FEDERICO II University of Naples

PhD PROGRAM IN NEUROSCIENCE

XXIX CYCLE

PhD Thesis:

IDENTIFICATION AND CHARACTERIZATION OF THE PROMOTER AND REGULATORY ELEMENTS OF THE CEREBRAL SODIUM/CALCIUM EXCHANGER ISOFORM 2 (NCX2) GENE IN RAT PHEOCHROMOCYTOMA CELLS

Coordinator: **Prof. Lucio Annunziato**

Tutor
Dr Pasquale Molinaro

Candidate
Dr Angelo Serani

ACADEMIC YEAR 2016-2017

Acknowledgment

I would like to express my gratitude to my tutor, Dr. Pasquale Molinaro, whose expertise, understanding, and patience guided me through this experience. I appreciate his vast knowledge and skill and his huge contribute to writing this reports. Most important, I'm grateful to him for supporting and letting my imagination free. He provided me with direction, technical support and became more of a mentor than a supervisor.

My sincere gratitude and respect to Prof. Lucio Annunziato for the help, the availability and access to all the best for the realization of this project.

My gratitude goes to the various members of the Department of Neuroscience, they provided a friendly and cooperative atmosphere at work and also useful feedback and insightful comments on my work.

I am deeply and forever indebted to my mother, my family, my friends and my beloved for their love, support and encouragement throughout my entire life.

Abstract

The isoform 2 (NCX2) of sodium–calcium exchanger family is involved in the regulation of the sodium and calcium homoeostasis of neuronal and glial cells. In particular, NCX2 participates in cytosolic Ca^{2+} clearance after spike and synaptic plasticity, whereas under pathophysiological conditions it exerts a neuroprotective effect in stroke.

To investigate the genetic regulation of NCX2, we identified its minimal promoter in the region of 1200 bp, localized between *slc8a2* and *kptn* genes. This sequence induced the transcription of a reporter gene in two different neuronal cell lines, PC12 and SHSY, that express the endogenous NCX2 mRNA. By contrast, the promoter failed to induce the transcription in BHK and U87 cell lines that do not express NCX2 under control conditions. These results reinforced the similarity observed in the transcription activity of the endogenous NCX2 promoter and the region we identified and cloned. In addition, we found, by *in silico* analysis, a great number of putative binding sites for transcription factors in the promoter sequence. These predicted sites showed high matrix identities and were found conserved among three considered species, including rat, mouse, and human genome. Interestingly, many of these TFs are expressed in the CNS where NCX2 is localized, suggesting a possible role in the regulation of the expression of this exchanger under physiological and pathophysiological

conditions. In particular, the overexpression of the transcription factors Sp1, Sp4, and CREB displayed a stimulatory effect on NCX2 promoter activity under control conditions as measured by luciferase reporter assay, whereas SREBP1 exerted an inhibitory effect. Notably, we found that Sp1 and Sp4 share the same molecular determinant localized very close to the transcription start site of the NCX2 promoter, while CREB1 exerted its effect at the beginning of the cloned sequence. We also evaluated the possible subregion where SREBP1 exerted its inhibitory effect on the transcription activity.

Interestingly, the transfection of these TFs in U87 cell line failed to express the endogenous NCX2, thus letting us hypothesize the participation of other regulatory mechanisms, such as epigenetic, in the downregulation of NCX2 transcription. In fact, we found some epigenetic differences between PC12 and L6 cells, expressing or not expressing, respectively, the endogenous NCX2 gene. In particular, we reported significant differences in CpG methylation in the above-mentioned cell lines on the antiporter promoter. In this regard, 5'-Azacytidine, by imparing genomic DNA methylation, induced the expression of NCX2 in U87 cells.

In conclusion, all these results showed that: (1) the basal expression of NCX2 is dependent on the methylation status of its promoter and (2) NCX2 expression can be up-regulated or down-regulated by several transcription factors found in the CNS.

Table of contents

List of Abbreviation

ADD1	Adipocyte Determination and Differentiation Factor 1
Ap-1, 2	Activating Protein 1, 2
ATF-1, 3	Activation Transcription Factor 1, 3
BREu	upstream TFIIB Recognition Element
bZIP	Basic Leucine Zipper Domain
cAMP	cyclic Adenosine Monophosphate
CGN	Cerebellar Granule Neurons
ChiP	Chromatin Immunoprecipitation assay
CNS	Central Nervous System
CRE	cAMP Response Elements
CREB	CRE Binding factor
CTCF	CCCTC-binding factor
D5D	Delta-5 Desaturase
D6D	Delta-6 Desaturase
DCE	Downstream Core Element
DG	Dentate Gyrus
DNA	Deoxyribonucleic acid
DNMT	DNA-methyl transferases

DPE	Downstream Promoter Element
DRE	Downstream Regulatory Element
DREAM	DRE Antagonist Modulator
EMSA	Electromobility Shift Assay
GABA-A	Gamma-Aminobutyric Acid A
GAPDH	Glyceraldehyde 3-Phosphate Dehydrogenase
GTF	General Transcription Factor
HD	Huntington's disease
HIF-1	Hypoxia-inducible Factor 1
Inr	Initiator
KID	kinase-Inducible Domain
LTP	Long Term Potentiation
MAOB	Mono-Amino-Oxidase B
MTE	Motif Ten Element
NCX	Sodium/Calcium Exchanger
NFY	Nuclear Factor Y
NK-kB	Nuclear Factor kappa-light-chain-enhancer of activated B cells
NMDA	N-Methyl-D-Aspartate
PCR	Polymerase Chain Reaction
PIC	Pre-Initiation Complex
PKA	Protein Kinase A

PPF	Paired-Pulse Facilitation
PTEN	Phosphatase and Tensin homolog
PUFA	Polyunsaturated Fatty Acid
PWM	Position Weight Matrix
qRT-PCR	Quantitative Real-Time Polymerase Chain Reaction
REST	RE1 Silencing Transcription Factor
SV40	Simian Virus 40
SCAP	SREBP Cleavage-Activating Protein
SCD1	Stearoyl-CoA Desaturase
Sp1, 3, 4	Specific Protein 1, 3, 4
SRE	Sterol Regulatory Element
SREBP	SRE Binding Protein
STAT3	Signal Transducer and Activator of Transcription 3
TAF	TBP Associated Factors
TBP	TATA Binding Protein
TFBS	Transcription Factor Binding Site
TMS	Transmembrane Segments
TSS	Transcription Starting Site
UTR	Untranslated Region
XCPE1	X Core Promoter Element 1

Chapter 1: Introduction

1.1 Hypothesis

This PhD thesis is based on the hypothesis that the promoter regulating the transcription of the gene encoding for the isoform 2 of the Na^{+}/Ca^{2+} exchanger protein (NCX2) is located in the intergenic region ranging between *slc8a2* and *Kptn* genes.

1.2 Objectives

- To identify the minimal promoter of *slc8a2* gene encoding for NCX2
- To identify several transcription factors involved in the regulation of NCX2 expression

1.3 Thesis organization

While the first chapter of this thesis was a brief introduction of the thesis, with the hypothesis and objectives of this research work, the second chapter of this thesis begins a detailed literature review of the sodium/calcium exchanger (NCX), particularly of the second isoform (NCX2), on its distribution, regulation and relevance under physiological and pathophysiological conditions. Chapter 3 presents a brief review of the eukaryotic transcription regulation and *in silico* approaches to identify gene promoters. Chapter 4 describes objectives and significance of this thesis. Chapter 5 presents the materials and methods used for

all experimental procedures. Chapter 6 describes the feature predictions on the putative NCX2 promoter, experiments validating this hypothesis and all the results obtained. Chapter 7 contains the discussion of the results obtained.

Chapter 2: Literature review on NCX2

2.1 Background

The sodium-calcium exchanger (NCX) is an antiporter of Na^+ and Ca^{2+} ions across the plasma membrane widely distributed in mammalian cells. It represents an essential component of Ca^{2+} homoeostasis under physiological and pathophysiological conditions in both cardiovascular and neurological tissues. Its activity was described for the first time in the late 1960's in the heart (Reuter and Seitz 1968) and in squid axon (Baker, Blaustein et al. 1969). However, the first Na^+-Ca^{2+} exchanger isoform was isolated (Philipson, Longoni et al. 1988), cloned and characterized (Nicoll, Longoni et al. 1990) 20 years later. Nowadays, three isoforms of NCX family have been identified, cloned and functionally characterized from several mammals including mouse (Conway, Wang et al. 2002), rabbit (Armoundas, Rose et al. 2007) guinea pig (Niggli and Lederer 1991) and dog (Sipido, Volders et al. 2002). These isoforms, named as NCX1, NCX2, and NCX3, are encoded by three different genes, *slc8a1*, *slc8a2*, and *slc8a3*, respectively, with a high grade of conservation among different mammalian species. NCX1, NCX2, and NCX3 share ~70% of amino acid sequence identity. NCX1 is expressed ubiquitously in the organism (Nicholas, Yang et al. 1998); NCX2 shows an expression restricted in central nervous system (CNS) (Li, Matsuoka et al. 1994); NCX3 is expressed in the CNS,

immune system and skeletal muscle *(Staiano, Granata et al. 2009)*. On the other hand, invertebrates seem to present only a single NCX ortholog, based upon investigation in arthropods (Dyck, Maxwell et al. 1998) and molluscs (He, Tong et al. 1998).

2.2 Molecular Biology and Topology of NCX2

Among the three exchangers, NCX1 is the most studied on molecular detail. Since NCX1 was studied in cardiomyocytes were it is highly expressed, this isoform is also referred as the cardiac isoform of the sodium/calcium exchanger. Initially, on the basis of the amino acid sequence, it was proposed a topological model of 12 hydrophobic transmembrane segments (TMS) separated by a large intracellular loop named 'f loop' (Philipson, Longoni et al. 1988; Nicoll,

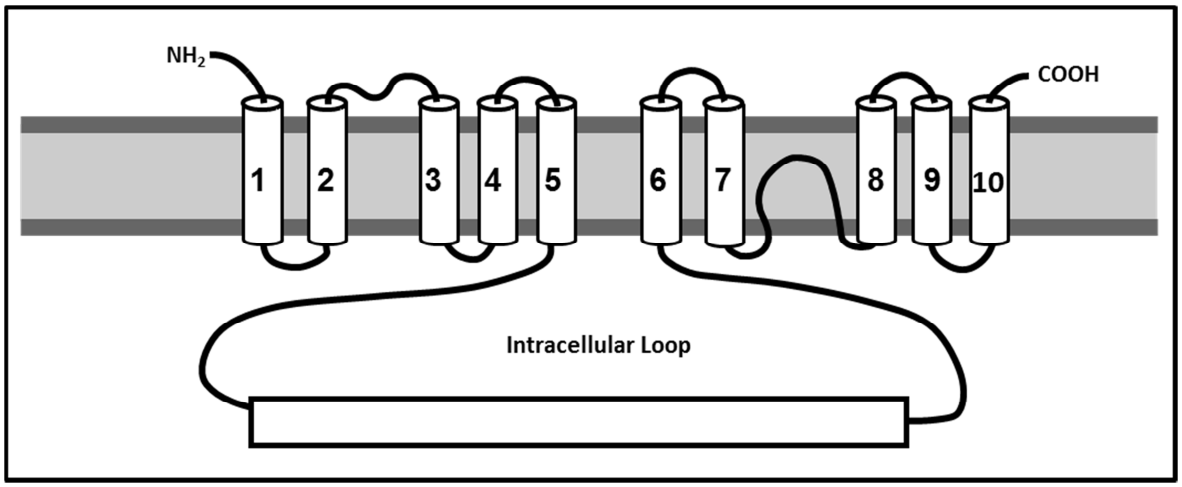


Figure 1: Cartoon representation of NCX topology. NCX is composted by 10 transmembrane segments (TMS). TMS 5 and 6 are separated by a citosolyc intracellular loop

Longoni et al. 1990). However, this model was subsequently corrected to 10 TMS based on the crystal structure of the archaebacterial ortholog NCX_Mj (Ren and Philipson 2013). In all NCX isoforms, TMS can be grouped in an amino-terminal domain, that comprises the first 5 TMS, and a carboxy-terminal domain that includes the last 5 TMS. Mutagenesis experiments indicate that specific regions of the transmembrane segments, named α_1 and α_2 repeats, are critical in the ion transport process (Nicoll, Quednau et al. 1996). In addition, all the exchanger isoforms show an initial hydrophobic segment that represents a signal peptide cleaved during initial processing in the endoplasmic reticulum (Durkin, Ahrens et al. 1991; Hryshko, Nicoll et al. 1993) and both N- and C-terminals exposed to the extracellular side. On the other hand, although specific experiments were not conducted on NCX2 and NCX3 isoforms, scientific community considers similar structures for these two isoforms since they share a high amino acid sequence identity and several common biophysical properties.

2.3 Brain distribution of NCX in Central Nervous System

In CNS, NCX1, NCX2 and NCX3 are widely expressed in neurons, glia, and microglia with several splicing variants that appear to be selectively expressed in different regions and cellular populations of the brain. In particular, NCX1 is highly expressed in pyramidal neurons of the motor cortex, while NCX2 is

expressed at increased cortical somatosensory level. In hippocampus NCX isoforms are expressed at good levels in dentate gyrus, CA1, CA2 and CA3 areas (Papa, Canitano et al. 2003). NCX isoforms are expressed at the mesencephalic level and in important areas for the control of extrapyramidal motor coordination of the basal ganglia. NCX1 signal is found in the compact part of the substantia nigra where the bodies of the dopaminergic cells are localized, while in the nucleus accumbens are expressed all isoforms of NCX (Canitano, Papa et al. 2002; Papa, Canitano et al. 2003).

2.4 NCX family genes

NCX family includes three different genes that encode three proteins: NCX1 (Nicoll, Longoni et al. 1990), NCX2 (Li, Matsuoka et al. 1994) and NCX3 (Nicoll, Quednau et al. 1996). NCX1 mRNA is found in 15 different alternative splicing forms that seems to be tissue or cell specific (Quednau, Nicoll et al. 1997; Philipson and Nicoll 2000), whereas NCX3 mRNA is found in 6 different splicing variants (Gabellini, Bortoluzzi et al. 2002). The region subjected to splicing variance is localized in the long intracellular loop f and comprises about 110 amino acids (Kofuji, Lederer et al. 1994). To date, alternative splicing forms of NCX2 were not identified.

Genes encoding for NCX1 are often long, about long 200 kb (Kofuji, Lederer et al. 1994; Scheller, Kraev et al. 1998), and are dispersed among different

species, since it located on chromosome 2 in human, chromosome 17 in mouse, chromosome 6 in rat. The entire gene is organized in twelve exons: first exon is a non-coding sequence and contains part of the gene promoter; the exon 2 together with exon 12 represents the major part of mRNA transcript (Kraev, Chumakov et al. 1996); exons 3-8, also indicated with the letters from A to F, are combined in different splicing isoforms to confer the tissue-specificity (Nakasaki, Iwamoto et al. 1993; Lee, Yu et al. 1994). In particular, exons A and B are mutually exclusive (Kofuji, Lederer et al. 1994), the presence of exon D is necessary for all the spliced isoforms, while the exons C, E and F may be present or not in the mature RNA sequence. Generally, in excitable cells such as cardiomyocytes and neurons, it is expressed exon A, while the exon B is expressed in non-excitable cells including kidney, stomach and skeletal muscle cells (Quednau, Nicoll et al. 1997).

The expression of NCX1 gene is controlled by three different tissue-specific promoters, named brain, heart and kidney promoters (Barnes, Cheng et al. 1997; Nicholas, Yang et al. 1998). The heart promoter does not possess the "TATA" box region, one of a hallmark of the presence of a putative promoter, but it is characterized by multiple transcriptional start sites. Such promoter possesses two consensus sequences for fos and jun (AP-1) transcription factors and two GATA-family binding sites (Nicholas, Yang et al. 1998). On the opposite, the promoter of the renal variant of NCX1 presents one "TATA box" and consensus

sequences for tissue-specific transcription factors. The promoter of brain variant is particularly rich in GC nucleotides and induces high level of expression in neuronal cells, and a lower level of expression almost in all other non-excitable tissues. Cerebral promoter does not possess the "TATA box" and contains many binding sites including Sp1, AP2, NF-kB, HIF-1, and REST consensus sequences (Nicholas, Yang et al. 1998; Sirabella, Secondo et al. 2009; Valsecchi, Pignataro et al. 2011; Formisano, Guida et al. 2015)

The gene of the third isoform of NCX is located on chromosome 12, 14 and 6 in mouse, human, and rat, respectively. It consists of 9 exons very similar to that observed in NCX1. In fact, NCX3 gene lacks the homologous exons 6-8, named D-F, if compared with NCX1 gene (Gabellini, Bortoluzzi et al. 2002; Gabellini, Bortoluzzi et al. 2003; Gomez-Villafuertes, Torres et al. 2005).

NCX3 possesses a single promoter whose 5'UTR region is rich in GC nucleotides and its minimal functional region contains all the specific sequences for the expression in neuronal and muscle cells (Gabellini, Bortoluzzi et al. 2003). The main consensus sequence for the tissue-specific expression of NCX3 is the cAMP response element (CRE), although alone is not enough to induce the transcription *in vivo* (Gomez, Gabellini et al. 2004). NCX3 promoter also features a double-inhibitory sequence known as downstream regulatory element (DRE). These two sequences confer an auto-regulatory mechanism to the gene inasmuch, when NCX3 activity is very high and thus intracellular calcium

concentration falls below a certain level, the affinity of DREAM for the DRE sequences increases, and consequently inhibits the transcription of NCX3 gene. (Gomez, Gabellini et al. 2004).

The gene coding for NCX2, instead, consists of 10 exons and has been localized on chromosome 7, 19 and 1, in mouse, human, and rat, respectively. The knowledge of the localization of NCX2 gene and the characterization of its regulatory properties was not accomplished yet (Kraev, Chumakov et al. 1996), and it is the object of the present study.

2.5 Relevance of Na^+/Ca^{2+} exchanger activity under physiological and pathophysiological conditions

The primary physiological role of NCX is Ca^{2+} efflux from the cell. This activity, known as "forward mode", contributes to the restoration of the intracellular level of Ca^{2+} after its physiological increase in concentration. This process is driven by the cellular Na^+ gradient produced by the Na^+/K^+ pump. According to the proposed model for NCX transport under physiological conditions, extracellular Na^+ binds to the transport site, is internalized and released inside the cell. Afterward, Ca^{2+} ion binds its translocation site, is internalised into the membrane and is released into the extracellular space (Hilgemann, Matsuoka et al. 1992). By contrast, NCX can also work in a

"reverse mode" extruding 3 ions of Na^+ and intruding 1 ion of Ca^{2+}. This mode is essential to restore the Na^+ concentration under some physiological or pathophysiological conditions, for instance, after a membrane potential depolarization. To date, classical exchange stoichiometry is 3:1 (Na^+:Ca^{2+}), but, however, this physiological rate might range from a minimum of 1:1 to a maximum of 4:1 (Kang and Hilgemann 2004).

Under pathophysiological conditions, NCX often works in the reverse mode, causing Ca^{2+} influx. This inward flow of Ca^{2+} can contribute to calcium overload, which may ultimately lead to myocardial damage resulting from arrhythmogenesis (Pogwizd 2003), cellular ultrastructural changes, disruption of membrane integrity and mitochondrial dysfunction (Bers 2008). Dysregulation of cellular Ca^{2+} homeostasis is also a common feature in a variety of neurological disorders, including stroke, epilepsy, and trauma, where activation of Ca^{2+}-dependent proteases, such as calpains, and activation of nitric oxide production lead to excitotoxic neuronal death (Araújo and Carvalho 2005; Kintner, Wang et al. 2007). Inhibition of the reverse mode of NCX may be useful in maintaining physiological levels of cytosolic Ca^{2+} and protecting cells from death (Matsuda, Arakawa et al. 2001)

Alteration in the regulation of ionic homoeostasis is one of the main event occurring during neuronal damage. In this regards, NCX plays an important role in maintaining homoeostasis of the intracellular Ca^{2+} and Na^+ ions under some

pathophysiological conditions. For these reasons, NCX is considered as a possible drug target to slow the course or to prevent the pathogenesis of some serious diseases characterized by loss of ionic homeostases, such as Alzheimer's, multiple sclerosis, epilepsy, ischemic heart, kidney, and brain diseases. In particular, during cerebral ischemia, there is a non-supply of oxygen and glucose that, in turn, leads to a marked reduction of ATP levels. ATP loss causes the block of the two ATP-dependent transport systems: the Na^+/K^+ ATPase and the Ca^{2+} ATPase. Consequently, there is a homoeostasis imbalance of these ions with progressive increase of their intracellular concentrations that cause necrosis or apoptosis of neurons after anoxia. Under these conditions, it is clear the importance of the role played by membrane proteins that do not make direct use of ATP. In fact, an increase in the activity of NCX may restrict or counteract the excessive accumulation of intracellular Ca^{2+} and Na^+ and, consequently, facilitate neuronal survival. In recent studies, it was shown that the knocking-down of NCX1 or NCX3 (Pignataro, Gala et al. 2004), or the deletion of NCX3 (Molinaro, Cuomo et al. 2008), causes a marked increase in infarct volume after cerebral ischemia, induced by middle cerebral artery occlusion. Even the use of inhibitors of NCX worsens the cerebral infarction, while the administration of compounds that enhance NCX activity reduces the area of cerebral infarction (Molinaro, Cuomo et al. 2008).

The absence of NCX2 produces a delay in the cytosolic Ca^{2+} after a depolarization event in hippocampal neurons. As consequence of this phenomenon, paired-pulse facilitation (PPF) and post-tetanic potentiation resulted enhanced. Furthermore, NCX2 knockout mice show an improved plasticity at the presynaptic level (Jeon, Yang et al. 2003). As the effect of this potentiation, NCX2 knock-out mice show an improved performance in learning and memory tasks that depends on from the hippocampus, such as water-maze and fear conditioning. This evidence suggests that NCX2 may play an important role in disease involving a cognitive impairment such as Alzheimer's and Parkinson's diseases. On the other hand, deletion of NCX2 increases the hippocampal susceptibility to the ischemic insult and a worsening of stroke outcome (Jeon, Chu et al. 2008).

Interestingly, NCX2 was found to be silenced or disrupted in glioma. Indeed, this tumour presents a frequent loss of heterozygosis in the long arm of chromosome 19 (19q13.3) where NCX2 is localized (Qu, Jiao et al. 2010). This genetic deletion is combined with an increasing of the methylation in this locus (Ceccarelli, Barthel et al. 2016). Take together this evidence suggests that the transcriptional silencing of NCX2 may be mediated by an epigenetic modification and that NCX2 may play a role to hamper the tumour development.

Chapter 3: Literature review on gene promoters

3.1 Eukaryotic regulation of the transcription

A gene is a unit of heredity that may influence the outcome of an organism's traits. These informative units are located at specific sites of a chromosome as a segment of DNA and they encode instructions necessary for the synthesis of proteins that are responsible for almost all biological processes and hence, phenotypic traits of an organism. Eukaryotic genes consist of alternating sequences of exons and introns. Typically, the extremities of a gene are characterized by sequences that are transcribed but not translated in proteins: 5' and 3' untranslated regions (UTRs). Region upstream the transcription starting site together with the 5' UTR is often involved in the regulation of the gene transcription and expression (Weaver and Hedrick 1997), and might contain the gene promoter. The promoter is influenced by several regulatory DNA-binding molecules. The outcome regulation of these elements depends on the specific affinities and a unique distribution of particular DNA sequences on the promoter (Segal and Widom 2009). These sites are structured to be recognized by transcription factors (TFs), a class of proteins or protein complexes with the ability to bind specific DNA sequences usually located in the gene promoter, called cis-acting elements.

3.2 The Eukaryotic Transcriptional Machinery

The machinery of gene transcription requires being accurate and fined controlled by the cell. A high number of factors have involved the composition of the complex that leads the RNA polymerase II to the transcription starting site (TSS) to initiate mRNA synthesis. These factors are grouped into general (or basic) transcription factors (GTFs), promoter-specific activator proteins (activators), and co-activators according to their activity or localization. GTFs are assembled to construct the basal transcription machinery necessary and sufficient to initiate the transcription. This transcription machinery includes RNA polymerase II and a variety of auxiliary components, such as TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH. These factors join the transcription in a precise sequence and finally assembly the pre-initiation complex (PIC), which directs RNA polymerase II to the TSS. The first step in PIC assembly is the binding of TFIID to the promoter. This factor is composed of different subunits consisting of TATA-box-binding protein (TBP) and a set of TBP associated factors (TAFs). The following steps include promoter melting, clearance, and escape, before the formation of a fully functional RNA polymerase II elongation complex. However, this transcription machinery does not initiate gene transcription. In fact, this control mechanism lays the scaffold structure composed by TFIID, TFIIE, and TFIIH on the binding sequence. Gene

transcription effectively starts when another RNA polymerase II-TFIIF-TFIIB complex is recruited (Weaver and Hedrick 1997).

3.3 Core Promoter

The minimum sequence that initiates DNA transcription is often located upstream the TSS and it named core promoter. This sequence frequently is ~40 nucleotides long and contains binding sites for transcription factors involved in the formation of PIC complex. These binding sites can include TATA (TATA-box), CAAT (CAAT-box) sequences, upstream or downstream TFIIB Recognition Element (BRE^u or BRE^d), Initiator (Inr), Motif Ten Element (MTE), Downstream Promoter Element (DPE), Downstream Core element (DCE), and X Core Promoter Element 1 (XCPE1). The Inr sequence is the most common cis-element, found in nearly half of the focused promoters. Its consensus sequence consists of YYANWYY (where Y is C or T, W is A or T and N is A, T, C or G) and the adenine represents the +1 from TSS. This motif showed a selective signal for the subunit TAF1 and TAF2 of the Transcription Factor II D (TFIID) (Chalkley and Verrijzer 1999) and can be associated to other core element motifs such as TATA-box, DPE or MTE. TATA-box was the first identified and the most investigated cis-acting element, its presence is restricted to the 10-15% of the mammal's promoters. The position of this motif depends on the tissue specificity of the promoter, but generally, is located at position -31 or -30 from

the +1 of the Inr. This motif is bound by the TATA-box Binding Protein (TBP), a subunit of transcription factor TFIID. The two sequences BRE^u and BRE^d participate in the composition of the initiator mechanism together with TATA boxes. These motifs are bound by TFIIB to complete the TFIIB-TBP promoter complex. Both DPE and MTE are signals for TFIID alternative to the TATA-box and the distance of those motifs from Inr sequence is generally crucial for the optimal construction of TFIID-based initiation machinery

Promoters provided with more alternative TSS are classified as “dispersed promoters”, whereas promoters that present a single TSS and either TATA or CAAT boxes are classified as “focused promoters”.

Dispersed promoters generally lack BRE, TATA, DPE, and MTE motifs

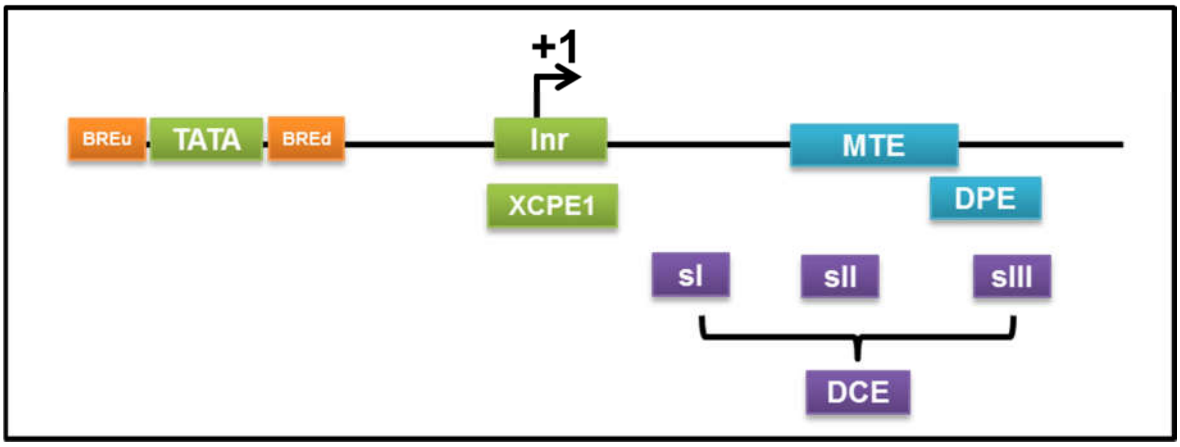


Figure 2. Cartoon representing the relative position of core promoter elements: BREu and BREd are located upstream and downstream of the TATA box, respectively. Inr and XCPE1 are located on the transcriptional starting site (TSS). MTE, DPE and the alternative initiation motif DCE are localized downstream the TSS.

(Juven-Gershon and Kadonaga 2010) and present multiple Inr combined with GC-boxes or other binding site for regulatory elements.

Focused promoters can present the combination of Inr and TATA-box, DPE or MTE motifs. Alternatively, focused promoter can lack Inr signal and can present DCE, XCPE1 or XCPE2 motifs in association or not with a TATA motif. However, the presence of DCE, XCPE1 or XCPE2 motifs is frequent in lower species, whereas these consensus sequences are present only in 1% of human promoters. Notably, DCE, XCPE1 and XCPE2 are mostly present in TATA-less human promoters. However, statistical analysis on 10,000 predicted promoters has shown that the focused promoters are not frequent as though before (Gershenzon and Ioshikhes 2005).

The mechanisms of gene transcription are different between focused and dispersed promoters. Their molecular dynamics are currently under investigations of basic research.

3.4 Proximal Promoter

The proximal promoter is generally defined as a regulatory sequence of 4 or 5 hundred bases that modulates gene transcription, and contains several specific transcription factor binding elements. Typically, this sequence confers the tissue specificity of a gene or allows genes to be selectively expressed in response to certain environmental conditions. TFs that bind these elements can deform

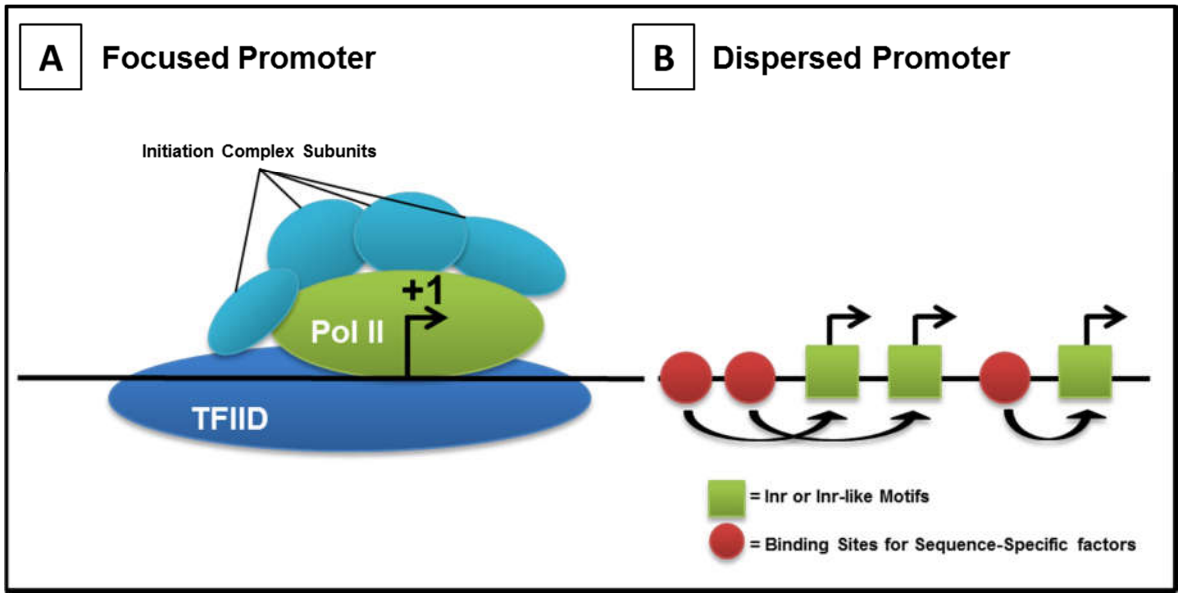


Figure 3. Cartoon representing focused (on the left) and dispersed (on the right) promoter. (A) Focused promoters are characterized by the presence of a single transcription starting site and an initiator complex (blue) for the RNA polimerase II (green). **(B)** dispersed promoters display multiple transcription starting sites with relative initiator motifs (green) and binding sequences for specific factors (red)

DNA, moving the distal promoter closer to the basal transcription complex and thus influence transcription machinery. Proximal promoters can enrol enhancers or silencers to regulate the transcriptional activity of the core promoter (Atchison 1988).

3.5 Distal regulatory elements

Distant regulatory elements are DNA sequences that participate in the regulation of gene transcription and that can be located thousands of base pairs distant from the TSS. In particular, distal regulatory elements can be located in the intergenic regions upstream of a gene, in introns, in a region downstream of the 3'UTR or even in a far chromosome region (Weaver and Hedrick 1997). Depending on

their activity and position, distant regulatory elements can be grouped in: enhancers, silencers, insulators and locus control regions. Enhancers amplifies the expression mediated by a single or more independent promoters. A prototype model of an enhancer consists of 10 to 15 transcription factor binding sequences (TFBS) of at least two or three cooperative transcription factor families. The composition of these enhancers is analogous of a proximal promoter. The interaction between the enhancer and the promoter is regulated by several control mechanisms. First of all, the efficiency of the stimulation strictly depends on the position of the enhancer from the target promoter. A second mechanism consists in the interference of a structure known as insulator. In addition, the pattern of transcription factors on the promoter must reflect the appropriate elements of interaction on the enhancer to allow the correct tethering of the structure. Mutations of enhancer sequences, or of their TFs, may result in some pathological conditions such as preaxial polydactyly (Lettice, Heaney et al. 2003) or Van Buchem diseases (Loots, Kneissel et al. 2005).

Silencer sequences, besides the negative effect on the transcription, are provided with many analogies with enhancers. Silencer can downregulate the promoter activity in a position- and orientation-dependent manner, as well as the enhancer. In addition, TFs that bind these structures are typically repressor or even can be both repressor or activator, depending on the target promoter.

Recent studies suggest that silencers can inhibit transcription by compromising the formation of the PIC complex (Chen and Hampsey 2002).

Insulators are DNA sequences that block the activation of a gene owing to a collateral unspecific activity of neighbours genes or stimulatory structures. They are usually composed of a cluster of binding sites for zinc finger proteins. CCCTC-binding factor (CTCF) is the only known factor mediating an insulator activity in vertebrates. Insulators may also act as a barrier preventing the spread of chromatin condensation to genes that need to remain actives. A typical insulator can present only one of the two types of control described above (Recillas-Targa, Pikaart et al. 2002).

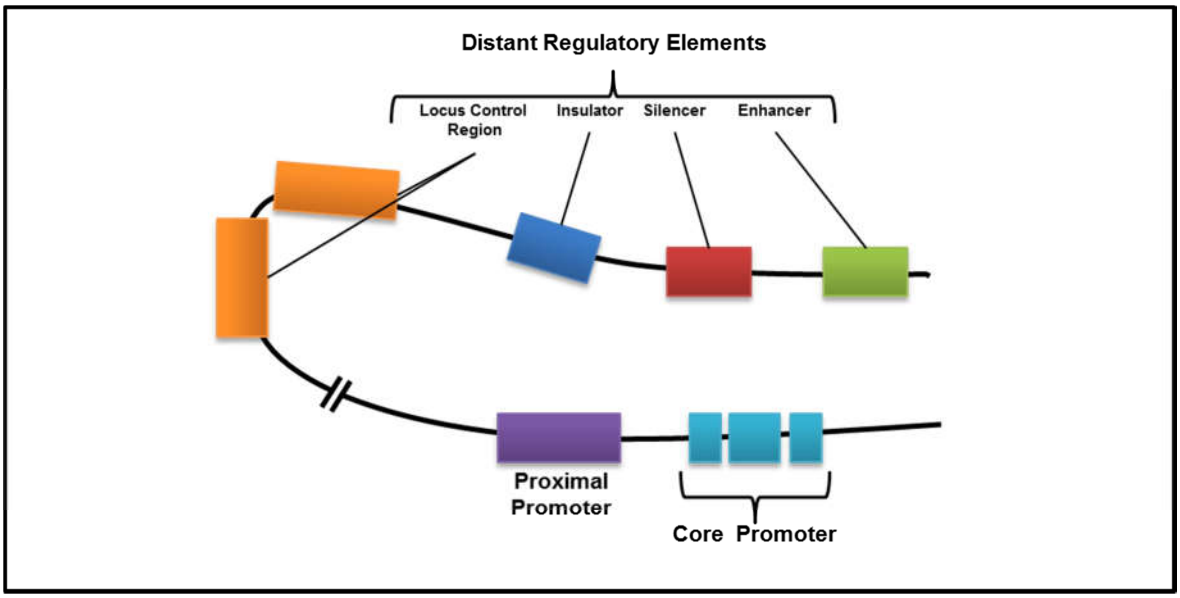


Figure 4: Cartoon representing the regulatory elements of a promoter: this representation displays a core promoter element (cyan), proximal promoter (purple), locus control region (orange), insulator (blue), silencer (red) and enhancer (green).

3.6 Transcription Factors

RNA polymerase II machinery initiates gene transcription by using core promoters. This reaction can be also obtained with high efficacy in a reaction solution *in vitro*. However, there are many regulatory proteins that can participate in the regulation of the transcription efficacy *in vivo*. This can be achieved by the interaction of TFs with proximal and/or distal promoters and thus by modulating the assembly of disassembly of the basal transcription mechanism. TFs are classified as activators, repressors or co-regulators depending on their mechanisms of interaction with the promoter and other proteins. In particular, TFs are considered activators or repressors according to their positive or negative effect on the transcription process. However, in most of the cases TF proteins can exert both activatory and inhibitory effects depending on specific cellular contexts. As regard co-regulators, these proteins modulate the effect of other TFs, or DNA binding proteins, via protein-protein interaction without a direct DNA bind. Co-regulators can be subclassified as co-activators or co-repressors depending on the effects of their interaction with the transcription efficiency. More important, they represent the link that mediates extra-cellular and intra-cellular signals to the transcription of target genes. These kinds of co-regulators exert their activity on the transcriptional machinery via enzymatic functions. For instance, co-regulators can change the chromatin environment of regulatory regions, or can mediate phosphorylation,

ubiquitinylation or sumoylation of target transcription factors (Lonard and O'Malley 2005).

3.6.1 cAMP response element binding protein (CREB)

CREB is a transcriptional factor that was first isolated in undifferentiated neuron-like PC12 cells (Montminy, Sevarino et al. 1986). This TF is a member of Basic Leucine Zipper Domain (bZIP) superfamily able to bind a DNA sequence consisting of the palindromic core 5'-TGACGTCA-3', known as cAMP response element (CRE) (Mayr and Montminy 2001). CREB family also includes the transcription factor CRE modulator (CREM) and Activation Transcription Factor 1 (ATF-1). Each member of this family is encoded by a different gene that contains at least two highly conserved regions. One region is composed of a ZIP domain followed by one or two structures that bind CRE motifs, whereas the central part of the protein is characterized by a kinase-inducible domain (KID) containing several protein kinase A (PKA) phosphorylatable residues. Frequently, the KID domain is delimited by two glutamine-rich regions (Q1 and Q2) that mediate the protein-protein interaction among CREB family members. These TFs can form both homodimers and heterodimers, however, the CREB homodimer complex presents the highest transcriptional efficacy (Dworkin and Mantamadiotis 2010).

Different metabolic pathways can activate CREB to enhance or to initialize the transcription of several target genes. CREB binds almost 4000 known promoters (Zhang, Odom et al. 2005), however, to date, it is still unknown how it is able to mediate distant responses from a very high number of different physiological and pathophysiological signallings.

Phosphorylation on the KID domain is the key mechanism of CREB regulation. Indeed, the counteracting effect of protein kinases and phosphatases regulate the activity of this TF in important processes in which it has a central role, including cell cycle (Liu, Begley et al. 2014), cell death (Martin 2010), DNA damage (Abreu, Kumar et al. 2013) and neurogenesis (Faigle and Song 2013). The CREB-mediated transcriptional regulatory machinery can require additional accessory proteins including PTEN or CREB-binding protein (CBP). In particular, CBP joins the machinery as chromatin remodeler, unwinding the promoter and providing the access to RNA polymerase II (Chan and La Thangue 2001). If this protein does not participate in the CREB-mediated transcription complex, the target genes remain transcriptionally silent (Merz, Herold et al. 2011).

CREB binding motif may present an additional regulatory element that consists in the presence a central CpG dinucleotide. The methylation of this group is able to sterically impair the binding to the CRE element, even when CREB is activated by phosphorylation (Ortega-Martínez 2015).

In last few decades, an increasing amount of studies demonstrated that CREB plays a pivotal role in several regulatory processes of the nervous system. In particular, during embryogenesis, this TF participates to the differentiation of the neuronal progenitor cells and to the development of brain. In the adult, CREB is involved in neuronal plasticity and consolidation of long-term memory (Yamashima 2012). CREB can be also activated in response to a harmful stimulus enhancing transcription of pro-survival genes. The protective effect of CREB-activated genes depends on phosphorylation span of the TF. For instance, in hippocampal dentate gyrus (DG), the increased permanence of CREB in the phosphorylated form causes more neuroprotection after cerebral ischemia as compared with CA1 area (Mabuchi, Kitagawa et al. 2001).

3.6.2 Specific Protein (SP) Family

Sp-family includes a wide group of transcription factors involved in the expression of genes that participate in the regulation of several biological functions including brain development, cell growth, apoptosis, and differentiation. Members of this family are generally able to bind GGGGCGGG (GC-box) and the related GGTGTGGGG (GT/CACC-box) sequences.

Members of this family are characterized by three C_2-H_2 zinc fingers in the C-terminus of the protein, whereas N-terminal region presents glutamine-rich

domains in proximity to the serine/threonine stretches that are involved in the binding with co-regulator proteins. The N-terminal region also contains an endoproteolytic cleavage site necessary for proteasome-dependent degradation. The C_2H_2-type zinc finger region is structured by 81 amino acids, mediates DNA binding, and is the most highly conserved part of the proteins. Among Sp family members, Sp1, Sp3, and Sp4 are most closely related each other than Sp2. In addition, the zinc fingers of Sp1/3/4 present higher affinity to GC-boxes if compared with Sp2 (Suske 1999).

3.6.2.1 Sp1

Sp1 was the first isoform of the Sp family to be identified and it is considered the prototypic member (Kadonaga, Carner et al. 1987). Sp1 gene encode for a 105 kDa protein that is ubiquitously expressed in the mammal cells. The expression of this protein is higher in the astrocytes than in the neurons in the

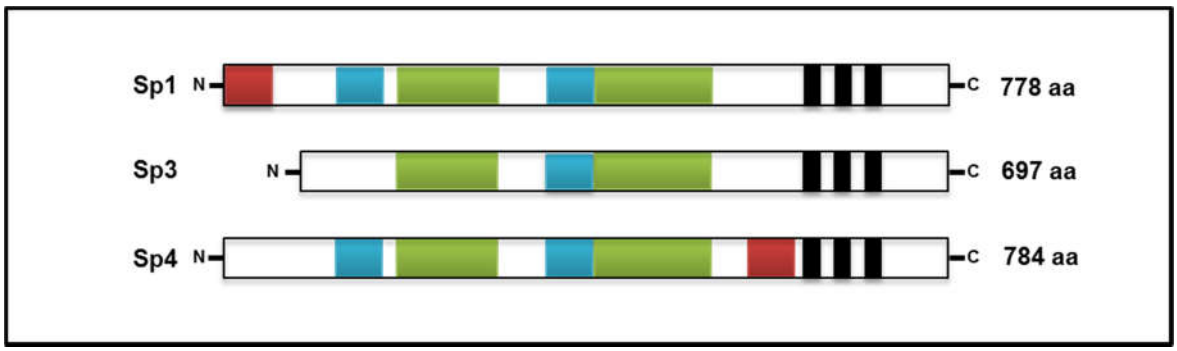


Figure 5 Topology of Sp1, Sp3 and Sp4: Coloured boxes indicate glutammine rich domains (green), serine/threonine residues (blue), inhibitory domains (red) and zinc finger domains (black). Amino acid length of the respective protein is indicated on the right.

brain. Sp1 presents a typical structure of Sp family, indeed, it has two activatory glutamine-rich regions in the center part of the protein and an inhibitory domain related to a proteasome-dependent degradation in the N-terminus part (Gill, Pascal et al. 1994; Murata, Kim et al. 1994; Su, Roos et al. 1999). Zinc fingers and the nearby area are necessary to bind multiple CG-boxes and form homo- or hetero-typic complexes with other transcription factors and co-regulators (Suske 1999).

To date, many mechanisms of DNA binding have been postulated for this TF. In particular, Sp1 can bind, with synergic activity, sequential consensus sequences of gene promoters as homo-, or hetero-multimer with Sp4 (Mastrangelo, Courey et al. 1991; Su, Roos et al. 1999). In addition, Sp1 seems to be regulated by the methylation status of target DNA. In fact, methylation of regulatory GC-boxes outside the binding site produces a significant decrease in the affinity of this TF. On the other hand, methylation of the CpG dinucleotide inside the binding motif does not reduce the Sp1 affinity for the site.

Basal activity of Sp1 participates in the transcription of housekeeping, tissue-specific and cell-cycle-related genes including core promoter transcription factors (Dunah, Jeong et al. 2002). Sp1 also exerts a role in development processes since knock-out animals present abnormal embryogenesis and die in the early stage of fetus formation (Marin, Karis et al. 1997). Among Sp1 target genes there are several receptors, such as NMDA and D_2 dopamine receptors,

and proteins regulating neurotransmission, such as MAO-B (Yajima, Lee et al. 1998; Chen, Kundakovic et al. 2007), that might exert a role in CNS disorders. Indeed, some human disorders can be related to Sp1 alteration or misregulation. A wide amount of studies suggests that Sp1 and its homologs, Sp3 and Sp4, participate to brain tumour development. In fact, their increased expression is frequently associated with a negative prognosis in several brain tumours such as glioma (Safe and Abdelrahim 2005). Sp1 also exerts a role in several neurodegenerative disorders such as Huntington's disease, (Dunah, Jeong et al. 2002), Alzheimer's (Santpere, Nieto et al. 2006), multiple sclerosis (Kristjansdottir, Sandling et al. 2008) and psychiatric disorders. In addition, Sp1 was found down-regulated in certain brain areas involved in the regulation of the social behaviour, such as striatum. On the other hand, Sp1 expression increases in the parietooccipital cortex, which misfunction in schizophrenia leads to hallucination and inability to planning (Ben-Shachar and Karry 2007).

Sp1 is activated during oxidative stress and hypoxia (Ryu, Lee et al. 2003; Miki, Ikuta et al. 2004) where it regulates the activity of neuroprotective and neurodetrimental genes. In particular, Sp1 increases transcription of CD39, that enhances cell survival (Eltzschig, Köhler et al. 2009), after hypoxic conditions, and enhances the transcription of the Sur1-regulated Ca^{2+}-ATP channel that leads to the enlargement of ischemic edema. Sp1 also act as co-regulator of ATF3, c-Jun, and STAT3 to activate injury-inducible genes in damaged neurons

(Kiryu-Seo, Kato et al. 2008). Despite the role of Sp1 under CNS pathophysiological conditions is constantly reinforced by new evidence, the mechanisms underlying its activity, especially in stroke, are not been determined yet.

3.6.2.2 Sp3

Sp3 was cloned for the first time in 1992 by two different groups (Kingsley and Winoto 1992; Hagen, Müller et al. 1994; Abreu, Kumar et al. 2013). As mentioned before, Sp3 presents high homology with Sp1 and the same activatory and inhibitory domains. Sp3 is expressed ubiquitously and knock-out mice present an impaired post-natal development and ossification process. In addition, this TF, as its homologous Sp1, is also involved in the regulation of brain-related genes with a key role in several neurological and behavioural disorders including D_2 (Yajima, Lee et al. 1998) dopamine receptors, MAO-B (Shih and Chen 2004) and GABA-A receptor subunit α4 (Ma, Song et al. 2004). Although literature presents less information on Sp3 involvement in brain disorders than is reported for Sp1, the investigation on the role of this protein is still at the beginning.

As regards the activity of Sp3, this isoform can act as repressor, rather than activator, of many target genes and is able to positively or negatively modulate the effect of other Sp-family members. For instance, in promoters that present multiple Sp1 sites, Sp3 can hamper the Sp1/4-mediated enhancement by two

ways: (1) it can compete for the same binding motif; (2) it can block the synergic Sp1 modulation by occupying a proximal regulatory GC-Box (Yu, Datta et al. 2003). However, Sp3 shows higher activatory effect, if compared with Sp1, on some gene promoters under specific conditions. For instance, p21 promoter is provided with six Sp1 binding sites, but Sp3 activates the gene transcription better than Sp1 (Gartel, Ye et al. 2001).

In addition, Sp3 is provided with an important post-transcriptional regulation mechanism, including sumoylation or acetylation at the level of regulatory domains, that participate to the repressor or activator activity of the TF (Braun, Koop et al. 2001; Ammanamanchi, Freeman et al. 2003). Furthermore, Sp3 can also be modulated by methylation as previously described for Sp1. All these mechanisms contribute to make the Sp3 activity dependent on the cellular, nuclear, and gene promoter environments.

3.6.2.3 Sp4

Sp4 was cloned together with Sp3 and share most of the structural characteristics of Sp1. The transactivation ability of this protein can be attributed to the N-terminal glutamine-rich domain but it presents different properties and mechanism of transactivation compared with the other family members. One of the basic differences is that Sp4 is unable to act synergically as homo-dimer, but it can bind Sp1 and strongly enhance its activatory effect (Hagen, Dennig et al.

1995). At difference with Sp1 and Sp3, the expression of Sp4 is restricted to the central nervous system, particularly in cerebellum and hippocampus. In addition, neurons express higher concentration of Sp4 compared with astrocytes (Mao, Yang et al. 2007).

Sp4 knock-out mice show low rate of survivability and a delayed physical and sexual development (Supp, Witte et al. 1996). On the other hand, Sp4 hypomorphic mice, with an expression reduction of 95-98% compared to the wild-type mice, show an enhanced vacuolization of the grey matter in hippocampus. As consequence, Sp4 is also involved in spatial learning and memory and LTP mechanisms (Zhou, Long et al. 2005; Zhou, Nie et al. 2010). In addition, these mice also show a reduction in the expression of the NR1 subunit of NMDA receptor. Thus, the relationship of Sp4 with this pathway suggest that this protein might be important for all brain disorders characterized by glutamate miss-regulation, such as bipolar disorders or schizophrenia (Braff, Geyer et al. 2001; Pinacho, Saia et al. 2015). Moreover, this protein plays a key role in the development of cerebellar granule neurons (CGN) by regulating target genes, such as the neurotrophin 3, that exert a crucial activity in the dendritic remodelling and the maturation of these cells during the morphogenesis (Ramos, Valín et al. 2009).

Overall, given the limited expression of Sp4 in CNS, and its importance under several neurophysiological functions and pathophysiological disorders, it

emerges as an interesting TF that deserves further investigation. On the other hand, NCX2 isoforms, displays similar aspects of Sp4 that suggest a possible correlation between these two proteins. These aspects include: (1) the expression limited in both neurons and glia of CNS; (2) participation in neurotransmission; (3) participation in long-term learning and memory; (4) participation in neurological cerebral ischemia and psychiatric disorders.

3.6.3 Sterol Regulatory Element Binding Protein (SREBP)

SREBPs are a group of DNA binding protein classified in the superfamily of Basic Helix-Loop-Helix Leucine Zipper transcription factors (bHLH-Zip) (Bengoechea-Alonso and Ericsson 2007). Furthermore, SREBPs present a particular bHLH that differs from the other members of the superfamily in which a tyrosine residue is a substitute to an arginine of the classical structure. This mutation make SREBP able to bind both Sterol Regulatory Element (SRE) and a motif is known as Enhancer Box (E-Box).

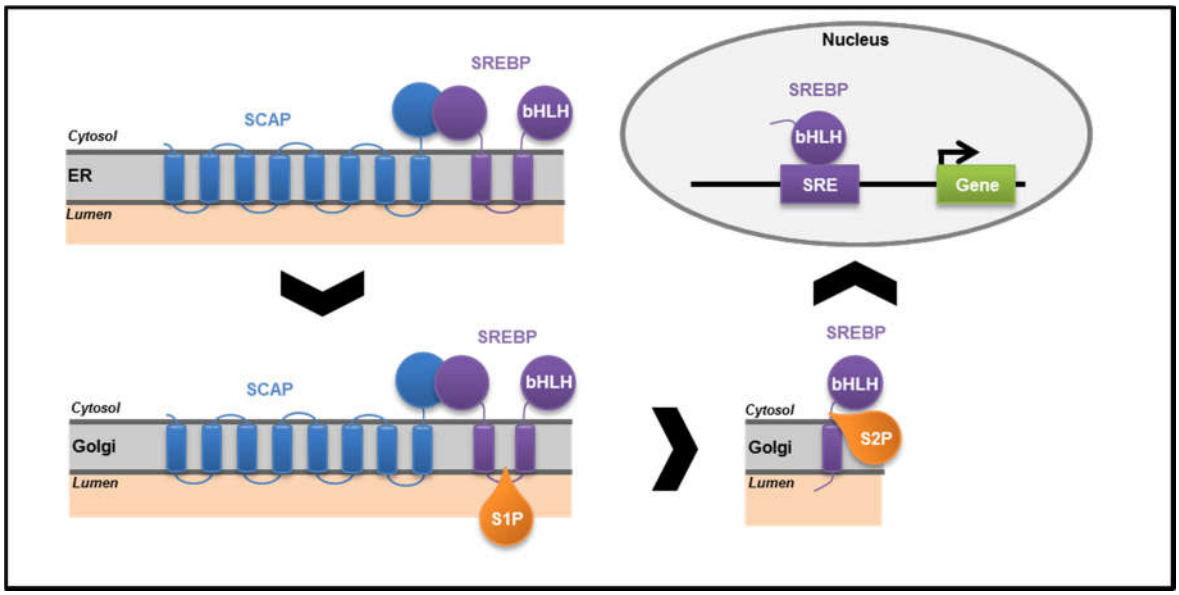


Figure 6 Fig6: Cartoon representing the post-tralational steps to activate SREBP: (1) SREBP (purple) is anchored to the ER membrane after its synthesis; (2) under appropriate stimulus SCAP (blue) mediates the cleavage by S1P Golgi protease (orange); (3) S2P Golgi protease (orange) release the active transcription factor migrating into the nucleus; (4) the active form of SREBP binds the target sequence SRE and exerts the regulatory effect on gene transcription

SREBPs family is basically composed by 2 isoforms, SREBP1 and SREBP2, with a restricted number of alternative-spliced forms (Hua 1993). SREBP1 gene presents two different transcription starting sites, encoding for SREBP1a and SREBP1c isoforms (Hua, Yokoyama et al. 1993). This latter is involved in adipogenesis in rodents, for this reason is also known as Adipocyte Determination and Differentiation Factor 1 (ADD1) in rodents (Eberlé, Hegarty et al. 2004). Beyond SREBP1a and SREBP1c variants, some other alternative spliced forms have been characterized, although they show less importance in the family (Nohturfft and Zhang 2009).

SREBPs are transmembrane TF localized on the nuclear membrane or on the endoplasmic reticulum. The overall structure of the family presents 5 domains: C- and N-terminus, DNA-binding site, two transmembrane segments. N-terminal region is the transactivation domain, whereas the C-terminal region is able to interact with other regulatory proteins. The DNA-binding domain is characterized by the bHLH-Zip sequence. The two transmembrane domains are connected by a cytoplasmic segment to anchor this protein to nuclear envelope or endoplasmic reticulum. Notably, the central part of the protein present two cleavage sites that are important for the release of the active form of the transcription factor. This activation requires the action of SREBP Cleavage-Activating Protein (SCAP) that mediates the cleavage by two different proteases, S1P and S2P.

Protein activation is dependent on several conditions, including cholesterol and triglyceride levels. In particular, when cholesterol concentrations are maintained under physiological levels, SREBP protein remains in an inactive form. By contrast, when cholesterol levels drop, SCAP triggers the translocation of the SREBP-SCAP complex to the Golgi to obtain the activation by S1P/S2P proteases. The released active form of SREBP moves into the nucleus where it regulates gene transcription (Fig. 6) (Gasic 1994).

SREBP is a weak gene activator but its activity is highly enhanced when it interacts with co-regulators (Magaña and Osborne 1996). In this regards, SREBP

reaches the highest transcriptional activity when the target promoter presents a binding sequence for co-regulators that is proximal to at least two consecutive SRE element (or SRE/E-box tandem). Most frequently, the co-regulators found to interact with SREB are Sp1, and Nuclear Factor Y (NFY). The binding motif on the first one must lay on a maximum distance of 10bp from the SRE sequence, whereas the NFY can range to 30 bp (Amemiya-Kudo, Shimano et al. 2000).

Despite principal SREBP target genes seems to be related to the cholesterol and fatty acids biosynthesis, in the last few years, ChiP experiments and RNA interference revealed a growing number of genes involved in other cellular processes than lipid metabolism, like cell cycle, proliferation and apoptosis (Reed, Charos et al. 2008; Rome, Lecomte et al. 2008). In fact, SREBP is highly expressed in neuron and glia where it participates in the myelination process (Camargo, Smit et al. 2009). In addition, SREBP1c mediates the transcription of polyunsaturated fatty acid (PUFA) regulators under pathophysiological conditions. In particular, SREBP1c can regulate stearoyl-CoA desaturase (SCD1), delta-5 desaturase (D5D) and (delta-6 desaturase D6D) that participate to the insulin-dependent control of PUFA metabolism. Abnormal regulation of this pathway is involved in neurodegenerative disease, such as Alzheimer's, in several neurological and behavioural disorders (Nakamura and Nara 2004), and in glioma (Ru, Hu et al. 2016). Moreover, SREBP translocation and activity are

impaired in Huntington's disease. This effect might be explained by two complementary hypothesis: (1) Huntingtin abnormalities may lead to the malfunction of the shuttling process (Ossareh-Nazari, Gwizdek et al. 2001); (2) SREBP might be misregulated by post-traslational mechanisms, such as SUMOylation, which are altered in Huntington's disease (Steffan, Agrawal et al. 2004).

3.7 Epigenetic Regulation

Epigenetics controls stably heritable phenotypes that does not result from changes in genomic DNA sequence. In fact, despite most of the cells in an organism share the same DNA, the expression of several genes might differ over the time by histone modification, DNA methylation, RNA interference and nucleosome positioning. Thanks to these marks, the regulation of target genes can occur without change the sequence of nucleotides. Since the embryogenesis, these mechanisms occur over the time and might endure even after cell division for the lifetime (Bird 2007). Moreover, environmental stimuli can modify these epigenetic mechanisms in the whole organism or in some cell groups (Görisch, Wachsmuth et al. 2005).

Histones are important for epigenetic mechanisms since thy bound the genomic DNA and cause chromatin packaging. Modifications of these proteins are able to tight or loosen DNA binding, and thus, they participate in the

regulation of gene promoter access for TFs. Histone modification involves almost 60 residues, most of which are located on the N-terminus tails. Modifications includes methylation, acetylation, citrullination, ADP-ribosylation phosphorylation, ubiquitination, and SUMOylation (Khorasanizadeh 2004; Kouzarides 2007; Christophorou, Castelo-Branco et al. 2014). Histone modifications produce different functional outcomes depending on the position where they occur along the gene. The number of modification and the modified residues are also important for the output effect. It was reported that tri-methylation of histone H3K4 marks an active gene promoter, whereas single methylation on the same residue is characteristic of distal regulatory regions. On the other hand, tri-methylation of H3K27 is peculiar of an inactive promoter (Bernstein, Meissner et al. 2007; Schneider and Grosschedl 2007).

Another common signalling tool to control gene expression is the DNA methylation. This epigenetic modification occurs at the cytosine base of a CpG, which is converted to a 5-methyl-cytosine by a DNA methyltransferases (DNMTs) enzyme. CpG residues are globally distributed in the entire genome and they are frequently grouped in sequences of 300-1,000 bp, known as CpG islands. For many years, methylation of these sequences was typically associated with gene silencing, however, in some cases, methylation plays a role in mediating gene expression (Suzuki and Bird 2008). A growing number of transcription factor have been found to bind sites that present or not a methylated

CpG sequence in the core of the binding site. These factors are able to give different output activities depending on the presence the dinucleotide in the binding motif. Thus, methylation regulates the interaction of a class of methyl-binding transcription factors with the DNA (Zhu, Wang et al. 2016). Interestingly, several studies are focusing their attention on the alteration of these pathways under some pathophysiological conditions.

All these epigenetic mechanisms are not independent, since they intertwine in a dynamic pattern of reaction that outcome to different cellular expression profile (Vaissière, Sawan et al. 2008). Condensed chromatin usually presents a combination of histone modifications and DNA methylation. It has been postulated that these two mechanisms can be influenced each other's (Nan, Ng et al. 1998; Reddington, Perricone et al. 2013). This may occur because DNMTs can be recruited by particular histone modification. Otherwise, methyl-CpG residues may indicate the locus that must be inactivated by DNA condensation (Ooi, Qiu et al. 2007; Cedar and Bergman 2009; Zhao, Rank et al. 2009). One of the canonical examples that show the importance of DNA methylation in gene expression is the inactivation of chromosome X during development.

Due to the critical role of epigenetic modifications, an amount of research has underlined the connection between dysregulations of these mechanisms and a wide range of diseases including lupus, muscular dystrophy, birth defects and

cancer. In particular, DNA methylation represents a key tool to suppress genes involved in tumor growth (Robertson 2005).

3.8 Strategies for identifying mammalian regulatory sequences

The analysis of regulatory motifs in genomic DNA acquires a particular relevance to understand the mechanisms that regulate gene expression. In the past, experimental determination of the sequences where the TFs bind DNA was not practical and efficient. However, the use of new techniques, including new generation sequencing, cDNA microarray, and chromatine immunoprecipitation (ChiP), offered a chance to discover new binding motifs, and their variants, extending the knowledge on the consensus sequence that a TF can bind.

One of the most common strategies to describe a TFBS is the position weight matrix (PWM) which is a probabilistic model that quantifyes the DNA binding preference of a TF. A PWM is generally learned from a collection of aligned DNA binding sites that come from sequencing techniques. Theoretically, the output of collected data can be formulated as a *maximum likelihood problem*. To solve this problem it can be supposed that sequencing of binding sites consist of random independent observations from a *product multinomial distribution*. In the distribution of all the reads, each entry is proportional to the observed count of its corresponding nucleotide at a corresponding position (Wilson and Chen 2007).

The first step is to collect a good number of sequencing for the predicted TF from novel read, a database or from literature (fig 7a). A way to describe this output is a *Consensus Sequence Model*: a motif based on the IUPAC nucleotide symbols for each position. However, as shown in the example reported in figure 7a and 7b, an output with a high varied sequence can influence the consensus output with the same weight of the others sequences, i.e. site 8 present only 50% of the conservation compared to the other sites.

A better way to describe the sequence output is to create a *Position Frequency Matrix* in which is reported the count of all the nucleotides in each position (Fig 7c). The frequency matrix is usually normalized and converted in a PWM (Fig 7d) with the following equations:

$$p(b,i) = \frac{f_{b,i} + s(b)}{N + \sum_{b' \in \{A,C;G;T\}} s(b')} \qquad W_{b,i} = log_2 \frac{p(b,i)}{p(b)}$$

$f_{b,i}$= counts of base b in position I; N= number of sites; $p(b,i)$= corrected probability of base b in position I; $s(b)$= pseudocount function; $p(b)$ = background probability of base b; $W_{b,i}$ = PWM value of base b in position i

In this way, a value can be attributed to each nucleotide of an unknown sequence. The sum of the values of all the positions can return as a *relative score* (fig 7e) compared to the score of the PWM (Wasserman and Sandelin 2004). The relative score count 1.00 when each position has the maximum value reported by the PWM.

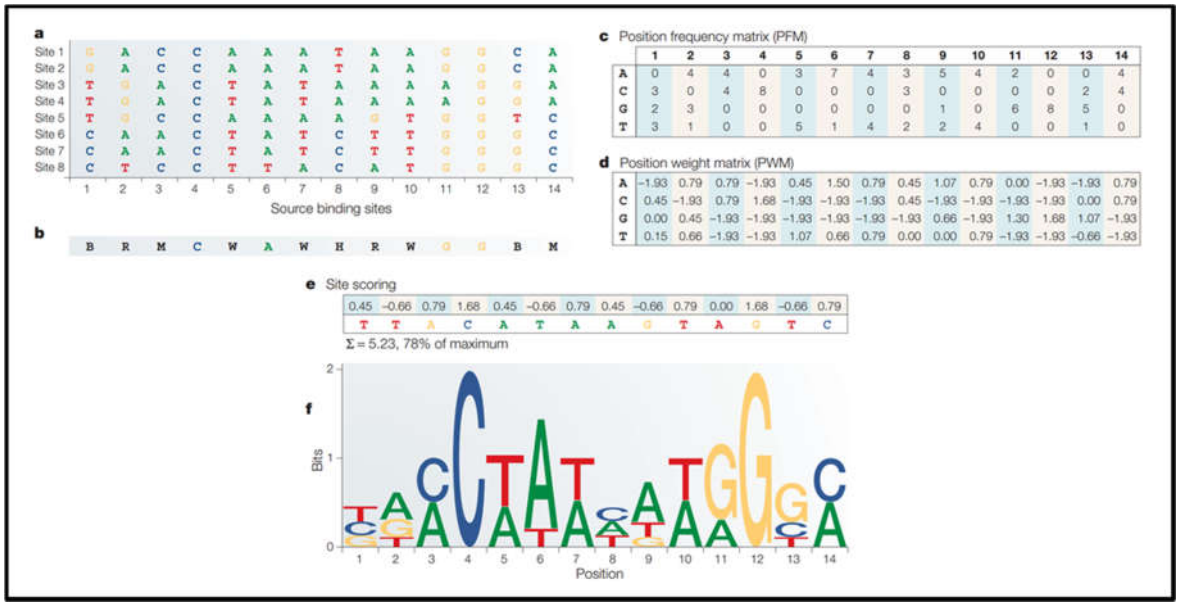


Figure 7 Building models for predicting transcription-factor binding sites: table from Wasserman and Sandelin (2004). (a) Data collected from literature and sequencing, (b) motif sequence derived from table a. (c) table displaying positional frequency matrix. (d) position weight matrix obtained from table c. (e) site scoring a representative unknown sequence. (f) Sequence logo that displays the weight of each nucleotide in the relative position.

To date, there is an increasing number of PWMs databases for many TF. TRANSFAC® (Matys, Fricke et al. 2003), JASPAR (Vlieghe, Sandelin et al. 2006) or MATINSPECTOR (Cartharius, Frech et al. 2005) also incorporate many internet-based software tools that allow to formulate a complete bioinformatic analysis of an unknown sequence.

These databases provide a strong likelihood for *in silico* analysis of a TFBS. However, some experimental discrepancy between the *in vivo* and *in silico* results can be observed. These discrepancies indicate that additional factors participates to the function of the regulatory sequences. Usually, two additional complementary observations regarding the characteristics of a promoter give substantial improvements in the prediction of functional binding sites.

Firstly, analysis of sequence conservation in the regulatory region can enhance the predictive specificity with matrix models. Secondly, cooperative interactions between TFs can contribute in gene regulation, and thus they must be considered in the computational algorithms to improve performance.

In this regard, some emerging user-orientated tools combine matrix-based site predictions with phylogenetic footprinting. In general, these tools require pairs of orthologous gene sequences. A database of PWM binding profiles, such as JASPAR, is used to predict binding sites within the conserved regions. This methods requires the most stringent parameters to predict TFBSs that are conserved among orthologous sequences (Wasserman and Sandelin 2004).

3.9 Concluding remarks

All these data suggest that that transcriptional regulation is a complex dynamic process. The interplay between the entire suite of core promoters, proximal regulatory elements, and distal regulatory elements, as well as their binding factors and cofactors contribute to the transcriptional output of a given promoter. Regulatory systems are also highly sensitive, since even single-nucleotide differences can result in a significant different effect on gene expression. Current endeavours aiming to annotate all the transcriptional regulatory elements in the human genome face considerable challenges. TFBSs are small, degenerate, often located distantly from the promoter upon which they

act, and are not always conserved through evolution. These properties make regulatory elements difficult to identify through computational means alone. Many experimental methods show binding of a transcription factor at a given site, but do not assess the functional significance of that binding. Functional assays that directly assess the regulatory capacity of a site are the best available tools, and the current challenge is to adapt these methods for their high-throughput usage to screen the entire human genome.

Chapter 4: Objectives

Since NCX2 antiporter exerts an important role under several physiological and pathophysiological conditions, the knowledge on its gene regulation might contribute to understand the etiopathogenetic of several neudegenerative and neurological diseases. In addition, the identification of mechanisms able to increase its expression and its activity might characterize new molecular targets useful for clinical prognosis and diagnosis, or might open new paths for the treatment of those diseases of CNS in which a dysregulation of Na+ and Ca2+ homeostasis occurs.

The identification of the NCX2 promoter will represent the first step to understand the molecular determinants that controls its transcription under physiological and pathophysiological conditions.

To this main aim, the present thesis was focused to:

- isolate the promoter of *slc8a2* gene encoding for NCX2
- characterize the minimal promoter of *slc8a2* gene encoding for NCX2
- identify several transcription factors involved in the regulation of NCX2 expression in CNS

- recognize the molecular determinants for the effect of transcription factors on NCX2 promoter
- find epigenetic mechanisms that might be involved in NCX2 expression in CNS

Chapter 5: Materials and Methods

5.1 In silico analysis

Rat intergenic region between *kptn* and *slc8a2* genes was analysed by comparing the putative transcription binding sites with those obtained from the orthologous region in human and mouse genome. In particular, we obtained the putative transcription binding sites for each species by using Genomatrix MatInspector (http://www.genomatix.de) and Jaspar (http://jaspar.genereg.net) websites with a threshold of 0.80 for both matrix similarity and core similarity scores. Afterwards, we selected those binding sites that are conserved among the species for the relative position and that present possible cofactors in their proximity.

5.2 Primer design and synthesis

In this work, to isolate SLC8A2 promoter we designed primers to amplify different regions reported in the following table:

Region	*Primer*	*Sequence*	*Amplified Length*
Whole intergenic Region	Forward	5'-AGCTATGAAGCCGTTCTCTGCTC-3'	5850
	Reverse	5'-GGATGTTGCTCAAAGGCACAGTAC-3'	
Region S1	Forward	5'-CATGCAAGCAAGGGAGGTGTC-3'	1072
	Reverse	5'-GTGGCTCAGTGGTTGAGCCC-3'	
Region S2	Forward	5'-ACACCTTCTGTGGGCTGTATAGCGG-3'	1258
	Reverse	5'-GTCTAACCCACAGCCTGCCCT-3'	
Region S3	Forward	5'-TCCTTCCTTCCCTACCTTGCC-3'	1476
	Reverse	5'-GCCCAGAGACAGACAGACAGACAT-3'	
Region S4	Forward	5'-AGCTATGAAGCCGTTCTCTGCTC-3'	1146
	Reverse	5'-AGCACTCCTACCTACAAGTGCACAG-3'	

Primers used to amplify candidate promoter regions were designed according to the melting temperature reported by Oligo-Calculator analysis tool, version 3.27 (http://biotools.nubic.northwestern.edu).

Primers for site-directed mutagenesis were designed with a Tm >78°C calculated according to manufacturer protocol (Agilent technologies, Italy). Briefly, we used the following formula:

$$Tm = 81.5 + 0.41(\%GC) - (675/N) - \%\ of\ Mismatch$$

where N is the primer length and %GC and % of mismatch are whole numbers. Primers were provided from Eurofins Genomics (Munich, Germany) in lyophilized form and eluted to obtain a stock concentration of 100 mM stored at

-20 °C and thaw in ice before use. See supplementary section for other primer sequences used in this work.

5.3 Polymerase chain reaction

PrimeSTAR GXL DNA Polymerase (Takara, cod. R050A) was used to obtain all PCR products for cloning experiments. According to the manufacturer protocol, the reaction mixture and thermocycler protocol used for the amplification contained:

Mix	
5X PrimeSTAR GLX Buffer	10 µl
dNTP Mixture (2.5 mM)	4 µl
Forward primer (10 pmol)	10 µl
Reverse primer (10 pmol)	10 µl
DNA Template (100ng/ul)	1 µl
PrimeSTAR GXL DNA Polymerase	1 µl
ddH_2O	14 µl

Thermocycler Protocol	
98 °C for 10 sec	30 Cycles
60 °C for 30 sec	
68 °C for 1 min/kb of amplicon	

5.4 Agarose gel electrophoresis

Electrophoresis was carried out in 1% agarose gel made in 1x Tris-acetate-EDTA buffer (40 mM Tris pH 7.6, 20 mM acetic acid, 1 mM EDTA) supplemented with ethidium bromide (250 ng/µl). PCR products were loaded on the gel with 1kb DNA ladder (Invitrogen) as a marker of molecular weight.

DNA was visualised using a UV transilluminator and photographed digitally using ChemiDoc™ XRS+ System with Image Lab™ Software (Bio-Rad).

5.5 DNA cloning in promoterless luciferase reporter vector

Genomic DNA was cloned in the pSC-b vector by using StrataClone Blunt PCR cloning kit (Agilent) by using the manufacturer protocol as illustrated in Fig. 8.

Briefly, PCR products were ligated with the StrataCLONE Blunt Vector Mix and added to an aliquot of STrataCLONE SOLO ultracompetent cells. Bacterial

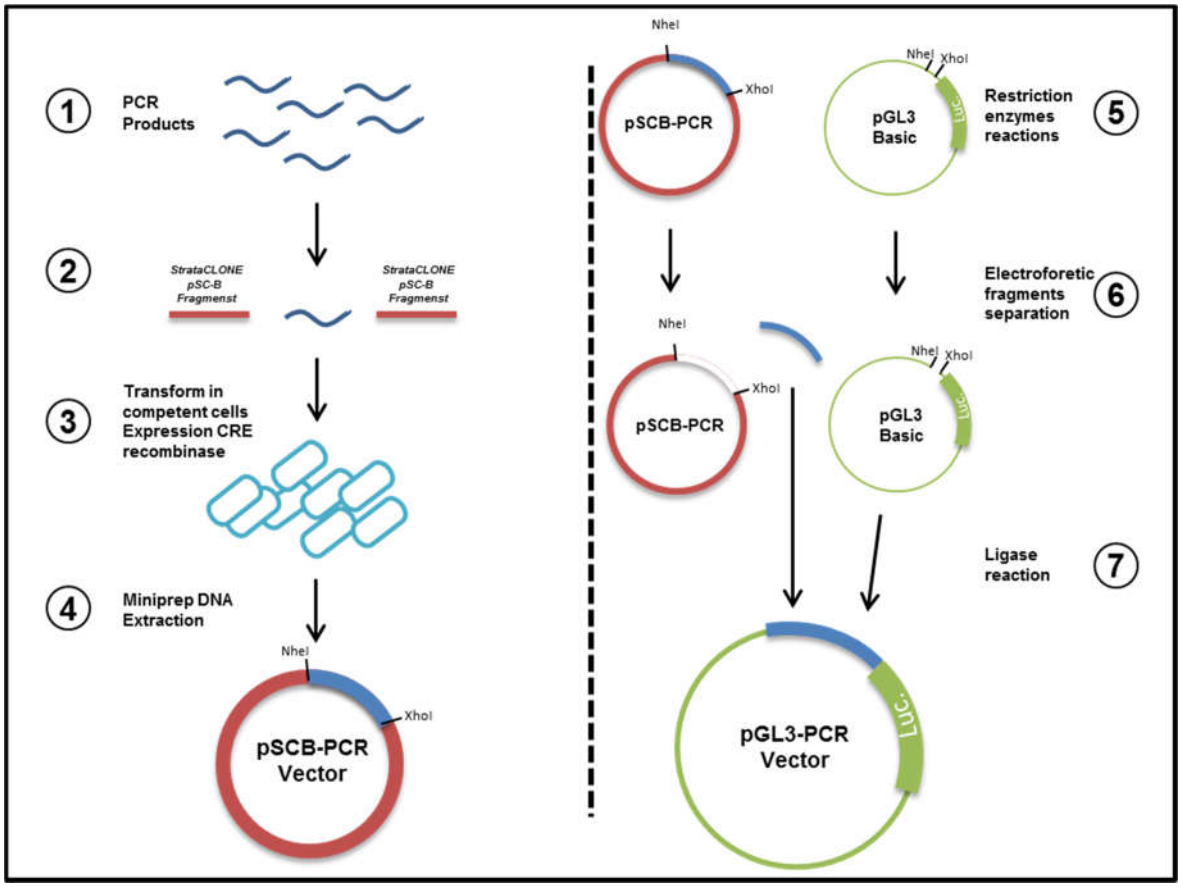


Figure 8 Schematic representation of insertion of PCR product in pGL3-basic vector.

cells were incubated in ice for 30 minutes, treated with a thermic shock (42°C for 30 seconds, 2 minutes in ice), resuspended in 300 µl of LB medium and then incubated for 1 hour at 37°C to express the antibiotic resistance of the internalized vector. Finally, cells were plated on LB-agar plates with ampicillin (50 µm/ml) and incubated overnight at 37°C. Single colonies were incubated in 5 ml of LB medium with ampicillin (50 µg/ml) overnight at 37°C.

DNA plasmid was extracted from bacteria cultures with PureYield™ Plasmid Maxi- and Mini-prep System (Promega, Italy). Positive vectors were identified by using enzymatic digestion followed by DNA sequencing (Microgem, Italy).

pSC-B-promoter and pGL3-Basic plasmids (Promega) were digested with the appropriate restriction enzymes. Fragments of the samples were electrophoretically separated on an agarose gel and the corresponding bands were extracted with StrataPrep DNA Gel Extraction Kit (Agilent) according to with the manufacturer protocol.

Promoter fragments and pGL3 fragment were ligated with T4 ligase enzyme and transformed into Subcloning Efficiency™ DH5α™ competent cells (Thermofisher, USA). Positive colonies were incubated in 400 ml of LB medium and plasmids were extracted from bacterial cells with PureYield™ Plasmid Maxi-Prep Kit (Promega)

5.6 Restriction Enzyme Digestion

All vectors were digested with appropriate restriction enzymes (Promega) to cut the insert. The recipe for the digestion reaction is shown in the following table:

Restriction Enzyme 1	0,5-2 µl
Restriction Enzyme 2	0,5-2 µl
Buffer 10X	5 µl
BSA 10X	5 µl
DNA Template	2-5 µg
H_2O	Up to 50 µl
Final Volume	50 µl

The reactions were performed by incubation at 37°C for 1 hour and then blocked in ice until loaded in agarose gel.

5.7 Site-direct Mutagenesis

All mutagenesis experiments were performed by using QuikChange Lightning Site-Direct Mutagenesis Kit (Agilent) according to manufacturing protocol by using the following PCR protocol

1. 95°C for 2 minutes
2. 95°C for 20 seconds
3. 60°C for 10 seconds
4. 68°C for 30 seconds
5. 68°C for 5 minutes

Steps from 2 to 4 were repeated 18 times

PCR products were incubated with Dpn I restriction enzyme for 5 minutes at 37°C. Subsequently, the transformation of Dpn-treated samples was performed into 45 ul XL10-Gold bacterial cells according to the above-mentioned protocol.

5.8 Isolation and retrotranscription of RNA

Total RNA was extracted from cells using TRIzol® (Thermofisher) reagent according to Maniatis protocol (Maniatis, Fritsch et al. 1982). Total RNA was eluted in 10 µl of nucleases free H_2O and 1/10 of the samples was loaded on a denaturating agarose gel to control and quantify RNA integrity. Two µg of total RNA were treated with DNase I (Sigma) and retrotranscribed by High-Capacity cDNA Reverse Transcription Kits (Life Technologies) according to the manufacturing protocol. Briefly, the following reaction mix was used to digest DNA

RNA	2 µg
DNase I (1 U/µL)	1 µl
DNase Buffer 10X	1 µl
H_2O	up to 10 µl

The mix was incubated for 15 minutes at room temperature and then was added 1 µl of the Stop Solution. Samples were heated at 70°C for 10 minutes and chilled in ice. Retrotranscription was performed by adding 10 µl of RT 2x mix and incubated at 25°C for 10 minutes, at 37°C for 2 hours and then at 85°C for 5 minutes.

5.9 Real-time PCR

Quantitative real-time PCR was carried out using 1/20 of the cDNAs with Fast SYBR Green60 Master Mix (Applied Biosystems). Samples were amplified simultaneously in triplicate in one assay as follows: 2 min @ 50º C, 10 min @ 95º C, 15 sec @ 95º C;1 min @ 60º C (7500 fast real-time PCR system, Applied Biosystems, Italy).

Steps 3 and 4 were repeated 40 and a single fluorescence measurement was obtained each cycle at step 4. qPCR data was collected using ABI Prism 7000 SDS software (Applied Biosystems). Differences in mRNA content between groups were calculated by using HPRT mRNA as the internal control. Data were analysed with Pfaffl method (Pfaffl 2001).

5.10 Cell Lines

Pheochromocytoma cell line (PC12) was maintained in culture in complete RPMI (500 ml RPMI (Invitrogen, Italy) 100 µg/ml penicillin 100 µg/ml

streptomycin 4 mM L-glutamine 10% of decomplemented fetal bovine serum, 5% of horse serum). When confluence reached 80%, cells were washed with PBS (0.2 g/l KCl, 8.0 g/l NaCl, 0.2 g/l KH2PO4, 1.44 g/l Na_2HPO_4, up to 500 ml of ddH_2O), detached from plates using Versene solution (0.2 mg/ml EDTA disodium salt in PBS) and a new sub-colony was established.

Neuroblastoma SH-SY5Y and Glioblastoma U87-MG cells were maintained in culture in complete DMEM medium (500 ml DMEM, 100 µg/ml penicillin, 100 µg/ml streptomycin, 4 mM L-glutamine. 10% of decomplemented fetal bovine serum), washed with PBS and detached from the plate using trypsin-EDTA solution.

Baby hamster kidney (BHK) cells were maintained in culture with medium containing 250 ml of DMEM, 250 ml of F12, 100 µg/ml penicillin, 100 µg/ml streptomycin, 4 mM L-glutamine, 10% of fetal bovine serum; washed with PBS and detached using Versene solution.

All cell lines were incubated in a humidified incubator at 37 ºC and 5 % CO_2.

Cells in their early passage were detached and counted. Then, 2 million cells were suspended in 1 ml of freezing solution (90% growth medium, 10% DMSO) and put in 1 ml cryogenic vials and stored at -80 ºC. For the thawing procedure, cells were rapidly thawed in a water bath at 37 ºC and transferred to a 15 ml tube with 7 ml of complete growth media. Then, cells were spun at 1,200 rpm (250 g) for 5 min, the media was removed and cells were suspended in the culture plate.

5.11 Eukaryotic transfections

All transfections were performed with Lipofectamine 2000 (Life technologies).

For the luciferase experiments, cells were plated in a 12-well plate at a concentration at planting time of 40,000 cells/well. After 24 h, cells were transfected with 1 µg of total plasmid DNA and 4 µl of Lipofectamine 2000 according to manufacturer protocol.

For the promoter individuation experiments 900 ng and 100 ng of pGL3-promoter and pRL-TK plasmids were used, respectively. 720 ng of pGL3-Promoter, 80ng of pRL-TK and 200ng of TF-cDNA expression plasmid were used in the analysis of the promoter regulation.

For the transfection of transcription factor cDNAs, cells were plated in 60 mm plate with a concentration of 150,000 cells/plate at plating time and transfected with 5 µg of plasmid DNA with 10 µl of Lipofectamine 2000.

For all the experiments, during transfection cell medium was substituted with reduced serum medium Opti-mem (Gibco) for 5 h. Afterwards, cells were incubated for 48 h with the culture medium until the experiments.

5.12 Gene Reporter luciferase Assay

Luciferase assay was performed with Dual-Luciferase® Reporter Assay System (Promega) according to the manufacturer protocol. Cells were collected

from the plate using provided lysis buffer, treated with two subsequent frost-thaw cycles (-80 °C for 40 minutes, 37 °C for 2 minutes) and centrifuged (12,000 g for 20 minutes at 4 °C). Bioluminescence measurement of *Pontius* luciferase was performed with Glomax® 20/20 luminometer by adding LAR substrate buffer to a protein sample. Next, *Renilla* luciferase luminescence was measured adding the Stop&GLOW buffer to inhibit *Photinus* luciferase and providing a correct substrate to *Renilla* luciferase. *Photinus* luminescence was normalized with *Renilla* luminescence and expressed as "Relative Luciferase Activity".

5.13 5'-Azacytidine (5-AZA-dC) administration

U87 cell line was treated once with 1, 5 or 10 μM at 5 h after plating time. mRNA was extracted from cells after 24h, 48h, 72h or 96h, and RT-PCR were performed using O-6-methylguanine-DNA methyltransferase (MGMT) mRNA expression as control (Cairns 2009). Finally, we used a treatment of 5 μM of 5Aza-dC for 72h in following experiments.

5.14 DNA bisulfite conversion and sequencing

In silico prediction of bisulfite-treated DNA sequence was performed with Bisulfite Primer Seeker tool (Zymoresearch) and primers were designed to amplify PCR products of 200-500bp. Bisulfite reaction was performed using

EpiJET Bisulfite Conversion Kit (Thermofisher Scientific) according to the manufacturer protocol. Converted genomic DNA was amplified by PCR reaction using Hot Start Taq DNA polymerase (Thermo Scientific). PCR products were electrophoretically separated on 1% agarose gel. Bands were extracted from the gel, purified and singly cloned in the pSC-B vector as described previously. pSC-B-fragment vectors were sequenced by Microgem s.r.l. (Naples, Italy)

5.16 Statistical Analysis

Data are expressed as mean (S.E.M.). Statistical comparisons between control and treated experimental groups were performed using the one-way analysis of variance, followed by Turkey test. Statistical comparisons of methylation between two cell lines were performed using χ^2 test. All tests were performed using GraphPad Prism (GraphPad Software, USA) and considered statistically significant with $p<0.05$.

Chapter 6: Results

6.1 Overview of the gene locus

Isoform 2 of sodium/calcium exchanger is encoded by *Slc8a2* gene. The locus in which NCX2 gene is located is highly conserved among different species. It includes five genes, *Napa*, *Kptn*, *Slc8a2*, *Meis3,* and *Dhx3,* with the same position and orientation in different mammal species (Fig. 9A).

We designed and analyzed the whole intergenic region and four smaller subregions to investigate the presence of cis-acting elements which localization can indicate a core promoter sequence between *Kptn* and *Slc8A2* genes. In addition, we compared our finding with the analysis of orthologous sequences in human and mouse DNA. Sequences alignment among these three species revealed that human exon 1 consist of a 379 bp sequence whereas, rat and mouse exon sequences are shorter (91 bp and 74 bp respectively) but highly conserved if compared to human sequence. *In silico* analysis revelaed the presence of two highly predictive initiator motifs. The first is located at the 5' end of the rat and mouse exon 1, whereas the second is located ~300 bp upstream of the first one. In addition, the second motif is located in a region of rat and mouse DNA corresponding to the beginning of exon 1 in human orthologous sequence (Fig. 9B).

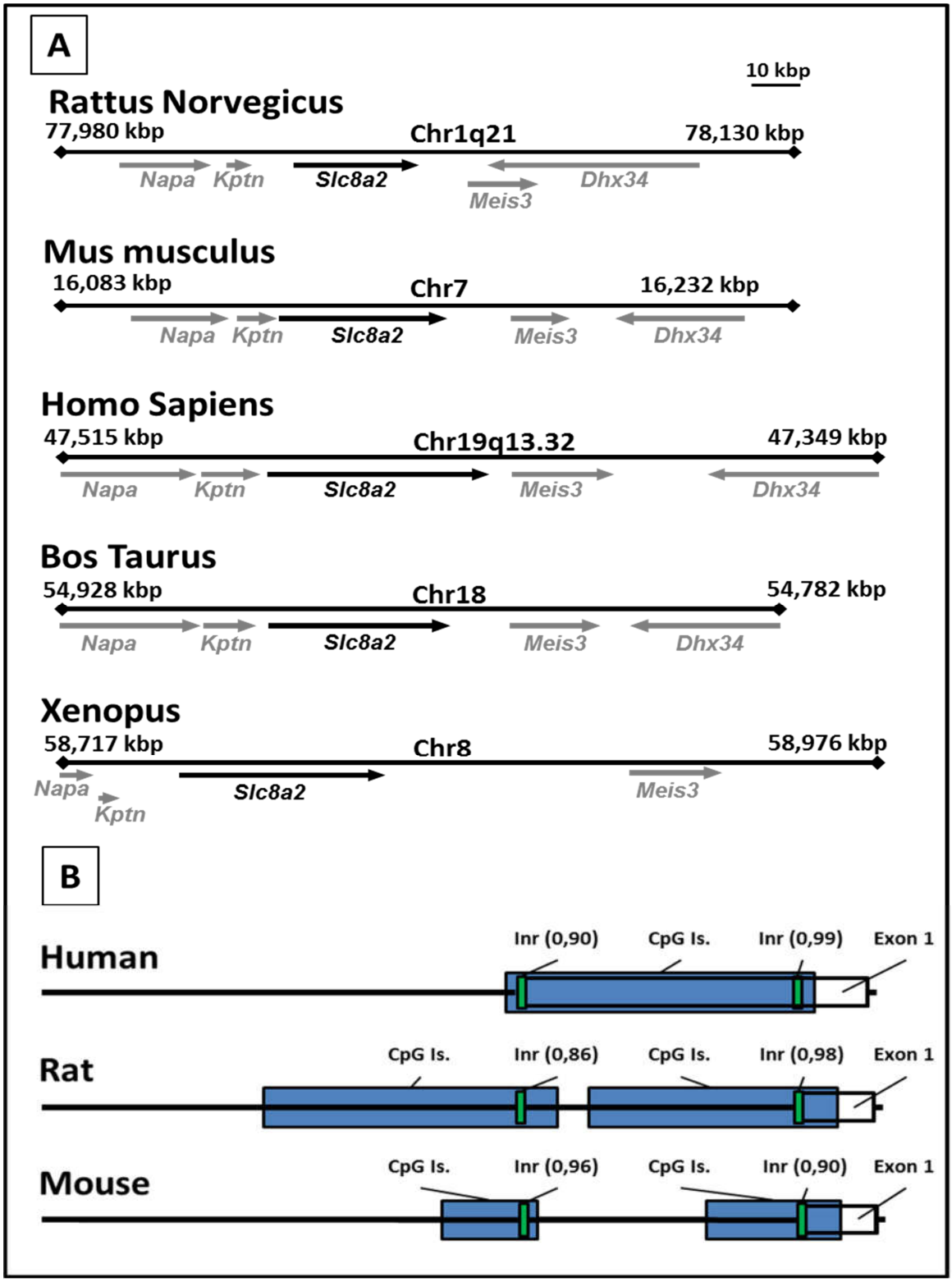


Figure 9 Relative position of the neighbour genes of NCX2 and putative promoter alignment: (A) Schematic representation of *Slc8a2* locus conserved in several mammals and lower species. (B) Schematic representation display the promoter features of *Slc8a2* 5'UTR region in human, rat and mouse. White boxes represent exon alignment, blue boxes represent CpG island and in green were displayed the Inr motifs conserved in all the tree sequences.

We found that 80% of the predicted GC-boxes are localized in the region upstream the exon 1. In rat, these motifs grouped in two CpG island, one of 364 nucleotides with a CG content of 61% and a second most dense island of 627 nucleotides with a CG content of 66% localized upstream the identified Inr (Supplementary Table S2, S3, S4).

6.2 Determination of Core Promoter Activity

All promoter constructs, including the intergenic region between *kptn* and *slc8a2* genes (-1929/+3912 bp), and its subregions S1(+2957/+3912 bp), S2 (+1295/2560 bp), S3(-709/+595 bp) and S4(-1929/-850 bp), were singly transfected in PC12 cells and bioluminescence was evaluated 48 h later. In all experiments, a pRL-TK vector carrying *Renilla* luciferase was cotransfected and used as internal control to normalize transfection efficiency and cell number variability. In addition, a reporter vector containing the Simian virus 40 (SV40) promoter (pI-S0) was used as positive control, whereas the promoterless empty vector pGL3-Basic was used as negative control to detect the background expression level of the luciferase reporter assay. Promoter activity for all constructs was expressed as arbitrary units.

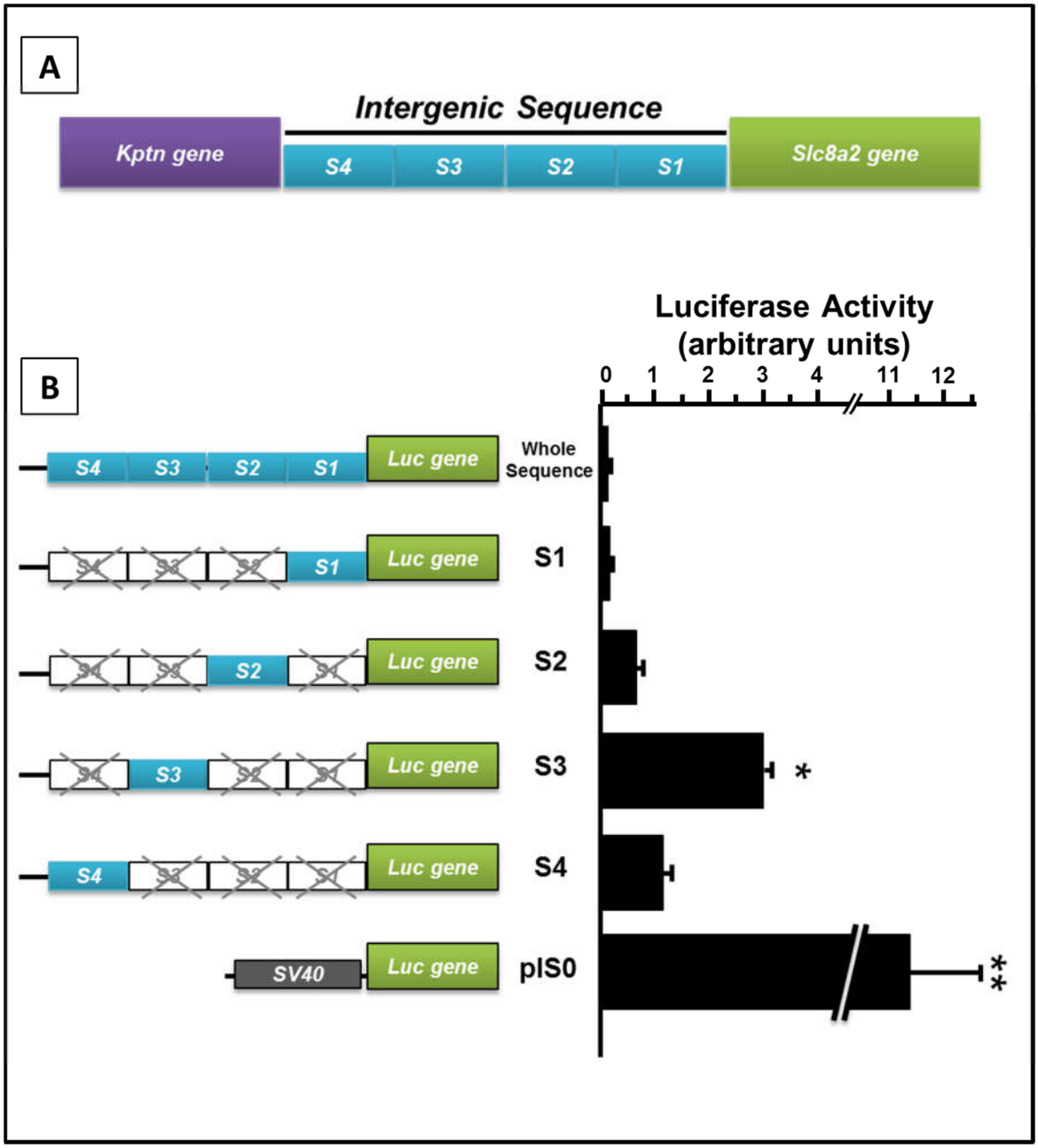


Figure 10 Identification of NCX2 promoter: (A) Schematic representation of the intergenic region between *Kptn* and *Slc8a2* genes in rattus norvegicus genome. S1, S2, S3, and S4 subregions are represented in cyan. (B) Transcriptional activity of several promoter regions of *slc8a2* gene measured as bioluminescence in pGL3basic plasmid. *, $p<0.05$ vs background signal (empty vector); **, $p<0.05$ vs S3 group.

Transfection of the luciferase constructs carrying the whole intergenic sequence (-1929/+3912), the subregions S1(+2957/+3912), S2(+1295/+2560) and S4(-1929/-850) displayed a weak transcriptional activity, whereas the construct including S3(-709/+595) sequence as promoter displayed a significant luciferase activity (Fig. 10B).

6.3 Comparison of promoter activities in several species

The promoter activity of the region S3(-709/+595) was investigated in several rodents and human cell lines expressing or not expressing the endogenous NCX2. In particular, both PC12 and SHSY cell lines showed a marked amount of NCX2 mRNA normalized for HPRT expression, whereas BHK and U87 cells did not show a significant signal of NCX2 mRNA as compared to the background signal (Fig. 11A). The transfection of pGL3-S3(-709/+595) showed a significant increase in luciferase activity as compared with the empty vector pGL3-basic in both cell lines expressing NCX2, PC12 and SHSY (Fig. 11B). On the other hand, transfection of pGL3-S3(-709/+595) did not show significant increase in luciferase activity in both cell lines that do not express the endogenous NCX2, BHK and U87 (Fig 11B). In addition, transfection of constructs including S3 subregions in both PC12 and SHSY cells showed an increase in luciferase activity from the whole S3 sequence until the subregion S3(+113/+595) (Fig 12). By contrast, the transfection of the plasmid carrying the

subregion +424/+595 of S3 dropped the transcriptional enhancement in both cell lines (Fig 12).

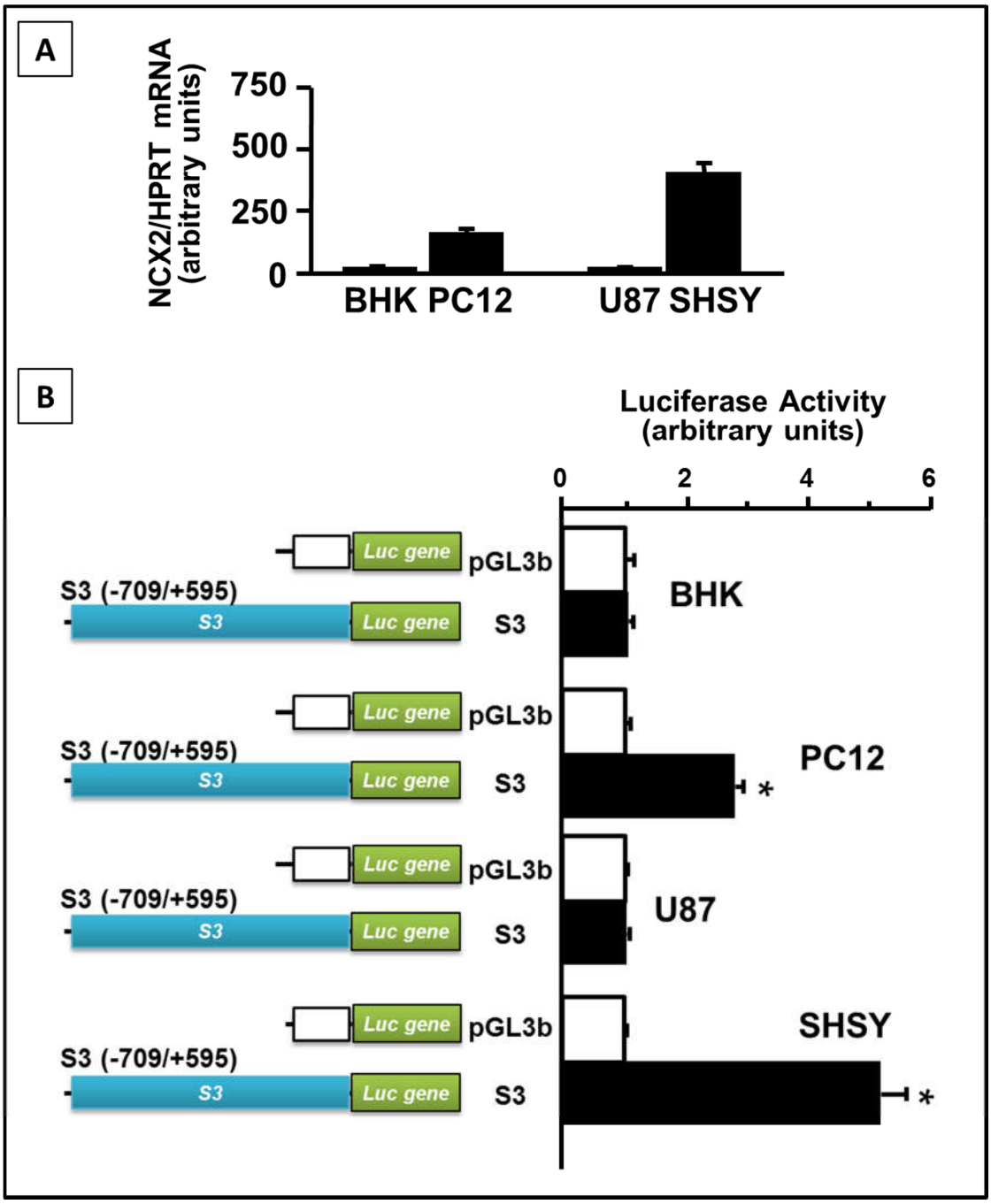


Figure 11 **Activity of NCX2 promoter in several cell lines:** (A) level of mRNA expression of NCX2 in several rodent and human cell lines. Data are normalized by HPRT expression level. (B) Enhancer activity of region S3 in cells expressing or not expressing NCX2. Data are reported luciferase activity in arbitrary units. *, $p<0.05$ vs respective empty pGL3basic control group.

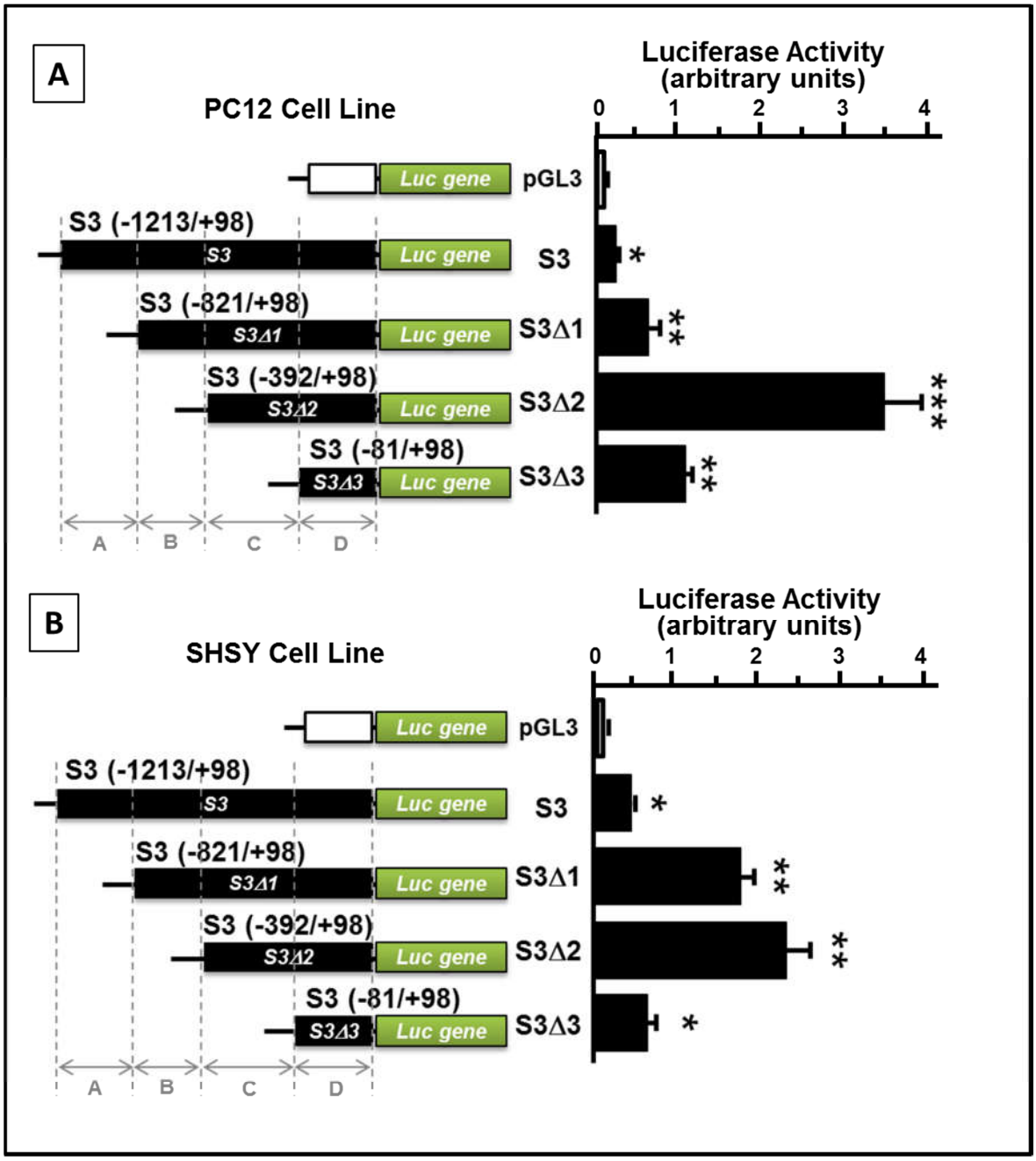


Figure 12 Luciferase activity of NCX2 promoter subregions in PC12 and SHSY cell lines: Subregion of NCX2 were transfected in PC12 (A) and in SHSY cell lines (B). *, $p<0.05$ vs pGL3basic group; **, $p<0.05$ vs S3 group, ***, $p<0.05$ vs S3Δ1 and S3Δ3 groups.

6.4 Analysis of cis-acting regulatory elements of S3 region

In silico analysis of rat S3 region revealed 788 putative transcription factor binding sequences with a threshold score above 80% (supplementary table 5) and 40 putative CpG methylation sites. In addition, these binding sites of rat S3 were compared with those obtained in the ortholog regions of mouse and human genome looking for conserved consensus sequences. In addition, we also considered the relative position of the consensus sequences from the transcription start site as a filter (supplementary table 5).

Epigenetic analysis revealed that both CpG islands of S3 region are methylated under control conditions in the PC12 cell line. In addition, these CpG islands were also found methylated in L6 cell line that does not express NCX2. However, the frequency of methylation was higher on CpG sites 4, 5, 29 and 33, and was lower on CpG sites 1, 23, 24 and 25, in PC12 cells as compared with L6 cells (Fig. 13B)

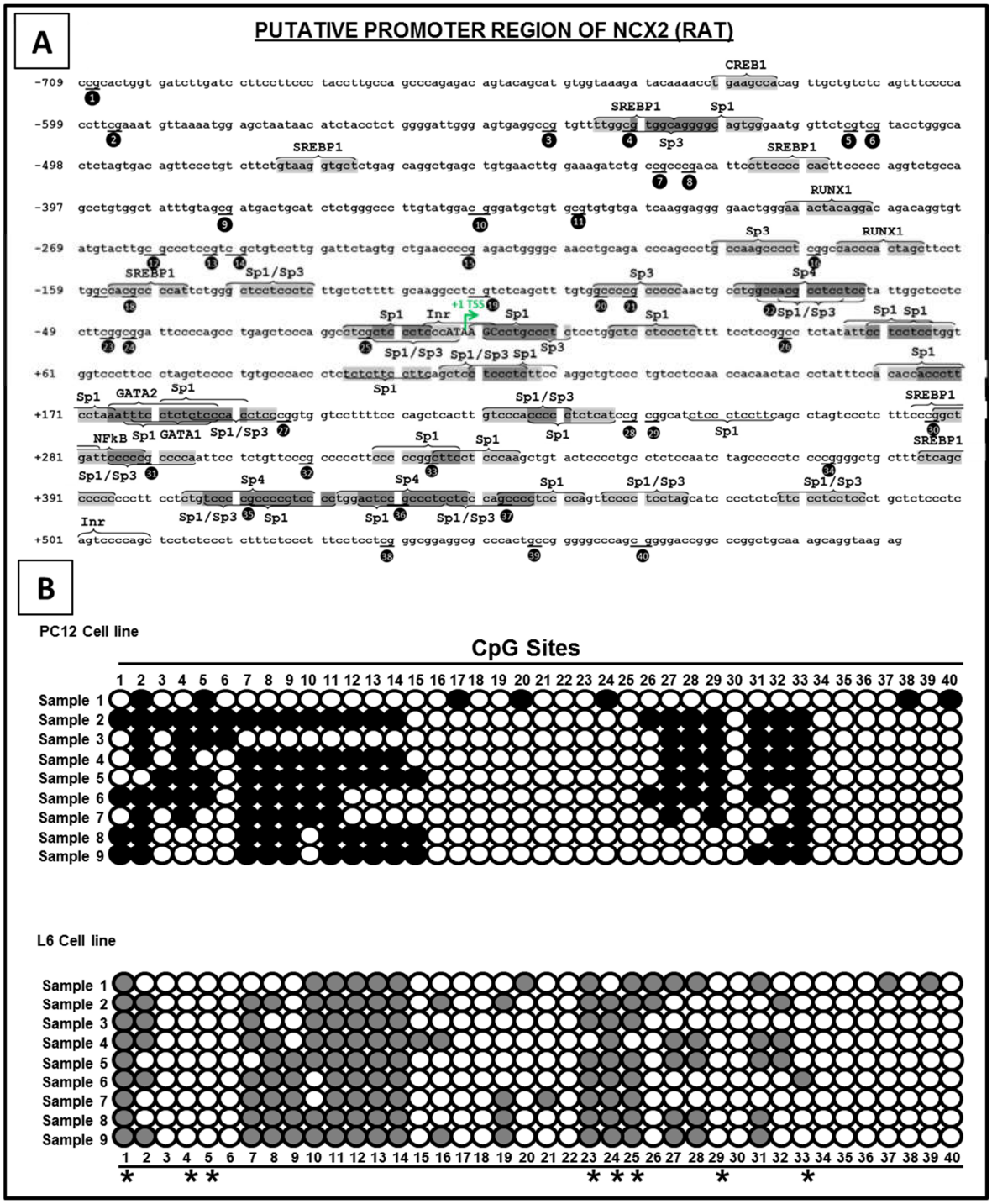


Figure 13. NCX2 promoter sequence and methylation status: (A) S3 sequence of rat *slc8a2* gene with several conserved consensus sequences for the transcription factors Sp1/3/4, GATA1/2, RUNX1, SREBP1, NFkB, CREB and the transcription initiator sequence Inr. Numbers in black circles indicate putative methylation sites. TSS is for transcription starting site. (B) Methylation status of CpG sites in PC12 and L6 cells, expressing and non-expressing NCX2, respectively. Numbers indicate the relative site in the sequence of panel A. Black-filled circle indicates methylated sites, open circle indicates not-methylated sites. *, $p<0.05$ vs respective site in PC12 cells

6.5 Effects of several transcription factors on NCX2 putative promoter

Selected transcription factors were selected to assess their effects on the transcriptional activity of the identified promoter region of NCX2. In particular, we selected several transcription factors on the basis of:

- The presence of almost one putative binding site on the promoter region with a high identity score on the known matrix (MatInspector, Jasper).
- The presence of almost one putative binding site conserved among the three considered species, rat, mouse, and human
- The expression in CNS

Results showed at least 9 transcription factor that might regulate the activity of the identified promoter of NCX2. Among these, the transfection of CREB1a, Sp1, or Sp4 enhanced the bioluminescence activity of pGL3-S3 in PC12 cells, whereas SREBP1 transfection abolished the luciferase activity of pGL3-S3 in same cells (Fig 14). On the other hand, the transfection of Sp3, RUNX1, Gata1/2, or NFkB1, although possess one or more putative conserved binding sites, did not show significant regulation of the identified promoter under our control conditions as compared with the empty vector of the transcription factors (Fig.14).

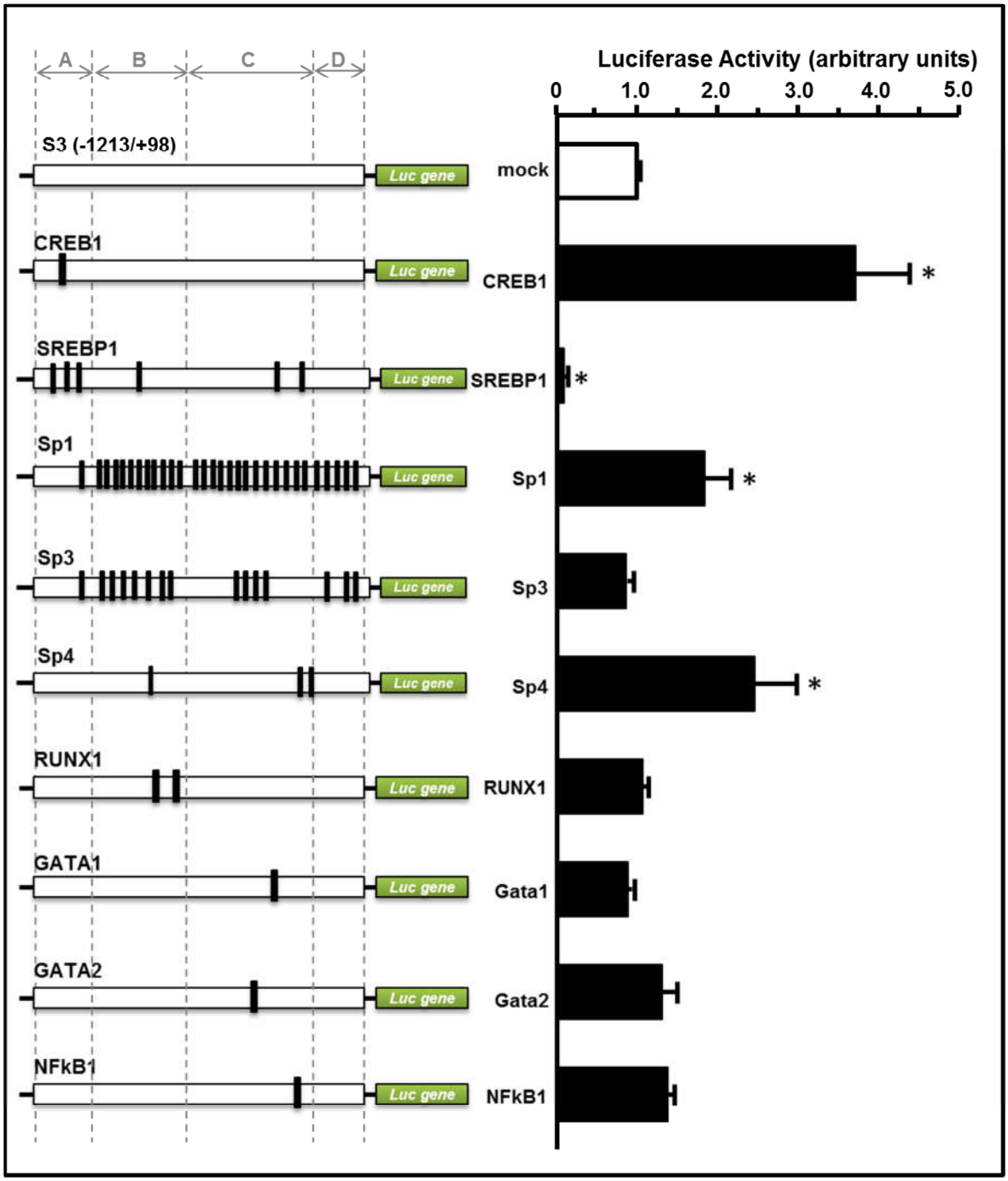


Figure 14 Effect of transcription factors on the identified promoter of NCX2: pGL3-S3 was cotransfected with single transcription factor in PC12 cell line. Luciferase activity was expressed as arbitrary units and normalized for the basal activity of pGL3-S3+mock. On the left, there is a representation of positions of binding sites (black box) for the transcription factor transfected. *, $p<0.05$ vs pGL3-S3+mock control group

6.6 CREB binding sites

The putative promoter of NCX2 cloned in pGL3 vector, presented one conserved binding site in the region A (-709/-319). The transfection of CREB1a increased more than 4 folds the bioluminescence activity of pGL3-S3 in PC12 cells under control conditions as compared with the transfection of empty vector. However, the transfection CREB1 did not affect the luciferase activity of the deletion-mutated pGL3-S3(-319/+595) and the deletion-mutated pGL3-S3(+113/+595), two deletion mutants that lack the region A or both A and B, respectively (Fig. 15).

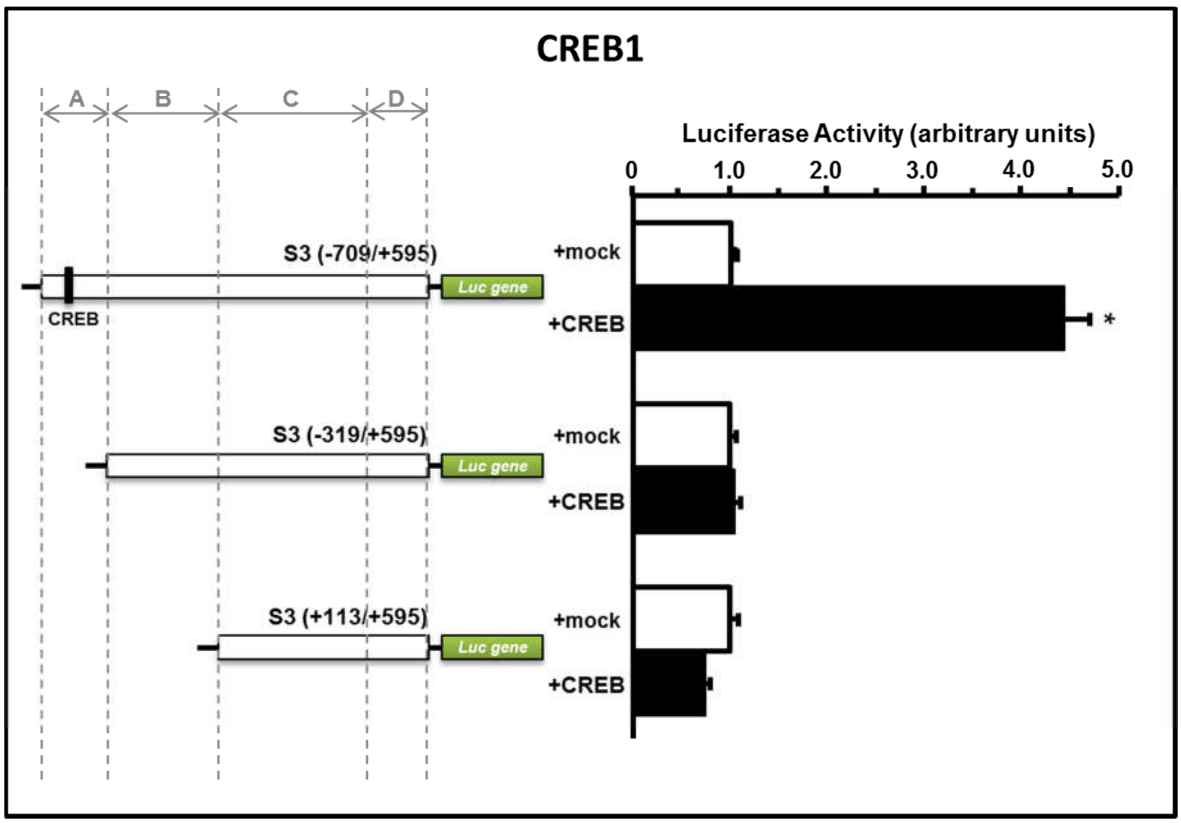


Figure 15. Activity of CREB on the identified promoter of NCX2: CREB presents one conserved putative binding site (black box). Luciferase activity was measured in PC12 cells and expressed as arbitrary units and normalized for the basal activity of the respective pGL3-S3+mock group. *, $p<0.05$ vs respective pGL3-S3+mock control group

6.7 Sp1 binding sites

Putative Sp1 binding sites are spread along the whole S3 promoter. In particular, one binding site is located in the region A, whereas 12 and 13 binding sites were located in the region B and C, respectively. Region D contains the last 5 conserved putative binding sites identified by *in silico* analysis.

The transfection of Sp1 increased 2 folds the bioluminescence activity of pGL3-S3 in PC12 cells under control conditions as compared with the cotransfection of pGL3-S3b and the empty vector of transcription factors (mock). The removal, by site-directed mutagenesis, of the region A or both A and B in pGL3-S3 construct did not affect the stimulatory activity of Sp1. By contrast, the removal of a wider region, that includes A, B and C, prevented the Sp1-mediated enhancement under same conditions. Furthermore, the disruption of the Sp1 site at +403 bp in pGL3-S3(+113/+595), which present the highest score in the region C, prevented the stimulatory effect exerted by Sp1 on the promoter (Fig 16). By contrast, the disruption of the second most probable Sp1 binding site at +319 bp in pGL3-S3(+113/+595), did not hamper the stimulatory effect of the transcriptional factor in PC12 cells under control conditions (Fig. 16).

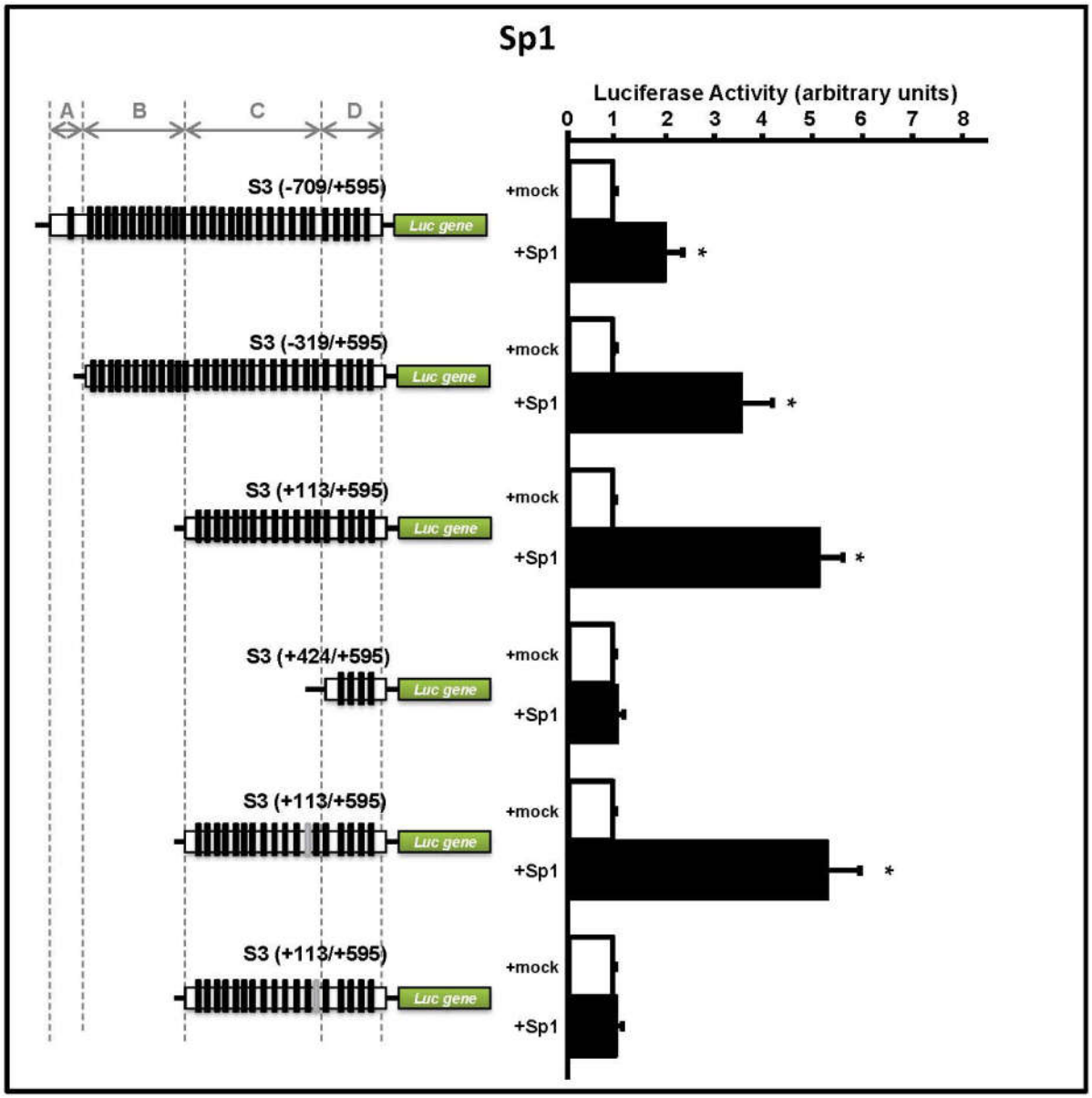


Figure 16 Activity of Sp1 on the identified promoter of NCX2: Sp1 presents several conserved putative binding sites on S3 (black box). Grey boxes indicate the disrupted binding sites in S3 promoter. Luciferase activity was measured in PC12 cells and expressed as arbitrary units and normalized for the basal activity of the respective pGL3-S3+mock group. *, $p<0.05$ vs respective pGL3-S3+mock control group

6.8 Sp4 binding sites

S3 region displayed 3 conserved binding sites for Sp4 located in the region B, C and D. Sp4 transfection increased 2 folds the transcriptional activity of pGL3-S3 in PC12 cells under control conditions as measured by luciferase assay. The removal, by site-directed mutagenesis, of either region A –S3(-319/+595)–, or both A and B –S3(+113/+595)–, in pGL3-S3 construct did not prevent the stimulatory activity of Sp4. By contrast, the removal of either region ranging from +376 to +465 in the deletion mutant pGL3-S3(-319/+595) or region ranging from +322 to +423 of the deletion mutant pGL3-S3(+113/+595), prevented the Sp4-mediated enhancement under same conditions. Furthermore, the disruption of the site at +403 bp of pGL3-S3(+113/+595), which presented the highest score in the region C, prevented the stimulatory effect mediated by Sp4 transfection (Fig. 17).

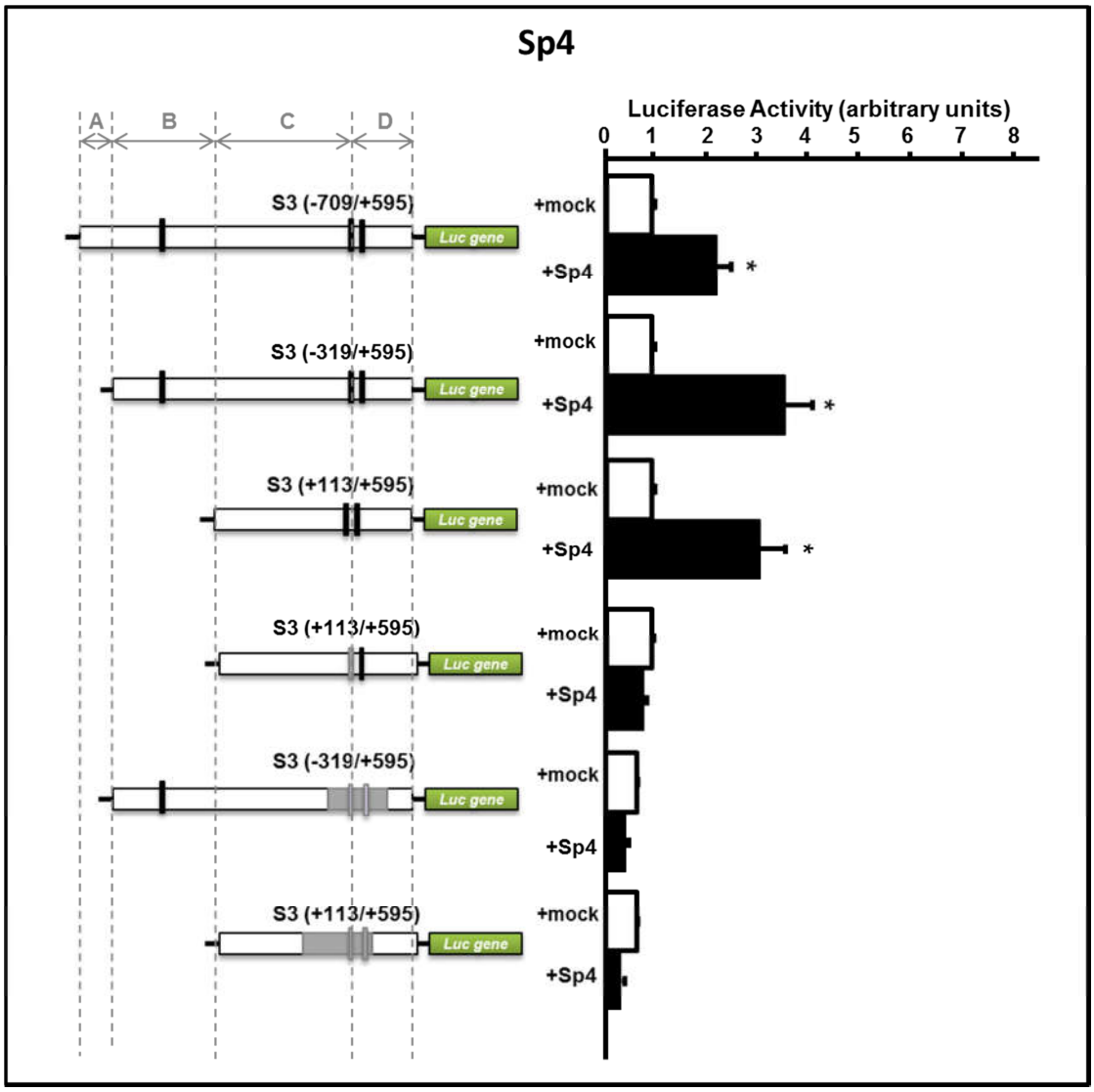


Figure 17 Activity of Sp4 on the identified promoter of NCX2: Sp4 presents three putative binding sites on S3 (black box). Grey boxes indicate the deleted sequences or the disrupted binding site for Sp4 in S3 promoter. Luciferase activity was measured in PC12 cells and expressed as arbitrary units and normalized for the basal activity of the respective pGL3-S3+mock group. *, $p<0.05$ vs respective pGL3-S3+mock control group.

6.9 SREBP binding sites

The putative promoter of NCX2 showed 6 conserved binding sites for SREBF1 transcription factor, 3 in the region A, 1 in the region B, and 2 in the region C. Transfection of SREBP1 markedly decreased the luciferase activity of both promoters, the entire S3 sequence and its deletion mutant pGL3-S3(+113/+595) (Fig. 18). This latter mutant lacks a region including A and B. On the other hand, SREBP-1 transfection failed to downregulate the bioluminescence of the deletion mutant pGL3-S3(+424/+595) construct under the same experimental conditions. In addition, the removal of a region from +322 to +423 of pGL3-S3(+113/+595), containing one of the two conserved SREBP1 binding sites, did not prevent the downregulation by the transfection of the relative transcription factor.

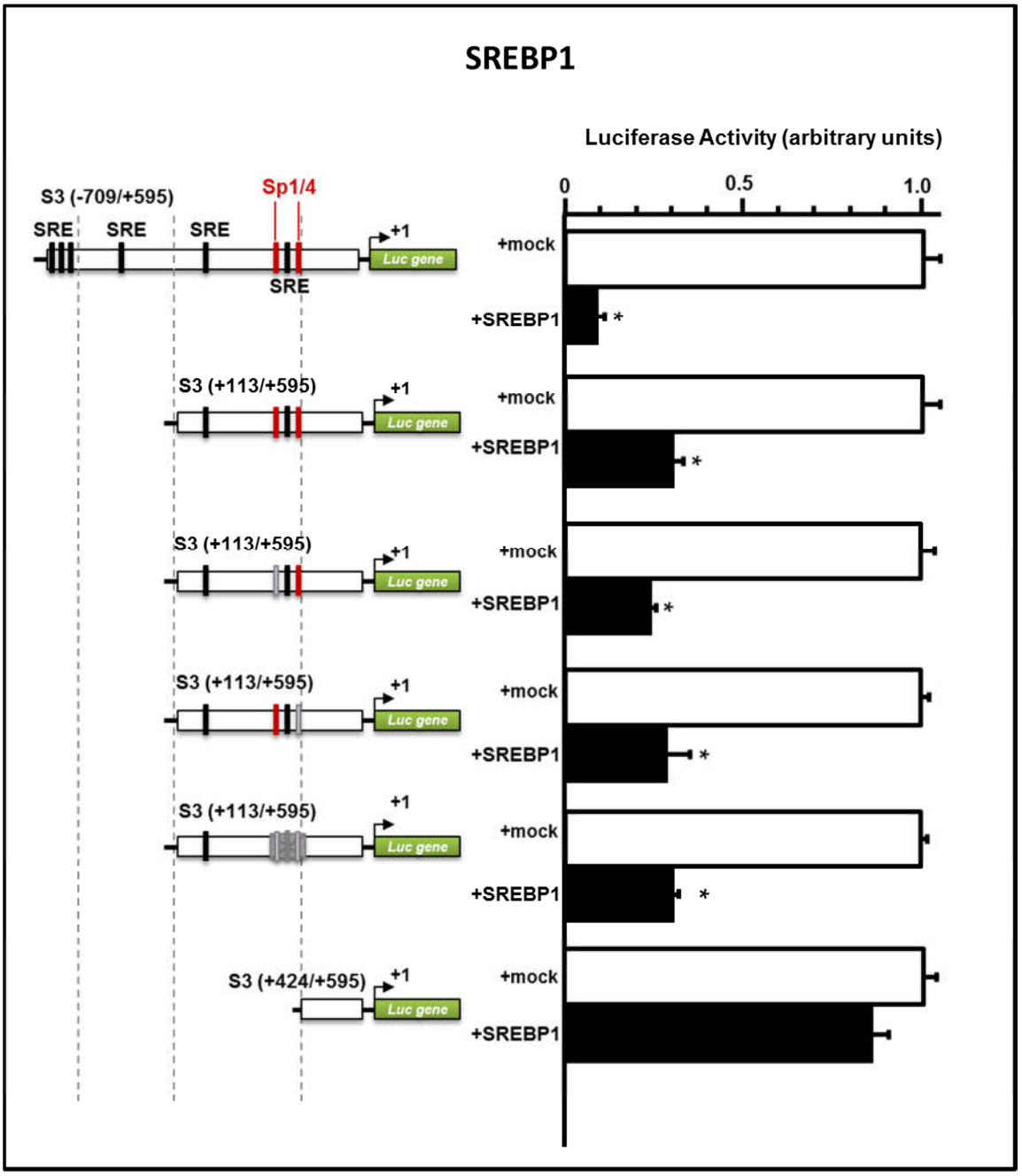


Figure 18 Activity of SREBP1 on the identified promoter of NCX2: SREBP1 present six putative binding sites on S3 (black box). Grey boxes indicate the deleted sequences or the disrupted binding site for Sp4 in S3 promoter. Red boxes indicate Sp1/4 binding sites that can represent putative co-regulators of SREBP1. Luciferase activity was measured in PC12 cells and expressed as arbitrary units and normalized for the basal activity of the respective pGL3-S3+mock group. *, $p<0.05$ vs respective pGL3-S3+mock control group.

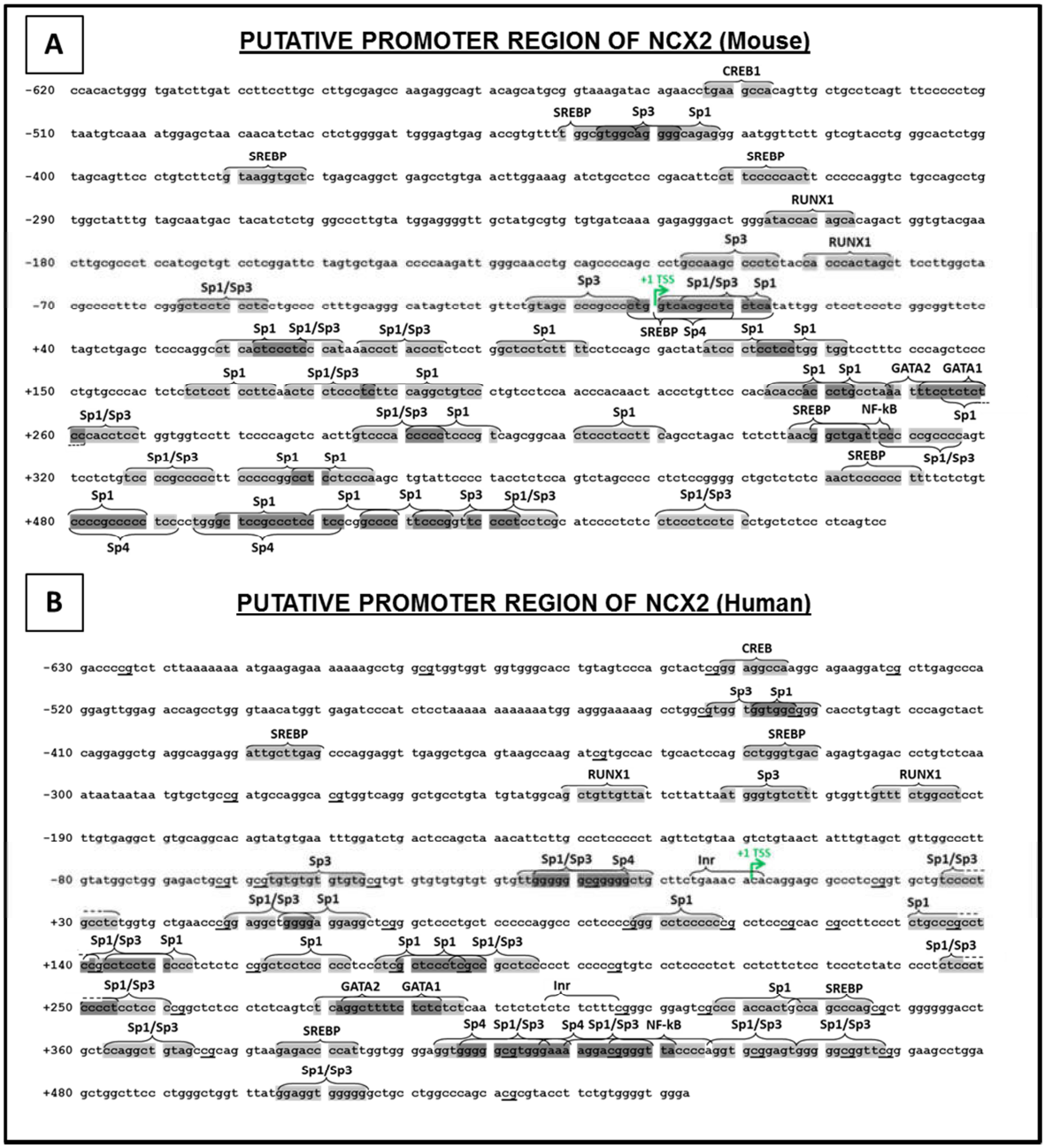


Figure 19 Ortholog sequences of mouse (in A) and human (in B) of the identified rat S3 region. several conserved consensus sequences for the transcription factors Sp1/3/4, GATA1/2, RUNX1, SREBP1, NFkB, CREB and the transcription initiator sequence Inr are showed in both sequences. TSS is for transcription starting site.

6.10 Role of Methylation in NCX2 expression

The effects of CREB1, Sp1 and Sp4 transfection on the transcription activity of the identified NCX2 promoter was also evaluated in U87 cells that do not express this exchanger under control conditions. The transfection of each above-mentioned transcription factors failed to induce a significant luciferase activity of rat pGL3-S3 in U87 cell line under control conditions (Fig. 20A). On the other hand, 72h and 96h of 5 μM 5'-Azacytidine, which prevents CpG methylation, induced a transcription of NCX2 mRNA in U87 cell line as measured by real-time PCR (Fig. 20B).

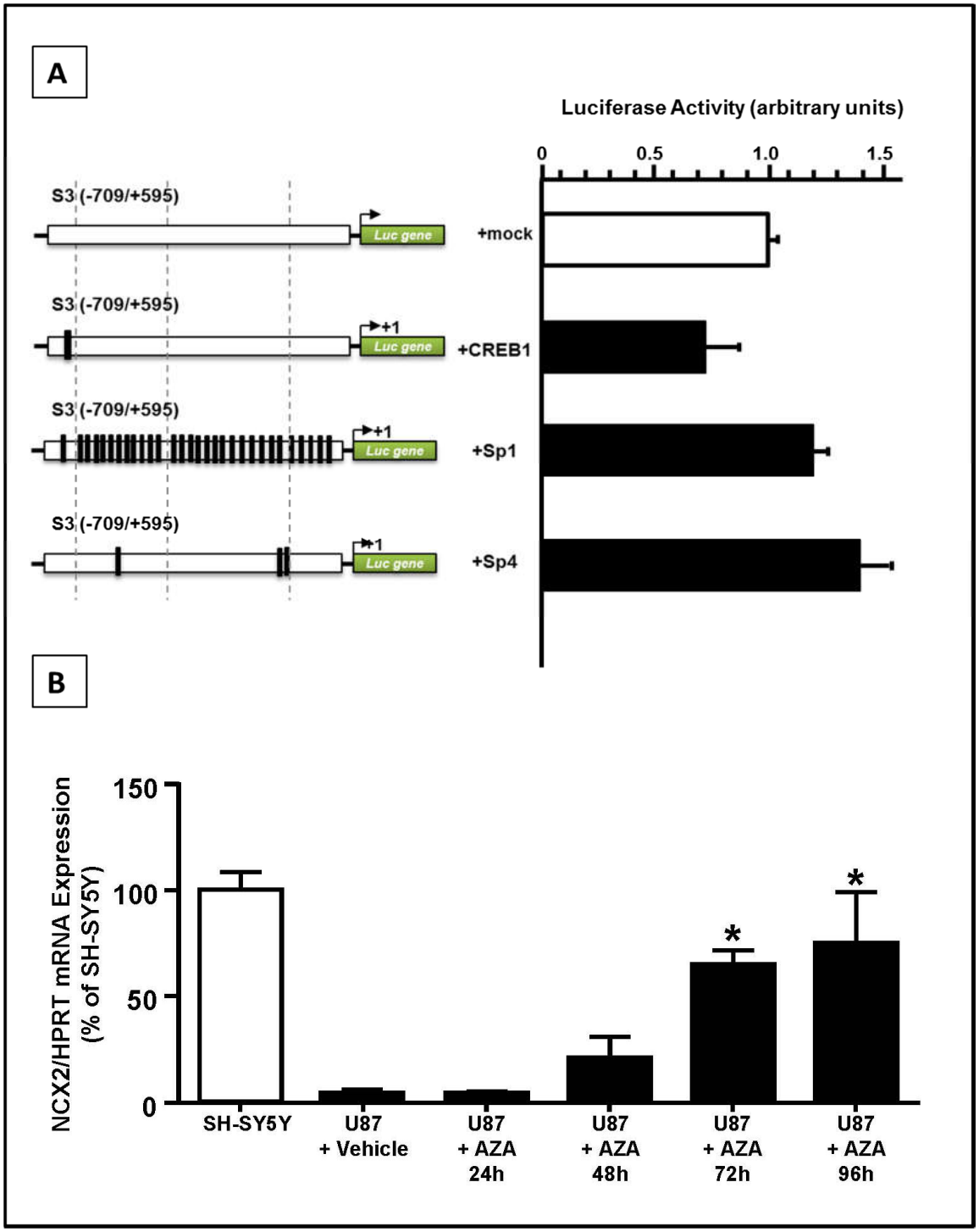


Figure 20. (A) Effect of CREB, Sp1, Sp4 on the identified promoter of NCX2. Luciferase activity was measured in U87 cells and expressed as arbitrary units and normalized for the basal activity of the respective pGL3-S3+mock group.
(B) Effect of 5-Azacytidine (AZA) on NCX2 mRNA levels. NCX2 mRNA was measured by real-time PCR in U87 and SH-SY5Y cells, normalized for hypoxanthine-guanine phosphoribosyltransferase (HPRT) signals, and expressed percentage of SHSY5Y values. *, $p<0.05$ vs U87+vehicle group

6.11 Summary of Results

These results showed that 4 transcriptional factors, expressed in CNS, are able to regulate the transcriptional activity of the identified NCX2 promoter under control conditions in cells that express the endogenous NCX2 gene. By contrast, these TFs failed to regulate the promoter activity of the identified S3 region in cells that do not express the endogenous gene of the exchanger. In particular, three TF increased significantly the luciferase activity of pGL-S3 with molecular determinants located in the regions A or C. In addition, the binding sequences for these TF were also studied. In particular, the molecular determinant for CREB1 activity was localized in the region A, possibly at -654 bp from the TSS, whereas Sp1 and Sp4 required the binding site at +403 bp, localized in the region C. SREBP1 was the only TF that showed an inhibitory effect on the luciferase activity. This effect was exerted in all deletion mutants of pGL3-S3 that contains region C. In addition, treatment with AZA, which prevents DNA methylation, reduced NCX2 silencing occurring in U87 cell line.

Chapter 7: Discussion

In this study, we reported the isolation and characterization of the promoter of rat NCX2 gene and the molecular determinants involved in the regulation of its expression in PC12 cells.

Results obtained in this thesis showed for the first time a regulatory region localized between *kptn* and *slc8a2* genes that is involved in the expression of NCX2 in CNS. In particular, this promoter sequence, that we named S3, was provided with two high identity matrix scores, above 0.80, of the initiator motif that facilitate the transcription by the RNA polymerase complex. These data were reinforced by 5' UTR sequences of mouse and human mRNA encoding for NCX2 reported in NCBI database. Indeed, both 5'UTR sequences of mouse and human mRNA overlapped with the DNA sequences downstream the observed Inr signals. Interestingly, these consensus sequences appeared conserved in rat genome where the 5'-UTR sequence of NCX2 mRNA was not reported yet. In this regard, from the orthologous comparison and from *in silico* analysis, it is conceivable to hypothesize that the two Inr sequences identified in rat *slc8a2* promoter might represent the effective transcription starting sites of this gene.

Notably, the identified promoter sequence lacks high predictive TATA-box, DPE, and MTE motifs in accordance with the mutually exclusive presence of TATA-box and Inr motif, as often sperimentally observed. These data, togheter

with the localization of the identified Inr motifs in high-dense CpG islands and the presence of many Sp1 binding sites, suggest that the promoter of NCX2 can be classified as dispersed (Juven-Gershon and Kadonaga 2010).

Regarding the intergenic regions surrounding the identified S3 sequence, bioinformatic analysis reported a great number of putative TF binding sites divided in four dense groups along the entire intergenic region. However, only region S3 appeared to promote a significant transcriptional luciferase activity under basal conditions even if compared with the promoter activity of the viral SV40 sequence. These data suggest that S3 was provieded with the most important transcription factors that maintain the basal expresison levels of NCX2 in the neuronal cell line PC12. In particular, among the 4 identified subregions of S3, A, B, C, and D, the basal expression levels of the transcriptional activity seemed to be related only to the presence of the subregion C. Indeed, deletion mutant that lacks the subregions A, B and C of S3 displayed a low trascriptional activity although the reduced length of the sequence. By contrast, the deletion mutant lacking A and B showed the highest luciferase activity among S3 and all its deletion mutans as observed in both PC12 and SHSY cell lines. These data suggested that the subregion of S3 including C and D might represent the minimal promoter sequence of NCX2. Indeed, deletion of the subregion A enahced, rather than inhibited, the basal transcription activity of S3 promoter in both neuronal cell lines, PC12 and SHSY. This might be explained by the

presence of binding sites for inhibitory transcription factors that we did not examine in this thesis (see supplementart table 5). On the other hand, the subregion B was ininfluent for the promoter activity of S3 in SHSY cell lines since the S3 mutants lacking the subregion A or both A and B showed the same luciferase activity. At difference with these data, the deletion of both subregions A and B increased the luciferase activity if compared with the activity of the deletion mutant lacking the subregion A in PC12 cells. This discrepancy might be partially explained by the difference in the expression levels of several transcription factors that substain the basal trascriptional activity of S3 promoter. In this regard, it is possible to hypothesize that PC12, but not SHSY, cells express inhibitory trascription factors with molecular determinants present in the subregion B. Supporting this hypothesis, SHSY cell line displays an increased amount of NCX2 mRNA compared with PC12 cells, as revealed by real-time PCR. Regarding the subregion D, we should consider that it showed a significant transcriptional activity that might be induced by the presence of several binding sites for transcriptional factors with high identity matrix scores that we did not exanimed in this thesis. On the other hand, this subregion did not contain the molecular determinants for the effect of the transcription factors that we used. By contrast, subregion C was the most interesting because it contained the regulatory binding sites of Sp1, Sp4 and SREBP1 transcription factors even under control conditions. Indeed, Sp1 failed to enhance the promoter activity of

NCX2 when region C was removed from S3 promoter. Furthermore, site-directed mutagenesis showed that the binding site in position +403/+413, localized in the region C, is the specific molecular determinant for the Sp1-mediated enhanced promoter activity in PC12 cells under control conditions. Likewise, the neuronal transcription factor Sp4 failed to enhance NCX2 promoter activity either when the region +376/+465 or +322/+423 were removed from S3. Also the disruption of the binding site in position +403/+413, by site-directed mutagenesis, prevented the Sp4-mediated stimulatory effect suggesting that this sequence is crucial for both transcription factors. Notably, this transcription binding site showed one of the highest matrix identity score in the whole S3 sequence. Interestingly, this subregion was also foundamental for the SREBP1-mediated downregulation of NCX2 promoter activity, since the removal of the region including A, B and C, but not the removal of the region including only A and B, prevented the SREBP1-mediated downregulation of NCX2 promoter activity. In fact, two putative SREBP1 binding sites were identified in this subregion. It was coinceivable to hypothesize that the sequence +388/+398 could represent the molecular determinant for SREBP1 since it was provided with a high matrix identity score and was in close proximity to two binding sites of its coregulator Sp1. Furthermore, one of these binding sites for coregulators was determinant for the effect of both Sp1 and Sp4 on S3 promoter. However, the disruption of either Sp1 binding sites or the deletion of a region

from +376 to +465 including the three sequences failed to prevent the SREBP1-inhibitory effect, suggesting that the most probable molecular determinant for SREBP1 is located in position +277/+287 in the subregion C. At difference with the above mentioned trascription factors, CREB1 showed to exert its activity only on the enteire S3 promoter, whereas the deletion of either region A, or both A and B, prevented its stimulatory effect on the identified NCX2 promoter. These data suggested that the molecular determinant for CREB1 is located in position -627/-615 in subregion A.

Interestingly, the overexpression of either Sp3, NFkB, GATA1/2, RUNX1 transcription factors failed to modify the promoter activity of NCX2 in PC12 cells under our experimental conditions, neverthless the promoter was provided with their conserved binding sites with high matrix identity scores. On the other hand, we can not exclude that these trascription factors might regulate NCX2 transcription under particular physiological or pathophysiological conditions that we did not exanimed. In addition, we also can not exclude that other transcription factors that have a conserved binding site on S3 can also partecipate in the regulation of the transcription of NCX2.

Another aspect that deserves attention is the reliability of S3 region as the promoter of NCX2. Indeed, we found many putative trascription binding sites that were conserved among three considered species, including rat, mouse and human DNA. In addition, it is worth of mention that the basal transcriptional

activity S3 was comparable with the level of expression of the endogenous NCX2 gene in neuronal and non-neutonal cells, including PC12, SHSY, U87, and BHK cell lines. On the other hand, the transfection of either Sp1, Sp4 and CREB1 failed to induce the transcription of S3 in U87 cell line. These data might be explained by several mechanisms of NCX2 silencing. In particular, it is known that glioblastoma cells, including U87, show an increase of SREBP1 (Ru, Hu et al. 2016) that exerted an inhibitory effect on NCX2 expression. In addition, glioblastoma cells show an increase in the methylation of NCX2 gene (Qu, Jiao et al. 2010). These aspects appear more interesting if we examine the effect of the demethylating agent Aza on the expression of NCX2 in U87 cell lines. Indeed, 72h of Aza induced the expression of NCX2 mRNA in U87 cells suggesting that the DNA methylation, occuring in glioblastoma, is involved in the silencing of this antiporter gene. Furthermore, several putative binding sites present on S3 sequence displayed methylable CpG dinuleotides that might interfere with the efficacy of the respective trascription factor to modify NCX2 expression. In this regards, rat myoblast cells that do not express the endogenous NCX2 gene showed at least significant difference in 8 CpG methylation if compared with rat pheocromocytoma cells.

Collectively, results obtained suggest that the region S3 represents the promoter of NCX2 gene. This promoter is enhanced by CREB1, Sp1, and Sp4, whereas it is inhibited by SREBP1 transcription factor in PC12 cells under

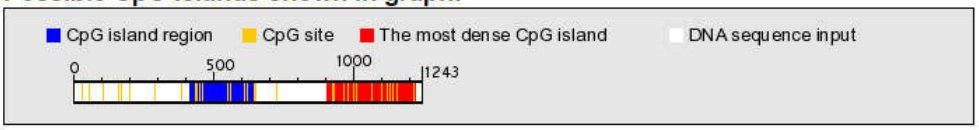
control conditions. Moreover, our findings also suggested that the epigenetic DNA methylation may be involved in the silencing of NCX2 in some cell lines, including glioblastoma and myoblastoma cells.

Chapter 8: Supplementary Materials

S-1: Analysis of CpG islands in rat NCX2 promoter sequence

The sequence you input has 1243 nucleotides in length.
There are 2 possible CpG island(s) found.

Possible CpG islands shown in graph:



Sequence of each CpG island:

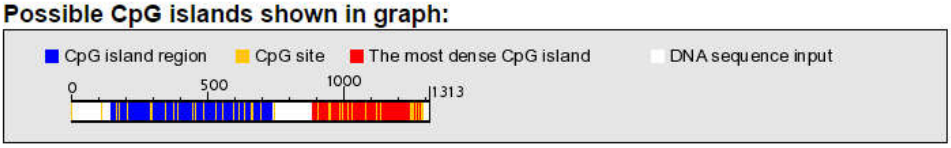
1. start site:417 end site:645 length:229 GC content:61%
ccacagcacagactggtgtacgaacttgcgccctccatcgctgtcctcggattctagtgctgaaccccaagattgggcaacctgcagc
cccagccctgccaagcccctctaccacccactagcttccttggctacgccccttccgggctcctccctcctgcccctttgcagggca
tagtctctgttctgtagccccgcccctggtcacgcctcctcatattggctcct

2. start site:909 end site:1215 length:307 GC content:67%
tcacttgtcccacccctcccgtcagcggcaactccctccttcagcctagactctcttaacggctgattcccccgccccagttcctct
gtccccgccccctttccccggcctcctcccaagctgtattcccctacctctccagtctagcccccctctccggggctgctctctcaact
ccccccttttctctgtccccgccccctccctgggctccgcccctcctcccggcccctttcccggttcccctcctcgcatcccctctcct
ccctcctccctgctctccctcagtcccctgctctccctcagtc

S-2: Analysis of CpG islands of mouse NCX2 promoter sequence

The sequence you input has 1313 nucleotides in length.
There are 2 possible CpG island(s) found.

Possible CpG islands shown in graph:



Sequence of each CpG island:

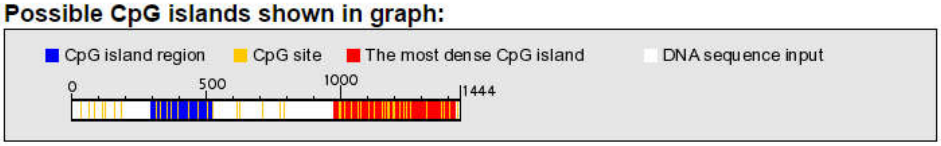
1. start site:149 end site:731 length:583 GC content:61%
ctggggattgggagtgaggccgtgttttggcgtggcaggggcagtgggaatggttctcgtcgtacctgggcactctagtgacagttcc
ctgtcttctgtaaggtgctctgagcaggctgagctgtgaacttggaaagatctgccgcccgacattccttcccccacttcccccaggt
ctgccagcctgtggctatttgtagcgatgactgcatctctgggcccttgtatggacgggatgctgtgcgtgtgtgatcaaggagggga
actgggaaactacaggacagacaggtgtatgtacttgcgcccctccgtcgctgtccttggattctagtgctgaaccccgagactggggc
aacctgcagacccagccctgccaagcccctcggccacccactagcttccttggccacgccccattctgggctcctccctcttgctctt
ttgcaaggcctcgtctcagctttgtggccccgcccccaactgcctggccacgcctcctcctattggctcctccttcggcggattcccc
agcctgagctcccaggcctcgctccctcccataagccctgccctctcctggctcc

2. start site:886 end site:1283 length:398 GC content:67%
atttcctctctcccacctcccggtggtcctttttcccagctcacttgtcccacccccctctcatccgcggcatctccctccttcagccta
gtccctctttcccggctgattcccccgccccaattcctctgttcccgcccccttcccccggcttcctcccaagctgtactcccctgcc
tctccaatctagcccccctcccgggggctgctttctcagccccccccctttcctctgtcccgcccccctcccctggactccgccctcctc
ccagcccctcccccagttcccctcctagcatcccctctcttccctcctccctgctctccctcagtcccccagctcctctccctctttctc
cctttcctcctcgggcggaggcgcccactgccgggggcccagcggg

S-3: Analysis of CpG islands of human NCX2 promoter sequence

The sequence you input has 1444 nucleotides in length.
There are 2 possible CpG island(s) found.

Possible CpG islands shown in graph:



Sequence of each CpG island:

1. start site:297 end site:517 length:221 GC content:53%
acattagaacatgccaggtgcggtggctctcacctgtgcggtggctctcacctgtaatctcagcgctttgggaggccgaggcagaagg
attgcttaagcccgggagttggagactagcctaggtaaaatagtgagaccccgtctcttaaaaaaaatgaagagaaaaaaagcctggc
gtggtggtggtgggcacctgtagtcccagctactcgggaggccaa

2. start site:978 end site:1421 length:444 GC content:69%
cttgtatggctgggagactgcgtgcgtgtgtgtgtgtgcgtgtgtgtgtgtgtgttgggggggcgggggctgcttctgaaacacaca
ggagcgccctccggtgctgtcccctgcctctggtgctgaacccggaggctggggaggaggctcgggctccctgctccccaggcccct
ccccgggcctcccccgcctcccgcaccgccttccctctgcccgcctccgcctcctcccccctctctccggctcctcccccctccctcgc
tccctcgccgcctcccctcccccgtgtccctccctctcctcttctcctccctctatccccctctccctcccctcctccccggctctc
cctctcagtctcaggcttttctctctctcaatctctctctctctttcgggcggagtcgcccaccactgccagcccagcgctgggggga
cctg

S-4: List of selected transcription factors binding sites.

TF	RAT			MOUSE			HUMAN		
	Sequence	Pos	Score	Sequence	Pos	Score	Sequence	Pos	Score
CREB1	tgaagcca	-627	83%	tgaagcca	-615	83%	ggaggcca	-670	80%
SREBF1	gccacgccaa	-532	76%	gccacgccaa	-523	76%	accacgccag	-567	82%
SP3	gcccctgccac	-527	80%	gcccctgccac	-518	80%	gccaccaccac	-562	82%
SP1	ccactgcccct	-522	82%	cctctgcccct	-513	87%	tgcccgccacc	-557	84%
SREBF1	agcaccttac	-461	76%	agcaccttac	-452	76%	ctcaagcaat	-508	76%
SREBF1	cttcccccac	-403	76%	cttcccccac	-393	76%	gtcacccagg	-448	76%
RUNX1	tcctgtagttt	-288	83%	tgctgtggtat	-278	91%	agctgttgtta	-360	82%
SP3	gccaagcccct	-187	87%	gccaagcccct	-178	87%	aagacacccat	-340	81%
RUNX1	gctagtgggtg	-172	80%	gctagtgggtg	-163	80%	gtttctggcct	-322	84%
SREBF1	cacgccccat	-152	76%	gtcacgcctc	-71	76%	ctgactccag	-271	75%
SP1	gctcctccctc	-137	94%	gctcctccctc	-128	94%	tgccctccccc	-250	93%
SP3	gctcctccctc	-137	80%	gctcctccctc	-128	80%	tgccctccccc	-250	82%
SP3	gccccgccccc	-92	97%	gccccgcccct	-83	94%	cacacacacac	-158	83%
SP4	ctggccacgcctcctcc	-75	92%	ctggtcacgcctcctca	-74	84%	cagcccccgcccccaa	-145	86%
SP1	gccacgcctcc	-72	87%	gtcacgcctcc	-71	85%	cccccgccccc	-142	98%
SP3	gccacgcctcc	-72	96%	gtcacgcctcc	-71	89%	cccccgccccc	-142	96%
SP1	acgcctcctcc	-69	84%	acgcctcctca	-68	80%	gcccccgcccc	-141	80%
SP1	ctcgctccctc	-13	82%	ctcactccctc	-13	82%	tcccctgcctc	-94	84%
SP1	ctccctcccat	-9	91%	ctccctcccat	-9	91%	cccctgcctct	-93	87%
SP3	ctccctcccat	-9	83%	ctccctcccat	-9	83%	cccctgcctct	-93	80%
SP1	agccctgccct	3	80%	accctaccctc	4	84%	ccccagcctcc	-70	86%
SP3	gccctgccctc	4	83%	accctaccctc	4	80%	ccccagcctcc	-70	82%
SP1	gctcctccctc	20	94%	gctcctctttt	20	80%	cctcctcccca	-64	92%
SP1	attcctcctcc	49	90%	atcctcctcc	45	92%	gggcctcccc	-21	85%
SP1	cctcctcctgg	52	80%	cctcctcctgg	48	80%	cccccgcctcc	-14	96%
SP1	tctcttccttc	97	82%	tctcctccttc	93	90%	tgcccgcctcc	13	92%
SP1	gctcctccctc	109	94%	actcctccctc	105	91%	cctcctcccc	25	95%
SP3	gctcctccctc	109	80%	actcctccctc	105	79%	cctcctcccc	25	84%
SP1	ctccctcttcc	113	82%	ctccctcttcc	109	82%	cctccccctct	28	83%
SP1	acaccaccctt	163	83%	acaccaccctg	162	81%	gctcctcccc	44	97%
SP1	acccttcctaa	168	79%	accctgcctaa	167	80%	cctccccctcc	47	84%
GATA2	aatttcctctctcc	178	80%	aatttcctctctcc	177	80%	tcaggcttttctct	161	82%
SP1	tttcctctctc	180	80%	tttcctctctc	179	80%	ctcgctccctc	58	82%
GATA1	ttcctctctcc	181	80%	ttcctctctcc	180	80%	ggcttttctct	164	83%
SP1	ctctctcccac	184	81%	ctctctcccac	183	81%	ctccctcgccg	62	81%
SP1	ctcccacctcc	188	92%	ctcccacctcc	187	92%	tcgccgcctcc	67	86%
SP3	ctcccacctcc	188	83%	ctcccacctcc	187	83%	tcgccgcctcc	67	80%
SP1	gtcccaccccc	224	96%	gtcccaccccc	223	96%	ctccctcccct	126	94%
SP3	gtcccaccccc	224	88%	gtcccaccccc	223	88%	ctccctcccct	126	84%
SP1	ccccctctcat	230	82%	cccctcccgt	229	86%	tcccctcctcc	131	95%
SP1	ctccctccttc	250	90%	ctccctccttc	249	90%	cgcccaccact	207	81%
SREBF1	atcagccggg	277	76%	atcagccgtt	275	79%	gccagcccag	218	78%
NFKB1	gggggaatcagcc	280	82%	gggggaatcagcc	278	82%	tggggtaaccccg	306	79%
SP1	cccccgcccca	288	95%	cccccgcccca	286	95%	cctcctccccg	134	92%
SP1	cccccggcttc	321	83%	cccccggcctc	319	85%	actccgcccga	196	81%
SREBF1	ctcagccccc	388	81%	ctccccccctt	391	76%	gagaccccat	266	76%

SP4	ctgtccccgcccctcc	407	89%	ctgtccccgcccctcc	405	89%	tttcccacgccccacc	285	86%
SP1	tccccgccccc	410	98%	tccccgccccc	408	98%	cccacgccccc	288	88%
SP3	tccccgccccc	410	94%	tccccgccccc	408	94%	cccacgccccc	288	100%
SP1	gccccctcccc	415	80%	gccccctcccc	413	80%	ttcccacgccc	290	83%
SP4	tggactccgccctcctc	426	79%	tgggctccgccctcctc	424	82%	gtaaccccgtccttttc	298	80%
SP1	actccgccctc	429	92%	gctccgccctc	427	95%	accccgtcctt	301	82%
SP1	ctcccagcccc	440	83%	ctcccggcccc	438	86%	actccgcacct	318	83%
SP3	tcccagcccct	441	81%	tcccggcccct	439	80%	cccactccgca	321	81%
SP1	gcccctcccca	446	95%	gccccttcccg	444	85%	cccccactccg	323	82%
SP1	tcccctcctag	459	89%	tcccctcctcg	457	92%	gaaccgccccc	329	83%
SP3	tcccctcctag	459	80%	tcccctcctcg	457	80%	gaaccgccccc	329	87%
SP1	ttccctcctcc	481	93%	ctccctcctcc	479	93%	cccccacctcc	376	94%
SP3	ttccctcctcc	481	79%	ctccctcctcc	479	82%	cccccacctcc	376	90%

S-5: List of conserved putative binding sites for transcription factors among rat, mouse and human promoter of NCX2

	RAT			MOUSE			HUMAN		
TF	**Sequence**	**Pos**	**Score**	**Sequence**	**Pos**	**Score**	**Sequence**	**Pos**	**Score**
OTX2	gtgatctt	-698	79%	gtgatctt	-682	79%	ttaagccc	-789	92%
FOXC1	aggatcaagat	-695	80%	aggatcaagat	-679	80%	aaaaaaaatga	-735	82%
FOXD2	atcaaga	-694	81%	atcaaga	-678	81%	gaaaaaa	-721	79%
FOXO4	atcaaga	-694	79%	atcaaga	-678	79%	gaaaaaa	-721	80%
PITX3	cttgatcct	-693	86%	cttgatcct	-677	86%	aaaaagcct	-718	79%
OTX1	ttgatcct	-692	85%	ttgatcct	-676	85%	aaaagcct	-717	81%
OTX2	ttgatcct	-692	88%	ttgatcct	-676	88%	aaaagcct	-717	81%
NFIX	cttgccagc	-673	96%	cgagccaag	-657	96%	cacgccagg	-712	92%
HES1	accacatgct	-651	83%	agcatgcggt	-639	79%	ggcacctgta	-694	82%
NEUROD	accacatgct	-651	93%	accgcatgct	-639	89%	ggcacctgta	-694	82%
TCFL5	accacatgct	-651	85%	accgcatgct	-639	85%	ctcccgagta	-676	81%
TCFL5	agcatgtggt	-651	85%	agcatgcggt	-639	85%	tactcgggag	-676	81%
NFIX	gaagccaca	-626	84%	gaagccaca	-614	84%	gaggccaag	-669	93%
NFIC	gtggct	-624	81%	gtggct	-612	81%	ttggcc	-667	94%
ID4	agcaactgtg	-621	82%	agcaactgtg	-609	82%	aactcctggg	-642	80%
SNAI2	cacagttgc	-621	81%	cacagttgc	-609	81%	ggctggtct	-630	81%
TCF3	cacagttgct	-621	81%	cacagttgct	-609	81%	aactcctggg	-642	86%
TCF4	cacagttgct	-621	81%	cacagttgct	-609	81%	aactcctggg	-642	84%
BHLHA	gcaactgt	-620	82%	gcaactgt	-608	82%	ccttctgc	-660	81%
RFX1	agttgctg	-618	85%	agttgctg	-606	85%	tgttaccc	-620	83%
JUN	ggaaactgagacag	-613	86%	ggaaactgaggca	-601	82%	catggtgagatccc	-614	81%
XBP1	ttcgaag	-595	97%	ctcgtaa	-585	80%	ctcaagc	-649	79%
TEAD3	acatttcg	-592	84%	acattacg	-576	84%	agatccca	-607	79%
CEBPB	atgttaaaat	-588	80%	attttgacat	-579	81%	cccatctcctaaaa	-603	81%

BARHL1	ttaaaatgga	-585	83%	tcaaaatgga	-576	83%	cctaaaaaaa	-596	82%
BARHL2	ttaaaatgga	-585	85%	tcaaaatgga	-576	85%	cctaaaaaaa	-596	81%
NOTO	ctccatttta	-584	81%	ctccattttg	-575	82%	ctccatttttt	-586	81%
YY1	taaaatggagct	-584	85%	caaaatggagct	-575	89%	taacatggtgag	-617	81%
SOX10	catttt	-583	81%	catttt	-574	81%	catggt	-614	80%
FOXD2	gctaata	-575	84%	aacaaca	-563	86%	aaaaatg	-586	83%
GSX1	tgttattagc	-575	88%	tgttgttagc	-566	79%	cctaaaaaaa	-596	81%
MNX1	gctaataaca	-575	89%	gctaacaaca	-566	81%	tctcctaaaa	-599	80%
STAT5	aatccccagag	-560	90%	aatccccagag	-551	90%	ttttttaggag	-598	83%
E2F6	gggagtgaggc	-549	85%	gggagtgagac	-540	85%	tggagggaaaa	-581	95%
SOX10	ccgtgt	-539	84%	ccgtgt	-530	84%	cttttt	-574	82%
HIC1	acgccaaaa	-534	81%	acgccaaaa	-525	81%	acgccaggc	-569	82%
HIC2	acgccaaaa	-534	80%	acgccaaaa	-525	80%	acgccaggc	-569	82%
THAP1	acgccaaaa	-534	80%	acgccaaaa	-525	80%	tttccctcc	-580	87%
NFIX	cacgccaaa	-533	94%	cacgccaaa	-524	94%	cacgccagg	-568	92%
HES1	gccacgccaa	-532	86%	gccacgccaa	-523	86%	accacgccag	-567	82%
MGA	tggcgtgg	-531	80%	tggcgtgg	-522	80%	tggcgtgg	-566	80%
ARNT	ggcgtg	-530	92%	ggcgtg	-521	92%	ggcgtg	-565	92%
EGR1	gcccctgccacgcc	-530	87%	gcccctgccacgc	-521	87%	ccgccaccaccac	-563	85%
PAX2	tgccacgc	-529	88%	tgccacgc	-520	88%	caccacgc	-564	81%
THAP1	ctgccacgc	-529	83%	ctgccacgc	-520	83%	ccgccacca	-558	86%
KLF16	gcccctgccac	-527	81%	gcccctgccac	-518	81%	gccaccaccac	-562	82%
SP2	ccactgcccctgcca	-526	80%	cctctgcccctgcc	-517	82%	tgcccgccacca	-561	81%
EGR1	ccactgcccctgcc	-525	87%	cctctgcccctgcc	-516	88%	tgcccgccaccac	-560	85%
NFIX	aggggcagt	-522	82%	aggggcaga	-513	82%	cccgccacc	-557	83%
E2F6	cagtgggaatg	-517	82%	cagagggaatg	-508	85%	tggcgggcacc	-555	84%
HIC2	attcccact	-516	88%	attccctct	-507	83%	gtgcccgcc	-554	97%
THAP1	attcccact	-516	81%	attccctct	-507	82%	gtgcccgcc	-554	90%
ZEB1	tcgtacctg	-499	89%	tcgtacctg	-490	89%	gggcacctg	-551	83%
ID4	cgtacctggg	-498	80%	cgtacctggg	-489	80%	tacaggtgcc	-550	84%
TCF3	cgtacctggg	-498	83%	cgtacctggg	-489	83%	ggcacctgta	-550	97%
TCF4	cgtacctggg	-498	86%	cgtacctggg	-489	86%	ggcacctgta	-550	97%
HIF1A	gtacctgg	-497	80%	gtacctgg	-488	80%	acaggtgc	-549	81%
SNAI2	cccaggtac	-497	85%	cccaggtac	-488	85%	tacaggtgc	-549	98%
HIC1	gtgcccagg	-494	86%	gtgcccagg	-485	86%	gtcccagct	-540	86%
HIC2	gtgcccagg	-494	96%	gtgcccagg	-485	96%	gtcccagct	-540	86%
MZF1	tgggca	-492	81%	tgggca	-483	81%	ctggga	-539	83%
NFIC	tgggca	-492	85%	tgggca	-483	85%	ctggga	-539	79%
ISL2	gcactcta	-489	87%	gcactctg	-480	85%	ctactcag	-533	81%
E2F6	gacagggaact	-476	80%	gacagggaact	-467	80%	aggcaggagga	-518	87%
MEIS3	aagacagg	-471	89%	aagacagg	-462	89%	gaggcagg	-519	81%
CEBPA	cttacagaaga	-467	80%	cttacagaaga	-458	80%	attgcttgagc	-508	82%
SNAI2	gtaaggtgc	-461	87%	gtaaggtgc	-452	87%	ggcaggagg	-517	79%
TCF4	agcaccttac	-461	79%	agcaccttac	-452	79%	tcctcctgcc	-517	82%
TFAP2A	tgctcagagca	-455	80%	tgctcagagca	-446	80%	gggctcaagca	-506	84%

TFAP2A	tgctctgagca	-455	80%	tgctctgagca	-446	80%	tgcttgagccc	-506	79%
TFAP2C	tgctcagagca	-455	81%	tgctcagagca	-446	81%	gggctcaagca	-506	83%
TFAP2C	tgctctgagca	-455	80%	tgctctgagca	-446	80%	tgcttgagccc	-506	80%
NRL	agcaggctgag	-448	80%	agcaggctgag	-439	80%	ttactgcagcc	-485	80%
ZFX	cacagctcagcctg	-446	81%	acaggctcagcct	-437	83%	cctgcctcagcctc	-525	85%
PITX1	ctcagcct	-445	79%	ctcagcct	-436	79%	ttgagccc	-503	84%
MEIS3	ttcacagc	-438	82%	ttcacagg	-428	82%	gtggcacg	-466	85%
ESRRG	ccaagttcac	-435	84%	ccaagttcac	-425	84%	ccaagatcgt	-473	80%
NFIA	ctttccaagt	-431	83%	ctttccaagt	-421	83%	cgtgccactg	-466	81%
NFIX	ctttccaag	-430	90%	ctttccaag	-420	90%	cgtgccact	-466	88%
NFIC	ttggaa	-429	93%	ttggaa	-419	93%	agtggcacg	-466	81%
NFIC	ttggaa	-429	93%	ttggaa	-419	93%	tgcactccag	-458	82%
TEAD1	atctttccaa	-429	80%	atctttccaa	-419	80%	gcactcca	-457	85%
TCF3	gatctgccgc	-422	80%	gatctgcctc	-412	84%	aggctggagt	-455	79%
NFIX	ggcggcaga	-420	84%	ggaggcaga	-410	85%	cactccagc	-456	83%
E2F6	gggaaggaatg	-408	83%	gggaaggaatg	-398	83%	aggctggagtg	-456	83%
MZF1	gggggа	-401	96%	gggggа	-391	96%	gggtga	-445	81%
EGR1	cccccacttccccc	-400	82%	cccccacttccccc	-390	82%	gtgctgccgatgcc	-407	80%
HIF1A	ggaagtgg	-397	80%	ggaagtgg	-387	80%	ccacgtgc	-391	97%
SMAD2	aggtctgccagcc	-386	84%	aggtctgccagcc	-376	84%	cactctgtcaccc	-445	86%
USF2	agcctgtggct	-377	82%	agcctgtggct	-367	82%	tcaaataa	-422	81%
NOTO	acaaatagcc	-370	79%	acaaatagcc	-360	79%	tcaaataata	-422	81%
MEF2A	gctacaaatagc	-369	91%	gctacaaatagc	-359	91%	aataataataat	-419	82%
MEF2B	gctacaaatagc	-369	92%	gctacaaatagc	-359	92%	aataataataat	-419	81%
MEF2D	gctacaaatagc	-369	91%	gctacaaatagc	-359	91%	aataataataat	-419	82%
RAX	tacaaatagc	-369	83%	tacaaatagc	-359	83%	ctcaaataat	-423	86%
ISL2	acaaatag	-368	79%	acaaatag	-358	79%	tcaaataa	-422	80%
ISX	acaaatag	-368	80%	acaaatag	-358	80%	tcaaataa	-422	82%
FOXD2	acaaata	-367	90%	acaaata	-357	90%	tcaaata	-422	82%
FOXL1	acaaata	-367	90%	acaaata	-357	90%	tcaaata	-422	83%
MEOX2	agtcatcgct	-360	80%	agcaatgact	-350	84%	aataataatg	-416	82%
PAX2	agtcatcg	-358	83%	agtcattg	-348	94%	tgtcaccc	-445	84%
JDP2	gatgactgcatc	-357	80%	aatgactacatc	-347	79%	cagagtgagaccc	-439	79%
SOX10	ccttgt	-338	92%	ccttgt	-328	92%	cattat	-412	81%
CDX1	tccatacaa	-336	80%	tccatacaa	-326	80%	ataataatg	-415	86%
TEAD1	agcatcccgt	-328	84%	cgcatagcaa	-313	80%	cacattatta	-414	85%
TEAD4	agcatcccgt	-328	79%	cgcatagcaa	-313	81%	cacattatta	-414	80%
HIF1A	gtgcgtgt	-318	90%	atgcgtgt	-308	87%	gcacgtgg	-391	96%
ARNT	tgcgtg	-317	100%	tgcgtg	-307	100%	cacgtg	-390	80%
NEUROD	cacacacgca	-317	81%	cacacacgca	-307	81%	accacgtgcc	-392	84%
TCFL5	cacacacgca	-317	80%	cacacacgca	-307	80%	accacgtgcc	-392	97%
TCFL5	tgcgtgtgtg	-317	80%	tgcgtgtgtg	-307	80%	ggcacgtggt	-392	97%
ARNT	cgtgtg	-315	84%	cgtgtg	-305	84%	cacgtg	-390	80%
ARNT	tgtgtg	-313	85%	tgtgtg	-303	85%	tgcctg	-376	82%
PAX2	gatcacac	-312	81%	gatcacac	-302	81%	ggtcaggg	-385	85%

ARID3A	atcaca	-311	79%	atcaca	-301	79%	attaat	-344	94%
GATA5	tgatcaca	-311	89%	tgatcaca	-301	89%	taataaga	-348	80%
LHX4	gtgatcaa	-310	84%	ttgatcac	-300	83%	cttattaa	-347	91%
LHX4	ttgatcac	-310	83%	gtgatcaa	-300	84%	ttaataag	-347	81%
LHX8	ttgatcac	-310	91%	ttgatcac	-300	91%	cttattaa	-347	84%
LHX8	gtgatcaa	-310	91%	gtgatcaa	-300	91%	ttaataag	-347	83%
UNCX	ttgatcac	-310	79%	ttgatcac	-300	79%	cttattaa	-347	89%
ARID3A	atcaag	-307	84%	atcaaa	-297	97%	attaat	-344	94%
HIC2	ttcccctcc	-302	80%	atcccagtc	-285	83%	ctgccatac	-367	82%
MYOD1	tacacctgtctgt	-278	94%	tacaccagtctgt	-268	82%	ggcagctgttgtt	-363	98%
SOX10	ctgtct	-276	82%	cagtct	-266	81%	tgttgt	-357	82%
TCF12	gacaggtgtat	-275	92%	gactggtgtac	-265	80%	aacagctgcca	-364	98%
TCF3	tacacctgtc	-275	97%	tacaccagtc	-265	86%	aacagctgcc	-363	94%
TCF4	tacacctgtc	-275	98%	tacaccagtc	-265	87%	aacagctgcc	-363	93%
PAX2	cgtcgctg	-250	85%	catcgctg	-240	79%	tgttattc	-354	82%
SNAI2	ttctagtgc	-234	79%	ttctagtgc	-224	79%	ggcagctgt	-363	79%
RFX1	ggttgccc	-210	89%	ggttgccc	-201	89%	tgttgtta	-357	82%
SNAI2	tgcaggttg	-207	92%	tgcaggttg	-198	92%	aacagctgc	-362	79%
GCM2	ctgcaggt	-205	88%	ctgcaggt	-196	88%	atgggtgt	-340	81%
KLF1	gacccagccct	-198	81%	agccccagccc	-190	81%	aaagacaccca	-339	83%
MEIS3	ttggcagg	-190	91%	ttggcagg	-181	91%	aagacacc	-337	85%
KLF4	ggggcttggc	-187	87%	ggggcttggc	-178	87%	tgggtgtctt	-339	81%
ELK4	tagcttccttg	-165	81%	tagcttccttg	-156	81%	ttgtttctggc	-324	79%
ERG	caaggaagct	-164	80%	caaggaagct	-155	80%	acagtatgtg	-290	81%
EGR1	gccacgccccattc	-154	82%	gctacgcccctttc	-145	85%	tgccctcccctag	-250	83%
KLF16	gccacgcccca	-154	91%	gctacgcccct	-145	89%	tgccctcccc	-250	83%
KLF4	ggggcgtggc	-154	98%	ggggcgtagc	-145	92%	ggggagggca	-250	80%
SP2	gccacgccccattct	-154	83%	gctacgcccctttc	-145	84%	tgccctcccctag	-250	87%
THAP1	gagcccaga	-142	82%	gagcccgga	-133	84%	ttgccctcc	-251	92%
NFIX	ggagcccag	-141	87%	ggagcccgg	-132	84%	gagggcaag	-252	85%
KLF16	gctcctccctc	-137	79%	gctcctccctc	-128	79%	gccctcccct	-249	80%
E2F4	aagagggagga	-135	79%	aggagggagga	-126	84%	tagggggaggg	-248	80%
E2F6	aagagggagga	-135	92%	aggagggagga	-126	95%	tagggggaggg	-248	83%
MZF1	ggagga	-135	84%	ggagga	-126	84%	ggggga	-245	96%
THAP1	cctccctct	-134	84%	cctccctcc	-125	85%	cctcccct	-247	80%
NFIA	cttgcaaaag	-121	81%	cctgcaaagg	-112	81%	agggccaaca	-209	83%
NFIX	ttttgcaag	-120	82%	ctttgcagg	-111	79%	ccagccata	-197	85%
SOX10	ctttgt	-99	100%	ctctgt	-95	86%	cagtct	-187	81%
KLF1	ggccccgcccc	-93	89%	agccccgcccc	-84	88%	acacacacaca	-159	79%
EGR1	gccccgccccaac	-92	88%	gccccgccctgg	-83	89%	cacacacacacac	-159	80%
KLF16	gccccgccccc	-92	98%	gccccgcccct	-83	94%	cacacacacac	-160	80%
SP8	gccccgccccca	-92	94%	gccccgcccctg	-83	84%	cacacgcacaca	-169	84%
THAP1	ccgccccca	-89	90%	ccgcccctg	-80	81%	tctcccagc	-193	80%
EGR1	tgcctggccacgcc	-78	81%	cccctggtcacgc	-77	79%	ccccaacacacac	-153	80%
KLF12	ggccacgcctcctcc	-73	87%	ggtcacgcctcctc	-72	80%	caacacacaca	-153	81%

KLF14	ggccacgcctcctc	-73	93%	ggtcacgcctcctc	-72	91%	gcccccgcccccc	-144	86%
EGR1	gccacgcctcctcc	-72	92%	gtcacgcctcctca	-71	89%	cccccgcccccca	-145	82%
KLF16	gccacgcctcc	-72	94%	gtcacgcctcc	-71	88%	ccccgcccccc	-143	81%
SP2	gccacgcctcctcct	-72	87%	gtcacgcctcctca	-71	84%	cccccgcccccca	-146	93%
SP8	gccacgcctcct	-72	94%	gtcacgcctcct	-71	86%	gcccccgccccc	-142	83%
MZF1	ggagga	-64	84%	tgagga	-63	88%	cggggg	-137	85%
FOXD2	gccaata	-58	87%	gccaata	-57	87%	acacaca	-120	83%
FOXL1	gccaata	-58	85%	gccaata	-57	85%	acacaca	-120	81%
FOXO4	gccaata	-58	79%	gccaata	-57	79%	aaacaca	-122	79%
EGR1	gctcctccttcggc	-53	85%	gctcctccctcggc	-52	88%	aagcagcccccgc	-139	88%
TFAP2A	ccgccgaaggag	-49	80%	ctccctcggcgg	-48	81%	caccggagggcg	-109	87%
INSM1	ggcctgggagct	-24	79%	ggcctgggagct	-24	79%	aggcaggggaca	-96	81%
TFAP2A	agctcccaggcc	-24	84%	ggcctgggagct	-24	81%	cgccctccggtg	-109	87%
TFAP2B	agctcccaggcc	-24	82%	ggcctgggagct	-24	80%	caccggagggcg	-109	86%
TFAP2B	ggcctgggagct	-24	80%	agctcccaggcc	-24	82%	cgccctccggtg	-109	84%
TFAP2C	agctcccaggcc	-24	82%	ggcctgggagct	-24	80%	caccggagggcg	-109	85%
TFAP2C	ggcctgggagct	-24	80%	agctcccaggcc	-24	82%	cgccctccggtg	-109	84%
ZFX	agctcccaggcctc	-24	90%	agctcccaggcct	-24	90%	ctgtcccctgcctc	-97	81%
MZF1	ctggga	-21	83%	ctggga	-21	83%	agggga	-94	95%
E2F6	gggagcgaggc	-15	85%	gggagtgaggc	-15	85%	cagaggcaggg	-92	79%
EGR1	gcctcgctccctcc	-15	88%	gcctcactccctcc	-15	87%	cccctgcctctggt	-93	82%
KLF16	ctccctcccat	-9	80%	ctccctcccat	-9	80%	ccccagcctcc	-70	82%
SP2	ctccctcccataagc	-9	83%	ctccctcccataaa	-9	82%	cccctgcctctggt	-93	82%
MZF1	ttatgggagg	-6	85%	ttatgggagg	-6	85%	ggaggctggg	-70	79%
EGR1	gccctgccctctcc	4	88%	accctaccctctcc	4	85%	cctcctccccagcc	-67	88%
KLF4	agggcagggc	4	88%	agggtagggt	4	85%	aggctgggga	-68	83%
SP2	gccctgccctctcct	4	89%	accctaccctctcct	4	82%	cctcctccccagcc	-68	84%
CTCF	gagccaggagaggg caggg	5	86%	gagccaggagagggt aggg	5	82%	gggcctgggggagca	-42	81%
EGR1	tgccctctcctggc	8	81%	taccctctcctggc	8	81%	gagcctcctcccca	-64	81%
MZF1	ggagga	22	84%	agagga	22	83%	ggagga	-61	84%
E2F6	cggaggaaaag	30	82%	tggaggaaaag	26	86%	agcagggagcc	-49	82%
TFAP2A	ggccggaggaa	33	81%	ttcctccagcg	29	82%	aggctcgggct	-56	87%
TFAP2B	ggccggaggaa	33	83%	ttcctccagcga	29	80%	agcccgagcct	-56	85%
TFAP2B	ttcctccggcc	33	82%	cgctggaggaa	29	79%	aggctcgggct	-56	83%
TFAP2C	ggccggaggaa	33	84%	cgctggaggaa	29	81%	aggctcgggct	-56	87%
TFAP2C	ttcctccggcc	33	81%	ttcctccagcg	29	79%	agcccgagcct	-56	88%
E2F4	aggccggagga	34	84%	tcgctggagga	30	82%	gcccggggagg	-28	81%
MZF1	ggagga	34	84%	ggagga	30	84%	gggggga	-39	96%
EGR1	attcctcctcctgg	49	85%	atccctcctcctgg	45	86%	gggcctccccccg	-21	82%
SP2	attcctcctcctggt	49	84%	atccctcctcctggt	45	87%	gggcctccccccg	-21	84%
MZF1	ggaggaggaa	50	88%	ggaggaggga	46	84%	ggcgggggga	-16	87%
MZF1	ggagga	54	84%	ggagga	50	84%	gggggg	-15	84%
GCM2	ctcctggt	56	81%	ctcctggt	52	81%	aggcgggg	-13	81%
GCM1	cctggtggtcc	58	79%	cctggtggtcc	54	79%	ggtgcgggagg	-8	81%

GCM2	ctggtggt	59	85%	ctggtggt	55	85%	gtgcggga	-6	90%
SPIB	tagggaa	70	84%	tggggaa	66	92%	tgcggga	-6	79%
PPARG	ggtgggcacaggg	79	84%	agtgggcacagg	75	81%	ggcgggcagagg	6	83%
NFIX	tgtgcccac	84	90%	tgtgcccac	80	90%	tctgcccgc	11	85%
HIC2	gtgcccacc	85	100%	gtgcccact	81	99%	ctgcccgcc	12	92%
THAP1	gtgcccacc	85	88%	gtgcccact	81	88%	cttccctct	5	88%
EGR1	caccctctctcttc	90	81%	ctctctcctccttc	90	85%	cccccgcctcccg	-14	86%
GATA4	ccctctctctt	92	84%	ctctctctcct	88	80%	ccttccctctg	4	80%
EGR1	ccttcagctcctcc	103	79%	ccttcaactcctcc	99	79%	ccgcctcccgcac	-11	87%
EGR1	gctcctccctcttc	109	86%	actcctccctcttc	105	85%	tgcccgcctccgc	13	93%
SP2	gctcctccctcttcc	109	89%	actcctccctcttcc	105	86%	tgcccgcctccgc	13	91%
E2F4	aagagggagga	111	79%	aagagggagga	107	79%	gaggggggagga	27	82%
E2F6	aagagggagga	111	92%	aagagggagga	107	92%	gaggggggagga	27	85%
MZF1	ggagga	111	84%	ggagga	107	84%	ggagga	27	84%
THAP1	cctccctct	112	84%	cctccctct	108	84%	cctcccct	28	80%
MZF1	gaagagggag	113	80%	gaagagggag	109	80%	agagggggag	29	91%
SP2	ctccctcttccaggc	113	80%	ctccctcttccagg	109	80%	cctccgcctcctcc	19	93%
GABPA	ctggaagaggg	115	82%	ctggaagaggg	111	82%	gcggaggcggg	15	81%
E2F6	tggttggagga	136	79%	gggttggagga	132	82%	gagccggagag	37	84%
KLF16	tccacaccacc	160	83%	cacacaccacc	159	84%	gctcctcccc	44	87%
MZF1	tgtgga	160	81%	gtggga	156	81%	gggggga	30	96%
KLF1	cacaccaccct	162	85%	cacaccaccct	161	85%	ggctcctcccc	43	81%
KLF4	agggtggtgt	163	89%	agggtggtgt	162	89%	ggggaggagc	44	87%
THAP1	ccacccttc	166	81%	ccaccctgc	165	79%	cctcccct	47	80%
E2F6	gagaggaaatt	178	80%	gagaggaaatt	177	80%	gcgagggaggg	52	94%
E2F6	gggagagagga	182	85%	gggagagagga	181	85%	gcgagggagcg	60	92%
EGR1	ctctctcccacctc	184	79%	ctctctcccacctc	183	79%	ccctcgccgcctcc	64	90%
E2F4	aggtgggagag	186	84%	aggtgggagag	185	84%	aggcggcgagg	65	82%
EGR1	ctcccacctcccgg	188	81%	ctcccacctcctgg	187	89%	tcgccgcctccccc	67	86%
SP2	ctcccacctcccggt	188	87%	ctcccacctcctgg	187	88%	tcgccgcctccccc	67	85%
MZF1	gtggga	189	81%	gtggga	188	81%	aggggg	76	83%
MZF1	ggaggtggga	189	84%	ggaggtggga	188	84%	ggaggggggag	74	97%
FIGLA	cccacctccc	190	80%	cccacctcct	189	82%	cccccgtgtc	82	79%
HIF1A	ggaggtgg	191	80%	ggaggtgg	190	80%	ccccgtgt	83	80%
ZBTB7B	aggaccaccggg	197	87%	aggaccaccagg	196	85%	tcgcccaccact	206	85%
ZBTB7C	aggaccaccggg	197	88%	aggaccaccagg	196	86%	tcgcccaccact	206	86%
ZBTB7A	aaggaccaccgg	198	85%	aaggaccaccag	197	82%	gtcgcccaccac	205	86%
PPARG	ctgggaaaaggacc	202	79%	ctggggaaagga	201	82%	gagggacacggg	81	81%
ARNTL	gctcacttgt	216	80%	gctcacttgt	215	80%	ggacacgggg	83	81%
USF1	ctcacttgtcc	217	81%	ctcacttgtcc	216	81%	cccccgtgtcc	82	81%
HES2	gtgggacaagtga	218	88%	gtgggacaagtga	217	88%	gagggacacggg	83	83%
ID2	acaagtga	218	84%	acaagtga	217	84%	acacgggg	83	82%
MLXIP	tcacttgt	218	84%	tcacttgt	217	84%	ccccgtgt	83	84%
ARNT	caagtg	219	83%	caagtg	218	83%	cccgtg	84	83%
ARNT	cacttg	219	83%	cacttg	218	83%	cacggg	84	83%

MEIS3	gggacaag	221	83%	gggacaag	220	83%	gggacacg	86	83%
E2F6	gggtgggacaa	222	83%	gggtgggacaa	221	83%	gggagggacac	87	87%
KLF14	tgtcccacccccctc	223	84%	tgtcccacccccctc	222	84%	tgtccctcccctct	88	82%
EGR1	gtcccacccccctct	224	93%	gtcccacccccctcc	223	95%	ctccctcccctcct	126	81%
KLF16	gtcccacccccc	224	89%	gtcccacccccc	223	89%	ctccctcccct	126	83%
KLF4	ggggtgggac	224	91%	ggggtgggac	223	91%	ggggagggag	126	80%
SP2	gtcccacccccctctc	224	93%	gtcccacccccctcc	223	96%	ctccctcccctcctc	126	90%
SP8	gtcccacccccct	224	84%	gtcccacccccct	223	84%	gtccctcccctc	89	80%
TBX5	gggtggga	225	87%	gggtggga	224	87%	aggggata	118	80%
EGR1	cccacccccctctca	226	81%	cacccccctcccgtc	227	81%	ccctcccctcctcc	128	83%
SP8	cccacccccctct	226	86%	cccacccccctcc	225	85%	tcccctcctccc	131	80%
SP2	ccccctctcatccgc	230	81%	ccccctcccgtcag	229	80%	tcccctcctccccg	131	90%
EGR1	cggcatctccctcc	244	82%	cggcaactccctcc	243	83%	tcccctcctccccg	131	84%
E2F6	aggagggagat	248	93%	aggagggagtt	247	90%	aggaggggagg	129	85%
MZF1	gaaggaggga	251	80%	gaaggaggga	250	80%	ggaggggagg	129	90%
MZF1	ggcgggggaa	286	91%	ggcgggggaa	284	91%	ggaggagggg	132	81%
MZF1	gggggga	287	96%	gggggga	285	96%	agggga	131	95%
EGR1	cccccgccccaatt	288	84%	cccccgccccagtt	286	85%	cctcctccccggct	134	86%
KLF16	cccccgcccca	288	87%	cccccgcccca	286	87%	tcccctcctcc	131	84%
KLF4	ggggcggggg	288	93%	ggggcggggg	286	93%	ggggaggagg	134	82%
MZF1	cggggg	288	85%	cggggg	286	85%	ggagga	136	84%
SP2	cccccgccccaattc	288	87%	cccccgccccagtt	286	88%	cctcctccccggct	134	83%
MZF1	agagga	301	83%	agagga	299	83%	cgggga	139	97%
EGR1	cctctgttcccgcc	302	79%	cctctgtcccccgcc	300	84%	ccggctctccctct	142	80%
E2F4	gggcgggaaca	306	98%	gggcggggaca	304	91%	agccggggagg	137	80%
HIC2	gttcccgcc	307	87%	gtccccgcc	305	89%	ctccccggc	138	82%
EGR1	ttcccgcccccttc	308	93%	tccccgcccccttc	306	94%	ggctctccctctca	144	83%
E2F4	ccgggggaagg	317	84%	ccgggggaagg	315	84%	gagagggagag	146	82%
E2F6	ccgggggaagg	317	80%	ccgggggaagg	315	80%	gagagggagag	146	93%
THAP1	cttcccccg	318	81%	cttcccccg	316	81%	tctccctct	147	83%
MZF1	gccgggggaa	319	81%	gccgggggaa	317	81%	tgagagggag	148	83%
SPIB	gggggaa	319	80%	gggggaa	317	80%	agaggga	149	84%
TFAP2A	ttcccccggct	319	85%	ttcccccggcc	317	84%	agtctcaggct	157	86%
TFAP2A	agccgggggaa	319	81%	aggccgggggga	318	81%	agcctgagact	157	85%
TFAP2A	ttcccccggctt	319	81%	ttcccccggcct	317	87%	cgcccgaaagag	191	81%
TFAP2B	ttcccccggct	319	83%	ttcccccggcc	317	80%	agtctcaggct	157	87%
TFAP2C	agccgggggaa	319	84%	ggccgggggaa	317	80%	agtctcaggct	157	86%
EGR1	ccccggcttcctcc	322	82%	ccccggcctcctcc	320	87%	cgcccaccactgc	207	82%
EGR1	tcccctgcctctcc	346	81%	tcccctacctctcc	344	81%	ccaccactgccag	210	82%
NFIX	ctctccaat	354	85%	ctctccagt	352	82%	actgccagc	215	94%
MZF1	tggaga	355	81%	tggaga	353	81%	tggggg	231	88%
NFIC	ttggag	356	88%	ctggag	354	84%	ctggag	243	84%
TFAP2A	ctccccggggct	371	89%	agccccggagag	369	81%	ccccccagcgct	226	85%
TFAP2B	agccccggggag	371	87%	agccccggagag	369	79%	ccccccagcgct	226	82%
TFAP2C	ctccccggggct	371	86%	agccccggagag	369	79%	ccccccagcgct	226	80%

EGR1	ccccccttcctct	395	81%	tctctgtccccgcc	402	83%	cccccacctcccc	279	86%
SOX10	ctctgt	405	86%	ctctgt	403	86%	cattgg	273	84%
HIC1	gtccccgcc	409	83%	gtccccgcc	407	83%	acgccccca	287	79%
HIC2	gtccccgcc	409	89%	gtccccgcc	407	89%	acgccccca	287	82%
KLF14	gtccccgccccctc	409	88%	gtccccgccccctc	407	88%	tcccacgccccca	286	89%
KLF16	tccccgccccc	410	93%	tccccgccccc	408	93%	cccacgccccc	288	97%
KLF4	ggggcgggga	410	93%	ggggcgggga	408	93%	ggggcgtggg	289	93%
SP2	tccccgccccctccc	410	97%	tccccgccccctcc	408	97%	cccccacctcccc	278	90%
SP8	tccccgccccct	410	91%	tccccgccccct	408	91%	cccacctcccca	279	81%
E2F6	gggaggggggcg	414	83%	gggaggggggcg	412	83%	gcgtgggaaaa	292	91%
MZF1	ggaggggggcg	414	86%	ggaggggggcg	412	86%	ggcgtgggaa	291	81%
SP8	gccccctcccct	415	83%	gccccctcccct	413	83%	cccccacctccc	281	81%
EGR1	ccccctccctgga	416	89%	ccccctccctgg	414	88%	ccttttcccacgcc	291	86%
KLF4	ggggaggggg	416	88%	ggggaggggg	414	88%	aggacggggt	302	84%
SP2	ccccctccctggac	416	88%	ccccctccctgg	414	89%	cccacgcccccac	284	83%
ZEB1	cctcccctg	419	87%	cctcccctg	417	87%	ccgcacctg	317	97%
MZF1	agggga	421	95%	agggga	419	95%	tggggt	313	81%
KLF1	gactccgccct	428	87%	ggctccgccct	426	88%	aaccccgtcct	302	80%
SP2	actccgccctcctcc	429	89%	gctccgccctcctc	427	92%	ttcccacgcccccа	286	80%
SP8	actccgccctcc	429	81%	gctccgccctcc	427	82%	cccacgcccccа	287	95%
THAP1	ccgccctcc	432	90%	ccgccctcc	430	90%	acgccccca	287	83%
SP2	cgccctcctcccagc	433	87%	cgccctcctcccg	431	88%	accccgtccttttcc	297	81%
MZF1	ggagga	438	84%	ggagga	436	84%	tgcgga	321	81%
KLF16	tcccagcccct	441	83%	tcccggcccct	439	81%	gaaccgcccccc	329	89%
KLF4	ggggctggga	441	82%	ggggccggga	439	81%	ggggcggttc	330	84%
SP2	tcccagcccctcccc	441	86%	tcccggccccttc	439	86%	cccccactccgca	319	79%
KLF4	ggggaggggc	446	94%	gggaaggggc	444	82%	ggagtggggg	324	81%
HIC2	gttcccctc	458	80%	ttcccctcc	456	80%	gcccccact	326	79%
SPIB	aggggaa	458	95%	aggggaa	456	95%	tgcggag	321	79%
SP2	tcccctcctagcatc	459	80%	tcccctcctcgcat	457	84%	gaaccgcccccac	325	84%
ETV5	aggggatgct	468	81%	aggggatgcg	466	82%	tcgggaagcc	338	81%
SPIB	aggggat	471	82%	aggggat	469	82%	tcgggaa	338	89%
SP2	tcccctctcttccct	472	81%	tcccctctcctccct	470	84%	cccccacctccata	372	84%
MZF1	ggagggaaga	479	80%	ggagggagga	477	84%	ggaggtgggg	376	81%
THAP1	cttccctcc	480	89%	cctccctcc	478	85%	cctccataa	372	83%
TBX15	aggaggga	482	79%	aggaggga	480	79%	aggtgggg	378	83%
TBX5	aggaggga	482	79%	aggaggga	480	79%	aggtgggg	378	86%
KLF4	agggaggagg	484	83%	agggaggagg	482	83%	gaggtggggg	377	84%
E2F6	agcagggagga	486	84%	agcagggagga	484	84%	aggtggggggc	378	80%
MZF1	ggagga	486	84%	ggagga	484	84%	tggggg	381	88%
THAP1	cctccctgc	487	82%	cctccctgc	485	82%	cagcccccc	382	80%
THAP1	tctccctca	496	88%	tctccctca	494	88%	tggcccagc	393	79%
GCM2	ctgaggga	498	81%	ctgaggga	496	81%	gtgctggg	396	87%

S-6: List of primers used in this work

PRIMER	SEQUENCE
Region S4 Forward	agctatgaagccgttctctgctc
Region S4 Reverse	agcactcctacctacaagtgcacag
Region S3 Forward	tccttccttccctaccttgcc
Region S3 Reverse	gcccagagacagacagacagacat
Region S2 Forward	acaccttctgtgggctgtatagcgg
Region S2 Reverse	gtctaacccacagcctgccct
Region S1 Forward	catgcaagcaagggaggtgtc
Region S1 Reverse	gtggctcagtggttgagccc
Pgl3-S3(-319/+595) Forward	cccagtgcaagtgcaggtgcgtgtgtgatcaaggagggg
Pgl3-S3(-319/+595) Reverse	cccctccttgatcacacacgcacctgcacttgcactggg
Pgl3-S3(+113/+595) Forward	ccagtgcaagtgcaggtgcccctcttccaggctgtcc
Pgl3-S3(+113/+595) Reverse	ggacagcctggaagaggggcacctgcacttgcactgg
Pgl3-S3(+424/+595) Forward	cccagtgcaagtgcaggtgcctggactccgccctc
Pgl3-S3(+424/+595) Reverse	gagggcggagtccaggcaacctggcacttgcactggg
Mouse Hprt Forward	agcgtcgtgattagcgatga
Mouse Hprt Reverse	tccaaatcctcggcataatga
Mouse Ncx2 Forward	agtggatgatgaagagtatgagaagaag
Mouse Ncx2 Reverse	ttggttgagtagcagagctgaga
Rat Hprt Forward	gaaagaacgtcttgattgttgaagatat
Rat Hprt Reverse	gagaggtccttttcaccagcaa
Rat Ncx2 Forward	cagcaccttctacgtggattacc
Rat Ncx2 Reverse	gccctcgctgtactcataatcg
Human Mgmt Forward	ggtccggagcccctgat
Human Mgmt Reverse	gctggtggaaataggcattca
Human Hprt Forward	aggaggccgcacacctttcc
Human Hprt Reverse	caaggcgtggctgggctctc
Human Ncx2 Forward	gaccagtcaacaggggacat
Human Ncx2 Reverse	cgaccttgaccatctttgga
A-F1	gagtttaggttttgttgtgatgatgg
A-F2	ggtttttgagtgttggggttatag
A-R1	caactcaacctactcaaaacacc
B-F3	tgtttttgtaaggtgttttgagtagg
B-F4	gagttgtgaatttggaaagatttgt
B-R2	cttaacaaaactaaatctacaaattaccc
C-F5	gggtaatttgtagatttagttttgttaag
C-R3	cacaaaaaaagtaaaaaaaaaaccacc
C-R4	caaaaacaacctaaaaaaaaaaaaaac
D-F6	agtttttttttttttaggttgtttttg
D-R5	gctaaactcttacctacttacaac

D-R6	ctcaaaaaaccccctacac
E-F7	ggttgtaaagtaggtaagagtttag
E-R7	cctcctccaaatccctatacatc
E-R8	ctataaacaaaaccaaccccaaaaaatc
Pgl3-S3(+113/+595)Mutsp1	ctgcctctccaatctagccactagtccggggctgctttctcag
Pgl3-S3(+113/+595)Mutsp1	ctgagaaagcagccccggactagtggctagattggagaggcag

9. Bibliography

Abreu, C. M., R. Kumar, et al. (2013). "Site-specific phosphorylation of the DNA damage response mediator rad9 by cyclin-dependent kinases regulates activation of checkpoint kinase 1." PLoS Genet **9**(4): e1003310.

Amemiya-Kudo, M., H. Shimano, et al. (2000). "Promoter analysis of the mouse sterol regulatory element-binding protein-1c gene." Journal of Biological Chemistry **275**(40): 31078-31085.

Ammanamanchi, S., J. W. Freeman, et al. (2003). "Acetylated sp3 is a transcriptional activator." Journal of Biological Chemistry **278**(37): 35775-35780.

Araújo, I. M. and C. M. Carvalho (2005). "Role of nitric oxide and calpain activation in neuronal death and survival." Current Drug Targets-CNS & Neurological Disorders **4**(4): 319-324.

Armoundas, A. A., J. Rose, et al. (2007). "Cellular and molecular determinants of altered Ca2+ handling in the failing rabbit heart: primary defects in SR Ca2+ uptake and release mechanisms." American Journal of Physiology-Heart and Circulatory Physiology **292**(3): H1607-H1618.

Atchison, M. L. (1988). "Enhancers: mechanisms of action and cell specificity." Annual review of cell biology **4**(1): 127-153.

Baker, P., M. Blaustein, et al. (1969). "The influence of calcium on sodium efflux in squid axons." The Journal of Physiology **200**(2): 431.

Barnes, K. V., G. Cheng, et al. (1997). "Cloning of cardiac, kidney, and brain promoters of the feline ncx1 gene." Journal of Biological Chemistry **272**(17): 11510-11517.

Ben-Shachar, D. and R. Karry (2007). "Sp1 expression is disrupted in schizophrenia; a possible mechanism for the abnormal expression of mitochondrial complex I genes, NDUFV1 and NDUFV2." PLoS One **2**(9): e817.

Bengoechea-Alonso, M. T. and J. Ericsson (2007). "SREBP in signal transduction: cholesterol metabolism and beyond." Current opinion in cell biology **19**(2): 215-222.

Bernstein, B. E., A. Meissner, et al. (2007). "The mammalian epigenome." Cell **128**(4): 669-681.

Bers, D. M. (2008). "Calcium cycling and signaling in cardiac myocytes." Annu. Rev. Physiol. **70**: 23-49.

Bird, A. (2007). "Perceptions of epigenetics." Nature **447**(7143): 396-398.

Braff, D. L., M. A. Geyer, et al. (2001). "Impact of prepulse characteristics on the detection of sensorimotor gating deficits in schizophrenia." Schizophrenia research **49**(1): 171-178.

Braun, H., R. Koop, et al. (2001). "Transcription factor Sp3 is regulated by acetylation." Nucleic acids research **29**(24): 4994-5000.

Cairns, P. (2009). "5′-azacytidine expression arrays." DNA Methylation: Methods and Protocols: 165-174.

Camargo, N., A. B. Smit, et al. (2009). "SREBPs: SREBP function in glia–neuron interactions." FEBS journal **276**(3): 628-636.

Canitano, A., M. Papa, et al. (2002). "Brain Distribution of the Na+/Ca2+ Exchanger-Encoding Genes NCX1, NCX2, and NCX3 and Their Related Proteins in the Central Nervous System." Annals of the New York Academy of Sciences **976**(1): 394-404.

Cartharius, K., K. Frech, et al. (2005). "MatInspector and beyond: promoter analysis based on transcription factor binding sites." Bioinformatics **21**(13): 2933-2942.

Ceccarelli, M., F. P. Barthel, et al. (2016). "Molecular profiling reveals biologically discrete subsets and pathways of progression in diffuse glioma." Cell **164**(3): 550-563.

Cedar, H. and Y. Bergman (2009). "Linking DNA methylation and histone modification: patterns and paradigms." Nature Reviews Genetics **10**(5): 295-304.

Chalkley, G. E. and C. P. Verrijzer (1999). "DNA binding site selection by RNA polymerase II TAFs: a TAFII250–TAFII150 complex recognizes the Initiator." The EMBO journal **18**(17): 4835-4845.

Chan, H. M. and N. B. La Thangue (2001). "p300/CBP proteins: HATs for transcriptional bridges and scaffolds." Journal of cell science **114**(13): 2363-2373.

Chen, B.-S. and M. Hampsey (2002). "Transcription activation: unveiling the essential nature of TFIID." Current biology **12**(18): R620-R622.

Chen, Y., M. Kundakovic, et al. (2007). "Induction of the reelin promoter by retinoic acid is mediated by Sp1." Journal of neurochemistry **103**(2): 650-665.

Christophorou, M. A., G. Castelo-Branco, et al. (2014). "Citrullination regulates pluripotency and histone H1 binding to chromatin." Nature **507**(7490): 104-108.

Conway, S. J., J. Wang, et al. (2002). "Role of Sodium-Calcium Exchanger (Ncx1) in Embryonic Heart Development." Annals of the New York Academy of Sciences **976**(1): 268-281.

Dunah, A. W., H. Jeong, et al. (2002). "Sp1 and TAFII130 transcriptional activity disrupted in early Huntington's disease." Science **296**(5576): 2238-2243.

Durkin, J. T., D. C. Ahrens, et al. (1991). "Purification and amino-terminal sequence of the bovine cardiac sodium-calcium exchanger: evidence for the presence of a signal sequence." Archives of biochemistry and biophysics **290**(2): 369-375.

Dworkin, S. and T. Mantamadiotis (2010). "Targeting CREB signalling in neurogenesis." Expert opinion on therapeutic targets **14**(8): 869-879.

Dyck, C., K. Maxwell, et al. (1998). "Structure-Function Analysis of CALX1. 1, a Na+-Ca2+ Exchanger fromDrosophila MUTAGENESIS OF IONIC REGULATORY SITES." Journal of Biological Chemistry **273**(21): 12981-12987.

Eberlé, D., B. Hegarty, et al. (2004). "SREBP transcription factors: master regulators of lipid homeostasis." Biochimie **86**(11): 839-848.

Eltzschig, H. K., D. Köhler, et al. (2009). "Central role of Sp1-regulated CD39 in hypoxia/ischemia protection." Blood **113**(1): 224-232.

Faigle, R. and H. Song (2013). "Signaling mechanisms regulating adult neural stem cells and neurogenesis." Biochimica et Biophysica Acta (BBA)-General Subjects **1830**(2): 2435-2448.

Formisano, L., N. Guida, et al. (2015). "Sp3/REST/HDAC1/HDAC2 complex represses and Sp1/HIF-1/p300 complex activates ncx1 gene transcription, in brain ischemia and in ischemic brain preconditioning, by epigenetic mechanism." Journal of Neuroscience **35**(19): 7332-7348.

Gabellini, N., S. Bortoluzzi, et al. (2002). "The human SLC8A3 gene and the tissue-specific Na+/Ca 2+ exchanger 3 isoforms." Gene **298**(1): 1-7.

Gabellini, N., S. Bortoluzzi, et al. (2003). "Control of the Na+/Ca2+ exchanger 3 promoter by cyclic adenosine monophosphate and Ca2+ in differentiating neurons." Journal of neurochemistry **84**(2): 282-293.

Gartel, A. L., X. Ye, et al. (2001). "Myc represses the p21 (WAF1/CIP1) promoter and interacts with Sp1/Sp3." Proceedings of the National Academy of Sciences **98**(8): 4510-4515.

Gasic, G. P. (1994). "Basic-helix-loop-helix transcription factor and sterol sensor in a single membrane-bound molecule." Cell **77**(1): 17-19.

Gershenzon, N. I. and I. P. Ioshikhes (2005). "Synergy of human Pol II core promoter elements revealed by statistical sequence analysis." Bioinformatics **21**(8): 1295-1300.

Gill, G., E. Pascal, et al. (1994). "A glutamine-rich hydrophobic patch in transcription factor Sp1 contacts the dTAFII110 component of the Drosophila TFIID complex and mediates transcriptional activation." Proceedings of the National Academy of Sciences **91**(1): 192-196.

Gomez-Villafuertes, R., B. Torres, et al. (2005). "Downstream regulatory element antagonist modulator regulates Ca2+ homeostasis and viability in cerebellar neurons." Journal of Neuroscience **25**(47): 10822-10830.

Gomez, R., N. Gabellini, et al. (2004). DREAM, a calcium-dependent transcriptional repressor, regulates NCX3 gene expression. EMBO Workshop on Calcium Signaling and Disease, Capri, Italy.

Görisch, S. M., M. Wachsmuth, et al. (2005). "Histone acetylation increases chromatin accessibility." Journal of cell science **118**(24): 5825-5834.

Hagen, G., J. Dennig, et al. (1995). "Functional analyses of the transcription factor Sp4 reveal properties distinct from Sp1 and Sp3." Journal of Biological Chemistry **270**(42): 24989-24994.

Hagen, G., S. Müller, et al. (1994). "Sp1-mediated transcriptional activation is repressed by Sp3." The EMBO journal **13**(16): 3843.

He, Z., Q. Tong, et al. (1998). "Cloning, expression, and characterization of the squid Na+–Ca2+ exchanger (NCX-SQ1)." The Journal of general physiology **111**(6): 857-873.

Hilgemann, D., S. Matsuoka, et al. (1992). "Steady-state and dynamic properties of cardiac sodium-calcium exchange. Sodium-dependent inactivation." The Journal of general physiology **100**(6): 905-932.

Hryshko, L. V., D. A. Nicoll, et al. (1993). "Biosynthesis and initial processing of the cardiac sarcolemmal Na+-Ca2+ exchanger." Biochimica et Biophysica Acta (BBA)-Biomembranes **1151**(1): 35-42.

Hua, X., C. Yokoyama, et al. (1993). "SREBP-2, a second basic-helix-loop-helix-leucine zipper protein that stimulates transcription by binding to a sterol regulatory element." Proceedings of the National Academy of Sciences **90**(24): 11603-11607.

Jeon, D., K. Chu, et al. (2008). "Na+/Ca 2+ exchanger 2 is neuroprotective by exporting Ca 2+ during a transient focal cerebral ischemia in the mouse." Cell calcium **43**(5): 482-491.

Jeon, D., Y.-M. Yang, et al. (2003). "Enhanced learning and memory in mice lacking Na+/Ca 2+ exchanger 2." Neuron **38**(6): 965-976.

Juven-Gershon, T. and J. T. Kadonaga (2010). "Regulation of gene expression via the core promoter and the basal transcriptional machinery." Developmental biology **339**(2): 225-229.

Kadonaga, J. T., K. R. Carner, et al. (1987). "Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain." Cell **51**(6): 1079-1090.

Kang, T. M. and D. W. Hilgemann (2004). "Multiple transport modes of the cardiac Na+/Ca2+ exchanger." Nature **427**(6974): 544-548.

Khorasanizadeh, S. (2004). "The nucleosome: from genomic organization to genomic regulation." Cell **116**(2): 259-272.

Kingsley, C. and A. Winoto (1992). "Cloning of GT box-binding proteins: a novel Sp1 multigene family regulating T-cell receptor gene expression." Molecular and Cellular Biology **12**(10): 4251-4261.

Kintner, D. B., Y. Wang, et al. (2007). "Role of membrane ion transport proteins in cerebral ischemic damage." Front Biosci **12**: 762-770.

Kiryu-Seo, S., R. Kato, et al. (2008). "Neuronal injury-inducible gene is synergistically regulated by ATF3, c-Jun, and STAT3 through the interaction with Sp1 in damaged neurons." Journal of Biological Chemistry **283**(11): 6988-6996.

Kofuji, P., W. Lederer, et al. (1994). "Mutually exclusive and cassette exons underlie alternatively spliced isoforms of the Na/Ca exchanger." Journal of Biological Chemistry **269**(7): 5145-5149.

Kouzarides, T. (2007). "Chromatin modifications and their function." Cell **128**(4): 693-705.

Kraev, A., I. Chumakov, et al. (1996). "The organization of the human gene NCX1 encoding the sodium–calcium exchanger." Genomics **37**(1): 105-112.

Kristjansdottir, G., J. K. Sandling, et al. (2008). "Interferon regulatory factor 5 (IRF5) gene variants are associated with multiple sclerosis in three distinct populations." Journal of medical genetics **45**(6): 362-369.

Lee, S.-L., A. Yu, et al. (1994). "Tissue-specific expression of Na (+)-Ca2+ exchanger isoforms." Journal of Biological Chemistry **269**(21): 14849-14852.

Lettice, L. A., S. J. Heaney, et al. (2003). "A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly." Human molecular genetics **12**(14): 1725-1735.

Li, Z., S. Matsuoka, et al. (1994). "Cloning of the NCX2 isoform of the plasma membrane Na (+)-Ca2+ exchanger." Journal of Biological Chemistry **269**(26): 17434-17439.

Liu, P., M. Begley, et al. (2014). "Cell-cycle-regulated activation of Akt kinase by phosphorylation at its carboxyl terminus." Nature **508**(7497): 541-545.

Lonard, D. M. and B. W. O'Malley (2005). "Expanding functional diversity of the coactivators." Trends in biochemical sciences **30**(3): 126-132.

Loots, G. G., M. Kneissel, et al. (2005). "Genomic deletion of a long-range bone enhancer misregulates sclerostin in Van Buchem disease." Genome research **15**(7): 928-935.

Ma, L., L. Song, et al. (2004). "Transcriptional regulation of the mouse gene encoding the α-4 subunit of the GABAA receptor." Journal of Biological Chemistry **279**(39): 40451-40461.

Mabuchi, T., K. Kitagawa, et al. (2001). "Phosphorylation of cAMP response element-binding protein in hippocampal neurons as a protective response after exposure to glutamate in vitro and ischemia in vivo." Journal of Neuroscience **21**(23): 9204-9213.

Magaña, M. M. and T. F. Osborne (1996). "Two tandem binding sites for sterol regulatory element binding proteins are required for sterol regulation of fatty-acid synthase promoter." Journal of Biological Chemistry **271**(51): 32689-32694.

Maniatis, T., E. F. Fritsch, et al. (1982). Molecular cloning: a laboratory manual, Cold Spring harbor laboratory Cold Spring Harbor, NY.

Mao, X., S. H. Yang, et al. (2007). "Glutamate receptor activation evokes calpain-mediated degradation of Sp3 and Sp4, the prominent Sp-family transcription factors in neurons." Journal of neurochemistry **100**(5): 1300-1314.

Marin, M., A. Karis, et al. (1997). "Transcription factor Sp1 is essential for early embryonic development but dispensable for cell growth and differentiation." Cell **89**(4): 619-628.

Martin, L. J. (2010). "Mitochondrial and cell death mechanisms in neurodegenerative diseases." Pharmaceuticals **3**(4): 839-915.

Mastrangelo, I. A., A. J. Courey, et al. (1991). "DNA looping and Sp1 multimer links: a mechanism for transcriptional synergism and enhancement." Proceedings of the National Academy of Sciences **88**(13): 5670-5674.

Matsuda, T., N. Arakawa, et al. (2001). "SEA0400, a novel and selective inhibitor of the Na+-Ca2+ exchanger, attenuates reperfusion injury in the in vitro and in vivo cerebral ischemic models." Journal of Pharmacology and Experimental Therapeutics **298**(1): 249-256.

Matys, V., E. Fricke, et al. (2003). "TRANSFAC®: transcriptional regulation, from patterns to profiles." Nucleic acids research **31**(1): 374-378.

Mayr, B. and M. Montminy (2001). "Transcriptional regulation by the phosphorylation-dependent factor CREB." Nature reviews Molecular cell biology **2**(8): 599-609.

Merz, K., S. Herold, et al. (2011). "CREB in adult neurogenesis–master and partner in the development of adult-born neurons?" European Journal of Neuroscience **33**(6): 1078-1086.

Miki, N., M. Ikuta, et al. (2004). "Hypoxia-induced activation of the retinoic acid receptor-related orphan receptor α4 gene by an interaction between hypoxia-inducible factor-1 and Sp1." Journal of Biological Chemistry **279**(15): 15025-15031.

Molinaro, P., O. Cuomo, et al. (2008). "Targeted disruption of Na+/Ca2+ exchanger 3 (NCX3) gene leads to a worsening of ischemic brain damage." Journal of Neuroscience **28**(5): 1179-1184.

Montminy, M. R., K. A. Sevarino, et al. (1986). "Identification of a cyclic-AMP-responsive element within the rat somatostatin gene." Proceedings of the National Academy of Sciences **83**(18): 6682-6686.

Murata, Y., H. G. Kim, et al. (1994). "Negative regulation of Sp1 trans-activation is correlated with the binding of cellular proteins to the amino terminus of the Sp1 trans-activation domain." Journal of Biological Chemistry **269**(32): 20674-20681.

Nakamura, M. T. and T. Y. Nara (2004). "Structure, function, and dietary regulation of Δ6, Δ5, and Δ9 desaturases." Annu. Rev. Nutr. **24**: 345-376.

Nakasaki, Y., T. Iwamoto, et al. (1993). "Cloning of the rat aortic smooth muscle Na+/Ca2+ exchanger and tissue-specific expression of isoforms." Journal of Biochemistry **114**(4): 528-534.

Nan, X., H.-H. Ng, et al. (1998). "Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex." Nature **393**(6683): 386-389.

Nicholas, S. B., W. Yang, et al. (1998). "Alternative promoters and cardiac muscle cell-specific expression of the Na+/Ca2+ exchanger gene." American Journal of Physiology-Heart and Circulatory Physiology **274**(1): H217-H232.

Nicoll, D. A., S. Longoni, et al. (1990). "Molecular cloning and functional expression of the cardiac sarcolemmal Na+-Ca2+ exchanger." Science **250**(4980): 562-566.

Nicoll, D. A., B. D. Quednau, et al. (1996). "Cloning of a third mammalian Na+-Ca2+ exchanger, NCX3." Journal of Biological Chemistry **271**(40): 24914-24921.

Niggli, E. and W. Lederer (1991). "Molecular operations of the sodium-calcium exchanger revealed by conformation currents." Nature **349**(6310): 621.

Nohturfft, A. and S. C. Zhang (2009). "Coordination of lipid metabolism in membrane biogenesis." Annual Review of Cell and Developmental **25**: 539-566.

Ooi, S. K., C. Qiu, et al. (2007). "DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA." Nature **448**(7154): 714-717.

Ortega-Martínez, S. (2015). "A new perspective on the role of the CREB family of transcription factors in memory consolidation via adult hippocampal neurogenesis." Frontiers in molecular neuroscience **8**.

Ossareh-Nazari, B., C. Gwizdek, et al. (2001). "Protein export from the nucleus." Traffic **2**(10): 684-689.

Papa, M., A. Canitano, et al. (2003). "Differential expression of the Na+-Ca2+ exchanger transcripts and proteins in rat brain regions." Journal of Comparative Neurology **461**(1): 31-48.

Pfaffl, M. W. (2001). "A new mathematical model for relative quantification in real-time RT–PCR." Nucleic acids research **29**(9): e45-e45.

Philipson, K. D., S. Longoni, et al. (1988). "Purification of the cardiac Na+ Ca2+ exchange protein." Biochimica et Biophysica Acta (BBA)-Biomembranes **945**(2): 298-306.

Philipson, K. D. and D. A. Nicoll (2000). "Sodium-calcium exchange: a molecular perspective." Annual review of physiology **62**(1): 111-133.

Pignataro, G., R. Gala, et al. (2004). "Two sodium/calcium exchanger gene products, NCX1 and NCX3, play a major role in the development of permanent focal cerebral ischemia." Stroke **35**(11): 2566-2570.

Pinacho, R., G. Saia, et al. (2015). "Transcription factor SP4 phosphorylation is altered in the postmortem cerebellum of bipolar disorder and schizophrenia subjects." European Neuropsychopharmacology **25**(10): 1650-1660.

Pogwizd, S. M. (2003). "Clinical potential of sodium-calcium exchanger inhibitors as antiarrhythmic agents." Drugs **63**(5): 439-452.

Qu, M., H. Jiao, et al. (2010). "Molecular genetic and epigenetic analysis of NCX2/SLC8A2 at 19q13. 3 in human gliomas." Neuropathology and applied neurobiology **36**(3): 198-210.

Quednau, B. D., D. Nicoll, et al. (1997). "Tissue specificity and alternative splicing of the Na+/Ca2+ exchanger isoforms NCX1, NCX2, and NCX3 in rat." American Journal of Physiology-Cell Physiology **272**(4): C1250-C1261.

Ramos, B., A. Valín, et al. (2009). "Sp4-dependent repression of neurotrophin-3 limits dendritic branching." Molecular and Cellular Neuroscience **42**(2): 152-159.

Recillas-Targa, F., M. J. Pikaart, et al. (2002). "Position-effect protection and enhancer blocking by the chicken β-globin insulator are separable activities." Proceedings of the National Academy of Sciences **99**(10): 6883-6888.

Reddington, J. P., S. M. Perricone, et al. (2013). "Redistribution of H3K27me3 upon DNA hypomethylation results in de-repression of Polycomb target genes." Genome biology **14**(3): R25.

Reed, B. D., A. E. Charos, et al. (2008). "Genome-wide occupancy of SREBP1 and its partners NFY and SP1 reveals novel functional roles and combinatorial regulation of distinct classes of genes." PLoS Genet **4**(7): e1000133.

Ren, X. and K. D. Philipson (2013). "The topology of the cardiac Na+/Ca 2+ exchanger, NCX1." Journal of molecular and cellular cardiology **57**: 68-71.

Reuter, H. and N. Seitz (1968). "The dependence of calcium efflux from cardiac muscle on temperature and external ion composition." The Journal of Physiology **195**(2): 451.

Robertson, K. D. (2005). "DNA methylation and human disease." Nature Reviews Genetics **6**(8): 597-610.

Rome, S., V. Lecomte, et al. (2008). "Microarray analyses of SREBP-1a and SREBP-1c target genes identify new regulatory pathways in muscle." Physiological genomics **34**(3): 327-337.

Ru, P., P. Hu, et al. (2016). "Feedback Loop Regulation of SCAP/SREBP-1 by miR-29 Modulates EGFR Signaling-Driven Glioblastoma Growth." Cell reports **16**(6): 1527-1535.

Ryu, H., J. Lee, et al. (2003). "Sp1 and Sp3 are oxidative stress-inducible, antideath transcription factors in cortical neurons." Journal of Neuroscience **23**(9): 3597-3606.

Safe, S. and M. Abdelrahim (2005). "Sp transcription factor family and its role in cancer." European journal of cancer **41**(16): 2438-2448.

Santpere, G., M. Nieto, et al. (2006). "Abnormal Sp1 transcription factor expression in Alzheimer disease and tauopathies." Neuroscience letters **397**(1): 30-34.

Scheller, T., A. Kraev, et al. (1998). "Cloning of the multipartite promoter of the sodium-calcium exchanger gene NCX1 and characterization of its activity in vascular smooth muscle cells." Journal of Biological Chemistry **273**(13): 7643-7649.

Schneider, R. and R. Grosschedl (2007). "Dynamics and interplay of nuclear architecture, genome organization, and gene expression." Genes & development **21**(23): 3027-3043.

Segal, E. and J. Widom (2009). "From DNA sequence to transcriptional behaviour: a quantitative approach." Nature Reviews Genetics **10**(7): 443-456.

Shih, J. and K. Chen (2004). "Regulation of MAO-A and MAO-B gene expression." Current medicinal chemistry **11**(15): 1995-2005.

Sipido, K. R., P. G. Volders, et al. (2002). "Altered Na/Ca exchange activity in cardiac hypertrophy and heart failure: a new target for therapy?" Cardiovascular research **53**(4): 782-805.

Sirabella, R., A. Secondo, et al. (2009). "Anoxia-induced NF-kB-dependent upregulation of NCX1 contributes to Ca2+ refilling into endoplasmic reticulum in cortical neurons." Stroke **40**(3): 922-929.

Staiano, R. I., F. Granata, et al. (2009). "Expression and function of Na+/Ca2+ exchangers 1 and 3 in human macrophages and monocytes." European journal of immunology **39**(5): 1405-1418.

Steffan, J. S., N. Agrawal, et al. (2004). "SUMO modification of Huntingtin and Huntington's disease pathology." Science **304**(5667): 100-104.

Su, K., M. D. Roos, et al. (1999). "An N-terminal region of Sp1 targets its proteasome-dependent degradation in vitro." Journal of Biological Chemistry **274**(21): 15194-15202.

Supp, D. M., D. P. Witte, et al. (1996). "Sp4, a member of the Sp1-family of zinc finger transcription factors, is required for normal murine growth, viability, and male fertility." Developmental biology **176**(2): 284-299.

Suske, G. (1999). "The Sp-family of transcription factors." Gene **238**(2): 291-300.

Suzuki, M. M. and A. Bird (2008). "DNA methylation landscapes: provocative insights from epigenomics." Nature Reviews Genetics **9**(6): 465-476.

Vaissière, T., C. Sawan, et al. (2008). "Epigenetic interplay between histone modifications and DNA methylation in gene silencing." Mutation Research/Reviews in Mutation Research **659**(1): 40-48.

Valsecchi, V., G. Pignataro, et al. (2011). "NCX1 is a novel target gene for hypoxia-inducible factor-1 in ischemic brain preconditioning." Stroke: STROKEAHA. 110.597583.

Vlieghe, D., A. Sandelin, et al. (2006). "A new generation of JASPAR, the open-access repository for transcription factor binding site profiles." Nucleic acids research **34**(suppl 1): D95-D97.

Wasserman, W. W. and A. Sandelin (2004). "Applied bioinformatics for the identification of regulatory elements." Nature Reviews Genetics **5**(4): 276-287.

Weaver, R. and P. Hedrick (1997). Genetics. Wm. C, Brown Publishers, USA.

Wilson, J. R. and G. S. Chen (2007). "Dirichlet-multinomial model with varying response rates over time." Journal of Data Science **5**(3): 413-423.

Yajima, S., S.-H. Lee, et al. (1998). "Sp family transcription factors regulate expression of rat D2 dopamine receptor gene." DNA and cell biology **17**(5): 471-479.

Yamashima, T. (2012). "'PUFA–GPR40–CREB signaling'hypothesis for the adult primate neurogenesis." Progress in lipid research **51**(3): 221-231.

Yu, B., P. K. Datta, et al. (2003). "Stability of the Sp3–DNA complex is promoter-specific: Sp3 efficiently competes with Sp1 for binding to promoters containing multiple Sp-sites." Nucleic acids research **31**(18): 5368-5376.

Zhang, X., D. T. Odom, et al. (2005). "Genome-wide analysis of cAMP-response element binding protein occupancy, phosphorylation, and target gene activation in human tissues." Proceedings of the National Academy of Sciences of the United States of America **102**(12): 4459-4464.

Zhao, Q., G. Rank, et al. (2009). "PRMT5-mediated methylation of histone H4R3 recruits DNMT3A, coupling histone and DNA methylation in gene silencing." Nature structural & molecular biology **16**(3): 304-311.

Zhou, X., J. Long, et al. (2005). "Reduced expression of the Sp4 gene in mice causes deficits in sensorimotor gating and memory associated with hippocampal vacuolization." Molecular psychiatry **10**(4): 393-406.

Zhou, X., Z. Nie, et al. (2010). "Reduced NMDAR1 expression in the Sp4 hypomorphic mouse may contribute to endophenotypes of human psychiatric disorders." Human molecular genetics: ddq298.

Zhu, H., G. Wang, et al. (2016). "Transcription factors as readers and effectors of DNA methylation." Nature Reviews Genetics **17**(9): 551-565.