Human promoter characterization

Vanessa Louis
Clark Atlanta University

Follow this and additional works at: http://digitalcommons.auctr.edu/dissertations

Part of the Arts and Humanities Commons

Recommended Citation
HUMAN ID3 PROMOTER CHARACTERIZATION

Advisor: Jaideep Chaudhary, Ph.D.

Thesis dated May 2010

The basic-helix-loop-helix (bHLH) protein family plays an innate role in cellular activities. A pivotal member of this family, the Inhibitor of differentiation (Id) proteins, is linked to cell proliferation, differentiation, and tumorigenesis via inhibiting the DNA binding of bHLH proteins. There are four isoforms of Id genes - Id1-4 are uniquely expressed. In prostate cancer cells, Id3 is up-regulated and targets cellular proliferation and metastasis. However, the mechanism that triggers Id3 expression in prostate cancer is unknown. Therefore, a series of experiments were performed in order to investigate the transcription factors that regulate the expression of Id3 promoter in prostate cancer.

The proximal 1 kb human Id3 promoter was characterized and validated by a combination of in silico and in vitro approaches. Collective results led to the identification of YY1, C/EPBa and C/EPBb as putative transcription factors that could regulate Id3 promoter in prostate cancer. It is believed that further in depth studies of these transcription factors will lead to the identification of the underlying molecular mechanism that regulates Id3 expression in prostate cancer.
HUMAN ID3 PROMOTER CHARACTERIZATION

A THESIS
SUBMITTED TO THE FACULTY OF CLARK ATLANTA UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE MASTER'S DEGREE IN BIOLOGICAL SCIENCES

BY
VANESSA LOUIS

DEPARTMENT OF BIOLOGICAL SCIENCES

ATLANTA, GEORGIA
MAY 2010
ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. Jaideep Chaudhary for an opportunity to work in his lab, and for his guidance and support throughout the length of this degree. I would also like to thank Drs. Deborah Cook, David Logan, and Cass Parker for being on my committee and giving helpful and needed suggestions on my project. Next, I would like to thank Kisha Scarlett and Jason Carey for the tremendous amount of assistance that they provided me with while doing this project. Also, I would like to thank Dr. Khan's lab for their generosity when our lab supplies were low. In addition, I would like to thank Alfred Ghogomu for not only ordering the supplies I needed but for his helpful and guidance.

Very special thanks to Edward Lockhart, Lena Somers, Corey Neal, Nicole Strong, Lindsey Walker, Ashley Evans, Shanora Gymph, Tia Harmon, and Mandy Lucas for providing various venues to garner support and stress relieving activities. Also, thanks to Dr. Isabella Finkelstein, Martha Edwards, and Jacqui Ryder for advisement and continued financial assistance. Most importantly, I would like to thank my love ones, James and Lynda Louis, Chrystal Louis, and Demar Greenslade for emotional and financial support, positive reinforcement, and motivation. Thank you for the continuous love and support and for believing in me.
TABLE OF CONTENTS

ACKNOWLEDGEMENTS .................................................................................. ii
LIST OF FIGURES ......................................................................................... v
LIST OF TABLES ............................................................................................ vii
LIST OF ABBREVIATIONS ........................................................................... viii
Chapter One. Introduction ............................................................................. 1
   Hypothesis ................................................................................................. 4
   Specific Aims ........................................................................................... 4
Chapter Two. Literature Review .................................................................. 5
   Gene Regulation ....................................................................................... 5
   Gene Promoters ...................................................................................... 5
   TATA Box and TBP .................................................................................. 6
   Transcription Initiation Complex .............................................................. 7
   Transcription Factors .............................................................................. 7
   Ids in Cancer ........................................................................................... 13
   Ids in Prostate Cancer ........................................................................... 20
Chapter Three. Materials and Methods ........................................................ 22
   Background (Databases) ......................................................................... 22
   Laboratory Experimental Design ............................................................... 23
   Id3 Promoter Cloning in Reporter Plasmids ............................................ 25
   Bioinformatics ...................................................................................... 30
Chapter Four. Results .................................................................................. 33
   Promoter Identification .......................................................................... 33
   Promoter Confirmation and Verification ................................................ 36
   Transfection ........................................................................................... 39
   Luciferase Assay Results ....................................................................... 40
   Bioinformatics ...................................................................................... 40
Chapter Five. Discussion .............................................................................. 64
   Future Directions ................................................................................... 68
   Conclusion ............................................................................................... 69
LIST OF FIGURES

Figure

1. Eukaryotic gene promoter ................................................................. 6
2. Basic Helix Loop Helix Binding interaction between classes .................. 11
3. Primer sequences located in the human Id3 promoter ................................ 34
4. Gel image of Id3 promoter region ..................................................... 35
5. Nested PCR confirming the human Id3 promoter region .......................... 36
6. EcoRI restriction enzyme digest on the pGEMT-Easy plasmid with the Id3 promoter insert ................................................................. 37
7. Sequenced Id3 promoter ................................................................. 38
8. Promoter sequence verification via the NCBI blast program .................... 38
9. Gel confirmation of pGLuc+Id3 plasmid .............................................. 39
10. Most successful luciferase assay ..................................................... 40
11. Partial list of the transcription factor binding sites found in the Id3 promoter .......... 41
12. Schematic diagram of promoter region with transcription factor binding sites provided by the TESS database ....................................................... 43
13. Human, mouse, and rat phylogenetic tree ........................................ 43
14. Conserved regions in the mouse and rat Id3 promoters ......................... 45
15. Conserved regions in the human and rat Id3 promoters ........................ 46
16. Conserved regions in the human and mouse Id3 promoters ..................... 47
17. Schematic diagram of the putative transcription factor binding sites in the murine 200 bp Id3 promoter ................................................................. 48
18. Schematic diagram of the conserved regions in the human and mouse Id3 promoters. ................................................................. 49
LIST OF FIGURES

Figure

19. Microarray expression levels of GAPDH in prostate cancer cell lines................. 50

20. Raw values of Id3 expression in microarray experiments across prostate cancer cell lines PC3, DU145, and LNCaP ........................................................................................................ 51

21. Relative expression level of transcription factors predicted to bind to Id3 Promoter (DU145) .......................................................................................................................... 52

22. Relative expression level of transcription factors predicted to bind to Id3 Promoter (PC3) .......................................................................................................................................... 53

23. Relative expression level of transcription factors predicted to bind to Id3 Promoter (LNCap). ........................................................................................................................................ 53

24. Microarray correlation chart .................................................................................. 62

25. Conserved and aligned transcription factors that have been found in either the human or mouse Id3 promoters or Id3 microarray data.......................................................... 63
LIST OF TABLES

Table

1. Location and length of PCR primers within the Id3 promoter sequence .................. 35
2. Expected band sizes of the promoter in PCR reactions ......................................... 35
3. Transcription factor binding sites location in Id3 sequence .................................. 42
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Adenine</td>
</tr>
<tr>
<td>ACS</td>
<td>American Cancer Society</td>
</tr>
<tr>
<td>AR</td>
<td>Androgen Receptor</td>
</tr>
<tr>
<td>bHLH</td>
<td>Basic Helix Loop Helix</td>
</tr>
<tr>
<td>bp(s)</td>
<td>Base Pair(s)</td>
</tr>
<tr>
<td>C</td>
<td>Cytosine</td>
</tr>
<tr>
<td>C/EBP</td>
<td>CCAAT/enhancer binding protein</td>
</tr>
<tr>
<td>C/EBPa</td>
<td>CCAAT/enhancer binding protein alpha</td>
</tr>
<tr>
<td>C/EBPb</td>
<td>CCAAT/enhancer binding protein beta</td>
</tr>
<tr>
<td>CDK</td>
<td>Cyclin Dependent Kinases</td>
</tr>
<tr>
<td>CIAP</td>
<td>Calf Intestine Alkaine Phosphate</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic Acid</td>
</tr>
<tr>
<td>Emc</td>
<td>Extramacrochaetae</td>
</tr>
<tr>
<td>Evi-1</td>
<td>Ecotropic viral integration site 1</td>
</tr>
<tr>
<td>G</td>
<td>Guanine</td>
</tr>
<tr>
<td>GAPDH</td>
<td>Glyceraldehyde 3-phosphate dehydrogenase</td>
</tr>
<tr>
<td>HLH</td>
<td>Helix Loop Helix</td>
</tr>
<tr>
<td>Id(s)</td>
<td>Inhibitor of Differentiation</td>
</tr>
<tr>
<td>MAPK</td>
<td>Map Kinase</td>
</tr>
<tr>
<td>NFKappaB</td>
<td>Nuclear Factor Kappa-light-chain-enhancer of Activated B Cells</td>
</tr>
<tr>
<td>NIT2</td>
<td>Nitrilase Family, Member 2</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>M</td>
<td>Mitosis</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger Ribonucleic Acid</td>
</tr>
<tr>
<td>mVista</td>
<td>Comparative Alignment of Multiple Sequence Database</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>PSA</td>
<td>Prostate Specific Antigen</td>
</tr>
<tr>
<td>rVista</td>
<td>Regulatory Comparative Alignment of Multiple Sequence Database</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic Acid</td>
</tr>
<tr>
<td>SRF</td>
<td>Serum response factor</td>
</tr>
<tr>
<td>T</td>
<td>Thymine</td>
</tr>
<tr>
<td>TBP</td>
<td>TATA Binding Protein</td>
</tr>
<tr>
<td>TCR</td>
<td>T Cell Receptor</td>
</tr>
<tr>
<td>TESS</td>
<td>Transcription Search Element Database</td>
</tr>
<tr>
<td>TSH</td>
<td>Thyroid Stimulating Hormone</td>
</tr>
<tr>
<td>UCSC</td>
<td>University of California at Santa Clara</td>
</tr>
<tr>
<td>UTR</td>
<td>Untranslated Region</td>
</tr>
<tr>
<td>VEGF</td>
<td>Vascularendothelial Growth Factor</td>
</tr>
<tr>
<td>YY1</td>
<td>Ying Yang 1</td>
</tr>
</tbody>
</table>
CHAPTER ONE
INTRODUCTION

The human prostate gland is approximately the size of a walnut and is located between the bladder and the penis. It is part of the male reproductive tract and functions to produce seminal fluid - a white secreted fluid which transports and nourishes sperm. The prostate gland is composed of fibrous, glandular, and muscular tissue and is commonly classified into four zones – the peripheral zone, the central zone, the transition zone, and the anterior fibromuscular zone. Of these zones, the peripheral zone has been noted to be more susceptible to carcinomas than any of the other three zones and is the site where prostate cancer is initiated (Vogelzang, et al. 2005).

Androgens play an integral role in both the development of the prostate and prostate cancer. The prostate gland is dependent on androgens for prostatic epithelium growth and prostate gland maintenance. Androgens are also vital to the onset of prostate cancer for many prostate tumors arise from androgen dependent secretory epithelial cells (Perry and Tindall 1996). Androgens are used throughout many stages of prostate cancer mainly to inhibit apoptosis and promote proliferation (Joly-Pharaboz, et al. 1995, Wang, et al. 1996). However, over time, tumors can become androgen independent. This independence is said to be responsible for most of the prostate cancer related deaths (Thomas and Sawyer 2001).

Androgen action is mediated by the androgen receptor (AR). The AR is a member of the nuclear receptor family of transcription factors and functions as a DNA binding,
androgen activated transcription factor (Schug and Overton 1997). Once bound to androgen, the AR translocates the nucleus and activates target genes required for proliferation, differentiation, apoptosis and secretion (Nieto, et al. 2007). AR expression is observed in all types of prostate cancer, including primary prostate cancer as well as hormone-refractory prostate cancer. AR’s ability to stimulate cyclin-dependent kinases (CDKs) regulates cellular proliferation in prostate cancer cells; however, the AR also possesses the ability to stimulate the expression of cell cycle inhibitors, subsequently leading to the downregulation of cellular proliferation (Newman and Zetter 2008).

Prostate cancer is the most common extracutaneous cancer and the second leading cause of cancer related deaths in American men (American Cancer Society 2008). The American Cancer Society (ACS) estimated that in 2008, slightly over 186,000 new prostate cancer cases were diagnosed in the United States. A higher incidence of prostate cancer cases are found in African-American men who are 1.5 times more likely to develop prostate cancer than their Caucasian counterparts and are two to three time more likely to die from the disease than Caucasians (American Cancer Society 2008). ACS estimates that in 2009, 27,130 African-American men will be diagnosed with prostate cancer and almost 4,000 men will die from this disease.

Prostate cancer detection can be delayed due to the fact that palpable masses detectable by physical exam may take years to form. Therefore, many men have prostate cancer and are asymptomatic. However, once the cancer cells metastasize to the bone, tumor growth occurs more rapidly. While there are risk factors – i.e. age, race, family history and genetics, and diet – associated with the development of prostate cancer, currently, there is no clear mechanism detailing how prostate cancer develops.
The basic helix loop helix (bHLH) protein family plays an integral role in cellular processes such as cell proliferation and differentiation (Olson 1992), apoptosis (Evan and Littlewood 1993), angiogenesis (Wenger and Gassmann 1997), cell lineage determination (Borycki and Emerson 1997), and sex determination (Benezra, et al. 1990). A key member of this family, the inhibitors of differentiation/DNA binding (Ids), lacks the basic DNA binding domain and forms a nonfunctional heterodimer with other members of bHLH family that inhibits DNA binding and thus prevents transcriptional activity (Benezra, et al. 1990). Ids have shown the ability to stimulate proliferation and inhibit differentiation and are overexpressed in a wide variety of human cancers including prostate cancer. These characteristics have lead researchers to believe that Ids play a key role in cancer initiation (Perk, Iavarone and Benezra 2005).

Ids have four isoforms (Id1-4). While there are several studies investigating the roles of Ids, there is a lack of published research on the Id3 gene. Id3 was first identified as a serum inducible immediate early gene in murine fibroblast cells (Christy, et al. 1991). The human Id3 is 1288 base pairs (bps) transcript composed of three exons and two introns. The first exon is 705 bps; this exon is located from base one to 705 (Trabosh, et al. 2009). The second exon is located between 706 and 790 and is 86 bps long (Trabosh, et al. 2009). The final exon is 498 bps long and located between base pairs 791 and 1288 (Trabosh, et al. 2009). Despite information being available on Id3’s coding sequence, there is not much information know about the promoter.

Researchers have demonstrated that Ids are expressed in various types of tissues; however, their expression levels vary between cell lines. In normal human prostate epithelial cells, Id3’s expression is down regulated; however, in prostate cancer cells, its
expression is up regulated (Asirvatham, Schmidt and Chaudhary, 2006). Most notably, this upregulation leads to increased cell proliferation, de-differentiation, increased metastasis, and cell survival (Li, Gerald and Benezra 2004, Asirvatham, Schmidt and Chaudhary 2006, Asirvatham, Carey and Chaudhary 2007). Therefore, it is of interest to understand how Id3 expression is up-regulated in prostate cancer. This transcriptional regulatory mechanism will provide an insight into how Id3 mediates prostate cancer cell survival.

**Hypothesis**

Microarray, RT-PCR data suggests that the human Id3 gene is up regulated in cancer. This is possible due to a transcriptional mechanism that may be common in highly metastatic prostate cancer cells. Understanding of this mechanism may provide insight into the molecular pathways involved in prostate cancer cell survival and metastasis.

**Specific Aims**

- To identify the human Id3 gene promoter regulatory domains through a comparative genomics approach
- To identify the putative transcription factor binding sites in the proximal 1 kb promoter of the human Id3 gene and evaluate their expression using microarray data.
- To determine the minimal promoter region required for optimum transcriptional activity
- To construct a potential transcriptional pathway that regulates Id3 gene expression.
There are several levels at which genes can be regulated: transcriptionally at the mRNA level, epigenetically (methylation) and protein stability. Genes are most promoter (Kingston and Narlikar 1999). Once the transcription preinitiation complex reverently regulated at the transcriptional level via activation or inactivation by specific transcription factors. During this process, chromatin is unwound by remodeling complexes that allow the basal transcriptional machinery – RNA Polymerase II, TATA box, and TATA box-binding protein (TBP) – access to the forms and binds to RNA polymerase II, this complex then binds to the promoter and directs RNA polymerase II to begin gene transcription.

**Gene Promoters**

Gene promoters are the regulatory regions of DNA typically located between -1000 to +50 base pairs relative to the transcription start site. The promoter of a gene encompasses key sequence elements necessary in order for a gene to be transcribed: TATA Box and TBP, transcription initiation complex and transcription factors (Figure 1).

Enhancer sequences also bind to activators and initiate gene transcription. Enhancer sequences are generally located either further upstream or downstream from the promoter. Their ability to function from such a far distance likely denotes the need for
multiple signals to determine whether a gene gets transcribe (Clark 2005). Additionally, enhancers can act to inhibit transcription; these enhancers are called silencers.

![Diagram](image)

**Figure 1.** Eukaryotic gene promoter: TFIID binds to the TATA box via TBP near the transcription initiation site. Next, other initiating transcription factor sites attach to the complex which guides and stabilizes the binding of RNA Polymerase II. Once this happens, more transcription factors bind to complete the mature transcription complex. Then ATP is converted to ADP and Pi in order to give the complex energy for transcription to take place. Then, transcription can occur (Picture taken from Kimball 2006).

**TATA Box and TBP**

The TATA box is the most widely-used core promoter motif throughout nature (Juven-Gershon, et al. 2008). It has a consensus TATAAA sequence approximately −32 or −29 positions relative to the transcriptional start site (Ponjavic, et al. 2006). The TATA box is recognized and bound by TBP.

The TBP is a highly conserved transcription factor whose binding to the TATA box is the rate-limiting step in transcription (Bjornsdottir and Myers 2008). The TBP positions RNA polymerase II over the transcription start site and aids in the formation of the transcription initiation complex.

Many gene promoters, such as Id3, do not contain a TATA box. TATA less genes often encode housekeeping enzymes, oncogenes and transcription factors (Azizkhan, et al. 1993). In these promoters, TFIIB and RNA polymerase II are believed to mediate transcription as well as TFIID and TBP- when available (Anish, et al. 2009). However,
in the TATA-less Histone H1 promoter, Isogai et al described a TBP related factor (TRF) believed to regulate promoter activity promoter in Drosophila (Isogai, et al. 2007). While discoveries have been made surrounding the function of TATA less promoters, the mechanism behind how TATA less promoters mediate transcription is still unknown.

*Transcription Initiation Complex*

The transcription initiation complex is comprised of two types of factors: general transcription factors and coactivators/corepressors. The general transcription factors include Polymerase II, TFIIB, TFIID, TFIIE, TFIIF, and TFIIFH. When all of these factors are attached and bound to DNA, they initiate transcription.

Coactivators and corepressors mediate the response to regulatory signals. Coactivators, such as PPARγ coactivator-1 (Li and JD 2009) and phosphoproteins (Weigel and Moore 2007), increase gene expression by binding to a protein activator which stabilizes the RNA polymerase holoenzyme and provides a faster route to the promoter; Corepressors, such as dCtBP, Groucho, and dSin3A (Cai and Laughon 2009) decrease gene expression via recruiting histone deacetylases after transcription which increases the positive charge on histones thus strengthening the interaction between histones and DNA, which making the DNA less accessible to transcription (Kornberg 2007).

*Transcription Factors*

Transcription factors are proteins that when bound to sequence specific DNA binding sites activate or repress gene transcription. This activity can be modulated by either altering the binding of RNA polymerase to DNA, or by DNA, protein, or histone modification (Locker 2000).
Transcription Factor Function

Transcription factors are vital to many biological functions. They regulate the initiation of transcription, and are important in many cellular processes such as: transcription regulation, cell fate determination and cellular differentiation, signal response, and cell cycle control. Transcription factors' regulation and expression are fostered through biological activities such as: transcription, translation, post transcriptional and translational modifications, splicing, RNA transportation to the cytoplasm, and mRNA degradation.

Transcription Factor Structure

Transcription factors are usually modular and comprised of two main domains: the DNA binding domain and the trans-activating domain. The DNA binding domain is the region of the transcription factor that binds to the DNA. Furthermore, the specific structure of the DNA binding domain determines the classification of the transcription factor. Trans-activating domains contain protein binding sites that allow for activation or suppression of gene transcription in conjunction with other transcription factors and corepressors. Trans-activating domains are classified based on the profuse presence of amino acids i.e. glutamine rich domains of SP1 and proline rich domain of NF1 (Krauss 2003).

DNA Binding Domains

DNA binding domains mediate the binding of transcription factors to DNA - which is required for activating transcriptional activity. This binding is sequence specific and
occurs at the major and occasionally minor groove of DNA or at the sugar-phosphate backbone of the DNA. There are several structural classes of DNA binding domains and transcription factors are grouped within these classes based upon characteristics such as consensus amino acid sequences, structural motifs, and functions. There are four super classes of DNA binding domains: Basic Domains, Zinc Coordinating DNA binding domains, Helix-Turn-Helix, and β Scaffold Factors with Minor Grooves (Sanderson and Walker 2000). While there are several sub classes of transcription factor binding domains within these super classes, the ones relevant to this research are:

**Zinc Finger Proteins**

Zinc-Finger Proteins are the most common transcription factor motif class found in eukaryotic organisms (Kang 2007). The classical zinc finger structure is comprised of five C2H2 zinc finger domains (thus called the C2H2 zinc finger). The C2H2 zinc finger domain is relatively small comprising of approximately 26 - 30 amino acid residues. In this domain, one zinc atom is bound by the side chains of two conserved cysteine and histidine residues, thus forming a finger like structure (Lee, et al. 1989, Klug and Rhodes 1987). These domains fold into a compact structure comprising of an irregular loop of two antiparallel β sheets and an α helix inserted into the major groove of DNA (Rhodes and Klug 1993). Some prominent transcription factors that have zinc finger domains are TFIIA, Sp1, Wt1, and YY1 (Schug and Overton 1997).

**Leucine Zipper Proteins**

Leucine Zipper proteins are characterized by four to five sequential hydrophobic leucine residues spaced every seventh position in its amino acid sequence. These proteins mediate DNA binding through the dimerization of two leucine rich amphipatic
b-helices (Passarge 2006). Once the helices dimerize, a coiled-coil is formed due to the hydrophobic interactions of each helix. When the helices disjoin, a Y-shaped structure which clutches the DNA molecule at two adjacent major grooves is formed (Strachan and Read 2004). Some prominent transcription factors that have leucine zipper domains are ATF, Fos, Jun, and members of the CEBP family (Schug and Overton 1997).

*Helix Loop Helix Protein*

The helix loop helix DNA-binding domain contains an N-terminal α helix, a middle loop region, and a C-terminal α helix. The two α helices vary in size. The larger N-terminal α helix contains basic residues that interact with DNA while the smaller C-terminal α helix region contains hydrophobic amino acids that may be required for binding specificity (Lodish, et al. 2003). Also, the N-terminal helix typically contains the DNA-binding regions while the C-terminal helix forms dimers by folding and packing against another helix. The loop is integral to helix stabilization (Kiewitz and Cabrele 2005). Helix loop helix proteins typically bind to a consensus sequence CACGTG named the E-box (Murre, et al. 1989); however, some helix loop helix proteins bind to a related sequence CACNAG named the N box (Massari and Murre 2000).

**Basic helix loop helix proteins (bHLH)**

The basic helix loop helix (bHLH) proteins play an innate role in regulating cell activities such as neurogenesis, myogenesis, cell proliferation and differentiation, cell lineage determination, sex determination, as well as other essential processes in organisms (Benezra, et al. 1990, Kadesch 1993). Their domains are highly conserved from yeast to mammals (Quong, et al. 1993). The basic domain facilitates DNA binding while the HLH domain mediates hetero/homodimerization, at specific hexanucleotide sequences,
between two bHLH proteins. The bHLH dimers usually bind to the E box or N box, both are transcription element binding sequences that lie upstream of transcribed DNA, and subsequently allow bHLH proteins to act as transcription factors. There are two main classes of bHLH proteins: the ubiquitously expressed proteins and the tissue specific proteins. Ubiquitously expressed E proteins are comprised of HEB, E2-2 and E2A splice variants E12 and E47 (Bain, et al. 2001, Edmondson and Olson 1993, Deed and Jasoik 1994, Murre, et al. 1989). E proteins are characterized by their non-tissue-specific expression patterns and their ability to form dimers with both tissue specific proteins and the inhibitors of differentiation proteins (Pan, et al. 1999).

![Diagram](image)

Figure 2. Basic Helix Loop Helix binding interaction between classes. (Top) The Ubiquitous E proteins (Class I) hetrodimerize with Tissues Specific proteins (Class II) and their basic regions bind to the E box binding site of DNA which activates transcription of cell specific gene and cellular differentiation. (Bottom) The Ubiquitous E proteins (Class I) dimerize with Id proteins that lack the basic binding domain thus blocking transcription of cell specific gene and inhibiting cellular differentiation (Picture taken from Meenakshil 2003).

The tissue specific bHLH proteins have unique dimerization properties. Despite appearing to homodimerize poorly (Chakraborty, et al. 1991), the tissue specific proteins preferentially heterodimerize with E proteins. These dimers have a high affinity to
binding to the E box (Benezra, et al. 1990, Cooper and Newburger 1998, Choi and Costa 1990) and regulating transcriptional activities in muscle, nerve, and brain tissues (Barone, et al. 1994, Jen, Manova and Benezra 1996, Kreider, et al. 1992). Some of the more prominent members of this protein class include MyoD, NeuroD, achaete-scute, and Hen (Tapscott and Weintraub 1991). Members of the MyoD family are specific to skeletal muscle and regulate muscle differentiation genes (Sirri, Leibovitch and Leibovitch 2003); NeuroD promotes neural cell differentiation (Ochocinska and Hitchcock 2009); achaete-scute and HEN/NSCL proteins are found in the nervous system and are thought to promote neurogenesis (Simionato, et al. 2008, reviewed by Tapscott and Weintraub 1991).

**Inhibitors of Differentiation Genes (Ids).** The inhibitors of differentiation / DNA binding family (Ids) are comprised of isoforms 1 to 4 (Id1-4). These genes are localized on human chromosomes 20q11, 2p25, 1p 36.13 – 36.12, and 6p 22-21, respectively (Mathew, et al. 1995). Ids 2-4 have three exons while Id1 has two (Nehlin, et al. 1997). Ids have a molecular weight ranging from 13 to 20 kDa (Coppé, Smith and Desprez 2003) and are thought to enter and exit through the nucleus via passive diffusion because of their size (Makitaa, et al. 2006).

Ids lack the basic region and act as dominant negative regulators of bHLH proteins through forming a nonfunctional heterodimer with E proteins and thereby inhibiting cell differentiation (Figure 2) (Benezra, et al. 1990, Mure 1994). Ids have a highly conserved HLH domain but divergent N- and C-terminal domains. The sequences outside their terminal domains are thought to help determine binding and tissue specificity. While the molecular function of the C terminal region has not been
established, the N terminal region has been linked to apoptotic induction and protein degradation (Florio, et al. 1998).


*Ids in Cancer*

**Differentiation**


Moreover, all four Id isoforms have some type of effect on differentiation. In HC11, mouse mammary epithelial cells, overexpression of Id4 not only increased cell
proliferation but inhibited lactogenic hormone-mediated differentiation as well; this was revealed by inhibition of beta-casein promoter activity and beta-casein expression (Shan, et al. 2003). In vivo, when lactating murine mammary cells were induced to differentiate, Id1 expression levels declined to undetectable levels; whereas, when Id1 expression was up regulated, the cells ability to differentiate became lost and the cells became invasive (Fong, et al. 2003).

Strikingly opposite effects have been shown in Id2. The overexpression of Id2 increased erythroid development while reduced Id2 levels hindered normal erythroid development - which is mediated via the transcription factor PU. (Ji, et al. 2008); Additionally, the same inhibitory effects were found in Id3 in thymocytes; however this regulation was mediated by TCR (Bain, et al. 2001).

Proliferation

Researchers have concluded that Ids have elevated expression in proliferating cells where as they have negative expression in non-proliferating cells. Even though the rationale as for why Id expression promotes proliferation in cancer cells is unknown, these expression tendencies have been widely noted. In prostate cancer, the loss of Id1 and Id3 expression by siRNA results in a loss of cellular proliferation (Asirvatham, 2006). Id4’s overexpression in rat mammary gland tumors has also shown an increase in cellular proliferation (Shan, et al. 2003). Additionally, Id2 not only enhances proliferation and survival of growth-arrested pancreatic beta cells but it induces expression of Bcl-2 and reduces the expression of p38 MAPK and NFkappaB (Hua and Sarvetnick 2007). Conversely, Id2 knockouts can re-establish the proliferative potential of tumor cells inhibited by removal of mutant p53 in pancreatic cancer (Yan, et al. 2008).
Wilson et al noted that Id3 overexpression in colorectal cancer directly corresponds with the loss of p53 expression. This overexpression resulted in uncontrolled cellular proliferation (Wilson, et al. 2001).

Metastasis

Cancer metastasis is a multi-step process that coincides with cellular proliferation. Ids are not only essential to proliferation but have integral roles in cell migration and invasion as well. In colorectal cancer, adenocarcinomas with lymphoid metastasis demonstrated greater Id1 expression than those without (Zhao, et al. 2008). An in vivo characterization of lung metastatic progression revealed that Id1 and Id3 facilitate sustained proliferation during the early stages of metastatic colonization (Gupta, et al. 2007). Shuno et al demonstrated, in pancreatic cancer, knockdown of both Id1 and Id3 expression decreased both cellular proliferation and migration (Shuno, et al. 2008); Tsuchiya et al exhibited these same results in gastric cancer cells (Tsuchiya, Okaji and Tsuno 2005). While in hepatocellular carcinomas, decreased Id2 expression was shown to induce metastasis in vitro by altering VEGF secretion (Tsunedomi, et al. 2008).

Apoptosis

Ids have been shown to promote apoptosis in many cell lines. In transgenic mice, not only does Id1 overexpression cause massive apoptosis in T cells, it blocks T cell development (Yang, Wang and Sun 2008). In immortalized keratinocytes, Id3’s overexpression was shown to induce a significant increase in UVB-induced apoptotic cells via a mitochondrial-caspase-9-mediated pathway (Simbulan-Rosenthal, et al. 2006). Also in B cells, the induction of Id3 expression in progenitor B cells was shown to induce apoptosis through a caspase-2-dependent mechanism (Lazorchak, et al. 2006). Others
have observed a downregulation of Id1 and caspase-3 expression, in prostate cancer cells, when treated with TNFα and in Poly (ADP-ribose) polymerase (PARP) (Wong, Wang and Ling 2004). This downregulation caused an increase in apoptosis. Additionally, in astrocytic cultures, Id4 expression induces apoptosis and is down-regulated by activation of the cAMP-dependent signal transduction pathway (Andres-Barquin, et al. 1999). Furthermore, Ids' ability to invoke cell death in various cell lines makes them novel targets for the treatment of various cancers.

Angiogenesis

Angiogenesis plays an important role in tumor progression and is critical for tumor survival, growth, invasion, and metastasis (Gabellini, Del Bufalo and Zupi 2006). Gao et al showed that tumors generated from bone marrow endothelial progenitor cells induced Id1 expression. This expression inhibited angiogenesis and prolonged the life of animals with tumors (Gao, et al. 2008). Ciarrocchi et al demonstrated that the mechanism behind this phenomena is due to Id1 reducing the expression of p21 (Ciarrocchi, et al. 2007).

In breast cancer cells, Jang et al determined that Id1 plays an integral role in tumor angiogenesis after correlating the degree of microvessel densities of 263 human breast cancers, 15 in situ lesions and 248 invasive cancers to Id1 expression levels. They noted that tumors with high Id1 expression levels also had high microvessel densities (Jang, et al. 2006). Furthermore, Henke et al constructed an antitumor agent, comprised of Hsp90 inhibitor 17-(allylamino)-17-demethoxygeldanamycin, that down regulated Id1 expression in tumor endothelium cells. In two different tumor models, they noted that the downregulation of Id1 inhibited metastasis and tumor growth (Henke, et al. 2008).
Raida noted that BMP-2 activation of the BMP pathway promotes vascularization and possibly angiogenesis via inducing the Id1 and p38 pathway (Raida, et al. 2005). Knockdown of Id1 and Id3 in MKN45 gastric cancer cells showed a decrease in these cells’ ability to proliferate and metastasize as well as their ability to bind to laminin (Tsuchiya, Okaji and Tsuno 2005). Furthermore, in breast cancer cells, Id1 and Id3 were highly expressed in the endothelial cells of tumor-infiltrating blood vessels; additionally, knockdown of Id1 and Id3 expression demonstrated angiogenesis inhibition (de Candia, Benera and Solit 2004). Vandeputte et al demonstrated that Id1-3’s expression levels are up regulated in neurological tumors and in their surrounding blood vessels. This upregulation lead to an increased proliferation in astrocytes tumor cells (Vandeputte, et al. 2002).

Vascular endothelial growth factor (VEGF) is associated with many types of tumors and plays an important role in angiogenesis (Toi, Matsumoto and Bando 2001). Studies have shown that the overexpression of Id1 in prostate cancer cells promotes angiogenesis via VEGF activation through autocrine signaling (Ling, Lau, et al. 2005). In gastric cancer, COX-2 was shown to stimulate VEGF and enhance endothelial cells proliferation by upregulating Id1’s expression levels (Lei, et al. 2007). Also, VEGF plays a central role in hepatocellular carcinoma growth and metastasis. Lee et. al showed that an increase VEGF secretion in Id1 transfectants induces morphologic change and proliferation of human umbilical vascular endothelial cells which promotes angiogenesis (Lee, et al. 2006).

Deregulated Id2 activity inhibits the Rb tumor suppressor pathway and promotes VEGF expression. This expression, in pituitary tumor cells, was shown to promote
cellular growth and angiogenesis (Lasorella, et al. 2005). Similarly, in neuroblastomas, this expression is mediated by N-Myc and also promotes angiogenesis (Lasorella, et al. 2005). Ids have an integral role in the angiogenic process in various cell lines. Understanding the mechanisms behind Ids ability to regulate the angiogenic process in these cell lines may provide insight into cancer cell survival.

Regulation of Ids’ Expression

Ids’ expression levels have been observed in several cancer cell lines. Various transcription factors are known to play a pivotal role in regulating Ids expression. Northern blot and Id1 immunohistochemistry experiments conducted by Kebebew et al showed Id1 was regulated by serum and thyroid-stimulating hormone (TSH) stimulation in human papillary thyroid cancer; Id1’s expression increased 1.5- and 4.0-fold, respectively (Kebebew, et al. 2003). Wong et al demonstrated that TNF-β regulates Id1 protein stability in prostate cancer cells. They observed a significant decrease in Id1 when DU145 and PC3 cells were exposed to TNF-β (Wong, Wang and Ling 2004).

Several studies have shown that Id1 expression is down-regulated by TGF-β in epithelial cells (Cao, et al. 2009, Liang, Brunicardi and Lin 2009), whereas it is up-regulated by BMP in a variety of cell types such as osteoblasts (Ogata, et al. 1993), embryonic stem cells (Hollnagel, et al. 1999) and neuroepithelial cells (Nakashima, et al. 2001). Liang et al found Id1 to be strongly up-regulated by TGF-β1 in the human mammary gland epithelial cell line MCF10A. Furthermore, they used Smad knockdown and knockout approaches to determine that Smad3 was responsible for mediating transcriptional activation of the Id1 gene (Liang, Brunicardi and Lin 2009).
In rat intestinal epithelial cells, TGF-β was shown to repress the expression of the Id gene family; Cao et al showed that apoptosis in RIE-1 cells was induced by the knockdown of Id1 and Id2 gene expression, whereas TGF-β-induced apoptosis resulted from the overexpression of Id2 (Cao, et al. 2009).

Id3’s overexpression also has been noted in several cancer cell lines. Deleu et al demonstrated that TSH stimulation mediated the upregulation of Id3 expression in dog thyrocytes (Deleu, et al. 2002). This result came by way of investigating Id3 as a potential prognostic marker of thyroid carcinomas. Also, insulin was shown induce theses same effects in dog thyrocytes (Deleu, et al. 2002) and preadipocytes (Inuzuka, et al. 1999). Additionally, BMPs were shown to increase Id3 expression in embryonic (Ying, et al. 2003), epithelial (Kowanetz, et al. 2004), and mesenchymal stem cells (Peng, et al. 2004) thus negatively affecting cell differentiation. However the length of Id3’s overexpression varied among cell lines.

Conversely, Id3 exhibits differing expression level responses in vascular smooth muscle cells. Angiotensin II, superoxide free radical, induces a rapid increase in Id3 expression which activates the p38, ERK, and calcium dependent pathways (Muller, et al. 2002). Concurringly, estradiol was shown to stimulate Id3’s expression levels in vascular smooth muscle cells also (Watanabe, et al., 2004). However, Gut-enriched Krüppel-like factor (GKFL)’s binding to the Id3 promoter negatively regulates Id3 expression in response to free radicals in vascular smooth muscle cells (Nickenig, et al. 2002). Additionally, in neuroblastoma cells, retinoic acid was shown to downregulate Id3 expression after 24 hours (López-Carballo, et al. 2002). Furthermore, understanding the
regulatory mechanisms of how Ids function will provide insight into how Ids mediate cancer cell survival.

*Ids in Prostate Cancer*

Id gene expression is prevalent in prostate cancer. This expression is deregulated and correlates to prostate cancer aggressiveness and Gleason score grading (Li, Gerald and Benezra 2004, Yuen, et al. 2006, Coppe, et al. 2004). An experiment conducted by Salomon et al demonstrated that just the overexpression of Ids is not sufficient enough for prostate cancer activity (Salomon, et al. 2009). Therefore, finding the underlying regulatory mechanisms involved in the regulation of Ids may lead to future therapeutic treatments and the possible alleviation of this disease.

Id4 is believed to be a tumor suppressor (Asirvatham, Schmidt and Chaudhary, 2006). Yuen et al evaluated 125 prostate tumor samples and found that Id4’s expression was down regulated in prostate cancers compared to nodular hyperplasia (Yuen, et al. 2006). This down regulated expression leads to a decrease in cellular proliferation and an increase in apoptosis (Yuen, et al. 2006, Carey, et al. 2009). However, the mechanism behind Id4’s regulation is still under investigation.

The upregulation of Id2 in LNCaP rendered cells highly aggressive with increased growth and invasion capabilities; contrarily, when Id2 was down regulated in PC3, cells were shown to be less aggressive and have minor growth capabilities (Coppe, et al. 2004). Asirvatham et al noted that decreased Id2 had no effect of proliferation; however, they did visualize an increase in apoptosis (Asirvatham, Schmidt and Chaudhary, 2006).

Id1 expression exhibited a similar expression patterns as Id2. Asirvatham et al demonstrated that Id1 and Id3 knockdown resulted in a loss of proliferation (Asirvatham, Schmidt and Chaudhary, 2006). Di et al showed that the loss of Id1 expression in NPTX
cells and cellular proliferation induced TGF-β growth arrest, senescence, cell cycle arrest – G2/M (Di, et al. 2006). Not only were Coppe et al results demonstrated in Id2, similar expression capabilities have been validated in Id1 as well. Studies have shown that Id1’s increased expression levels during human prostate cancer progression lead to an increase in cell proliferation and invasiveness (Yuen, et al. 2006, Coppe, et al. 2004). It was believed that matrix metalloproteinases (MMPs) regulated these expression patterns in Id1 (Ouyang, Wang and Lee, et al. 2001, Fong, et al. 2003) through controlling secretion and activation of MMP expression (Coppe, et al. 2004). However, Darby et al discovered these regulatory activities were actually caused by Bone morphogenetic protein-6’s (BMP-6) activation of both MMP and Id1 (Darby, et al. 2008).

The overexpression of Id1 has been shown to activate several molecular pathways involved in prostate cancer. The p16(INK4a)/pRB pathway was shown to be inactivated by cellular proliferation caused by the overexpression of Id1 (Ouyang, Wang and Ling, et al. 2002). Ling et al discovered that Id1 expression is essential to the activation of the Raf/MEK1/2 pathway (Ling, Wang, et al., 2002) and that the activation of the NFκappaB pathway was shown to promote cell proliferation and hinder apoptosis via Id1 expression (Ling, Wang, et al., 2003). Investigating the mechanisms that regulate Ids’ expression in prostate cancer will lead to the identification of potential prognostic markers, the understanding of cancer cell survival, and potential strategies to combat the prostate cancer.
CHAPTER THREE
MATERIALS AND METHODS

Background (Databases)

The Human Id3 Promoter Sequence and Transcription Factor Binding Sites

The human Id3 promoter was located using the University of California at Santa Cruz Database (UCSC) genome browser (http://genome.ucsc.edu). From the UCSC homepage, the Human Genome Browser tab was selected and the human Id3 gene was searched. Next, the DNA tab was selected and the parameters for promoter position were chosen (1500 extra bases upstream and 50 extra bases downstream). Due to the orientation of the Id3 gene, the reverse complement sequence option was selected.

Transcription Factor Binding Sites in Id3 Promoter

The potential transcription factor binding sites for the Id3 gene were identified using The Transcription Search Element Database (TESS) (http://www.cbil.upenn.edu/cgi-bin/tess/tess).

At the TESS database website, the promoter sequence that was obtained from the UCSC database was entered into the sequence field and submitted. The query returned a datasheet that was formatted by its log-likelihood score and log-likelihood score derivates (which will be discussed in the results section), that generated a list of potential transcription factor binding sites.
Laboratory Experimental Design

Polymerase Chain Reaction

Primer Design

Polymerase Chain Reaction (PCR) primers were designed with the Invitrogen Perfect Design website (https://tools.invitrogen.com/content.cfm?pageid=9716) in order to amplify Id3's promoter region by PCR. The 1.5 kb promoter sequence from the UCSC website was entered in FASTA format into the design field. The default parameters for the minimum, optimal, and maximum primer size, temperature, GC percentage, and product size were selected. The resulting primer matches were compared with the original sequences to ensure primer uniqueness and specificity via NCBI's BLAST program. Subsequent primers were synthesized commercially.

Genomic PCR

A preliminary PCR was performed via testing different primer pairings from the six primers (three forward primers and three reverse primers) that were obtained from the Invitrogen primer design database.

<table>
<thead>
<tr>
<th>Forward</th>
<th>Reverse</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) 5' ACTGAAGAGCCTGTCCTTGA 3'</td>
<td>5' GCTCAAAAGATCTGGGTT 3'</td>
</tr>
<tr>
<td>(2) 5' TGTGAGATTCCTGGAAGCAGGA 3'</td>
<td>5' GGAAAAAGCAAATTCTGGGAAG 3'</td>
</tr>
<tr>
<td>(3) 5' AATCAAGTCCAGGTGGCCT 3'</td>
<td>5' TATCAGCGCTTCCTCATTCT 3'</td>
</tr>
</tbody>
</table>

The PCR was performed on genomic DNA isolated from DU145 prostate cancer cell line using the Invitrogen PCR kit. The master mix for the PCR contained: 10x PCR Buffer (2.5 μl), 5x Master Taq (2.5 μl), 10mm dNTP (0.5 μl), dH2O (16.4 μl), Taq
Polymerase (0.1 µl), DU145 genomic DNA (5 µl), and PCR Primers (0.5 µM each) for a final total of 25 µl. PCR conditions were set to run for 60 seconds at 94°C on the denaturing cycle, 60 seconds at 58°C on the annealing cycle, and ten minutes at 72°C extended annealing cycle, followed by 4°C infinite chilling. Afterwards, the PCR sample was analyzed on an agarose gel using an ultraviolet transilluminator; and, extracted from the gel using the Eppendorf Gel Clean-up kit.

Promoter Confirmation

Nested PCR

Nested PCR was performed on the purified PCR product yielded above to confirm the presence of the amplified promoter region. This PCR was performed using the same procedure stated above. However, DNA was produced from the above PCR product while the primers consisted of mixed pairs (F2, R2; F3, R1; F2, R1; F3, R2) that were found on the outside and within the promoter sequence.

Restriction Digest

The proximal 1kb promoter nucleotide sequence was inserted into the NEB cutter database (http://tools.neb.com/NEBcutter2/index.php) in order to identify the restriction enzyme sites. An EcoRI site was found at the 600 bp within the Id3 promoter sequence. In order to confirm the promoter, a restriction enzyme digest was performed on the PCR product using the following procedure. 10 µg of plasmid DNA, 2 µl of deionized water, 2 µl of the restriction enzyme EcoRI, 1 µl of restriction enzyme Multi-Core Buffer. The reaction was incubated at 37°C in water bath for two hours and then placed in a 65°C water bath for ten minutes to inactivate the enzyme.
**Id3 Promoter Cloning in Reporter Plasmids**

**pGEMT-Easy Cloning**

After the promoter sequence was confirmed, via the Nested PCR, the promoter was ligated into the pGEMT-Easy shuttle vector to generate a greater volume of the promoter and to perform further experimentation. This ligation was performed using the Promega Ligation Kit. Reactions were set up in the following manner: one ligation reaction per sample to be cloned, one negative control, and one positive control.

<table>
<thead>
<tr>
<th></th>
<th>Sample</th>
<th>Positive</th>
<th>Background</th>
</tr>
</thead>
<tbody>
<tr>
<td>2X rapid ligation buffer</td>
<td>5 µl</td>
<td>5 µl</td>
<td>5 µl</td>
</tr>
<tr>
<td>pGem-T Easy Cloning Vector</td>
<td>1 µl</td>
<td>1 µl</td>
<td>1 µl</td>
</tr>
<tr>
<td>PCR Product</td>
<td>3 µl</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Control insert DNA</td>
<td>-</td>
<td>2 µl</td>
<td>-</td>
</tr>
<tr>
<td>T4 ligase</td>
<td>1 µl</td>
<td>1 µl</td>
<td>1 µl</td>
</tr>
<tr>
<td>Deionized Water</td>
<td>-</td>
<td>1 µl</td>
<td>3 µl</td>
</tr>
</tbody>
</table>

Reactions were mixed in a 0.5 ml tube and placed in a refrigerator (at 4°C) overnight.

**Transformation**

Each ligation mixture was equilibrated to room temperature for one minute and added to the bottom of a sterile 5 ml round-bottomed tube that was pre-cooled on ice. 50 µl of pUC19 competent cells were added to each round-bottomed tube and left on ice. After 20 minutes, the tubes were placed in a 42°C waterbath for 45 seconds in order to heat shock the cells and returned to ice for two minutes.

Next, 950 µl of Super Optimal Broth were added to each transformation mixture and placed in an incubator shaker (37°C) for 90 minutes. Midway through this process, Luria
broth (LB)/Ampicillin (Amp)/ Isopropyl β-D-1-thiogalactopyranoside (IPTG)/ 5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside (X-Gal) (LAIX) plates were dried out as they were incubated at 37°C in an inverted position. Upon completion of incubation, 150 ul of the transformation cultures were plated onto the LAIX plates to induce bacteria growth. These plates were sealed with parafilm, and placed in 37°C incubator overnight.

**Blue White Colony Screening**

pUC 19 plasmid DNA contains the LacZ gene which codes for the b-galactosidase enzyme. IPTG is used to activate the transcription of Id3 promoter insert in the pGEMT-T Easy Plasmid and the b-galactosidase enzyme. b-galactosidase metabolizes X-gal therefore, if the gene is unsuccessfully transcribed, the bacteria colony displays the color blue.

After the overnight incubation period, several white colonies were picked using sterile pipette tips, placed into tubes containing 5 mL of LB + Amp liquid media, and incubated on a 37°C shaker overnight. This culture was used to extract plasmid DNA. Once the mini prep was performed, a maxi prep was conducted using the remaining 3 mL of the colony/LB+ Amp mixture. Both procedures were performed using the Mo Bio Labs Maxi prep kit.

**Sequencing**

In order to obtain the nucleotide composition, the PGEMT-Easy + Id3 promoter plasmid was sequenced at Morehouse School of Medicine’s Core Molecular Biology Laboratory using the T7 primer.
Restriction Enzyme Digests and Dephosphorylation

Restriction enzyme digest

Two restriction enzyme digests were conducted. The first digest was performed on the pGEMT-Easy plasmid that contained the Id3 promoter. This digestion was used to validate the Id3’s promoter orientation inside the pGEMT-Easy plasmid. The second digest was performed in order to remove the Id3 promoter from the pGEMT-Easy plasmid and to dephosphorylate the promoter’s phosphate ends. Ultimately, this process would allow for the Id3 promoter to be ligated into the pGLuc luciferase vector so that promoter activity could be determined via luciferase assay.

In the first digest, 5 μg of plasmid DNA (pGEMT-Easy + Id3), 2 μl of deionized water, 1 μl of EcoRI, 1 μl of HindIII, 1 μl of restriction enzyme Buffer E were pipetted into a 5 mL tube and placed in a 37°C water bath. After two hours, the enzymes were inactivated by placing the reaction in a 65°C water bath for 10 minutes. This product was inserted into an agarose gel and analyzed using an ultraviolet transilluminator after ethidium bromide staining.

The second digest combined 20 μg of DNA, 10 μl of EcoRI, and 10 μl of buffer H into a 0.5 mL microcentrifuge tube. The tube was placed into a 37°C water bath for 3 hours and then placed in a 65°C water bath for 10 minutes. The digest was split into two tubes. One tube contained 70 μl while the other tube contained 30 μl. The 10 μl from the 30 μl tube was run on a gel while the other 20 μl were placed in a freezer further analysis. The tube that contained 70 μl of the enzyme product was used for the ligation reaction described below.
Dephosphorylation reaction

*The Calf Intestine Alkaline Phosphatase (CIAP)* reaction was performed in order to ligate the promoter into the pGLuc luciferase plasmid. First, 1ul of CIAP buffer was diluted in 9 μl of 10X CIAP buffer. Then, 3ul of that diluted CIAP mixture, along with 10ul of the 10X CIAP buffer and 17 μl of deionized water were added to the restriction enzyme product tube. The tube was placed in a 37°C water bath for 30 minutes. After the tubes were removed, 3ul of additional diluted CAIP buffer was added to the mixture and placed back into the 37 °C water bath for an additional 30 minute and finally placed into 65 °C waterbath for fifteen minutes.

Afterwards, a ligation and a transformation were performed using the standard protocols (instead of the pGem-T Easy Vector, the pGLuc was used) and kits described above. The newly formed pGLuc + Id3 plasmids were analyzed on an agarose gel to verify its size and also digested with EcoRI.

*Sequencing*

In order to obtain the nucleotide composition, the pGLuc + Id3 promoter plasmid was sequenced at Morehouse School of Medicine’s Core Molecular Biology Laboratory using the pGLuc — Basic Forward Sequencing primer and the pGLuc Basic Reverse primer.

*Cell Culture*

Human prostate cancer cell lines DU145, PC3, and LNCaP and cervical cancer cell line HeLa were obtained from American Type Culture Collection (ATCC, Rockville, MD). The DU145 and PC3 cells were cultured in F12-BCS-A (Ham’s F12 (Gibco, Carlsbad, CA) medium containing 10% Bovine Calf Serum (HyClone) with appropriate antibiotics (pen/strep, fungizone, and gentamycin); LNCaP cells were cultured in RPMI-
1640 media containing 10% FBS (Hyclone, Logan, UT) with appropriate antibiotics (pen/strep, fungizone, and gentamycin); HeLa cells were cultured in F-12/DMEM media containing 10% FBS (Hyclone, Logan, UT) with appropriate antibiotics (Penicillin/streptomycin (pen/strep), fungizone, and gentamycin). All cells were cultured at 37 °C in a fully humidified atmosphere containing 5% CO2.

Transfection

In order to determine promoter activity, the pGLuc+Id3 promoter luciferase plasmid and the PCMV control plasmids were transfected in sub-confluent (60%) DU145, LNCaP, PC3, and HeLa cells. All cells were grown in twenty-four well plates using TransIT-prostate transfection reagent (Mirus Bio). Each component of the DNA transfection mixture (2.5 μl TransIT-prostate reagent, 0.5 μg of the pGLuc+Id3 plasmid DNA, 2.5 μl prostate Boost reagent) and 50ul of serum free media were added to wells. The media was changed once after an overnight incubation period. Forty-eight hours after the transfection, the cells were harvested for the luciferase assay.

Luciferase Assay

Luciferase Assays are used to measure the regulatory potential of a DNA promoter sequence via bioluminescence. Light is emitted by luciferin oxidation. The intensity of the light proportionally correlates to the amount of luciferase molecules present in the sample which in turns depends upon the light activity of the promoter. In this experiment, the pGLuc+Id3 plasmid is expected to drive luciferease expression which in turn will catalyze the luciferin oxidation. This experiment was conducted after the plasmids were transfected in prostate cancer cells.
Gaussia Luciferase Assay Kit (New England Bio Labs)

A 1X Gaussia Luciferase assay working solution was prepared by adding 50 µl of the GLuc Substrate to 5 ml of the GLuc Assay Buffer. Then, 20 µl of the culture supernatant were pipetted into a microtiter plate well. Afterwards, 50 µl of the 1X GLuc assay working solution was added to the wells of the cells transfected with pGLuc+Id3 plasmid and PCMV control. Luminescence was promptly measured.

Luciferase Assay System (Promega)

First, 20 µl of cell lysate were added to respective wells in the cell culture plate. Then, 100 µl of Luciferase Assay Reagent were added to each well containing the cell lysate and luminescence was measured using two second measurement delays followed by a ten second measurement reading of luciferase activity.

Bioinformatics

Transcription Factor Expression Data

The raw microarray data, on transcription factor expression in Id3 expression in prostate cancer from Chaudhary’s lab, were used to analyze and compare transcription factor expression levels within DU145, LNCaP, and PC3 cells. The data returned a Microsoft Excel spreadsheet which contained Affymetrix gene identifiers (AGI), transcription factor expression data (raw intensity), the absent or present call, and a description of the transcription factor being expressed.

In order to decipher this information, first, the transcription factor binding sites that were obtained from the TESS database were searched in the description fields of the microarray data. If found, the transcription factor was placed into The National Center for Biotechnology Information (NCBI) Gene database.
and its AGI was searched via the Ensembl Database. Once the AGI was validated at Ensembl, the transcription binding site location was determined using the TESS database and its name and microarray score (intensity data) were placed into a spreadsheet. Calculations were performed in order to determine the transcription factors that were being expressed and their expression levels in different prostate cancer cell lines—DU145, LNCaP, and PC3. Data was normalized by using the equation (intensity/Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) intensity/the number of GAPDH entries))*1000. (GAPDH is constitutively expressed gene and 1000 is the preferred normalized data reference). This normalization approach allowed us to compare microarray data from different cell lines.

Vista Sequence Alignment

Using software tools provided by Vista Database, a comparative sequence analysis was used to determine the conserved sites within the human, rat, and mouse 1d3 promoter. The rat and mouse were chosen because they are most widely studied as a representation of human disease model and have validated genomes that share the same number of genes as humans. Since conserved regions within genes denote similar organizational patterns among species, it is believed that if these sequences share similar regulatory regions in their promoter then their promoters are likely to be regulated by the same mechanism.

In this analysis, each proximal 1kb promoter sequence was identified and saved into FASTA format in separate text files. From the submission page, each sequence file was submitted. Upon completion, mVista returned options to view the phylogenetic tree.
of the three sequences, the conserved regions between all three sequences, and the
aligned and conserved transcription factor binding sites among all three sequences.
CHAPTER FOUR

RESULTS

Promoter Identification

Id3 Proximal Promoter Sequence (with Primer Information)

After the correct orientation of the promoter was determined, the proximal 1.5 kb promoter sequence was retrieved via the UCSC database. PCR primers were designed within this sequence to amplify the full length of the Id3 promoter in order to determine the minimal promoter region required for optimum transcriptional activity. Three primer sets were designed, Figure 3. Combinations of these primers were used in a series of genomic PCRs in order to amplify the Id3 promoter. Since a minimum of 1000 base pairs was the target region for the promoter, the Forward 1 primer was not used in the PCR experiment; however this primer was important for the nested PCR verification of the promoter sequence.

Table 1 shows the length of each primer sequence shown in Figure 3. It also shows the prospective PCR band sizes of mixed and matched forward and reversed primers that will be used in a PCR in order to validate the proximal human Id3 promoter region.

Also, the figure highlights the actual promoter region that was found using the Forward primer sequence: TGTGAGATTCTTGAAGCAGGA – F2 – and the Reverse primer sequence: GGAAAAAGCAATTCTGGAAG – R2 - which is validated in Figure 4.
PCR of Id3 Promoter with Primers

Figure 4 shows the results from the PCRs using the various combinations of the primers described in Table 2. The PCR product sizes in the resulting picture were cross referenced with the expected results shown in Figure 3. The band size in lane two falls within the 1000 bp range which correlates with its expected size of 1038 bp with forward and reverse primers of F2 and R2. Therefore, the conclusion was made that the Id3 promoter was in lane two and was approximately 1000 bps long, Figure 4.

Id3 Human Promoter Sequence

![Id3 promoter sequence diagram]

Figure 3. Primer sequences located in the human Id3 promoter. PCR primers were developed using the 1.5 kb proximal promoter shown above. The red, blue, and yellow bolded coded sequences represent the forward primers while the green, orange, and purple bolded sequences represent the reverse primers. The black bolded sequence represents the ATG start site that is required for the initiation of translation. The promoter region which is shown in grey was identified using the primers F2 and R2.
Table 1. Location and length of PCR primers within the Id3 promoter sequence.

<table>
<thead>
<tr>
<th>Forward</th>
<th>Reverse</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. 25 – 45 (20)</td>
<td>1. 1494 – 1510 (17)</td>
</tr>
<tr>
<td>2. 500 – 521 (21)</td>
<td>2. 1518 – 1538 (21)</td>
</tr>
<tr>
<td>3. 800 – 818 (19)</td>
<td>3. 1597 – 1616 (20)</td>
</tr>
</tbody>
</table>

Table 2. Expected band sizes of the promoter in PCR reactions. Column 1 denotes the primers pairs that were used in order to determine the promoter region. The second Column represents the starting location of the forward primer and ending location of the reverse primer within the sequence shown in figure 3. Column three represents the expected band size of the promoter which can be found by subtracting the reverse primer from the forward primer.

<table>
<thead>
<tr>
<th>PCR Primers</th>
<th>Amplified Sequence Length</th>
<th>Product Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. F2, R1</td>
<td>500-1510</td>
<td>1010</td>
</tr>
<tr>
<td>2. F2, R2</td>
<td>500-1538</td>
<td>1038</td>
</tr>
<tr>
<td>3. F2, R3</td>
<td>500-1616</td>
<td>1116</td>
</tr>
<tr>
<td>4. F3,R2</td>
<td>800-1538</td>
<td>738</td>
</tr>
</tbody>
</table>
Figure 1. Gel Image of Id3 Promoter Region. Genomic PCRs were performed on the DU145 human prostate cancer DNA cell line using several PCR primers as describe in Table 1. The expected result of the PCR experiment was an approximate band size of 1kb, which was obtained using the Forward 2 and Reverse 2 primers in lane 3. Lane 1 was comprised of the 1kb base pair ladder while lanes 2-5 are PCR reactions with varying primers – F2, R1: F2, R2; F2, R3; and F3, R2. The locations of these primers are denoted in figure 3.

**Promoter Confirmation and Verification**

Nested PCR

The PCR product in lane 3, Figure 1, was extracted, gel purified, and used as template in nested PCR in order to confirm that the product was actually the Id3 promoter. As shown in figure 5, the nested PCR using the primers forward 1 and reverse 3 on the Id3 promoter with the F2 and R2 primers yielded the approximate 710 bp band size that was expected. Therefore, the region between the F2 and R2 primers was confirmed as the Id3 promoter.
Figure 5: Nested PCR Confirming the Human Id3 Promoter Region. A nested PCR was performed using DNA from the 1kb PCR product yielded in Figure 4 and the various primers that surrounded that region. Since the band was made from the F2,R2 product, this band should be nested outside of the F3,R1 primer — as shown above.

Restriction Digest

After the nested PCR confirmed the amplified region, the 1039 bp PCR amplified promoter was ligated into the pGEM-T Easy shuttle vector for sequencing and further analysis. In order to confirm that the promoter was inserted into the plasmid, a restriction enzyme digest using the EcoRI enzyme was performed (shown in Figure 6). The pGEM-T Easy plasmid is 3015 bps long and flanked by EcoRI sites at 52 and 70 bps. Hence digestion of the ligated plasmid should render a 1038 bp band — as seen below.
Figure 6. EcoRI restriction enzyme digest on the pGEMT-Easy plasmid with the Id3 promoter insert. An restriction enzyme digest was conducted in order to determine if the Id3 plasmid was properly inserted into the pGEMT-Easy plasmid. The pGEMT-Easy plasmid is 3015 bps long and has EcoRI sites at 52 and 70 – each site is located near both ends of the plasmid insertion site. A band size of 1kb is expected if the promoter is ligated properly inside of the pGEMT-Easy vector. In this figure, in lane 1 is the 1kb DNA ladder, followed by the restriction enzyme digest product that shows a promoter band size of 1kb.

Sequencing

The pGEMT-Easy + Id3 plasmid was sequenced using the T7 primer at Morehouse School of Medicine’s Molecular Biology Sequencing Lab. The nucleotide sequence returned is shown below in Figure 7.

Figure 7. Sequenced Id3 promoter. Nucleotide sequence of the human Id3 promoter that was ligated into the pGEMT-Easy plasmid. Promoter was sequenced using a T7 primer.
This sequence was aligned with its reference sequence – provided by the gene database – using the NCBI align two sequences BLAST tool. As shown in Figure 8, Blast results showed 100% identity with the Id3 promoter. These experiments confirmed that the cloned sequenced was the Id3 promoter.

Sequence 1: 1.5 kb Promoter Sequence
Length: 1680 (1...1680)

Sequence 2: Sequenced Promoter Plasmid
Length: 988 (1...988)

Score = 1708 bits (888), Expect = 0.0
Identities = 968/1012 (95%), Gaps = 6/1012 (0%)

Figure 8. Promoter sequence verification via the NCBI Blast program. The sequenced insert from the PGEMT Easy plasmid was blasted against the initial 1.5kb Id3 promoter sequence in order to determine if the sequences were the same. In order for sequences to be considered a match, an identity percentage of 88% or higher must be returned. The Id3 promoter and its reference sequence returned a score of 99% which fully confirmed validity of the 1kb the promoter sequence.

Plasmid Verification

After the promoter sequence was verified, the phosphate ends of the promoter were dephosphorylated so that the promoter could be ligated into the pGLuc lucifearse plasmid – thus, making the pGLuc+Id3 plasmid. Upon completion, the pGLuc+Id3 plasmid was verified via digestion using EcoRI and HindIII, shown in figure 9. Plasmid verification was conducted to in order to ensure the continuity of the pGLuc+Id3 plasmid before transfections were performed.
Figure 9. Gel confirmation of pGLuc+Id3 plasmid. In order to confirm that the pGLuc+Id3 plasmid was ready for transfection, the pGLuc+Id3 plasmid was digested with EcoRI and HindIII. Both EcoRI and HindIII returned the expected bind sizes of 1000 and 200 respectively, shown in lanes 1 and 2. Luciferase Plasmids with Id3 Promoter Insert (pGLuc+Id3) were ran on the gel as controls as well, shown in lanes 3 and 4. The last lane contains the molecular DNA ladder.

Once the restriction digests verified the correct size of the pGLuc+Id3 plasmid, the pGLuc+Id3 and the PCMV control plasmid were transfected into human prostate cancer cells lines in the presence and absence of serum.

Transfection

Approximately forty-eight hours after transfections were completed, luciferase assays were performed in order to measure the regulatory activity of each promoter. However, the initial readings for luciferase activity in all plasmids and cell lines were extremely low. Therefore, pGL3 control and enhancer plasmids were transfected into an easier transfecutable HeLa cell line and used as a control for both transfection and luciferase activity.
Luciferase Assay Results

Figure 10. Most successful luciferase assay. Luciferase plasmids (pGLuc, pGLuc+Id3) and luciferase plasmid controls (pCMV, pGL3-Control, pGL3-Enhancer) were transfected into DU145 and HeLa cells in the presence and absence of serum. Under each plasmid is the plasmid’s perceptive luminometer reading for luminescence activity. The pCMV-GLuc control’s luminescence activity corresponds to New England Bio Labs expected plasmid activity of 10000; however, the pGLuc+Id3’s reading of 32.48 is not accurate enough to correspond with adequate promoter activity.

All luciferase assays were unsuccessful because luminometer readings were extremely low, even in the controls. The best reading was found in figure 10 where only the pCMV-Gluc control plasmid expressed high levels of lumination. Therefore, a bioinformatics approach was employed to further investigate the potential mechanism regulating Id3’s promoter activity.

Bioinformatics

TESS Data

All of the bioinformatics experiments conducted aided in finding the potential transcription binding sites in the human Id3 promoter that transcriptionally regulate Id3 in prostate cancer. The most helpful tool in the process was the TESS database which was used to identify potential transcription factor binding sites and their location within the Id3 promoter (Figure 11).
Figure 11. Partial list of the transcription factor binding sites found in the Id3 promoter. This diagram is a partial list of transcription factors binding sites and their position in the Id3 promoter returned from the TESS database search. The factor lane denotes which factor (s) the model represents; The model lane represents the site string or weight matrix that was used in order to pick the perspective site; the Beg field denotes the start of the site in the query sequence – numbered from 1; The Sns lane denotes the sense of the site: N – normal and R – reverse complement; the Len lane denotes the length of the site; The Seq lane denotes the matching portion of the query sequence colored or cased to indicate mismatches; La denotes the log-likelihood score, higher is better; La/ denotes the La/Len, higher is better, max is 2.0; Lq denotes La/L_M, where L_M is the maximum La possible for the site model, higher is better, best is 1.0; Ld denotes the L_M – La, 0 is the best, higher is worse.

In order to make the transcription factor list manageable and ultimately more functional, the following parameters were set: The La value, which represents the log-likelihood score, was formatted in decreasing order starting from 22 because higher numbers are indicative of accuracy; La/ value, which denotes La/Len, was formatted in decreasing order starting from 2 because higher numbers are indicative of accuracy; The Lq value, which represents La/L_M where L_M is the maximum La possible for the site model, was formatted in decreasing order from 1 because 1 is more indicative of accuracy; and the Ld value, which denotes the L_M – La, was formatted in increasing order starting from 0 because lower numbers are indicative of accuracy.
Transcription Factor Sites Found on the Id3 Promoter

After formatting the list of transcription factor bindings sites based on the transcription factor’s La, La’, Lq, and Ld values, the top sites listed in the spreadsheet are believed to provide the most accurate list of potential transcription factors found within the Id3 promoter (Table 3). A schematic showing the potential transcription factor sites as generated from the TESS analysis is shown in Figure 12.

Table 3. Transcription factor binding sites’ location in Id3 sequence (TESS Database). Lanes 1 and 4 represent the list of potential transcription factor binding sites. Lanes 2 and 5 represents the location of the binding sites relative to the sequence input into the TESS Database. Lanes 3 and 6 represents the binding sites’ location in the Id3 promoter.

<table>
<thead>
<tr>
<th>Transcription factor</th>
<th>TESS Position</th>
<th>Promoter Position</th>
<th>Transcription factor</th>
<th>Tess Position</th>
<th>Promoter Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP-1</td>
<td>63</td>
<td>867</td>
<td>c/ebp beta</td>
<td>690</td>
<td>240</td>
</tr>
<tr>
<td>CP2</td>
<td>85</td>
<td>845</td>
<td>evi-1</td>
<td>775</td>
<td>155</td>
</tr>
<tr>
<td>NF Kappa B</td>
<td>90</td>
<td>840</td>
<td>Tcf-2alpha</td>
<td>816</td>
<td>114</td>
</tr>
<tr>
<td>YY1</td>
<td>106</td>
<td>824</td>
<td>Srf</td>
<td>824</td>
<td>106</td>
</tr>
<tr>
<td>c/ebp alpha</td>
<td>107</td>
<td>823</td>
<td>sp1</td>
<td>835</td>
<td>95</td>
</tr>
<tr>
<td>Tcf-4e</td>
<td>177</td>
<td>753</td>
<td>c-ets-2</td>
<td>847</td>
<td>83</td>
</tr>
<tr>
<td>TCF-3</td>
<td>183</td>
<td>747</td>
<td>e2f+p107</td>
<td>876</td>
<td>54</td>
</tr>
<tr>
<td>Tcf-4e</td>
<td>185</td>
<td>745</td>
<td>sp1</td>
<td>891</td>
<td>39</td>
</tr>
<tr>
<td>c/ebp alpha</td>
<td>207</td>
<td>723</td>
<td>Tcf-2alpha</td>
<td>909</td>
<td>21</td>
</tr>
<tr>
<td>MGRF</td>
<td>289</td>
<td>641</td>
<td>Tcf-4e</td>
<td>909</td>
<td>21</td>
</tr>
<tr>
<td>c/ebp alpha</td>
<td>389</td>
<td>541</td>
<td>Tcf-2alpha</td>
<td>914</td>
<td>16</td>
</tr>
<tr>
<td>c/ebp alpha</td>
<td>681</td>
<td>249</td>
<td>POU1F1a</td>
<td>924</td>
<td>6</td>
</tr>
</tbody>
</table>
Figure 12. Schematic diagram of promoter region with transcription factor binding sites provided by the TESS database. The presence of putative transcription factor binding sites in the human ld3 genes was identified using the TESS database. The locations of these binding sites in the proximal ld3 promoter (−930 to 0) are indicated schematically in the diagram.

Sequence Alignment

Since ld genes share a similar genomic organization, including exon/intron boundaries, in their coding regions that are consistent with evolution from common ancestral ld genes, it is believed that ld3 regulatory mechanism in various species is potentially conserved. Sequence conservation and alignment analyses were conducted on the proximal 1kb human, mouse, and rat promoters in order to determine the regulatory domains within the promoters. The alignment of all three sequences produced an evolutionary tree, shown in Figure 13, which illustrates that all three promoters are derived from common ancestry.

Figure 13. Human, mouse, and rat phylogenetic tree. This tree shows the evolutionary relationship between the Human, Mouse, and Rat 1kb proximal promoter regions.
Since the human, rat, and mouse sequences share similar sequence conservation, the conserved regions within the promoters were analyzed. The conservation parameter was set to only return regions that were 70% or more conserved. All three alignments showed two major conserved regions between all three promoters (figures 14 – 16). Overall, the mouse and the rat promoters are the most conserved and over 95% of its non-coding sequence is conserved which is shown in the promoter regions of 1000 to 650 (Figure 14). In Figure 15, the alignment of the human and rat promoters showed that the region between 380 and 485 is almost 100% conserved, while Figure 16 shows the human and mouse promoters have a more expansive range of conservation from approximately the 80th base pair to the 500th. In both analyses, the nucleotide regions of 410 to about 825 were equally conserved.

Conserved Regions in Human and Mouse

Yeh and Lim conducted a systematic promoter characterization on the murine Id3 promoter and investigated the mechanisms involved in its regulation (Yeh and Lim 2000). They cloned the 1kb murine promoter region and determined that a 180 bp proximal Id3 promoter fragment was sufficient for full transcriptional activity in the C2C12 myogenic cell line. Within this region, they identified putative transcription factor binding sites that regulate murine Id3 promoter activity, (Figure 17).
Figure 14. Conserved regions in the mouse and rat ld3 promoters. The highlighted regions in both sections of the figure indicate the two conserved regions within the proximal 1kb mouse and rat ld3 promoters. The top portion of the figure is a pictorial view of the conserved regions in the promoters while bottom portion shows the covered regions within the nucleotide alignment of the two promoters.
Figure 15. Conserved regions in the human and rat Id3 promoters. The highlighted regions in both sections of the figure indicate the two conserved regions within the proximal 1kb human and rat Id3 promoters. The top portion of the figure is a pictorial view of the conserved regions in the promoters while bottom portion shows the covered regions within the nucleotide alignment of the two promoters.
Figure 16. Conserved regions in the human and mouse Id3 promoters. The highlighted regions in both sections of the figure indicate the two conserved regions within the proximal 1kb human and mouse Id3 promoters. The top portion of the figure is a pictorial view of the conserved regions in the promoters while bottom portion shows the covered regions within the nucleotide alignment of the two promoters.
Figure 17. Schematic diagram of the putative transcription factor binding sites in the murine 200 bp Id3 promoter (Yeh and Lim 2000). The putative transcription factors and their locations within the proximal 200 bp region of mouse Id3 promoter were determined by analyzing the results of a computationally based search of the TRANSFAC database.

In order to compare Yeh and Lim results to our findings, first, a schematic diagram was designed based on the Vista human and mouse alignments, shown in Figure 18. This diagram inverted the Vista alignment’s conservation schematic so that the conserved regions could be viewed relative to the genes’ ATG start site.

Based on the schematic diagram in Figure 17, the mouse promoter’s regulatory region extends from -200 bps of the ATG translational start site. However, its conservation with the human promoter begins 169 bps from the ATG start site. According to Yeh and Lim, only transcription factors PEBP2 and ATF are located within this twenty-one bp region. However, when these transcription factors were crossed referenced with the full list of human Id3 transcription factors retrieved from the TESS Database, these factors were not found. Nonetheless, there were multiple occurrences of the Sp1 transcription factor found within both the full list and the formatted list of transcription factors retrieved from the TESS Database.
Microarray Data

Prostate cancer cell line microarray data for the expression of Id3 and potential transcription factors that are predicted to bind to Id3 were analyzed in order to determine expression levels in prostate cancer cell lines, whether the putative transcription factor binding sites shown in Figure 12 are expressed, and whether their expression levels are statistically significant enough to aid in Id3’s upregulation in prostate cancer cells.

The expression of the constitutively expressed gene, GAPDH, was used to normalize the expression levels within each microarray data set. GAPDH functions as a known housekeeping gene and serves as a baseline control in order to normalize gene expression levels in the comparisons of gene expression data (Barber, et al. 2005). Two forms of GAPDH – GAPDH1 and GAPDH2 – were present in each data set of the cell line microarray expression data.

Figure 19 depicts each GAPDH expression levels in the three prostate cell lines – LNCap, DU145, and PC3.
Figure 19. Microarray expression levels of GAPDH in prostate cancer cell lines. GAPDH expression (raw value), in the respective microarray chips, was used to normalize expression levels of transcription factors in this study.

The mean of the GAPDH expression scores was used to normalize Id3’s expression levels in prostate cancer cell lines LNCaP, DU145, and PC3. The data below indicates that Id3 is expressed in all three cancer cell lines. It also depicts that in the least aggressive prostate cancer cell line, LNCap, Id3 expression is four times higher than the highly metastatic PC3 cell line.
Figure 20. Raw values of Id3 expression in microarray experiments across prostate cancer cell lines PC3, DU145, and LNCaP. Id3’s expression is observed in the microarray data set. Its expression was normalized by dividing its expression levels by its mean GAPDH score.

Transcription factors retrieved from the microarray data were validated via the Ensembl Database (http://www.ensembl.org/index.html) by way of searching the Gene database (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene). Once the transcription factor was validated, statistical analysis were performed in order to normalize its microarray score using GAPDH values. Data was normalized by using the following equation:

\[
\frac{\text{microarray expression data}}{\text{GAPDH values/number of GAPDH entries}} \times 1000
\]

(1000 is the preferred normalized data reference).

Once data was normalized, the transcription factors’ microarray values were evaluated and scores higher than 100 in DU145 and PC3 and 1800 in LNCap were deemed significant to the possible regulation of Id3’s expression level in its perspective prostate
cancer cell line. Transcription factor expression levels in DU145, PC3, and LNCaP prostate cancer cell lines retrieved from microarray data are depicted in figures 21 – 23 respectively.

Results showing the overexpression of transcription factor levels in figures 21 – 23 combined with the identification of putative transcription factors within the promoter construction shown in Figure 12 suggest that investigating transcription factors YY1, SP1, SRY, SRF, NIT2, NFKappaB, HSF, Evi-1, C/EBPb, and C/EBPa may have a potential role in the transcriptional mechanism that regulates Id3 expression prostate cancer cells.

Figure 21. Relative expression level of transcription factors predicted to bind to Id3 promoter (DU145). Transcription factor binding sites expression levels were gathered from microarray expression data in DU145 cells. Once microarray data was normalized, the transcription factor binding sites’ microarray signal scores were evaluated. Transcription Factor scores higher than 100 deemed the transcription factor significant to Id3’s regulation.
Predicted To Bind to the Id3 Promoter

Figure 22. Relative expression level of transcription factors predicted to bind to Id3 promoter (PC3). Transcription factor binding sites expression levels were gathered from microarray expression data in PC3 cells. Once microarray data was normalized, the transcription factor binding sites’ microarray signal scores were evaluated. Transcription Factor scores higher than 100 deemed the transcription factor significant to Id3’s regulation.

Transcription Factors Expression in Id3 Gene

Figure 23. Relative expression level of transcription factors predicted to bind to Id3 promoter (LNCap). Transcription factor binding sites expression levels were gathered from microarray expression data in LNCap cells. Once microarray data was normalized, the transcription factor binding sites’ microarray signal scores were evaluated. Transcription Factor scores higher than 1800 deemed the transcription factor significant to Id3’s regulation.
Background information on each transcription factor was taken from data compiled from the TESS Database (Schug and Overton 1997). This database obtains transcription factor data via published findings in peer reviewed scientific journals.

**Ying Yang 1**

Yin Yang 1 (YY1) is an ubiquitously expressed transcription factor belonging to the GLI-Kruppel class of zinc finger proteins. YY1 plays an integral role in normal biological activities such as embryogenesis, differentiation, DNA replication, and cellular proliferation. YY1 regulates the genes involved in these activities by its ability to initiate, activate, or repress transcription usually through mechanistic means of direct activation or repression, indirect activation or repression via cofactor recruitment, or activation or repression by disruption of binding sites or conformational DNA changes (Shi, et al. 1991, Deng, Wan and Sui 2007). Overexpression and/or activation of YY1 is associated with unchecked cellular proliferation, resistance to apoptotic stimuli, initiation of tumorigenesis and metastatic potential (Sui 2009).

YY1 expression was observed in all three cancer cell lines. This normalized level ranges from moderately high in DU145 and LNCaP to extremely high in PC3. The statistical data of YY1 in all three cells lines yield expressions levels of 600.6, 719.8, and 9814.4 in DU145, PC3, and LNCaP cancer lines respectively; these levels of expression are greater than the microarray standard threshold. This data directly correlates to YY1’s known overexpression/activation capabilities of initiating tumors and its metastatic capabilities. Therefore, YY1 expression may play a role in Id3 upregulation in cancer cells and its activation should be investigated for possible tumor suppressor capabilities.
**Sp1**

Sp1 is a zinc finger transcription factor that is localized in the nucleus and important to biological processes such as regulating gene expression and the cell cycle, hormonal activation, apoptosis, angiogenesis, and oncogenesis (Chu and Ferro 2005, Jiang, et al. 2007). Sp1 binds to DNA and enhances gene transcription via binding to GC-rich decanucleotide recognition elements within promoter sequences (Jiang, et al. 2007). Sp1 has been identified in the promoter region of numerous genes affiliated with prostate cancer.

Sp1 expression in LNCaP and DU145 is extremely low while its expression level in PC3 is moderately low. The statistical data of Sp1 in all three cells lines yield expressions levels of 7.6, 34.6, and 200 in DU145, PC3, and LNCaP cancer lines respectively; these levels of expression are lower than the microarray standard threshold. Despite the fact that Sp1 has been shown to regulate several prostate cancer genes in which it is expressed and has numerous sites with the murine and human Id3 promoters including sites within 100 bps of the ATG start site, its expression level does not seem to be significant enough to have impact on the human Id3 promoter’s expression in prostate cancer.

**Serum Response Factor**

Serum response factor (SRF) is a member of the MCM1, Agamous, Deficiens, and SRF (MADS) box class of the beta-Scaffold transcription factor family and can act as either a transcription activator or repressor. SRF induces an optimal TFIID conformation for recruitment of the transcription preinitiation complex and binds to the serum response element in the promoter region of target genes (Norman, et al. 1988, Papavassiliou 1994).
It is integral in processes such as apoptosis, cell differentiation, cell cycle regulation, and cell growth (Patten, et al. 2004, Parlakian, et al. 2004, Mampuru, et al. 1996). SRF is an important determinant of AR activity which is critical for prostate cancer progression. Loss of SRF has been shown to impede prostate cancer cell proliferation (Heemers, et al. 2007).

SRF has moderately low expression levels in all three prostate cancer cell lines. The statistical data of SRF in all three cells lines yield expressions levels of 23.7, 32.5, and 473.7 in DU145, PC3, and LNCaP cancer lines respectively; these levels of expression are lower than the microarray standard threshold. Despite the fact that Id3 is an immediate-early gene and SRF is known to have regulatory roles within these types of genes, SRF’s low expression level indicates that its role in Id3 regulation is minimal.

*Nitrilase Family, Member 2*

The nitrilase family, member 2 (NIT2) protein is a member of the nitrilase super family of proteins. It is ubiquitously expressed in multiple tissues and has been shown to play an important role in cell growth inhibition potentially via tumor suppression (Lin, et al. 2007, Krasnikov, et al. 2009). Its reduced expression in human epithelial cells is associated with accelerated proliferation, in vitro, while abnormal NIT2 expression is linked to malignant progression in humans (Krasnikov, et al. 2009).

NIT2’s expression level in PC3 and LNCaP are highly elevated while its level in DU145 is only slightly high. The statistical data of NIT in all three cells lines yield expressions levels of 208.3, 546.1, and 12917 in DU145, PC3, and LNCaP cancer lines respectively; these levels of expression are significantly higher than the standard microarray thresholds. These raised expression levels correlate to NIT2’s malignant
tendencies. These levels show that NIT2's expression may play a role in Id3 regulation in prostate cancer progression.

**Nuclear Factor Kappa - Light - Chain - Enhancer of Activated B Cells**

The nuclear factor kappa-light-chain-enhancer of activated B cells (NFkappaB) transcription factor is a member of the beta-Scaffold transcription factor family. It exists in the cytoplasm of many eukaryotic cells as either a hetero or homodimer with other structurally related proteins of the Rel family. Once NFKappaB is activated, it regulates genes that control cell proliferation, angiogenesis, invasiveness, and cell survival (Lerebours, et al. 2008). However, incorrect regulation of NFKappaB has been linked to cancer, inflammatory and autoimmune diseases, viral infections, and improper immune development.

NFkappaB has a moderate expression level in Id3 regulation in DU145, LNCaP, and PC3 cells. The statistical data of NFkappaB in all three cell lines yield expressions levels of 290.6, 171.2, and 4118.7 in DU145, PC3, and LNCaP cancer lines respectively; these levels of expression are higher than the microarray standard threshold in PC3 and DU145 cells and significantly higher in LNCap cells. These levels of expression indicate that NFkappaB activity likely poses significant influence on Id3 regulation through means of improper regulation possibly due to mutation.

**Ecotropic Viral Integration Site 1**

Ecotropic viral integration site 1 (Evi-1) is a transcriptional repressor that has a zinc finger motif. It causes repression via its association with histone deacetylases and its recruitment of chromatin-modifying enzymes that create a heterochromatin-like structure that causes transcription repression for extended periods of time (Vinatzer, et al. 2001 ).
Not only does Evi-1 play a critical role in the development of a variety of tissues, it also has been shown to be an oncprotein inappropriately expressed in acute myeloid leukemia where it participates in its (leukemia) progression (Kurokawa, et al. 2000, Izutsu, et al. 2001).

Evi-1 expression level is low in PC3 and LNCaP and is dismally low in DU145. The statistical data of Evi-1 in all three cells lines yield expressions levels of 29.3, 28.4, and 160.7 in DU145, PC3, and LNCaP cancer lines respectively; these levels of expression are significantly lower than the microarray standard thresholds. Even though there is a possibility of Evi-1 being inappropriately expression, due its low statistical data, it is more than likely Evi-1 is playing no integral role in Id3’s expression in prostate cancer.

CP2

CP2 is an ubiquitously expressed member of the grainyhead-like class of beta-Scaffold family of transcription factors. It is a major factor in the regulation of globin expression and cell cycle progression and is linked to erythroid differentiation (Chae and Kim 2003). Evolutionary evidence has shown that CP2 to be an ancestor of the tumor supressor p53 (Kokoszynska, et al. 2008).

CP2 is actively expressed in DU145 and PC3 while its expression in LNCaP is extremely low. The statistical data of CP2 in all three cells lines yield expressions levels of 185.6, 277.4, and 199.4 in DU145, PC3, and LNCaP cancer lines respectively; in the DU145 and PC3 cell lines, these levels of expression are higher than the microarray standard threshold, while in the LNCaP cell line, this level is significantly lower. This data suggests that CP2 may have a functional transcription factor in Id3’s regulation and should be further investigated.
CCAAT/Enhancer Binding Protein Beta

CCAAT/enhancer binding protein beta (C/EBPβ) is a member of the basic leucine zipper transcription factor family. It heterodimerizes with other CCAAT/enhancer binding proteins - CEBP-alpha, CEBP-delta, and CEBP-gamma – and homodimerizes with other specific DNA regulatory regions. C/EBPβ has an important role in the regulation of genes involved in immune and inflammatory responses (Gade, et al. 2008). It also acts antagonistically to the p53 tumor suppressor gene (Ewing, et al. 2008, Reisman and Boggs 2007) and is critical in the survival of Wilms tumor cells (Li, et al. 2005).

C/EBPβ is highly expressed in all three cell lines. The statistical data of C/EBPβ in all three cells lines yield expressions levels of 8963, 575.4, and 10283 in DU145, PC3, and LNCaP cancer lines respectively; these levels of expression are significantly higher than the microarray standard thresholds. C/EBPB’s high/overexpressed in Id3’s microarray data and its ability to function as a protooncogene makes the likely proponent of Id’s upregulation in prostate cancer.

CCAAT/Enhancer Binding Protein Alpha

CCAAT/enhancer binding protein alpha (C/EBPa) is a member of the basic leucine zipper transcription factor family and is sequestered in the cytosol. It forms heterodimers with related proteins CEBP-beta and CEBP-gamma and binds as a homodimer to certain promoters and enhancers. It negatively regulates cell proliferation and induces terminal differentiation of various cell types (Seifeddine, Fulchignon-Lataud and Massaad-Massade 2009). It also functions as a tumor suppressor through its strong antiproliferative actions on cell cycle regulatory proteins (Kumagai, et al. 2009). C/EBP-
alpha is highly expressed at the growth arrest stage in prostate development and regulates androgen receptor signaling (Yin, Lowery and Glass, 2009).

C/EBPa is moderately expressed in all three cancer cell lines. The statistical data of C/EBPa in all three cells lines yield expressions levels of 50.7, 73.8, and 2258.3 in DU145, PC3, and LNCaP cancer lines respectively; these levels of expression are lower than the microarray standard threshold in DU145 and PC3 cells but higher than the LNCaP threshold. However, the microarray data also shows that the C/EBP induced levels in DU145 and PC3 are higher than standard threshold of 100. This may be due to C/EBPa and C/EBPb dimerizing 245bps from Id3’s translational start site. While C/EBPa by itself is not important to Id3’s regulation, its interaction with C/EBPb may make C/EBPa integral to Id3 and its regulation.

Correlation between Id3 and Putative Transcription Factor Expression in Microarray Data

In an attempt to identify potential regulatory patterns that may give insight on how the expression of Id3, in Figure 20, correlates to the expression of the transcription factors identified in the microarray data, scores were given to transcription factors based upon their expression levels in the three prostate cancer cell lines. The transcription factors and Id3 were given a score of (4) if their expression level was high; (3) if their expression level was moderate; (2) if their expression level was low; and (1) if their expression level was extremely low. The results are graphed in Figure 24.

Figure 24 indicates that C/EBP induced’s expression level is closely correlated to Id3’s expression level than any other transcription factor. From this chart, both Id3 and CEBP’s LNCaP levels are low and PC3 levels are moderate. Id3’s DU145 level is high
while CEPB’s DU145 level is moderate. Based on this information, C/EBP induced is more than likely functioning as a repressor of Id3 activity.

Conserved and Aligned Transcription Factor Binding Sites in Mouse and Human Id3 Promoters

Lastly, a final conservation analysis was conducted in order to determine if all of the transcription factor binding sites found in human Id3 promoter from the TESS database, 200 bp murine Id3 promoter region, and the microarray expression data are conserved. All common transcription factors from the above experiments were analyzed for conservation and alignment using the 1kb mouse and human Id3 promoters in the rVista database.

Figure 25 depicts both the conserved and aligned transcription factors in the mouse and human Id3 promoter regions. The conserved binding sites are defined to be predicted binding sites located in the sequence fragments conserved between two species at the level of over 80% over a 24 bp window. Aligned binding sites are those where core positions of the potential binding sites on the sequences corresponded to each other in the alignment. The green tick marks represent conserved binding sites while the red tick marks represent the aligned sites.

In Figure 25, the transcription factors that are aligned and show the greatest degree of conservation are CEBP, PBX1, SRY, and YY1. However, since SRY and PBX1 binding sites are not observed in either promoter computational analysis, these sites are eliminated for being potentially viable candidates regulating Id3 promoter activity. Therefore, based on all the gathered data, investigating the only consistent transcription
Figure 24. Microarray correlation chart. Transcription factors and Id3 were given scores of either (4) if their expression level was high; (3) if their expression level was moderate; (2) if their expression level was low; and (1) if their expression level was extremely low. These scores were used to help predict patterns specific to Id3 regulation in prostate cancer cells. The only significant correlation shown to be found was C/EBP induced expression is correlated to Id3 activity. Therefore, the dimers formed between members of the C/EBP family (in conjunction with some other transcription factor) may be acting as a repressor to adversely affect Id3 expression.
Figure 25. Conserved and aligned transcription factors that have been found in either the human or mouse Id3 promoters or Id3 microarray data. (Top) Transcription factor binding site alignment within the Human and Mouse 1kb promoter regions (red) and the tentative conserved transcription factor binding sites within the Human and Mouse 1kp promoter regions (green). (Bottom) conserved regions within the Human and Mouse 1kb promoter.

Therefore, based on all the gathered data, investigating the only consistent transcription factors of YY1 and the CEBP interaction between factors CEBPa and CEBPb may lead to discovering the potential mechanism of Id3 regulation in prostate cancer.
CHAPTER FIVE
DISCUSSION

From this research, the 1kb proximal promoter of the human Id3 gene was identified and confirmed using three different techniques: Nested PCR, Restriction Enzyme Digest, and DNA Sequencing. After the promoter was identified and cloned, it was transfected into three different prostate cancer cell lines and unsuccessfully measured for promoter activity via luciferase assays.

There are several suspected reasons that the luciferase assays were unsuccessful. One reason for the low lumination levels could be due to low transfection rates. The luciferase assay had to be completed within a forty-eight hour period without changing media. Under those conditions, there was no experiment that could be conducted to test transfection efficiency. Another, suspected reason that the luciferase assays may not have worked is because cells and plasmids may have been contaminated due to factors such as preparation and storage of plasmids, the air filtration in the hoods during cell culture, transportation of cells and plasmids through and to other facilities, and minimal exposure to light sources during luciferase assay preparation time. With one, if not all, of these factors in play, further experimentation could not be conducted to determine the minimal promoter region needed for transcriptional activity and thus which transcription factors are influencing Id3 promoter activity. Therefore, a bioinformatics approach was employed to investigate the potential transcription factors regulating Id3 promoter activity.
This approach began with the Id3 promoter sequence being queried through the TESS database which returned a list of potential transcription factor binding sites and their locations. Transcription factors on this list were formatted based upon their log likelihood scores which allowed for the most accurate scores to be shown first. Then, a human and mouse Id3 promoter conservation analysis was conducted. This list yielded two conserved regions: (1) mouse -929 to -484: human -916 to -502; and (2) mouse -415 to -169: human -389 to -189. The second conserved region encompasses part of the 200 bp characterized mouse promoter that Lim and Wu discovered to be sufficient for murine Id3 promoter activity. Therefore, the small region of 20 bps between where the vista analysis begins and Wu and Lim promoter characterization ends was analyzed to determine if similar transcription factors were located in those regions.

Since there were no similar transcription factors located in that region, Id3 microarray data was analyzed to determine the potential transcription factors regulating Id3 activity. This analysis lead to the identification of nine potential transcription factors: YY1, SP1, SRF, NIT2, NFKappaB, HSF, Evi-1, C/EBPb and C/EBPa. These transcription factors were chosen based upon their presence in both the TESS database/promoter characterization pictorial and microarray data/microarray graphs. Despite the fact the SP1 and NFKappaB are known exhibit regulatory roles on the promoters of genes involved in cancer initiation and development, three transcription factors – YY1, C/EBPb and C/EBPa – were indentified to be the potential regulators of Id3 in prostate cancer via either their extremely high presence in the microarray data or based upon a pattern that correlated the microarray data of Id3 expression to transcription factors expression.
YY1 is a transcription factor that differs from many others because it has the ability to initiate and regulate transcription through the activation and repression of cellular and viral genes (Gordon, et al. 2006, Donohoe, et al. 1999, Nguyen, et al. 2004). Despite there being a dearth of information regarding the regulation of YY1 activity, investigators have shown that an increase in promoter activity in some genes correlated with an increase in YY1 expression (Gordon, et al. 2006).

YY1’s expression levels have been noted in several human cancer cell lines including prostate cancer. Seligon et al conducted a study that used tissue microarrays to investigate YY1’s expression and localization in 1364 tissue samples from 246 prostate cancer patients who lacked hormones and had radical prostatectomies performed. (Seligson, et al. 2005). In this study, they showed that the staining intensity and frequency measures for YY1 nuclear and cytoplasmic expression were higher in neoplastic tissues and predominantly elevated in early malignancy samples compared to matched benign cells. Furthermore, they were able to show that YY1 expression is predominantly elevated in early malignancy, as well as in tumors of intermediate to high morphologic grade (Seligson, et al. 2005).

Concurrently, the microarray data from Chaudhary’s Lab showed that YY1 is not only detected in all three prostate cancer cells lines but its expression is up regulated in all three cell lines also. YY1’s high expression leads to the notion that it may be playing a distinct role in Id3 upregulation in prostate cancer; however, further experimentation is needed to validate this position and determine the extinct of its role in prostate cancer. More than likely, YY1 is working in conjunction with some other transcription factor(s) to influence Id3 activity.
In several types of tumors, CEPBa has shown to have decreased expression and/or a mutation. However, ironically similarly to YY1, there is not much information known regarding the regulation of CCAAT/ enhancer proteins in prostate cancer. Yen et al conducted an experiment where they examined the growth properties of prostate cancer cells stably expressing C/EBPalpha. They noticed that the forced expression of C/EBPa in prostate cancer cell lines accelerated cell growth, stimulated cells into the S and G2 phases of cell cycle, and enhanced anchorage-independent colony formation (Yin, Lowery and Glass, 2009). Conversely, the overexpression of C/EBPb has been shown in many tumors. The most significant investigation regarding C/EBPb and prostate cancer was conducted by Wang et al. They showed that C/EBPb was overexpressed in proliferative inflammatory atrophy lesions and concluded that chronic inflammation is integral in the stimulation of C/EBPb expression in prostate epithelium (Wang, et al. 2007).

From the Id3 microarray data, not only was C/EBP induced expression shown to correlate with Id3 expression, C/EBP induced was shown to be more than likely functioning as a repressor of Id3 activity. Furthermore, C/EBPa was moderately expressed in all three prostate cancer cells (yet its expression was only significant in low metastatic LNCaP cells) and C/EBPb was overly expressed in all three cell lines. These results are consistent with the research findings mentioned above. Even though more experiments are necessary to determine whether these findings are related and even significant to Id3’s regulation in prostate cancer, it is believed that C/EBPa and C/EBPb both are playing integral roles in Id3 regulation.
Future Direction

Authenticate Promoter Sequence

Despite the human Id3 promoter region being verified by three different methods, negligible promoter expression was exhibited during every luciferase assay experiment. Therefore, the first course of action to furthering the research performed in this thesis would be to ensure that the Id3 promoter sequence is correct. The most probable courses of action are either to start over with fresh sequences and plasmids to ensure there are no containments in any of the plasmids or just to reclone the Id3 plasmid into pGEMT-Easy plasmid to make sure there are no containments in either of those plasmids. These methods would be used as a verification tool that ensures that all plasmids are correct and usable.

Site Directed Mutagenesis

After validating the promoter, the next experiment that should be conducted is a site directed mutagenesis of the identified transcription factors thought to be affecting Id3 promoter activity. In this procedure, chemically modified anti strand DNA fragments are constructed against each transcription factor binding site sequence. These fragments will insert a mutation, via PCR, into the Id3 promoter at the specified binding sites of YY1, C/EBP, C/EBPb, and SP1, individually. Once the mutation is cloned into the promoter, luciferase assays are conducted in order to determine how the mutation in the specified transcription factor drives Id3 promoter activity and which mutation has the most effect on Id3 promoter activity.
Deletion Mutagenesis

The next step to fully characterizing the Id3 promoter is the creation of deletion mutants. In this experiment, DNA sequences are removed from either end of the Id3 promoter using endonucleases. Luciferase assays are conducted on the mutants produced from this experiment and aid in determining the minimal region needed for optimal promoter activity. Additionally, this experiment combined with the site directed mutagenesis experiment are used to determine which binding site promotes transcriptional activity.

Chromatin Immunoprecipitation Assay

Lastly, a chromatin immunoprecipitation assay will be conducted in order to identify the position of the transcription factors that were identified in the experiments above. The technique involves crosslinking of proteins with DNA, fragmentation and preparation of soluble chromatin followed by immunoprecipitation with an antibody recognizing the protein of interest.

Conclusion

Gene regulation in eukaryotic organisms is achieved through many complex mechanisms, most of which are not well understood. However, the work conducted on the Id3 promoter has laid the foundation for further investigation and understanding of the various mechanisms for its role in prostate cancer. Through this research, the promoter region of Id3 was located and verified. Transcription factors YY1, C/EPBa, and C/EPBb have all been identified as potential Id3 regulators via comparative promoter analysis and microarray data. In summary, while more work is required to validate the
mechanism that regulates Id3's expression in prostate cancer, the work presented in this thesis provides an excellent starting point for further examination.
APPENDIX A:

TRANSCRIPTION FACTOR BINDING SITES AND THEIR AFFYMETRIX GENE IDS

<table>
<thead>
<tr>
<th>Transcription factors</th>
<th>Affymetrix gene id</th>
<th>Transcription factors</th>
<th>Affymetrix gene id</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADR1</td>
<td>HFH-2</td>
<td>AP-1</td>
<td>HSF</td>
</tr>
<tr>
<td>AML-1a</td>
<td>210365_at</td>
<td>(Jun)</td>
<td>201465_s_at</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Fos)</td>
<td>209189_at</td>
</tr>
<tr>
<td>AP-1</td>
<td>HSF</td>
<td>Brn-2</td>
<td>MIG1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C/EBP (induced)</td>
<td>221249_s_at</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C/EBPa</td>
<td>204039_at</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C/EBPb</td>
<td>212501_at</td>
</tr>
<tr>
<td>Cap</td>
<td>200625_s_at</td>
<td>Cap</td>
<td>200651_at</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCAAT</td>
<td>Pbx-1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CdxA</td>
<td>214971_s_at</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CF2-II</td>
<td>SP1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c-Myb</td>
<td>214732_at</td>
</tr>
<tr>
<td>CP2</td>
<td>209338_at</td>
<td>CRE-BP</td>
<td>205931_s_at</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Croc</td>
<td>STRE</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dfd</td>
<td>Th1/E4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Evi-1</td>
<td>215851_at</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GATA-1</td>
<td>210446_at</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GATA-x</td>
<td>XFD-1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GCN4</td>
<td>YY1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hb</td>
<td>200047_s_at</td>
</tr>
</tbody>
</table>
APPENDIX B:
CORRECT NUMBERING OF THE HUMAN ID3 SEQUENCE

-1493 TCCAGGGTTT TTTGGCTTT TTTTAAGCT AAGGCTCTTG GTTTAAAAA ACGTCCGGG
-1433 AGGAGGTTT TTTGGGCA AAGAGCCCTG TGGTTGATG GAAAAAGGCA AGGCTCCGAG
-1373 AGGCCCTTT GTATATATG TAGTCGACA TTGCGCAATG CTAAAAGGC TTTCTTGGGC
-1313 CCTTGAGCA TTTTGCCTA CAGATCCAA TTTCGCAAC CTTGCACTT CTCAATTGG
-1253 AAGGATTCT AGAATATCAG TTGAGGACAG CACTTTTTT ATCTCAAGG CAAITAAAGC
-1193 CAGAATACCT GGGTTGCTC AGAGATTACG GACAAATAP TGTGCGGCT AGCCTAGA
-1133 AGCCAGGGTG GCCTAGTTCC AGTCACTGTC CTTTCACCC CTCTGCGCAG TCGGGCGG
-1073 TCTAGGCCCC TTTGCGCTTG TAACTAGTG CACAGATG CAACTCTTG TTTCTTGGT
-1013 TATTCCTGCT CCCCCCAATT CTGAGATTT TGAAGGCAAG AGCCACTGTT TTAAGCGGAG
-953 GACGCTCATA ATATATTCCAG ATGATTTCA TAGGAAGCA ACCTGGGGA GAAAGCAGCT
-893 GTTCTATTTC CAAAGGGGAG AGCCAGACGT TGGAGTGGC GCTCCCCAAT TCTCTAGGGA
-833 TGCAGCTAGA AGGCACAGGA AGAACTCAGG ATACCTGCGT AGCCTCGTTT GTTGAATAC
-773 AGGCCTTGGT TGCCGCCTCC TCTGTGGCTT TCTGTGGCTT TCTGAGGGA GGGCGAGCTA
-713 GTCTACGACA GCTACAGAT GC CTTCCCAATG GCGRAAGGGA
-653 CAGACATAGC AGCAGTGGAT TTACCATCTT TTTTTAACA GAGCTCGTGG GAAAGTGAAC
-593 TGGTGCGTTT GTTGCACATA GTTGCACATA GTTGCACATA GAAATGAGTG TAAAGCTGGT
-533 GACCTCTCAA GCCCTAGCTT GCTATATTT ATGCTCTAGT GAGATTTAG GGGCGAGGGA
-473 TGCAGGCTCT CTTTTTCCCT ATGGCAAAAAG CAGGGGGCTA AGCTTTGCT TTTCCCGG
-413 TCCTCAGAAA ATATATGGCT CTGTGCTGCT GTCTCATCTT GCGCCCACTTG
-353 TCTTCAGAAA ATATATGGCT CTGTGCTGCT GTCTCATCTT GCGCCCACTTG
-293 AGCCGGTACGC GCAATAGCTG TTTCTCTATG GAAATGCGAG CTTCACGCAA ATAAATATTT
-233 TTTCCCTGG GTTCAAGATG AATCGCGATG GCGAGCTGGT TTGCGGCATC GAGGAGGCA
-173 GAAAGGGGAA AGACTGCGGC ACCATTTTCT CAAACCTTT GTCTGAAA PTCTGCGTT
-633 TCTATGGATT GGGGAGCTG TCTGTGGCTT TCTGTGGCTT TCTGAGGGA

96 AATGCCGGAC GCCTAGGTTT CCGTGGCTCT ATACCTTTT TATAGTTAAA AATGCGGGGTT ATGCGGGGTT TCGACGTTTTA
### APPENDIX C:

**FORMATTED TESS DATABASE RESULTS**

<table>
<thead>
<tr>
<th>Factor</th>
<th>Model</th>
<th>Beg</th>
<th>Sns Len</th>
<th>Sequence</th>
<th>L a</th>
<th>L a/</th>
<th>L q</th>
<th>L d</th>
</tr>
</thead>
<tbody>
<tr>
<td>T00759 Sp1</td>
<td>R04741</td>
<td>39</td>
<td>R 11</td>
<td>GGCGGCCTG</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T01059 MN1a</td>
<td>R08281</td>
<td>343</td>
<td>R 10</td>
<td>GCTCCTTTTT</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00395 Hb</td>
<td>R03023</td>
<td>68</td>
<td>R 10</td>
<td>CTTGAAAAAA</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00029 AP-1</td>
<td>R08494</td>
<td>867</td>
<td>R 9</td>
<td>ATGATTCAAT</td>
<td>18</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00497 MBP-1</td>
<td>R01060</td>
<td>840</td>
<td>R 9</td>
<td>GGGAATCCC</td>
<td>18</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00107 C/EbPa, alpha</td>
<td>R02583</td>
<td>823</td>
<td>R 9</td>
<td>ATTAGAAA</td>
<td>18</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00107 C/EbPa1</td>
<td>R02585</td>
<td>719</td>
<td>R 9</td>
<td>TGGTCAAT</td>
<td>18</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00479 Ly-1</td>
<td>R03504</td>
<td>647</td>
<td>N 9</td>
<td>TTCAGGAGG</td>
<td>18</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00273 Ev-1</td>
<td>M00011</td>
<td>155</td>
<td>N 9</td>
<td>ACAAGATATA</td>
<td>16.8</td>
<td>1.87</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00843 Ttk69</td>
<td>R02057</td>
<td>913</td>
<td>N 8</td>
<td>GCAGGTGCC</td>
<td>16</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00410 IGF-1</td>
<td>R00849</td>
<td>620</td>
<td>N 8</td>
<td>CAGGTGCC</td>
<td>16</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00596 NF-kappa E1</td>
<td>R00897</td>
<td>620</td>
<td>N 8</td>
<td>ATTAGAAA</td>
<td>16</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00410 IGF-1</td>
<td>R00849</td>
<td>542</td>
<td>N 8</td>
<td>TGTAAGA</td>
<td>16</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00821 TFI-D</td>
<td>R01167</td>
<td>281</td>
<td>R 8</td>
<td>ATTAATCA</td>
<td>18</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00333 GR</td>
<td>R01119</td>
<td>274</td>
<td>N 9</td>
<td>AGTGTCTC</td>
<td>16</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00106 C/EbPa, alpha</td>
<td>R03657</td>
<td>249</td>
<td>N 9</td>
<td>CCTCCCTC</td>
<td>16</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T01100 SEF1</td>
<td>R09807</td>
<td>439</td>
<td>N 9</td>
<td>ATATTTAWW</td>
<td>16</td>
<td>1.87</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00152 CP2</td>
<td>R00444</td>
<td>908</td>
<td>N 7</td>
<td>AGCCACT</td>
<td>14</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00270 EF-2</td>
<td>R01371</td>
<td>881</td>
<td>N 7</td>
<td>ATAAATAA</td>
<td>14</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00122 c-Fos T00131 c-Jun T001876</td>
<td>R01876</td>
<td>1866</td>
<td>N 7</td>
<td>TGATTCA</td>
<td>14</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00273 gammaC-ATAT</td>
<td>R00563</td>
<td>847</td>
<td>N 7</td>
<td>CTATGGG</td>
<td>14</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T02878 TCF-4E</td>
<td>R08643</td>
<td>753</td>
<td>R 7</td>
<td>AGCAAGAG</td>
<td>14</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T02878 TCF-4E T02905 LEF-1</td>
<td>R08642</td>
<td>1745</td>
<td>R 7</td>
<td>ATCAAGA</td>
<td>14</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T02857 TCF-3</td>
<td>R08594</td>
<td>745</td>
<td>N 7</td>
<td>ATCAAGA</td>
<td>14</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00123 c-Fos T00131 c-Jun T001195</td>
<td>R01195</td>
<td>505</td>
<td>R 7</td>
<td>TGACCCA</td>
<td>14</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00706 PuF</td>
<td>R02239</td>
<td>495</td>
<td>N 7</td>
<td>GGGTGCGG</td>
<td>14</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00706 CAC-binding protein</td>
<td>R04250</td>
<td>494</td>
<td>N 7</td>
<td>GGTTGGG</td>
<td>14</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00797 TBP</td>
<td>R00061</td>
<td>438</td>
<td>N 7</td>
<td>TATTTAA</td>
<td>14</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00706 CAC-binding protein T001251</td>
<td>359</td>
<td>R 7</td>
<td>GGGGCAG</td>
<td>14</td>
<td>2</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00617 NF-Zz T00618 NF-Zc</td>
<td>R01926</td>
<td>352</td>
<td>R 7</td>
<td>AGGGTCC</td>
<td>14</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00923 Zta</td>
<td>R00301</td>
<td>288</td>
<td>R 7</td>
<td>TTGCTCA</td>
<td>14</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00270 ETF</td>
<td>R01371</td>
<td>280</td>
<td>N 7</td>
<td>ATAAATAA</td>
<td>14</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00276 TCF-4E</td>
<td>R08643</td>
<td>268</td>
<td>N 7</td>
<td>TTGCTCA</td>
<td>14</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00923 Zta</td>
<td>R00301</td>
<td>266</td>
<td>R 7</td>
<td>TTGCTCA</td>
<td>14</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00794 TBP</td>
<td>R04100</td>
<td>49</td>
<td>N 7</td>
<td>TTGCTCA</td>
<td>14</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00706 PuF</td>
<td>R02239</td>
<td>36</td>
<td>N 7</td>
<td>GGTTGGG</td>
<td>14</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00706 CAC-binding protein</td>
<td>R04290</td>
<td>35</td>
<td>N 7</td>
<td>GGTTGGG</td>
<td>14</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00287 TCF-4C</td>
<td>R06643</td>
<td>21</td>
<td>N 7</td>
<td>TTGCTCA</td>
<td>14</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00212 E2BP</td>
<td>R02140</td>
<td>716</td>
<td>N 8</td>
<td>TGAAAYAY</td>
<td>14</td>
<td>1.75</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T00377 HOX-A5</td>
<td>R02638</td>
<td>632</td>
<td>R 10</td>
<td>RMTAAATNR14</td>
<td>14</td>
<td>1.4</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>_00000 T3r-beta</td>
<td>100031 (T3 771</td>
<td>9</td>
<td>GAGGTTGCAC 13.8</td>
<td>1.53</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00077 CACCC-binding_factor Q00019 (2)</td>
<td>495</td>
<td>R 8</td>
<td>GGGTGCGG</td>
<td>13.5</td>
<td>1.68</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00077 CACCC-binding_factor Q00019 (2)</td>
<td>495</td>
<td>R 8</td>
<td>GGGTGCGG</td>
<td>13.5</td>
<td>1.68</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 MN1b</td>
<td>100351 (M 659</td>
<td>8</td>
<td>TTTCTAC 13.1</td>
<td>1.63</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00479 EF-3 (2)</td>
<td>R02122</td>
<td>542</td>
<td>N 8</td>
<td>TGGTGGWW13</td>
<td>1.62</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00104 C/EbPa, alpha</td>
<td>R02132</td>
<td>541</td>
<td>N 8</td>
<td>TGGTGGWW13</td>
<td>1.62</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00051 ATF T00052 ATF-a T00126</td>
<td>142</td>
<td>R 8</td>
<td>TGGTGGG</td>
<td>13.5</td>
<td>1.68</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 Ftz.1</td>
<td>100242 (Ftz 181</td>
<td>18</td>
<td>GCAATTA 12.9</td>
<td>1.61</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 HNF-5</td>
<td>100100 (H 283</td>
<td>8</td>
<td>CAAATTA 12.5</td>
<td>1.56</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 NF-InsE3</td>
<td>100069 (NF 620</td>
<td>N 8</td>
<td>CAGGTGCC</td>
<td>12.3</td>
<td>1.54</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 E-box-factor I06102 (E 1620</td>
<td>200</td>
<td>R 8</td>
<td>CAGGTGCC</td>
<td>12.3</td>
<td>1.54</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00918 Zeste T02100 Zeste</td>
<td>R04948</td>
<td>927</td>
<td>N 6</td>
<td>TGGTGGWW13</td>
<td>1.62</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00321 GCN4</td>
<td>R00832</td>
<td>866</td>
<td>N 6</td>
<td>TGGTGGWW13</td>
<td>1.62</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T01035 Isl-1 T01956 Isl-1 T02195 R03916</td>
<td>861</td>
<td>R 6</td>
<td>CATTAG</td>
<td>12</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00539 NF-1</td>
<td>R00802</td>
<td>856 R 6</td>
<td>GGAAAG 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>--------</td>
<td>--------</td>
<td>-------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00539 NF-1</td>
<td>R00802</td>
<td>812 R 6</td>
<td>GGAAAG 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00033 AP-2alpha T00035 AP-2</td>
<td>R01054</td>
<td>786 R 6</td>
<td>TCCCCA 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00321 GCN4</td>
<td>R00832</td>
<td>747 R 6</td>
<td>GAATCA 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00333 GR</td>
<td>R00429</td>
<td>724 N 6</td>
<td>TGTCTT 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00033 AP-2alpha T00035 AP-2</td>
<td>R08502</td>
<td>699 N 8</td>
<td>TTTGCC 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00113 c-Ets-2</td>
<td>R04339</td>
<td>683 N 8</td>
<td>TTCCTT 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00539 NF-1</td>
<td>R00802</td>
<td>680 N 6</td>
<td>TTCTCC 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00204 E12 T01532 myogenin</td>
<td>R08512</td>
<td>666 R 6</td>
<td>CAGATG 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00140 c-Myc T00876 USF2 TX</td>
<td>R09477</td>
<td>1620 R 6</td>
<td>CAGGTTG 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00753 Sp1</td>
<td>R01540</td>
<td>612 R 6</td>
<td>CTGCC 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00333 GR</td>
<td>R03539</td>
<td>592 N 6</td>
<td>GG ACA 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00639 NF-1</td>
<td>R00802</td>
<td>539 R 6</td>
<td>GGAAAG 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00422 IRF-1 T00426 IRF-2</td>
<td>R00917</td>
<td>536 N 6</td>
<td>AAGTCA 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00113 c-Ets-2</td>
<td>R04343</td>
<td>524 N 6</td>
<td>GGAAAG 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00759 Sp1</td>
<td>R08207</td>
<td>494 N 6</td>
<td>GGTGGA 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00333 GR</td>
<td>R03539</td>
<td>457 R 6</td>
<td>TGCC 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00701 PTF1-beta</td>
<td>R01835</td>
<td>455 R 6</td>
<td>TCCCAT 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T01532 myogenin</td>
<td>R08512</td>
<td>452 N 6</td>
<td>TCCCAT 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00333 GR</td>
<td>R03535</td>
<td>448 N 6</td>
<td>TGCC 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00137 c-Myb</td>
<td>R04082</td>
<td>412 N 6</td>
<td>ATGCAA 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00333 GR</td>
<td>R01815</td>
<td>410 N 6</td>
<td>TGAACT 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T02657 TCF-3</td>
<td>R08594</td>
<td>403 R 6</td>
<td>ACAAAG 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00140 c-Myc</td>
<td>R02893</td>
<td>395 R 6</td>
<td>CAACTG 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00238 EII</td>
<td>R01342</td>
<td>389 R 6</td>
<td>TGCA 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00753 Sp1</td>
<td>R01540</td>
<td>358 N 6</td>
<td>GGCAG 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00033 AP-2alpha T00035 AP-2</td>
<td>R01054</td>
<td>333 R 6</td>
<td>TCCCCA 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00335 GR</td>
<td>R01313</td>
<td>303 N 6</td>
<td>TCTCTT 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00337 GR alpha T00696 PR T00973</td>
<td>272 R 6</td>
<td>TGGTCT 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T0042 AR T00335 GR T00697</td>
<td>R01101</td>
<td>1272 N 6</td>
<td>TGGTCT 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00539 NF-1</td>
<td>R00802</td>
<td>253 N 6</td>
<td>TCCCTT 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00918 Zeste T02100 Zeste</td>
<td>R04948</td>
<td>233 R 6</td>
<td>ATCTCA 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T01649 HES-1</td>
<td>R04468</td>
<td>156 R 6</td>
<td>CACAG 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00305 GATA-1 T01302 GATA-1</td>
<td>R03884</td>
<td>152 N 6</td>
<td>AGATAA 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00918 Zeste T02100 Zeste</td>
<td>R04947</td>
<td>133 N 6</td>
<td>TGAGTC 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00694 PPAR-alpha T00991 PP</td>
<td>R02233</td>
<td>124 N 6</td>
<td>AGSTCA 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T04014 SF-1</td>
<td>R08684</td>
<td>124 R 6</td>
<td>AGTCA 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00475 Lva</td>
<td>R01135</td>
<td>115 N 6</td>
<td>GAACAG 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T0033 AP-2alpha T00035 AP-2</td>
<td>R01054</td>
<td>104 N 6</td>
<td>TGGCGA 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00769 Sp1</td>
<td>R02860</td>
<td>95 N 6</td>
<td>AGCCGG 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00113 c-Ets-2</td>
<td>R04339</td>
<td>83 N 6</td>
<td>TCCCTT 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00475 Lva</td>
<td>R01135</td>
<td>16 R 6</td>
<td>CTGTTC 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00630 POU3F2 T01524 POJ3F R04327</td>
<td>882 N 7</td>
<td>MATWAAT 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00630 POU3F2 T01524 POJ3F R04327</td>
<td>281 N 7</td>
<td>MATWAAT 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00630 POU3F2 T01524 POJ3F R04327</td>
<td>179 N 7</td>
<td>MATWAAT 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00630 POU3F2 T01524 POJ3F R04327</td>
<td>419 N 9</td>
<td>MATNWAAT 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00913 Y1</td>
<td>R03154</td>
<td>249 N 10</td>
<td>CCCNCCNNCC 12 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00278 delta factor</td>
<td>R04142</td>
<td>821 R 11</td>
<td>WNNNANANAA 12 1.99 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00278 delta factor T00915 Y1</td>
<td>R04142</td>
<td>766 R 11</td>
<td>WNNNANANAA 12 1.99 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00278 delta factor T00915 Y1</td>
<td>R04142</td>
<td>481 R 11</td>
<td>WNNNANANAA 12 1.99 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00278 delta factor T00915 Y1</td>
<td>R04142</td>
<td>441 N 11</td>
<td>CWTNNTTTN 12 1.99 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 HOXD10</td>
<td>I00179</td>
<td>8 R 11</td>
<td>GTTTTAG 11.7 1.46 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 Malt</td>
<td>I00377</td>
<td>8 R 11</td>
<td>CCTCTCTCTT 11.6 1.16 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 RAR-gamma</td>
<td>IO0404</td>
<td>8 R 7</td>
<td>GAGTGA 11.1 1.58 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 OBF</td>
<td>I00210</td>
<td>8 R 7</td>
<td>ATGCAA 11.1 1.58 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00788 T-Ag</td>
<td>R01244 ( ) 445 N 5 GCCCTC</td>
<td>10 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00691 POU1F1a</td>
<td>R00609 ( ) 437 R 5 ATTTA</td>
<td>10 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00764 Sp1</td>
<td>R08166 ( ) 398 R 5 GCACA</td>
<td>10 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00764 Sp1</td>
<td>R08166 ( ) 391 N 5 TGTGC</td>
<td>10 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00762 Sp1</td>
<td>R01021 ( ) 390 R 5 GTGCA</td>
<td>10 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00765 SRF (504 AA)</td>
<td>R09635 ( ) 386 R 5 ATATA</td>
<td>10 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00915 YY1</td>
<td>R00603 ( ) 383 N 5 TATTT</td>
<td>10 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00691 POU1F1a</td>
<td>R00609 ( ) 382 R 5 ATTTA</td>
<td>10 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00788 T-Ag</td>
<td>R01372 ( ) 359 N 5 GGGGC</td>
<td>10 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00261 ER-alpha T00467 LF-A1</td>
<td>R01171 ( ) 1358 N 5 GGGCA</td>
<td>10 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00333 GR</td>
<td>R01813 ( ) 355 N 5 CAGAG</td>
<td>10 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00172 CTCF</td>
<td>R02137 ( ) 353 R 5 GAGGG</td>
<td>10 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00392 H4TF-2</td>
<td>R00681 ( ) 350 N 5 GTGCC</td>
<td>10 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00302 GAL4</td>
<td>R00495 ( ) 345 R 5 AGGCT</td>
<td>10 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00788 T-Ag</td>
<td>R01372 ( ) 316 N 5 TGGGC</td>
<td>10 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00915 YY1</td>
<td>R00603 ( ) 282 R 5 AAATA</td>
<td>10 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00691 POU1F1a</td>
<td>R00609 ( ) 279 N 5 TAAAT</td>
<td>10 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00172 CTCF</td>
<td>R02137 ( ) 249 N 5 CCCTC</td>
<td>10 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00302 GAL4</td>
<td>R00492 ( ) 246 R 5 TCCCT</td>
<td>10 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00752 Sp1</td>
<td>R01021 ( ) 220 R 5 GCGCA</td>
<td>10 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00104 C/EBPalpha T00105 C/E</td>
<td>R02132 ( ) 181 R 5 GCAAT</td>
<td>10 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00172 CTCF</td>
<td>R02137 ( ) 165 N 5 CCCTC</td>
<td>10 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00333 GR</td>
<td>R01813 ( ) 163 R 5 CTCTG</td>
<td>10 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00261 ER-alpha</td>
<td>R01194 ( ) 159 N 5 GTGCA</td>
<td>10 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00302 GAL4 T00794 TBP</td>
<td>R00494 ( ) 1151 N 5 GATAA</td>
<td>10 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T01274 ABF2</td>
<td>R03826 ( ) 151 R 5 GATAA</td>
<td>10 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00163 CREB</td>
<td>R02710 ( ) 142 N 5 TGACG</td>
<td>10 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00261 ER-alpha</td>
<td>R01194 ( ) 123 N 5 GGTCA</td>
<td>10 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00082 CBF (1) T00083 CBF (2)</td>
<td>R00231 ( ) 1106 N 6 ATTGG</td>
<td>10 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00081 alpha-CBF T00100 CUT1</td>
<td>R00339 ( ) 1105 N 5 ATTGG</td>
<td>10 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00074 gammaCAC1 T00075 ga</td>
<td>R00659 ( ) 136 R 5 GGGTG</td>
<td>10 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00891 POU1F1a</td>
<td>R00609 ( ) 6 R 5 ATTTA</td>
<td>10 2 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T04362 IPF1</td>
<td>R04228 ( ) 861 N 6 CATYAS</td>
<td>10 1.67 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00865 YY1</td>
<td>R02133 ( ) 816 N 6 ANATGG</td>
<td>10 1.67 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00865 YY1</td>
<td>R02133 ( ) 665 N 6 ANATGG</td>
<td>10 1.67 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00865 YY1</td>
<td>R02133 ( ) 453 R 6 CCATNT</td>
<td>10 1.67 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00865 YY1</td>
<td>R02133 ( ) 441 R 6 CCATNT</td>
<td>10 1.67 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00136 c-Myb T00137 c-Myb</td>
<td>R02206 ( ) 395 N 6 YAACKG</td>
<td>10 1.67 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00267 GATA-1 T00304 GATA-1R20156</td>
<td>1152 N 6 WAGATAR</td>
<td>10 1.67 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00320 GCF</td>
<td>R02159 ( ) 695 N 7 SGGSSCC</td>
<td>10 1.63 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00033 AP-2alpha T00034 AP-2</td>
<td>R02121 ( ) 496 R 8 SSSNKKGG</td>
<td>10 1.25 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00033 AP-2alpha T00034 AP-2</td>
<td>R02121 ( ) 37 R 8 SSSNKKGG</td>
<td>10 1.25 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00017 C/EBPbeta T00459 C/E</td>
<td>R02216 ( ) 859 N 9 TKNNGNAAK</td>
<td>10 1.11 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00017 C/EBPbeta T00459 C/E</td>
<td>R02216 ( ) 843 N 9 TKNNGNAAK</td>
<td>10 1.11 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00017 C/EBPbeta T00459 C/E</td>
<td>R02216 ( ) 823 R 9 MTTCNNMA</td>
<td>10 1.11 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00017 C/EBPbeta T00459 C/E</td>
<td>R02216 ( ) 719 N 9 TKNNGNAAK</td>
<td>10 1.11 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00017 C/EBPbeta T00459 C/E</td>
<td>R02216 ( ) 542 N 9 TKNNGNAAK</td>
<td>10 1.11 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00017 C/EBPbeta T00459 C/E</td>
<td>R02216 ( ) 518 N 9 TKNNGNAAK</td>
<td>10 1.11 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00017 C/EBPbeta T00459 C/E</td>
<td>R02216 ( ) 466 N 9 TKNNGNAAK</td>
<td>10 1.11 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00017 C/EBPbeta T00459 C/E</td>
<td>R02216 ( ) 240 N 9 TKNNGNAAK</td>
<td>10 1.11 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 TFII-I</td>
<td>I00216 (TF 530) R 6 AGAGAG</td>
<td>9.93 1.65 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 GATA-1</td>
<td>I00109 (GA152) N 6 AGATAA</td>
<td>9.85 1.64 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 MBF-1</td>
<td>I00083 (ME819) R 6 CCAAAA</td>
<td>9.85 1.64 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 MBF-1</td>
<td>I00083 (ME423) R 6 CCAAAA</td>
<td>9.85 1.64 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 GATA-3</td>
<td>I00108 (GA152) N 6 AGATAA</td>
<td>9.85 1.64 1 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Name</td>
<td>Start</td>
<td>End</td>
<td>Type</td>
<td>Description</td>
<td>Score 1</td>
<td>Score 2</td>
<td>Score 3</td>
<td>Score 4</td>
</tr>
<tr>
<td>-----------------</td>
<td>-------</td>
<td>------</td>
<td>------</td>
<td>-------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>_00000 CP1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 CP1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 CP1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 CP1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 PR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 MIG1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 MIG1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 NF1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 RFX2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 NF-X3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 NF-X3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 CBP/CRF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 CBP/CRF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 CBP/CRF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 CBP/CRF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 CBP/CRF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 ER</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 ER</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 ER</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00395 Hb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00395 Hb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00395 Hb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00395 Hb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00395 Hb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00395 Hb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00395 Hb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00395 Hb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00395 Hb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00395 Hb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00395 Hb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00395 Hb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00395 Hb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T00395 Hb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 T3R-alpha</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 T3R-alpha</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 Cad</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 Cad</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 Cad</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 Cad</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 Cad</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 Cad</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 Cad</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_00000 Cad</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sequence</td>
<td>M00028 (IS208)</td>
<td>L00266 (twi832)</td>
<td>N00266 (twi782)</td>
<td>I00266 (CA618)</td>
<td>C00266 (CA494)</td>
<td>I00266 (CA35)</td>
<td>I00266 (LB676)</td>
<td>I00191 (LB621)</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>----------------</td>
<td>-----------------</td>
<td>----------------</td>
<td>----------------</td>
<td>----------------</td>
<td>--------------</td>
<td>--------------</td>
<td>--------------</td>
</tr>
<tr>
<td>T00386 HSTF</td>
<td>M00028 (IS208)</td>
<td>R</td>
<td>5</td>
<td>TTTCT</td>
<td>7.82</td>
<td>1.56</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>_00000 twi</td>
<td>L00266 (twi832)</td>
<td>R</td>
<td>6</td>
<td>CACTTG</td>
<td>7.81</td>
<td>1.3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>_00000 twi</td>
<td>N00266 (twi782)</td>
<td>N</td>
<td>6</td>
<td>CAAAATG</td>
<td>7.81</td>
<td>1.3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>_00000 CAC-binding</td>
<td>I00266 (CA618)</td>
<td>R</td>
<td>5</td>
<td>GGTGG</td>
<td>7.66</td>
<td>1.53</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>_00000 CAC-binding</td>
<td>C00266 (CA494)</td>
<td>R</td>
<td>5</td>
<td>GGTGG</td>
<td>7.66</td>
<td>1.53</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>_00000 CAC-binding</td>
<td>I00266 (CA35)</td>
<td>R</td>
<td>5</td>
<td>GGTGG</td>
<td>7.66</td>
<td>1.53</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>_00000 LBP-1</td>
<td>I00191 (LB676)</td>
<td>N</td>
<td>5</td>
<td>CCTGG</td>
<td>7.52</td>
<td>1.5</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>_00000 LBP-1</td>
<td>I00191 (LB621)</td>
<td>R</td>
<td>5</td>
<td>CCAGG</td>
<td>7.52</td>
<td>1.5</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>_00000 LBP-1</td>
<td>I00191 (LB347)</td>
<td>R</td>
<td>5</td>
<td>CCAGG</td>
<td>7.52</td>
<td>1.5</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>T00385 HSF1</td>
<td>M00029 (F271)</td>
<td>R</td>
<td>5</td>
<td>GTTCT</td>
<td>7.38</td>
<td>1.48</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>T00788 T-Ag</td>
<td>Q00168 (-93)</td>
<td>N</td>
<td>5</td>
<td>GCGGC</td>
<td>0.5</td>
<td>1.3</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
BIBLIOGRAPHY

American Cancer Society. “What Are the Key Statistics About Prostate Cancer.” 
http://www.cancer.org/docroot/CRI/content/CRI_2_4_1X_What_are_the_key_statistics_for_prostate_cancer_36.asp?sitearea= (accessed March 27, 2009).


Darby, S, SS Cross, NJ Brown, FC Hamdy, and CN Robson. 2008. “BMP-6 over-expression in prostate cancer is associated with increased Id-1 protein and a more invasive phenotype.” *J Patho*: 394-404.


Jen, Y, K Manova, and R Benezra. 1996. “Expression Patterns of Id1, Id2, and Id3 are highly related but distinct from that of Id4 during mouse embryogenesis.” *Developmental Dynamics*: 235-252.


Yin, H, M Lowery, and J Glass. 2009. “In prostate cancer C/EBPalpha promotes cell growth by the loss of interactions with CDK2, CDK4, and E2F and by activation of AKT.” *Prostate*: 1001-16.

Yin, H, M Lowery, and J Glass. 2009. “In prostate cancer C/EBPalpha promotes cell growth by the loss of interactions with CDK2, CDK4, and E2F and by activation of AKT.” *Prostate*: 1001-16.


Yin, H, M Lowery, and J Glass. 2009. “In prostate cancer C/EBPalpha promotes cell growth by the loss of interactions with CDK2, CDK4, and E2F and by activation of AKT.” *Prostate*: 1001-16.

Yin, H, M Lowery, and J Glass. 2009. “In prostate cancer C/EBPalpha promotes cell growth by the loss of interactions with CDK2, CDK4, and E2F and by activation of AKT.” *Prostate*: 1001-16.

