T.C.
DOKUZ EYLÜL UNIVERSITY
IZMIR INTERNATIONAL BIOMEDICINE AND GENOME INSTITUTE

TRANSCRIPTIONAL REGULATION OF CCAAT ENHANCER-BINDING PROTEIN ALPHA (CEBPA)

ÇAĞRI BULDU

DEPARTMENT OF MOLECULAR BIOLOGY AND GENETICS

MASTER OF SCIENCE THESIS

İZMİR – 2019

THESIS CODE: DEU.HSI.MSc-2016850021

T.C.
DOKUZ EYLÜL UNIVERSITY
IZMIR INTERNATIONAL BIOMEDICINE AND GENOME INSTITUTE

TRANSCRIPTIONAL REGULATION OF CCAAT ENHANCER-BINDING PROTEIN ALPHA (CEBPA)

DEPARTMENT OF MOLECULAR BIOLOGY AND GENETICS

MASTER OF SCIENCE THESIS

ÇAĞRI BULDU

Thesis Supervisor:
Asst. Prof. Dr. HANI ALOTAIBI

The research of the thesis was supported and funded by TÜBİTAK projects, numbered **114Z245** and **117Z223** respectively.

THESIS CODE: DEU.HSI.MSc-2016850021

KOMİTE ONAYI

Dokuz Eylül Üniversitesi İzmir Uluslararası Biyotıp ve Genom Enstitüsü
Genom Bilimleri ve Moleküler Biyoteknoloji Anabilim Dalı,
Moleküler Biyoloji ve Genetik Yüksek Lisans programı öğrencisi Çağrı Buldu
'TRANSCRIPTIONAL REGULATION OF CCAAT ENHANCER-BINDING PROTEIN ALPHA (CEBPA)' konulu Yüksek Lisans tezini
29/ 07 / 2019 tarihinde başarılı olarak tamamlamıştır.

Dr. Öğr. Üye. Hani ALOTAİBİ
BAŞKAN

Doç.Dr. Şerif ŞENTÜRK
ÜYE

Dr. Öğr. Üye. Ayşe Banu DEMİR
ÜYE

Doç.Dr. Gökhan KARAKÜLAH
YEDEK ÜYE

Dr. Öğr. Üye. Cihangir YANDIM
YEDEK ÜYE

COMMITTEE APPROVAL

Dokuz Eylül University İzmir International Biomedicine and Genome Enstitute Department of Genomics and Molecular Biotechnology, Molecular Biology and Genetics graduate program Master of Science student Çağrı Buldu has successfully completed his Master of Science thesis titled **'TRANSCRIPTIONAL REGULATION OF CCAAT ENHANCER-BINDING PROTEIN ALPHA (CEBPA)'** on the date of 29 / 07 / 2019.

Assist. Prof. Dr. Hani ALOTAİBİ
CHAIR

Assoc. Prof. Dr. Şerif ŞENTÜRK
MEMBER

Assist. Prof. Dr. Ayşe Banu DEMİR
MEMBER

Assoc. Prof. Dr. Gökhan KARAKÜLAH
SUBSTITUTE MEMBER

Assist. Prof. Dr. Cihangir YANDIM
SUBSTITUTE MEMBER

TABLE OF CONTENTS: PAGE #

INDEX OF TABLES

INDEX OF FIGURES

LIST OF ABBREVIATIONS

EMT –	Epithelial-Mesenchymal Transition
MET –	Mesenchymal-Epithelial Transition
E-cad –	Epithelial Cadherin; E-cadherin
N-cad –	Neural Cadherin; N-cadherin
Cdh1 –	Cadherin-1 (E-cad gene name)
Cebpa –	CCAAT Binding Protein Alpha
TF(s) –	Transcription Factor(s)
TFBS(s) –	Transcription Factor Binding Site(s)
kb –	Kilobase
bp –	Basepair
CNS –	Conserved Sequence
TGFβ –	Transforming Growth Factor Beta
HGF –	Hepatocyte Growth Factor
EGF –	Epidermal Growth Factor
CDS –	Coding Sequence; Protein coding sequence
DBD –	DNA binding domain
bZIP –	Basic Leucine Zipper
kDa –	Kilodalton
GMP –	Granulocyte-monocyte progenitors
HCC –	Hepatocellular Carcinoma
UVB –	Ultraviolet B
RNA-seq –	RNA Sequencing
NMuMG –	Normal Mouse Mammary Gland
GEO –	Gene Expression Omnibus
NCBI –	National Center for Biotechnology Information
DMEM –	Dulbecco's Modified Eagle's Medium
FBS –	Fetal Bovine Serum
PBS –	Phosphate-Buffered Saline
NEAA –	Non-essential Aminoacids
BSL2 –	Biological Safety Level 2
siRNA –	Small interfering RNA
cDNA –	Complementary DNA
BME –	Beta-mercaptoethanol; 2-mercaptoethanol
rDNAse –	Recombinant DNAse
qPCR –	Quantitative PCR; Realtime PCR
GAPDH –	Glyceraldehyde-3-Phosphate Dehydrogenase
PT72 –	Post-treatment 72 hours
chr7 –	Chromosome 7
TFDB –	Transcription Factor Database
UCSC –	University of California Santa Cruz
LICR –	Ludwig Institute for Cancer Research
Caltech –	California Institute of Technology
Scr siRNA –	Scrambled siRNA
Up1-3 –	Upstream CNS regions 1 to 3
D1-5 –	Downstream CNS regions 1 to 5
UTR –	Untranslated Region
TSS –	Transcription Start Site
Pol II –	RNA Polymerase II
Jarid2 –	Jumonji And AT-Rich Interaction Domain Containing 2
PRC2 –	Polycomb repressive complex 2
AML –	Acute myeloid leukemia

ACKNOWLEDGEMENTS

First of all, for all her support and encouragement, I am deeply grateful to my mother. Without her, I wouldn't be the person here I am now in the world of Academia.

Secondly, my biggest gratitude is to my thesis advisor Hani Alotaibi, and to my lab mates. I greatly appreciate their guidance, help, and support throughout this process,

To Dokuz Eylül University, IBG center, and IBG institute for supporting our projects with grants and internal funds, and to TUBITAK for granting our projects and for the generous scholarships;

For their help and guidance, to other labs and PIs in IBG, especially, Şentürk Lab, Koçer Lab, Atabey Lab, Meuwissen Lab, Diril Lab, Erdal Lab, and to Öztürk Lab; and my friends in Bioinformatics department, Doğa Eskier, Ahmet Bursalı and Deniz Doğan;

To my friends in iBG who helped me in one way or another, either physically or mentally in these 3 years; especially to Canan Büşra Ersayar, Elif Gökçe Hayden, Ece Cesur, İrem Artun, Nergiz Uçmaz, and Arzu Akça;

And to my overseas friends who supported me and cheered for me in darkest times; Lauren Brown, and Christine Hoplite, I would like to show my greatest appreciation.

ABSTRACT

Comparative genomics is a well-known method with diverse applications. By comparing intergenic DNA sequences among different species, one can identify potential regulatory sequences such as enhancers, by assuming that conserved function would lead to conserved expression, thus conserved regulation.

EMT and MET are crucial biological processes where the cells can alter their expressed genes and consequently shift between epithelial and mesenchymal phenotypes. Since the E-cadherin expression -which is the identifier marker for epithelial cells- sits on the center of the questions that need to be answered, any factors that were known to play a role affecting the E-cad expression needs to be addressed in order to reveal the mechanisms of MET process. Cebpa was shown to be correlated with E-cadherin in MET in our previous project, where Cebpa has a high correlation with epithelial phenotype, and increased expression alongside with Cdh1 (E-cad) in MET, and decreased levels in EMT. In the view of such evidence, it was hypothesized that Cebpa might have an essential role in EMT-MET processes and its regulation and function can be crucial in MET.

The in-vitro experiments of Cebpa knockdown in MET also showed supporting results for the hypothesis. When Cebpa was silenced in mesenchymal state cells, they were not able to return to their epithelial state in MET, the process was blocked. Therefore the regulation of Cebpa was needed to be addressed more. In order to identify putative regulators of Cebpa, cis-acting elements and trans-acting factors were identified among conserved regions of Cebpa 23kb region and common TFBSs were observed. GEO datasets of publicly available gene expression profiles of the same tissue where Cebpa was known to be expressed in high levels were analyzed with bioinformatics tools and common TFs were identified among the datasets. The results of conserved regions and GEO datasets were compared and putative regulator(s) of Cebpa was selected.

According to the findings of this study, Jarid2 is a possible candidate of Cebpa regulation since it was one of the conserved TFs both in location and in sequence information. The two factors are both have a role in differentiation, transcriptional repression and early development, and mutation in Prc2 which is highly interacting with Jarid2 also causes mutation in Cebpa that leads a way to Acute Myeloid Leukemia (AML).

Keywords: Transcriptional regulation, EMT, MET, Cebpa, Jarid2

ÖZET

Karşılaştırmalı genomik, çeşitli uygulama yöntemlerini barındıran iyi bilinen bir metottur. Farklı türler arasındaki intergenik DNA sekans dizilerini karşılaştırarak, korunmuş fonksiyonun korunmuş gen ifadesine götüreceği ve bunun da korunmuş regülasyonu göstereceği varsayılarak, bir genin enhancer ve benzeri potansiyel regüle edici sekanslarını tanımlayabilmek mümkündür.

EMT ve MET, hücrelerin ifade edilen genlerini değiştirmelerinde ve düzenli olarak epitelyal ve mezenkimal fenotipler arasında geçiş yapabilmeleri için önemli biyolojik olaylardır. Epitelyal hücrelerin tanınmasında rol oynayan bir markör olan E-kaderin gen ifadesi, cevaplanması istenilen soruların merkezinde durmaktadır. E-kaderin ifadesine etki ettiği bilinen her türlü faktör MET'nin mekanizmalarını aydınlatmada önemlidir. Bir önceki projemizde, Cebpa ifadesinin MET prosesinde E-kaderin ile korelasyon içinde olduğu görülmüştür. Cebpa epitelyal fenotip ile yüksek korelasyon göstermiş ve MET'de E-kaderin ifadesi yükselirken, Cebpa gen ifade düzeyi de artımıştır, EMT'de ise düşmüştür. Bu bilgilerin ışığında, Cebpa geninin EMT-MET olaylarında önemli bir role sahip olabileceği ve MET'de Cebpa'nın regülasyon ve fonksiyonunun kritik öneme sahip olabileceği hipotezi kurulmuştur.

Cebpa'nın susturulduğu in-vitro deneyler kurulan hipoteze destek veren sonuçlar göstermiştir. Mezenkimal durumdaki hücrelerde Cebpa susturulduğunda, MET prosesinin bloke olduğu ve hücrelerin epitelyal duruma geri dönemedikleri görülmüştür. Bu sebeple Cebpa'nın regülasyonunun ortaya çıkarılması ihtiyacı doğmuştur. Cebpa'nın potansiyel regülatörlerini tanımlayabilmek için, Cebpa'nın 23kb uzunluğundaki gen bölgesindeki korunmuş sekans dizilerinde cis-acting elementler ve trans-acting faktörler analiz edilmiş ve ortak transkripsiyon faktörleri tanımlanmıştır. Gen ifade profillerinin halka açık olarak paylaşıldığı GEO veri tabanından Cebpa'nın yüksek düzeyde ifade edildiği aynı dokuya ait veri setleri biyoinformatik yöntemler ile analiz edilmiş ve ortak transkripsiyon faktörleri tanımlanmıştır. Korunmuş bölgelerden elde edilen sonuçlar ile veri setleri karşılaştırılmış ve Cebpa'nın olası regülatörleri seçilmiştir.

Araştırmanın bulguları doğrultusunda, hem lokasyon hem de sekans dizisi olarak korunmuş transkripsiyon faktörlerinden biri olan Jarid2, Cebpa'nın potansiyel bir regülatörü olarak tanımlanmıştır. Bu iki faktör de, farklılaşma, transkripsiyonel baskılama ve erken gelişimde rol oynamaktadır. Jarid2 ile yüksek etkileşimi olan Prc2 genindeki mutasyon ise, Cebpa'nın mutasyonuna yol açarak Akut Miyeloid Lösemi'ye (AML) sebep olmaktadır.

Anahtar sözcükler: Transkripsiyonel regülasyon, EMT, MET, Cebpa, Jarid2

1. INTRODUCTION AND AIM

1.1 Aim of the Study

Main aim: To study the potential role of Cebpa in MET and identify the putative regulators of Cebpa.

Specific aim: To find cis-acting elements and trans-acting factors regulating Cebpa.

1.2 Hypothesis of the Study

The expression behavior of Cebpa gene changes under different conditions such as EMT and MET. Cebpa has high correlation with epithelial phenotype especially in MET. Because of this, the regulation and function Cebpa could be critical for MET.

2. GENERAL INFORMATION

2.1 Epithelial-Mesenchymal Transition (EMT)

When it comes to studying the more complex multicellular organisms, the ability of differentiation and capacity of changing cellular morphology are crucial to understand how an organism is capable of organizing these specialized functions coming from a single cell from an embryo to a whole functioning living being (Frankfurt, 1996; Hindley and Philpott, 2012). Functions of the organisms are dependent on the factors and interactions between the product of the genes, environmental factors and ontogenetic processes (Raff, 1996; West-Eberhard, 2003). The cause and effect of these processes are the main questions and in the focus of epigenetic studies (Hallgrimsson and Hall, 2011). The research on this area is generally done to understand better how the process of phenotype development occurs and how evolutionary modifications are at play to change and modify the function that has a role in these systems of the organisms. The primary role in this mechanism is the development of plasticity, which corresponds to the ability of the organism to modify its phenotypic features in a regulatable way to adapt to the changes in the environment and develop a system to answer to those changes accordingly. This phenotypical plasticity allows the cells to have the ability to transition between different morphologies. At this point, it is immensely crucial for multicellular organisms to ensure the occurrence of plasticity and secure the fate of phenotype of the cells.

Since it was classified in the 19th century, there are two primary phenotypic cellular morphologies which are called epithelial and mesenchymal cellular morphology (Duval M., 1889) which both are essential during early development (Rossant and Tam, 2009), but not permanent throughout the lifetime of the organism once generated (Maccarty and Caylor, 1922; Slack and Tosh, 2001). Epithelial-to-Mesenchymal Transition (EMT) and Mesenchymal-to-Epithelial Transition (MET) make up the whole backbone of these dynamic processes (Somarelli, Shetler et al., 2016).

The possibility of the cells that can lose their epithelial state and gain more of a mesenchymal phenotype was first observed by Elizabeth Hay in 1980s when she was studying with chicken embryos. When it was first described, EMT was defined to be the process of differentiation of normal cells during early development of the organism, where epithelial cells are in polarized nature and highly bound to the matrix and with neighboring cells, and mesenchymal cells with no distinct apicobasal polarity (Ozdamar et al., 2005; Aigner et al., 2007; Moreno-Bueno, Portillo and Cano, 2008) and without the bonds with the matrix (Le Bras, Taubenslag and Andl, 2012) and have extended mobility, and it was called “Epithelial-Mesenchymal Transformation” (Hay, 1995). Later, the term was renamed to “transition” to emphasize the reversibility of the process (Kalluri and Nielson, 2003).

In EMT, the cells gain an advanced migratory capacity and invasiveness, resistance of apoptosis (Kalluri and Nielson, 2003), and an advantageous ability to cross the extracellular matrix and blood-tissue barrier and join the bloodstream (Thiery, 2009) which plays a crucial role firstly in embryogenesis (Lamouille, Xu et al. 2014), wound healing (Kim et al., 2017) and angiogenesis (Ribatti D., 2017). To make it more organized, EMT can be classified in three types (Figure 1.1), Type 1 EMT occurs in embryonic development, Type 2 EMT takes place in tissue repair, inflammation and fibrosis, and Type 3 EMT which is involved in cancer progression (Kalluri and Weinberg, 2009; Zeisberg M., 2009). Micalizzi et al. also describes that EMT is the first and the most critical step of metastasis where the cells can migrate from a tumor and relocate at a different site in the organism (Micalizzi et al., 2010).

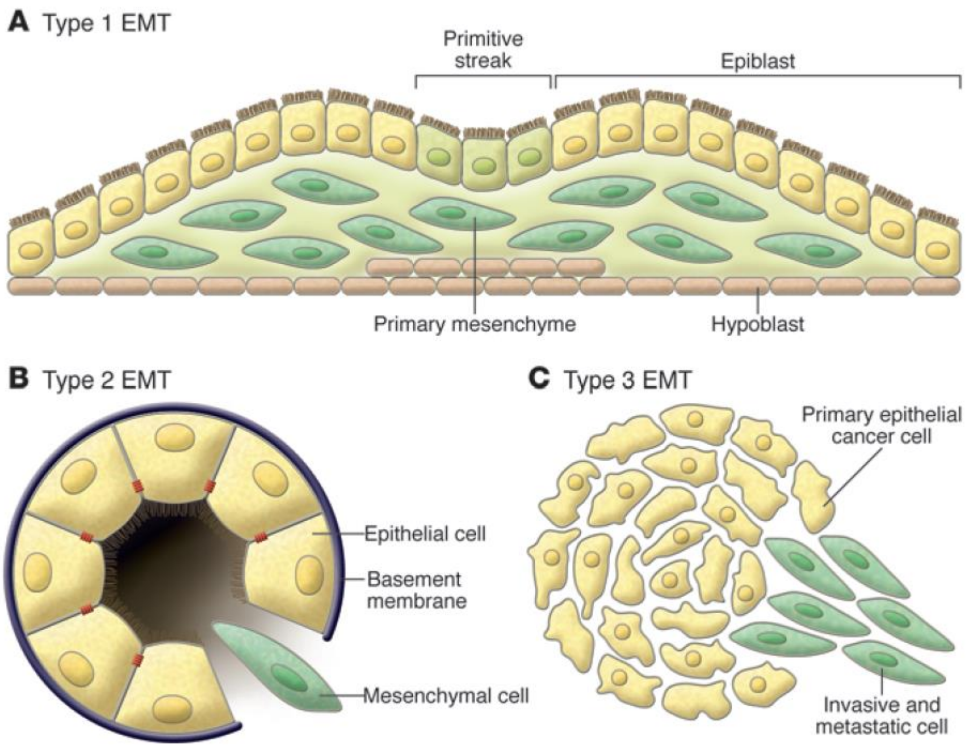


Figure 1.1 Three types of EMT. A) Type 1 EMT gives rise to mesoderm and endoderm and to mobile neural crest cells. Specifically epiblast gives rise to primary mesenchyme via EMT. The primary mesenchyme can be re-induced to give rise to secondary epithelia via MET. B) Type 2 EMT is expressed over extended periods of time and in the context of inflammation can destroy an affected organ if inflammation is not attended or removed. C) Type 3 EMT can transform into cancer that later can lead to invasion and metastasis (Kalluri and Weinberg, 2009).

Numerous molecular processes were defined to initiate EMT, which include activation of transcription factors, reorganization of cytoskeleton and expression of cell-surface proteins, production of extracellular matrix-degrading enzymes, and changes in the expression of specific miRNAs (Kalluri and Weinberg, 2009). All these different factors that are involved in the process are also used as biomarkers to identify this transition (Figure 1.2)

Transforming Growth Factor Beta (TGFβ) is the most known EMT inducer and acts to induce other EMT-related transcription factors (Illman et al., 2006; Xu, Lamouille and Derynck, 2009). EMT can also be induced by hepatocyte growth factor (HGF), epidermal growth factor (EGF) and fibroblast growth factor (FGF) (Grünert, 2003; Hay, 1995; Huber, 2005). When EMT is induced with abnormal signaling of TGFβ it can make way to harmful circumstances such as tumorigenesis and fibrosis (Tan, Olsson and Moustakas, 2015; Gilbert, Vickaryous and Viloria-Petit, 2016).

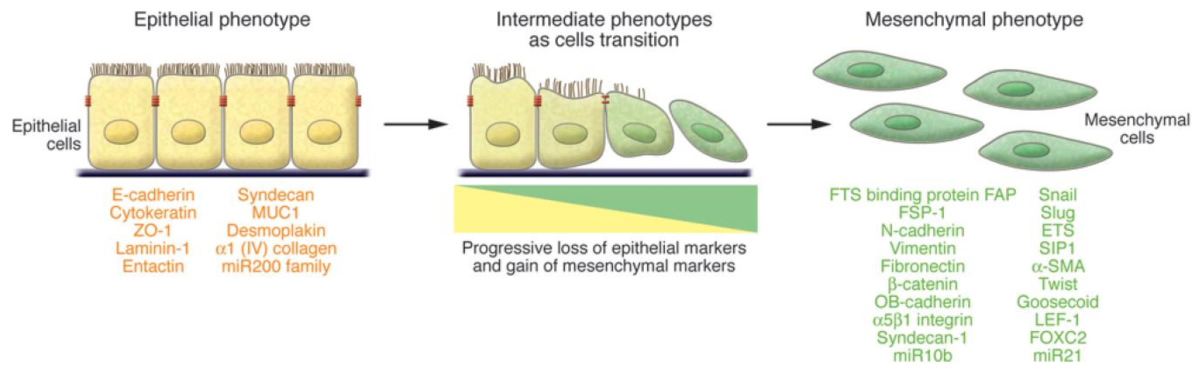


Figure 1.2 EMT process. Epithelial and mesenchymal markers commonly used were shown. In epithelial phenotype E-cadherin, cytokeratin, Laminin etc. markers were present. As cells go through EMT progressive loss of epithelial markers and gain of mesenchymal markers occur. In mesenchymal phenotype N-cadherin, vimentin, fibronectin, B-catenin etc. markers were present (Kalluri and Weinberg, 2009).

The changes affecting this transition were found to be cellular modifications and differentiation of gene expressions with other underlying events of epigenetic reprogramming (Reik, Dean and Walter, 2001). This transition then found out to be coordinated by different groups of transcription factors such as Snai1/Snail (Thuault et al., 2008), Snai2/Slug (Lin and Baritaki, 2010), Twist1/2 (Lo. H.W., 2007), and Zeb1/2 (Jung et al., 2008; Lamouille, Xu and Derynck, 2014). E-cad repressors can also be grouped by their activity on the E-cad promoter whether they affect the transcription directly or indirectly. Snail, Zeb, and Klf8 factors bind and repress the promoter activity (Peinado et al., 2007; Wand et al., 2007), whereas the factors such as Twist and Foxc2 repress the transcription indirectly (Yang and Weinberg, 2008; Sobrado et al., 2009). This induced differentiation leads the way for the organization of specialized tissues and organ systems in various organisms. The characteristics of the transition is observed by the increased expression of the mesenchymal markers such as N-cadherin (N-cad), Fibronectin and Vimentin with gaining a more invasive fibroblastoid phenotype (Wells and Yates, 2008; Micalizzi and Farabaugh, 2010; Garside and Chang, 2012) and by downregulation of Epithelial cadherin (E-cadherin, E-cad) in plasma membrane which is a dominant suppressor of tumor invasion and metastasis (Peinado, Portillo and Cano, 2004; Nakamura and Tokura, 2011). A switch from E-cad to N-cad promotes disassembly of adherens junctions (Scarpa and Szabo, 2015) and this allows the release of associated proteins such as β -catenin, which can upregulate EMT genes (Samuel and Lopez, 2011).

Since EMT is such a crucial process to assess, maintain and regulate the livelihood of the organism from the first stages of development to the last of cell fate in apoptosis, it is not surprising to see that the pathways which have a role in this process such as Snail / Slug, Twist, Six1 (McCoy et al., 2009), TGFβ, NFkB (Julien et al., 2007; Strippoli et al., 2008) and Wnt/β-catenin (Thiery and Sleeman, 2006) are highly conserved among different species (Micalizzi et al., 2010).

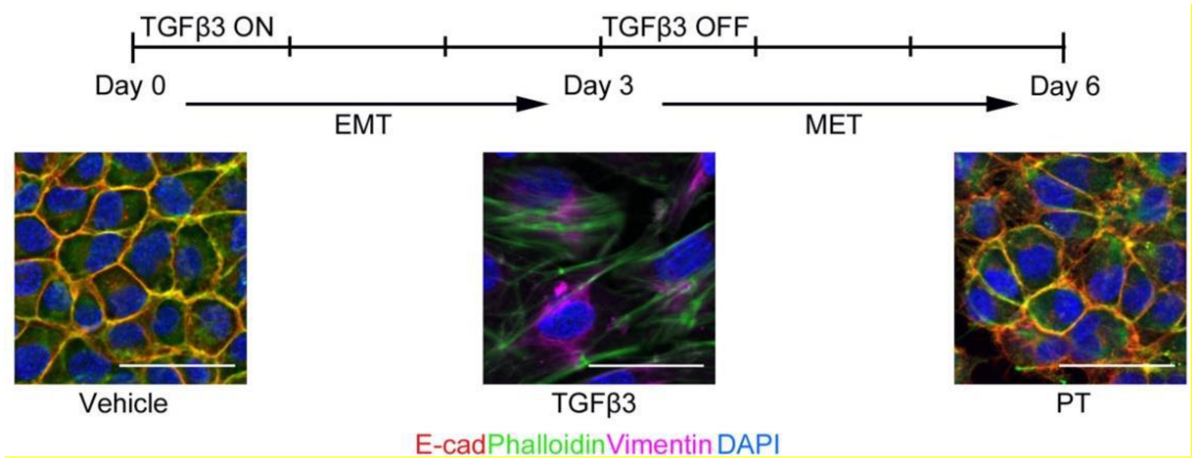


Figure 1.3 EMT-MET processes visualized via immunofluorescent staining. Day 0 Vehicle represents epithelial NMuMG cells and epithelial marker E-cad (red) is seen in the membrane along with actin staining phalloidin (green). After treatment with TGFβ for 72 hours in Day 3 (TGFβ3), cells gain mesenchymal phenotype and mesenchymal marker vimentin can be seen around the nuclei in purple, and epithelial marker E-cad cannot be seen. When TGFβ was removed from the medium, the cells grow in normal medium for 72 hours, in PT condition which represents Post-treatment, cells were shown to turn back to epithelial state and epithelial marker E-cad (red) can be seen in the membrane just like in Vehicle. (Alotaibi lab project #114Z245, 2017)

2.2 Mesenchymal-Epithelial Transition (MET)

Mesenchymal-to-Epithelial Transition (MET) is first thought to be the opposite of EMT, where unpolarized motile cells were phenotypically going through a progressive change to gain back their tightly-bound polarized morphology. Just like in EMT, the process of MET is also crucial in developmental, regenerative and pathological conditions in multicellular organisms (Antonello et al., 2015). The evolution of a complex organism from a single cell is highly associated with the process of MET (Valentine, 1978), since reorganization of cells and tissues from loose structures to specialized compartments in the organism can only be possible if the system could overturn and rearrange itself again to build more detailed and somewhat functional structures. The two main layers of a multicellular organism, epithelial ectoderm and mesenchymal mesoderm layers (Arnold and Robertson, 2009), were all we need to look at to understand how the machinery works and decide on the fate of the cells and revolve around the formations of more complex structures.

Firstly, the EMT process gives rise to all mesenchymal cells in early embryogenesis. However, just like it was thought to be before, MET is not the exact opposite of EMT when it comes to gene regulation but rather a complementary one. Even though the phenotypical properties of the cells were back into their epithelial state, the machinery behind the process is quite different and yet to be determined wholly. The relocalization of E-cad to the membrane, the reformation of cell-matrix adhesions, and reorganization of the cytoskeleton along with apicobasal polar morphology, and with the downregulation of E-cad and upregulation of mesenchymal factors are not enough to explain how MET works (Kim, Jackson and Davidson, 2017). From start to finish, both EMT and MET are controlled by many transcriptional, post-transcriptional and post-translational regulators on each step (Lamouille et al., 2013).

MET can be divided into four cellular processes where 1) triggering developmental and micro-environmental cues, 2) establishment of cell polarity, 3) spread of MET through tissue, and 4) stabilization of new tissue architecture. The external forces and matrix stiffness along with intercellular tension affects the rate of MET in a tissue-wide manner. With junction and adhesion stabilization, the return of the cells to the epithelial state are secured (Kim, Jackson and Davidson, 2017).

MET process has recently gained the deserved attention (Celia-Terrassa et al., 2012; Kalluri et al., 2009; Kumarswamy et al., 2012). The mesenchyme has the properties of a relaxed organization and characterized by motile flexibility and individualism. MET has been observed to occur in development and neoplastic transformation. The studies about MET are mostly around and about stem cell reprogramming (Takahashi et al., 2006; Downing et al., 2013; Choi et al., 2015) and secondary tumor formation (Chao et al., 2012; Atsuta et al., 2015; Shibue et al., 2011). However, there are also present studies of MET on regeneration, organogenesis (Radice et al., 1997; Trepat et al., 2009; Vasilyev et al., 2012) and somitogenesis (Latimer et al., 2010; Davidson et al., 2004; Nakaya et al., 2004). Lastly, recent studies that research the regulation of MET by miRNA have been on the rise (Liu et al., 2011, Renner et al., 2012). Even though it may seem like that MET is the reverse-EMT, it is not the same in terms of mechanisms. A number of studies on cancer showed what could be called reverse-EMTs, about hormones (Planas-Silva et al., 2007) and growth factors (Gal et al., 2008), however when it comes to development of progenitor cell populations, the transitions do not appear to be reverse of each other and it depends more on the change in microenvironment and the progression toward terminal differentiation. Still the studies are not enough to assess all the changes and differences in the mechanisms driving MET, and more research is needed to uncover specific relationships.

2.3 Gene Regulation, Expression and Transcription Factors

The genome of the cell is where all the information of compulsorily necessary data is stored for the healthy function of the living organism and where we can reach, observe and analyze the occurrences and reactions of the complete construction from a single cell to the whole organism. The genome in eukaryotes consists of the DNA that carries the necessary information to synthesize RNAs and many different protein molecules. The genes are essential elements of the genome where the protein-coding sequences (CDS) of the chromosomes are located inside the DNA strand. Those CDS regions are responsible for the production of their proteins, and those proteins have many different roles in regulating other genes that have functions in many pathways to maintain the healthy state of the organism.

Not all the genes of the cell are expressed at the same time nor in the same type of cells. The different sets of genes that are expressed during the different stages of the cell cycle or in the different type of cells are what makes the complexity of the organism when it’s about understanding the broader knowledge of how a gene (a piece of DNA) functions and takes up a role in order to achieve a specific outcome. If we consider the importance of non-coding regions of the DNA that are in the focus of the scientific community more in the last decade, the picture gets a more broad perspective in which we have to have a dire need to understand better.

As for the gene expression in multicellular eukaryotes, first we need to understand what transcription means. The central dogma of the cell cycle tells us that for a protein to be produced, first it needs to undergo of some changes from DNA to protein where the mechanism involves the reading of the information from the piece of DNA and transcribe it to RNA and translate and turn it into a functional protein. In this process, the coding (exons) and non-coding (introns) parts of the DNA are separated from each other during the transcription of mRNA. What we call gene expression is the process of this information in DNA that is spliced and constructed to include only the exons of the sequence which are then transcribed into mRNA and can be measured by various experimental methods to assess the level of expression of a gene inside the cell.

The regulation of the genes can be dependent on various factors, such as the size and the structure of the genome, the type of the cell, whether it has nuclear transcription and/or cytoplasmic translation, presence or absence of histone modifications, or the placement of promoters and enhancers of the gene, and post-translational modifications of proteins, to name a few. (Phillips, T., 2008) The diversity of these elements all play a role in the regulation of a gene and the way we observe it with the genes that are turned on or off with the intention of specific functions in a cell create a variety and all is possible by differential gene regulation. (Pierce, B., 2005)

The transcriptional regulation is simply, the combination of constitutive properties of DNA and the interactions of proteins which we call Transcription Factors (TFs). TFs are proteins that bind to specific sequences on the DNA to activate and sometimes to repress the rate of gene transcription. They use different mechanisms to regulate gene expression such as, affecting the binding of the RNA polymerase II on the promoter, recruit co-activators or co-repressors to the core promoter (Xu L, 1999), acetylation and deacetylation of histone proteins (Gill G, 2001, Lee, T.I., 2000). They have consisted of different domains, which are DNA-binding domain (DBD), transcriptional activation domain (Trans-activating, TAD) (Latchman DS, 1997), and response elements (Wang JC, 2005). The DNA-binding domains are which have a higher affinity for the specific sequences on the DNA. These sequences are highly conserved among species and were used to categorize the known transcription factors into families such as SOX proteins, and POU factors (Remenyi, A., 2004).

Most of the times the family members of TFs interact with each other or with different family members, or act together as in groups (Remenyi, A., 2004) to influence a coordinated action on the fate of the cells in such as cell division, differentiation, embryonic development, cell migration and in apoptosis as a result of regulation of gene activity in respect to specific stimuli. On this matter, the individual members of the TF family can adapt to certain types of conformation and function depending on their specific interaction partners and the sequence of the DNA binding site (Lefstin, J.A., 1998).

The TFs can also be grouped by their 3D protein structures of DNA-binding domains such as, helix-turn-helix, helix-loop-helix, basic leucine-zipper, homeodomain proteins and, zinc finger proteins. These structural motifs cause TFs to have specificity for their binding sequences. (Littlewood TD., 1995, Vinson C., 2002, Wintjens R., 1996, Laity JH., 2001, Gehring WJ., 1994, Bürglin TR., 2016).

2.4 CCAAT Enhancer-binding Protein-alpha (Cebpa)

The cellular functions of the organism are controlled and regulated by mechanisms which accompany the change in gene expression, differentiation, and proliferation, that were all maintained on the transcriptional level. Those target genes have regulatory sequences which are called cis-acting elements that are present on their promoter and enhancer regions, where the trans-acting transcription factors interact with the sequence and control these specialized cellular functions (Ramji and Foka, 2002).

One family of transcription factors called CCAAT enhancer-binding protein (CEBP) were identified as master regulators in many various cellular processes in the last three decades, where diseases and functional problems occur if they are defective. The first time one of the transcription factors in this family were identified in the late 1980s (Landschulz et al., 1988), it led to the discovery of basic-leucine zipper (bZIP) class of DNA-binding and dimerization domain (Landschulz et al., 1988; Landschulz et al., 1989; Vinson et al., 1989; Agre et al., 1989).

This first transcription factor that was identified was named CEBPA as the first of its family. In the 90s, five more of this family members were identified and found to be also contain a conserved bZIP domain at the C-terminus (Akira et al., 1990; Poli et al., 1990; Descombes et al., 1990; Chang et al., 1990; Roman et al., 1990; Cao et al., 1991; Williams et al., 1991; Ron and Habener, 1992; Tsukada et al., 2011). The function of this CEBP family now has been studied in the last three decades and many essential roles have been found to take place in cellular responses, control of cellular growth, in immune and inflammatory events and cellular differentiation in numerous diseases, and their expression levels change significantly in different pathophysiological conditions with the aid of extracellular triggers (Ko et al., 2015; Okuma et al., 2015; Choi et al., 2016; Wang et al., 2018).

The one we are concerned about this six-membered CEBP family, Cebpa is an intronless gene, longest of its family with three transactivation domains, one regulatory domain and one bZIP domain (Osada et al., 1996; Pei and Shih, 1991). The protein has two isoforms which are 42kDa and 30kDa in length, for short, p42 and p30, respectively (Calkhoven et al., 2000) (Figure 2.1).

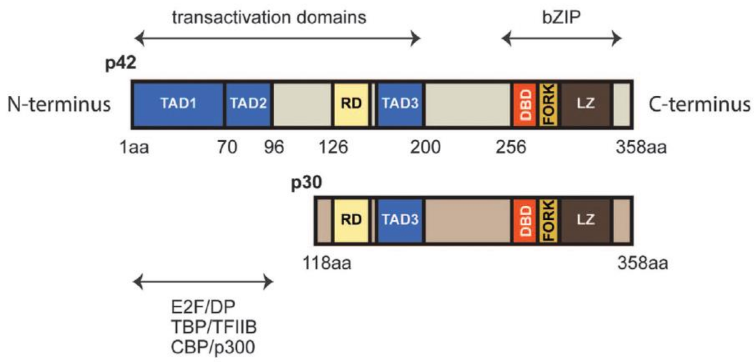


Figure 2.1 Protein isoforms of Cebpa with 42 kDa and 30 kDa in length, short for p42 and p30. Both isoforms have basic leucine zipper (bZIP) domain at C-terminus, and transactivating domains in N-terminus (Calkhoven et al., 2000).

The gene is expressed mostly at high levels in skin, liver, adipose tissue, intestines, lung, peripheral blood, adrenal gland and placenta (Cao et al., 1991; Williams et al., 1991; Antonson and Xanthopoulos, 1995; Lekstrom and Xanthopoulos, 1998) and the most mRNA levels were detected in liver and adipose tissue in terminally differentiated cells (Cao et al., 1991; Williams et al., 1991; Lekstrom and Xanthopoulos, 1998). (Figure 2.2)

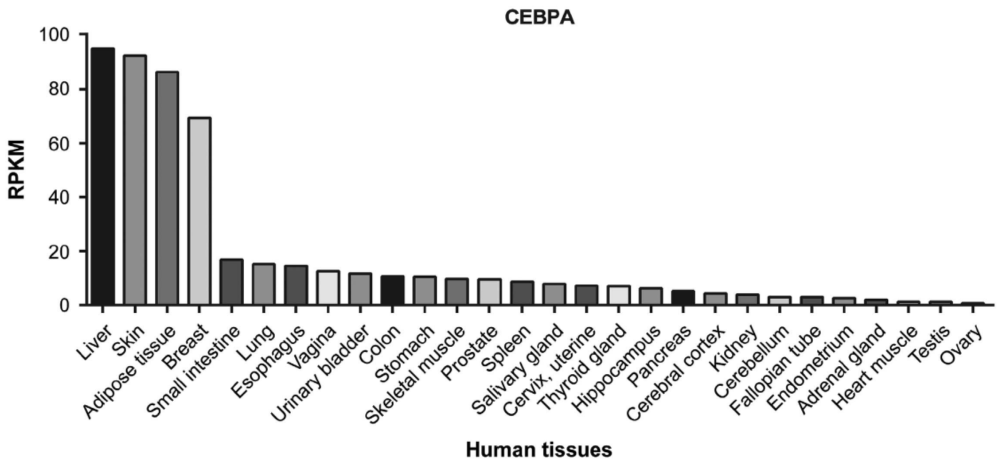


Figure 2.2 CEBPA mRNA levels in 28 different human tissues. Most expressed tissues are liver, skin, adipose tissue and breast tissue. (Lourenço and Coffer, 2017. Adapted from www.proteinatlas.org)

The function of this gene has been investigated by numerous methods using analysis of promoter regions, overexpression or inhibition and using knock-out mouse models. The roles of the gene and its TF mostly take place in differentiation, inflammatory response, liver regeneration and metabolism. When the target genes were investigated for the CEBP family, CEBPA was found to have a direct role on differentiation of hepatocytes (Sato et al., 2006; Takayama et al., 2014), differentiation of adipocytes (Hu et al., 2015; Kim and Nam, 2017), and differentiation of hematopoietic cells (Ye et al., 2013; Bararia et al., 2016; Zjablovskaja et al., 2017) and as well as in lung and myeloid cells (Poli et al., 1990; Descombes et al., 1990; Chang et al., 1990; Nerlov and Ziff, 1994; Wedel and Lömsziegler-Heitbrock, 1995; Lee et al., 1997; Wang et al., 1995; Roesler, 2001; Kimura et al., 1998; Darlington et al., 1998; Rosen et al., 2000; Lane et al., 1999; Mason et al., 1998; Hwang et al., 1996; Elberg et al., 2000; Poli, V., 1998; Bretz et al., 1994; Zhang et al., 1998; Zhang et al., 1997; Flodby et al., 1996; Cassel et al., 2000; Sugahara et al., 2001; Chew et al., 2007; Kanamori et al., 2017; Guo et al., 2015; Li et al., 2017; Khanna-Gupta et al., 2003; Avellino and Delwel, 2017). Also during hematopoiesis, Cebpa was recently found to be expressed by hematopoietic stem cells, granulocyte-monocyte progenitors (GMPs), neutrophils and macrophages (Wu et al., 2016 (http://biogps.org/#goto=genereport&id=12606); Lattin et al., 2008) and found to be enough to specify macrophages since B-cells can be transdifferentiated into them by expressing Cebpa ectopically (Xie et al., 2004). Also its effects were demonstrated to be a tumor suppressor by repressing cell proliferation and promoting terminal differentiation in some tissues such as in lung, skin cancer where it was continuously silenced by promoter hypermethylation or by mutation, and mutation of this gene was found to be causing Acute Myeloid Leukemia (AML) (Sato et al., 2013; Lourenço and Coffer, 2017; Kirstetter et al., 2008; Tada et al., 2006; Loomis et al., 2007). The oncogenic role of CEBPA was also studied in Hepatocellular carcinoma (HCC), hepatoblastoma, prostate cancer and B-cell precursor acute lymphoblastic leukemia (Hu et al., 2019; Yin et al., 2009; Chapiro et al., 2006; Lu et al., 2010; Lu et al., 2015; Cast et al., 2017). In Yoon and Smart's study with keratinocytes involving ultraviolet B (UVB) assisted DNA damage, CEBPA was found to have antiproliferative function, which the radiation causes DNA damage and activates p53-dependent G1 checkpoint in cell cycle and arrests the progression during repair, where CEBPA expression was strongly induced by UVB irradiation in cultured keratinocytes and mouse skin. CEBPA knockdown reduces UVB-induced cell cycle arrest, and it was shown that induction of CEBPA in p53-mutant cells was impaired (Yoon and Smart, 2004). So, in addition to regulating terminal differentiation, during DNA repair CEBPA can be induced by stress signals that block cell proliferation and this could add to its role in tumor suppression (Johnson, 2005).

2.4.1 Preliminary Data on Cebpa in MET

According to our preliminary data from the TUBITAK project #114Z245 named "Identification of Transcriptional Network Controlling MET", Cebpa was found to be significantly crucial in MET. In this project, nine datasets that specifically have epithelial and mesenchymal samples from GEO database were analyzed and a list of epithelial-specific TFs was obtained. The RNA-seq data of MET in NMuMG cells also showed the TFs that were active in MET.

Our RNA-seq data showed the correlation of Cebpa with Cdh1 during a time course with 11 time points, with R-score of 0.9485 and p-value of 0,8621.10-5 (Figure 2.3) (Eskier, Karakulah, and Alotaibi, submitted, unpublished data).

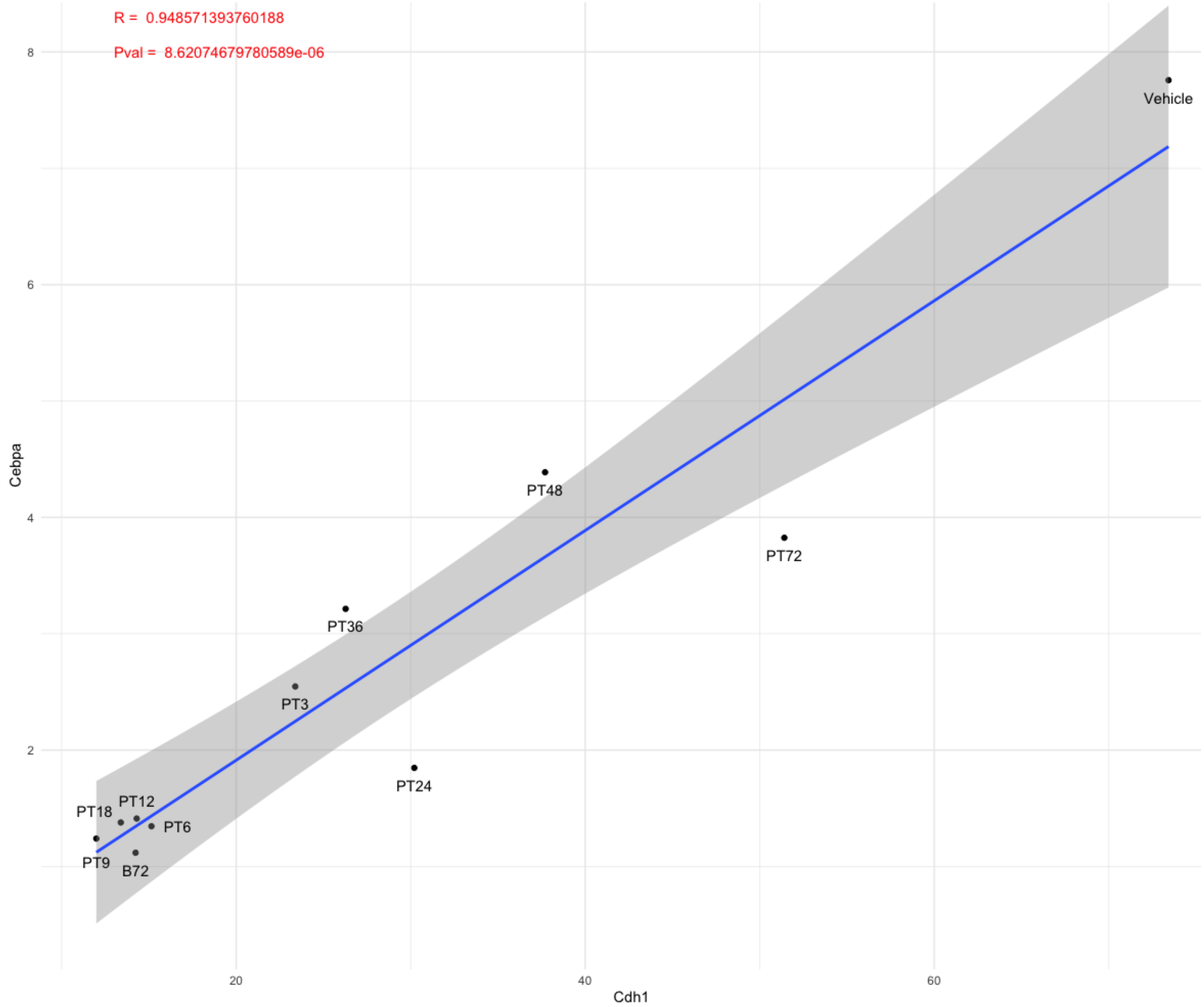


Figure 2.3 RNA-seq data of Cebpa correlation with Cdh1 during EMT-MET in 11 time points (Eskier, Karakulah and Alotaibi, submitted unpublished data). Vehicle = Normal epithelial NMuMG cells, PT = post-treatment after TGFβ removal with corresponding timepoints, B72 = TGFβ treatment for 72 hours. When TGFβ was removed from the medium, cells gain back their epithelial state as shown in progressive time points in PT condition.

This evidence data was one of the main reasons why we wanted to understand the regulation of Cebpa. Later with other experiments, the correlation of Cebpa with Cdh1 was also shown with additional data in the results section.

3. MATERIALS AND METHODS

3.1 Type of the Study

This study pursues an analytical and computational approach with in-vitro-supported experiments.

3.2 Time and Place of the Study

This study was conducted at İzmir International Biomedicine and Genome Center on dates of between June 2018 - July 2019.

3.3 Population and Samples of the Study

None of the experiments of the study were conducted on humans. The study was done in in-vitro using commercial cell lines and with in-silico methods using publicly available databases.

3.4 Variables of the Study

The NMuMG cell line model of EMT and MET is the independent variable. The siRNAs that were used to transfect the cells are dependent variables. The binding sites of transcription factors that were identified on the conserved regions of Cebpa are dependent variables. Gene Expression Omnibus (GEO) datasets obtained from National Center for Biotechnology Information (NCBI) database that consist of array and sequence based data of publicly available curated gene expression profiles are dependent variables.

3.5 Tools for Data Collection

3.5.1 Normal Murine Mammary Gland (NMuMG) Cell Line

The NMuMG cell line that were used in this study were obtained commercially (ATCC: The Global Bioresource Center/ CRL-1636).

3.5.2 Cell Culture

Cell culture experiments that were done to cultivate the cells using High Glucose Dulbecco's Modified Eagle's Medium (DMEM) (Gibco 41965-039) with addition of 10% Fetal Bovine Serum (FBS) (Gibco 1347D), 1% Non-essential aminoacids (NEAA) (Gibco 11140-035) and 1% Penicillin/Streptomycin antibiotics (Pen/Strep) (Gibco 15140-122) on 100mm cell culture plates. The preparation of complete medium was shown in Table 3.1.

Table 3.1 Preparation of Complete DMEM. The medium consists of 10% Fetal bovine serum, 1% Non-essential aminoacids and 1% antibiotics

Contents of the Medium	Concentrations/mL
DMEM	88%
FBS	10%
NEAA	1%
Pen/Strep	1% (0.1 mg/mL)

Incubation of the cells was done in CO_2 cell culture incubators (Eppendorf New Brunswick Galaxy 170S) at 37^oC with 10% CO_2. Collection and treatment of the cells were done using BSL-2 biosafety cabinets. According to the confluency of the cell culture plate that the cells were grown in, when the confluency reached around 80%, subculture of the cells was done. The procedure is as follows:

- Aspirate the medium from the plate
- Wash the cells with 10 mL 1X PBS twice
- Aspirate the PBS from the plate
- Add 1 mL of 0.25% Trypsin-EDTA onto the cells
- Incubate at 37^oC inside the incubator for 1-2 minutes until the cells are detached from the surface of the plate
- Add 9 mL of Complete DMEM on top of the cells to de-activate Trypsin-EDTA
- Collect the cells in a tube and centrifuge at 1200 rpm for 2 minutes
- Aspirate the supernatant
- Flick the tube gently to soften the cell pellet
- Add 10 mL of Complete DMEM and mix the content of the tube by pipetting
- Prepare a new plate with 9 mL of Complete DMEM
- Add 1 mL of the cell mixture into the plate

In every new subculture of NMuMG cells, 1µL/mL (10µg/mL) of insulin (Sigma-Aldrich/ I9278-5ML) was added fresh into the medium. The follow-up observations of the cells were done using inverted phase-contrast microscopy.

3.5.3 EMT-MET

NMuMG cells were treated with 5ng/mL TGFβ3 for 48 hours to change their phenotypes into mesenchymal cells. After EMT, the cells would go back to their epithelial state after the removal of TGFβ3 with washing the cells with 1X PBS for three times and changing the media with the addition of insulin.

3.5.4 Gene Silencing with siRNAs

Epithelial NMuMG cells were kept under regular culture conditions and the cells that were younger than sixth passage were chosen to use in gene silencing experiments. The cells were transfected with siRNAs using Lipofectamine RNAiMAX Transfection Reagent (#1836079 Invitrogen) to assess the difference of expression levels of Cebpa on MET. The siRNAs that were used in knock-down experiments are shown below in Table 3.2 and transfection calculation for siRNAs is shown in Table 3.3.

Table 3.2 siRNAs that were used in gene knock-down experiments. Scrambled siRNA represents Silencer select negative control siRNA.

siRNA	Company	Ref. Number
Scrambled (Scr) siRNA	ThermoFisher	4390844
Cebpa siRNA 1	Qiagen	SI00948311
Cebpa siRNA 2	Qiagen	SI00948318
Cebpa siRNA 3	Qiagen	SI00948328
Cebpa siRNA 4	Qiagen	SI00948332

As for the protocol:

- NMuMG cells were collected from the plate just as they would had been while subculturing and were counted.
- The cells were seeded into 6-well plates as 100,000 cells/well
- 1ul/ml of insulin and 0,5ul/ml TGFβ3 were added on each well
- Plates were put into 37^{o}C 10% CO_2 incubator for 48 hours for EMT
- After 48 hours when cells were in mesenchymal state, siRNA transfection were done using Lipofectamine RNAiMAX (Table## for transfection calculation)

- After transfection, cells were incubated at 37^{o}C for 48 hours and the media changed with complete DMEM
- After media change, cells were incubated at 37^{o}C for 24 hours and collected when incubation was done at 72nd hour as PT72 condition was met.
- The cells were then collected for RNA isolation.
- RNAs were then turned into cDNAs and analyses were done in qPCR for the expression of silenced genes.

Table 3.3 Transfection calculation for siRNAs. The concentration of transfection reagent and siRNA volumes were calculated according to the plate volume of manufacturers.

Plates	Surface Area	Plating volume	Dilution	RNA	Lipofectamine	siRNA volume
24 well	1	500 µL	2 x 50 µL	50 µM	1,5 µL	2,5 µL
12 well	2	1000 µL	2 x 100 µL	50 µM	3 µL	5 µL
6 well	5	2500 µL	2 x 250 µL	50 µM	7,5 µL	12,5 µL

3.5.6 RNA isolation

The RNA isolation were done using NucleoSpin RNA kit (Macherey-Nagel #740955.50) as described below.

- The cells in the 6-well plate were washed with 1X PBS for 2 times.
- 3,5 µL BME (Betamerchaptoethanol) added onto 350 µL of RA1 buffer per well and this would be the lysis buffer to detach the cells from the bottom surface of the wells.
- The cells were collected from the wells and put into the pink-columned filter microcentrifuge tubes.
- Centrifuged at 11000 G for 1 minute.
- Supernatant was kept and pink columns were discarded.
- 350 µL 70% Ethanol was added onto the supernatants and pipetted for 5-6 times.
- The mixture was loaded onto new blue-columned filter tubes.
- Centrifuged at 11000 G for 30 seconds.
- The blue column were taken out and put into new lidless tubes and the remaining solution was discarded.

- 350 μL of MDB buffer was added onto the blue column and centrifuged at 11000 G for 1 minute.
- The remaining liquid at the bottom of the tube were discarded and the tube with the column was centrifuged again at 11000 G for 30 seconds to get rid of the excess alcohol.
- DNAse mix were preaped as follows;
- 10 μL rDNAse
- 90 μL Reaction buffer
- From this 100 uL mix/well, 95 μL of it were added onto the column in the tubes and incubated at room temperature for 15 minutes.
- 200 μL of RAW2 buffer were added onto the columns and centrifuged at 11000 G for 30 seconds.
- The columns from the tubes then placed onto new lidless microcentrifuge tubes.
- 600 μL of RA3 buffer were added onto the columns and centrifuged at 11000 G for 30 seconds.
- The remaining liquid at the bottom of the tubes were discarded and the tubes were centrifuged again at 11000 G for 30 seconds.
- 250 μL of RA3 buffer were added onto the columns and centrifuged at 11000 G for 2 minutes.
- The columns were taken out and put into new microcentrifuge tubes with lids.
- Appropriate amount of Nuclease-free water were added onto the columns and centrifuged at 11000 G for 1 minute to elute the RNAs.
- RNA concentrations were measured with spectrophotometer (NanoDrop).

3.5.7 cDNA Synthesis

cDNA synthesis were done using Maxima First Strand cDNA Synthesis Kit (Thermo Fisher #K1642). The protocol is as described as in Table 3.4 below.

Table 3.4 cDNA synthesis mix. The mix was prepared according to the protocol of manufacturer. Template RNA was used in 1 ug concentration.

5X Reaction Mix	4 µL
Maxima Enzyme Mix	2 µL
Template RNA	1 µg
Nuclease-free water	up to 20 µL
Total volume	20 µL

RNAs that were measured were prepared as 1ug per tube for the use of cDNA synthesis.The mixture above was prepared for each sample individually and put into RNA-free tubes and kept on ice. The thermo-cycler was set to appropriate conditions;

- 25^{o}C for 10 minutes
- 50^{o}C for 30 minutes
- 85^{o}C for 5 minutes
- 4^{o}C at infinite

After the synthesis reaction was done, cDNAs could be used directly for qPCR or could be stored at -20^{o}C for up to a month or at -86^{o}C for longer periods of time.

3.5.8 qPCR (Real-Time PCR)

Probes for the qPCR that were used were listed at Table 3.5 below.

Table 3.5 qPCR probes. Cdh1 probe 18 was used to detect Cdh1 expression levels. Cebpa probe 67 was used to detect Cebpa expression levels. GADPH was used as housekeeping gene and was detected by probe 9.

Genes	Probes	Vendor	Ref. Number
GAPDH	Universal ProbeLibrary Probe #9	Roche	04685075001
Cdh1	Universal ProbeLibrary Probe #18	Roche	04686918001
Cebpa	Universal ProbeLibrary Probe #67	Roche	04688660001

The cDNAs that would be used for qPCR were diluted 1/20 prior to use and kept at 4^{o}C. The qPCR mix were prepared as described at Table 3.6 below and distributed at each well of 96-well qPCR plate (MicroAmp Fast Optical 96-well Reaction plate from Applied Biosystems #4346906).

Table 3.6 qPCR master mix

TaqMan Mastermix	6.25 µL
Forward Primer	0.5 µL
Reverse Primer	0.5 µL
Probe	0.25 µL
Nuclease-free water	2.5 µL
Total volume	10 µL

After putting the mixture in each well, 2.5 µL of cDNA were added onto their own wells that were corresponded to each of them.

3.5.9 Immunoflourescent Staning

In order to conduct fluorescence staining, cells were prepared on 12-well plates and EMT process was started for 48 hours. After 48 hours when cells were in mesenchymal state, siRNA transfection was done. For staning E-cadherin (E-cadherin; BD #610181) and Phalloidin (Thermo Fisher; #A22287 Alexa Fluor 647) antibodies were used. The protocol is as follows;

- Round coverslips were sterilized and inserted into each well of a 12-well plate.
- The cells were seeded as 40.000 cells/well upon TGFβ introduction.
- In 48th hour, the cells were transfected with siRNAs when they were in mesenchymal state
- After transfection, cells were incubated at 37^{o}C for 48 hours and the media changed with complete DMEM
- After media change, cells were incubated at 37^{o}C for 24 hours and the incubation was done at 72nd hour as PT72 condition was met.
- The coverslips were layed down on parafilm with their cell-side down
- The cells were washed with 1X PBS for three times under the edge of the coverslips
- Fixed with 4% formaldehyde for 10 minutes.
- Washed with 1X PBS for three times
- Permeabilized with 0.5% Triton X100 for 5 minutes
- Washed with 1X PBS for three times
- The primary antibody was applied for 1 hour (1:200)

- Washed with 1X PBS for three times
- The secondary antibody was applied for 1 hour (1:500)
- Washed with 1X PBS for three times
- Stained with DAPI for 2 minutes
- Washed with ultrapure water for one time
- Fixed on slides with mounting medium
- Imaging was done with Confocal microscopy (Zeiss, LSM880)

3.6 Tools and Databases

3.6.1 mVista

The tool from Joint Genome Institute of the University of California, *mVista*, was used to visualize, determine and analyze the conserved regions of Cebpa among different species.

The location of the gene region of Cebpa in *Mus musculus* on chromosome 7 with latest mouse genome mm10 (December 2011) release as follows; **"chr7:35,112,762-35,130,690"**

In order to visualize the alignment of conserved regions, five more species were added below the mouse gene region, and total in six species (mouse, human, rhesus, chimp, dog, cow) Cebpa conserved regions were analyzed.

3.6.2 GEO datasets

In order to compare and analyze the TFs that have correlation with Cebpa expression, Gene Expression Omnibus (GEO) database was used. The GEO profiles, series accession numbers, platforms and species are as follows in Table 3.7.

Table 3.7 The GEO datasets that were used in bioinformatics analysis. *Affymetrix Mouse Genome 430 2.0

Series Accession	Platform	Species	Cells/Tissue
GSE10175	GPL1261*	*Mus musculus*	Skin
GSE8733	GPL1261*	*Mus musculus*	Skin
GSE35160	GPL1261*	*Mus musculus*	Skin

The chosen GEO datasets have 36 samples in total. Raw data that were obtained in .CEL format were analyzed in R studio (Version 1.2.1335) with "simpleaffy", "genefilter", “gcrma”, “DGCA” and “annotate” packages. Annotation data were summoned from Affimetrix mouse genome 430 2.0 array annotation database “mouse4302.db”. List of mouse transcription factors was retrieved from Animal TFDB3.0 (Hu et al., 2019)

The number of the total probe sets that are in the datasets are 3334. The probe sets that were able to pass the threshold for correlation p-value <0.05 were 228 in total.

3.6.3 UCSC Genome Browser

In order to fact-check and determine the location of the sequences, and also for determining the present histone modification markers of the conserved sequences, UCSC Genome browser from the University of California Santa Cruz was used.

The location of the Cebpa region in 23kb length which were used to visualize histone modifications are as follows; **“chr7:35,894,947-35,918,597”** in mouse mm9 genome (July 2007) release.

The histone markers in the “Expression and Regulation” section of the genome browser that were used in visualization were selected from LICR histone and Caltech histone, respectively.

3.6.4 MatInspector

MatInspector tool on the Genomatix software suite (Cartharius et al., 2005) were used to search for transcription factor binding sites by using each of the conserved regions sequence data. Matrix family were selected “matches to individual matrices”, since individual family members were aimed to be identified rather than only families. The library version of the tool was selected as the latest release 11.0 (September 2017). Matrix group was selected as “Vertebrates”, core similarity was selected as 0.70, and matrix similarity was optimized.

3.7 Planning and Calendar of the Study

June’18-August’18	Sep’18 - Feb’19	Feb-June’19	July 2019
Literature reading	In-vitro analysis	In-silico analysis	Writing

3.8 Ethical Committee

There was no need to apply for ethical committee since the study was done using in-silico methods and with non-tumorigenic mouse mammary gland cell line.

4. RESULTS

4.1 Cebpa knockdown on MET

The NMuMG cells were cultured in the presence of TGFβ for 72 hours in order to gain their mesenchymal morphology. After that, at 72nd hour, NMuMG cells were transfected with control and Cebpa siRNAs and incubated for an additional 72 hours after the removal of TGFβ. The expression levels of both Cebpa and Cdh1 were calculated. Statistical significance determined using the Holm-Sidak method with unpaired T-test on GraphPad Prism 7.0. First, the most effective of siRNAs were selected to transfect the cells. (Figure 4.1). After determining the most effective Cebpa siRNA, the experiments were continued using Cebpa siRNA #3. As shown in the figure below (Figure 4.2), without the presence of Cebpa, the expression of Cdh1 was also reduced significantly.

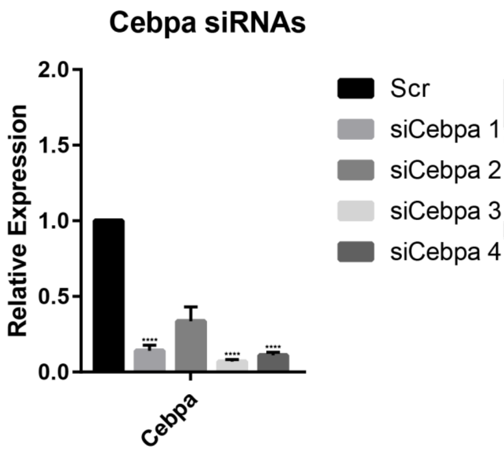


Figure 4.1 Relative expression of different Cebpa siRNAs. siCebpa 3 was selected for further experiments as it was the most effective. (**** $p<0.0001$)

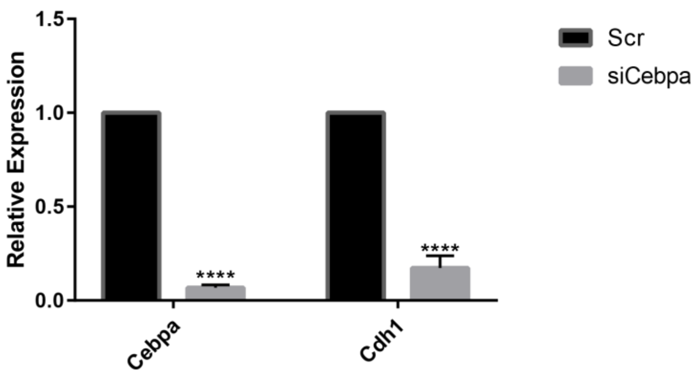


Figure 4.2 Relative expressions of Cebpa and Cdh1 when Cebpa was knocked-down. The expression of Cebpa was shown to drop when Cebpa was silenced. The expression of E-cad

(Cdh1) was also shown to be dropped significantly when Cebpa was silenced. (**** p<0.0001)

In order to investigate the morphological changes in the cells in contrast to epithelial NMuMG (Figure 4.3), microscopy images were taken after siRNA transfection at 72nd hour (Figure 4.4)

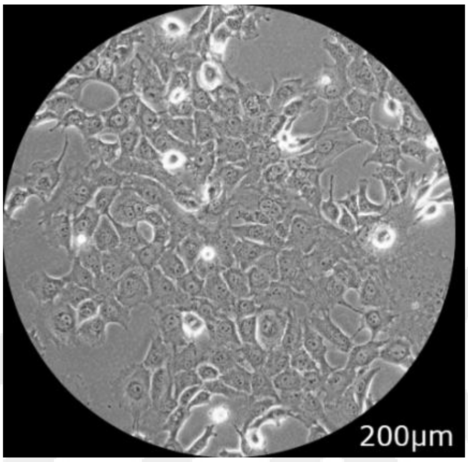


Figure 4.3 Normal epithelial NMuMG cells

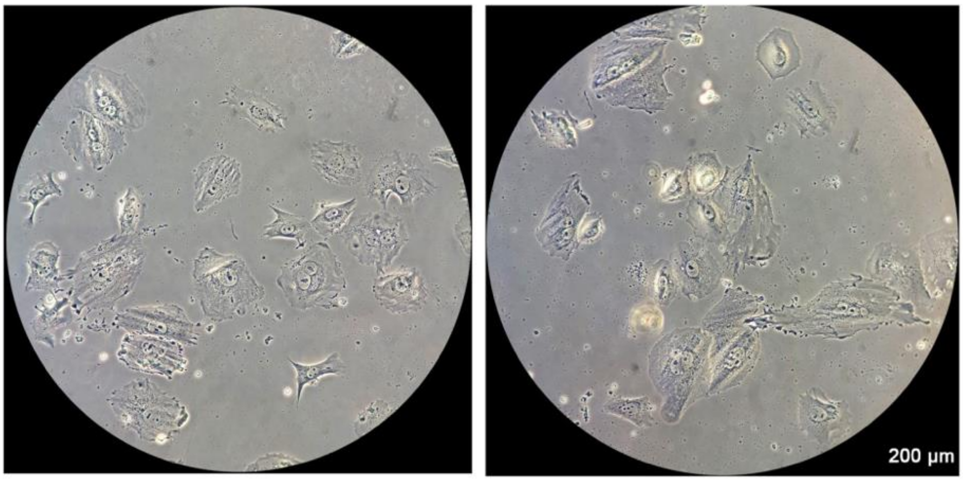


Figure 4.4 NMuMG cells in MET when Cebpa was silenced

To confirm the phenotype of the cells after Cebpa silencing and to examine the position of actin and the levels of E-cad localization, immunofluorescence staining was done using E-cadherin, and phalloidin antibodies. (Figure 4.5)

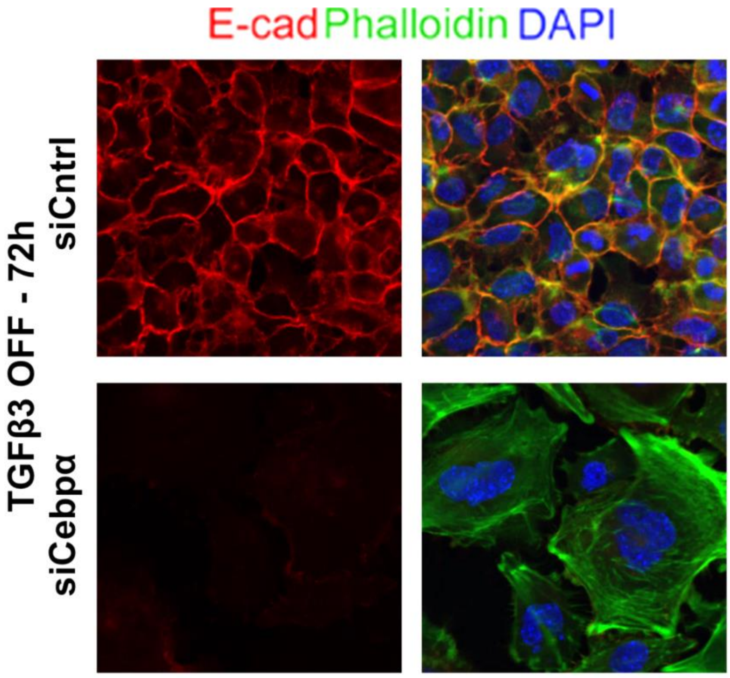


Figure 4.5 Confocal microscopy of Immunoflourescence staining of NMuMG in MET when Cebpa was silenced. Cells transfected with control siRNA (Scr) were able to go back to epithelial state in MET. Cells that were transfected with Cebpa siRNA, they were not able to go back to epithelial state, as it is evident in the absence of epithelial markers in in PT72 condition. E-cad staining (red), actin staning with phalloidin (green), nuclei was tagged with DAPI (blue).

As we investigate the results of Cebpa knockdown, confocal microscopy showed that when Cebpa was silenced, the cells were locked in mesenchymal state. Even though TGFβ was removed from the medium for 72 hours, the cells were not able to return to their epithelial state, as it is evident by the absence of E-cad expression in the cytoplasm and the position of actin in stress fibers. This observation is in agreement with mRNA levels examined by qPCR (Figure 4.1). The absence of E-cadherin neither in the cytoplasm nor in the membrane explains the blockage of the MET process.

4.2 Identification of the Conserved Sequences of Cebpa

Conservations were checked based on the mouse genome, aligned with the human, chimp, rhesus, dog, and cow genomes. Approximately 10kb portion on both sides of the entire conserved regions were selected including the upstream and downstream of the gene based on the latest mouse genome mm10 (December 2011) release. Total of eight conserved regions were selected with mVista tool, which included three upstream and five downstream of the gene, respectively (Fig.4.6 and 4.7)

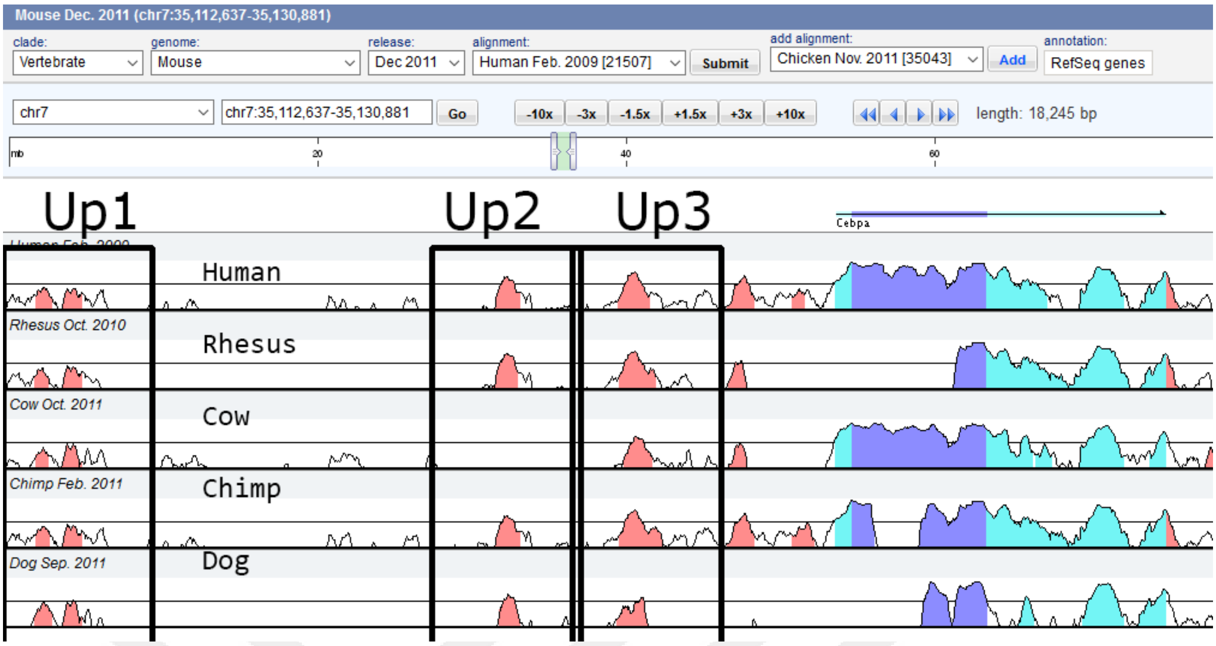


Figure 4.6 Conservation among species of Upstream Cebpa regions. From top to bottom, sequences were aligned against Mouse genome as follows; human, rhesus, cow, chimp, dog. Conserved sequences (red), UTR regions (cyan), gene body (blue).

Figure 4.7 Conservation among species of Downstream Cebpa regions. From top to bottom, sequences were aligned against Mouse genome as follows; human, rhesus, cow, chimp, dog. Conserved sequences (red), UTR regions (cyan), gene body (blue).

4.3 Analysis of the Conserved regions

Conserved intergenic regions on the aligned sequences were analyzed based on their conservation identities for each of the species. (Figures 4.8 to 4.12). The highest conservation on 5 of the species except cow was detected on Downstream 3 region with the highest score of 88.8% sequence conservation on the dog. Morever, the lowest conservation on 5 of the species except cow was detected on Upstream 1 region with the score of 67.2% sequence conservation on the mouse. Cow’s highest conservation was detected on Downstream 2 region with a score of 81.9%, and the lowest conservation was seen on Downstream 4 region with a score of 68.7%.

***** Conserved Regions - Mouse Dec. 2011 chr7 (Human Feb. 2009 chr19) *****

35112884	(33802302)	to	35113011	(33802169)	=	137bp	at	67.2%	intergenic
35113084	(33802036)	to	35113253	(33801876)	=	170bp	at	70.6%	intergenic
35116545	(33796496)	to	35116763	(33796284)	=	219bp	at	74.9%	intergenic
35117547	(33795177)	to	35117803	(33794920)	=	261bp	at	75.5%	intergenic
35122746	(33789786)	to	35123112	(33789423)	=	376bp	at	75.8%	intergenic
35123699	(33788632)	to	35124018	(33788306)	=	335bp	at	74.6%	intergenic
35124215	(33787829)	to	35124454	(33787601)	=	243bp	at	74.9%	intergenic
35124502	(33787554)	to	35124817	(33787225)	=	333bp	at	78.4%	intergenic
35126998	(33784666)	to	35127099	(33784568)	=	103bp	at	70.9%	intergenic
35127111	(33784556)	to	35127640	(33784017)	=	542bp	at	84.5%	intergenic
35127796	(33783862)	to	35128041	(33783618)	=	249bp	at	73.9%	intergenic
35128751	(33782685)	to	35129182	(33782250)	=	443bp	at	71.1%	intergenic
35129233	(33782213)	to	35129477	(33781975)	=	248bp	at	75.8%	intergenic
35130121	(33781339)	to	35130621	(33780837)	=	512bp	at	78.5%	intergenic

Figure 4.8 Mouse conservation against human on CNS regions

***** Conserved Regions - Mouse Dec. 2011 chr7 (Rhesus Oct. 2010 chr19) *****

35112881	(39936900)	to	35112987	(39936795)	=	107bp	at	70.1%	intergenic
35113087	(39936627)	to	35113253	(39936470)	=	167bp	at	71.9%	intergenic
35116539	(39932524)	to	35116746	(39932325)	=	208bp	at	77.4%	intergenic
35117555	(39931187)	to	35117849	(39930891)	=	300bp	at	76.3%	intergenic
35122752	(39926136)	to	35123106	(39925789)	=	362bp	at	77.1%	intergenic
35123693	(39925003)	to	35124027	(39924660)	=	352bp	at	74.7%	intergenic
35124207	(39924194)	to	35124454	(39923957)	=	251bp	at	75.3%	intergenic
35124498	(39923914)	to	35124819	(39923578)	=	339bp	at	78.2%	intergenic
35127111	(39920824)	to	35127640	(39920287)	=	542bp	at	83.4%	intergenic
35127796	(39920132)	to	35128036	(39919893)	=	244bp	at	75.4%	intergenic
35128835	(39918900)	to	35129167	(39918563)	=	343bp	at	73.2%	intergenic
35129228	(39918516)	to	35129511	(39918241)	=	286bp	at	77.6%	intergenic
35130121	(39917638)	to	35130613	(39917145)	=	504bp	at	78.2%	intergenic

Figure 4.9 Rhesus conservation scores against mouse on CNS regions

```
***** Conserved Regions - Mouse Dec. 2011 chr7 (Chimp Feb. 2011 chr19) *****

35112884  (38472279)  to  35112997  (38472164)  =     117bp  at  68.4%  intergenic
35113087  (38472010)  to  35113253  (38471853)  =     167bp  at  71.3%  intergenic
35116561  (38466382)  to  35116763  (38466186)  =     203bp  at  75.9%  intergenic
35117555  (38465069)  to  35117893  (38464728)  =     348bp  at  74.7%  intergenic
35122746  (38459404)  to  35123112  (38459041)  =     376bp  at  75.8%  intergenic
35123693  (38458187)  to  35124018  (38457853)  =     343bp  at  74.3%  intergenic
35124203  (38457388)  to  35124454  (38457148)  =     255bp  at  75.7%  intergenic
35124502  (38457101)  to  35124819  (38456770)  =     335bp  at  78.5%  intergenic
35126998  (38454193)  to  35127099  (38454095)  =     103bp  at  70.9%  intergenic
35127111  (38454083)  to  35127640  (38453544)  =     542bp  at  84.3%  intergenic
35127796  (38453389)  to  35128036  (38453150)  =     244bp  at  74.2%  intergenic
35128750  (38452213)  to  35129167  (38451790)  =     430bp  at  71.9%  intergenic
35129233  (38451739)  to  35129473  (38451503)  =     246bp  at  75.2%  intergenic
35130121  (38450863)  to  35130621  (38450361)  =     512bp  at  78.5%  intergenic
```

Figure 4.10 Chimp conservation against mouse on CNS regions

```
***** Conserved Regions - Mouse Dec. 2011 chr7 (Dog Sep. 2011 chr1) *****

35112882 (118841256)  to  35113011 (118841391)  =     138bp  at  69.6%  intergenic
35113085 (118841525)  to  35113250 (118841681)  =     166bp  at  74.7%  intergenic
35116601 (118845270)  to  35116756 (118845423)  =     156bp  at  82.7%  intergenic
35117541 (118846550)  to  35117791 (118846799)  =     262bp  at  71.4%  intergenic
35122761 (118852130)  to  35123077 (118852457)  =     328bp  at  76.2%  intergenic
35123699 (118853226)  to  35123954 (118853484)  =     269bp  at  77.3%  intergenic
35124206 (118853728)  to  35124381 (118853904)  =     177bp  at  83.6%  intergenic
35124520 (118854017)  to  35124796 (118854311)  =     295bp  at  75.9%  intergenic
35127125 (118855732)  to  35127503 (118856108)  =     379bp  at  88.9%  intergenic
35127949 (118856413)  to  35128058 (118856528)  =     116bp  at  73.3%  intergenic
35129252 (118857570)  to  35129510 (118857823)  =     260bp  at  73.5%  intergenic
35130121 (118858373)  to  35130554 (118858805)  =     446bp  at  73.5%  intergenic
```

Figure 4.11 Dog conservation against mouse on CNS regions

```
***** Conserved Regions - Mouse Dec. 2011 chr7 (Cow Oct. 2011 chr18) *****

35112891  (43018049)  to  35112987  (43017949)  =     101bp  at  71.3%  intergenic
35113085  (43017800)  to  35113250  (43017644)  =     166bp  at  73.5%  intergenic
35117598  (43011247)  to  35117814  (43011034)  =     218bp  at  72.5%  intergenic
35122801  (43005930)  to  35123087  (43005640)  =     293bp  at  80.9%  intergenic
35123697  (43004943)  to  35123922  (43004710)  =     239bp  at  73.2%  intergenic
35124217  (43004409)  to  35124381  (43004236)  =     177bp  at  81.9%  intergenic
35124537  (43004088)  to  35124807  (43003815)  =     279bp  at  79.9%  intergenic
35127105  (43002309)  to  35127589  (43001824)  =     488bp  at  81.6%  intergenic
35127856  (43001549)  to  35128065  (43001328)  =     224bp  at  73.2%  intergenic
35128799  (43000571)  to  35128943  (43000432)  =     150bp  at  68.7%  intergenic
35129261  (42999928)  to  35129516  (42999679)  =     263bp  at  73.8%  intergenic
35130183  (42999165)  to  35130426  (42998915)  =     254bp  at  79.1%  intergenic
```

Figure 4.12 Cow conservation against mouse on CNS regions

4.4 Identification of Cebpa Promoter

One kb upstream of Cebpa (Figure 4.13) was selected, and conservation among the species mouse, human, rhesus, cow, and the chimp was identified. The functional elements were shown on the promoter region located at “chr7:35,118,293-35,119,292” on mouse genome mm10 December 2011 release.

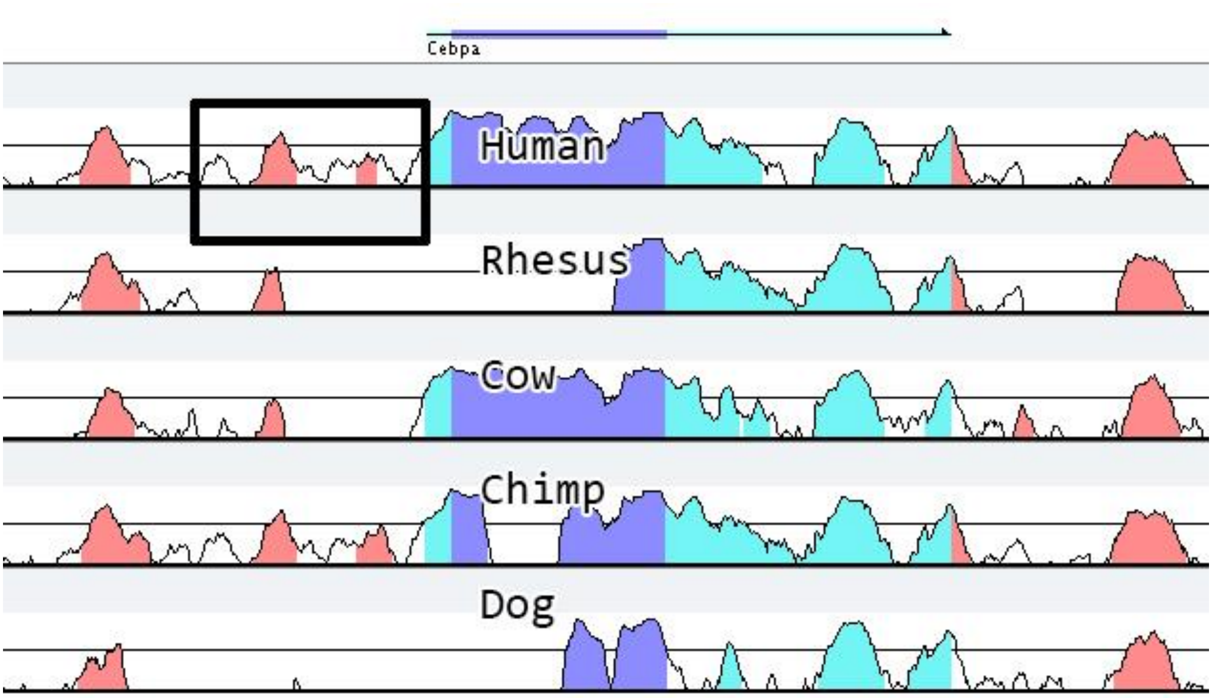


Figure 4.13 1 kb Cebpa promoter region on conserved species. From top to bottom, sequences were aligned against Mouse genome as follows; human, rhesus, cow, chimp, dog.

The sequence of the region was obtained from UCSC genome browser. Conservation identities are shown below between mouse versus selected vertebrates.

******* Conserved Regions - Mouse Dec. 2011 vs Selected vertebrates *******

Human	35118424	(33794256)	to	35118634	(33794055)	=	211bp	at	73.9%	intergenic
Rhesus	35118429	(39930275)	to	35118585	(39930116)	=	161bp	at	73.3%	intergenic
Cow	35118440	(43010357)	to	35118592	(43010206)	=	154bp	at	70.1%	intergenic
Chimp	35118424	(38464159)	to	35118634	(38463958)	=	211bp	at	74.4%	intergenic

In order to show transcription start site (TSS), 200bp was added on 3’ end. The conserved regions and functional elements were shown in the sequence below (Figure 4.14). The top 10

TFBSs on 1 kb promoter region of Cebpa were shown in Table 4.1. Please refer to the appendix for the full table (Table A1).

```
   1 tccccgattc aagttcactc ctctccaaca ccctgccctc tcgcggcccc tgtgcgctct
  61 ccttagggtc cttttccgcg aggctcagag gacccaccgg ctggtcccgg gcgtggggtg
 121 gtggtgtccc gaacacttga ctagagtgct ccacgctggg tagcaacgtc tgcctggtaa
 181 gcctagcaat cctatcgctc tggcctggag acgcaatgaa aaagaaagtt ttccagccta
 241 ggcgagtgga cgagccaggt ccaccaggct ccggggttag cggccgcctc cgctccccgc
 301 cgggtcctag cgccctacta ctctgaagag cccgtgggac cctgtagttc tagagaagct
 361 gggcgaaaga gaggtgctct gcctggaagc cgtggggtcg cgtggagttc agagaaaaag
 421 acgcacaatc tctgcgctcc cggcctcgcc actcggcggt gcgcgctagg ttgctggtcc
 481 aaagcagtct ccaacctccc gcgccgcggc tctccgccac agccttttag aaatccgggt
 541 gggagacagg cctagtccca gcttttaaca caagtctgca ctacggtagc tcaaaaccaa
 601 cattctctct ccaaacgctc cccaacctcc acctcccctc gctcggcctc tagatgctcc
 661 cgggctccct agtgttggct ggaagtgggt gacttagagg cttaaaggag gggcgcctaa
 721 ccacggacca cgtgtgtgcg ggggcgacag cgccgccggg gtggggctga gcgctgcaag
 781 ccgggttcgc cttgcagcgc aggagtcagt gggcgttgcg ccacgatctc tctccactag
 841 cactatgctc cccactcacc gccttggaaa gtcacaggag aaggcgggct ctaagaccca
 901 gcaggcacca tcctactggc gccttcgatc cgagacccgt ttggacacca gggggcgatg
 961 cccgaccctc tataaaagcg gtccccgcgc gggcctggcc ATTCGCGACC CGAAGCTGCG
1021 CGGGCGCGAG CCAGTTGGGG CACTGGGTGG GCGGCGGCGA CAGCGGCGCC ACGCGCAGGC
1081 TGGAGGCCGC CGAGGCTCGC CATGCCGGGA GAACTCTAAC TCCCCCATGG AGTCGGCCGA
1141 CTTCTACGAG GTGGAGCCGC GGCCCCCGAT GAGCAGTCAC CTCCAGAGCC CCCCGCACGC
```

Figure 4.14 Conserved sequences and functional elements present on the 1kb promoter of Cebpa. Conserved sequences (blue), TATA box (pink), GC box (Sp1 box) (yellow), TSS (green) were shown.

Table 4.1 Top 10 TFBSs on Cebpa 1 kb promoter region

Matrix	Matrix sim.	Sequence
O$TF2B/BRE.01	1	ccgCGCC
O$VTBP/VTATA.02	1	ctctaTAAAagcggtcc
V$EBOX/USF.01	1	gcacaCACGtggtccgt
V$MZF1/MZF1.01	1	gtGGGGagcat
V$NKXH/NKX25.01	1	ctagtcAAGTgttcgggac
V$PLAG/PLAG1.02	1	caGGGGgcgatgcccgaccctct
V$TAIP/CSRNP1.01	1	AGAGtgc
V$TAIP/CSRNP3.01	1	aGAGTgc
V$ZF02/ZKSCAN3.01	1	ctcagCCCCaccccggcggcgct
V$KLFS/IKLF.01	0.997	cgccgGGGTgggctgagc

4.5 Transcription Factor Binding Sites on Cebpa 23kb region

Each of the eight conserved regions of approximately 20kb of upstream and downstream of Cebpa region were analyzed on *MatInspector* and the analysis of transcription factor binding site predictions showed a vast number of possible transcription factor binding sites (TFBSs). I opted to divide them into sections which in each of them only consists of the species and the sequence information of the region along with their predicted TFBSs.

4.5.1 Mus musculus and Conserved Regions and Their TFBSs

4.5.1.1 Upstream CNS1 Region

This region consists of 2 intergenic sequences, 137bp, and 170bp in length. The region was aligned mouse against the human genome (Figure 4.15). In MatInspector, the two sequences were submitted separately to identify the TFBSs and the lists were merged after analyzing.

```
***** Conserved Regions - Mouse Dec. 2011 chr7 (Human Feb. 2009 chr19) *****

35112884 (33802302) to 35113011 (33802169) = 137bp at 67.2% intergenic
35113084 (33802036) to 35113253 (33801876) = 170bp at 70.6% intergenic
35116545 (33796496) to 35116763 (33796284) = 219bp at 74.9% intergenic
35117547 (33795177) to 35117803 (33794920) = 261bp at 75.5% intergenic
35122746 (33789786) to 35123112 (33789423) = 376bp at 75.8% intergenic
35123699 (33788632) to 35124018 (33788306) = 335bp at 74.6% intergenic
35124215 (33787829) to 35124454 (33787601) = 243bp at 74.9% intergenic
35124502 (33787554) to 35124817 (33787225) = 333bp at 78.4% intergenic
35126998 (33784666) to 35127099 (33784568) = 103bp at 70.9% intergenic
35127111 (33784556) to 35127640 (33784017) = 542bp at 84.5% intergenic
35127796 (33783862) to 35128041 (33783618) = 249bp at 73.9% intergenic
35128751 (33782685) to 35129182 (33782250) = 443bp at 71.1% intergenic
35129233 (33782213) to 35129477 (33781975) = 248bp at 75.8% intergenic
35130121 (33781339) to 35130621 (33780837) = 512bp at 78.5% intergenic
```

Figure 4.15 Cebpa Upstream CNS1 region conservation identities shown on mouse against the human genome

After filtering out the repeating entries of the databases and arranging the TFs from the highest matrix similarity, we got a table of 64 TFBS predictions. The top 10 results were shown as in the table below (Table 4.2). Please refer to appendix for the full table (Table A2).

Table 4.2 Top 10 TFBSs on *Mus musculus* Upstream CNS1 region

Matrix	strand	Matrix sim.	Sequence
V$PLAG/PLAG1.02	(+)	1	gaGGGGggcctctttccagtca
V$NOLF/EBF1.01	(-)	0.996	atgtggTCCCcagggcttggct
V$RUSH/SMARCA3.02	(+)	0.993	aggcACTTttt
V$MOKF/MOK2.02	(-)	0.988	gaacaaaaaagtgCCTTggaa
V$NEUR/NGN_NEUROD.01	(+)	0.988	accaCATCtgcctga
V$ZF07/ZNF263.02	(-)	0.988	cccctcCTCCcctgg
V$NOLF/OLF1.02	(-)	0.977	atgtggTCCCcagggcttggct
V$SMAD/SMAD.01	(-)	0.976	cttGTCTgtcc
V$AP2F/AP2.02	(+)	0.97	tctGCCTgaggggg
V$AP2F/TCFAP2A.02	(+)	0.964	agccCCTGgggacca

In this region, the TFs that have a binding site mainly revolves around hormone receptor binding sites, myc-associated zing finger proteins, HMG-box proteins, activator protein (AP) and its TFs, and octamer-binding sites.

4.5.1.2 Upstream CNS2 Region

This region consists of 1 intergenic sequence, 219bp in length with 74.9% conservation identity. The region was aligned mouse against the human genome (Figure 4.16).

```
***** Conserved Regions - Mouse Dec. 2011 chr7 (Human Feb. 2009 chr19) *****

35112884  (33802302)  to  35113011  (33802169)  =      137bp  at  67.2%  intergenic
35113084  (33802036)  to  35113253  (33801876)  =      170bp  at  70.6%  intergenic
35116545  (33796496)  to  35116763  (33796284)  =      219bp  at  74.9%  intergenic
35117547  (33795177)  to  35117803  (33794920)  =      261bp  at  75.5%  intergenic
35122746  (33789786)  to  35123112  (33789423)  =      376bp  at  75.8%  intergenic
35123699  (33788632)  to  35124018  (33788306)  =      335bp  at  74.6%  intergenic
35124215  (33787829)  to  35124454  (33787601)  =      243bp  at  74.9%  intergenic
35124502  (33787554)  to  35124817  (33787225)  =      333bp  at  78.4%  intergenic
35126998  (33784666)  to  35127099  (33784568)  =      103bp  at  70.9%  intergenic
35127111  (33784556)  to  35127640  (33784017)  =      542bp  at  84.5%  intergenic
35127796  (33783862)  to  35128041  (33783618)  =      249bp  at  73.9%  intergenic
35128751  (33782685)  to  35129182  (33782250)  =      443bp  at  71.1%  intergenic
35129233  (33782213)  to  35129477  (33781975)  =      248bp  at  75.8%  intergenic
35130121  (33781339)  to  35130621  (33780837)  =      512bp  at  78.5%  intergenic
```

Figure 4.16 Cebpa Upstream CNS2 region conservation identities shown on mouse against the human genome

After filtering out the repeating entries of the databases and arranging the TFs from the highest matrix similarity, we got a table of 111 TFBS predictions. The top 10 results were shown as in the table below (Table 4.3). Please refer to the appendix for the full table (Table A3).

Table 4.3 Top 10 TFBSs on *Mus musculus* Upstream CNS2 region

Matrix	strand	Matrix sim.	Sequence
V$MYOD/TCF12.01	(-)	1	gaacagcaGGTGcaaag
V$MZF1/MZF1.01	(-)	1	gtGGGGagagc
V$TALE/TGIF.01	(+)	1	ctccaacctGTCAaaca
V$FKHD/FOXJ3.01	(-)	0.999	cccaggTAAAcaaacct
V$TALE/MEIS1.01	(+)	0.999	ctccaaccTGTCaaaca
V$FKHD/FOXP1.01	(-)	0.996	cccaggtAAACaaacct
V$HAND/TAL1_E2A.02	(-)	0.996	gagggaaCAGCaggtgcaaag
V$FKHD/FREAC2.01	(-)	0.994	cccaggTAAAcaaacct
V$NEUR/MASH1.01	(+)	0.994	tttgcacCTGCtgtt

V$SF1F/SF1.01	(-)	0.993	tggcCAAGgccatcc

In this region, mainly TFs that are associated with hepatic nuclear factors and forkhead-domain proteins, GATA binding proteins, hormone binding sites, KLF binding sites, MYOD family TFs, SOX proteins, TALE family TFs are common.

4.5.1.3 Upstream CNS3 Region

This region consists of 1 intergenic sequence, 261bp in length with 75.5% conservation identity. The region was aligned mouse against the human genome (Figure 4.17).

```
***** Conserved Regions - Mouse Dec. 2011 chr7 (Human Feb. 2009 chr19) *****

35112884 (33802302) to 35113011 (33802169) =  137bp at 67.2% intergenic
35113084 (33802036) to 35113253 (33801876) =  170bp at 70.6% intergenic
35116545 (33796496) to 35116763 (33796284) =  219bp at 74.9% intergenic
35117547 (33795177) to 35117803 (33794920) =  261bp at 75.5% intergenic
35122746 (33789786) to 35123112 (33789423) =  376bp at 75.8% intergenic
35123699 (33788632) to 35124018 (33788306) =  335bp at 74.6% intergenic
35124215 (33787829) to 35124454 (33787601) =  243bp at 74.9% intergenic
35124502 (33787554) to 35124817 (33787225) =  333bp at 78.4% intergenic
35126998 (33784666) to 35127099 (33784568) =  103bp at 70.9% intergenic
35127111 (33784556) to 35127640 (33784017) =  542bp at 84.5% intergenic
35127796 (33783862) to 35128041 (33783618) =  249bp at 73.9% intergenic
35128751 (33782685) to 35129182 (33782250) =  443bp at 71.1% intergenic
35129233 (33782213) to 35129477 (33781975) =  248bp at 75.8% intergenic
35130121 (33781339) to 35130621 (33780837) =  512bp at 78.5% intergenic
```

Figure 4.17 Cebpa Upstream CNS3 region conservation identities shown on mouse against the human genome

After filtering out the repeating entries of the databases and arranging the TFs from the highest matrix similarity, we got a table of 121 TFBS predictions. The top 10 results were shown as in the table below (Table 4.4). Please refer to appendix for the full table (Table A4).

Table 4.4 Top 10 TFBSs on *Mus musculus* Upstream CNS3 region

Matrix	strand	Matrix sim.	Sequence
V$TAIP/CSRNP1.01	(-)	1	AGAGtga
V$TAIP/CSRNP3.01	(-)	1	aGAGTga
V$FKHD/HFH3.01	(-)	0.994	actaaatAAACaacctg
V$FKHD/HFH8.01	(-)	0.99	actaaatAAACaacctg
V$FKHD/FHXA.01	(-)	0.988	actaaatAAACaacctg
V$BRAC/EOMES.03	(-)	0.987	taaacaacctggTGTGagcagaggatcgc
V$BCDF/CRX.01	(-)	0.985	tactTAATcaaattct
V$FKHD/FREAC7.01	(-)	0.984	actaaaTAAAcaacctg

V$NEUR/NGN_NEUROD.01	(+)	0.982	cgtgCATCtggccca
V$SAL1/SALL1.01	(-)	0.975	gcATAAaaaagat

In this region, mainly TFs that are associated with Forkhead-box proteins, CART homeobox proteins, hepatic nuclear factor proteins, LIM homeobox TFs, HOX homeobox proteins are common.

4.5.1.4 Downstream CNS1 Region

This region consists of 1 intergenic sequence, 376bp in length. The region was aligned mouse against the human genome (Figure 4.18).

```
***** Conserved Regions - Mouse Dec. 2011 chr7 (Human Feb. 2009 chr19) *****

35112884 (33802302) to 35113011 (33802169) =   137bp at 67.2% intergenic
35113084 (33802036) to 35113253 (33801876) =   170bp at 70.6% intergenic
35116545 (33796496) to 35116763 (33796284) =   219bp at 74.9% intergenic
35117547 (33795177) to 35117803 (33794920) =   261bp at 75.5% intergenic
35122746 (33789786) to 35123112 (33789423) =   376bp at 75.8% intergenic
35123699 (33788632) to 35124018 (33788306) =   335bp at 74.6% intergenic
35124215 (33787829) to 35124454 (33787601) =   243bp at 74.9% intergenic
35124502 (33787554) to 35124817 (33787225) =   333bp at 78.4% intergenic
35126998 (33784666) to 35127099 (33784568) =   103bp at 70.9% intergenic
35127111 (33784556) to 35127640 (33784017) =   542bp at 84.5% intergenic
35127796 (33783862) to 35128041 (33783618) =   249bp at 73.9% intergenic
35128751 (33782685) to 35129182 (33782250) =   443bp at 71.1% intergenic
35129233 (33782213) to 35129477 (33781975) =   248bp at 75.8% intergenic
35130121 (33781339) to 35130621 (33780837) =   512bp at 78.5% intergenic
```

Figure 4.18 Cebpa Downstream CNS1 region conservation identities shown on mouse against the human genome

After filtering out the repeating entries of the databases and arranging the TFs from the highest matrix similarity, we got a table of 116 TFBS predictions. The top 10 results were shown as in the table below (Table 4.5). Please refer to the appendix for the full table (Table A5).

Table 4.5 Top 10 TFBSs on *Mus musculus* Downstream CNS1 region

Matrix	strand	Matrix sim.	Sequence
V$PLAG/PLAG1.02	(-)	1	aaGGGGgaattcaaggatcattc
V$ZF02/ZBTB7.03	(-)	0.998	agggcCCCCtcccccgttcagcc
V$FKHD/FOXP2.01	(-)	0.997	tccctgtaAACAagaaa
V$KLFS/BKLF.02	(+)	0.997	aacgggGGAGgggccctc
V$EGRF/CKROX.01	(+)	0.996	gaacggGGGAgggggccct
V$FAST/FAST1.03	(+)	0.996	ttcagtgtgGATTccag

V$MZF1/MZF1.02	(-)	0.995	tgGGGGaaggg
V$ZTRE/ZTRE.03	(-)	0.993	ccCTCCcccgttcagcc
V$FKHD/FOXO1.01	(-)	0.991	tccctgtaAACAagaaa
V$FKHD/FOXP1.01	(-)	0.991	tccctgtAAACaagaaa

In this region, mainly TFs that are associated with hepatic nuclear factors, nuclear receptors, KLF TFs, SPI proteins, epidermal growth factor receptors are common.

4.5.1.5 Downstream CNS2 Region

This region consists of 3 intergenic sequences, 335bp, 243bp and 333bp in length. The region was aligned mouse against the human genome (Figure 4.19). In MatInspector the three sequences were submitted separately to identify the TFBSs.

```
***** Conserved Regions - Mouse Dec. 2011 chr7 (Human Feb. 2009 chr19) *****

35112884 (33802302) to 35113011 (33802169) =    137bp at 67.2% intergenic
35113084 (33802036) to 35113253 (33801876) =    170bp at 70.6% intergenic
35116545 (33796496) to 35116763 (33796284) =    219bp at 74.9% intergenic
35117547 (33795177) to 35117803 (33794920) =    261bp at 75.5% intergenic
35122746 (33789786) to 35123112 (33789423) =    376bp at 75.8% intergenic
35123699 (33788632) to 35124018 (33788306) =    335bp at 74.6% intergenic
35124215 (33787829) to 35124454 (33787601) =    243bp at 74.9% intergenic
35124502 (33787554) to 35124817 (33787225) =    333bp at 78.4% intergenic
35126998 (33784666) to 35127099 (33784568) =    103bp at 70.9% intergenic
35127111 (33784556) to 35127640 (33784017) =    542bp at 84.5% intergenic
35127796 (33783862) to 35128041 (33783618) =    249bp at 73.9% intergenic
35128751 (33782685) to 35129182 (33782250) =    443bp at 71.1% intergenic
35129233 (33782213) to 35129477 (33781975) =    248bp at 75.8% intergenic
35130121 (33781339) to 35130621 (33780837) =    512bp at 78.5% intergenic
```

Figure 4.19 Cebpa Downstream CNS2 region conservation identities shown on mouse against the human genome

After filtering out the repeating entries of the databases and arranging the TFs from the highest matrix similarity, we got a table of 297 TFBS predictions. The top 10 results were shown as in the table below (Table 4.6). Please refer to the appendix for the full table (Table A6).

Table 4.6 Top 10 TFBSs on *Mus musculus* Downstream CNS2 region

Matrix	strand	Matrix sim.	Sequence
V$SMAD/SMAD.01	(-)	1	tttGTCTggct
V$PLAG/PLAG1.02	(-)	1	gaGGGGgcgggaagcctgggagg
V$SORY/SOX12.01	(+)	1	ccagaACAAttaaacagccagat
V$TAIP/CSRNP1.01	(-)	1	AGAGtgg

V$TAIP/CSRNP3.01	(-)	1	aGAGTgg
V$ZTRE/ZTRE.04	(-)	1	ctgGGAGgggtggacac
V$KLFS/KLF2.01	(+)	1	agcagGGGTggggtttggg
V$E2FF/E2F1.01	(-)	1	agggGGCGggaagcctg
V$E2FF/E2F4.01	(-)	1	aggggGCGGgaagcctg
V$FKHD/FOXP1_ES.01	(-)	1	aaataaaAACAaaagtg

Mainly on this region, hormone receptors, STAT proteins, GATA proteins, HOX proteins, MYOD family TFs, POU factors, SOX proteins, ETS family, HOX homeobox proteins, SP1 family, KLF TFs, E2F family, bZIP factors, binding sites of heterodimers of various factors, HMG box factors and their heterodimers involved in Wnt signaling pathway, Forkhead box proteins, zinc finger TFs, LIM homeobox factors, octamer binding sites, hepatic nuclear factors are common.

4.5.1.6 Downstream CNS3 Region

This region consists of 3 intergenic sequences, 103bp, 542bp and 249bp in length. The region was aligned mouse against the human genome (Figure 4.20). In MatInspector the three sequences were submitted separately to identify the TFBSs.

```
***** Conserved Regions - Mouse Dec. 2011 chr7 (Human Feb. 2009 chr19) *****

35112884 (33802302) to 35113011 (33802169) =     137bp at 67.2% intergenic
35113084 (33802036) to 35113253 (33801876) =     170bp at 70.6% intergenic
35116545 (33796496) to 35116763 (33796284) =     219bp at 74.9% intergenic
35117547 (33795177) to 35117803 (33794920) =     261bp at 75.5% intergenic
35122746 (33789786) to 35123112 (33789423) =     376bp at 75.8% intergenic
35123699 (33788632) to 35124018 (33788306) =     335bp at 74.6% intergenic
35124215 (33787829) to 35124454 (33787601) =     243bp at 74.9% intergenic
35124502 (33787554) to 35124817 (33787225) =     333bp at 78.4% intergenic
35126998 (33784666) to 35127099 (33784568) =     103bp at 70.9% intergenic
35127111 (33784556) to 35127640 (33784017) =     542bp at 84.5% intergenic
35127796 (33783862) to 35128041 (33783618) =     249bp at 73.9% intergenic
35128751 (33782685) to 35129182 (33782250) =     443bp at 71.1% intergenic
35129233 (33782213) to 35129477 (33781975) =     248bp at 75.8% intergenic
35130121 (33781339) to 35130621 (33780837) =     512bp at 78.5% intergenic
```

Figure 4.20 Cebpa Downstream CNS3 region conservation identities shown on mouse against the human genome

After filtering out the repeating entries of the databases and arranging the TFs from the highest matrix similarity, we got a table of 334 TFBS predictions. The top 10 results were shown as in the table below (Table 4.7). Please refer to the appendix for the full table (Table A7).

Table 4.7 Top 10 TFBSs on *Mus musculus* Downstream CNS3 region

Matrix	strand	Matrix sim.	Sequence
V$PLAG/PLAG1.02	(+)	1	aaGGGGgtgtcagggactgggta
V$CART/S8.01	(+)	1	cactgTAATtggctttggccc
V$TAIP/CSRNP1.01	(-)	1	AGAGtgc
V$TAIP/CSRNP3.01	(-)	1	aGAGTga
V$MZF1/MZF1.01	(+)	1	gtGGGGatgag
V$FKHD/FOXP1_ES.01	(-)	1	caaataaAACAtgtttg
V$OCT1/OCT1.02	(+)	1	tttATGCaaatcgta
V$FKHD/HNF3.01	(-)	0.999	gtgaggcAAACacattt
V$ZF11/ZBTB3.01	(-)	0.997	gccaGCCAgag
V$EBOX/MNT.01	(-)	0.997	cccaaCACGtgcacagg

Mainly on this region, lymphocyte differentiation regulators, hormone receptors, Distal-less homeobox, GATA proteins, HOX proteins, Grhl TFs, HOM family proteins, CART family proteins, MYOD family proteins, p53 family proteins, IRF factors, myocyte enhancer factors, SOX proteins, ETS family proteins, hepatic nuclear factors, Iroquois homeobox, orphan receptors, bZIP factors, CEBP family TFs, NFY factors, Forkhead box proteins, TCFAP family factors, EBOX factors, octamer binding sites, PAX proteins, HBOX family factors, are common.

4.5.1.7 Downstream CNS4 Region

This region consists of 2 intergenic sequences, 443bp and 248bp in length. The region was aligned mouse against the human genome (Figure 4.21). In MatInspector the three sequences were submitted separately to identify the TFBSs.

```
***** Conserved Regions - Mouse Dec. 2011 chr7 (Human Feb. 2009 chr19) *****

35112884 (33802302) to 35113011 (33802169) =   137bp at 67.2% intergenic
35113084 (33802036) to 35113253 (33801876) =   170bp at 70.6% intergenic
35116545 (33796496) to 35116763 (33796284) =   219bp at 74.9% intergenic
35117547 (33795177) to 35117803 (33794920) =   261bp at 75.5% intergenic
35122746 (33789786) to 35123112 (33789423) =   376bp at 75.8% intergenic
35123699 (33788632) to 35124018 (33788306) =   335bp at 74.6% intergenic
35124215 (33787829) to 35124454 (33787601) =   243bp at 74.9% intergenic
35124502 (33787554) to 35124817 (33787225) =   333bp at 78.4% intergenic
35126998 (33784666) to 35127099 (33784568) =   103bp at 70.9% intergenic
35127111 (33784556) to 35127640 (33784017) =   542bp at 84.5% intergenic
35127796 (33783862) to 35128041 (33783618) =   249bp at 73.9% intergenic
35128751 (33782685) to 35129182 (33782250) =   443bp at 71.1% intergenic
35129233 (33782213) to 35129477 (33781975) =   248bp at 75.8% intergenic
35130121 (33781339) to 35130621 (33780837) =   512bp at 78.5% intergenic
```

Figure 4.21 Cebpa Downstream CNS4 region conservation identities shown on mouse against the human genome

After filtering out the repeating entries of the databases and arranging the TFs from the highest matrix similarity, we got a table of 276 TFBS predictions. The top 10 results were shown as in the table below (Table 4.8). Please refer to the appendix for the full table (Table A8).

Table 4.8 Top 10 TFBSs on *Mus musculus* Downstream CNS4 region

Matrix	strand	Matrix sim.	Sequence
V$ZFHX/DELTAEF1.01	(-)	1	gactcACCTgact
V$ZFXY/ZFX.01	(-)	1	ggGGCCtggcc
V$TAIP/CSRNP1.01	(+)	1	AGAGtgt
V$TAIP/CSRNP3.01	(+)	1	aGAGTgt
V$PARF/TEF.01	(+)	1	gaagtttatGTAAttat
V$TALE/TGIF.01	(+)	1	cagaaatgtGTCAgagg
V$MZF1/MZF1.01	(+)	1	gtGGGGacgat
V$ZF02/ZKSCAN3.01	(-)	1	atcgtCCCCacctccccacggca
V$ETSF/ETV1.02	(+)	0.997	acaaggcaGGAAggcagctct
V$TEAF/TEAD4.01	(-)	0.996	tgacATTCctggc

Mainly in this region, TF that are AP1F-AP1R family, EGRF family, retinoid receptors, CART family, hepatic nuclear factors, ETS family, HOX proteins, TALE family, bZIP factors, AML-related factors, LIM homeobox factors, EBOX factors, HBOX family proteins, PAX proteins common.

4.5.1.8 Downstream CNS5 Region

This region consists of one intergenic sequence, 512bp in length with 78.5% conservation identity. The region was aligned mouse against the human genome (Figure 4.22).

```
***** Conserved Regions - Mouse Dec. 2011 chr7 (Human Feb. 2009 chr19) *****

35112884  (33802302)  to  35113011  (33802169)  =     137bp  at  67.2%  intergenic
35113084  (33802036)  to  35113253  (33801876)  =     170bp  at  70.6%  intergenic
35116545  (33796496)  to  35116763  (33796284)  =     219bp  at  74.9%  intergenic
35117547  (33795177)  to  35117803  (33794920)  =     261bp  at  75.5%  intergenic
35122746  (33789786)  to  35123112  (33789423)  =     376bp  at  75.8%  intergenic
35123699  (33788632)  to  35124018  (33788306)  =     335bp  at  74.6%  intergenic
35124215  (33787829)  to  35124454  (33787601)  =     243bp  at  74.9%  intergenic
35124502  (33787554)  to  35124817  (33787225)  =     333bp  at  78.4%  intergenic
35126998  (33784666)  to  35127099  (33784568)  =     103bp  at  70.9%  intergenic
35127111  (33784556)  to  35127640  (33784017)  =     542bp  at  84.5%  intergenic
35127796  (33783862)  to  35128041  (33783618)  =     249bp  at  73.9%  intergenic
35128751  (33782685)  to  35129182  (33782250)  =     443bp  at  71.1%  intergenic
35129233  (33782213)  to  35129477  (33781975)  =     248bp  at  75.8%  intergenic
35130121  (33781339)  to  35130621  (33780837)  =     512bp  at  78.5%  intergenic
```

Figure 4.22 Cebpa Downstream CNS5 region conservation identities shown on mouse against the human genome

After filtering out the repeating entries of the databases and arranging the TFs from the highest matrix similarity, we got a table of 163 TFBS predictions. The top 10 results were shown as in the table below (Table 4.9). Please refer to the appendix for the full table (Table A9).

Table 4.9 Top 10 TFBSs on *Mus musculus* Downstream CNS5 region

Matrix	strand	Matrix sim.	Sequence
V$SMAD/GC_SBE.01	(-)	1	ctgggCTCCag
V$NEUR/MASH1.01	(+)	1	cacacacCTGCtgtg
V$TAIP/CSRNP1.01	(+)	1	AGAGtga
V$TAIP/CSRNP3.01	(+)	1	aGAGTga
V$EVI1/MEL1.02	(+)	1	gcaagatGATGagtggt
V$FKHD/FOXP1_ES.01	(+)	1	gttcgaaAACAcaacca
V$ZF02/ZNF219.01	(-)	0.996	ccgcaCCCCcccccaaccttag
V$ZICF/ZIC2.02	(-)	0.995	tcacaCAGCaggtgt
V$HAND/TAL1_E2A.02	(-)	0.995	ggtcacaCAGCaggtgtgtgt
V$ZF11/ZBTB3.01	(+)	0.993	gccaGCCAggt

Mainly in this region, early growth response elements, MYB TF family, SOX proteins, hepatic leukemia factors, HOX family, CEBP family, KLF family, Forkhead box proteins, AP1 and AP2 families, and zinc finger proteins are common.

4.5.2 Homo sapiens Conserved Regions and Their TFBSs

Since *Mus musculus* conservation was aligned against the human genome, the figures will be the same with the same percentage of conservation identity and with the same sequence lengths among these two species. The TFBSs lists will differ in content since the sequences are not 100% identical.

4.5.2.1 Upstream CNS1 Region

This region consists of 2 intergenic sequences, 137bp and 170bp in length. The region was aligned mouse against the human genome (Figure 4.23). In MatInspector the two sequences were submitted separately to identify the TFBSs.

```
***** Conserved Regions - Mouse Dec. 2011 chr7 (Human Feb. 2009 chr19) *****

35112884 (33802302) to 35113011 (33802169) =    137bp at 67.2% intergenic
35113084 (33802036) to 35113253 (33801876) =    170bp at 70.6% intergenic
35116545 (33796496) to 35116763 (33796284) =    219bp at 74.9% intergenic
35117547 (33795177) to 35117803 (33794920) =    261bp at 75.5% intergenic
35122746 (33789786) to 35123112 (33789423) =    376bp at 75.8% intergenic
35123699 (33788632) to 35124018 (33788306) =    335bp at 74.6% intergenic
35124215 (33787829) to 35124454 (33787601) =    243bp at 74.9% intergenic
35124502 (33787554) to 35124817 (33787225) =    333bp at 78.4% intergenic
35126998 (33784666) to 35127099 (33784568) =    103bp at 70.9% intergenic
35127111 (33784556) to 35127640 (33784017) =    542bp at 84.5% intergenic
35127796 (33783862) to 35128041 (33783618) =    249bp at 73.9% intergenic
35128751 (33782685) to 35129182 (33782250) =    443bp at 71.1% intergenic
35129233 (33782213) to 35129477 (33781975) =    248bp at 75.8% intergenic
35130121 (33781339) to 35130621 (33780837) =    512bp at 78.5% intergenic
```

Figure 4.23 Cebpa Upstream CNS1 region conservation identities shown on mouse against the human genome

After filtering out the repeating entries of the databases and arranging the TFs from the highest matrix similarity, we got a table of 64 TFBS predictions. The top 10 results were shown as in the table below (Table 4.10). Please refer to the appendix for the full table (Table A10).

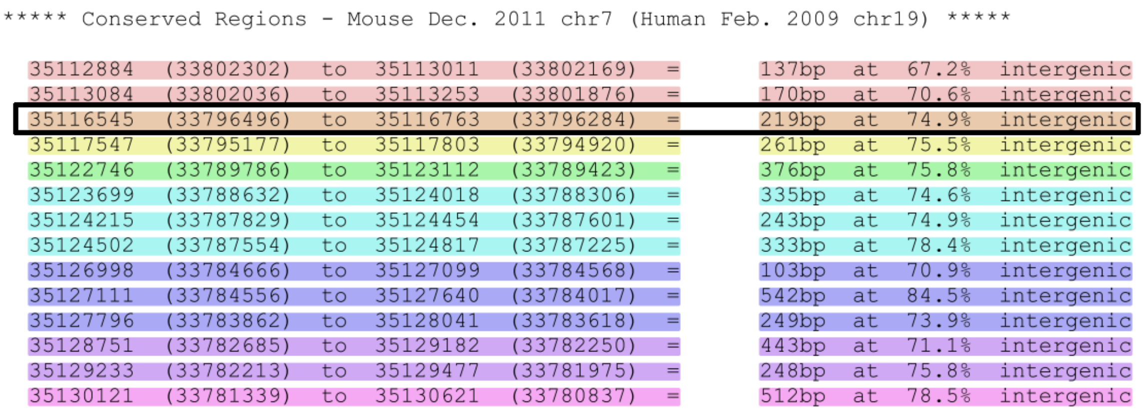
Table 4.10 Top 10 TFBSs on *Homo sapiens* Upstream CNS1 region

Matrix	matrix sim.	Sequence
V$MOKF/MOK2.02	1	gaggaaaatggggCCTTtttt
V$TALE/TGIF.01	1	taactgcctGTCAgtgg
V$NEUR/NGN_NEUROD.01	0.987	aatcCATCtgtgttc
V$IKRS/IK2.01	0.983	ctctGGGAatggg
V$ZFXY/ZFX.01	0.983	agGGCCtggag
V$SMAD/SMAD.01	0.976	cttGTCTgtcc
V$SORY/SOX6.01	0.974	tccctACAAaggttcttcttgct
V$AP2F/TCFAP2A.02	0.972	agcccCCTGggggcca
V$NFAT/NFAT.01	0.971	tagggaGGAAaatggggcc
V$HOXF/NANOG.01	0.97	ctctgggAATGggcagaac

Mainly in this region, hormone receptors, neural differentiation genes, STAT proteins, SOX proteins, TALE family, are common.

4.5.2.2 Upstream CNS2 Region

This region consists of 1 intergenic sequence, 219bp in length with 74.9% conservation identity. The region was aligned mouse against the human genome (Figure 4.24).



```
***** Conserved Regions - Mouse Dec. 2011 chr7 (Human Feb. 2009 chr19) *****

35112884 (33802302) to 35113011 (33802169) = 137bp at 67.2% intergenic
35113084 (33802036) to 35113253 (33801876) = 170bp at 70.6% intergenic
35116545 (33796496) to 35116763 (33796284) = 219bp at 74.9% intergenic
35117547 (33795177) to 35117803 (33794920) = 261bp at 75.5% intergenic
35122746 (33789786) to 35123112 (33789423) = 376bp at 75.8% intergenic
35123699 (33788632) to 35124018 (33788306) = 335bp at 74.6% intergenic
35124215 (33787829) to 35124454 (33787601) = 243bp at 74.9% intergenic
35124502 (33787554) to 35124817 (33787225) = 333bp at 78.4% intergenic
35126998 (33784666) to 35127099 (33784568) = 103bp at 70.9% intergenic
35127111 (33784556) to 35127640 (33784017) = 542bp at 84.5% intergenic
35127796 (33783862) to 35128041 (33783618) = 249bp at 73.9% intergenic
35128751 (33782685) to 35129182 (33782250) = 443bp at 71.1% intergenic
35129233 (33782213) to 35129477 (33781975) = 248bp at 75.8% intergenic
35130121 (33781339) to 35130621 (33780837) = 512bp at 78.5% intergenic
```

Figure 4.24 Cebpa Upstream CNS2 region conservation identities shown on mouse against the human genome

After filtering out the repeating entries of the databases and arranging the TFs from the highest matrix similarity, we got a table of 85 TFBS predictions. The top 10 results were shown as in the table below (Table 4.11). Please refer to the appendix for the full table (Table A11).

Table 4.11 Top 10 TFBSs on *Homo sapiens* Upstream CNS2 region

Matrix	strand	matrix sim.	Sequence
V$PLAG/PLAG1.02	(-)	1	aaGGGGgaacccttatctcccag
V$TAIP/CSRNP1.01	(-)	1	AGAGtgg
V$TAIP/CSRNP3.01	(-)	1	aGAGTgg
V$ZF02/ZKSCAN3.01	(+)	1	ctaggCCCCaccctccctcaccc
V$GATA/GATA4.01	(+)	0.998	gggaGATAagggt
V$KLFS/BKLF.01	(-)	0.996	agggaGGGTgggcctaga
V$MZF1/MZF1.02	(-)	0.995	agGGGGaaccc
V$NEUR/MASH1.01	(+)	0.994	ccttcacCTGCagtg
V$SF1F/SF1.01	(-)	0.993	tggcCAAGgccactg
V$GATA/GATA1.01	(+)	0.992	gggaGATAagggt

In this region mainly, GATA proteins, p53 family, SOX proteins, early growth response elements, estrogen receptors, hepatic nuclear factors, KLF family and zinc finger proteins are common.

4.5.2.3 Upstream CNS3 Region

This region consists of 1 intergenic sequence, 261bp in length with 75.5% conservation identity. The region was aligned mouse against the human genome (Figure 4.25).

```
***** Conserved Regions - Mouse Dec. 2011 chr7 (Human Feb. 2009 chr19) *****

35112884  (33802302)  to  35113011  (33802169)  =     137bp  at  67.2%  intergenic
35113084  (33802036)  to  35113253  (33801876)  =     170bp  at  70.6%  intergenic
35116545  (33796496)  to  35116763  (33796284)  =     219bp  at  74.9%  intergenic
35117547  (33795177)  to  35117803  (33794920)  =     261bp  at  75.5%  intergenic
35122746  (33789786)  to  35123112  (33789423)  =     376bp  at  75.8%  intergenic
35123699  (33788632)  to  35124018  (33788306)  =     335bp  at  74.6%  intergenic
35124215  (33787829)  to  35124454  (33787601)  =     243bp  at  74.9%  intergenic
35124502  (33787554)  to  35124817  (33787225)  =     333bp  at  78.4%  intergenic
35126998  (33784666)  to  35127099  (33784568)  =     103bp  at  70.9%  intergenic
35127111  (33784556)  to  35127640  (33784017)  =     542bp  at  84.5%  intergenic
35127796  (33783862)  to  35128041  (33783618)  =     249bp  at  73.9%  intergenic
35128751  (33782685)  to  35129182  (33782250)  =     443bp  at  71.1%  intergenic
35129233  (33782213)  to  35129477  (33781975)  =     248bp  at  75.8%  intergenic
35130121  (33781339)  to  35130621  (33780837)  =     512bp  at  78.5%  intergenic
```

Figure 4.25 Cebpa Upstream CNS3 region conservation identities shown on mouse against the human genome

After filtering out the repeating entries of the databases and arranging the TFs from the highest matrix similarity, we got a table of 136 TFBS predictions. The top 10 results were shown as in the table below (Table 4.12). Please refer to the appendix for the full table (Table A12).

Table 4.12 Top 10 TFBSs on *Homo sapiens* Upstream CNS3 region

Matrix	strand	matrix sim.	Sequence
V$WHNF/WHN.01	(-)	0.986	ttgACGCtttt
V$ZTRE/ZTRE.04	(-)	0.983	acgGGAGggagtaaaaa
V$FKHD/HFH3.01	(-)	0.981	gttaaatAAATatcccg
V$FKHD/FREAC7.01	(-)	0.976	agctcaTAAAcaagata
V$FKHD/HFH8.01	(-)	0.961	agctcatAAACaagata
V$MEF2/MEF2A.01	(+)	0.96	cgtcaagctaaAAATatcttgtt
V$BCDF/CRX.01	(-)	0.958	cgcttTAATcaagttcc
V$HAND/TH1E47.01	(-)	0.952	gagtaaaaaCCAGatacgtta
V$HOMF/HHEX.01	(-)	0.952	gtcccagtgatTAATagcc
V$FKHD/HNF3B.01	(-)	0.95	gttaaataAATAtcccg

Mainly in this region, homeobox proteins, AP1R TFs, CART family, Forkhead box proteins, HOX proteins, IRF proteins, LIM homeobox, Myocyte specific enhancer factors, hepatic nuclear factors, PAX proteins are common.

4.5.2.4 Downstream CNS1 Region

This region consists of 1 intergenic sequence, 376bp in length with 75.8% conservation identity. The region was aligned mouse against the human genome (Figure 4.26).

```
***** Conserved Regions - Mouse Dec. 2011 chr7 (Human Feb. 2009 chr19) *****

35112884 (33802302) to 35113011 (33802169) =  137bp at 67.2% intergenic
35113084 (33802036) to 35113253 (33801876) =  170bp at 70.6% intergenic
35116545 (33796496) to 35116763 (33796284) =  219bp at 74.9% intergenic
35117547 (33795177) to 35117803 (33794920) =  261bp at 75.5% intergenic
35122746 (33789786) to 35123112 (33789423) =  376bp at 75.8% intergenic
35123699 (33788632) to 35124018 (33788306) =  335bp at 74.6% intergenic
35124215 (33787829) to 35124454 (33787601) =  243bp at 74.9% intergenic
35124502 (33787554) to 35124817 (33787225) =  333bp at 78.4% intergenic
35126998 (33784666) to 35127099 (33784568) =  103bp at 70.9% intergenic
35127111 (33784556) to 35127640 (33784017) =  542bp at 84.5% intergenic
35127796 (33783862) to 35128041 (33783618) =  249bp at 73.9% intergenic
35128751 (33782685) to 35129182 (33782250) =  443bp at 71.1% intergenic
35129233 (33782213) to 35129477 (33781975) =  248bp at 75.8% intergenic
35130121 (33781339) to 35130621 (33780837) =  512bp at 78.5% intergenic
```

Figure 4.26 Cebpa Downstream CNS1 region conservation identities shown on mouse against the human genome

After filtering out the repeating entries of the databases and arranging the TFs from the highest matrix similarity, we got a table of 148 TFBS predictions. The top 10 results were shown as in the table below (Table 4.13). Please refer to the appendix for the full table (Table A13).

Table 4.13 Top 10 TFBSs on *Homo sapiens* Downstream CNS1 region

Matrix	strand	matrix sim.	Sequence
V$PLAG/PLAG1.02	(+)	1	gaGGGGgtggggcgaccctcctg
V$TAIP/CSRNP1.01	(-)	1	AGAGtgg
V$TAIP/CSRNP3.01	(-)	1	aGAGTgg
V$ZF02/ZKSCAN3.01	(-)	1	ggtcgCCCCacccccctcccgtcc
V$ZF02/ZNF300.01	(+)	1	cctttccCCCCagaacaggtcct
V$MZF1/MZF1.02	(-)	0.998	gaGGGGaagga
V$ZFXY/ZFX.01	(+)	0.993	gaGGCCtgggg
V$KLFS/IKLF.01	(+)	0.991	ggaggGGGTggggcgaccc
V$KLFS/KLF2.01	(+)	0.991	ggaggGGGTggggcgaccc
V$SMAD/SMAD3.01	(+)	0.991	tgtGTCTggcc

In this region mainly, AP2 family TFs, E2F family TFs, EBOX factors, early growth response elements, ETS family, GLIS family zinc finger proteins, KLF family, SMAD TFs, SP1 family, and zinc finger proteins are common.

4.5.2.5 Downstream CNS2 Region

This region consists of 3 intergenic sequences, 335bp, 243bp and 333bp in length. The region was aligned mouse against the human genome (Figure 4.27). In MatInspector the three sequences were submitted separately to identify the TFBSs.

```
***** Conserved Regions - Mouse Dec. 2011 chr7 (Human Feb. 2009 chr19) *****

35112884 (33802302) to 35113011 (33802169) =    137bp at 67.2% intergenic
35113084 (33802036) to 35113253 (33801876) =    170bp at 70.6% intergenic
35116545 (33796496) to 35116763 (33796284) =    219bp at 74.9% intergenic
35117547 (33795177) to 35117803 (33794920) =    261bp at 75.5% intergenic
35122746 (33789786) to 35123112 (33789423) =    376bp at 75.8% intergenic
35123699 (33788632) to 35124018 (33788306) =    335bp at 74.6% intergenic
35124215 (33787829) to 35124454 (33787601) =    243bp at 74.9% intergenic
35124502 (33787554) to 35124817 (33787225) =    333bp at 78.4% intergenic
35126998 (33784666) to 35127099 (33784568) =    103bp at 70.9% intergenic
35127111 (33784556) to 35127640 (33784017) =    542bp at 84.5% intergenic
35127796 (33783862) to 35128041 (33783618) =    249bp at 73.9% intergenic
35128751 (33782685) to 35129182 (33782250) =    443bp at 71.1% intergenic
35129233 (33782213) to 35129477 (33781975) =    248bp at 75.8% intergenic
35130121 (33781339) to 35130621 (33780837) =    512bp at 78.5% intergenic
```

Figure 4.27 Cebpa Downstream CNS2 region conservation identities shown on mouse against the human genome

After filtering out the repeating entries of the databases and arranging the TFs from the highest matrix similarity, we got a table of 283 TFBS predictions. The top 10 results were shown as in the table below (Table 4.14). Please refer to the appendix for the full table (Table A14).

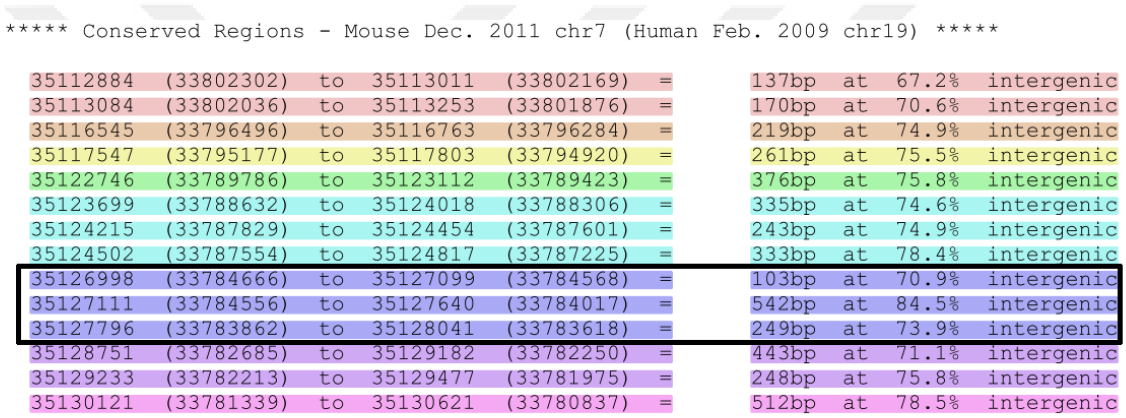
Table 4.14 Top 10 TFBSs on *Homo sapiens* Downstream CNS2 region

Matrix	strand	matrix sim.	Sequence
V$CREB/CREB2.01	(+)	1	agtgttTGACgtaagctgagg
V$FKHD/FOXP1_ES.01	(-)	1	aaataaaAACAaaccga
V$KLFS/KLF2.01	(+)	1	ggcagGGGTggggctgtga
V$NEUR/MASH1.01	(+)	1	ctccccacCTGCtcct
V$SMAD/SMAD.01	(-)	1	tttGTCTggct
V$ZF02/ZKSCAN3.01	(+)	1	tctggCCCCacccactaggaaca
V$ZTRE/ZTRE.04	(-)	1	ctgGGAGggcagacac
V$HOMF/HHEX.01	(-)	0.998	atctggcttttTAATtgct
V$MZF1/MZF1.02	(+)	0.998	gaGGGGaaggg
V$MOKF/MOK2.02	(-)	0.997	ggtgcactgtgggCCTTgcaa

In this region mainly, HOX proteins, SMAD proteins, NK6 homeobox proteins, hormone receptors, Distal-less homeobox proteins, v-Myb, STAT proteins, GATA proteins, myc-associated zinc finger proteins, MYOD family proteins, SOX proteins, HOX proteins, CREB family, SP1 family, CEBP family, KLF family, Forkhead box proteins, AP2 family TFs, zinc finger domains are common.

4.5.2.6 Downstream CNS3 Region

This region consists of 3 intergenic sequences, 103bp, 542bp and 249bp in length. The region was aligned mouse against the human genome (Figure 4.28). In MatInspector the three sequences were submitted separately to identify the TFBSs.


***** Conserved Regions - Mouse Dec. 2011 chr7 (Human Feb. 2009 chr19) *****

35112884	(33802302)	to	35113011	(33802169)	=	137bp	at	67.2%	intergenic
35113084	(33802036)	to	35113253	(33801876)	=	170bp	at	70.6%	intergenic
35116545	(33796496)	to	35116763	(33796284)	=	219bp	at	74.9%	intergenic
35117547	(33795177)	to	35117803	(33794920)	=	261bp	at	75.5%	intergenic
35122746	(33789786)	to	35123112	(33789423)	=	376bp	at	75.8%	intergenic
35123699	(33788632)	to	35124018	(33788306)	=	335bp	at	74.6%	intergenic
35124215	(33787829)	to	35124454	(33787601)	=	243bp	at	74.9%	intergenic
35124502	(33787554)	to	35124817	(33787225)	=	333bp	at	78.4%	intergenic
35126998	(33784666)	to	35127099	(33784568)	=	103bp	at	70.9%	intergenic
35127111	(33784556)	to	35127640	(33784017)	=	542bp	at	84.5%	intergenic
35127796	(33783862)	to	35128041	(33783618)	=	249bp	at	73.9%	intergenic
35128751	(33782685)	to	35129182	(33782250)	=	443bp	at	71.1%	intergenic
35129233	(33782213)	to	35129477	(33781975)	=	248bp	at	75.8%	intergenic
35130121	(33781339)	to	35130621	(33780837)	=	512bp	at	78.5%	intergenic

Figure 4.28 Cebpa Downstream CNS3 region conservation identities shown on mouse against the human genome

After filtering out the repeating entries of the databases and arranging the TFs from the highest matrix similarity, we got a table of 476 TFBS predictions. The top 10 results were shown as in the table below (Table 4.15). Please refer to the appendix for the full table (Table A15).

Table 4.15 Top 10 TFBSs on *Homo sapiens* Downstream CNS3 region

Matrix	strand	matrix sim.	Sequence
V$CART/S8.01	(+)	1	ctgttTAATtggccttagccc
V$FKHD/FOXP1_ES.01	(+)	1	tatgcaaAACAagtagc
V$FKHD/HNF3.01	(-)	1	gtgaagcAAACacacat
V$HOMF/MSX.01	(+)	1	agctgttTAATtggcctta
V$MZF1/MZF1.01	(+)	1	gtGGGGatggg
V$PLAG/PLAG1.02	(+)	1	aaGGGGgcgtcagggctgggtgt
V$SMAD/GC_SBE.01	(+)	1	cttggCTCCag

V$TAIP/CSRNP1.01	(-)	1	AGAGtga
V$TAIP/CSRNP3.01	(-)	1	aGAGTga
V$TALE/TGIF.01	(+)	1	actgagtctGTCAgttt

Mainly in this region, CEBP family, PPAR family, Distal-less homeobox factors, GRHL family, GATA family, HOX family, zinc finger proteins, HOM family TFs, CART family, ESSR family, nuclear factors, p53 family, HOX family, IRF family, SOX family, ETF family TFs, hepatic nuclear factors, RORA family, NFKB family, TALE family, AML family, KLF family, AP1R family, CAAT family, HMG box TFs, Forkhead box proteins, LIM homeobox proteins, OCT1 family, and PAX family are common.

4.5.2.7 Downstream CNS4 Region

This region consists of 2 intergenic sequences, 443bp and 248bp in length. The region was aligned mouse against the human genome (Figure 4.29). In MatInspector the three sequences were submitted separately to identify the TFBSs.

```
***** Conserved Regions - Mouse Dec. 2011 chr7 (Human Feb. 2009 chr19) *****

 35112884  (33802302)  to  35113011  (33802169)  =      137bp  at  67.2%  intergenic
 35113084  (33802036)  to  35113253  (33801876)  =      170bp  at  70.6%  intergenic
 35116545  (33796496)  to  35116763  (33796284)  =      219bp  at  74.9%  intergenic
 35117547  (33795177)  to  35117803  (33794920)  =      261bp  at  75.5%  intergenic
 35122746  (33789786)  to  35123112  (33789423)  =      376bp  at  75.8%  intergenic
 35123699  (33788632)  to  35124018  (33788306)  =      335bp  at  74.6%  intergenic
 35124215  (33787829)  to  35124454  (33787601)  =      243bp  at  74.9%  intergenic
 35124502  (33787554)  to  35124817  (33787225)  =      333bp  at  78.4%  intergenic
 35126998  (33784666)  to  35127099  (33784568)  =      103bp  at  70.9%  intergenic
 35127111  (33784556)  to  35127640  (33784017)  =      542bp  at  84.5%  intergenic
 35127796  (33783862)  to  35128041  (33783618)  =      249bp  at  73.9%  intergenic
 35128751  (33782685)  to  35129182  (33782250)  =      443bp  at  71.1%  intergenic
 35129233  (33782213)  to  35129477  (33781975)  =      248bp  at  75.8%  intergenic
 35130121  (33781339)  to  35130621  (33780837)  =      512bp  at  78.5%  intergenic
```

Figure 4.29 Cebpa Downstream CNS4 region conservation identities shown on mouse against the human genome

After filtering out the repeating entries of the databases and arranging the TFs from the highest matrix similarity, we got a table of 294 TFBS predictions. The top 10 results were shown as in the table below (Table 4.16). Please refer to the appendix for the full table (Table A16).

Table 4.16 Top 10 TFBSs on *Homo sapiens* Downstream CNS4 region

Matrix	strand	matrix sim.	Sequence
V$NEUR/MASH1.01	(+)	1	gtgccacCTGCatgt
V$IKRS/LYF1.01	(+)	1	gtcTGGGagaggt
V$MZF1/MZF1.01	(+)	1	gtGGGGacgat
V$NKXH/NKX25.01	(+)	1	cccctcAAGTgcaattccc
V$TALE/TGIF.01	(+)	1	ctgaaatgtGTCAgagg
V$ZFXY/ZFX.01	(-)	1	ggGGCCtggcc
V$ZF02/ZKSCAN3.01	(-)	1	atcgtCCCCacctcccctcggct
V$SORY/SOX6.01	(-)	0.999	cttgaACAAaggagcacaggcct
V$NKXH/TTF1.01	(+)	0.997	cccctCAAGtgcaattccc
V$EBOX/MNT.01	(+)	0.996	gcagcCACGtgccacct

In this region, mainly, AP1 and AP2 family, CAAT family, EBOX family, ETS family, BHLH family, AhR nuclear translocator homodimers, HOX proteins, KLF family, SOX family, and zinc finger proteins are common.

4.5.2.8 Downstream CNS5 Region

This region consists of one intergenic sequence, 512bp in length with 78.5% conservation identity. The region was aligned mouse against the human genome (Figure 4.30).

***** Conserved Regions - Mouse Dec. 2011 chr7 (Human Feb. 2009 chr19) *****

35112884	(33802302)	to	35113011	(33802169)	=	137bp	at	67.2%	intergenic
35113084	(33802036)	to	35113253	(33801876)	=	170bp	at	70.6%	intergenic
35116545	(33796496)	to	35116763	(33796284)	=	219bp	at	74.9%	intergenic
35117547	(33795177)	to	35117803	(33794920)	=	261bp	at	75.5%	intergenic
35122746	(33789786)	to	35123112	(33789423)	=	376bp	at	75.8%	intergenic
35123699	(33788632)	to	35124018	(33788306)	=	335bp	at	74.6%	intergenic
35124215	(33787829)	to	35124454	(33787601)	=	243bp	at	74.9%	intergenic
35124502	(33787554)	to	35124817	(33787225)	=	333bp	at	78.4%	intergenic
35126998	(33784666)	to	35127099	(33784568)	=	103bp	at	70.9%	intergenic
35127111	(33784556)	to	35127640	(33784017)	=	542bp	at	84.5%	intergenic
35127796	(33783862)	to	35128041	(33783618)	=	249bp	at	73.9%	intergenic
35128751	(33782685)	to	35129182	(33782250)	=	443bp	at	71.1%	intergenic
35129233	(33782213)	to	35129477	(33781975)	=	248bp	at	75.8%	intergenic
35130121	(33781339)	to	35130621	(33780837)	=	512bp	at	78.5%	intergenic

Figure 4.30 Cebpa Downstream CNS5 region conservation identities shown on mouse against the human genome

After filtering out the repeating entries of the databases and arranging the TFs from the highest matrix similarity, we got a table of 195 TFBS predictions. The top 10 results were shown as in the table below (Table 4.17). Please refer to the appendix for the full table (Table A17).

Table 4.17 Top 10 TFBSs on *Homo sapiens* Downstream CNS5 region

Matrix	strand	matrix sim.	Sequence
V$MZF1/MZF1.01	(+)	1	gtGGGGagcgt
V$NEUR/MASH1.01	(+)	1	cacacacCTGCtgtg
V$SMAD/GC_SBE.01	(-)	1	ctgggCTCCag
V$TALE/TGIF.01	(-)	1	gacctggctGTCAaaga
V$ZF02/ZKSCAN3.01	(-)	1	acgctCCCCaccttggtggccag
V$EVI1/MEL1.02	(-)	0.996	ggtgtgtGATGaggcca
V$HAND/TAL1_E2A.02	(-)	0.995	ggtcacaCAGCaggtgtgtga
V$ZICF/ZIC2.02	(-)	0.995	tcacaCAGCaggtgt
V$ZF5F/ZF5.01	(+)	0.994	gtgtgtGCGCcctgg
V$ZF11/ZBTB3.01	(+)	0.993	gacaGCCAggt

Mainly in this region, AP1 and AP2 family, CEBP family, E2F family, GRHL family, GLI family, hormone receptors, KLF family, MYOD family, hepatic nuclear factors, SOX family, STAT family, and ZIC family are common.

4.6 The TFs that have correlation with Cebpa

Three of the GEO data sets that were from the cells/tissues to have high expression of Cebpa were selected and analyzed. The TFs with the highest correlation with Cebpa with correlation p-values < 0,05 were filtered and selected from the results. The top 10 results of each dataset are below with their corresponding information (Table 4.18, 4.19, 4.20). Please refer to appendix for the full tables (Table A18, A19, A20).

Table 4.18 Top 10 GSE8733 TFs with highest correlation with Cebpa

ProbeIDs	Gene1	ProbeIDs	Gene2	Corr_pVal
1417728_at	**Mbd3**	1418982_at	**Cebpa**	0.000558504
1440156_s_at	**Tox2**	1418982_at	**Cebpa**	0.001650935
1420500_at	**Dnajc1**	1418982_at	**Cebpa**	0.003004278
1437567_at	**Zfp772**	1418982_at	**Cebpa**	0.005147655
1423416_at	**Smarcc1**	1418982_at	**Cebpa**	0.005410294
1423260_at	**Id4**	1418982_at	**Cebpa**	0.00652402
1436364_x_at	**Nfix**	1418982_at	**Cebpa**	0.007991639
1455790_at	**E2f2**	1418982_at	**Cebpa**	0.008238733
1433508_at	**Klf6**	1418982_at	**Cebpa**	0.008363104
1420952_at	**Son**	1418982_at	**Cebpa**	0.008567128

Table 4.19 Top 10 GSE35160 TFs with highest correlation with Cebpa

	ProbeIDs	Gene1	ProbeIDs	Gene2	Corr_pVal
1	1417621_at	**Nfatc1**	1418982_at	**Cebpa**	0.001004402
2	1428662_a_at	**Hopx**	1418982_at	**Cebpa**	0.00238143
3	1449732_at	**Zscan21**	1418982_at	**Cebpa**	0.004832033
4	1443591_at	**Dlx4**	1418982_at	**Cebpa**	0.007558002
5	1417127_at	**Msx1**	1418982_at	**Cebpa**	0.007571702
6	1429610_a_at	**Zfp511**	1418982_at	**Cebpa**	0.009244521
7	1418939_at	**Hlx**	1418982_at	**Cebpa**	0.012850828
8	1418241_at	**Usf2**	1418982_at	**Cebpa**	0.014084548
9	1422865_at	**Runx1**	1418982_at	**Cebpa**	0.014296301
10	1448877_at	**Dlx2**	1418982_at	**Cebpa**	0.016758872

Table 4.20 Top 10 GSE10175 TFs with highest correlation with Cebpa

	ProbeIDs	Gene1	ProbeIDs	Gene2	Corr_pVal
1	1422102_a_at	**Stat5b**	1418982_at	**Cebpa**	0.001391707
2	1452174_at	**Srebf2**	1418982_at	**Cebpa**	0.001526449
3	1426790_at	**Ssrp1**	1418982_at	**Cebpa**	0.001909138
4	1438325_at	**Mecom**	1418982_at	**Cebpa**	0.002082493
5	1455210_at	**Zhx2**	1418982_at	**Cebpa**	0.002772165
6	1417549_at	**Zfp68**	1418982_at	**Cebpa**	0.002941861
7	1420973_at	**Arid5b**	1418982_at	**Cebpa**	0.002952893
8	1454773_at	**Rxra**	1418982_at	**Cebpa**	0.004009393
9	1434246_at	**L3mbtl3**	1418982_at	**Cebpa**	0.004431938
10	1419150_at	**Myf6**	1418982_at	**Cebpa**	0.0050273

4.7 Identification of Possible Enhancer Region of Cebpa

The histone modification markers were investigated to identify the possible enhancer region of Cebpa gene. In the downstream of Cebpa where the "Downstream CNS3" region is located between "ch7:35126998-33783618" sites on mouse mm9, histone modification markers that are tell-tale signs of enhancer sites were detected as shown in the figure below (Figure 4.31);

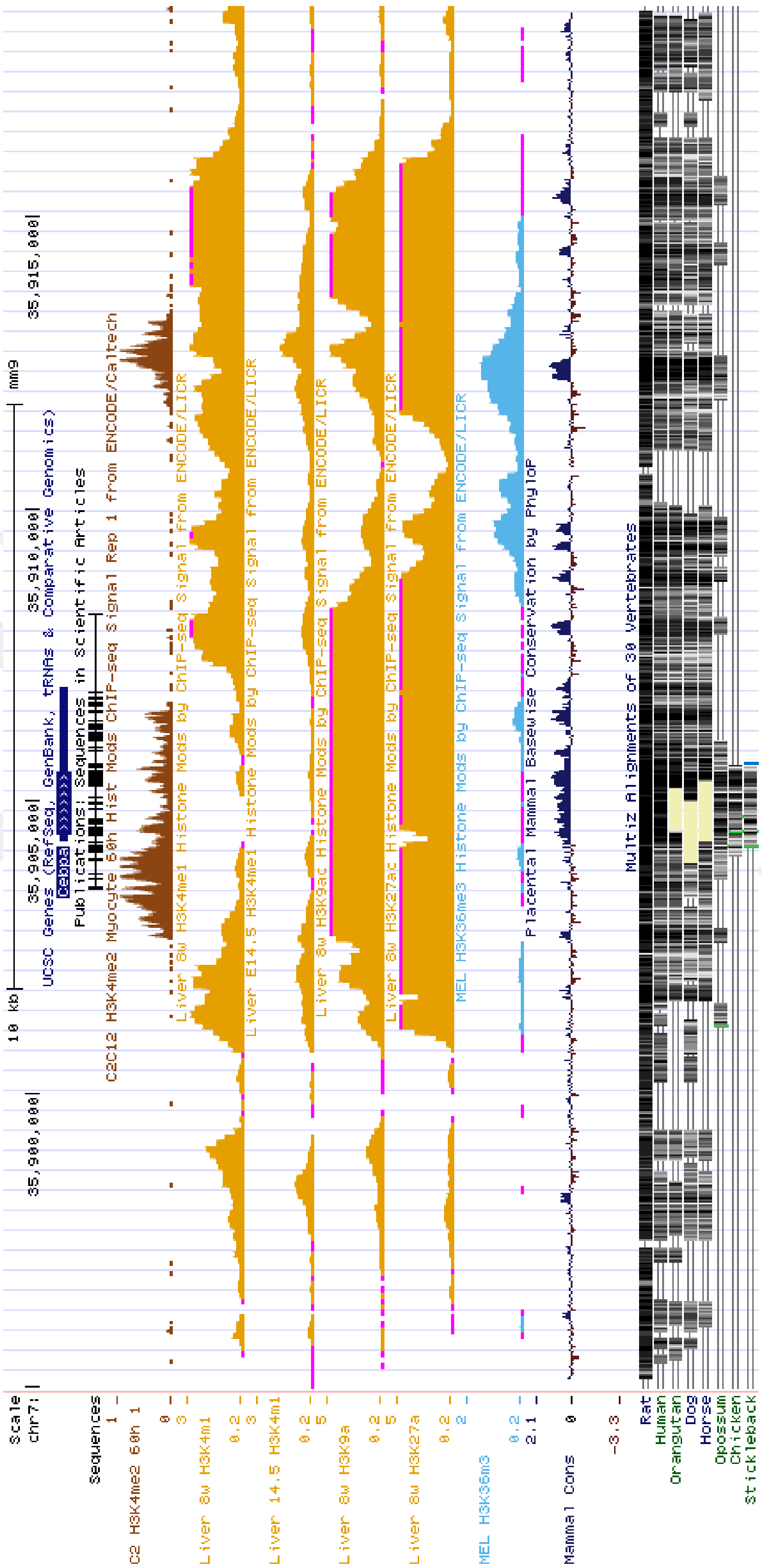


Figure 4.31 Histone markers on CNS regions of Cebpa

Throughout the whole ~20kb region of Cebpa, H3K9ac, H3K4me2 and H3K27ac modifications were present. H3K4m1 was active in the whole region except the gene body. So now when we look at the Downstream CNS3 region (Figure 4.32),

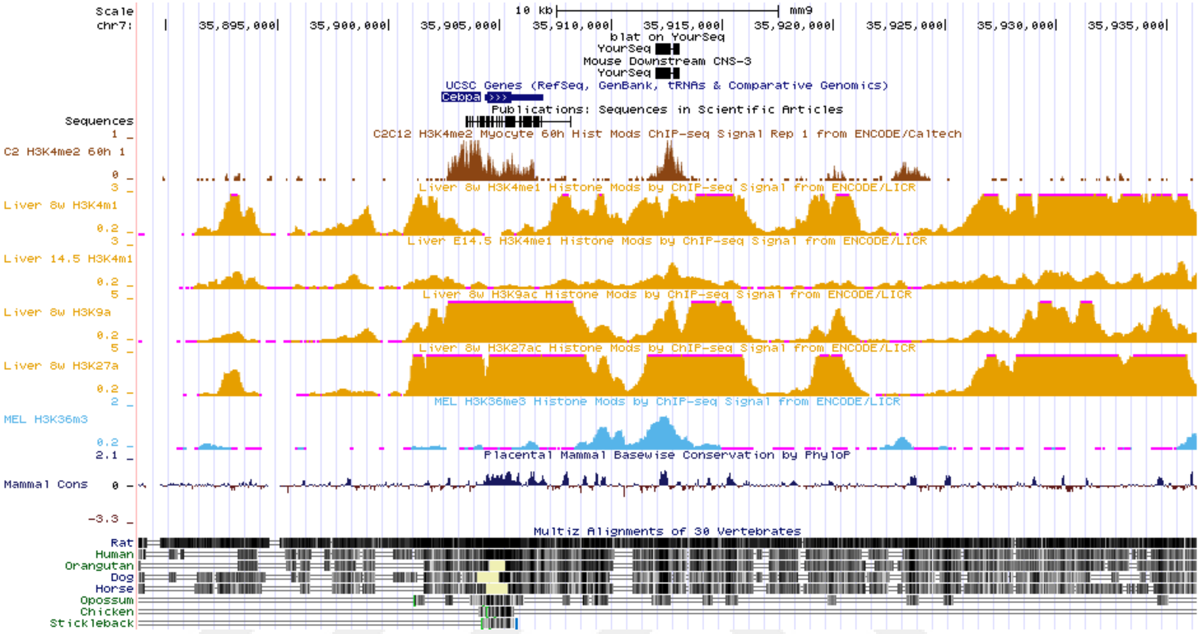


Figure 4.32 Histone markers visible on Cebpa Downstream CNS3 region

we see that H3K4me2, H3K4me1, H3K9a, H3K27a and H3K36m3 were active in this region. If we scan the literature for what they are responsible for, we can summarize their roles in this specific region.

H3K4me2 is generally associated with transcriptional activity (Kooistra and Helin, 2012). It overlaps on average 90% of TF binding sites, and those regions with stronger H3K4me2 signals have higher TF binding sites than the weaker ones. Moreover, with H3K27ac, it can significantly reduce false-positive TF binding site predictions (Wang, Li, and Hu, 2014).

H3K4me1 modification at promoters is linked to gene repression, but in other distal regions it can signify enhancers (Cheng et al., 2014), and strong levels of H3K4me1 along with H3K27ac and as well as recruitment of RNA polymerase II (Pol II) and the histone acetyl-transferase, p300 distinguish the presence of enhancers (Blum et al., 2012; Visel et al., 2008; Heintzman et al., 2007; Creyghton et al., 2011; Heintzman et al., 2009; Rada-Iglesias et al., 2011; Zentner et al., 2011).

H3K9ac, along with H3K27ac is associated with enhancers and promoters of active genes (Gates et al., 2017).

H3K36me3 modification is deposited on histones as they are displaced by RNA polymerase II during transcription. H3K36me3 then serves as a mark for HDACs to bind and deacetylate the histones, preventing run-away transcription in the wake of Pol II (Carrozza et al., 2005; Joshi and Struhl, 2005) and H3K36 is used in conjunction with H3K9 to repressive aberrant transcription (Bartke et al., 2010).

H3K27ac is one of the most well-known active enhancer histone markers (Heintzman et al., 2007). Also, recently, it is found to distinguish active enhancers from inactive/poised enhancer elements containing H3K4me1 alone (Creyghton et al., 2010). Also, H3K27ac is associated with active promoters in mammalian cells (Wang et al., 2008).

Summarizing all of the information available on these chromatin data, Downstream CNS3 region of Cebpa is a strong candidate of a possible enhancer region.

4.8 Identification of Possible Regulators of Cebpa

The results from the database investigation for cis-acting elements and trans-acting factors, common TFs were identified from the lists of results to address the question of what might be the possible regulators of Cebpa.

Firstly, conserved CNS regions from all the selected species with conserved binding sequences were analyzed. Only the ones that show conservation among the species on the same CNS region were selected. In 8 CNS regions, 55 TFs in total were found to be in common. (Table 4.21)

Table 4.21 TFs that are in common in selected species in Cebpa CNS regions

Cebpa CNS region	Number of TFs	TF IDs
Upstream CNS1	1	TCFAP2A
Upstream CNS2	6	NFIB, SOX11, TCFCP2L1, SOX2, GATA6, HNF1B
Upstream CNS3	2	LMX1A, IRF7
Downstream CNS1	2	KLF5, SPIB
Downstream CNS2	14	GLI3, GATA6, CEBPB, SOX11, KLF5, SALL2, TCFCP2L1, HHEX, SOX2, TCF7, STAT5A, SALL1, CEBPD, KLF4
Downstream CNS3	18	POU2F3, GRHL3, MEF2C, JARID2, DBP, CEBPB, CEBPA, CEBPD, CDX2, HMGA1, ISX, ARNT, HNF3B, PBX1, TCF7,

		GRHL2, SPI1, PPARG, HLXB9
Downstream CNS4	5	JUND, NFIB, TGIF, FOSL1, LMX1A
Downstream CNS5	6	CEBPB, ZIC3, ESRRA, FOSL1, STAT5A, TCFAP2A

Secondly, on all of the GEO datasets 296 TFs in total were analyzed for conservation among Cebpa CNS regions both for location and for sequence conservation. 1 of the GEO TFs was found to be in common (Figure 4.33 and Table 4.22).

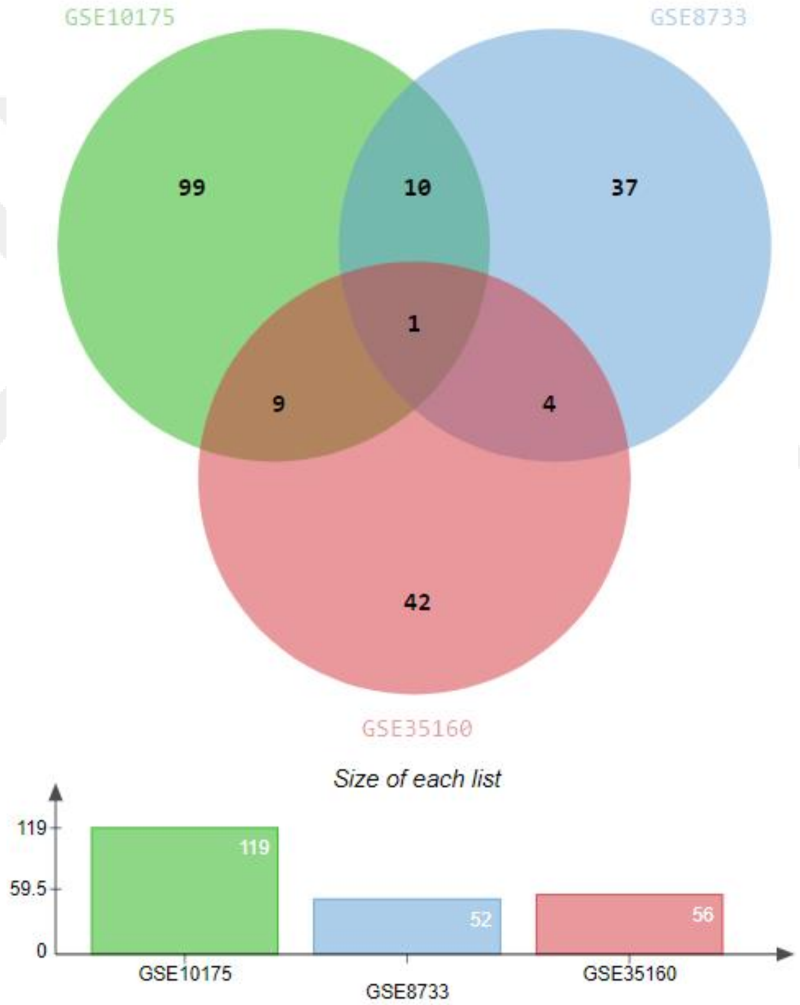


Figure 4.33 Venn diagram of TFs from GEO datasets with Cebpa correlation.

The results were analyzed and the TFs that were common among any of the two GEO datasets were identified (Table 4.22).

Table 4.22 TFs which are common in any of the two GEO datasets

GSE10175-GSE35160	**GSE35160-GSE8733**	**GSE10175-GSE8733**
Hinfp	Hopx	E2f2
Hlx	Jarid2	Foxo4
Jarid2	Klf5	Jarid2
Klf7	Zfp410	Kmt2c
Nfatc1	Zfp706	Mecp2
Tsc22d1		Son
Zfp263		Sp1
Zfp282		Stat5b
Zfp53		Tcf12
Zfp930		Trp63
		Zbtb20

The common TFs were checked for their sequence conservation identities. The ones that are common in more than three species in the same position and the same CNS region were selected. Below are the sequence information with binding sites alongside with JASPAR binding site predictions (Figures 4.34 to 4.44)

```
KLF5 – Upstream CNS-2
Mouse    AGCTAGGGAGGGGCCTAGA
Human    AGGGAGGGTGGGGCCTAGA
Chimp    AGGGAGGGTGGGGCCTAGA
Dog      TGCTGGGGAGGGGCCTAGA
IDENTITY  *   *** **********

KLF5 – Downstream CNS-1
Mouse    AACGGGGGAGGGGGCCCTC
Human    AGGACGGGAGGGGGTGGGG
Chimp    AGGACGGGAGGGGGTGGGG
Rhesus   AGGACGGGAGGGGGTGGGG
Dog      GGAGAGGGCGGGGGCCGGG
Cow      AATGGGGGAGGGGCGCCCC
IDENTITY      *** ****
```

Figure 4.34 KLF5 transcription factor conserved binding sites among species on Upstream CNS2 and Downstream CNS1 regions of Cebpa

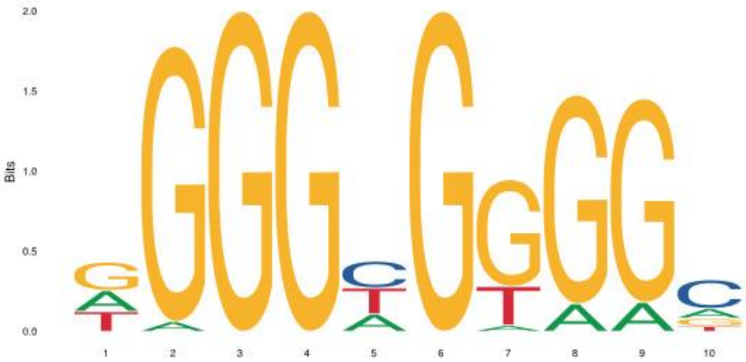


Figure 4.35 KLF5 transcription factor binding site profile from JASPAR

```
SP1 - Downstream CNS-1
Mouse     AACGGGGGAGGGGCCC
Human     GGAGGGGGTGGGGCGAC
Chimp     GGAGGGGGTGGGGCGAC
Rhesus    GGAGGGGGTGGGGCGAC
Dog       GGCCGGGGCGGGAAGGG
Cow       AATGGGGGAGGGGCGCC
IDENTITY      **** ***
```

Figure 4.36 SP1 transcription factor conserved binding sites among species on Downstream CNS1 region of Cebpa

Figure 4.37 SP1 transcription factor binding site profile from JASPAR

```
KLF7 – Downstream CNS-1
Human     GGAGGGGGTGGGGCGACCC
Chimp     GGAGGGGGTGGGGCGACCC
Rhesus    GGAGGGGGTGGGGCGACCC
Dog       GGAGAGGGCGGGGGCCGGG
IDENTITY  **** *** ****
```

```
KLF7 – Downstream CNS-2
Mouse     TTAATGGGTGGGGCCAGAA
Human     CTAGTGGGTGGGGCCAGAA
Chimp     CTAGTGGGTGGGGCCAGAA
Rhesus    CTAGTGGGTGGGGCCAGAA
Dog       CTCATGGGTGGGGCCAGAA
Cow       CAGATGGGTGGGGCCAGAA
IDENTITY      ***************
```

Figure 4.38 KLF7 transcription factor conserved binding sites among species on Downstream CNS1 and Downstream CNS2 regions of Cebpa

```
ZFP410 – Downstream CNS-2
Mouse     TGTTCTGGGATGGGG
Human     TGCTCTGGGATGGGG
Chimp     TGCTCTGGGATGGGG
Rhesus    TGCTCTGGGATGGGG
Dog       TGCTCTGGGATGGGG
Cow       TGCTCTGGGATGGGG
IDENTITY  ** ************
```

Figure 4.39 ZFP410 transcription factor conserved binding sites among species on Downstream CNS2 region of Cebpa

```
KLF5 – Downstream CNS-2
Mouse     TTAATGGGTGGGGCCAGAA
Human     CTAGTGGGTGGGGCCAGAA
Chimp     CTAGTGGGTGGGGCCAGAA
Rhesus    CTAGTGGGTGGGGCCAGAA
Dog       CTCATGGGTGGGGCCAGAA
Cow       CAGATGGGTGGGGCCAGAA
IDENTITY      ***************
```

Figure 4.40 KLF5 transcription factor conserved binding sites among species on Downstream CNS2 region of Cebpa

Figure 4.41 KLF5 transcription factor binding site profile from JASPAR

```
TCF12 – Downstream CNS-5
Mouse      ACACAGCAGGTGTGTGT
Human      ACACAGCAGGTGTGTGA
Chimp      ACACAGCAGGTGTGTGA
Rhesus     ACACAGCAGGTGTGTGT
Dog        ACACAGCAGGTGTGCGG
IDENTITY   ************** *
```

Figure 4.42 TCF12 transcription factor conserved binding sites among species on Downstream CNS5 region of Cebpa

Figure 4.43 TCF12 transcription factor binding site profile from JASPAR

```
JARID2 – Downstream CNS-3
Mouse         TTTGATTTTAATCTGATATTT
Human         TTTGATTTTAATCTGATATTT
Chimp         TTTGATTTTAATCTGATATTT
Rhesus        TTTGATTTTAATCTGATATTT
Dog           TTTGATTTTAATCTGATATTT
Cow           TTTGATTTTAATCTGATATTT
IDENTITY      *********************
```

Figure 4.44 JARID2 transcription factor conserved binding sites among species on Downstream CNS3 region of Cebpa

In those GEO analysis results, one TF were found to be common in all datasets. The correlation values of the TF were shown below with corresponding GSE accession number (Table 4.23).

Table 4.23 TFs which are common in all three GEO datasets

GEO Acces.	ProbeIDs	Gene1	ProbeIDs	Gene2	Correlation_pVal
GSE10175	1450710_at	**Jarid2**	1418982_at	**Cebpa**	0.0259607085121916
GSE35160	1450710_at	**Jarid2**	1418982_at	**Cebpa**	0.0377190203002264
GSE8733	1450710_at	**Jarid2**	1418982_at	**Cebpa**	0.0160946241438491

Aside from the TFs that are in common in all the selected species, other TFs were only conserved in 3-5 species on the same CNS regions different than what were shown in the Table 4.21, such as; Cebpg, Creb3l2, Irf1, Creb1, E2f1, Gli3. Since the criteria were to capture the TFs that were conserved both in position and in sequence among the species, they were not included in the table.

After analyzing all the data from the results, we have candidates of Cebpa regulators which were discussed at the next section.

5. DISCUSSION

The processes of EMT and MET in normal conditions of the organism are critical for development and physiological events such as wound healing, maintaining cellular phenotypical features and ensure plasticity (Chaffer et al., 2007; Kalluri, 2009; Lamouille et al., 2014). In addition to this, tumor cells also capture and use these mechanisms in order to gain a more aggressive phenotype in pathological processes such as tumorigenesis and metastasis (Heerboth et al. 2015). This metastatic ability of tumor cells to relocate in near or far tissues is one of the major cause of deaths by cancer. Even though the nature of metastasis is not fully unveiled yet by the scientific community, EMT-MET processes are thought to have a leading role in exposing this (Spano et al., 2012; Seyfried and Huysentruyt 2013).

The EMT process was regulated by the factors known as EMT inducers that repress the E-cadherin function such as Snai1/Snail, Snai2/Slug, Twist1/2, Zeb, and KLF8 (Thuault et al., 2008; Lin and Baritaki, 2010; Lo. H.W., 2007; Jung et al., 2008; Lamouille, Xu and Derynck, 2014; Peinado et al., 2007; Wand et al., 2007), whether by cellular modifications and differentiation of gene expressions directly, or indirectly by repressing the transcription (Yang and Weinberg, 2008; Sobrado et al., 2009). These factors are responsible for downgrading the E-cad expression rapidly in EMT and increasing it back again swiftly in MET (Voutsadakis, 2016). Even though MET can look like the reverse of EMT, it is regulated by more complex and dynamic mechanisms. Although it is not fully known how the mechanism works on MET and the EMT inducers might play a role in activating E-cad expression only to an unknown degree as of yet (Kim et al.; 2017).

Since the E-cad expression is on the center of these questions that need to be answered, any factors that were known to play a role affecting the E-cad expression needs to be addressed in order to reveal the secrets of MET process. To answer these questions, our lab is trying to unveil the network nodes of MET by various approaches and Cebpa is one of the TFs that are found to be correlated with E-cad expression. This is one of the reasons why the regulation of Cebpa is needed to be identified in more detail.

The results of our last project showed that in RNA-seq data of MET, Cebpa has a high correlation with epithelial phenotype. Cebpa expression decreases with EMT and increases in MET alongside with Cdh1. In the view of such evidence, it was hypothesized that Cebpa might have an essential role in EMT-MET processes and its regulation and function can be crucial in MET.

The in-vitro experiments of Cebpa knockdown in MET also showed supporting results for the hypothesis. When Cebpa was silenced in mesenchymal state cells, they were not able to return to their epithelial state in MET, the process was blocked. Therefore the regulation of Cebpa was needed to be addressed more.

In order to identify the possible regulators of Cebpa, three main criteria were followed;

Conserved function → Conserved regulation → Conserved binding site

First, the conserved sequences were selected among the six vertebrate species in mVista. This tool is selected because of its up-to-date interface, versatile usage, and ability to visualize aligned sequences in great detail. Additional sequences were selected from the upstream and downstream of Cebpa to include both cis-acting elements and trans-acting factors present in the sequence. The conserved TFBSs were detected in CNS regions, and each of the species were scanned for the common TFs that are both in the same CNS and have conserved binding sequences.

After obtaining the complete list of TFBSs, GEO expression profile database was searched for the array data on cells and tissues that are known to have high Cebpa expression. Among those expression profiles, three GEO datasets of the same tissue were selected as candidates to compare with TFBSs results.

When we look at the TFBSs at the conserved sequences in eight CNS regions, we see a pattern in common TFs among species. Sox, Hox, Klf, Forkhead box, hepatic nuclear factors, and hormone receptors are common in almost all CNS regions. Since Cebpa was known to have roles in differentiation, inflammatory response, liver regeneration and metabolism (Takayama et al., 2014; Zjablovskaja et al., 2017), it is not surprising to see those TFBSs. In transcriptional regulation, TF complexes play a more prominent role in developmental processes rather than single ones. (Reményi et al., 2004; Verger and Duterque-Coquillaud, 2002). However, when we look at the fewer TFs that are common and have binding sites on Cebpa region, it will show us a more specific story about its regulation. After analyzing the current lists of TFs, we have one possible candidate for Cebpa regulation, which is Jarid2.

Jarid2 is a DNA-binding protein that mainly acts as a transcriptional repressor. It interacts with the Polycomb repressive complex 2 (PRC2) which plays an essential role in regulating gene expression during embryonic development. Jarid2 is required for recruitment of PRC2 to chromatin in embryonic stem cells (Peng et al., 2009; Shen et al., 2009; Landeira et al., 2010; Li et al., 2010; Pasini et al., 2010; Landeira and Fisher, 2011; Holoch and Margueron, 2017). It is also a regulator of histone methyltransferase complexes which have a role in embryonic development, including heart and liver development, neural tube fusion process and hematopoiesis. In embryonic stem cells, it associates with the PRC2 complex and inhibits trimethylation of H3K27me3, thereby playing a pivotal role in the differentiation of embryonic stem cells and normal development (Pasini et al., 2010, Al-Raawi et al., 2018).

Since Cebpa is a significant cause of Acute myeloid leukemia (AML) when it is mutated (Sato et al., 2013; Lourenço and Coffer, 2017; Kirstetter et al., 2008; Tada et al., 2006; Loomis et al., 2007), the pathways and regulators factors might be interacting or effecting one another, because in AML, PRC2 mutations leads the way to Cebpa mutations by impairing maturation of hematopoietic cells (Sauntharararajah and Maciejewski, 2012). The role of Cebpa in tumor suppression by repressing cell proliferation and promoting terminal differentiation might be a cause of being targeted by the effectors of embryonic development repressive factors such in the example, Jarid2.

6. CONCLUSION AND SUGGESTIONS

Cebpa is one of the transcription factors that regulate MET process, and its relationship or partnership with the other EMT-MET-specific factors will help to shed light on the MET process.

Cebpa has more than three regulators which were found common among GEO datasets and TFBSs since the study was only conducted with three gene expression profiles on the same tissue where Cepba is known to be expressed in high levels. More sample size and different cells/tissues should be analyzed in order to map the specific regulators for different biological or pathophysiological conditions.

Jarid2 is a putative regulator of Cebpa.

The role of Cebpa in MET needs to be explained more in detail with extensive in-vitro experiments on MET to support the results of this study.

The tumor suppressor role of Cebpa should be studied in accordance with blockage of MET when Cebpa was silenced since it could be significant in metastatic processes.

7. REFERENCES

Agre, P., Johnson, P. F., and McKnight, S.L., Cognate DNA binding specificity retained after leucine zipper exchange between GCN4 and C/EBP. Science, 1989;246, 922-926

Aigner, K., Dampier, B., Descovich, L., Mikula, et al. The transcription factor ZEB1 (deltaEF1) promotes tumour cell dedifferentiation by repressing master regulators of epithelial polarity. Oncogene, 2007;26(49), 6979–6988.

Akira S, Isshiki H, Sugita T, et al. A nuclear factor for IL-6 expression (NF-IL6) is a member of a C/EBP family. EMBO J. 1990;9(6):1897–1906.

Alotaibi, H., Basilicata MF, Shehwana H, Kosowan T et al. Enhancer cooperativity as a novel mechanism underlying the transcriptional regulation of E-cadherin during mesenchymal to epithelial transition, Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms,2015; 1849(6), pp. 731–742.

Al-Raawi, D., Jones, R., Wijesinghe, S., Halsall, et al. A novel form of JARID2 is required for differentiation in lineage-committed cells. The EMBO Journal, 2018; e98449.

Antonello ZA, Reiff T, Dominguez M. Mesenchymal to epithelial transition during tissue homeostasis and regeneration: Patching up the Drosophila midgut epithelium. Fly (Austin). 2015;9(3):132–137.

Antonson, P., and Xanthopoulos, K. G. Molecular Cloning, Sequence, and Expression Patterns of the Human Gene Encoding CCAAT/Enhancer Binding Protein α (C/EBPα). Biochemical and Biophysical Research Communications, 1995; 215(1), 106–113.

Arnold S Robertson E, Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. Nature Reviews Molecular Cell Biology 2009; vol: 10 (2) pp: 91-103

Atsuta Y, Takahashi Y. FGF8 coordinates tissue elongation and cell epithelialization during early kidney tubulogenesis. Development, 2015;142:2329-37.

Avellino, R., and Delwel, R. Expression and regulation of C/EBPα in normal myelopoiesis and in malignant transformation. Blood, 2017;129(15), 2083–2091. doi:10.1182/blood-2016-09-687822

B. Hallgrimsson and B. K. Hall, Epigenetics: Linking Genotype and Phenotype in Development and Evolution, University of California Press, 2011

Bararia, D., Kwok, H. S., Welner, R. S., Numata, A. et al. Acetylation of C/EBPα inhibits its granulopoietic function. Nature Communications, 2016; 7, 10968.

Bartke, T., Vermeulen, M., Xhemalce, B., Robson, S.C. et al. Nucleosome-interacting proteins regulated by DNA and histone methylation. Cell, 2010;143, 470-484.

Blum R, Vethantham V, Bowman C, Rudnicki M, et al. Genome-wide identification of enhancers in skeletal muscle: the role of MyoD1. Genes Dev. 2012 Dec 15; 26(24):2763-79.

Bretz JD, Williams SC, Baer M, Johnson PF, et al. C/EBP-related protein 2 confers lipopolysaccharide-inducible expression of interleukin 6 and monocyte chemoattractant protein 1 to a lymphoblastic cell line. Proc Natl Acad Sci U S A. 1994;91(15):7306–7310.

Brivanlou, A. H. Signal Transduction and the Control of Gene Expression. Science, 2002;295(5556), 813–818.

Bürglin TR,and Affolter M. Homeodomain proteins: an update. Chromosoma. 2016;125 (3): 497–521.

Calkhoven CF, Müller C, Leutz A. Translational control of C/EBPalpha and C/EBPbeta isoform expression. Genes Dev. 2000;14(15):1920–1932.

Calo E, Wysocka J. Modification of enhancer chromatin: what, how, and why?. Mol Cell. 2013;49(5):825–837.

Cao, Z., Umek, R. M., and McKnight, S. L. Regulated expression of three C/EBP isoforms during adipose conversion of 3T3-L1 cells. Genes & Development, 1991;5(9), 1538–1552. doi:10.1101/gad.5.9.1538

Carrozza, M.J., Li, B., Florens, L., Suganuma, T., Swanson, et al. Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. Cell,2005; 123, 581-592.

Cartharius K, Frech K, Grote K, Klocke B, et al. MatInspector and beyond: promoter analysis based on transcription factor binding sites Bioinformatics,2005; 21, 2933-42

Cassel, T. N., Nordlund-Möller, L., Andersson, O., Gustafsson, J.-Å., et al. C/EBP α and C/EBP δ Activate the Clara Cell Secretory Protein Gene through Interaction with Two Adjacent C/EBP-Binding Sites. American Journal of Respiratory Cell and Molecular Biology, 2000;22(4), 469–480.

Cast, A., Valanejad, L., Wright, M., Nguyen, P. et al. C/EBPα-dependent preneoplastic tumor foci are the origin of hepatocellular carcinoma and aggressive pediatric liver cancer. Hepatology, 2018;67(5), 1857–1871.

Celia-Terrassa T, Meca-Cortes O, Mateo F, de Paz AM, et al. Epithelial–mesenchymal transition can suppress major attributes of human epithelial tumor-initiating cells. J Clin Invest 2012;122(5):1849–68.8

Chang, C. J., Chen, T. T., Lei, H. Y., Chen, D. S., et al. Molecular cloning of a transcription factor, AGP/EBP, that belongs to members of the C/EBP family. Molecular and Cellular Biology, 1990; 10(12), 6642–6653.

Chao Y, Wu Q, Acquafondata M, Dhir R, et al. Partial Mesenchymal to Epithelial Reverting Transition in Breast and Prostate Cancer Metastases. Cancer Microenvironment 2012;5:19-28.

Chapiro, E., Russell, L., Radford-Weiss, I., Bastard, C., et al. Overexpression of CEBPA resulting from the translocation t(14;19)(q32;q13) of human precursor B acute lymphoblastic leukemia. Blood, 2006;108(10), 3560–3563.

Cheng J, Blum R, Bowman C, Hu D. et al. A role for H3K4 monomethylation in gene repression and partitioning of chromatin readers. Mol Cell. 2014;53(6):979–992.

Chew, C.-H., Chew, G.-S., Najimudin, N., and Tengku-Muhammad, T. S. Interleukin-6 inhibits human peroxisome proliferator activated receptor alpha gene expression via CCAAT/enhancer-binding proteins in hepatocytes. The International Journal of Biochemistry & Cell Biology, 2007;39(10), 1975–1986.

Choi B, Park KS, Kim JH, Ko KW, et al. Stiffness of Hydrogels Regulates Cellular Reprogramming Efficiency Through Mesenchymal-to-Epithelial Transition and Stemness Markers. Macromolecular bioscience 2015.

Choi, S. K., Park, S., Jang, S., Cho, H. et al. Cascade regulation of PPARγ2 and C/EBPα signaling pathways by celastrol impairs adipocyte differentiation and stimulates lipolysis in 3T3-L1 adipocytes. Metabolism, 2016;65(5), 646–654.

Creyghton MP, Cheng AW, Welstead GG, Kooistra T, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc Natl Acad Sci U S A. 2010;107(50):21931–21936.

Darlington, G. J., Ross, S. E., and MacDougald, O. A. The Role ofC/EBPGenes in Adipocyte Differentiation. Journal of Biological Chemistry, 1998;273(46), 30057–30060.

Davidson LA, Keller R, and Desimone DW. Assembly and remodeling of the fibrillar fibronectin extracellular matrix during gastrulation and neurulation in Xenopus laevis. Dev Dyn 2004;231:888-95.

Descombes, P., Chojkier, M., Lichtsteiner, S., Falvey, E., et al. LAP, a novel member of the C/EBP gene family, encodes a liver-enriched transcriptional activator protein. Genes & Development,1990; 4(9), 1541–1551.

Downing TL, Soto J, Morez C, Houssin T, et al. Biophysical regulation of epigenetic state and cell reprogramming. Nat Mater 2013;12:1154-62.

Duval, M. and 1844-1907. Atlas d'embryologie'. G. Masson. 1889; Available at: http://agris.fao.org/agris-search/search.do?recordID=US201300442359, Reached on May 2019

Elberg, G. Modulation of the Murine Peroxisome Proliferator-Activated Receptor gamma 2 Promoter Activity by CCAAT/Enhancer Binding Proteins. Journal of Biological Chemistry, 2000;275(36):27815-22

Evan D. Rosen, Christopher J. Walkey, et al. Transcriptional regulation of adipogenesis. Genes Dev. 2000;14: 1293-1307.

Flodby, P., Barlow, C., Kylefjord, H., Åhrlund-Richter, L., et al. Increased Hepatic Cell Proliferation and Lung Abnormalities in Mice Deficient in CCAAT/Enhancer Binding Protein α. Journal of Biological Chemistry, 1996;271(40), 24753–24760.

Frankfurt, U. Cell Growth and Differentiation From the Perspectine of Dynamical Organization of Cellular and Subcellular Processes‘, Science, 1996;64(1), pp. 55–79.

Gal A, Sjoblom T, Fedorova L, Imreh S, et al. Sustained TGF beta exposure suppresses Smad and non-Smad signalling in mammary epithelial cells, leading to EMT and inhibition of growth arrest and apoptosis. Oncogene 2008;27:1218-30.

Garside VC, Chang AC, Karsan A, Hoodless PA. Co-ordinating Notch, BMP, and TGF-β signaling during heart valve development. Cell Mol Life Sci. 2013;70(16):2899–2917.

Gates, L. A., Shi, J., Rohira, A. D., Feng, Q. et al. Acetylation on histone H3 lysine 9 mediates a switch from transcription initiation to elongation. Journal of Biological Chemistry, 2017;292(35), 14456–14472.

Gehring WJ, Affolter M, Bürglin T. "Homeodomain proteins". Annual Review of Biochemistry. 1994;63: 487–526.

Gilbert, R., Vickaryous, M. and Viloria-Petit, A. Signalling by Transforming Growth Factor Beta Isoforms in Wound Healing and Tissue Regeneration, Journal of Developmental Biology, 2016;4(2), p. 21.

Gill G. "Regulation of the initiation of eukaryotic transcription". Essays in Biochemistry. 2001;37: 33–43.

Godbey, W. T. Genes. An Introduction to Biotechnology, 2014;65–105.

Grünert S. Jechlinger M. Beug H, Diverse cellular and molecular mechanisms contribute to epithelial plasticity and metastasis. Nature Reviews Molecular Cell Biology 2003;vol: 4 (8) pp: 657-665

Guo, L., Li, X., and Tang, Q.-Q. Transcriptional Regulation of Adipocyte Differentiation: A Central Role for CCAAT/Enhancer-binding Protein (C/EBP) β. Journal of Biological Chemistry, 2014;290(2), 755–761.

Hay ED. An overview of epithelio-mesenchymal transformation. Acta Anat. 1995;154:8–20.

Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, et al. Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature. 2009; 459(7243):108-12.

Heintzman ND, Stuart RK, Hon G, Fu Y, et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet. 2007; 39(3):311-8.

Hindley, C. and Philpott, A. Co-ordination of cell cycle and differentiation in the developing nervous system.‘, The Biochemical journal. Portland Press Ltd, 2012;444(3), pp. 375–82.

Holoch D, and Margueron R. Mechanisms regulating PRC2 recruitment and enzymatic activity. Trends Biochem Sci, 2017;42: 531–542

Hu, H.; Miao, Y.R.; Jia, L.H.; Yu, Q.Y.; et al. AnimalTFDB 3.0: a comprehensive resource for annotation and prediction of animal transcription factors. Nucleic Acids Res 2019, 47, D33-D38.

Hu, X.-X., Feng, J., Huang, X.-W., Lu, P.-Z., et al. Histone deacetylases up-regulate C/EBPα expression through reduction of miR-124-3p and miR-25 in hepatocellular carcinoma. Biochemical and Biophysical Research Communications. 2019; 514(3):1009-1016.

Hu, Y.-J., Belaghzal, H., Hsiao, W.-Y., Qi, J., et al. Transcriptional and post-transcriptional control of adipocyte differentiation by Jumonji domain-containing protein 6. Nucleic Acids Research, 2015;43(16), 7790–7804.

Huber M. Kraut N. Beug H, Molecular requirements for epithelial–mesenchymal transition during tumor progression, Current Opinion in Cell Biology 2005 vol: 17 (5) pp: 548-558

Hwang CS, Mandrup S, MacDougald OA, Geiman DE, Lane MD. Transcriptional activation of the mouse obese (ob) gene by CCAAT/enhancer binding protein alpha. Proc Natl Acad Sci U S A. 1996;93(2):873–877.

Illman SA, Lehti K, Keski-Oja J, Lohi J, Epilysin (MMP-28) induces TGF-β mediated epithelial to mesenchymal transition in lung carcinoma cells, Journal of Cell Science 2006;119: 3856-3865;

J.P. Thiery, H. Acloque, R.Y.J Huang, M.A Nieto, Epithelial-Mesenchymal in development and disease, Cell 2009;139. 871-890

Johnson, P. F. Molecular stop signs: regulation of cell-cycle arrest by C/EBP transcription factors. Journal of Cell Science, 2005;118(12), 2545–2555.

Joshi, A.A., and Struhl, K. Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. Mol. Cell 2005;20, 971-978.

Julien S. Puig I. Caretti E. Bonaventure J. et. al. Activation of NF-kappaB by Akt upregulates Snail expression and induces epithelium mesenchyme transition. Oncogene 2007; vol: 26 (53) pp: 7445-7456

Jung H. Lee K. Park S. Park J. et. al. TMPRSS4 promotes invasion, migration and metastasis of human tumor cells by facilitating an epithelial–mesenchymal transition, Oncogene 2008; vol: 27 (18) pp: 2635-2647

Kalluri R, and Neilson EG. Epithelial-mesenchymal transition and its implications for fibrosis. J Clin Invest. 2003;112(12):1776–1784.

Kalluri R, and Weinberg RA. The basics of epithelial-mesenchymal transition [published correction appears in J Clin Invest. 2010 May 3;120(5):1786]. J Clin Invest. 2009;119(6):1420–1428.

Kanamori, Y., Murakami, M., Sugiyama, M., Hashimoto, et al. Interleukin-1β (IL-1β) transcriptionally activates hepcidin by inducing CCAAT enhancer-binding protein δ (C/EBPδ) expression in hepatocytes. Journal of Biological Chemistry, 2017;292(24), 10275–10287.

Khanna-Gupta, A. Chromatin immunoprecipitation (ChIP) studies indicate a role for CCAAT enhancer binding proteins alpha and epsilon (C/EBPalpha and C/EBPepsilon) and CDP/cut in myeloid maturation-induced lactoferrin gene expression. Blood, 2003;101(9), 3460–3468.

Kim HY, Jackson TR, Davidson LA. On the role of mechanics in driving mesenchymal-to-epithelial transitions. Semin Cell Dev Biol. 2017;67:113–122.

Kim, D. Xing T. Yang Z. Dudek R. et al. Epithelial Mesenchymal Transition in Embryonic Development, Tissue Repair and Cancer: A Comprehensive Overview‘, Journal of Clinical Medicine, 2017;7(1), p. 1.

Kim, I.-H., and Nam, T.-J. Enzyme-treated Ecklonia cava extract inhibits adipogenesis through the downregulation of C/EBPα in 3T3-L1 adipocytes. International Journal of Molecular Medicine, 2017;39(3), 636–644.

Kimura, T., Christoffels, V. M., Chowdhury, S., Iwase, K.,et al. Hypoglycemia-associated Hyperammonemia Caused by Impaired Expression of Ornithine Cycle Enzyme Genes in C/EBPα Knockout Mice. Journal of Biological Chemistry, 1998;273(42), 27505–27510.

Kirstetter, P., Schuster, M. B., Bereshchenko, O., Moore, S., et al. Modeling of C/EBPα Mutant Acute Myeloid Leukemia Reveals a Common Expression Signature of Committed Myeloid Leukemia-Initiating Cells. Cancer Cell, 2008;13(4), 299–310.

Ko, C.-Y., Chang, W.-C., and Wang, J.-M. Biological roles of CCAAT/Enhancer-binding protein delta during inflammation. Journal of Biomedical Science, 2015;22(1), 6.

Kooistra, S. M., and Helin, K. Molecular mechanisms and potential functions of histone demethylases. Nature Reviews Molecular Cell Biology, 2012;13(5), 297–311.

Kumarswamy R, Mudduluru G, Ceppi P, Muppala S, et al. MicroRNA-30a inhibits epithelial-tomesenchymal transition by targeting Snai1 and is downregulated in non-small cell lung cancer. Int J Cancer 2012;130(9):2044–53.

Laity JH, Lee BM, Wright PE (February 2001). "Zinc finger proteins: new insights into structural and functional diversity". Current Opinion in Structural Biology. 11 (1): 39–46.

Lamouille, S., et al. Regulation of epithelial-mesenchymal and mesenchymal-epithelial transitions by microRNAs. Curr Opin Cell Biol 2013;25(2): 200-207.

Lamouille, S., Xu, J. and Derynck, R. Molecular mechanisms of epithelial–mesenchymal transition, Nature Reviews Molecular Cell Biology. Nature Publishing Group, 2014;15(3), pp. 178–196.

Landeira D, Fisher AG. Inactive yet indispensable: the tale of Jarid2. Trends Cell Biol 2011;21: 74–80

Landeira D, Sauer S, Poot R, Dvorkina M, et al. Jarid2 is a PRC2 component in embryonic stem cells required for multi-lineage differentiation and recruitment of PRC1 and RNA Polymerase II to developmental regulators. Nat Cell Biol 2010;12: 618–624

Landschulz, W. H., Johnson, P. F., Adashi, E. Y., Graves, B. J. et al. L.Isolation of a recombinant copy of the gene encoding C/EBP. Genes Dev. 1988;2, 786-800

Landschulz, W. H., Johnson, P. F., and McKnight, S. L., The DNA binding domain of the rat liver nuclear protein C/EBP is bipartite. Science 1989;243, 1681-1688

Lane, M. D., Tang, Q.-Q., and Jiang, M.-S. Role of the CCAAT Enhancer Binding Proteins (C/EBPs) in Adipocyte Differentiation. Biochemical and Biophysical Research Communications, 1999;266(3), 677–683.

Latchman DS. Transcription factors: an overview. The International Journal of Biochemistry & Cell Biology. 1997;29 (12): 1305–12.

Latimer A, and Jessen JR. Extracellular matrix assembly and organization during zebrafish gastrulation. Matrix Biology 2010;29:89-96.

Lattin JE, Schroder K, Su AI, Walker JR. et al. Expression analysis of G Protein-Coupled Receptors in mouse macrophages. Immunome Res. 2008;4:5. Published 2008 Apr 29.

Le Bras, G. F., Taubenslag, K. J. and Andl, C. D. The regulation of cell-cell adhesion during epithelial-mesenchymal transition, motility and tumor progression., Cell adhesion & migration. Taylor & Francis, 6(4), 2012;pp. 365–73.

Lee YH, Sauer B, Johnson PF, Gonzalez FJ. Disruption of the c/ebp alpha gene in adult mouse liver. Mol Cell Biol. 1997;17(10):6014–6022.

Lefstin, J.A. and Yamamoto, K.R. Allosteric effects of DNA on transcriptional regulators. Nature 1998;392, 885–888.

Lekstrom-Himes, J., and Xanthopoulos, K.G. Biological Role of the CCAAT/Enhancer-binding Protein Family of Transcription Factors. Journal of Biological Chemistry, 1998;273(44), 28545–28548.

Li G, Margueron R, Ku M, Chambon P, et al. Jarid2 and PRC2, partners in regulating gene expression. Genes Dev 2010;24: 368–380

Lin K, Baritaki S, Militello L, Malaponte G, et al. The Role of B-RAF Mutations in Melanoma and the Induction of EMT via Dysregulation of the NF-κB/Snail/RKIP/PTEN Circuit. Genes Cancer. 2010;1(5):409–420.

Littlewood TD, Evan GI. Transcription factors 2: helix-loop-helix. Protein Profile. 1995;2 (6): 621–702.

Liu X, Wang C, Chen Z, Jin Y, et al. MicroRNA-138 suppresses epithelial–mesenchymal transition in squamous cell carcinoma cell lines. Biochem J 2011;440(1):23–31.

Lo HW, Hsu SC, Xia W, Cao X. et al. Epidermal growth factor receptor cooperates with signal transducer and activator of transcription 3 to induce epithelial-mesenchymal transition in cancer cells via up-regulation of TWIST gene expression. Cancer Res. 2007;67(19):9066–9076.

Loomis KD, Zhu S, Yoon K, Johnson PF, et al. Genetic ablation of CCAAT/enhancer binding protein alpha in epidermis reveals its role in suppression of epithelial tumorigenesis. Cancer Res. 2007;67(14):6768–6776.

Lourenço, A. R., and Coffer, P. J. A tumor suppressor role for C/EBPα in solid tumors: more than fat and blood. Oncogene, 2017;36(37), 5221–5230.

Lu GD, Ang YH, Zhou J, Tamilarasi J, et al. CCAAT/enhancer binding protein α predicts poorer prognosis and prevents energy starvation-induced cell death in hepatocellular carcinoma [published correction appears in Hepatology. 2019 Jun;69(6):2720]. Hepatology. 2015;61(3):965–978.

Lu, G., Leung, C. H., Yan, B., Tan, C. et al. C/EBPα Is Up-regulated in a Subset of Hepatocellular Carcinomas and Plays a Role in Cell Growth and Proliferation. Gastroenterology, 2010;139(2), 632–643.e4.

M. J. West-Eberhard, Developmental Plasticity and Evolution, Oxford University Press, New York, NY, USA, 2003.

Maccarty, W. C. and Caylor, H. D. Metaplasia in ovarian dermoids and cystadenomas: REPORT OF THREE CASES., Annals of surgery. Lippincott, Williams, and Wilkins, 1922;76(2), pp. 238–45.

Mason, M. M., He, Y., Chen, H., Quon, M. J., et al. Regulation of Leptin Promoter Function by Sp1, C/EBP, and a Novel Factor1. Endocrinology, 1998;139(3), 1013–1022.

McCoy EL, Iwanaga R, Jedlicka P, Abbey N. et al. Six1 expands the mouse mammary epithelial stem/progenitor cell pool and induces mammary tumors that undergo epithelial-mesenchymal transition. J Clin Invest. 2009;119(9):2663–2677.

Micalizzi D, Farabaugh S, Ford H. Epithelial-mesenchymal transition in cancer: parallels between normal development and tumor progression. J Mammary Gland Biol Neoplasia. 2010; 15 (2): 117-34.

Moreno-Bueno, G., Portillo, F. and Cano, A. Transcriptional regulation of cell polarity in EMT and cancer‘, Oncogene. Nature Publishing Group, 2008;27(55), pp. 6958–6969.

Nakamura M, Tokura Y, Epithelial-mesenchymal transition in the skin. Journal of Dermatological Science, 2011 vol: 61 (1) pp: 7-13

Nakaya Y, Kuroda S, Katagiri YT, Kaibuchi K, et al. Mesenchymal-epithelial transition during somitic segmentation is regulated by differential roles of Cdc42 and Rac1. Developmental cell 2004;7:425-38.

Nerlov, C., and Ziff, E. B. Three levels of functional interaction determine the activity of CCAAT/enhancer binding protein-alpha on the serum albumin promoter. Genes & Development, 1994;8(3), 350–362.

Okuma, T., Hirata, M., Yano, F., Mori, D., et al. Regulation of mouse chondrocyte differentiation by CCAAT/enhancer-binding proteins . Biomedical Research, 2015;36(1), 21–29.

Osada, S., Yamamoto, H., Nishihara, T., and Imagawa, M. DNA Binding Specificity of the CCAAT/Enhancer-binding Protein Transcription Factor Family. Journal of Biological Chemistry, 1996;271(7), 3891–3896.

Ozdamar, B. Bose, R. Barrios-Rodiles, M. Wang, H. et al. Regulation of the polarity protein Par6 by TGFbeta receptors controls epithelial cell plasticity.‘, Science (New York, N.Y.). American Association for the Advancement of Science, 2005;307(5715), pp. 1603–9.

Pasini D., Cloos P.A., Walfridsson J., Olsson L., et al. "ARID2 regulates binding of the Polycomb repressive complex 2 to target genes in ES cells. Nature 2010;464:306-310

Pei DQ, and Shih CH. An "attenuator domain" is sandwiched by two distinct transactivation domains in the transcription factor C/EBP. Mol Cell Biol. 1991;11(3):1480–1487.

Peinado H. Olmeda D. Cano A, Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype. Nature Reviews Cancer 2007;vol: 7 (6) pp: 415-428

Peinado, H., Portillo, F. and Cano, A. Transcriptional regulation of cadherins during development and carcinogenesis‘, International Journal of Developmental Biology, 2004;48(5–6), pp. 365–375.

Peng JC, Valouev A, Swigut T, Zhang J, et al. Jarid2/Jumonji coordinates control of PRC2 enzymatic activity and target gene occupancy in pluripotent cells. Cell 2009;139: 1290–1302

Bardou P, Mariette J, Escudié F, Djemiel C, et al. jvenn: an interactive Venn diagram viewer. BMC Bioinformatics 2014;15:293

Phillips, T. Regulation of transcription and gene expression in eukaryotes. Nature Education 2008;1(1):199

Planas-Silva MD, and Waltz PK. Estrogen promotes reversible epithelial-to-mesenchymal-like transition and collective motility in MCF-7 breast cancer cells. The Journal of steroid biochemistry and molecular biology 2007;104:11-21.

Poli, V. The Role of C/EBP Isoforms in the Control of Inflammatory and Native Immunity Functions. Journal of Biological Chemistry, 1998;273(45), 29279–29282.

Poli, V., Mancini, F. P., and Cortese, R. IL-6DBP, a nuclear protein involved in interleukin-6 signal transduction, defines a new family of leucine zipper proteins related to CEBP. Cell, 1990;63(3), 643–653.

Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, et al. A unique chromatin signature uncovers early developmental enhancers in humans. Nature, 2011; 470(7333):279-83.

Radice GL, Rayburn H, Matsunami H, Knudsen KA, et al. Developmental defects in mouse embryos lacking N-cadherin. Dev Biol 1997;181:64-78.

Raff R, The Shape of Life: Genes, Development, and the Evolution of Animal Form, University Of Chicago Press, 1996.

Ramji, Dipak P, and Pelagia Foka. CCAAT/enhancer-binding proteins: structure, function and regulation. The Biochemical journal 2002;vol. 365,Pt 3: 561-75.

Reik, W., Dean, W. and Walter, J. Epigenetic reprogramming in mammalian development.‘, Science (New York, N.Y.). American Association for the Advancement of Science, 2001;293(5532), pp. 1089–93.

Reményi A., Schöler H. R., Wilmanns M. Combinatorial control of gene expression. Nat. Struct. Mol. Biol. 2004;11, 812–815.

Remenyi, A., Scholer, H., and Wilmanns, M. Combinatorial control of gene expression. Nature Structural and Molecular Biology 2004;11, 812–815

Renner M, Czwan E, Hartmann W, Penzel R, et al. MicroRNA profiling of primary high-grade soft tissue sarcomas. Genes Chromosomes Cancer 2012;51(11):982–96.

Roesler, W. J. The role of C/EBP in nutrient and hormonal regulation of gene expression. Annual Review of Nutrition, 2001;21(1), 141–165.

Roman, C., Platero, J. S., Shuman, J., and Calame, K. Ig/EBP-1: a ubiquitously expressed immunoglobulin enhancer binding protein that is similar to C/EBP and heterodimerizes with C/EBP. Genes & Development, 1990;4(8), 1404–1415.

Ron, D., and Habener, J. F. CHOP, a novel developmentally regulated nuclear protein that dimerizes with transcription factors C/EBP and LAP and functions as a dominant-negative inhibitor of gene transcription. Genes & Development, 1992;6(3), 439–453.

Rossant, J. and Tam, P. P. L. Blastocyst lineage formation, early embryonic asymmetries and axis patterning in the mouse, Development, 2009;136(5), pp. 701–713.

Samuel MS, Lopez JI, McGhee EJ, Croft DR, et al. Actomyosin-mediated cellular tension drives increased tissue stiffness and β-catenin activation to induce epidermal hyperplasia and tumor growth. Cancer Cell. 2011;19(6):776–791.

Sato, A., Yamada, N., Ogawa, Y., and Ikegami, M. CCAAT/Enhancer-Binding Protein-α Suppresses Lung Tumor Development in Mice through the p38α MAP Kinase Pathway. PLoS ONE, 2013;8(2), e57013.

Sato, Y., Miyake, K., Kaneoka, H., and Iijima, S. Sumoylation of CCAAT/Enhancer-binding Protein α and Its Functional Roles in Hepatocyte Differentiation. Journal of Biological Chemistry, 2006;281(31), 21629–21639.

Saunthararajah, Y., and Maciejewski, J. Polycomb segment myeloid malignancies. Blood, 2012;119(5), 1097–1098.

Scarpa E, Szabó A, Bibonne A, Theveneau E, et al. Cadherin Switch during EMT in Neural Crest Cells Leads to Contact Inhibition of Locomotion via Repolarization of Forces. Dev Cell. 2015;34(4):421–434.

Shen X, Kim W, Fujiwara Y, Simon MD, et al. Jumonji modulates polycomb activity and self-renewal versus differentiation of stem cells. Cell 2009;139: 1303–1314

Shibue T, Weinberg RA. Metastatic colonization: Settlement, adaptation and propagation of tumor cells in a foreign tissue environment. Seminars in Cancer Biology 2011;21:99-106.

Slack, J. M. . and Tosh, D. Transdifferentiation and metaplasia - switching cell types, Current Opinion in Genetics & Development. Elsevier Current Trends, 2001;11(5), pp. 581–586.

Sobrado V. Moreno-Bueno G. Cubillo E. Holt L. et. al., The class I bHLH factors E2-2A and E2-2B regulate EMT, Journal of Cell Science 2009;vol: 122 (7) pp: 1014-1024

Somarelli, J. A., et al. Mesenchymal-Epithelial Transition in Sarcomas Is Controlled by the Combinatorial Expression of MicroRNA 200s and GRHL2. Mol Cell Biol 2016;36(19): 2503-2513.

Strippoli R, Benedicto I, Pérez Lozano ML, Cerezo A, et al. Epithelial-to-mesenchymal transition of peritoneal mesothelial cells is regulated by an ERK/NF-kappaB/Snail1 pathway. Dis Model Mech. 2008;1(4-5):264–274.

Sugahara, K., Iyama, K.-I., Kimura, T., Sano, K., et al. Mice lacking CCAAT/enhancer-binding protein-? show hyperproliferation of alveolar type II cells and increased surfactant protein mRNAs. Cell and Tissue Research, 2001;306(1), 57–63.

Tada, Y., Brena, R. M., Hackanson, B., Morrison, C., et al. Epigenetic Modulation of Tumor Suppressor CCAAT/Enhancer Binding Protein α Activity in Lung Cancer. JNCI: Journal of the National Cancer Institute, 2006;98(6), 396–406.

Takahashi K, and Yamanaka S,. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. Cell 2006;126:663-76.

Takayama, K., Kawabata, K., Nagamoto, Y., Inamura, M., et al. CCAAT/enhancer binding protein-mediated regulation of TGF receptor 2 expression determines the hepatoblast fate decision. Development, 2013;141(1), 91–100.

Tan, E. J., Olsson, A. K. and Moustakas, A. Reprogramming during epithelial to mesenchymal transition under the control of TGFbeta, Cell Adh Migr, 2015;9(3), pp. 233–246.

Thiery J. Sleeman J., Complex networks orchestrate epithelial–mesenchymal transitions, Nature Reviews Molecular Cell Biology 2006 vol: 7 (2) pp: 131-142

Thuault S, Tan EJ, Peinado H, Cano A, et al. HMGA2 and Smads co-regulate SNAIL1 expression during induction of epithelial-to-mesenchymal transition. J Biol Chem. 2008;283(48):33437–33446.

Tong IL and Richard AY, Transcription of Eukaryotic Protein-coding Genes. Annu Rev Genet. 2000;34:77-137.

Trepat X, Wasserman MR, Angelini TE, Millet E, et al. Physical forces during collective cell migration. Nature Physics 2009;5:426-30.

Tsukada, J., Yoshida, Y., Kominato, Y., and Auron, P. E. The CCAAT/enhancer (C/EBP) family of basic-leucine zipper (bZIP) transcription factors is a multifaceted highly-regulated system for gene regulation. Cytokine, 2011;54(1), 6–19.

Valentine JW, The evolution of multicellular plants and animals. Scientific American 1978 vol: 239 (3) pp: 140-6, 148-9, 152-3

Vasilyev A, Liu Y, Hellman N, Pathak N, et al. Mechanical stretch and PI3K signaling link cell migration and proliferation to coordinate epithelial tubule morphogenesis in the zebrafish pronephros. PloS one 2012;7:e39992.

Verger A., Duterque-Coquillaud M. When Ets transcription factors meet their partners. BioEssays 2002;24, 362–370.

Vinson C, Myakishev M, Acharya A, Mir AA, et al. Classification of human B-ZIP proteins based on dimerization properties. Molecular and Cellular Biology. 2002;22 (18): 6321–35.

Vinson, C.R., Sigler, P.B., and McKnight, S.L., Scissor-grip model for DNA recognition by a family of leucine zipper proteins. Science 1989;246, 911-916

Visel A, Prabhakar S, Akiyama JA, Shoukry M, et al. Ultraconservation identifies a small subset of extremely constrained developmental enhancers. Nat Genet. 200;40(2):158-60.

W. H. Freeman and Company. Pierce, B. Genetics: A conceptual approach. 2nd Edition. 2005

Wang JC Finding primary targets of transcriptional regulators. Cell Cycle. 2005;4 (3): 356–8.

Wang X. Zheng M. Liu G. Xia W. et. al. Kruppel-Like Factor 8 Induces Epithelial to Mesenchymal Transition and Epithelial Cell Invasion, Cancer Research 2007 vol: 67 (15) pp: 7184-7193

Wang Z, Zang C, Rosenfeld JA, Schones DE, et al. Combinatorial patterns of histone acetylations and methylations in the human genome. Nat Genet. 2008; 40(7):897-903.

Wang, N., Finegold, M., Bradley, A., Ou, C., et al. Impaired energy homeostasis in C/EBP alpha knockout mice. Science, 1995;269(5227), 1108–1112.

Wang, Y., Li, X., and Hu, H. H3K4me2 reliably defines transcription factor binding regions in different cells. Genomics, 2014;103(2-3), 222–228.

Wang, Z.-H., Gong, K., Liu, X., Zhang, Z., et al. C/EBPβ regulates delta-secretase expression and mediates pathogenesis in mouse models of Alzheimer's disease. Nature Communications, 2018; 9(1).

Wedel, A., and Lömsziegler-Heitbrock, H. W. The C/EBP Family of Transcription Factors. Immunobiology, 1995;193(2-4), 171–185.

Wells A, Yates C, Shepard CR. E-cadherin as an indicator of mesenchymal to epithelial reverting transitions during the metastatic seeding of disseminated carcinomas. Clin Exp Metastasis. 2008;25(6):621–628.

Williams, S. C., Cantwell, C. A., and Johnson, P. F. A family of C/EBP-related proteins capable of forming covalently linked leucine zipper dimers in vitro. Genes & Development, 1991;5(9), 1553–1567.

Wintjens R, Rooman M. Structural classification of HTH DNA-binding domains and protein-DNA interaction modes. Journal of Molecular Biology. 1996;262 (2): 294–313.

Wu C, Jin X, Tsueng G, Afrasiabi C, et al. BioGPS: building your own mash-up of gene annotations and expression profiles. Nucleic Acids Res. 2016;44(D1):D313–D316.

Xie, H., Ye, M., Feng, R., and Graf, T. Stepwise Reprogramming of B Cells into Macrophages. Cell, 2004;117(5), 663–676.

Xu J, Lamouille S, Derynck R. TGF-beta-induced epithelial to mesenchymal transition. Cell Res. 2009;19(2):156–172.

Xu L, Glass CK, Rosenfeld MG. Coactivator and corepressor complexes in nuclear receptor function. Current Opinion in Genetics & Development. 1999;9 (2): 140–7.

Yang J, Weinberg R, Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis. Developmental cell 2008 vol: 14 (6) pp: 818-29

Ye, M., Zhang, H., Amabile, G., Yang, H., et al. C/EBPa controls acquisition and maintenance of adult haematopoietic stem cell quiescence. Nature Cell Biology, 2013;15(4), 385–394.

Yin, H., Lowery, M., and Glass, J. In prostate cancer C/EBPα promotes cell growth by the loss of interactions with CDK2, CDK4, and E2F and by activation of AKT. The Prostate, 2009;69(9), 1001–1016.

Yoon, K., and Smart, R. C. C/EBP Is a DNA Damage-Inducible p53-Regulated Mediator of the G1 Checkpoint in Keratinocytes. Molecular and Cellular Biology, 2004;24(24), 10650–10660.

Zeisberg M, Neilson EG. Biomarkers for epithelial-mesenchymal transitions. J Clin Invest. 2009;119(6):1429–1437.

Zentner GE, Tesar PJ, Scacheri PC. Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. Genome Res. 2011; 21(8):1273-83.

Zhang DE, Zhang P, Wang ND, Hetherington CJ, et al. Absence of granulocyte colony-stimulating factor signaling and neutrophil development in CCAAT enhancer binding protein alpha-deficient mice. Proc Natl Acad Sci USA. 1997;94(2):569–574.

Zhang P, Iwama A, Datta MW, Darlington GJ, et al. Upregulation of interleukin 6 and granulocyte colony-stimulating factor receptors by transcription factor CCAAT enhancer binding protein alpha (C/EBP alpha) is critical for granulopoiesis. J Exp Med. 1998;188(6):1173–1184.

Zjablovskaja, P., Kardosova, M., Danek, P., Angelisova, P., et al. EVI2B is a C/EBPα target gene required for granulocytic differentiation and functionality of hematopoietic progenitors. Cell Death and Differentiation, 2017;24(4), 705–716.

8. APPENDIX

Table A1. Complete List of TFBSs on Cebpa promoter

Matrix	Matrix sim.	Sequence
O$TF2B/BRE.01	1	ccgCGCC
O$VTBP/VTATA.02	1	ctctaTAAAagcggtcc
V$EBOX/USF.01	1	gcacaCACGtggtccgt
V$MZF1/MZF1.01	1	gtGGGGagcat
V$NKXH/NKX25.01	1	ctagtcAAGTgttcgggac
V$PLAG/PLAG1.02	1	caGGGGgcgatgcccgaccctct
V$TAIP/CSRNP1.01	1	AGAGtgc
V$TAIP/CSRNP3.01	1	aGAGTgc
V$ZF02/ZKSCAN3.01	1	ctcagCCCCaccccggcggcgct
V$KLFS/IKLF.01	0.997	cgccgGGGTggggctgagc
O$VTBP/VTATA.01	0.995	ctctaTAAAagcggtcc
V$KLFS/BKLF.01	0.994	cgccgGGGTggggctgagc
V$ZF02/ZNF300.01	0.993	gctgtcgCCCCcgcacacacgtg
V$ZF5F/ZF5.01	0.993	cggtgcGCGCtaggt
V$ZFXY/ZFX.01	0.993	cgGGCCtggcc
V$MYT1/MYT1.02	0.992	agaAAGTtttcca
V$SAL4/SALL4.01	0.992	tTGGAgagaga
V$EBOX/MNT.01	0.991	cggacCACGtgtgtgcg
V$HIFF/HIF1.02	0.991	cgcacacaCGTGgtccg
V$EBOX/NMYC.01	0.99	cggaccaCGTGtgtgcg
V$HAND/TAL1_E2A.02	0.99	taagaccCAGCaggcaccatc
V$EBOX/CMYC.01	0.988	cggaccaCGTGtgtgcg
V$ZICF/ZIC2.02	0.988	agacccCAGCaggcaccatc
V$EBOX/MYCMAX.03	0.987	ggcgccaCGCGcaggct
V$KLFS/KLF2.01	0.986	cgccgGGGTggggctgagc
V$MYT1/MYT1L.01	0.986	agaaAGTTttcca
V$MOKF/MOK2.02	0.984	ctctccttagggtCCTTttcc
V$BEDF/ZBED4.01	0.982	ggtGGGCggcggcga
V$GLIF/ZIC2.01	0.982	accaccaCCCCacgccc
V$KLFS/BKLF.02	0.982	cgccggGGTGgggctgagc
V$KLFS/GKLF.02	0.981	agaggcttAAAGgaggggc
V$MIZ1/MIZ1.01	0.981	cctgcCCTCtc
V$NKXH/NKX25.05	0.981	ggcggTGAGtggggagcat
V$E2FF/E2F1.01	0.978	acccGGCGggagcgga
V$HESF/DEC2.01	0.978	ggaccaCGTGtgtgc
V$ZICF/ZIC3.03	0.976	agacccCAGCaggcaccatc
V$MAZF/MAZ.01	0.973	ccgagcGAGGggaggtggaggttggggagcg
V$CIZF/NMP4.01	0.972	agAAAAagacg
V$HIFF/HRE.02	0.972	acggaccaCGTGtgtgc
V$FKHD/HFH2.01	0.971	tcaaaaccAACAttctc
V$STAT/STAT3.02	0.971	cggcTTCCaggcagagcac
V$ZICF/ZIC1.01	0.969	agacccCAGCaggcaccatc
V$HICF/HIC1.02	0.968	cagTGCCccaact
V$EBOX/NMYC.02	0.967	cggaccaCGTGtgtgcg
V$ETSF/FLI1.03	0.967	tgttggctGGAAgtgggtgac

V$OSRF/OSR2.01	0.967	ctacgGTAGctca
V$EBOX/CMYC.02	0.965	cggaccaCGTGtgtgcg
V$EGRF/WT1.01	0.965	gtgtgCGGGggcgacagcg
V$SAL2/SALL2.01	0.964	ccggGTGGgag
V$EGRF/EGR2.02	0.963	ctgggTGGGcggcggcgac
V$NKXH/NKX32.01	0.963	ggcggtgAGTGgggagcat
V$SMAD/GC_SBE.01	0.963	gaaggCGCCag
V$CP2F/CP2.01	0.962	aaCTGGctcgcgcccgcgc
V$AHRR/AHRARNT.03	0.961	ctggtcccggGCGTgggtggtggt
V$SMAD/SMAD.01	0.96	aacGTCTgcct
V$WHNF/WHN.01	0.96	gagACGCaatg
V$GLIF/GLI3.02	0.959	ccgcCCACccagtgccc
V$OSRF/OSR1.01	0.959	ctacgGTAGctca
V$IRFF/IRF3.01	0.958	gacgcaatgaaaaaGAAAgttttcc
V$ZICF/ZIC3.01	0.958	agacccCAGCaggcaccatc
V$MYBL/CMYB.01	0.957	ccagtgcccCAACtggctcgc
V$ETSF/GABPA.02	0.956	tgttggctGGAAgtgggtgac
V$SF1F/FTF.01	0.956	tttcCAAGgcggtga
V$BCDF/PTX1.01	0.955	gccgCTAAccccggagc
V$EBOX/MAX.01	0.955	cggacCACGtgtgtgcg
V$IRFF/IRF7.01	0.955	gacgcaatGAAAaagaaagttttcc
V$BZIP/CEBPE_ATF4.01	0.954	cgtggcGCAAcgc
V$STAT/STAT5.01	0.954	ggagTTCAgagaaaaagac
V$SF1F/SF1.01	0.953	gctgCAAGgcgaacc
V$E2FF/E2F6.01	0.952	gcggcGCGGgaggttgg
V$HIFF/ARNT.01	0.952	cgcacacaCGTGgtccg
V$KLFS/EKLF.02	0.95	cgcccggGGTGggctgagc
V$NACA/NACA1.01	0.95	agcgCAGAgattg
V$ETSF/ETS1.01	0.949	tgttggctGGAAgtgggtgac
V$HIFF/CLOCK_BMAL1.01	0.949	acggacCACGtgtgtgc
V$ZF02/ZBTB7.03	0.949	ctcagCCCCaccccggcggcgct
V$ZF03/ZNF217.01	0.949	GAATgttggtttt
V$EBOX/ATF6.01	0.948	ggcgCCACgcgcaggct
V$NFAT/NFAT5.02	0.948	ggctGGAAaactttctttt
V$XBBF/RFX5.01	0.948	gctttggaccAGCAaccta
V$ETSF/FLI1.02	0.946	tgttggctGGAAgtgggtgac
V$NRF1/NRF1.01	0.946	gctcGCGCccgcgcagc
V$HIFF/ARNTL.01	0.945	tctagtcaAGTGttcgg
V$SP1F/GC.01	0.943	cgcccggGGTGggctga
V$MYBL/VMYB.05	0.941	ctggtgtccaAACGggtctcg
V$AHRR/AHRARNT.01	0.94	ccgtggggtcgCGTGgagttcagag
V$NDPK/NM23.01	0.94	ccGGGGtggggctgagc
V$BEDF/ZBED4.02	0.937	tgtgtgcGGGGgcga
V$CHRF/CHR.01	0.937	gaacTTGAatcgg
V$ETSF/ETV3.02	0.937	tgttggctGGAAgtgggtgac
V$CP2F/TCFCP2L1.01	0.936	aaCTGGctcgcgcccgcgc
V$AP2F/TCFAP2A.01	0.935	ggacccggCGGGgag
V$MYBL/MYBL1.02	0.935	ccagtgccccAACTggctcgc
V$ZF02/ZBP89.01	0.934	ctctaactCCCCcatggagtcgg
V$ETSF/CETS1P54.01	0.933	tgttggCTGGaagtgggtgac

V$PLAG/PLAG1.01	0.933	taGAGGcttaaaggagggcgcc
V$SP1F/SP4.01	0.933	ccagggGGCGatgcccg
V$BTBF/KAISO.02	0.932	cctctCGCGgcccct
V$HESF/HES1.01	0.932	ctggctcGCGCccgc
V$ETSF/ELF2.01	0.931	tgttggctGGAAgtgggtgac
V$ETSF/ELK1.03	0.931	tgttggctGGAAgtgggtgac
V$ETSF/GABP.01	0.931	tgttggctGGAAgtgggtgac
V$KLFS/KLF12.01	0.931	ccagccgGTGGgtcctctg
V$CDXF/CDX2.03	0.93	gaccgcttTTATagagggt
V$HESF/DEC1.02	0.929	ggaccaCGTGtgtgc
V$ABDB/HOXD13.01	0.928	accctctaTAAAagcgg
V$EBOX/USF.04	0.928	gcacaCACGtggtccgt
V$ETSF/SPDEF.01	0.928	tcccaccCGGAtttctaaaag
V$HIFF/HRE.01	0.928	acggaccaCGTGtgtgc
V$LTFM/LACTOFERRIN.01	0.926	gGCACtggg
V$SP1F/SP1.03	0.926	tgggtGGGCggcggcga
V$HOMF/TLX1.01	0.925	tccaaggCGGTgagtgggg
V$EBOX/MYCMAX.01	0.923	gcacaCACGtggtccgt
V$XBBF/RFX3.02	0.923	gctctggcctgGAGAcgca
V$ZF07/ZNF263.01	0.923	cctccacCTCCcctcgc
V$HESF/HELT.01	0.922	actcCACGcgacccc
V$MAZF/MAZR.01	0.922	gccggggtGGGGctgagcgctgcaagccggg
V$IRFF/ISRE.02	0.921	gacgcaatgaaaaaGAAAgttttcc
O$MTEN/HMTE.01	0.92	gaAGCTgggcgaaagagaggt
V$NF1F/NFIB.01	0.918	cagtgggcgttgcGCCAcgat
V$AP1F/JUNB.01	0.917	ctctaaGTCAccc
V$INSM/INSM1.01	0.916	tgtgcGGGGgcga
V$HOMF/HMX3.01	0.915	ttggctggAAGTgggtgac
V$ZF02/ZNF219.01	0.913	tgtcgCCCCcgcacacacgtggt
V$ETSF/ELF1.01	0.912	tgttggctGGAAgtgggtgac
V$CTCF/CTCFL.01	0.911	ccgtttggacaccagggGGCGatgccc
V$GATA/GATA2.02	0.911	gagcGATAggatt
V$MYBL/VMYB.02	0.911	ctggtgtccaAACGggtctcg
V$GLIF/GLI3.01	0.91	ccgcCCACccagtgccc
V$AP2F/AP2.01	0.908	ctgGCCTggagacgc
V$NOLF/OLF1.02	0.908	acagggTCCCacgggctcttcag
V$GCMF/GCM1.02	0.907	tcccaCCCGgatttc
V$AHRR/NXF_ARNT.01	0.906	tgcctggaagcCGTGgggtcgcgtg
V$EGRF/EGR2.01	0.906	cactGGGTgggcggcggcg
V$HIFF/HIF1.01	0.904	cgcacacACGTggtccg
V$ZF04/ZID.01	0.904	ggGCTCtaagacc
V$IRFF/IRF2.01	0.903	gacgcaatgaaaaaGAAAgttttcc
V$KLFS/EKLF.01	0.901	cagcgccgccgGGGTgggg
V$ZF10/PRDM14.01	0.901	gtcTTAGagcccgcc
V$GLIF/GLIS2.01	0.9	gcgaCCCCacggcttcc
V$GLIF/GLI1.01	0.899	ccgcccaCCCAgtgccc
V$SP1F/SP1.01	0.899	cgccgGGGTggggctga
V$XBBF/RFXDC2.01	0.899	cacgctgggtagCAACgtc
V$YBXF/YB1.01	0.899	gcgaaTGGCcagg
V$NOLF/EBF1.01	0.898	acagggTCCCacgggctcttcag

V$SP1F/SP1.02	0.897	tgggtGGGCggcggcga
O$VTBP/MTATA.01	0.896	ctctaTAAAagcggtcc
V$CREB/XBP1.01	0.896	cacggaccACGTgtgtgcggg
V$GLIF/GLI2.01	0.895	ccgcCCACccagtgccc
V$MYBL/VMYB.03	0.895	ctggtgtccaAACGggtctcg
V$E2FF/E2F2.01	0.893	gcggtGCGCgctaggtt
V$STAT/STAT6.01	0.891	cagcTTCTctagaactaca
V$EGRF/CKROX.01	0.89	cggcgcGGGAggttggaga
V$ETSF/ELK1.04	0.89	tgttggctGGAAgtgggtgac
V$PRDF/PRDM1.02	0.888	tgaaaaaGAAAgtttcca
V$IRFF/IRF1.01	0.887	gacgcaatgaaaaaGAAAgttttcc
V$MYOD/MYOD.01	0.887	gcgaGCCAgttgggca
V$CDEF/CDE.01	0.885	gcggCGCGggagg
V$GLIF/GLIS3.01	0.885	ccaaCCTCccgcgccgc
V$SREB/SREBP.02	0.885	ccaCCACcccacgcc
V$EGRF/EGR1.01	0.884	gtgtgcgggGGCGacagcg
V$RBPF/RBPJK.01	0.883	cccgTGGGaccct
V$PAX6/PAX6.02	0.881	tcgcgaatggCCAGgcccg
V$ETSF/ETV4.01	0.88	tcccaccCGGAtttctaaaag
V$KLFS/KLF7.02	0.88	aagctgGGCGaaagagagg
V$ETSF/ELK4.02	0.879	tgttggctGGAAgtgggtgac
V$IRFF/IRF2.02	0.879	gacgcaatGAAAaagaaagttttcc
V$HDBP/HDBP1_2.01	0.878	cccaccCCGGcggcgctgt
O$TF3C/TFIIIC.01	0.877	GGTTggagact
V$E2FF/E2F7.02	0.877	gcggcgcGGGAggttgg
V$NFKB/NFKAPPAB65.01	0.875	ctgtgactTTCCaag
V$EGRF/NGFIC.01	0.874	cactGGGTgggcggcggcg
V$DMTF/DMP1.01	0.873	cacccGGATttctaa
V$GCMF/GCM1.01	0.873	taactCCCCcatgga
V$PLAG/PLAGL1.01	0.872	gttagGCGCccctcctttaagcc
V$EGRF/EGR1.04	0.871	cgcggcgcgGGAGgttgga
V$STAT/STAT.01	0.871	ggagttcagAGAAaaagac
V$NFKB/NFKAPPAB.02	0.87	ctGTGActttccaag
V$ZF15/ZSCAN10.01	0.869	gctcgcgccCGCGcagc
V$E2FF/E2F3.01	0.868	gcggtGCGCgctaggtt
V$GLIF/GLI1.02	0.867	ccgcCCACccagtgccc
V$STAT/STAT5A.01	0.867	ggagTTCAgagaaaaagac
V$BRAC/TBR1.01	0.866	gtgttggagagGAGTgaacttgaatcggg
V$EGRF/EGR3.01	0.865	cactGGGTgggcggcggcg
V$HOMF/HMX2.02	0.865	ctctctccAAACgctcccc
V$NFKB/NFKAPPAB50.01	0.865	gggGCGAtgcccgac
V$RREB/RREB1.01	0.864	cCCCAacctccacct
V$XBBF/RFX3.01	0.863	cacgctgggtagCAACgtc
V$PLZF/PLZF.01	0.862	atcTAGAggccgagcgagggg
V$BNCF/BNC.01	0.858	ggggtggtggTGTCccgaa
V$HNF1/HMBOX.01	0.856	gcactctaGTCAagtgt
O$XCPE/XCPE1.01	0.855	ggCCGGgagcg
V$CEBP/CEBPG.01	0.855	ggcgTTGCgccacga
V$HAML/AML3.01	0.854	gtccGTGGttaggcg
V$XBBF/RFX4.01	0.854	cacgctgggtagCAACgtc

V$AP1F/FOSL1.01	0.853	ctctaAGTCaccc
V$HESF/BHLHB2.01	0.853	cacacACGTggtccg
V$CTCF/CTCF.01	0.852	cgcccgcgcagcttCGGGtcgcgaatg
V$ZF22/ZNF704.01	0.852	gaCCAGccggtgg
V$NFKB/HIVEP1.01	0.851	ctGTGActttccaag
V$GREF/GRE.01	0.85	tgggaccctgtaGTTCtag
V$HOXF/HOXB8.01	0.85	agacgcaATGAaaaagaaa
V$ZF20/ZNF771.01	0.849	gccgctaACCC
V$RXRF/RAR_RXR.03	0.848	cgtggGGTCgcgtggagttcagaga
V$RXRF/RXRA.01	0.848	ttgcagcgcaggAGTCagtgggcgt
V$EBOX/USF2.01	0.847	aactccaCGCGaccccca
V$ETSF/ERG.01	0.847	tgttggcTGGAagtgggtgac
V$NR2F/TR2.01	0.847	tggggtcgcgtggaGTTCagagaaa
V$ZF57/ZFP57.01	0.846	gctTGCAgcgctc
V$ZF26/ZNF396.01	0.845	ctgggCGAAag
V$EBOX/USF1.02	0.843	ctagTCAAgtgttcggg
V$HNFP/MIZF.01	0.843	ggCAGAcgttgctacccag
V$ETSF/ETV1.01	0.842	tgttggcTGGAagtgggtgac
V$PEG3/PEG3.01	0.842	caccgccgagTGGCg
V$ETSF/ELF4.01	0.841	tgttggctGGAAgtgggtgac
V$NOLF/OLF1.01	0.841	gccgacTCCAtggggagttaga
V$PARF/HLF.01	0.841	gatcgtggcGCAAcgcc
V$ETSF/EHF.01	0.84	tgttggcTGGAagtgggtgac
V$EREF/ER.03	0.839	tcgggacaccacCACCcca
V$P53F/TP63.01	0.838	ctaccgtagtgcagaCTTGtgttaa
V$YY1F/YY2.02	0.838	gcctggCCATtcgcgacccgaag
V$NR2F/HNF4A.03	0.836	agtggacgagccagGTCCaccaggc
V$PAX5/PAX5.01	0.834	ctagcgCGCAccgccgagtggcgaggccg
V$ETSF/ETV5.01	0.831	tcccaccCGGAtttctaaaag
V$PRDF/BLIMP1.01	0.83	tgaaaaaGAAAgttttcca
V$ZTRE/ZTRE.01	0.829	cgccgccgGGGTggggc
V$ETSF/FLI1.01	0.828	tcgcgaCCCGaagctgcgcgg
V$LTSM/LTSM.01	0.825	cagcaggcacCATCc
V$IRFF/ISRE.01	0.824	gacgcaatgaaaaaGAAAgtttcc
V$LEFF/LEF1.03	0.824	ctcggaTCGAaggcgcc
V$ETSF/ETV6.01	0.823	tgttggctGGAAgtgggtgac
V$GRHL/GRHL3.01	0.822	gggaccGCTTtta
V$XBBF/RFX3.03	0.818	gacgTTGCtacccagcgtg
V$MYBL/MYBL1.01	0.816	ctggtgtccaaACGGgtctcg
V$ETSF/GABPA.01	0.815	tgttggcTGGAagtgggtgac
V$E2FF/E2F.01	0.814	cctaggctgGAAAactt
V$ZF07/ZNF444.01	0.814	acctCCACctcccctcg
V$HOMF/HMX1.01	0.813	ttcttttTCATtgcgtctc
V$PRDM/PRDM5.01	0.813	ctggaagccgTGGGgtcgcgtggagttcaga
V$SNAP/PSE.01	0.811	cTCGCctaggctggaaaac
V$ETSF/GABPB1.01	0.81	tgttggcTGGAagtgggtgac
V$HOMF/HMX2.03	0.81	ttctttTTCAttgcgtctc
V$STEM/OCT3_4.01	0.81	ttcttttTCATtgcgtctc
V$ZF02/ZNF202.01	0.809	cgctccCCAAcctccacctcccc
V$NGRE/IR2_NGRE.01	0.805	gaCTCCatgggggag

V$HNFP/HINFP.01	0.803	ggTGGAcctggctcgtcca
V$MYT1/MYT1.01	0.801	ttcAAGTtcactc
V$SP1F/SP2.01	0.801	cgccgggtgGGGCtga
V$PAX9/PAX9.01	0.8	cctcgccactcggcGGTGcgc
V$ZF16/PRDM15.01	0.8	aaaggctgtggCGGA
V$GLIF/ZIC3.02	0.799	tccgcTCCCcgccgggt
V$ZF01/ZBRK1.01	0.799	ccgggagcGCAGagattgtgcgtct
V$ZF41/ZFP41.01	0.798	ctctaaCTCCccc
V$OCT1/OCT.01	0.796	gcaATGAaaaagaaa
V$PAX3/PAX3.01	0.793	cTCGCcactcggcggtgcg
V$STAF/STAF.01	0.789	tctccaaacgctcCCCAacctccacctcccc
V$BCL6/BCL6.02	0.787	agcctttTAGAaatccg
V$E2FF/E2F7.01	0.784	tgttGGCTggaagtggg
V$STAT/STAT1.01	0.784	cggcttccaGGCAgagcac
V$ZF12/ZFP652.01	0.784	cttagaggcTTAAag
V$RXRF/VDR_RXR.03	0.782	cgcgaGGCTcagaggacccaccggc
V$STAF/ZNF76_143.01	0.781	gcctctagatgctCCCGggctccctagtgtt
O$MTEN/DMTE.01	0.779	caatcctATCGctctggcctg
V$HUB1/ZNF282.01	0.778	TTTCtctgaactcca
V$XBBF/MIF1.01	0.777	cacgctgggtaGCAAcgtc
V$AHRR/AHRARNT.02	0.775	tccagcctgcGCGTggcgccgctgt
V$DEAF/NUDR.01	0.775	gccTCGGcggcctccagcc
V$RXRF/VDR_RXR.06	0.773	cgtggggtcgcgtggAGTTcagaga
V$XBBF/RFX4.03	0.769	gacGTTGctacccagcgtg
V$NR2F/NR2F6.01	0.765	ccgcgaggctcaGAGGacccaccgg
V$CREB/TAXCREB.02	0.761	agatcgTGGCgcaacgcccac
V$GLIF/GLIS1.01	0.76	cagccccacCCCGgcgg
V$MOKF/MOK2.01	0.76	tccccactcaccgCCTTggaa
V$ZICF/ZIC5.01	0.758	ttggacaccaGGGGgcgat
V$NR2F/TR4.02	0.754	cttggaAAGTcacaggagaaggcgg
V$ESRR/ESRRA.04	0.753	gagtggacgagcCAGGtccacca

Table A2. Complete table of TFBSs of *Mus musculus* Cebpa Upstream CNS1 region

Matrix	strand	Matrix sim.	Sequence
V$PLAG/PLAG1.02	(+)	1	gaGGGGgggcctctttccagtca
V$NOLF/EBF1.01	(-)	0.996	atgtggTCCCcaggggcttggct
V$RUSH/SMARCA3.02	(+)	0.993	aggcACTTttt
V$MOKF/MOK2.02	(-)	0.988	gaacaaaaaagtgCCTTggaa
V$NEUR/NGN_NEUROD.01	(+)	0.988	accaCATCtgcctga
V$ZF07/ZNF263.02	(-)	0.988	cccctcCTCCcctgg
V$NOLF/OLF1.02	(-)	0.977	atgtggTCCCcaggggcttggct
V$SMAD/SMAD.01	(-)	0.976	cttGTCTgtcc
V$AP2F/AP2.02	(+)	0.97	tctGCCTgaggggg
V$AP2F/TCFAP2A.02	(+)	0.964	agcccCCTGgggacca
V$AP2F/TCFAP2C.01	(+)	0.961	tctGCCTgaggggg
V$SF1F/FTF.01	(+)	0.957	cttcCAAGgcacttt
V$MAZF/MAZR.01	(+)	0.949	tgagggGGGGcct
V$MYOD/TCFE2A.02	(-)	0.945	ctcaggcaGATGtggtc
V$AP2F/TCFAP2E.01	(+)	0.932	tctgcctgAGGGggg
V$BEDF/ZBED4.02	(+)	0.931	tgcctgaGGGGgggc
V$MAZF/MAZ.01	(+)	0.93	agggGAGGagggg
V$CARE/CARF.01	(-)	0.928	ggaaaGAGGcc
V$LTFM/LACTOFERRIN.01	(+)	0.928	gGCACtttt
V$ZF02/ZBTB7.01	(-)	0.927	gaggccCCCCctcaggcagatgt
V$SORY/SOX4.01	(-)	0.922	gctgaACAAaaaagtgccttgga
V$MYOD/E47.01	(-)	0.92	ctcagGCAGatgtggtc
V$OCT1/OCT2.02	(+)	0.917	ggtaTGCAgaacagg
V$STEM/OCT3_4.02	(-)	0.916	ctgttctGCATacccctcc
V$PAX6/PAX4_PD.01	(+)	0.913	tagGCACcgctgcattcag
V$ESRR/ESRRA.01	(+)	0.906	aacaggacagacAAGGtcgctgc
V$GLIF/GLIS2.01	(-)	0.905	aggcCCCCcctcaggca
V$HNF1/HMBOX.01	(-)	0.904	aaatacttGTTAaatga
V$INSM/INSM1.01	(+)	0.9	tgaggGGGGgcct
V$SORY/SOX11.01	(-)	0.9	gctgaACAAaaaagtgccttgga
V$GCMF/GCM1.01	(-)	0.897	gccccCCCTcaggca
V$SORY/HBP1.01	(-)	0.895	gaacacagAATGaactatgtggc
V$RXRF/RAR_RXR.03	(+)	0.888	acataGTTCattctgtgttcattta
V$AP2F/TCFAP2B.01	(+)	0.887	caaGCCCctggggac
V$GLIF/GLIS3.01	(-)	0.886	aggcCCCCcctcaggca
V$RP58/RP58.01	(+)	0.884	accaCATCtgcct

V$GREF/GRE.01	(-)	0.869	cttgtctgtcctGTTCtgc
V$OCT1/POU3F3.01	(-)	0.863	gttctGCATacccct
V$EGRF/EGR1.03	(+)	0.859	gcctgagGGGGggcctctt
V$NFAT/NFAT5.01	(-)	0.859	gactGGAAagaggcccccc
V$GREF/PRE.01	(-)	0.857	cttgtctgtccTGTTctgc
V$IRXF/IRX4.01	(-)	0.857	aataCTTGttaaa
V$NKXH/NKX31.01	(+)	0.856	tttaacAAGTatttgaggc
V$STAT/STAT6.01	(-)	0.851	aaagTGCCttggaaggagg
V$EVI1/EVI1.02	(-)	0.838	ctgagAAGAgagcctca
V$ZF01/SZF1.01	(+)	0.838	ccaGGGGaggaggggtatgcagaac
V$FAST/FAST1.01	(+)	0.835	ttctgtgttCATTtaac
V$HNF1/TCF2.01	(-)	0.829	aaatacttGTTAaatga
O$VTBP/ATATA.01	(+)	0.822	aagtattTGAGgctctc
V$GREF/ARE.01	(-)	0.819	cttgtctgtccTGTTctgc
V$NR2F/COUPTFII.01	(-)	0.806	agcagctacccagaGGTCctcctag
V$DICE/DICE.01	(+)	0.805	ggctCTCTtctcagc
V$SP1F/SP2.01	(+)	0.803	cctgaggggggGGCCtct
V$ZF13/ZNF435.01	(-)	0.802	ctTGTTaaatgaacacaga
V$BCL6/BCL6.02	(-)	0.799	gtcctccTAGAactcag
V$NR2F/TR2.01	(+)	0.796	atagttcattctgtGTTCatttaac
V$OCT1/OCT.01	(+)	0.793	ggtATGCagaacagg
V$ZF16/PRDM15.01	(-)	0.79	aaaaagtgcctTGGA
V$PAX9/PAX9.01	(-)	0.788	agactgaatgcagcGGTGcct
V$DMRT/DMRT1.01	(+)	0.779	gctttctgccacaTAGTtcat
V$BRN5/BRN5.01	(+)	0.768	ctgccaCATAgttcattctgtgt
V$SNAP/PSE.01	(-)	0.757	cTCCCctggactgggacag
V$ZTRE/ZTRE.01	(+)	0.755	cagtccagGGGAggagg
V$PRDM/PRDM4.01	(+)	0.75	gcctcTTTCcagtcagccctaggcaccgc

Table A3. Complete table of TFBSs of *Mus musculus* Cebpa Upstream CNS2 region

Matrix	strand	Matrix sim.	Sequence
V$MYOD/TCF12.01	(-)	1	gaacagcaGGTGcaaag
V$MZF1/MZF1.01	(-)	1	gtGGGGagagc
V$TALE/TGIF.01	(+)	1	ctccaacctGTCAaaca
V$FKHD/FOXJ3.01	(-)	0.999	cccaggTAAAcaaacct
V$TALE/MEIS1.01	(+)	0.999	ctccaaccTGTCaaaca
V$FKHD/FOXP1.01	(-)	0.996	cccaggtAAACaaacct
V$HAND/TAL1_E2A.02	(-)	0.996	gagggaaCAGCaggtgcaaag
V$FKHD/FREAC2.01	(-)	0.994	cccaggTAAAcaaacct
V$NEUR/MASH1.01	(+)	0.994	tttgcacCTGCtgtt
V$SF1F/SF1.01	(-)	0.993	tggcCAAGgccatcc
V$SORY/SOX3.01	(-)	0.992	tggtaaCAAAggacaagggtgtt
V$FKHD/FOXP2.01	(-)	0.991	cccaggtaAACAaacct
V$KLFS/BKLF.02	(-)	0.991	agctagGGAGggcctaga
V$SORY/SOX2.01	(-)	0.991	tggtaACAAaggacaagggtgtt
V$FKHD/FOXJ1.01	(-)	0.99	cccaggtaAACAaacct
V$GATA/GATA1.06	(-)	0.99	cctaGATAaggct
V$SORY/SOX6.01	(-)	0.988	tggtaACAAaggacaagggtgtt
V$FKHD/HNF3.01	(-)	0.986	cccaggtAAACaaacct
V$ZTRE/ZTRE.03	(+)	0.985	ccCTCCctagctctccc
V$TALE/MRG2.01	(+)	0.984	ctccaacctGTCAaaca
V$GATA/GATA2.03	(-)	0.983	cctaGATAaggct
V$MYOD/TCFE2A.01	(-)	0.983	gaacagcaGGTGcaaag
V$GATA/GATA4.01	(-)	0.98	cctaGATAaggct
V$SORY/SOX4.01	(-)	0.98	tggtaACAAaggacaagggtgtt
V$GATA/GATA6.02	(-)	0.977	cctaGATAaggct
V$FKHD/FKHRL1.01	(-)	0.976	cccaggtaAACAaacct
V$FKHD/FOXO1.01	(-)	0.974	cccaggtaAACAaacct
V$TALE/MRG1.01	(+)	0.974	ctccaacctGTCAaaca
V$SORY/SOX11.01	(-)	0.97	tggtaACAAaggacaagggtgtt
V$FKHD/HNF3B.03	(-)	0.967	cccaggTAAAcaaacct
V$MYOD/SGN1.01	(-)	0.967	gaacagCAGGtgcaaag
V$SF1F/FTF.02	(-)	0.965	tggccAAGGccatcc
V$FKHD/FOXA1.01	(-)	0.961	cccaggtaAACAaacct
V$KLFS/GKLF.02	(-)	0.961	ctggtaacAAAGgacaagg
V$KLFS/IKLF.01	(-)	0.961	agctaGGGAggggcctaga
V$TALE/PKNOX1.01	(+)	0.954	ctccaacctGTCAaaca

V$SPZ1/SPZ1.01	(-)	0.952	aGGAGggaaca
V$GREF/GRE.03	(+)	0.951	tttgcacctgcTGTTccct
V$FKHD/FOXK1.01	(-)	0.946	cccaggtaAACAaacct
V$XBBF/RFX4.02	(+)	0.945	ccttgtcctttGTTAccag
V$GATA/GATA.01	(-)	0.944	cctaGATAaggct
V$SORY/SOX9.02	(-)	0.943	tggtaACAAaggacaagggtgtt
V$KLFS/EKLF.01	(-)	0.941	atctacagacaGGGTtggg
V$ZFHX/AREB6.01	(+)	0.936	tgtttACCTggga
V$GATA/GATA3.01	(+)	0.934	gtgaGATAaagca
V$ZF02/ZBTB7.03	(+)	0.934	ctaggCCCCtccctagctctccc
V$TALE/PKNOX2.01	(+)	0.932	ctccaacctGTCAaaca
V$LEFF/LEF1.01	(-)	0.93	ctggtaaCAAAggacaa
V$XBBF/RFX3.02	(+)	0.929	ccttgtcctttGTTAccag
V$GREF/PRE.01	(-)	0.927	ggagcatggccTGTTctgg
V$NF1F/NF1.04	(-)	0.927	ggttgggttgatggCCAAggc
V$GREF/ARE.04	(+)	0.925	tttgcacctgctGTTCcct
V$ESRR/ESRRA.03	(-)	0.924	gggttgatggccAAGGccatccc
V$FKHD/FREAC3.01	(-)	0.923	cccagGTAAacaaacct
V$NR2F/HNF4A.01	(-)	0.922	gttctggtaaCAAAggacaagggtg
V$TALE/TGIF1.01	(+)	0.921	ctccaacctGTCAaaca
V$TALE/TGIF2.01	(+)	0.92	ctccaacctGTCAaaca
V$MAZF/MAZR.01	(-)	0.917	tagggaGGGGcct
V$GREF/GRE.02	(-)	0.914	aggGAACagcaggtgcaaa
V$SP1F/GC.01	(-)	0.914	agctagGGAGggccta
V$FKHD/XFD1.01	(-)	0.912	cccaggTAAAcaaacct
V$FKHD/FREAC4.01	(-)	0.911	cccaggtaAACAaacct
V$ZF05/ZFP410.01	(+)	0.911	ttacctgGGATggcc
V$NF1F/NFIB.01	(-)	0.909	ggttgggttgatgGCCAaggc
V$FKHD/FOXL1.01	(-)	0.906	cccaggTAAAcaaacct
V$MYOD/MYF5.01	(-)	0.904	agggaaCAGCaggtgca
V$FKHD/HFH1.01	(-)	0.898	cccaggTAAAcaaacct
V$GREF/GRE.01	(-)	0.897	ggagcatggcctGTTCtgg
V$GRHL/GRHL2.01	(-)	0.896	caAACCtgcttta
V$LTSM/LTSM.01	(+)	0.896	gatccagcctTATCt
V$ZICF/ZIC2.02	(-)	0.894	gggaaCAGCaggtgc
V$ESRR/ERR.01	(-)	0.893	gggttgatggccAAGGccatccc
V$GATA/GATA5.01	(-)	0.89	cctaGATAaggct
V$STAT/STAT.01	(-)	0.89	cctgttctgGTAAcaaagg

V$FKHD/XFD3.01	(-)	0.889	cccaggtaAACAaacct
V$GREF/MRE.01	(-)	0.889	agggaacagcagGTGCaaa
V$PBXC/PBX1_MEIS1.01	(-)	0.889	agggTGTTtgacaggtt
V$BARB/BARBIE.01	(+)	0.883	agatAAAGcaggttt
V$PERO/PPARG.03	(-)	0.883	tgttctggtaacAAAGgacaagg
V$ZICF/ZIC3.01	(-)	0.881	gggaaCAGCaggtgc
V$CP2F/TCFCP2L1.01	(-)	0.873	gcCTAGataaggctggatc
V$ESRR/ESRRB.01	(-)	0.872	gggttgatggccaAGGCcatccc
V$NBRE/NBRE.01	(-)	0.869	aacaAAGGacaaggg
V$NR2F/COUPTFII.01	(-)	0.868	gttctggtaacaaaGGACaagggtg
V$GRHL/GRHL1.01	(-)	0.866	ttgacaGGTTgga
V$GRHL/GRHL3.01	(-)	0.853	ttgacaGGTTgga
V$NR2F/HNF4.01	(-)	0.85	acagcaggtgCAAAggtggagaggc
V$SRFF/SRF.02	(+)	0.85	gttacCAGAacaggccatg
V$SREB/SREBP.02	(+)	0.849	agcTCTCcccacact
V$RXRF/THRA.01	(-)	0.848	agcatggcctgttcTGGTaacaaag
V$NGRE/IR1_NGRE.01	(+)	0.847	atgctccgTGAGata
V$HOXH/MEIS1A_HOXA9.01	(-)	0.845	TGACaggttggagtg
V$ABDB/HOXC9.01	(+)	0.843	cttggccaTCAAcccaa
V$RXRF/THRB.01	(-)	0.834	agcatGGCCtgttctggtaacaaag
V$NR2F/COUP.01	(-)	0.83	acagcaggtgcaaAGGTggagaggc
V$NR2F/HPF1.01	(-)	0.816	gttctggtaacAAAGgacaagggtg
V$PAX5/PAX5.01	(-)	0.816	tctacaGACAgggttgggttgatggccaa
V$RREB/RREB1.01	(+)	0.814	cCCCAcactccaacc
V$MYOD/TCF21.01	(-)	0.811	agggaacaGCAGgtgca
V$HOXH/MEIS1B_HOXA9.01	(-)	0.793	TGACaggttggagtg
V$HNF1/TCF2.01	(+)	0.791	tgtcctttGTTAccaga
V$GABF/GAGA.01	(-)	0.787	tggggAGAGctagggaggggcctag
V$HOXC/PBX1.01	(-)	0.787	ggttgGGTTgatggcca
V$STAF/STAF.01	(+)	0.784	ctccctagctctcCCCAcactccaacctgtc
V$SORY/SOX10.03	(-)	0.781	gcTGGAtctacagacagggttgg
O$VTBP/ATATA.01	(-)	0.78	gcctagaTAAGgctgga
V$PERO/PPAR_RXR.01	(-)	0.779	gtaacaaaggacaaGGGTgtttg
V$ZF13/ZNF435.01	(+)	0.778	ttTGTTaccagaacaggcc
V$NR2F/NR2F6.01	(-)	0.771	acagcaggtgcaAAGGtggagaggc
V$PRDM/PRDM4.01	(+)	0.752	agcctTATCtaggcccctccctagctctc
V$MYBL/MYBL2.01	(+)	0.75	caacctgtcAAACacccttgt

Table A4. Complete table of TFBSs of *Mus musculus* Cebpa Upstream CNS3 region

Matrix	strand	Matrix sim.	Sequence
V$TAIP/CSRNP1.01	(-)	1	AGAGtga
V$TAIP/CSRNP3.01	(-)	1	aGAGTga
V$FKHD/HFH3.01	(-)	0.994	actaaatAAACaacctg
V$FKHD/HFH8.01	(-)	0.99	actaaatAAACaacctg
V$FKHD/FHXA.01	(-)	0.988	actaaatAAACaacctg
V$BRAC/EOMES.03	(-)	0.987	taaacaacctggTGTGagcagaggatcgc
V$BCDF/CRX.01	(-)	0.985	tactTAATcaaattct
V$FKHD/FREAC7.01	(-)	0.984	actaaaTAAAcaacctg
V$NEUR/NGN_NEUROD.01	(+)	0.982	cgtgCATCtggccca
V$SAL1/SALL1.01	(-)	0.975	gcATAAaaaagat
V$FKHD/HFH1.01	(-)	0.967	actaaaTAAAcaacctg
V$HOMF/BARX2.01	(-)	0.966	cagtgatTAATggctcaac
V$HAML/AML1.01	(-)	0.965	cactGCGGttacagt
V$RBP2/PLU1_JARID1B.01	(+)	0.962	GCACagcag
V$MYBL/CMYB.02	(+)	0.961	tctttactgTAACcgcagtgg
V$HOXC/HOXA9.01	(-)	0.96	tcagtGATTaatggctc
V$MEF2/MEF2A.01	(+)	0.96	catcaagctaaAAATatcttttt
V$WHNF/WHN.01	(-)	0.958	aggACGCatcc
V$FKHD/HNF3B.02	(-)	0.956	actaaaTAAAcaacctg
V$YY1F/REX1.01	(+)	0.955	gttgagCCATtaatcactgagaa
V$ETSF/PDEF.01	(+)	0.953	aaagtacaGGATgcgtccttc
V$HOMF/HHEX.01	(-)	0.953	atcctgtacttTAATcaaa
V$HOXC/HOXB4.02	(-)	0.952	tcagTGATtaatggctc
V$HOXF/NANOG.01	(+)	0.951	tatgttgAATGagaatttg
V$FKHD/XFD2.01	(-)	0.95	actaaaTAAAcaacctg
V$HESF/HES1.01	(+)	0.948	gagtcgcGTGCatct
V$MEF2/MEF2.07	(+)	0.948	catcaagctaaAAATatcttttt
V$SREB/SREBP.04	(+)	0.942	tgcTCACaccaggtt
V$FKHD/FREAC2.01	(-)	0.94	actaaaTAAAcaacctg
V$ETSF/ETS1.01	(+)	0.938	aaagtacaGGATgcgtccttc
V$FKHD/FOXJ3.01	(-)	0.938	actaaaTAAAcaacctg
V$NKX6/NKX61.01	(+)	0.938	gccaTTAAtcactga
V$LHXF/LMX1A.02	(-)	0.936	ctcagtgaTTAAtggctcaactg
V$ETSF/CETS1P54.01	(+)	0.932	aaagtaCAGGatgcgtccttc
V$FAST/FAST1.01	(+)	0.931	caggttgttTATTtagt
V$BRNF/BRN3.02	(+)	0.929	gagccatTAATcactgaga

V$HOXC/MEIS1.03	(-)	0.929	tcagtGATTaatggctc
V$SORY/HMGA.01	(+)	0.925	ttgagccattAATCactgagaat
V$CART/PHOX2.01	(-)	0.923	tactttTAATcaaattctcatt
V$CHRF/CHR.01	(-)	0.923	gctttTTGAacaga
V$GREF/ARE.02	(-)	0.923	aggacgcatcctGTACttt
V$FKHD/FOXJ1.01	(-)	0.917	actaaataAACAacctg
V$FKHD/FKHRL1.01	(-)	0.915	actaaataAACAacctg
V$CAAT/CAAT.01	(+)	0.914	tgagCCATtaatcac
V$HOXF/HOXD8.01	(-)	0.914	ctcagtgATTAatggctca
V$HOXC/HOXC9.02	(-)	0.912	tcagtgaTTAAtggctc
V$CP2F/CP2.01	(-)	0.91	acCTGGtgtgagcagagga
V$FKHD/FOXP2.01	(-)	0.909	actaaataAACAacctg
V$HOXC/HOX_PBX.01	(-)	0.907	tcagTGATtaatggctc
O$VTBP/VTATA.01	(-)	0.906	gtgaaTAAAaagcagat
V$NKX6/NKX63.01	(+)	0.905	gccatTAATcactga
V$ABDB/HOXC9.01	(+)	0.903	ttgagccaTTAAtcact
V$PLZF/PLZF.01	(+)	0.903	aagTACAggatgcgt
V$FKHD/FOXO1.01	(-)	0.902	actaaataAACAacctg
V$SORY/HBP1.01	(+)	0.901	atatgttgAATGagaatttgatt
V$DICE/DICE.01	(+)	0.888	tgcgCTCTgcaccga
V$FKHD/FOXL1.01	(-)	0.888	actaaaTAAAcaacctg
V$MTF1/MTF-1.01	(+)	0.888	ttatGCGCtctgcac
V$HOXF/HOXB8.01	(+)	0.887	ttgagccATTAatcactga
V$BARB/BARBIE.01	(-)	0.885	ataaAAAGcagatac
V$LHXF/ISL2.01	(+)	0.883	gcagttgagccATTAatcactga
V$CART/PROP1.02	(+)	0.882	aatgagaatttgATTAaagta
V$STAT/STAT6.01	(-)	0.882	cgacTTCGtctgaaggacg
V$ABDB/HOXB9.02	(-)	0.881	agagcgcaTAAAaaaga
V$GRHL/GRHL2.01	(-)	0.881	acAACCtggtgtg
V$MEF2/RSRFC4.01	(+)	0.881	catcaagctaaaaATATcttttt
V$FKHD/FOXK1.01	(-)	0.878	actaaataAACAacctg
V$CART/CART1.01	(-)	0.877	tactttTAATcaaattctcatt
V$FKHD/FREAC3.01	(-)	0.877	atacaCTAAataaacaa
V$IRFF/IRF7.01	(+)	0.877	aatatgttGAATgagaatttgatta
V$BRNF/BRN2.01	(-)	0.873	ctCATTcaacatattctca
V$HOMF/HMX1.01	(-)	0.873	cagtgatTAATggctcaac
V$HOMF/HMX3.03	(-)	0.871	cagtgatTAATggctcaac
V$LHXF/ISL1.01	(-)	0.871	ctcagtgatTAATggctcaactg

V$ABDB/HOXB13.01	(-)	0.87	agagcgcaTAAAaaaga
V$OCT1/OCT1.06	(+)	0.87	tgaatgagAATTtga
V$ABDB/HOXA9.02	(-)	0.865	agagcgcaTAAAaaaga
V$LEFF/LEF1.04	(+)	0.865	gagccaTTAAtcactga
V$LEFF/TCF7L1.01	(+)	0.863	atttgatTAAAgtacag
V$PARF/VBP.01	(+)	0.863	ttctttactGTAAccgc
V$PEG3/PEG3.01	(-)	0.863	cagtgattaaTGGCt
V$HOXF/HOXB7.01	(-)	0.862	ctcagtGATTaatggctca
V$LEFF/TCF7.01	(+)	0.86	atttgatTAAAgtacag
V$FKHD/XFD3.01	(-)	0.859	acagagtgAATAaaaag
V$LHXF/LHX6.01	(-)	0.858	cctgtacttTAATcaattctca
V$HOMF/HMX2.03	(-)	0.854	cagtgaTTAAtggctcaac
V$FKHD/FHXB.01	(+)	0.853	aagctaAAAAtatcttt
V$CART/RHOX6.01	(-)	0.85	ctcagTGATtaatggctcaac
V$CDXF/CDX2.02	(+)	0.85	tctgctttTTATtcactct
V$YY1F/YY2.02	(+)	0.849	gttgagCCATtaatcactgagaa
V$HOXC/PBX_HOXA9.01	(-)	0.847	tcagTGATtaatggctc
V$LHXF/LMX1B.02	(+)	0.846	atgagaatttGATTaaagtacag
V$CART/MIXL1.01	(-)	0.844	tactTTAATcaaattctcatt
V$HOXF/HOXC4.01	(-)	0.842	tactttAATCaaattctca
V$LEFF/TCF7L2.01	(+)	0.842	atttgatTAAAgtacag
V$HZIP/HOMEZ.01	(+)	0.84	aaaatATCTttttta
V$LHXF/LHX3.03	(-)	0.834	cctgtacttTAATCaaattctca
V$ABDB/HOXC13.02	(-)	0.833	agagcgcaTAAAaaaga
V$LHXF/LHX5.01	(-)	0.833	cctgtacttTAATCaaattctca
V$STEM/OCT3_4.01	(+)	0.833	ttcaaaaGCATcaagctaa
V$CART/UNCX4.01	(-)	0.829	tactttAATCaaattctcatt
V$PDX1/PDX1.01	(+)	0.827	ttgagccatTAATcactga
V$CART/PHOX2B.01	(-)	0.824	tactttAATCaaattctcatt
V$IRFF/ISGF3G.01	(+)	0.824	atgcgctctgcaccGAAAatgaagc
V$ETSF/FLI1.01	(+)	0.823	tctgcaCCGAaaatgaagcag
V$FKHD/FREAC4.01	(-)	0.82	actaaataAACAacctg
V$DEAF/NUDR.01	(+)	0.819	aagTCGGttctttactgta
V$EVI1/EVI1.05	(-)	0.819	taaaaaaGATAttttta
V$PAX7/PAX7.01	(+)	0.817	agaatttGATTaaag
V$PTF1/PTF1.01	(-)	0.815	ctcaACTGcttcattttcggt
V$HOXH/MEIS1B_HOXA9.01	(+)	0.807	TGAGaatttgattaa
V$DUXF/DUX4.01	(+)	0.806	tttGATTaaagtaca

V$BRNF/BRN4.02	(+)	0.805	gagccaTTAAtcactgaga
V$BRN5/BRN5.03	(-)	0.803	tTAATcaaattctcattcaacat
V$SORY/SOX1.04	(-)	0.802	atcAAATtctcattcaacatatt
V$HOXC/PBX1.01	(-)	0.795	tcagtGATTaatggctc
V$SORY/SOX21.03	(-)	0.783	agtGAATaaaaagcagatacact
V$HBOX/EN1.01	(+)	0.775	aatgagaaTTTGattaaag
V$DMRT/DMRT2.01	(-)	0.769	ctcaactgcTTCAttttcggt
V$AIRE/AIRE.01	(+)	0.757	gtTTATttagtgtat
V$ZTRE/ZTRE.01	(+)	0.755	ctggcccaGAGAgcgat

Table A5. Complete table of TFBSs of *Mus musculus* Cebpa Downstream CNS1 region

Matrix	strand	Matrix sim.	Sequence
V$PLAG/PLAG1.02	(-)	1	aaGGGGgaattcaaggatcattc
V$ZF02/ZBTB7.03	(-)	0.998	agggcCCCCtcccccgttcagcc
V$FKHD/FOXP2.01	(-)	0.997	tccctgtaAACAagaaa
V$KLFS/BKLF.02	(+)	0.997	aacgggGGAGggggccctc
V$EGRF/CKROX.01	(+)	0.996	gaacggGGGAggggccct
V$FAST/FAST1.03	(+)	0.996	ttcagtgtgGATTccag
V$MZF1/MZF1.02	(-)	0.995	tgGGGGaaggg
V$ZTRE/ZTRE.03	(-)	0.993	ccCTCCccgttcagcc
V$FKHD/FOXO1.01	(-)	0.991	tccctgtaAACAagaaa
V$FKHD/FOXP1.01	(-)	0.991	tccctgtAAACaagaaa
V$PURA/PURALPHA.01	(+)	0.991	ggAGGCggtggct
V$HAML/AML2.01	(+)	0.984	cttTGTGgttaagat
V$FKHD/HNF3.01	(-)	0.983	tccctgtAAACaagaaa
V$ZF07/ZNF263.01	(-)	0.982	ggccccCTCCcccgt
V$MIZ1/MIZ1.01	(+)	0.981	actgcCCTCtc
V$FKHD/FREAC2.01	(-)	0.978	tccctgTAAAcaagaaa
V$BCDF/PTX1.01	(-)	0.977	cagcCTAAgccctgttc
V$CIZF/NMP4.01	(-)	0.976	agAAAAactgg
V$ZF02/ZBP89.01	(-)	0.976	ggccccctCCCCcgttcagccta
V$SORY/SOX6.01	(-)	0.974	taaccACAAagcctggtcacact
V$EGRF/WT1.02	(+)	0.972	gaacgggGGAGggggccct
V$HAML/AML3.02	(+)	0.969	ctttGTGGttaagat
V$FKHD/FKHRL1.01	(-)	0.967	tccctgtaAACAagaaa
V$KLFS/GKLF.02	(-)	0.966	ctgggagcAAAGgggaat
V$NFAT/NFAT.01	(+)	0.965	ggcagaGGAAaggccaaag
V$TALE/MEIS1.01	(+)	0.963	gcttgcccTGTCagcct
V$FKHD/FOXA1.01	(-)	0.962	tccctgtaAACAagaaa
V$NFKB/CREL.01	(-)	0.961	tttggcctTTCCtct
V$ETSF/SPI1.02	(-)	0.96	gcagtgggGGAAggggttggt
V$KLFS/IKLF.01	(+)	0.958	aacggGGGAgggggccctc
V$FKHD/HNF3B.04	(-)	0.957	tccctgtaAACAagaaa
V$SORY/SOX9.02	(-)	0.952	taaccACAAagcctggtcacact
V$FKHD/FOXJ3.01	(-)	0.951	tccctgTAAAcaagaaa
V$SF1F/FTF.02	(-)	0.951	ctgccAAGGgcagcc
V$NACA/NACA1.01	(-)	0.95	aagaCAGAgaggg
V$CEBP/CEBPB.01	(-)	0.949	gtggttagGAAAgga

V$NR2F/HNF4G.01	(+)	0.945	gaggaaaggccAAAGcccagtgagg
V$KLFS/KLF6.01	(+)	0.943	tgaacgGGGGagggggccc
V$KLFS/EKLF.01	(-)	0.941	gaactcagagaGGGTttgc
V$NR2F/HNF4.01	(+)	0.937	gaggaaaggcCAAAgcccagtgagg
V$FKHD/FOXJ1.01	(-)	0.936	tccctgtaAACAagaaa
V$HAML/AML1.01	(+)	0.936	ctttGTGGttaagat
V$NR2F/HNF4A.01	(+)	0.935	gaggaaaggcCAAAgcccagtgagg
V$NFKB/NFKAPPAB65.01	(-)	0.932	tttggcctTTCCtct
V$ESRR/ESRRA.03	(-)	0.931	ctttcctctgccAAGGcagcca
V$KLFS/KLF12.01	(-)	0.931	gagggcaGTGGgggaaggg
V$MYBL/VMYB.05	(+)	0.931	gcttaggctgAACGggggagg
V$TEAF/TEAD.01	(-)	0.93	gatCATTccctgt
V$BEDF/ZBED4.02	(+)	0.928	cgggggaGGGGgccc
V$STAT/STAT5B.01	(+)	0.925	tttgctcccAGAAcagggc
V$ZF02/ZNF219.01	(-)	0.925	gagggCCCCctcccccgttcagc
V$AHRR/NXF_ARNT.01	(+)	0.924	ctggggcaggcCGTGgcttgcaggc
V$PLAG/PLAGL1.01	(-)	0.923	ggaggGCCCcctcccccgttcag
V$HASF/HAS.01	(-)	0.921	agcCACGgcct
V$MAZF/MAZR.01	(+)	0.92	gggggaGGGGgcc
V$PAX6/PAX4_PD.01	(-)	0.917	aggGCAGccatgggtgagg
V$PERO/PPARG.02	(+)	0.915	agaggaaaggccAAAGcccagtg
V$SP1F/SP1.03	(+)	0.912	aacggGGGAgggggccc
V$P53F/P53.02	(+)	0.911	gtgaggaggcggtggCTTGccctgt
V$EGRF/EGR1.04	(+)	0.91	cagtgaggaGGCGgtggct
V$MAZF/MAZ.01	(+)	0.91	ggggGAGGgggcc
V$INSM/INSM1.01	(+)	0.906	tgaacGGGGgagg
V$FKHD/XFD3.01	(-)	0.905	tccctgtaAACAagaaa
V$P53F/TP63.02	(-)	0.903	gcctgcaagccacggCCTGccccag
V$FKHD/FOXK1.01	(-)	0.902	tccctgtaAACAagaaa
V$BRAC/TBX21.01	(-)	0.9	ggaaggggttGGTGggacatcttaaccac
V$OAZF/ROAZ.01	(-)	0.899	ggGCAGccatgggtgag
V$RBPF/RBPJK.01	(+)	0.899	agccTGGGaacta
V$SP1F/GC.01	(+)	0.899	aacgggGGAGggggccc
V$ESRR/ERR.01	(-)	0.897	ctttcctctgccAAGGcagcca
V$FKHD/XFD2.01	(-)	0.895	tccctgTAAAcaagaaa
V$LEFF/LEF1.01	(-)	0.893	ctgggagCAAAgggga
V$NDPK/NM23.01	(+)	0.893	cgGGGGagggggccctc
V$FKHD/FOXL1.01	(-)	0.891	tccctgTAAAcaagaaa

V$ETSF/SPIB.01	(-)	0.89	gaaaaactGGAAtccacactg
V$NBRE/NBRE.01	(+)	0.882	aggaAAGGccaaagc
V$NR2F/HPF1.01	(+)	0.882	gaggaaaggccAAAGcccagtgagg
V$RXRF/THR.01	(-)	0.881	caggtggttaggaAAGGagagcagg
V$AP1R/VMAF.01	(+)	0.879	tggcttgccctGTCAgcctgggaac
V$AP2F/TCFAP2B.01	(-)	0.877	ccaGCCTctaggcag
V$EREF/ER.03	(-)	0.877	gtgggacatcttAACCaca
V$CAAT/ACAAT.01	(+)	0.875	cccaCCAAccccttc
V$EVI1/EVI1.06	(-)	0.875	taaacaAGAAaaactgg
V$ETSF/GABP.01	(-)	0.874	gcagtgggGGAAggggttggt
V$GREF/PRE.01	(-)	0.873	agcctaagcccTGTTctgg
V$FKHD/FREAC3.01	(-)	0.872	tccctGTAAacaagaaa
V$TALE/TGIF1.01	(+)	0.872	gcttgccctGTCAgcct
V$NR2F/ARP1.01	(-)	0.87	agggggaattcaagGATCattccct
V$SORY/HBP1.01	(+)	0.87	ttacagggAATGatccttgaatt
V$IKRS/IK3.01	(+)	0.866	tacagGGAAtgat
V$XBBF/RFX2.02	(+)	0.855	ctgtcagcctggGAACtag
V$ETSF/SPIC.01	(-)	0.853	gcaaagggGGAAttcaaggat
V$SRFF/SRF.02	(+)	0.852	gctccCAGAacagggctta
V$HNF1/HNF1.04	(+)	0.851	gctttgtgGTTAagatg
V$PAX9/PAX9.01	(+)	0.838	cctcctcacccatgGCTGccc
V$PAX6/PAX6.04	(+)	0.831	tccTCACccatggctgccc
V$RREB/RREB1.01	(+)	0.82	cCCCAgcaaaccctc
V$CREB/CREB3L2.01	(-)	0.819	gggtgagGAGGtggcaaggca
V$NR2F/PNR.01	(-)	0.817	agggggaatTCAAggatcattccct
V$HNF1/TCF2.01	(+)	0.814	gctttgtgGTTAagatg
V$CTCF/CTCF.01	(+)	0.802	ggcttaggctgaacGGGGgaggggcc
V$PAX5/PAX5.01	(+)	0.802	ccaaagCCCAgtgaggaggcggtggcttg
V$FKHD/FREAC4.01	(-)	0.8	tccctgtaAACAagaaa
V$RXRF/THRB.03	(+)	0.8	tggctgcccttggcagAGGAaaggc
V$PARF/TEF_HLF.01	(+)	0.798	tctccTTTCctaaccac
V$HOXH/MEIS1B_HOXA9.01	(-)	0.79	GGACatcttaaccac
V$RXRF/RXR_RXR.01	(-)	0.789	ggggaattcaagGATCattccctgt
V$NR2F/TR4.02	(+)	0.787	gaggaaAGGCcaaagcccagtgagg
V$NR2F/TR2.01	(-)	0.781	caagggcagccatgGGTGaggaggt
V$PAX3/PAX3.01	(+)	0.778	cTCCTcacccatggctgcc
V$AP1R/MARE_ARE.01	(+)	0.774	tggcttgccctGTCAgcctgggaac
V$AHRR/AHRARNT.02	(+)	0.771	gcaggccgtgGCTTgcaggccccag

V$RORA/RORA.02	(-)	0.77	ctctgccaaGGGCagccatgggtga
V$ZF12/ZNF652.02	(-)	0.76	gggaagGGGTtggtg
V$PRDM/PRDM5.01	(-)	0.757	agggcagccATGGgtgaggaggtggcaag
V$SNAP/PSE.02	(-)	0.757	gacatCTTAaccacaaagc

Table A6. Complete table of TFBSs of *Mus musculus* Cebpa Downstream CNS2 region

Matrix	strand	Matrix sim.	Sequence
V$SMAD/SMAD.01	(-)	1	tttGTCTggct
V$PLAG/PLAG1.02	(-)	1	gaGGGGgcgggaagcctgggagg
V$SORY/SOX12.01	(+)	1	ccagaACAAttaaacagccagat
V$TAIP/CSRNP1.01	(-)	1	AGAGtgg
V$TAIP/CSRNP3.01	(-)	1	aGAGTgg
V$ZTRE/ZTRE.04	(-)	1	ctgGGAGggtggacac
V$KLFS/KLF2.01	(+)	1	agcagGGGTggggtttggg
V$E2FF/E2F1.01	(-)	1	agggGGCGggaagcctg
V$E2FF/E2F4.01	(-)	1	aggggGCGGgaagcctg
V$FKHD/FOXP1_ES.01	(-)	1	aaataaaAACAaaagtg
V$ZF02/ZKSCAN3.01	(-)	1	ccaaaCCCCacccctgctccagg
V$GATA/GATA2.03	(-)	0.998	gagaGATAagggt
V$GATA/GATA4.01	(-)	0.998	gagaGATAagggt
V$CEBP/CEBPB.02	(-)	0.998	cctctTGTGaaatca
V$MOKF/MOK2.02	(-)	0.997	ggcgtacagtgggCCTTgcaa
V$BRAC/TBX5.01	(-)	0.996	gctcatggcaGGTGtcagggccaaacccc
V$SMAD/SMAD3.01	(-)	0.995	tttGTCTggct
V$GATA/GATA1.07	(-)	0.995	gagaGATAagggt
V$HOMF/MSX.01	(-)	0.995	ggctgttTAATtgttctgg
V$CART/S8.01	(-)	0.995	ctgttTAATtgttctgggatg
V$MZF1/MZF1.03	(+)	0.995	gaGGGGagggg
V$FKHD/HNF3.01	(+)	0.995	tggctgcAAACagagcc
V$ZF11/ZBTB3.01	(+)	0.994	cacaGCCAgcc
V$SP1F/SP4.01	(-)	0.994	ggagggGGCGggaagcc
V$EVI1/MEL1.02	(-)	0.994	ccttcaaGATGaggggt
V$E2FF/E2F6.01	(-)	0.991	aggggGCGGgaagcctg
V$HOMF/HHEX.01	(-)	0.99	atctggctgttTAATtgtt
V$SAL2/SALL2.01	(+)	0.989	agggGTGGggt
V$CEBP/CEBPD.01	(-)	0.989	cctctTGTGaaatca
V$CEBP/CEBPA.01	(-)	0.988	cctcttgtGAAAtca
V$KLFS/BKLF.02	(+)	0.988	aagaggGGAGggacccct
V$AP1R/NRL.02	(-)	0.988	ttgtgaaatcAGCAgcttgtggagg
V$SORY/SOX5.02	(+)	0.987	ccagaACAAttaaacagccagat
V$GATA/GATA6.02	(-)	0.986	gagaGATAagggt
V$SORY/SOX15.01	(+)	0.986	ccagaACAAttaaacagccagat
V$SATB/SATB1.01	(-)	0.984	aatAATAaataaaaa

V$IKRS/IK2.01	(+)	0.983	ctctGGGAaaacg
V$LTSM/LTSM.03	(+)	0.983	ATCTcctgggaatct
V$HOMF/NOBOX.01	(-)	0.983	ggctgttTAATtgttctgg
V$PURA/PURALPHA.01	(-)	0.982	ggAGGGggcggga
V$NEUR/NGN_NEUROD.01	(-)	0.982	attcCATCtggctgt
V$ZFHX/SIP1.01	(+)	0.982	ctgacACCTgcca
V$KLFS/KLF6.01	(-)	0.982	aaggagGGGGcgggaagcc
V$E2FF/E2F7.02	(-)	0.982	agggggcGGGAagcctg
V$SAL1/SALL1.01	(-)	0.981	taATAAataaaaa
V$MYOD/TCF12.01	(-)	0.98	tcatggcaGGTGtcagg
V$SORY/SOX5.01	(+)	0.98	ccagaaCAATtaaacagccagat
V$ETSF/SPI1.02	(+)	0.979	ttgggagtGGAAgtgggcaga
V$KLFS/IKLF.01	(-)	0.979	ttaatGGGTggggccagaa
V$KLFS/KLF7.02	(+)	0.979	ccccggGGCGtatggaagc
V$CSEN/DREAM.01	(-)	0.979	gtGTCAgggcc
V$BRAC/TBX6.01	(-)	0.977	gctcatggcAGGTgtcagggccaaacccc
V$NEUR/NEUROD1.02	(-)	0.977	attcCATCtggctgt
V$SORY/SOX30.01	(+)	0.977	ccagaACAAttaaacagccagat
V$ZF02/ZBTB7.03	(-)	0.977	ggggtCCCCtcccctcttgtgaa
V$NEUR/ATOH1.01	(-)	0.976	attcCATCtggctgt
V$SORY/SOX13.01	(+)	0.974	ccagaACAAttaaacagccagat
V$EVI1/EVI1.07	(-)	0.974	ccttcAAGAtgaggggt
V$SORY/SOX7.01	(+)	0.973	ccagaACAAttaaacagccagat
V$SP1F/SP1.01	(-)	0.973	ggaggGGGCgggaagcc
V$HOMF/BARX2.01	(-)	0.972	ctgttctTAATgggtgggg
V$HOXF/NANOG.01	(-)	0.972	tgttcttAATGggtggggc
V$TALE/MEIS1.01	(+)	0.972	cttaacgcTGTCacgga
V$ZF08/ZNF354C.01	(-)	0.971	aggttCCACct
V$LHXF/LMX1B.01	(-)	0.971	aaagctagaaaaaTAATaaataa
V$SORY/SOX3.01	(-)	0.97	taaaaaCAAAagtgaagagaaga
V$YY1F/YY2.01	(-)	0.969	tgggttCCATatttgtctggctc
V$LHXF/LMX1A.02	(-)	0.967	ctggctgtTTAAttgttctggga
V$SF1F/FTF.01	(+)	0.965	cttgCAAGgcccact
V$ETSF/ETV1.02	(+)	0.964	tgggcagaGGAAggctgggga
V$KLFS/EKLF.02	(-)	0.962	ttaatgGGTGgggccagaa
V$BEDF/ZBED4.02	(-)	0.962	ggaaggaGGGGgcgg
V$NOLF/EBF1.01	(+)	0.961	caggccTCCCcggggcgtatgga
V$KLFS/GKLF.02	(-)	0.96	ctttcgttAAAGgctctgt

V$NKX6/NKX61.02	(-)	0.958	ctgtTTAAttgttct
V$GATA/GATA3.01	(-)	0.958	gagaGATAagggt
V$MYOD/E47.01	(-)	0.958	tcatgGCAGgtgtcagg
V$NEUR/NEUROG.01	(-)	0.957	attCCATctggctgt
V$MYOD/SGN1.01	(-)	0.957	tcatggCAGGtgtcagg
V$LEFF/LEF1.02	(+)	0.957	tcactatCAAAgcatcc
V$ZF02/ZBP89.01	(+)	0.957	gcttcccgCCCCctccttcccag
V$ZF07/ZNF263.01	(-)	0.956	ggtcccCTCCcctct
V$SIXF/SIX4.01	(-)	0.956	tggcagGTGTcaggg
V$MYOD/TCFE2A.02	(+)	0.955	aacagccaGATGgaatc
V$NFAT/NFAT.01	(+)	0.955	aagctaGGAAagaaaccaa
V$HEAT/HSF2.02	(+)	0.954	gcattaataacgCGAAcaatctgtg
V$SMAD/GC_SBE.01	(-)	0.953	ctgggCTCCtg
V$RORA/RORA1.01	(+)	0.953	catcttgaaGGTCaaggggagcaga
V$NBRE/NBRE.01	(+)	0.953	cttgAAGGtcaaggg
V$HBOX/GSH2.01	(-)	0.953	tgttctTAATgggtggggc
V$AP1F/AP1.02	(-)	0.952	ggctGAGTaatgg
V$AP1F/AP1.03	(-)	0.952	ggcTGAGtaatgg
V$ETSF/PEA3.01	(+)	0.951	tgggcagAGGAaggctgggga
V$HICF/HIC1.02	(+)	0.951	ctgTGCCcacatc
V$EGRF/EGR1.03	(-)	0.95	gaaggagGGGGcgggaagc
V$DLXF/DLX3.01	(-)	0.95	ctggctgttTAATtgttct
V$MAZF/MAZ.01	(+)	0.95	agggGAGGggacc
V$ESRR/ESRRB.01	(+)	0.95	cccctcatcttgaAGGTcaaggg
V$MYT1/MYT1L.01	(+)	0.95	cgaaAGCTgtaag
V$ARID/BRIGHT.01	(+)	0.949	gaacaATTAaacagccagatg
V$MYRF/MYRF.01	(-)	0.949	gggtgCCAGaatg
V$FKHD/FOXP1.01	(+)	0.948	tggctgcAAACagagcc
V$RBPF/RBPJK.02	(+)	0.947	tctcTGGGaaaac
V$HOMF/HMX3.03	(-)	0.946	ggctgttTAATtgttctgg
V$ESRR/ESRRA.03	(+)	0.944	cccctcatcttgAAGGtcaaggg
V$FKHD/FOXA1.01	(+)	0.944	tggctgcaAACAgagcc
V$GREF/GRE.03	(-)	0.942	cctgcacagatTGTTcgcg
V$SORY/SRY.01	(-)	0.942	taaaaACAAaagtgaagagaaga
V$KLFS/KKLF.01	(+)	0.942	cctggagcaGGGGtgggt
V$HOMF/HMX2.03	(-)	0.94	ggctgtTTAAttgttctgg
V$GATA/GATA.01	(-)	0.939	gagaGATAagggt
V$AP2F/TCFAP2A.01	(+)	0.939	cctccccgGGGCgta

V$DLXF/DLX2.01	(-)	0.938	ctggctgtttAATTgttct
V$ETSF/ERG.02	(+)	0.937	gtaagctaGGAAagaaaccaa
V$MAZF/MAZR.01	(+)	0.936	aggggaGGGGacc
V$HAND/TH1E47.01	(+)	0.935	caaacagagCCAGacaaatat
V$LEFF/TCF7L2.02	(+)	0.935	tcactatcAAAGcatcc
V$DLXF/DLX1.01	(-)	0.934	ctggctgtttAATTgttct
V$NR2F/COUP.02	(+)	0.934	atcttgaaggTCAAggggagcagaa
V$NACA/NACA1.01	(-)	0.934	gggcCAGAgggga
V$GREF/ARE.04	(-)	0.933	cctgcacagattGTTCgcg
V$IKRS/IK1.01	(+)	0.932	ctctGGGAaaacg
V$AP2F/TCFAP2C.01	(+)	0.932	cctCCCCggggcgta
V$BRNF/BRN3.03	(+)	0.931	aacgcatTAATaacgcgaa
V$HOMF/HMX1.01	(-)	0.93	ggctgttTAATtgttctgg
V$SORY/SOX2.01	(-)	0.93	taaaaACAAaagtgaagagaaga
V$NDPK/NM23.01	(+)	0.928	caGGGGtggggtttggg
V$CP2F/CP2.01	(-)	0.928	ggCTGGctgtggctgtggc
V$STAT/STAT5B.01	(+)	0.927	ccccatcccAGAAcaatta
V$HEAT/HSF1.05	(+)	0.927	caagtggtaacatcttCCAGagcac
V$SORY/SOX9.01	(+)	0.927	ccagaaCAATtaaacagccagat
V$ZF05/ZFP410.01	(-)	0.926	tgttctgGGATgggg
V$SP1F/GC.01	(+)	0.925	agcaggGGTGggtttg
V$AP1R/MAFA.01	(-)	0.924	cctcttgtgaaatcAGCAgcttgtg
V$AP2F/AP2.02	(-)	0.924	aggGCCAgaggggac
V$BCDF/PCE1.01	(-)	0.923	ctgttTAATtgttctgg
V$ABDB/HOXD13.01	(-)	0.921	ataataaaTAAAaacaa
V$TALE/TGIF2LX.01	(+)	0.92	ctgacacctGCCAtgag
V$INSM/INSM1.01	(-)	0.919	tgggtGGGGccag
V$ZF02/ZNF219.01	(+)	0.919	tcccgCCCCctccttcccagccc
V$DMTF/DMP1.02	(+)	0.918	ggggcGTATggaagc
V$GLIF/GLIS2.01	(+)	0.918	gggaCCCCtcatcttga
V$EBOX/MYCMAX.03	(+)	0.917	cctgccaTGAGccactg
V$EREF/ESR2.01	(+)	0.915	gaagGTCAaggggagcaga
V$CDEF/CDE.01	(-)	0.915	tgttCGCGttatt
V$SORY/SOX4.01	(-)	0.915	taaaaACAAaagtgaagagaaga
V$LEFF/TCF7L1.01	(+)	0.915	tcactatCAAAgcatcc
V$HOXC/HOXC9.02	(+)	0.914	gtttttaTTTAttattt
V$HOMF/BARHL2.01	(-)	0.913	ggctgttTAATtgttctgg
V$BARB/BARBIE.01	(+)	0.913	aacgAAAGctgtaag

V$LHXF/ISL2.01	(+)	0.912	tggccccacccATTAagaacagg
V$CDXF/CDX2.03	(+)	0.911	gtttttatTTATtattttt
V$OAZF/ROAZ.01	(+)	0.911	gaGCACccatgggtctt
V$ETSF/GABPA.02	(+)	0.91	ttgggagtGGAAgtgggcaga
V$CP2F/TCFCP2L1.01	(-)	0.908	ggCAGGttccacctgggct
V$SORY/HMGA.01	(-)	0.908	tcgcgttattAATGcgttttccc
V$BCL6/BCL6.04	(-)	0.907	gttTTCCcagagattcc
V$NKXH/NKX25.02	(-)	0.903	ctgttTAATtgttctggga
V$BRNF/TST1.01	(+)	0.903	gaacaATTAaacagccaga
V$MEF3/SIX.01	(-)	0.903	agcTCGGtgcca
V$MYBL/VMYB.05	(+)	0.902	cagagcctttAACGaaagctg
V$EBOX/MAX.03	(-)	0.902	caaacCCCGtgttcaca
V$EGRF/CKROX.01	(+)	0.901	caagagGGGAggggacccc
V$SORY/SOX11.01	(-)	0.9	taaaaACAAagtgaagagaaga
V$PERO/PPARG.02	(-)	0.899	cttgcaagggccAGAGgggacag
V$GATA/GATA5.01	(-)	0.899	gagaGATAagggt
V$HOMF/MSX1.01	(+)	0.899	cccagaacAATTaaacagc
V$NOLF/OLF1.02	(+)	0.899	caggccTCCCcggggcgtatgga
V$ETSF/GABP.01	(+)	0.896	ttgggagtGGAAgtgggcaga
V$FKHD/FOXO1.01	(+)	0.896	aacaattaAACAgccag
V$FKHD/FOXP2.01	(+)	0.896	aacaattaAACAgccag
V$FKHD/HFH1.01	(+)	0.894	aacaatTAAAcagccag
V$ABDB/HOXB9.02	(+)	0.893	cagaacaaTTAAacagc
V$LEFF/TCF7.01	(+)	0.893	tcactatCAAAgcatcc
V$ESRR/ERR.01	(+)	0.892	cccctcatcttgAAGGtcaaggg
V$FKHD/FREAC2.01	(+)	0.892	aacaatTAAAcagccag
V$FKHD/XFD2.01	(-)	0.892	aaataaTAAAtaaaaac
V$SRFF/SRF.04	(-)	0.891	ggttccataTTTGtctggc
V$TALE/TGIF1.01	(+)	0.89	cttaacgctGTCAcgga
V$RORA/RORA.01	(+)	0.887	catcttgaagGTCAaggggagcaga
V$GREF/MRE.01	(-)	0.886	cctgcacagattGTTCgcg
V$HOMF/MSX3.01	(-)	0.886	ggctgttTAATtgttctgg
V$NKX6/NKX63.01	(+)	0.884	cgcatTAATaacgcg
V$RREB/RREB1.01	(-)	0.884	tCCCAaaccccaccc
V$BRAC/EOMES.02	(-)	0.882	gctcatggcaGGTGtcagggccaaacccc
V$EREF/ER.04	(+)	0.882	gaagGTCAaggggagcaga
V$AIRE/AIRE.01	(-)	0.882	cgTTATtaatgcgtt
V$GLIF/GLI2.01	(+)	0.88	cagaGCACccatgggtc

V$HOMF/MSX2.02	(-)	0.879	ggctgtttAATTgttctgg
V$AP1R/MAFK.01	(-)	0.878	cctcttgtgaaatcAGCAgcttgtg
V$ETSF/NRF2.01	(-)	0.877	acagcaaaGGAAgagaaaagg
V$YY1F/REX1.01	(+)	0.877	ccccacCCATtaagaacaggcct
V$ARID/ARID5A.01	(+)	0.876	gacaaATATggaacccattac
V$STAT/STAT5A.01	(-)	0.876	cgttTTCCcagagattcca
V$GCMF/GCM1.01	(+)	0.876	ctgtgCCCAcatcca
V$CHRE/CHREBP_MLX.01	(-)	0.876	CACGgggagtggatgtg
V$ABDB/HOXC9.01	(-)	0.875	tcgcgttaTTAAtgcgt
V$IRFF/IRF4.03	(+)	0.875	tgtaagctaggaaaGAAAccaattc
V$BRNF/BRN2.01	(+)	0.874	ctCATCttgaaggtcaagg
V$HNF1/HNF1.04	(-)	0.874	accacttgGTTAaggaa
V$HOXC/HOXA9.01	(+)	0.873	gttttTATTtattattt
V$HAND/TAL1BETAE47.01	(+)	0.873	aaacagcCAGAtggaatctct
V$ABDB/HOXB13.01	(-)	0.872	ataataaaTAAAaacaa
V$HAND/TAL1ALPHAE47.01	(+)	0.872	aaacagcCAGAtggaatctct
V$PLAG/PLAGL1.01	(-)	0.871	tgaggGGTCccctcccctcttgt
V$FKHD/FOXL1.01	(-)	0.871	aaataaTAAAtaaaaac
V$GLIF/GLI1.02	(+)	0.871	cagaGCACccatgggtc
V$NGRE/IR2_NGRE.01	(-)	0.87	gaTTCCcaggagatc
V$OCT1/OCT1.06	(+)	0.87	ggtatgctAAGTctg
V$GREF/PRE.01	(-)	0.869	ccggggaggccTGTTctta
V$HAML/AML3.01	(+)	0.869	gggtGGGGtttggga
V$FAST/FAST1.03	(-)	0.868	ttgatagtgAATTggtt
V$GLIF/GLI3.01	(+)	0.867	cagaGCACccatgggtc
V$ETSF/SPIC.01	(+)	0.866	ttgggagtGGAAgtgggcaga
V$IKRS/IK3.01	(+)	0.865	tcctgGGAAtctg
V$TALE/MRG1.02	(+)	0.864	ctgaCACCtgccatgag
V$CTCF/CTCF.05	(+)	0.864	tgctgatttcacaAGAGggaggggac
V$HOXF/HOX1-3.01	(+)	0.86	cgcatTAATaacgcgaaca
V$AP1R/NFE2L2.01	(-)	0.86	tggctggctGAGTaatgggttccat
V$MTF1/MTF-1.02	(-)	0.859	ttatccaCACGggga
V$SP1F/TIEG.01	(-)	0.858	acacGGGGagtggatgt
V$AP1R/MAFF.01	(+)	0.858	caagctGCTGatttcacaagagggg
V$ATBF/ATBF1.01	(+)	0.857	tttgtttttATTTatt
V$HOXF/HOXC8.01	(+)	0.857	cgcattaATAAcgcgaaca
V$PAX3/PAX3.02	(+)	0.857	gatgTCACccttatctctc
V$PEG3/PEG3.01	(-)	0.857	cagtggctcaTGGCa

V$OVOL/OVOL1.01	(+)	0.855	gtcaccCTTAtctct
V$AP1R/NFE2.01	(-)	0.855	tggctggCTGAgtaatgggttccat
V$HAND/TAL1BETAITF2.01	(+)	0.855	aaacagcCAGAtggaatctct
O$XCPE/XCPE1.01	(-)	0.853	ggGCGGgaagc
V$PRDF/BLIMP1.01	(-)	0.853	aggaagaGAAAagggcttc
V$CART/RHOX6.01	(-)	0.852	tcgcgTTATtaatgcgttttc
V$AP1R/BACH1.02	(-)	0.851	tggctggCTGAgtaatgggttccat
V$E2FF/E2F.02	(-)	0.849	tgtcagggcCAAAcccc
V$HOXC/HOXB4.02	(-)	0.848	tggcTGAGtaatgggtt
V$AP1R/MAFB.01	(-)	0.847	cctcttgtgaaaTCAGcagcttgtg
V$HOMF/TLX2.01	(-)	0.846	gcgttaTTAAtgcgttttc
V$HOXF/HOXB8.01	(+)	0.846	cagaacaATTAaacagcca
V$FKHD/FHXB.01	(-)	0.845	agaaaaATAAtaaataa
V$STAT/STAT6.01	(+)	0.843	ggatCTCCtgggaatctgt
V$AP1R/MARE.03	(-)	0.842	tggctgGCTGagtaatgggttccat
V$LHXF/LHX9.01	(+)	0.84	atcccagaacAATTaaacagcca
V$NFKB/HIVEP1.01	(-)	0.839	ttGGGAtggtcctgt
V$HNF1/HMBOX.01	(-)	0.839	accacttgGTTAaggaa
V$CREB/ATF.02	(+)	0.836	gggaggTGACttctttgccac
V$E2FF/E2F1_DP1.01	(-)	0.832	agggGGCGggaagcctg
V$PRDM/PRDM4.01	(+)	0.832	ggcccTTGCaaggcccactgtacgccctg
V$BRN5/BRN5.04	(+)	0.831	tgggaaaacgcATTAataacgcg
V$IRFF/ISGF3G.01	(+)	0.829	tgtaagctaggaaaGAAAccaattc
V$HOXF/HOXD8.01	(-)	0.829	tcgcgttATTAatgcgttt
V$HOXF/HOXC5.01	(+)	0.828	cagaacAATTaaacagcca
V$HNF1/TCF2.01	(-)	0.828	accacttgGTTAaggaa
V$RXRF/RAR_RXR.01	(-)	0.827	gcaagggccagAGGGgacagcacag
V$OCT1/OCT2.01	(+)	0.827	ggtATGCtaagtctg
V$LHXF/ISL1.01	(-)	0.826	tcgcgttatTAATgcgttttccc
V$OCT1/POU2F3.01	(+)	0.826	ggtATGCtaagtctg
V$HOXF/HOXA4.01	(-)	0.825	aaaaatAATAaataaaaac
V$E4FF/E4F.01	(+)	0.825	cggATGTcaccct
V$PIT1/PIT1.02	(+)	0.822	acgcaTTAAtaacgc
V$HNF6/OC2.01	(-)	0.821	aaaaatAATAaataaaa
V$HOXH/MEIS1B_HOXA9.01	(-)	0.821	TGACagcgttaagct
V$RXRF/CAR_RXR.01	(-)	0.819	cacttGGTTaaggaagatcagaatt
V$RXRF/VDR_RXR.05	(-)	0.819	cggGGAGtggatgtgggcacagcaa
V$EGRF/NGFIC.01	(-)	0.818	ggaaGCCTgggaggggtgg

V$ABDB/HOXD11.01	(-)	0.818	agctttcgTTAAaggct
V$HAND/LYL1_E12.01	(-)	0.818	gcctgcacaGATTgttcgcgt
V$BRNF/BRN4.02	(+)	0.816	aacgcaTTAAtaacgcgaa
V$HNF6/HNF6.01	(-)	0.811	aaaaataaTAAAtaaaa
V$PRDF/PRDM1.01	(+)	0.81	tgggagtGGAAgtgggcag
V$RXRF/PXR_RXR.01	(+)	0.809	ggaagGGTGaggagtgtccacccct
V$HESF/BHLHB2.02	(-)	0.808	tccacaCGGGgagtg
V$SNAP/PSE.01	(+)	0.804	cTCCCcgtgtggataatag
V$OCT1/OCT.01	(+)	0.804	ggtATGCtaagtctg
V$HBOX/EN1.01	(-)	0.804	tctggctgTTTAattgttc
V$CEBP/CEBPE.01	(-)	0.803	cctcttgtGAAAtca
V$ETSF/GABPB1.01	(+)	0.802	gctgtcaCGGAtgtcaccctt
V$ZF02/ZNF202.01	(-)	0.801	ccactcCCAAaccccaccccctgc
V$NR2F/TR2_TR4.01	(-)	0.8	tggctcatggcaggTGTCagggcca
V$ZF01/ZBRK1.01	(+)	0.8	caaggggaGCAGaaattctgatctt
V$E2FF/E2F4_DP2.01	(-)	0.799	aggggGCGGgaagcctg
V$PARF/TEF_HLF.01	(+)	0.798	tgatcTTCCttaaccaa
V$HOXH/MEIS1A_HOXA9.01	(-)	0.795	TGACagcgttaagct
V$DMRT/DMRT4.01	(+)	0.793	ccaagtggtAACAtcttccag
V$HOXC/PBX_HOXA9.01	(+)	0.792	gtttTTATttattattt
V$PAX5/PAX5.01	(+)	0.791	gcctggAGCAggggtggggtttgggagtg
V$NR2F/HPF1.01	(-)	0.79	aaaggaagagaAAAGggcttccata
V$ZF12/ZFP652.01	(-)	0.789	gtgacagcgTTAAgc
V$PAX7/PAX7.01	(+)	0.783	ttttattTATTattt
O$VTBP/ATATA.01	(+)	0.781	gcgtatgGAAGccttt
V$EGRF/EGR3.01	(-)	0.773	ggaaGCCTgggaggggtgg
V$PDX1/PDX1.01	(+)	0.773	aaaacgcatTAATaacgcg
V$NR2F/TR4.02	(-)	0.768	ggccagAGGGgacagcacagactta
V$PTF1/PTF1.01	(+)	0.763	acaaGCTGctgatttcacaag
V$ZF13/ZNF435.01	(-)	0.758	cgTGTTcacagattcccag
V$RXRF/THRB.02	(+)	0.755	cctctggcccttgcaaGGCCcactg
V$STAT/STAT3.01	(+)	0.755	ggatCTCCtgggaatctgt
V$HAND/HAND2_E12.01	(+)	0.752	aaacagccagaTGGAatctct

Table A7. Complete table of TFBSs of *Mus musculus* Cebpa Downstream CNS3 region

Matrix	strand	Matrix sim.	Sequence
V$PLAG/PLAG1.02	(+)	1	aaGGGGgtgtcagggactgggta
V$CART/S8.01	(+)	1	cactgTAATtggctttggccc
V$TAIP/CSRNP1.01	(-)	1	AGAGtgc
V$TAIP/CSRNP3.01	(-)	1	aGAGTga
V$MZF1/MZF1.01	(+)	1	gtGGGGatgag
V$FKHD/FOXP1_ES.01	(-)	1	caaataaAACAtgtttg
V$OCT1/OCT1.02	(+)	1	tttATGCaaatcgta
V$FKHD/HNF3.01	(-)	0.999	gtgaggcAAACacattt
V$ZF11/ZBTB3.01	(-)	0.997	gccaGCCAgag
V$EBOX/MNT.01	(-)	0.997	cccaaCACGtgcacagg
V$EVI1/MEL1.02	(-)	0.996	ccatgatGATGaggtac
V$FKHD/FOXP2.01	(+)	0.996	gaaaggtaAACAggcca
V$RU49/RU49.01	(+)	0.994	cAGTAcc
V$CEBP/CEBPE.02	(+)	0.993	cgtgttggGCAAcca
V$NKXH/NKX25.05	(+)	0.992	ctcctTGAGtgatttacaa
V$GATA/GATA1.06	(+)	0.991	tgccGATAaggaa
V$HAML/AML1.01	(+)	0.991	aactGTGGtcaactt
V$CEBP/CEBPB.02	(+)	0.991	ggagtTGTGaaaggt
V$ZF57/ZFP57.01	(-)	0.989	tgtTGCCgcacca
V$FKHD/FOXO1.01	(+)	0.989	gaaaggtaAACAggcca
V$RUSH/SMARCA3.02	(-)	0.986	cagaACTTgtg
V$HOXC/HOXC9.02	(+)	0.986	ccactgaTTTAtaggga
V$FKHD/FOXP1.01	(+)	0.986	gaaaggtAAACaggcca
V$HBOX/GSH2.01	(+)	0.986	tcccacTAATttgatttta
V$BZIP/CEBPE_ATF4.01	(-)	0.985	agtgagGCAAaca
V$MYOD/MYF6.01	(+)	0.984	agacaACAGgtggcaag
V$RBPF/RBPJK.02	(-)	0.984	ttagTGGGaaaat
V$HOXC/MEIS1.03	(+)	0.982	ccactGATTtataggga
V$IKRS/IK2.01	(-)	0.981	cagtGGGAatgtg
V$MEF2/MEF2C.01	(+)	0.981	caacattttaaAAATagactcgc
V$FKHD/ILF1.01	(+)	0.981	gaaaggtaAACAggcca
V$EBOX/MYCMAX.02	(+)	0.981	ctgtgcaCGTGttgggc
V$EBOX/NMYC.01	(-)	0.981	cccaacaCGTGcacagg
V$EBOX/USF.01	(+)	0.981	ctgtgCACGtgttgggc
V$OCT1/POU2F3.01	(+)	0.98	tttATGCaaatcgta
V$SORY/SOX5.01	(+)	0.979	ccagaaCAATttgaaggcactca

V$EVI1/MEL1.03	(-)	0.979	ccatgatGATGaggtac
V$CSEN/DREAM.01	(+)	0.979	gtGTCAgggac
V$AP2F/TCFAP2A.02	(+)	0.979	attcCCTGggggcag
V$SORY/SOX12.01	(+)	0.978	ccagaACAAtttgaaggcactca
V$HESF/DEC1.02	(-)	0.978	ccaacaCGTGcacag
V$HOMF/MSX.01	(+)	0.977	ttcccacTAATttgatttt
V$HOMF/MSX2.01	(+)	0.977	tatagcCTAAtttagcctc
V$ETSF/SPI1.02	(+)	0.977	tgataagtGGAAgtttcctgt
V$BCDF/CRX.01	(+)	0.977	gatttTAATctgatatt
V$CEBP/CEBPD.01	(+)	0.977	ggagtTGTGaaaggt
V$OCT1/OCT2.02	(+)	0.977	tttaTGCAaatcgta
V$HOMF/NOBOX.01	(+)	0.976	ggcactgTAATtggctttg
V$CEBP/CEBPA.01	(+)	0.976	ggagttgtGAAAggt
V$FKHD/FOXA1.01	(+)	0.976	gaaaggtaAACAggcca
V$DLXF/DLX3.01	(-)	0.975	tagtcactgTAATtgcccc
V$FKHD/HNF3B.04	(+)	0.975	gaaaggtaAACAggcca
V$HIFF/HRE.02	(+)	0.973	cctgtgcaCGTGttggg
V$SORY/SOX15.01	(+)	0.97	ccagaACAAtttgaaggcactca
V$FKHD/FREAC2.01	(+)	0.97	gaaaggTAAAcaggcca
V$GATA/GATA.01	(+)	0.969	tgccGATAaggaa
V$HOMF/BSX.01	(+)	0.969	gccggggcAATTacagtga
V$PAX3/PAX3.03	(-)	0.968	aaggtCATGgcagctcccc
V$HOXC/HOXA9.01	(+)	0.967	ccactGATTtataggga
V$MEF2/MEF2A.01	(+)	0.965	caacattttaaAAATagactcgc
V$SORY/SOX13.01	(+)	0.965	ccagaACAAtttgaaggcactca
V$EBOX/CMYC.01	(+)	0.965	ctgtgcaCGTGttgggc
V$HESF/DEC2.01	(-)	0.964	ccaacaCGTGcacag
V$MYOD/MYOGENIN.01	(+)	0.963	agacaACAGgtggcaag
V$PAX5/PAX2.02	(+)	0.963	agtgactaagcaTGACaaatcactctcag
V$KLFS/GKLF.02	(-)	0.961	tttgcataAAAGgaaactg
V$LTSM/LTSM.03	(-)	0.959	ATCAgattaaaatca
V$MEF2/SL1.01	(-)	0.959	agcgagtCTATttttaaaatgtt
V$SORY/SOX30.01	(+)	0.959	ccagaACAAtttgaaggcactca
V$MEF3/SIX2.02	(+)	0.959	cgtTCAGtttcct
V$AP2F/TCFAP2C.01	(+)	0.958	attCCCTggggggcag
V$SIXF/SIX4.01	(+)	0.958	aaggggGTGTcaggg
V$HIFF/ARNT.01	(+)	0.957	cctgtgcaCGTGttggg
V$ESRR/ESRRB.01	(-)	0.956	aacctgatggtgaAGGTcatggc

V$HICF/HIC1.02	(-)	0.955	ggtTGCCcaacac
V$EBOX/MAX.02	(+)	0.955	ctgtgCACGtgttgggc
V$NEUR/ATOH1.01	(-)	0.954	ttgcCACCtgttgtc
V$NF1F/NF1.04	(-)	0.953	ctcctgggccaaagCCAAtta
V$HOMF/HHEX.01	(+)	0.952	taatttgatttTAATctga
V$SORY/SOX7.01	(+)	0.952	ccagaACAAtttgaaggcactca
V$ETSF/ELF5.01	(+)	0.952	tgataagtGGAAgtttcctgt
V$CAAT/NFY.04	(-)	0.952	aaagCCAAttacagt
O$VTBP/MTATA.01	(-)	0.952	ccctaTAAAtcagtggg
V$ESRR/ESRRA.03	(-)	0.951	aacctgatggtgAAGGtcatggc
V$NBRE/NBRE.01	(-)	0.95	ggtgAAGGtcatggc
V$SORY/SOX9.02	(-)	0.949	tccccACAAtggtatgagaggaa
V$AP2F/TCFAP2E.01	(-)	0.949	ccagccagAGGCaaa
V$CAAT/CAAT.01	(-)	0.947	aaagCCAAttacagt
V$IRXF/IRX5.01	(-)	0.946	aaaaCATGtttgc
V$HAML/AML2.01	(+)	0.946	aacTGTGgtcaactt
V$DLXF/DLX5.01	(-)	0.945	tagtcactgtAATTgcccc
O$VTBP/VTATA.01	(-)	0.945	ccctaTAAAtcagtggg
V$ZF02/ZNF219.01	(-)	0.945	tgacaCCCCcttgccacctgttg
V$CHRF/CHR.01	(+)	0.944	caatTTGAaggca
V$RORA/RORA1.01	(-)	0.944	gatggtgaaGGTCatggcagctccc
V$TEAF/TEAD4.01	(+)	0.942	ccacATTCccact
V$IKRS/IK1.01	(-)	0.941	cagtGGGAatgtg
V$MYBL/VMYB.05	(-)	0.941	aaggaaactgAACGgcagagg
V$GATA/GATA2.01	(+)	0.941	tgccGATAaggaa
V$ETSF/ERG.02	(-)	0.941	gcataaaaGGAAactgaacgg
V$GFI1/GFI1.02	(-)	0.941	ataAATCagtgggaa
V$DLXF/DLX2.01	(-)	0.94	tagtcactgtAATTgcccc
V$CLOX/CDP.02	(-)	0.94	ccaaagcCAATtacagtgccgcc
V$KLFS/KLF12.01	(-)	0.94	taaatcaGTGGgaatgtgg
V$BRNF/BRN2.01	(-)	0.938	cgAATTtaaatatcagat
V$CAAT/NFYA.01	(-)	0.937	aaagCCAAttacagt
V$CARE/CARF.01	(-)	0.936	acagtGAGGca
V$KLFS/BTEB3.01	(+)	0.936	tcatcatcatGGAGttgtg
V$IRFF/ISRE.02	(-)	0.935	gatttgcataaaagGAAActgaacg
V$HICF/HIC1.01	(-)	0.935	gagTGCCacccag
V$PLZF/PLZF.02	(+)	0.934	ccaTACAgtcctggc
V$PAX5/PAX8.01	(-)	0.933	aacattaagacATGAgtgccttcaaattg

V$CART/ISX.01	(-)	0.932	aaatcaAATTagtgggaaaat
V$BRNF/BRN2.03	(+)	0.932	actaatttgATTTtaatct
V$ZF03/ZNF217.01	(+)	0.931	GAATactggatcc
V$STAT/STAT5B.01	(+)	0.931	atttatcccAGAAcaattt
V$STEM/OCT3_4.02	(-)	0.931	acgatttGCATaaaaggaa
V$ABDB/HOXD13.01	(-)	0.93	gaatcaaaTAAAacatg
V$SORY/HMGA.01	(-)	0.93	attccctataAATCagtgggaat
V$GATA/GATA6.01	(+)	0.929	tgccGATAaggaa
V$ETSF/SPIB.01	(+)	0.929	tgataagtGGAAgtttcctgt
V$ETSF/SPIC.01	(+)	0.929	tgataagtGGAAgtttcctgt
V$CEBP/CEBP.02	(+)	0.929	ggagttgtGAAAggt
V$ARID/JARID2.01	(+)	0.928	tttgatTTTAatctgatattt
V$ABDB/HOXC13.01	(-)	0.928	ggaatcaaaTAAAacat
V$ABDB/HOXD10.01	(-)	0.928	gaatcaaaTAAAacatg
V$HIFF/CLOCK_BMAL1.01	(+)	0.928	cctgtgCACGtgttggg
V$P53F/P53.03	(-)	0.927	gggaatcaaataaaaCATGtttgca
V$ZICF/ZIC3.03	(+)	0.927	agcctCAGCtggact
V$AP2F/AP2.01	(+)	0.927	tttGCCTctggctgg
V$GATA/GATA4.01	(+)	0.926	tgccGATAaggaa
V$CART/CART1.01	(+)	0.926	cccacTAATttgattttaatc
V$CDXF/CDX2.03	(+)	0.926	ccactgatTTATagggaat
V$SORY/HMGIY.01	(+)	0.925	tccaAATTttcccactaatttga
V$EBOX/USF.04	(+)	0.925	tgagaCACTtggcccaa
V$NF1F/NFIB.01	(-)	0.924	ctcctgggccaaaGCCAatta
V$DLXF/DLX4.01	(-)	0.923	tagtcactgtAATTgcccc
V$GATA/GATA5.01	(+)	0.923	tgccGATAaggaa
V$DLXF/DLX1.01	(-)	0.922	tagtcactgtAATTgcccc
V$ABDB/HOXB13.01	(-)	0.922	gaatcaaaTAAAacatg
V$BCDF/PCE1.01	(+)	0.921	tagcccTAATttagcctc
V$AP1F/JUNB.01	(-)	0.921	tgcttaGTCActg
V$AP1R/MARE_ARE.01	(-)	0.921	tgtcatgcttaGTCActgtaattgc
V$CP2F/TCFCP2L1.01	(+)	0.92	taCTGGatccctcttgccg
V$AP1R/MARE.03	(-)	0.92	tgtcatGCTTagtcactgtaattgc
V$MEF2/MEF2.02	(+)	0.919	caacatttaaaAATAgactcgc
V$HESF/HELT.01	(+)	0.918	tgtgCACGtgttggg
O$VTBP/VTATA.02	(-)	0.918	ttgcaTAAAaggaaact
V$NBRE/NUR77.01	(+)	0.917	gtgaaagGTAAacag
V$HOMF/BARHL1.01	(+)	0.916	ggcactgTAATtggctttg

V$CART/MIXL1.01	(+)	0.916	cccacTAATttgattttaatc
V$IRFF/ISGF3G.01	(-)	0.915	gatttgcataaaagGAAActgaacg
V$HAML/AML3.02	(+)	0.915	aactGTGGtcaactt
V$TEAF/TEAD.01	(+)	0.914	ccaCATTcccact
V$IRXF/IRX4.01	(+)	0.913	caaaCATGtttta
V$GFI1/GFI1B.01	(-)	0.912	gtaAATCactcaagg
V$SNAI/SCRT2.01	(+)	0.911	acaaCAGGtgg
V$AP1R/TCF11MAFG.01	(-)	0.911	tgtcatgcttaGTCActgtaattgc
V$IRFF/ISRE.01	(-)	0.91	gatttgcataaaagGAAActgaacg
V$MEF2/RSRFC4.01	(+)	0.909	caacatttttaaaaATAGactcgc
V$GREF/GRE.01	(-)	0.908	tgccttcaaattGTTCtgg
V$INSM/INSM1.01	(+)	0.908	tccctGGGGgcag
V$MYOD/MYF5.01	(+)	0.908	cccagaCAACaggtggc
V$MEF3/SIX.01	(+)	0.908	ccaTCAGgttagg
V$FKHD/FOXJ3.01	(+)	0.908	gaaaggTAAAcaggcca
V$BRNF/TST1.01	(+)	0.906	gggcaATTAcagtgactaa
V$HOXC/PBX_HOXA9.01	(+)	0.906	ccacTGATttataggga
V$ZICF/ZIC2.02	(+)	0.905	agcctCAGCtggact
V$PARF/DBP.01	(+)	0.903	tccttTTATgcaaatcg
V$CREB/ATF.01	(-)	0.903	cacttcTGACgggtcctaacc
V$ESRR/ERR.01	(-)	0.902	aacctgatggtgAAGGtcatggc
V$HMTB/MTBF.01	(+)	0.902	tgatATTTa
V$NKXH/TTF1.02	(+)	0.9	aacttctAGTGgctttcct
V$AP1R/NFE2.02	(-)	0.9	tgtcatgcttaGTCActgtaattgc
V$OCT1/OCT.01	(+)	0.9	tttATGCaaatcgta
V$HBOX/GBX1.01	(+)	0.9	tcccactAATTtgatttta
V$IRFF/IRF7.01	(-)	0.898	tatgagagGAACaggaaacttccac
V$HOXC/HOXB4.02	(+)	0.897	ccacTGATttataggga
V$NFAT/NFAT5.02	(-)	0.897	agtgGGAAaatttggaaca
V$IRXF/IRX2.01	(-)	0.897	aaaaCATGtttgc
V$HBOX/GSH1.01	(+)	0.897	attttcccacTAATttgat
V$HOMF/HMX3.01	(+)	0.896	cccgtcagAAGTggcctcc
V$HOMF/MSX3.01	(+)	0.896	ggcactgTAATtggctttg
V$FKHD/FOXJ1.01	(+)	0.896	gaaaggtaAACAggcca
V$FKHD/FOXK1.01	(+)	0.896	gaaaggtaAACAggcca
V$GRHL/GRHL1.01	(-)	0.895	ggtactGGTTttg
V$CART/PHOX2.01	(+)	0.894	cccacTAATttgattttaatc
V$CP2F/CP2.02	(+)	0.894	tACTGgatccctcttgccg

V$CART/PROP1.02	(+)	0.893	cccactaatttgATTTtaatc
V$IRFF/IRF1.01	(-)	0.893	gatttgcataaaagGAAActgaacg
V$TEAF/TEF1.01	(+)	0.892	ccaCATTcccact
V$HBOX/GBX2.01	(+)	0.892	tcccactAATTtgatttta
V$GREF/ARE.03	(-)	0.891	tgccttcaaattGTTCtgg
V$AP2F/TCFAP2B.01	(+)	0.891	attCCCTgggggcag
V$NOLF/EBF1.01	(-)	0.889	gtcctgCCCCcagggaatcagaa
V$IRXF/IRX3.01	(-)	0.889	aaaaCATGtttgc
V$AP1F/AP1.02	(+)	0.889	ccttGAGTgattt
V$NKX1/NKX11.01	(-)	0.889	aaatcaAATTagtggga
V$NKXH/NKX24.01	(+)	0.888	aacttctAGTGgctttcct
V$GLIF/GLI3.02	(-)	0.888	agtgCCACccagccagc
V$CREB/XBP1.01	(+)	0.886	tcctgtgcACGTgttgggcaa
V$FKHD/FREAC4.01	(+)	0.886	gaaaggtaAACAggcca
V$RORA/REV-ERBA.01	(-)	0.885	tgggaatgtgGTCAcacgaatttta
V$GREF/PRE.01	(-)	0.884	tgccttcaaatTGTTctgg
V$RORA/RORA.01	(-)	0.884	gatggtgaagGTCAtggcagctccc
V$NOLF/OLF1.01	(-)	0.883	ggcagcTCCCcaaagagctggta
V$HOXC/HOX_PBX.01	(+)	0.882	ccacTGATttataggga
V$HBOX/EN2.01	(+)	0.882	tcccactAATTtgatttta
V$GRHL/GRHL2.01	(+)	0.881	aaAACCagtacct
V$GRHL/GRHL3.01	(-)	0.881	ggtactGGTTttg
V$P53F/TP63.02	(-)	0.881	tctgacgggtcctaaCCTGatggtg
V$FKHD/FOXL1.01	(+)	0.881	gaaaggTAAAcaggcca
V$IRXF/IRX6.01	(-)	0.879	aaaaCATGtttgc
V$LHXF/LHX4.01	(+)	0.879	ttttcccactAATTtgattttaa
V$LHXF/LHX2.01	(-)	0.878	attaaaatcaAATTagtgggaaa
V$LEFF/TCF7.01	(-)	0.877	tttgcatAAAAggaaac
V$NR2F/HNF4.01	(-)	0.876	actcctgggcCAAAgccaattacag
V$BRAC/TBR1.01	(+)	0.875	tcatcatggagTTGTgaaaggtaaacagg
V$HOXF/HOXD1.01	(+)	0.875	cactgtAATTggctttggc
V$ARID/ARID5A.01	(+)	0.874	atctgATATttaaaattcgtg
V$HESF/BHLHB2.02	(+)	0.872	tgtgcaCGTGttggg
V$FKHD/HFH1.01	(+)	0.872	gaaaggTAAAcaggcca
V$PAXH/PAX4.02	(-)	0.868	aatcaAATTagtggg
V$LHXF/ISL1.01	(-)	0.866	tagtcactgTAATtgccccggca
V$IRFF/IRF3.01	(-)	0.865	gatttgcataaaagGAAActgaacg
V$LEFF/TCF7L1.01	(-)	0.865	ttaaaatCAAAttagtg

V$HOXF/HOXC8.01	(-)	0.864	cactgtaATTGccccggca
V$IKRS/IK3.01	(+)	0.863	tatagGGAAtaag
V$CART/RAX.01	(-)	0.863	aaatcaAATTagtgggaaaat
V$STEM/OSNT.01	(+)	0.863	tgactaaGCATgacaaatc
V$AP1R/NFE2L2.01	(-)	0.863	tgtcatgctTAGTcactgtaattgc
V$MTF1/MTF-1.02	(-)	0.862	ttttgcaTACGattt
V$AP1R/BACH1.02	(-)	0.862	tgtcatgCTTAgtcactgtaattgc
V$FKHD/FKHRL1.01	(+)	0.862	gaaaggtaAACAggcca
V$P53F/TP63.01	(+)	0.861	tgctcgctgttaagaAATGtgtttg
V$HOXF/HOXA6.01	(-)	0.861	cactgtAATTgccccggca
V$NFKB/HIVEP1.01	(+)	0.861	tgGGGAgctgccatg
V$PERO/PPARG.03	(+)	0.86	catggagttgtgAAAGgtaaaca
V$CART/PRRX2.01	(-)	0.86	aaatcaAATTagtgggaaaat
V$HOXF/HOXA5.01	(+)	0.857	cactgtAATTggctttggc
V$HBOX/EMX2.01	(+)	0.856	tcccacTAATttgatttta
V$CART/ESX1.01	(+)	0.855	cccactAATTtgattttaatc
V$HOXC/PBX1.01	(+)	0.855	tccttGAGTgatttaca
V$NR2F/ARP1.01	(+)	0.855	cccaggagtccacaGGACagagcat
V$HOXF/HOXB6.01	(+)	0.855	cggggcAATTacagtgact
V$SRFF/SRF.02	(-)	0.855	atttgCATAaaaggaaact
V$HOXF/HOXD3.01	(-)	0.854	cactgtAATTgccccggca
V$HOMF/HMX1.01	(-)	0.853	gtcactgTAATtgccccgg
V$CEBP/CEBPG.01	(+)	0.853	cgtgTTGGgcaacca
V$BRN5/BRN5.04	(+)	0.852	accagtacctcATCAtcatggag
V$FAST/FAST1.03	(-)	0.852	tgtcctgtgGACTcctg
V$ZF01/SZF1.01	(-)	0.852	ccaGGGAatcagaacttgtgaggct
V$ZF01/ZBRK1.01	(+)	0.852	actgggtaCCAGctctttggggagc
V$IRFF/IRF2.01	(-)	0.851	gatttgcataaaagGAAActgaacg
V$CART/SHOX2.01	(+)	0.85	tagcctAATTtagcctcacaa
V$E2FF/E2F.02	(-)	0.849	ctcctgggcCAAAgcca
V$FKHD/FHXB.01	(+)	0.849	tatcccAGAAcaatttg
V$LHXF/LHX8.01	(-)	0.849	attaaaatcaAATTagtgggaaa
V$HBOX/VAX1.01	(-)	0.849	caaagccAATTacagtgcc
V$HOMF/BARHL2.01	(+)	0.848	ggcactgTAATtggctttg
V$GATA/GATA3.03	(+)	0.847	tgccGATAaggaa
V$HOXF/HOXB3.01	(-)	0.847	aaatcAAATtagtgggaaa
V$GREF/MRE.01	(-)	0.846	ccaggactgtatGGTCcag
V$LEFF/LEF1.04	(-)	0.846	tccctaTAAAtcagtgg

V$LEFF/TCF7L2.01	(-)	0.846	tttgcatAAAAggaaac
V$SIXF/SIX6.01	(-)	0.846	actcgggAATCaaat
V$NEUR/NEUROD1.01	(-)	0.845	ttgcCACCtgttgtc
V$NKXH/NKX26.01	(+)	0.845	aacttctAGTGgctttcct
V$HBOX/HLXB9.01	(+)	0.845	tcccactAATTgatttta
V$PDX1/IPF1.01	(+)	0.844	gcggcactgTAATtggctt
V$HBOX/EVX1.01	(+)	0.844	tcccactaATTTgatttta
V$RXRF/RXRA.01	(+)	0.843	ataaggaactgtGGTCaacttctag
V$ETSF/ETS2.01	(-)	0.843	gaggaacAGGAaacttccact
V$TALE/MEIS1.02	(+)	0.843	ggatgagctGTGAagct
V$CART/PHOX2A.01	(-)	0.842	aaatcaAATTagtgggaaaat
V$PBXC/PBX1_MEIS1.02	(+)	0.842	tgagTGATttacaattt
V$SIXF/SIX1.01	(+)	0.842	aaggggggTGTCaggg
V$HBOX/MEOX1.01	(+)	0.842	ccggggcAATTacagtgac
V$PAX5/PAX5.03	(+)	0.842	ccgtgCCTCtgccgttcagtttcctttta
V$HOXF/HOXA2.01	(+)	0.841	cactgTAATtggctttggc
V$HOXF/HOXA3.02	(+)	0.841	cggggcaATTAcagtgact
V$PARF/TEF_HLF.01	(-)	0.839	acacaTTTCttaacagc
V$HOXF/HOXB8.01	(+)	0.839	cggggcaATTAcagtgact
V$FKHD/XFD3.01	(+)	0.839	gaaaggtaAACAggcca
V$RXRF/PXR_RXR.01	(-)	0.837	gtgaaGGTCatggcagctccccaaa
V$CREB/BATF3.01	(-)	0.836	tgatGATGaggtactggtttt
V$CART/XVENT2.01	(-)	0.834	aatttTAAAtatcagattaaa
V$TALE/TGIF2.01	(+)	0.832	ggatgagctGTGAagct
V$PRDF/PRDM1.02	(+)	0.832	gataagtGGAAgttcctg
V$GCMF/GCM1.03	(-)	0.831	cttatTCCCtataaa
V$PAX7/PAX7.01	(+)	0.829	tttatttGATTcccg
V$HOXF/HOXA1.01	(+)	0.828	cactgtAATTggctttggc
V$HEAT/HSF1.04	(-)	0.827	atgagaggaacaggaaacTTCCact
V$NR2F/COUP.01	(+)	0.827	cccaggagtccacAGGAcagagcat
V$NFKB/NFKAPPAB.02	(+)	0.826	gtGGAAgtttcctgt
V$HBOX/EVX2.01	(+)	0.824	tcccacTAATttgatttta
V$BRN5/POU6F2.01	(-)	0.823	ttaaatatcagATTAaaatcaaa
V$DMRT/DMRT7.01	(+)	0.823	tgtgTTTGcctcactgttttg
V$CART/ALX4.01	(-)	0.822	aggctaAATTaggctatattt
V$CHRE/CHREBP_MLX.01	(+)	0.822	CAGGtggcaaggggggtg
V$HNF1/HNF1.01	(+)	0.82	tCTTAatgttttaaata
V$RXRF/RXR_RXR.01	(-)	0.819	tccagcttcacaGCTCatccccaca

V$ZF16/PRDM15.01	(-)	0.819	gaagccaggacTGTA
V$HOMF/HMX2.03	(+)	0.818	ggcactGTAAttggctttg
V$HNF6/HNF6.01	(-)	0.815	gattaaaaTCAAattag
O$XCPE/XCPE1.01	(+)	0.814	gtGCGGcaaca
V$NR2F/PNR.01	(-)	0.814	ggtccagctTCACagctcatcccca
V$PRDF/PRDM1.01	(+)	0.814	gagttgtGAAAggtaaaca
V$SORY/SOX1.03	(-)	0.812	agtCACTgtaattgccccggcag
V$MYOD/TCF21.01	(-)	0.809	cccaaagaGCTGgtacc
V$CTCF/CTCF.01	(+)	0.809	acaacaggtggcaaGGGGgtgtcaggg
V$ETSF/GABPB1.01	(+)	0.807	gacccgtCAGAagtggcctcc
V$CREB/E4BP4.01	(+)	0.806	ttcctttatGCAAatcgtat
V$E2FF/E2F.01	(-)	0.806	aattagtggGAAAattt
V$NR2F/COUPTFII.01	(+)	0.805	atggagttgtgaaaGGTAaacaggc
V$DMRT/DMRT4.01	(-)	0.803	aaaatttggAACAaattgcct
V$SP1F/SP2.01	(+)	0.803	ctacctggcgGCACtgt
V$PAX6/PAX6.05	(-)	0.801	agctgagGCTTggttgccc
V$E2FF/E2F4_DP2.01	(-)	0.801	aattaGTGGgaaaattt
V$RXRF/CAR_RXR.01	(-)	0.8	gtgaaGGTCatggcagctccccaaa
V$DMRT/DMRT3.01	(+)	0.798	gcaatttgtTCCAaattttcc
V$PRDM/PRDM4.01	(+)	0.785	aacaaTTTGaaggcactcatgtcttaatg
V$SNAP/PSE.02	(-)	0.783	atttgCATAaaaggaaact
V$DMRT/DMRT1.01	(+)	0.782	tgtgtttgcctcaCTGTtttg
V$CART/HESX1.01	(+)	0.773	cccacTAATttgattttaatc
V$STAF/STAF.01	(-)	0.772	ttcacagctcatcCCCAcaatggtatgagag
V$HUB1/ZNF282.01	(+)	0.772	TTTAgcctcacaagt
V$NR2F/TR2_TR4.01	(+)	0.771	cccaggagtccacaGGACagagcat
V$STAF/ZNF76_143.01	(+)	0.768	taattggctttggCCCAggagtccacaggac
V$SORY/SOX21.03	(+)	0.763	tgtGAAAggtaaacaggccaggc
V$SORY/SRY.04	(-)	0.763	agtcACTGtaattgccccggcag
V$ZF05/ZNF410.02	(-)	0.757	tgtTCTGggataaat
V$NR2F/TR4.01	(+)	0.756	atggagttgtgaaAGGTaaacaggc
V$DUXF/DUX4.01	(+)	0.751	cttGAGTgatttaca

Table A8. Complete table of TFBSs of *Mus musculus* Cebpa Downstream CNS4 region

Matrix	strand	Matrix sim.	Sequence
V$ZFHX/DELTAEF1.01	(-)	1	gactcACCTgact
V$ZFXY/ZFX.01	(-)	1	ggGGCCtggcc
V$TAIP/CSRNP1.01	(+)	1	AGAGtgt
V$TAIP/CSRNP3.01	(+)	1	aGAGTgt
V$PARF/TEF.01	(+)	1	gaagtttatGTAAttat
V$TALE/TGIF.01	(+)	1	cagaaatgtGTCAgagg
V$MZF1/MZF1.01	(+)	1	gtGGGGacgat
V$ZF02/ZKSCAN3.01	(-)	1	atcgtCCCCacctccccacggca
V$ETSF/ETV1.02	(+)	0.997	acaaggcaGGAAggcagctct
V$TEAF/TEAD4.01	(-)	0.996	tgacATTCctggc
V$ETSF/ETV3.02	(+)	0.995	caagcacaGGAAgcgactgga
V$AP1R/MARE.01	(+)	0.993	cccggctgTCAGcacagggcagcag
V$CART/S8.01	(+)	0.992	ttatgTAATtatgtggggcag
V$EVI1/MEL1.02	(+)	0.992	gaggccaGATGagcatc
V$HAND/TAL1_E2A.02	(+)	0.991	gggatgaCAGCaggtattgag
V$NKXH/NKX25.05	(+)	0.99	ggtatTGAGtgggactgca
V$AP1F/JUNB.01	(+)	0.989	aggtgaGTCAcag
V$BRAC/TBX20.02	(-)	0.987	tctgctgccctgtgctGACAgccgggctg
V$ZFHX/ZEB1.01	(-)	0.986	gactcACCTgact
V$MZF1/MZF1.03	(+)	0.986	gtGGGGaggtg
V$DLXF/DLX2.01	(+)	0.985	aagtttatgtAATTatgtg
V$AP1F/FOSL1.01	(+)	0.984	aggtgAGTCacag
V$AP1R/BACH2.01	(+)	0.984	gagtcaggTGAGtcacaggccaggc
V$ZFHX/SIP1.01	(-)	0.983	cccacACCTggag
V$EGRF/WT1.01	(+)	0.982	tggcgTGGGggcaggagca
V$DLXF/DLX1.01	(+)	0.982	aagtttatgtAATTatgtg
V$NEUR/NGN_NEUROD.01	(-)	0.982	tgctCATCtggcctc
V$HOXF/HOXC8.01	(-)	0.982	ccacataATTAcataaact
V$AP1F/AP1.01	(+)	0.981	aggtgAGTCacag
V$SPZ1/SPZ1.01	(-)	0.981	aGGAGggtaga
V$NF1F/NF1.03	(+)	0.979	tttccctgtcactGCCAggaa
V$ETSF/ETS1.01	(+)	0.979	caagcacaGGAAgcgactgga
V$HAML/AML1.02	(-)	0.979	gcatGTGGtgtgctc
V$AP1F/BATF.01	(-)	0.978	ctgtgaCTCAcct
V$TEAF/TEAD.01	(-)	0.978	tgaCATTcctggc
V$DLXF/DLX3.03	(-)	0.975	tgccccacatAATTacata
V$DLXF/DLX4.01	(-)	0.975	tgccccacatAATTacata
V$NF1F/NFIB.01	(+)	0.973	acgatggctccgaGCCAgaaa
V$AP1F/JUND.01	(+)	0.973	aggtgAGTCacag
V$KLFS/EKLF.02	(-)	0.973	gtcctgGGTGtggaggagc
V$ZFHX/AREB6.02	(-)	0.972	gactCACCtgact
V$TALE/MEIS1.01	(+)	0.972	tctttcccTGTCactgc
V$AP1F/FOSL2.01	(+)	0.971	aggtgAGTCacag
V$MYBL/CMYB.02	(-)	0.969	tttgctcatTAACtgaaactg
V$ETSF/FLI1.02	(+)	0.969	caagcacaGGAAgcgactgga
V$KLFS/BKLF.01	(-)	0.969	gtcctGGGTgtggaggagc
V$LHXF/LMX1A.02	(+)	0.969	agtttcagTTAAtgagcaaactc
V$ETSF/PEA3.01	(+)	0.968	caagcacAGGAagcgactgga
V$STAT/STAT3.02	(-)	0.967	gacaTTCCtggcagtgaca
V$CEBP/CEBPB.01	(+)	0.966	gaactttgGAAAtct
V$NKRF/NKRF.01	(+)	0.966	aacTCCTcaga
V$HESF/DEC2.01	(-)	0.964	tcctcaCATGagatt
V$NKXH/NKX32.01	(+)	0.96	ggtattgAGTGggactgca

V$BRN5/BRN5.04	(-)	0.96	ggagtttgctcATTAactgaaac
V$WHNF/WHN.01	(+)	0.959	gggACGCgaag
V$STAT/STAT6.01	(+)	0.958	aaacTTCCctagaacggag
V$CREB/E4BP4.01	(+)	0.958	cgaagtttatGTAAttatgtg
V$ZFHX/AREB6.01	(-)	0.956	gactcACCTgact
V$HEAT/HSF2.02	(+)	0.956	ctctgctcgaggAGAAatctcatgt
V$HAML/AML2.01	(-)	0.956	gcaTGTGgtgtgctc
V$AP1F/JUNDM2.02	(+)	0.955	aggtgaGTCAcag
V$PAX3/PAX3.03	(-)	0.955	gctctCATGgtgacattcc
V$ETSF/ELF2.01	(+)	0.952	caagcacaGGAAgcgactgga
O$INRE/DINR.01	(+)	0.952	ttTCAGttaat
V$AHRR/AHRARNT.03	(+)	0.951	gtgggtgctgGCGTggggcaggag
V$ETSF/CETS1P54.01	(+)	0.951	caagcaCAGGaagcgactgga
V$MITF/MIT.01	(+)	0.951	aatctCATGtgagga
V$HOXF/HOXA3.02	(+)	0.95	ttatgtaATTAtgtggggc
V$HOXF/HOXC6.01	(-)	0.95	ccacatAATTacataaact
V$HOXF/NANOG.01	(+)	0.95	ttcagttAATGagcaaact
V$HIFF/ARNTL.01	(+)	0.946	tggagtcaGGTGagtca
V$AP2F/TCFAP2A.02	(-)	0.945	tgggCCTGaggaatt
V$HOXF/HOXA5.01	(+)	0.944	ttatgtAATTatgtggggc
V$PARF/VBP.01	(+)	0.943	gaagtttatGTAAttat
V$SORY/SOX3.01	(+)	0.941	gcagcaGAAAggacaggggacaa
V$OCT1/OCT1.03	(+)	0.941	tttatgtaATTAtgt
V$GREF/ARE.02	(-)	0.94	cagtcgcttcctGTGCttg
V$DLXF/DLX5.01	(-)	0.94	tgccccacatAATTacata
V$MYBL/VMYB.05	(+)	0.939	acttccctagAACGgagggag
V$HOXF/HOXD3.01	(-)	0.939	ccacatAATTacataaact
V$NFAT/NFAT5.02	(+)	0.937	ctttGGAAatctaccctcc
V$HBOX/MEOX1.01	(+)	0.937	tttatgtAATTatgtgggg
V$CARE/CARF.01	(-)	0.936	acagtGAGGcc
V$ETSF/ERG.02	(+)	0.936	acaaggcaGGAAggcagctct
V$PAX5/PAX2.02	(+)	0.936	ggactgcaggccTGACatccttagctcct
V$TEAF/TEF1.01	(-)	0.935	tgaCATTcctggc
V$MITF/CLEAR.01	(+)	0.934	ggagTCAGgtgagtc
V$ETSF/ETS2.01	(+)	0.933	caagcacAGGAagcgactgga
V$ETSF/GABPA.02	(+)	0.932	caagcacaGGAAgcgactgga
V$NR2F/HNF4G.01	(-)	0.932	ggtagatttccAAAGttctttcttg
V$ZF07/ZNF263.01	(+)	0.932	tctacccCTCCtcttt
V$NEUR/NEUROG.01	(-)	0.931	tgcTCATctggcctc
V$HOMF/NOBOX.01	(-)	0.928	ccccacaTAATtacataaa
V$EBOX/CMYC.02	(-)	0.928	agtccccCGTGcaactg
V$GLIF/ZIC2.01	(-)	0.927	cccacctCCCCacggca
V$ETSF/ELK1.03	(+)	0.926	caagcacaGGAAgcgactgga
V$NR2F/HNF4A.01	(-)	0.925	ggtagatttcCAAAgttctttcttg
V$BRAC/EOMES.02	(+)	0.923	acaagctccaGGTGtgggcgagtcccaga
V$MYBL/VMYB.01	(-)	0.923	tttgctcattAACTgaaactg
V$NACA/NACA1.01	(+)	0.923	aggcCAGAtgagc
V$HBOX/GSH2.02	(-)	0.923	cccacatAATTacataaac
V$HOMF/HMX3.02	(+)	0.922	gcagtttcagTTAAtgagc
V$PAX6/PAX4_PD.01	(+)	0.921	gttGCACgggggactgaat
V$NFKB/NFKAPPAB.01	(-)	0.921	ctGGGActcgcccac
V$HAML/AML3.02	(-)	0.921	gcatGTGGtgtgctc
V$EBOX/MAX.03	(-)	0.921	agtccCCCGtgcaactg
V$MYBL/VMYB.02	(+)	0.92	acttccctagAACGgagggag
V$ETSF/SPI1.04	(+)	0.919	tcatgtgaGGAAggtggagtc

V$PDX1/IPF1.01	(+)	0.918	aagtttatgTAATtatgtg
V$HBOX/VAX1.01	(+)	0.916	tttatgtAATTatgtgggg
V$HOXF/HOXA1.01	(-)	0.915	ccacatAATTacataaact
V$HOMF/MSX1.01	(+)	0.914	gtttatgtAATTatgtggg
V$HOXF/HOXA6.01	(-)	0.914	ccacatAATTacataaact
V$HNF1/HNF1.04	(+)	0.914	cagtttcaGTTAatgag
V$MYRF/MYRF.01	(-)	0.913	cagtgACAGggaa
V$HOXF/HOXB5.01	(+)	0.913	ttatgTAATtatgtggggc
V$KLFS/KKLF.01	(+)	0.913	ctgtgccgtGGGGaggtgg
V$HBOX/VAX2.01	(+)	0.913	tttatgtaATTAtgtgggg
V$ETSF/NRF2.01	(+)	0.911	caagcacaGGAAgcgactgga
V$MAZF/MAZ.01	(+)	0.91	tgggGAGGtgggg
V$ZF02/ZNF219.01	(-)	0.91	tcctgCCCCcacgccagcaccca
V$PAX6/PAX6.02	(+)	0.908	tgggcgagtcCCAGagcag
V$CART/SHOX2.01	(+)	0.908	ttatgtAATTatgtggggcag
V$HOXF/HOXB6.01	(-)	0.908	ccacatAATTacataaact
V$TALE/TGIF1.01	(-)	0.908	atacctgctGTCAtccc
V$HOXF/HOXD1.01	(+)	0.907	ttatgtAATTatgtggggc
V$ABDB/HOXC9.01	(-)	0.906	tttgctcaTTAActgaa
V$CART/MIXL1.01	(-)	0.906	ccacaTAATtacataaacttc
V$AP1R/NFE2.01	(+)	0.906	gagtcagGTGAgtcacaggccaggc
V$ETSF/GABP.01	(+)	0.905	caagcacaGGAAgcgactgga
V$HOXF/HOXC5.01	(-)	0.905	ccacatAATTacataaact
V$PARF/HLF.01	(+)	0.904	gaagtttatGTAAttat
V$ZF02/ZBTB7.03	(-)	0.904	atcgtCCCCacctccccacggca
V$EVI1/EVI1.07	(-)	0.902	ccctcAAGAtgctcatc
V$INSM/INSM1.01	(+)	0.901	tgcacGGGGgact
V$NOLF/EBF1.01	(-)	0.9	attcagTCCCccgtgcaactgca
V$HOXF/HOXB3.01	(+)	0.899	ttatgTAATtatgtggggc
V$MYT1/MYT1.02	(-)	0.899	ccaAAGTtctttc
V$EGRF/EGR1.02	(+)	0.898	tggcgtggGGGCaggagca
V$TALE/PKNOX2.01	(-)	0.898	atacctgctGTCAtccc
V$HOXF/HOXB7.01	(+)	0.897	ttatgtAATTatgtggggc
V$HAND/PARAXIS.01	(+)	0.897	agcacACCAcatgcagttgca
V$CP2F/TCFCP2L1.01	(+)	0.896	gaCTGGattccacttgccc
V$CART/ALX3.01	(-)	0.895	ccacatAATTacataaacttc
V$ETSF/ELF1.01	(+)	0.893	caagcacaGGAAgcgactgga
V$HOXF/HOXA2.01	(+)	0.893	ttatgTAATtatgtggggc
V$STAT/STAT1.02	(+)	0.891	tcactgccaGGAAtgtcac
V$MAZF/MAZR.01	(+)	0.89	ggaggtGGGGacg
V$BRNF/BRN2.04	(+)	0.89	gtttatgTAATtatgtggg
V$TALE/MRG2.01	(-)	0.89	atacctgctGTCAtccc
V$BARB/BARBIE.01	(-)	0.889	agggAAAGaggaggg
V$NGRE/IR1_NGRE.01	(+)	0.889	ctgctcgaGGAGaaa
V$HOMF/BARX2.02	(+)	0.888	gtttatgtAATTatgtggg
V$AP1R/BACH1.01	(+)	0.888	gagtcaggTGAGtcacaggccaggc
V$CART/ESX1.01	(-)	0.887	ccacatAATTacataaacttc
V$SP1F/SP1.01	(+)	0.887	cgtggGGGCaggagcac
V$HOXF/HOXB4.01	(+)	0.885	ttatgtAATTatgtggggc
V$PAX5/PAX8.01	(-)	0.885	gcttcctgtgcTTGAcagtgaggcccaag
V$CART/RHOX6.01	(+)	0.884	ttatgTAATtatgtggggcag
V$MYBL/VMYB.03	(+)	0.883	acttccctagAACGgagggag
V$XBBF/RFX3.02	(-)	0.882	ctgctctcatgGTGAcatt
V$PAXH/PAX6_HD.01	(+)	0.882	tatgtAATTatgtgg
V$HBOX/HLXB9.01	(-)	0.882	cccacatAATTacataaac

V$PBXC/PBX1_MEIS1.02	(-)	0.881	cctgTGCTtgacagtga
V$CART/PRRX1.01	(+)	0.879	ttatgtAATTatgtggggcag
V$SP1F/TIEG.01	(+)	0.879	tgctCGGGtgtgtgaga
V$EGRF/NGFIC.01	(+)	0.878	gctgGCGTgggggcaggag
V$NR2F/HNF4.03	(+)	0.878	gcagaaaggaCAGGgacaagctcc
V$PAXH/PAX4.02	(+)	0.877	tatgtAATTatgtgg
V$HBOX/EVX1.01	(-)	0.876	cccacataATTAcataaac
V$PARF/TEF_HLF.01	(-)	0.875	cataaTTACataaactt
V$AP1R/NFE2L2.01	(+)	0.874	gagtcaggtGAGTcacaggccaggc
V$HBOX/EN1.02	(+)	0.874	tttatgtAATTatgtgggg
V$HNF1/HMBOX.01	(-)	0.874	tctttcttGTTAgctcc
V$PRDF/BLIMP1.01	(-)	0.872	tgacaggGAAAgaggaggg
V$NBRE/NBRE.01	(+)	0.872	cagaAAGGacagggg
V$CART/VSX1.01	(-)	0.871	ccacatAATTacataaacttc
V$XBBF/RFX2.02	(-)	0.871	cagtccccgtgCAACtgc
V$HOXF/HOXC4.01	(+)	0.871	ttatgtAATTatgtggggc
V$CART/ALX4.01	(-)	0.87	ccacatAATTacataaacttc
V$IRFF/IRF4.03	(-)	0.87	tttgctcattaactGAAActgctct
V$ETSF/ELK3.01	(+)	0.87	caagcacAGGAagcgactgga
V$YY1F/REX1.01	(-)	0.87	aatcagCCATtgctgggcctgag
V$IRFF/ISRE.01	(-)	0.869	tttgctcattaactGAAActgctct
V$SORY/HBP1.01	(+)	0.868	tttcagttAATGagcaaactcct
V$PRDF/PRDM1.02	(-)	0.868	tgacaggGAAAgaggaggg
V$LHXF/LHX3.02	(+)	0.868	cgcgaagtttatgTAATtatgtg
V$AP1R/MAFB.01	(+)	0.867	acagcccggctgTCAGcacagggca
V$CART/CART1.01	(-)	0.866	ccacaTAATtacataaacttc
V$HOXF/HOXA4.01	(-)	0.866	ccacatAATTacataaact
V$CHRE/CHREBP_MLX.01	(-)	0.866	CATGtggtgtgctcctg
V$LHXF/LHX6.01	(+)	0.866	aagtttatgTAATtatgtggggc
V$HOXH/MEIS1B_HOXA9.01	(-)	0.864	TGACacatttctggc
V$DICE/DICE.01	(+)	0.863	ggctGTCAgcacagg
V$NKX6/NKX63.01	(+)	0.863	tcagtTAATgagcaa
V$NR2F/TR2_TR4.01	(+)	0.861	gcagaaaggacaggGGACaagctcc
V$LHXF/LHX9.01	(-)	0.861	tgccccacatAATTacataaact
V$PTF1/PTF1.01	(-)	0.86	cacaCCTGgagcttgtcccct
V$CART/PROP1.02	(-)	0.86	ccacataattacATAAacttc
V$IRFF/IRF5.01	(-)	0.86	tttgctcattaactGAAActgctct
V$CART/PHOX2A.01	(-)	0.859	ccacatAATTacataaacttc
V$NF1F/NF1.01	(+)	0.859	acgATGGctccgagccagaaa
V$EGRF/EGR2.01	(+)	0.858	gctgGCGTgggggcaggag
V$PERO/PPARG.03	(+)	0.857	acagggcagcagAAAGgacaggg
V$EGRF/EGR3.01	(+)	0.856	gctgGCGTgggggcaggag
O$XCPE/XCPE1.01	(+)	0.855	ggGCAGgagca
V$ZF01/SZF1.01	(-)	0.855	agaGGGAatcagccattgctgggcc
V$HESF/BHLHB2.01	(+)	0.853	gagtcAGGTgagtca
V$RXRF/RXRA.01	(+)	0.852	agaaaggacaggGGACaagctccag
V$HESF/DEC1.01	(+)	0.852	atctCATGtgaggaa
V$EBOX/USF2.01	(-)	0.852	tgactcaCCTGactcca
V$ETSF/ELK4.02	(+)	0.851	caagcacaGGAAgcgactgga
V$NFKB/HIVEP1.01	(-)	0.851	ctGGGActcgcccac
V$TALE/TGIF2LX.01	(-)	0.851	cctacaccgGGCAagtg
V$LHXF/LHX4.01	(+)	0.851	aagtttatgtAATTatgtggggc
V$NKX6/NKX61.03	(-)	0.85	ccacaTAATtacata
V$GCMF/GCM1.03	(-)	0.85	cttgtCCCCtgtcct
V$IKRS/IK3.01	(+)	0.848	tgccaGGAAtgtc

V$HBOX/EMX2.01	(-)	0.847	cccacaTAATtacataaac
V$PARF/DBP.01	(+)	0.846	gaagtTTATgtaattat
V$IRFF/ISGF3G.01	(-)	0.841	tttgctcattaactGAAActgctct
V$SIXF/SIX1.01	(-)	0.841	cagagggAATCagcc
V$NEUR/NEUROD1.01	(-)	0.84	tgctCATCtggcctc
V$CART/ARX.01	(-)	0.84	ccacatAATTacataaacttc
V$PAX5/PAX5.01	(+)	0.84	tcctcaGACAaggcaggaaggcagctctt
V$TALE/PKNOX1.01	(-)	0.838	atacctgctGTCAtccc
V$EBOX/USF1.02	(+)	0.838	ggagTCAGgtgagtcac
V$ABDB/HOXC13.02	(-)	0.837	tttgctcaTTAActgaa
V$BRNF/BRN3.03	(+)	0.837	tttcagtTAATgagcaaac
V$TALE/TGIF2.01	(-)	0.837	atacctgctGTCAtccc
V$HNF6/HNF6.01	(-)	0.836	agtcccacTCAAtacct
V$EVI1/EVI1.06	(+)	0.836	ctaacaAGAAagaactt
V$CAAT/ACAAT.01	(+)	0.834	cccaGCAAtggctga
V$LHXF/LHX3.01	(+)	0.834	aagtttatGTAAttatgtggggc
V$RORA/REV-ERBA.02	(+)	0.832	ccagaaatgtGTCAgaggccagatg
V$NFKB/NFKAPPAB50.01	(+)	0.832	tggGCGAgtcccaga
V$HAND/HEN1.01	(+)	0.832	gtggagtCAGGtgagtcacag
V$LHXF/LHX1.01	(+)	0.832	aagtttatgtAATTatgtggggc
V$LHXF/LMX1B.02	(+)	0.832	aagtttatgtAATTatgtggggc
O$VTBP/LTATA.01	(+)	0.832	agaTAAAaggagttctt
V$SNAI/SCRT2.01	(+)	0.831	gctcCAGGtgt
V$HNF1/TCF2.01	(-)	0.831	gaagtttgGTTAagaac
V$CART/PHOX2B.01	(+)	0.83	ttatgtAATTatgtggggcag
V$NOLF/OLF1.01	(-)	0.827	cccaccTCCCcacggcacagggg
V$HEAT/HSF1.03	(+)	0.825	agggagctaacaAGAAagaactttg
V$NR2F/ARP1.01	(+)	0.822	gcagaaaggacaggGGACaagctcc
V$LTSM/LTSM.01	(-)	0.821	tacctgctgtCATCc
V$DMRT/DMRT5.01	(+)	0.82	ggagagtGTCActgtgtgggt
V$RREB/RREB1.01	(-)	0.815	cCCCAcgccagcacc
V$ZF01/ZBRK1.01	(-)	0.815	agagggaaTCAGccattgctgggcc
V$RXRF/THRA.01	(+)	0.81	cctgtgccgtggggAGGTggggacg
V$NKXH/NKX23.01	(-)	0.81	tttccaaaGTTCtttcttg
V$NR2F/TR2.01	(+)	0.809	tcaggtgagtcacaGGCCaggcccc
V$NR2F/COUPTFII.01	(+)	0.803	agaaatgtgtcagaGGCCagatgag
V$CTCF/CTCF.01	(+)	0.803	tcccagagcagcccGGGGagagtgtca
O$VTBP/ATATA.01	(+)	0.797	atgtaatTATGtggggc
V$RXRF/RAR_RXR.01	(+)	0.796	aggtgagtcacAGGCcaggcccctg
V$RXRF/RXR_RXR.01	(+)	0.793	aaatgtgtcagaGGCCagatgagca
V$BCL6/BCL6.01	(-)	0.791	tccTTCCtagcatggct
V$ZF02/ZNF202.01	(-)	0.79	cctgccCCCAcgccagcacccac
V$ZTRE/ZTRE.01	(+)	0.789	gcagcccgGGGAgagtg
V$HOXH/MEIS1A_HOXA9.01	(-)	0.788	TGACacatttctggc
V$TALE/MRG1.02	(-)	0.785	ttgtCCCCtgtcctttc
V$BCL6/BCL6.02	(-)	0.784	tccttccTAGCatggct
V$SNAP/PSE.01	(+)	0.782	gTCACcatgagagcagttt
V$ESRR/ESRRG.01	(+)	0.782	tcaggtgagtcacaGGCCaggcc
V$RXRF/VDR_RXR.06	(+)	0.78	agtcaggtgagtcacAGGCcaggcc
V$AHRR/AHR.01	(+)	0.78	gtgggtgctgGCGTggggcaggag
V$BRN5/POU6F2.01	(-)	0.776	ggagtttgctcATTAactgaaac
V$NR2F/HNF4.02	(-)	0.776	ggtagatttccAAAGttctttcttg
V$NR2F/TR4.02	(+)	0.765	gcagaaAGGAcaggggacaagctcc
V$NRSF/NRSF.02	(+)	0.764	cttagctcctccACACccaggacagcccggc
V$RORA/RORA.02	(+)	0.762	gtggagtcaGGTGagtcacaggcca

V$NR2F/NR2F6.01	(+)	0.761	agaaatgtgtcaGAGGccagatgag
V$OAZF/ROAZ.01	(+)	0.757	gaGCAGcccggggagag
V$SORY/SOX10.03	(+)	0.754	gaGGAAggtggagtcaggtgagt

Table A9. Complete table of TFBSs of *Mus musculus* Cebpa Downstream CNS5 region

Matrix	strand	Matrix sim.	Sequence
V$SMAD/GC_SBE.01	(-)	1	ctgggCTCCag
V$NEUR/MASH1.01	(+)	1	cacacacCTGCtgtg
V$TAIP/CSRNP1.01	(+)	1	AGAGtga
V$TAIP/CSRNP3.01	(+)	1	aGAGTga
V$EVI1/MEL1.02	(+)	1	gcaagatGATGagtggt
V$FKHD/FOXP1_ES.01	(+)	1	gttcgaaAACAcaacca
V$ZF02/ZNF219.01	(-)	0.996	ccgcaCCCCccccccaaccttag
V$ZICF/ZIC2.02	(-)	0.995	tcacaCAGCaggtgt
V$HAND/TAL1_E2A.02	(-)	0.995	ggtcacaCAGCaggtgtgtgt
V$ZF11/ZBTB3.01	(+)	0.993	gccaGCCAggt
V$AP1F/AP1.01	(+)	0.991	cggtgAGTCagaa
V$MYOD/TCF12.01	(-)	0.989	acacagcaGGTGtgtgt
V$FKHD/ILF1.01	(+)	0.989	tgttcgaaAACAcaacc
V$KLFS/KLF2.01	(-)	0.987	ccagtGGGTgggaaggcag
V$ZICF/ZIC1.01	(-)	0.986	tcacaCAGCaggtgt
V$AP1F/JUNB.01	(+)	0.984	cggtgaGTCAgaa
V$NKXH/NKX25.05	(+)	0.983	gatgaTGAGtggtgtgttc
V$ZFHX/ZEB1.01	(+)	0.983	cacacACCTgctg
V$AP1R/BACH2.01	(+)	0.983	gggtgcggTGAGtcagaaacaaaca
V$IKRS/IK2.01	(-)	0.982	ctctGGGAatccc
V$MYOD/E47.01	(-)	0.981	acacaGCAGgtgtgtgt
V$SORY/SOX5.01	(-)	0.981	ggtgaaCAATctagctggcaggt
V$AP1F/BATF.01	(-)	0.981	ttctgaCTCAccg
V$ZFXY/ZFX.01	(-)	0.98	atGGCCtgggc
V$NF1F/NFIB.01	(+)	0.977	atcctggcttcttGCCAgcca
V$ZICF/ZIC3.01	(-)	0.976	tcacaCAGCaggtgt
V$HAML/AML2.01	(-)	0.976	tctTGTGgttgtgtt
V$BEDF/ZBED4.02	(+)	0.975	ggttgggGGGGgggt
V$ZF02/ZBP89.01	(-)	0.975	gcacccccCCCCcaaccttagtg
V$EVI1/EVI1.07	(+)	0.972	gaggcAAGAtgatgagt
V$AP1F/FOSL1.01	(+)	0.971	cggtgAGTCagaa
V$EGRF/WT1.01	(-)	0.97	gtgtgTGGGggccagggta
V$FKHD/HFH3.01	(+)	0.97	cagaaacAAACaagcca
V$IKRS/IK1.01	(-)	0.969	ctctGGGAatccc
V$XBBF/RFX5.01	(-)	0.969	gggcactgtgAGCAacatc
V$HAML/AML3.02	(-)	0.968	tcttGTGGttgtgtt
V$CEBP/CEBPA.01	(-)	0.967	ggccttgtGTAAcca
V$CEBP/CEBPD.01	(-)	0.967	ggcctTGTGtaacca
V$NF1F/NF1.03	(+)	0.965	aaacaagccacctGCCAgcta
V$ZTRE/ZTRE.03	(-)	0.965	caCTCCcaacccttcac
V$CEBP/CEBPB.01	(+)	0.965	cgggttgaGTAAttc
V$SAL2/SALL2.01	(-)	0.964	gtggGTGGgaa
V$MAZF/MAZR.01	(+)	0.962	ggggggGGGGtgc
V$AP1F/JUND.01	(+)	0.961	cggtgAGTCagaa
V$ZTRE/ZTRE.04	(+)	0.961	ttgGGAGtgagagggct
V$AP1F/FOSL2.01	(+)	0.96	cggtgAGTCagaa
V$CLOX/CDPCR3HD.01	(-)	0.957	gtctcatgGATGccctgttgcca
V$P53F/P53.03	(+)	0.957	aggctgatcccaggaCAAGtcacag
V$RBPF/RBPJK.02	(-)	0.957	tgggTGGGaaggc
V$AP2F/TCFAP2A.01	(+)	0.955	attcccagAGGCaag
V$NEUR/ATOH1.01	(+)	0.954	gggcCACCtggccac
V$SORY/SOX30.01	(-)	0.954	ggtgaACAAtctagctggcaggt

V$KLFS/IKLF.01	(+)	0.95	ggttgGGGGgggggtgcgg
V$GLIF/ZIC2.01	(-)	0.949	cgcacccCCCCcccaac
V$SORY/SOX15.01	(-)	0.947	ggtgaACAAtctagctggcaggt
V$KLFS/BKLF.02	(+)	0.946	ggttggGGGGggggtgcgg
V$KLFS/KLF6.01	(+)	0.946	gttgggGGGGgggtgcggt
V$ZF02/ZBTB7.01	(-)	0.946	caccccCCCCccaaccttagtgg
V$FKHD/FOXP1.01	(+)	0.945	tgttcgaAAACacaacc
V$GLIF/GLIS3.01	(-)	0.945	cgcaCCCCcccccaac
V$AP2F/TCFAP2C.01	(+)	0.945	attCCCAgaggcaag
V$SPZ1/SPZ1.01	(+)	0.944	gGGGGggtgcg
V$PAX3/PAX3.03	(-)	0.944	catctCATGgagtctcatg
V$HEAT/HSF1.05	(-)	0.943	agccaggatctgggctCCAGaagca
V$KLFS/KKLF.01	(+)	0.938	gttggggggGGGGtgcggt
V$SORY/SOX7.01	(-)	0.937	ggtgaACAAtctagctggcaggt
V$PARF/HLF.01	(-)	0.932	aggccttgtGTAAccag
V$HAND/SCX.01	(-)	0.932	tagctggcaggTGGCttgttt
V$AP1F/JUNDM2.02	(-)	0.931	ttctgaCTCAccg
V$FKHD/FOXA1.01	(+)	0.931	tgttcgaaAACAcaacc
V$SORY/SRY.01	(-)	0.93	ggtgaACAAtctagctggcaggt
V$NDPK/NM23.01	(+)	0.929	tgGGGGggggtgcggt
V$SREB/SREBP.01	(-)	0.929	gacTCACcgcacccc
V$MYBL/VMYB.05	(+)	0.928	gagaacagggAACGggccacc
V$CARE/CARF.01	(-)	0.926	acagaGAGGct
V$INSM/INSM1.01	(-)	0.925	tgtgtGGGGgcca
V$MYOD/MYOD.01	(-)	0.925	agctGGCAggtggcttg
V$EBOX/USF.04	(+)	0.924	cgggcCACCtggccact
V$KLFS/KLF12.01	(-)	0.922	gctcccaGTGGgtgggaag
V$SORY/SOX12.01	(-)	0.92	ggtgaACAAtctagctggcaggt
V$FKHD/HFH2.01	(+)	0.919	cagaaacaAACAagcca
V$SORY/SOX13.01	(-)	0.918	ggtgaACAAtctagctggcaggt
V$IKRS/IK3.01	(-)	0.917	ctctgGGAAtccc
V$GLIF/GLIS2.01	(-)	0.913	cgcaCCCCccccccaac
V$HICF/HIC1.01	(+)	0.912	tctTGCCagccag
V$AP2F/TCFAP2E.01	(+)	0.912	attcccagAGGCaag
V$HOMF/BARX1.01	(-)	0.91	aggcagggAATTactcaac
V$AP1R/MARE.02	(+)	0.91	ccatacgctggtttGGCAgcctctc
V$AHRR/NXF_ARNT.01	(+)	0.908	gagcgtgagacCGTGtataccctgg
V$DMTF/DMP1.02	(-)	0.906	ccagcGTATggtcac
V$NFKB/NFKAPPAB50.01	(+)	0.906	gccGGGAttcccaga
V$SORY/SOX9.01	(-)	0.905	ggtgaaCAATctagctggcaggt
V$MYBL/VMYB.02	(+)	0.904	gagaacagggAACGggccacc
V$NGRE/IR1_NGRE.01	(-)	0.904	actctcttGGAGagg
V$EGRF/EGR1.04	(+)	0.902	ggttgggggGGGGtgcgg
V$P53F/TP63.01	(+)	0.901	gcctctctgttcgagCATGtctaaa
V$AP1R/NFE2.01	(-)	0.9	tttgtttCTGActcaccgcaccccc
V$PARF/VBP.01	(-)	0.897	aggccttgtGTAAccag
V$SP1F/SP1.01	(+)	0.896	gttggGGGGggggtgcg
V$SP1F/GC.01	(+)	0.895	ggttggGGGGggggtgc
V$STEM/OCT3_4.01	(-)	0.893	cacaataGCATcacctagg
V$SP1F/SP4.02	(+)	0.893	ctaaaaGGTGtgccggg
V$BCL6/BCL6.04	(+)	0.892	ggaTTCCcagaggcaag
V$E2FF/E2F7.02	(-)	0.885	agtgggtGGGAaggcag
V$DLXF/DLX4.01	(+)	0.884	cgggttgagtAATTccctg
V$AP1R/BACH1.01	(+)	0.882	gggtgcggTGAGtcagaaacaaaca

V$NEUR/NEUROD1.01	(+)	0.88	aagcCACCtgccagc
V$MYBL/VMYB.03	(+)	0.879	gagaacagggAACGggccacc
V$SP1F/TIEG.01	(-)	0.879	agcaGGTGtgtgtgggg
V$HNF1/HMBOX.01	(+)	0.878	tggtgctgGTTAcacaa
V$FKHD/FKHRL1.01	(-)	0.876	gagaggtgAACAatcta
V$RREB/RREB1.01	(+)	0.873	cCCCAcacacacctg
V$ZF04/ZID.01	(-)	0.872	ggGCTCcagaagc
V$HESF/HES1.02	(-)	0.871	ctggCAGGtggcttg
V$PARF/DBP.01	(+)	0.868	ctaggTGATgctattgt
V$PLAG/PLAGL1.02	(+)	0.867	ggggGGGGtgcggtgagtcagaa
V$AP1R/NFE2L2.01	(+)	0.867	gggtgcggtGAGTcagaaacaaaca
V$HAND/HAND2_E12.01	(-)	0.865	tagtggccaggTGGCccgttc
V$SNAI/SCRT2.01	(+)	0.862	gcaaCAGGgca
V$PARF/TEF_HLF.01	(-)	0.862	gggaaTTACtcaacccg
V$AP2F/TCFAP2B.01	(-)	0.862	cttGCCTctgggaat
V$HOXF/HOXA3.01	(+)	0.861	ttgagTAATtccctgcctg
V$HOXF/HOXB6.01	(-)	0.861	gcagggAATTactcaaccc
V$HOMF/TLX1.01	(+)	0.859	gggggtgCGGTgagtcaga
V$PARF/TEF.01	(-)	0.859	aggccttgtGTAAccag
V$HOXF/HOXC8.01	(-)	0.856	gcagggaATTActcaaccc
V$STAT/STAT5A.01	(+)	0.855	gggaTTCCcagaggcaaga
V$CEBP/CEBPG.01	(+)	0.855	cgggTTGAgtaattc
V$RXRF/VDR_RXR.01	(+)	0.85	ggttgggagtgaGAGGgctgcagga
V$FKHD/FREAC2.01	(-)	0.849	gagaggTGAAcaatcta
V$HBOX/MEOX1.01	(+)	0.849	gttgagtAATTccctgcct
V$FKHD/XFD3.01	(-)	0.845	gagaggtgAACAatcta
V$ZF02/ZNF202.01	(-)	0.844	actcacCGCAcccccccccaac
V$HOXF/HOXA5.01	(-)	0.843	gcagggAATTactcaaccc
V$CREB/E4BP4.01	(-)	0.841	caggccttgtGTAAccagcac
V$TALE/TGIF2LX.01	(-)	0.84	aacacaccaCTCAtcat
V$FKHD/FOXK1.01	(-)	0.838	gagaggtgAACAatcta
V$HOXF/HOXA6.01	(-)	0.835	gcagggAATTactcaaccc
V$ZTRE/ZTRE.01	(+)	0.835	cacccactGGGAgcgtg
V$HOXF/HOXB8.01	(-)	0.834	gcagggaATTActcaaccc
V$BHLH/BHLHA15.01	(+)	0.834	tgaccaTACGctg
V$EGRF/EGR2.01	(-)	0.827	gtgtGTGTgggggccaggg
V$FAST/FAST1.01	(-)	0.821	gtggttgtgTTTTcgaa
V$AP1R/VMAF.01	(-)	0.821	gcctgggcactGTGAgcaacatctc
V$CTCF/CTCFL.01	(-)	0.821	tctgggctccagaagcaGGCAgggaat
V$GCF2/LRRFIP1.01	(+)	0.82	tatACCCtggcccccacac
V$RXRF/RAR_RXR.01	(+)	0.818	agccaggtcctAGGTgatgctattg
V$PTF1/PTF1.01	(+)	0.812	cacaCCTGctgtgtgaccata
V$NR2F/TR4.02	(+)	0.812	ccagccAGGTcctaggtgatgctat
V$CTCF/CTCF.04	(-)	0.812	tctgggctccagaagcAGGCagggaat
V$FKHD/FREAC4.01	(-)	0.812	cttgtgtaACCAgcacc
V$PAX5/PAX5.03	(-)	0.806	cctggGCACtgtgagcaacatctcatgga
V$NRSF/NRSF.02	(-)	0.804	cttgtgtaaccaGCACcatggcctgggcact
V$EGRF/NGFIC.01	(-)	0.801	gtgtGTGTgggggccaggg
V$GLIF/ZIC3.02	(+)	0.799	cttccCACCcactggga
V$CEBP/CEBPE.01	(+)	0.791	cgggttgaGTAAttc
V$PERO/PPAR_RXR.01	(+)	0.788	gccagccaggtcctAGGTgatgc
V$SNAP/PSE.01	(+)	0.788	cTCTCcaagagagtgaagg
V$NR2F/TR2_TR4.01	(+)	0.786	gagaacagggaacgGGCCacctggc
V$ESRR/ESRRA.04	(+)	0.784	tgctggttacacAAGGcctgaac

V$DMRT/DMRT1.01	(+)	0.779	cctgccagctagaTTGTtcac
V$AP1R/MARE_ARE.01	(+)	0.764	acacctgctgtGTGAccatacgctg
V$MYBL/MYBL2.01	(+)	0.751	tcacagtgaGAACagggaacg

Table A10. Complete table of TFBSs of *Homo sapiens* Cebpa Upstream CNS1 region

Matrix	matrix sim.	Sequence
V$MOKF/MOK2.02	1	gaggaaaatggggCCTTtttt
V$TALE/TGIF.01	1	taactgcctGTCAgtgg
V$NEUR/NGN_NEUROD.01	0.987	aatcCATCtgtgttc
V$IKRS/IK2.01	0.983	ctctGGGAatggg
V$ZFXY/ZFX.01	0.983	agGGCCtggag
V$SMAD/SMAD.01	0.976	cttGTCTgtcc
V$SORY/SOX6.01	0.974	tccctACAAaggttcttcttgct
V$AP2F/TCFAP2A.02	0.972	agcccCCTGggggcca
V$NFAT/NFAT.01	0.971	tagggaGGAAaatggggcc
V$HOXF/NANOG.01	0.97	ctctgggAATGgcagaac
V$NEUR/NEUROD1.02	0.969	aatcCATCtgtgttc
V$SORY/SOX3.01	0.968	tccctaCAAAggttcttcttgct
V$E2FF/E2F4.01	0.966	gggtgGCGGggtagcag
V$MYBL/CMYB.02	0.964	tggggggccaTAACtgcctgtc
V$NEUR/ATOH1.01	0.964	aatcCATCtgtgttc
V$TALE/MEIS1.01	0.96	taactgccTGTCagtgg
V$CSEN/DREAM.01	0.956	ctGTCAgtggc
V$SPZ1/SPZ1.01	0.955	aGGAGggtggc
V$ARID/ARID5A.01	0.952	aacaaATATttgaggctctcg
V$STAT/STAT3.02	0.951	tcagTTCCaggagggtggc
V$SORY/SOX2.01	0.944	tccctACAAaggttcttcttgct
V$KLFS/EKLF.01	0.937	cagttccaggaGGGTggcg
V$EBOX/ATF6.01	0.936	cccgCCACcctcctgga
V$ETSF/ERG.02	0.933	tgtagggaGGAAaatggggcc
V$IKRS/IK1.01	0.933	ctctGGGAatggg
V$TALE/PKNOX2.01	0.933	taactgcctGTCAgtgg
V$KLFS/KLF12.01	0.916	cctgtcaGTGGcctctttg
V$BHLH/MESP1_2.01	0.913	aacaCAGAtggat
V$NOLF/OLF1.02	0.904	ttatggCCCCcaggggcttggct
V$TALE/MRG2.01	0.895	taactgcctGTCAgtgg
V$NOLF/EBF1.01	0.893	ttatggCCCCcaggggcttggct
V$BCL6/BCL6.04	0.887	ccaTTCCcagagctagg
V$HAND/TAL1BETAE47.01	0.885	acgaacaCAGAtggatttgac
V$HNF6/HNF6.02	0.884	cagtcaaatCCATctgt
V$HAND/TAL1ALPHAE47.01	0.88	acgaacaCAGAtggatttgac
V$TALE/TGIF1.01	0.879	taactgcctGTCAgtgg
V$HAND/TAL1BETAITF2.01	0.877	acgaacaCAGAtggatttgac
V$GREF/GRE.01	0.869	cttgtctgtcctGTTCtgc
V$HAND/LYL1_E12.01	0.862	acgaacacaGATGgatttgac
V$AP2F/TCFAP2B.01	0.861	ggcCCCCaggggctt
V$HAND/TAL1BETAHEB.01	0.86	acgaacaCAGAtggatttgac
V$ETSF/ETS2.01	0.858	caaacgcAGGAaccttgtctg
V$STAT/STAT5A.01	0.858	cccaTTCCcagagctaggg
V$GREF/PRE.01	0.857	cttgtctgtccTGTTctgc
V$HOXC/HOX_PBX.01	0.843	acacAGATggatttgac
V$TALE/PKNOX1.01	0.839	taactgcctGTCAgtgg
V$MYBL/MYBL1.02	0.838	tgggggccatAACTgcctgtc
V$GCF2/LRRFIP1.01	0.833	gctACCCcgccaccctcct
V$ZF01/ZBRK1.01	0.831	gcggggtaGCAGttttctgcagtca
V$NEUR/NEUROD1.01	0.83	aatcCATCtgtgttc
V$SNAP/PSE.02	0.83	gttacCCTAaaggcactga
V$FAST/FAST1.01	0.823	gcgtttgtgGAGTtggc
V$GREF/ARE.01	0.819	cttgtctgtccTGTTctgc

V$TALE/TGIF2.01	0.819	taactgcctGTCAgtgg
V$PLAG/PLAGL1.02	0.816	tggcGGGGtagcagttttctgca
V$CLOX/CLOX.01	0.815	gtcaaATCCatctgtgttcgttc
V$CTCF/CTCF.01	0.801	tagccaagccccctgGGGGccataactg
V$ESRR/ESRRA.04	0.79	aacaggacagacAAGGttcctgc
V$STAT/STAT1.01	0.779	tcagttccaGGAGggtggc
V$STAF/STAF.01	0.776	ctctttggagttaCCCTaaaggcactgagca
V$SORY/SOX9.03	0.774	gaAAAAaggccccatttttcctcc
V$ZF13/ZNF435.01	0.756	ttTGTTgaacgaacacaga
V$RXRF/CAR_RXR.01	0.752	ggaatGGGCagaacaggacagacaa
V$ZTRE/ZTRE.02	0.752	caAGCCcctgggggcca

Table A11. Complete table of TFBSs of *Homo sapiens* Cebpa Upstream CNS2 region

Matrix	strand	matrix sim.	Sequence
V$PLAG/PLAG1.02	(-)	1	aaGGGGgaacccttatctcccag
V$TAIP/CSRNP1.01	(-)	1	AGAGtgg
V$TAIP/CSRNP3.01	(-)	1	aGAGTgg
V$ZF02/ZKSCAN3.01	(+)	1	ctaggCCCCaccctccctcaccc
V$GATA/GATA4.01	(+)	0.998	gggaGATAagggt
V$KLFS/BKLF.01	(-)	0.996	agggaGGGTggggcctaga
V$MZF1/MZF1.02	(-)	0.995	agGGGGaaccc
V$NEUR/MASH1.01	(+)	0.994	ccttcacCTGCagtg
V$SF1F/SF1.01	(-)	0.993	tggcCAAGgccactg
V$GATA/GATA1.01	(+)	0.992	gggaGATAagggt
V$SREB/SREBP.04	(+)	0.991	cccTCACcccactct
V$ZTRE/ZTRE.03	(+)	0.988	ccCTCCctcaccccact
V$KLFS/KLF2.01	(-)	0.986	agggaGGGTggggcctaga
V$KLFS/IKLF.01	(-)	0.984	agggaGGGTggggcctaga
V$IKRS/LYF1.01	(+)	0.983	gccTGGGagataa
V$GATA/GATA6.02	(+)	0.982	gggaGATAagggt
V$SORY/SOX6.01	(-)	0.982	tggtaACAAagaacgagatggtt
V$KLFS/BKLF.02	(-)	0.979	agggagGGTGgggcctaga
V$EGRF/WT1.02	(-)	0.978	gggtgagGGAGggtggggc
V$SORY/SOX4.01	(-)	0.97	tggtaACAAagaacgagatggtt
V$SPZ1/SPZ1.01	(-)	0.97	aGGAGggaatg
V$NKXH/NKX32.01	(-)	0.966	aatggtaAGTGcaaaggtg
V$GATA/GATA2.01	(-)	0.965	cctaGATAaggct
V$SF1F/FTF.02	(-)	0.963	tggccAAGGccactg
V$NF1F/NFIB.01	(-)	0.962	ggctgggctgatgGCCAaggc
V$NKXH/NKX31.04	(-)	0.958	aatggtaAGTGcaaaggtg
V$SAL2/SALL2.01	(-)	0.957	gaggGTGGggc
V$HOXF/NANOG.01	(-)	0.957	aggagggAATGgtaagtgc
V$KLFS/EKLF.02	(-)	0.957	agggagGGTGgggcctaga
V$NF1F/NF1.04	(-)	0.954	ggctgggctgatggCCAAggc
V$EGRF/CKROX.01	(-)	0.953	gggtgaGGGAgggtggggc
V$SORY/SOX11.01	(-)	0.952	tggtaACAAagaacgagatggtt
V$GATA/GATA3.01	(+)	0.945	gggaGATAagggt
V$XBBF/RFX4.02	(+)	0.945	tctcgttctttGTTAccag
V$RORA/RORA1.01	(+)	0.945	aaatctggaGGTCaaaagcctggga
V$SORY/SOX2.01	(-)	0.944	tggtaACAAagaacgagatggtt
V$GATA/GATA.01	(+)	0.939	gggaGATAagggt
V$HAND/PARAXIS.01	(+)	0.935	accagAACAcatgcactggca
V$ESRR/ESRRA.03	(-)	0.932	gggctgatggccAAGGccactgc
V$MITF/MIT.01	(-)	0.932	cagtgCATGtgttct
V$EGRF/EGR2.02	(-)	0.931	gggtgAGGGagggtggggc
V$SP1F/GC.01	(-)	0.926	agggagGGTGgggccta
V$XBBF/RFX3.02	(+)	0.923	tctcgttctttGTTAccag
V$ZF02/ZBTB7.03	(+)	0.918	ctaggCCCCaccctccctcaccc
V$MAZF/MAZR.01	(-)	0.915	gagggtGGGGcct
V$ZF02/ZNF219.01	(+)	0.913	gggttCCCCcttcacctgcagtg
V$NDPK/NM23.01	(-)	0.91	ggAGGGtggggcctaga
V$ESRR/ERR.01	(-)	0.91	gggctgatggccAAGGccactgc
V$BHLH/MESP1_2.01	(+)	0.91	agaaCACAtgcac
V$DMTF/DMP1.02	(+)	0.906	tctgcGGATccagcc
V$LEFF/LEF1.01	(-)	0.906	ctggtaaCAAAgaacga
V$AP1R/MARE.02	(-)	0.905	gataaggctggatcCGCAgacaggg
V$TEAF/TEAD.01	(+)	0.902	tacCATTccctcc

V$LTSM/LTSM.01	(+)	0.896	gatccagcctTATCt
V$NKXH/BAPX1.01	(-)	0.891	aatggtAAGTgcaaaggtg
V$STAT/STAT.01	(-)	0.89	tgtgttctgGTAAcaaaga
V$GATA/GATA5.01	(-)	0.89	cctaGATAaggct
V$EGRF/EGR1.04	(-)	0.878	gtgagggagGGTGgggcct
V$NEUR/NEUROD1.01	(+)	0.875	tggcCATCagcccag
V$CP2F/TCFCP2L1.01	(-)	0.873	gcCTAGataaggctggatc
V$ESRR/ESRRB.01	(-)	0.871	gggctgatggccaAGGCcactgc
V$GCMF/GCM1.01	(+)	0.866	accctCCCTcacccc
V$P53F/TP63.02	(+)	0.865	gaacacatgcactggCAGGctctcc
V$RP58/RP58.01	(+)	0.864	ctaaAATCtggag
V$P53F/P53.08	(-)	0.859	gagagcctgccagtgCATGtgttct
V$ZF13/ZNF435.01	(+)	0.858	ttTGTTaccagaacacatg
V$RXRF/VDR_RXR.01	(-)	0.851	agtggggtgaggGAGGgtggggcct
V$NR2F/HNF4.01	(-)	0.846	atggtaagtgCAAAggtggagagcc
V$PLAG/PLAGL1.02	(-)	0.843	gaagGGGGaacccttatctccca
V$OVOL/OVOL1.01	(-)	0.84	ggaaccCTTAtctcc
V$HDBP/HDBP1_2.01	(+)	0.84	ctccagCCGGccaaccatc
V$RXRF/VDR_RXR.05	(-)	0.838	agtGGGGtgagggagggtggggcct
V$NR2F/COUP.01	(-)	0.827	atggtaagtgcaaAGGTggagagcc
V$PAX5/PAX5.01	(-)	0.825	ggtaagTGCAaaggtggagagcctgccag
V$NR2F/HNF4A.04	(-)	0.823	atggtaagtgCAAAGgtggagagcc
V$SRFF/SRF.05	(-)	0.82	gggcctagatAAGGctgga
V$CREB/TAXCREB.01	(-)	0.817	gtggggTGAGggagggtgggg
V$RREB/RREB1.01	(+)	0.811	cCCCAccctccctca
V$SP1F/SP2.01	(-)	0.809	agacagggctGGGCtga
V$ZF02/ZNF202.01	(+)	0.799	gccccaCCCTccctcaccccact
V$HNF1/TCF2.01	(+)	0.794	cgttctttGTTAccaga
V$NR2F/NR2F6.01	(+)	0.792	atctggaggtcaAAAGcctgggaga
O$VTBP/ATATA.01	(-)	0.78	gcctagaTAAGgctgga
V$NR2F/TR4.02	(+)	0.767	atctggAGGTcaaaagcctgggaga
V$PRDM/PRDM4.01	(+)	0.752	agcctTATCtaggccccaccctccctcac

Table A12. Complete table of TFBSs of *Homo sapiens* Cebpa Upstream CNS3 region

Matrix	strand	matrix sim.	Sequence
V$WHNF/WHN.01	(-)	0.986	ttgACGCtttt
V$ZTRE/ZTRE.04	(-)	0.983	acgGGAGggagtaaaaa
V$FKHD/HFH3.01	(-)	0.981	gttaaatAAATatcccg
V$FKHD/FREAC7.01	(-)	0.976	agctcaTAAAcaagata
V$FKHD/HFH8.01	(-)	0.961	agctcatAAACaagata
V$MEF2/MEF2A.01	(+)	0.96	cgtcaagctaaAAATatcttgtt
V$BCDF/CRX.01	(-)	0.958	cgcttTAATcaagttcc
V$HAND/TH1E47.01	(-)	0.952	gagtaaaaaCCAGatacgtta
V$HOMF/HHEX.01	(-)	0.952	gtcccagtgatTAATagcc
V$FKHD/HNF3B.01	(-)	0.95	gttaaataAATAtcccg
V$HOXC/HOXA9.01	(-)	0.948	ccagtGATTaatagccc
V$MEF2/MEF2.07	(+)	0.948	cgtcaagctaaAAATatcttgtt
V$EVI1/EVI1.06	(-)	0.947	taaacaAGATattttta
V$MEF3/SIX2.02	(-)	0.943	taaTCAAgttcct
V$FKHD/FREAC2.01	(-)	0.939	agctcaTAAAcaagata
V$GATA/GATA.01	(-)	0.939	accaGATAcgtta
V$NKX6/NKX61.01	(+)	0.938	gctaTTAAtcactgg
V$KLFS/BKLF.02	(-)	0.933	tggacgGGAGggagtaaaa
V$LHXF/LMX1A.02	(-)	0.933	actgcgctTTAAtcaagttccta
V$ABDB/HOXC13.02	(-)	0.932	agagctcaTAAAcaaga
V$ZICF/ZIC2.02	(+)	0.929	tggctCAGCaagcag
V$BRNF/BRN3.03	(+)	0.927	gggctatTAATcactggga
V$HOXC/HOXB4.02	(-)	0.926	ccagTGATtaatagccc
V$LTFM/LACTOFERRIN.01	(+)	0.926	gGAACttga
V$SORY/HMGA.01	(+)	0.925	ttgggctattAATCactgggact
V$BTBF/KAISO.01	(+)	0.924	tttaCTGCgatcgtc
V$HOXC/MEIS1.03	(-)	0.922	ccagtGATTaatagccc
V$CART/PHOX2.01	(-)	0.918	cgcttTAATcaagttcctatt
V$MYBL/CMYB.01	(-)	0.918	ttaatagccCAACtgcttcgt
V$EGRF/CKROX.01	(-)	0.916	ttggacGGGAgggagtaaa
V$FKHD/FOXJ3.01	(-)	0.916	agctcaTAAAcaagata
V$NKX6/NKX63.01	(+)	0.916	gctatTAATcactgg
V$ZICF/ZIC3.01	(+)	0.916	tggctCAGCaagcag
V$FKHD/FOXJ1.01	(-)	0.914	agctcataAACAagata
V$FKHD/XFD2.01	(-)	0.91	agctcaTAAAcaagata
V$ZICF/ZIC1.01	(+)	0.907	tggctCAGCaagcag
V$ABDB/HOXA13.01	(-)	0.906	agagctcaTAAAcaaga
V$NR2F/ARP1.01	(-)	0.906	gttttcggtccagaGCTCataaaca
V$FKHD/FOXP2.01	(-)	0.903	agctcataAACAagata
V$FKHD/FREAC3.01	(-)	0.899	tcgcaGTAAagaagcag
V$FKHD/FOXO1.01	(-)	0.895	agctcataAACAagata
V$NKX6/NKX61.03	(+)	0.895	gctatTAATcactgg
V$GRHL/GRHL2.01	(-)	0.893	aaAACCagatacg
V$MYBL/MYBL1.02	(-)	0.892	ttaatagcccAACTgcttcgt
V$ABDB/HOXB9.01	(-)	0.891	agagctcaTAAAcaaga
V$FKHD/FKHRL1.01	(-)	0.891	agctcataAACAagata
V$CREB/CREB.02	(-)	0.89	tttttagctTGACgcttttgg
V$ABDB/HOXC9.01	(-)	0.888	cccagtgaTTAAtagcc
V$HOXC/HOX_PBX.01	(-)	0.886	ccagTGATtaatagccc
V$MYBL/VMYB.04	(+)	0.882	atatttatttAACGtatctgg
V$MEF2/RSRFC4.01	(+)	0.881	cgtcaagctaaaaATATcttgtt
V$NR2F/HNF4.03	(-)	0.879	ttcatctgggCAACgcacactgcgc
V$HOXF/HOXB8.01	(+)	0.878	ttgggctATTAatcactgg

V$IRFF/IRF7.01	(+)	0.877	actatgttGAATaggaacttgatta
V$PARF/TEF.01	(+)	0.877	gatatttatTTAAcgta
V$PARF/VBP.01	(+)	0.869	gatatttatTTAAcgta
V$FKHD/HFH1.01	(-)	0.868	agctcaTAAAcaagata
V$OCT1/OCT1.03	(+)	0.868	ggaacttgATTAaag
V$NR2F/COUP.01	(-)	0.866	gttttcggtccagAGCTcataaaca
V$ABDB/HOXC10.01	(-)	0.862	agagctcaTAAAcaaga
V$AP1R/BACH1.02	(-)	0.861	ctgcttgCTGAgccaaatgcaggac
V$CDXF/CDX2.02	(+)	0.861	tatcttgtTTATgagctct
V$SRFF/SRF.04	(-)	0.861	aaaaccagaTACGttaaat
V$LHXF/LHX6.01	(-)	0.859	actgcgcttTAATcaagttccta
V$AP1R/MAFF.01	(-)	0.858	ctgcttGCTGagccaaatgcaggac
V$CREB/E4BP4.01	(+)	0.855	ggatatttatTTAAcgtatct
V$LEFF/LEF1.04	(+)	0.855	gggctaTTAAtcactgg
V$HBOX/VAX2.01	(-)	0.852	tcccagtgATTAatagccc
V$HOXF/HOXB7.01	(+)	0.852	gctattAATCactgggact
V$HNF6/HNF6.01	(-)	0.851	acgttaaaTAAAtatcc
V$CART/CART1.02	(-)	0.85	tacgttAAATaaatatcccga
V$E2FF/E2F.02	(-)	0.849	ttgctgagcCAAAtgca
V$AP1R/NFE2L2.01	(-)	0.848	ctgcttgctGAGCcaaatgcaggac
V$BRN5/BRN5.04	(-)	0.848	gtccagagctcATAAacaagata
V$CART/OTP.01	(-)	0.848	tacgttAAATaaatatcccga
V$LTSM/LTSM.03	(-)	0.848	ATCAagttcctattc
V$RXRF/RXRA.01	(-)	0.847	tttcggtccagaGCTCataaacaag
V$HOXF/HOXC4.01	(-)	0.845	cgctttAATCaagttccta
V$HNF1/HNF1.04	(-)	0.844	ccagatacGTTAaataa
V$CART/ALX4.01	(-)	0.843	cgctttAATCaagttcctatt
V$CART/PROP1.01	(-)	0.843	tacgttAAATaaatatcccga
V$CDXF/CDX1.02	(+)	0.843	tatcttgTTTAtgagctct
V$ABDB/HOXA11.01	(-)	0.842	agagctcaTAAAcaaga
V$CART/RHOX6.01	(+)	0.842	gctatTAATcactgggactat
V$CART/UNCX4.01	(-)	0.842	tacgttAAATaaatatcccga
V$FKHD/FOXK1.01	(-)	0.842	agctcataAACAagata
V$LHXF/LMX1B.02	(+)	0.842	cgggatatttATTTaacgtatct
V$CALM/CAMTA.01	(-)	0.841	cacacTGCG
V$FKHD/FHXB.01	(+)	0.841	aagctaAAAAtatcttg
V$LEFF/TCF7L2.01	(+)	0.841	acttgatTAAAgcgcag
V$GLIF/GLIS1.01	(-)	0.84	ccgagcttcCACGacga
V$LHXF/LHX5.01	(-)	0.84	cagatacgttAAATaaatatccc
V$ABDB/HOXC11.01	(-)	0.839	agagctcaTAAAcaaga
V$HOMF/HMX1.01	(-)	0.839	cagtgatTAATagcccaac
V$HNF1/TCF2.01	(+)	0.838	gatatttaTTTAacgta
V$HZIP/HOMEZ.01	(-)	0.837	actgcTTCGttttcg
V$OVOL/OVOL1.01	(-)	0.836	agatacGTTAaataa
V$CART/PHOX2B.01	(-)	0.834	tacgttAAATaaatatcccga
V$PTF1/PTF1.01	(-)	0.834	cccaACTGcttcgttttcggt
V$AP1R/NFE2.02	(-)	0.833	ctgcttgctgaGCCAaatgcaggac
V$LHXF/LHX3.03	(-)	0.833	cagatacgttAAATaaatatccc
V$LHXF/ISL1.01	(-)	0.832	cccagtgatTAATagcccaactg
V$NEUR/NEUROD1.01	(-)	0.831	agcccAACTgcttcg
V$FKHD/XFD3.01	(-)	0.827	tcgcagtaAAGAagcag
V$LHXF/LHX1.01	(+)	0.827	cgggatatttATTTaacgtatct
V$GRHL/GRHL3.01	(+)	0.826	gtatctGGTTttt
V$XBBF/MIF1.01	(+)	0.825	tatgttgaataGGAActtg
V$PAX5/PAX5.03	(-)	0.824	catagTCCCagtgattaatagcccaactg

V$XBBF/RFX3.03	(-)	0.824	caagTTCCtattcaacata
V$AP1R/MARE.03	(-)	0.823	ctgcttGCTGagccaaatgcaggac
V$HOXF/HOXD8.01	(+)	0.823	ttgggctATTAatcactgg
V$ABDB/HOXD11.01	(-)	0.822	agagctcaTAAAcaaga
V$AP1R/MAFK.02	(+)	0.822	cctgcatttggctCAGCaagcagct
V$PAX7/PAX7.01	(+)	0.819	ggaacttGATTaaag
V$ABDB/HOXC12.01	(-)	0.818	agagctcaTAAAcaaga
V$DICE/DICE.01	(+)	0.816	tgagCTCTggaccga
V$HOXF/HOXA4.01	(-)	0.816	tacgttAAATaaatatccc
V$PAX2/PAX2.01	(+)	0.815	tgtgcgttgcccagatgAAACtg
V$ABDB/HOXD12.01	(-)	0.814	agagctcaTAAAcaaga
V$ZF12/ZFP652.01	(+)	0.812	ccaaaagcgTCAAgc
V$HOXC/PBX_HOXA9.01	(-)	0.81	ccagTGATtaatagccc
V$PARF/TEF_HLF.01	(+)	0.803	gatatTTATttaacgta
V$BRNF/BRN4.02	(+)	0.801	gggctaTTAAtcactggga
V$DMRT/DMRT1.01	(+)	0.794	aatcactgggactATGTtgaa
V$EVI1/MEL1.01	(-)	0.791	taaacAAGAtattttta
V$IRFF/IRF4.02	(+)	0.788	tctggaccGAAAacgaagcagttgg
V$HOXC/PBX1.01	(-)	0.785	ccagtGATTaatagccc
V$NR2F/HNF4A.04	(-)	0.785	gttttcggtccAGAGctcataaaca
V$E2FF/E2F4_DP2.01	(+)	0.782	caaaaGCGTcaagctaa
V$HNF1/HNF1.02	(+)	0.781	ggaTATTtatttaacgt
V$IRFF/IRF8.01	(+)	0.781	tctggaccGAAAacgaagcagttgg
V$XBBF/RFX4.03	(+)	0.781	tatGTTGaataggaacttg
V$PBXC/PBX1_MEIS1.02	(-)	0.774	ccagTGATtaatagccc
V$PDX1/PDX1.01	(+)	0.773	ttgggctatTAATcactgg
V$ZBED/ZBED1.01	(+)	0.767	acatCGGGatatt
V$GZF1/GZF1.01	(-)	0.766	TGCGctttaatca

Table A13. Complete table of TFBSs of *Homo sapiens* Cebpa Downstream CNS1 region

Matrix	strand	matrix sim.	Sequence
V$PLAG/PLAG1.02	(+)	1	gaGGGGgtggggcgaccctcctg
V$TAIP/CSRNP1.01	(-)	1	AGAGtgg
V$TAIP/CSRNP3.01	(-)	1	aGAGTgg
V$ZF02/ZKSCAN3.01	(-)	1	ggtcgCCCCacccccctcccgtcc
V$ZF02/ZNF300.01	(+)	1	cctttccCCCCagaacaggtcct
V$MZF1/MZF1.02	(-)	0.998	gaGGGGaagga
V$ZFXY/ZFX.01	(+)	0.993	gaGGCCtgggg
V$KLFS/IKLF.01	(+)	0.991	ggaggGGGTggggcgaccc
V$KLFS/KLF2.01	(+)	0.991	ggaggGGGTggggcgaccc
V$SMAD/SMAD3.01	(+)	0.991	tgtGTCTggcc
V$PURA/PURALPHA.01	(+)	0.989	ggAGGGggtgggg
V$KLFS/BKLF.01	(+)	0.988	ggaggGGGTggggcgaccc
V$SMAD/SMAD.01	(+)	0.987	tgtGTCTggcc
V$KLFS/KLF6.01	(+)	0.986	cgggagGGGGtggggcgac
V$KLFS/GKLF.03	(+)	0.983	cgggaggggGTGGggcgac
V$ZTRE/ZTRE.04	(+)	0.983	acgGGAGggggtggggc
V$MIZ1/MIZ1.01	(+)	0.981	cctgcCCTCac
V$E2FF/E2F4.01	(-)	0.978	acgcgGCGGggctctgc
V$ETSF/GABPA.02	(-)	0.978	ggggggccGGAAggggtgcca
V$CIZF/NMP4.01	(-)	0.976	agAAAAactgg
V$E2FF/E2F1.01	(-)	0.974	acgcGGCGgggctctgc
V$EGRF/WT1.01	(-)	0.974	cagagTGGGgggccggaag
V$ETSF/ELK1.03	(-)	0.973	ggggggccGGAAggggtgcca
V$AP2F/TCFAP2C.02	(+)	0.972	gaggCCTGgggcagg
V$ETSF/FLI1.03	(-)	0.972	ggggggccGGAAggggtgcca
V$ZF02/ZBP89.01	(-)	0.971	tcgccccaCCCCctcccgtccta
V$BCDF/PTX1.01	(+)	0.97	cttgCTAAgccccagca
V$HESF/HELT.01	(-)	0.97	ggggCACGcggcggg
V$EBOX/ATF6.01	(+)	0.969	cccgCCACcctgggaac
V$ETSF/ELF1.01	(-)	0.968	ggggggccGGAAggggtgcca
V$ETSF/ELK4.02	(-)	0.968	ggggggccGGAAggggtgcca
V$EBOX/CMYC.02	(+)	0.967	cccgccgCGTGccccga
V$ETSF/SPI1.02	(-)	0.967	agaggaggGGAAggagtgagg
V$NF1F/NF1.03	(-)	0.967	ggccggaaggggtGCCAgcta
V$AP2F/TCFAP2A.02	(+)	0.966	gaggCCTGgggcagg
V$HBOX/GSH2.01	(+)	0.966	cagtgcTAATtccagtttt
V$ZF07/ZNF263.01	(+)	0.965	cttccccCTCCtctct
V$SAL2/SALL2.01	(+)	0.962	ggggGTGGggc
V$ETSF/ETV3.02	(-)	0.961	ggggggccGGAAggggtgcca
V$MYRF/MYRF.01	(-)	0.96	gggtgCCAGctaa
V$EBOX/MYCMAX.03	(+)	0.958	gaggccaCGAGcccagt
V$SPZ1/SPZ1.01	(-)	0.958	aGGAGggtcgc
V$ZF02/ZBTB7.03	(-)	0.957	ggtcgCCCCacccccctcccgtcc
V$HOMF/MSX2.01	(+)	0.956	tcagtgCTAAttccagttt
V$BEDF/ZBED4.02	(+)	0.955	gacgggaGGGGgtgg
V$AHRR/AHRARNT.03	(+)	0.954	ttacctgcctGCGTggttagctggc
V$SMAD/SMAD4.01	(+)	0.954	tgtGTCTggcc
V$NACA/NACA1.01	(-)	0.95	aggaCAGAgtggg
V$ZFHX/AREB6.01	(+)	0.95	gctttACCTgcct
V$DLXF/DLX3.02	(+)	0.949	gttcagtgcTAATtccagt
V$EGRF/EGR1.04	(+)	0.949	acgggagggGGTGgggcga
V$HESF/HES1.02	(-)	0.949	ggggCACGcggcggg
V$ETSF/NRF2.01	(-)	0.948	ggggggccGGAAggggtgcca

V$EGRF/EGR2.02	(-)	0.946	cagagTGGGggccggaag
V$EBOX/NMYC.02	(+)	0.945	cccgccgCGTGccccga
V$KLFS/EKLF.02	(+)	0.943	ggagggGGTGgggcgaccc
V$AP2F/TCFAP2E.01	(+)	0.942	gaggcctgGGGCagg
V$ETSF/ETS1.01	(-)	0.942	ggggggccGGAAgggtgcca
V$ETSF/GABP.01	(-)	0.942	ggggggccGGAAgggtgcca
V$SP1F/GC.01	(+)	0.942	ggagggGGTGgggcgac
V$MAZF/MAZ.01	(-)	0.94	agagGAGGggaag
V$ZF02/ZNF219.01	(-)	0.939	ccccaCCCCctcccgtcctagga
V$GLIF/GLIS3.01	(+)	0.934	ccggCCCCccactctgt
V$MAZF/MAZR.01	(+)	0.934	gggggtGGGGcga
V$NDPK/NM23.01	(+)	0.934	gaCGGGaggggggtgggg
V$P53F/P53.03	(+)	0.934	ccccgaggcctggggCAGGtcgcag
V$EGRF/CKROX.01	(+)	0.932	taggacGGGAgggggtggg
V$HIFF/HIF1.02	(+)	0.931	cctgcctgCGTGgttag
V$KLFS/KKLF.01	(+)	0.93	ggaggggggtGGGGcgaccc
V$THAP/THAP1.01	(-)	0.93	agtgagGGCAg
V$EBOX/MAX.03	(+)	0.927	gaggcCACGagcccagt
V$AP2F/AP2.02	(-)	0.925	cagGCCTcgggcac
V$ETSF/CETS1P54.01	(-)	0.923	ggggggCCGGaaggggtgcca
V$ETSF/ELF2.01	(-)	0.922	ggggggccGGAAgggtgcca
V$XBBF/RFX1.02	(+)	0.922	ccgccaccctgGGAAcaga
V$KLFS/KLF7.01	(+)	0.921	ggagggGGTGgggcgaccc
V$AP2F/TCFAP2B.01	(-)	0.917	cagGCCTcggggcac
V$ETSF/SPIB.01	(-)	0.917	gaaaaactGGAAttagcactg
V$ETSF/ELK3.01	(-)	0.916	ggggggcCGGAaggggtgcca
V$BCDF/PCE1.01	(+)	0.909	agtgcTAATtccagttt
V$BRNF/TST1.01	(-)	0.909	ctggaATTAgcactgaact
V$GLIF/GLIS2.01	(+)	0.908	ccggCCCCccactctgt
V$AHRR/NXF_ARNT.01	(-)	0.907	ccaagggcggcCGTGggtgaggagg
V$ETSF/ETV4.01	(-)	0.902	ggggggcCGGAaggggtgcca
V$P53F/TP63.02	(+)	0.902	tggggcaggtcgcagCTTGctaagc
V$E2FF/E2F7.02	(-)	0.899	ttctgggGGGAaagggg
V$RBPF/RBPJK.01	(+)	0.899	accccTGGGaacag
V$SP1F/SP1.01	(+)	0.899	ggaggGGGTggggcgac
V$EBOX/USF.03	(+)	0.896	cccgccgCGTGccccga
V$PAX6/PAX6.04	(+)	0.896	cttCCCCtcctctctgctg
V$XBBF/RFX2.02	(+)	0.892	ccgccaccctggGAACaga
V$ETSF/ETV5.01	(-)	0.886	ggggggcCGGAaggggtgcca
V$SP1F/SP4.01	(+)	0.886	ggagggGGTGgggcgac
V$ZICF/ZIC3.01	(+)	0.886	ggtcgCAGCttgcta
V$ETSF/EHF.01	(-)	0.882	ggggggcCGGAagggggtgcca
V$ETSF/ETV1.01	(-)	0.881	ggggggcCGGAaggggtgcca
V$ETSF/ERG.01	(-)	0.88	ggggggcCGGAagggggtgcca
V$CTCF/CTCF.01	(+)	0.879	gtcctaggacgggaGGGGgtggggcga
V$GCMF/GCM1.01	(+)	0.876	tcctgCCCTcactcc
V$EVI1/EVI1.06	(-)	0.875	caaacaAGAAaaactgg
V$GREF/PRE.01	(-)	0.874	gtcctaggaccTGTTctgg
V$ETSF/ELF3.01	(-)	0.87	gaaaaactGGAAttagcactg
V$E2FF/E2F3.02	(+)	0.869	aggcgGCGCctgccccg
V$SREB/SREBP.02	(-)	0.868	gggTCGCcccacccc
V$GLIF/GLI3.01	(+)	0.866	ccggCCCCccactctgt
V$NRF1/NRF1.02	(-)	0.86	ggcggGGCAggcgccgc
V$HDBP/HDBP1_2.01	(+)	0.858	ccccttCCGGccccccact
V$RREB/RREB1.01	(-)	0.854	cCCCAcccccctcccg

V$GREF/GRE.01	(-)	0.853	gtcctaggacctGTTCtgg
V$ZF15/ZSCAN10.01	(+)	0.852	ggcggcgccTGCCccgc
V$ETSF/ELF4.01	(-)	0.851	ggggggccGGAAggggtgcca
V$ETSF/SPIC.01	(-)	0.851	gaaaaactGGAAttagcactg
V$SRFF/SRF.02	(+)	0.847	cccccCAGAacaggtccta
V$PDX1/IPF1.01	(-)	0.846	gaaaaactgGAATtagcac
V$HAML/AML3.01	(+)	0.844	ctgcGTGGttagctg
V$HOXF/HOXA5.01	(-)	0.844	aactggAATTagcactgaa
V$ETSF/ETV6.01	(-)	0.839	gggggggccGGAAggggtgcca
V$HOXF/HOXA2.01	(-)	0.833	aactgGAATtagcactgaa
V$NR2F/HNF4.03	(-)	0.832	ggcggccagaCACAggacagagtgg
V$RXRF/RAR_RXR.03	(-)	0.831	gaggcGGCCagacacaggacagagt
V$XBBF/RFX3.01	(-)	0.831	gctggggcttagCAAGctg
V$XBBF/RFX4.01	(-)	0.83	gctggggcttagCAAGctg
V$CHRE/CHREBP_MLX.01	(-)	0.829	CACGcggcggggctctg
V$GCF2/LRRFIP1.01	(+)	0.829	agaGCCCcgccgcgtgccc
V$NOLF/OLF1.01	(+)	0.828	cccсttTCCCcccagaacaggtc
V$PAX5/PAX5.01	(-)	0.828	tctgttCCCAgggtggcggggcaggcgcc
V$RXRF/THRA.01	(-)	0.827	gcactgaactcagaGGGTttgctgg
V$SP1F/SP2.01	(-)	0.827	gcacgcggcgGGGCtct
V$PRDM/PRDM5.01	(-)	0.826	tgggggggaaAGGGggaagccagggatcag
V$NR2F/ARP1.01	(-)	0.824	agggggaagccaggGATCagccccc
V$ETSF/GABPB1.01	(-)	0.819	gggggggcCGGAagggggtgcca
V$E2FF/E2F7.01	(-)	0.817	ttctGGGGggaaagggg
V$GLIF/GLI1.02	(+)	0.817	ccggCCCCccactctgt
V$OCT1/OCT1.06	(+)	0.817	tcagtgctAATTcca
V$AHRR/AHR.01	(-)	0.816	tcccagggtgGCGGggcaggcgccg
V$EGRF/NGFIC.01	(-)	0.816	gacaGAGTgggggccgga
V$HNF1/TCF2.01	(+)	0.802	cctgcgtgGTTAgctgg
V$RORA/RORA.02	(+)	0.798	gtgccccgaGGCCtggggcaggtcg
V$PTF1/PTF1.01	(-)	0.797	gcaaGCTGcgacctgccccag
V$CREB/CREB3L2.01	(-)	0.789	ccagggtGGCGgggcaggcgc
V$GLIF/ZIC3.02	(+)	0.786	ccggcCCCCcactctgt
V$NR2F/TR2_TR4.01	(-)	0.786	agggggaagccaggGATCagccccc
V$RXRF/THRB.02	(+)	0.782	actctgtcctgtgtctGGCCgcctc
V$STAF/ZNF76_143.01	(+)	0.781	cagcttgctaagcCCCAgcaaaccctctgag
V$NR2F/TR4.02	(+)	0.77	cagaacAGGTcctaggacgggaggg
V$NR2F/NR2F6.01	(+)	0.767	cagaacaggtccTAGGacgggaggg
V$HUB1/ZNF282.01	(+)	0.761	TTTCccccccagaaca
V$ZF12/ZNF652.02	(-)	0.758	gggaaaGGGGgaagc

Table A14. Complete table of TFBSs of *Homo sapiens* Cebpa Downstream CNS2 region

Matrix	strand	matrix sim.	Sequence
V$CREB/CREB2.01	(+)	1	agtgttTGACgtaagctgagg
V$FKHD/FOXP1_ES.01	(-)	1	aaataaaAACAaaccga
V$ZF02/ZKSCAN3.01	(+)	1	tctggCCCCacccactaggaaca
V$HOMF/HHEX.01	(-)	0.998	atctggctttTAATtgct
V$BRAC/TBX5.01	(-)	0.996	gctcatcgcaGGTGtcagggccaaacccc
V$CSEN/DREAM.01	(-)	0.996	aaGTCAggggt
V$AP2F/TCFAP2C.02	(-)	0.995	cctgCCTGgggctcc
V$CART/S8.01	(-)	0.995	cttttTAATtgctctgggatg
V$HOMF/MSX.01	(-)	0.995	ggcttttTAATtgctctgg
V$MYOD/MYOGENIN.03	(+)	0.995	ctcagaCAGCtgcagcc
V$DLXF/DLX3.01	(-)	0.992	ctggcttttTAATtgctct
V$HOMF/NOBOX.01	(-)	0.992	ggcttttTAATtgctctgg
V$KLFS/KLF2.01	(-)	0.991	ctagtGGGTggggccagaa
V$EVI1/MEL1.02	(+)	0.99	cacctgcGATGagccac
V$MYOD/TCF12.01	(-)	0.99	aaggagcaGGTGggaga
V$MOKF/MOK2.02	(-)	0.988	aggcctgtgggtgCCTTggag
V$SORY/SOX6.01	(-)	0.987	gctaaACAAagtcaggggtttct
V$KLFS/GKLF.03	(-)	0.985	gagaggggcGTGGgggagg
V$NEUR/MASH1.01	(+)	0.985	cagacagCTGCagcc
V$NKXH/NKX25.05	(-)	0.985	ccccaTGAGtggccccttc
V$SMAD/SMAD3.02	(-)	0.985	tttGTCTggct
V$IKRS/IK2.01	(+)	0.983	ctctGGGAaaacg
V$ZTRE/ZTRE.04	(+)	0.983	aggGGAGggaagctggg
V$CREB/ATF2.01	(+)	0.982	agtgttTGACgtaagctgagg
V$NEUR/NGN_NEUROD.01	(-)	0.982	attcCATCtggcttt
V$ZFHX/ZEB1.01	(+)	0.982	ctcccACCTgctc
V$HIFF/HRE.02	(+)	0.981	tctgtgaaCGTGgggtt
V$GATA/GATA3.01	(-)	0.98	aggaGATAacagt
V$ZFXY/ZFX.01	(+)	0.98	tgGGCCtggtg
V$NEUR/ASCL1.01	(+)	0.978	gctgcaGCTGcactg
V$PURA/PURALPHA.01	(+)	0.978	ggAGGGgaagggg
V$BRAC/TBX6.01	(-)	0.977	gctcatcgcAGGTgtcagggccaaacccc
V$DMTF/DMP1.02	(-)	0.977	gcagcGGATgtgggc
V$GATA/GATA1.06	(+)	0.977	ggaaGATAccgat
V$NEUR/NEUROD1.02	(-)	0.977	attcCATCtggcttt
V$MYOD/TCF21.02	(+)	0.976	taggtccAGCTgcagct
V$NEUR/ATOH1.01	(-)	0.976	attcCATCtggcttt
V$AP4R/AP4.03	(+)	0.975	aggtccaGCTGcagctg
V$GATA/GATA2.02	(-)	0.975	aggaGATAacagt
V$NEUR/OLIG2.01	(+)	0.974	gctgcaGCTGcactg
V$GATA/GATA4.01	(-)	0.973	aggaGATAacagt
V$FKHD/FREAC7.01	(-)	0.971	aaaataTAAAtaaaaac
V$HICF/HIC1.01	(-)	0.971	ctaTGCCagccaa
V$SAL1/SALL1.01	(-)	0.971	atATAAataaaaa
V$HOMF/HMX3.03	(-)	0.97	ggcttttTAATtgctctgg
V$KLFS/IKLF.01	(-)	0.97	gagctGGGAggggcagaca
V$HOMF/HMX1.01	(-)	0.969	ggcttttTAATtgctctgg
V$MYOD/TCFE2A.03	(+)	0.969	cagctgCAGCtgcactg
V$CREB/CREB2CJUN.01	(+)	0.968	agtgtttgACGTaagctgagg
V$RUSH/SMARCA3.01	(-)	0.968	ccACATttgtc
V$MZF1/MZF1.03	(-)	0.966	gaGGGGacagt
V$ETSF/SPI1.02	(+)	0.964	actcatggGGAAggacaaggg

V$HOMF/BARX2.01	(-)	0.964	ggcttttTAATtgctctgg
V$SMAD/SMAD.01	(-)	0.963	gctGTCTgagt
V$MEF3/SIX2.02	(-)	0.961	gccTCAGcttacg
V$BZIP/CEBPE_ATF4.01	(+)	0.96	tttgacGTAAgct
V$HOMF/HMX2.03	(-)	0.96	ggctttTTAAttgctctgg
V$TALE/MEIS1.01	(+)	0.96	cttagtgcTGTCacgga
V$WHNF/WHN.01	(-)	0.958	aggACGCatct
V$AP2F/TCFAP2A.02	(-)	0.957	cctgCCTGgggctcc
V$LEFF/LEF1.02	(+)	0.957	tcactatCAAAgcatcc
V$SAL2/SALL2.01	(+)	0.957	agggGTGGggc
V$GATA/GATA6.02	(-)	0.956	aggaGATAacagt
V$SIXF/SIX4.01	(-)	0.956	tcgcagGTGTcaggg
V$AHRR/AHRARNT.03	(-)	0.955	gcagagagggGCGTgggggaggggg
V$MYOD/TCFE2A.02	(+)	0.955	aaaagccaGATGgaatc
V$NEUR/ASCL2.01	(+)	0.955	ggtccaGCTGcagct
V$ZF02/ZNF219.01	(+)	0.955	cccctCCCCcacgcccctctctg
V$HEAT/HSF2.02	(+)	0.954	gcattaataacaCGAAcaatctgtg
V$HBOX/GSH2.01	(-)	0.953	gcttttTAATtgctctggg
V$SMAD/GC_SBE.01	(-)	0.953	tggggCTCCtg
V$EBOX/ATF6.01	(+)	0.952	ggggCCACtcatgggga
V$ARID/BRIGHT.01	(+)	0.951	gagcaATTAaaaagccagatg
V$HOMF/BSX.01	(+)	0.951	cccagagcAATTaaaaagc
V$ZF11/ZBTB3.01	(+)	0.951	cccaGGCAgga
V$AP2F/TCFAP2E.01	(-)	0.949	cctgcctgGGGCtcc
V$HOXF/NANOG.01	(-)	0.949	tctgagtAATGggttccac
V$KLFS/EKLF.02	(+)	0.949	ggcaggGGTGggctgtga
V$BEDF/ZBED4.02	(-)	0.948	ggggcgtGGGGgagg
V$HOXF/HOXB8.01	(+)	0.948	cagagcaATTAaaaagcca
V$HAND/TH1E47.01	(+)	0.947	cctctgctgCCAGatgcgtcc
V$MYRF/MYRF.01	(-)	0.946	gggtgCCAGaacg
V$SF1F/FTF.01	(-)	0.946	cttgCAAGggctgga
V$SORY/SOX3.01	(+)	0.946	gccagaCAAAtgtggaacccatt
O$VTBP/VTATA.01	(-)	0.945	aaataTAAAtaaaaaca
V$CREB/ATF.02	(+)	0.944	agtgttTGACgtaagctgagg
V$SORY/SOX30.01	(+)	0.944	cacgaACAAtctgtgttggccgg
V$RBPF/RBPJK.02	(+)	0.943	ctccTGGGaatct
V$CREB/ATF1.02	(+)	0.942	agtgttTGACgtaagctgagg
V$FKHD/FOXJ1.01	(-)	0.942	cttggctaAACAaagtc
V$SPZ1/SPZ1.01	(+)	0.94	gGGAGggaagc
V$ZF07/ZNF263.01	(-)	0.938	gcttccCTCCcctca
V$CEBP/CEBP.02	(+)	0.937	tgtttgacGTAAgct
V$KLFS/BKLF.02	(+)	0.936	gtgaggGGAGggaagctgg
V$PARF/VBP.01	(+)	0.936	gtgtttgacGTAAgctg
V$AP1R/MAFA.01	(-)	0.935	gggctggccagggcTGCAgctgtct
V$LEFF/TCF7L2.02	(+)	0.935	tcactatcAAAGcatcc
V$MYOD/E47.01	(-)	0.935	gggctGCAGctgtctga
V$SORY/SOX2.01	(-)	0.935	gctaaACAAagtcaggggtttct
V$YBXF/YB1.01	(-)	0.935	atgagTGGCccct
V$MYBL/VMYB.05	(-)	0.934	tgggtgccagAACGgggctgg
V$SORY/SOX13.01	(+)	0.934	cacgaACAAtctgtgttggccgg
V$AP4R/AP4.01	(+)	0.933	tcagaCAGCtgcagccc
V$LHXF/LMX1A.02	(+)	0.933	aaacggcaTTAAtaacacgaaca
V$PAX3/PAX3.03	(+)	0.933	ccactCATGgggaaggaca
V$AP1F/AP1.02	(-)	0.932	gtctGAGTaatgg

V$IKRS/IK1.01	(+)	0.932	ctctGGGAaaacg
V$SORY/SOX4.01	(-)	0.932	gctaaACAAagtcaggggtttct
V$SORY/SOX5.02	(+)	0.932	cacgaACAAtctgtgttggccgg
V$ABDB/HOXB9.02	(+)	0.931	cagagcaaTTAAaaagc
V$CEBP/CEBPB.02	(+)	0.928	acaaaTGTGgaaccc
V$DLXF/DLX1.01	(-)	0.928	ctggcttttttAATTgctct
V$SORY/SOX15.01	(+)	0.928	cacgaACAAtctgtgttggccgg
V$CREB/CREB1.02	(+)	0.927	agtgtTTGAcgtaagctgagg
V$BTBF/KAISO.01	(-)	0.925	gtttCTGCtatccct
V$BCDF/PCE1.01	(-)	0.924	cttttTAATtgctctgg
V$BRNF/TST1.01	(+)	0.924	gagcaATTAaaaagccaga
V$FKHD/FOXJ3.01	(-)	0.923	cttggcTAAAcaaagtc
V$HASF/HAS.01	(+)	0.923	cccCACGcccc
V$ABDB/HOXD13.01	(-)	0.922	aatataaaTAAAaacaa
V$DLXF/DLX5.01	(-)	0.922	ctggctttttAATTgctct
V$CEBP/CEBPD.01	(+)	0.921	acaaaTGTGgaaccc
V$KLFS/KLF7.01	(-)	0.921	ctagtgGGTGgggccagaa
V$MYOD/MYOD.02	(-)	0.921	acagtgCAGCtgcagct
V$NEUR/NEUROG.01	(+)	0.921	tctCCATttgctgtg
V$NF1F/NF1.04	(+)	0.921	ctggtggctgcagaCCAAgcc
V$SORY/SOX7.01	(+)	0.921	cacgaACAAtctgtgttggccgg
V$FKHD/XFD2.01	(-)	0.92	aaaataTAAAtaaaaac
V$BHLH/MESP1_2.01	(-)	0.919	acagCAAAtggag
V$ZF05/ZFP410.01	(-)	0.919	tgctctgGGATgggg
V$CREB/CREB.03	(+)	0.918	agtgttTGACgtaagctgagg
V$FKHD/FOXL1.01	(-)	0.918	aaaataTAAAtaaaaac
V$KLFS/KKLF.01	(+)	0.918	ccccaggcaGGGGtgggc
V$AP2F/AP2.01	(-)	0.916	cctGCCTggggctcc
V$HOMF/BARHL2.01	(-)	0.916	ggcttttTAATtgctctgg
V$LEFF/TCF7L1.01	(+)	0.915	tcactatCAAAgcatcc
V$LHXF/ISL2.01	(+)	0.915	atcccagagcaATTAaaaagcca
V$GREF/GRE.03	(-)	0.914	ccaacacagatTGTTcgtg
V$SORY/SOX11.01	(-)	0.914	gctaaACAAagtcaggggtttct
V$SORY/SOX12.01	(+)	0.914	cacgaACAAtctgtgttggccgg
V$BRNF/BRN3.03	(+)	0.912	acggcatTAATaacacgaa
V$FKHD/XFD1.01	(-)	0.912	aaaataTAAAtaaaaac
V$HAND/SCX.01	(-)	0.912	ctcatcgcaggTGTCagggcc
V$INSM/INSM1.01	(-)	0.912	tgggaGGGGcaga
V$CAAT/CAAT.01	(+)	0.911	ggggCCACtcatggg
V$HOXC/HOXC9.02	(+)	0.91	gttttaTTTAtatttt
V$SP1F/GC.01	(-)	0.908	gagctgGGAGggcaga
V$BCL6/BCL6.04	(-)	0.907	gttTTCCcagagattcc
V$FKHD/FREAC2.01	(-)	0.907	cttggcTAAAcaaagtc
V$NDPK/NM23.01	(-)	0.906	agTGGGtggggccagaa
V$EBOX/MAX.03	(+)	0.905	tccccCACGcccctctc
V$AHRR/NXF_ARNT.01	(-)	0.904	gcagagaggggCGTGggggaggggg
V$MAZF/MAZ.01	(+)	0.904	acagGAGGggaag
V$MYOD/MYF5.01	(-)	0.903	acagtgCAGCtgcagct
V$HMTB/MTBF.01	(-)	0.902	cggtATCTt
V$P53F/TP63.02	(-)	0.902	cctggcaggttcctgCCTGggaagg
V$NF1F/NFIB.01	(+)	0.901	ctggtggctgcagACCAagcc
V$OVOL/OVOL1.01	(+)	0.9	attactGTTAtctcc
V$GLIF/GLI3.02	(+)	0.899	ggggCCACtcatgggga
V$ETSF/SPIB.01	(+)	0.898	acaaatgtGGAAcccattact

V$HOMF/MSX1.01	(+)	0.898	cccagagcAATTaaaaagc
V$ETSF/GABPA.02	(+)	0.895	acaggaggGGAAggggccact
V$HOMF/MSX3.01	(-)	0.895	ggcttttTAATtgctctgg
V$FKHD/FOXP2.01	(-)	0.894	cttggctaAACAaagtc
V$ZF02/ZBTB7.03	(-)	0.894	cagctTCCCtcccctcacagccc
V$HOMF/BARHL1.01	(-)	0.893	ggcttttTAATtgctctgg
V$LEFF/TCF7.01	(+)	0.893	tcactatCAAAgcatcc
V$BRNF/BRN3.02	(+)	0.891	acggcatTAATaacacgaa
V$E2FF/E2F7.02	(+)	0.89	gaggggaGGGAagctgg
V$DLXF/DLX4.01	(-)	0.889	ctggcttttTAATTgctct
V$HOMF/MSX2.02	(-)	0.889	ggcttttTAATTgctctgg
V$TALE/TGIF1.01	(+)	0.889	cttagtgctGTCAcgga
V$BARB/BARBIE.01	(-)	0.888	aagcAAAGctcttgg
V$EGRF/CKROX.01	(+)	0.885	tgtgagGGGAgggaagctg
V$PLAG/PLAG1.01	(+)	0.884	aaGGGGccactcatggggaagga
V$BRAC/EOMES.02	(-)	0.882	gctcatcgcaGGTGtcagggccaaacccc
V$MAZF/MAZR.01	(-)	0.882	gctggaGGGGaca
V$ABDB/HOXC9.01	(-)	0.881	tcgtgttaTTAAtgccg
V$E4FF/E4F.01	(+)	0.881	ttgACGTaagctg
V$KLFS/GKLF.01	(-)	0.878	aacaaagtcAGGGgtttct
V$CP2F/TCFCP2L1.01	(+)	0.876	ctCCGGgtttggctggcat
V$GCMF/GCM1.01	(+)	0.876	ctgtgCCCAcatccg
V$FKHD/FKHRL1.01	(-)	0.875	cttggctaAACAaagtc
V$GREF/ARE.04	(-)	0.874	ccaacacagattGTTCgtg
V$ETSF/GABP.01	(+)	0.873	acaggaggGGAAggggccact
V$SP1F/SP4.01	(-)	0.873	gagaggGGCGtggggga
V$NKX6/NKX61.03	(+)	0.872	ggcatTAATaacacg
V$ABDB/HOXA9.02	(+)	0.871	cagagcaaTTAAaaagc
V$FKHD/HNF3B.03	(-)	0.871	aaaataTAAAtaaaaac
V$HOXF/HOX1-3.01	(-)	0.871	gttatTAATgccgttttcc
V$NKX6/NKX63.01	(-)	0.871	cttttTAATtgctct
V$SP1F/TIEG.01	(-)	0.869	gagaGGGGcgtggggga
O$VTBP/LTATA.01	(-)	0.868	aaaTATAaataaaaaca
V$HAND/PARAXIS.01	(-)	0.868	ggcacAGCAaatggagagaaa
V$NGRE/IR2_NGRE.01	(-)	0.868	ccCTCCtgtgagacc
V$NKRF/NKRF.01	(-)	0.868	aatTGCTctgg
V$SRFF/SRF.04	(-)	0.867	ggttccacaTTTGtctggc
V$EGRF/NGFIC.01	(-)	0.865	agggGCGTgggggaggggg
V$IKRS/IK3.01	(+)	0.865	tcctgGGAAtctg
V$STAT/STAT5A.01	(-)	0.865	cagaTTCCcaggagatcca
V$HAML/AML3.01	(+)	0.864	acgtGGGGtttggcc
V$PAX6/PAX6.05	(-)	0.864	tccttccCCATgagtggcc
V$CREB/ATF6.02	(+)	0.862	agtgtttGACGtaagctgagg
V$FKHD/FOXK1.01	(-)	0.862	aaaatataAATAaaaac
V$SP1F/SP1.02	(+)	0.861	ggcagGGGTggggctgt
V$YY1F/YY2.02	(+)	0.861	tggaacCCATtactcagacagct
V$NOLF/OLF1.01	(-)	0.86	acagatTCCCaggagatccagga
V$EGRF/EGR3.01	(-)	0.859	agggGCGTgggggaggggg
V$EGRF/EGR2.01	(-)	0.857	agggGCGTgggggaggggg
V$HOXF/HOXC8.01	(+)	0.857	ggcattaATAAcacgaaca
V$CTCF/CTCFL.01	(-)	0.856	cagcagctcctggaggaGGCCtggcag
V$HOXF/HOXC6.01	(+)	0.856	cagagcAATTaaaaagcca
V$PARF/TEF.01	(-)	0.856	tcagcttacGTCAaaca
V$HEAT/HSF1.01	(-)	0.855	gtgggtggggccAGAAggctagaaa

V$AP1F/FOSL1.01	(+)	0.854	cgatgAGCCactg
V$PRDF/PRDM1.02	(-)	0.853	atggagaGAAAggtgtct
V$PARF/HLF.01	(+)	0.852	gtgtttgacGTAAgctg
V$ZF01/SZF1.01	(-)	0.852	cctGGGAaggagcaggtgggagagc
V$HAND/TAL1BETAITF2.01	(+)	0.851	aaaaagcCAGAtggaatctct
V$HBOX/VAX2.01	(+)	0.851	ccagagcaATTAaaaagcc
V$GATA/GATA5.01	(-)	0.85	aggaGATAacagt
V$E2FF/E2F.02	(-)	0.849	tgtcagggcCAAAcccc
V$ATBF/ATBF1.01	(+)	0.848	gtttgttttATTTata
V$RP58/RP58.01	(-)	0.847	gacgCATCtggca
V$CTCF/CTCF.05	(+)	0.844	ttgtttagccaagAGGGgcctccaagg
V$HOXC/HOXB4.02	(-)	0.844	tgtcTGAGtaatgggtt
V$CART/RHOX6.01	(-)	0.843	tcgtgTTATtaatgccgtttt
V$FKHD/FREAC4.01	(-)	0.843	cttggctaAACAaagtc
V$LTSM/LTSM.03	(+)	0.843	ATCCcagagcaatta
V$STAT/STAT6.01	(+)	0.843	ggatCTCCtgggaatctgt
V$CREB/CREB3.01	(+)	0.842	agtgtttgACGTaagctgagg
V$FKHD/HFH1.01	(-)	0.841	cttggcTAAAcaaagtc
V$P53F/P53.05	(+)	0.841	cctcCAAGgcacccacaggccttct
V$CREB/JUNDM2.01	(+)	0.839	agtgtttgACGTaagctgagg
V$LHXF/ISL1.01	(-)	0.839	tcgtgttatTAATgccgttttcc
V$ETSF/ETV5.01	(-)	0.837	caggcagCGGAtgtgggcaca
V$GREF/MRE.01	(-)	0.835	ccaacacagattGTTCgtg
V$HOXF/HOXC5.01	(+)	0.835	cagagcAATTaaaaagcca
V$CREB/E4BP4.01	(+)	0.834	agtgtttgacGTAAgctgagg
V$PLAG/PLAGL1.02	(+)	0.834	gccgGGGGaacccaggagcccca
V$LHXF/LHX9.01	(+)	0.833	atcccagagcAATTaaaaagcca
V$PRDM/PRDM4.01	(+)	0.832	agcccTTGCaaggcccacagtgcaccctg
V$GLIF/GLI1.02	(+)	0.83	ggggCCACtcatgggga
V$NR2F/HPF1.01	(+)	0.829	ttgcaaggcccACAGtgcaccctgg
V$ETSF/ETV1.01	(-)	0.826	caggcagCGGAtgtgggcaca
V$HOXF/HOXD8.01	(-)	0.825	tcgtgttATTAatgccgtt
V$DEAF/NUDR.01	(-)	0.824	gaaTCGGtatcttccctgg
V$EGRF/EGR1.01	(-)	0.824	aggggcgtgGGGGaggggg
V$RORA/REV-ERBA.02	(-)	0.822	gctaaacaaaGTCAggggtttctgc
V$FKHD/XFD3.01	(-)	0.821	cttggctaAACAaagtc
V$LHXF/LHX1.01	(-)	0.821	ctggcttttTAATTgctctggga
V$HAND/LYL1_E12.01	(-)	0.816	gccaacacaGATTgttcgtgt
V$PIT1/PIT1.02	(+)	0.816	cggcaTTAAtaacac
V$EVI1/EVI1.05	(+)	0.814	cagggaaGATAccgatt
V$BRNF/BRN4.02	(+)	0.813	acggcaTTAAtaacacgaa
V$HBOX/EN1.01	(-)	0.813	tctggcttTTTAattgctc
V$RXRF/THRA.01	(+)	0.813	cactgggcctggtgAGGAgtgtctg
V$HAND/HEN1.02	(-)	0.812	agggctgcaGCTGtctgagta
V$MYBL/MYBL1.01	(-)	0.811	tccaggagataACAGtaatat
V$RREB/RREB1.01	(-)	0.806	cCCCAcccctgcctg
V$SP1F/SP2.01	(+)	0.801	ggcagggtgGGGCtgt
V$PTF1/PTF1.01	(+)	0.799	tgcaGCTGcactgtcccctcc
V$E2FF/E2F1_DP2.01	(-)	0.794	aaatGGAGagaaagggt
V$SNAP/PSE.02	(-)	0.79	ctcagCTTAcgtcaaacac
V$SORY/SOX9.03	(-)	0.79	agAGAAagggtgtctatgccagc
V$ZF01/ZBRK1.01	(-)	0.79	ccagggctGCAGctgtctgagtaat
V$STAF/STAF.01	(+)	0.789	agcccttgcaaggCCCAcagtgcaccctgga
V$AIRE/AIRE.01	(-)	0.785	tgTTATtaatgccgt

V$ESRR/ESRRA.04	(-)	0.785	tgtgggccttgcAAGGgctggag
V$HOXH/MEIS1A_HOXA9.01	(-)	0.784	TGGCaggttcctgcc
V$PBXC/PBX1_MEIS1.03	(+)	0.783	cctgtgctTGACaatag
V$STEM/OSNT.01	(+)	0.783	tgcctgtGCTTgacaatag
V$MEF2/MEF2.01	(-)	0.781	aatataaaTAAAaacaaaccgaa
V$STAF/ZNF76_143.01	(+)	0.78	cagctgcactgtcCCCTccagcccttgcaag
V$BRN5/BRN5.03	(-)	0.767	gTAATgggttccacatttgtctg
V$RXRF/THRB.03	(-)	0.766	ccctgggcctcagcttACGTcaaac
V$PRDM/PRDM5.01	(-)	0.765	tggggctccTGGGttcccccggccaacac
V$ZF02/ZNF202.01	(+)	0.762	ccctccCCCAcgcccctctctgc
V$PDX1/PDX1.01	(+)	0.76	aaacggcatTAATaacacg
V$OAZF/ROAZ.01	(+)	0.756	tgGCACccaagagcttt
V$HAND/HAND2_E12.01	(+)	0.752	aaaaagccagaTGGAatctct
V$MYBL/MYBL2.01	(+)	0.752	tcacggataTTACtgttatct
V$STAT/STAT3.01	(-)	0.75	agatTCCCaggagatccag

Table A15. Complete table of TFBSs of *Homo sapiens* Cebpa Downstream CNS3 region

Matrix	strand	matrix sim.	Sequence
V$CART/S8.01	(+)	1	ctgttTAATtggccttagccc
V$FKHD/FOXP1_ES.01	(+)	1	tatgcaaAACAagtagc
V$FKHD/HNF3.01	(-)	1	gtgaagcAAACacacat
V$HOMF/MSX.01	(+)	1	agctgttTAATtggcctta
V$MZF1/MZF1.01	(+)	1	gtGGGGatggg
V$PLAG/PLAG1.02	(+)	1	aaGGGGgcgtcagggctgggtgt
V$SMAD/GC_SBE.01	(+)	1	cttggCTCCag
V$TAIP/CSRNP1.01	(-)	1	AGAGtga
V$TAIP/CSRNP3.01	(-)	1	aGAGTga
V$TALE/TGIF.01	(+)	1	actgagtctGTCAgttt
V$ZF02/ZKSCAN3.01	(-)	1	cccatCCCCacccccttctctg
V$SORY/SOX6.01	(+)	0.998	acaggACAAaggatttatcccgg
V$CEBP/CEBPB.02	(+)	0.996	ggagtTGTGaaagat
V$EVI1/MEL1.02	(+)	0.996	atcccgtGATGaggatg
V$SF1F/SF1.01	(-)	0.996	gaggCAAGgccagga
V$AP1R/MARE.01	(-)	0.995	ccacccagTCAGcaggaggcaaggc
V$HOMF/HHEX.01	(+)	0.994	tggcagctgttTAATtggc
V$SORY/SOX3.01	(+)	0.994	acaggaCAAAggatttatcccgg
V$AP1R/NRL.02	(-)	0.993	ccacccagtcAGCAggaggcaaggc
V$RUSH/SMARCA3.02	(-)	0.993	tagcACTTgtg
V$DLXF/DLX3.01	(-)	0.992	tagtcacggTAATtgctct
V$KLFS/KLF2.01	(+)	0.991	aaaggGGGTggggatgggt
V$NF1F/NF1.03	(-)	0.991	caattaaacagctGCCAggta
V$HAND/TAL1_E2A.02	(-)	0.99	acccagtCAGCaggaggcaag
V$CEBP/CEBPB.01	(+)	0.989	catgttggGCAAtgg
V$HOMF/NOBOX.01	(+)	0.989	agctgttTAATtggcctta
V$CEBP/CEBPE.02	(+)	0.987	catgttggGCAAtgg
V$ZTRE/ZTRE.03	(+)	0.987	ccCTCCcgcagatgcgg
V$HBOX/GSH2.01	(+)	0.986	tcccacTAATttgatttta
V$HOXC/HOXC9.02	(+)	0.986	ccactgaTTTAtaggga
V$BZIP/CEBPE_ATF4.01	(-)	0.985	agtgaaGCAAaca
V$RBPF/RBPJK.02	(-)	0.984	ttagTGGGaaaat
V$ZTRE/ZTRE.04	(-)	0.984	gcgGGAGggcatggcat
V$NEUR/ASCL2.01	(-)	0.983	taaacaGCTGccagg
V$ZFHX/AREB6.04	(+)	0.983	ttcccGTTTctcc
V$CEBP/CEBPD.01	(+)	0.982	ggagtTGTGaaagat
V$HOXC/MEIS1.03	(+)	0.982	ccactGATTtataggga
V$PAX3/PAX3.03	(+)	0.982	catgtCATGgagttgtgaa
V$IKRS/IK2.01	(-)	0.981	cagtGGGAatgtg
V$MEF2/MEF2C.01	(+)	0.981	cagtgttttaaAAATagactcat
V$MZF1/MZF1.03	(+)	0.981	gtGGGGatggg
V$CEBP/CEBPA.01	(+)	0.98	ggagttgtGAAAgat
V$KLFS/IKLF.01	(+)	0.98	aaaggGGGTggggatgggt
V$EVI1/MEL1.03	(+)	0.979	tgatgagGATGgcccat
V$KLFS/BKLF.01	(+)	0.978	aaaggGGGTggggatgggt
V$MYOD/MYOGENIN.03	(-)	0.978	attaaaCAGCtgccagg
V$BCDF/CRX.01	(+)	0.977	gatttTAATctgatatt
V$FKHD/FREAC7.01	(+)	0.977	gaaagaTAAAcagggaa
V$HAML/AML1.02	(+)	0.977	cactGTGGtctgtca
V$SORY/SOX5.01	(+)	0.977	ccggaaCAATctgaaggcactgt
V$HOMF/BSX.01	(+)	0.976	gctagagcAATTaccgtga
V$HOMF/MSX2.01	(+)	0.975	ttcccaCTAAtttgatttt

V$KLFS/BKLF.02	(+)	0.975	aaagggGGTGggatgggt
V$KLFS/KLF6.01	(+)	0.975	agaaagGGGGtggggatgg
V$MIZ1/MIZ1.01	(+)	0.975	catgcCCTCcc
V$NEUR/ASCL1.01	(-)	0.974	taaacaGCTGccagg
V$SORY/SOX2.01	(+)	0.974	acaggACAAaggatttatcccgg
V$GLIF/GLI3.02	(-)	0.972	ggggCCACccagtcagc
V$HAML/AML1.01	(+)	0.972	cactGTGGtctgtca
V$HOXF/NANOG.01	(-)	0.972	tacagtgAATGagtctatt
V$OCT1/OCT1.02	(+)	0.972	tgtATGCaaaacaag
V$CSEN/DREAM.01	(+)	0.97	gcGTCAgggct
V$TALE/MEIS1.01	(-)	0.969	gagtgatcTGTCatgct
V$SP1F/SP4.01	(+)	0.968	caagggGGCGtcagggc
V$HOXC/HOXA9.01	(+)	0.967	ccactGATTtataggga
V$LHXF/LMX1A.02	(+)	0.967	gcagctgtTTAAttggccttagc
V$HESF/DEC2.01	(-)	0.966	tggccaCATGaattt
V$MYOD/MYF6.01	(-)	0.966	attaaACAGctgccagg
V$MYOD/MYOD.02	(+)	0.966	acctggCAGCtgtttaa
V$MEF2/MEF2A.01	(+)	0.965	cagtgttttaaAAATagactcat
V$SAL2/SALL2.01	(+)	0.964	ggggGTGGgga
V$HAML/AML2.01	(+)	0.963	cacTGTGgtctgtca
V$MYOD/TCFE2A.03	(-)	0.963	attaaaCAGCtgccagg
V$PAX5/PAX2.02	(+)	0.963	cgtgactaagcaTGACagatcactcccag
V$ESRR/ESRRA.03	(-)	0.962	aacctgatgctgAAGGtcacggc
V$LTSM/LTSM.03	(+)	0.962	ATCCcggaacaatct
V$NDPK/NM23.01	(+)	0.962	agGGGGtggggatgggt
V$ZF57/ZFP57.01	(-)	0.962	cacTGCCgcacca
V$FKHD/FOXA1.01	(+)	0.961	aatttgcaAACAcgttt
V$KLFS/GKLF.02	(-)	0.961	tttgcataAAAGgaaactg
V$NFKB/CREL.01	(+)	0.961	agcggcctTTCCcgt
V$HOMF/TLX1.01	(-)	0.96	ttagtcaCGGTaattgctc
V$MEF2/SL1.01	(-)	0.959	aatgagtCTATttttaaaacact
V$NBRE/NBRE.01	(-)	0.959	gctgAAGGtcacggc
V$ESRR/ESRRB.01	(-)	0.958	aacctgatgctgaAGGTcacggc
V$GATA/GATA1.03	(+)	0.958	gaaaGATAaacag
V$GATA/GATA4.01	(+)	0.958	gaaaGATAaacag
V$GATA/GATA2.03	(+)	0.957	gaaaGATAaacag
V$CEBP/CEBPA.02	(+)	0.956	ggagttgtGAAAgat
V$SF1F/FTF.02	(-)	0.956	gaggcAAGGccagga
V$SORY/SOX9.02	(+)	0.956	acaggACAAaggatttatcccgg
V$ETSF/PDEF.01	(-)	0.953	tgatgccaGGATggcacggcc
V$HOMF/BARX1.01	(-)	0.953	gtcacggtAATTgctctag
V$MEF3/SIX2.02	(+)	0.953	gctTCAGtttcct
O$VTBP/MTATA.01	(-)	0.952	ccctaTAAAtcagtggg
V$FKHD/HFH8.01	(+)	0.952	gaaagatAAACagggaa
V$SORY/SOX30.01	(+)	0.952	ccggaACAAtctgaaggcactgt
V$RBPF/RBPJK.01	(-)	0.951	ttagTGGGaaaat
V$BCDF/PCE1.01	(+)	0.95	ctgttTAATtggcctta
V$GATA/GATA6.02	(+)	0.95	gaaaGATAaacag
V$MYT1/MYT1L.01	(+)	0.95	ggaaAGCTtatga
V$SORY/HBP1.01	(-)	0.95	gtacagtgAATGagtctattttt
V$NF1F/NF1.04	(-)	0.949	ctcctgggctaaggCCAAtta
V$ZF02/ZBTB7.03	(-)	0.949	cccatCCCCacccccctttctctg
V$CLOX/CDP.02	(-)	0.948	ctaaggcCAATtaaacagctgcc
V$DLXF/DLX5.01	(-)	0.948	tagtcacggtAATTgctct

V$FKHD/FOXP1.01	(-)	0.948	gtgaagcAAACacacat
V$ARID/BRIGHT.01	(-)	0.947	ggccaATTAaacagctgccag
V$EGRF/CKROX.01	(-)	0.946	atctgcGGGAgggcatggc
V$GATA/GATA2.01	(+)	0.946	gaaaGATAaacag
O$VTBP/VTATA.01	(-)	0.945	ccctaTAAAtcagtggg
V$MYBL/VMYB.05	(-)	0.945	ggatgggagaAACGggaaacg
V$OCT1/OCT2.02	(+)	0.945	tttaTGCAaaatgta
V$CAAT/NFY.04	(-)	0.944	aaggCCAAttaaaca
V$AP2F/TCFAP2C.02	(+)	0.942	cttgCCTGgagccag
V$HICF/HIC1.01	(+)	0.942	ccgTGCCatcctg
V$TEAF/TEAD4.01	(+)	0.942	ccacATTCccact
V$AP2F/TCFAP2A.02	(-)	0.941	ttgtCCTGtggcctc
V$DLXF/DLX2.01	(-)	0.941	tagtcacggtAATTgctct
V$ETSF/ERG.02	(-)	0.941	gcataaaaGGAAactgaagct
V$GFI1/GFI1.02	(-)	0.941	ataAATCagtgggaa
V$IKRS/IK1.01	(-)	0.941	cagtGGGAatgtg
V$P53F/P53.02	(+)	0.941	accccccaagaatgcCATGccctcc
V$AP2F/AP2.01	(+)	0.94	cttGCCTggagccag
V$HAML/AML3.02	(+)	0.94	cactGTGGtctgtca
V$KLFS/KLF12.01	(-)	0.94	taaatcaGTGGgaatgtgg
V$PLZF/PLZF.02	(-)	0.94	aagTACAgtgaatga
V$AP4R/AP4.02	(+)	0.938	cctggcAGCTgtttaat
V$GATA/GATA.01	(+)	0.937	gaaaGATAaacag
V$GATA/GATA1.05	(+)	0.937	gaaaGATAaacag
V$RORA/RORA1.01	(-)	0.937	gatgctgaaGGTCacggcggctccc
V$EREF/ESR2.01	(-)	0.936	gaagGTCAcggcggctccc
V$KLFS/BTEB3.01	(+)	0.936	agcatgtcatGGAGttgtg
V$MITF/MIT.01	(+)	0.936	aaattCATGtggcca
V$SORY/SOX5.02	(+)	0.936	ccggaACAAtctgaaggcactgt
V$IRFF/ISRE.02	(-)	0.935	attttgcataaaagGAAActgaagc
V$HAND/SCX.01	(+)	0.934	taaaattcatgTGGCcacatt
V$NFKB/NFKAPPAB65.01	(+)	0.934	agcggcctTTCCcgt
V$BRNF/BRN2.03	(+)	0.932	acacgtttATTTgattcc
V$CART/ISX.01	(-)	0.932	aaatcaAATTagtgggaaaat
V$ETSF/ELK1.02	(+)	0.932	tttatcccGGAAcaatctgaa
V$SORY/SOX15.01	(+)	0.932	ccggaACAAtctgaaggcactgt
V$STEM/OCT3_4.02	(-)	0.931	acatttGCATaaaaggaa
V$PAX5/PAX8.01	(+)	0.93	cgtgactaagcATGAcagatcactcccag
V$SORY/HMGA.01	(-)	0.93	attccctataAATCagtgggaat
V$SORY/SOX7.01	(+)	0.93	ccggaACAAtctgaaggcactgt
V$ABDB/HOXC13.01	(-)	0.928	ggaatcaaaTAAAacgt
V$ARID/JARID2.01	(+)	0.928	tttgatTTTAatctgatattt
V$DLXF/DLX1.01	(+)	0.928	gcagctgtttAATTggcct
V$TALE/MRG2.01	(-)	0.928	gagtgatctGTCAtgct
V$ABDB/HOXD10.01	(-)	0.927	gaatcaaaTAAAacgtg
V$ABDB/HOXD13.01	(-)	0.927	gaatcaaaTAAAacgtg
V$CAAT/CAAT.01	(-)	0.927	aaggCCAAttaaaca
V$ESRR/ESRRA.01	(-)	0.927	aacctgatgctgAAGGtcacggc
V$NKX6/NKX61.01	(+)	0.927	ctgtTTAAttggcct
V$BEDF/ZBED4.02	(+)	0.926	agagaaaGGGGtgg
V$CART/CART1.01	(+)	0.926	cccacTAATttgattttaatc
V$CDXF/CDX2.03	(+)	0.926	ccactgatTTATagggaat
V$CEBP/CEBP.02	(+)	0.926	ggagttgtGAAAgat
V$GATA/GATA1.07	(+)	0.926	gaaaGATAaacag

V$ZICF/ZIC2.02	(+)	0.926	actccCAGCacggtg
V$FKHD/HNF3B.02	(-)	0.925	gtgaagCAAAcacacat
V$HOMF/NOBOX.02	(+)	0.925	agctgtttAATTggcctta
V$SORY/HMGIY.01	(+)	0.925	tccaAATTttcccactaatttga
V$TALE/TGIF2LX.01	(-)	0.925	taaacagctGCCAggta
V$CARE/CARF.01	(-)	0.924	tgaatGAGGca
V$GATA/GATA1.04	(+)	0.924	gaaaGATAaacag
V$MEF2/MEF2.02	(+)	0.924	cagtgtttaaAATAgactcat
V$SORY/SOX13.01	(+)	0.924	ccggaACAAtctgaaggcactgt
V$BARB/BARBIE.01	(+)	0.923	agagAAAGggggtgg
V$NKX6/NKX61.02	(+)	0.923	ctgtTTAAttggcct
V$E2FF/E2F6.01	(-)	0.922	catctGCGGagggcat
V$FKHD/XFD2.01	(+)	0.922	gaaagaTAAAcagggaa
V$GATA/GATA3.01	(+)	0.922	gaaaGATAaacag
V$NOLF/OLF1.02	(-)	0.922	ccctctTCCCatgggccatcctc
V$CREB/ATF.01	(-)	0.921	cagcccTGACgccccttgac
V$HASF/HAS.01	(-)	0.921	tggCACGgccc
V$OCT1/POU3F3.01	(-)	0.921	atttTGCATaaaagg
V$ZF02/ZNF219.01	(-)	0.92	tgacgCCCCcttgacccatcccc
V$IRXF/IRX5.01	(-)	0.919	ccaaCATGtgcac
O$VTBP/VTATA.02	(-)	0.918	ttgcaTAAAaggaaact
V$AP1F/JUNB.01	(-)	0.918	tgcttaGTCAcgg
V$EBOX/USF.04	(+)	0.918	tgagaCACTtggctcca
V$ESRR/ERR.01	(-)	0.918	aacctgatgctgAAGGtcacggc
V$NF1F/NFIB.01	(-)	0.918	ctcctgggctaagGCCAatta
V$NOLF/EBF1.01	(-)	0.918	ccctctTCCCatgggccatcctc
V$ABDB/HOXB13.01	(-)	0.917	gaatcaaaTAAAacgtg
V$AP1R/MARE.03	(-)	0.917	tgtcatGCTTagtcacggtaattgc
V$AP1R/MARE_ARE.01	(-)	0.917	tgtcatgcttaGTCAcggtaattgc
V$SORY/SOX12.01	(+)	0.917	ccggaACAAtctgaaggcactgt
V$CART/MIXL1.01	(+)	0.916	cccacTAATttgattttaatc
V$FKHD/FREAC2.01	(+)	0.916	gaaagaTAAAcagggaa
V$HESF/HELT.01	(-)	0.916	tgtgCACGggaaagg
V$FKHD/HNF3B.04	(+)	0.915	aatttgcaAACAcgttt
V$IRFF/ISGF3G.01	(-)	0.915	atttgcataaaagGAAActgaagc
V$MAZF/MAZR.01	(+)	0.915	gggggtGGGGatg
V$SORY/SOX4.01	(+)	0.915	acaggACAAaggatttatcccgg
V$TALE/MEIS1.02	(-)	0.915	gagtgatctGTCAtgct
V$FKHD/HFH1.01	(+)	0.914	gaaagaTAAAcagggaa
V$HOMF/LBX2.01	(+)	0.914	agctgtttAATTggcctta
V$PERO/PPARG.02	(+)	0.914	gcccaggaggccACAGgacaaag
V$SP1F/GC.01	(+)	0.914	aaagggGGTGggatgg
V$TEAF/TEAD.01	(+)	0.914	ccaCATTcccact
V$AP1R/BACH2.01	(+)	0.913	ggtggggaTGGGtcaaggggcgtc
V$FKHD/FOXP2.01	(+)	0.913	gaaagataAACAgggaa
V$HOMF/MSX2.02	(+)	0.913	agctgtttAATTggcctta
V$IRFF/IRF7.01	(-)	0.913	cataaaagGAAActgaagctcggag
V$MYBL/VMYB.03	(-)	0.913	ggatgggagaAACGggaaacg
V$CAAT/NFY.01	(-)	0.912	aaggCCAAttaaaca
V$CREB/ATF1.02	(-)	0.912	cagcccTGACgccccttgac
V$GLIF/GLI3.01	(-)	0.912	ggggCCACccagtcagc
V$BRNF/BRN2.01	(-)	0.911	tgAATTttaaatatcagat
V$CAAT/NFYA.01	(-)	0.911	aaggCCAAttaaaca
V$GFI1/GFI1B.01	(-)	0.911	ataAATCagtgggaa

V$TALE/PKNOX1.01	(-)	0.911	gagtgatctGTCAtgct
V$BRNF/TST1.01	(-)	0.91	ggccaATTAaacagctgcc
V$EREF/ER.04	(-)	0.91	gaagGTCAcggcggctccc
V$IRFF/ISRE.01	(-)	0.91	attttgcataaaagGAAActgaagc
V$AP1R/TCF11MAFG.01	(-)	0.909	tgtcatgcttaGTCAcggtaattgc
V$FKHD/FOXO1.01	(+)	0.909	gaaagataAACAgggaa
V$MEF2/MEF2.06	(+)	0.909	cagtgttttaaAAATagactcat
V$DLXF/DLX4.01	(-)	0.908	tagtcacggtAATTgctct
V$LTFM/LACTOFERRIN.01	(-)	0.908	aGCACttgt
V$MEF3/SIX.01	(+)	0.908	gcaTCAGgttagg
V$SORY/SOX11.01	(+)	0.908	acaggACAAaggatttatcccgg
V$HOXC/PBX_HOXA9.01	(+)	0.906	ccacTGATttataggga
V$ETSF/ELF5.01	(+)	0.905	gcagatgcGGAAatgtggctg
V$ETSF/SPI1.03	(+)	0.905	gcagatgcGGAAatgtggctg
V$GATA/GATA6.01	(+)	0.905	gaaaGATAaacag
V$MEF2/RSRFC4.01	(+)	0.905	cagtgtttaaaaATAGactcat
V$MYBL/VMYB.01	(-)	0.905	ggatgggagaAACGggaaacg
V$SORY/SOX9.01	(+)	0.905	ccggaaCAATctgaaggcactgt
V$TALE/TGIF1.01	(-)	0.905	gagtgatctGTCAtgct
V$TALE/MRG1.01	(-)	0.904	gagtgatctGTCAtgct
V$NKXH/NKX25.02	(+)	0.903	ctgttTAATtggccttagc
V$PARF/DBP.01	(+)	0.903	tccttTTATgcaaaatg
V$ZICF/ZIC1.01	(+)	0.903	actccCAGCacggtg
V$HMTB/MTBF.01	(+)	0.902	tgatATTTa
V$ZICF/ZIC3.01	(+)	0.902	actccCAGCacggtg
V$AHRR/NXF_ARNT.01	(+)	0.901	aagggagccgcCGTGaccttcagca
V$DLXF/DLX3.03	(-)	0.901	tagtcacggtAATTgctct
V$HIFF/ARNT.01	(-)	0.901	aaataaaaCGTGtttgc
V$AP1R/NFE2.02	(-)	0.9	tgtcatgcttaGTCAcggtaattgc
V$HBOX/GBX1.01	(+)	0.9	tcccactAATTtgatttta
V$HOMF/HMX3.01	(+)	0.899	agcatcacAAGTgctactt
V$ETSF/ELF3.01	(+)	0.898	gcagatgcGGAAatgtggctg
V$HOMF/HMX2.03	(+)	0.898	agctgtTTAAttggcctta
V$HBOX/GSH1.01	(+)	0.897	attttcccacTAATttgat
V$HOXC/HOXB4.02	(+)	0.897	ccacTGATttataggga
V$MEF2/MEF2.05	(+)	0.897	cagtgtttaaaAATAgactcat
V$MYOD/MYOD.01	(+)	0.897	acctGGCAgctgtttaa
V$NFAT/NFAT5.02	(-)	0.897	agtgGGAAaatttggaaca
V$NR2F/HNF4.03	(+)	0.897	cccaggaggcCACAggacaaaggat
V$EBOX/USF.03	(+)	0.896	cccatccCGTGatgagg
V$SP1F/SP1.01	(+)	0.896	aaaggGGGTggggatgg
V$HOMF/BARHL1.01	(+)	0.895	agctgttTAATtggcctta
V$HOMF/MSX3.01	(+)	0.895	agctgttTAATtggcctta
V$PLAG/PLAG1.01	(+)	0.895	gtGGGGatgggtcaagggggcgt
V$CART/PHOX2.01	(+)	0.894	cccacTAATttgattttaatc
V$CREB/CREB.02	(-)	0.894	acccagcccTGACgcccccctt
V$GLIF/GLI2.01	(-)	0.894	ggggCCACccagtcagc
V$OCT1/OCT1.05	(-)	0.894	ggCATGgcattcttg
V$CART/PROP1.02	(+)	0.893	cccactaatttgATTTtaatc
V$HBOX/MEOX1.01	(+)	0.893	ctagagcAATTaccgtgac
V$IRFF/IRF1.01	(-)	0.893	attttgcataaaagGAAActgaagc
V$OCT1/OCT.01	(+)	0.893	tgtATGCaaaacaag
V$GLIF/GLI1.01	(-)	0.892	ggggccaCCCAgtcagc
V$HBOX/GBX2.01	(+)	0.892	tcccactAATTtgatttta

V$HEAT/HSF2.01	(+)	0.892	aggatttatcccgGAACaatctgaa
V$TEAF/TEF1.01	(+)	0.892	ccaCATTcccact
V$HOMF/HMX1.01	(+)	0.891	agctgttTAATtggcctta
V$KLFS/KLF7.02	(+)	0.891	caagggGGCGtcagggctg
V$ETSF/SPIB.01	(-)	0.89	gagaaacgGGAAacgcctcct
V$HOMF/HMX3.03	(+)	0.89	agctgttTAATtggcctta
V$NGRE/IR1_NGRE.01	(+)	0.89	agcctcctGGAGcga
V$SORY/HBP1.02	(-)	0.89	tacagtgAATGagtctattttta
V$GREF/GRE.03	(-)	0.889	tgccttcagatTGTTccgg
V$NKX1/NKX11.01	(-)	0.889	aaatcaAATTagtggga
V$HOMF/BARX2.02	(+)	0.888	gctagagcAATTaccgtga
V$DLXF/DLX1.02	(+)	0.887	ttgctagagcAATTaccgt
V$MEF2/MEF2.07	(+)	0.887	cagtgtttttaaAAATagactcat
V$TALE/PKNOX2.01	(-)	0.886	gagtgatctGTCAtgct
V$AP1R/MAFB.01	(-)	0.885	ggggccacccagTCAGcaggaggca
V$ETSF/ETV6.01	(+)	0.885	gcagatgcGGAAatgtggctg
V$RORA/RORA.01	(-)	0.885	gatgctgaagGTCAcggcggctccc
V$LTSM/LTSM.01	(+)	0.884	catcaaaccaAATCc
V$PLZF/PLZF.01	(-)	0.883	aagTACAgtgaatga
V$AP1R/MAFF.01	(-)	0.882	tgtgatGCTGaatgaggcatatatt
V$HBOX/EN2.01	(+)	0.882	tcccactAATTtgatttta
V$HOXC/HOX_PBX.01	(+)	0.882	ccacTGATttataggga
V$ETSF/ETV3.01	(+)	0.881	gcagatgCGGAaatgtggctg
V$P53F/TP63.02	(-)	0.881	tcagacgggtcctaaCCTGatgctg
V$PBXC/PBX1_MEIS1.02	(+)	0.88	tggcTGATttacagcgg
V$LHXF/LHX4.01	(+)	0.879	ttttcccactAATTtgattttaa
V$OCT1/OCT1.01	(+)	0.879	ttTATGcaaaatgta
V$OCT1/OCT2.01	(+)	0.879	tgtATGCaaaacaag
V$LHXF/LHX2.01	(-)	0.878	attaaaatcaAATTagtgggaaa
V$AP1R/NFE2.01	(-)	0.877	tgtcatgCTTAgtcacggtaattgc
V$LEFF/TCF7.01	(-)	0.877	tttgcatAAAAggaaac
V$BRAC/TBR1.01	(+)	0.875	atgtcatggagTTGTgaaagataaacagg
V$OCT1/POU2F3.01	(+)	0.875	tgtATGCaaaacaag
V$ARID/ARID5A.01	(+)	0.874	atctgATATttaaaattcatg
V$ETSF/GABP.01	(+)	0.874	ggcccatgGGAAgaggccgt
V$HOMF/HMX2.02	(-)	0.874	atcaaataAAACgtgtttg
V$IRXF/IRX4.01	(-)	0.874	ccaaCATGtgcac
V$NR2F/COUP.02	(+)	0.874	tggggatgggTCAAgggggcgtcag
V$EBOX/USF.01	(+)	0.873	aaattCATGtggccaca
V$NF1F/NF1.01	(-)	0.873	gatTTGGttgatgccaggat
V$NKX1/NKX12.01	(-)	0.873	cacggtAATTgctctag
V$SNAI/SCRT2.01	(+)	0.873	gcatCAGGtta
V$ETSF/ELK4.01	(+)	0.872	tttatccCGGAacaatctgaa
V$IRXF/IRX2.01	(+)	0.872	tgcaCATGttggg
V$IRXF/IRX3.01	(-)	0.872	ccaaCATGtgcac
V$HOMF/BARHL2.01	(-)	0.871	gtcacggTAATtgctctag
V$IRFF/IRF3.01	(-)	0.871	ggatgggagaaacgGGAAacgcctc
V$P53F/TP63.01	(+)	0.871	tactcgctgttaggaAATGtgtgtt
V$TALE/TGIF2.01	(-)	0.871	gagtgatctGTCAtgct
V$CAAT/NFY.02	(-)	0.87	aaggCCAAttaaaca
V$BZIP/CEBPE_ATF4.02	(+)	0.869	ttttatGCAAaat
V$GLIF/GLI1.02	(-)	0.868	ggggCCACccagtcagc
V$NF1F/NF1.02	(+)	0.868	atccTGGCatcaaaccaaatc
V$PAXH/PAX4.02	(-)	0.868	aatcaAATTagtggg

V$HOXF/HOXC8.01	(-)	0.867	cacggtaATTGctctagca
V$HOXF/HOXD3.01	(-)	0.867	cacggtAATTgctctagca
V$CAAT/NFY.05	(-)	0.866	aaggCCAAttaaaca
V$AP1R/BACH1.02	(-)	0.865	tgtcatgCTTAgtcacggtaattgc
V$CEBP/CEBPG.01	(-)	0.865	accaTTGCccaacat
V$LEFF/TCF7L1.01	(-)	0.865	ttaaaatCAAAttagtg
V$PRDF/PRDM1.02	(+)	0.865	aaacaggGAAAgcttatga
V$NR2F/TR2_TR4.01	(+)	0.864	cccaggaggccacaGGACaaaggat
O$YTBP/SPT15.01	(-)	0.863	tgaggcaTATAtttcag
V$CART/RAX.01	(-)	0.863	aaatcaAATTagtgggaaaat
V$GATA/GATA3.03	(+)	0.863	gaaaGATAaacag
V$GRHL/GRHL2.01	(-)	0.863	ctAACCtgatgct
V$HOXF/HOXA3.01	(+)	0.863	ctgttTAATtggccttagc
V$IKRS/IK3.01	(+)	0.863	tatagGGAAtaag
V$AP1R/NFE2L2.01	(-)	0.862	tgtcatgctTAGTcacggtaattgc
V$HOXF/HOXA6.01	(-)	0.862	cacggtAATTgctctagca
V$LHXF/ISL1.01	(-)	0.862	tagtcacggTAATtgctctagca
V$AP1R/VMAF.01	(-)	0.861	ggggccacccaGTCAgcaggaggca
V$AP1R/MAFK.02	(+)	0.86	tatatgcctcattCAGCatcacaag
V$CART/PRRX2.01	(-)	0.86	aaatcaAATTagtgggaaaat
V$ETSF/ELK3.01	(+)	0.859	tttatccCGGAacaatctgaa
V$HBOX/VAX2.01	(-)	0.859	taaggccaATTAaacagct
V$HOXF/HOXB6.01	(-)	0.859	cacggtAATTgctctagca
V$CHRE/CHREBP_MLX.01	(-)	0.858	CACGggaaaggccgctg
V$RXRF/RXRA.01	(+)	0.858	ggggtggggatgGGTCaagggggcg
V$IRXF/IRX6.01	(-)	0.857	ccaaCATGtgcac
V$ZF05/ZFP410.01	(-)	0.857	catcacgGGATggga
V$CART/SHOX2.01	(-)	0.856	aaggccAATTaaacagctgcc
V$HBOX/EMX2.01	(+)	0.856	tcccacTAATttgatttta
V$HOXF/HOXA5.01	(-)	0.856	cacggtAATTgctctagca
V$AP1F/FOSL2.01	(-)	0.855	gaatgAGGCatat
V$CART/ESX1.01	(+)	0.855	cccactAATTtgattttaatc
V$SRFF/SRF.02	(-)	0.855	ttttgCATAaaaggaaact
V$STEM/OCT3_4.01	(+)	0.855	tcattcaGCATcacaagtg
V$GCMF/GCM1.02	(+)	0.854	gtggcCCCGcactga
V$NR2F/COUP.01	(+)	0.854	cccaggaggccacAGGAcaaaggat
V$STEM/OSNT.01	(+)	0.854	tcattcaGCATcacaagtg
V$CEBP/CEBPE.01	(+)	0.853	catgttggGCAAtgg
V$NR2F/HNF4.01	(+)	0.853	ggccacaggaCAAAggatttatccc
V$PBXC/PBX1_MEIS1.01	(+)	0.853	tggcTGATttacagcgg
V$HOXF/HOXA3.02	(-)	0.851	aaggccaATTAaacagctg
V$IRFF/IRF2.01	(-)	0.851	attttgcataaaagGAAActgaagc
V$NR2F/TR4.02	(+)	0.851	cccaggAGGCcacaggacaaaggat
V$HBOX/VAX1.01	(-)	0.85	tcacggtAATTgctctagc
V$YY1F/YY1.01	(+)	0.85	gccgtgCCATcctggcatcaaac
V$E2FF/E2F.02	(+)	0.849	tgtctgctcCAAAggga
V$LHXF/LHX8.01	(-)	0.849	attaaaatcaAATTagtgggaaa
V$OCT1/OCT1.04	(+)	0.849	tgTATGcaaaacaag
V$CTCF/CTCF.05	(-)	0.848	ccacccagtcagcAGGAggcaaggcca
V$HOXF/HOXB3.01	(-)	0.847	aaatcAAATtagtgggaaa
V$LEFF/LEF1.04	(-)	0.846	tccctaTAAAtcagtgg
V$LEFF/TCF7L2.01	(-)	0.846	tttgcatAAAAggaaac
V$PBXC/PBX1_MEIS1.03	(+)	0.846	tggctgatTTACagcgg
V$HBOX/HLXB9.01	(+)	0.845	tcccactAATTtgatttta

V$HOXF/HOXB8.01	(+)	0.845	tagagcaATTAccgtgact
V$FKHD/FOXK1.01	(+)	0.844	gaaagataAACAgggaa
V$HBOX/EVX1.01	(+)	0.844	tcccactaATTTgatttta
V$OCT1/OCT1.06	(-)	0.844	tgtttgcaAATTgta
V$ETSF/ETS2.01	(+)	0.843	cgctgttAGGAaatgtgtt
V$CALM/CAMTA.01	(-)	0.842	ctcagTGCG
V$CART/PHOX2A.01	(-)	0.842	aaatcaAATTagtgggaaaat
V$HAML/AML3.01	(+)	0.842	cactGTGGtctgtca
V$EBOX/MNT.01	(-)	0.841	cccaaCATGtgcacggg
V$GRHL/GRHL3.01	(+)	0.841	caaacaCGTTtta
V$HOXF/HOXA2.01	(-)	0.841	cacggTAATtgctctagca
V$P53F/P53.05	(-)	0.841	gaggCAAGgccaggatttggtttga
V$AP1R/BACH1.01	(-)	0.84	atgctgaaTGAGgcatatatttcag
V$DMTF/DMP1.01	(-)	0.839	tgccaGGATggcacg
V$PARF/TEF_HLF.01	(-)	0.839	acacaTTTCctaacagc
V$DMRT/DMRT7.01	(+)	0.838	tgtgTTTGcttcactcttttg
V$RXRF/RAR_RXR.01	(+)	0.838	caggaggccacAGGAcaaaggattt
V$PERO/PPARG.03	(+)	0.835	gcccaggaggccACAGgacaaag
V$CART/VSX1.01	(-)	0.834	aaggccAATTaaacagctgcc
V$CART/XVENT2.01	(-)	0.834	aatttTAAAtatcagattaaa
V$ETSF/SPIC.01	(+)	0.834	gcagatgcGGAAatgtggctg
V$NOLF/OLF1.01	(-)	0.833	ggcggcTCCCtttggagcagaca
V$NR2F/HNF4A.04	(+)	0.833	acatgttgggcAATGgtctcagccg
V$NR2F/NR2F6.01	(+)	0.833	acatgttgggcaATGGtctcagccg
V$ZF01/SZF1.01	(+)	0.833	ataGGGAataagccctacctggcag
V$CART/ALX4.01	(+)	0.831	ctgtttAATTggccttagccc
V$CTCF/CTCF.04	(-)	0.831	ccacccagtcagcaggAGGCaaggcca
V$GCMF/GCM1.03	(-)	0.831	cttatTCCCtataaa
V$RREB/RREB1.01	(-)	0.83	aCCCAtccccacccc
V$CTCF/CTCFL.01	(-)	0.828	ccacccagtcagcaggaGGCAaggcca
V$ETSF/ELK1.01	(-)	0.828	gagaaacgGGAAacgcctcct
V$NR2F/ARP1.01	(+)	0.828	cccaggaggccacaGGACaaaggat
V$PAX7/PAX7.01	(+)	0.828	tttatttGATTcctg
V$PERO/PPAR_RXR.01	(+)	0.828	cacatgttgggcaaTGGTctcag
V$CAAT/NFY.03	(-)	0.825	ccacCCAGtcagcag
V$PAX6/PAX6.05	(-)	0.825	tatttcaGCATgacagacc
V$FAST/FAST1.02	(+)	0.824	ggaaaTGTGtgtttgct
V$HBOX/EVX2.01	(+)	0.824	tcccacTAATttgatttta
O$XCPE/XCPE1.01	(-)	0.823	gtGCGGggcca
V$BRN5/POU6F2.01	(-)	0.823	ttaaatatcagATTAaaatcaaa
V$LEFF/LEF1.03	(-)	0.823	tttgcaTAAAaggaaac
V$SRFF/SRF.05	(-)	0.821	ttttgcataaAAGGaaact
V$DMRT/DMRT3.01	(-)	0.819	tgttttgcaTACAttttgcat
V$HESF/BHLHB2.02	(-)	0.818	tgtgcaCGGGaaagg
V$MYOD/TCF21.01	(+)	0.818	acctggcaGCTGtttaa
V$ZF12/ZFP652.01	(-)	0.816	gcaaaagagTGAAgc
V$HNF6/HNF6.01	(-)	0.815	gattaaaaTCAAattag
V$MYBL/MYBL1.01	(-)	0.815	catgcttagtcACGGtaattg
V$STAT/STAT1.01	(-)	0.811	attgttccgGGATaaatcc
V$RORA/REV-ERBA.02	(-)	0.81	gatgctgaagGTCAcggcggctccc
V$CREB/E4BP4.01	(+)	0.809	ttcctttatGCAAaatgtat
V$HUB1/ZNF282.01	(+)	0.809	TTTCtcccatcccgt
V$PRDM/PRDM5.01	(+)	0.809	ggggtggggATGGtcaagggggcgtcag
V$BRN5/BRN5.01	(-)	0.807	gttttgCATAcattttgcataaa

V$SORY/SOX1.03	(+)	0.807	attCACTgtacttgaaggcaatt
V$CREB/TAXCREB.02	(-)	0.806	ccccctTGACccatcccccacc
V$E2FF/E2F.01	(-)	0.806	aattagtggGAAAattt
V$FKHD/FREAC4.01	(+)	0.806	gaaagataAACAgggaa
V$NGRE/IR2_NGRE.01	(-)	0.805	gcCTCCtggctaag
V$HBOX/EN1.01	(+)	0.804	ggcagctgTTTAattggcc
V$DMRT/DMRT4.01	(-)	0.803	aaaatttggAACAaattgcct
V$HEAT/HSF1.04	(+)	0.802	aggatttatcccggaacaATCTgaa
V$PAX5/PAX5.03	(-)	0.802	aaaggCCGCtgtaaatcagccacatttcc
V$BRNF/BRN3.01	(-)	0.801	gccacatgaATTTtaaata
V$E2FF/E2F4_DP2.01	(-)	0.801	aattaGTGGgaaaattt
V$ETSF/SPI1.05	(+)	0.801	gcagatgcGGAAatgtggctg
V$PAX2/PAX2.01	(-)	0.801	gaaactgaagctcggagAAACtg
V$NR2F/PNR.01	(-)	0.799	acatgaattTTAAatatcagattaa
V$ZF02/ZNF202.01	(-)	0.799	ccttgaCCCAtcccccacccccctt
V$IRFF/IRF8.01	(-)	0.795	cataaaagGAAActgaagctcggag
V$NFAT/NFATC1.01	(-)	0.795	cttcccATGGgccatcctc
O$VTBP/ATATA.01	(-)	0.79	atatattTCAGcatgac
V$BRN5/BRN5.03	(+)	0.79	cTAATttgattttaatctgatat
V$CTCF/CTCF.03	(-)	0.79	ccacccagtcagcaggAGGCaaggcca
V$HOXH/MEIS1B_HOXA9.01	(-)	0.789	TGACagactcagtgc
V$GLIF/ZIC3.02	(+)	0.787	cttgcCTCCtgctgact
V$RXRF/THRB.03	(+)	0.787	taattggccttagcccAGGAggcca
V$ZBED/ZBED1.02	(-)	0.787	aTGTGgccacatg
V$TALE/MRG1.02	(-)	0.786	taaaCAGCtgccaggta
V$HOXC/PBX1.01	(+)	0.785	ccactGATTtataggga
V$SNAP/PSE.02	(-)	0.783	ttttgCATAaaaggaaact
V$RXRF/VDR_RXR.03	(+)	0.782	gggatGGGTcaaggggcgtcaggg
V$DEAF/NUDR.01	(-)	0.779	cctTCAGattgttccggga
V$ZF12/ZNF652.02	(+)	0.776	ggggatGGGTcaagg
V$HOXH/MEIS1A_HOXA9.01	(+)	0.774	TGGCagctgtttaat
V$MOKF/MOK2.01	(+)	0.774	aaccaaatcctggCCTTgcct
V$CART/HESX1.01	(+)	0.773	cccacTAATttgattttaatc
V$PTF1/PTF1.01	(+)	0.772	ggcaGCTGtttaattggcctt
V$STAF/ZNF76_143.01	(+)	0.77	taattggccttagCCCAggaggccacaggac
V$NR2F/TR4.01	(+)	0.767	cccaggaggccacAGGAcaaaggat
V$RORA/RORA.02	(+)	0.763	gtggggatgGGTCaaggggcgtca
V$SORY/SOX9.08	(+)	0.762	gagCAATtaccgtgactaagcat
V$PAX3/PAX3.01	(-)	0.76	tTAGTcacggtaattgctc
V$SORY/SOX21.03	(+)	0.758	tgtGAAAgataaacagggaaagc
V$PAX9/PAX9.02	(+)	0.757	gggcgtCAGGgctgggtgtct
V$SORY/SRY.04	(+)	0.756	attcACTGtacttgaaggcaatt
V$DUXF/DUX4.01	(-)	0.752	cagGATTtggtttga
V$BRAC/MGA.01	(-)	0.751	acatttcctaacagcgagtagACACcgtg
V$PERO/PPARA.01	(+)	0.75	cacatgttgggcAATGgtctcag
V$RORA/REV-ERBA.03	(+)	0.75	gtggggatggGTCAaggggcgtca

Table A16. Complete table of TFBSs of *Homo sapiens* Cebpa Downstream CNS4 region

Matrix	strand	matrix sim.	Sequence
V$NEUR/MASH1.01	(+)	1	gtgccacCTGCatgt
V$IKRS/LYF1.01	(+)	1	gtcTGGGagaggt
V$MZF1/MZF1.01	(+)	1	gtGGGGacgat
V$NKXH/NKX25.01	(+)	1	cccctcAAGTgcaattccc
V$TALE/TGIF.01	(+)	1	ctgaaatgtGTCAgagg
V$ZFXY/ZFX.01	(-)	1	ggGGCCtggcc
V$ZF02/ZKSCAN3.01	(-)	1	atcgtCCCCacctcccctcggct
V$SORY/SOX6.01	(-)	0.999	cttgaACAAaggagcacaggcct
V$NKXH/TTF1.01	(+)	0.997	cccctCAAGtgcaattccc
V$EBOX/MNT.01	(+)	0.996	gcagcCACGtgccacct
V$SORY/SOX3.01	(-)	0.996	cttgaaCAAAggagcacaggcct
V$EBOX/MYCMAX.03	(+)	0.995	gcagccaCGTGccacct
V$NF1F/NF1.03	(+)	0.995	tcttccctgccctGCCAggaa
V$CART/S8.01	(-)	0.995	cttcaTAATtgtggccctcct
V$ZF02/ZNF300.01	(-)	0.994	acccacaCCCCagcctctcctct
V$AP1F/JUNB.01	(+)	0.994	gcatgaGTCAcag
V$AP4R/AP4.03	(-)	0.993	gagcccaGCTGatctgc
V$EBOX/USF.01	(-)	0.993	ggtggCACGtggctgca
V$EBOX/CMYC.01	(+)	0.991	gcagccaCGTGccacct
V$ETSF/ETV1.02	(+)	0.991	caagcacaGGAAgctgttgga
V$TEAF/TEAD4.01	(-)	0.991	tagcATTCctggc
V$HESF/HES1.02	(-)	0.99	gtggCACGtggctgc
V$EBOX/USF.02	(+)	0.99	gcagcCACGtgccacct
V$AP1F/AP1.01	(+)	0.99	gcatgAGTCacag
V$NKXH/NKX25.05	(+)	0.99	catctTGAGtgaggcagtg
V$SORY/SOX4.01	(-)	0.989	cttgaACAAaggagcacaggcct
V$INSM/INSM1.01	(+)	0.988	tgcctGGGGgcgg
V$SORY/SOX2.01	(-)	0.988	cttgaACAAaggagcacaggcct
V$HIFF/ARNT.01	(-)	0.987	aggtggcaCGTGgctgc
V$EBOX/NMYC.01	(+)	0.987	gcagccaCGTGccacct
V$MOKF/MOK2.02	(-)	0.986	ttggtttctagaaCCTTttcc
V$AP1F/FOSL1.01	(+)	0.986	gcatgAGTCacag
V$IKRS/IK2.01	(-)	0.986	acttGGGAattgc
V$RUSH/SMARCA3.01	(-)	0.986	ccCCATatagc
V$EBOX/CMYC.02	(+)	0.985	gcagccaCGTGccacct
V$EBOX/MAX.03	(+)	0.985	gcagcCACGtgccacct
V$HESF/DEC1.02	(+)	0.984	cagccaCGTGccacc
V$ETSF/ETV3.02	(+)	0.984	caagcacaGGAAgctgttgga
V$SORY/SOX11.01	(-)	0.983	cttgaACAAaggagcacaggcct
V$HIFF/HRE.03	(+)	0.981	tgcagccACGTgccacc
V$AP1F/BATF.01	(-)	0.981	ctgtgaCTCAtgc
V$MZF1/MZF1.03	(+)	0.981	gaGGGGaggtg
V$SMAD/SMAD3.02	(+)	0.981	cttGTCTggga
V$AP2F/TCFAP2C.02	(+)	0.981	gctgCCTGggggcgg
V$HIFF/HRE.02	(-)	0.98	aggtggcaCGTGgctgc
V$NKXH/NKX25.02	(-)	0.98	cttcaTAATtgtggccctc
V$MYOD/TCF12.01	(-)	0.98	cacatgcaGGTGgcacg
V$AP1R/BACH2.01	(+)	0.98	tggaggcaTGAGtcacaggccaggc
V$CHOP/CHOP.01	(+)	0.979	aagtGCAAttccc
V$EBOX/NMYC.02	(+)	0.978	gcagccaCGTGccacct
V$XBBF/RFX5.01	(-)	0.978	tgccccatatAGCAacatc
V$BEDF/ZBED4.01	(+)	0.978	tggGGGCggtgcgtg

V$AP1F/FOSL2.01	(+)	0.977	gcatgAGTCacag
V$AP1F/JUND.01	(+)	0.977	gcatgAGTCacag
V$HESF/DEC2.01	(-)	0.974	gtggcaCGTGgctgc
V$HOXC/MEIS1.03	(-)	0.973	ccaatGATTtaaggcct
V$CAAT/CAAT.01	(-)	0.972	gaagCCAAtgattta
V$AP1F/JUNDM2.02	(+)	0.972	gcatgaGTCAcag
V$EBOX/MYCMAX.02	(-)	0.971	ggtggcaCGTGgctgca
V$MYOD/TCF21.02	(-)	0.971	cgagcccAGCTgatctg
V$MAZF/MAZ.01	(+)	0.97	agccGAGGggagg
V$HOMF/BARX2.01	(-)	0.968	ctcttcaTAATtgtggccc
V$EBOX/USF.03	(-)	0.968	ggtggcaCGTGgctgca
V$NACA/NACA1.01	(+)	0.967	aggaCAGAggaga
V$TEAF/TEAD.01	(-)	0.967	tagCATTcctggc
V$HESF/HELT.01	(-)	0.966	gtggCACGtggctgc
V$KLFS/GKLF.02	(-)	0.966	ccttgaacAAAGgagcaca
V$KLFS/KLF6.01	(+)	0.966	tgcctgGGGGcggtgcgtg
V$KLFS/BKLF.01	(+)	0.964	ggctgGGGTgtgggtgact
V$AHRR/AHRARNT.03	(+)	0.963	tgggggcggtGCGTgagccacgctg
V$HIFF/CLOCK_BMAL1.01	(-)	0.962	aggtggCACGtggctgc
V$ETSF/FLI1.02	(+)	0.962	caagcacaGGAAgctgttgga
V$AP1F/AP1.03	(-)	0.962	ctgTGACtcatgc
V$STAT/STAT3.02	(-)	0.961	agcaTTCCtggcagggcag
V$IKRS/IK1.01	(-)	0.961	acttGGGAattgc
V$KLFS/EKLF.02	(+)	0.959	ggctggGGTGtgggtgact
V$HOMF/NOBOX.01	(-)	0.959	ctcttcaTAATtgtggccc
V$SF1F/SF1.01	(+)	0.958	ttgcCAAGgcaaagg
V$ETSF/PEA3.01	(+)	0.956	caagcacAGGAagctgttgga
V$CAAT/NFY.04	(-)	0.955	gaagCCAAtgattta
V$EBOX/USF.04	(-)	0.955	ggtggCACGtggctgca
V$AP1R/BACH2.02	(+)	0.954	tggaggcaTGAGtcacaggccaggc
V$HOXC/HOXC9.02	(-)	0.953	ccaatgaTTTAaggcct
V$EBOX/ATF6.01	(+)	0.953	tgagCCACgctggaggg
V$PAX5/PAX2.02	(+)	0.953	gaggtggaggcaTGAGtcacaggccaggc
V$ZF11/ZBTB3.01	(-)	0.953	cccaGGCAgcg
V$HOMF/HHEX.01	(+)	0.952	gtaattcaagtTAATaagc
V$YY1F/YY1.02	(-)	0.952	gctgccCCATatagcaacatcac
V$HESF/HES1.01	(+)	0.951	cagccacGTGCcacc
V$SP1F/SP4.01	(+)	0.951	cctgggGGCGgtgcgtg
V$SP1F/SP1.03	(+)	0.948	cctggGGGCggtgcgtg
V$AP2F/TCFAP2A.02	(+)	0.948	gctgCCTGggggcgg
V$ETSF/CETS1P54.01	(+)	0.946	caagcaCAGGaagctgttgga
V$HIFF/HIF1.02	(+)	0.946	tgcagccaCGTGccacc
V$AP1F/AP1.02	(+)	0.946	gcatGAGTcacag
V$CLOX/CDP.02	(-)	0.944	aagaagcCAATgatttaaggcct
V$HEAT/HSF1.05	(+)	0.944	ccaacaggaaaaggttCTAGaaacc
V$ZF02/ZBP89.01	(-)	0.944	acgcaccgCCCCcaggcagcgct
V$SP1F/TIEG.01	(+)	0.943	ggctGGGGtgtgggtga
V$ETSF/ERG.02	(+)	0.941	cgccaacaGGAAaaggttcta
V$MITF/MIT.01	(+)	0.941	acctgCATGtgatct
V$HAND/TWIST.01	(-)	0.941	caggtggCACGtggctgcacc
V$RBPF/RBPJK.02	(-)	0.941	aactTGGGaattg
V$MYOD/E47.01	(-)	0.94	cacatGCAGgtggcacg
V$SREB/SREBP.01	(-)	0.94	tgcTCACctcacctc
V$HAND/SCX.01	(-)	0.939	tcacatgcaggTGGCacgtgg

V$ZF02/ZNF219.01	(-)	0.939	cagcaCCCCcacatcgcccctct
V$ETSF/FLI1.03	(+)	0.938	caagcacaGGAAgctgttgga
V$SP1F/GC.01	(+)	0.938	cctgggGGCGgtgcgtg
V$GLIF/ZIC2.01	(-)	0.938	cgcaccgCCCCcaggca
V$ETSF/ETS1.01	(+)	0.937	caagcacaGGAAgctgttgga
V$GFI1/GFI1.02	(+)	0.937	ttaAATCattggctt
V$SP1F/SP1.02	(+)	0.937	cctggGGGCggtgcgtg
V$PLAG/PLAG1.01	(+)	0.936	gaGAGGggcgatgtggggtgct
V$EBOX/MAX.01	(+)	0.935	gcagcCACGtgccacct
V$ETSF/ETS2.01	(+)	0.934	caagcacAGGAagctgttgga
V$LHXF/LMX1A.02	(+)	0.933	aattcaagTTAAtaagcaaactc
V$AP2F/AP2.01	(+)	0.933	gctGCCTggggcgg
V$AP2F/TCFAP2A.01	(+)	0.933	gctgcctgGGGGcgg
V$CAAT/NFY.01	(-)	0.932	gaagCCAAtgattta
V$AP2F/AP2.02	(+)	0.932	gctGCCTggggcgg
V$KLFS/KKLF.01	(+)	0.93	ggagaggctGGGGtgtggg
V$GREF/ARE.02	(-)	0.929	caacagcttcctGTGCttg
V$DLXF/DLX1.01	(-)	0.928	ccctcttcatAATTgtggc
V$DLXF/DLX3.01	(-)	0.928	ccctcttcaTAATtgtggc
V$HIFF/HRE.01	(+)	0.928	tgcagccaCGTGccacc
V$CHRF/CHR.01	(-)	0.927	taacTTGAattac
V$ABDB/HOXB9.01	(+)	0.927	ggaggcctTAAAtcatt
V$CARE/CARF.01	(+)	0.926	tgagtGAGGca
V$ZF07/ZNF263.01	(-)	0.925	ccccacCTCCcctcg
V$ETSF/ELK1.03	(+)	0.924	caagcacaGGAAgctgttgga
V$NF1F/NFIB.01	(-)	0.924	tttcaggcttggaGCCAtcgt
V$EGRF/WT1.01	(+)	0.923	cgatgTGGGggtgctggtg
V$DLXF/DLX2.01	(-)	0.922	ccctcttcatAATTgtggc
V$GCMF/GCM1.02	(-)	0.922	ctgcaCCCGcaccag
V$EBOX/MAX.02	(-)	0.922	ggtggCACGtggctgca
V$SREB/SREBP.04	(-)	0.922	tgcTCACctcacctc
V$HASF/HAS.01	(+)	0.921	agcCACGtgcc
V$ZF07/ZNF263.02	(-)	0.921	ggcctcCTCCtacct
V$CAAT/NFYA.01	(-)	0.92	gaagCCAAtgattta
V$AP1R/BACH1.01	(+)	0.92	tggaggcaTGAGtcacaggccaggc
V$AP2F/TCFAP2B.01	(-)	0.919	accGCCCccaggcag
V$NFKB/NFKAPPAB.01	(-)	0.917	ctGGGAgtcacccac
V$PAX6/PAX4_PD.01	(+)	0.916	gggGCAGcgctgcctgggg
V$ETSF/ELF2.01	(+)	0.915	caagcacaGGAAgctgttgga
V$NKX6/NKX61.01	(+)	0.915	caagTTAAtaagcaa
V$SP1F/SP1.01	(+)	0.915	cctggGGGCggtgcgtg
V$ETSF/ELK1.02	(+)	0.914	caagcacaGGAAgctgttgga
V$NKRF/NKRF.01	(+)	0.914	aacTCCTcagg
V$AP2F/TCFAP2E.01	(+)	0.913	aactcctcAGGCaag
V$HOXC/PBX_HOXA9.01	(-)	0.912	ccaaTGATttaaggcct
V$LTFM/LACTOFERRIN.01	(-)	0.911	tGCACttga
V$GATA/GATA2.02	(-)	0.91	ttgaGATCacatg
V$HOMF/HMX3.01	(-)	0.91	gccttggcAAGTggaatcc
V$EBOX/MYCMAX.01	(-)	0.909	ggtggCACGtggctgca
V$LEFF/LEF1.01	(-)	0.909	ccttgaaCAAAggagca
V$HIFF/ARNTL.01	(-)	0.908	aggtggcaCGTGgctgc
V$BCDF/PCE1.01	(-)	0.907	cttcaTAATtgtggccc
V$PERO/PPARG.02	(+)	0.907	agcagaaaggacAGAGgagaggc
V$PAX6/PAX6.02	(+)	0.906	ggcatgagtcACAGgccag

V$RBPF/RBPJK.01	(-)	0.906	aactTGGGaattg
V$CAAT/NFY.03	(-)	0.905	gaagCCAAtgattta
V$MYBL/CMYB.01	(-)	0.904	agtggaatcCAACagcttcct
V$RXRF/VDR_RXR.01	(+)	0.904	gccaaggcaaagGAGGgccacaatt
V$ZF02/ZBTB7.03	(-)	0.904	atcgtCCCCacctcccctcggct
V$HOXC/HOXA9.01	(-)	0.903	ccaatGATTtaaggcct
V$MTF1/MTF-1.01	(-)	0.903	ggctGCACccgcacc
V$EVI1/EVI1.07	(-)	0.902	cactcAAGAtgctcacc
V$BRNF/BRN4.01	(+)	0.9	taccacgagggTAATtcaa
V$BRN5/BRN5.04	(-)	0.9	ggagtttgcttATTAacttgaat
V$CAAT/NFY.02	(-)	0.899	gaagCCAAtgattta
V$MEF3/MEF3.01	(-)	0.897	ggcTCAGgggcct
V$EBOX/USF2.01	(+)	0.896	gcagccaCGTGccacct
V$CTCF/CTCF.01	(+)	0.896	gggcagcgctgcctGGGGgcggtgcgt
V$HESF/HES7.01	(+)	0.895	cagccaCGTGccacc
V$EGRF/EGR1.02	(+)	0.892	ctgcctggGGGCggtgcgt
V$TEAF/TEF1.01	(-)	0.89	tagCATTcctggc
V$MAZF/MAZR.01	(+)	0.89	ggaggtGGGGacg
V$GFI1/GFI1B.01	(+)	0.889	ttaAATCattggctt
V$HESF/BHLHB2.01	(+)	0.886	cagccACGTgccacc
V$ETSF/GABP.01	(-)	0.886	cagggcagGGAAgagcagggt
V$STAT/STAT1.02	(+)	0.886	gccctgccaGGAAtgctac
V$YBXF/YB1.01	(+)	0.886	atcatTGGCttct
V$ETSF/NRF2.01	(+)	0.885	caagcacaGGAAgctgttgga
V$HAND/PARAXIS.01	(-)	0.885	ttgagATCAcatgcaggtggc
V$GCMF/GCM1.03	(-)	0.885	gctgcCCCAtatagc
V$DLXF/DLX3.03	(-)	0.883	ccctcttcatAATTgtggc
V$HEAT/HSF1.04	(-)	0.883	gggttggttctagaaccTTTTcct
V$DLXF/DLX4.01	(-)	0.882	ccctcttcatAATTgtggc
V$GRHL/GRHL1.01	(-)	0.882	ctccccGGTTgct
V$EGRF/EGR1.04	(+)	0.882	ccgaggggaGGTGgggacg
V$HOXC/HOXB4.02	(-)	0.881	ccaaTGATttaaggcct
V$XBBF/RFXDC2.01	(+)	0.881	tgactcccagagCAACcgg
V$NBRE/NBRE.01	(+)	0.87	cagaAAGGacagagg
V$IKRS/IK3.01	(-)	0.87	acttgGGAAttgc
V$HNF1/HNF1.04	(+)	0.869	taattcaaGTTAataag
V$HOXF/HOXB5.01	(+)	0.868	caagtTAATaagcaaactc
V$MTF1/MTF-1.02	(-)	0.868	ggctgcaCCCGcacc
V$ETSF/ELK3.01	(+)	0.867	caagcacAGGAagctgttgga
V$EVI1/EVI1.06	(-)	0.867	cagacaAGATtgctcct
V$XBBF/RFX4.01	(+)	0.866	tgactcccagagCAACcgg
V$XBBF/RFX3.01	(+)	0.865	tgactcccagagCAACcgg
V$HOXF/HOXA2.01	(+)	0.861	caagtTAATaagcaaactc
V$CP2F/CP2.02	(-)	0.861	cACTGcctcactcaagatg
V$HBOX/VAX1.01	(-)	0.86	gtttgctTATTaacttgaa
V$PERO/PPARG.03	(+)	0.86	ctgaaatgtgtcAGAGgtgaggt
V$HESF/DEC1.01	(-)	0.859	gtggCACGtggctgc
V$HOXF/HOXA3.02	(-)	0.858	tttgcttATTAacttgaat
V$CAAT/NFY.05	(-)	0.858	gaagCCAAtgattta
V$HESF/BHLHB2.02	(-)	0.857	gtggcaCGTGgctgc
V$HOXF/HOXA5.01	(+)	0.856	caagttAATAagcaaactc
V$HEAT/HSF2.03	(+)	0.856	ggaaaaggttctAGAAaccaaccct
V$EVI1/EVI1.02	(-)	0.856	cagacAAGAttgctcct
V$EGRF/EGR1.03	(+)	0.855	cgatgtgGGGGtgctggtg

V$HDBP/HDBP1_2.01	(+)	0.853	ctagcgCCGGggccgccaa
V$HOXF/HOXC8.01	(-)	0.853	aacttgaATTAccctcgtg
V$RXRF/PXR_RXR.01	(+)	0.853	acgagGGTAattcaagttaataagc
V$OCT1/OCT1.03	(-)	0.851	gtttgcttATTAact
V$SREB/SREBP.02	(-)	0.849	tgcTCACctcacctc
V$LEFF/TCF7L2.01	(-)	0.848	ccttgaaCAAAggagca
V$PAX5/PAX5.01	(+)	0.846	tcctcaGGCAaggaaggcaggcagctctt
V$ZF02/ZNF202.01	(-)	0.845	tgcacccCGCAccagcacccccac
V$LEFF/LEF1.03	(-)	0.845	ccttgaACAAaggagca
V$PDX1/IPF1.01	(+)	0.844	aattcaagtTAATaagcaa
V$PLAG/PLAGL1.02	(+)	0.844	tgtgGGGGtgctggtgcgggtgc
V$FKHD/FREAC3.01	(+)	0.842	ttcaaGTTAataagcaa
V$HEAT/HSF1.01	(-)	0.842	gggttggtttctAGAAcctttttcct
V$RREB/RREB1.01	(-)	0.841	cCCCAcatcgcccct
V$EVI1/EVI1.05	(-)	0.841	cagacaaGATTgctcct
V$XBBF/RFX2.02	(+)	0.84	tgactcccagagCAACcgg
V$NR2F/HNF4.03	(+)	0.839	aggcatgagtCACAggccaggcccc
V$RORA/REV-ERBA.02	(+)	0.839	cctgaaatgtGTCAgaggtgaggtg
V$CHRE/CHREBP_MLX.01	(+)	0.838	CACGtgccacctgcatg
V$HOXF/HOXA6.01	(-)	0.838	aacttgAATTaccctcgtg
V$ABDB/HOXC9.01	(-)	0.838	tttgcttaTTAActtga
V$EVI1/MEL1.01	(-)	0.838	cagacAAGAttgctcct
V$HNF1/HNF1.01	(+)	0.833	aGTTAataagcaaactc
V$BRNF/BRN2.04	(+)	0.832	ttcaagtTAATaagcaaac
V$HNF1/HMBOX.01	(+)	0.832	taattcaaGTTAataag
V$NFKB/NFKAPPAB.02	(-)	0.832	ctGGGAgtcacccac
V$E4FF/E4F.01	(+)	0.831	gtgATGTtgctat
V$DICE/DICE.01	(+)	0.83	tggcTTCTtcaccct
V$ETSF/ETV5.01	(+)	0.83	caagcacAGGAagctgttgga
V$NFAT/NFATC1.01	(-)	0.829	tttgccTTGGcaagtggaa
V$RXRF/RAR_RXR.01	(+)	0.828	aaatgtgtcagAGGTgaggtgagca
V$HOXF/HOXA1.01	(+)	0.825	caagttAATAagcaaactc
V$HOXH/MEIS1B_HOXA9.01	(-)	0.825	TGACacatttcaggc
V$EREF/ER.03	(-)	0.823	gggagtcacccaCACCcca
V$RXRF/CAR_RXR.01	(+)	0.821	acgagGGTAattcaagttaataagc
V$HOXF/HOXA4.01	(+)	0.821	caagttAATAagcaaactc
V$CTCF/CTCFL.01	(-)	0.821	acgcaccgccccaggcAGCGctgccc
V$NGRE/IR1_NGRE.01	(+)	0.82	cttgtctgGGAGagg
V$ZTRE/ZTRE.01	(-)	0.82	cgccccaGGCAgcgct
V$HAND/HAND2_E12.01	(-)	0.819	gcacccccacaTCGCccctct
V$PAX9/PAX9.01	(-)	0.817	cgcccccaggcagcGCTGccc
V$BCL6/BCL6.02	(-)	0.816	tggtttcTAGAacctttt
V$HNF1/TCF2.01	(+)	0.816	taattcaaGTTAataag
V$ETSF/FLI1.01	(+)	0.813	caagcaCAGGaagctgttgga
V$BRN5/POU6F2.01	(-)	0.812	tgaagaagccaATGAtttaaggc
V$NR2F/TR4.02	(+)	0.811	gcagaaAGGAcagaggagaggctgg
V$AP1R/NFE2L2.01	(+)	0.811	tggaggcatGAGTcacaggccaggc
V$CREB/TAXCREB.01	(-)	0.811	cgtggcTCACgcaccgccccc
V$ZF01/ZBRK1.01	(-)	0.808	aaaggagcACAGgccttggtgaact
V$AIRE/AIRE.01	(-)	0.803	gcTTATtaacttgaa
V$BCL6/BCL6.01	(-)	0.803	gcaTTCCtggcagggca
V$NR2F/COUPTFII.01	(+)	0.803	tgaaatgtgtcagaGGTGaggtgag
V$EVI1/EVI1.03	(-)	0.802	cagacAAGAttgctcct
O$XCPE/XCPE1.01	(+)	0.801	gtGCGGgtgca

V$RXRF/RXR_RXR.01	(+)	0.8	gcatgagtcacaGGCCaggcccctg
O$MTEN/DMTE.01	(+)	0.797	cttccctAGCGccggggccgc
V$HNF1/HNF1.02	(+)	0.794	agtTAATaagcaaactc
V$NRF1/NRF1.01	(+)	0.788	cctaGCGCcggggccgc
V$CREB/TAXCREB.02	(-)	0.788	cgtggcTCACgcaccgccccc
V$PBXC/PBX1_MEIS1.03	(+)	0.786	ggcatgagTCACaggcc
V$PAX3/PAX3.01	(+)	0.785	tTCTTcaccctgcagatca
V$XBBF/MIF1.01	(+)	0.784	ctcggtgcaggGCAGcaga
V$STAT/STAT1.01	(-)	0.784	agcattcctGGCAgggcag
V$CTCF/CTCF.03	(-)	0.781	tgagatcacatgcaggTGGCacgtggc
V$XBBF/RFX5.02	(-)	0.781	ccTTTGccttggcaagtgg
V$PBXC/PBX1_MEIS1.01	(+)	0.781	ggcaTGAGtcacaggcc
V$ETSF/SPI1.05	(+)	0.78	ttgagtgaGGCAgtgatgttg
V$PBXC/PBX1_MEIS1.02	(+)	0.779	ggcaTGAGtcacaggcc
V$PERO/PPAR_RXR.01	(+)	0.777	agcagaaaggacagAGGAgaggc
V$AIRE/AIRE.02	(-)	0.777	cttggtgaacTTGGg
V$GLIF/ZIC3.02	(-)	0.775	ccgccCCCAggcagcgc
V$HBOX/EN1.01	(+)	0.774	gagggtaaTTCAagttaat
V$HOXH/MEIS1A_HOXA9.01	(-)	0.772	TGACacatttcaggc
V$XBBF/RFX4.03	(+)	0.764	ctcGGTGcagggcagcaga
V$STAF/ZNF76_143.01	(-)	0.761	acatcgcccctctCCCCggttgctctgggag
V$SORY/SOX8.02	(+)	0.761	catGAGTcacaggccaggcccct
V$GMEB/GMEB2.01	(-)	0.759	aGCGTggctcacgca
V$NRSF/NRSF.02	(+)	0.758	cggtgcagggcaGCAGaaaggacagaggaga
V$NRSF/NRSF.01	(+)	0.756	cggtgcagggcAGCAgaaaggacagaggaga
V$NR2F/TR4.01	(+)	0.754	tgaaatgtgtcagAGGTgaggtgag
V$GTBX/ZEC.01	(-)	0.751	agagcagggttGGTTtcta

Table A17. Complete table of TFBSs of *Homo sapiens* Cebpa Downstream CNS5 region

Matrix	strand	matrix sim.	Sequence
V$MZF1/MZF1.01	(+)	1	gtGGGGagcgt
V$NEUR/MASH1.01	(+)	1	cacacacCTGCtgtg
V$SMAD/GC_SBE.01	(-)	1	ctgggCTCCag
V$TALE/TGIF.01	(-)	1	gacctggctGTCAaaga
V$ZF02/ZKSCAN3.01	(-)	1	acgctCCCCaccttggtggccag
V$EVI1/MEL1.02	(-)	0.996	ggtgtgtGATGaggcca
V$HAND/TAL1_E2A.02	(-)	0.995	ggtcacaCAGCaggtgtgtga
V$ZICF/ZIC2.02	(-)	0.995	tcacaCAGCaggtgt
V$ZF5F/ZF5.01	(+)	0.994	gtgtgtGCGCcctgg
V$ZF11/ZBTB3.01	(+)	0.993	gacaGCCAggt
V$E2FF/E2F1.01	(-)	0.992	agtgGGCGggaaggcag
V$ZFHX/ZEB1.01	(+)	0.991	tacccACCTgcca
V$AP1F/JUNB.01	(+)	0.989	gcgtgaGTCAcaa
V$MEF3/SIX.01	(-)	0.989	gaaTCAGgttcct
V$MYOD/TCF12.01	(-)	0.989	acacagcaGGTGtgtga
V$ZFXY/ZFX.01	(-)	0.989	agGGCCtgggc
V$AP1R/BACH2.01	(+)	0.988	ggggagcgTGAGtcacaaacaagct
V$MZF1/MZF1.03	(+)	0.986	gtGGGGagcgt
V$ZICF/ZIC1.01	(-)	0.986	tcacaCAGCaggtgt
V$E2FF/E2F6.01	(-)	0.984	agtggGCGGgaaggcag
V$AP1F/FOSL1.01	(+)	0.983	gcgtgAGTCacaa
V$MEF3/SIX2.02	(-)	0.983	gaaTCAGgttcct
V$IKRS/IK2.01	(-)	0.982	ctctGGGAatctc
V$ZF08/ZNF354C.01	(+)	0.982	gagctCCACca
V$AP1F/AP1.01	(+)	0.981	gcgtgAGTCacaa
V$MYOD/E47.01	(-)	0.981	acacaGCAGgtgtgtga
V$E2FF/E2F7.02	(-)	0.979	agtgggcGGGAaggcag
V$ZICF/ZIC3.01	(-)	0.976	tcacaCAGCaggtgt
V$STAT/STAT3.02	(+)	0.975	ctgaTTCCaggacgagtca
V$AP1F/BATF.01	(-)	0.974	ttgtgaCTCAcgc
V$AP1F/JUND.01	(+)	0.972	gcgtgAGTCacaa
V$EVI1/EVI1.07	(+)	0.972	gaggcAAGAtgatgaat
V$CSEN/DREAM.01	(-)	0.97	gcGTCAgggcc
V$AP1F/FOSL2.01	(+)	0.968	gcgtgAGTCacaa
V$KLFS/BKLF.01	(+)	0.968	actggGGGTgtgggactgt
V$TALE/MEIS1.01	(-)	0.968	gacctggcTGTCaaaga
V$CEBP/CEBPA.01	(-)	0.967	gcccttgtGTAAcca
V$CEBP/CEBPD.01	(-)	0.967	gccctTGTGtaacca
V$WHNF/WHN.01	(+)	0.967	ctgACGCtggt
V$CEBP/CEBPB.01	(+)	0.966	agggttggGTAAtac
V$NF1F/NF1.03	(+)	0.965	aagctacccacctGCCAgcca
V$AHRR/AHRARNT.03	(+)	0.964	tgtctgagagGCGTgccgagattcc
V$EVI1/MEL1.03	(+)	0.962	ggtagagGATGgtagga
V$MYBL/CMYB.02	(+)	0.961	agaacagagCAACtgcaggaa
V$SP1F/SP4.02	(+)	0.961	ctgagaGGCGtgccgag
V$SPZ1/SPZ1.01	(-)	0.959	aGGTGggtagc
V$HOXF/NANOG.01	(+)	0.958	gatgatgAATGgtgtccct
V$AP1F/AP1.03	(-)	0.955	ttgTGACtcacgc
V$AP2F/TCFAP2A.01	(+)	0.955	attcccagAGGCaag
V$IKRS/IK1.01	(-)	0.954	ctctGGGAatctc
V$KLFS/EKLF.02	(+)	0.954	actgggGGTGtgggactgt
V$XBBF/RFX5.01	(+)	0.954	tccagaacagAGCAactgc
V$NEUR/ATOH1.01	(+)	0.953	aggcCACCtggccac

V$AP1R/BACH2.02	(+)	0.951	ggggagcgTGAGtcacaaacaagct
V$PAX3/PAX3.03	(-)	0.951	agcttCATGgtggagctca
V$ZICF/ZIC3.03	(-)	0.951	tcacaCAGCaggtgt
V$AP1F/JUNDM2.02	(+)	0.95	gcgtgaGTCAcaa
V$CEBP/CEBPB.02	(-)	0.95	gccctTGTGtaacca
V$SF1F/FTF.01	(+)	0.948	ccacCAAGgtgggga
V$LEFF/LEF1.02	(-)	0.947	tggctgtCAAAgaccta
V$AP2F/TCFAP2C.01	(+)	0.945	attCCCAgaggcaag
V$AP1F/AP1.02	(+)	0.943	gcgtGAGTcacaa
V$SP1F/TIEG.01	(+)	0.943	actgGGGGtgtgggact
V$SORY/SOX9.02	(+)	0.942	gttacACAAgggctgggcctgcc
V$KLFS/EKLF.01	(-)	0.941	aggggcaggcaGGGTatta
V$HEAT/HSF1.05	(+)	0.94	gatgaatggtgtccctCCAGaacag
V$AHRR/AHRARNT.01	(+)	0.939	tgtctgagaggCGTGccgagattcc
V$CEBP/CEBPA.02	(-)	0.939	gcccttgtGTAAcca
V$MYOD/MYF5.01	(+)	0.939	acagagCAACtgcagga
V$SP1F/SP1.03	(-)	0.939	ccagtGGGCgggaaggc
V$P53F/P53.03	(+)	0.937	gcctctctgtttgagCATGtctgag
V$CEBP/CEBP.02	(+)	0.935	agggttggGTAAtac
V$ZF02/ZBP89.01	(-)	0.935	gtcccacaCCCCcagtgggcggg
V$PARF/HLF.01	(-)	0.932	agcccttgtGTAAccag
V$EBOX/USF.04	(+)	0.93	caggcCACCtggccacc
V$KLFS/KKLF.01	(+)	0.927	caccaaggtGGGGagcgtg
V$TALE/MRG2.01	(-)	0.927	gacctggctGTCAaaga
V$CARE/CARF.01	(-)	0.926	acagaGAGGct
V$MYOD/MYOD.01	(-)	0.925	ggctGGCAggtgggtag
V$ZF03/ZNF217.01	(+)	0.925	GAATggtgtccct
V$TALE/TGIF1.01	(-)	0.923	gacctggctGTCAaaga
V$CREB/ATF.01	(+)	0.917	aggcccTGACgctggttacac
V$KLFS/KLF12.01	(-)	0.916	acccccaGTGGgcgggaag
V$SP1F/SP4.01	(-)	0.915	ccagtgGGCGggaaggc
V$AP2F/TCFAP2E.01	(+)	0.912	attcccagAGGCaag
V$BRNF/TST1.01	(+)	0.912	tgggtAATAccctgcctgc
V$SP1F/SP1.01	(-)	0.91	ccagtGGGCgggaaggc
V$TALE/PKNOX2.01	(-)	0.907	gacctggctGTCAaaga
V$AHRR/NXF_ARNT.01	(+)	0.906	aaggtggggagCGTGagtcacaaac
V$HICF/HIC1.01	(+)	0.906	accTGCCagccag
V$MYBL/CMYB.01	(+)	0.904	agaacagagCAACtgcaggaa
V$XBBF/RFX3.02	(-)	0.898	gcctcctcatgGTGActcg
V$GLIF/GLIS3.01	(-)	0.897	tctaCCTCccagccttg
V$PARF/VBP.01	(-)	0.897	agcccttgtGTAAccag
V$CTCF/CTCFL.01	(-)	0.896	gtcccacaccccagtgGGCGggaagg
V$P53F/TP63.01	(+)	0.896	gcctctctgtttgagCATGtctgag
V$BCL6/BCL6.04	(+)	0.892	agaTTCCcagaggcaag
V$CTCF/CTCF.05	(-)	0.89	gtcccacaccccAGTGggcgggaagg
V$GREF/ARE.04	(+)	0.889	gtgggactgtgtGTGCgcc
V$HOXC/HOXB4.02	(+)	0.889	aagaTGATgaatggtgt
V$GLIF/GLI1.01	(-)	0.887	tctacctCCCAgccttg
V$PAX3/PAX3.01	(-)	0.886	cTCGGcacgcctctcagac
V$AP1R/BACH1.01	(+)	0.882	ggggagcgTGAGtcacaaacaagct
V$MTF1/MTF-1.01	(+)	0.882	gtgtGCGCcctggcc
V$P53F/P53.05	(+)	0.881	caaaCAAGctacccacctgccagcc
V$E2FF/E2F3.01	(-)	0.88	ccaggGCGCacacacag
V$GREF/GRE.02	(-)	0.879	ggcGCACacacagtcccac
V$PAX3/PAX3.02	(-)	0.878	cgggTCACacagcaggtgt

V$HNF1/HMBOX.01	(+)	0.877	tgacgctgGTTAcacaa
V$IKRS/IK3.01	(-)	0.877	ctctgGGAAtctc
V$TALE/MEIS1.02	(-)	0.876	gacctggctGTCAaaga
V$BRAC/TBX21.01	(-)	0.875	tcacacagcaGGTGtgtgatgaggccagg
V$SP1F/SP1.02	(-)	0.875	ccagtGGGCgggaaggc
V$NEUR/NEUROD1.01	(+)	0.874	aggcCACCtggccac
V$TALE/MRG1.01	(-)	0.874	gacctggctGTCAaaga
V$TALE/TGIF2.01	(-)	0.874	gacctggctGTCAaaga
V$CTCF/CTCF.04	(-)	0.871	gtcccacacccccagtGGGCgggaagg
V$GREF/ARE.03	(+)	0.871	gtgggactgtgtGTGCgcc
V$GRHL/GRHL2.01	(+)	0.87	ggAACCtgattcc
V$GCMF/GCM1.02	(-)	0.867	gcacaCCAGcatcac
V$PARF/TEF_HLF.01	(-)	0.866	gggtaTTACccaaccct
V$GREF/MRE.01	(+)	0.865	gtgggactgtgtGTGCgcc
V$GRHL/GRHL1.01	(+)	0.865	ggcacaGGTTtgg
V$GLIF/GLI3.01	(+)	0.864	ctggCCACcaaggtggg
V$STEM/OCT3_4.01	(-)	0.864	cacaccaGCATcacctggg
V$E2FF/E2F2.01	(-)	0.863	ccaggGCGCacacacag
V$XBBF/RFXDC2.01	(+)	0.863	tccagaacagagCAACtgc
V$AP2F/TCFAP2B.01	(-)	0.862	cttGCCTctgggaat
V$PARF/TEF.01	(-)	0.859	agcccttgtGTAAccag
V$CEBP/CEBPG.01	(+)	0.856	agggTTGGtaatac
V$GRHL/GRHL3.01	(+)	0.855	ggcacaGGTTtgg
V$HOXC/HOX_PBX.01	(+)	0.855	aagaTGATgaatggtgt
V$NR2F/HNF4.03	(+)	0.855	atgaggaggcCAGGaccaggccac
V$STAT/STAT5A.01	(+)	0.855	gagaTTCCcagaggcaaga
O$XCPE/XCPE1.01	(-)	0.853	ggGCGGgaagg
V$MYBL/MYBL1.02	(+)	0.852	agaacagagcAACTgcaggaa
V$PTF1/PTF1.01	(+)	0.852	cacaCCTGctgtgtgacccgg
V$ZF04/ZID.01	(-)	0.851	ggGCTCcaggggc
V$HAND/HEN1.01	(+)	0.85	caggtccCAGGtgatgctggt
V$ETSF/ETS2.01	(+)	0.847	caactgcAGGAacctgattcc
V$GLIF/GLI2.01	(+)	0.846	ctggCCACcaaggtggg
V$TALE/PKNOX1.01	(-)	0.844	gacctggctGTCAaaga
V$CTCF/CTCF.03	(-)	0.842	gtcccacacccccagtGGGCgggaagg
V$CREB/E4BP4.01	(-)	0.841	cagcccttgtGTAAccagcgt
V$GREF/PRE.01	(-)	0.841	tgcagttgctcTGTTctgg
V$NOLF/OLF1.01	(+)	0.841	gccaggTCCCaggtgatgctggt
V$TALE/TGIF2LX.01	(-)	0.84	gggacaccaTTCAtcat
V$XBBF/RFX4.01	(+)	0.839	tccagaacagagCAACtgc
V$NFKB/NFKAPPAB50.01	(-)	0.838	tcaGGGAggccctgg
V$CAAT/ACAAT.01	(+)	0.834	gccaCCAAggtgggg
V$RREB/RREB1.01	(-)	0.834	tCCCAcaccccccagt
V$XBBF/RFX3.01	(+)	0.834	tccagaacagagCAACtgc
V$E2FF/E2F1_DP1.01	(-)	0.832	agtgGGCGggaaggcag
V$BCL6/BCL6.03	(-)	0.83	gctttctTGGAcaggcg
V$GLIF/GLI1.02	(+)	0.83	ctggCCACcaaggtggg
V$AP1R/TCF11MAFG.01	(-)	0.827	acctgtgccggGTCAcacagcaggt
V$E2FF/E2F4_DP2.01	(-)	0.826	agtggGCGGgaaggcag
V$AP1R/VMAF.01	(-)	0.821	gcctgggcactGTGAgcagcttcat
V$FKHD/FREAC4.01	(-)	0.821	cttgtgtaACCAgcgtc
V$DICE/DICE.01	(+)	0.82	tggcCTCAtcacaca
V$EREF/ER.02	(-)	0.814	cctgGGCActgtgagcagc
V$NR2F/TR4.02	(+)	0.812	acagccAGGTcccaggtgatgctgg
V$E2FF/E2F1_DP2.01	(-)	0.811	agtgGGCGggaaggcag

V$RXRF/RAR_RXR.01	(+)	0.811	agccaggtcccAGGTgatgctggtg
V$NRF1/NRF1.02	(+)	0.81	cagtgCCCAggccctga
V$AP1R/NFE2.02	(-)	0.807	acctgtgccggGTCAcacagcaggt
V$ZTRE/ZTRE.01	(+)	0.806	cgcccactGGGGgtgtg
V$NGRE/IR2_NGRE.01	(-)	0.803	ggCTCCaggggcagg
V$RXRF/THRB.02	(-)	0.798	tccctggcctcctcatGGTGactcg
V$HAND/HAND2_E12.01	(-)	0.797	tggtggccaggTGGCctggtc
V$GLIF/ZIC3.02	(+)	0.795	cttccCGCCcactgggg
V$STAT/STAT1.01	(+)	0.794	ctgattccaGGACgagtca
V$CEBP/CEBPE.01	(+)	0.791	agggttggGTAAtac
V$ESRR/ESRRA.04	(+)	0.789	cgctggttacacAAGGgctgggc
V$BCL6/BCL6.02	(+)	0.788	cctgtccAAGAaagcac
V$OAZF/ROAZ.01	(+)	0.788	taGGACccagggttggg
V$PAX6/PAX6.05	(+)	0.788	tgtttgaGCATgtctgaga
V$GLIF/GLIS1.01	(+)	0.787	acgagtcacCATGagga
V$HEAT/HSF1.03	(+)	0.784	ctcgcctgtccaAGAAagcacaagg
V$PERO/PPAR_RXR.01	(+)	0.781	gacagccaggtcccAGGTgatgc
V$E2FF/E2F7.01	(-)	0.778	agtgGGCGggaaggcag
V$NR2F/NR2F6.01	(+)	0.777	gagcccagctccTAGGtctttgaca
V$PRDM/PRDM5.01	(-)	0.777	ctgggctccAGGGgcaggcagggtattac
V$TALE/MRG1.02	(-)	0.774	gggaCACCattcatcat
V$MOKF/MOK2.01	(-)	0.772	accagcgtcagggCCTGggca
V$ZTRE/ZTRE.02	(-)	0.768	caCCCCcagtgggcggg
V$NR2F/TR4.01	(+)	0.766	acagccaggtcccAGGTgatgctgg
O$TELO/ZSCAN4.01	(-)	0.765	cGCACacacagtccc
V$BCL6/BCL6.01	(+)	0.765	agaTTCCcagaggcaag
V$BRAC/MGA.01	(+)	0.765	tgtgcgccctggcctcatcacACACctgc
V$RXRF/THRB.03	(-)	0.764	gactcgtcctggaatcAGGTtcctg
V$SORY/SOX18.03	(+)	0.758	gatGAATggtgtccctccagaac
V$SORY/SOX17.02	(+)	0.756	gatGAATggtgtccctccagaac
V$MYBL/MYBL2.01	(+)	0.753	caactgcagGAACctgattcc
V$SORY/SOX3.04	(+)	0.753	gatGAATggtgtccctccagaac

Table A18. Complete table of GSE8733 TFs correlated with Cebpa

ProbeIDs	Gene1	ProbeIDs	Gene2	Corr_pVal
1417728_at	**Mbd3**	1418982_at	**Cebpa**	0.000558504
1440156_s_at	**Tox2**	1418982_at	**Cebpa**	0.001650935
1420500_at	**Dnajc1**	1418982_at	**Cebpa**	0.003004278
1437567_at	**Zfp772**	1418982_at	**Cebpa**	0.005147655
1423416_at	**Smarcc1**	1418982_at	**Cebpa**	0.005410294
1423260_at	**Id4**	1418982_at	**Cebpa**	0.00652402
1436364_x_at	**Nfix**	1418982_at	**Cebpa**	0.007991639
1455790_at	**E2f2**	1418982_at	**Cebpa**	0.008238733
1433508_at	**Klf6**	1418982_at	**Cebpa**	0.008363104
1420952_at	**Son**	1418982_at	**Cebpa**	0.008567128
1460564_at	**Zfp280b**	1418982_at	**Cebpa**	0.009359608
1446057_at	**Rlf**	1418982_at	**Cebpa**	0.010085094
1438861_at	**Bnc2**	1418982_at	**Cebpa**	0.012204088
1438443_at	**Zbtb20**	1418982_at	**Cebpa**	0.012843492
1459581_at	**Trp63**	1418982_at	**Cebpa**	0.013729815
1423809_at	**Tcf19**	1418982_at	**Cebpa**	0.014761267
1440125_at	**Zfp729a**	1418982_at	**Cebpa**	0.015119386
1450710_at	**Jarid2**	1418982_at	**Cebpa**	0.016094624
1435802_at	**Zbtb45**	1418982_at	**Cebpa**	0.016328991
1448977_at	**Tfap2c**	1418982_at	**Cebpa**	0.016782588
1438930_s_at	**Mecp2**	1418982_at	**Cebpa**	0.016909881
1422103_a_at	**Stat5b**	1418982_at	**Cebpa**	0.017820121
1416724_x_at	**Tcf4**	1418982_at	**Cebpa**	0.018254165
1438255_at	**Foxn3**	1418982_at	**Cebpa**	0.018779953
1431232_a_at	**Mga**	1418982_at	**Cebpa**	0.018915012
1422485_at	**Smad4**	1418982_at	**Cebpa**	0.019643627
1419642_at	**Purb**	1418982_at	**Cebpa**	0.022273575
1418180_at	**Sp1**	1418982_at	**Cebpa**	0.023118288
1439619_at	**Tcf12**	1418982_at	**Cebpa**	0.023649992
1451518_at	**Zfp709**	1418982_at	**Cebpa**	0.026394547
1434178_at	**Kmt2c**	1418982_at	**Cebpa**	0.027437064
1451151_s_at	**Zfp410**	1418982_at	**Cebpa**	0.031206491
1460231_at	**Irf5**	1418982_at	**Cebpa**	0.031447563
1439314_at	**Clock**	1418982_at	**Cebpa**	0.03177044
1451716_at	**Mafb**	1418982_at	Cebpa	0.032190115
1428662_a_at	**Hopx**	1418982_at	Cebpa	0.03268151
1424233_at	**Meox2**	1418982_at	Cebpa	0.033379236
1416982_at	**Foxo1**	1418982_at	Cebpa	0.033654214
1425565_at	**Rest**	1418982_at	Cebpa	0.034203279
1422018_at	**Hivep2**	1418982_at	Cebpa	0.035618338
1425930_a_at	**Mlx**	1418982_at	Cebpa	0.035915609
1426587_a_at	**Stat3**	1418982_at	Cebpa	0.036431175
1419555_at	**Elf5**	1418982_at	Cebpa	0.036728545
1456753_at	**Foxo4**	1418982_at	**Cebpa**	0.038899698

1451739_at	**Klf5**	1418982_at	**Cebpa**	0.041403184
1448957_at	**Rbpj**	1418982_at	**Cebpa**	0.043432576
1438238_at	**2010315B03Rik**	1418982_at	**Cebpa**	0.044328825
1416505_at	**Nr4a1**	1418982_at	**Cebpa**	0.045874734
1456177_x_at	**Zfp706**	1418982_at	**Cebpa**	0.046539787
1424301_at	**Zfp219**	1418982_at	**Cebpa**	0.048293111
1421233_at	**Pknox1**	1418982_at	**Cebpa**	0.049524283
1426377_at	**Zfp281**	1418982_at	**Cebpa**	0.04980269

Table A19. Complete table of GSE35160 TFs correlated with Cebpa

ProbeIDs	Gene1	ProbeIDs	Gene2	Corr_pVal
1417621_at	**Nfatc1**	1418982_at	**Cebpa**	0.001004402
1428662_a_at	**Hopx**	1418982_at	**Cebpa**	0.00238143
1449732_at	**Zscan21**	1418982_at	**Cebpa**	0.004832033
1443591_at	**Dlx4**	1418982_at	**Cebpa**	0.007558002
1417127_at	**Msx1**	1418982_at	**Cebpa**	0.007571702
1429610_a_at	**Zfp511**	1418982_at	**Cebpa**	0.009244521
1418939_at	**Hlx**	1418982_at	**Cebpa**	0.012850828
1418241_at	**Usf2**	1418982_at	**Cebpa**	0.014084548
1422865_at	**Runx1**	1418982_at	**Cebpa**	0.014296301
1448877_at	**Dlx2**	1418982_at	**Cebpa**	0.016758872
1427027_a_at	**Gtf3a**	1418982_at	**Cebpa**	0.01778048
1419303_at	**Heyl**	1418982_at	**Cebpa**	0.018354342
1419355_at	**Klf7**	1418982_at	**Cebpa**	0.019283716
1449559_at	**Msx2**	1418982_at	**Cebpa**	0.020164708
1454758_a_at	**Tsc22d1**	1418982_at	**Cebpa**	0.020884625
1440255_at	**Hinfp**	1418982_at	**Cebpa**	0.024852241
1434689_at	**Zfp637**	1418982_at	**Cebpa**	0.026945719
1418025_at	**Bhlhe40**	1418982_at	**Cebpa**	0.031321922
1454734_at	**Lef1**	1418982_at	**Cebpa**	0.03236389
1448519_at	**Tead2**	1418982_at	**Cebpa**	0.032589675
1428342_at	**Rcor3**	1418982_at	**Cebpa**	0.032796446
1455459_at	**Prdm15**	1418982_at	**Cebpa**	0.03388661
1422970_at	**Mxd3**	1418982_at	**Cebpa**	0.034872099
1451151_s_at	**Zfp410**	1418982_at	**Cebpa**	0.03605094
1439789_at	**Ebf1**	1418982_at	**Cebpa**	0.036630547
1422698_s_at	**Jarid2**	1418982_at	**Cebpa**	0.03771902
1418517_at	**Irx3**	1418982_at	**Cebpa**	0.038192408
1448890_at	**Klf2**	1418982_at	**Cebpa**	0.040089482
1424297_at	**Zfp282**	1418982_at	**Cebpa**	0.040765955
1421163_a_at	**Nfia**	1418982_at	**Cebpa**	0.043471831
1426422_at	**Cc2d1a**	1418982_at	**Cebpa**	0.045609674
1416331_a_at	**Nfe2l1**	1418982_at	**Cebpa**	0.049040763

Table A20. Complete table of GSE10175 TFs correlated with Cebpa

ProbeIDs	Gene1	ProbeIDs	Gene2	Corr_pVal
1422102_a_at	**Stat5b**	1418982_at	**Cebpa**	0.001391707
1452174_at	**Srebf2**	1418982_at	**Cebpa**	0.001526449
1426790_at	**Ssrp1**	1418982_at	**Cebpa**	0.001909138
1438325_at	**Mecom**	1418982_at	**Cebpa**	0.002082493
1455210_at	**Zhx2**	1418982_at	**Cebpa**	0.002772165
1417549_at	**Zfp68**	1418982_at	**Cebpa**	0.002941861
1420973_at	**Arid5b**	1418982_at	**Cebpa**	0.002952893
1454773_at	**Rxra**	1418982_at	**Cebpa**	0.004009393
1434246_at	**L3mbtl3**	1418982_at	**Cebpa**	0.004431938
1419150_at	**Myf6**	1418982_at	**Cebpa**	0.0050273
1455228_at	**Nsd2**	1418982_at	**Cebpa**	0.005134533
1438816_at	**Ahctf1**	1418982_at	**Cebpa**	0.006091356
1450644_at	**Zfp36l1**	1418982_at	**Cebpa**	0.00649331
1425261_at	**Cebpg**	1418982_at	**Cebpa**	0.008494243
1420663_at	**Zbtb7b**	1418982_at	**Cebpa**	0.008779665
1439209_at	**Tcf12**	1418982_at	**Cebpa**	0.008825241
1427518_at	**Zfp930**	1418982_at	**Cebpa**	0.008870363
1439128_at	**Zbtb20**	1418982_at	**Cebpa**	0.009072175
1452069_a_at	**Thap7**	1418982_at	**Cebpa**	0.009388571
1456060_at	**Maf**	1418982_at	**Cebpa**	0.010003663
1455323_at	**Rbak**	1418982_at	**Cebpa**	0.010004502
1419354_at	**Klf7**	1418982_at	**Cebpa**	0.010159427
1449414_at	**Zfp53**	1418982_at	**Cebpa**	0.010211561
1419984_s_at	**Zfp644**	1418982_at	**Cebpa**	0.010684828
1434752_at	**Zfp207**	1418982_at	**Cebpa**	0.010731015
1420174_s_at	**Tax1bp1**	1418982_at	**Cebpa**	0.01093285
1418939_at	**Hlx**	1418982_at	**Cebpa**	0.010953864
1451506_at	**Mef2c**	1418982_at	**Cebpa**	0.01106426
1454758_a_at	**Tsc22d1**	1418982_at	**Cebpa**	0.011140275
1454884_at	**Zbtb46**	1418982_at	**Cebpa**	0.011319307
1436434_at	**E2f2**	1418982_at	**Cebpa**	0.012040883
1438218_at	**Zbtb24**	1418982_at	**Cebpa**	0.012060248
1434896_at	**Zfp955a**	1418982_at	**Cebpa**	0.012679762
1436382_at	**Zbtb12**	1418982_at	**Cebpa**	0.012781815
1418158_at	**Trp63**	1418982_at	**Cebpa**	0.0129853
1457193_at	**Kmt2c**	1418982_at	**Cebpa**	0.013510029
1456753_at	**Foxo4**	1418982_at	**Cebpa**	0.014145437
1420864_at	**Zbtb14**	1418982_at	**Cebpa**	0.014359793
1427844_a_at	**Cebpb**	1418982_at	**Cebpa**	0.014617703
1434355_at	**Zfp617**	1418982_at	**Cebpa**	0.015095229
1432143_a_at	**Hbp1**	1418982_at	**Cebpa**	0.015275689
1450762_s_at	**Zfp24**	1418982_at	**Cebpa**	0.015446998
1437061_at	**Mbd1**	1418982_at	**Cebpa**	0.015459427
1426867_at	**Dmrt2**	1418982_at	**Cebpa**	0.015504829

1427325_s_at	**Akna**	1418982_at	**Cebpa**	0.016279264
1425297_at	**Zfp825**	1418982_at	**Cebpa**	0.016612898
1444041_at	**AU041133**	1418982_at	**Cebpa**	0.017673156
1434318_a_at	**Tfe3**	1418982_at	**Cebpa**	0.019030957
1449515_at	**Zfp292**	1418982_at	**Cebpa**	0.01943651
1451012_a_at	**Ybx3**	1418982_at	**Cebpa**	0.019820378
1439181_at	**Zfp658**	1418982_at	**Cebpa**	0.021483446
1418180_at	**Sp1**	1418982_at	**Cebpa**	0.021492143
1436554_at	**Ikzf5**	1418982_at	**Cebpa**	0.021721469
1437757_at	**Hinfp**	1418982_at	**Cebpa**	0.022997771
1437924_at	**Son**	1418982_at	**Cebpa**	0.023301596
1452381_at	**Creb3l2**	1418982_at	**Cebpa**	0.023358794
1434777_at	**Mycl**	1418982_at	**Cebpa**	0.023509593
1455153_at	**Zfp236**	1418982_at	**Cebpa**	0.023781175
1426297_at	**Tcf3**	1418982_at	**Cebpa**	0.02428637
1418782_at	**Rxrg**	1418982_at	**Cebpa**	0.024810634
1418044_at	**Hmgxb3**	1418982_at	**Cebpa**	0.025438464
1428479_at	**Nfatc1**	1418982_at	**Cebpa**	0.025602335
1423477_at	**Zic1**	1418982_at	**Cebpa**	0.025688358
1450710_at	**Jarid2**	1418982_at	**Cebpa**	0.025960709
1434343_at	**Zfp954**	1418982_at	**Cebpa**	0.026023212
1440890_a_at	**Zfp809**	1418982_at	**Cebpa**	0.026587286
1451730_at	**Zfp62**	1418982_at	**Cebpa**	0.026845037
1440091_at	**Meis2**	1418982_at	**Cebpa**	0.0281983
1424787_a_at	**Nrf1**	1418982_at	**Cebpa**	0.029203673
1425349_a_at	**Myef2**	1418982_at	**Cebpa**	0.029857178
1416029_at	**Klf10**	1418982_at	**Cebpa**	0.030054433
1419155_a_at	**Sox4**	1418982_at	**Cebpa**	0.030560247
1455638_at	**Zfp319**	1418982_at	**Cebpa**	0.03082476
1418301_at	**Irf6**	1418982_at	**Cebpa**	0.031132501
1421646_a_at	**Pias3**	1418982_at	**Cebpa**	0.031290677
1429177_x_at	**Sox17**	1418982_at	**Cebpa**	0.031615017
1438930_s_at	**Mecp2**	1418982_at	**Cebpa**	0.032693723
1425058_at	**Zfp472**	1418982_at	**Cebpa**	0.032974521
1425528_at	**Prrx1**	1418982_at	**Cebpa**	0.03324433
1431127_at	**Zbtb43**	1418982_at	**Cebpa**	0.033507183
1450780_s_at	**Hmga2**	1418982_at	**Cebpa**	0.033726519
1452560_a_at	**Nfya**	1418982_at	**Cebpa**	0.034404149
1434332_at	**Zzz3**	1418982_at	**Cebpa**	0.035255209
1452876_x_at	**2610044O15Rik8**	1418982_at	**Cebpa**	0.035402367
1450333_a_at	**Gata2**	1418982_at	**Cebpa**	0.035450734
1426625_at	**Zfp623**	1418982_at	**Cebpa**	0.035878416
1434569_at	**Tada2b**	1418982_at	**Cebpa**	0.035970663
1416722_at	**Hmg20a**	1418982_at	**Cebpa**	0.036823032
1460647_a_at	**Nr2f6**	1418982_at	**Cebpa**	0.03754912
1449069_at	**Zfp148**	1418982_at	**Cebpa**	0.037751336

1426018_a_at	**Sox6**	1418982_at	**Cebpa**	0.038732715
1419266_at	**Nfyb**	1418982_at	**Cebpa**	0.039285467
1450507_at	**Pou2f3**	1418982_at	**Cebpa**	0.039489439
1422693_a_at	**Sub1**	1418982_at	**Cebpa**	0.039720647
1424091_at	**Zfp868**	1418982_at	**Cebpa**	0.03984608
1449117_at	**Jund**	1418982_at	**Cebpa**	0.040287941
1453105_at	**Zfp263**	1418982_at	**Cebpa**	0.04051186
1426563_at	**Zfp553**	1418982_at	**Cebpa**	0.040737093
1435984_at	**Zfp40**	1418982_at	**Cebpa**	0.040824101
1438211_s_at	**Dbp**	1418982_at	**Cebpa**	0.04236853
1440870_at	**Prdm16**	1418982_at	**Cebpa**	0.04306093
1421708_a_at	**Stat6**	1418982_at	**Cebpa**	0.043804146
1453266_at	**Zbtb4**	1418982_at	**Cebpa**	0.044573415
1428796_at	**Bbx**	1418982_at	**Cebpa**	0.044593035
1448436_a_at	**Irf1**	1418982_at	**Cebpa**	0.04468955
1424297_at	**Zfp282**	1418982_at	**Cebpa**	0.044725711
1453097_a_at	**Ubtf**	1418982_at	**Cebpa**	0.044812122
1450482_a_at	**Pitx2**	1418982_at	**Cebpa**	0.045050045
1418417_at	**Msc**	1418982_at	**Cebpa**	0.045217934
1455260_at	**Lcorl**	1418982_at	**Cebpa**	0.045497597
1416638_at	**Sall2**	1418982_at	**Cebpa**	0.045593359
1441112_at	**Zbed6**	1418982_at	**Cebpa**	0.045710068
1416830_at	**Glmp**	1418982_at	**Cebpa**	0.045923304
1419960_at	**Cphx3**	1418982_at	**Cebpa**	0.047494805
1434101_at	**Nfib**	1418982_at	**Cebpa**	0.047593835
1427233_at	**Tshz1**	1418982_at	**Cebpa**	0.0479962
1457708_at	**Mbd4**	1418982_at	**Cebpa**	0.048103445
1452497_a_at	**Nfatc3**	1418982_at	**Cebpa**	0.048594333
1436883_at	**Mbtps2**	1418982_at	**Cebpa**	0.049563751

ÖZGEÇMİŞ
Çağrı, Buldu;

Doğum Yılı:	1987
Yazışma Adresi:	Dokuz Eylül Üniversitesi İzmir Biyotıp ve Genom ve Merkezi, 35340 Balçova/İZMİR
Telefon:	+90 554 884 1124
Faks:	-
e-posta:	cagribuldu@gmail.com

EĞİTİM BİLGİLERİ

Ülke	Üniversite	Fakülte/Enstitü	Öğrenim Alanı	Derece	Mezuniyet Yılı
Türkiye	Anadolu Üniversitesi	Fen Fakültesi	Biyoloji Bölümü	Lisans	2015
Türkiye	Dokuz Eylül Üniversitesi	İzmir Biyotıp ve Genom Enstitüsü	Moleküler Biyoloji ve Genetik	Yüksek Lisans	2019