

FUNCTIONAL ANNOTATION OF ALTERNATIVELY SPLICED ISOFORMS OF
HUMAN G PROTEIN-COUPLED RECEPTORS (GPCRs)

by

DOROTHY ARABA HAMMOND

(Under the Direction of YING XU)

ABSTRACT

The functional differences that exist between alternatively spliced isoforms of genes can be substantial; it could be variations of the function of the “primary” full-length protein, its antagonist function or not directly related to its function. Regardless of the functions of spliced isoforms there is a need for proper annotation. G protein-coupled receptors (GPCRs) are important trans-membrane proteins with important biological and pharmaceutical implications. Incorrect alternative splicing has been associated with a number of human diseases including diabetes and cancers. In this dissertation is presented a framework and results for annotating transcripts of GPCR genes that have originated from alternative splicing of their pre-mRNA.

This was achieved by first assessing the genomic landscape of the GPCRs to determine the type of alternative splicing event that exist and the localization of such events. In addition, the influences of exon-intron structures of GPCRs on the alternative splicing of the genes are evaluated. Then, the molecular level functions are ascertained via motif identification. Furthermore, the cellular state in which

each isoform could function was determined using tissues of expression from RNA-sequencing data. Lastly, isoform function is inferred from *cis* regulatory elements predicted to influence the transcription of the isoform in the specific cellular state. The results are presented here and made publicly available in a PHP/MySQL relational database.

Results from the research analyses indicate that the genomic structures of GPCRs are complex; their exon-intron structures are not arbitrarily arranged; rather they are ordered in such a manner as to ease the alternative splicing of the gene for functional purposes. Also, that GPCR isoforms expression patterns differ across tissues. Additionally it was evident that GPCR isoforms have multiple cellular pathways in which they may be involved depending on the cellular conditions. Overall, 1044 human GPCR isoforms were annotated. These results are now publicly available via the Internet in the database called the G protein-coupled receptor transcriptional annotation of genes (*gpcrTAG*) database. The results from these analyses offer the promise of transcriptional drug control, as well as offer the framework to study functional diversity between isoforms of other genes families.

INDEX WORDS: GPCR, G protein-coupled receptor, RNA-seq, membrane protein, trans-membrane proteins, gene structure, alternative splicing, functional prediction, gene regulation, differential gene expression, next-generation sequencing, transcriptomic data analysis

FUNCTIONAL ANNOTATION OF ALTERNATIVELY SPLICED ISOFORMS OF
HUMAN G PROTEIN-COUPLED RECEPTORS (GPCRs)

by

DOROTHY ARABA HAMMOND

B.A., Kwame Nkrumah University of Science and Technology, Kumasi, Ghana, 2002

M.P.H., Texas A&M University, College Station, Texas, 2006

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial
Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2014

© 2014

DOROTHY ARABA HAMMOND

All Rights Reserved

FUNCTIONAL ANNOTATION OF ALTERNATIVELY SPLICED ISOFORMS OF
HUMAN G PROTEIN-COUPLED RECEPTORS (GPCRs)

by

DOROTHY ARABA HAMMOND

Major Professor: YING XU

Committee: LIMING CAI
NATARAJAN KANNAN

Electronic Version Approved:

Julie Coffield
Interim Dean of the Graduate School
The University of Georgia
December 2014

DEDICATION

To my husband, for all his support, encouragement, love and patience

To my boys (Jack & Joel), just because I love you

To my father, for his encouragements, support, and for being my personal
manuscript reviewer

To my mother, for all the love and support

To my Aunt Emelia, for your desire to see me succeed

To my grandparents, you said it. See what we did

Most importantly, to my HEAVENLY FATHER to whom I owe it all

ACKNOWLEDGEMENTS

A special heartfelt appreciation goes to Prof. Ying Xu for his direction and supervision throughout my study here at the University of Georgia. Your work ethics is encouraging. I feel privileged to have benefited from your guidance and mentorship. Thank you for the knowledge and experience.

Additionally, my sincere appreciation goes to my advisory committee members Liming Cai, Ph.D., David Puett, Ph.D., and Natarajan Kannan, Ph.D., for your support through it all. Another sincere appreciation also goes to Jonathan Arnold, Ph.D. for his encouragement. Lastly, to the Institute of Bioinformatics and to Ms. Joan Yantko, you are a treasure.

Another heartfelt gratitude goes to Victor Olman, Ph.D., former member of the University of Georgia community for all the statistical support.

Financial support came from:

- The National Institute of Health – Supplementary Grant (R01 Grant # GM07533101)
- The National Institute of General Medical Sciences, National Institute of Health, Ruth L. Krirschstein Predoctoral Fellowship F31 (Grant # GM089093)

*Trust in the Lord with all your heart and lean not in your own understanding. In all your ways acknowledge Him, and He will make your path straight
Proverbs 3:5&6*

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	v
LIST OF TABLES	ix
LIST OF FIGURES	x
CHAPTER	
1 INTRODUCTION	1
1.1 RESEARCH OBJECTIVE	1
1.2 G PROTEIN-COUPLED RECEPTORS (GPCR)	1
1.3 ALTERNATIVE SPLICING	10
1.4 STATE OF THE ART IN ANNOTATING ALTERNATIVELY SPLICED ISOFORMS AND CHALLENGES FACED	14
1.5 OVERVIEW OF CHAPTERS: SUMMARY OF THE APPROACH PRESENTED IN THIS DISSERTATION	17
REFERENCES	20
2 FUNCTIONAL UNDERSTANDING OF THE DIVERSE EXON-INTRON STRUCTURES OF THE HUMAN GPCR.....	27
ABSTRACT.....	28
2.1 INTRODUCTION	28
2.2 DATA COLLECTION AND ANALYSIS	30
2.3 RESULTS	34

2.4 DISCUSSIONS.....	38
2.5 CONCLUDING REMARKS	39
REFERENCES	41
3 FUNCTIONAL ANNOTATION OF ALTERNATIVELY SPLICED ISOFORMS OF HUMAN G PROTEIN-COUPLED RECEPTORS	43
3.1 INTRODUCTION	44
3.2 GPCR DATA RETRIEVAL AND PROCESSING.....	45
3.3 ANNOTATION OF MOLECULAR FUNCTION	46
3.4 ANNOTATION OF CELLULAR FUNCTION	55
3.5 INTEGRATION OF MOLECULAR- AND CELLULAR-LEVEL FUNCTION	74
3.6 DISCUSSION.....	74
3.7 CONCLUSION.....	79
REFERENCES	80
4 THE <i>gpcrTAG</i> DATABASE.....	83
4.1 INTRODUCTION	83
4.2 DATABASE CONTENT.....	84
4.3 DATABASE DESIGN.....	85
4.4 DATABASE IMPLEMENTATION	87
4.5 ACCESSING DATABASE	87
4.6 DISCUSSION	90
4.7 FUTURE IMPROVEMENTS	91
REFERENCES	92

5 PERSPECTIVES AND CONCLUSIONS.....	93
5.1 SUMMARY	93
5.2 FUTURE DIRECTIONS	95
5.5 CONCLUSIONS.....	96
APPENDICES	
A Appendix Tables	97
B Links to Supplementary Tables	126

LIST OF TABLES

	Page
Table 1.1: Five major families of GPCRs.....	2
Table 1.2: Physiological roles of GPCRs	6
Table 3.1: Motif information from database search results	53
Table 3.2: New identified motifs in GPCR families and subfamilies.....	54
Table 4.1: Database Statistics	85

LIST OF FIGURES

	Page
Figure 1.1: GPCR structure represented by Rhodopsin.....	7
Figure 1.2: Schematic representation of the types of alternative splicing	12
Figure 1.3: Schematic representation of homology and fold recognition methods	16
Figure 1.4: Flowchart illustrating the framework presented.....	19
Figure 2.1: Splicing graph view of transcripts from alternatively spliced isoform of human gene DRD2.....	36
Figure 2.2: Graphical representation of spliced isoforms of GPRC5C	37
Figure 3.1: Distribution of motifs identified through database search	52
Figure 3.2: Classic schematic representation of promoter identification	58
Figure 3.3: Selection of isoform specific transcription factor	60
Figure 3.4: Heatmap.....	64
Figure 3.5: Frequency distribution of isoform expression.....	65
Figure 3.6: Frequency distribution of transcription factor expression.....	65
Figure 3.7: Frequency distribution of cellular processes	67
Figure 3.8: Frequency distribution of transcription factors	70-72
Figure 3.9: Transcription factor combinations and the isoforms they regulate	73
Figure 4.1: Relationship of tables	86
Figure 4.2: Screenshot of search page	88
Figure 4.3: Screenshot of function search result page	89

CHAPTER 1

INTRODUCTION

1.1 Research Objective

The objective of this research is to annotate the functions of alternatively spliced isoforms of human G protein-coupled receptors.

1.2 G Protein-Coupled Receptors

G protein-coupled receptors (GPCRs) are cell-surface receptors that signal through the heterotrimeric guanine nucleotide-binding regulator proteins (G $\alpha\beta\gamma$). They represent one of the largest families of trans-membrane proteins encoded in human and other complex organisms. GPCRs mediate several physiological and pathophysiological effects: including cardiac and vascular activity, neurotransmission, sensory responses, metabolism, pain control, and endocrine signaling [1, 2], thus, making them of great interest in human therapeutic applications [3, 4]. Due to this multifaceted character, GPCRs selectively sense and bind a variety of extracellular stimuli, such as odorants, lipids, cytokines, neurotransmitters, ions (Ca^{2+}), photons, amines, hormones, peptides etc. [5, 6], with agonistic properties initiating the signal transduction cascade by triggering conformational changes in the receptor that promote the relevant G protein activation, and subsequent corresponding cellular response process [7].

1.2.1 Functional Families of GPCR

Analysis of the human genome revealed at least 800 unique GPCRs [8]. Based on certain conserved features and common functional characteristics, GPCRs are classified into multilevel hierarchical families using various approaches; sequence-based (e.g. GPCRDB) [9], hidden markov model (e.g. HMMTOP) [10, 11], and support vector machine (SVM) (e.g. PRED-GPCR) [12]. GPCRs are classified into 5 classes/families: class A, B, C, D, and E [8] (Table 1.1). A large fraction of them have unknown physiologic functions [13]. Each of these families, however, has unique family traits that make them stand out from the larger superfamily as a whole.

Table 1.1: Five Major Families of GPCRs

Families	Family Name	Number of members
<i>Class A (family A)</i>	Rhodopsin-like	~701
<i>Class B (family B)</i>	Adhesion	~33
<i>Class C (family C)</i>	Secretin-like	~15
<i>Class D (family D)</i>	Metabotropic-Glutamate-like	~15
<i>Class E (family E)</i>	Fizzled/taste	~24

*Detailed listing of families with sub-families are given in appendix

Class A: Rhodopsin family

Considered to be the largest and most studied family of all GPCRs, rhodopsin-like is made up of a wide range of functional proteins, which though similar, show considerable diversity at the sequence level. Based on their sequence similarity they have been separated into distinct subfamilies (see appendix). They are made up of over 700 receptors that include hormone, neurotransmitters, and light receptors. Although their activating ligands display wide structural variation, the receptors themselves comprise of relatively similar sequences. This family is the

first GPCR for which high-resolution crystal structure was obtained. The current understanding is that rhodopsin-like GPCRs represent probably the most ancient GPCRs and the other GPCR genes are the evolutionary descendants of this family through gene and chromosomal duplications as well as other evolutionary events [14].

Rhodopsin has been intensely studied in part due to its stability and abundance. With the amount of effort put into researching the rhodopsin-like family, there still remain many unresolved issues including: (i) the trans-membrane helix movement responsible for the signal transduction process (which is in part addressed in recent publication [15]), (ii) the stoichiometry of coupling to G proteins and their mode of activation, and (iii) the role of receptor oligomerization. Now that a high-resolution activated structure is available, most of these issues can start to be addressed.

Class B: Adhesion family

Adhesion family of GPCRs is considered to be the second largest family of GPCRs with approximately 33 members identified in the human genome. They are defined by the existence of a large extracellular region linked to a trans-membrane seven (TM7) moiety via a GPCR proteolytic site (GPS)-containing stalk region. Phylogenetic analysis of the trans-membrane domains reveals that the adhesion class of GPCRs forms a distinct family [16]. This family is characterized by rather long N-termini with multiple functional domains commonly seen in proteins such as tyrosine kinases, epidermal growth factor- (EGF) like, integrins and cadherins. They have a complex genomic structure with multiple introns and splice variants, and a seven trans-membrane region that is clearly unique from other GPCRs [17]. Several

members of this family are known to play a role in the immune system, and it has become evident that many also have important roles in the central nervous system (CNS).

The large extracellular domain responsible for ligand binding may pose a number of difficulties, the simplest being binding specificity. Theoretically speaking, with such a large domain there might be more than one potential binding site. Modeling of such a complex molecule is not an easy task. It is imperative during such modeling to identify all like motifs in the domain and do site-specific mutations and model the responses separately as well as together. It is necessary to also model multiple sites together to assess the possibility of a synergistic effect in response. It is not uncommon for the effect of two reactions on a molecule to be greater than the effect of each reaction individually or the sum of the individual effect. The presence of one reactant may enhance the effect of a second.

Class C: Secretin family

Secretin-like family is extremely important in hormonal homeostasis and the treatment of many metabolic conditions (diabetes and obesity) as well as nervous system disorders (migraine and anxiety). They include receptors for 15 peptide hormones that are grouped into multiple subfamilies based on their physiological role (Vasodilation, stress response, calcium homeostasis, growth hormone release etc.) [18]. The main difference between the secretin-like family and the rest of the superfamily is the very large amino termini. There has been a recent addition of 2 crystal structures to this family, namely the human glucagon receptor (GCR) [19], PDB ID: 4L6R, and the human corticotropin-releasing factor receptor 1 (CRF1) [20] PDB ID: 4K5Y.

Class D: Metabotropic-Glutamate family

The metabotropic-Glutamate-like family of GPCRs is recognized for their role in modulating synaptic transmission and neuronal excitement signaling. Although they are structurally similar to other proteins in the superfamily, they have significantly low sequence similarity with other GPCRs.

Class E: Fizzled family

The fizzled family of GPCRs is one that has had very little attention since the beginning. Receptors in this family contain a cysteine-rich extracellular domain that has been implicated in the Wnt signaling pathway [21]. Fizzled proteins are believed to also play important roles in embryonic development and formation of neural synapses.

Physiological Functions of GPCRs

Other physiological functions associated with GPCRs are described in Table 1.2 below. A large fraction of GPCRs however, have unknown physiological functions. The current estimate is that nearly a third of them have unknown or unclear physiologic function [22]. With the amount of time and effort dedicated to GPCR research both in academia and industry, the number of known physiological functions can only be projected to increase.

Table 1.2: Physiological roles of GPCRs

Physiological function	Description
Visual sense	Photoisomerization reaction that translate electromagnetic radiation into cellular signals
Sense of taste	Taste mediated release of gustducin in response to sweet or bitter tasting properties
Sense of smell	Receptors of olfactory epithelium bind odorants and pheromones
Behavioral and mood regulation	Receptors bind neurotransmitters such as serotonin, dopamine, GABA, and glutamate
Immune system activity and inflammation	Chemokine receptors binds ligands that mediate intercellular communication between immune system cells, histamine receptors bind inflammatory mediators
Autonomic nervous system	Regulation of blood pressure, heart rate, and digestive systems
Homeostasis modulation	Water balance

1.2.2 Structural Features of GPCR

GPCRs are located in the lipid bilayer. They share a common structural signature of seven hydrophobic trans-membrane (TM) α -helices that thread back and forth through the lipid bilayer, which allows for the GPCR to have an extracellular N-terminus (for glycosylation and ligand binding), an intracellular C-terminus (for phosphorylation), three extracellular loops (also for binding), three intracellular loops, and an “eighth” cytoplasmic structure (which aids in signal transduction) (Figure 1.1).

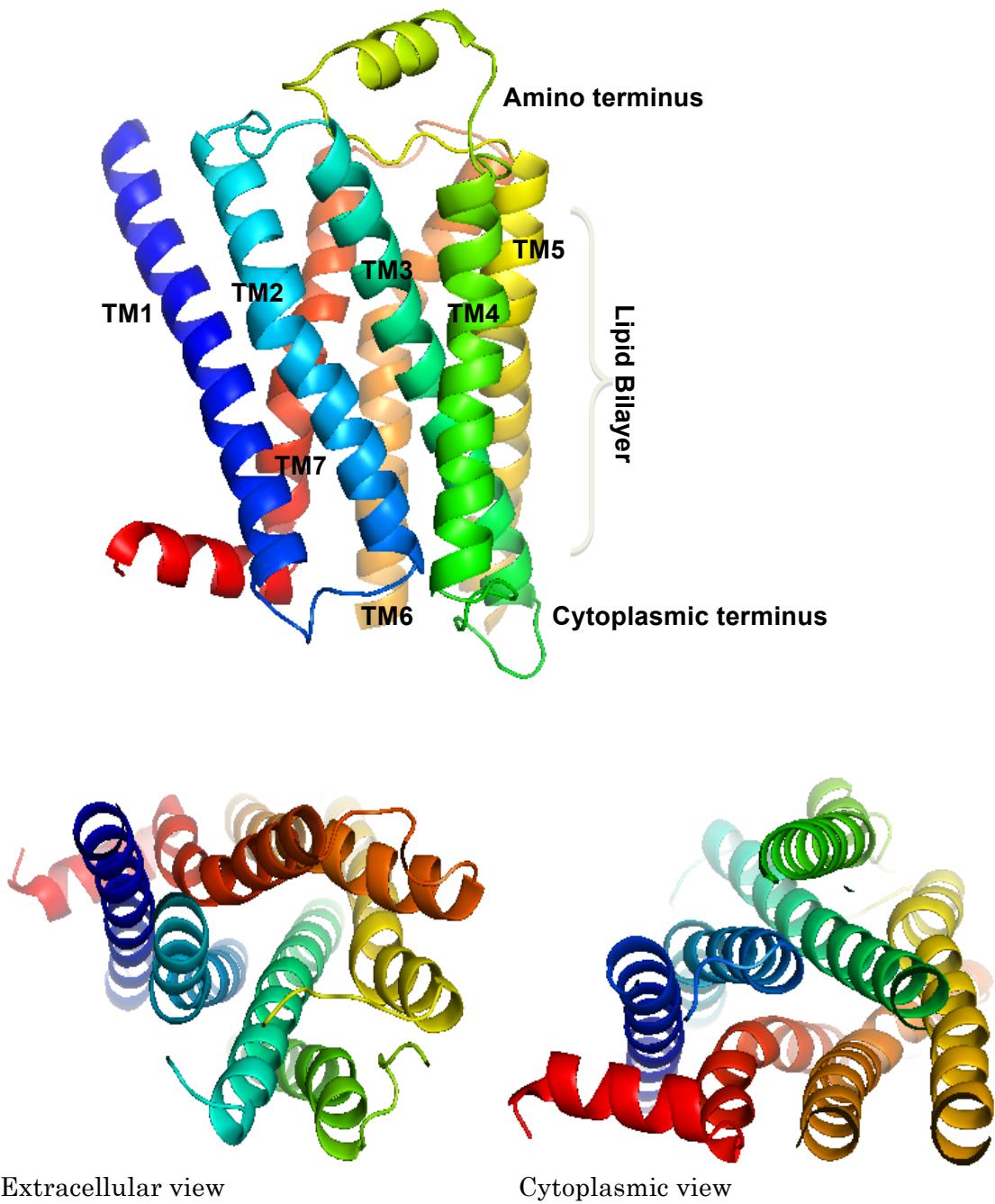


Figure 1.1: a) Rhodopsin structure from the PDB 2RH1 showing the seven transmembrane helices, the 3 intracellular and 3 extracellular loops that connect them. Also shown are the extracellular and cytoplasmic views of the same structure.

GPCRs have sequence homology among the TM segments. The most structural diversity is observed in the carboxyl termini, the amino termini, and the intracellular loop spanning TM5 and TM6. The amino termini have the most sequence diversity across all GPCRs. In rhodopsin, the sequence is relatively short (approximately 10 to 50 amino acids), hormone receptors and the glutamate family receptors have much longer sequences (approximately 350 to 600 amino acids). The adhesion family receptors have the largest amino terminal domains. The wide range of sequence diversity among these GPCR families makes them a dreadful task to model them computationally.

The first insights into the structure of GPCRs came from cryoelectron microscopy of two-dimensional crystals of bovine rhodopsin from Gebhard Schertler's group [23-25]. Though the resolutions of these structures were limiting (ranging from 5Å to 9Å), it uncovered the general architecture of the seven transmembrane helices of GPCRs. Due to the conformational complexity of ligand-activated GPCRs it is of no surprise that it took the scientific community this long to obtain 3D activated structure. Currently, there are about twelve unique high resolution inactive-state structures of GPCRs including: the human β_2 AR (PDBID 2RH1) bound to the high-affinity inverse agonists carazolol [26-28] and timolol [29]; avian β_1 AR (PDBID 2VT4) bound to the antagonist cyanopindolol [30]; the human A_{2A} adenosine receptor (PDBID 3EML) bound to the antagonist ZM241385 [31]; and bovine rhodopsin[32-34] (PDBID 1F88) containing the covalently bound inverse agonist 11-*cis* retinal. These four proteins have a similar overall architecture, yet the divergences are still high enough to signify important differences in helical packing interactions [35].

1.2.3 GPCR Activation and Response Interaction with G proteins

Following the ligand binding and subsequent conformational change, the $G\alpha$ and $G\beta\gamma$ dissociates from the $G\alpha\beta\gamma$ complex allowing them to freely interact and modulate the activity of downstream elements of the signaling cascades; adenylyl cyclase, phospholipases, mitogen-activated protein kinases (MAPKs), or calcium and potassium ion channels. The signal transduction process is strictly regulated by receptor posttranslational modifications. Chief among them is the receptor phosphorylated by GPCR-specific and -nonspecific kinases that modulates subsequent interactions with several intracellular proteins involved in receptor internalization and down regulation [36, 37] or promoting growth factor receptor transactivation [38]. Additional regulatory mechanisms ensue from the interplay of G protein subunits with regulators of G protein signaling (RGS) [39].

Several experimental works have shown that the binding of external stimuli in the form of ligand or other signal mediators creates a conformational change in the receptor, which results in the activation of a G protein. The precise manner of signal transduction through the lipid bilayer however, is not entirely understood.

Typically, the role of coupling the binding of agonist to the activation of specific heterotrimeric G protein leads to the modulation of downstream effector proteins. An example of such signaling response is a liver cells response to the hormone epinephrine [40]. In times of stress, epinephrine is released into the blood stream from the adrenal gland. Upon reaching the liver, epinephrine binds to β -adrenergic receptors (a GPCR) on the surface of the liver cells. Binding of epinephrine causes a conformational change in the receptor, allowing it to associate with a trimeric G_{sa} protein. Binding to the receptor creates a change in the G_{sa}

subunit, resulting in GDP displacement and GTP binding. The G_{α} – GTP complex dissociates from the $G_{\beta\gamma}$ subunit, then binds and activates adenylate cyclase. Activated adenylate cyclase catalyzes the formation of cyclic AMP (cAMP) from ATP. The effects of cAMP are mediated through cAMP dependant protein kinase (cAPK). cAMP bind to regulatory subunits of cAPK which release and activates the catalytic subunits. The catalytic subunit enters the nucleus and phosphorylates one of its multiple targets, such as the CREB transcription factor. Phosphorylated CREB bind cAMP response elements (CRE sequences), with the co-activator P₃₀₀/CBP stimulate transcription of CRE regulated genes. In the liver, CRE containing genes encode enzymes involved in glucose production. The end result is increased levels of glucose necessary to energize tissues during the stress response.

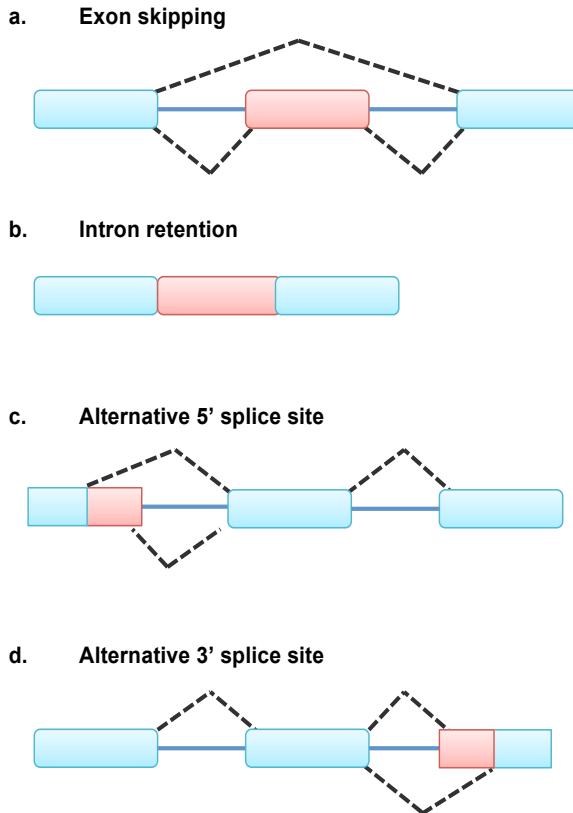
1.3 Alternative Splicing

Alternative splicing is a cellular process by which multiple mRNAs can be produced from the same precursor mRNA (pre-mRNA). The inclusion of various exons, retention of some introns, the use of alternative 5', alternative 3', as well as in some instances the use of mutually exclusive exons, alternative promoters and alternative polyadenylation in the mRNA results in the generation of different isoforms from a single gene, which may differ in structure, function, localization, and/or other properties. This process has been named as the main influential factor in the disparity that exist between the number of proteins (estimated at about 100,000) and the number of identified genes (approximately 24,000) in the human genome [41]. Also, it is estimated that ~92% of human genes [42, 43] undergo alternative splicing. Equally, proteomic analysis indicates that the majority of the

functional protein species in human cells are in spliced forms [44]. Disrupted alternative splicing has also been implicated in a number of human diseases [45, 46], thus, further emphasizing the importance of alternative splicing in the human genome.

1.3.1 Types of Alternative Splicing Events

Alternative splicing events include several different types of events that can generally be classified into four main groups, namely: exon skipping – the alternative splicing event in which the whole exon along with its flanking introns is spliced out of a transcript (Figure 1.2 a); intron retention – the splicing event in which an intron remains in the mature mRNA transcript (Figure 1.2 b); alternative 5' splice site (Figure 1.2 c) and alternative 3' splice site (Figure 1.2 d) – the splicing event in which two or more splice sites are recognized at one end of an exon [47]. Other less frequently observed splicing events in eukaryotes are the mutually exclusive events, alternative promoter usage, and alternative polyadenylation.

**Figure 1.2:**

Schematic representation of the different types of alternative splicing events identified in GPCRs. Blue rectangles represent consecutive exons, pink rectangles show alternative splicing regions, solid lines represent introns, while dash lines represent splicing options.

1.3.2 Functional Consequences of Alternative Splicing

The functional differences that exist among alternatively spliced isoforms of genes can be substantial; it could be variations of the function of the primary full-length protein, its antagonist function or even not directly related to its function. A classic example is the alternative splicing of the calcitonin/CGRP gene to produce the calcitonin, which is produced by thyroid cells in response to high blood calcium levels and the calcitonin gene-related peptide (CGRP) in neural cells. The calcitonin/CGRP gene contains 6 exons; in thyroid cells the mRNA does not contain the 5th and 6th exons, while in the neuron cells, they include the exons 5 and 6. However, exon 4 is spliced out in the CGRP transcript version of the gene. In both transcripts, exon 1 is not translated, exon 2 and 3 together encode a signal peptide,

which serves as tags to ensure the correct transportation of the protein to the Golgi apparatus after translation, which is removed from the protein after it is formed, exon 4 therefore is the calcitonin coding exon, and 5 and 6 together encodes the CGRP [48]. Another example of diverse function of alternatively spliced isoforms is the multiple functions of the G protein-coupled receptor kinases (GRK) [49]. GRK compose of seven isoforms (GRK1-GRK7) [50]. Each variant shows differential expression patterns, structure, and function. GRK1, GRK4, and GRK7 are expressed in limited tissues (retina [51], testis [52], and retina [53], respectively), while GRK2, GRK3, GRK5, and GRK6 are ubiquitously expressed throughout the body [54, 55]. Also, the amino-terminal domain of GRK2 interacts with the G protein $\beta\gamma$ subunit, whereas that of GRK4, GRK5, and GRK6 interacts with phosphatidylinositol 4,5-bisphosphate (PIP2) [56-58]. As these examples show, alternatively spliced isoforms of genes contain unique functionality, which may be as subtle as expressing in a different tissue, or as drastic as that of the Calcitonin/CGRP gene. Either way, they each need to be properly annotated. The current state of human GPCR research, however lacks annotation for their isoforms. We therefore seek with this research to bridge this gap.

1.3.3 Alternative Splicing Events in GPCRs

Initially, GPCRs were thought to be intronless in their open reading frames (ORF) and hence did not undergo alternative splicing [59]. This argument was flawed however since not all alternative splicing events is dependent on the presence of introns, such as the alternative 5' and 3' splicing. The current knowledge estimates that ~50% of GPCR genes undergo alternative splicing [41, 60, 61]. Presently, there are rarely annotations publicly available on alternatively spliced

isoforms of GPCRs, where such information exist, they often take on the annotation of the “primary” isoform, which is usually the isoform that was first sequenced.

1.4 State of the Art in Annotating Alternatively Spliced Isoforms and the Challenges Faced

Over the last decade, computational biology has offered invaluable tools that provide insight into the function of proteins based on their sequence contents, structure, evolutionary history, and association with other proteins. These computational methods have numerous limitations that become compounded when considering proteins that have originated from alternative splicing of the pre-mRNA. Discussed below are some of the problems encountered by these conventional tools in functional predictions of alternative spliced proteins.

1.4.1 Function Inferred from Sequence

The most commonly used computational technique in inferring protein function is the sequence homology; examples include BLAST [62] and HMMER [63]. Such inferences, however, must be arrived at with caution since homology only implies similar evolutionary origin and does not guarantee retained functional property. Also, the kind of evolutionary relationship (orthologous versus paralogous) that exists between the two must be evaluated since orthologous relationship tends to conserve function better than paralogous ones [64]. Additionally, no similarity/identity threshold is considered safe enough to ensure functional conservation, even reasonably small differences can drastically change the function of the protein causing a total loss of function, slight change in function, or acquisition of a new function altogether [65-67].

Where sequence homology fails or simply to add validity to sequence homology annotation, there is the use of conserved motifs as functional characterization tool such as PROSITE [68], BLOCKS [69], InterPro [70], and PRINTS [71]. Common conserved motifs however do not always imply same function.

1.4.2 Function Inferred from Structure

The three-dimensional structure information of a protein is key in predicting its function since proteins live and function in their three-dimensional state. Structural information is computationally derived in three main ways: homology modeling, fold recognition, and *ab initio* prediction.

As with sequence homology, homology modeling and fold recognition (Figure 1.4) are riddled with the same challenges. Thus, proteins with the same general structural architecture, and even conserved functional residues can have unrelated functions, while others with profoundly different structures may perform the same function [72-74]. Yet, structural prediction via homology modeling and fold recognition is helpful in revealing potential evolutionary relationship, suggest function, and/or indicate evolutionary convergence.

Ab-initio methods (meaning from “first principles of physics” [75, 76]) rely solely on the laws of physics without prior knowledge or assumptions of how the protein’s final structure should be. Examples of *ab-initio* structural techniques are molecular dynamic simulation and Monte Carlo sampling. In recent years, a combinatory approach of combining the best parts of two or more methods has also proven successful.

To a large extent, all of the above-discussed computational tools have had great success in predicting the molecular functions of cytosolic proteins; yet, minimal success is seen in membrane proteins. Reasons for this include lack of matching templates, difficulties in crystallizing membrane proteins, time constraint, and difficulty in simulating the lipid environment in the case of *ab initio*.

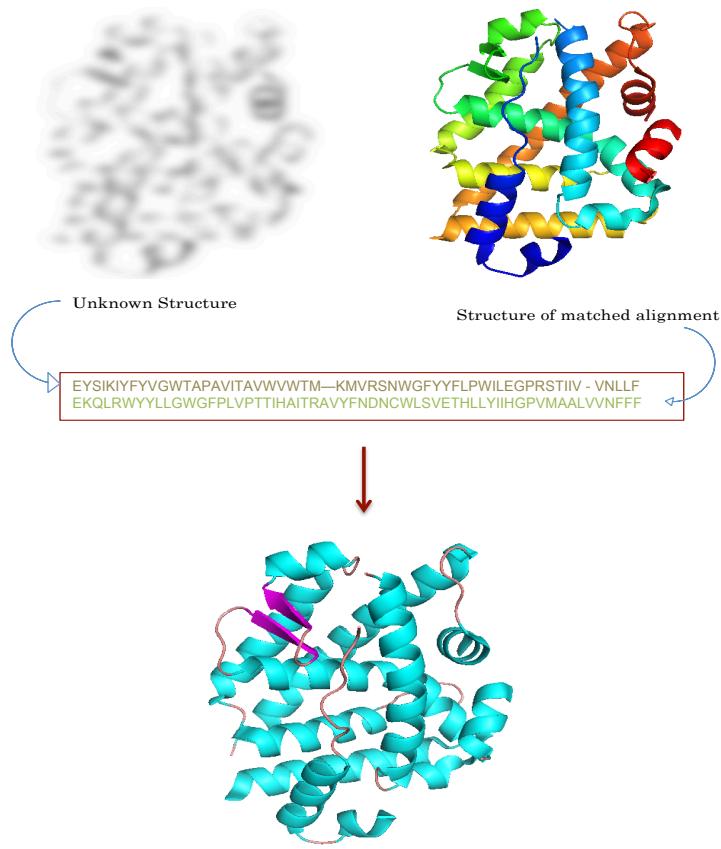


Figure 1.3: Schematic representation of homology modeling and fold recognition methods of protein structure prediction

1.4.3 Challenges faced with Alternative Spliced Isoform Annotations Using These Conventional Computational Tools

The first problem with using conventional computational methods in

predicting the functions of alternatively spliced isoforms is the lack of an evolutionary relationship between isoforms. For homology-based approaches, the two proteins (query and target) must be evolutionarily related, but spliced isoforms are not. Secondly, functional motif-based prediction methods, which characterize areas of conserved and potentially functional sites of proteins, may not be directly applicable in this instance either since functional preservation of such motifs is highly (local) structure-dependent [77]. Thus, the tertiary structure of a functional motif may change due to the change in the global structure of the splicing isoform from the “primary” full-length protein, hence altering or even possibly losing the function of the motif. A single motif is insufficient to properly capture the function of a protein and therefore additional expression information is needed to accurately annotate them. Also, whether molecular level functional annotation, which is provided from discussed computational methods, is sufficient to functionally distinguish isoforms is another issue to be considered.

1.5 Overview of Chapters: Summary of the Approach Presented in this Dissertation

As discussed above, the conventional approach of annotating proteins is often not applicable when considering proteins that have originated from alternative splicing. In the following chapters of this dissertation is presented an approach that avoids many of the pitfalls of using the conventional annotation approach.

The remaining chapters of this dissertation are divided into four. Chapters 2 and 3 cover the approach/framework used, its implementation and results on 1044 GPCR isoforms, chapter 4 presents a database built to share the results from the analysis in chapters 2 and 3 publicly, and lastly chapter 5 will pull in all the prior

chapters and discuss the pros and cons of our approach, draw conclusions, and give future directions of the research.

The first step in properly annotating alternatively spliced isoforms is to assess the genomic structure of the gene of interest, determine the type of splicing events occurring and its localization on the gene structure. Therefore, in chapter 2 the following tasks are performed:

1. Assess the genomic structure of the GPCRs,
2. Determined the type and location of alternative splicing event that occur in GPCR genes,
3. Determine exon boundaries and their constraints on the alternative splicing of the genes, and lastly
4. Determine the potential impact of spliced exons on the functioning of the resulting proteins by identifying conserved motifs in such exons.

We tested the hypothesis that exon boundaries were randomly determined.

Chapter 3 was aimed at annotating the cellular and molecular functions of isoform. Using a motif-based approach, molecular function information were obtained for all research data in two ways; one, motif finding by database search, and two finding new motifs for our dataset. Cellular functional annotation had to be divided into 3 phases. Phase 1: identification of transcription factors that may bind the promoters of the isoforms, Phase 2: determine the tissues in which isoform and transcription factor may be expressed, and Phase 3: identify the common expression pattern between isoform and transcription factor and from it infer isoform cellular function. Outline is summarized in Figure 1.4 below.

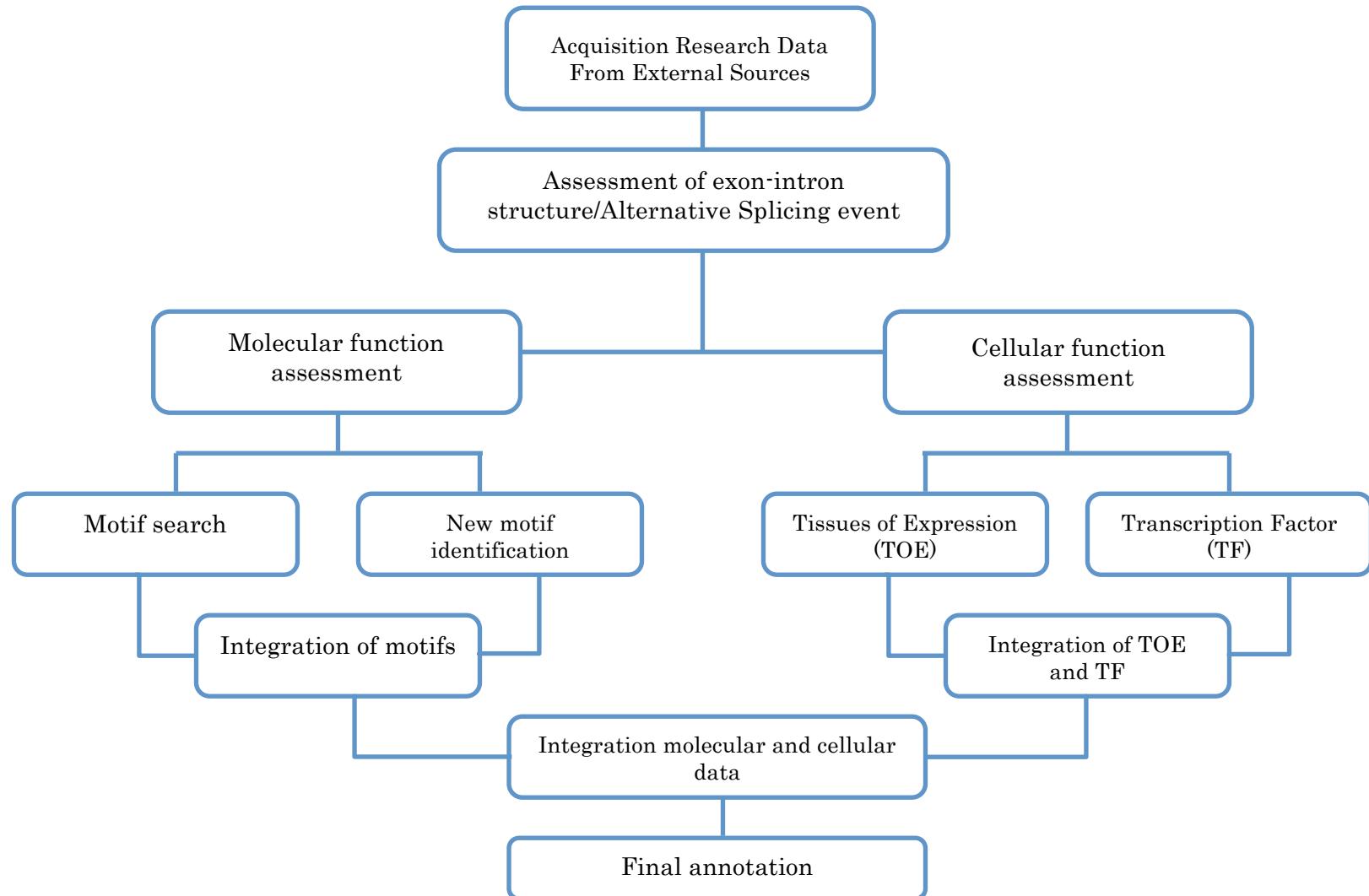


Figure 1.4: flowchart illustrating the framework presented in this dissertation

REFERENCES

1. Jacoby E, Bouhelal R, Gerspacher M, Seuwen K: **The 7 TM G-protein-coupled receptor target family.** *ChemMedChem* 2006, **1**(8):761-782.
2. Lefkowitz RJ: **Seven transmembrane receptors: something old, something new.** *Acta Physiol (Oxf)* 2007, **190**(1):9-19.
3. Wise A, Gearing K, Rees S: **Target validation of G-protein coupled receptors.** *Drug Discov Today* 2002, **7**(4):235-246.
4. Pahikkala T, Airola A, Pietila S, Shakyawar S, Szwajda A, Tang J, Aittokallio T: **Toward more realistic drug-target interaction predictions.** *Brief Bioinform* 2014.
5. Hargrave PA, McDowell JH: **Rhodopsin and phototransduction: a model system for G protein-linked receptors.** *FASEB J* 1992, **6**(6):2323-2331.
6. Ji TH, Grossmann M, Ji I: **G protein-coupled receptors. I. Diversity of receptor-ligand interactions.** *J Biol Chem* 1998, **273**(28):17299-17302.
7. Gether U: **Uncovering molecular mechanisms involved in activation of G protein-coupled receptors.** *Endocr Rev* 2000, **21**(1):90-113.
8. Fredriksson R, Lagerstrom MC, Lundin LG, Schioth HB: **The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paragon groups, and fingerprints.** *Mol Pharmacol* 2003, **63**(6):1256-1272.
9. Vroeling B, Sanders M, Baakman C, Borrmann A, Verhoeven S, Klomp J, Oliveira L, de Vlieg J, Vriend G: **GPCRDB: information system for G protein-coupled receptors.** *Nucleic Acids Res* 2011, **39**:D309-D319.
10. Tusnady GE, Simon I: **The HMMTOP transmembrane topology prediction server.** *Bioinformatics* 2001, **17**(9):849-850.
11. Wistrand M, Kall L, Sonnhammer EL: **A general model of G protein-coupled receptor sequences and its application to detect remote homologs.** *Protein science : a publication of the Protein Society* 2006, **15**(3):509-521.
12. Papasaikas PK, Bagos PG, Litou ZI, Promponas VJ, Hamodrakas SJ: **Pred-Gpcr: Gpcr Recognition and Family Classification Server.** *Nucleic Acids Res* 2004, **32**:W380-W382.

13. Howard AD, McAllister G, Feighner SD, Liu Q, Nargund RP, Van der Ploeg LH, Patchett AA: **Orphan G-protein-coupled receptors and natural ligand discovery.** *Trends Pharmacol Sci* 2001, **22**(3):132-140.
14. Rompler H, Staubert C, Thor D, Schulz A, Hofreiter M, Schoneberg T: **G protein-coupled time travel: evolutionary aspects of GPCR research.** *Molecular interventions* 2007, **7**(1):17-25.
15. Scheerer P, Park JH, Hildebrand PW, Kim YJ, Krauss N, Choe HW, Hofmann KP, Ernst OP: **Crystal structure of opsin in its G-protein-interacting conformation.** *Nature* 2008, **455**(7212):497-502.
16. Bjarnadottir TK, Fredriksson R, Schioth HB: **The adhesion GPCRs: a unique family of G protein-coupled receptors with important roles in both central and peripheral tissues.** *Cell Mol Life Sci* 2007, **64**(16):2104-2119.
17. Pickering C, Hagglund M, Szmydynger-Chodobska J, Marques F, Palha JA, Waller L, Chodobski A, Fredriksson R, Lagerstrom MC, Schioth HB: **The Adhesion GPCR GPR125 is specifically expressed in the choroid plexus and is upregulated following brain injury.** *BMC Neurosci* 2008, **9**:97.
18. Hollenstein K, de Graaf C, Bortolato A, Wang M-W, Marshall FH, Stevens RC: **Insights into the structure of class B GPCRs.** *Trends Pharmacol Sci* 2014, **35**(1).
19. Siu FY, He M, de Graaf C, Han GW, Yang D, Zhang Z, Zhou C, Xu Q, Wacker D, Joseph JS *et al*: **Structure of the human glucagon class B G-protein-coupled receptor.** *Nature* 2013, **499**(7459):444-449.
20. Hollenstein K, Kean J, Bortolato A, Cheng RK, Dore AS, Jazayeri A, Cooke RM, Weir M, Marshall FH: **Structure of class B GPCR corticotropin-releasing factor receptor 1.** *Nature* 2013, **499**(7459):438-443.
21. Huang HC, Klein PS: **The Frizzled family: receptors for multiple signal transduction pathways.** *Genome Biol* 2004, **5**(7):234.
22. Chung S, Funakoshi T, Civelli O: **Orphan GPCR research.** *Brit J Pharmacol* 2008, **153 Suppl 1**:S339-346.
23. Krebs A, Villa C, Edwards PC, Schertler GF: **Characterisation of an improved two-dimensional p22121 crystal from bovine rhodopsin.** *J Mol Biol* 1998, **282**(5):991-1003.

24. Krebs A, Edwards PC, Villa C, Li J, Schertler GF: **The three-dimensional structure of bovine rhodopsin determined by electron cryomicroscopy.** *J Biol Chem* 2003, **278**(50):50217-50225.
25. Schertler GF: **Structure of rhodopsin.** *Eye* 1998, **12** (Pt 3b):504-510.
26. Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, Choi HJ, Kuhn P, Weis WI, Kobilka BK *et al*: **High-resolution crystal structure of an engineered human beta₂-adrenergic G protein-coupled receptor.** *Science* 2007, **318**(5854):1258-1265.
27. Rasmussen SG, Choi HJ, Rosenbaum DM, Kobilka TS, Thian FS, Edwards PC, Burghammer M, Ratnala VR, Sanishvili R, Fischetti RF *et al*: **Crystal structure of the human beta₂ adrenergic G-protein-coupled receptor.** *Nature* 2007, **450**(7168):383-387.
28. Rosenbaum DM, Cherezov V, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, Choi HJ, Yao XJ, Weis WI, Stevens RC *et al*: **GPCR engineering yields high-resolution structural insights into beta₂-adrenergic receptor function.** *Science* 2007, **318**(5854):1266-1273.
29. Hanson MA, Cherezov V, Griffith MT, Roth CB, Jaakola VP, Chien EY, Velasquez J, Kuhn P, Stevens RC: **A specific cholesterol binding site is established by the 2.8 Å structure of the human beta₂-adrenergic receptor.** *Structure* 2008, **16**(6):897-905.
30. Warne T, Serrano-Vega MJ, Baker JG, Moukhametzianov R, Edwards PC, Henderson R, Leslie AG, Tate CG, Schertler GF: **Structure of a beta₁-adrenergic G-protein-coupled receptor.** *Nature* 2008, **454**(7203):486-491.
31. Jaakola VP, Griffith MT, Hanson MA, Cherezov V, Chien EY, Lane JR, Ijzerman AP, Stevens RC: **The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist.** *Science* 2008, **322**(5905):1211-1217.
32. Palczewski K, Kumasaka T, Hori T, Behnke CA, Motohashima H, Fox BA, Le Trong I, Teller DC, Okada T, Stenkamp RE *et al*: **Crystal structure of rhodopsin: A G protein-coupled receptor.** *Science* 2000, **289**(5480):739-745.
33. Okada T, Sugihara M, Bondar AN, Elstner M, Entel P, Buss V: **The retinal conformation and its environment in rhodopsin in light of a new 2.2 Å crystal structure.** *J Mol Biol* 2004, **342**(2):571-583.

34. Li J, Edwards PC, Burghammer M, Villa C, Schertler GF: **Structure of bovine rhodopsin in a trigonal crystal form.** *J Mol Biol* 2004, **343**(5):1409-1438.
35. Rosenbaum DM, Rasmussen SG, Kobilka BK: **The structure and function of G-protein-coupled receptors.** *Nature* 2009, **459**(7245):356-363.
36. Tsao P, Cao T, von Zastrow M: **Role of endocytosis in mediating downregulation of G-protein-coupled receptors.** *Trends Pharmacol Sci* 2001, **22**(2):91-96.
37. Kohout TA, Lefkowitz RJ: **Regulation of G protein-coupled receptor kinases and arrestins during receptor desensitization.** *Mol Pharmacol* 2003, **63**(1):9-18.
38. Hur EM, Kim KT: **G protein-coupled receptor signalling and cross-talk: achieving rapidity and specificity.** *Cell Signal* 2002, **14**(5):397-405.
39. Ishii M, Kurachi Y: **Physiological actions of regulators of G-protein signaling (RGS) proteins.** *Life Sci* 2003, **74**(2-3):163-171.
40. Lodish HF: **Molecular cell biology**, 4th edn. New York: W.H. Freeman; 2000.
41. Modrek B, Lee C: **A genomic view of alternative splicing.** *Nat Genet* 2002, **30**(1):13-19.
42. Hallegger M, Llorian M, Smith CW: **Alternative splicing: global insights.** *Fews J* 2010, **277**(4):856-866.
43. Kelemen O, Convertini P, Zhang Z, Wen Y, Shen M, Falaleeva M, Stamm S: **Function of alternative splicing.** *Gene* 2013, **514**(1):1-30.
44. Collins FS, Lander ES, Rogers J, Waterston RH, Conso IHGS: **Finishing the euchromatic sequence of the human genome.** *Nature* 2004, **431**(7011):931-945.
45. Tazi J, Bakkour N, Stamm S: **Alternative splicing and disease.** *Biochim Biophys Acta* 2009, **1792**(1):14-26.
46. Wang GS, Cooper TA: **Splicing in disease: disruption of the splicing code and the decoding machinery.** *Nat Rev Genet* 2007, **8**(10):749-761.

47. Keren H, Lev-Maor G, Ast G: **Alternative splicing and evolution: diversification, exon definition and function.** *Nat Rev Genet* 2010, **11**(5):345-355.
48. Yeakley JM, Hedjran F, Morfin JP, Merillat N, Rosenfeld MG, Emeson RB: **Control of calcitonin/calcitonin gene-related peptide pre-mRNA processing by constitutive intron and exon elements.** *Mol Cell Biol* 1993, **13**(10):5999-6011.
49. Watari K, Nakaya M, Kurose H: **Multiple functions of G protein-coupled receptor kinases.** *J Mol Signal* 2014, **9**(1):1.
50. Pitcher JA, Freedman NJ, Lefkowitz RJ: **G protein-coupled receptor kinases.** *Annu Rev Biochem* 1998, **67**:653-692.
51. Weller M, Virmaux N, Mandel P: **Light-stimulated phosphorylation of rhodopsin in the retina: the presence of a protein kinase that is specific for photobleached rhodopsin.** *Proceedings of the National Academy of Sciences of the United States of America* 1975, **72**(1):381-385.
52. Premont RT, Macrae AD, Stoffel RH, Chung N, Pitcher JA, Ambrose C, Inglese J, MacDonald ME, Lefkowitz RJ: **Characterization of the G protein-coupled receptor kinase GRK4. Identification of four splice variants.** *The Journal of biological chemistry* 1996, **271**(11):6403-6410.
53. Hisatomi O, Matsuda S, Satoh T, Kotaka S, Imanishi Y, Tokunaga F: **A novel subtype of G-protein-coupled receptor kinase, GRK7, in teleost cone photoreceptors.** *Fefs Lett* 1998, **424**(3):159-164.
54. Benovic JL, DeBlasi A, Stone WC, Caron MG, Lefkowitz RJ: **Beta-adrenergic receptor kinase: primary structure delineates a multigene family.** *Science* 1989, **246**(4927):235-240.
55. Benovic JL, Gomez J: **Molecular cloning and expression of GRK6. A new member of the G protein-coupled receptor kinase family.** *The Journal of biological chemistry* 1993, **268**(26):19521-19527.
56. Willets JM, Challiss RA, Nahorski SR: **Non-visual GRKs: are we seeing the whole picture?** *Trends Pharmacol Sci* 2003, **24**(12):626-633.
57. Eichmann T, Lorenz K, Hoffmann M, Brockmann J, Krasel C, Lohse MJ, Quitterer U: **The amino-terminal domain of G-protein-coupled receptor kinase 2 is a regulatory Gbeta gamma binding site.** *The Journal of biological chemistry* 2003, **278**(10):8052-8057.

58. Premont RT, Macrae AD, Aparicio SA, Kendall HE, Welch JE, Lefkowitz RJ: **The GRK4 subfamily of G protein-coupled receptor kinases. Alternative splicing, gene organization, and sequence conservation.** *The Journal of biological chemistry* 1999, **274**(41):29381-29389.
59. Minneman KP: **Splice variants of G protein-coupled receptors.** *Mol Interv* 2001, **1**(2):108-116.
60. Nakao M, Barrero RA, Mukai Y, Motono C, Suwa M, Nakai K: **Large-scale analysis of human alternative protein isoforms: pattern classification and correlation with subcellular localization signals.** *Nucleic Acids Res* 2005, **33**(8):2355-2363.
61. Johnson JM, Castle J, Garrett-Engele P, Kan Z, Loerch PM, Armour CD, Santos R, Schadt EE, Stoughton R, Shoemaker DD: **Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays.** *Science* 2003, **302**(5653):2141-2144.
62. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool.** *J Mol Biol* 1990, **215**(3):403-410.
63. Bateman A, Birney E, Durbin R, Eddy SR, Finn RD, Sonnhammer EL: **Pfam 3.1: 1313 multiple alignments and profile HMMs match the majority of proteins.** *Nucleic Acids Res* 1999, **27**(1):260-262.
64. Theissen G: **Secret life of genes.** *Nature* 2002, **415**(6873):741.
65. Jeffery CJ: **Molecular mechanisms for multitasking: recent crystal structures of moonlighting proteins.** *Curr Opin Struc Biol* 2004, **14**(6):663-668.
66. Bateman OA, Purkiss AG, van Montfort R, Slingsby C, Graham C, Wistow G: **Crystal structure of eta-crystallin: adaptation of a class 1 aldehyde dehydrogenase for a new role in the eye lens.** *Biochemistry-US* 2003, **42**(15):4349-4356.
67. Jeffery CJ: **Moonlighting proteins.** *Trends Biochem Sci* 1999, **24**(1):8-11.
68. Hulo N, Sigrist CJ, Le Sauvage V, Langendijk-Genevaux PS, Bordoli L, Gattiker A, De Castro E, Bucher P, Bairoch A: **Recent improvements to the PROSITE database.** *Nucleic Acids Res* 2004, **32**(Database issue):D134-137.

69. Pietrokovski S, Henikoff JG, Henikoff S: **The Blocks database--a system for protein classification.** *Nucleic Acids Res* 1996, **24**(1):197-200.
70. Hunter S, Jones P, Mitchell A, Apweiler R, Attwood TK, Bateman A, Bernard T, Binns D, Bork P, Burge S *et al*: **InterPro in 2011: new developments in the family and domain prediction database.** *Nucleic Acids Res* 2012, **40**(D1):D306-D312.
71. Attwood TK: **The PRINTS database: a resource for identification of protein families.** *Brief Bioinform* 2002, **3**(3):252-263.
72. Bartlett GJ, Borkakoti N, Thornton JM: **Catalysing new reactions during evolution: economy of residues and mechanism.** *J Mol Biol* 2003, **331**(4):829-860.
73. Sadowski MI, Jones DT: **The sequence-structure relationship and protein function prediction.** *Curr Opin Struc Biol* 2009, **19**(3):357-362.
74. Punta M, Ofran Y: **The rough guide to in silico function prediction, or how to use sequence and structure information to predict protein function.** *PLoS Comput Biol* 2008, **4**(10):e1000160.
75. Hardin C, Pogorelov TV, Luthey-Schulten Z: **Ab initio protein structure prediction.** *Curr Opin Struc Biol* 2002, **12**(2):176-181.
76. Ishida T, Nishimura T, Nozaki M, Inoue T, Terada T, Nakamura S, Shimizu K: **Development of an ab initio protein structure prediction system ABLE.** *Genome Inform* 2003, **14**:228-237.
77. Menon R, Roy A, Mukherjee S, Belkin S, Zhang Y, Omenn GS: **Functional implications of structural predictions for alternative splice proteins expressed in Her2/neu-induced breast cancers.** *J Proteome Res* 2011, **10**(12):5503-5511.

CHAPTER 2

FUNCTIONAL UNDERSTANDING OF THE DIVERSE EXON-INTRON

STRUCTURES OF THE HUMAN GPCR¹

¹Hammond DA, Olman V, Xu Y. 2014. *J Bioinform Comput Biol*, **12**(1):1350019.
Reprinted here with the permission of the publisher

Abstract

The GPCR genes have a variety of exon-intron structures even though their proteins are all structurally homologous. We have examined human GPCR genes with at least two functional protein isoforms, totaling 199, aiming to gain an understanding of what may have contributed to the large diversity of the exon-intron structures of the GPCR genes. The 199 genes have a total of 808 known protein in splicing isoforms with experimentally verified functions. Our analysis reveals that 1,301 (80.6%) adjacent exon-exon pairs out of the total of 1,613 in the 199 genes have either exactly one exon skipped or the intron in-between retained in at least one of the 808 protein splicing isoforms. This observation has a statistical significance *p*-value of $2.051762^* e^{-09}$, assuming that the observed splicing isoforms are independent of the exon-intron structures. Our interpretation of this observation is that the exon boundaries of the GPCR genes are not randomly determined; instead they may be selected to facilitate specific alternative splicing for functional purposes.

2.1 Introduction

It is a well-accepted principle in genomics that the selection of specific exons and in some instances retention of introns in a gene results in the formation of specific isoforms of the gene. This chapter focuses on exploring the general architecture of GPCR genes and examine how the exon-intron structures influence it's alternative splicing. It will be demonstrated how the exon-intron boundaries in GPCRs are not randomly determined but are arranged in order to facilitate easy

formation of alternatively spliced isoforms.

We carried out statistical analysis on 808 alternatively spliced isoforms of human GPCR with experimentally validated functional protein product and at least two validated spliced isoforms, which came from 199 genes [1]. The idea of our analysis was to show statistically that there is a strong correlation between the specific way in splitting the 199 GPCR coding messages into the current sets of exon-exon boundaries and the 808 known splicing isoforms. While this correlation does not necessarily imply causality relationships, we speculate that the constraints cast by the splicing isoforms needed by the diverse GPCR functions may have contributed to the splitting of the GPCR coding messages into the current exon-exon boundaries.

The following tasks were conducted to accomplish the aim of this chapter:

- i. Survey the general exon-intron structure of GPCRs,
- ii. Determine the type and localization of alternative splicing events that occur in GPCR gene,
- iii. Determine exon boundaries and its constraints on the alternative splicing of the genes,
- iv. Determine the potential impact of spliced exons on the functioning of the resulting proteins by identifying conserved motifs in such exons
- v. Validate the relationship observed between exon boundaries and alternative slicing event by calculating a correlation between them along with a p-value.

2.2 Data Collection and Analysis

2.2.1 Dataset

From the HUGO Gene Nomenclature Committee (HGNC) database [2], 205 human non-olfactory GPCR genes with multiple transcripts were retrieved. The select genes were then filtered to retain only those with at least two protein products and are not candidates for nonsense-mediated decays based on the ensEMBL (ensEMBL 71 (e! 71) – April 2013)[1], database content. This resulted in a total of 199 genes producing 808 experimentally verified alternatively spliced isoforms of GPCRs, which formed the data used for this study. From the ensEMBL database we also retrieved genomic data (localization data on exons, introns, transcription start sites (TSS), and regulatory sequence data) on each associated isoform. These 199 genes covered four out of the five major families of the GPCR genes: glutamate, rhodopsin, taste/fizzle and secretin families ([Supplementary Table 2.1](#)).

2.2.2 Data Analysis

In this section, all the needed analysis to aid in finding out how the exon-intron structures affects the occurrence of specific alternative splicing events are explained. We began by identifying the secondary structure features of the exons using the hidden markov model prediction tool for predicting trans-membrane protein topology, TMHMM program [3]. This provided data on features such as start and end positions of trans-membrane helices, and loops, in translated and un-translated exons. The presence of conserved sequence motifs on each exon was determined using the expectation maximization technique as used by the MEME [4] motif finding tool due to its robustness and ease of implementation. Additionally, a

search of protein domains, and functional site was conducted with a search of transcript sequences against the PROSITE database [5].

Subsequently, an assessment of the genomic structure of isoforms was conducted as follows. First, considered the general architecture of the genomic structure of isoforms. Then, identified the type of splicing event(s) that resulted in each isoform, taking note of their location on the gene. For each pair of exons (exon boundaries), we counted the number with an alternative splicing event in any isoform. A splicing event of exon removal, intron retention, alternative 3' splice site selection, alternative 5' splice site selection, or mutually exclusive is counted only once per gene per exon. Thus, even if the event occurs in multiple isoforms in the same exon or intron it is counted as one event occurrence in the gene.

For decades, a number of user-friendly algorithms have been designed to predict the entire topology of trans-membrane proteins from their amino acid sequences. Most of them are web-based with the option to obtain stand-alone versions for academic use. A predicted topology describes all the possible trans-membrane helices and their location on the sequence, the orientation of the helices in the membrane, and the location of the N-terminal. In the beginning, topology prediction tools relied mainly on the hydrophobicity properties of the residues, and the use of “positive inside” rule [6]; e.g. TopPred [7], DAS-TMfilter [8], and SOSUI [9]. In these algorithms, a threshold of hydrophobicity is set below which a segment is considered not to be a trans-membrane. A downside to this approach is the potential to miss actual trans-membrane helices that fall just short of the threshold.

Other method employed in the prediction of transmembrane topology is the use of various machine-learning techniques; PHD [10], and MEMSAT3 [11] that

utilizes neural networks (NN) in combination with evolutionary information; other methods combine a number machine learning methods together in an attempt to improve prediction accuracy, MemBrain [12].

There is also the use of hidden markov model (HMM) based methods for topology prediction. It has been shown from review and whole-genome studies to be one of the best methods to use. Its success can be attributed to its ability to incorporate hydrophobicity, charge bias, helix lengths, and grammatical constraints into one model for which algorithm for parameter estimation and prediction already exist. Popular HMM-based tools include TMHMM [3], HMMTOP [13], THUMBUP [14], and Phobius [15].

Compared with individual methods, consensus methods have been shown to yield better results than individual methods because they exploit the advantages of integrated individual methods [16, 17]. Due to their advantages as discussed above, the HMM method of topology prediction was preferred over other methods. The TMHMM program was chosen because compared to other HMM based methods, TMHMM designate seven possible different HMM state: 1) core TM helix, 2) amino-tail, 3) cytoplasmic-tail, 4) cytoplasmic loops, 5) short and, 6) long extracellular loops, and 7) globular-domain-like structures in the loops. Additionally, TMHMM program was preferred because it is estimated to identify ~97-98% of all transmembrane helices; secondly, it can discriminate between soluble and membrane proteins with both specificity and sensitivity to be better than 99%; and lastly it is regularly maintained and updated [3].

There was the use of the expectation maximization technique to determine the presence of conserved motifs. This was preferred because one, it is by far the fastest and well-studied algorithm for learning mixture models; and two, the

algorithm maximizes only the likelihood without biasing the means towards zero, or bias the cluster sizes to have specific structures that might or might not apply.

2.2.3 Hypothesis testing

Thereafter, we assessed the statistical significance of the observed events occurring across an exon boundary. We tested the hypothesis that exon boundaries with splicing signals are independent of each other. For each isoform, let S be the number of exon clusters, with a *cluster* being the maximum run of consecutive exons without an alternative splicing event. Based on the null hypothesis all R-exon isoforms of the E-exon genes should have the equal probability to be formed. Then the density of S can be calculated using the following formula based on the Theory of Series, which is a well-established model for studying series of runs since the 1940's [18].

$$P(S | E, R) = \binom{R-1}{S-1} \binom{E-R+1}{S} / \binom{E}{R}$$

with E being the number of exons in the underlying gene and R the number of exons present in the current isoform. The p -value for each observed isoform can be calculated using

$$\sum_{S \leq S_{data}} P(S | E_{data}, R_{data})$$

with S_{data} , E_{data} and R_{data} being the S , E and R values for the current isoform, respectively. We then calculate the total p -value as the geometric average of the p -values across all the 808 isoforms, giving rise to a p -value = $2.051762 * e^{-09}$ for the whole set of splicing isoforms. This implies that our null hypothesis is false, hence indicating that there is a strong relationship between the current exon boundaries in

the 199 GPCR genes and the specific set of the 808 observed splicing isoforms. Note that this specific set of 808 splicing isoforms already largely defines the exon boundaries as the above formulation requires the consistency between the exon boundaries defined in the genes and the segments of the messages included in the splicing isoforms when we consider only two types of splicing events, whole exon skipping and intronic region retention.

2.3 Results

It was observed that the genomic structures of GPCRs are quite complex. They contain multiple exons intercepted by intronic regions. Exon start and end positions (exon boundaries) differ amongst isoforms as illustrated by the example in Figure 2.1 (GPCR gene DRD2 and its known alternatively spliced isoforms). A simple examination of their gene structures reveals that the number of exons across different GPCR genes ranges from 2 to 42 (average 14 exons per gene) with the exon length ranging from as low as 15 bps to well over 10,000 bps. Introns also ranged in length from 4 base pairs to 80,000 base pairs. 98% of splicing event observed in the genes was exon skipping and alternative 5' and 3'. These events were predominantly localized in the un-translated regions (UTR) of the genes and as signal peptides (Figure 2.2). Less than 1% of observed splicing events was an intron retention event.

The 199 genes under consideration contained 1,613 exon boundaries. 1,203 boundaries had either an exon skipped on either side of the boundary or an alternative 5' or 3' splice site event, and 98 had the intron between them retained. Thus, 80.7% of exon boundaries in our dataset were used for an alternative splicing event.

The statistical significance of this observation, assuming that the splicing events (exon-skipping, alternative 5' and 3', and intron-retention) were random, was 2.051762×10^{-9} , using the method given in section 2.2.3 above. This high statistical significance indicates that there is a strong correlation between all the 808 known splicing isoforms and the exon boundaries in the 199 GPCRs.

Furthermore, the prediction of secondary structure features of exons revealed that in 100% of the time, alternative splicing events that occur in translated exons of the gene mapped to the loop regions of the proteins. In 17% of the times they also covered helical regions.

Functional motif search found that 74.7% of exons with alternative splicing event contain functional motifs. Detailed data for these statistics are given in [Supplementary Table 2.2](#).

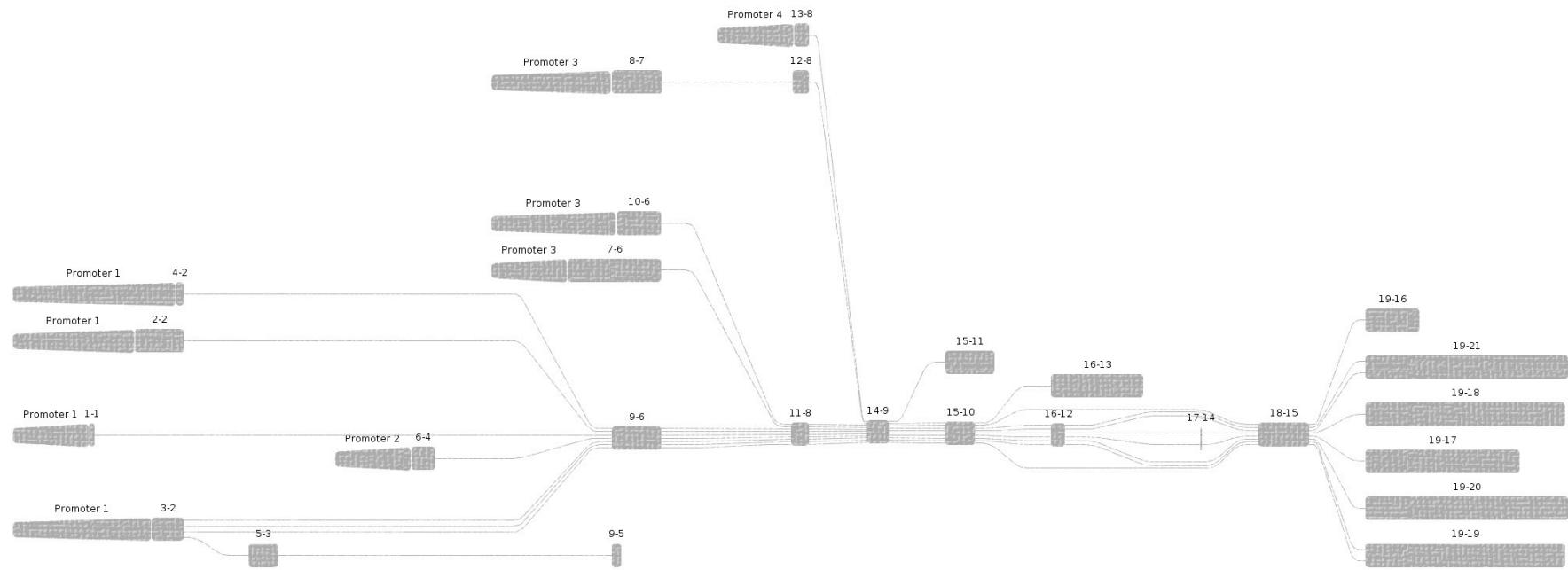


Figure 2.1: Splicing graph view of transcripts from alternatively spliced isoform of genes

Graphical representation of the locus from the human gene DRD2. In this view, shared exons are merged. Paths through the splicing graph describe isoforms. Multiple identical splice junctions connect two or more exons, if they are in multiple isoforms. The positions of the exons are preserved. Thus, exons that start or stop at the same position are aligned accordingly. To obtain a condensed view, introns are not drawn to scale. Two numbers describe each exon: the first number describes the relative start position and the second number describes the relative stop position of the exon. If two exons have the same start position on the genome, they both get the same number. An example, the exon 2-1 and 3-1 end in the same position, but 2-1 start at an early position on the genome before exon 3-1.

Image generated by the Genomatix Transcriptome Viewer

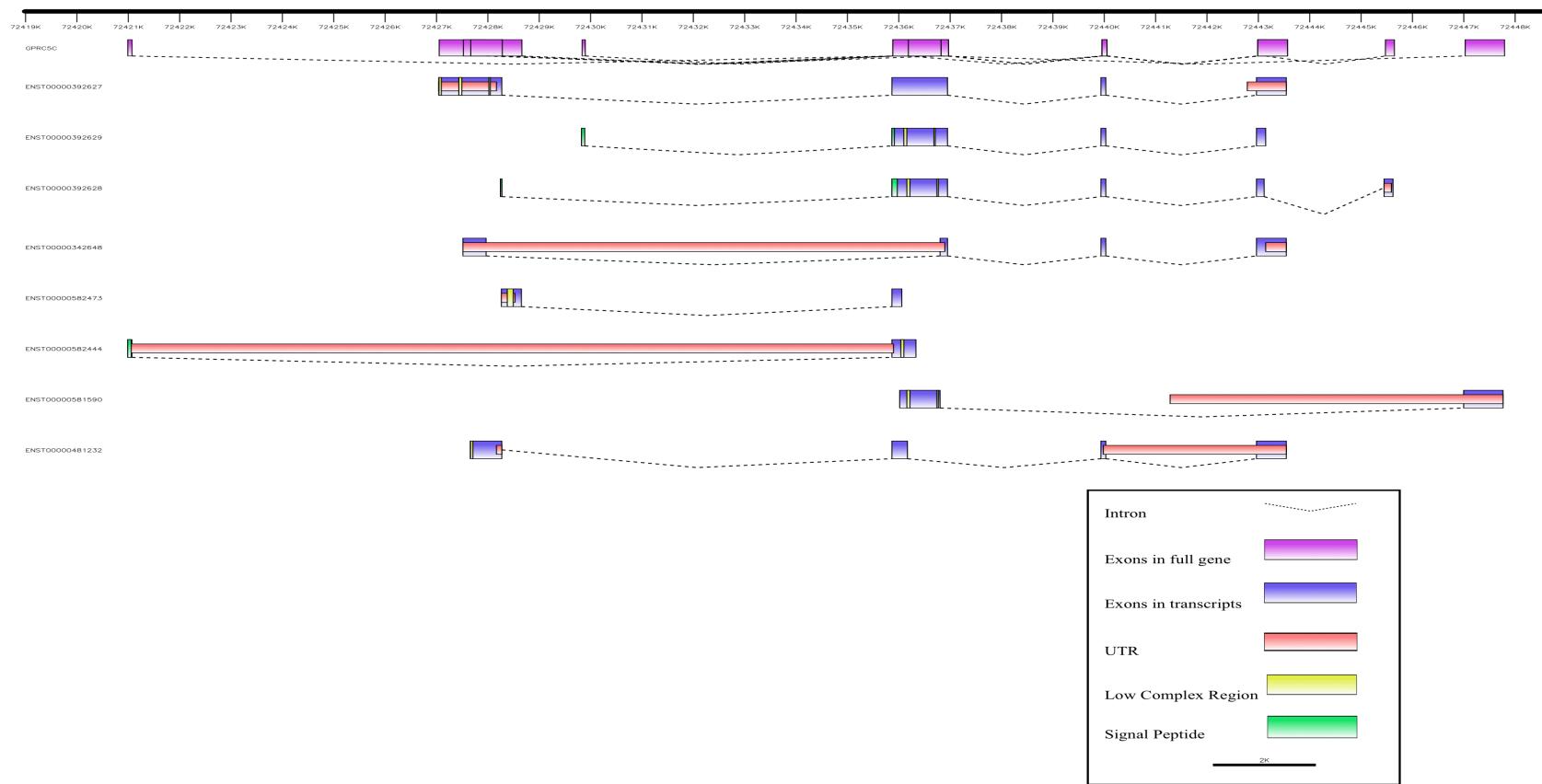


Figure 2.2: A graphical representation of splicing variants of GPRC5C gene. Shaded bars indicate exons and dash lines indicate intronic regions. Top purple gene structure shows the entire gene region. Bottom turquoise graphs show the selected exons for the various mRNAs isoforms with resulting protein products for the gene. Each isoform is represented by its exon-intron structures exhibiting the inclusion and/or exclusion of certain exons. See legend for detailed color annotation.

2.4 Discussions

The purpose of this study was to assess the exon-intron structure of GPCRs and determine if they had any functional influence on alternative splicing of the genes. The results distinctly suggest that the exon-intron structures of GPCRs are not randomly determined, but are arranged in such a manner as to facilitate easy alternative splicing for functional purposes. This is supported by the findings that ~80% of exon boundaries were used in an alternative splicing event. The results also show that 74.7% of exon used in an alternative splicing event contains conserved motifs. Hence leading us to speculate that these alternative splicing events are also for functional purposes, since an isoform missing an exon containing a functional motif will not have use of the function of that motif while an isoform with the exons containing such motifs will potentially have use of that function.

It was also obvious that GPCRs have very complex genomic structure. This is evident by the number of exons per GPCR gene, shown to range anywhere from 2 to 42 exons (average 14 exons per GPCR gene). Also, conforming to eukaryotic research data, the dominant alternative slicing event observed in GPCRs were the exon skipping and alternative 5' and alternative 3' events.

The results also implicitly suggest that almost 20% of exon boundaries were not used in an alternative slicing event. Reasons for this could be that not all possible alternatively spliced isoforms of the genes were considered. Thus, exons that were not observed to be involved in an alternative splicing events were as a result of limited sampled dataset. Another reason may be that not all exons must be involved in an alternative splicing event. This may be nature's way of preserving the integrity of the gene.

Overall, this study has given information on GPCR gene structure and how it affects the alternative splicing of the gene. It however limited the work to verified alternative spliced isoforms of GPCR. It may be that if we had considered all possible isoforms of the genes, the number of exons involved in a splicing event would have increased. It is also unclear if the same relationship can be observed in other gene families. As more isoform data become available on GPCRs and other gene families, this study can be replicated to determine if the observed relationship between exon-intron structure and alternative splicing events still hold.

2.5 Concluding Remarks

For the first time, evidence have been given to show the type of alternative splicing event that occur in GPCRs and how that may influence the general functioning of their proteins. With a statistical significance *p-value* of $2.051762 * e^{-09}$, we observed that ~80% of adjacent exon boundaries in GPCRs are correlated with known splicing event. Also, that spliced exons contain important motifs suggesting that the alternative splicing is for functional purposes. We can therefore conclude that the genomic structures of GPCRs are arranged in such a manner as to facilitate easy alternative splicing of the genes for functional purposes. Clearly this warrants further and detailed phylogenetic analyses to confirm the roles of splicing isoforms in the determination of the exon boundaries, which represents a highly important evolutionary and fundamental biological question. While the data for GPCR splicing isoforms are clearly incomplete, we believe that the dataset we used is large enough, and the statistical significance is high enough to warrant further study along this

line of research in elucidation of the fundamental causes in the selection of the specific exon boundaries.

Funding

This work was supported by National Institute of General Medical Sciences, National Institute of Health [grant number GM089093] Ruth L. Kirschstein Predoctoral Fellowship.

REFERENCES

1. Strozzi F, Aerts J: **A Ruby API to query the Ensembl database for genomic features.** *Bioinformatics* 2011, **27**(7):1013-1014.
2. Gray KA, Daugherty LC, Gordon SM, Seal RL, Wright MW, Bruford EA: **Genenames.org: the HGNC resources in 2013.** *Nucleic Acids Res* 2013, **41**(Database issue):D545-552.
3. Krogh A, Larsson B, von Heijne G, Sonnhammer ELL: **Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes.** *J Mol Biol* 2001, **305**(3):567-580.
4. Bailey TL, Elkan C: **Fitting a mixture model by expectation maximization to discover motifs in biopolymers.** *Proc Int Conf Intell Syst Mol Biol* 1994, **2**:28-36.
5. Sigrist CJ, de Castro E, Cerutti L, Cuche BA, Hulo N, Bridge A, Bougueleret L, Xenarios I: **New and continuing developments at PROSITE.** *Nucleic Acids Res* 2013, **41**(Database issue):D344-347.
6. Heijne G: **The distribution of positively charged residues in bacterial inner membrane proteins correlates with the trans-membrane topology.** *The EMBO journal* 1986, **5**(11):3021-3027.
7. von Heijne G: **Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule.** *J Mol Biol* 1992, **225**(2):487-494.
8. Cserzo M, Eisenhaber F, Eisenhaber B, Simon I: **TM or not TM: transmembrane protein prediction with low false positive rate using DAS-TMfilter.** *Bioinformatics* 2004, **20**(1):136-137.
9. Hirokawa T, Boon-Chieng S, Mitaku S: **SOSUI: classification and secondary structure prediction system for membrane proteins.** *Bioinformatics* 1998, **14**(4):378-379.

10. Rost B, Casadio R, Fariselli P: **Refining neural network predictions for helical transmembrane proteins by dynamic programming.** *Proc Int Conf Intell Syst Mol Biol* 1996, **4**:192-200.
11. Jones DT: **Improving the accuracy of transmembrane protein topology prediction using evolutionary information.** *Bioinformatics* 2007, **23**(5):538-544.
12. Shen H, Chou JJ: **MemBrain: improving the accuracy of predicting transmembrane helices.** *PLoS One* 2008, **3**(6):e2399.
13. Tusnady GE, Simon I: **The HMMTOP transmembrane topology prediction server.** *Bioinformatics* 2001, **17**(9):849-850.
14. Zhou H, Zhou Y: **Predicting the topology of transmembrane helical proteins using mean burial propensity and a hidden-Markov-model-based method.** *Protein science : a publication of the Protein Society* 2003, **12**(7):1547-1555.
15. Kall L, Krogh A, Sonnhammer EL: **A combined transmembrane topology and signal peptide prediction method.** *J Mol Biol* 2004, **338**(5):1027-1036.
16. Nilsson J, Persson B, von Heijne G: **Consensus predictions of membrane protein topology.** *Fefs Lett* 2000, **486**(3):267-269.
17. Ikeda M, Arai M, Lao DM, Shimizu T: **Transmembrane topology prediction methods: a re-assessment and improvement by a consensus method using a dataset of experimentally-characterized transmembrane topologies.** *In Silico Biol* 2002, **2**(1):19-33.
18. Mood AM: **The distribution theory of runs.** Princeton,; 1940.

CHAPTER 3
FUNCTIONAL ANNOTATION OF ALTERNATIVELY SPLICED
ISOFORMS OF HUMAN G PROTEIN-COUPLED RECEPTORS ¹

¹Hammond DA, and Xu Y

To be submitted to *Nucleic Acids Research* 2015

3.1 Introduction

A principal element of any genome project is the annotation of identified genes. Since it is from functional annotation that the biology of the gene and hence the organism is elucidated. Typically, the uncharacterized genes/proteins are compared to a set of already functionally characterized genes/protein. This is usually done on the peptide level as protein sequences are more conserved than their corresponding nucleotide sequence. As explained in chapter 1, homology-based methods of functional annotation do not work for spliced isoforms due to the lack of homology between isoforms. The motif-based approach, although may not provide distinctively unique functional annotations for the spliced isoforms due to their “similar” amino acid content, can still provide some useful family function information. In this chapter, we seek to annotate the functions of alternatively spliced isoforms of GPCRs.

Evidence from genome-wide studies and proteomics studies has shouldered the responsibility of functional complexity in the human genome on alternative slicing. A large percentage of the diversities that exist between gene transcripts come from either tissue-specific or stage-specific expression. Such conditional expression must be appropriately evaluated in order to functionally distinguish one isoform from another, especially those that have originated from alternative splicing.

Due to their ligand binding diversity, GPCRs are of great pharmaceutical interest. These interests though, have been focused largely on the tertiary structure identification, ligand binding potential, and understanding of their mechanism of activation, with little attention given to its alternative splicing and how it influences

its primary functions. In the unlikely case that functional information is given on such isoforms, they often inherit molecular level function of their “primary” full-length protein (which is usually the transcript that was first sequenced).

As discussed in earlier chapters, for alternatively spliced isoforms whose splicing events occur in the un-translated regions of the genes, molecular level annotations alone are often not sufficient to distinctively annotate their primary function from other isoforms from the same gene. Also, it has been discussed why the use of the conventional functional methods is incapable of providing the annotation as well. In this chapter therefore, we introduce a framework of computational methods that will best provide distinct primary cellular level functional annotation for GPCR isoforms. We however begin this chapter by annotating molecular level functions for isoforms in our dataset then continue on with the cellular level annotation. Lastly, we integrate both molecular level annotation and cellular level annotations to distinctively distinguish 1044 GPCR isoforms.

3.2 GPCR Data Retrieval and Processing

Human GPCR genes data were obtained from the HUGO Gene Nomenclature Committee (HGNC) database, which were supplemented with data from the IUPHAR-DB [1]. The HGNC database has GPCRs classified into the following classes: rhodopsin-like, secretin, glutamate, and fizzled, plus a fifth unclassified group. IUPHAR-DB was used to supplement the data from HGNC to ensure that our select dataset is a good representation of the complete GPCR family. IUPHAR-DB was chosen because it is one of the most up-to-date publicly available databases on GPCRs and other membrane proteins that are of pharmaceutical interest. A total of

436 non-olfactory GPCR genes are available in these two databases. After filtering to remove duplicate genes, 394 non-olfactory GPCR genes were retained. Gene transcripts were then obtained from the ensEMBL (e! 75 February 2014). A total of 1,390 alternatively spliced isoforms of these genes expressed in various tissue types were retrieved. They were further filtered to remove isoforms without protein product and those that are candidates for nonsense mediated decay, resulting in the retention of 1,044 transcripts that have resulted from alternative splicing of 332 genes, which makes up the dataset for the research analysis in this chapter (Supplementary Table 3.1).

3.3 Annotation of Molecular Function

The majority of proteins consist of multiple conserved patterns or motifs. These motifs are made up of any conserved element of sequence, which comprises of a short sequence of contiguous residues or in a more distributed pattern. Functionally related sequences are believed to share similar distribution patterns of critical functional residues that are not necessarily contiguous. In this section, we explored the existence of functional motifs in our dataset in two main ways: first, was to use our sequences to query motif databases in search of patterns that can help suggest functional activity; and secondly define new patterns specific to our set of sequences.

3.3.1 Outline of Framework

Molecular level functional information

1. Motif finding by database search
 - a. Used protein sequences of our isoforms to query motif databases: PROSITE, PRINTS, and BLOCKS.

- b. Consolidate similar functions from these databases into one functional motif based on its description of function
- 2. Finding new motifs specific to our dataset
 - a. Using multiple sequence alignment by log-expectation algorithm sequence alignment are generated
 - b. Selected candidates motifs, defined as conserved patterns observed within alignments to be at least 3 residues long and not more than 35 residues and were present in a set of at least 3 isoform sequences were selected as candidates for consideration as potential functional motifs.
 - c. Tested the hypothesis that each residue position in identified conserved pattern can occur at random in all isoforms.

3.3.2 Motif Finding by Database Search

The leading databases that maintain collections of patterns or motifs associated with a function are PROSITE [2], PRINTS [3], and BLOCKS [4]. Using protein sequences of our dataset as inputs, these databases were queried for predefined functional conserved patterns using the InterProScan [5] server. InterPro does not contribute any new information to the search. Its power comes from having all these databases in one place providing a range of evidence for a protein annotation.

PROSITE database contains a collection of pattern descriptions that are usually associated with a biochemical function. These signatures, represented as regular expression patterns, are generated from curated multiple sequence alignments and describe conserved positions within a domain family. For each pattern, signature profiles are developed, which assign a score to each of the 20

amino acids at each position of the pattern according to the frequency of which each amino acid is found at a particular position. Further, alternative protein structure-based profiles and methods involving hidden Markov models have been employed in PROSITE to identify patterns.

PRINTS on the other hand maintain collections of protein fingerprints. The concept behind fingerprints is that a protein can be represented by several conserved motifs. A fingerprint is defined as an ordered list of these motifs that describes a protein family. The database is built from SwissProt and TrEMBL [6] datasets. Each entry is associated with bibliographic information, functional descriptions, lists of matching sequences and comments. The current version of PRINTS (version 42.0) contains 2,156 fingerprints encoding 12,444 single motifs.

Similar to PRINTS, the BLOCKS database contains a list of motifs that are representative for a family. It is generated from protein family databases such as PFAM [7], PRINTS, and ProDom [8]. BLOCKS identify local motifs within given protein families but do not find new protein families since it uses domain families of existing databases as input. The BLOCKS database can be searched with user protein sequences via the BLIMPS program [9]. The current (and last) version of BLOCKS database version 14.3 contains 29,068 blocks representing 5900 families.

Searches were done using all default parameters except for the exclusion of motifs with a high probability of occurrence. The motifs classified to have a high probability of occurrence in these databases are mostly patterns that are found to be abundant in many protein sequences. Which often include commonly found post-translational modifications and some other compositionally biased regions. While it is generally useful to note their presence, they serve no practical purpose in this research, hence their exclusion. Motifs from the three database sources were

consolidated using the description and name of each motif. Due to the overlaps that exist between these databases, integration of results was smooth.

Harnessing all the information that these different databases can provide – from motif patterns, motif profiles, protein fingerprints etc.; the motifs identified were used to infer molecular level function to our dataset. Results are provided below.

3.3.3 Finding New Pattern Specific to Our Dataset

Data Preprocessing

The data used for this section of analysis were the protein sequences of the 1044 isoforms of the 332 human GPCR genes. Sequences were verified to contain at least 10 amino acids in length. This was done on the assumption that, isoforms with sequence length less than 10 amino acids does not contain enough amino acids to yield useful functional motif information.

Multiple Sequence Alignment

Using a multiple sequence comparison by log-expectation algorithm [10] sequence alignment was generated. Conserved patterns observed within alignments to be at least 3 residues long and not more than 35 residues and were present in a set of at least 3 sequences were selected as candidates for consideration as potential functional motifs. Validation statistics for selected motifs were then calculated to distinguish random pattern occurrences from an actual potential functional motif.

Validation of Motif

Using the algorithm described below we determine the probability that an identified motif can occur by chance.

Let N be the number of proteins in the family (subset) and a_{ij} is a symbol of amino acid (or $a_{ij} = ' \sim '$ for deletion) in i^{th} protein in j^{th} position, $i = 1 \dots N$, $j = 1 \dots L$ (L is the length of alignment), and I_F be the set of conservative positions for all proteins. Let K be the number of subset proteins, and I_S be the set of conservative positions for the set of proteins in family or subfamily. We answer the question whether the motif presented by I_S is a signature of the family or subfamily. Assume indices of these proteins are $1 \dots K$ and $I_F \cap I_S = \emptyset$, or otherwise $I_S = I_S \setminus I_F$, since the signature should not contain elements of common motif from the original set. For each position we calculate a probability (P) that for a random subset of proteins from the original family, the highest frequency of an amino acid will not be lower than the same value in subfamily. Then using threshold, λ , we find all positions (R) with $P \leq \lambda$ and among them the number of positions (T) from I_S . The final *P-value* is calculated as a probability of having $\geq T$ positions in subfamily motif if the motif was formed by chance. This probability is defined from the formula of hyper geometric distribution as

$$P\text{-value} = \sum_{t \geq T} \frac{C_R^{M-t} C_R^t}{C_R^M}$$

where $M = |I_S|$ - number of elements in I_S .

In other words we test the probability that by randomly selecting a number of sequences within the alignment of sequences the probability of a set of amino acids occurring as a conserved pattern will be less than or equal to $\lambda \leq 0.05$.

Inferred Function

Molecular functional inferences made for each identified motifs were based on public record search of functions that are associated with each family and subfamily motif found. Records used included published research articles, RefSeq database records [11], Gene Ontology Annotations (GO) [12], and information accumulated in the InterPro database [5]. Common functional information identified for the set of isoforms grouped by our motif search algorithm is considered to potentially be the function of the identified motif.

Identified motifs were searched against PROSITE and PFAM databases in search of matches to any known pattern or profile presently available in those database, as well as to aid in categorizing functions for the predicted motifs.

3.3.4 Results from Molecular Function Analysis

Database search

A search for functional motifs through the use of existing motifs databases provided molecular level functional information on isoforms. Majority of the information were family and subfamily information, which helps in classifying the isoforms into the five GPCR families and their subsequent subfamilies (Figure 3.1).

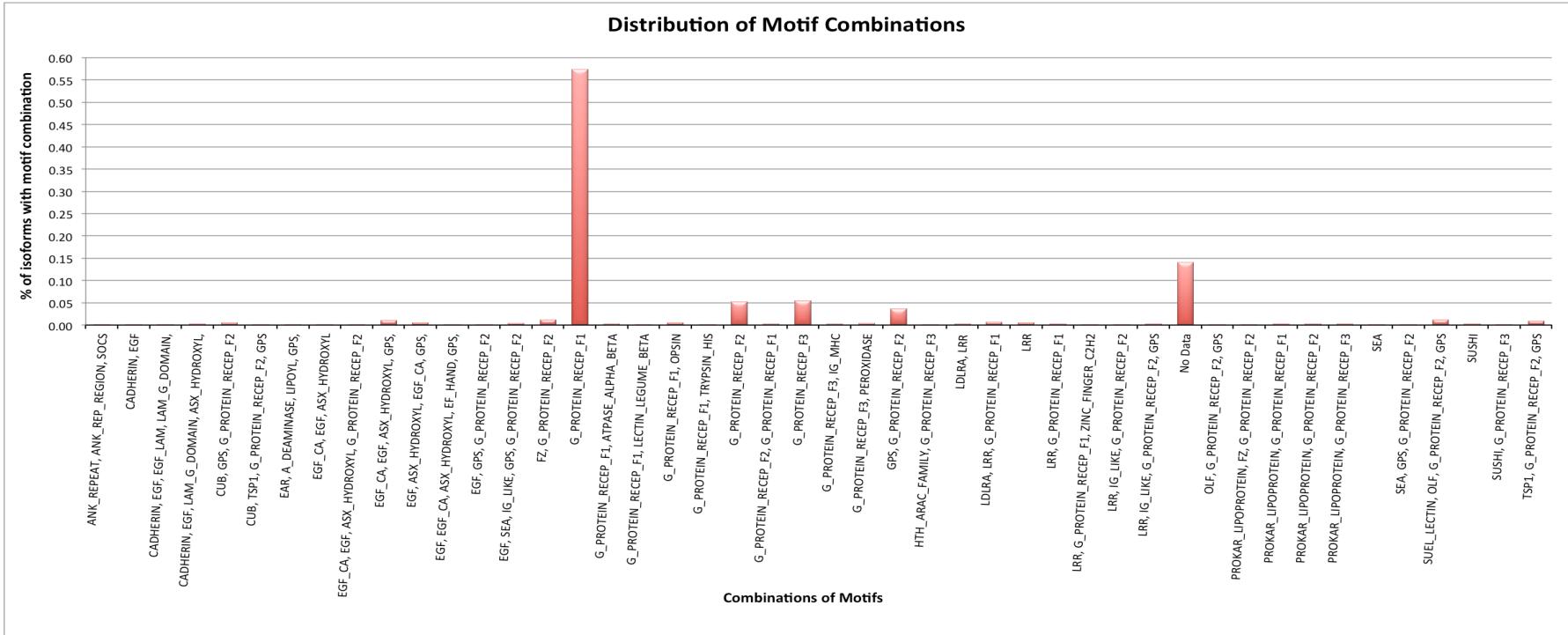


Figure 3.1: Distribution of motifs identified through database search

This figure shows the percentage of isoforms with identified combinations of motifs from database search. It can be observed that motifs identified through this search matched mostly motif and patterns identified to be family-specific signatures. From this figure also, it is evident that there are a set of other non family-specific motifs that are often associated with specific families of GPCR than others, which aids in the molecular functional annotation. See supplementary files for raw values used to produce this figure.

Various combinations of GPCR family-specific and non-GPCR family-specific motifs were observed in the database search results (as illustrated in Figure 3.1). The combinations of these motifs are used to infer potential molecular function for each isoform. It was observed that ~68% of isoforms contained only family motif signature, and 16% contain motif with family-specific signature and non family-specific signature. A few motifs combinations did not include GPCR signature. Motif distribution can be found in Table 3.1 below.

Table 3.1: Functional motifs information identified through database search

Motif	Description	% In Dataset
GPCR family signature	Isoforms that contained motifs that aided in assigning a particular family or subfamily of GPCR to the isoform	87%
Other non GPCR family	Non GPCR family motif information	32%
GPCR family only	Isoforms that contained only specific GPCR family signature	68%
GPCR family + non family	Isoforms with family-specific and non family-specific motifs	16%
Total with motif data	Total of isoforms with any motif data	89%
No Data	Where no motif that met the set threshold was identified an isoform	11%

Also, the database motif search results separated the 1044 isoforms into 45 categories based on the presence of specific motifs plus 1 set with no motif data (Supplementary table 3.2). Clearly, further functional classification is necessary to functionally distinguish one isoform from another.

Table 3.2: New identified motifs in GPCR families and subfamilies

Family	Sub-family	Motifs	Significance	Functional_Annotations
Rhodopsin	Amine	W-x(2)-[WY]-x-[CNS]-S-x(2)-N-P-x(2)-[HY]-[ACGPTV]	1.29E-02	
	Adrenoceptor	D-V-L-x(0,1)-C-x(0,1)-T-[AS]-S-I-x(2)-L-C-x-[IL]-[AS]-[ILV]-D-R-Y	1.41E-03	Mediates its action by association with G proteins that activate a phosphatidylinositol-calcium second messenger system. Its effect is mediated by G(q) and G(11) proteins. Nuclear ADRA1A-ADRA1B heterooligomers regulate phenylephrine(PE)-stimulated ERK signaling in cardiac myocytes.
	Dopamine	D-[IV]-M-x(0,1)-C-x(0,1)-T-A-S-I-[FL]-N-L-C-[AV]-I-S-[IV]-D-R	6.07E-04	
	Trace	F-K-[AQ]-L-H-[ST]-P-T-N-[FW]-L-[ILV]-x-S-[LM]-A-x-[ATV]-D-[FM]	4.83E-03	Neurotransmitters in invertebrates, the extent to which they function as true neurotransmitters in vertebrates has remained speculative
	Lysosphingolipid	W-[FL]-[AL]-R-[EQ]-G-x(4)-[AST]-L-[AGST]-A-S-x(2)-[NST]-L-L	6.33E-03	Act as signalling molecules in addition to their roles as intermediates in membrane biosynthesis. They have roles in the regulation of cell growth, differentiation, apoptosis and development
	Sphingosine	W-F-[AL]-R-E-G-x(2)-F-x-[AT]-L-[AGST]-A-S-[TV]-x-S-L-L	1.74E-03	Cell growth related effects
Nucleotide-like				Energy metabolism and vasodilation. Vasodilation paths differ for adenosine and purinoceptors. Purinoceptor for instance induces vasodilation through endothelium-dependent release of nitric oxide.
	Adenosine	L-[IV]-[FL]-T-x-[AS]-S-I-[FLM]-[AS]-L-L-A-[IV]-A-[IV]-D-R-Y-[IL]	5.89E-02	Depending on where in the body they are found their functional effects may include bronchoconstriction, immunosuppression, inhibition of platelet aggregation, cardiac depression, stimulation of nociceptive afferents, inhibition of neurotransmitter release and inhibition of the release of other factors, e.g. hormones, when found in the peripheral tissues. In the central nervous system (CNS), adenosine reducing motor activity, depresses respiration, induces sleep and relieves anxiety
Peptide	Melanocortin	I-A-[AV]-D-R-Y-[FLV]-[ST]-I-F-[HY]-A-L-[QR]-Y-H-x-I-[MV]-T	2.77E-02	Involved in hair color determination
	Neuropeptide	I-[AGV]-N-L-[AS]-[FLV]-[AS]-D-x-[LM]-[IMV]-[ACNTV]-x-[FLM]-[CN]-x(2)-[FL]-T		Modulate hypothalamic-pituitary-adrenal axis activity, binds neuropeptide
	Opioid	G-S-[KR]-E-K-D-R-[NS]-L-R-R-I-T-R-[LM]-V-L-V-V-V	3.62E-05	Mediate analgesia, hypothermia, respiratory depression, miosis, bradycardia, nausea, euphoria and physical dependence, beta-endorphin being the most potent endogenous ligand; delta-receptors mediate analgesia; and kappa-opioid receptors are believed to mediate analgesia, sedation, miosis and diuresis, dynorphin being the most potent endogenous ligand.
	Orexigenic	V-T-N-x-F-I-x-[NS]-L-[AS]-[ILV]-[AS]-D-[LV]-L-[IV]-[GT]-x(2)-C	6.51E-02	
	Proteinase	M-Y-[ACG]-S-[IV]-L-[FL]-[LM]-[AT]-[ACV]-[ILV]-S-[ILV]-[DNQ]-R-[FY]-x-[AV]-[ILV]-V	1.32E-03	
	Galanin	P-F-[QT]-A-x-[ILV]-Y-[APT]-L-[DP]-[AGT]-W-[LV]-[FL]-G-[ADS]-[FL]-[ILMV]-C-K	2.03E-02	Regulate the secretion of various hormones such as pituitary growth hormone
Prostanoid		I-[ILV]-D-P-W-[IV]-[FY]-[AIL]-[IL]-[FL]-R-x(2)-[LV]-[FL]	5.99E-02	Receptor for prostaglandin. The activity of these receptor is mediated by G(q) proteins which activate a phosphatidylinositol-calcium second messenger system. May play a role as an important modulator of renal function.
	Olfactory	K-x(0,2)-A-[FL]-[ACNST]-T-C-[GS]-S-H-[LV]-x-[LV]-V-x-[ILM]-[FWY]-[FY]-[GS]	5.06E-03	Involved in its function as odorant receptor

New Motif Identification

Fifteen family-specific motifs were identified using the methods detailed in section 3.3.3 above. 4 of the 15 motifs identified were family level motifs, while the remaining 11 were subfamily level motifs. All motifs are shown in Table 3.2 immediately above. Due to lack of sufficient data on other families of GPCRs, identified motifs belong to the rhodopsin-like GPCR family. As more data on other families and their subsequent subfamilies become available, this work can be extended to identify family-specific motifs for these other families as well.

The significant p-values of identified motifs reported (Table 3.2) led to the rejection the null hypothesis that these conserved motifs can occur randomly. A search of identified motifs against current version of PROSITE, and BLOCKS databases of motifs found no existing motif match. Indicating that identified motifs are not currently available in those public databases. Hence, these 15 motifs are presumed to be newly discovered.

3.4 Annotation of Cellular Function

3.4.1 Outline of Framework

1. Identify and process dataset to use in the analysis
2. Retrieve promoter sequences of the isoform dataset
3. Find transcription factors (TFs) that may bind those promoters
4. Determine expression data (tissues of expression) on isoforms and transcription factors.
5. Find common expression pattern between isoforms and TFs
6. Infer isoform function based on transcriptional influence of the TFs

3.4.2 Additional Dataset for Cellular level Functional Analysis

Expression Data (RNA sequencing Data)

Tissues of expression data were obtained from the Human BodyMap project 2.0, which came from transcription profiling by high throughput sequencing from Illumina's HiSeq 2000 technology. They were processed through the TopHat 2 with quantification by cufflinks [13], mapped against Ensembl release 75. This data contain expression level read details for isoforms across individual and mixture of 16 human tissues – adipose, adrenal, brain, breast, colon, heart, kidney, liver, lung, lymph, prostate, skeletal muscle, ovary, testes, thyroid, and white blood cells. The raw data can be downloaded at the following link:

<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE30611>

By virtue of the RNA data sequencing process, the read counts contain some inherent biases including; influence of sequencing depth, dependences on gene length, and differences in count distributions among samples. The first two biases were addressed using the read per kilobase per million (RPKM) method as suggested by Montazavi et al 2008 [14] for each transcript in each of the 16 tissues. This was calculated as:

$$RPKM = \frac{\text{number of reads of the region (transcript)}}{\left(\text{total reads}/1,000,000\right) * \left(\text{region length}/1000\right)}$$

After removing transcripts with zero RPKM across all tissues we corrected for biases that may result from differences in count distribution among samples (tissues) using the Upper Quartile (UQ) normalization method [15]. This was achieved as following:

$$RPKM_{norm(i,j)} = \frac{RPKM_{(i,j)}}{U_i}$$

Where:

i = i^{th} tissue, varying from tissue 1 to 16

j = transcript at j^{th} position

$RPKM_{norm(i,j)}$ = normalized expression value of the j^{th} transcript in the i^{th} tissue

$RPKM_{(i,j)}$ = original expression (raw read count) of j^{th} transcript in i^{th} tissue

U_i = Upper Quartile expression value in the i^{th} tissue

(Supplementary Table 3.3)

3.4.3 Analysis of Data

Promoter Retrieval and Analysis

Using the ensEMBL transcript identifiers of the 1,044 isoforms, promoter sequences were retrieved from the genomatix promoter library. The genomatix promoter library currently contains promoters for 104,473 experimentally determined promoters built on NCBI build 37/hg19. The promoters in the library are evaluated by setting the promoter for each transcript to be 500bp upstream and 100bp downstream of the transcription start site (TSS). Promoters of two or more transcripts are merged into one promoter when they belong to the same locus and the promoter and first exons of the two transcripts overlap (Figure 3.2). TSS is defined as the region of genomic sequence for which experimental evidence for transcription initiation is available (from individual full-length cDNA and from Cap-Analysis Gene Expression - CAGE tags).

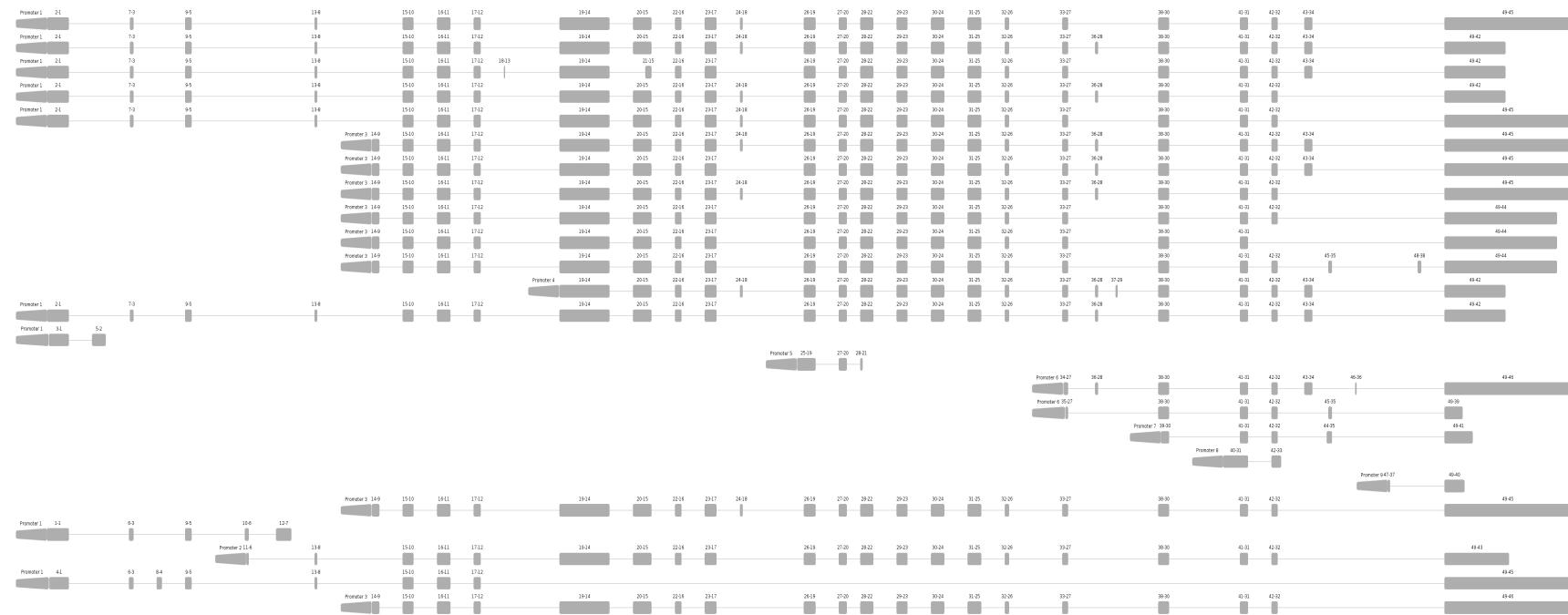


Figure 3.2: Classic schematic representation of promoter identification based on the TSS

A graphical representation of the locus from the human gene LPHN2 with 24 isoforms and 9 promoters. Promoters and exons are lined up to show where they are the same. Introns are not drawn to scale to allow for this depiction of aligned exons where they are the same. Two numbers describe each exon: the first number describes the relative start position and the second number describes the relative stop position of the exon. If two exons have the same start position on the genome, they both get the same number. An example, the exon 2-1 and 3-1 end in the same position, but 2-1 start at an early position on the genome before exon 3-1.

Prediction of Transcription Factor Binding Sites (TFBS)

TFBSs were predicted using the MatInspector [16] on the promoter sequences of the 1,044 transcripts. Prediction analyses were done on the current vertebrates library of 932 TFBS matrices in 188 matrix families. MatInspector uses a position weight matrix (PWM) to describe the binding site of each TF. Unlike the IUPAC consensus, a weight matrix pattern definition represents the complete nucleotide distribution for each position, which allows for the quantification of the similarity between the weight matrix and the potential binding site detected in the sequence.

Results from MatInspector provided summary match statistics details for each TFBS position in each of the promoters, description of the family of transcription factors that can potentially bind to those sites, the binding site matrix representing the family at that position, and matrix similarity scores that measures the similarity of the sequence at the position to the matching binding site matrix. TFBS matrices with similarity scores of at least 95% were considered as candidates that may be of potential transcriptional influence and therefore were selected for further analysis.

Identification of Isoform Specific Relevant Transcription Factors

Common tissues of expression pattern between isoforms and their potential binding factors were used to screen for isoform specific binding sites. For each tissue in which an isoform has some expression, transcription factors (TFs) that have expression in those same tissues are considered to have binding potential with the isoform. Thus, for an isoform with expression in N tissues, TFs with expression in each of the same N tissues are considered as having binding potential and hence, possible regulatory influence on the isoform (Figure 3.3 below).

TF_family	TF_matrixID	isoformID	adipose	adrenal	brain	breast	colon	heart	kidney	liver	lung	lymph	ovary	prostate	skeletal muscle	testes	thyroid	white blood	Sum of Expression
		ENST00000297146	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	2
CDXF	CDX1		0	1	0	0	1	0	0	0	0	0	0	0	0	1	0	0	3
CDXF	CDX2		0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
CEBP	CEBPD		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	16
GATA	GATA1		0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	2
GATA	GATA2		1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	14
GFI1	GFI1B		0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	2
NR2F	HNF4A		0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	3
ABDB	HOXC13		0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	2
HEAT	HSF2		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	16
MYT1	MYT1		0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
HOXF	NANOG		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
XBBF	RFX5		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	16
SORY	SOX6		1	0	1	1	1	1	1	0	1	1	0	1	1	1	0	0	12
ETSF	SPI1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	16

Figure 3.3: An example explaining the selection of isoform specific transcription factors

Using the isoform with ID ENST00000297146 and the predicted transcription factors that may potentially bind to the promoter of this isoform, we considered the tissues in which the isoform showed expression from the RNA-seq analysis. For analysis purposes, expression values were converted into binary numbers (0,1 – where 0 means there is no expression in the tissue and one mean expression in the tissue). From those results, isoform ENST00000297146 is expressed in 2 tissues (Adipose and thyroid – highlighted vertically in green). Also the tissues in which the transcription factors may be express in are considered. 6 TFs out of the 14 TFs predicted to potentially bind the promoter of this isoform actually have a chance of binding. The TFs, which we selected as possibly having regulatory influence on the isoform, was the 2 shaded rows, SOX6 and GATA2 because they meet the criteria for selection. Thus, they have expression in the same tissues as the isoform, have no expression in some of the tissues that the isoform have no quantified expression in, and the TFs are not ubiquitously expressed in all 16 tissues.

Other conditions used in the selection of the isoform specific transcription factor(s) are: 1) to select the transcription factors that has as limited expression in all tissues as the isoform under consideration. Thus, for instance, the TF should have as high number of no expression in the tissues that the isoform has no expression in. So in the example in figure 3.2, the 4 TFs with expression in all 16 tissues were not considered as functionally informative TFs and hence were not considered in the final annotation of the isoform's function. 2) TFs with at least N-1 expressions in the same tissues of expression as the isoform are considered as a potential regulator of the isoform. 3) In instances where no suitable TF meets the last two conditions, TFs with expression in all tissues are considered for that isoform

Functional Inference from Transcription Factors

Functional inferences were made based on the transcriptional influence that an identified TF have on their gene when they bind to their promoters. Thus for instance, the transcription factor SOX6 (Sex determining region Y – box 6) from the SORY (sex/testis determining and related HMG box factors) family of transcription factors is known to influence the genes they regulate in various developmental processes, specifically in neurogenesis and skeleton formation. We conclude therefore that an isoform that we determined to be potentially regulated by an SOX6 transcription factor may be involved in cellular processes that lead to neurogenesis and skeleton formation.

Functional information on transcription factors were obtained from mutagenesis studies and data from Entrez Gene [17] and UniProtKB/SwissProt [18].

Validation of Tools

A number of genomatix proprietary commercial tools are used to obtain promoter information as well as transcription factor binding module data. Although these tools have been widely used in a number of researches, little public knowledge is made available on the robustness of the techniques they employ. Therefore, we tested the tools to prove their predictive power and accuracy.

Using a sample of experimentally verified transcription factors and their target genes of expression, we assessed the accurate predictive power of the genomatix tools used for this research. Over the last decade or so, the MatInspector and other genomatix tools such as the gene2promoter and the ModelInspector have been cited and used in hundreds of bioinformatics research (search Google Scholar with keywords such as MatInspector, gene2promoter). However, none clearly reported the accurate predictive power of these tools, which we defined as the number of experimentally determined TFBS on a promoter that MatInspector and other genomatix tools can accurately predict. Therefore to validate the accuracy of our annotations we assessed the accurate predictive power as an indicator of how much reliable our final annotations are. We did this by selecting a list of transcription factors and their experimentally determined target genes, and then used the gene2promoter to retrieve the promoter sequences of those genes. We then run the MatInspector on those set of sequences to assess how accurately it can predict those TFBS on their promoter.

From the Broad Institutes' Molecular Signatures Database (MSigDB) [19] we obtained 1,116 genes and their *cis*-regulatory elements (transcription factors) shown in microarray data to regulate such genes (Supplementary Table 3.4). In all, 173

transcription factors were targeted to the 1,116 genes. From the MatInspector output, we determined the number of experimentally verified TFBS that MatInspector is able to accurately detect (results are provided below).

3.4.4 Results from Cellular Function Analysis

Tissue of Expression Data Results

The RNA sequencing data contained expression reads on 16 tissues. Its analysis revealed that 712 isoforms out of the 1044 isoforms under consideration had expression in at least one of the 16 tissues and 213 transcription factor matrices from 93 transcription factor families. The normalized RPKM on the 712 isoforms ranged from 0.01rpkm to 3001rpkm. Isoforms from the same gene showed differences in expression pattern across the 16 tissues, e.g. Figure 3.4.

In the 712 isoforms, isoform expressions were widespread across all tissues. However as evident by the U-shaped frequency distribution of Figure 3.5, GPCR isoform expression were present more on the extremities of the distribution. Thus, more of the isoforms are either expressed in very few select tissues or are expressed in a large number of multiple tissues. Implying that some GPCR isoforms are tissues-specific while other isoforms are ubiquitously expressed across multiple tissues. Normalized expression RPKM values on 712 isoforms are available in appendix A (Appendix_Table 3.1).

An expression frequency distribution figure was also generated for the expression data on transcription factors. Unlike the isoform expressions that showed some noticeable distribution pattern, the expression frequencies on the transcription

factors appeared to be well distributed. However, majority of the transcription factors are expressed in all 16 tissues types, Figure 3.6. Transcription factor expressions are available in the appendix A (Appendix_Table 3.2).

Using tissues of expression values, the 712 isoforms separated into 459 expression groups (see appendix A – Appendix_Table 3.3). Group with the highest number of isoforms was the group with expression in all 16 tissues. From this table, isoforms can be matched to specific select tissues of expression and vice versa. Without considering isoforms that are ubiquitously expressed in the 16 tissues, approximately 74% (529 isoforms) have co-expressed isoforms. That is, these 529 isoforms share expression pattern with other isoform(s).

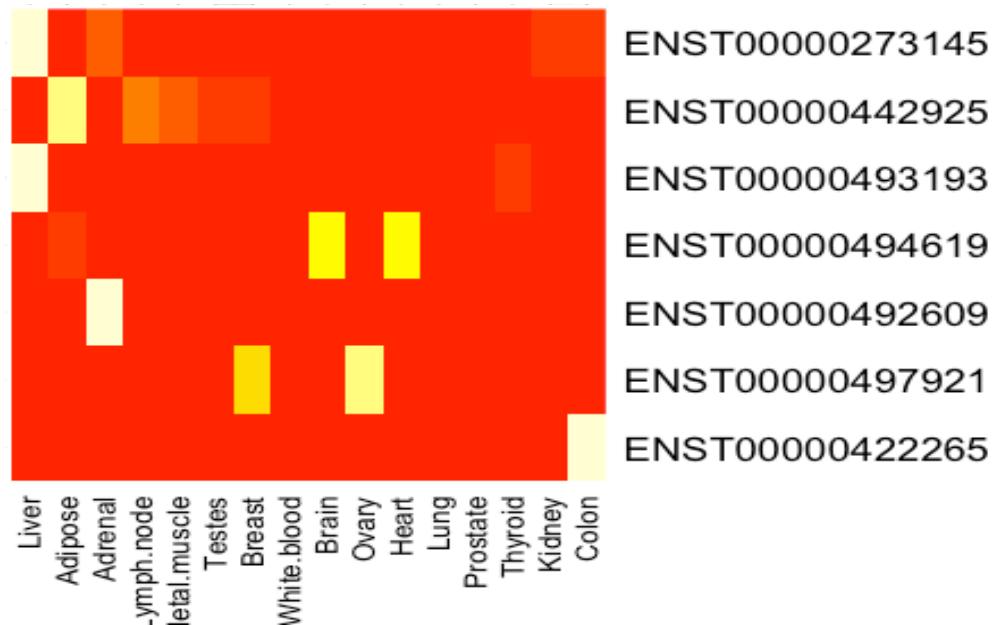


Figure 3.4: Heatmap

Shows varying expression pattern across isoforms belonging to the GPCR gene ACKR2 (Full heatmap of all isoforms in supplementary data). Expression intensity low expression to high expression (Red (low) → pale (high))

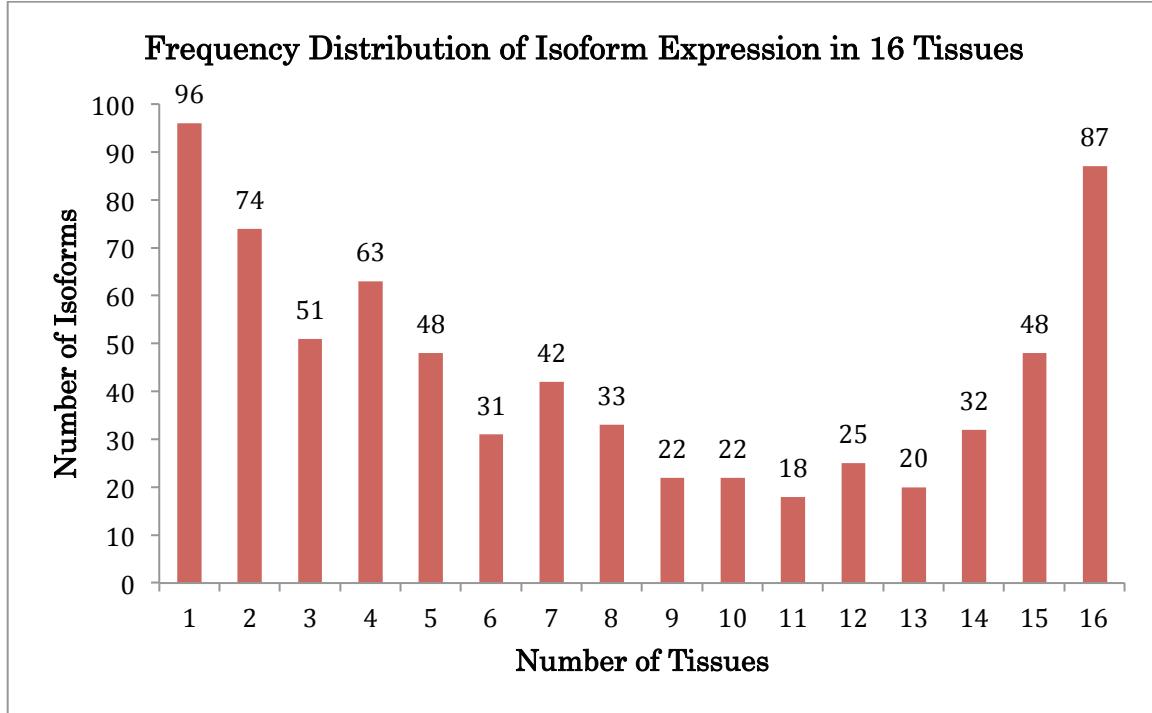


Figure 3.5: showing the frequency distribution of the number of isoforms and the number of tissues in which they are expressed.

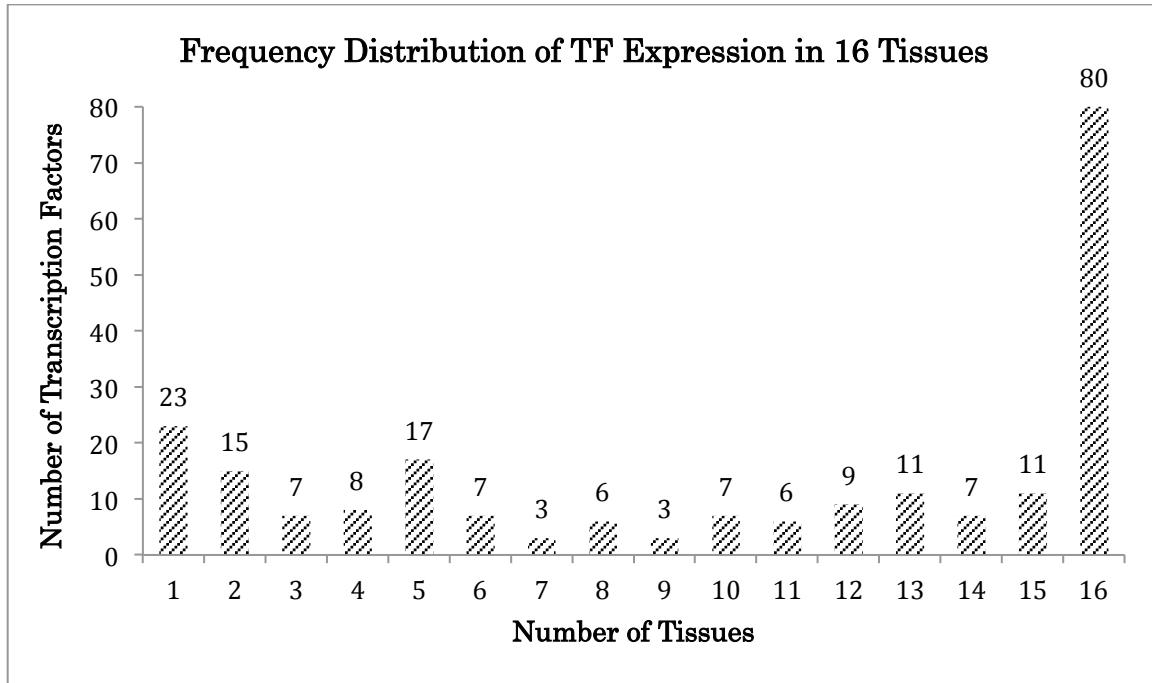


Figure 3.6: showing the frequency distribution of the number of transcription factors and the number of tissues in which they are expressed.

Promoter Analysis Results

Evaluations of promoter information on the study dataset revealed that the 1044 isoforms contain 561 promoters. 219 (39.04%) of them regulate multiple transcripts while the remaining 342 (60.96%) of the promoters goes to individual transcripts. Implying that majority of GPCR isoforms have their own promoters. Also, GPCR gene loci contain multiple TSS (promoters). Transcripts that share the same promoter share the same combinations of TFs that may potentially bind the transcripts (Supplementary Table 3.5).

Transcripts expression pattern analysis revealed differences in expression pattern for gene transcripts that share the same promoter. Although transcripts sharing the same promoters have the same potential TF binding sites, actual TF binding differs for each transcript due to their expression in different tissues. Thus, leading us to the conclusion that though transcripts may share the same TF binding sites, their pattern and timing of expression may differ hence resulting in different functional paths.

Isoform Annotation Results

Cellular level annotations were attained for 68% (712 out of the 1044) of isoforms in the dataset (Annotation Table – in supplementary data). We found that each isoform is functionally unique. They each have their own unique tissues of expression, and specific cellular pathways in which they may be involved. Evaluation of the annotations also reveals that single isoforms might be involved in multiple cellular processes depending on the cellular condition. The multiple transcription factors that may regulate the isoforms, which are shown from

mutagenesis studies and public data search to regulate very different cellular paths, as well as the multiple tissues in which the isoforms express support this assertion.

The remaining 332 isoforms also received molecular level annotation and promoter data but could not process them any further for cellular function due to the lack of expression data on them (Supplementary Table 3.4).

Summary evaluation of the number of functional pathways that isoforms may be involved ranges from 1 to 18 cellular processes as shown in Figure 3.7. It is evident that GPCR isoforms although shown to have multiple cellular pathways that they may be involved; many have few select pathways, indicated by the high number of isoforms with fewer than 6 cellular pathways as shown by the distribution in Figure 3.7.

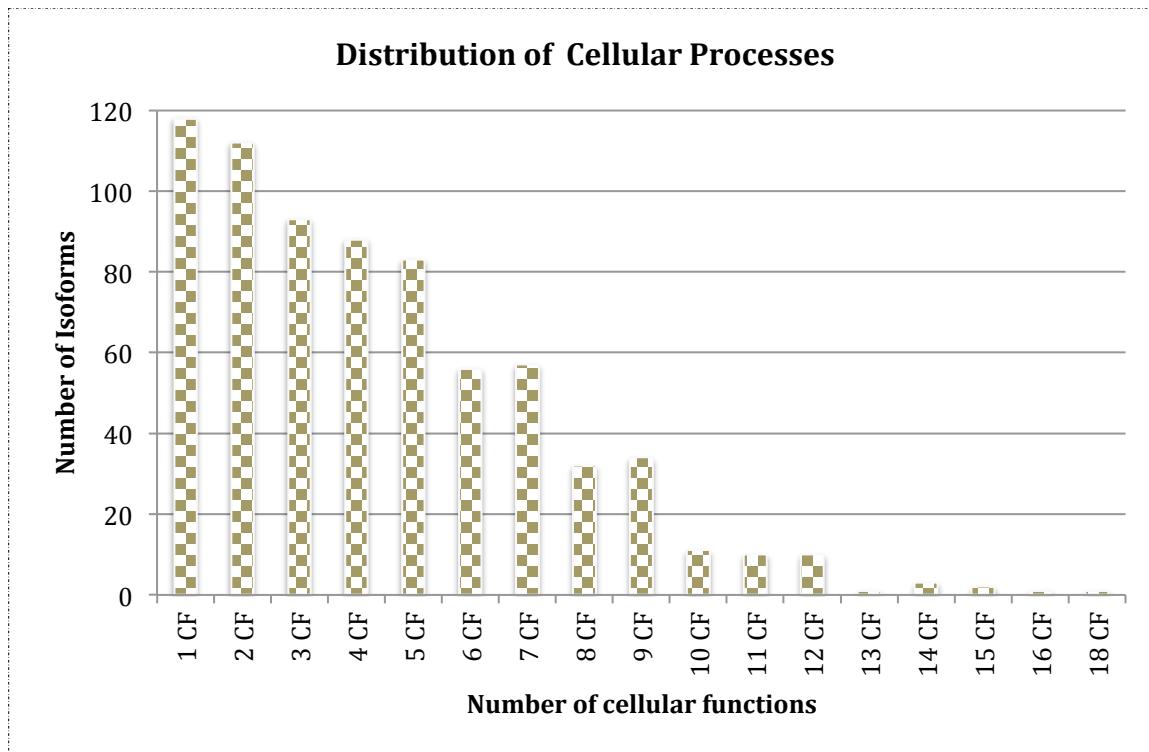


Figure 3.7: Distribution of number of cellular processes (CF) and the number of isoforms.

Additionally, the results show that majority of GPCR isoforms are regulated by only a few number of transcription factors, eight to be exact (E2F1 – E2F transcription factor 1, ERG - v-ets avian erythroblastosis virus E26 oncogene homolog, ETV1 – ets (E Twenty-six) variant 1, HIC1 – hypermethylated in cancer 1, PLAG1 – pleiomorphic adenoma gene 1, SALL2 - spalt-like transcription factor 2, ZKSCAN3 - zinc finger with KRAB and SCAN domains 3, and ZNF300 - zinc finger protein 300) (Figure 3.8) erythroblast transformation-specific (ETS) family of transcription factors. This shows that in the broader spectrum, majority of GPCR isoforms are involved in cellular growth, survival, angiogenesis, migration, proliferation, and differentiation pathways. Specifically, these transcription factors have been reported to regulate genes that are involved in biological processes that include: E2f1 – G1/S transition of mitotic cell cycle, Notch signaling pathways, and in apoptotic process; ETV1 - mechanosensory behavior, muscle organ development, and peripheral nervous system neuron development; HIC1 - intrinsic apoptotic signaling pathway in response to DNA damage, and negative regulation of Wnt signaling pathway; PLAG1 – multicellular organism growth, and gland morphogenesis; and ZKSCAN3 – lysosome organization, autophagy, and cellular senescence. These observations go to endorse the importance of GPCRs in many survival-dependent biological processes.

Furthermore, it is observed that an equally large number of isoforms also have some specific regulators that allow their function to include slightly different biological process than these described. Often, transcription factors do not work alone but in combination with other factors to initiate transcription. Assessments of

the combinations of transcription factors predicted to regulate isoforms are shown in Figure 3.9 and also available in Appendix_Table 3.4. It can be observed from the distribution as well as from the Appendix_Table 3.4 that many (85%) of the TF combinations regulate one isoform. Thus, further lending validation to the fact that GPCR isoforms are functionally different from other isoforms especially those from the same gene loci.

Predictive Power of MatInspector

MatInspector accurately predicted all (100%) of the TF binding sites shown in microarray analysis to regulate specific genes in 37% of the assessment dataset. 14% of the assessment dataset were predicted at between 90-99% accuracy. 28% at between 80-89%. All together, the ability of MatInspector to predict an actual TF binding site is calculated to be approximately 80%. Meaning that the predictions generated from the MatInspector output is about 80% reliable.

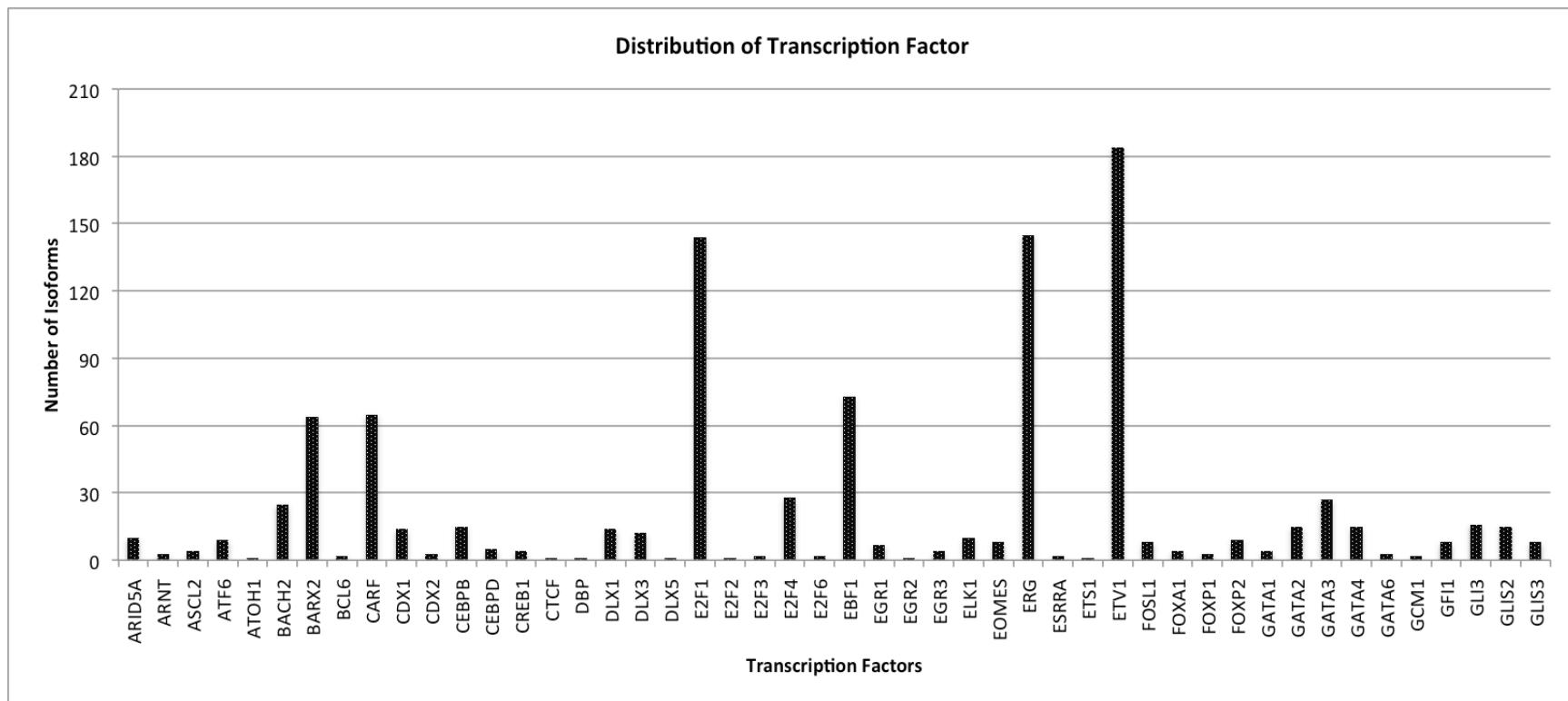


Figure 3.8 a: Part a of the distribution of transcription factors and the number of GPCR isoforms they regulate

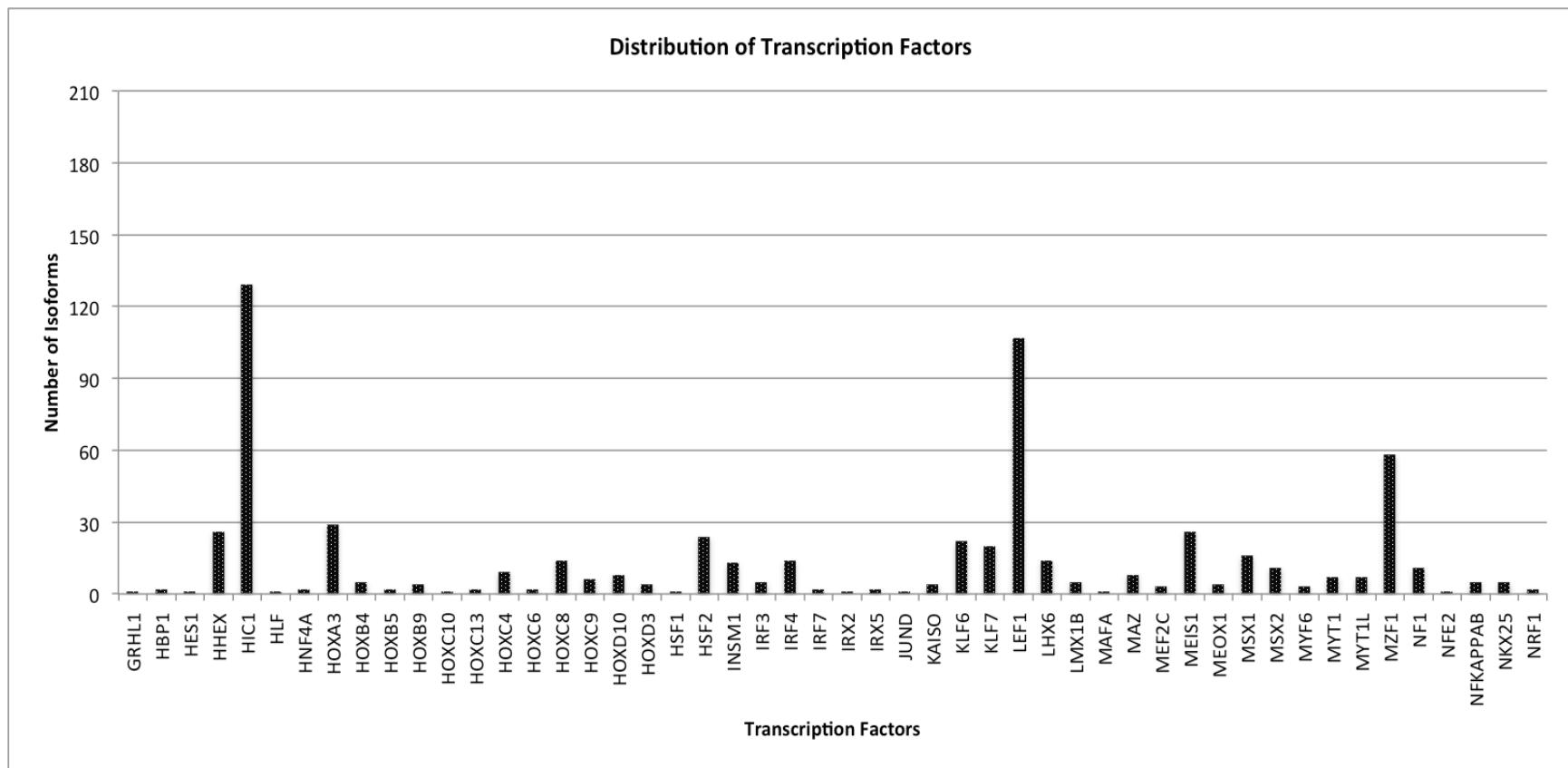


Figure 3.8 b: Part b of the distribution of transcription factors and the number of GPCR isoforms they regulate

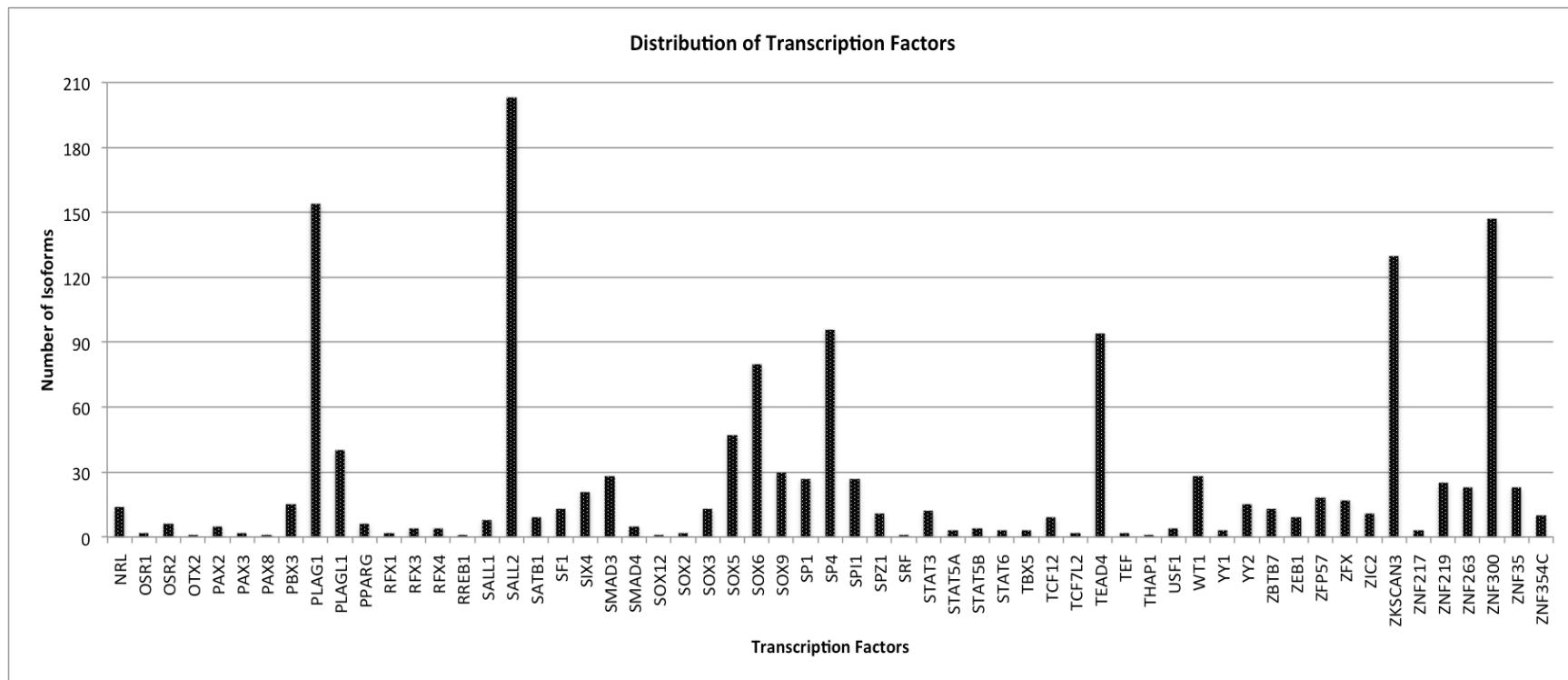


Figure 3.8 c: Part c of the distribution of transcription factors and the number of GPCR isoforms they regulate

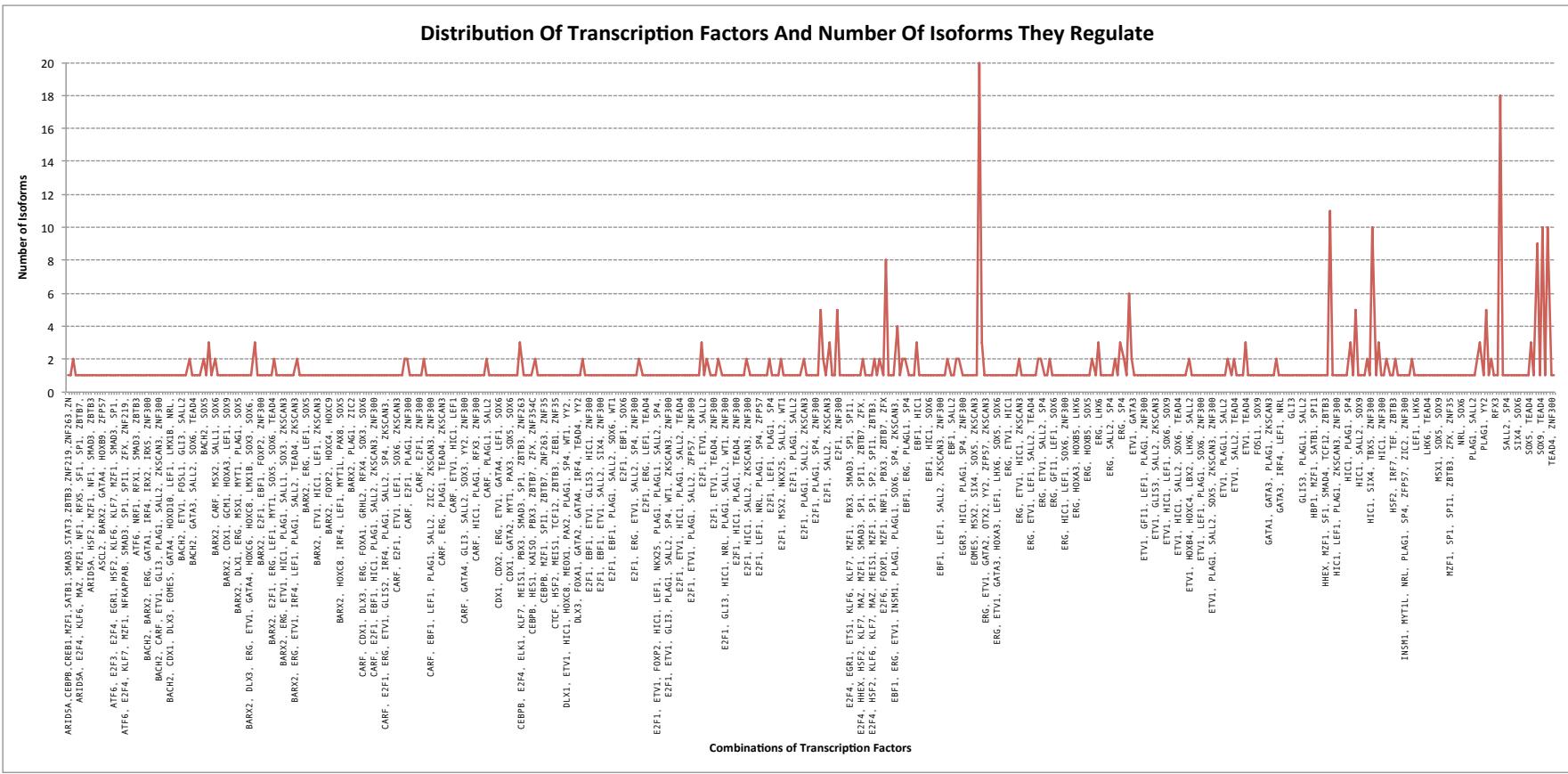


Figure 3.9: Distribution of combinations of transcription factors and the number of isoforms they regulate.

The figure shows the distribution of the various predicted combinations of transcription factors and the number of isoforms they regulate.

3.5 Integration of Molecular and Cellular Function

From the molecular level annotations obtained from motif-based methods discussed in section 3.3 above, family and subfamily information for isoforms were attained. From family information, the type of possible interactions that the gene participates in is identified. This information though valuable, did not provide distinct separation as to the function of independent isoforms from its “sister” isoforms.

Cellular level annotations, which were acquired through the processes discussed in section 3.4 above, provided pathway information for each isoform. They predicted the specific cellular conditions in which an isoform may express and gave the potential pathway in which that isoform may be involved.

By combining the two levels of annotation (molecular and cellular), the functions of isoforms become definitive. In this work, 68% of the dataset received this higher level of annotation because cellular functions were uncovered for them. For the remaining 32%, only molecular level annotations were possible. Even with only molecular level annotations, those 32% provided more information on the isoforms than is currently available. Overall this framework has added the much needed information on GPCR isoforms.

3.6 Discussion

Results from this research strongly suggest that each alternatively spliced isoform of GPCR is functionally distinct from other isoforms. This assertion is supported by the findings that (a) the cellular paths of each isoform are distinctive,

(b) isoforms have different tissues of expression, and (c) they each have unique sets of transcription factors that may regulate them.

The main task of this chapter was to identify the molecular and cellular level differences between alternatively spliced isoforms of GPCRs. Results from evaluating molecular level differences between isoforms, did not offer unique functional details between isoforms as would have been anticipated. Rather, isoforms were grouped under various GPCR families and subfamilies, with additional information on other molecular functional influences as suggested by the present of specific motifs. The inability of molecular level annotation to offer functional distinction between isoforms may have been due to the absence of detailed functional motifs. It may also be due to the very low residue differences that exist between alternatively spliced isoforms from the same gene loci as discussed in chapter 2. The initial thought had been that alternatively spliced isoforms of genes would have enough protein sequence similarity to its “sister” isoforms and yet enough different attributes to distinguish its function from others by means of functional motifs. Conversely, this was not the observed relationship.

Cellular level functions on the other hand, offered enough information to allow for the identification of distinct functional differences between isoforms. First, identification of isoform specific promoter information began the separation of isoforms. The 1044 isoforms were separated into 561 promoter groups from the 58 GPCR families and subfamilies of the molecular annotation. Even at this point, there still remain isoforms that shared common promoters, implying that they shared the same potential transcription factor binding sites and hence may not be functionally different from another isoform sharing the same promoter. It became necessary to identify actual binding sites for each isoform to further separate them.

Based on the understanding that transcription factors are condition specific genes that are only active when the cellular condition is conducive, we considered the use of tissues of expression to identify the isoform specific transcription factors from all available TFs for the promoter. Thus leading to the obtaining of tissues of expression information from RNA-seq data. Tissue of expression data could have come from microarray data, but RNA-seq data was chosen because one, it contains all the information on alternatively spliced variants. Secondly, it can more accurately identify differentially expressed genes, especially those that tend to express at a relatively low levels or extremely high levels. Also in genes where changes in expression may be small but statistically significant, such as those often observed with transcription factors, RNA-seq is the best technique to utilize. Additionally, data from RNA-seq was preferred because it can be replicated due to its digital nature and is less prone to random errors (experimental environment factors).

712 out of the 1044 isoforms offered expression read values in at least one tissue. Those isoforms therefore received additional evaluation beyond the promoter analysis. For those isoforms for which no expression reads in any of the 16 tissues were found, we speculated that they might not be expressed in any of the tissues under consideration but in other tissues that were not considered for this research. Another notable observation from the expression reads were the varying intensities of reads. After ruling out the possibility of biases in the measurements, we speculate that low expression reads may be to allow for the binding of factors to prevent the factor from binding to another promoter. Thus, the isoform functions as a repressor by competing to bind rather than actually binding to activate any downstream effector. Undoubtedly, this claim needs further studies.

Now, the inclusion of tissues of expression data allowed us to further distinguish one functional isoforms from another. At this level of annotation, all 712 isoforms received distinct cellular paths from their sister isoforms although they may have some processes in common. Additional evidence also exists to show that single isoforms may be involved in multiple cellular processes depending on the cellular condition.

The question now remains as to why this framework was employed to achieve the purpose of this study instead of the conventional computational approach to functional annotation. Over the years, various computational methods including pairwise similarity comparison, and functional motif based, have been used successfully in spite of its inherent limitations to annotate many proteins. However, when it comes to proteins that have originated as a result of alternative splicing of the pre-mRNA, these techniques fall short of adequately achieving this purpose. The reasons include that one, homology based methods are built on the premise of evolutionary relationship between genes, but spliced isoforms are not homologous. Secondly, spliced forms of many genes do not always splice out the major functional exons as it is in the case of the calcitonin/CGRP gene, but rather the splicing is subtle (changes in the un-translated regions and in signal peptide coding exons) [20]. These subtle changes may lead to the expression of a specific splice form of the protein in a different tissue (e.g. GRK proteins), or at different biological conditions or cellular state. Exploration of the genomic landscape of GPCRs as discussed in chapter 2 have shown that ~69% of the alternative splicing that occurred in GPCR are in the UTRs (splicing that does not change the resulting amino acid content), hence the use of protein level functional motifs in annotating them are less effective.

The tools used for our analysis were selected due to specific advantages that they possessed. For instance, the MatInspector was preferred over other TFBS prediction tools such AliBaba2 [21], and P-match [22], because although they employ the use of position weight matrices (PWM) as does the MatInspector, they do not fully harness and embrace the power that it offers. Secondly, both are rarely maintained, and are dependent on outdated PWM libraries. MatInspector on the other hand makes full use of the position weight matrices, which allows for the quantification of the similarity between the weight matrix and the potential binding site detected in the sequence [16] as well as maintain an active library of current PWM. It is worth noting that transcription factor binding sites only has the potential to bind their corresponding proteins and can occur anywhere in the genome. However, it is the context (sites within regulatory regions, tissue type etc.) that differentiates a functional binding site affecting gene regulation from a mere physical binding site [23, 24]. Computationally, predicted binding sites regardless of their location, can only infer the binding potential. It is through a wet-lab experiment with defined settings that claims of functionality can be made. Therefore, to determine that the resulting transcription factor binding sites are actual functional site, we used only sequences within regulatory regions of our transcripts in finding TF binding sites. Also, all reported sites were verified with reported experimental research on these sites.

Ultimately, no computational work can trump a wet-lab experiment, but serves as a good starting point. All inferences made from our analysis points to possible association rather than an absolute proof of association. The quality of results depends on the quality of the sequence data; therefore, advances in better sequence mapping techniques will increase the precision of the prediction.

3.7 Conclusion

Using the framework propose in this dissertation, we have predicted the various cellular and molecular functional processes in which GPCR isoforms are involved in. Based on the results from implementing the procedures outlined in the framework, we conclude that majority of GPCR isoform expressions are tissue-specific. We identified that each variant of GPCR isoform has unique cellular process they mediate. Furthermore, GPCR isoforms contain identifiable distinct features that make them function differently from other isoforms, such as multiple *cis* regulatory binding domains. This research work offers the promise of isoform-specific ligand binding (drug targeting), as well as, a good starting point for researchers interested in pathway analysis and alternative spliced isoform signaling. It must be noted however that results from this research comes from computational prediction validated by highly significant statistical power. Yet, experimental validation of our findings is the ultimate.

REFERENCES

1. Sharman JL, Benson HE, Pawson AJ, Lukito V, Mpamhanga CP, Bombail V, Davenport AP, Peters JA, Spedding M, Harmar AJ *et al*: **IUPHAR-DB: updated database content and new features.** *Nucleic Acids Res* 2013, **41**(D1):D1083-D1088.
2. Sigrist CJA, Cerutti L, de Castro E, Langendijk-Genevaux PS, Bulliard V, Bairoch A, Hulo N: **PROSITE, a protein domain database for functional characterization and annotation.** *Nucleic Acids Res* 2010, **38**:D161-D166.
3. Attwood TK, Coletta A, Muirhead G, Pavlopoulou A, Philippou PB, Popov I, Roma-Mateo C, Theodosiou A, Mitchell AL: **The PRINTS database: a fine-grained protein sequence annotation and analysis resource--its status in 2012.** *Database (Oxford)* 2012, **2012**:bas019.
4. Pietrokovski S, Henikoff JG, Henikoff S: **The Blocks database--a system for protein classification.** *Nucleic Acids Res* 1996, **24**(1):197-200.
5. Hunter S, Jones P, Mitchell A, Apweiler R, Attwood TK, Bateman A, Bernard T, Binns D, Bork P, Burge S *et al*: **InterPro in 2011: new developments in the family and domain prediction database (vol 40, pg D306, 2011).** *Nucleic Acids Res* 2012, **40**(10):4725-4725.
6. Bairoch A, Apweiler R: **The SWISS-PROT protein sequence data bank and its new supplement TREMBL.** *Nucleic Acids Res* 1996, **24**(1):21-25.
7. Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J *et al*: **The Pfam protein families database.** *Nucleic Acids Res* 2012, **40**(D1):D290-D301.
8. Bru C, Courcelle E, Carrere S, Beausse Y, Dalmar S, Kahn D: **The ProDom database of protein domain families: more emphasis on 3D.** *Nucleic Acids Res* 2005, **33**(Database issue):D212-215.
9. Wallace JC, Henikoff S: **PATMAT: a searching and extraction program for sequence, pattern and block queries and databases.** *Computer applications in the biosciences : CABIOS* 1992, **8**(3):249-254.

10. Edgar RC: **MUSCLE: a multiple sequence alignment method with reduced time and space complexity.** *Bmc Bioinformatics* 2004, **5**:1-19.
11. Pruitt KD, Brown GR, Hiatt SM, Thibaud-Nissen F, Astashyn A, Ermolaeva O, Farrell CM, Hart J, Landrum MJ, McGarvey KM *et al*: **RefSeq: an update on mammalian reference sequences.** *Nucleic Acids Res* 2014, **42**(Database issue):D756-763.
12. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT *et al*: **Gene ontology: tool for the unification of biology. The Gene Ontology Consortium.** *Nat Genet* 2000, **25**(1):25-29.
13. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L: **Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks.** *Nat Protoc* 2012, **7**(3):562-578.
14. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B: **Mapping and quantifying mammalian transcriptomes by RNA-Seq.** *Nat Methods* 2008, **5**(7):621-628.
15. Bullard JH, Purdom E, Hansen KD, Dudoit S: **Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments.** *Bmc Bioinformatics* 2010, **11**:94.
16. Cartharius K, Frech K, Grote K, Klocke B, Haltmeier M, Klingenhoff A, Frisch M, Bayerlein M, Werner T: **MatInspector and beyond: promoter analysis based on transcription factor binding sites.** *Bioinformatics* 2005, **21**(13):2933-2942.
17. Brown GR, Hem V, Katz KS, Ovetsky M, Wallin C, Ermolaeva O, Tolstoy I, Tatusova T, Pruitt KD, Maglott DR *et al*: **Gene: a gene-centered information resource at NCBI.** *Nucleic Acids Res* 2014.
18. **Ongoing and future developments at the Universal Protein Resource.** *Nucleic Acids Res* 2011, **39**(Database issue):D214-219.
19. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES *et al*: **Gene set enrichment analysis: a knowledge-based approach for**

- interpreting genome-wide expression profiles.** *Proceedings of the National Academy of Sciences of the United States of America* 2005, **102**(43):15545-15550.
20. Hammond DA, Olman V, Xu Y: **Functional understanding of the diverse exon-intron structures of human GPCR genes.** *J Bioinform Comput Biol* 2014, **12**(1):1350019.
 21. Grabe N: **AliBaba2: context specific identification of transcription factor binding sites.** *In Silico Biol* 2002, **2**(1):S1-15.
 22. Chekmenev DS, Haid C, Kel AE: **P-Match: transcription factor binding site search by combining patterns and weight matrices.** *Nucleic Acids Res* 2005, **33**(Web Server issue):W432-437.
 23. Kodadek T: **Mechanistic parallels between DNA replication, recombination and transcription.** *Trends Biochem Sci* 1998, **23**(2):79-83.
 24. Elkon R, Linhart C, Sharan R, Shamir R, Shiloh Y: **Genome-wide in silico identification of transcriptional regulators controlling the cell cycle in human cells.** *Genome Res* 2003, **13**(5):773-780.

CHAPTER 4

THE *gpcrTAG* DATABASE

4.1 Introduction

Alternative splicing (AS) events are very important cellular control mechanism because they increase protein function diversity, and have been suggested to be the major contributor to the functional complexity of mammalian and other genomes. Examples of the effects of AS include variations of the function of the primary full-length protein, its antagonist function or even not directly related to its function at all. With the accelerating progress in proteomics, RNA-sequencing, and other high-throughput sequencing technologies, the door to wealth of biological information is widely open. Particularly in isoform function annotation. Harnessing and using the information from these new resources can improve our understanding of protein function, especially AS isoforms.

Publicly available databases on gene annotations, i.e. UniProt/SwissProt [1] database, EMBL [2], GenBank [3], and Gene Ontology (GO) [4] include as much information on alternative splicing of genes as is available, allowing for easy retrieval by researchers. GPCRs do not have enough annotation on spliced isoforms, especially on how each isoform functions differently from its “sister” isoform (isoforms from the same gene loci). Even GPCR-specific databases such as IUPHAR-DB [5], GPCRDB [6], GPCR-RD [7] all lack annotations on isoforms. We have

therefore designed the *gpcrTAG* database to make the annotations discussed in the earlier chapters of this dissertation publicly and easily accessible to all.

In this database, G Protein-Coupled Receptor Transcriptional Annotation of Genes (*gpcrTAG*) we collected knowledge on the current data on alternative spliced isoforms of GPCRs from external data sources, and then applied the computational framework discussed in chapter 3 to comprehensively predict the cellular functions of 1044 select GPCR isoforms. This web interface of the research was created to facilitate easy public access to the research result content, as well as to facilitate easy information inference from the predicted functions of GPCR isoforms.

4.2 Database Content

The *gpcrTAG* database contains two types of data; one is the data imported from external sources such as SwissProt, EMBL, and RefSeq (Gene ID, transcript/isoform ID, gene name and description etc.); and secondly, data that were as results of our analysis, such as cellular function, tissues of expression, and transcription factor data. Table 4.1 summarizes the statistics of the data contained in *gpcrTAG*. Tissues of expression data were obtained from the analysis of RNA-seq data from the Human BodyMap project 2.0 as discussed in chapter 3. Transcription factor data came from the detailed processing of predictions made by MatInspector. Function annotation information is from the reported annotation discussed in chapter 3 as well.

Table 4.1: Statistics on the December 2014 release of *gpcrTAG* database

Statistics on the December 2014 release of <i>gpcrTAG</i> database	
GPCR data contained	332 genes
	1044 isoforms
Cellular function annotation	712 isoforms
Transcription factor	712 isoforms verified
	332 isoforms unverified
Expression data	16 tissues
Isoform tissues of expression	712 isoforms
Transcription factor tissues of expression	213 TF matrices
	93 TF families

4.3 Database Design

We designed a relational database to store and disseminate the records of data from our analysis. One advantage of using a relational database design for *gpcrTAG* is its ability to ensure data integrity and accuracy. The database is organized into five tables, namely, the fundamental table, promoter table, transcription factor table, tissues of expression table, and function table. From these tables we are able to offer functional annotation, transcription factor, and tissues of expression data to users. Relationships that exist between tables are illustrated in Figure 4.1.

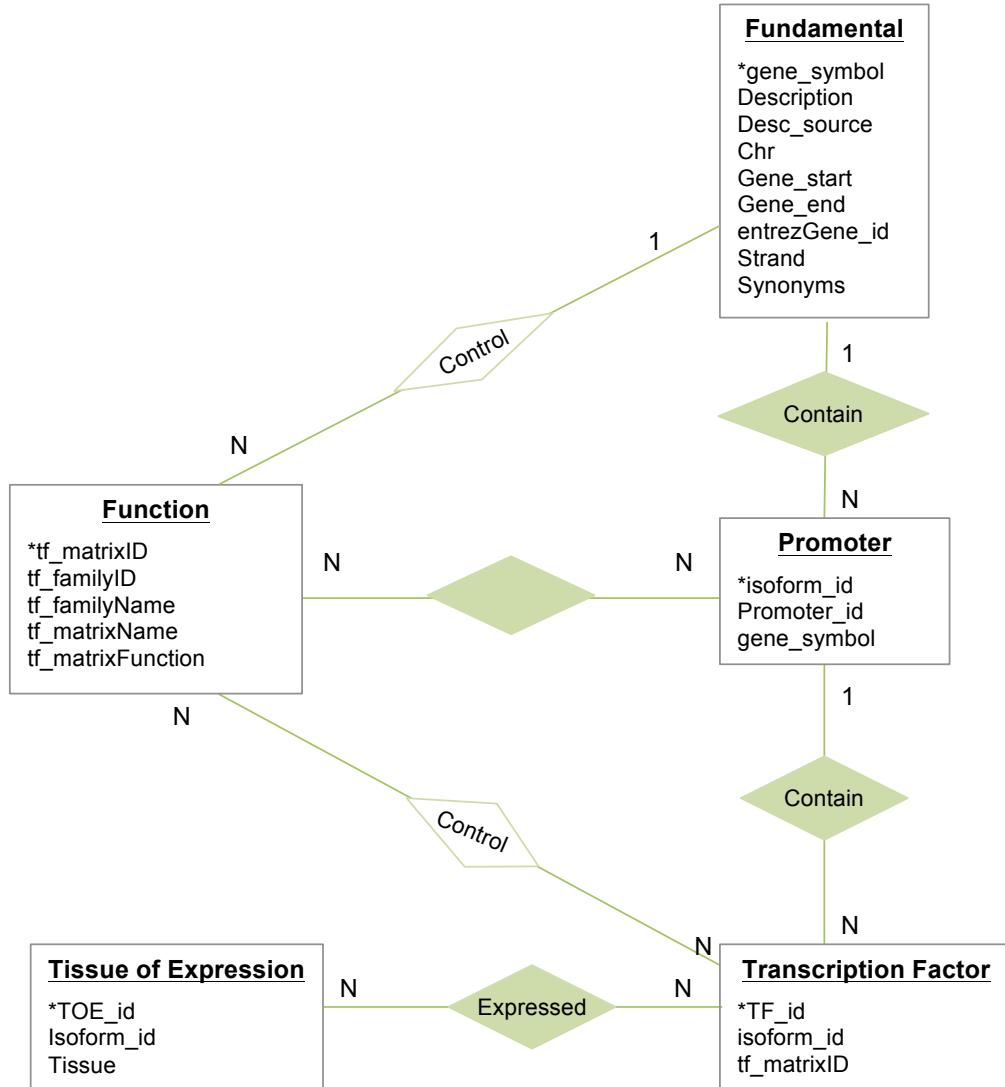


Figure 4.1: Schematic representation of the relationship between tables in the *gpcrTAG* database. The relationships that exist between tables are shown in the one-to-many (1-N) or many-to-many (N-N) relationship. Full explanations of table attributes are given in the supplementary data.

*Table Key

4.4 Database Implementation

The system, *gpcrTAG* database presents the data in a simplistic format. Primarily, this relational database focuses on integrating data between GPCRs and their transcription factor, expression pattern across 16 tissues, and cellular function. User can browse, retrieve, and query the database. To the best of our knowledge this is the first publicly available database on cellular function annotation, transcriptional regulation, and tissue of expression on human G protein-coupled receptor isoform.

The database is represented online using a PHP server side scripting language, HTML5, CSS3, java scripts and jQuery libraries. By harnessing the responsive nature of CSS3 gpcrTAG database can be viewed on all media types (minimum width of 150px and maximum width of 600px). MySQL relational database is used to store all data, which runs in a Linux server on 16 Intel Xeon CPU with 2.4 GHz and 16GB of memory. The database can be access at the following location <http://www.gpcrtag.com>

4.5 Accessing Database

At the moment, we have implemented and will continue to develop three main basic information system tools browse, retrieve and query to provide access to the database content. Browsing, the data is organized in an alphabetical order and the easiest way to access the database content is via the browse page of ordered list based on HGNC gene symbol. For records on specific gene, users can access gene description, available isoform identifications, and functional annotations for all associated isoforms.

Retrieval, most of the data can be retrieved using the links provided in the help page to download content. It must be noted that not all of the database content is downloadable at this time.

Query, a basic query of the database via search page with user defined HGNC gene symbol, ensEMBL transcript identifier, and ENTREZ identifier is the most effective way to obtain isoform specific information on function, tissues of expression, and transcription factors (Figure 4.1). Users must choose the appropriate type of identifier they wish to use as well as input the identifier in the provided space to obtain desired output from the database. The database is setup in this way to limit phishing and embedding of html entities into the database.

Search Database

Function Search

- ✓ Choose... e.g. 3350, DARC, ENST00000409489
- Gene ID
- HGNC Symbol
- ensEMBL Transcript

Tissues of Expression

Choose... e.g. 113, CCR5, ENST00000399036

Transcription factors

Choose... e.g. S1PR5, ENST00000276393

Buttons: Search, Clear All

Figure 4.2: Screenshot of query page of *gpcrTAG*

From user-defined query as illustrated in Figure 4.2 above, Figure 4.3 shows an example of expected output from such search.

gpcrTAG

[Back to Search](#)

Your search for function for ENST00000409489 returned 7 results

Gene Symbol : ADCYAP1R1
 Transcript ID : ENST00000409489
 Functional Annotation : involved in the embryonic formation of blood or blood cells

Gene Symbol : ADCYAP1R1
 Transcript ID : ENST00000409489
 Functional Annotation : activation of IL4 promoter

Gene Symbol : ADCYAP1R1
 Transcript ID : ENST00000409489
 Functional Annotation : immune activation by HTLV-1 LTR

Gene Symbol : ADCYAP1R1
 Transcript ID : ENST00000409489
 Functional Annotation : skin immunity

Gene Symbol : ADCYAP1R1
 Transcript ID : ENST00000409489
 Functional Annotation : Constitutive expression of genes that play significant role in physiologic lymphocyte functions at mucosal site

Gene Symbol : ADCYAP1R1
 Transcript ID : ENST00000409489
 Functional Annotation : facilitate the recruitment of coactivators

Gene Symbol : ADCYAP1R1
 Transcript ID : ENST00000409489
 Functional Annotation : involved in the mediation of the transcriptional activation of the human BAX promoter which accelerate apoptosis

Figure 4.3: Showing sample functional annotations with ensEMBL ID ENST00000409489

Additionally, *gpcrTAG* supports user information inference based on the derived functional data supplied by the database. Such as querying for isoforms regulated by specific sets of cis regulator(s), and isoforms involved in certain pathways and are expressed in select tissues. Thus, with a few lines of MySQL code, various conditions can be set to obtain selective information. Here is an example of a MySQL code to select only isoforms that has expression in either brain or heart and its regulation involves the SOX5 transcription factor.

```

SELECT DISTINCT

`tbl_tranFactor`.`isoform_id`, `tbl_tranFactor`.`tf_matrixID`, `tbl_tissueOfexp`.`Tissue`

FROM `tbl_tranFactor` 

JOIN `tbl_tissueOfexp` ON `tbl_tranFactor`.`isoform_id` = `tbl_tissueOfexp`.`isoform_id` 

WHERE (`tbl_tranFactor`.`tf_matrixID` = 'SOX5' AND `tbl_tissueOfexp`.`Tissue` = 'Brain')

OR (`tbl_tranFactor`.`tf_matrixID` = 'SOX5' AND `tbl_tissueOfexp`.`Tissue` = 'Heart');

```

The results query returns the isoform with ID ENST00000283303. Thus, from all 1044 GPCR isoforms currently in the database only one met the criteria set forth by the query. A web interface of such query will be available soon.

4.6 Discussion

gpcrTAG is intended to offer a convenient and user-friendly annotation on alternatively spliced isoforms of GPCRs to users via the Internet. This database provides a service that will be useful to researchers interested in drug targeting and design, as well as those involved in pathway analysis. For drug designers, it may offer the opportunity for isoform specific drug targeting and hence improving drug efficacy and potentially minimizing drug side effects as a much effective target/ligand can be designed. An exclusive feature of *gpcrTAG* is its collection of isoform specific tissues of expression data, which was obtained from the RNA-sequencing data from the Human BodyMap Project 2.0, and the isoform specific transcription factor information.

Some notable features of *gpcrTAG* include (i) the annotations presented in *gpcrTAG* are not inferred from mere similarity of sequence or structure of a characterized protein, but is annotated based on its own merits of expression and

transcription factor binding influences. (ii) The data is presented in an easy-to-manage browse manner, so that users may not need to have any prior knowledge on specific GPCR to gain information. (iii) The only GPCR database by far to provide comprehensive data on *cis* regulatory elements of GPCRs and expression data.

Although *gpcrTAG* provide a unique service to the GPCR research community, we recognize the need to make the database more open to allow for users to obtain automated functional annotations beyond the 1044 isoforms provided. This is a future endeavor that will be undertaken. This current version of *gpcrTAG* however, is focused on making the annotations of the 1044 isoform publicly available.

4.7 Future Improvements

We hope to be able to soon extend the data content to include more human GPCR isoforms than is currently available in the database. Also, we would include links to GPCR structural features such as PDB structures and include secondary structure feature information. Much later on, we would also like to add GPCR data for other organisms, and include links to alternative splicing resources.

REFERENCES

1. Magrane M, Consortium U: **UniProt Knowledgebase: a hub of integrated protein data.** *Database (Oxford)* 2011, **2011**:bar009.
2. Flicek P, Amode MR, Barrell D, Beal K, Billis K, Brent S, Carvalho-Silva D, Clapham P, Coates G, Fitzgerald S *et al*: **Ensembl 2014.** *Nucleic Acids Res* 2014, **42**(Database issue):D749-755.
3. Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW: **GenBank.** *Nucleic Acids Res* 2013, **41**(Database issue):D36-42.
4. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT *et al*: **Gene ontology: tool for the unification of biology. The Gene Ontology Consortium.** *Nat Genet* 2000, **25**(1):25-29.
5. Sharman JL, Benson HE, Pawson AJ, Lukito V, Mpamhangwa CP, Bombail V, Davenport AP, Peters JA, Spedding M, Harmar AJ *et al*: **IUPHAR-DB: updated database content and new features.** *Nucleic Acids Res* 2013, **41**(D1):D1083-D1088.
6. Vroeling B, Sanders M, Baakman C, Borrmann A, Verhoeven S, Klomp J, Oliveira L, de Vlieg J, Vriend G: **GPCRDB: information system for G protein-coupled receptors.** *Nucleic Acids Res* 2011, **39**:D309-D319.
7. Zhang J, Zhang Y: **GPCRRD: G protein-coupled receptor spatial restraint database for 3D structure modeling and function annotation.** *Bioinformatics* 2010, **26**(23):3004-3005.

CHAPTER 5

PERSPECTIVE AND CONCLUSIONS

5.1 Summary

The research work presented in this dissertation has helped to advance the understanding of the functional diversity that exist between alternatively spliced isoforms of G protein-coupled receptors. Beginning with the general exon-intron structures of GPCRs and how they influence the alternative splicing events of the gene. The genomic structure of GPCRs were shown to be quite complex, and that the diverse exon-intron structures were so ordered to facilitate alternative splicing of the gene for functional purposes. Also, it was shown that majority of the splicing events that occur in GPCRs occur in un-translated regions of the gene, which results in minimal or no change in the resulting amino acid content of the resulting protein product. These findings are of immerse importance because it will shape the way alternatively spliced isoforms of GPCRs are eventually annotated. The work on motif finding for molecular level annotations showed that due to the localization of splicing events on GPCR genes as reported in the first part of this dissertation, molecular level annotation alone would not be sufficient to annotate the functional diversity of GPCR isoforms. However, there still exist some useful information that can be derived from molecular annotation via motif identification for instance. Results from the work done in isoform expression found expression diversity among GPCR isoforms, hinting to the possible functional diversity to be expected among the isoforms. The cellular level annotation work found that GPCR isoforms each have at

least one unique biological process that it is involved in. Further emphasizing the relevance and need of alternative splicing events in the human and other complex organisms. Lastly, results of all the discussed works are publicly available via the web through the website www.gpcrtag.com

I have taken a set of existing computational methods, which independently work to attain different objectives and integrated them into a framework that can successfully be used to i) functionally distinguish one alternatively spliced isoform from another, ii) predicted a set of isoforms that may be involved in a specific pathway, as well as iii) predicted isoform specific transcription factors that may be involved in the regulation of the isoform.

This work has been exciting and challenging at the same time. The most challenging part had been the selection of appropriate transcription factor that may regulate an isoform relative to other transcription factors that also have the potential to bind. The exciting part is that after all the kinks have been worked out; the framework presented here can be implemented in other gene families. Another advantage of the framework is that, it is versatile. Thus, it is not restricted to any one specific algorithm.

It must be noted that the reported annotations are limited by publicly available data on expression and transcription factor influences. However, as more data become available, the framework presented here can provide detailed cellular functions on human GPCR transcripts. Ultimately, no computational work can trump that of experiments; we therefore need additional experimental work to adequately validate the reported results. We look forward to seeing biologist and other researchers taking on the task of furthering the work presented in this dissertation.

5.2 Future Directions

5.2.1 Submitting Annotations to Public Databases

Although we have made our dataset and results publicly available, another way of making annotations available to a wider audience is to submit them to large bioinformatics data centers such as Ensembl, GenBank etc., which we intend to do. Enquiries have already been made on submitting the tissues of expression results to EMBL. As for the annotation results, we are yet to identify the appropriate agency to submit to.

5.2.2 Applying the Framework to Update Annotations

Many of the reported annotations on gene transcripts were annotated years ago before the advantage of high-throughput sequencing and therefore can benefit from modern tools and datasets such as the RNA-seq data. I therefore would want to explore the possibilities of using the framework to bring new life into old and potentially outdated annotations in public databases. This will be a large undertaking and will require the development of additional comparative modeling tools to identify and report differences between existing transcripts and new ones.

5.2.3 Opportunity For Further Expression Analysis

Recently, May 2014, there was a new addition of RNA-seq data covering 27 tissues, which included the original 16 analyzed in this research, plus an additional 11. This gives the prospect of potentially extending the expression data to include these other tissues in the near future.

5.3 Conclusion

Overall, this work provides knowledge on how human GPCR gene structure influence its alternative splicing, provide detailed knowledge on tissues in which human GPCR isoforms are expressed, and of how human GPCR isoforms function through the influence of binding transcription factor. This research may provide insightful resource for researchers interested in gene regulation. By identifying the transcript specific cellular processes that specific GPCR transcripts are involved in, it also may allow for transcript specific ligand binding studies (drug targeting). It may also be a good starting point for those interested in pathway analysis and alternative spliced isoform signaling.

APENDIX A

Contain tables referred to in the dissertation as Appendix_Table

Appendix_Table 3.1: Normalized RPKM values for isoforms

HGNC Symbol	ensemblTransID	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymph node	Ovary	Prostate latal muscle	Testes	Thyroid	White blood	
ACKR2	ENST00000273145	0	21.456067	0.675	0	11.823899	0	11.156627	94.939394	0	0	0	0	0	7.2911392	0	
ACKR2	ENST00000422265	0	0	0	0	0.2515723	0	0	0	0	0	0	0	0	0	0	
ACKR2	ENST00000442925	36.6	0.460251	0	9.4362416	0	3.4957983	0	2.9090909	5.5505618	17.59253499	4.3432343	6.3255814	12.875	9.115192	0	0.2931937
ACKR2	ENST00000492609	0	1.4058577	0	0	0	0	0	0	0	0	0	0	0	0	0	
ACKR2	ENST00000493193	0	0	0.05	0	0	0	0	0	9.3333333	0	0	0	0	0.3806344	0.8607595	
ACKR2	ENST00000494619	0.325	0	1.975	0.1208054	0.0251572	1.9327731	0	0	0	0	0.2244224	0	0	0.0133556	0	
ACKR2	ENST00000497921	0	0	0	0.2013423	0	0	0	0	0	0	0.290429	0	0	0	0	
ACKR3	ENST00000272928	210.4625	25.305439	2.4	69.959732	25.962264	40.722689	4.8674699	111.30303	25.662921	12.33592535	4.6468647	35.395349	107.875	7.8464107	30.101266	28.984293
ACKR3	ENST00000447924	15.3875	151.61506	1.7	28.120805	19.522013	3.6302521	5.5060241	16.515152	40.865169	85.30015552	7.5379538	18.651163	53.875	2.0767947	34.164557	4.4188482
ACKR4	ENST00000249887	16.5875	9.2384937	3.425	10.966443	0.3018868	8.7394958	2.373494	0.1818182	13.752809	3.185069984	0.0792079	5.1162791	0	0.4073456	3.2151899	0
ADCYAP1R1	ENST00000304166	14.3375	2.5104603	87.875	11.275168	8.490566	5.9159664	0.4337349	0.6666667	1.5168539	1.841368585	10.217822	4.6395349	0.5	3.4724541	2.1772152	0.104712
ADCYAP1R1	ENST00000396211	0	0	1.825	0	0	0	0	0	8.6404494	0	0	0	0	0	0	
ADCYAP1R1	ENST00000409363	0	0	0	0	0	0	0	0	0	0	0	0	31.5	0	0	
ADCYAP1R1	ENST00000409489	0	1.2217573	0	0	0	0	0	0	0	0	0	0	0	0	0	
ADCYAP1R1	ENST00000431811	0	0	0	0.2818792	0	3.6638655	0	0	0	0	0	0	0	0.0066778	0	
ADORA1	ENST00000309502	1.425	0.6443515	0	9.7181208	0.163522	4.4537815	1.7831325	0.6969697	0	0	0.4488449	0.9651163	0.5	9.6293823	4.1012658	0
ADORA1	ENST00000337894	0	0	17.7	0	0.3522013	0	1	0	0.5168539	0.385692068	1.1881188	0	0	1.8430718	3.1772152	0
ADORA1	ENST00000367235	0	0	0	0	0	0	0	0	0	0	0.5412541	0	0	0	0	0.1884817
ADORA1	ENST00000367236	0	0	0.5	0	0	0	0	0	0	0	0	0	1.5	0.066778	0.2911392	0
ADORA2A	ENST0000037539	3.8125	5.497908	0	0.5100671	0	9.5966387	11.518072	41.545455	11.617978	3.707620529	14.587459	2.627907	0	5.4357262	0	13.109948
ADORA2A	ENST00000424232	0	0	0	3.3825503	0	0	0	0	0	0	0	2.3366337	0	0	0	5.9113924
ADORA2A	ENST00000436735	2.7125	0.0585774	0	0	0	0	0	9.4939759	0	0	0.410575428	1.1881188	0.1627907	0	0	0.278481
ADORA2A	ENST00000439591	0	0	0	0	0	0	0	0	26.921348	0	0	0	0	0	0	
ADORA2A	ENST00000444262	0	0	0	0	0	0	0	0	0	0	0.9636964	0	0	2.4240401	0.0379747	59.874346
ADORA2A	ENST00000464977	0	0.7949791	0	1.5436242	0	0	0	0	0	0	0	0	0	0	0	
ADORA2A	ENST00000472248	0	0	0	0	0	0	0	0	18.424242	0	0	0	0	0	0	
ADORA2A	ENST00000486108	0	0.9539749	0	0	2.7295597	0.1008403	0	7.030303	1.0224719	0	0	0	0	0.9549249	2.4683544	
ADORA2A	ENST00000496258	0	0.5857741	0	0	0	0	0	0	15.2120121	0.6629213	0	0	0	0	0	
ADORA2A	ENST00000496497	1.05	0	0	0	0	0	0	5.8554217	0	0	0	3.1023102	0	0	1.7028381	5.1265823
ADORA2B	ENST00000304222	0.225	0.2175732	1.625	0.1342282	0.754717	0.6386555	0.4216867	1.4545455	0.3370787	0.043545879	1.2541254	0.2093023	0.25	0.2136895	0.1392405	1.5287958
ADORA3	ENST00000241356	4.875	0.334728	17.625	0.8993289	1.3207547	0.7394958	1.2289157	4.2121212	0.1011236	5.306376361	0.330033	2.1860465	1.75	2.5175292	1.4810127	3.2879581
ADORA3	ENST00000369716	0.425	0.167364	0.125	0.8724832	0.6918239	0.1008403	0.6626506	1.3636364	0	0.277760498	0	0	0	0	0.5063291	0
ADORA3	ENST00000369717	0.2875	0.1422594	2.1	0	0	0	0	0	0	0	1.629860031	0	0	0	10.964942	0
ADORA3	ENST00000442484	0	0	0	0	0	0	0.3193277	0	0	0	0	0	0	0	0	
ADORA3	ENST00000443498	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.6282723	
ADORA3	ENST00000463993	0.4125	0	5.4	0	0	0	0.3529412	0	3.6060606	0	0.715396579	0.6336634	0.9767442	0	1.8764608	0.0886076
ADRA1A	ENST00000276393	1.05	0.4686192	3.8	41.946309	0.9433962	108.42017	3.8915663	170.27273	16.370787	3.869362364	0.7656766	8.5581395	0	0.3405676	3.8987342	0
ADRA1A	ENST00000354550	9.1625	4.6443515	3.175	22.389262	1.0943396	11.84874	0	16.363636	1.5393258	0	5.2673267	0	27.75	2.2570952	0	0
ADRA1A	ENST00000380582	4.025	3.0041841	1.425	13.087248	1.7484277	12.588235	1.4457831	11.606061	1.5168539	0	3.9207921	0.4651163	0	1.6560935	1.9240506	0
ADRA1A	ENST00000380586	3.5875	1.9246862	0	5.7986577	3.7358491	3.4957983	0.2168675	404.48485	0.7191011	0	2.6006601	0.255814	0	0.5208681	0.8481013	0
ADRA1A	ENST00000380587	0	0.2677824	0	0	0.754717	0.2689076	0	0	0	0	0	0	0	1.4023372	0	5.4886911
ADRA2C	ENST00000330055	2.325	1.3723849	0.75	1.261745	0.7169811	1.512605	3.6024096	0.0606061	0.2921348	1.561430793	0.8448845	1	6.5	0.8013356	0.5316456	0
AGTR1	ENST00000349243	3.9125	0	0	1.8389262	0	0	5.3253012	0	0	4.547433904	0	5.627907	0	0	0	
AGTR1	ENST00000402260	341.975	14.786611	0.25	85.557047	168.67925	44.420168	122.74699	967.63636	17.640449	33.43079316	16.910891	23.918605	260.875	60.534224	84.848101	0
AGTR1	ENST00000404754	37.225	0	0	10.926175	43.396226	17.07563	16.759036	160.42424	13.606742	4.416796267	0	49.802326	291.125	0	12.556962	0
AGTR1	ENST00000461609	0	0	0	0	0.1132075	0	0	0	0	0	0	14.616279	0	0	2.7341772	0
AGTR1	ENST00000474935	6.275	0	0.175	9.2885906	0	0	0	5.2727273	0	0	0	0	0	0.4683544	0	
AGTR1	ENST00000475347	4.6625	0	0	3.1946309	4.6792453	0	0	17.121212	0	0	0.2244224	0	0	0	14.43038	0
AGTR1	ENST00000497524	0	0	0	0	47.949686	0	0	18.818182	0	0	0	0	3.2790698	0	2.4440735	0
AGTR2	ENST00000371906	0	0.0502092	0	0	0.0628931	0.1512605	0.1084337	0	0.6966292	0.080870918	0.0924092	0.0697674	0	0.0601002	0.0379747	0
AVPR1A	ENST00000299178	5.7125	0.5104603	0.15	2.0402685	0.9433962	0.487395	0.7831325	10.151515	0.0674157	0.479004666	0.330033	0.3372093	0.5	0.2604341	0.9493671	0.0628272
AVPR1B	ENST00000367126	0	0	0.05	0	0.0251572	0	0.0240964	0	0	0	0	0	0	0.0126582	0	

HGNC Symbol	ensemblTransID	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymph node	Ovary	Prostate latal muscle	Testes	Thyroid	White blood	
AVPR2	ENST00000337474	0	2.4937238	0	4.4161074	1.9245283	0.3529412	0	0	0.0337079	2.189735614	0	1.2674419	4.625	0	0	0
AVPR2	ENST00000358927	2.5375	0	0	0	0	0	2.4819277	0	0	0	0	0	0.360601	0.0379747	0	
BAI1	ENST00000323289	3.55	4.9958159	6.3	5.2348993	8.5157233	9.7983193	4.7349398	12.454545	3	2.67496112	6.9438944	4.1744186	10.625	4.39399	3.0126582	5.0890052
BAI2	ENST00000257070	0	0.0083682	0	0	0	0	0	0	0	0	0	0	0.4006678	0	0	
BAI2	ENST00000373655	0	0	0	0	0	0.0504202	0	0	0.0786517	0	0.2640264	0	0	0	0.2911392	
BAI2	ENST00000373658	0	0	0	0.0939597	0	0	0	0	0	0	0	0	0	0	0	
BAI2	ENST00000398538	0.1875	0	0	0	0.1509434	0.1680672	0.0963855	0	0	0	0.8184818	0	0	0	0	
BAI2	ENST00000398547	0	0	3.075	0	0.0125786	0	0	0	0	0	0	0	0	0	0	
BAI2	ENST00000420125	0.125	0.2259414	0.975	0	0	0	0.0120482	0	0	0.161741835	0.8316832	0	0.125	0.2136895	0.4810127	0.0628272
BAI2	ENST00000436464	0	0	0	0	0	0	0	0	0	0.024883359	0.1188119	0.0348837	0	0	0.0506329	0
BAI3	ENST00000370598	0.9	0.6276151	8.4	0.7516779	10.591195	10.487395	0.5060241	2.4545455	2.0449438	0.447900467	0.2376238	8.0930233	4.875	5.6227045	1.8101266	0.2303665
BDKRB1	ENST00000216629	0.05	0.0334728	0.075	0.2416107	0.1383648	0.0840336	0.9638554	0.0606061	0.1348315	0.018662519	0.1848185	0.5116279	0	0.3873122	0.1139241	0
BDKRB1	ENST00000306005	5.4375	1.1799163	1.55	6.1073826	6.0880503	1.1092437	9.626506	5.2727273	2.8202247	0.702954899	2.5742574	5.7674419	4.875	1.8764608	2.2278481	0.6701571
BRS3	ENST00000370648	0	0.1171548	5.025	0	0	0.2891566	0	0	0.00622084	0.3168317	0.0348837	0	0.360601	0.1139241	0	
C3AR1	ENST0000037637	1.4875	0.7029289	2.325	0.5100671	0.2389937	2.3529412	1.4457831	3.9393939	1.9775281	0.690513219	0.6336634	1.0813953	2.75	0.293823	0.556962	17.696335
C5AR1	ENST00000355085	11.35	0.4686192	5.325	1.6241611	3.0188679	2.6722689	4.4698795	6.5151515	6.7865169	2.090202177	2.4026403	1.1046512	6	1.0217028	0.3417722	64.753927
C5AR2	ENST00000252767	0.4125	0.0585774	0.075	0.1208054	0.0251572	0.0504202	0.1566265	0.3939394	0.1910112	0.099533437	0.0606066	0.0465116	0	0.0601002	0.0886076	2.9319372
CALCR	ENST00000359558	0.125	0	0.15	0.1610738	0	0.1008403	7.3493976	0	0.2022472	0	0.1716172	0.1046512	0	0.2470785	0	0
CALCR	ENST00000360249	0	0	0	0	0	0	0	0	0	0	0	0	7.75	0	0	0
CALCR	ENST00000421592	0	0	0	0	0	0	0	0	0	0.080870918	0	0	0	0	0	0
CALCR	ENST00000426151	0	0	0	0	0	0	0	0	0	0	0	0	0	0.1803005	0.0379747	0
CALCRL	ENST00000392370	252.35	7.958159	1.65	31.879195	10.943396	56.991597	10.457831	0	151.1236	10.38880249	5.4125413	11.94186	33.125	6.5642738	24.037975	12.879581
CALCRL	ENST00000409998	0	9.7238494	1	0	0	0	1.5783133	24.848485	0	0	0	0	0	0.1469115	1.5189873	2.1361257
CALCRL	ENST00000410068	13.875	0.5020921	0	14.845638	4	3.4957983	0	0	20.213483	4.037325039	5.2277228	0.0116279	78.625	1.2554257	0	0.4188482
CALCRL	ENST00000447403	0	0	1.35	0	0	0	0	0	0	0	0	0	0	0	0	0
CALCRL	ENST00000461244	0.15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CALCRL	ENST00000474212	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.460733	0
CALCRL	ENST00000479784	0	8.3598326	0	0	0	0.4537815	0	0	0	0	0	0	0	0	0.9620253	0
CALCRL	ENST00000485973	0.4	2.8786611	0.625	0.3758389	0	0.7563025	1.2048193	0	9.4382022	1.312597201	4.1320132	0.5581395	0	0.0801336	0.8227848	0
CASR	ENST00000296154	0	0	0.175	0	0	0	72.759036	0	0	0	0	0	0	0.2871452	11.670886	0
CASR	ENST00000490131	0	0.0502092	0	0	0	0	3.0240964	0	0.3932584	0	0	0	0	0	2.4050633	0.3350785
CASR	ENST00000490186	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0601002	0.1898734	0
CASR	ENST00000498619	0	0	0	0	0	0	17.86747	0	0	0	2.6006601	0	0	0	12.367089	0
CCKAR	ENST00000295589	0.15	0	0	0	0.2767296	0.0168067	0.060241	0	0	0	0	0	0.0930233	0	0.033389	0.164557
CCKBR	ENST00000334619	0.05	0	2.55	0.0671141	0.0754717	0	0.0963855	0	0.0224719	0	0.0924092	0.0697674	0	0.0601002	0.0506329	0
CCR1	ENST00000296140	1.0375	1.2217573	4.75	1.6510067	0.2012579	1.1932773	2.1686747	4.8484848	1.6629213	2.077760498	1.4785479	0.2209302	3.125	0.7145242	0.6708861	39.308901
CCR10	ENST00000332438	0.1125	0.334728	1.075	0.6845638	0.2389937	0.1176471	0.3975904	0.1515152	0.6516854	0.255054432	0.3564356	0.2674419	0.75	0.0534224	0.2025316	0.5026178
CCR2	ENST00000400888	0.4625	14.953975	0.175	1.2885906	1.7987421	6.487395	10.819277	23.393939	7.3033708	103.8009331	9.8745875	1.2325581	1.625	5.2487479	3.5063291	1363.5602
CCR2	ENST00000421659	0	0	0	0	0	0	0	0	1.3595506	0	0	0	0	0	0	0
CCR2	ENST00000445132	0.6875	0.8870293	0	0.9530201	0	5.3613445	0.5301205	7.8484848	1.7078652	8.939346812	4.0132013	0.744186	3.125	0.6944908	1.4303797	120.31414
CCR3	ENST00000357422	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.2405063	1.0471204
CCR3	ENST00000395940	0	0	0	0.0805369	0	0	0	0	1.3595506	0	0.0132013	0	0	0	0	13.256545
CCR3	ENST00000395942	0	0.0669456	0	0.0402685	0	0	0.060241	0.0606061	0.2022472	0.255054432	0	0	0	0	0.2531646	0
CCR3	ENST00000452454	0	0	0	0.2818792	0	0	0	0	0	0	0	0	0	0	0	0.565445
CCR4	ENST00000309593	0.1125	0.4267782	0.15	0.0536913	0.0251572	0.0504202	0.0240964	0.1212121	0.6179775	1.219284603	0.039604	0.0232558	2.25	0.0400668	0.0253165	8.0209424
CCR5	ENST00000292303	0	0	0	0	0	0	0	4.6969697	5.9438202	0	0	0.9186047	0	0	0.6455696	0
CCR5	ENST00000343801	0	0	1.45	0	0	0	0	0.5454545	0	0	0	0	0	0	0	0
CCR5	ENST00000445772	0.1625	0	0	0	0	0	0	0.6506024	2.3636364	0.3820225	8.149300156	0.7392739	0.4302326	0	0	0.3037975
CCR6	ENST00000341935	0	3.0794979	0	0	0	0.9747899	0.3855422	0	0	1.00777605	0.9108911	0	0	0.7946578	0	5.6335079
CCR6	ENST00000349984	0.8625	7.9414226	0.625	0	1.3584906	0	1.4578313	0	0	8.821150855	0.2112211	0.372093	4.375	3.1051753	6.3417722	33.277487
CCR6	ENST00000400926	0	18.384937	0	2.3758389	0	0	0	6.1212121	0.5955056	6.569206843	0.5940594	0	0	0.0400668	1.2025316	24.691099
CCR7	ENST00000246657	0.2375	6.1589958	0.55	0.1208054	0.1886792	0.487395	0.0843373	0.3939394	0.4382022	1.953343701	0.4224422	0.2790698	0	0.1268781	0.2025316	26.492147
CCR8	ENST00000326306	0	0	0	0.1342282	0	0	0	0	0	1.076205288	0	0	0	0	0	3.2041885
CCR9	ENST00000355983	0	0	0	0	0	0	0	0	0	0	0.0697674	0	0	0	0	0

HGNC Symbol	ensemblTransID	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymph node	Ovary	Prostate	Testes	Thyroid	White blood		
CCR9	ENST00000395963	0	3.4058577	0.125	0	0.6540881	0.1008403	0.8433735	0	0	1.088646967	0.5412541	0	10.75	0	0	11.371728	
CCR9	ENST00000422395	0.1125	0	0	0.1208054	0	0	0.0963855	0	0.2359551	0	0	0	0	0.2151899	0		
CCR9	ENST00000463197	0	0.0669456	0	0	0	0	0	0	0	0	0	0	0	0	0		
CCL2	ENST00000357392	0	2.8535565	0	2.3624161	4.0880503	0	0	0	0	0	0	0	0	0	16		
CCL2	ENST00000399036	77.375	2.125523	2.7	8.7651007	9.9496855	5.7478992	3.0240964	7.3030303	15.05618	1.65474339	0.7656766	5.8139535	12.75	0.327212	2.6708861	3.6230366	
CCL2	ENST00000400880	7.7625	0	0	1.4228188	0	0.7058824	0.120482	0	1.3707865	4.180404355	0	1.0465116	0	0	0.0253165	0	
CCL2	ENST00000400882	8.025	0	0	7.5436242	0	0.5546218	0	0	5.7640449	0	0	0	0	0.5742905	0		
CCL2	ENST00000433848	5.3875	0.0502092	0	0	4.3522013	0	0	0	0	0	0	0	13.75	0	0	5.4031414	
CD97	ENST00000242786	0	0	5.325	0	0	0	0	7.1818182	0	0	0	0	0	0	0		
CD97	ENST00000357355	0	20.627615	1.025	0	0	0	21.337349	0	10.898876	0	0	36.651163	178	6.9048414	31.924051	243.51832	
CD97	ENST00000358600	123.55	234.17573	7.825	74.926175	112.57862	96.857143	52.795181	81.151515	215.70787	123.2597201	101.41254	143.83721	995.375	67.48581	151.77215	2022.555	
CELSR1	ENST00000262738	1	1.8912134	0.275	0.8053691	0.327044	5.8991597	2.746988	6.8484848	1.8202247	1.101088647	4.660066	2.0116279	0	1.2086811	3.6582278	1.0052356	
CELSR1	ENST00000395964	0	1.5230126	0	0	0	3.3613445	3.1807229	0	3.4719101	0	7.1947195	0	0	0.0066778	4.7468354	0	
CELSR2	ENST00000271332	0.65	0.3263598	8.3	1.6912752	1.0440252	2.9411765	3.0481928	8.8181818	0.4606742	0.597200622	3.4587459	1.5348837	23.625	1.8163606	9.7088608	3.078534	
CELSR3	ENST00000164024	0.525	1.9246862	1.575	0.1342282	0.5408805	0.3193277	0	0.0909091	0.5505618	2.127527216	0.6468647	0.744186	0.5	3.1051753	0.9113924	1.382199	
CHRM1	ENST00000306960	0.0125	0.0083682	6.5	0.0536913	0.0503145	0.1512605	0.060241	0	0.0674157	0.00622084	0.0660066	0.5930233	0	0.1202003	0.1518987	0.0209424	
CHRM2	ENST00000320658	0.0375	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
CHRM2	ENST00000401861	0	0	0	0	0	3.1764706	0	0	0	0	0	0	0	0	0		
CHRM2	ENST00000402486	0	0	0	0	2.3773585	13.243697	0	0	0	0	0	0	0	0	0		
CHRM2	ENST00000445907	0	0.0083682	3.7	0	24.679245	164.58824	0.0240964	6.6363636	1.0786517	0.018662519	0.1056106	33.55814	0	0.2337229	0.0379747	0	
CHRM2	ENST00000453373	0	0	0	0	0	0	0	0	0	0.130637636	0	0	0	0.0400668	0		
CHRM2	ENST00000480591	0	0	0	0	0.4025157	9.9159664	0	0	0	0.037325039	0	1.4651163	0	0.0400668	0		
CHRM2	ENST00000481598	0	0	0	0	0.0125786	0	0	0	0	0	0	0	0	0	0		
CHRM3	ENST00000255380	0.8375	0.2761506	11.5	0.5369128	4.7924528	0.0672269	1.3975904	3.2121212	0.247191	0.167962675	15.617162	14.186047	0	0.6944908	1.3164557	0.0418848	
CHRM3	ENST0000048020	0.1125	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
CHRM4	ENST00000433765	0.025	0.0502092	0.45	0.0536913	0.1509434	0.0504202	0.0240964	0	0	0	0.0792079	0.1744186	1.25	0.3338898	0.0253165	0.104712	
CHRM5	ENST00000383263	0	0	13.05	0	0	0	0	0	0	0	0	0	0	0	0		
CMKLR1	ENST00000312143	41.6375	0	8.525	46.805369	6.5157233	1.8319328	18.674699	26.151515	15.516854	13.20684292	28.831683	18.011628	0	2.6844741	33.037975	112.71204	
CMKLR1	ENST00000397688	0	20.502092	0	0	0	0	0	0	0	0	0	0	0	0	0		
CMKLR1	ENST00000412676	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4.4188482		
CNR1	ENST00000369499	105.5875	0	4.9	0	76.163522	10.957983	5.8915663	3.4545455	0	21.9533437	2.4158416	7.0465116	0	19.719533	0	0.6701571	
CNR1	ENST00000369501	38.725	8.7447699	4.35	12.456376	7.9622642	0	0	0	3.3370787	3.856920684	0	6.0697674	134.875	2.8981636	9.8481013	0	
CNR1	ENST00000428600	0	0	0	8.5503356	0	7.3613445	0	0	0	6.2178218	0	10.375	4.1669449	8.7341772	0.7120419		
CNR2	ENST00000374472	0.0375	10.51046	0.025	0.0268456	0.0503145	0.0336134	0.0120482	0.8787879	0.1573034	7.925349922	0.0660066	0.0581395	0.125	0.1202003	0.8607595	15.853403	
CRHR1	ENST00000293493	0	0	0.775	0	0	0	0	0	0	0	0.9108911	0	0	0.3005008	0		
CRHR1	ENST00000314537	0	0	0	0	0.3892617	0	0	0	0	0	0	0	0	0	0.2405063		
CRHR1	ENST00000339069	0	0.1422594	0	0	0	0	0	0	0	0	0	0	0	0			
CRHR1	ENST00000352855	0	0	0.575	0	0	0	0	0	0	0	0	0	0	0.4808013	0		
CRHR2	ENST00000341843	0	0	0	0	0	0	1.1932773	0	0	0	0	0	0	0.0534224	0		
CRHR2	ENST00000348438	0.0375	0.1506276	0	0.2013423	0	0	0	0	0	0	0	0	0	0	0		
CRHR2	ENST00000423776	0	0.0753138	0	0	0	2.2352941	0	0	0	0	0.0924092	0.0465116	0.625	0	0.7594937	0	
CRHR2	ENST00000452278	0	0	0	0	0	0	0	0	0	0.6072607	2.0232558	0	0.3205342	0.2151899	0		
CRHR2	ENST00000462882	0	0.2259414	0	0.9530201	0	0	0	0	0	0	0.1046512	0	0	0	0		
CRHR2	ENST00000471646	0.1	0.5439331	1.55	0.4161074	0.1006289	0.605042	0	0	0	0.290429	0.3372093	0	0.0467446	0.2658228	0		
CX3CR1	ENST00000358309	0	0	0	0	0	0	0	0	8.6666667	0	0	12.475248	0	0	7.3188648	0	1050.9948
CX3CR1	ENST00000399220	0	0	0	0	0.1257862	0	0	0	7.494382	2.0155521	0	0	0	1.9113924	10.324607		
CX3CR1	ENST00000412814	0	0	11.175	0	0.1257862	0.0168067	0	0.4242424	0	0	0	0	0	0.0759494	21.005236		
CX3CR1	ENST00000435290	0.075	0	0	0	0	0	0.0168067	0	1.0561798	0	0.2640264	0	0	0	1.382199		
CXCR1	ENST00000295683	23.65	0.2175732	0.05	0.4832215	1.0943396	1.2268908	0.0481928	3.2424242	21.730337	3.384136858	0.0264026	1.3953488	8.875	0.7746244	0.1139241	1688.8586	
CXCR2	ENST00000318507	28.2	0.9874477	0.1	5.0201342	1.4591195	1.1932773	0.2289157	9.2727273	19.921348	3.720062208	0.039604	2.8837209	3.25	0.6410684	0.1012658	2958.911	
CXCR2	ENST00000415392	0.0125	0	0.175	0	0	0	0	0	0	8.255054432	0	0	0	0	0.3670886	0	
CXCR2	ENST00000428565	0	0	0	0	0	0	0	0	9.6629213	0	0.1056106	0	0	0	0	203.97906	
CXCR2	ENST00000453237	0	0	0	0	0.2515723	0.2857143	0	0	0	0	0	0	2	0.0467446	0	0	
CXCR3	ENST00000373691	0	5.0041841	0	0	0	0	0	0.2727273	0	1.542768274	0.9372937	0	0	0	1.4177215	3.3926702	

HGNC Symbol	ensemblTransID	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymph node	Ovary	Prostate latal muscle	Testes	Thyroid	White blood		
CXCR3	ENST00000373693	0.15	6.4518828	0	0.738255	0	0.3697479	0.4096386	0	0.7303371	5.611197512	0	0.0930233	0	0.2804674	0		
CXCR4	ENST00000241393	246.9	433.57322	4.325	35.610738	33.559748	7.6302521	11.46988	45.878788	48.921348	238.2830482	40.158416	19.546512	197.75	26.530885	27.21519		
CXCR4	ENST00000409817	1.65	124.27615	0.525	2.2684564	0	1.1092437	0.0481928	6.5454545	15.280899	33.57387247	2.2178218	0	0	2.130217	8.3164557	17.947644	
CXCR5	ENST00000292174	0.125	4.0753138	0.025	1.1812081	0.6289308	0.4033613	0.373494	0.1818182	0.1797753	2.712286159	0.1716172	2.3604651	22.25	0.39399	0.721519	7.7277487	
CXCR6	ENST00000304552	0	10.259414	0	2.1208054	0.0251572	3.2773109	1.0722892	17.121212	4.505618	12.54121306	0.9372937	0.9767442	0	0.1268781	1.7468354	16.544503	
CXCR6	ENST00000438735	0	0.5439331	0	0	0	0.4369748	0	0	1.494382	0	0	0	13.5	1.1352254	0	0	
CXCR6	ENST00000457814	0.1875	0	0.025	0	0	0.4033613	0.1807229	0	0.8089888	14.24572317	0	0.2093023	0	0	0	4.8167539	
CXCR6	ENST00000458629	0.2125	2.6861925	0.1	0	2.5408805	1.7815126	0	0.8181818	3.3932584	1.418351477	1.0429043	0	2.875	0	1	0	
CYSLTR1	ENST00000373304	14.6875	12.853556	0.875	9.5033557	2.7798742	4.8403361	4	5.3636364	2.6404494	12.52877138	2.3630363	3.9186047	9	1.6227045	1.6962025	116.60733	
CYSLTR1	ENST00000493254	2.5875	1.7154812	0	0.2416107	1.3333333	3.3445378	0.060241	0	1.6741573	1.275272162	1.7161716	1.5465116	0	0.9148581	1.0253165	3.2670157	
CYSLTR2	ENST00000282018	0.075	0.3430962	1	0.6174497	0.2641509	2.8235294	0.4096386	0.6060606	0.0786517	0.155520995	0.7656766	0.9302326	0	0.1001669	0.0632911	2.9319372	
DARC	ENST00000368121	294.3625	65.556485	2.375	128.33557	29.672956	16.05042	3.3493976	0.030303	159.08989	76.85225505	11.419142	95.27907	30	28.954925	12.265823	0	
DARC	ENST00000368122	16.4625	1.790795	23.525	8.5637584	10.716981	1.2941176	0	0.1818182	21.101124	12.16174184	0	0	11.25	5.015025	2.1518987	0	
DARC	ENST00000435307	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.1675393	
DRD1	ENST00000393752	0.9625	1.7322176	2.7	7.7181208	2.6666667	0	0.4819277	1.8484848	1.4269663	0.926905132	0.6732673	3.6395349	0	1.081803	0.2405063	0	
DRD2	ENST00000346454	0	0.2343096	0	2.1744966	0.9056604	0	0	0	0.3932584	0.223950233	0.7656766	0.5930233	0	2.9582638	0.7721519	0	
DRD2	ENST000003055319	0	0	0.4	0	0	0	0	0	0	0	0	0	0	0	0	0	
DRD2	ENST00000362072	0.85	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
DRD3	ENST00000383673	0	0.0920502	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5026178	
DRD3	ENST00000467632	0	0	0	0	0	0	0	0	0	0	0	0	0	0.1001669	0	0	
DRD4	ENST00000176183	0.075	0.125523	0.4	0.0939597	0.1257862	0.0336134	0.060241	0.030303	0.0898876	0.018662519	0.1584158	0.2906977	1	0.0801336	0.1012658	0.0209424	
DRD5	ENST00000304374	0	0	0.825	0	0.0125786	0.0672269	0	0	0	0	0.0264026	0	0	0.066778	0	0	
EDNRA	ENST00000324300	17.55	56.393305	1.925	28.845638	16.767296	55.294118	3.8554217	4.6060606	46.651685	11.76982893	64.646865	92.813953	0	58.550918	14.683544	0	
EDNRA	ENST00000339690	0	0	0	0	0	4.9915966	0	0	0	28.33592535	0	0	166.625	0	0	0	
EDNRA	ENST00000358556	92.45	86.953975	3.575	144.56376	68.566038	172.68908	21.024096	0	135.42697	13.06998445	368.09241	127.60465	119.375	189.37563	44.860759	1.1308901	
EDNRB	ENST00000342486	0	125.04603	27.55	76.52349	433.13208	735.34454	796.85542	647.51515	1297.5506	136.9953344	15.722772	103.67442	806.75	34.323873	278.56962	0.3141361	
EDNRB	ENST00000446573	47.5	61.690377	0	3.7583893	19.962264	0	10.819277	206.30303	0	22.92379471	8.2508251	13.697674	83	13.736227	5.164557	1.1937173	
ELTD1	ENST00000370742	177.35	25.589958	3.8	65.932886	57.63522	35.731092	19.807229	70.878788	65.651685	27.22239502	23.80198	24.523256	144.25	24.207012	26.278481	0	
ELTD1	ENST00000401034	0	0	0	0	0	0	0	0	0	0	0	0	8.875	0	0	0	
EMR1	ENST00000250572	1	6.4686192	1.125	0.0671141	0.3396226	0.8235294	0.1807229	0	0	3.290824261	2.5214521	0.8488372	0	0	0.1265823	139.62304	
EMR1	ENST00000312053	1.45	3.832636	0	0.3624161	0	0	0	0	0	0	0	0	0	1.9365609	0	113.50785	
EMR1	ENST00000381404	0	0	0.175	0	0	0	0	8.8787879	0	0	0	0	0	0	0	10.659686	
EMR1	ENST00000381407	0	0	0	0	0	0	0	0	12.764045	0	0	0	0	0	0	0	
EMR2	ENST00000315576	0	0	0	1.114094	3.4968553	0	0	0	0	0	0	0	0	0	0	0	
EMR2	ENST00000346057	3.6125	0	0	6.442953	0	0	0	0	0	0	0	0	44	0	3.2658228	0	
EMR2	ENST00000353876	2.4125	0	5.3	0	3.8742138	15.647059	1.5060241	11.939394	6.7078652	2.052877138	3.8151815	0	0	0.5542571	0	352.20942	
EMR2	ENST00000360222	18.15	13.280335	3.8	0	0	0	2.4819277	11.545455	0	15.80093313	6.2310231	11.406977	0	2.3372287	2.6962025	281.90576	
EMR2	ENST00000392962	0	24.410042	4.325	83.100671	7.9874214	32.218487	0	1.8484848	14.831461	0	41.623762	18.639535	329.375	12	72.189873	13.445026	
EMR2	ENST00000392964	0	0	0	0	0.9559748	0	5.060241	0	0	0	0	0	0	0	0	0	
EMR2	ENST00000392965	3.0375	22.577406	0	0	0	0	13.445783	7.8787879	8.2022472	0	0	0	6.127907	0	2.0701169	2.6835443	0
EMR2	ENST00000392967	1.1375	5.958159	0	0	0	0	0	0	16.842697	0	0	0	0	1.1419032	0	157.48691	
EMR3	ENST00000253673	2.375	3.7991632	0	0	0	0.487395	0	0	3.6067416	0.99533437	0.4224422	0	0	0	0.4936709	268.77487	
EMR3	ENST00000344373	0	0	0.4	0	0.1509434	0.6218487	0	3.4545455	0	0	0	0	0	0	0	201.75916	
EMR3	ENST00000443157	0	0	0	0	9.0604027	0	0	0	0.5617978	0	0	0	0	0	0	0	
F2R	ENST00000319211	48.275	57.129707	1.65	72.885906	85.962264	124	90.578313	320.39394	186.1573	91.47744946	57.108911	56	290	46.357262	58.392405	277.7801	
F2RL1	ENST00000296677	0.1375	0.2677824	0.275	0.1879195	1.8616352	0.302521	3.2650602	3.3636364	1.0337079	0.055987558	0.1716172	1.0348837	1.125	0.2136895	1.0253165	4.921466	
F2RL2	ENST00000296641	0.3875	0.251046	0.1	0.5369128	0.7044025	0.1512605	0.6626506	14.424242	0.3483146	0.242612753	0.4620462	0.2674419	0.375	1.7295492	1.7721519	3.078534	
F2RL3	ENST00000248076	61.1375	0.6861925	0.25	0.1610738	3.4968553	0.184874	0.2409639	0.6363636	5.9101124	0.485225505	0.1320132	0.627907	6.875	0.1869783	7.2658228	0.1256545	
FFAR1	ENST00000246553	0	0.0251046	0.025	0	0	0	0	0	0.00622084	0.1452145	0	0	0	0.0126582	0		
FFAR2	ENST00000246549	0.2875	0.0502092	0	0.1744966	0.0251572	0.0168067	0.0240964	0.030303	0.6516854	0.00622084	0	0.0581395	0.25	0.0267112	0.0253165	5.8429319	
FFAR3	ENST00000327809	0.625	0.0585774	0.2	0.2684564	0.8427673	0.0840336	0	0.1818182	0.4157303	0.087091757	0.1320132	0.1046512	1	0.033389	0.1392405	0.0837696	
FFAR3L	ENST00000454971	51.7125	2.3430962	0	20.550336	55.509434	0.7058824	0	4.8181818	1.8314607	1.01399689	4.8976898	7.1627907	6.125	0.4674457	2.5949367	1.3403141	
FPR1	ENST00000304748	6.1375	0.9372385	10.15	0.3758389	0.7421384	4.1512605	1.4337349	3.5151515	8.7865169	2.388802488	0.3564356	0.8953488	2.125	0.2871452	0.3544304	114.99476	
FPR2	ENST00000340023	1.05	0.2845188	0.75	0.1208054	0.1006289	0.6890756	0.253012	0.9393939	1.7752809	0.559875583	0.3564356	0.1395349	0.75	0.1469115	0.0506329	34.534031	

HGNC Symbol	ensemblTransID	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymph node	Ovary	Prostate	Testes	Thyroid	White blood		
FPR3	ENST00000339223	2.1	1.4560669	5.075	4.4697987	1.3584906	1.7647059	1.3614458	2.4545455	1.5505618	0.230171073	1.8745875	1.0232558	1.875	0.5275459	0.4810127	4.0418848	
FSHR	ENST00000304421	0	0	0.025	0	0	0.1680672	0.0481928	0	0	0	0.1188119	0.1046512	0	0.0801336	0	0	
FSHR	ENST00000346173	0	0	0	0	0	0	0	0	0	0	0	0	0	0.2003339	0	0	
FSHR	ENST00000406846	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0400668	0	0	
FSHR	ENST00000419927	0	0	0	0	0	0	0	0	0	0	0.4752475	0	0	0.2470785	0	0	
FSHR	ENST00000469138	0	0	0	0	0	0	0.4096386	0	0	0	0	0	0	0.0868114	0	0	
FZD1	ENST00000287934	2.325	1.790795	2.075	4.4161074	4.1132075	7.1260504	8.3373494	2	3.0337079	1.250388802	9.029703	3.5232558	8.25	5.4223706	26.227848	2.8900524	
FZD10	ENST00000229030	0.15	0.292887	0.15	0.0939597	0.0503145	0.0504202	0.0240964	0	0.1123596	0.037325039	0.3036304	0.1976744	0.375	0.0934891	0.2151899	0	
FZD2	ENST00000315323	0.0875	0.0753138	0.1	0.0536913	0.2138365	0.5378151	0.7951807	0.2727273	0.2921348	0.298600311	0.6732673	0.3023256	4.125	0.1803005	0.0886076	0.5026178	
FZD3	ENST00000240093	7.075	11.087866	29.075	8.7785235	38.075472	14.857143	27.626506	10.272727	6.7303371	11.34059098	75.854785	24.883721	70.875	42.217028	27.012658	12.335079	
FZD5	ENST00000295417	0.6375	0.2761506	0.75	1.261745	2.8176101	0.2352941	1.9277108	12.969697	1	0.093312597	0.5544554	0.4534884	1.375	0.1669449	1.2531646	0.104712	
FZD6	ENST00000358755	5.8	3.7154812	3.25	4.8993289	3.408805	4.0504202	7.1084337	2.4242424	5.2022472	2.295489891	4.7656766	6.0930233	13.625	1.5893155	5.278481	1.2356021	
FZD7	ENST00000286201	0.8875	1.0209205	1.2	4.7114094	5.2578616	1.7815126	2.1084337	0.2424242	0.8651685	0.684292379	5.3069307	9.5465116	12.25	1.1352254	1.7721519	0.104712	
FZD8	ENST00000374694	0.8125	1.5983264	0.8	1.261745	2.327044	0.3361345	3.4975904	1	0.4719101	1.374805599	1.3993399	1.4302326	10.875	0.4874791	1.443038	0.104712	
FZD9	ENST00000344575	0.075	0.0167364	0.55	0.0805369	0.0503145	0.0336134	0.0240964	0.0909091	0.0224719	0.031104199	0.039604	0.0116279	5.625	0.1602671	0.1265823	0.0418848	
GABBR1	ENST00000355973	0.525	0	0	0	0	0	0	3.8181818	0	0	0	0	0	1.6708861	0	0	
GABBR1	ENST00000376977	2.075	0	0	0	0	0	3.9156627	0	1.2921348	1.405909798	2.4422442	0	0	7.8664441	4.9746835	0	
GABBR1	ENST00000377012	0.875	13.481172	1.075	2.6442953	11.572327	35.378151	3.0722892	1.8181818	2.8651685	10.19595645	12.50165	3.9186047	0	2.0166945	1.4177215	1.6963351	
GABBR1	ENST00000377016	0	0	0	0	1.2704403	0	0	0	0	0	0	0	0	0	0	0	
GABBR1	ENST00000377034	0	0	0	0	0	39.798319	0	0	0	0	0	0	22.5	0	0	0	
GABBR1	ENST00000462632	0	0	0	0	0	0	0	0	0	0	0	0	0	0.1268781	0	0	
GABBR1	ENST00000467259	3.125	5.6569038	22.825	12.61745	10.654088	9.5630252	5.2048193	1.3030303	5.9101124	2.550544323	25.306931	17.860465	55	3.9933222	20.35443	3.5183246	
GABBR1	ENST00000472823	0	0	49.925	0	0	0	0	0	0	0.3932584	0	2.1122112	7.1511628	0	0	0	0
GABBR2	ENST00000259455	2.8625	0.083682	174.25	0.4832215	0.0754717	6.7563025	1.3012048	0.5151515	0.6067416	0.454121306	0.0660066	0.4651163	0.375	2.4841402	15.468354	0.0418848	
GALR1	ENST00000299727	0.6125	0.2175732	1.025	0.8322148	0.1761006	0.1176471	0.1566265	0.0606061	0.0224719	0.018662519	0.5940594	0.5348837	0.375	0.0934891	0.0632911	0.0418848	
GALR2	ENST00000329003	0	0.0083682	0	0.0134228	0.327044	0	0.0120482	0.0606061	0.011236	0.031104199	0.0132013	0.0116279	0	0.0133556	0.0126582	0.0209424	
GALR3	ENST00000249041	0	0	0.025	0.0134228	0.0125786	0	0.0120482	0	0	0	0	0.0116279	0	0.0133556	0.0126582	0	
GCCR	ENST00000400723	0.0375	0.2677824	0	0	0.0251572	0	1.686747	20.181818	0	0.024883359	0	0.2674419	0	0.3005008	0.0886076	0	
GHSR	ENST00000427970	0	0	0.2	0	0	0	0	0	0	0	0	0	0	0.0200334	0	0	
GIPR	ENST00000263281	15.225	1.2635983	1.325	3.5436242	5.1320755	1.8151261	0.5421687	0.7272727	6.4157303	4.89580933	6.2442244	2.7093023	1.375	2.203673	0.3670886	0	
GLP1R	ENST00000373256	0.025	0.2594142	0.875	0.2684564	0.0251572	1.5294118	0.0722892	0.2727273	0.1460674	0.136858476	0.0132013	0.0232558	0	0.0934891	0.6455696	0	
GLP2R	ENST00000262441	10.2875	1.3723849	1.05	10.496644	84.062893	2.7394958	0.060241	11.727273	3.3483146	7.944012442	1.359736	85.290698	6.875	2.8247078	1.3037975	0	
GNRHR	ENST00000226413	0	0.5271967	0.075	0	0	0	1.4698795	0	0.6404494	0	0	0.3023256	22.25	0.3539232	1.1392405	0	
GNRHR	ENST00000420975	0.575	0.292887	0.35	1.2885906	0.5534591	1.5630252	0.626506	0.8181818	0	0.367029549	1.2541254	0.3023256	1.125	0.5008347	0.1139241	1.8848168	
GPBAR1	ENST00000479077	2.1	0.167364	0.075	0.7919463	0	0.1680672	0	0	0.1011236	0.311041991	0	0.1511628	0	0.066778	0	1.9057592	
GPER1	ENST00000297469	0	1.5732218	0	0.9127517	0.8679245	0	2.6385542	46.424242	1.9438202	0.976671851	0	0.744186	0	0	0	0	
GPER1	ENST00000397088	1.9	7.2133891	3.95	22.120805	6.9056604	7.6806723	4.6024096	4.9090909	0	5.393468118	3.7887789	9.4302326	14.875	3.8731219	2.8607595	0	
GPER1	ENST00000397092	2.6	0.6192469	0	3.0201342	0	0	0	32.757576	0	0	0	0	0	0	10.341772	3.078534	0
GPER1	ENST00000401670	0	0	0	7.0067114	16.427673	0	0	0	0	5.903576983	0	9.8488372	0	0	0	0	0
GPER1	ENST00000413368	0	0	0.1	0.9127517	0	0.789916	0	0	0	0.230171073	0.5808581	0	25	0.0601002	0	0	
GPR1	ENST00000407325	0.25	0	0.2	0.295302	0	0.4033613	0	0	0	0	0	0.6162791	0.5	0.0467446	0	0	
GPR1	ENST00000437420	0.125	2.9121339	0	0.9798658	0	0	0	0	0	0.261275272	3.5643564	0.3139535	0	0.1402337	0	0	
GPR1	ENST00000439932	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.4683544	0	0
GPR1	ENST00000442134	0	0	0	0	0	0	0	0	0	3.794712286	0	0.1395349	0	0.0534224	0	0	
GPR1	ENST00000447845	0.025	4.1589958	0.725	0	0.1761006	0.7226891	0	0	0.1685393	0	0	0	2.375	0	0	0	
GPR1	ENST00000451790	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.2658228	0	0
GPR101	ENST00000298110	0	0	0.25	0	0	0	0	0	0	0	0.1188119	0	0	0.0400668	0	0	
GPR110	ENST00000283297	0	0	0	0	0	0	1.6626506	0	0	0	0	8.0465116	0	0	4.1392405	0	0
GPR110	ENST00000371243	0	0	0	0.885906	0	0	0	0	0.247191	0	0	0.2790698	0	0	0	0	0
GPR110	ENST00000371253	0	0	0	0.3758389	0	0	21.554217	0	0.4157303	0.043545879	0.1452145	3.3604651	0	2.8380634	12.037975	0	
GPR110	ENST00000419892	0.0625	0.0920502	0	0	0	0	1.8554217	0.1212121	0.3258427	0	0.0528053	1.2790698	0	0.0267112	0	0	
GPR110	ENST00000471487	0	0	0	0.0536913	0	0	0	0	0	0	0	0.0348837	0	0	0.7721519	0	0
GPR110	ENST00000474800	0	0.1171548	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
GPR110	ENST00000477771	0	0	0	0	0	0	0	0.0843373	0	0	0	0	0	0	0.2151899	0	0

HGNC Symbol	ensemblTransID	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymph node	Ovary	Prostate latal muscle	Testes	Thyroid	White blood	
GPR110	ENST00000491283	0	0	0	0	0	0	0	0	0	0	0.2209302	0	0	0	0	
GPR111	ENST00000283303	0	0	0.075	0	0	3.2941176	0	0	0	0	1.3069307	2.2209302	0	0	0	
GPR111	ENST00000296862	0	0	0	0	0	0	0.4939759	0	0	0	0	0	0.9549249	0.3291139	0	
GPR111	ENST00000327753	0	0	0	0	0	1.6638655	0	0	0	0	0.4620462	2.5232558	0	0.1068447	0.2025316	
GPR111	ENST00000371220	0.225	0	0	0	0	0	0.2289157	0	0	0	0	0	0	0	0	
GPR111	ENST00000467205	0	0	0	0	0	0	0.1084337	0	0.505618	0	0.1320132	0	0	0.6076795	0	
GPR112	ENST00000287534	0	0.0334728	0	0.0402685	0	0	0	0	0	0.043545879	0	0.1395349	0	0	0	
GPR112	ENST00000394141	0	0	0	0	0.0503145	0	0	0	0.0898876	0	0.1188119	0	0	0.0934891	0	
GPR112	ENST00000394143	1.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
GPR112	ENST00000412101	0.6875	0	0	0	0	0	0	0	0	0	0.1980198	0.3255814	0	0.293823	0.0506329	
GPR113	ENST00000333478	0	0	0	0	0	0	0.2771084	0	0.0449438	0.037325039	0	0	0	0	0.4188482	
GPR113	ENST00000435303	0	0.0251046	0	0.0268456	0	0.0336134	0.060241	0	0	0	0	0	0	0	0	
GPR113	ENST00000447444	0	0.1171548	0.075	0	0	0.1008403	0.1204819	0	0	0.00622084	0	0	0	0.0868114	0.1012658	
GPR114	ENST00000340339	0	1.7656904	0	1.7315436	0	0	0	0	0	2.33281493	2.3894389	1.6627907	0	3.5125209	0	
GPR114	ENST00000349457	0	0	0.35	0	0	0	4.1204819	1.7575758	0	0	0	0	0	0	0	
GPR114	ENST00000394361	0	0	0	0	0	0	0	0	0	0	1.1353135	0	0	0	0	
GPR116	ENST00000265417	114.775	37.874477	0	82.33557	59.069182	341.78151	155.18072	208.42424	195.70787	26.7496112	19.234323	30.069767	0	37.315526	60.392405	
GPR116	ENST00000362015	241.3	33.76569	1.75	65.691275	25.031447	488.11765	76.60241	58.939394	154.97753	14.17729393	9.7161716	27.581395	193.25	7.0984975	44.822785	
GPR116	ENST00000456426	17.825	17.456067	3.5	41.261745	30.08805	621.21008	2.8313253	12.757576	1.7191011	1.038880249	2.4422442	0	655.5	0	25.556962	
GPR119	ENST00000276218	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0200334	0.0126582	
GPR12	ENST00000381436	0.15	0	4.65	0	0	0	0	0.2121212	0	0	0.330033	0.0930233	0	0.9348915	0	
GPR12	ENST00000405846	0	0	19.2	2.6040268	0.7044025	8.8067227	0	0	0	0	0.0792079	0	0	0	0	
GPR124	ENST00000315215	1.4375	96.217573	3.5	0.0134228	10.993711	21.142857	36.144578	46.606061	38.640449	50.18351477	0.1716172	186.24419	71.875	28.520868	4.6582278	
GPR124	ENST00000412232	0	0	2.8	0	0	0	0.7590361	0	0	0	0	0	0	0	0	
GPR125	ENST00000334304	14.3	0.6276151	4.875	15.395973	12.08805	40.285714	4.9518072	6.9393939	3.3707865	1.611197512	10.943894	6.1395349	453.5	5.6093489	13.544304	
GPR126	ENST00000230173	0	0	0	0	5.1572327	37.697479	0	52.848485	0	0	0	0	0	0	6.2278481	
GPR126	ENST00000296932	5.325	3.2217573	0.25	2.8456376	0	0	5.4096386	109.48485	0	3.931570762	0.9372937	0.8488372	0	3.5525876	0	
GPR126	ENST00000367608	5.375	1.0962343	0	2.1610738	35.572327	7.8151261	15.891566	748.39394	70.651685	1.866251944	3.1683168	47.093023	0	15.465776	1.6835443	
GPR126	ENST00000367609	0	0	0	2.6040268	10.327044	0	0	0	12.258427	0	0.5280528	9.0813953	35.875	0	0	0.3560209
GPR126	ENST00000415128	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0126582	0	
GPR128	ENST00000273352	0	0	0	0	0.754717	0	0	97.878788	0	0	0	0	0	0	0	
GPR132	ENST00000329797	12.9375	26.276151	1.45	0	3.509434	0	0	0	8.9775281	7.601866252	4.0132013	4.4069767	0	1.4023372	4.1265823	
GPR133	ENST00000261654	8.425	0.5104603	5.15	10	9.6981132	36.168067	4.5783133	15.848485	8.3707865	1.835147745	17.359736	8.872093	37.25	4.2804674	9.4177215	
GPR135	ENST00000395116	8.4875	4.5188285	7.35	14.228188	9.672956	6.2689076	10.048193	7.1818182	7.1460674	0.690513219	2.310231	32.44186	93.875	10.310518	13.392405	
GPR135	ENST00000481661	3.7375	0	0	0	0	0	3.939759	0	0	0	0	0	0	0	0	
GPR141	ENST00000343425	0.6375	0.0669456	0.3	0.9127517	0	0	0.0963855	1.8484848	0.4494382	0	0.2376238	0.0348837	0	0.6811352	0	
GPR141	ENST00000447769	0	0	0	0	0	1.1596639	0	0	0.5617978	0	0	0	0	0	21.905759	
GPR141	ENST00000450180	0.0125	0	0	0.1208054	0	0	0.626506	0	0.5280899	0.111975117	0	0	0	0	6.5968586	
GPR141	ENST00000467202	0.175	0.1924686	0	0	0	0	0	1.6969697	0	0	0	0	0	0	2.1151832	
GPR142	ENST00000335666	0.9375	3.4895397	0	0	0.6163522	0	0	0	0	1.237947123	26.20462	0.8488372	0	0	8.0886076	
GPR144	ENST00000334810	0	0	0	0.0939597	0	0	0	0	0	0	0	0	0	0	0	
GPR146	ENST00000297468	63	149.39749	34.925	270.65772	92.767296	84.789916	40.301205	1501.8182	223.39326	93.12597201	28.066007	137.69767	754.75	16.253756	6.3037975	
GPR146	ENST00000397095	0	58.117155	0	137.04698	0	75.344538	0	0	0	49.8911353	6.7062706	0	360.5	0	0	
GPR146	ENST00000444847	26.275	0.4100418	0.125	0	4.7421384	0	15.433735	0	5.3483146	0.34836703	0	0	175.875	0.7145242	0	
GPR146	ENST00000474396	4.525	5.2552301	0.35	4.2416107	1.4591195	2.3697479	0.2891566	1.030303	0.7078652	0.447900467	0.8448845	0.9883721	4	1.0283806	0.6329114	
GPR146	ENST00000481403	2.7375	0.4853556	0.05	3.3825503	2.2515723	13.932773	2.1445783	3.2121212	0.0786517	1.088646967	0.2772277	1.5465116	10.25	0.2070117	0.6962025	
GPR149	ENST00000389740	0	0	0.5	0	0	0	0	0	0.3258427	0	0	0.5116279	0	0	0	
GPR15	ENST00000284311	0.05	0.1338912	0.1	0.0402685	0.0503145	0	0	0	0.2247191	0.043545879	0.0264026	0.0116279	0	0	0	1.6544503
GPR150	ENST00000380007	0.25	0.0167364	0.2	0.0268456	0.0377358	0.0336134	0	0.0606061	0.0337079	0.018662519	0.039604	0.0697674	0.25	0.0133556	0.0253165	
GPR151	ENST00000311104	0	0	0.075	0	0	0	0	0	0	0	0	0	0	0	0	
GPR152	ENST00000312457	0.0125	0.1841004	0	0	0	0	0	0	0.011236	0.00622084	0	0	0	0.0066778	0.0126582	
GPR153	ENST00000377893	1.025	0.3096234	0.525	0.5234899	0.4528302	0.1512605	0.1686747	2.7575758	0.988764	0.34836703	1.8349835	0.3488372	1.375	0.2337229	0.4177215	
GPR156	ENST00000315843	0	0	0	0	0	0	0	0	0	0	5.3465347	0	0	3.8130217	0	
GPR156	ENST00000461057	0	0.6610879	0.525	0	0	0.3193277	0.5783133	0	0	0	0	0	0	0.0886076	0	

HGNC Symbol	ensemblTransID	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymph node	Ovary	Prostate fetal muscle	Testes	Thyroid	White blood	
GPR156	ENST00000464295	0.075	0.3430962	0.275	0	0.1006289	0.0504202	0.2650602	0	0	0	2.0066007	2.2093023	0	3.5459099	0	0
GPR156	ENST00000481907	0	0.1171548	0.65	0	0	0	0	0	0.4494382	0	0	0	0	0.7011686	0.1898734	0
GPR156	ENST00000495912	0.225	0.125523	0.05	0.5100671	1.5597484	0	0	0.1235955	0	0.1320132	4.3139535	0	0.9482471	1.0126582	0	
GPR157	ENST00000377411	4.4625	1.8912134	0.325	5.8657718	3.0188679	14.302521	2.5421687	16.030303	8.2921348	2.195956454	7.669967	3.3604651	258.375	2.6911519	14.316456	6.408377
GPR157	ENST00000414642	0	0	0	0	0	0	0	0.7878788	0.2247191	0.466562986	0	0.1395349	0	0	0	0
GPR157	ENST00000466131	0.0625	0	0.375	0.2013423	0	0.3865546	0.0240964	0.3636364	0.0561798	0.118195956	0.2508251	0	1	0.293823	0.3417722	0
GPR158	ENST00000376351	0	0	15.2	0	2.427673	0	0	0	0.752809	0.839813375	0.3432343	0.7790698	0	0.6944908	0.3670886	0
GPR160	ENST00000355897	83.225	26.217573	22.525	20.832215	169.58491	23.764706	172.37349	84.424242	23.202247	41.80404355	10.547855	95.232558	71.125	40.126878	32.329114	191.32984
GPR160	ENST00000473675	4.6	0.292887	0	0	0	0	3.5180723	0	32.955056	0	0.1188119	0.5581395	0	0	0	0.9842932
GPR160	ENST00000482710	0.6125	0	2.65	0	0	0	9.8674699	0	0	0	0.0792079	0.7790698	0	0	0	1.4450262
GPR160	ENST00000492492	0.775	0.5439331	0	0.6577181	0.6540881	0	0.3614458	0.3636364	0	1.872472784	1.5577558	0	9	1.3489149	0.2531646	0
GPR161	ENST00000271357	0	0	0.75	1.5973154	1.7106918	0	0.9518072	0.8787879	1.7977528	0	0	0	3	0	1.8860759	0
GPR161	ENST00000361697	0	0	0	0	0	0	0	0	0.258427	0	0	0	0	0	0	0
GPR161	ENST00000367835	0	0	0	0	0.163522	0.1176471	0	0	0.2696629	0.516329705	0.9636964	0	0	0.3806344	0	0
GPR161	ENST00000367836	1.6125	0.0251046	0	0.1610738	0.591195	0	0	0	0.186625194	0.1452145	0.1395349	0	0.1669449	0	0	0
GPR161	ENST00000367838	3.2125	1.1380753	0	0	0	0.4033613	0	0	0	0	0.1848185	1.4651163	0	3.278798	0	0
GPR161	ENST00000493800	0	0.2175732	3.375	4.3489933	0	0.23529412	0.0361446	0.2121212	0	0.566096423	0.2508251	0.6744186	0	0	0	0
GPR162	ENST00000311268	1.325	7.0460251	77.875	0	5.5471698	6.0504202	3.4096386	1.5757576	13.730337	2.283048212	21.676568	2.5813953	0.5	1.3021703	2.9620253	3.4554974
GPR162	ENST00000382315	0	0	0	3.6778523	4.327044	0	0	0	0	0	0.3537958	0.4302326	0	1.5759599	2.3670886	36.020942
GPR162	ENST00000428545	1.675	7.7656904	52.3	6.3758389	23.345912	2.6218487	3.1445783	10.393939	31.033708	2.967340591	14.138614	2.5697674	0	6.1168614	1.7848101	117.96859
GPR17	ENST00000272644	0	0	0.175	0	0	1.8823529	0	0.5757576	0.2921348	0.01244168	0.039604	0.1976744	0	0	0.2278481	0
GPR17	ENST00000393018	2.7625	1.0041841	0	3.5167785	0	1.5798319	0	2.1818182	0	0.230171073	0	0.6627907	0	0.3472454	1.9367089	0
GPR17	ENST00000423019	0	0	0	0	0	0	0	0	0	0	0	0	0	0.033389	0	0
GPR17	ENST00000486700	7.025	0.6778243	1.425	0.5637584	0.8176101	0.2521008	0.0240964	0.7272727	0	0	0	0	0	0.0467446	0	0
GPR17	ENST00000496086	0.6125	2.2259414	0	0	3.4339623	1.1596639	1.4216867	0.5151515	1.1685393	0	0.6732673	0.1976744	0	0.687813	0	0
GPR171	ENST00000309180	1.0125	7.497908	0.15	1.6912752	1.5471698	1.7983193	1.3253012	3.5757576	4.9101124	11.11041991	0.2376238	0.6046512	3.25	0.0400668	1.0379747	60.251309
GPR171	ENST00000480322	0	1.58159	0	0	0	1.394958	0	0	0	0	2.1518152	0	0	0	0	1.9895288
GPR173	ENST0000032582	1.2375	3.958159	13.975	0	0	0.21680672	0	0	0.8764045	0	5.9933993	0	0	0	0	0
GPR173	ENST00000375466	4.075	2.9539749	13.95	1.704698	1.8490566	1.4621849	2.3373494	0.4848485	0.7752809	1.648522551	21.940594	2.7093023	0	2.0701169	1.8101266	0.2513089
GPR174	ENST00000276077	0	0.2761506	0	0.0134228	0	0.0168067	0	0.0909091	0.3146067	1.536547434	0.0528053	0.0116279	0.25	0.0133556	0.0126582	4.0837696
GPR176	ENST00000299092	3.3625	2.0920502	7.275	2.3221477	2	0.6890756	1.7951807	1.3030303	2.3595506	1.741835148	1.0033003	1.2093023	5.375	1.4824708	1.2911392	0.0837696
GPR179	ENST00000342292	0	0	1	0	0	0	0	0	0.0898876	0	0	0.6046512	15.375	1.1686144	0.5822785	0.1256545
GPR179	ENST00000494542	0	0	0	0	0	0	0	0	0	0	0	0.4651163	0	0	0	0
GPR18	ENST00000340807	0	4.2845188	0	0	0	0	0	0	0	0	0	0	0	5.9899833	0	0
GPR18	ENST00000397470	0.65	5.4728033	0	2.2818792	3.0062893	0	3.8192771	0	2.9550562	24.82115086	0	0.5348837	11.5	0.1335559	0.1392405	101.08901
GPR18	ENST00000397473	0	8.6192469	0	0	0	1.7478992	0	1.1212121	0	0	2.310231	0	0	19.739566	0	0
GPR18	ENST00000416594	0	0	0	0	0	0	0	2.8484848	0	0.958009331	0.3168317	0	0	0	0	52.460733
GPR182	ENST00000300988	0.125	1.2970711	0.325	0	0.0503145	0.0840336	0.1807229	0.2575756	0.1910112	0.143079316	0.039604	0.0348837	1.5	0.0400668	0.1518987	0.2094241
GPR183	ENST00000376414	3.175	13.447699	1.8	2.8053691	6.490566	4.0504202	2.3855422	4.1515152	5.9213483	5.972006221	2.7590759	5.5232558	21.25	1.3288815	0.7468354	43.47644
GPR19	ENST00000332427	0.85	1.5062762	12.65	3.5704698	9.8867925	1.4789916	1.746988	37.090909	2.2359551	5.144635426	3.5379538	6.4767442	42.125	2.0367279	4.8227848	13.465969
GPR20	ENST00000377741	0.05	0.0585774	0.15	0.0536913	0.2389937	0.3529412	0.2650602	0.7272727	0.2921348	0.205287714	0.3828383	0.0930233	0	0.1535893	0.0632911	1.0052356
GPR21	ENST00000373642	2.025	0.2259414	3.025	0.3624161	3.1194969	0.4201681	0.20963855	0	0.2494382	0.242612753	0	1.7906977	0	0	1.9240506	0.3979058
GPR22	ENST00000304402	0.325	0.167364	27.05	0.0805369	0	66.655462	0.2168675	0.5757576	0	0	0.0132013	0.0697674	0	0.0066778	0.0632911	0
GPR22	ENST00000496754	0	0	37.2	0.2416107	0.3144654	185.29412	3.4457831	0	0	1.287713841	0	0	0	0.1402337	0.0379747	0
GPR25	ENST00000304244	0	0.0083682	0	0	0	0	0	0.011236	0.00622084	0	0	0	0.0066778	0	0.1256545	
GPR26	ENST00000284674	0	0.0753138	1.1	0	0.1257862	0	0	0	0.00622084	0	0.0116279	0	0.1268781	0	0	
GPR27	ENST00000304411	0.7375	0.251046	8	1.2080537	0.9811321	0.8235294	1	0.2121212	0.2359551	0.01244168	1.8481848	0.9302326	1.125	1.7829716	3.2531646	9.1099476
GPR3	ENST00000374024	0.3125	0.041841	0.25	0.0402685	0.0125786	0	0.0240964	0.030303	0.2022472	0.024883359	0.039604	0.0232558	0.375	0.0133556	0.0759494	0
GPR31	ENST00000366834	0	0	0.05	0	0	0	0	0	0	0	0.0132013	0	0	0.066778	0	0
GPR34	ENST00000378138	0	0	0	0	0	0	0	25.096386	45.363636	1.8202247	7.079315708	0	5.244186	21.625	0	11.582278
GPR34	ENST00000378142	31.775	18.661088	29.575	74.33557	3.7358491	27.529412	3.5421687	20.30303	2.3370787	1.685847589	6.0594059	1.4186047	0	12.948247	4.1392405	38.094241
GPR35	ENST00000319838	4.875	0.5188285	0.35	1.6778523	1.8742138	0.4201681	0.2168675	0.1515152	0	0.734059098	0.8580858	0.1627907	0	0.1202003	1.3670886	42.282723
GPR35	ENST00000403859	0	0	0	0	0	0	0	0.8313253	0.3939394	0	0	0	0	0.3472454	0	0
GPR35	ENST00000407714	0.2875	0.3179916	0.125	0	1.0440252	0	0	0	0	0.093312597	0.7128713	0	0	0	0	5.7382199

HGNC Symbol	ensemblTransID	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymph node	Ovary	Prostate latal muscle	Testes	Thyroid	White blood		
GPR35	ENST00000430267	0.2625	0.1589958	0	0.2013423	2.0125786	0	0	0	0.3370787	0.080870918	0.8052805	0.1162791	10.625	0.4006678	0.556962	1.1308901	
GPR35	ENST00000438013	0	0	0	0	0	0	0	0	0.2921348	0	0	0.0116279	0	0	0	0	
GPR37	ENST00000303921	0.0625	0.041841	33.625	0.1342282	0.1509434	0.1344538	0.1084337	6.8787879	0.0224719	0.087091757	0.0264026	0.3023256	1	1.4290484	0.0506329	0	
GPR37L1	ENST00000367282	0.025	0.0334728	6.725	0.0134228	0.0125786	0.0168067	0	0	0.0224719	0.099533437	0.0264026	0.0465116	0.125	0.0200334	0.0759494	0	
GPR39	ENST00000329321	0	0	0.1	1.1409396	0.2641509	0	2.1927711	12.848485	1.3258427	0	0.290429	0.9534884	0	0.2604341	0.6582278	0	
GPR4	ENST00000323040	10.175	2.1087866	5.75	1.1812081	6	5.4957983	1.4819277	1.6969697	18.617978	2.152410575	0.7260726	1.0232558	18.125	0.4607679	1.278481	0	
GPR45	ENST00000258456	0	0.0585774	0.15	0.1073826	0.0377358	0.0336134	0.1084337	0.0606061	0	0	0.0264026	0.1627907	0	0.033389	0.0253165	0.104712	
GPR50	ENST00000218316	0.0375	0	0	0.0805369	0	0	0.1325301	0	0	0	0	0	0	0.066778	0	0	
GPR52	ENST00000367685	0	0.0083682	1	0	0.0125786	0.2857143	0	0	0	0	0	0	0	0	0.0253165	0	
GPR55	ENST00000392039	0	0	0	0	0	0	0	0	0	0.32970451	0	0	0	0.5275459	0.0886076	5.0052356	
GPR55	ENST00000392040	0	0.9790795	0.075	0.1879195	0.0503145	0	0	0.1212121	0.2808989	0.609642302	0.0264026	0	0	1.7629382	0.1265823	16.565445	
GPR55	ENST00000438398	0.0375	0	0.125	0	0	0.0336134	0	0	0.5168539	11.80093313	0.039604	0	0	0	1.7591623	0	
GPR55	ENST00000444078	0.3375	0	0.3	0	0.3018868	0	0.0361446	0	0.0337079	0	0	0.1162791	0	0	0	0	
GPR6	ENST00000275169	0	0	6.45	0	0	0	0	0	0	0	0	0	0	0	0	0	
GPR61	ENST00000404129	0.0125	0	0.475	0	0.0503145	0.0336134	0	0	0	0.01244168	0	0	0.25	0	0.721519	0	
GPR61	ENST00000469383	0	0	0.25	0.1073826	0	0	0.0722892	0.030303	0.0224719	0	0.0792079	0.0581395	0	0	0	0.0628272	0
GPR62	ENST00000322241	0.05	0.0920502	1.2	0.1208054	0.0754717	5.7478992	1.3373494	0.3636364	0.1460674	1.580093313	4.290429	2.872093	1.25	1.7495826	2.3924051	2.0314136	
GPR63	ENST00000229955	1.2375	0	1.225	0	1.1572327	0	0	0.8787879	0	1.940902022	1.5841584	0	0	7.345576	0.8987342	0	
GPR63	ENST00000417980	0	0	0	0.3087248	0	0	0	0	0	0	0	0	2.255814	5.875	0	0	
GPR64	ENST00000340581	0	0	0.225	0	1.3836478	0.4201681	0.0963855	0	0	0	0	0	0.9186047	12.625	0	0	
GPR64	ENST00000354791	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
GPR64	ENST00000356606	19.3875	1.8828452	0.175	2.6845638	0	0	0	0	0	2.836702955	3.7887789	1.244186	0	55.171953	1.1139241	0	
GPR64	ENST00000357544	3.375	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
GPR64	ENST00000357991	0	0	0	0	0	0	0.5210084	0.4216867	0	0.7078652	0	0	0	0	0	0	
GPR64	ENST00000379873	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
GPR65	ENST00000267549	32.9875	112.27615	10.05	51.543624	17.647799	28.521008	60.554217	51.181818	95.629213	68.34214619	56.765677	29.127907	25.625	11.325543	7.3797468	3001.9058	
GPR68	ENST00000238699	0.1625	3.6066946	1.775	0.3489933	0.3144654	0.3361345	0.7228916	1	2.8202247	0.821150855	0.8052805	1.1162791	2.375	0.6143573	0.8481013	6.5759162	
GPR75	ENST00000352846	0	0	0	0	0	0	0	4.3975904	0	0	0.00622084	0	0	0	0	0	
GPR75	ENST00000394705	1.875	0.4853556	5.5	0.9798658	0.2012579	1.5630252	0.5060241	1.5757576	0.1011236	0.584758942	0.4752475	0.3372093	0.5	0.3138564	1.0759494	2.1989529	
GPR75	ENST00000406625	3.025	0.5523013	1.025	0.590604	0	0.6722689	0	0	0.7640449	1.293934681	0	0	0	0	0	0	
GPR75	ENST00000459916	0	0	0	0	0	0.5714286	0.6987952	0	1.3595506	0	0	0.4767442	0	0	0	0	
GPR75	ENST00000498475	0	0	2.45	0	0	0	0	0	6.9438202	0	4.5412541	0.2209302	6.75	1.836394	0	5.0890052	
GPR78	ENST00000382487	0	0	0.075	0	0	0	0	0	0	0	0	0	0	0	0	0	
GPR82	ENST00000302548	1.0375	1.3305439	0.5	0.9798658	0	0.302521	1.8192771	8.7272727	0.3033708	0.715396579	0.3828383	0.255814	0	0.3205342	0.7341772	5.0680628	
GPR83	ENST00000243673	0.15	0.0585774	0.975	0.1342282	0.0377358	0.0168067	0.1445783	0.2424242	0.0449438	0.031104199	0.0792079	0.0813953	0.5	0.0601002	0.1772152	0	
GPR84	ENST00000267015	0.1	0.0502092	0.425	0.1208054	0	0.1176471	0.1084337	0.2727273	0.1685393	0.093312597	0.1584158	0.1046512	0	0.0934891	0.1012658	0.6492147	
GPR85	ENST00000297146	3.2625	0	0	0	0	0	0	0	0	0	0	0	0	0	0.1898734	0	
GPR85	ENST00000424100	0.1375	1.2970711	9.425	2.5637584	1.0062893	2.1680672	0.5662651	1.969697	2.3033708	2.133748056	1.8877888	1.1860465	0	3.8330551	0	0.3350785	
GPR85	ENST00000438062	0	3.0962343	0	0	0	0	0	0	0.6363636	0	0.4884488	5.4302326	0	0	0	0	
GPR85	ENST00000449591	0	0.9874477	2.375	0	0	3.0084034	5.2771084	10.818182	9.3370787	0	0	0	0	0.9949917	0	0	
GPR85	ENST00000449735	0	0	3.125	0	2.7924528	0	0	0	0	0	0	0.4534884	0	0	0.1772152	0	
GPR85	ENST00000487573	0	0.0167364	0	0	0	0	0	0	0	0	0	0	0	1.081803	0	0	
GPR87	ENST00000260843	0.0125	0.0083682	0.75	0.2147651	0	0	0.0481928	0.030303	0.2022472	0	0.0264026	1.0581395	0	0.066778	0.5443038	0.0209424	
GPR88	ENST00000315033	0.325	0.6945607	3.125	0.1342282	0.7672956	0.6218487	0	12	0.1573034	0.553654743	7.2211221	0.9186047	4.5	0.4741235	0.8734177	1.2146597	
GPR97	ENST00000327655	0	0	5.525	0	0	0	0	0.1566265	0	0	0	0	0	0	0	3.1623037	
GPR97	ENST0000033493	23.95	2.8451883	1.4	0	8.4150943	16.537815	0	0	11.280899	2.382581649	6.8646865	2.5232558	123.125	0	11.227848	37.842932	0
GPR98	ENST00000405460	0	0	0.425	1.261745	0.2012579	0	0	0	0.6292135	0	0.3960396	1.4534884	0	0	0	13	0
GPRC5A	ENST00000014914	2.05	0.1757322	0.3	0.885906	1.5849057	0.3697479	0.8795181	0.1212121	32.179775	0.279937792	2.9966997	3.755814	4.625	0.4340568	8.1772152	0.1675393	
GPRC5B	ENST00000300571	22.9	13.76569	313.375	12.993289	9.6981132	23.747899	16.277108	1.3333333	3.5842697	8.827371695	15.933993	7.1860465	215.375	7.8530885	16.164557	0.5026178	
GPRC5C	ENST00000392627	3.7	0.7531381	0	6.4697987	2.3522013	4.0840336	21.879518	20.30303	2.0898876	2.618973561	5.3069307	11.290698	65.75	7.0851419	22.443038	0	
GPRC5C	ENST00000392629	0	4.4518828	0	0	0	0	0	0	0	0	49.914191	0	0	0	0	0	
GPRC5D	ENST00000228887	0	0	0	0	0	0	0	1.7575758	0	0	0	0	0	0	0	0	
GPRC5D	ENST00000396333	0.9625	0.7949791	0.05	2.0268456	0.2515723	0.9243697	0.5421687	0	3.4606742	1.486780715	0.2508251	2.3372093	0	2.1235392	3.9493671	1.9895288	
GPRC6A	ENST00000310357	0	0	0	0	0	0	0.0963855	0	0	0	0	0	0	0	0	0	

HGNC Symbol	ensemblTransID	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymph node	Ovary	Prostate fetal muscle	Testes	Thyroid	White blood	
GPRC6A	ENST00000368549	0	0	0.075	0	0	0	0.6144578	0	0	0	0.255814	0	0.1001669	0	0	
GRM1	ENST00000282753	0.0625	0.1338912	3.425	0	0	38.084034	2.3493976	0.3030303	0.0224719	0	0.2906977	0	0	0.4177215	0	
GRM1	ENST00000361719	0	0	0	0	0	0.2147651	0.0503145	0	0	0	0	0	0.6277129	0	0	
GRM1	ENST00000392299	0	0	0	0	0	0	0	0	0	0	0.4356436	0	0	0	0	
GRM2	ENST00000296479	0	0	0	0	0	0.327044	0	0	0	0	0	0	0	0	0	
GRM2	ENST00000395052	0	0	0	0	0	0	0	0	0	0	0.5016502	0.0813953	0	0	0	
GRM2	ENST00000419928	0	0	0	0	0	0	0	0	0	0	0.174183515	0.1848185	0.0581395	0	0.0759494	
GRM2	ENST00000442933	0	0.292887	1.425	0	0	0	0	0	0	0	0.039604	0	0	0	0	
GRM2	ENST00000464585	0.175	1.9330544	1.425	0.1610738	0.3522013	0	0.0843373	0	0.2134831	0.223950233	0.3036304	0.1744186	0	2.7646077	0.164557	0.7539267
GRM2	ENST00000475478	0	0	0	0	0	0.1610738	0	0	0.0843373	0	0.5505618	0	0.1320132	0	0	
GRM3	ENST00000361669	0	0.1841004	67.275	0.147651	0.0377358	0.1008403	0.1204819	0	0	0	0	0	1.375	0.2136895	0	0
GRM3	ENST00000394720	0.0625	0.0167364	0	0.0805369	0	0	0	0.3030303	0	0	0	0	0	0.0126582	0	0
GRM3	ENST00000441140	0	0	45.475	0	0	0	0	0	0	0	0	0.1320132	0	0	0	0
GRM4	ENST00000374177	0	0.0669456	0	0.2416107	0.2264151	0	0	0	0.0224719	0	0	0.0581395	0	1.509182	0	0
GRM4	ENST00000374181	0	0	0.45	0	0	0	0	0	0	0	0	0.2772277	0	0	0.5542571	0.0759494
GRM4	ENST00000455714	0	0.0167364	47.675	0	0	0	0	0	0	0	0	0	0.0116279	0	0.4808013	0
GRM5	ENST00000305432	0	0	0	0	0.1208054	0	0	0	0	0	0	0.0924092	0	0	0	0
GRM5	ENST00000393294	0	0	0	0	1.081761	1.0252101	0	0	0	0	0	0	0	0.1335559	0	0
GRM5	ENST00000418177	0	0	2.75	0	0	0	0	0	0	0	0	0	0	0	0	0
GRM5	ENST00000455756	0.0875	0	24.325	0.0536913	0.0880503	0.0336134	1.1084337	0	1.0898876	0	0	0.1511628	0	0.6611018	0.0379747	0
GRM6	ENST00000231188	0	0	0.225	0	0.0754717	0.0504202	0.2168675	0	0	0	0	0.5280528	0	0	0.3071786	0
GRM7	ENST00000357716	1.7875	0.6276151	0	0.1208054	0.0754717	0	0	0.2424242	0.5168539	0	1.3729373	0.2790698	0	0.3338898	0.1898734	0
GRM7	ENST00000389335	0	0	0.775	0	0	0	0	0	0	0	0	0	0	0	0	0
GRM7	ENST00000389336	0	0	0	0	0	0	0	0	0	0	0	0.1627907	0	0	0	0
GRM7	ENST00000403881	0	0	0	0	1.6603774	0	0	0	0	0	0	0	0	0	0	0
GRM7	ENST00000440923	0	0	0	0	0	0	0	0	0	0	0	0	0	2.4974958	0	0
GRM7	ENST00000467425	0	0	0	0	0	0	0	0	0	0	0	0.1976744	0	0.6477462	0	0
GRM8	ENST00000339582	0	0	4.55	0	0	0	0	0	2.5151515	0	0	0	0	0	0.427379	0
GRM8	ENST00000341617	0	0	0	0	0	0	0	0	26.878788	0.1797753	0	0.8316832	0	0	0.3672788	0
GRM8	ENST00000358373	0	0	1.55	0	0	0	0	0	0	0	0	0	0	0.9949917	0	0
GRM8	ENST00000405249	0	0	0	0	0	0	0	0	2.8915663	0	0	0	0	0	0.0886076	0
GRM8	ENST00000412160	0	0	0	0	0	0	19.831933	0.626506	3.3333333	0	0	0	0	0	0	0
GRM8	ENST00000444921	0	0	0	0	0	0	0	0.373494	0	0.1123596	0	0.4224422	0	0	0.4874791	0
GRM8	ENST00000457830	0	0	0	0	1.2327044	0	0	0	0	0	0	0.0924092	0	0	0	0
GRM8	ENST00000472701	0	0	0	0.0536913	0	0	0	0	0	0	0	0	0	0	0	0
GRPR	ENST00000380289	0.1125	0.2008368	0.525	0.1610738	0.3396226	0.2689076	0.1566265	0.4848485	0	0.074650078	0.1980198	0.3488372	3.875	0.1936561	0.1012658	0.6492147
HCAR1	ENST00000356987	11.125	5.5230126	0.4	30.375839	1.7106918	0.7226891	8.0120482	0.8787879	0.4157303	0.690513219	7.1815182	8.6511628	0	5.1953255	8.0759494	0.6492147
HCAR1	ENST00000432564	18.4125	0.7866109	0.15	98.993289	3.672956	0.1680672	4.5903614	0	0.0898876	0.796267496	2.9570957	4.9418605	1.125	1.4490818	2.6708861	0
HCAR1	ENST00000436083	0	0.041841	0.125	0	0	0.7058824	0	0	0.4157303	0	0	0	0	0	0	0
HCRTR1	ENST00000373705	0	0	0.075	0	0	0	0	0	0	0	0	0.0132013	0.0232558	0	0	0
HCRTR1	ENST00000373706	0	0	0.175	0	0	0.0168067	0.0240964	0	0	0	0	0	0.375	0.0133556	0	0
HCRTR1	ENST00000403528	0	0.2677824	0	0.0805369	0	0	0	0	0	0.080870918	0	0	0	0	0.4936709	0
HCRTR1	ENST00000468521	0	0	0	0	0	0	0	0	0	0	0	0	0.0232558	0	0	0.2531646
HCRTR1	ENST00000485464	0	0	0	0	0	0	0	0.0120482	0	0	0	0	0	0	0	0
HCRTR2	ENST00000370862	0.025	0.0251046	5.825	0.7651007	0.0125786	0.0336134	0.1445783	0.0606061	0.0224719	0.024883359	0.0264026	0.0232558	0.375	0.066778	0.3797468	0.0209424
HRH1	ENST00000397056	0	0	1.55	3.0469799	3.5471698	0	2.3253012	0	1.9662921	6.606531882	3.1023102	1.2325581	16.25	1.1886477	0	0
HRH1	ENST00000413416	0.175	0.0753138	0.825	0.0805369	0	0	0	0	0	0	0.1320132	0	0	0.0734558	0	0
HRH1	ENST00000431010	0	0	0	0	52.138365	2.789916	0	0	0	0	10.47589425	36.924092	0	78.375	0	0.7329843
HRH1	ENST00000438284	0	2.0334728	0	0	16.377358	2.4033613	1.1445783	5.1212121	11.393258	0	3.5643564	2.255814	0	0	0.9367089	1.1727749
HRH2	ENST00000231683	3.6875	0.5523013	2.475	0.3087248	0.2138365	1.1596639	0	0	2.7865169	1.468118196	0	0	0	0	0	7.5811518
HRH2	ENST00000377291	3.1125	0.7364017	0.25	2.5234899	0	51.529412	0	0	16.955056	1.754276827	0	0	0	0	0	15.706806
HRH3	ENST00000317393	0	0	0	0	0	0	0	0	0	0.32348367	0	0	0	0	0	0
HRH3	ENST000003040177	0	0	10.425	0	0	0	0	0	0	0	0.1848185	0.4302326	0	1.7963272	0	0
HRH4	ENST00000256906	0	0	0	0	3.0062893	0	1.3493976	0	0	0.248833593	0.990099	0	46.125	2.7646077	0	17.486911

HGNC Symbol	ensemblTransID	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymph node	Ovary	Prostate testal muscle	Testes	Thyroid	White blood	
HRH4	ENST00000426880	4.075	0.7531381	0	3.5973154	0	0.3361345	0	1.4242424	0.7865169	2.102643857	0	2.9534884	0	0.3472454	2.1772152	19.518325
HTR1A	ENST00000323865	0	0	0.325	0	0	0	0	0	0	0	0	0	0	0.0267112	0	0
HTR1B	ENST00000369947	0.025	0.0083682	0.15	0.0536913	0.0251572	0	0.0481928	0	0.011236	0.043545879	0	0.1046512	0.25	0.0400668	0.0126582	0
HTR1D	ENST00000314113	0	0	0	0	0.0251572	0	0	0	0	0.01244168	0	0.0232558	0	0.0066778	0	0
HTR1D	ENST00000374619	0.0125	0.0083682	0.7	0	0	0	0	0	0.1685393	0.055987558	0.3432343	0.0581395	0	0.2203673	0	0
HTR1E	ENST00000305344	0	0	0.375	0	0	0	0.1325301	0	0	0	0	0	0	0.7746244	0	0
HTR1E	ENST00000369584	0	0	1.925	0	0.2264151	0	0	0	0	0	14.79868	0.1627907	0	0.3071786	0	0
HTR1F	ENST00000319595	0.0625	0.0251046	0.275	0.0536913	0.0251572	0.0840336	0.1566265	0	0.0224719	0.055987558	0.039604	0.0581395	0.375	0.0801336	0.0506329	0.0418848
HTR2A	ENST00000378688	0.15	0.0669456	7.225	2.3355705	0.591195	0.3361345	0.3975904	0.5757576	0.1348315	0.136858476	0.2376238	0.3255814	10.25	0.3138564	0.2405063	0.7329843
HTR2B	ENST00000258400	4.5625	1.8912134	0.375	4.1073826	2.6163522	2.2857143	4.8915663	3.6666667	1.3033708	0.970451011	1.9009901	1.9767442	114.5	1.7495826	4.1265823	12.816754
HTR2C	ENST00000371950	0.175	0	10.85	0	0	0	0	0	0	0	0	0	0	0.2136895	0	0
HTR2C	ENST00000371951	0	0	0	0	0	0	0	0.2048193	0	0	0	0.0930233	0	0.0801336	0.0759494	0
HTR4	ENST00000314512	0	0	0.125	0	0	0	0	0	0	0.441679627	0	0	0	0	0.6962025	0
HTR4	ENST00000354217	0	0.334728	0	2.1886792	0.8067227	0	0	0	0.5168539	0	0	0	0	0	0	0
HTR4	ENST00000360693	0	0	0.4	0	0	0	0	0	0	0	0	0	0	0	0	0
HTR4	ENST00000362016	0	0	0	0	0	0.0672269	0	2.0909091	0	0	0	0.5	0	0.0734558	0.2025316	0
HTR4	ENST00000377888	0	0	0	0	0.8176101	0	0	0	0	0	0.2772277	0	0	0	0	0
HTR5A	ENST00000287907	0	0	21.65	0	0.163522	0	0	0.5757576	0	0	0	0	0	0.5542571	0	0
HTR6	ENST00000289753	0.0375	0.1338912	0.15	0.2416107	0.0754717	0	0.0120482	0	0	0.01244168	0.369637	0.0116279	0.625	0.0066778	0.0126582	0.1675393
HTR7	ENST00000277874	0.8	1.6485356	6.05	1.0067114	1.3584906	1.1092437	0.9036145	0.9393939	0.2134831	0.951788491	1.8745875	4.6162791	4.75	23.619366	3.3544304	6.4502618
HTR7	ENST00000336152	0	0	0	0	0	0	0	0	0	0	0	0	0	4.78798	0	0
HTR7	ENST00000371719	0	0	0.7	0	0	0	0	0.9393939	0	0	0	0	0	0.2404007	0	0
HTR7	ENST00000371721	0	0	0	0	0	0	0	0	0.4382022	0	0	0	0	0	0	0
KISS1R	ENST00000234371	0	0.0502092	0.3	0.0402685	0	0	0.0963855	0	0.0224719	0.043545879	0	0	0	0	0.3560209	0
LGR4	ENST00000389858	0	18.158996	0	116.30872	75.044025	138.4874	293.51807	493.24242	12.078652	11.94401244	30.693069	70.034884	189.125	53.582638	0	9.5706806
LGR5	ENST00000266674	2.3375	0.5606695	13.4	1.0067114	1.9874214	0	0.3373494	0.8484848	0.5280899	0.65318818	5.320132	0.255814	376.75	0.8414023	1.5063291	0
LGR6	ENST00000255432	0	0	0	0	0	0.5882353	0	0	0	0	0	0	0	0.9113924	0	0
LGR6	ENST00000367278	0	0	0.25	0	0	0	0.5662651	0	0.9550562	0.534992224	0	0.8255814	1.125	0	0.6202532	2.2827225
LGR6	ENST00000439764	0	0	0	0	0	0	0	0	0	0	0	0	0	0.9767442	0	0
LGR6	ENST00000487787	0.1625	0.5188285	0	0	0	1.3277311	0	0	0	0	0	1.650165	0	0	0.4340568	1.2911392
LHCGR	ENST00000294954	0	0	0	1.7449664	0	0	0	0	0	0	0.2772277	0	0	0	0	0
LHCGR	ENST00000344775	0.4625	0	0	8.3892617	0	0	0.0963855	0	0	0	0	0	0	0.9816361	0.0886076	0
LHCGR	ENST00000401907	0.8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LHCGR	ENST00000403273	0	0	0	0	0.1761006	0	0	0	0.0786517	0	0	0.127907	0	0	0.3670886	0
LHCGR	ENST00000405626	0	0	0	0	0	0.1566265	0	0	0	0	0.2772277	0	0	0	0	0
LHCGR	ENST00000428232	0.7375	0	0	3.033557	0.0503145	0.184874	0	0	0	0	0.0264026	0	0	0.2270451	0.1518987	0
LPAR1	ENST00000358883	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LPAR1	ENST00000374430	0	0	0	0	0	51.361345	0	0	0	0	0	0	0	160.75	0	0.89.193717
LPAR1	ENST00000374431	0	0	0	0	2.4528302	0	0	10.727273	13.730337	0	0	0	0	3.5392321	6.7594937	0
LPAR1	ENST00000441240	44.15	10.594142	49.175	8.6845638	23.006289	0	2.8915663	2	19.853933	3.564541213	4.3036304	8.7674419	49.625	15.405676	15.746835	0
LPAR2	ENST00000407877	0.1875	0.8284519	0.675	0.8187191	0.3144654	0.1008403	0.5060241	0.3333333	0.7640449	1.138413686	0.7392739	0.8953488	0	1.9432387	0.556962	2.0942408
LPAR3	ENST00000370611	0	0	1	0	0	0.8907563	0.0843373	0.0909091	0.1685393	0	0	0.1511628	0	0	0.2278481	0
LPAR3	ENST00000440886	0.0625	0.083682	1.95	0.1879195	0.0377358	10.369748	0.0240964	0	5.6179775	0.018662519	13.280528	3.372093	0	9.6293823	0.0632911	0.1884817
LPAR3	ENST00000491034	0	0.0251046	0.075	0.0134228	0.0251572	0	0	0.0606061	1.7078652	0.105754277	0	0	0	0	0	0
LPAR5	ENST00000329858	2.7625	6.0167364	6.35	5.1946309	3.1069182	23.07563	1.8433735	15.363636	3.2022472	8.292379471	2.6930693	1.1627907	0	0	11.810127	16.335079
LPAR5	ENST00000431922	0	19.514644	0	0	8.6289308	0	3.4457831	0	0	6.482115086	4.3960396	4.2790698	0	8.2003339	0	79.790576
LPAR6	ENST00000345941	170.7375	0	6.825	226.73826	58.226415	0	84.180723	32.272727	0	75.92534992	8.9636964	31.337209	0	24.267112	36.632911	101.82199
LPAR6	ENST00000378434	0.9	5.1213389	0	0	0	65.411765	0	0	0	0	0	31.5	7.2721202	0	0	0
LPAR6	ENST00000465365	0	0	0	3.7718121	0	0.7394958	0	0	0	0	1.7425743	0	0	3.7662771	0	0
LPHN1	ENST00000340736	17.675	1.8828452	0	18.630872	0	0	11.674699	8.1515152	13.247191	5.642301711	0	3.4883721	3.5	0	0	25.089005
LPHN1	ENST00000361434	28.0375	42.251046	87.925	96.416107	62.742138	72.319328	31.73494	0.4545455	8.6404494	8.783825816	79.260726	31.139535	3.125	55.232053	133.35443	11.497382
LPHN2	ENST00000271029	0	0	0	0	0	0	0	0.3939394	0	0	0	0	0	0	0	0
LPHN2	ENST00000319517	0	0	0.15	0	0	0	0	0.8484848	0	0	0	1.5841584	0	0	0	0
LPHN2	ENST00000335786	0	0	0	0	0	0	0	0.030303	1.1685393	0	0	0	0	0	0	0

HGNC Symbol	ensemblTransID	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymph node	Ovary	Prostate latal muscle	Testes	Thyroid	White blood		
LPHN2	ENST00000359929	0	0.41841	0	0.6442953	1.7610063	1.8487395	4.373494	0	57.696629	5.953343701	0	9.7674419	2.375	0	8.5063291	0	
LPHN2	ENST00000370713	15.1875	11.221757	0	5.0738255	8.8427673	31.764706	9.0481928	8	53.865169	3.129082426	17.425743	0	18	6.6310518	15.544304	0	
LPHN2	ENST00000370727	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3.0253165	0	
LPHN2	ENST00000370730	0	0.0920502	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
LPHN2	ENST00000402328	0.7875	5.1631799	0	0	4.918239	5.9327731	5.0481928	0.6666667	0	0.398133748	0.8844884	2.5813953	0	1.2954925	0	0	
LPHN2	ENST00000449420	6.9625	0	0.1	0.2550336	6.1132075	0	1	0	0	0.043545879	6.3366337	0	0	3.1853088	15.620253	0	
LPHN2	ENST00000469377	0	0	0.05	0	0	1.0084034	0	0	0	0	0.5412541	0	0	0.3138564	0	0	
LPHN2	ENST00000473719	0	0.3682008	0	0	0	0.6554622	0	0	0.2022472	0	0	0	0	0	0	0	
LPHN2	ENST00000492666	0	0	0.075	0	0.0503145	0.9579832	0.0963855	0.0909091	0.7640449	0	0.1188119	0	0	0	0	0	
LTB4R	ENST00000345363	79.3625	16.60251	0	71.315436	4.9559748	3.9663866	29.120482	35.272727	17.269663	3.421461897	65.570957	12.162791	35.875	0	72.278481	89.13089	
LTB4R	ENST00000396782	3.75	1.1046025	1.575	0.5771812	2.9056604	0.5546218	5.8313253	4.7575758	1.8988764	2.923794712	2.2706271	0.3139535	0	4.0801336	1.6455696	245.50785	
LTB4R	ENST00000396789	24.3125	16.058577	1.05	8.4161074	1.8993711	3.5294118	15.385542	356.57576	16.573034	2.849144635	12.792079	12.197674	24.875	2.3105175	5.3291139	89.612565	
MAS1	ENST00000252660	0	0	0.275	0	0	0	0	0	0	0	0	0	0	0.0066778	0	0	
MAS1L	ENST00000377127	0	0.0083682	0	0	0	0.0168067	0.0240964	0	0.011236	0	0	0	0.0348837	0	0	0	
MC2R	ENST00000327606	0.2375	0.0920502	0.075	0	0	0.3193277	0	0	0.4382022	0.130637636	0	0.2093023	0	0.721202	0	0	
MC2R	ENST00000399821	0	0	0	0	0	0	0	0	0	0	0	0	0	4.3873122	0	0	
MC4R	ENST00000299766	0	0.1087866	0.725	0	0	0	0.0481928	0.0909091	0	0.124416796	0.0132013	0	0	0.0467446	0	0	
MC5R	ENST00000324750	0.1125	0	0	0.0402685	0	0	0.0120482	0	0	0	0.039604	0	0	0	0.0126582	0	
MCHR1	ENST00000249016	0	2.7531381	26.925	0	1.8742138	1.6302521	0.6144578	155.51515	0.7191011	1.449455677	16.554455	7.7674419	0	1.1085142	0.164557	0	
MCHR1	ENST00000498400	0	0	0	0.1342282	0	0	0	0	0.0561798	0	0	0	0	0	0	0	
MCHR2	ENST00000281806	0	0	1.9	0	0	0	0	0	0	0	0	0	0	8.5141903	0	0	
MCHR2	ENST00000369212	0	0	0	0	0	0	0	0	0	0	0	0	0.372093	0	0	0	
MCHR2	ENST00000445970	0	0	0.025	0.2684564	0	0	0	0.2727273	0	0	0	0	0	0	0	0	
MLN R	ENST00000218721	0	0	0	0	0	0.907563	0	0	0.0674157	0	0.0264026	0	2.25	0.5742905	0	0	
MLNR	ENST00000398307	0	0	0	0	0	0	0	0	0	0	0	0	0	3.5443038	0	0	
MRGPRD	ENST00000309106	0	0	0	0	0	0	0	0	0	0	0.0132013	0	0	0	0	0	
MRGPRE	ENST00000389832	0.75	0.3514644	0	0	2.1886792	1.1092437	0	0	0	0.516329705	0.7788779	1.2209302	0	0	0	3.1265823	0
MRGPRE	ENST00000436689	0	0	0	0	0	0	0	0	0	0	0	0	0	1.0016694	0	0	
MRGPFR	ENST00000309099	33.25	226.7364	0	2.295302	318.7044	0	35.277108	14.363636	72.955056	49.13219285	245.61056	397.30233	0	132.1202	32.101266	0	
MRGPFR	ENST00000320913	0	0	0.1	0	0	0	0	0	0	0	0	0	136.125	0	0	0	
MRGPFR	ENST00000441623	107.375	0	0.85	114.48322	27.672956	16.302521	19.253012	0	3.9775281	60.04354588	134.91749	43.965116	302	42.864775	27.392405	0	
MRGPRG	ENST00000323214	0	0	0	0	0	0	0	0	0	0	0	0	0	0.066778	0	0	
MRGPRX1	ENST00000302797	0	0	0	0	0	0	0	0	0	0	0	0	0	0.1335559	0	0	
MRGPRX2	ENST00000329773	0	0.0167364	0	0	0	0.0168067	0	0	0	0.01244168	0	0.0232558	0	0.033389	0.0379747	0.104712	
MRGPRX3	ENST00000396275	0.1875	0.1171548	0	0	0.5786164	0	0.0963855	0	0	0	0	0.0930233	0	1.803005	0.1392405	0	
MRGPRX4	ENST00000314254	0	0	0	0.0134228	0	0	0	0	0	0	0	0	0	0	0	0	
MTNR1A	ENST00000307161	0.0125	0.0251046	0.05	0.0134228	0.0503145	0	0.1927711	0.030303	0	0.00622084	0	0	0	0.0734558	0.0126582	0	
MTNR1B	ENST00000257068	0	0	0.05	0	0	0.0168067	0	0	0	0	0	0	0	0	0	0	
NMBR	ENST00000258042	0	0	0.25	0	0.2893082	0	0	0	0	0.4620462	0.0813953	0	1.115192	0.278481	0	0	
NMR1	ENST00000305141	0.325	0.2343096	0.025	0.8590604	0.0377358	0.1680672	0.0481928	0.3030303	0.2022472	0.06842938	0.0528053	0.1976744	0.5	0.1268781	0.0759494	1.3403141	
NMR2	ENST00000255262	0.075	0	0.625	0.1342282	0.0125786	0.1008403	0.0963855	0	0.0449438	0.00622084	0.0264026	0.0813953	0	0.1936561	0.1898734	0	
NPBWR1	ENST0000031251	0.0625	0	0	0.025	0	0	0	0.1325301	0.1818182	0	0	0	0.0348837	0	0	0	
NPBWR2	ENST00000369768	0	0	0.025	0	0	0	0	0	0	0	0	0	0	0.033389	0	0	
NPFFR1	ENST00000277942	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.2405063	0	
NPFFR2	ENST00000344413	0	0	0	0	0	0.487395	0.6746988	0	0.1123596	0	0.9504951	0	0	0	1.4556962	0	
NPFFR2	ENST00000358749	0	0	0	0	0.6289308	0	0	0	0	0	0	0	3.2325581	0	2.130217	0.9367089	
NPFFR2	ENST00000395999	0	0	0	0	0	0	0	0	0	0	0	0	0	2.4641068	0	0	
NPSR1	ENST00000360581	0	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	
NPY1R	ENST00000296533	4.25	3.9079498	1.375	35.275168	0.3773585	0.3361345	5.0963855	0.6666667	0.5280899	2.961119751	0.6468647	0.8837209	7	0.9015025	0.8860759	0	
NPY2R	ENST00000329476	0	0	2.675	8.295302	11.333333	0	0	0.3939394	0	3.135303266	0.1584158	1.627907	0	21.84975	0.2658228	0	
NPY4R	ENST00000374312	0	0	0	2.3624161	5.4339623	0	0	0	0	0	0	0	1.7093023	0	0.0534224	1.3291139	0
NPY4R	ENST00000395716	2.4625	0.6527197	0.15	0	13.698113	3.4957983	0	1.7878788	0.4382022	0.516329705	1.6237624	0	0	0	0	0	0
NPY5R	ENST00000338566	0.4375	0.167364	1.025	2.5234899	0.2893082	0.0672269	0.9036145	0.0606061	0.3483146	0.205287714	0.1056106	0.1395349	0	0.6210351	0.1265823	0	
NTSR1	ENST00000370501	1.5	0	0.525	0	54.90566	0	0	0	0	0.790046656	0	1.0232558	0	0.6210351	0	3.9790576	

HGNC Symbol	ensemblTransID	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymph node	Ovary	Prostate latal muscle	Testes	Thyroid	White blood	
NTSR2	ENST00000306928	0	0	27.3	0	0	0.1008403	0	0	0	0.037325039	3.4191419	0.0465116	8	0.2470785	0.721519	0
OPN3	ENST0000031838	0	0	0	0	0	2.6554622	0.6385542	50.454545	0	0	3.0231023	0	0	1.2621035	0	15.141361
OPN3	ENST00000366554	1.9	1.3723849	4.45	3.8657718	1.8742138	1.7142857	3.1807229	23.939394	7.4606742	6.115085537	8.990099	1.1627907	6.5	1.8430718	8.9113924	30.743456
OPN3	ENST00000469376	5.3875	0.0167364	0	0	2.1383648	0	0	0	0.4494382	2.979782271	1.9273927	0.8255814	0	2.4373957	3.7088608	0
OPN3	ENST00000490673	3.1125	1.1631799	42.775	6.7114094	4.2138365	1.2605042	9.8915663	15.515152	8.0224719	7.919129082	0.3168317	2.5465116	3	7.3722871	4.6455696	22.701571
OPN5	ENST00000371211	0	0	0	0	0	0	0	0	0	0	0	0	0	0.4674457	0	0
OPN5	ENST00000489301	0	0	0	0	0	0	0	0	0	0	0	0	0	0.1135225	0.4683544	0
OPRD1	ENST00000234961	0.025	0.0251046	0.225	0.0268456	0.2389937	0.0168067	0.0240964	0.030303	0.0224719	0.024883359	0.0264026	0.0348837	1.375	0.3405676	0.0379747	0.0418848
OPRL1	ENST00000265572	0	0	7.3	0	0	0	0	0	0	0	0	0	0	3.4924875	0.8987342	0
OPRL1	ENST00000336866	1.8875	4.7949791	0.375	0	0	0	0	0	0	1.835147745	0	0	0	0	33.947644	0
OPRL1	ENST00000349451	0	0	0.25	0	0	0	0	0	0	0	9.4521452	0	0	5.7028381	0	0
OPRL1	ENST00000355631	0	0	0	6.2550336	0	0	0	0	0	0	0	0	0	9.4490818	0	0
OPRM1	ENST00000229768	0	0	0	0	0	0	0	0	0	0	0	0	0	0.6677796	0	0
OPRM1	ENST00000330432	0	0	0.35	0	0	0	0	0	0	0	0	0	0	0	0	0
OPRM1	ENST00000370749	0	0	0	0	0	0	0	0	0.2247191	0	0.2508251	0	0	0	5.1518987	0.104712
OPRM1	ENST00000434900	0	0	0	0	0.0754717	0	0	0	0	0.093312597	0	0	0	1.5893155	0	0
OXER1	ENST00000378661	0.0875	0.0753138	0.075	0.0402685	0.2264151	0.0336134	0.2168675	2.5454545	0.1011236	0.037325039	0.3828383	0.1511628	1.5	0.1202003	0.1392405	0.5026178
OXGR1	ENST00000298440	0.15	0.1087866	0.7	0.590604	0.2012579	0.1008403	1.4578313	2.121212	0.258427	0.11975117	0.2244224	0.3255814	0	0.2203673	0.1518987	0
OXTR	ENST00000316793	0.4375	1.0376569	3.8	81.516779	0.9685535	0.3697479	3.3493976	1	5.3033708	0.734059098	13.465347	1.7209302	2.5	1.6026711	3.0759494	1.026178
OXTR	ENST00000431493	0.05	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OXTR	ENST00000449615	0	0	0	0.9127517	0	0	0	0	0	0	1.5577558	0	0	0	0	0
OXTR	ENST00000474615	0	0	3.375	0.4832215	0	2.0336134	0	0	0	0	0	0	0	0	0	0
P2RY1	ENST00000305097	0.6625	0.1589958	4.5	0.8993289	3.7106918	3.1764706	0	0.7878788	4.3033708	0	5.2805281	1.8953488	0	0.9616027	0.6202532	1.2146597
P2RY10	ENST00000171757	0	2.6192469	0	0	0	0	0	0	1.0561798	7.944012442	0	0	0	0	0	27.979058
P2RY10	ENST00000461541	0	1.2217573	0	0	0	0.0336134	0	0.6363636	0	0	0	0	0	0	0	0
P2RY10	ENST00000475374	0.0125	1.6736402	0	0	0	0	0	0	2.3033708	0	0	0.0348837	0	0.033389	0	7.3926702
P2RY11	ENST00000321826	158.225	156.97071	29.425	95.463087	54.339623	28.184874	83.385542	166.63636	23.719101	71.52099533	53.927393	57.232558	883.75	92.707846	106.20253	200.08377
P2RY13	ENST00000325602	6.375	27.39749	4.8	10.442953	0	9.6806723	4.3253012	105.36364	10.044944	10.24572317	9.7821782	3.9418605	8	2.7111853	2.164557	1270.8272
P2RY14	ENST00000309170	12.6875	5.4895397	4.6	28.308725	91.119497	41.865546	3.7831325	1.7272727	49.988764	5.972006221	2.4950495	69.127907	142.125	3.9866444	13.708861	44.146597
P2RY14	ENST00000424796	0	0	0	4.1208054	3.6226415	3.0756303	0	1.1573034	0	0	1	0	0	5.556962	0	
P2RY14	ENST00000494668	0	0	0	0	0	0	0	0	0	0	0	0	13.25	0	0	0
P2RY2	ENST00000311131	0	0	0	0	0	0	0.3865546	0	0	0.080870918	2.9306931	0	0	0	3.2405063	0
P2RY2	ENST00000393596	2.3625	2.3849372	1.2	0	0	6.9579832	2.5060241	0	0	0	6.6976744	408	2.4507513	42.468354	0	
P2RY2	ENST00000393597	3.6	1.497908	0	11.718121	0.8553459	48.773109	2.939759	41.727273	18.988764	3.732503888	9.9405941	0.8255814	526.125	0	123	26.95288
P2RY4	ENST00000374519	0	0.125523	0	0.0939597	0	0	0	0	0	0	0.0264026	0.0813953	0	0	0.0379747	0
P2RY6	ENST00000349767	0	7.3472803	0	1.9463087	0	8.2521008	1.686747	0	0	0	0	0.9069767	0	0	0	0
P2RY6	ENST00000393590	4.8375	6.4435146	0.9	0.10469799	0.0754717	3.2436975	7.7590361	1.8181818	1.8764045	3.415241058	1.4389439	3.627907	0	1.442404	4.4556962	0.6701571
P2RY6	ENST00000393591	0	0	0	0	1.8238994	0	1.7951807	0	0	0	0	0	0	0	0	0
P2RY6	ENST00000393592	0	0	0	3.0201342	0	0	0	0	0.3146067	9.598755832	11.009901	0	22.5	3.7929883	2.1772152	5.0052356
P2RY8	ENST00000381297	8.1875	34.845188	0.325	6.8993289	2.2893082	5.394958	3.4337349	7.2121212	5.247191	25.60497667	4.5412541	1.2906977	37.375	0.8681135	8.8101266	332.18848
PRLHR	ENST00000369169	0	0	0.275	0	0	0	0	0	0	0	0	0.3255814	0	0	0	0
PROKR1	ENST00000303786	1.7625	0	0	0.1744966	0.2389937	0	0	0	0	0	0	0	0	0	0.0837696	0
PROKR1	ENST00000394342	0	0.0334728	0	0.6577181	0.0503145	0.2352941	0	0	0	0.049766719	0.0132013	0.244186	0	0.2871452	0	0
PROKR2	ENST00000217270	0.075	0.1087866	0.525	0.6442953	0.163522	0.8907563	0.1807229	1.0909091	0.2134831	0.043545879	0.2508251	0.1395349	0	0.1736227	0.0506329	0.0837696
PTAFR	ENST00000305392	0.25	0.041841	0	0	0	0	0	0	0	0	0.127907	0.625	0	0	0	0
PTAFR	ENST00000373857	12.7375	4.7029289	0.975	8.2684564	1.7106918	2.2016807	1.8674699	13.878788	6.3033708	4.049766719	6.0858086	2.0348837	1.875	1.3222037	2.164557	395.91623
PTGDR	ENST00000306051	0.1375	0.5355649	0.325	0.6040268	0.2264151	0.2184874	0.373494	0.0909091	1	0.410575428	0.8052805	0.7906977	2.625	0.0534224	0.1265823	1.0052356
PTGER1	ENST00000292513	0.025	0.1841004	0.075	0.0134228	0.0628931	0.0336134	1.1325301	0.0606061	0.1011236	0.087091757	0.039604	0.0930233	0.125	0.0133556	0.0126582	0
PTGER2	ENST00000245457	3.3875	1.0209205	1.1	0.6711409	1.1320755	0.4537815	1.6987952	0.6666667	2.1123596	1.076205288	0.620462	1.1860465	1.125	0.1202003	0.1772152	9.5078534
PTGER3	ENST00000306666	15.2	1.1129707	0.025	79.973154	1.1572327	1.0588235	61.879518	0.6363636	0.0224719	0.715396579	2.2442244	1.5697674	2.375	2.4106845	0.3417722	0
PTGER3	ENST00000351052	0	0	5.325	0	0	0	1.5662651	0	0	0	0	0	0	0	0	0
PTGER3	ENST00000356595	0.1	0	0	0	0	0	0	0	0	0	0.0232558	0	0	0	0	0
PTGER3	ENST00000361210	0	0	0	0	0.1383648	0	0	0	0	0	0.1320132	0	0	0	0	0
PTGER3	ENST00000370924	52.2375	0.5857741	0.35	53.637584	7.672956	1.8487395	52.433735	0	1.8539326	2.233281493	3.5511551	8.2325581	0.5	3.5325543	2.7974684	0.0837696

HGNC Symbol	ensemblTransID	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymph node	Ovary	Prostate latal muscle	Testes	Thyroid	White blood		
PTGER3	ENST00000370931	0.6	0	0	0	0.1257862	0.0840336	0	0	0	0	0.244186	0	0	0.0632911	0		
PTGER3	ENST00000370932	0	0	0	0	0	0	0	0	0	0.298600311	0	0.0581395	0	0	0		
PTGER3	ENST00000414819	0	0	0	0	0	0	0.0722892	0	0	0	0	0	0	0	0		
PTGER3	ENST00000460330	0	0.0167364	0	0	0	0	0	0	0	0	0	0	0	0	0.4188482		
PTGER3	ENST00000479353	0.6625	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
PTGER4	ENST00000302472	0.6	2.8702929	0.2	4.3892617	2.6918239	0.9747899	2.7349398	0.3939394	3.258427	1.493001555	0.3564356	1.1046512	3.5	0.5141903	0.8481013	10.450262	
PTGFR	ENST00000370757	0.5875	2.0251046	0	5.4899329	0.4779874	1.0084034	5.0481928	0	0	2.494556765	8.4488449	6.4651163	0	0	1.0379747	0.3769634	
PTGFR	ENST00000370758	22.0625	0	0.425	0	4.9811321	0	3.0120482	4.3333333	4.7640449	1.710730949	18.719472	13.569767	18.625	3.8130217	1.4050633	0	
PTGFR	ENST00000497923	0	0	0	0.9530201	0	0	0	0	0	0	0.0528053	0	0	0	0.7120419		
PTGIR	ENST00000291294	0.7375	2.3514644	0.775	0.2684564	0.7295597	0.3697479	0.8554217	1.2121212	9.6516854	1.045101089	0.2508251	0.2325581	4.75	0.0868114	0.7721519	1.8219895	
PTH1R	ENST00000313049	2.0125	2.1087866	0	2.7785235	0	0	0	0	0	0	0	0	0	0.8213689	2.9873418	0	
PTH1R	ENST00000418619	0	0.5020921	0	0	1.3207547	0	0	0	0	0	0	0	0	0	0	0	
PTH1R	ENST00000422115	0	0	0.925	1.5302013	0	0	29.084337	9.6060606	1.7303371	0.192846034	1.9537954	2.9186047	0.25	0.4073456	0.6835443	0	
PTH1R	ENST00000427125	0.275	0.3849372	0	0.6308725	0.6792453	0	24.963855	4.7878788	1.4382022	0.279937792	0	1.8837209	3.75	0.4674457	0	0	
PTH1R	ENST00000428220	0	0	2.05	0	0	2.789916	193.83133	11.333333	4.3595506	1.474339036	0	2.9651163	7.375	0	0	0	
PTH1R	ENST00000430002	0	0	0.8	0	0	0	12.987952	0	0.6516854	1.00155521	5.7557756	0	0	0	0	0	
PTH2R	ENST00000272847	1.15	0	0.775	4.3221477	0.0880503	0	0.3253012	0.1212121	0	0	0.0132013	0.0465116	0	0.1268781	0	0	
QRFP1	ENST00000394427	0	0	1.475	0	0	0	0	0	0	0	0	0	0	0.4140234	0	0	
RXFP1	ENST00000307765	0	0	9.225	0	0	0	0	2.6969697	0	0	0	0	0	0	0	0	
RXFP1	ENST00000343542	0	0	0	0	0.7421384	1.0252101	0	0	0	0	0.01244168	0	0	0	0	0	
RXFP1	ENST00000423548	0	0.3933054	0	0	0	0	0	0	0	0	0.4092409	0	0	0	0.1012658	0	
RXFP1	ENST00000448688	0	0.3682008	0	0	0	0.487395	11.855422	0	11.808989	1.101088647	6.7458746	0.8372093	0	4.2470785	0	0	
RXFP1	ENST00000470033	0	0.460251	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
RXFP1	ENST00000471616	0	0	0	0	0	0	0	0	0	0	0.6976744	0	0	0	0	0	
RXFP2	ENST00000298386	0	0	0	0	0	0	0.313253	0	0	0.373250389	0	0	0	2.263773	0	5.0052356	
RXFP2	ENST00000380314	0	0.1506276	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
RXFP3	ENST00000330120	0	0	0	0	0.0125786	0	0	0	0	0.018662519	0	0	0	0	0	0	
RXFP4	ENST00000368318	0	0.0167364	0	0	0.0125786	0	0	0	0	0	0	0	0	0	0	0	
S1PR1	ENST00000305352	421.6625	80.016736	34.725	100.55034	113.2956	96.907563	23.903614	128.87879	272.52809	52.41679627	29.742574	34.581395	266.25	21.409015	86.860759	378.57592	
S1PR3	ENST00000358157	45.425	92.702929	35.95	13.436242	81.823899	399.2437	471.93976	115.30303	83.449438	79.65163297	46.983498	107.83721	309.875	55.439065	43.227848	199.20419	
S1PR3	ENST00000375846	0.425	3.1966527	0.05	0.5369128	0.3773585	0.9747899	1.8674699	0	0.741573	1.387247278	0.9504951	0.4767442	6	0.1936561	0.9493671	0.7120419	
S1PR4	ENST00000246115	0.475	2.1171548	0	0.4295302	0.1132075	0.0840336	0.1807229	0.3636364	0.07078652	0.958009331	0.1056106	0.1976744	0.75	0.033389	0.2531646	26.450262	
S1PR5	ENST00000334340	2.0625	22.552301	11.5	3.2080537	1.572327	0	2.686747	5.5757576	10.483146	12.06842924	6.7194719	12.023256	0	5.2888147	0	773.3822	
SCTR	ENST0000019103	0.7	0	0.125	0	0	0.1176471	2.1325301	0	0.3146067	0	1.0693069	0	0	0.1602671	0.3797468	0	
SMO	ENST00000249373	5.5625	4.167364	4.6	11.90604	6.327044	5.8319328	18.843373	70.30303	3.2358427	5.275272162	36.831683	18.197674	12	12.414023	7.0253165	0.1465969	
SSTR1	ENST00000267377	8.8375	0	12.35	0	0	0	0	35.33333	2.3146067	0	5.2013201	6.6046512	0	0	3.164557	0	
SSTR2	ENST00000315332	82.875	11.422594	4.375	30.147651	0	13.932773	0	0	0.4719101	38.69362364	5.8613861	0	53	0	0	0	0
SSTR2	ENST00000357585	67.9875	13.129707	9.125	10.107383	19.056604	11.294118	45.349398	18.757576	9.2359551	9.493001555	8.4752475	6.8604651	91.25	9.1686144	11.202532	0	
SSTR3	ENST00000328544	0	3.0041841	24.2	0.1879195	0	0	0	0	0.5730337	2.581648523	94.574257	0.0930233	0	15.712855	0.1139241	22.34555	
SSTR3	ENST00000402501	0.0125	0	1.675	0.0268456	0.1886792	0	0	0	0	0.0924092	0	0	0	0	0	0	
SSTR4	ENST00000255008	0	0	0.125	0	0	0	0	0.1011236	0	0	0	0	0	0.2270451	0.443038	0	
SSTR5	ENST00000293897	0	0	0.025	0	0	19.94958	5.8915663	0	0	0.503888025	0	0	0	0	0	0	
SSTR5	ENST00000397547	0	0	0.325	0	0	132.47059	1.5060241	0	0	0	4.744186	30.875	0.9415693	0.2025316	0	0	
TACR1	ENST00000305249	29.7875	2.3263598	0.3	4.3221477	1.1698113	0.4537815	0.4096386	0	0.6966292	0.33592535	0.4356436	1.7325581	4	0.5008347	0.1265823	0	
TACR1	ENST00000409848	1.25	1.1548117	0	0	3.2578616	0	0.3975904	0	0	0	0.9302326	0	0	0.1772152	0	0	
TACR2	ENST00000373306	0	0.5439331	0	1.2483221	289.16981	0.6554622	0.373494	0.5151515	0.3033708	0	6.0066007	132.82558	0	2.0033389	0	0	
TACR2	ENST00000373307	7.1375	3.0292887	2.1	3.6107383	59.220126	137.68067	0	0	1.741573	0.765163297	2.3234323	33.546512	160.375	1.1018364	42.43038	1.1727749	
TACR3	ENST00000304883	0.025	0.0502092	0.4	0.0536913	0.0251572	0.1176471	0.1204819	0.030303	0.1910112	0.01244168	0.039604	0.2674419	0.125	0.0868114	0.2025316	0.0209424	
TAS1R1	ENST00000333172	0	0	0.05	0	0	0	0.0361446	0	0	0	0.0792079	0	0	0.9215359	0.0759494	0	
TBXA2R	ENST00000375190	3.55	29.564854	2.15	4.6845638	41.622642	18.672269	44.746988	333.87879	14.842697	24.52877138	50.653465	3.9651163	7.5	28.313856	5.4810127	6.7015707	
TBXA2R	ENST00000411851	0	0	0	9.0872483	0	0	0	0	19.146067	0	49.240924	42.953488	0	0	4.0126582	2.8691099	0
TRHR	ENST00000311762	0.0625	0.0753138	0.25	0.0671141	0	0.2016807	0.0481928	0.0909091	0	0	0.039604	0.0348837	0	0	0.0126582	0.0209424	0
TSHR	ENST00000298171	2.4875	0.2343096	0.325	0.885906	0	0.8571429	1	3.1818182	0	0.466562986	0.7656766	0.2325581	0	0.1335559	2053.3924	1.7801047	
TSHR	ENST00000342443	0	0.1338912	0	1.1006711	0	0	0	1.8181818	0	0.510108865	3.0363036	0	0	0	76.924051	0	0

HGNC Symbol	ensemblTransID	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymph node	Ovary	Prostate	Ilet muscle	Testes	Thyroid	White blood
VIPR1	ENST00000325123	2.575	2.7782427	0	0	3.5849057	1.2941176	3.7349398	151	50.988764	7.838258165	0	8.2209302	0	2.5843072	4.3797468	0
VIPR1	ENST00000433647	0	0	0	2.1342282	0	3.2941176	0	0	0	0	0	0	0	0	0	0
VIPR1	ENST00000436487	0	0.2175732	0	0	0	0	0	0	0	0	0	0	0	0	0	0
VIPR1	ENST00000438259	11.2375	4.4769874	5.15	12.791946	6.3144654	0	0	150.36364	78.932584	0	0.4488449	4.9883721	22	0.6076795	21.164557	5.8429319
VIPR1	ENST00000439910	0	0	0	0	0	0	0	0	0	0	0.0792079	0	0	0	0	0
VIPR1	ENST00000443646	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.9685864
VIPR1	ENST00000446673	0	0	0	0	0	0	0	0	0	0	0.3168317	0	0	0	0	0
VIPR1	ENST00000487545	0	0	0	0	0	0.0481928	0	0	0	0	0.0348837	0	0	0	0	0
VIPR2	ENST00000262178	0.6	1.748954	2.625	1.1006711	4.7924528	13.563025	0.3975904	0.4242424	0.4831461	3.950233281	9.4785479	0.1046512	3.625	2.851419	0.8481013	1.2565445
VIPR2	ENST00000377633	0	0	0	0	0.7924528	3.8823529	1.746988	0	0	0	0	0	2.125	0	1	0
VIPR2	ENST00000402066	0.025	0	0	0	0	0	0	0	1.2247191	0.609642302	0	0	0	0.2337229	0.2531646	0
XCR1	ENST00000309285	0.2375	1.3221757	0	1.1543624	0	0.5882353	0.6987952	1.7878788	1.4044944	9.206842924	0.7392739	0.1162791	1.25	0.0934891	0.3417722	2.1151832

Appendix_Table 3.2: Transcription factor expression across the 16 tissues

TF_Identifier	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymph	Ovary	Prostate	Skeletal Muscle	Testes	Thyroid	White blood
HOXB3	Adipose	Adrenal		0 Breast	Colon		0 Kidney	0 Lung	Lymphnode	Ovary	Prostate		0 Testes	Thyroid	White blood	0
HOXB4	Adipose	Adrenal		0 Breast	Colon		0 Kidney	0 Lung	Lymphnode	Ovary	Prostate		0 Testes	Thyroid	White blood	0
HOXB5	Adipose	Adrenal		0 Breast	Colon		0 Kidney	Liver	Lung	Lymphnode	Ovary	Prostate		0 Testes	0	0
HOXB6	Adipose	Adrenal		0 Breast	Colon		0 Kidney	0 Lung	Lymphnode	Ovary	Prostate		0 Testes	0	0	0
HOXB8	Adipose		0	0 Breast	Colon		0 Kidney	0	0	0	0	0 Prostate		0 Testes	0	0
HOXB9	Adipose	0 Adrenal		0	0 Colon		0 Kidney	0	0	0	0	0 Prostate		0 Testes	0	0
HOXC10	Adipose	Adrenal		0 Breast	Colon		0 Kidney	0	0 Lymphnode	Ovary	Prostate		0 Testes	0	0	0
HOXC11	Adipose	Adrenal		0	0	0	0	0	0	0	0	0 Skeletal Muscle		0	0	0
HOXC12	Adipose		0	0	0	0	0	0	0	0	0	0 Testes		0	0	0
HOXC13	Adipose	0 Adrenal		0	0	0	0	0	0	0	0	0 Testes		0	0	0
HOXC4	Adipose	Adrenal		0 Breast	Colon		0 Kidney	0	0 Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	0 White blood		
HOXC5	Adipose	Adrenal		0 Breast	0		0 Kidney	0	0 Lymphnode	Ovary	0	0 Testes		0	0	0
HOXC6	Adipose	Adrenal		0 Breast	Colon		0 Kidney	0	0 Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	0	0	0
HOXC8	Adipose	Adrenal		0 Breast	0		0 Kidney	0	0 Lymphnode	Ovary	0	Skeletal Muscle	Testes	0	0	0
HOXC9	Adipose	Adrenal		0 Breast	Colon		0 Kidney	0	0 Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	0	0	0
HOXD1	Adipose		0	0	0 Colon		0 Kidney	0 Lung	Lymphnode	Ovary	Prostate	0 Testes	0	0	0	0
HOXD10	Adipose	0 Adrenal		0	0 Colon		0 Kidney	0	0 Lymphnode		0 Prostate	0 Testes		0	0	0
HOXD11	0	0	0	0	0 Colon		0 Kidney	0	0	0	0 Prostate	0	0	0	0	0
HOXD13	0	0	0	0	0 Colon		0	0	0 Lymphnode	Ovary	0 Prostate	0	0	0	0	0
HOXD3	Adipose	Adrenal		0 Breast	Colon		0 Kidney	0	0 Lymphnode	Ovary	Prostate	0 Testes		0	0	0
HOXD8	Adipose	Adrenal		0 Breast	Colon		0 Kidney	0 Lung	Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	Thyroid	White blood	
HSF1	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	Thyroid	White blood
HSF2	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	Thyroid	White blood
INSM1	0	0 Brain	0	0	0	0	0	0	0	0	0	0	0 Testes	0	0	0
IRF1	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	Thyroid	White blood
IRF2	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	Thyroid	White blood
IRF3	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	Thyroid	White blood
IRF4	Adipose	Adrenal		0 Breast	0	0	0	0 Lung	Lymphnode	Ovary	0	0 Testes	0	0 White blood		
IRF5	Adipose	Adrenal	Brain	Breast	0	0 Kidney	Liver	Lung	Lymphnode	Ovary	Prostate	0 Testes	0	0 White blood		
IRF6	Adipose	Adrenal		0 Breast	Colon	Heart	Kidney	Liver	Lung	Lymphnode	Ovary	Prostate	0 Testes	Thyroid	0	0
IRF7	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	Thyroid	White blood
IRX2	Adipose		0	0 Breast	0	0 Heart	Kidney	0 Lung	Lymphnode	Ovary	0 Prostate	0 Testes	Thyroid	0	0	0
IRX3	Adipose	Adrenal		0 Breast	0	0 Heart	Kidney	0 Lung	Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	Thyroid	0	0
IRX4	0	0	0 Breast	0	0 Heart	0	0	0 Lung	Lymphnode	Ovary	0	0 Testes	0	0	0	0
IRX5	0	0	0 Breast	0	0 Heart	0	0	0 Lung	Lymphnode	Ovary	0	0 Testes	0	0	0	0
IRX6	Adipose		0	0 Breast	0	0 Heart	Kidney	0	0	0	0	0 Skeletal Muscle	0	0	0	0
ISL1	0	0	0	0	0	0	0	0	0	0	0	0 Prostate	0 Testes	0	0	0
ISL2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0 White blood
ISX	0	0	0	0	0 Colon	0	0	0	0	0	0	0	0	0	0	0
JARID2	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	Thyroid	White blood
JUND	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	Thyroid	White blood
KAISO	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	Thyroid	White blood
KLF6	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	Thyroid	White blood
KLF7	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	Thyroid	White blood
LBX2	Adipose	Adrenal		0 Breast	Colon		0 Kidney	Liver	0 Lung	Lymphnode	Ovary	0 Prostate	0 Testes	Thyroid	White blood	
LEF1	Adipose	Adrenal	Brain	Breast	Colon		0 Kidney	0 Lung	Lymphnode	Ovary	Prostate	0	0 Testes	Thyroid	White blood	
LHX1	0	0	0	0	0	0	0 Kidney	0 Lung	0	0	0	0	0	0	0	0
LHX2	0	0	0 Brain	0	0	0	0 Kidney	0 Liver	0	0	0	0	0	0	0	0
LHX4	0	0	0	0	0	0	0 Kidney	0	0	0 Lymphnode	Ovary	0	0	0	0	0
LHX6	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	0 Lung	Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	Thyroid	0	0
LHX8	0	0	0	0	0	0	0 Kidney	0	0	0	0	0	0	0	0	0
LHX9	0	0	0	0	0	0	0 Kidney	0	0	0	0 Ovary	0	0	0	0	0
LMX1A	0	0	0	0	0	0	0 Kidney	0	0	0	0	0	0	0	0	0
LMX1B	0	0	0	0 Breast	0	0 Kidney	0	0	0	0	0	0	0	0	0	0
MAFA	0	0	0	0	0	0	0 Kidney	0	0	0	0	0 Skeletal Muscle	0	0	0	0
MAFB	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	Thyroid	White blood
MAFK	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	Thyroid	White blood
MAX	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	Thyroid	White blood
MAZ	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	Thyroid	White blood
MEF2C	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	Thyroid	White blood
MEIS1	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	Thyroid	White blood
MEOX1	Adipose	Adrenal		0 Breast	Colon	Heart	0	0	0 Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	Thyroid	White blood	
MXL1	0	0 Adrenal		0	0	0	0	0	0 Lymphnode	Ovary	0 Prostate	Skeletal Muscle	Testes	Thyroid	0	0
MSX1	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	0 Lung	Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	Thyroid	0	0
MSX2	0	0	0 Breast	0	0 Heart	0	0	0 Lung	Lymphnode	Ovary	0 Prostate	Skeletal Muscle	Testes	Thyroid	0	0
MYBL1	0	0 Adrenal	Brain	Breast	Colon	Heart	0	0 Lung	Lymphnode	Ovary	0 Prostate	Skeletal Muscle	Testes	Thyroid	0	0 White blood
MYF5	0	0	0	0	0	0	0 Kidney	0	0	0	0	0 Skeletal Muscle	0	0	0	0
MYF6	0	0	0	0 Breast	0	0	0	0	0	0	0	0 Skeletal Muscle	Testes	Thyroid	0	0
MYT1	0	0	0 Brain	0	0	0	0 Kidney	0	0	0	0	0	0	0	0	0
MYT1L	0	0	0 Brain	0	0	0	0 Kidney	0	0	0	0	0	0	0	0	0
MZF1	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	Thyroid	White blood
NEUROD1	0	0	0 Brain	0	0	0 Kidney	0	0 Lung	Lymphnode	Ovary	0 Prostate	Skeletal Muscle	Testes	Thyroid	0	0
NF1	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	Thyroid	White blood
NFAT5	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	Thyroid	White blood
NFE2	Adipose		0	0	0	0	0 Kidney	0 Lung	Lymphnode	Ovary	0 Prostate	0	0 Thyroid	White blood		
NFKAPPA-B	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	Thyroid	White blood
NKX22	0	0	0 Brain	0	0	0	0 Kidney	0	0	0	0	0	0	0	0	0
NKX23	0	0	0	0	0	0	0 Colon	0	0	0	0	0	0	0	0	0
NKX24	0	0	0	0	0	0	0 Kidney	0	0	0	0	0	0	0	0	0
NKK31	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	Thyroid	White blood
NRF1	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	Thyroid	White blood
NRL	0	0	0	0	0	0	0 Kidney	0 Liver	0	0	0	0	0	0	0	0
OSR1	Adipose	Adrenal		0 Breast	Colon	Heart	Kidney	0 Lung	Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	Thyroid	0	0
OSR2	Adipose	Adrenal		0 Breast	Colon	Heart	Kidney	0 Lung	Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	Thyroid	0	0
OTX2	0	0 Brain	0	0	0	0	0 Kidney	0	0	0	0	0	0	0	0	0
OVL1	0	0	0	0	0	0	0 Kidney	0	0	0	0	0	0	0	0	0
PAX1	0	0	0	0	0	0	0 Kidney	0	0	0	0	0	0	0	0	0
PAX2	0	0	0	0	0	0	0 Kidney	0	0	0	0	0	0	0	0	0
PAX3	0	0	0	0	0	0	0 Kidney	0	0	0	0	0	0	0	0	0
PAX5	0	0 Adrenal	0	0	0	0	0 Kidney	0	0	0	0 Lymphnode	0	0	0	0	0 White blood
PAX6	0	0	0 Brain	0	0	0	0 Kidney	0	0	0	0 Ovary	0	0	0	0	0
PAX8	0	0	0	0	0	0	0 Kidney	0	0	0	0 Ovary	0	0	0	0	0
PAX9	0	0	0	0	0	0	0 Kidney	0	0	0	0 Lung	0	0	0	0	0
PBX1	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymphnode	Ovary	Prostate	Skeletal Muscle	Testes	Thyroid	White blood
PBX3	Adipose	Adrenal	Brain	Breast	Colon	Heart	Kidney	Liver	Lung	Lymphnode	Ovary</td					

Appendix Table 3.3: Shows the tissues in which isoforms are expressed

Tissues of Expression	Isoform ID (ensemblTransID)	# of Isoforms
Adipose	ENST00000320658, ENST00000357544, ENST00000362072, ENST00000394143, ENST00000401907, ENST00000431493, ENST00000448020, ENST00000461244, ENST00000479353	9
Adipose, Adrenal, Brain, Breast, Colon, Heart, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Skeletal muscle, Testes, Thyroid	ENST00000243673, ENST00000262441, ENST00000263281, ENST00000292513, ENST00000296533, ENST00000303921, ENST00000306666, ENST00000323040, ENST00000330055, ENST00000334304, ENST00000337585, ENST00000362015, ENST00000370742, ENST00000402260, ENST00000474396, ENST00000481403	16
Adipose, Adrenal, Brain, Breast, Colon, Heart, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Skeletal muscle, Testes, Thyroid	ENST00000368121	1
Adipose, Adrenal, Brain, Breast, Colon, Heart, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Skeletal muscle, Testes, Thyroid, White blood	ENST0000014914, ENST00000176183, ENST00000234961, ENST00000238699, ENST00000240093, ENST00000241356, ENST00000245457, ENST00000248076, ENST00000249373, ENST00000258400, ENST00000259455, ENST00000261654, ENST00000262178, ENST00000267549, ENST00000271332, ENST00000272928, ENST00000277874, ENST00000286201, ENST00000287934, ENST00000291294, ENST00000292174, ENST00000295417, ENST00000295683, ENST00000296140, ENST00000296641, ENST00000296677, ENST00000297468, ENST00000299092, ENST00000299178, ENST00000299727, ENST00000300571, ENST00000302472, ENST00000304166, ENST00000304222, ENST00000304411, ENST00000304748, ENST00000304883, ENST00000305141, ENST00000305352, ENST00000306005, ENST00000306051, ENST00000307637, ENST00000309170, ENST00000309180, ENST00000315215, ENST00000315325, ENST00000316793, ENST00000318507, ENST00000319211, ENST00000321826, ENST00000322241, ENST00000323289, ENST00000330953, ENST00000332427, ENST00000332438, ENST00000339223, ENST00000340023, ENST00000344575, ENST00000355085, ENST00000355897, ENST00000358157, ENST00000358600, ENST00000358755, ENST00000361434, ENST00000366554, ENST00000370598, ENST00000370862, ENST00000373304, ENST00000373857, ENST00000374472, ENST00000374694, ENST00000375190, ENST00000376414, ENST00000377411, ENST00000377893, ENST00000378661, ENST00000378688, ENST00000381297, ENST00000394705, ENST00000395116, ENST00000396789, ENST00000399036, ENST00000400888, ENST00000447924, ENST00000467259, ENST00000490673	87
Adipose, Adrenal, Brain, Breast, Colon, Heart, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Testes, Thyroid	ENST00000216629, ENST00000249887, ENST00000276393, ENST00000298440, ENST00000324300, ENST00000338566, ENST00000373256	7
Adipose, Adrenal, Brain, Breast, Colon, Heart, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Testes, Thyroid, White blood	ENST00000217270, ENST00000246657, ENST00000255380, ENST00000257267, ENST00000262738, ENST00000282018, ENST00000356987, ENST00000375466, ENST00000377012, ENST00000377741, ENST00000378142, ENST00000393590, ENST00000396782, ENST00000407877, ENST00000428545	15
Adipose, Adrenal, Brain, Breast, Colon, Heart, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Testes, Thyroid	ENST00000424100	1
Adipose, Adrenal, Brain, Breast, Colon, Heart, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Thyroid, White blood	ENST000000329858	1
Adipose, Adrenal, Brain, Breast, Colon, Heart, Kidney, Liver, Lung, Lymph node, Ovary, Skeletal muscle, Thyroid, White blood	ENST000000456426	1
Adipose, Adrenal, Brain, Breast, Colon, Heart, Kidney, Liver, Lung, Ovary, Prostate, Testes, Thyroid	ENST000000380582	1
Adipose, Adrenal, Brain, Breast, Colon, Heart, Kidney, Liver, Lymph node, Ovary, Prostate, Skeletal muscle, Testes, Thyroid	ENST000000397088	1
Adipose, Adrenal, Brain, Breast, Colon, Heart, Kidney, Liver, Lymph node, Ovary, Prostate, Skeletal muscle, Testes, Thyroid, White blood	ENST000000380289, ENST00000420975	2
Adipose, Adrenal, Brain, Breast, Colon, Heart, Kidney, Liver, Lymph node, Ovary, Prostate, Testes, Thyroid, White blood	ENST000000319838	1
Adipose, Adrenal, Brain, Breast, Colon, Heart, Kidney, Liver, Lymph node, Thyroid	ENST000000369716	1
Adipose, Adrenal, Brain, Breast, Colon, Heart, Kidney, Liver, Testes	ENST000000486700	1
Adipose, Adrenal, Brain, Breast, Colon, Heart, Kidney, Lung, Lymph node, Ovary, Prostate, Skeletal muscle, Testes, Thyroid	ENST000000229030, ENST00000305249, ENST00000432564	3
Adipose, Adrenal, Brain, Breast, Colon, Heart, Kidney, Lung, Lymph node, Ovary, Prostate, Skeletal muscle, Testes, Thyroid, White blood	ENST000000319595, ENST00000358556, ENST00000370924, ENST00000375846, ENST00000392370	5
Adipose, Adrenal, Brain, Breast, Colon, Heart, Kidney, Lung, Lymph node, Ovary, Prostate, Testes, Thyroid, White blood	ENST000000306960, ENST00000396333, ENST00000440886	3
Adipose, Adrenal, Brain, Breast, Colon, Heart, Kidney, Lung, Lymph node, Prostate, Thyroid, White blood	ENST000000373642	1
Adipose, Adrenal, Brain, Breast, Colon, Heart, Kidney, Lymph node, Ovary, Prostate, Thyroid, White blood	ENST000000250572	1
Adipose, Adrenal, Brain, Breast, Colon, Heart, Kidney, Lung, Lymph node, Ovary, Prostate, Skeletal muscle, Testes, Thyroid, White blood	ENST000000433765	1
Adipose, Adrenal, Brain, Breast, Colon, Heart, Kidney, Lung, Lymph node, Ovary, Prostate, Skeletal muscle, Testes, Thyroid, White blood	ENST000000164024, ENST00000315033, ENST00000327809, ENST00000380007	4
Adipose, Adrenal, Brain, Breast, Colon, Heart, Liver, Lung, Lymph node, Skeletal muscle, Testes, Thyroid	ENST000000368122	1
Adipose, Adrenal, Brain, Breast, Colon, Heart, Liver, Lung, Ovary, Prostate, Testes, Thyroid, White blood	ENST000000305097	1
Adipose, Adrenal, Brain, Breast, Colon, Heart, Liver, Lung, Ovary, Skeletal muscle, Testes	ENST000000354550	1
Adipose, Adrenal, Brain, Breast, Colon, Heart, Lung, Lymph node, Ovary, Prostate, Skeletal muscle, Testes, Thyroid	ENST000000367282	1
Adipose, Adrenal, Brain, Breast, Colon, Heart, Lung, Lymph node, Ovary, Prostate, Skeletal muscle, Testes, Thyroid, White blood	ENST000000373307	1
Adipose, Adrenal, Brain, Breast, Colon, Heart, Lung, Lymph node, White blood	ENST000000231683	1
Adipose, Adrenal, Brain, Breast, Colon, Heart, Ovary, Prostate, Testes, Thyroid	ENST000000471646	1
Adipose, Adrenal, Brain, Breast, Colon, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Skeletal muscle, Testes, Thyroid	ENST000000266674, ENST00000374024, ENST00000441240	3
Adipose, Adrenal, Brain, Breast, Colon, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Testes, Thyroid	ENST000000393752	1
Adipose, Adrenal, Brain, Breast, Colon, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Skeletal muscle, Testes, Thyroid, White blood	ENST000000333430	1
Adipose, Adrenal, Brain, Breast, Colon, Kidney, Liver, Lymph node, Testes, Thyroid	ENST000000307161	1
Adipose, Adrenal, Brain, Breast, Colon, Kidney, Lung, Lymph node, Ovary, Prostate, Testes, Thyroid, White blood	ENST000000464585	1
Adipose, Adrenal, Brain, Breast, Colon, Kidney, Lung, Lymph node, Ovary, Prostate, Skeletal muscle, Testes, Thyroid	ENST000000369947	1
Adipose, Adrenal, Brain, Breast, Colon, Kidney, Lung, Lymph node, Ovary, Prostate, Skeletal muscle, Testes, Thyroid, White blood	ENST000000289753	1
Adipose, Adrenal, Brain, Breast, Colon, Liver, Lung, Ovary, Prostate, Skeletal muscle, Testes, Thyroid, White blood	ENST000000438259	1
Adipose, Adrenal, Brain, Breast, Colon, Lung, Lymph node, Ovary, Prostate, White blood	ENST000000284311	1
Adipose, Adrenal, Brain, Breast, Colon, Lung, Lymph node, Prostate, Skeletal muscle, Testes, Thyroid	ENST000000369501	1
Adipose, Adrenal, Brain, Breast, Colon, Lung, Ovary, Prostate, Testes, Thyroid	ENST000000495912	1
Adipose, Adrenal, Brain, Breast, Heart, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Skeletal muscle, Testes, Thyroid, White blood	ENST000000325602	1
Adipose, Adrenal, Brain, Breast, Heart, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Testes, Thyroid, White blood	ENST000000267015, ENST00000302548	2

Tissues of Expression	Isoform ID (ensemblTransID)	# of Isoforms
Adipose, Adrenal, Brain, Breast, Heart, Kidney, Liver, Lung, Lymph node, Ovary, Testes, Thyroid, White blood	ENST00000409817	1
Adipose, Adrenal, Brain, Breast, Heart, Kidney, Liver, Lymph node, Ovary, Prostate, Testes, Thyroid, White blood	ENST00000298171	1
Adipose, Adrenal, Brain, Breast, Heart, Kidney, Liver, Ovary, Prostate, Testes, Thyroid	ENST00000304402	1
Adipose, Adrenal, Brain, Breast, Heart, Kidney, Liver, Ovary, Prostate, Thyroid, White blood	ENST00000311762	1
Adipose, Adrenal, Brain, Breast, Heart, Kidney, Lung, Lymph node, Ovary, Prostate, Testes, Thyroid, Testes, Thyroid	ENST00000485973	1
Adipose, Adrenal, Brain, Breast, Heart, Lung, Lymph node	ENST00000406625	1
Adipose, Adrenal, Brain, Breast, Heart, Lung, Lymph node, Ovary, Skeletal muscle	ENST00000315332	1
Adipose, Adrenal, Brain, Breast, Heart, Lung, Lymph node, Prostate, Testes, White blood	ENST00000479077	1
Adipose, Adrenal, Brain, Breast, Heart, Lung, Lymph node, White blood	ENST00000377291	1
Adipose, Adrenal, Brain, Breast, Kidney, Liver, Lung, Ovary, Prostate, Testes, Thyroid, White blood	ENST00000260843	1
Adipose, Adrenal, Brain, Breast, Kidney, Liver, Lung, Ovary, Prostate, Testes, White blood	ENST00000334425	1
Adipose, Adrenal, Brain, Breast, Kidney, Liver, Lymph node, Ovary, Prostate, Testes	ENST00000296932	1
Adipose, Adrenal, Brain, Breast, Lymph node, Ovary, Prostate, Testes, Thyroid	ENST00000356606	1
Adipose, Adrenal, Brain, Breast, Ovary, Testes	ENST00000413416	1
Adipose, Adrenal, Brain, Colon, Heart, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Skeletal muscle, Testes, Thyroid, White blood	ENST00000300098, ENST00000311268	2
Adipose, Adrenal, Brain, Colon, Heart, Kidney, Ovary, Prostate, Testes	ENST00000464295	1
Adipose, Adrenal, Brain, Colon, Heart, Liver, Lung, Lymph node, Ovary	ENST00000395716	1
Adipose, Adrenal, Brain, Colon, Heart, Liver, Lung, Lymph node, Ovary, Skeletal muscle, Thyroid	ENST00000458629	1
Adipose, Adrenal, Brain, Colon, Heart, Lung, Lymph node, Ovary, Prostate, Skeletal muscle, Thyroid, White blood	ENST00000333493	1
Adipose, Adrenal, Brain, Colon, Heart, Lung, Skeletal muscle	ENST00000447845	1
Adipose, Adrenal, Brain, Colon, Kidney, Lung, Lymph node, Skeletal muscle, Testes	ENST00000444847	1
Adipose, Adrenal, Brain, Colon, Kidney, Lymph node, Ovary, Prostate, Skeletal muscle, Testes, Thyroid, White blood	ENST00000349984	1
Adipose, Adrenal, Brain, Colon, Lung, Lymph node, Ovary, Prostate, Testes, Thyroid, White blood	ENST00000329797	1
Adipose, Adrenal, Brain, Colon, Lymph node, Ovary, White blood	ENST00000407714	1
Adipose, Adrenal, Brain, Heart, Kidney, Liver, Lung, Prostate, Thyroid	ENST00000282753	1
Adipose, Adrenal, Brain, Heart, Kidney, Prostate, Skeletal muscle, Testes, Thyroid	ENST00000393596	1
Adipose, Adrenal, Brain, Heart, Lung, Lymph node, Prostate, Testes	ENST00000327606	1
Adipose, Adrenal, Brain, Heart, Lung, Ovary	ENST00000332582	1
Adipose, Adrenal, Brain, Kidney, Liver, Lymph node, Ovary, Prostate, Testes, Thyroid, White blood	ENST00000360222	1
Adipose, Adrenal, Brain, Kidney, Lymph node, Ovary, Skeletal muscle, Testes, Thyroid, White blood	ENST00000420125	1
Adipose, Adrenal, Brain, Lung, Lymph node, Ovary, Prostate, Testes	ENST00000374619	1
Adipose, Adrenal, Brain, Lymph node, Testes	ENST00000369717	1
Adipose, Adrenal, Brain, Lymph node, White blood	ENST00000336866	1
Adipose, Adrenal, Breas	ENST00000348438	1
Adipose, Adrenal, Breast, Colon, Heart, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Skeletal muscle, Testes, Thyroid	ENST00000392627	1
Adipose, Adrenal, Breast, Colon, Heart, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Thyroid, White blood	ENST00000246115	1
Adipose, Adrenal, Breast, Colon, Heart, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Skeletal muscle, Thyroid, White blood	ENST00000345363, ENST00000393597	2
Adipose, Adrenal, Breast, Colon, Heart, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Testes, Thyroid	ENST00000265417, ENST00000367608	2
Adipose, Adrenal, Breast, Colon, Heart, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Testes, Thyroid, Thyroid	ENST00000370713	1
Adipose, Adrenal, Breast, Colon, Heart, Kidney, Liver, Lung, Lymph node, Ovary, Skeletal muscle, Testes, Thyroid	ENST00000246549	1
Adipose, Adrenal, Breast, Colon, Heart, Kidney, Liver, Lung, Ovary, Prostate, Testes, Thyroid	ENST00000380586	1
Adipose, Adrenal, Breast, Colon, Heart, Kidney, Liver, Ovary, Prostate, Skeletal muscle, Testes, Thyroid	ENST00000309502	1
Adipose, Adrenal, Breast, Colon, Heart, Kidney, Lung, Lymph node, Ovary, Prostate, Testes, Thyroid, Testes, Thyroid	ENST00000493254	1
Adipose, Adrenal, Breast, Colon, Heart, Kidney, Lung, Lymph node, Ovary, Prostate, Testes, Thyroid, White blood	ENST00000370757	1
Adipose, Adrenal, Breast, Colon, Heart, Liver, Lung, Lymph node, Ovary, Prostate, Skeletal muscle, Testes, Thyroid, White blood	ENST00000454971	1
Adipose, Adrenal, Breast, Colon, Heart, Lung, Lymph node, Ovary, Prostate, Skeletal muscle, Testes, White blood	ENST00000410068	1
Adipose, Adrenal, Breast, Colon, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Testes, Thyroid	ENST00000309099	1
Adipose, Adrenal, Breast, Colon, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Skeletal muscle, Testes	ENST00000427125	1
Adipose, Adrenal, Breast, Colon, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Skeletal muscle, Testes, Thyroid, White blood	ENST00000446573	1
Adipose, Adrenal, Breast, Colon, Kidney, Liver, Lung, Lymph node, Ovary, Skeletal muscle, Testes, Thyroid	ENST00000492492	1
Adipose, Adrenal, Breast, Colon, Kidney, Lung, Lymph node, Prostate, Skeletal muscle, Testes, Thyroid, White blood	ENST00000397470	1
Adipose, Adrenal, Breast, Colon, Liver, Lung, Ovary, Prostate, Testes, Thyroid	ENST00000357716	1
Adipose, Adrenal, Breast, Colon, Lung, Lymph node, Ovary, Prostate, Skeletal muscle, Testes, Thyroid, White blood	ENST00000430267	1
Adipose, Adrenal, Breast, Colon, Lymph node, Ovary, Prostate, Testes	ENST00000367836	1
Adipose, Adrenal, Breast, Heart, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Skeletal muscle, Testes, Thyroid, White blood	ENST00000309285, ENST00000445132	2
Adipose, Adrenal, Breast, Heart, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Testes, Thyroid, White blood	ENST00000337539	1
Adipose, Adrenal, Breast, Heart, Kidney, Lung, Lymph node, Prostate, Testes, White blood	ENST00000373693	1
Adipose, Adrenal, Breast, Heart, Liver, Lung, Lymph node, Ovary, Prostate, Skeletal muscle, Testes, White blood	ENST00000442925	1
Adipose, Adrenal, Breast, Heart, Liver, Lung, Lymph node, Prostate, Testes, Thyroid, White blood	ENST00000426880	1
Adipose, Adrenal, Breast, Heart, Liver, Lymph node, Prostate, Testes, Thyroid	ENST00000393018	1
Adipose, Adrenal, Breast, Kidney, Liver, Lung, Lymph node, Prostate, Skeletal muscle, White blood	ENST00000340736	1
Adipose, Adrenal, Breast, Liver, Thyroid	ENST00000394720	1
Adipose, Adrenal, Breast, Liver, Thyroid, White blood	ENST00000397092	1
Adipose, Adrenal, Breast, Lymph node, Ovary, Prostate, Testes	ENST00000437420	1
Adipose, Adrenal, Breast, Testes, Thyroid	ENST00000313049	1
Adipose, Adrenal, Breast, Testes, White blood	ENST00000312053	1
Adipose, Adrenal, Colon, Heart, Kidney, Liver, Lung, Lymph node, Prostate, Testes, Thyroid	ENST00000325123	1
Adipose, Adrenal, Colon, Heart, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Testes	ENST00000496086	1
Adipose, Adrenal, Colon, Heart, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Testes	ENST00000402328	1
Adipose, Adrenal, Colon, Heart, Lymph node, Ovary, Prostate, Thyroid	ENST00000389832	1

Tissues of Expression	Isoform ID (ensemblTransID)	# of Isoforms
Adipose, Adrenal, Colon, Kidney, Liver, Lymph node, Prostate, Testes, Thyroid	ENST00000400723	1
Adipose, Adrenal, Colon, Kidney, Prostate, Testes, Thyroid	ENST00000396275	1
Adipose, Adrenal, Colon, Kidney, Prostate, Thyroid	ENST00000409848	1
Adipose, Adrenal, Colon, Lung, Lymph node, Ovary, Prostate, Testes, Thyroid	ENST00000469376	1
Adipose, Adrenal, Colon, Lymph node, Ovary, Prostate, Thyroid, White blood	ENST00000335666	1
Adipose, Adrenal, Colon, Skeletal muscle, White blood	ENST00000433848	1
Adipose, Adrenal, Heart, Lung, Lymph node, Ovary, Thyroid, White blood	ENST00000253673	1
Adipose, Adrenal, Heart, Ovary, Prostate, Testes	ENST00000367838	1
Adipose, Adrenal, Heart, Ovary, Testes, Thyroid	ENST00000487787	1
Adipose, Adrenal, Heart, Skeletal muscle, Testes	ENST00000378434	1
Adipose, Adrenal, Kidney, Liver, Lung, Ovary, Prostate, Testes	ENST00000419892	1
Adipose, Adrenal, Kidney, Liver, Lung, Prostate, Testes, Thyroid	ENST00000392965	1
Adipose, Adrenal, Kidney, Lung, Ovary, Prostate, White blood	ENST00000473675	1
Adipose, Adrenal, Kidney, Lymph node, Ovary, Prostate, Thyroid	ENST00000436735	1
Adipose, Adrenal, Liver, White blood	ENST00000467202	1
Adipose, Adrenal, Lung, Lymph node, Testes, Thyroid	ENST00000312457	1
Adipose, Adrenal, Lung, Prostate, Testes, White blood	ENST00000475374	1
Adipose, Adrenal, Lung, Testes, White blood	ENST00000392967	1
Adipose, Adrenal, Prostate, Skeletal muscle	ENST00000305392	1
Adipose, Brain, Breast, Colon, Heart, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Testes, Thyroid, White blood	ENST00000312143	1
Adipose, Brain, Breast, Colon, Heart, Kidney, Lung, Lymph node, Ovary, Prostate, Skeletal muscle, Testes, Thyroid	ENST00000441623	1
Adipose, Brain, Breast, Colon, Heart, Kidney, Lung, Lymph node, Ovary, Prostate, Testes, Thyroid	ENST00000255262	1
Adipose, Brain, Breast, Colon, Heart, Kidney, Lung, Prostate, Testes, Thyroid	ENST00000455756	1
Adipose, Brain, Breast, Colon, Heart, Ovary, Testes	ENST00000494619	1
Adipose, Brain, Breast, Colon, Kidney, Liver, Lymph node, Ovary, Prostate, Testes, Thyroid, White blood	ENST00000345941	1
Adipose, Brain, Breast, Colon, Kidney, Liver, Ovary, Prostate, Testes	ENST00000272847	1
Adipose, Brain, Breast, Colon, Kidney, Lung, Ovary, Prostate, Testes, Thyroid	ENST00000334619	1
Adipose, Brain, Breast, Colon, Kidney, Lymph node, Ovary, Testes, Thyroid	ENST00000449420	1
Adipose, Brain, Breast, Colon, Ovary	ENST00000402501	1
Adipose, Brain, Breast, Heart, Kidney, Liver, Lung, Lymph node, Ovary, Skeletal muscle, Testes, Thyroid	ENST00000466131	1
Adipose, Brain, Breast, Heart, Kidney, Lung, Ovary, Prostate, Testes	ENST00000359558	1
Adipose, Brain, Breast, Heart, Prostate, Skeletal muscle, Testes	ENST00000407325	1
Adipose, Brain, Breast, Liver, Thyroid	ENST00000474935	1
Adipose, Brain, Colon, Heart, Kidney, Liver, Lung, Lymph node, Ovary, Testes, White blood	ENST00000353876	1
Adipose, Brain, Colon, Heart, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Testes, White blood	ENST00000369499	1
Adipose, Brain, Colon, Heart, Lymph node, Skeletal muscle, Thyroid	ENST00000404129	1
Adipose, Brain, Colon, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Skeletal muscle, Testes, Thyroid	ENST00000370758	1
Adipose, Brain, Colon, Kidney, Lung, Prostate	ENST00000444078	1
Adipose, Brain, Colon, Liver, Lymph node, Ovary, Testes, Thyroid	ENST00000229955	1
Adipose, Brain, Colon, Lymph node, Prostate, Testes, White blood	ENST00000370501	1
Adipose, Brain, Heart, Kidney, Lung, Lymph node, Prostate, White blood	ENST00000457814	1
Adipose, Brain, Heart, Kidney, Lung, Ovary, Testes, Thyroid	ENST00000019103	1
Adipose, Brain, Heart, Liver, Lymph node, Ovary, Prostate, Testes, Thyroid	ENST00000463993	1
Adipose, Brain, Heart, Lung, Lymph node, Ovary, White blood	ENST00000438398	1
Adipose, Brain, Kidney, Ovary, Prostate, White blood	ENST00000482710	1
Adipose, Brain, Liver, Lung, Ovary, Prostate, Thyroid	ENST00000267377	1
Adipose, Brain, Liver, Ovary, Prostate, Testes	ENST00000381436	1
Adipose, Brain, Lymph node, Thyroid	ENST00000415392	1
Adipose, Brain, Testes	ENST00000371950	1
Adipose, Breast, Colon, Heart, Kidney, Liver, Lung, Lymph node, Prostate, Skeletal muscle, Thyroid	ENST00000404754	1
Adipose, Breast, Colon, Heart, Ovary, Testes, Thyroid	ENST00000428232	1
Adipose, Breast, Colon, Liver, Ovary, Thyroid	ENST00000475347	1
Adipose, Breast, Colon, White blood	ENST00000305786	1
Adipose, Breast, Heart, Kidney, Lung, Lymph node, Prostate, Thyroid	ENST00000400880	1
Adipose, Breast, Heart, Lung, Testes	ENST00000400882	1
Adipose, Breast, Kidney, Lung, Lymph node, White blood	ENST00000450180	1
Adipose, Breast, Kidney, Lung, Thyroid	ENST00000422395	1
Adipose, Breast, Kidney, Lymph node, Prostate	ENST00000349243	1
Adipose, Breast, Kidney, Ovary, Thyroid	ENST00000324750	1
Adipose, Breast, Kidney, Testes	ENST00000218316	1
Adipose, Breast, Kidney, Testes, Thyroid	ENST00000344775	1
Adipose, Breast, Skeletal muscle, Thyroid	ENST00000346057	1
Adipose, Colon, Heart, Kidney, Ovary	ENST00000398538	1
Adipose, Colon, Heart, Kidney, Prostate, Testes, Thyroid	ENST00000295589	1
Adipose, Colon, Heart, Prostate, Thyroid	ENST00000370931	1
Adipose, Heart, Lung, Ovary, White blood	ENST00000435290	1
Adipose, Kidney	ENST00000371220, ENST00000481661	2
Adipose, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Thyroid	ENST00000445772	1
Adipose, Kidney, Liver, Prostate	ENST00000331251	1
Adipose, Kidney, Lung, Lymph node, Ovary, Testes, Thyroid	ENST00000376977	1
Adipose, Kidney, Ovary, Testes, Thyroid	ENST00000496497	1
Adipose, Kidney, Testes, Thyroid	ENST00000358927	1
Adipose, Liver, Thyroid	ENST00000355973	1
Adipose, Lung, Lymph node, Testes, Thyroid	ENST00000402066	1
Adipose, Ovary, Prostate, Testes, Thyroid	ENST00000412101	1
Adipose, Prostate	ENST00000356595	1
Adipose, Thyroid	ENST00000297146	1
Adrenal	ENST00000339069, ENST00000370730, ENST00000380314, ENST00000397688, ENST00000409489, ENST00000436487, ENST00000463197, ENST00000470033, ENST00000474800, ENST00000492609	10
Adrenal, Brain, Breast, Colon, Heart, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Skeletal muscle, Testes, Thyroid, White blood	ENST00000334286	1
Adrenal, Brain, Breast, Colon, Heart, Kidney, Liver, Ovary, Prostate, Testes, Thyroid, White blood	ENST00000258456	1
Adrenal, Brain, Breast, Colon, Heart, Kidney, Skeletal muscle, Testes	ENST00000361669	1
Adrenal, Brain, Breast, Colon, Heart, Liver, Lung, Ovary, Prostate, Skeletal muscle, Testes, Thyroid, White blood	ENST00000392962	1
Adrenal, Brain, Breast, Colon, Liver, Lung, Lymph node	ENST00000491034	1
Adrenal, Brain, Breast, Colon, Liver, Lung, Lymph node, Ovary, Testes, Thyroid, White blood	ENST00000392040	1
Adrenal, Brain, Breast, Heart, Kidney, Liver, Lymph node, Ovary, Prostate	ENST00000493800	1
Adrenal, Brain, Breast, Kidney, Lung, Lymph node, White blood	ENST00000234371	1
Adrenal, Brain, Breast, Lung, Lymph node, Ovary, Prostate, Testes, Thyroid, White blood	ENST00000328544	1
Adrenal, Brain, Colon, Heart, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Testes, Thyroid	ENST00000249016, ENST00000445907	2
Adrenal, Brain, Colon, Heart, Kidney, Lymph node, Ovary, Skeletal muscle, White blood	ENST00000395963	1

Tissues of Expression	Isoform ID (ensemblTransID)	# of Isoforms
Adrenal, Brain, Colon, Heart, Thyroid	ENST00000367685	1
Adrenal, Brain, Colon, Kidney, Liver, Thyroid	ENST00000273145	1
Adrenal, Brain, Colon, Lymph node, Prostate, Testes	ENST00000284674	1
Adrenal, Brain, Heart, Kidney, Liver, Lung, Testes	ENST00000449591	1
Adrenal, Brain, Heart, Kidney, Lymph node, Testes, Thyroid	ENST00000447444	1
Adrenal, Brain, Heart, Kidney, Thyroid	ENST00000461057	1
Adrenal, Brain, Heart, Lung	ENST00000436083	1
Adrenal, Brain, Kidney, Liver, Lymph node, Ovary, Testes	ENST00000299766	1
Adrenal, Brain, Kidney, Liver, Testes, Thyroid, White blood	ENST00000409998	1
Adrenal, Brain, Kidney, Lung, Prostate, Skeletal muscle, Testes, Thyroid, White blood	ENST00000357355	1
Adrenal, Brain, Kidney, Lymph node, Ovary, Prostate, Testes, Thyroid	ENST00000370648	1
Adrenal, Brain, Lung, Testes, Thyroid	ENST00000484007	1
Adrenal, Brain, Lymph node, Ovary, Thyroid	ENST00000246553	1
Adrenal, Brain, Ovary	ENST00000442933	1
Adrenal, Brain, Prostate, Testes	ENST00000455714	1
Adrenal, Breast	ENST00000464977	1
Adrenal, Breast, Colon, Heart, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Skeletal muscle, Testes, White blood	ENST00000389858	1
Adrenal, Breast, Colon, Heart, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Testes, Thyroid, White blood	ENST00000304552	1
Adrenal, Breast, Colon, Heart, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Testes, Thyroid, White blood	ENST00000373306	1
Adrenal, Breast, Colon, Heart, Kidney, Lung, Lymph node, Prostate, Skeletal muscle, Thyroid	ENST00000359929	1
Adrenal, Breast, Colon, Heart, Lung, Lymph node, Prostate, Skeletal muscle	ENST00000337474	1
Adrenal, Breast, Colon, Heart, Lymph node, Ovary, Prostate, Testes	ENST00000394342	1
Adrenal, Breast, Colon, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Testes, Thyroid, White blood	ENST00000329003	1
Adrenal, Breast, Colon, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Testes, Thyroid	ENST00000297469	1
Adrenal, Breast, Colon, Lung, Lymph node, Ovary, Prostate, Testes, Thyroid	ENST00000346454	1
Adrenal, Breast, Colon, Lung, Prostate, Testes	ENST00000374177	1
Adrenal, Breast, Colon, White blood	ENST00000357392	1
Adrenal, Breast, Heart, Kidney	ENST00000435303	1
Adrenal, Breast, Heart, Kidney, Prostate	ENST00000349767	1
Adrenal, Breast, Heart, Liver, Lung, Lymph node, Ovary, Prostate, Skeletal muscle, Testes, Thyroid, White blood	ENST00000276077	1
Adrenal, Breast, Heart, Lymph node, Ovary, Skeletal muscle	ENST00000397095	1
Adrenal, Breast, Kidney, Liver, Lung, Lymph node, Thyroid	ENST00000395942	1
Adrenal, Breast, Liver, Lung, Lymph node, Ovary, Testes, Thyroid, White blood	ENST00000400926	1
Adrenal, Breast, Liver, Lymph node, Ovary, Thyroid	ENST00000342443	1
Adrenal, Breast, Lymph node, Ovary, Prostate, Testes, White blood	ENST00000340339	1
Adrenal, Breast, Lymph node, Prostate	ENST00000287534	1
Adrenal, Breast, Lymph node, Thyroid	ENST00000403528	1
Adrenal, Breast, Ovary, Prostate, Thyroid	ENST00000374519	1
Adrenal, Breast, Prostate	ENST00000462882	1
Adrenal, Colon	ENST00000368318, ENST00000418619	2
Adrenal, Colon, Heart, Kidney, Liver, Lung, Ovary, Prostate, Thyroid, White blood	ENST00000438284	1
Adrenal, Colon, Heart, Kidney, Lung, Lymph node, Ovary, Prostate, Testes, Thyroid	ENST00000371906	1
Adrenal, Colon, Heart, Liver, Lung, Testes, Thyroid	ENST00000486108	1
Adrenal, Colon, Heart, Lung	ENST00000354217	1
Adrenal, Colon, Heart, Testes, White blood	ENST00000380587	1
Adrenal, Colon, Kidney, Lymph node, Ovary, Prostate, Testes, White blood	ENST00000431922	1
Adrenal, Heart, Kidney, Lung, Lymph node, Ovary, Prostate, Testes	ENST00000448688	1
Adrenal, Heart, Kidney, Lung, Ovary, Testes, Thyroid	ENST00000395964	1
Adrenal, Heart, Kidney, Lung, Prostate	ENST00000377127	1
Adrenal, Heart, Kidney, Lymph node, Ovary, Testes, White blood	ENST00000341935	1
Adrenal, Heart, Liver	ENST00000461541	1
Adrenal, Heart, Liver, Ovary, Testes	ENST00000397473	1
Adrenal, Heart, Lung	ENST00000473719	1
Adrenal, Heart, Lung, Skeletal muscle, Testes	ENST00000438735	1
Adrenal, Heart, Lymph node, Prostate, Testes, Thyroid, White blood	ENST00000329773	1
Adrenal, Heart, Ovary, Prostate, Skeletal muscle, Thyroid	ENST00000423776	1
Adrenal, Heart, Ovary, White blood	ENST00000480322	1
Adrenal, Heart, Thyroid	ENST00000479784	1
Adrenal, Kidney, Lung, Prostate, Thyroid, White blood	ENST00000490131	1
Adrenal, Liver, Lung, Prostate	ENST00000496258	1
Adrenal, Liver, Lymph node, Ovary, Thyroid, White blood	ENST00000373691	1
Adrenal, Liver, Ovary, Prostate	ENST00000438062	1
Adrenal, Lung, Lymph node, Testes, White blood	ENST00000304244	1
Adrenal, Lung, Lymph node, White blood	ENST00000171757	1
Adrenal, Ovary	ENST00000392629	1
Adrenal, Ovary, Thyroid	ENST00000423548	1
Adrenal, Testes	ENST00000257070, ENST00000340807, ENST00000487573	3
Adrenal, White blood	ENST00000383673, ENST00000460330	2
Brain	ENST00000275169, ENST00000311104, ENST00000330432, ENST00000355319, ENST00000360581, ENST00000360693, ENST00000382487, ENST00000383263, ENST00000389335, ENST00000418177, ENST00000447403	11
Brain, Breast, Colon, Heart, Kidney, Lymph node, Testes, Thyroid	ENST00000496754	1
Brain, Breast, Colon, Heart, Ovary	ENST00000405846	1
Brain, Breast, Colon, Kidney, Liver, Lung, Ovary, Prostate, Testes, Thyroid	ENST00000329321	1
Brain, Breast, Colon, Kidney, Liver, Lung, Skeletal muscle, Thyroid	ENST00000271357	1
Brain, Breast, Colon, Kidney, Lung, Lymph node, Ovary, Prostate, Skeletal muscle, Testes	ENST00000397056	1
Brain, Breast, Colon, Kidney, Prostate, Testes, Thyroid	ENST00000249041	1
Brain, Breast, Colon, Liver, Lymph node, Ovary, Prostate, Testes, Thyroid	ENST00000329476	1
Brain, Breast, Colon, Lung, Ovary, Prostate, Thyroid	ENST00000405460	1
Brain, Breast, Heart	ENST00000474615	1
Brain, Breast, Heart, Lymph node, Ovary, Skeletal muscle, Testes	ENST00000413368	1
Brain, Breast, Kidney, Liver, Lung, Lymph node, Ovary, Prostate, Skeletal muscle, Testes, Thyroid	ENST00000422115	1
Brain, Breast, Kidney, Liver, Lung, Ovary, Prostate, White blood	ENST00000469383	1
Brain, Breast, Liver	ENST00000445970	1
Brain, Colon	ENST00000398547	1
Brain, Colon, Heart, Kidney, Liver, Lung, Ovary	ENST00000492666	1
Brain, Colon, Heart, Kidney, Ovary, Testes	ENST00000231188	1
Brain, Colon, Heart, Kidney, Prostate, Skeletal muscle	ENST00000340581	1
Brain, Colon, Heart, Liver, Thyroid, White blood	ENST00000412814	1
Brain, Colon, Heart, Liver, White blood	ENST00000344373	1
Brain, Colon, Heart, Ovary, Testes	ENST00000304374	1
Brain, Colon, Kidney, Lung, Lymph node, Ovary, Testes, Thyroid	ENST00000337894	1
Brain, Colon, Kidney, Thyroid	ENST00000367126	1
Brain, Colon, Liver, Testes	ENST00000287907	1
Brain, Colon, Lung, Lymph node, Ovary, Prostate, Testes, Thyroid	ENST00000376351	1
Brain, Colon, Ovary, Prostate, Testes	ENST00000369584	1
Brain, Colon, Ovary, Prostate, Testes, Thyroid	ENST00000258042	1
Brain, Colon, Prostate, Thyroid	ENST00000449735	1

Tissues of Expression	Isoform ID (ensemblTransID)	# of Isoforms
Brain, Heart	ENST00000257068	1
Brain, Heart, Kidney, Liver, Lung, Lymph node, Prostate, Skeletal muscle	ENST00000428220	1
Brain, Heart, Kidney, Liver, Lung, Prostate, Thyroid	ENST00000370611	1
Brain, Heart, Kidney, Lymph node	ENST00000293897	1
Brain, Heart, Kidney, Ovary, Prostate, Testes	ENST00000304421	1
Brain, Heart, Kidney, Prostate, Skeletal muscle, Testes, Thyroid	ENST000003397547	1
Brain, Heart, Kidney, Skeletal muscle, Testes	ENST00000373706	1
Brain, Heart, Liver, Lung, Lymph node, Ovary, Prostate, Thyroid	ENST00000272644	1
Brain, Heart, Lymph node, Ovary, Prostate, Skeletal muscle, Testes, Thyroid	ENST00000306928	1
Brain, Heart, Ovary, Prostate	ENST00000283303	1
Brain, Heart, Ovary, Testes	ENST00000469377	1
Brain, Kidney	ENST00000351052, ENST00000412232	2
Brain, Kidney, Liver	ENST00000349457	1
Brain, Kidney, Lung, Lymph node, Ovary	ENST00000430002	1
Brain, Kidney, Lung, Lymph node, Prostate, Skeletal muscle, Thyroid, White blood	ENST00000367278	1
Brain, Kidney, Ovary, Testes, Thyroid	ENST00000333172	1
Brain, Kidney, Prostate, Testes	ENST00000368549	1
Brain, Kidney, Testes	ENST00000305344	1
Brain, Kidney, Testes, Thyroid	ENST00000296154	1
Brain, Kidney, White blood	ENST00000327655	1
Brain, Liver	ENST00000242786, ENST00000307765, ENST00000343801	3
Brain, Liver, Ovary	ENST00000319517	1
Brain, Liver, Prostate, White blood	ENST00000381404	1
Brain, Liver, Testes	ENST00000339582, ENST00000371719	2
Brain, Liver, Testes, Thyroid	ENST00000493193	1
Brain, Lung	ENST00000396211	1
Brain, Lung, Ovary, Prostate	ENST00000472823	1
Brain, Lung, Ovary, Prostate, Skeletal muscle, Testes, White blood	ENST00000498475	1
Brain, Lung, Prostate	ENST00000389740	1
Brain, Lung, Prostate, Skeletal muscle, Testes, Thyroid, White blood	ENST00000342292	1
Brain, Lung, Testes, Thyroid	ENST00000255008	1
Brain, Lymph node, Thyroid	ENST00000314512	1
Brain, Ovary	ENST00000441140	1
Brain, Ovary, Prostate	ENST00000373705	1
Brain, Ovary, Prostate, Testes	ENST00000340177	1
Brain, Ovary, Testes	ENST00000293493, ENST00000298110, ENST00000349451, ENST00000366834, ENST00000369768	5
Brain, Ovary, Testes, Thyroid	ENST00000374181	1
Brain, Prostate	ENST00000369169	1
Brain, Prostate, Testes	ENST00000352855	1
Brain, Skeletal muscle	ENST00000320913	1
Brain, Skeletal muscle, Testes, Thyroid	ENST00000367236	1
Brain, Testes	ENST00000252660, ENST00000281806, ENST00000323865, ENST00000358373, ENST00000394427, ENST00000427970	6
Brain, Testes, Thyroid	ENST00000265572	1
Breast	ENST00000314254, ENST00000334810, ENST00000373658, ENST00000472701	4
Breast, Colon	ENST00000315576	1
Breast, Colon, Heart, Lung, Prostate, Thyroid	ENST00000424796	1
Breast, Colon, Lung, Ovary, Prostate, Skeletal muscle, White blood	ENST00000367609	1
Breast, Colon, Lymph node, Prostate	ENST00000401670	1
Breast, Colon, Ovary, Prostate, Testes, Thyroid, White blood	ENST00000382315	1
Breast, Colon, Prostate, Testes, Thyroid	ENST00000374312	1
Breast, Colon, Testes	ENST00000361719	1
Breast, Heart	ENST00000433647	1
Breast, Heart, Ovary, Skeletal muscle, Testes, Thyroid, White blood	ENST00000428600	1
Breast, Heart, Ovary, Testes	ENST00000465365	1
Breast, Heart, Testes	ENST00000431811	1
Breast, Kidney, Lung, Lymph node, Ovary, Prostate, Testes, Thyroid	ENST00000371253	1
Breast, Kidney, Lung, Ovary	ENST00000475478	1
Breast, Lung	ENST00000443157, ENST00000498400	2
Breast, Lung, Lymph node, Ovary, Skeletal muscle, Testes, Thyroid, White blood	ENST00000393592	1
Breast, Lung, Ovary, Prostate, Thyroid, White blood	ENST00000411851	1
Breast, Lung, Ovary, White blood	ENST00000395940	1
Breast, Lung, Prostate	ENST00000371243	1
Breast, Lymph node, White blood	ENST00000326306	1
Breast, Ovary	ENST00000294954, ENST00000305432, ENST00000449615, ENST00000497921	4
Breast, Ovary, Thyroid	ENST00000424232	1
Breast, Ovary, White blood	ENST00000497923	1
Breast, Prostate, Skeletal muscle	ENST00000417980	1
Breast, Prostate, Thyroid	ENST00000471487	1
Breast, Testes	ENST00000355631	1
Breast, Thyroid	ENST00000314537	1
Breast, White blood	ENST00000452454	1
Colon	ENST00000296479, ENST00000377016, ENST00000403881, ENST00000422265, ENST00000481598	5
Colon, Heart	ENST00000402486	1
Colon, Heart, Kidney, Ovary, Testes, Thyroid, White blood	ENST00000444262	1
Colon, Heart, Kidney, Skeletal muscle, Thyroid	ENST00000377633	1
Colon, Heart, Liver, Thyroid	ENST00000230173	1
Colon, Heart, Lung, Lymph node, Ovary, Testes	ENST00000367835	1
Colon, Heart, Lymph node	ENST00000343542	1
Colon, Heart, Lymph node, Ovary, Skeletal muscle, White blood	ENST00000431010	1
Colon, Heart, Lymph node, Prostate, Testes	ENST00000480591	1
Colon, Heart, Skeletal muscle, Testes	ENST00000453237	1
Colon, Heart, Testes	ENST00000393294	1
Colon, Kidney	ENST00000392964, ENST00000393591	2
Colon, Kidney, Lymph node, Ovary, Skeletal muscle, Testes, White blood	ENST00000256906	1
Colon, Liver	ENST00000273352	1
Colon, Liver, Lung, Testes, Thyroid	ENST00000374431	1
Colon, Liver, Prostate, Testes	ENST00000497524	1
Colon, Lung, Lymph node, Thyroid, White blood	ENST00000399220	1
Colon, Lung, Ovary, Testes	ENST00000394141	1
Colon, Lung, Prostate, Thyroid	ENST00000403273	1
Colon, Lymph node	ENST00000330120	1
Colon, Lymph node, Prostate, Testes	ENST00000314113	1
Colon, Lymph node, Testes	ENST00000434900	1
Colon, Ovary	ENST00000361210, ENST00000377888, ENST00000457830	3
Colon, Prostate, Testes, Thyroid	ENST00000358749	1
Colon, Prostate, Thyroid	ENST00000461609	1
Heart	ENST00000401861, ENST00000442484	2
Heart, Kidney, Liver	ENST00000412160	1
Heart, Kidney, Liver, Ovary, Testes, White blood	ENST00000331838	1
Heart, Kidney, Lung	ENST00000379873	1
Heart, Kidney, Lung, Ovary, Thyroid	ENST00000344413	1

Tissues of Expression	Isoform ID (ensemblTransID)	# of Isoforms
Heart, Kidney, Lung, Prostate	ENST000004579916	1
Heart, Liver	ENST00000357991	1
Heart, Liver, Prostate, Testes, Thyroid	ENST00000362016	1
Heart, Lung, Ovary, Skeletal muscle, Testes	ENST00000218721	1
Heart, Lung, Ovary, Thyroid	ENST00000373655	1
Heart, Lung, White blood	ENST00000447769	1
Heart, Lymph node, Ovary, Thyroid	ENST00000311131	1
Heart, Lymph node, Skeletal muscle	ENST00000339690	1
Heart, Ovary, Prostate, Testes, Thyroid	ENST00000327753	1
Heart, Skeletal muscle	ENST00000377034	1
Heart, Skeletal muscle, White blood	ENST00000374430	1
Heart, Testes	ENST00000341843	1
Heart, Thyroid	ENST00000255432	1
Kidney	ENST00000310357, ENST00000414819, ENST00000485464	3
Kidney, Liver, Lung, Lymph node, Prostate, Skeletal muscle, Thyroid	ENST00000378138	1
Kidney, Liver, Testes	ENST00000403859	1
Kidney, Lung, Lymph node, White blood	ENST00000333478	1
Kidney, Lung, Ovary, Testes	ENST00000444921, ENST00000467205	2
Kidney, Lymph node	ENST00000352846	1
Kidney, Lymph node, Testes, White blood	ENST00000298386	1
Kidney, Ovary	ENST00000405626	1
Kidney, Ovary, Thyroid	ENST00000498619	1
Kidney, Prostate	ENST00000487545	1
Kidney, Prostate, Testes, Thyroid	ENST00000371951	1
Kidney, Prostate, Thyroid	ENST00000283297	1
Kidney, Testes	ENST00000469138	1
Kidney, Testes, Thyroid	ENST00000298682	1
Kidney, Thyroid	ENST00000405249, ENST00000477771	2
Liver	ENST00000228887, ENST00000271029, ENST00000354791	3
Liver, Lung	ENST00000335786	1
Liver, Lung, Lymph node, Prostate	ENST00000414642	1
Liver, Lung, Ovary, Testes	ENST00000341617	1
Liver, Lung, Prostate, Thyroid	ENST00000292303	1
Liver, Lymph node, Ovary, White blood	ENST00000416594	1
Liver, Ovary, Testes, White blood	ENST00000358309	1
Liver, Prostate, Testes	ENST00000472248	1
Lung	ENST00000361697, ENST00000371721, ENST00000381407, ENST00000421659, ENST00000439591	5
Lung, Ovary, Thyroid, White blood	ENST00000337049	1
Lung, Ovary, White blood	ENST00000428565	1
Lung, Prostate	ENST00000438013	1
Lymph node	ENST00000317393, ENST00000421592	2
Lymph node, Ovary, Prostate, Thyroid	ENST00000419928, ENST00000436464	2
Lymph node, Prostate	ENST00000370932	1
Lymph node, Prostate, Testes	ENST00000442134	1
Lymph node, Testes	ENST00000453373	1
Lymph node, Testes, Thyroid, White blood	ENST00000392039	1
Ovary	ENST00000309106, ENST00000392299, ENST00000394361, ENST00000439910, ENST00000446673	5
Ovary, Prostate	ENST00000395052	1
Ovary, Prostate, Testes, Thyroid	ENST00000452278	1
Ovary, Testes	ENST00000315843, ENST00000419927	2
Ovary, White blood	ENST00000367235	1
Prostate	ENST00000355983, ENST00000358883, ENST00000369212, ENST00000389336, ENST00000439764, ENST00000471616, ENST00000491283, ENST00000494542	8
Prostate, Testes	ENST00000467425	1
Prostate, Thyroid	ENST00000468575	1
Skeletal muscle	ENST00000360249, ENST00000401034, ENST00000409363, ENST00000494668	4
Testes	ENST00000229768, ENST00000302797, ENST00000332314, ENST00000336152, ENST00000346173, ENST00000371211, ENST00000395998, ENST00000399821, ENST00000406846, ENST00000423019, ENST00000436689, ENST00000440923, ENST00000462632, ENST00000467632	14
Testes, Thyroid	ENST00000276218, ENST00000426151, ENST00000489301, ENST00000490186	4
Thyroid	ENST00000277942, ENST00000370727, ENST00000398307, ENST00000415128, ENST00000439932, ENST00000451790	6
Thyroid, White blood	ENST00000357422	1
White blood	ENST00000412676, ENST00000435307, ENST00000443498, ENST00000443646, ENST00000474212	5
Adrenal, Brain, Kidney, Lung, Prostate, Skeletal muscle, Testes, Thyroid	ENST00000226413	1
Number of categories		459

Appendix_Table 3.4: Combinations of predicted TFs and the isoforms they regulate

Transcription factors	IsoformID	# of isoforms
ARID5A, CEBPB, CREB1, MZF1, SATB1, SMAD3, STAT3, ZBTB3, ZNF219, ZNF263, ZNF353	ENST00000261654	1
ARID5A, CEBPB, KLF6, MEF2C, MEIS1, MZF1, NF1, SATB1, SMAD3, SP1, STAT5B, ZBTB3, ZNF219	ENST00000272928	1
ARID5A, CEBPD, HHEX, HSF2, IRF3, MZF1, ZNF219	ENST00000378138, ENST00000378142	2
ARID5A, E2F4, KLF6, MAZ, MZF1, NF1, RFX5, SF1, SP1, ZBTB7, ZNF263, ZNF354C	ENST00000392627	1
ARID5A, HHEX, MEF2C, MZF1, SP1	ENST00000402328	1
ARID5A, HHEX, MEIS1, PBX3, SMAD4, SP1, STAT5B, ZBTB3, ZFX, ZNF219, ZNF35	ENST00000402328	1
ARID5A, HSF2, MZF1, NF1, SMAD3, ZBTB3	ENST00000402328	1
ARNT, HSF2, IRF7, STAT5A	ENST00000400926	1
ARNT, KLF6, MZF1, SP1, ZNF354C	ENST000004345941	1
ARNT, MZF1, SP1, ZFX	ENST00000474396	1
ASCL2, BARX2, GATA4, HOXB9, ZFP57	ENST000004093294	1
ASCL2, EOMES, LEF1, NRL, SALL2, ZNF300	ENST00000412676	1
ASCL2, ETV1, HIC1, LEF1, SALL2	ENST000004330120	1
ASCL2, ETV1, HIC1, TEAD4, ZKSCAN3	ENST000004392964	1
ATF6, E2F3, E2F4, EGR1, HSF2, KLF6, KLF7, MZF1, SMAD3, SP1, STAT5B, ZBTB3, ZFX, ZNF219, ZNF263	ENST000004354791	1
ATF6, E2F4, KAI5, KLF7, NKAPPAB, PBX3, SP1, ZBTB3, ZFX	ENST000004344575	1
ATF6, E2F4, KLF6, KLF7, MZF1, SP1, SP1, ZBTB3, ZNF219	ENST000004393597	1
ATF6, E2F4, KLF6, MZF1, SP1, TCF12, ZBTB3, ZBTB7	ENST00000423289	1
ATF6, E2F4, KLF7, MZF1, NKAPPAB, SMAD3, SP1, SP1, ZFX, ZNF219, ZNF263, ZNF35	ENST000004292513	1
ATF6, KLF1, KLF6, MAZ, MEIS1, MZF1, SP1, TCF7L2, ZBTB3, ZEB1, ZFX, ZNF219	ENST00000422115	1
ATF6, HSF2, KLF6, MZF1, RFX1, SMAD3, SP1	ENST000004395116	1
ATF6, MZF2, SF1, YY1, ZBTB3	ENST000004356666	1
ATF6, NRFL1, RFX1, SMAD3, ZBTB3	ENST000004332427	1
ATOH1, ETV1, SOX6, WT1, ZNF300	ENST000004393591	1
BACH2	ENST00000479077	1
BACH2, BARX2, DLX1, ERG, FOSL1, HOXA3, HOXB9, HOXC9, SOX6	ENST000004283297	1
BACH2, BARX2, ERG, GATA1, IRF4, IRX2, IRX5, ZNF300	ENST00000421659	1
BACH2, BARX2, FOSL1, HOXA4, LEF1, WT1	ENST000004309106	1
BACH2, BARX2, HOXA3, SOX6	ENST000004268549	1
BACH2, CARF	ENST00000443498	1
BACH2, CARF, ETV1, GLI3, PLAG1, SALL2, ZKSCAN3, ZNF300	ENST000004374181	1
BACH2, CARF, HIC1	ENST000004387373	1
BACH2, CARF, HOXA3, HOX8, MSX1	ENST000004434146	1
BACH2, CARF, TEAD4	ENST000004369717	1
BACH2, CDX1, DLX3, EOMES, GATA4, HOXD10, LEF1, LMX1B, NRL, SALL1, SOX3	ENST00000441843	1
BACH2, CDX1, DLX3, ERG, GATA4, GRHL2, HIC1, HOXA3, HOXA9, HOXD10, LEF1, RFX4, SOX3, SOX6, SPZ1, YY2, ZIC2, ZNF300	ENST00000435843	1
BACH2, E2F1, ETV1, GLI2, HIC1, INSM1, MAFA, PLAG1, SALL2, TEAD4, ZKSCAN3, ZNF300	ENST00000420913	1
BACH2, ERG, HIC1, HOXA3, LEF1, ZNF300	ENST000004095912	1
BACH2, ETV1, FOSL1, GLI3, SALL2	ENST000004305392	1
BACH2, ETV1, HOXA9, LEF1, SOX6	ENST00000443157	1
BACH2, ETV1, ZNF300	ENST00000402501	1
BACH2, FOSL1, GATA6, TEAD4, ZNF300	ENST000004296862, ENST00000467205	2
BACH2, GATA3, SALL2, SOX6, TEAD4	ENST000004374430	1
BACH2, HIC1, LEF1	ENST000004260843	1
BACH2, LEF1	ENST000004253673	1
BACH2, LEF1, TEAD4	ENST000004389740	1
BACH2, SOX5	ENST000004283303, ENST00000327753	2
BARX2, CARF, E2F1, EBF1, HIC1, HOXA3, ZNF300	ENST000004344413	1
BARX2, CARF, ETV1, HIC1, LEF1, PLAG1	ENST000004333478, ENST00000435303, ENST00000447444	3
BARX2, CARF, HIC1, IRF4, SOX6	ENST000004326306	1
BARX2, CARF, MSX2, SALL1, SOX6	ENST00000439932, ENST00000451790	2
BARX2, CDX1, ERG, GATA3, HOXD10, IRF4, LEF1, TEAD4, YY2	ENST000004395999	1
BARX2, CDX1, ETV1, GATA4, GRHL2, SOX5, SOX6	ENST000004349900	1
BARX2, CDX1, ETV1, HOXA4, LBX2, LMX1B, SOX3, SOX5, ZNF300	ENST000004340807	1
BARX2, CDX1, GM1, HOXA3, LEF1, SOX9	ENST000004370730	1
BARX2, DLX1, DLX3, ERG, GATA3, HOXA3, LH6X, MYT1L, NRL, SOX9	ENST000004489301	1
BARX2, DLX1, DLX5, SALL2, SOX5, ZKSCAN3	ENST000004360581	1
BARX2, DLX1, E2F1, EOMES, GATA3, HOXA6, HOX8, INSM1, LMX1B, MSX1, SALL2, SOX5, SPZ1, ZFP57, ZKSCAN3, ZNF300	ENST000004394427	1
BARX2, DLX1, ERG, MSX1, MYT1, PLAG1, SALL2, SOX5	ENST000004383263	1
BARX2, DLX1, FOXP2, GATA2, HOXA3, HOXD10, SOX5	ENST000004461609	1
BARX2, DLX1, HOXA3, LEF1, SOX9, YY2	ENST000004370727	1
BARX2, DLX3, E2F1, ERG, ETV1, MYT1L, PLAG1, SPZ1, YY2, ZFP57, ZKSCAN3	ENST000004233865	1
BARX2, DLX5, ERG, ETV1, GATA4, HOXA6, HOX8, HOXA9, LMX1B, SOX3, SOX6, ZKSCAN3, ZNF300	ENST000004305432	1
BARX2, DLX5, ERG, GATA3, HOXA3, LH6X, MYT1L, NRL, SOX9	ENST000004371211	1
BARX2, DLX5, ETV1, GATA3, GATA4, HOXA8, HOXD3, SPZ1, TEAD4	ENST000004346173, ENST00000406846, ENST00000419927	3
BARX2, E2F1, EBF1, ETV1, LEF1, PLAG1, SALL2, SP4, ZNF300	ENST000004234371	1
BARX2, E2F1, EBF1, FOXP2, ZNF300	ENST000004361697	1
BARX2, E2F1, EBF1, HOXA3, LEF1, MSX1, SOX6, ZNF300	ENST000004367835	1
BARX2, E2F1, EBF1, MSX1, SP4, ZNF300	ENST00000481907	1
BARX2, E2F1, ERG, HOXA9, LEF1, SOX5, TEAD4	ENST00000470033	1
BARX2, E2F1, ERG, LEF1, MYT1, SOX5, SOX6, TEAD4	ENST000004307765	1
BARX2, E2F1, ERG, LEF1, SOX5, SOX6, TEAD4	ENST000004435452, ENST00000423548	2
BARX2, EOMES, ERG, ETV1, GATA4, GRHL2, IRF4, LEF1, PLAG1, SALL2, TEAD4, ZKSCAN3	ENST00000447800	1
BARX2, ERG, ETV1, HIC1, NRL, PLAG1, SALL1, SOX3, ZKSCAN3	ENST00000405626	1
BARX2, ERG, ETV1, HIC1, PLAG1, SALL1, SOX3, ZKSCAN3	ENST000004294954	1
BARX2, ERG, ETV1, HIC1, PLAG1, ZKSCAN3	ENST00000403273	1
BARX2, ERG, ETV1, HIC1, ZKSCAN3	ENST000004344775	1
BARX2, ERG, ETV1, HNF4A, LEF1, PLAG1, SALL2, TEAD4, ZKSCAN3	ENST00000477771	1
BARX2, ERG, ETV1, IRF4, LEF1, PLAG1, SALL2, TEAD4, ZKSCAN3	ENST000004371243	1
BARX2, ERG, ETV1, LEF1, PLAG1, SALL2, TEAD4, ZKSCAN3	ENST000004371253, ENST00000471487	2
BARX2, ERG, ETV1, PLAG1, ZNF300	ENST000004373705	1
BARX2, ERG, GATA3, HOXD10, LEF1, TEAD4	ENST000004358749	1
BARX2, ERG, LEF1, SOX5	ENST000004395942	1
BARX2, ERG, MSX2, PLAG1, SALL2, SOX6, TEAD4, WT1, ZKSCAN3	ENST000004421134	1
BARX2, ERG, PLAG1, SP4	ENST00000447615	1
BARX2, ETV1, FOXP2, HOXA9, IRX5, MSX2, PLAG1, SALL1, SOX5	ENST000004045249	1
BARX2, ETV1, HOXA9, PLAG1, SALL1, SOX5	ENST000004399220	1
BARX2, ETV1, HOXA9, PLAG1, SALL1, SOX5	ENST00000444921	1
BARX2, ETV1, PLAG1, SALL1, SOX5	ENST000004341617	1
BARX2, ETV1, SOX6	ENST00000415392	1
BARX2, FOXP2, HOXA4, HOXA9	ENST000004394141	1
BARX2, GATA4, LEF1, SOX9	ENST00000469377	1
BARX2, HNF4G, HOXA3, HOXA8, SOX6	ENST000004310357	1
BARX2, HOXA3, LEF1, SOX9	ENST000004473719	1
BARX2, HOXA8, IRF4, LEF1, MYT1L, PAX8, SOX5	ENST000004366834	1
BARX2, HOXD10	ENST000004455714	1
BARX2, LEF1, SOX5, YY2	ENST000004399821	1
BARX2, PLAG1	ENST000004370648	1
BARX2, PLAG1, ZIC2	ENST000004181177	1
BCL6, CREB1, MEIS1, STAT3, STAT5A, STAT5B, TCF12, ZBTB3	ENST000004305249	1
BCL6, HHEX, MZF1, SMAD3, SP1	ENST000004306051	1
CARF	ENST000004467632	1
CARF, CDX1, DLX3, ERG, FOXA1, GRHL2, RFX4, SOX3, SOX6	ENST000004467632	1
CARF, CDX1, ERG, GATA1	ENST000004383673	1
CARF, DLX1, HOXB4, HOXA4, LBX2, OSR2, SALL2, SIX4, ZKSCAN3	ENST000004355983	1
CARF, E2F1, EBF1, ETV1, PLAG1, SALL2, ZNF300	ENST000004367126	1
CARF, E2F1, EBF1, HIC1, PLAG1, SALL2, ZKSCAN3, ZNF300	ENST000004328544	1
CARF, E2F1, EBF1, RXF3, ZNF300	ENST000004374312	1
CARF, E2F1, EBF1, SALL2, SP4, ZKSCAN3, ZNF300	ENST000004382315	1
CARF, E2F1, ERG, ETV1, GATA2, PLAG1, SALL2, SOX5, SOX6, ZKSCAN3	ENST000004394720	1
CARF, E2F1, ERG, ETV1, GLI2, IRF4, PLAG1, SALL2, SP4, ZKSCAN3, ZNF300	ENST00000402066	1
CARF, E2F1, ERG, ETV1, SALL2, ZKSCAN3	ENST00000439764	1
CARF, E2F1, ERG, GLI3, ZNF300	ENST000004231188	1
CARF, E2F1, ETV1, GLI2, GLI3, PLAG1, SALL2	ENST000004332582	1

Transcription factors	IsoformID	# of Isoforms
CARF, E2F1, ETV1, LEF1, SOX6, ZKSCAN3	ENST00000398538	1
CARF, E2F1, ETV1, SALL2	ENST00000417980	1
CARF, E2F1, GATA3, GLIS3, SALL2, SIX4, ZKSCAN3	ENST00000218316	1
CARF, E2F1, MSX1	ENST00000349767, ENST00000393592	2
CARF, E2F1, PLAG1, ZNF300	ENST00000496258, ENST00000496497	2
CARF, E2F1, PLAG1L, SOX6, ZNF300	ENST00000419928	1
CARF, E2F1, PLAG1L, ZNF300	ENST00000464585	1
CARF, E2F1, SALL2, ZKSCAN3, ZNF300	ENST00000424232	1
CARF, E2F1, ZNF300	ENST00000329797	1
CARF, EBF1, GATA1, LEF1, SALL2, SOX6, ZKSCAN3	ENST00000498400	1
CARF, EBF1, LEF1, PLAG1, SALL2, SIX4, SPZ1, WT1, ZIC2, ZKSCAN3, ZNF300	ENST00000440923, ENST00000467425	2
CARF, EBF1, LEF1, PLAG1, SALL2, SIX4, WT1, ZKSCAN3, ZNF300	ENST00000389336	1
CARF, EBF1, LEF1, PLAG1, SALL2, ZIC2, ZKSCAN3, ZNF300	ENST00000389335	1
CARF, EBF1, PLAG1, SALL2, SP4, ZNF300	ENST00000358927	1
CARF, ERG, ETV1, PLAG1, SALL2, SOX6, SP4, ZFP57, ZNF300	ENST00000382487	1
CARF, ERG, PLAG1, SIX4, TEAD4, YY2, ZKSCAN3	ENST00000472248	1
CARF, ERG, PLAG1, TEAD4, ZKSCAN3	ENST00000436738	1
CARF, ETV1, GLI3, NRL, SALL2, ZKSCAN3, ZNF300	ENST00000327655	1
CARF, ETV1, GLI3, SALL2, ZNF300	ENST00000349457	1
CARF, ETV1, GLIS2, HIC1, PLAG1, RFX3, ZNF300	ENST00000494542	1
CARF, ETV1, HIC1, LEF1,	ENST00000457814	1
CARF, ETV1, HLF, PLAG1, SALL2	ENST00000333172	1
CARF, ETV1, PLAG1	ENST000003111131	1
CARF, GATA2, HIC1, LEF1, SALL2, ZKSCAN3, ZNF300	ENST00000341938	1
CARF, GATA4, GLI3, SALL2, SOX3, YY2, ZNF300	ENST00000394361	1
CARF, GH11, HOXB4, HOXC4, LBX2, OSR2, SALL2, ZKSCAN3	ENST00000463197	1
CARF, GLI3, SALL2, ZNF300	ENST00000340339	1
CARF, HIC1, MSX1, SOX5	ENST00000284311	1
CARF, HIC1, PLAG1, RFX3, ZNF300	ENST00000342292	1
CARF, HIC1, SP4	ENST00000392965	1
CARF, HOXB4, LBX2, OSR2, SALL2, SIX4, SOX6, ZKSCAN3	ENST00000422395	1
CARF, LEF1, ZKSCAN3	ENST00000420125	1
CARF, PLAG1, SALL2	ENST00000298171, ENST00000393596	2
CARF, PLAG1L, SALL2, ZKSCAN3	ENST00000342443	1
CARF, TEAD4	ENST00000373693	1
CARF, TEAD4, ZNF300	ENST00000373691	1
CDX1, CDX2, ERG, ETV1, GATA4, LEF1, SOX6	ENST00000273352	1
CDX1, CDX2, LEF1, SOX2	ENST00000422265	1
CDX1, ERG, GF1, GLI3, HOXA10, NKX25, PLAG1, SOX3, YY2	ENST00000490186	1
CDX1, GATA2, HOXC13, SOX6	ENST00000487573	1
CDX1, GATA4, MYT1, PAX3, SOX5, SOX6	ENST00000252560	1
CDX1, GATA3, LEF1, SOX5, SOX6	ENST00000469138	1
CDX2, ERG, GATA3, GLI2	ENST00000315576	1
CEBPB, CREB1, HHEX, NF1, SF1, SMAD3, USF1	ENST00000330953	1
CEBPB, E2F4, ELK1, KLF7, MEIS1, PBX3, SMAD3, SP1, ZBTB3, ZNF263	ENST00000349243, ENST00000407454, ENST00000497524	3
CEBPB, EGR1, KLF6, MZF1, SRF, TCF7L2, ZBTB3, ZEB1, ZNF219	ENST00000309502	1
CEBPB, ELK1, HHEX, HSF2, MZF2, SATB1, SMAD3, YY1, ZNF35	ENST00000307637	1
CEBPB, ELK1, HHEX, HSF2, MZF2, SP1, ZNF35	ENST00000318507	1
CEBPB, HES1, KAISO, PBX3, ZBTB7, ZFX, ZNF354C	ENST00000240093	1
CEBPB, HHEX, MZF1, SATB1	ENST00000226413, ENST00000420975	2
CEBPB, HSF2, MEIS1, MZF1, SMAD3, SP1, STAT3, ZNF263	ENST00000473675	1
CEBPB, MZF1, SP1, ZBTB7, ZNF263, ZNF35	ENST00000217270	1
CEPBD, ELK1, KLF6, MZF2, ZNF219, ZNF35, ZNF354C	ENST00000255758	1
CEPBD, FOXP1, HHEX, HSF2, JUND, KAISO, MZF1, SF1, SMAD3, SP1, ZNF263, ZNF35	ENST00000309178	1
CREB1, FOXP1, HSF1, HSF2, MZF2, SP1, ZEB1	ENST00000442925	1
CTCF, HSF2, MEIS1, TCF12, ZBTB3, ZEB1, ZNF35	ENST00000367282	1
DBP, HIC1	ENST00000454971	1
DLX1, ERG, ETV1, LEF1, LHX6, MYT1L, RFX4, SOX5, SOX6	ENST00000447403	1
DLX1, ETV1, GATA3, HOXA3, HOXC4, LBX2, LMX1B, MSX1, MSX2, PLAG1	ENST00000276218	1
DLX1, ETV1, HIC1, HOXC8, MEOX1, PAX2, PLAG1, SP4, WT1, YY2, ZNF300	ENST00000377888	1
DLX1, ETV1, HIC1, MEOX1, PLAG1, SP4, YY2, ZFP57, ZNF300	ENST00000314512	1
DLX1, ETV1, PLAG1, SP4, ZFP57, ZNF300	ENST00000360693	1
DLX1, GATA4, INSM1, MSX2, SALL2, SOX6, TEAD4, ZIC2	ENST00000257089	1
DLX3, FOXA1, GATA2, GATA4, IRF4, TEAD4, YY2	ENST00000441140	1
DLX3, HOXA3, HOXC8, SOX3, SOX5	ENST00000293493	1
E2F1	ENST00000397095, ENST00000450460	2
E2F1, E2F2, EOMES, MSX2, SALL2, WT1	ENST00000453373	1
E2F1, EBF1, ETV1, GLIS3, HIC1, ZNF300	ENST00000370501	1
E2F1, EBF1, ETV1, HIC1, PLAG1, PLAG1L, SALL2, WT1, ZKSCAN3, ZNF300	ENST00000340171	1
E2F1, EBF1, ETV1, HIC1, PLAG1, PLAG1L, SALL2, ZFP57, ZKSCAN3, ZNF300	ENST00000317393	1
E2F1, EBF1, ETV1, PLAG1, SALL2, ZNF300	ENST00000339582	1
E2F1, EBF1, ETV1, SALL2, SIX4, ZNF300	ENST00000472701	1
E2F1, EBF1, GLIS2, SALL2, SP4, ZKSCAN3, ZNF300	ENST00000405846	1
E2F1, EBF1, MSX1, SP4	ENST00000464295	1
E2F1, EBF1, PLAG1, PLAG1L, SALL2, SP4, ZKSCAN3, ZNF300	ENST00000430002	1
E2F1, EBF1, PLAG1, SALL2, SOX6, ZKSCAN3, ZNF300	ENST00000377034	1
E2F1, EBF1, PLAG1, SALL2, ZKSCAN3, ZNF300	ENST00000376977	1
E2F1, EBF1, PLAG1, SALL2, ZKSCAN3, ZNF300	ENST00000472823	1
E2F1, EBF1, SALL2, SOX6, ZKSCAN3, ZNF300	ENST000003377016	1
E2F1, EBF1, SOX6	ENST00000303786	1
E2F1, EBF1, ZNF300	ENST00000395964	1
E2F1, EGR2, EGR3, ETV1, HIC1, SALL2, SP4, ZNF300	ENST00000414642	1
E2F1, EGR3, FOXA1, HIC1, INSM1, PLAG1, SP4, ZIC2, ZNF300	ENST00000242786	1
E2F1, ERG, ETV1, SALL2, SP4, ZNF300	ENST00000444078	1
E2F1, ERG, ETV1, SALL2, ZKSCAN3	ENST00000362027	1
E2F1, ERG, ETV1, SALL2, ZKSCAN3, ZNF300	ENST00000346454, ENST00000355319	2
E2F1, ERG, HOXB9, LEF1, SOX6, TEAD4	ENST00000471616	1
E2F1, ERG, LEF1, TEAD4	ENST00000448688	1
E2F1, ERG, PLAG1, SALL2, SP4	ENST00000438398	1
E2F1, ERG, SALL2, SP4, ZNF300	ENST00000392040	1
E2F1, ETV1	ENST00000230173	1
E2F1, ETV1, FOXP2, HIC1, LEF1, NKX25, PLAG1, PLAG1L, SALL2, SP4, ZKSCAN3	ENST00000277942	1
E2F1, ETV1, FOXP2, IRF4, LEF1, ZKSCAN3, ZNF300	ENST00000481661	1
E2F1, ETV1, GATA2, GLIS3, PLAG1, RREB1, SALL2, WT1, ZKSCAN3	ENST00000392629	1
E2F1, ETV1, GLI3, PLAG1, SALL2, SP4, WT1, ZKSCAN3, ZNF300	ENST00000421592	1
E2F1, ETV1, GLI3, PLAG1, SALL2, ZKSCAN3, ZNF300	ENST00000358883	1
E2F1, ETV1, GLI3, SALL2, SOX6, SP4, ZNF300	ENST00000374431	1
E2F1, ETV1, HIC1	ENST00000267377	1
E2F1, ETV1, HIC1, LEF1, PLAG1, SIX4, WT1, ZKSCAN3	ENST00000369768	1
E2F1, ETV1, HIC1, PLAG1, SALL2, TEAD4	ENST00000272644	1
E2F1, ETV1, INSM1, MYF6, PLAG1, SALL2, WT1, ZKSCAN3, ZNF300	ENST00000426151	1
E2F1, ETV1, MSX1, SALL2, SP4, ZKSCAN3, ZNF300	ENST00000367685	1
E2F1, ETV1, PLAG1, PLAG1L, ZKSCAN3, ZNF300	ENST000003034374	1
E2F1, ETV1, PLAG1, SALL2, ZFP57, ZNF300	ENST00000352855	1
E2F1, ETV1, PLAG1, SALL2, ZKSCAN3	ENST00000367838	1
E2F1, ETV1, PLAG1, SALL2, ZNF300	ENST00000314537	1
E2F1, ETV1, PLAG1, ZNF300	ENST00000415128	1
E2F1, ETV1, SALL2, SOX6, ZKSCAN3, ZNF300	ENST00000229955, ENST00000359558, ENST00000369501	3
E2F1, ETV1, SALL2, SOX9, ZKSCAN3, ZNF300	ENST00000369502	1
E2F1, ETV1, SALL2, ZKSCAN3	ENST00000359573, ENST00000367836	2
E2F1, ETV1, TEAD4	ENST00000287907	1
E2F1, ETV1, ZNF300	ENST00000436487	1
E2F1, ETV1, ZTF1, TEAD4, ZNF300	ENST00000464971	1
E2F1, GLI3	ENST00000439971, ENST00000446673	2
E2F1, GLS2, SALL2, SOX6, SP4	ENST00000367235	1
E2F1, GLS2, SALL2, SOX6, SP4	ENST00000381436	1

Transcription factors	IsoformID	# of Isoforms
E2F1, GLI3, PLAGL1, SALL2, ZKSCAN3, ZNF300	ENST00000329476	1
E2F1, HIC1, PLAG1, SPZ1, TEAD4, ZNF300	ENST00000229768	1
E2F1, HIC1, PLAG1, TEAD4, ZNF300	ENST00000337049	1
E2F1, HIC1, PLAG1, ZNF300	ENST00000397688	1
E2F1, HIC1, SALL2, TEAD4	ENST00000305097	1
E2F1, HIC1, SALL2, WT1, ZKSCAN3, ZNF300	ENST00000296479, ENST00000475478	2
E2F1, HIC1, SALL2, ZKSCAN3, ZNF300	ENST00000412232	1
E2F1, HOXA4, INSM1, PLAG1, PLAGL1, PPARG, WT1, ZNF300	ENST00000401670	1
E2F1, INSM1, PLAG1, WT1, ZKSCAN3, ZNF300	ENST00000349451	1
E2F1, LEF1, NRL, PLAG1, SP4, ZFP57	ENST00000352846	1
E2F1, LEF1, PLAG1	ENST0000030019103	1
E2F1, LEF1, PLAG1, SALL2, ZKSCAN3	ENST00000305344	1
E2F1, LEF1, PLAG1, SALL2, ZKSCAN3	ENST00000369584	1
E2F1, LEF1, PLAG1, SP4	ENST00000459916, ENST00000498475	2
E2F1, LEF1, PLAG1, ZKSCAN3, ZNF300	ENST00000255008	1
E2F1, LEF1, SP4	ENST00000408626	1
E2F1, LEF1, WT1	ENST00000369169	1
E2F1, MSX2, NKX25, SALL2, WT1	ENST00000401881, ENST00000402486	2
E2F1, OSR2, SALL2, TEAD4, ZKSCAN3, ZNF300	ENST00000334619	1
E2F1, PAZ2, PLAG1, SALL2	ENST000004686521	1
E2F1, PLAG1	ENST000004098619	1
E2F1, PLAG1, SALL2	ENST00000357991	1
E2F1, PLAG1, SALL2, WT1	ENST00000355631	1
E2F1, PLAG1, SALL2, WT1, ZKSCAN3, ZNF300	ENST00000379873	1
E2F1, PLAG1, SALL2, ZKSCAN3, ZNF300	ENST00000275169, ENST00000411851	2
E2F1, PLAG1, SALL2, ZKSCAN3, ZNF300	ENST000003536866	1
E2F1, PLAG1, SP4, WT1, ZFP57, ZKSCAN3, ZNF300	ENST00000257070	1
E2F1, PLAG1, SP4, ZFP57, ZKSCAN3, ZNF300	ENST000003598547	1
E2F1, PLAG1, SP4, ZNF300	ENST00000373655	1
E2F1, PLAG1, TEAD4, ZNF300	ENST00000330432	1
E2F1, SALL2	ENST00000320658, ENST00000397092, ENST00000480591,	5
E2F1, SALL2, SOX6	ENST00000481598, ENST00000486108	2
E2F1, SALL2, ZKSCAN3	ENST00000357544	1
E2F1, SALL2, ZKSCAN3, ZNF300	ENST00000356606, ENST00000380587, ENST00000407714	3
E2F1, SP4, ZKSCAN3, ZNF300	ENST00000373658	1
E2F1, TEAD4, ZKSCAN3	ENST00000268572, ENST0000031049, ENST00000367609,	1
E2F1, ZNF300	ENST00000449846, ENST00000444262	5
E2F3, E2F4, HSF2, KLF7, MZF1, SP1, STAT3, ZNF219, ZNF263	ENST00000287324	1
E2F4, E2F6, EGR1, ELK1, KLF6, MEF2C, MZF1, PBX3, SP1, SPI1, ZBTB7, ZFX, ZNF219, ZNF263	ENST00000428220	1
E2F4, EGR1, ESRRA, KLF6, MZF1, PBX3, SMAD3, SP1, SPI1, STAT6, ZBTB3, ZEB1, ZFX, ZNF219, ZNF263	ENST00000263281	1
E2F4, EGR1, ETS1, KLF6, KLF7, MZF1, PBX3, SMAD3, SP1, SPI1, STAT6, ZBTB3, ZEB1, ZFX, ZNF219, ZNF263	ENST00000375190	1
E2F4, EGR1, HSF2, KLF6, KLF7, MZF1, PBX3, SMAD3, SMAD4, SP1	ENST00000303921	1
E2F4, EGR1, KLF6, KLF7, MZF1, SP1, SPI1, STAT3, ZBTB3, ZBTB7, ZNF219, ZNF263, ZNF35	ENST00000334304	1
E2F4, EGR1, KLF6, KLF7, MZF1, SP1, SPI1, ZNF219, ZNF263	ENST00000358157	1
E2F4, ELK1, HHEX, KLF6, MAZ, MEIS1, MZF1, SMAD3, SP1, ZBTB7, ZEB1, ZFX, ZNF219, ZNF263	ENST00000447924	1
E2F4, HHEX, HSF2, KLF7, MAZ, MZF1, SMAD3, SP1, SPI1, ZBTB7, ZFX, ZNF263	ENST00000334286, ENST00000446573	2
E2F4, HHEX, KLF7, MAZ, MZF1, NFKAPPAB, PBX3, SPI1, ZBTB7, ZNF219, ZNF263, ZNF354C	ENST00000246549	1
E2F4, HHEX, KLF7, MEIS1, MZF1, NFKAPPAB, PBX3, SPI1, ZBTB7, ZNF219, ZNF263	ENST00000304166, ENST00000409363	2
E2F4, HHEX, KLF7, MEIS1, SPI1, TCF12, USF1, ZBTB3, ZNF354C	ENST00000374794	1
E2F6, FOXP1, MZF1, NRF1, PBX3, ZBTB7, ZFX	ENST00000355897, ENST00000492492	2
EBF1	ENST00000305141	1
EBF1, EGR3, ERG, HIC1, LEF1, ZKSCAN3	ENST00000241393, ENST00000262738, ENST00000271332, ENST0000029661, ENST00000299727, ENST00000300571,	8
EBF1, EGR, ETV1, INSM1, PLAG1, PLAGL1, SOX6, SP4, ZKSCAN3, ZNF300	ENST00000307161	1
EBF1, EGR, ETV1, PLAG1, PLAGL1, SOX6, SP4, ZKSCAN3	ENST00000355082	1
EBF1, EGR, ETV1, PLAG1, SP4	ENST00000351052	1
EBF1, EGR, ETV1, PLAG1, PLAGL1, SOX6, SP4, ZKSCAN3, ZNF300	ENST00000356595, ENST00000361210, ENST00000370932,	4
EBF1, EGR, ETV1, PLAG1, PLAGL1, SP4, ZKSCAN3, ZNF300	ENST00000448819	1
EBF1, EGR, ETV1, PLAG1, PLAGL1, SP4, ZKSCAN3	ENST00000460330	1
EBF1, EGR, ETV1, PLAG1, SALL2, SOX6, SP4, ZKSCAN3	ENST00000370931, ENST00000479353	2
EBF1, EGR, ETV1, PLAG1, SP4	ENST00000306666, ENST00000370924	2
EBF1, ETV1, HIC1, INSM1, LEF1, PAX3, PLAG1	ENST00000423019	1
EBF1, ETV1, HIC1, TEAD4	ENST00000369716	1
EBF1, ETV1, SALL2, ZNF300	ENST00000329321	1
EBF1, HIC1	ENST00000246115, ENST00000337474, ENST00000402260	3
EBF1, HIC1, HOXA10, LEF1, SALL2, SOX5, SOX6, ZKSCAN3	ENST00000341113	1
EBF1, HIC1, LEF1, SALL2, SOX5, ZKSCAN3	ENST00000374619	1
EBF1, HIC1, PLAG1, SALL2, SP4, ZKSCAN3	ENST00000289753	1
EBF1, HIC1, SOX6	ENST00000377653	1
EBF1, HIC1, TEAD4	ENST00000241356	1
EBF1, LEF1	ENST00000241356	1
EBF1, LEF1, SALL2	ENST00000357716	1
EBF1, LEF1, SALL2, ZKSCAN3, ZNF300	ENST00000403881	1
EBF1, PLAG1, SALL2	ENST00000412160	1
EBF1, PLAG1, ZNF300	ENST00000457830	1
EBF1, PLAG1, SP4	ENST00000295417, ENST00000427125	2
EBF1, SALL2	ENST00000249016	1
EBF1, SALL2, SP4	ENST000002428545	1
EBF1, TEAD4	ENST00000014914, ENST00000311268	2
EGR3, HIC1, PLAG1, SP4, ZNF300	ENST00000243673, ENST00000249373	2
ELF2, KASO, MZF1, NF1, SMAD3, ZBTB3, ZFX, ZNF35	ENST00000357355	1
ELK1, HSF2, MAZ, MZF1, SF1, SPI1, STAT3, ZNF263, ZNF35	ENST00000248076	1
EOMES, FOXA1, GFI1, PAX2, PLAG1	ENST00000282753	1
EOMES, MSX2, SIX4, SOX6, ZKSCAN3	ENST00000462632	1
EOMES, MYT1, SALL2, SIX4, SPZ1, WT1, ZKSCAN3	ENST00000452454	1
EOMES, MYT1, SALL2, SIX4, SPZ1, WT1, ZKSCAN3	ENST00000427970	1
ERG	ENST00000258400, ENST00000267549, ENST00000282018, ENST00000291294, ENST00000296641, ENST00000304402, ENST00000305180, ENST0000034286, ENST00000323040, ENST00000349057, ENST00000349157, ENST00000349159, ENST00000349160, ENST00000370789, ENST00000370713, ENST00000373304, ENST00000373307, ENST00000396782, ENST00000407325, ENST00000445132	20
ERG, ETV1	ENST00000249887, ENST00000424798, ENST00000426880	3
ERG, ETV1, GATA2, OTX2, YY2, ZFP57, ZKSCAN3	ENST00000281806	1
ERG, ETV1, GATA2, PLAGL1, SALL2, SOX6	ENST00000361668	1
ERG, ETV1, GATA2, ZKSCAN3	ENST00000369212	1
ERG, ETV1, GATA3, HOXA3, HOXD3, IRF4, LEF1, LHX6, SOX5, SOX6	ENST00000461244	1
ERG, ETV1, GATA3, HOXA3, LEF1, LHX6, SOX5, SOX6	ENST00000479784	1
ERG, ETV1, GATA3, PLAG1, SALL2, SOX6, SP4, ZFP57	ENST00000339690	1
ERG, ETV1, GATA6, LEF1	ENST00000329773	1
ERG, ETV1, GCM1, HIC1, PLAG1, TEAD4, ZKSCAN3	ENST00000409489	1
ERG, ETV1, HIC1	ENST00000255262	1
ERG, ETV1, HIC1, INSM1, MYT1, PLAG1, SOX6, TBX5, TEAD4, ZKSCAN3	ENST00000396211	1
ERG, ETV1, HIC1, LEF1, PLAG1, SALL2, SIX4, ZKSCAN3, ZNF300	ENST00000431811	1
ERG, ETV1, HIC1, SP4	ENST00000444847	1
ERG, ETV1, HIC1, ZKSCAN3,	ENST00000401907, ENST00000428232	2
ERG, ETV1, LEF1	ENST00000409988	1
ERG, ETV1, LEF1, LHX6	ENST00000485973	1
ERG, ETV1, LEF1, PLAG1, SALL2, TEAD4, ZKSCAN3	ENST000004491283	1
ERG, ETV1, LEF1, SALL2, TEAD4	ENST00000419692	1
ERG, ETV1, PLAG1	ENST00000337306	1
ERG, ETV1, PLAG1, SALL2	ENST00000367236	1

Transcription factors	IsoformID	# of Isoforms
ERG, ETV1, SALL2	ENST00000265417, ENST00000377291	2
ERG, ETV1, SALL2, SP4	ENST00000324300, ENST00000493800	2
ERG, ETV1, SP4	ENST00000271357	1
ERG, ETV1, ZKSCAN3	ENST00000445970	1
ERG, GATA3, GLIS2	ENST00000346057, ENST00000392967	2
ERG, GF11, LEF1, SOX6	ENST00000335786	1
ERG, GLB3, PLAG1	ENST00000296154	1
ERG, GLIS2, LEF1, PLAG1, SALL2, SOX9, ZFP57	ENST00000436083	1
ERG, GLIS2, SAI1, SOX9	ENST00000432564	1
ERG, HIC1, LEF1, SOX6, ZNF300	ENST00000461057	1
ERG, HIC1, SOX9, TEAD4	ENST000003380289	1
ERG, HIC1, SP4	ENST00000481403	1
ERG, HOXA3	ENST00000359929	1
ERG, HOXA3, HOXB5, LHX6	ENST00000450180	1
ERG, HOXA3, IRF4, MSX2, NRL, PAX2, PLAG1, SIX4, SOX9, WT1, YY2, ZIC2, ZKSCAN3	ENST00000302797	1
ERG, HOXA3, LEF1, MSX1, TEAD4	ENST00000428565	1
ERG, HOXA3, SOX5	ENST00000437420	1
ERG, HOXB5, SOX5	ENST00000467202	1
ERG, LEF1	ENST00000493284	1
ERG, LEF1, MSX1	ENST00000371906, ENST00000480322	2
ERG, LEF1, SOX6	ENST00000319517	1
ERG, LHX6	ENST00000392370, ENST00000410068, ENST00000447769	3
ERG, OSR2	ENST00000331251	1
ERG, SALL2	ENST00000393590	1
ERG, SALL2, SOX9	ENST00000356987	1
ERG, SALL2, SP4	ENST00000358556	1
ERG, SALL2, TEAD4	ENST00000447845	1
ERG, SOX6	ENST00000256906, ENST00000271029	2
ERG, SOX6, ZKSCAN3	ENST00000496754	1
ERG, SP4	ENST00000262178, ENST00000316793, ENST00000431493	3
ERG, TEAD4	ENST00000319595, ENST00000332438	2
ESRRA, HHEX, HSF2, MZF1, SP1, TCF12, ZBTB3, ZBTB7, ZNF35	ENST00000375846	1
ETV1	ENST00000284674, ENST00000296932, ENST00000344373, ENST00000367606, ENST00000389832, ENST00000461541	6
ETV1, GATA3	ENST00000177177, ENST00000475374	2
ETV1, GATA3, IRF4, SALL2, SOX6	ENST00000361719	1
ETV1, GATA3, SALL2, SP4, ZKSCAN3	ENST00000374177	1
ETV1, GF11, HIC1, SOX6, ZNF300	ENST00000418619	1
ETV1, GF11, LEF1, PLAG1, ZNF300	ENST00000381407	1
ETV1, GLI3, HIC1, PLAG1, SALL2, TEAD4, ZKSCAN3	ENST00000304244	1
ETV1, GLI3, SALL2	ENST00000490131	1
ETV1, GLS2, HIC1, SALL2, TEAD4	ENST00000309099	1
ETV1, GLS3, SALL2, ZKSCAN3	ENST00000246553	1
ETV1, HIC1, HNF4G, SALL2, SOX6, ZKSCAN3	ENST00000403859	1
ETV1, HIC1, HOXA3	ENST00000440082	1
ETV1, HIC1, LEF1	ENST00000373642	1
ETV1, HIC1, LEF1, SOX6, SOX9	ENST00000314254	1
ETV1, HIC1, MEOX1, PLAG1, SP4	ENST00000362016	1
ETV1, HIC1, MEOX1, PLAG1, SP4, ZNF300	ENST00000354217	1
ETV1, HIC1, MYT1, PLAG1, TEAD4	ENST00000358373	1
ETV1, HIC1, SALL2, SOX6, TEAD4	ENST00000486700	1
ETV1, HIC1, SALL2, TEAD4	ENST00000496086	1
ETV1, HIC1, SALL2, ZKSCAN3	ENST00000438013	1
ETV1, HOXA3	ENST00000312457	1
ETV1, HOXB4, HOXC4, LBX2, LHX6, SALL2	ENST00000287534, ENST00000412101	2
ETV1, INSM1, LEF1, OSR1, PLAG1, SALL2, TEAD4, ZFP57, ZKSCAN3	ENST00000371950	1
ETV1, LEF1	ENST00000409848	1
ETV1, LEF1, OSR1, PLAG1, SALL2, TEAD4, WT1, ZKSCAN3	ENST00000371951	1
ETV1, LEF1, PLAG1, SOX6, ZNF300	ENST00000392039	1
ETV1, LEF1, ZNF300	ENST00000381404	1
ETV1, PLAG1, SALL2	ENST00000398307	1
ETV1, PLAG1, SALL2, SOX5, ZKSCAN3, ZNF300	ENST00000337894	1
ETV1, PLAG1, SOX6	ENST00000397473	1
ETV1, PLAG1, SP4	ENST00000436464	1
ETV1, PLAG1, TEAD4, ZNF300	ENST00000293897	1
ETV1, PLAG1, SALL2	ENST00000400723	1
ETV1, PLAG1, SP4	ENST00000272847	1
ETV1, PPARG	ENST00000357392, ENST00000400880	2
ETV1, SALL2, SOX6	ENST00000412814	1
ETV1, SALL2, TEAD4	ENST00000373256, ENST00000373306	2
ETV1, SALL2, ZKSCAN3	ENST00000394143	1
ETV1, SIX4, ZFP57	ENST00000436680	1
ETV1, SOX3, SPZ1, YY2	ENST00000332314	1
ETV1, TEAD4	ENST00000304421, ENST00000325123, ENST00000453237	3
ETV1, ZKSCAN3, ZNF300	ENST00000368318	1
FOSL1	ENST00000374024	1
FOSL1, HIC1	ENST00000377127	1
FOSL1, SOX9	ENST00000395716	1
FOXP2, GATA2, HOXA3, SOX5	ENST00000475347	1
FOXP2, GATA4, HOXC10, PLAG1, SOX3	ENST00000298110	1
FOXP2, SOX5	ENST00000474935	1
GATA1, GATA3, PLAG1, ZKSCAN1	ENST00000435307	1
GATA2, HIC1, LEF1, SALL2, ZKSCAN3	ENST00000349984	1
GATA2, SOX6	ENST00000297146	1
GATA2, TEAD4	ENST00000255432, ENST00000487787	2
GATA3, IRF4, LEF1, NRL	ENST00000474212	1
GATA3, LEF1	ENST00000397470	1
GATA3, SIX4	ENST00000401034	1
GF11, NFE2, TEAD4	ENST00000357422	1
GLI3	ENST00000370611	1
GLI3, HIC1, SALL2, TEAD4	ENST00000441623	1
GLI3, LEF1	ENST00000311104	1
GLI3, PLAG1, SALL2	ENST00000397547	1
GLI3, PLAG1, SALL2	ENST00000338566	1
GRHL1, PLAG1, SP4, ZNF300	ENST00000371721	1
GRHL2, LBX2, LEF1, PLAG1, SALL2	ENST00000292303	1
HPB1, HHEX, MEIS1, MZF1, ZNF263	ENST00000300098	1
HPB1, MZF1, SATB1, SP1	ENST00000494619	1
HHEX, MEIS1, MZF1, RFX8, SF1	ENST00000378688	1
HHEX, MEIS1, PBX3, STAT3, ZNF219	ENST00000376414	1
HHEX, MEIS1, PBX3, ZNF35	ENST00000339223	1
HHEX, MZF1, SF1, SMAD4, TCF12, ZBTB3	ENST00000354550	1
HHEX, MZF1, SMAD4, STAT3, YY1, ZBTB3	ENST00000309170	1
HIC1	ENST00000258456, ENST00000266674, ENST00000292174, ENST00000296533, ENST00000297468, ENST00000306005, ENST00000312143, ENST00000327809, ENST00000381297, ENST00000397088, ENST00000458629	11
HIC1, LEF1, PLAG1	ENST00000435290	1
HIC1, LEF1, PLAG1, ZKSCAN3, ZNF300	ENST00000449615	1
HIC1, PLAG1, SALL2, SOX6, ZNF300	ENST00000258042	1
HIC1, PLAG1, SALL2, ZNF300	ENST00000376351	1
HIC1, PLAG1, SOX9	ENST00000308928	1
HIC1, PLAG1, SP4	ENST00000497923	1
HIC1, PLAG1, SP4	ENST00000255380, ENST00000267015, ENST00000319211	3
HIC1, PLAG1, SP4	ENST00000176183	1
HIC1, SAL1	ENST00000319638, ENST00000369947, ENST00000377741, ENST00000393018, ENST00000423776	5
HIC1, SALL2, SOX9	ENST00000329850	1

Transcription factors	IsoformID	# of Isoforms
HIC1, SALL2, TBX5, ZKSCAN3, ZNF300	ENST00000462882	1
HIC1, SALL2, TEAD4, ZKSCAN3	ENST00000492609	1
HIC1, SALL2, ZKSCAN3, ZNF300	ENST000003248438, ENST00000452278	2
HIC1, SIX4, TBX5, ZNF300	ENST00000334810	1
HIC1, SP4	ENST00000164024, ENST00000238699, ENST00000286201, ENST00000302472, ENST00000358600, ENST00000370757, ENST00000377411, ENST00000380007, ENST00000392962, ENST00000466131	10
HIC1, SP4, ZNF300	ENST00000396275	1
HIC1, TEAD4	ENST00000262441, ENST00000299092, ENST00000324750	3
HIC1, ZNF300	ENST00000397278	1
HOXA9	ENST00000397056	1
HOXA9, HOXA8, SOX5	ENST00000339069, ENST00000371220	2
HOXA9, LEF1	ENST00000442933	1
HSF2, IRF7, TEF, ZBTB3	ENST00000299766	1
HSF2, KLF6, NF1, SMAD3, STAT3, STAT6, ZNF35	ENST00000362015, ENST00000456426	2
HSF2, MEIS1, MZF1, SMAD3, STAT3, ZFX, ZNF219	ENST00000216629	1
INSM1, MYT1L, NRL, PLAG1, SOX3, SP4, ZFP57, ZIC2, ZNF300	ENST00000336152	1
INSM1, MYT1L, NRL, PLAG1, SP4, ZFP57, ZIC2, ZNF300	ENST00000237179	1
IRF3, MZF1, SATB1, SP1, ZEB1	ENST00000295683	1
IRF4, PLAG1, SALL2, SOX6	ENST00000392299	1
LEF1	ENST00000250572, ENST00000482710	2
LEF1, LHX6	ENST00000231683	1
LEF1, PLAG1, SOX6	ENST00000416594	1
LEF1, SALL2, TEAD4	ENST00000469383	1
LEF1, SOX2	ENST00000497921	1
LHX6, TEAD4	ENST00000428600	1
MAZ, MEIS1, MZF1, SF1, SP1, TCF12, ZEB1, ZFX	ENST000003368121	1
MEIS1, STAT3, ZEB1, ZNF219	ENST00000445907	1
MSX1, NRL, SOX6, SOX9	ENST00000298386	1
MSX1, SOX5, SOX9	ENST000003380314	1
MYF6	ENST00000360249	1
MZF1, NF1, SATB1, SOX12, SP1, STAT5A	ENST00000311762	1
MZF1, SMAD3, SMAD4, SP1, ZNF35	ENST00000374472	1
MZF1, SP1, SP1, ZBTB3, ZFX, ZNF35	ENST00000304222	1
MZF1, ZFX, ZNF219, ZNF263, ZNF35	ENST00000378661	1
NF1, PBX3, SF1, SMA03, ZNF263, ZNF35	ENST00000389858	1
NKX25	ENST00000442484	1
NRL, SOX6	ENST00000493193	1
OSR2, SOX5	ENST00000449020	1
PAX3, PLAG1, SALL2	ENST000004485464	1
PLAG1	ENST00000431922	1
PLAG1, SALL2	ENST00000218721	1
PLAG1, SALL2, SOX5, ZNF300	ENST00000439591	1
PLAG1, SIX4, SOX5, SOX6, TEAD4, ZNF300	ENST00000433647, ENST00000487545	2
PLAG1, SP4	ENST00000338052, ENST00000449591, ENST00000449735	3
PLAG1, YY2	ENST00000359305	1
PLAGL1	ENST00000259456, ENST00000330055, ENST00000340736, ENST00000245363, ENST00000361434	5
PLAGL1, PPARG	ENST00000315215	1
PPARG	ENST00000339036, ENST00000433848	2
RFX3	ENST00000334425	1
RFX4	ENST000002433801	1
SALL2	ENST00000246657, ENST00000278393, ENST00000298440, ENST00000304552, ENST00000306960, ENST00000333493, ENST00000375466, ENST00000377102, ENST00000380582, ENST00000380586, ENST00000393752, ENST00000395963, ENST00000407877, ENST00000409817, ENST00000430267, ENST00000433765, ENST00000438284, ENST00000463993	18
SALL2, SOX9	ENST00000331638	1
SALL2, SP4	ENST00000333430	1
SALL2, TEAD4	ENST00000404129	1
SALL2, ZNF300	ENST00000297469	1
SIX4, SOX5, ZKSCAN3	ENST000003395940	1
SIX4, SOX6	ENST00000494668	1
SIX4, ZKSCAN3	ENST00000374519	1
SOX5	ENST00000445772	1
SOX5, SOX9	ENST00000491034	1
SOX5, TEAD4	ENST00000327604	1
SOX6	ENST00000228887, ENST00000273145, ENST00000449420	3
SOX6, SP4	ENST00000340581	1
SOX9	ENST00000245457, ENST00000249041, ENST00000366554, ENST00000368122, ENST00000369499, ENST00000370742, ENST00000440886, ENST00000490673, ENST00000492666	9
SOX9, TEAD4	ENST00000370862	1
SP4	ENST00000246657, ENST00000277674, ENST00000305352, ENST00000309285, ENST00000315323, ENST00000357585, ENST00000358755, ENST00000394705, ENST00000424100, ENST00000441240	10
SP4, TEAD4	ENST00000377893	1
TEAD4	ENST00000276077, ENST00000296140, ENST00000304411, ENST00000304983, ENST00000315033, ENST00000315332, ENST00000373857, ENST00000378434, ENST00000400888, ENST00000438259	10
TEAD4, ZNF300	ENST00000465365	1
ZNF300	ENST00000329003	1

APENDIX B

Supplementary Tables

All supplementary materials can be found at the links provided below. They are presented in a PDF format. Interested party can contact the author of this dissertation for a downloadable spreadsheet format copy of all supplementary materials.

<http://csbl.bmb.uga.edu/publications/materials/dhammond/dissertationSupMaterials.html>