a complex containing ChREBP, HNF4 alpha, and CBP. The Federation of American Societies for Experimental Biology Journal, 23, 2855-2865

Carlezon, W. A., Jr., Duman, R. S., and Nestler, E. J. (2005). The many faces of CREB. Trends in Neuroscience, 28, 436−445.

Cassiman D, Sinelli N, Bockx I, Vander Borght S, Petersen B, De Vos R, van Pelt J, Nevens F, Libbrecht L, Roskams T (2007). Human hepatic progenitor cells express vasoactive intestinal peptide receptor type 2 and receive nerve endings. Liver International, 27 (3), 323-328.

Cha J and Repa J J (2007). The Liver X Receptor (LXR) and Hepatic Lipogenesis. The Journal of Biological Chemistry, 282 (1), 743–751

Chew LJ, Yuan X, Scherer SE, Qie L, Huang F, Hayes WP, Gallo V (2001). Characterization of the rat GRIK5 kainate receptor subunit gene promoter and its intragenic regions involved in neural cell specificity. Journal of Biological Chemistry, 276 (45), 42162-42171.

Chini B A, Fiedler M A, Milligan L, Hopkins T, and Stark JM (1998). Essential Roles of NF-kB and C/EBP in the Regulation of Intercellular Adhesion Molecule-1 after Respiratory Syncytial Virus Infection of Human Respiratory Epithelial Cell Cultures. Journal of Viology, 72 (2) 1623–1626

Ciccarelli E, Svoboda M, De Neef P, Di Paolo E, Bollen A, Dubeaux C, Vilardaga JP, Waelbroeck M and Robberecht P (1995). Pharmacological properties of two recombinant spice variants of the PACAP type 1 receptor, transfect and stably expressed in CHO cells. European Journal of Pharmacology, 288, 259-267.

Cutler DJ, Haraura M, Reed HE, Shen S, Sheward WJ, Morrison CF, Marston HM, Harmar AJ, Piggins HD (2003). "The mouse VPAC2 receptor confers suprachiasmatic nuclei cellular rhythmicity and responsiveness to vasoactive intestinal polypeptide in vitro". European Journal of Neuroscience, 17 (2), 197–204.

Daniels R, Kinis T, Serhal PM, Monk M (1995). Expression of the myotonin protein kinase gene in preimplantation human embryos. Human Molecular Genetics, 4, 389-393.

Daniels R, Lowell S, Bolton V, Monk M (1997). Transcription of tissue specific genes in human preimplantation embryos. Human reproduction, 12**,** 2251-2256.

Delgado M, Ganea D (2001). Cutting edge: is vasoactive intestinal peptide a type 2 cytokine? Journal of Immunology, 166, 2907–2912.

Dey, R.D., Shannon, W.A., Jr and Said, S.J, (1981). Localisation of VIP immuno-reactive nerves in airways and pulmonary vessels of dogs, cats and human subjects. Cell and Tissue Research, 220, 231-238.

Dickinson T, Fleetwood-Walker SM (1999). VIP and PACAP: very important in pain? Trends in Pharmacological Sciences, 20, 324– 329.

Dickson L and Finlayson K (2009). VPAC and PAC receptors: From ligands to function. Pharmacology and therapeutics, 121, 294-316.

Ding C, Racusen LC, Wilson PD, Burrow CR, Levine MA (1995). Identification of an alternative spliced form of PTH/PTHrP receptor mRNA in immortalized renal tubular cells. Journal of Bone and Mineral Research,10 (supplementary-1) S484.

Dorsam G and Goetzl E J (2002). Vasoactive intestinal peptide receptor-1 is a novel gene target of the hemolymphopoietic transcription factor Ikaros. The Journal of Biological Chemistry, 277, 13488-13493.

Duckles, S.P and Said, S.I. (1982). Vasoactive intestinal peptide as a neurotransmitter in the cerebral circulation. European journal of pharmacology, 78 (3), 371-374.

Eddy S R (1995). Multiple alignment using Hidden markov models. Proceedings of International conference on Intelligent systems for molecular biology, 3, 114-120.

Eftimi R, Brennert H R and Buonanno A (1991). Myogenin and MyoD join a family of skeletal muscle genes regulated by electrical activity. Proceedings of National Academy of the Sciences. United States of America, 88, 1349-1353.

Foretz M, Pacot C, Dugail I, Lemarchand P, Guichard C, Lièpvre X L, Berthelier-Lubrano C, Spiegelman B, Kim J B, Ferré P and Foufelle F (1999). ADD1/SREBP-1c is required in the activation of hepatic lipogenic gene expression by glucose. Molecular and Cellular Biology, 19 (5), 3760-3768.

Forman BM, Tontonoz P, Chen J, Brun RP, Speigelman BM, Evans RM (1995). 15-deoxy 12,14-prostaglandin J2 is a ligand for the adipocyte determination factor PPARgamma. Cell, 83, 803-812.

Frech K, Herrmann G, Werner T (1993). Computer-assisted prediction, classification, and delimitation of protein binding sites in nucleic acids. Nucleic Acids Research, 21, 1655-1664

Fredriksson, R., Lagerstrom, M.C., Lundin, L.G. and Schioth, H.B. (2003). The G-protein-coupled receptors in the human genome form five main families. Phylogenetics analysis, paralogon groups, and fingerprints. Molecular Pharmacology, 63, 1256-1272.

Ganley AR, and Kobayashi T (2007). Phylogenetic footprinting to find functional DNA elements. Methods Molecular Biology. 395, 367-380.

Gao F and Zhang CT (2006). GC-Profile, a web-based tool for visualizing and analyzing the variation of GC content in genomic sequences. Nucleic Acids Research. 2006; 34 (Web Server issue),W686-W691.

Galvagni F, Cantini M, and Oliviero S (2002). The Utrophin Gene Is Transcriptionally Up-regulated in Regenerating Muscle. Journal of Biological Chemistry, 277 (21), 19106–19113.

Gardiner-Garden M and Frommer M (1987). CpG islands in vertebrate genomes. Journal of Molecular Biology.196 (2), 261-282.

Gardiner-Garden M, Frommer M (1994). Transcripts and CpG islands associated with the pro-opiomelanocortin gene and other neurally expressed genes. Journal of Molecular Endocrinology 12, 366-382.

Geiger A, Decaux J F, Burcelin R, Le Cam A, Salazar G, Charron MJ, Girard J and Kervran A (2000). Structural and Functional Characterizations of the 5' Flanking Region of the Mouse Glucagon receptor Gene, Comparison with the Rat Gene. Biochemical and Biophysical Research Communications 272**,** 912–921.

Geiger A, Salazar G and Kervran A (2001). Role of the Sp family of transcription factors on glucagon receptor gene expression. Biochemical and Biophysical Research Commuications. 285, 838–844

Ghatei MA; Takahashi K; Suzuki Y; Gardiner J; Jones PM; Bloom SR. (1993). Distribution, molecular characterization of pituitary adenylate cyclase-activating polypeptide and its precursor encoding messenger RNA in human and rat tissues. Journal of Endocrinology, 136, 159-166.

Gourlet P, Vertongen P, Vandermeers A, Vandermeers-Piret MC, Rathe J, De Neef P and Robberecht P (1997). The long acting vasoactive Intestinal polypeptide agonist Ro 25-1553 is highly selective of the VIP2 receptor subclass. Peptides, 18, 403-408.

Gozes I and Brenneman DE (1989). VIP, Molecular biology and neurobiological function. Molecular Neurobiology, 3, 202–235.

Gozes I, Fridkinb, M, Hill J.M and Brenneman D.E, (1999), Pharmaceutical VIP, prospects and problems, Current Medicinal Chemistry, 6, 1019–1034.

Grammatopoulos, D.K, Dai, Y., Randeva, H.S., Levine, M.A., Karteris, E., Easton, A.J., Hillhouse, E.W. (1999). A novel spliced variant of the type 1 corticotropin-releasing hormone receptor with a deletion in the seventh transmembrane domain present in the human pregnant term myometrium and fetal membranes, Molecular Endocrinology. 13, 2189–2202.

Gressens P, Hill JM, Gozes I (1993). Growth factor function of vasoactive intestinal peptide in whole cultured mouse embryos. Nature, 362, 155–158.

Grinninger C, Wang W, Oskoui KB, Voice JK and Goetzl EJ (2004). A Natural Variant Type II G Protein-coupled Receptor for Vasoactive Intestinal Peptide with Altered Function. Journal of Biological Chemistry, 279, 40259-40262

Groneberg D A, Hartmann P, Dinh Q T, and Fischer A (2001). Expression and Distribution of Vasoactive Intestinal Polypeptide Receptor VPAC2 mRNA in Human Airways. Laboratory Investigation 81 (5), 749-755.

Handen JS, Rosenberg HF (1997). Intronic enhancer activity of the eosinophil derived neurotoxin (RNS2) and eosinophil cationic protein (RNS3) genes is mediated by an NFAT-1 consensus binding sequence. Journal of Biological Chemistry, 272:1665–1669.

Harmar A J, Arimura A, Gozes I, Journot L, Laburthe M and Pisegna J R, Rawlings S R, Robberecht P, Said S I, Sreedharan S P, Wank S A, Washchek J A (1998). Nomenclature of receptors for vasoactive intestinal peptide and pituitary adenylate-activating polypeptide. Pharmacological Reviews, 50, 265–270.

Harmar A J (2001). Family-B G-protein-coupled receptors. Genome Biology, 2 (12), 3013.1-3013.10.

Harmar AJ, Marston HM, Shen S, Spratt C, West KM, Sheward WJ, Morrison CF, Dorin JR, Piggins HD, Reubi JC, Kelly JS, Maywood ES, Hastings MH. (2002). "The VPAC2 receptor is essential for circadian function in the mouse suprachiasmatic nuclei.". Cell, 109 (4), 497–508.

Hampsey, M., (1998). Molecular genetics of the RNA polymerase II general transcriptional machinery. Microbiology and Molecular Biology Reviews, 62, 465–503.

Hendrich B and Bird A (1998). Identification and Characterization of a Family of Mammalian Methyl-CpG Binding Proteins. Molecular Cellular Biology, 18 (11), 6538–6547.

Hellqvist M, Mahlapuu M, Samuelsson L, Enerba S, and Carlsson P (1996). Differential Activation of Lung-specific Genes by Two Forkhead Proteins, FREAC-1 and FREAC-2. The Journal of biological chemistry, 271 (8), 4482–4490.

Helmer K S, Chang L, Cui Y, and Mercer D W (2002). Induction of NF-kB, IkB alpha, and iNOS in Rat Gastric Mucosa during Endotoxemia. Journal of Surgical Research, 104**,** 46–52.

Hezareh M, Journot L, Bépoldin L, Schlegel W, Rawlings SR (1996) PACAP/VIP receptor subtypes, signal transducers, and effectors in pituitary cells. Annals of Newyork Academy of Sciences, 805, 315-327

Hirose M, Hashimoto H, Shintani N, Nakanishi M, Arakawa N, Iga J, Niwa H, Miyazaki J, Baba A (2005). Differential expression of mRNAs for PACAP and its receptors during neural differentiation of embryonic stem cells. Regulatory peptides, 126 (1-2), 109-113.

Ho P.K, Rango S.M, Fongo, Heidi S.T. Kai, Elisa H.Y. Lau, Elly S.W. Ngan, Calvin U, Cotton, Billy K.C. Chow (1999). The human secretin receptor gene, genomic organization and promoter characterization. Federation of European Biochemical Societies Letters, 455, 200-214.

Ho, SH; So, GMK; Chow, KL (2001). "Postembryonic expression of Caenorhabditis elegans mab-21 and its requirement in sensory ray differentiation". Developmental Dynamics, 221 (4), 422–430.

Horton J D, Goldstein J L and Brown M S (2002). SREBPs, activators of the complete program of cholesterol and fatty acid synthesis in the liver. Journal of Clinical Investigation, 109, (9) 1125–1131

Howard ML, Davidson EH (2004). cis-Regulatory control circuits in development. Developmental Biology, 271,109–118.

Hughes JE, Srinivasan S, Lynch KR, Proia RL, Ferdek P, Hedrick CC (2008). Sphingosine-1-phosphate induces an anti-inflammatory phenotype in macrophages. Circulation Research 102 (8), 950-958.

Inagaki, N., Yoshida, H., Mizuta,M., Mizuno, N., Fujii, Y., Gonoi, T., et al. (1994). Cloning and functional characterization of a third pituitary adenylate cyclase-activating polypeptide receptor subtype expressed in insulin-secreting cells. Proceedings of National Academy of Sciences United States of America, 91, 2679−2683.

Itoh N, Obata K, Yanaihara N and Okamoto H (1983). Human pre-pro-vasoactive intestinal polypeptide contains a novel PHI-27- like peptide, PHM-27. Nature,. 304, (5926), 547–549.

Jenuwein T, Forrester WC, Fernandez Herrero LA, Laible G, Dull M, Grosschedl R.(1997). Extension of chromatin accessibility by nuclear matrix attachment regions. Nature, 85, 269 –272.

Jeon, JS; Lee, S; Jung, KH; Jun, SH; Kim, C; An, G (2000). "Tissue-Preferential Expression of a Rice α-Tubulin Gene, OsTubA1, Mediated by the First Intron". Plant Physiology 123 (3), 1005–1014.

Jonsson, JJ; Foresman, MD; Wilson, N; McIvor, RS (1990). "Intron requirement for expression of the human purine nucleoside phosphorylase gene". Nucleic Acids Research, 20 (12), 3191–3198.

Jossic-Corcos C L, Duclos S, Ramirez L C, Zaghini I, Chevillarda G, Martinb P, Pineaub T, Bournot P (2004). Effects of peroxisome proliferator-activated receptor a activation on pathways contributing to cholesterol homeostasis in rat hepatocytes. Biochimica et Biophysica Acta, 1683, 49– 58.

Joyeux, A., Balauer, P., Germain, P., Boussioux, A.M., Pons, M., Nicholas, J.C., (1997). Engineered cell lines as a tool for monitoring biological activity of hormone analogs. Analytical Biochemistry, 249, 119–130.

Karin M, Ben-Neriah Y (2000). Phosphorylation meets ubiquitination, the control of NF-[kappa] B activity. Annual Review of Immunology, 18,621-663.

Kim J B and Spiegelman B M (1996). ADD1/SREBP1 promotes adipocyte differentiation and gene expression linked to fatty acid metabolism. Genes and Development, 10, 1096-1107

Kim, Y.J., Björklund, S., Li, Y., Sayre, M.H. and Kornberg, R.D. (1994) A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. Cell 77, 599-608.

Kim, J.L., Nikolov, D.B. and Burley, S.K. (1993) Co-crystal structure of TBP recognizing the minor groove of a TATA element. Nature 365, 520-527.

Klaman LD, Thorley-Lawson DA (1995). Characterization of the CD48 gene demonstrates a positive element that is specific to Epstein-Barr virus-immortalized B-cell lines and contains an essential NF-kappa B site. Journal of Virology, 69(2), 871-81.

Kochetov A V (2008). Alternative translation start sites and hidden coding potential of eukaryotic mRNAs. Bioassays, 30, 683-691.

Kolakowski LF Jr (1994). GCRDb, a G-protein-coupled receptor database. Receptors Channels 2 (1), 1–7.

Kornberg RD (2007). "The molecular basis of eukaryotic transcription". Proceedings of National Academy of Sciences United States of America, 104 (32), 12955–61.

Kraus J, Börner C, Giannini E, Höllt V.(2003). The role of nuclear factor kappaB in tumor necrosis factor-regulated transcription of the human mu-opioid receptor gene. Molecular Pharmacology, 64 (4), 876-884.

Kulkarni-Narla A, Beitz AJ, and Brown DR. (1999). Catecholaminergic, cholinergic and peptidergic innervation of gut-associated lymphoid tissue in porcine jejunum and ileum. Cell and Tissue Research, 298, 275–286.

Kundu TK, Rao MR. (1999). CpG islands in chromatin organization and gene expression. Journal of Biochemistry, 125 (2), 217-222.

Knuppel, R., Dietze, P., Lehnberg, W., Frech, K. and Wingender, E. (1994). TRANSFAC retrieval program: a network model database of eukaryotic transcription regulating sequences and proteins. Journal of Computational Biology, 1, 191–198.

Laburthe M, Couvineau A and Marie J C, (2002). VPAC receptors for VIP and PACAP, Receptors Channels. 8, 137–153.

Laburthe M, Kitabgi P, Couvineau A and Amiranoff B (1993). Peptide receptors and signal transduction in the digestive tract. Handbook of Experimental Pharmacology, 106, 133–176.

Laburthe M, Couvineau A, Gaudin P, Maoret J J, Fessard R C and Nicole P (1996). Receptors for VIP, PACAP, secretin, GRF, glucagon, GLP-1, and other members of their new family of G protein-linked receptors, structure-function relationship with special reference to the human VPAC1 receptor. Annals of New York Academy of Sciences, 805, 94–109.

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W et al (2001). The International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. Nature, 409**,** 860–921.

Lara-Marquez ML, O'Dorisio MS, Karacay B (2000). Vasoactive intestinal peptide (VIP) receptor type 2 (VPAC2) is the predominant receptor expressed in human thymocytes. Annals of New York Academy of Science, 921, 45-54.

Lara-Marquez M, O'Dorisio M, O'Dorisio T, Shah M, Karacay B (2001) Selective gene expression and activation-dependent regulation of vasoactive intestinal peptide receptor type 1 and type 2 in human T cells. Journal of Immunology, 166 (4), 2522-2530.

Larsen, F., Gundersen, G., Lopez, R. and Prydz, H. (1992). CpG islands as gene markers in the human genome. Genomics, 13, 1095-1107.

Lee S, Xu H, and Montell C (2004). Rhodopsin kinase activity modulates the amplitude of the visual response in Drosophila. Proceedings of National Academia of Science. United States of America. 101 (32), 11874–11879.

Lewin B Genes 9th edition (2008) Jones & Bartlett publishers.

Levinson DF, Duan J, Oh S, Wang K, Sanders AR, Shi J, Zhang N, Mowry BJ, Olincy A, Amin F, Cloninger CR, Silverman JM, Buccola NG, Byerley WF, Black DW, Kendler KS, Freedman R, Dudbridge F, Pe'er I, Hakonarson H, Bergen SE, Fanous AH, Holmans PA, Gejman PV (2011). "Copy Number Variants in Schizophrenia: Confirmation of Five Previous Findings and New Evidence for 3q29 Microdeletions and VIPR2 Duplications". American Journal of Psychiatry, 168 (3), 302–316

Li Q, Notkins AL, Lan MS (1997). Molecular characterisation of the promoter region of neuroendocrine tumour marker IA-1. Biochemical and Biophysical Research Communication, 236 (3), 776-781.

Li X X and Liao W S (1991). Expression of rat serum amyloid A1 gene involves both C/EBP-like and NF kappa B-like transcription factors. The Journal of Biological Chemistry, 266, 15192-15201.

Lin Y.J., Seroude L., Benzer S (1998). Extended life-span and stress resistance in the Drosophila mutant methuselah. Science, 282, 943–946.

Loots GG (2008) Genomic identification of regulatory elements by evolutionary sequence comparison and functional analysis. Advances in genetics, 61, 269-293

Lu S Y, Rodriguez M, and Liao W S (1994). YY1 represses rat serum amyloid A1 gene transcription and is antagonized by NF-kappa B during acute-phase response. Molecular and Cellular Biology, 14 (9), 6253–6263.

Lutz EM, Sheward WJ, West KM, Morrow JA, Fink G and Harmar AJ (1993). The VIP2 receptor, Molecular characterisation of a cDNA encoding a novel receptor for vasoactive intestinal peptide. Federation of European Biochemical Societies Letters, 334, 3–8.

Lutz EM, Shen S, Mackay M, West K, Harmar AJ (1999). Structure of the human VIPR2 gene for vasoactive intestinal peptide receptor type 2 Federation of European Biochemical Societies Letters, 458 (2), 197-203.

Macleod D, Charlton J, Mullins J, Bird A (1994). Sp1 sites in the mouse aprt gene promoter are required to prevent methylation of the CpG island. Genes and Development, 8, 2282-2292.

Macleod D, Ali R R and Bird A (1998) An Alternative Promoter in the Mouse Major Histocompatibility Complex Class II I-Aβ Gene: Implications for the Origin of CpG Islands. Molecular and Cellular Biology 18(8), 4433-4443.

MacKenzie CJ, Lutz EM, Johnson MS, Robertson DN, Holland PJ, Mitchell R (2001) Mechanisms of phospholipase C activation by the vasoactive intestinal polypeptide/pituitary adenylate cyclase-activating polypeptide type 2 receptor. Endocrinology, 142(3):1209-17.

Macaya G, Thiery JP, Bernardi G. (1976). An approach to the organization of eukaryotic genomes at a macromolecular level. Journal of Molecular Biology, 108, 237–254.

Maget B, Tastenoy M and Svoboda M, (1994), Sequencing of eleven introns in genomic DNA encoding rat glucagon receptor and multiple alternative splicing of its mRNA. Federation of European Biochemical Societies Letters, 351, 271–275.

Majewski J and Ott J (2002). Distribution and Characterization of Regulatory Elements in the Genome. Genome research, 12, 1827–1836.

Malbon CC, Wang H, Moon RT. (2001). Wnt signaling and heterotrimeric G-proteins, strange bedfellows or a classic romance? Biochemical and Biophysical Research Communications, 287, 589-593.

Malhotra RK, Wakade TD and Wakade AR (1988). Vasoactive intestinal polypeptide and muscarine mobilize intracellular Ca2+ through breakdown of phosphoinositides to induce catecholamine secretion. Role of IP3 in exocytosis. Journal of Biological Chemistry, 263, 2123–2126.

Manen D, Pougeon M, Damay P and Geiselmann J (1997). A sensitive reporter gene system using bacterial luciferase based on a series of plasmid cloning vectors compatible with derivatives of pBR322. Gene 186, 197–200.

Margana RK, Boggaram V (1996). Rabbit surfactant protein B gene: structure and functional characterization of the promoter. American Journal of Physiology, 270 (4 Pt 1), L601-L612.

Margana RK and Boggaram V (1997). Functional Analysis of Surfactant Protein B (SP-B) Promoter Sp1, Sp3, TTF-1, and HNF-3a Transcription factors are necessary

for lung cell specific activation of SP-B gene transcription. The Journal of biological chemistry, 272 (5), 3083–3090.

Mayor C, Brudno M, Schwartz J R, Poliakov A, Rubin E M, Frazer KA, Pachter LS and Dubchak I (2000). VISTA, Visualizing global DNA sequence alignments of arbitrary length. Bioinformatics, 16 (11),1046-1047.

McCuaig, K A, Clarke J C and White J H (1994). Molecular cloning of the gene encoding the mouse parathyroid hormone/parathyroid hormone-related peptide receptor. Proceedings of National academy of Sciences, United States of America, 91, 5051-5055.

McCuaig KA, Lee HS, Clarke JC, Assar H, Horsford J, White JH (1995). Parathyroid hormone/parathyroid hormone related peptide receptor gene transcripts are expressed from tissue-specific and ubiquitous promoters, 23 (11), 1948-1955.

McEwen D G and Ornitz DM (1998). Regulation of the Fibroblast Growth Factor Receptor 3 Promoter And Intron I Enhancer by Sp1 Family Transcription Factors. The Journal of Biological Chemistry, 273 (9), 5349–5357.

McLellan AS, Kealey T, Langlands K (2006). An E box in the exon 1 promoter regulates insulin-like growth factor-I expression in differentiating muscle cells. American Journal of Physiology Cell Physiology, 291 (2), C300-C307.

Merino VF, Silva JA Jr, Araújo RC, Avellar MC, Bascands JL, Schanstra JP, Paiva AC, Bader M, Pesquero JB. (2005). Molecular structure and transcriptional regulation by nuclear factor-kappaB of the mouse kinin B1 receptor gene. Journal of Biological chemistry, 386 (6), 515-22

Miller AL, Verma D, Grinninger C, Huang MC, Goetzl EJ (2006). Functional splice variants of the type II G protein-coupled receptor (VPAC2) for vasoactive intestinal peptide in mouse and human lymphocytes. Annals of the New York Academy of Sciences, 1070, 422-426.

Minagawa M, Kwan MY, Bettoun JD, Mansour FW, Dassa J, Hendy G N, Goltzman D and White J H **(**2000). Dissection of differentially regulated (G+C)-rich promoters of the human parathyroid hormone (PTH)/PTH-related peptide receptor gene. Endocrinology, 141, 2410–2421

Miyata A, Arimura A, Dahl RR, Minamino N, Uehara A, Jiang L, Culler MD and Coy DH (1989). Isolation of a novel 38 residue-hypothalamic polypeptide which stimulates adenylate cyclase in pituitary cells. Biochemical and Biophysical Research Communications, 164, 567–574.

Miyata A, Jiang L, Dahl RD, Kitada C, Kubo K, Fujino M, Minamino N and Arimura A (1990). Isolation of a neuropeptide corresponding to the N-terminal 27 residues of the pituitary adenylate cyclase-activating polypeptide with 38 residues (PACAP38). Biochemical Biophysical Research Communications, 170, 643–648.

Morin S, Pozzulo G, Robitaille L, Cross J and Nemer M (2005). MEF2-dependent Recruitment of the HAND1 Transcription Factor Results in Synergistic Activation of Target Promoters. Journal of Biological Chemistry 280, 32272-32278.

Morrison R F and Farmer S R (2000). Hormonal Signalling and Transcriptional Control of Adipocyte Differentiation Journal of Nutrition. 130, 3116S-3121S

Mount D.W. 2004. Bioinformatics. Cold Spring Harbor Laboratory Press, 692.

Nagy P, Bisgaard HC, and Thorgeirsson SS (1994). Expression of hepatic transcription factors during liver development and oval cell differentiation. Journal of Cell Biology, 126, 223–233.

Naylor LH (1999). Reporter gene technology: the future looks bright. Biochemical Pharmacology, 58, 749-757.

Ni A, Chao L, Chao J (1998). Transcription factor nuclear factor kappaB regulates the inducible expression of the human B1 receptor gene in inflammation. Journal of Biological Chemistry, 273 (5), 2784-2791.

Nordström, K.J.V., Lagerström, M.C., Wallér, L.M.J., Fredriksson, R., Schiöth, H.B. (2009). The Secretin GPCRs descended from the family of Adhesion GPCRs, Molecular Biology and Evolution, 26, 71–84

Novina C D and Roy AL (1996). Core promoters and transcriptional control. Trends in genetics, 12, (9) 351-355.

Odenwald, W. F., Rasband, W., Kuzin, A. and Brody, T. (2005). *EvoPrinter*, A multigenomic comparative tool for rapid identification of functionally important DNA. Proceedings of National. Academy of Sciences. 102, 14700-14705.

Ogi K, Miyamoto Y, Masuda Y, Habata Y, Hosoya M, Ohtaki T, Masuo Y, Onda H and Fujino M (1993). Molecular cloning and functional expression of a cDNA encoding a human pituitary adenylate cyclase-activating polypeptide receptor. Biochemical and Biophysical Research Communications, 196, 1511–1521.

Oguchi H, Kimura S (1995). Transcription factors and thyroid disease. Molecular Medicine, 32, 518-526

Ohtaki T, Masuda Y, Ishibashi Y, Kitada C, Arimurasa A, and Fujino M (1993). Purification and Characterization of the Receptor for Pituitary Adenylate Cyclase-activating Polypeptide. The Journal of Biological Chemistry, 268 (35), 26650-26657.

Ott CJ, Blackledge1 NP , Leir SH, and Harris A (2009). Novel regulatory mechanisms for the CFTR gene. Biochemistry Society Transactions, 37 (Pt 4), 843–848.

Ottaway CA (1987). Selective effects of vaso-active intestinal peptide on the mitogenic response of murine T cells.. Immunology, 62, 291–297.

Ovcharenko I, Loots G, Hardison R, Miller W and Stubbs L (2004). zPicture, Dynamic Alignment and Visualization Tool for Analyzing Conservation Profiles. Genome Research, 14, 472-477.

Palamarchuk A, Zanesi N, Aqeilan R I, Efanov A, Maximov V, Santanam U, Hagan J P, Croce C M, and Pekarsky Y (2006). Tal1 Transgenic Expression Reveals Absence of B Lymphocytes. Cancer Research, 66 (12), 6014-6017.

Palmiter, RD; Sandgren, EP; Avarbock, MR; Allen, DD; Brinster, RL (1991). "Heterologous introns can enhance expression of transgenes in mice". Proceedings of National academy of Sciences, 88 (2), 478–482.

Parham K.L., Zervou S., Karteris E., Catalano R.D., Old R.W and Hillhouse E.W (2004). Promoter analysis of human corticotrophin releasing factor type-1 receptor and regulation by CRF and Urocortin. Endocrinology, 145, 3971-3983.

Pazzagli M, Devine J H, Peterson D.O and Baldwin T.O. (1992). Use of bacterial and firefly luciferases as reporter genes in DEAE-dextran-mediated transfection of mammalian cells. Analytical Biochemistry, 204**.** 315–323.

Pearse A G E, Polak J M, Bloom S R (1977). The newer gut hormones, cellular sources, physiology and clinical aspects. Gastroenterology, 72, 746-761.

Pedersen A G and Nielsen H (1997). Neural network prediction of translation initiation sites in eukaryotes, perspectives for EST and genome analysis. Institute of Structural Molecular Biology, 5, 226-233.

Petersenn.S, Rasch A C, Heyens M and Schulte H M (1998). Structure and regulation of the human growth hormone releasing hormone receptor gene. Molecular Endocrinology, 12 (2) 233-247.

Pierce RA, Moore CH, and Arikan M C (2006). Positive transcriptional regulatory element located within exon 1 of elastin gene. American Journal of Physiology - Lung cellular and molecular Physiology, 291(3), L391-L399.

Porntadavity S, Xu Y, Kiningham K, Rangnekar VM, Prachayasittikul V, St Clair DK. (2001). TPA-activated transcription of the human MnSOD gene: role of transcription factors Sp-1 and Egr-1. DNA and Cell Biology, 20 (8), 473-481.

Przywara DA, Guo X, Angelilli ML, Wakade TD and Wakade AR (1996). Pituitary Adenylyl Cyclase-Activating Polypeptide and Nerve Growth Factor Use the Proteasome to Rescue Nerve Growth Factor-Deprived Sympathetic Neurons Cultured from Chick Embryos. Journal of Biological Chemistry, 271,10545–10550.

Pugh B.F and Tjian R (1990). Mechanism of transcriptional activation by Sp1, evidence for co-activators. Cell, 61, 1187–1197.

Pugh B.F and Tjian R (1991). Transcription from a TATA less promoter requires a multisubunit TFIID complex. Genes and Development, 5, 1935-1945.

Rachdi L, Marie JC, Scharfmann R (2003). Role for VPAC2 receptor-mediated signals in pancreas development. Diabetes. 52 (1), 85-92.

Rawlings S.R and Hezareh M, (1996). Pituitary adenylate cyclase-activating polypeptide (PACAP) and PACAP/vasoactive intestinal polypeptide receptors, actions on the anterior pituitary gland. Endocrinology Reviews, 17, 4–29.

Ray DW, Ren SG, Melmed S.(1996). Leukemia inhibitory factor (LIF) stimulates proopiomelanocortin (POMC) expression in a corticotroph cell line. Role of STAT pathway. Journal of Clinical Investigation, 97 (8), 1852-1859.

Reese MG (2001). Application of a time-delay neural network to promoter annotation in the Drosophila melanogaster genome. Computers and Chemistry, 26 (1), 51-56.

Reichlin S (1988). Neuroendocrine significance of vasoactive intestinal polypeptide. Annals of New York Academy of Sciences, 527, 431-449.

Rose, AB; Last, RL (2003). "Introns act post-transcriptionally to increase expression of the Arabidopsis thaliana tryptophan pathway gene PAT1". The Plant Journal 11 (3), 455–464.

Ryan A K, Rosenfeld MG (1997). POU domain family values: flexibility, partnerships, and developmental codes. Genes and Development, 11, 1207-1225.

Sabourin JC , Kern AS, Gregori C, Porteu A, Cywiner C, Chatelet FP, Kahn A, and Pichard AL (1996). An Intronic Enhancer Essential for Tissue-specific Expression of the Aldolase B Transgenes. Journal of Biological Chemistry, 271 (7), 3469–3473.

Said S I and Mutt V (1970). Polypeptide with broad biological activity, Isolation from small intestine. Science, 169**,** 1217–1218.

Said SI. Vasoactive intestinal peptide (1986). Journal of Endocrinological Investigation, 9, 191–200.

Said S I and Mutt V (1972). Isolation from porcine-intestinal wall of a vasoactive octacosapeptide related to secretin and to glucagon. European Journal of Biological, Chemistry 28**,** 199–204.

Said S I (1991), Vasoactive intestinal polypeptide: Biological role in health and disease. Trends in endocrine and metabolism, 2, 107-112.

Said S I (1996). Vasoactive intestinal peptide and nitric oxide, Divergent roles in relation to tissue injury. Annals of New York Academy of Sciences, 805**,** 379–388.

Said S I. (2000). VIP and PACAP in pain and inflammation. Trends in Pharmacological Sciences, 21:57.

Sandelin A, Carninci P, Lenhard B, Ponjavic J, Hayashizaki Y, Hume DA, (2007). Mammalian RNA polymerase II core promoters, insights from genome-wide studies. Natural Review Genetics, 8 (6), 424-436.

Schwartz C.J., Kimberg D.V., Sheerin H.E., Field M, Said S.I (1974). Vasoactive intestinal peptide stimulation of adenylate cyclise and active electrolyte secretion in intestinal mucosa. Journal of Clinical Investigation, 54 (3), 536-544.

Schweizer-Groyer G, Cadepond F, Jibard N, Neau E, Segard-Maurel I, Baulieu E E, and Groyer A (1997). Stimulation of transcription in vitro from a liver-specific promoter by human glucocorticoid receptor (hGRalpha). Biochemistry Journal, 324, 823–831.

Smale S T and Baltimore D (1989). The ‘initiator’ as a transcription control element. Cell, 57, 103–113.

Smale, S.T., (1997). Transcriptional initiation from TATA-less promoters within eukaryotic protein-encoding genes. Biochimica. Biophysica Acta, 1351, 73–88.

Smale S T and Kadonaga, T (2003). "The RNA polymerase II core promoter". Annual review of biochemistry, 72, 449–479

Sorg O and Magistretti PJ (1992). Vasoactive intestinal peptide and noradrenaline exert long-term control on glycogen levels in astrocytes, Blockade by protein synthesis inhibition. Journal of Neuroscience, 12**,** 4923–4931.

Spitzer JA, Zheng M, Kolls JK, Vande Stouwe C, Spitzer JJ.(2002). Ethanol and LPS modulate NF-kappaB activation, inducible NO synthase and COX-2 gene expression in rat liver cells in vivo. Frontiers in Bioscience, 7, a99- a108.

St Clair DK, Porntadavity S, Xu Y, Kiningham K (2002). Transcription regulation of human manganese superoxide dismutase gene. Methods in Enzymology, 349, 306-312.

Staal FJ, Meeldijk J, Moerer P, Jay P, van de Weerdt BC, Vainio S, Nolan GP, Clevers H (2001). Wnt signaling is required for thymocyte development and activates Tcf-1 mediated transcription. European Journal of Immunology, 31 (1), 285-293.

Steel G, Lutz EM (2007). Characterisation of the mouse vasoactive intestinal peptide receptor type 2 gene, Vipr2, and identification of a polymorphic LINE-1-like sequence that confers altered promoter activity. Journal of Neuroendocrinology, 19 (1), 14-25.

Suto C.M and Ignar D.M (1997). Selection of an optimal reporter gene for cell-based high throughput screening assays, Journal of Biomolecular Screening, 2,7–9.

Su J, Zhang Y, Lv J, Liu H, Tang X, Wang F, Qi Y, Feng Y and Li X (2009). CpG_MI: a novel approach for identifying functional CpG islands in mammalian genomes. Nucleic acids research, 38(1), e6.

Sun L, Liu A and Georgopoulos K (1996). Zinc finger-mediated protein interactions modulate Ikaros activity, a molecular control of lymphocyte Development. The EMBO Journal, 15 (19) 5358-5369.

Sun S, Hong J, Ying C, Zang Q, Liu X and Zhang J Z (2006). Altered expression of vasoactive intestinal peptide receptors in T lymphocytes and aberrant Th1immunity in multiple sclerosis. International Immunology, 18 (12), 1691–1700

Svoboda, M., Ciccarelli, E., Tastenoy, M., Robberecht, P., and Christophe, J. (1993). A cDNA construct allowing the expression of rat hepatic glucagon receptors. Biochemical Biophysical Research Communications, 192**,** 135–142.

Svoboda, M., Tastenoy, M., Van Rampelbergh, J., Goossens, J. F., De Neef, P., Waelbroeck,M., et al. (1994). Molecular cloning and functional characterization of a human VIP receptor from SUP-T1 lymphoblasts. Biochemical Biophysical Research Communication, 205, 1617−1624.

Takai D, Jones PA (2002). Comprehensive analysis of CpG islands in human chromosomes 21 and 22. Proceedings of National Academy of Sciences United States of America,. 99 (6), 3740–3745.

Tang T, Shi Y, Opalenik S R, Brantley-Sieders D M, Chen J, Davidson J M and Brandt S J (2006). Expression of the TAL1/SCL transcription factor in physiological and pathological vascular processes. Journal of Pathology, 210, 121–129

Terzidou V, Lee Y, Lindström T, Johnson M, Thornton S, Bennett PR (2006). Regulation of the human oxytocin receptor by nuclear factor-kappaB and CCAAT/enhancer-binding protein-beta. Journal of Clinical Endocrinology and Metabolism, 91 (6), 2317-2326

Tischkau S A, Mitchell J W, Tyan S, Buchanan GF and Gillette M U (2003). Ca^{2+}/cAMP Response Element-binding Protein (CREB)-dependent Activation of*Per1* Is Required for Light-induced Signaling in the Suprachiasmatic Nucleus Circadian Clock. The Journal of Biological Chemistry, 278, 718-723

Thompson J.D., Higgins D.G., Gibson T.J (1994). CLUSTAL W, improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research, 22 (22), 4673–4680.

Tone Y, Kojima Y, Furuuchi K, Brady M, Yashiro-Ohtani Y, Tykocinski ML, Tone M (2007). OX40 gene expression is up-regulated by chromatin remodeling in its promoter region containing Sp1/Sp3, YY1, and NF-kappa B binding sites. Journal of Immunology. 179 (3),1760-1767

Travis A, Amsterdam A, Belanger C, and Grosschedl R (1991) LEF-1, a gene encoding a lymphoid-specific protein with an HMG domain, regulates T-cell receptor alpha enhancer function. Genes & Development, 5, 880-894.

Tsunoda T and Takagi T (1999). Estimating transcription factor bindability on DNA. Bioinformatics, 15 (7/8) 622-630.

Umezawa A, Yamamoto H, Rhodes K, Klemsz MJ, Maki RA, Oshima RG (1997). Methylation of an ETS site in the intron enhancer of the keratin 18 gene participates in tissue-specific repression. Molecular Cell Biology, 17, 4885–4894.

Usdin TB, Bonner TI, Mezey E (1994).Two receptors for vasoactive intestinal polypeptide with similar specificity and complementary distributions. Endocrinology, 135 (6), 2662-2680.

Unson, C. G., Cypess, A. M., Kim, H. N., Goldsmith, P. K., Carruthers, C. J. L.,Merrifield, R. B., & Sakmar, T. P. (1995). Characterization of Deletion and Truncation Mutants of the Rat Glucagon Receptor. Seven Transmembrane Segments are Necessary for Receptor Transport to the Plasma Membrane and Glucagon Binding. Journal of Biological Chemistry, 270, 27720-27727.

Vacic V, McCarthy S, Malhotra D, Fiona Murray, Hsun-Hua Chou, Aine Peoples, Vladimir Makarov, Seungtai Yoon, Abhishek Bhandari, Roser Corominas, Lilia M. Iakoucheva, Olga Krastoshevsky, Verena Krause, Vero´nica Larach-Walters, David K. Welsh, David Craig, John R. Kelsoe, Elliot S. Gershon, Suzanne M. Leal, Marie Dell Aquila, Derek W. Morris, Michael Gill, Aiden Corvin, Paul A. Insel, McClellan J, King MC, Karayiorgou M, Levy D L, DeLisi LE & Sebat J(2011) Duplications of the neuropeptide receptor gene VIPR2 confer significant risk for schizophrenia. Nature (471), 499–501

Vaudry D, Gonzalez B J, Basille M, Yon L, Fournier A and Vaudry H, (2000). Pituitary adenylate cyclase-activating polypeptide and its receptors, from structure to functions. Pharmacological Reviews, 52, 269–324.

Vaudry H and Laburthe M (2006). VIP, PACAP, and related peptides. From gene to therapy. Annals of NewYork Academy of Sciences, 1070, 17-18.

Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, et al. (2001). The sequence of the human genome. Science 291, 1304–1351.

Voigtlander T, Ganten D, Bader M (1999). Transcriptional Regulation of the Rat Renin Gene by Regulatory Elements in Intron I. Hypertension, 33 (part II), 303-311.

Waschek JA (1995). Vasoactive intestinal peptide: an important trophic factor and developmental regulator? Developmental Neuroscience, 17, 1– 7.

Wang X, Briggs MR, Hua X, Yokoyama C, Goldstein JL, Brown MS (1993). Nuclear protein that binds sterol regulatory element of low density lipoprotein receptor promoter. II. Purification and characterization. Journal of Biological Chemistry, 268, 14497-14504.

Wang X, Yang Y, Adamo ML (1997). Characterization of the rat insulin-like growth factor I gene promoters and identification of a minimal exon 2 promoter. , 138 (4), 1528-1536.

Wang Z, Lu Z, Guo A (2000). The role of 5'proximal fragment of POMC gene in the regulation of transcription in AtT20 cells. Zhongguo Yi Xue Ke Xue Yuan Xue Bao, 22 (2), 144-148.

Welsh.S and Kay.SA (1997). Reporter gene expression for monitoring gene transfer. Current opinion in Biotechnology, 8, 617-622.

Wehr, R and Gruss, P (1996). Pax and vertebrate development. International Journal of Developmental Biology, 40, 369-377.

Williams D A, Nienhuis A W, Hawley R G, and Smith F O (2000). Gene therapy. Hematology, 1, 376-393

Wingender E (2008). The TRANSFAC project as an example of framework technology that supports the analysis of genomic regulation. Briefings in Bioinformatics, 9 (4), 326-332.

Wisor JP, Takahashi JS (1997). Regulation of the VGF gene in the golden hamster suprachiasmatic nucleus by light and by the circadian clock. Journal of Comparative Neurology, 378 (2), 229-238

Wood KV (1995). Marker proteins for gene expression. Current Opinion in Biotechnology, 6, 50–58.

Wray G A, Hahn M.W, Abouheif E, Balhoff J P, Pizer M, Rockman M.V and Romano L A (2003). The Evolution of Transcriptional Regulation in Eukaryotes. Molecular Biology Evolution, 20 (9) 1377-1419.

Yada T, Sakurada M, Ihida K, Nakata M, Murata F, Arimura A and Kikuchi M (1994). Pituitary adenylate cyclase-activating polypeptide is an extraordinarily potent intra-pancreatic regulator of insulin secretion from islet beta-cells. Journal of Biological Chemistry, 269**,**1290–1293.

Yavatkar A S, Lin Y, Ross J, Fann Y, Brody T and Odenwald W F (2008). Rapid detection and curation of conserved DNA via enhanced-BLAT and EvoPrinterHD analysis. Biomedical Central Genomics, 9**,**106

Yokoyama C, Wang X, Briggs MR, Admon A, Wu J, Hua X, Goldstein JL, Brown MS (1993). SREPB-1, a basic-helix-loop-helixleucine zipper protein that controls transcription of the low density lipoprotain receptor gene. Cell, 75, 187-197.

Young, D.C., Kingsley S.D, Ryan, K.A and Dutko, R.J (1993). "Selective Inactivation of Eukaryotic B-Galactosidase in Assays for Inhibitors of HIV-1 TAT Using Bacterial b-Galactosidase as a Reporter Enzyme". Analytical Biochemistry, 215, 24-30.

Yoshikawa K, Aizawa T (1988). Enkephalin precursor gene expression in postmeiotic germ cells. Biochemical and Biophysical Research Communications, 51, 664-671.

Zhang L, Kasif S, Cantor C R and Broude N E (2004). GC/AT-content spikes as genomic punctuation marks. Proceedings of National Academy of Sciences United States of America, 101 (48), 16855–16860.

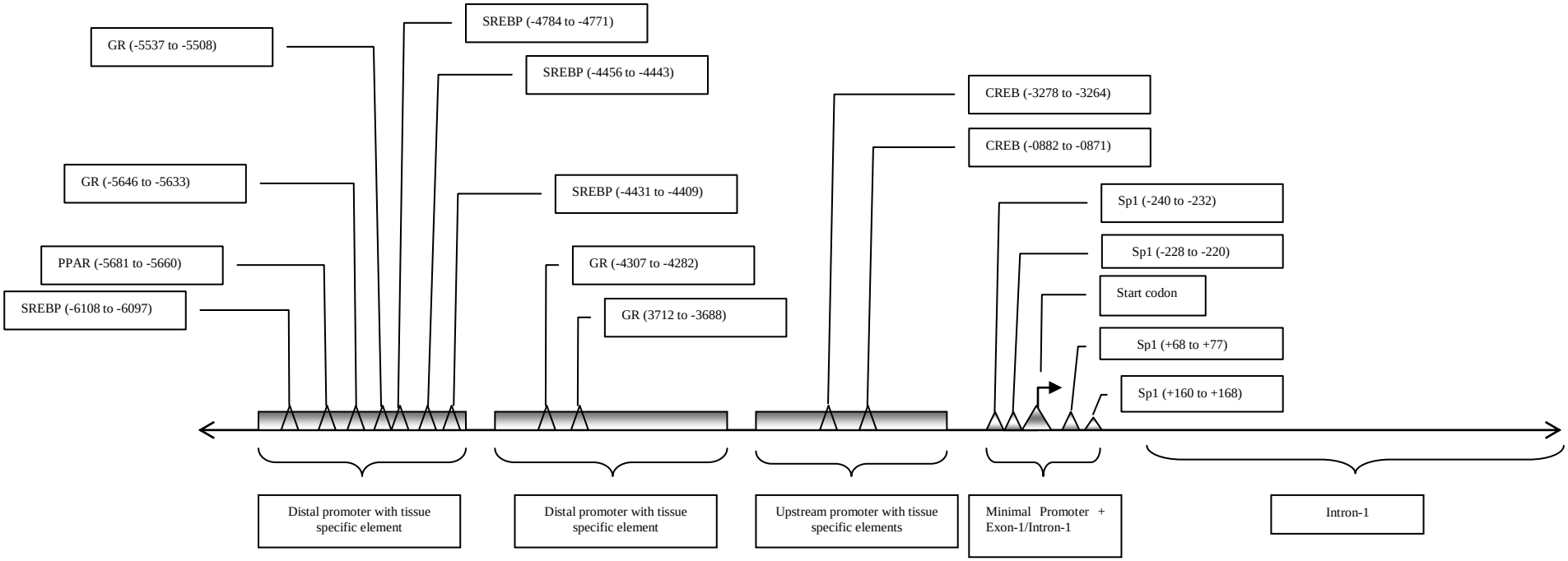
Zhang P, O'Loughlin L, Brindley D N, and Reue1 K (2008). Regulation of lipin-1 gene expression by glucocorticoids during adipogenesis. Journal of Lipid Research, 49, 1519–1528.

Zhang ZH, Wu SD, Gao H, Shi G, Jin JZ, Kong J, Tian Z, Su Y (2006). "Expression of pituitary adenylate cyclase-activating polypeptide 1 and 2 receptor mRNA in gallbladder tissue of patients with gallstone or gallbladder polyps". World Journal of Gastroenterol. 12 (9), 1468–1471

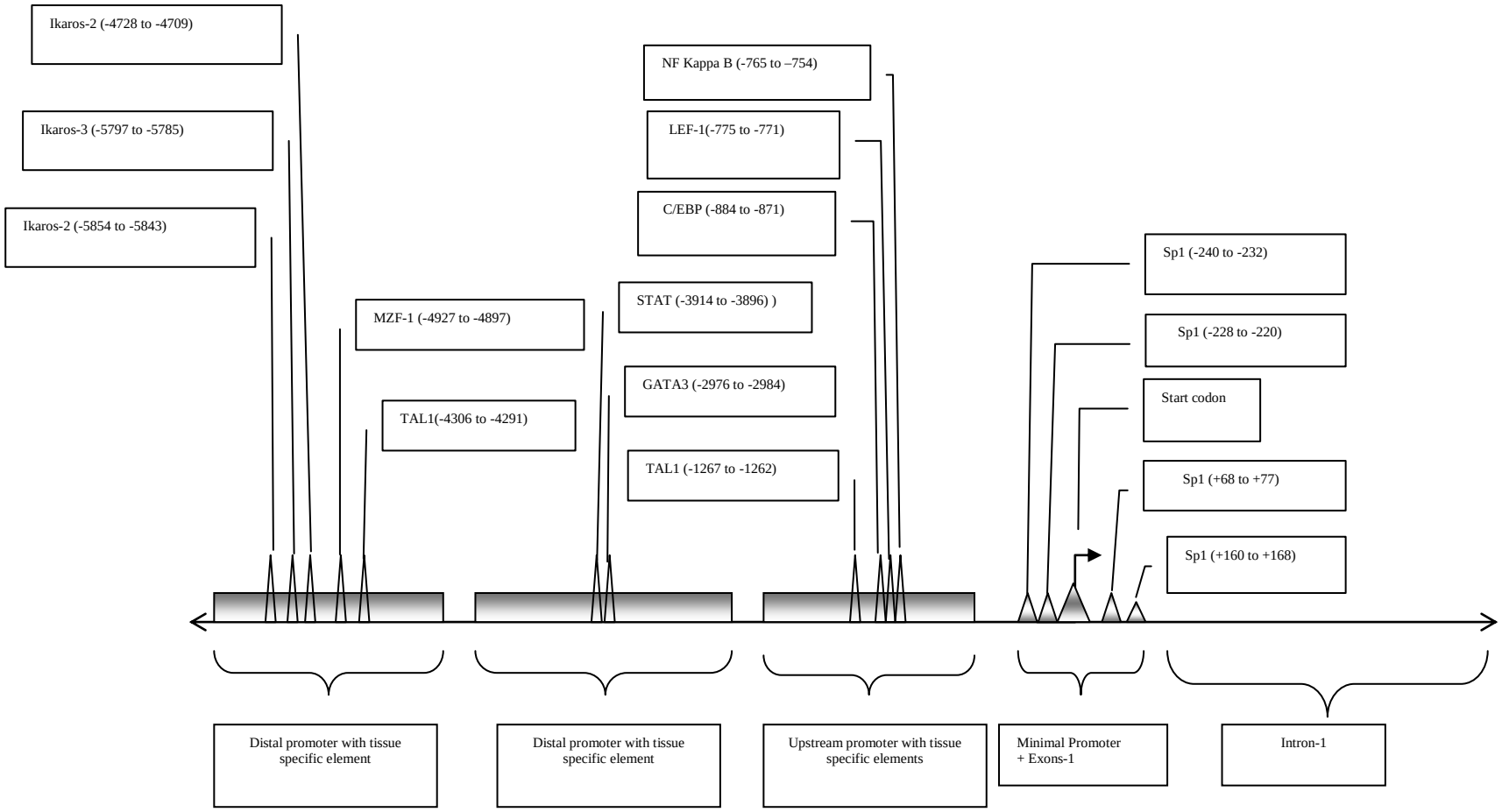
Zolnierowicz S, Cron P, Solinas-Toldo S, Fries R, Linn H Y, and Hemming BA (1994). Isolation, Characterization, and Chromosomal Localization of the Porcine Calcitonin Receptor Gene. Identification of two variants of the receptor generated by alternative splicing.The Journal of Biological Chemistry, 269, 19530-19538.

Appendix

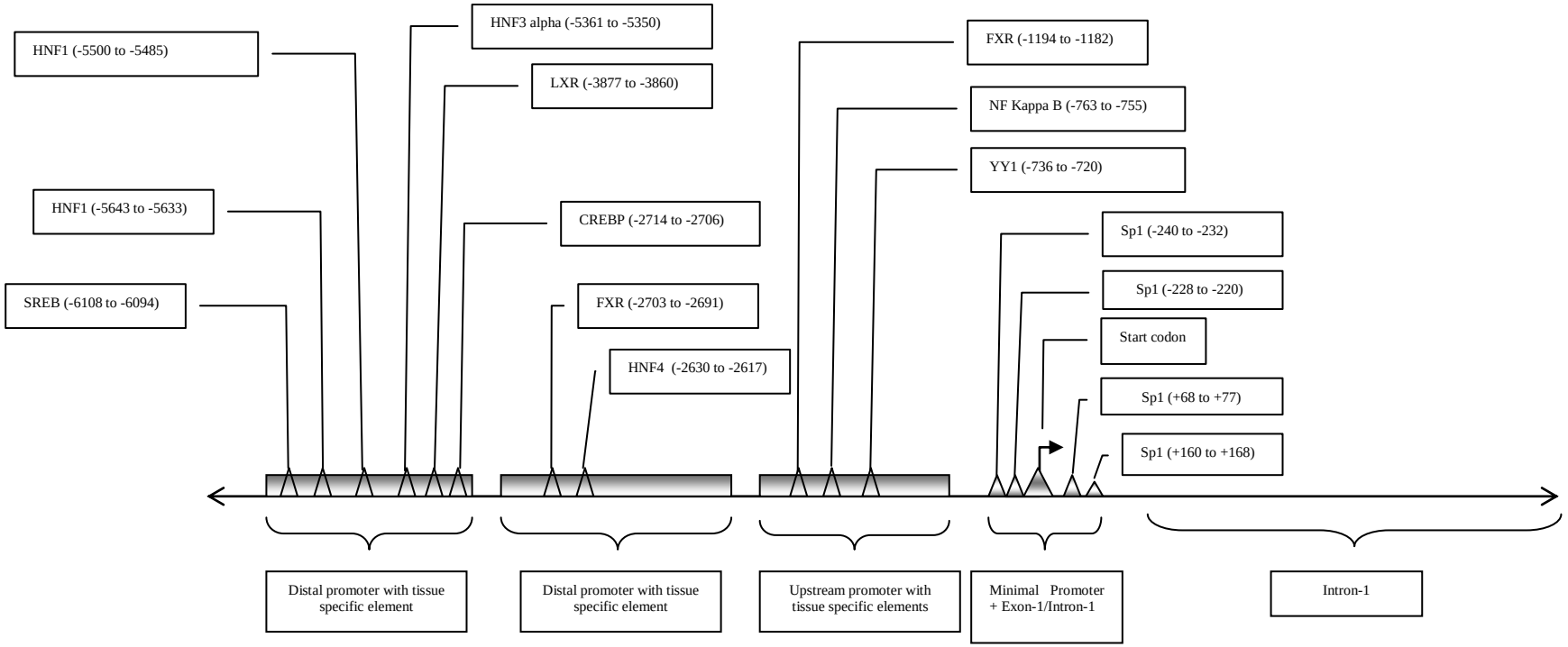
This section contains schematic representations of data derived from multiple sequence alignment of the VIPR2 gene of evolutionarily (closely and distantly) related species. The alignment includes the ~6kb upstream region and ~2.5kb downstream region in relation to translation start site of the VIPR2 gene which contains evolutionarily resilient elements required for the basal and tissue specific regulation of the VPAC2 receptor expression.



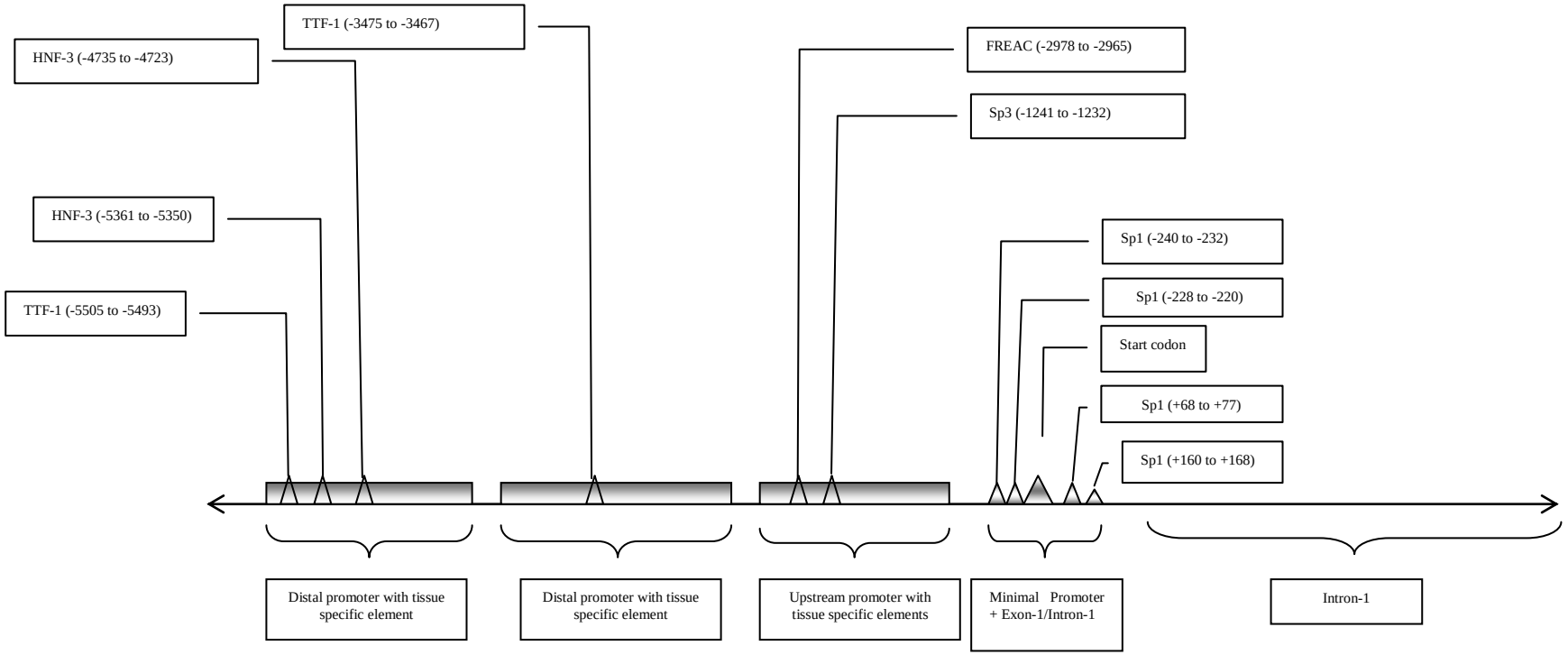
Evolutionary conserved Minimal promoter and Tissue specific elements required for the transcription regulation of human VIPR2 gene in the Adipocytes.



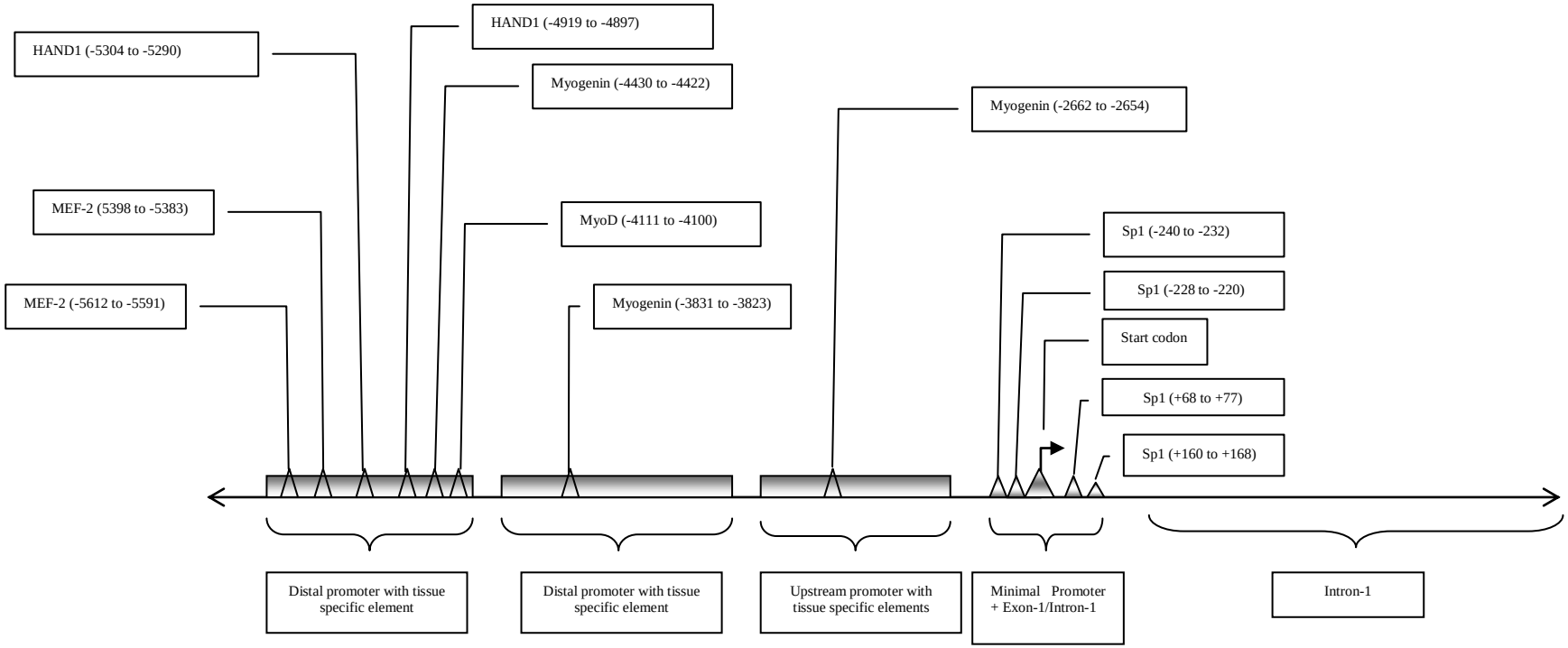
Evolutionary conserved Minimal promoter and Tissue specific elements required for the transcription regulation of human VIPR2 gene in the activated T cells.



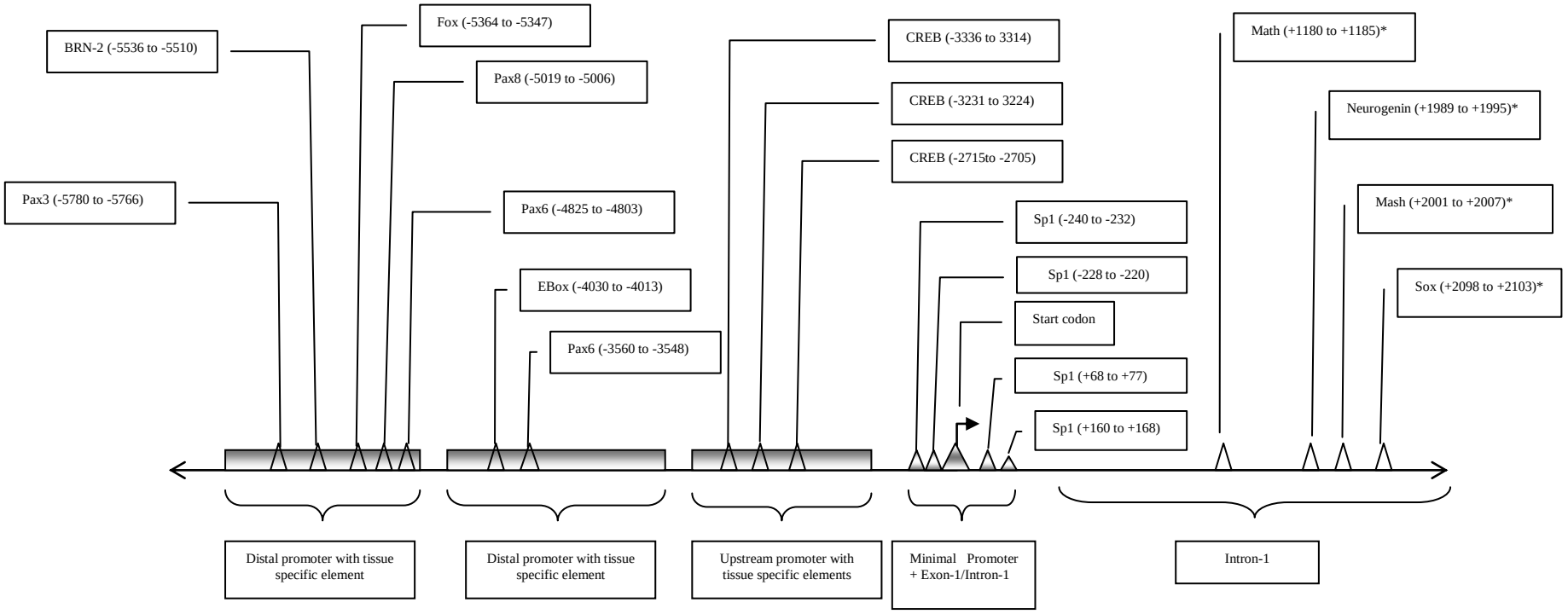
Evolutionary conserved Minimal promoter and Tissue specific elements required for the transcription regulation of human VIPR2 gene in the Hepatocytes



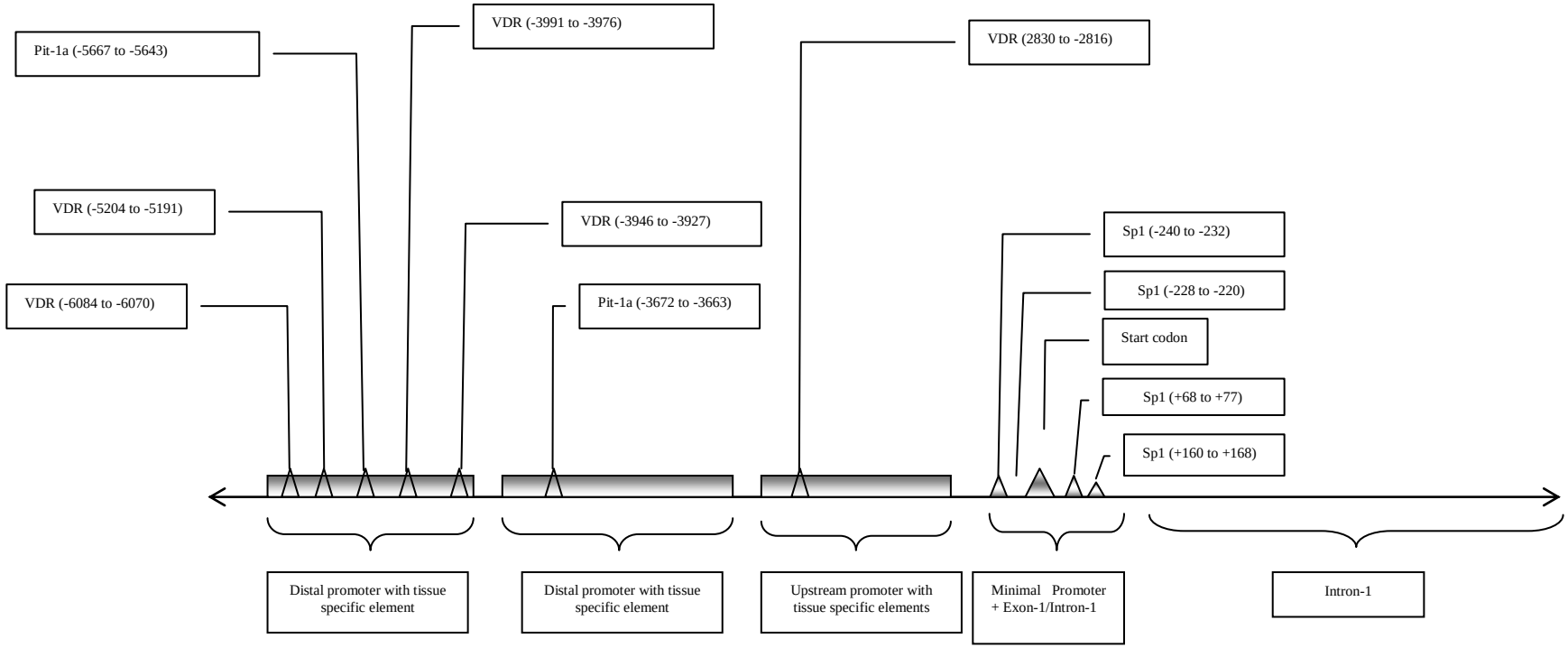
Evolutionary conserved Minimal promoter and Tissue specific elements required for the transcription regulation of human VIPR2 gene in the lung epithelial cells.



Evolutionary conserved Minimal promoter and Tissue specific elements required for the transcription regulation of human VIPR2 gene in the myocytes.



Evolutionary conserved Minimal promoter and Tissue specific elements required for the transcription regulation of human VIPR2 gene in the neurons. *Elements found to be conserved among Rhesus monkey, cow, dog and cat VIPR2 genes.



Evolutionary conserved Minimal promoter and Tissue specific elements required for the transcription regulation of human VIPR2 gene in the Pituitary cells