PhD PROGRAM IN TRANSLATIONAL
AND MOLECULAR MEDICINE
DIMET
University of Milano-Bicocca
School of Medicine and Faculty of Science

Functional characterization of SMARCA2 gene

Coordinator: Prof. Andrea Biondi
Tutor: Prof. Silvia Barabino

Dr. Aurora Rigamonti
Matr. No. 063184

XXVIII Cycle
Academic Year 2014-2015

Table of contents

Chapter 1.

General Introduction

1.1 Transcriptional and post-transcriptional mechanisms that influence mammalian complexity.

Comparative genome analyses indicate that increases in gene number do not account for increases in morphological and behavioural complexity. For example, the simple nematode worm, *Caenorhabditis elegans*, possesses nearly 20,000 genes but lacks the full range of cell types and tissues seen in the fruitfly Drosophila, which contains fewer than 14,000 genes. Indeed, the revelation that the human genome contains only, 30,000 protein-coding genes precipitated a frenzy of speculation regarding the molecular basis of organismal complexity (Levine 2003). The flow of genomic information, from DNA to RNA to protein, is a highly complex process in mammalian cells (Moore and Proudfoot 2009).

Alternative transcription and alternative splicing (AS) are the main mechanisms that generate proteomic diversity in multicellular eukaryotes, particularly in mammals. The architecture of gene loci, location of transcription initiation and termination sites, and the presence of splicing sites and regulatory elements

define the structures of pre mRNA and mature transcripts. Alternative transcription initiation and termination shape the structures of 5’ and 3’ terminal regions of transcripts and play distinct and independent roles in the regulation of expression of polymorphic mammalian genes (Shabalina et al., 2014). Recent global analyses suggest that pre-mRNA splicing is predominantly co-transcriptional in different organisms (Brugiolo et al.,2013). The two processes are functionally coupled, and the prevalence of different types of alternative events differs between functional regions (e.g untranslated regions (UTRs) and protein coding sequences (CDSs)) of transcripts. In particular, AS is common in the 5’ UTRs and protein CDS. In contrast, AS is rare in 3’ UTRs. The Alternative transcription initiation and termination events occur primarily in the 5’UTRs and 3’UTRs, respectively. (Resch et al., 2009).

Expression of mRNA in animals is a complex and intricately controlled process during which the transcriptional apparatus closely cooperates with pre-mRNA processing machinery. Numerous biochemical and cytological experiments indicate that in eukaryotes transcription and mRNA processing including capping, splicing, and polyadenylation/cleavage form a network of elaborately regulated and coupled processes that

occur together within nuclear ‘‘gene expression factories’’ (Bentley 2005). These findings suggest intriguing possibility that alternative events occurring at different levels of gene expression and transcript processing might not be independent (Shabalina et al., 2010).

Alternative transcription initiations: Alternative Promoter (AP).

The mRNA isoforms are usually transcribed by alternative promoters by inclusion of alternative first exons. Several genome-wide analyses indicate that 30–50% of the human and about one half of the mouse genes have multiple alternative promoters (Beak et al.,2007).

A global analysis of mammalian promoters concluded that alternative promoters are over-represented among genes involved in transcription regulation and development and single promoter gene are active in broad range of tissues and are more likely to be involved in general cellular processes, such as RNA processing, DNA repair, and protein biosynthesis (Baek et al., 2007).

The transcription initiation by RNA polymerase II (PolII) in the core promoter region requires many proteins,

including many general transcription factors (TFs). Mammalian organisms have evolved cell type–specific components of the general TFs as a mechanism to accommodate more complex programs of tissue-specific and promoter-selective transcription, in which the basal transcriptional machinery and the core promoter sequences display a high degree of diversity and selectivity in the regulation of corresponding gene isoforms. For example, the alternative use of eight promoters, distributed over a 93-kb regulatory region, forms the basis for differential regulation of CYP19A1 expression by various hormones, growth factors and cytokines in a tissue-specific manner (Fig.1). Instead of duplicating genes, tandem alternative core promoters would allow a given cell type to regulate a single gene that may be expressed at different times during development or in different cell types by using functionally specialized transcription initiation complexes.

Alternative promoter (AP) use results in mRNAs that eighter have distinct 5’ UTRs but encode identical proteins, or different protein isoforms with distinct functional activities.

In mammals, several genes with widely spaced alternative first exons produce distinct mRNA isoforms with heterogeneous 5’ UTRs, which can affect the

stability or translation efficiency of the RNA variants, yet encode identical proteins. For example, RUNX1 mRNA initiated from the proximal promoter has a long 5’ UTR, which contains a functional internal ribosome entry site and mediates cap-independent translation (Pozner et al, 2000). The OTX2 gene (orthodenticle homeobox 2), implicated in medulloblastoma brain tumour is transcribed by three alternative promoters with tissue-specific expression. These APs encode identical protein isoforms giving rise to mRNA that differ in their 5’ UTRs (Courtois et al., 2003).

Alternative promoters can also generate mRNA isoforms that encode distinct proteins, sometimes having opposing biological activities, if the translation start site of the corresponding mRNA isoform exists within the first exon. For example *GNAS* locus could generate three different proteins both in mouse and human: Gsa (G protein a-subunit, ubiquitously expressed), XLas (extra large Gsa isoform, primarily expressed in neuroendocrine tissues) and NESP55 (neuroendocrine-specific protein of 55 kDa) thank to alternative promoters. Gsa and XLas proteins have different roles in the cell: Gsa-specific knockout mice develop increased fat mass, glucose intolerance, insulin resistance and hypertriglyceridemia, whereas XLas deficiency leads to a perinatal suckling defect and a

lean phenotype with increased insulin sensitivity. The opposite metabolic effects of Gsa and XLas deficiency are associated with decreased and increased sympathetic nervous system activity, respectively. (Davuluri et al., 2008).

Alternative Transcription Terminations: Alternative Cleavage and Poliadenilation.

Alternative polyadenylation (APA) is a widespread phenomenon, generating mRNAs with alternative 3’ ends. APA contributes to the complexity of the transcriptome by generating isoforms that differ either in their coding sequence or in their 3’ untranslated regions (UTRs), thereby potentially regulating the function, stability, localization and translation efficiency of target RNAs (Elkron et al.,2013).

The poly(A) site (PAS) is defined by multiple cis-elements. For canonical mammalian PASs, these cis-elements include an AAUAAA hexamer, a U- or GU-rich downstream element (DSE), and other auxiliary sequences including the U-rich upstream stimulatory element (USE). PAS recognition relies on several multisubunit protein complexes.

The polyadenylation reaction requires the assembly of a rather large number of interacting protein factors that

recognize a relatively simple set of *cis*-acting signal sequence elements in the mRNA precursor. Cleavage stimulation factor (CstF) is one of the essential polyadenylation factors. CstF is active most likely as a dimer with each subunit consisting of three protein factors called CstF-77, CstF-64 and CstF-50. CstF-64 interacts directly with the downstream located GU-rich *cis*-acting element. Both CstF-50 and CstF-77 subunits interact specifically with the C-terminal domain (CTD) of the RNA polymerase II largest subunit (RNAP II LS), likely facilitating the RNAP II-mediated activation of 3'-end processing (Hirose and Manley, 1998) Furthermore, 3' end processing can be repressed after DNA damage as a result of an interaction between CstF-50 and BRCA1-associated RING domain protein (BARD1) (Kleiman and Manley, 1999) and of the proteasome-mediated degradation of RNAP II (Kleiman et al., 2005).

For a large portion of eukaryotic genes, their mRNAs have multiple alternative ends that are formed by cleavage/polyadenylation at distinct sites, a phenomenon known as APA. Lutz and Moreira (2011) described three different types of polyadenylation and alternative polyadenylation in higher eukaryotic mRNAs (Fig.2). In Type I polyadenylation, only one polyadenylation signal is present in the 3' UTR, thus

resulting in only one mRNA isoform. In Type II alternative polyadenylation, more than one polyadenylation signal is present but are only present in a common terminal exon. In this type of alternative polyadenylation, more than one resulting mRNA is produced, but with no effect on the encoded protein; however, due to possible alteration of mRNA stability/translatability/other downstream effects the choice of APA signals may affect the amount of protein produced if alternative polyadenylation signals are chosen. Type III alternative polyadenylation involves alternative polyadenylation signals that are present in upstream introns or exons, thus invoking alternative splicing along with alternative polyadenylation. They have further defined here Type III polyadenylation as Type IIIi (intronic alternative polyadenylation signal) or Type IIIe (exonic alternative polyadenylation signal). These types of alternative polyadenylation may or may not result in different protein products being produced, depending on the stability of the mRNA, the presence of an in-frame stop codon, and the overall translation competence of the mRNA.

Alternative Splicing (AS).

Splicing involves the recognition of exons and introns followed by exons being joined together and introns being removed. There are different types of alternative splicing. These include exon skipping (or cassette exon), mutually exclusive exon splicing, alternative 3’ or 5’ splice site usage, and intron retention (Fig.3). Among these different modes of alternative splicing, exon skipping is most common, accounting for 40% of the entire alternative splicing events. This mode of alternative splicing produces different messenger RNAs that translate into protein isoforms with distinct coding sequences.

Splicing occurs through the concerted actions of multi-subunit complexes. Splice sites are recognized by the spliceosome, a catalytic splicing machine comprised of five small nuclear ribonucleoprotiens (snRNPs) and over 100 individual proteins. The recognition of exons is determined by consensus sequences at the 5’ and 3’ exon boundary and a polypyrimidine track located 20 – 40 nucleotides upstream of the 3’ splice site. Alternatively spliced exons often contain weak splice sites that diverge from the consensus sequence and are poorly recognized by the spliceosome. The recognition of these weak exons and hence alternative splicing is facilitated by trans-acting splicing regulators bound on

cis-acting premRNA sequences. The cis-acting elements can be located in the variable exons and adjacent introns and, depending on their functions, these regulatory elements are categorized as Exonic Splicing Enhancers (ESEs), Exonic Splicing Silencers (ESSs), or Intronic Splicing Enhancers and Silencers (ISEs and ISSs). The most well-characterized classes of splicing regulators are the families of SR proteins (serine/arginine-rich proteins) and hnRNPs (heterogeneous nuclear ribonuceoproteins). The precise function of these splicing factors can be influenced by the location and sequence context of cis-elements that recruit them. In many cases, these splicing factors exert a combinatorial effect of positive and negative regulations for the control of alternative splicing (Liu and Cheng, review 2013).

Interconnections between AP, AS and APA.

The architecture of gene loci, location of transcription initiation and termination sites, and the presence of splicing sites and regulatory elements define the structures of pre mRNA and mature transcripts.

Alternative transcription initiation (ATI) and termination (ATT) shape the structures of 5’ and 3’ terminal regions of transcripts and play distinct and independent roles in

the regulation of expression of polymorphic mammalian gene (Fig.4). At the transcription initiation step, the selection of alternative initiation sites by the RNA polymerase defines the choice of splicing sites. Transcription from alternative promoters and transcription start sites creates distinct 5' terminal exons and changes splicing patterns in the 5' transcript regions. Moreover, differential occupation of promoters and enhancers by certain transcription factors and coregulators can define the choice of exons that are included into the mRNA and trigger exon skipping (Kornblihtt et al., 2004), especially in 5' UTRs. Alternative exons created by ATI in the 5' regions of transcripts typically connect to the coding sequence, do not affect splicing patterns in the coding sequence and often extend the encoded protein at the N-terminus, with the reading frame preserved (Landry et al., 2003). The ATI group is enriched for genes that are preferentially involved in development, signal transduction and apoptosis, in contrast to the ATT group that is enriched for more broadly expressed genes preferentially involved in cellular processes and organization, protein modification and regulation of metabolism.

In contrast to ATI, ATT strongly depends on upstream splicing events and therefore is tightly coupled with AS

in the coding regions (Shabalina et al., 2010). In mechanistic terms, this coupling appears to be due in part to the ability of RNA polymerase II (PolII) to bind some of the pre-mRNA processing factors which then travel along with PolII to their targets in a complex known as 'the mRNA factory' (Kornblihtt et al., 2013).There is growing evidence that splicing and polyadenylation and cleavage factors can be recruited to the C-terminal domain of the largest subunit of PolII which participates in the key pre mRNA processing reactions. As transcription proceeds, exons are defined along the pre-mRNA by communication between the 5' and 3' splice sites and in cooperation with PolII. At the end of the transcription process, the splicing and cleavage machineries cooperate to define the poly(A) site and the terminal exon (Kornblihtt et al., 2013). Reporter experiments have indicated that the choice of poly(A) sites depends on the rate of transcription and that RNA polymerase II is more likely to pause at proximal poly(A) sites of highly expressed genes and at distant poly(A) sites of low-expressed genes (Almada et al., 2013). Furthermore, there are indications that selection of the poly(A) site might occur prior to splicing of the terminal exon (Yamashita et al.,2010). Details of this process have been recently studied using a mammalian in vitro transcription and splicing system

(Rigo et al.,2008). It has been shown that communication between the 3’ splice site of the terminal exon and the poly(A) site is establish hedearly in the transcription process, followed with the poly(A) site cleavage, polyadenylation at the 3’end of the transcript, and concluded with splicing of the last exon. The strong positive correlation between AS events in the coding sequence and ATT observed at the global transcriptome level (Shabalina et al.,2010) indicates that the coupling of transcription termination to splicing is a common and wide-spread phenomenon that occurs during transcription of thousands of human genes.

1.2 Epigenetic control of Transcription.

Transcription is not only controlled by transcription factors and the transcription apparatus, but also at the chromatin level.

The fundamental unit of chromatin, the nucleosome, is regulated by protein complexes that can mobilize the nucleosome or modify its histone components. Gene activation is accompanied by recruitment of ATP-dependent chromatin remodelling complexes of the SWI/SNF family, which mobilize nucleosomes to

facilitate access of the transcription apparatus and its regulators to DNA. In addition, there is recruitment, by transcription factors and the transcription apparatus, of an array of histone-modifying enzymes that acetylate, methylate, ubiqutinylate, and otherwise chemically modify nucleosomes in a stereotypical fashion across the span of each active gene.

These modifications provide interaction surfaces for protein complexes that contribute to transcriptional control. Enzymes that remove these modifications are also typically present at the active genes, producing a highly dynamic process of chromatin modification as RNA polymerase is recruited and goes through the various steps of initiation and elongation of the RNA species. Repressed genes are embedded in chromatin with modifications that are characteristic of specific repression mechanisms. One type of repressed chromatin, which contains nucleosome modifications generated by the Polycomb complex (e.g., histone H3K27me3), is found at genes that are silent but poised for activation at some later stages of development and differentiation. Another type of repressed chromatin is found in regions of the genome that are fully silenced, such as that containing retrotransposons and other repetitive elements. The mechanisms that silence this latter set of genes can

involve both nucleosome modifications (e.g., histone H3K9me3) and DNA methylation. (Lee et Yang, 2014).

The chromatin.

The function of chromatin is to efficiently package DNA into a small volume to fit into the nucleus of a cell and protect the DNA structure and sequence. Packaging DNA into chromatin allows for mitosis and meiosis, prevents chromosome breakage and controls gene expression and DNA replication.

Although the numbers and sizes of chromosomes vary considerably between different species, their basic structure is the same in all eukaryotes. The DNA of eukaryotic cells is tightly bound to small basic proteins (histones) that package the DNA in an orderly way in the cell nucleus. The basic structural unit of chromatin, the nucleosome was described by Roger Kornberg in 1974. Electron microscopy revealed that chromatin fibers have a beaded appearance, with the beads spaced at intervals of approximately 200 base pairs, which were called nucleosomes. These particles has shown that they contain 146 base pairs of DNA wrapped 1.65 times around a histone core consisting of two molecules each of H2A, H2B, H3, and H4 (the core histones). One molecule of the fifth histone, H1, is bound to the DNA as it enters each nucleosome core

particle. This forms a chromatin subunit known as a chromatosome, which consists of 166 base pairs of DNA wrapped around the histone core and held in place by H1 (a linker histone). The packaging of DNA with histones yields a chromatin fiber approximately 10 nm in diameter that is composed of chromatosomes separated by linker DNA segments averaging about 80 base pairs in length. Packaging of DNA into such a 10-nm chromatin fiber shortens its length approximately sixfold. The chromatin can then be further condensed by coiling into 30-nm fibers, the structure of which still remains to be determined. The extent of chromatin condensation varies during the life cycle of the cell. Euchromatin is relatively decondensed and distributed throughout the nucleus. In this state genes could be transcribed. In contrast to euchromatin, heterochromatin is in a very highly condensed state and transcriptionally inactive. (Cooper et al.,2000).

Chromatin remodelling complexes.

Chromatin is a dynamic structure that not only helps to package the entire eukaryotic genome into the confines of the nucleus but also regulates the accessibility of DNA for transcription, recombination, DNA repair and replication. Although the structure of nucleosomes

appears rigid at the cytological level these repeating subunits of chromatin are very dynamic (Luger et al.2012). The assembly and the compaction of chromatin are regulated by multiple mechanisms, including DNA modifications (for example, cytosine methylation and cytosine hydroxymethylation), post-translational modifications (PTMs) of histones (for example, phosphorylation, acetylation, methylation and ubiquitylation), the incorporation of histone variants (for example, H2A.Z and H3.3), ATP-dependent chromatin remodelling and non-coding RNA (ncRNA)-mediated pathways (Chenetet et al.,2014).

Histone Post-Translational Modifications (PTMs).

PTMs of histones may either directly affect chromatin compaction and assembly or serve as binding sites for effector proteins, including other chromatin-modifying or chromatin-remodelling complexes, and ultimately influence transcription initiation and/or elongation. Most, if not all, histone PTMs are reversible. Many enzymes that are involved in their addition and removal have been identified. These include histone acetyltransferases (HATs; also known as lysine acetyltransferases) and histone deacetylases (HDACs; also known as lysine deacetylases); lysine methyltransferases (KMTs) and lysine demethylases

(KDMs); and ubiquitylation enzymes (that is, E1, E2 and E3 enzymes) and deubiquitylases (DUBs). These enzymes often exist in multisubunit complexes and modify specific residues either on the amino-terminal tails or within the globular domains of core histones (H2A, H2B, H3 and H4). For example, in the two repressive Polycomb group (PcG) protein complexes, Polycomb repressive complex 1 (PRC1) contains either ring finger protein 1A (RING1A) or RING1B, both of which catalyse the monoubiquitylation of histone H2A at lysine 119 (H2AK119ub1), and PRC2 contains enhancer of zeste 2 (EZH2), which catalyses the trimethylation of H3K27 (H3K27me3).

ATP-dependent chromatin-remodelling complexes

In addition to PTMs of histones, chromatin compaction is also regualted by ATP-dependent chromatin-remodelling complexes that use energy from ATP hydrolysis to exchange histones and to reposition or evict nucleosomes. Approximately 30 genes that encode the ATPase subunits have been identified in mammals. On the basis of the sequence and the structure of these ATPases, chromatin-remodelling complexes are divided into four main families: SWI/SNF, ISWI, CHD and INO80 complexes (Ho et al., 2010). Many histone modifiers and chromatin

remodellers have been implicated in stem cell pluripotency, cellular differentiation and development.
All the remodelers contain an ATPase subunit characterized by an ATPase domain which is splitted in to units: a DExx-box and a Helicase, spaced by a linker (Tang et al., 2010). The other domains which are adjacent to ATPase domain differ from one to another in the four families of chromatin remodelers.
I will now discuss the peculiar features that distinguish the ISWI, CHD and INO80 families, and next I will focus specifically on the SWI/SNF family.

ISWI family remodelers:

The ISWI ("Imitation SWItch") remodelers contain two to four subunits. They were initially identified in an *in vitro* screening aimed to test the nucleosome remodeling activity of *Drosophila* embryo extracts (Tsukiyama et al., 1995). The ISWI catalytic subunit, in addition to the ATPase domain, contains a SANT and a SLIDE domain, which together form a nucleosome recognition motif that binds to unmodified histone tails (Clapier et al., 2009). The ISWI ATPase subunit forms at least three different chromatin remodeling complexes, termed ACF, CHRAC and NURF, which were initially identified in *Drosophila*. The three homologous human ISWI complexes differ for the number of

subunits incorporated, and, depending on that, for their relative enzymatic activity. Additional subunits incorporated in the ISWI remodelers comprise DNA-binding/ histone fold domains (CHRAC 15-17), bromodomains (BPTF and ACF1) and DNA-binding domain (HMG1(Y)). The resulting enzymatic activity of the ISWI complexes ranges from chromatin assembly, to chromatin remodeling, to the maintenance of the euchromatin and of the heterochromatin, and finally to the assistance to the RNA Polymerase to facilitate or repress the transcription of target genes. The different activities depend from the attendant subunits (Clapier et al., 2009; Hargreaves et al., 2011).

CHD family remodelers:

The CHD ("Chromodomain, Helicase, DNA binding") complexes contain from 1 to 10 subunits. They were firstly purified from *Xenopus leavis* (Marfella et al., 2007). The ATPase incorporated in the CHD chromatin remodeling complexes contains two characteristic chromodomain, arranged in tandem in the N-terminal region. This catalytic subunit is monomeric in the lower eukaryotes, but forms large complexes in vertebrates. Attendant subunits comprise DNA-binding and SANT domains-containing proteins, histone deacetylases (HDAC 1/2) and methyl-CpG binding proteins (MBD).

The human CHD family contains nine members, named CHD1-9, which differ for the ATPase subunit incorporated. As a matter of fact, in human there are nine different CHD ATPase subunits, which are further categorized in three sub-groups basing on their domain composition: CHD 1-2 ATPases contain a C-terminal DNA-binding domain, CHD3-4 ATPases lack the DNA-binding domain but have two N-terminal PHD fingers, while CDH5-9 have additional domains (Hargreaves et al., 2011). CHD family members are mostly studied for their roles in differentiation and in the regulation of genome stability (Marfella et al.,2007).

INO80 family remodelers:

The INO80 ("INOsitol requiring 80") family contains more than 10 subunits, and they were firstly purified from *Saccharomyces cerevisiae* in a screening aimed to identify the regulators of phosphlipid biosynthesis (Ebbert et al., 1999). The human orthologs of the yeast complex contains the Ino80, SRCAP and SWR1 ATPases. The defining feature of the ATPase subunit incorporated in the INO80 remodelers is the presence of a "split" ATPase domain. The two sub-domains of the ATPase are divided by a long spacer region, to which the helicase-related Rvb1/2 and ARP proteins bind. The INO80 family members have different activities in

human cells ranging from promotion of transcription to DNA repair. The SWR1 member is unique in its ability to restructure the nucleosome: as a matter of fact, it can replace the canonical H2A-H2B dimer with the variant dimer H2A.Z-H2B dimer, specifically in the chromosome regions that mark transcription start sites (Clapier et al., 2009).

The SWI/SNF complex.

The SWI/SNF (mating type SWItch or SWItching defective/Sucrose NonFermenting) chromatin remodeling complexes were initially described in Saccharomyces cerevisiae as positive regulators of HO and SUC2 genes (Lander et al., 1992). Shortly after the identification of this complex in yeast, homologues complexes were identified in fruit fly and human. As in the *S.cerevisiae* protein, these complexes contains the ATPase subunit, which is homologous to the yeast SWI2/SNF2 protein, as well as other attendant subunits, corresponding to the yeast SNF5, SWI3 and SWP73 subunits (Muchardt et al., 1999). Together, these four proteins form the core of the eukaryotic SWI/SNF chromatin remodeling complex. (Phelan et al., 1999). On the other hand, some other proteins which are contained in the human SWI/SNF complexes are not present in other species. For example, BAF57, a

protein contained in the human BAF complexes, has a counterpart in Drosophila but is not present in yeasts (Papoulas et al., 1998).

As reported by many papers, the human SWI/SNF complex can also be named BAF, which is the acronym of Brahma Associated Factors, by the name of one of the two alternative ATPases present in the complex, Brahma. As a consequence, the protein that are incorporated in the human complex are named BAF, but, alternatively, they could also retain the yeast homologs name or could be identified by the name SMARC (SWI/SNF-related, Actin containing, Regulators of Chromatin). Depending on the different eukaryotic species, the SWI/SNF complex could contain from 8 to 14 subunits, generating a huge macromolecular complex. For example, a typical human BAF complex is composed by five yeast ortholog subunits (one of the two alternative ATPase subunits BRM or Brg1, BAF155/170, BAF60a/b/c, Baf53a/b and BAF47), plus several unique, human-restricted subunits (such as BAF57, BAF250a/b, BAF200, BAF45a/b/c/d) and monomeric nuclear actin. This huge "macromolecular machine" reaches the molecular weight of 2 MDa, which is larger than the calculated molecular weights of the known subunits, indicating that several additional interactors have yet to be identified (Hargreaves et al.,

2011). It is long known that, following biochemical purification of the SWI/SNF-BAF complexes, it is possible to obtain several fractions, each one representing a single particular complex. Having identified the four constitutive subunits, the combinatorial assembly of the others attendant subunits generates a huge amount of possible combinations, each one corresponding to one remodeler complex that may or may not be present in a particular cell of a tissue or in specific differentiation stage. The combinatorial generation of the different complexes reflects a specialization in the tasks accomplished by the different SWI/SNF-BAF multi-subunits enzymes (Wang et al., 1996). One peculiar mechanism of catalysis distinguishes the SWI/SNF remodelers from the other three families of chromatin remodeling complexes. It has been recently observed that SWI/SNF remodelers use the energy of ATP hydrolysis to catalyze a two step reaction whose outcome consists in the eviction of one entire nucleosome (Fig.5). In the first reaction, the SWI/SNF complex induces the exit of the histone dimer H2A-H2B from one nucleosome, while in the second reaction substitutes the lost histone dimer with one identical dimer from the neighbouring nucleosome. The nucleosome which loses the dimer now lacks the H2A-

H2B, and needs to get a new dimer that can be obtained from the next neighbouring nuclesome, as so the two reaction start again. The coupling of these two reactions creates a "wave" of chromatin remodeling, whose aftermath consists in the localized displacement and in the movement of nucleosomes in a particular region of the chromosome (Dechassa et al., 2010). This observation indicates that, at least in vitro, SWI/SNF is able to catalyze the eviction of one nucleosome by exploiting the presence of another nucleosome localized in close proximity. This neighbouring nucleosome is requested for the catalysis to increase the processivity of the wave of displacement. When the "wave" of remodeling reaches a DNA region in which no more nucleosomes are present, the reaction stops, leaving a nucleosome with the H3 and H4 dimers lacking the H2A-H2B dimer. The H3 and H4 dimers are ejected from the DNA, so that the final outcome is the shifting of the nucleosomes in one direction and the ejection of one single nucleosome (Liu et al., 2010).

The two alternative ATPase subunits which are incorporated in the human SWI/SNF-BAF complexes are named Brahma (BRM, hBRM, SNF2α) and Brg1 (Brahma-Related Gene 1, hBrg1 SNF2β). These two proteins are 75% identical (Kadam et al., 2003) and

elute in the same fraction upon purification by conventional chromatography (Muchardt et al., 1999).

BRM & BRG1.

The evolutionarly divergence between the two ATPases BRM and Brg1 is present starting from mouse (mBRM and mBrg1). Other eukaryotic species, such as yeasts and *Drosophila*, have only one SWI/SNF ATPase subunit, and this observation, together with the high level of sequence homology, indicates that the two ATPase genes derive from the duplication of one single gene (Muchardt et al., 1999; Clapier et al., 2009; Hargreaves et al., 2011). The two human genes, *SMARCA2* and *SMARCA4* code for two proteins of about 200 kDa, respectively encoding BRM and Brg1. The two genes are mostly co-expressed in many cells, even if it is also important to note that some differences in the relative expression of the two genes have been observed during physiological and pathological processes. A comparison of the two proteins indicates an high grade of homology, as reported in the schematic structure of the two proteins.

The two proteins share a common ATPase domain, which includes a Helicase-SANT (HSA) domain. In addition, they also share a common E7 sequence, a

protein module that interacts with Retinoblastoma protein Rb (Kadam et al., 2003). The peculiarity of both these enzymes consists in the presence of a C-terminal Bromodomain. Firstly identified in the fruit flyhomolog (Lander et al., 1992), this domain is unique to the SWI/SNF family of ATPase subunits. In the others families of remodelers, the Bromodomain is not present in the ATPase subunits, but is provided by attendant proteins. The bromodomain allows the interactions with the acetylated histone tails, an interaction important for the recrutiment of the SWI/SNF-BAF complexes in specific regions of the chromatin. It has also been demonstrated that the bromodomain can interact with specific acetylated histone residues, to promote localized gene activation (Clapier et al., 2009), and that bromodomain deletions inactivate the chromatin-targeting activity exerted by this protein module (Winston et al., 1999). Moreover, it has been demonstrated that the bromodomain can also interact with regulators of histone acetylation, such as HATs and HDACs, thus creating an auto-regulatory loop (Hargreaves et al., 2011).

Even if BRM and Brg share a high grade of homology in the genes and in the polypetide sequences, the differences between them are more marked than their

similarities. Experimental evidencies sustain the hypothesis that BRM and Brg1 are not functionally and genetically redundant (Flowers et al., 2009). They differs in co-attivators, in their ATPase level activity and for their antagonist role during brain development.

BRM and Brg1 interacts with different co-activators.

It has been demonstrated that BRM and Brg1 interact with different proteins thanks to a specific N-terminal region. Brg1 contains in this region a zinc finger domain that allows this protein to interact with zinc finger-containing proteins (such as transcription factors of the KLF and GATA families), while the same region in BRM allows the interactions between BRM and ankyrin repeats-containing proteins (such as proteins of the Notch pathways). These different interactions are functional to localize both the BRM- and Brg1-containing BAF complexes and the correct interactors on specific promoter and/or chromosome regions (Kadam et al., 2003).

BRM and Brg1 have different levels of ATPase activity.

BRM and Brg1 are incorporated in the SWI/SNF-BAF complexes in a mutually exclusive fashion and in 1:1 stoichiometry (one ATPase subunit per single complex). The two different SWI/SNF-BAF complexes are

endowed with ATPase-depedent chromatin remodeling activity and are able to increase the accessibility of the DNA to nucleases in an *in vitro* assay. However, the complexes differ not only in subunit composition but also in the ATPase activity. The Brg1 containing complexes show high ATPase activity, while BRM containing complexes show a lower (a five fold decrease) enzymatic activity (Sif et al., 2001).

BRM and Brg1 play antagonistic roles during brain development.

The knock out of Brg1 in mice causes a very severe phenotype: the Brg1-/- die during the peri-implantation stage (E3.5-E5.5), because of an impairment in the formation of the inner cell mass and the trophectoderm. The peri-implantation lethal phenotype of mice lacking Brg1 indicates that esBAF (ES cell-specific BAF) has an essential role in early embryonic development (Bultman et al., 2000). Cre-mediated deletion of Brg1 impairs selfrenewal of cultured ES cells but not mouse embryonic fibroblasts (MEFs) (Ho et al., 2009b) or glial cells (Lessard et al., 2007). Brg1-deleted ES cells first lose proliferative capacity and, after several passages, lose the expression of ES cell markers Oct4, Sox2 and Nanog; they are also defective in forming ectodermal and mesodermal lineages during

spontaneous differentiation, indicating a loss of pluripotency.

The specific configuration of subunits in esBAF is crucial for the ES cell state. BRM, the alternative ATPase subunit, is not found in esBAF and does not compensate for the absence of Brg1 (Bultman et al., 2000); moreover, Brm−/− mice are viable despite being slightly larger than normal (Reyes et al., 1998).

During development of the brain neural precursors must downregulate the pluripotency gene network promoted in ES cells while remaining proliferative. They must also suppress the neuronal differentiation program until mitotic exit. The chromatin landscape associated with this particular state requires the neural progenitor-specific BAF (npBAF) complex (Lessard et al., 2007). Starting from this stage the neuronal BAF complex can have either Brg1 or BRM as the core ATPase.

1.3 Oxidative stress in Neurodegenerative disorders.

Oxidative stress is characterized by the overproduction of reactive oxygen species, which can induce mitochondrial disfunctions. DNA mutations, damage the mitochondrial respiratory chain, alter membrane permeability, and influence Ca^{2+} homeostasis and mitochondrial defense systems. Oxidative stress and mitochondrial damage have been implicated in the pathogenesis of several neurodegenerative diseases, including Alzheimer's Disease (AD), Parkinson's Disease (PD) and Amyotrophic Lateral Sclerosis (ALS). Despite these diseases show different clinical symptoms and pathology, increasing evidence suggests that mitochondrial damage plays an important role in their pathogenesis (Guo et al.,2013).

In will now discuss oxidative stress in ALS to better indroduce the starting mitochondrial stress cell models of my work.

Mitochondrial damage and ALS

Amyotrophic lateral sclerosis (ALS) is a lethal disorder characterized by subtle onset of focal weakness, typically in the limbs but sometimes in bulbar muscles,

which progresses to paralysis of almost all skeletal muscles. There is significant clinicopathological and genetic overlap between ALS and frontotemporal lobar dementia (FTLD). In ALS, death from respiratory paralysis is typically within five years. The cellular pathology is focal at onset and spreads in a pattern suggesting successive involvement of contiguous neuronal populations. Death of motoneurons occurs in conjunction with deposition of aggregated proteins in motoneurons and oligodendrocytes, and neuroinflammation. Whereas most cases of ALS are sporadic (SALS), about 10% are inherited, usually dominantly (familial ALS [FALS]).

Approximately 20% of familial ALS cases are due to mutations in the gene encoding superoxide dismutase 1(SOD1)(Ferri et al., 2006). The expression of mutant superoxide dismutase 1 is not only associated with mitochondrial morphological changes, but also with mitochondrial dysfunction. Postmortem tissue, as well as biopsy samples of a variety of tissues, show mitochondrial ab-normalities in both sporadic and familial ALS patients.

Mitochondrial alterations may represent an early event triggering the onset of ALS, rather than simply a by product of cell degeneration (Rosen et al.,1993). In the mutant superoxide dismutase 1 mouse model,

mitochondrial abnormalities appear before symptoms of paralysis and motor neuron degeneration. At the time of disease onset, mitochondrial respiration and adenosine triphosphate synthesis are defective in the brain and spinal cord of G93A mutant superoxide dismutase 1 transgenic mice (Corti et al., 2009).

Epigenetic study in Amyotrophic lateral sclerosis (ALS).

A simple view is that there can only be two causes of ALS: genetic variation and environmental exposures. A third possibility is also important to consider and is known as epigenetic variation: environmental exposures, random processes or genetic variation result in changes to DNA, principally, but not exclusively, by methylation. The study of epigenetic factors is in its infancy but making rapid progress (Al-Chalabi A. et al., 2013).

It is known that ALS has a genetic component because a family history of ALS is reported by about 5% of patients (Byrne et al.,2011); twin studies show that the heritability is about 60% (Al-Chalabi et al.,2010), and disease-causing mutations have been identified in up to 60% of those with a family history, and between 5% and 15% of those without to date (Andersen et

al.,2011). One problem is that genes described as ALS genes may produce atypical phenotypes including very young age of onset, very slow progression, or purely upper or lower motor neuron signs, while other genes regarded as typical ALS genes may give rise to ALS with frontotemporal dementia (ALS-FTD).

There is a rapidly evolving body of evidence that ALS is associated with complex and nuanced alterations in RNA regulatory networks (Strong et al., 2010). These include processes associated with transcriptional regulation, post-transcriptional processing, deployment and function of different subclasses of non-coding RNAs, local control of translation and additional epigenetic processes. Epigenetic processes include DNA methylation, histone, nucleosome and higher-order chromatin remodeling (Mehler et al., 2008).

Recently Figueroa-Romeo and colleagues (2012) identified ALS-dependent methylation dysregulation of several genes previously implicated in neuronal development, differentiation, and proliferation.

There are many genes involved in ALS. A database of these genes (ALSoD) (Abel et al., 2009) with their associated phenotypes and a measure of their credibility is available at the World Federation of Neurology and European Network for the Cure of ALS

(ENCALS) project website http://alsod.iop. kcl.ac.uk, generously supported by ALSA, ALS Therapy Alliance, MNDA and ALS Canada.

Now I will focusing on SOD1, the most known protein that has been linked to fALS.

SOD1 protein

The identification of mutations in the superoxide dismutase 1 (SOD1) gene in 1993 triggered the first major wave of molecular research in ALS. SOD1 is a ubiquitously expressed protein that catalyses the detoxification of superoxide. More than 160 mutations in SOD1 associate exclusively with ALS. Nearly all are dominant missense mutations. Transgenic mice overexpressing mutant SOD1 display an ALS-like phenotype and remain a cornerstone of ALS research (Gurney et al.,1994). One of the most common human mutations present in the SOD1 enzyme and linked to ALS is the substitution of the Glycine in position 93 with an Alanine (G93A) It has been proposed that the pathological effect of this mutation does not cause a loss of function, but rather a gain of function, by which the protein acquires some toxic properties. Mutant SOD1 toxicity is mediated through several mechanisms, prominently including protein misfolding and oligomerization. Downstream effects of this gain-of-

function toxicity include impaired mitochondrial metabolism, axonal degeneration, axonal transport failure, excitotoxicity, proteasomal disruption, and endoplasmic reticulum stress. Intriguingly, recent studies have implicated wildtype SOD1 toxicity in the genesis of SALS (Synofzik et al., 2012). Wild-type SOD1 in SALS can assume aberrant conformations that resemble those of mutant ALS; misfolded wild-type SOD1 can reproduce some forms of cytotoxicity induced by mutant SOD1. It has also been postulated that wildtype SOD1, once misfolded, can propagate intercellular pathology in a prion-like manner in which there is induction of toxic misfolding in otherwise normal SOD1 upon exposure to misfolded SOD1 (Munch et al., 2011).

1.4 Scope of the thesis

My PhD project was aimed to investigate the putative roles played by the human protein Brahma (BRM, SNF2α) in neuronal cells in culture exposed to mitochondrial stress.

Chapter 1: General introduction
This chapter underlines the most recent topics regarding the gene expression regulation, the

epigenetic control of transcription, the oxidative stress and their possible links with Amyotropic Lateral Sclerosis (ALS) pathology.

Chapter 2: Oxidative stress controls the choice of alternative last exons via a Brahma-BRCA1-CstF pathway. The chromatin-remodeling factor Brahma modulate the choice of alternative 3' terminal exons in a class of transcripts encoding proteins involved in axon growth and guidance. BRM seems to skip proximal last exons recruiting the BARD1/BRCA1 ubiquitin ligase complex. This complex ubiquitins CstF50 and the result is an inhibition of transcript cleavage at the proximal poly(A) site and a shift towards inclusion of the distal terminal exon.

Chapter 3: Functional characterization of *SMARCA2* isoforms. In this chapter I report the identification and the characterization of a class of "short" isoforms of Brahma (BRMS) and of the *SMARCA2* alternative promoter that controls their transcription.

Chapter 4: Conclusions and future perspectives .
The last chapter summarizes the results obtained and underlines the possible future perspectives, focusing

the attention on the putative translational applications of my research.

1.5 Figures

Figure 1:

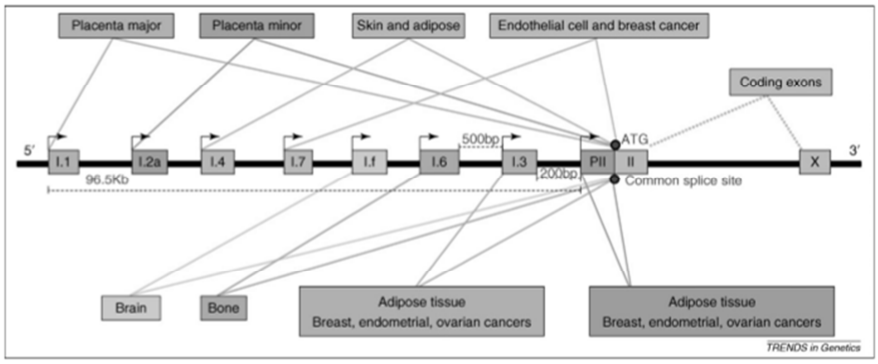


Figure 1: Alternative Promoters usage in different tissues.

Genomic organization of widely spaced alternative promoters of CYP19A1. Alternative first exons I.1, I.2a, I.4, I.7, I.f, I.6 and I.3 splice with Ex-2 to encode the 50 untranslated regions of the aromatase P450 mRNA in various tissues (indicated in boxes connected by splicing pattern). From Davuluri et al., 2008.

Figure 2:

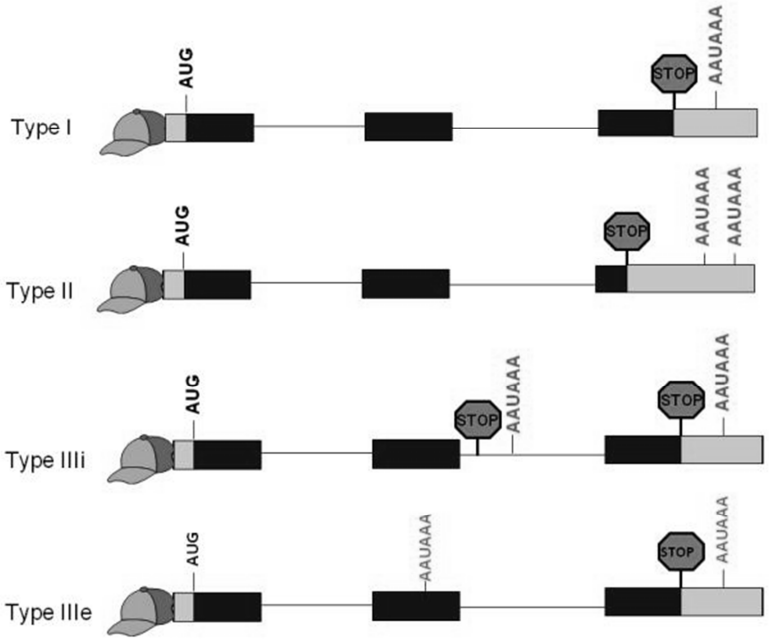


Figure 2: Schematic of polyadenylation events*.*

Grey boxes, untranslated regions; black boxes, coding regions; lines, introns. mRNAs with Type I polyadenylation only have one polyadenylation signal in the 3’ most exon. mRNAs with Type II polyadenylation have more than one polyadenylation signal in the 3’ most exon. Type III polyadenylation is alternative splicing coupled with alternative polyadenylation; Type IIIi signals have one or more polyadenylation signals in upstream introns; Type IIIe signals have one or more polyadenylation signals in upstream exons. From Lutz and Moreira, 2011.

Figure 3:

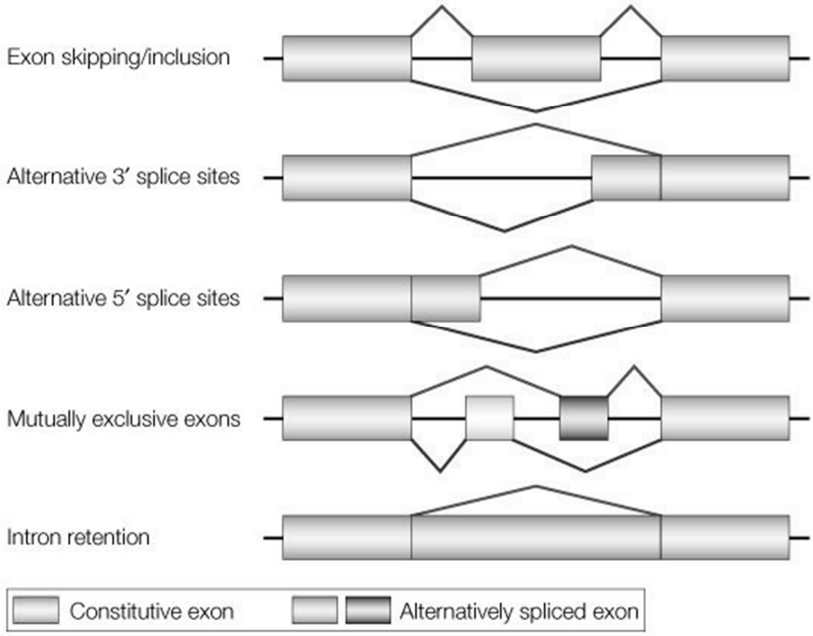


Figure 3: *A schematic representation of alternative splicing.*

The figure illustrates different types of alternative splicing: exon inclusion or skipping, alternative splice-site selection, mutually exclusive exons, and intron retention. For an individual pre-mRNA, different alternative exons often show different types of alternative-splicing patterns. From Cartegni et al., 2002.

Figure 4:

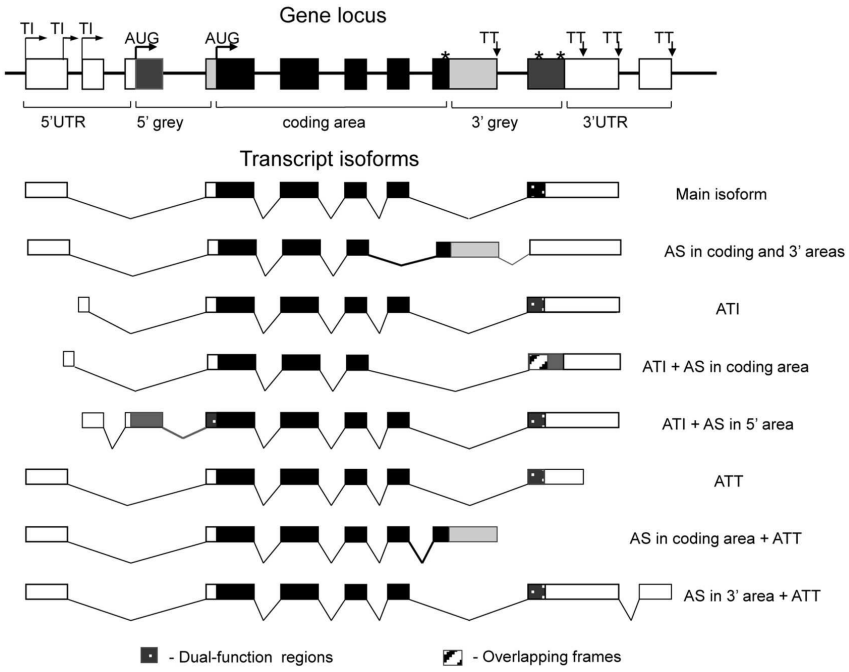


Figure 4: Anatomy of mammalian transcripts: functional domains, constitutive and alternative nucleotides and alternative events.

TI, transcription initiation site; AUG, translation initiation site; TT, transcription termination site; translation termination site; ATI, alternative transcription initiation; AS, alternative splicing; ATT, alternative transcription termination. Protein-coding regions are filled by black (in cCDSs) or by dark grey (in grey areas). UTRs are shown in white (for UTRs) and in light grey (for grey areas). From Svetlana A. Shabalina et al. , 2014.

Figure 5:

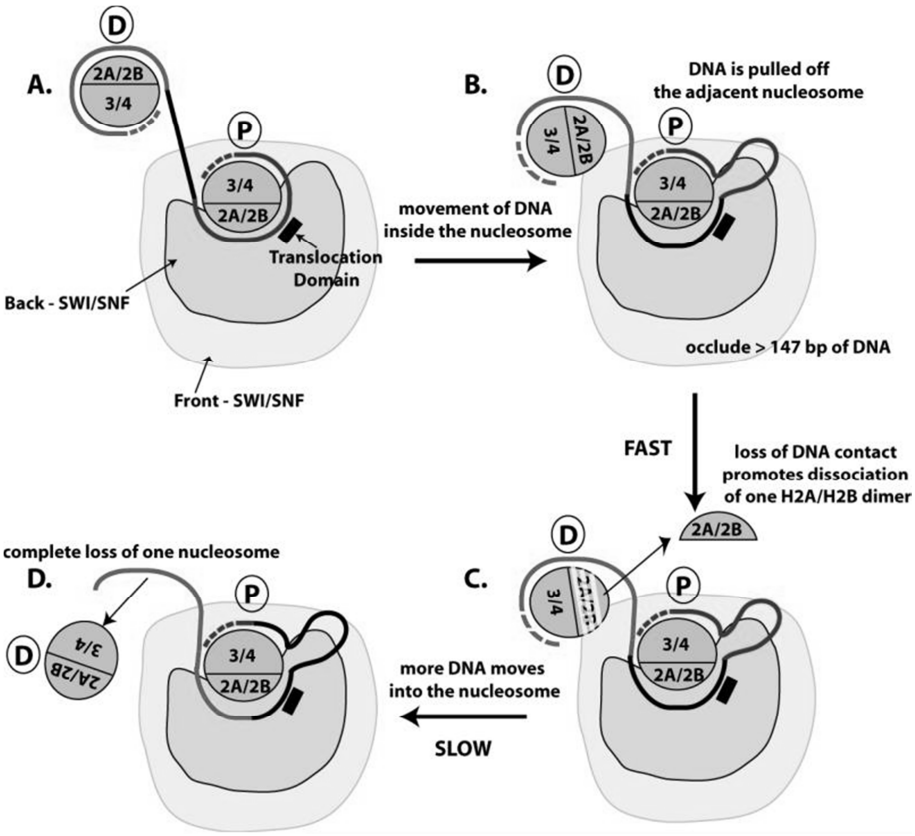


Figure 5: Model for SWI/SNF-Mediated Disassembly of Nucleosomes.

(A) The approximate dimensions of SWI/SNF with respect to nucleosomes is depicted using two shades of blue for the two walls (front and back) that surround the trough of SWI/SNF where the nucleosome is thought to be bound. The location of the DNA translocase domain is shown as a filled black rectangle. The DNA initially associated with the nucleosome that is bound to SWI/SNF is highlighted in red; whereas the linker DNA and DNA associated with the other

nucleosome are black and green, respectively. H2A/H2B (2A/2B) dimers and H3/H4 (3/4) tetramers in histone octamers (gray circle) are shown. The circled P and D refer to the proximal and distal nucleosomes, respectively.

(B) Nucleosomal DNA is moved around the SWI/SNF-bound histone octamer by pulling linker DNA into the nucleosome. DNA bulges are created that propagate around the nucleosome and leave at the opposite nucleosome entry site. The second nucleosome is brought into close proximity to the SWI/SNF surface and can cause DNA to be displaced from the surface of one of the H2A/H2B dimers.

(C) Loss of DNA contact with one H2A/H2B dimer promotes dissociation of the dimer from the histone octamer in a fast kinetic step.

(D) Pulling more DNA away from the second nucleosome in order to move more DNA through the SWI/SNF-bound nucleosome is rate-limited by the stability of the interaction of the H3/H4 tetramer with DNA. Once DNA is transiently released from the tetramer SWI/SNF is poised to pull the extra DNA into and around the same SWI/SNF-bound nucleosome. From Dechassa et al., 2010.

1.6 References

Abel, O., Powell, JF., Andersen, PM., Al-Chalabi, A. (2012). ALSoD: 9. a user-friendly online bioinformatics tool for amyotrophic lateral sclerosis genetics . Human Mutation.

Al-Chalabi, A., Fang, F., Hanby, MF., Leigh, PN., Shaw, CE., Ye, W., et al .(2010). An estimate of amyotrophic lateral sclerosis heritability using twin data. Journal of Neurology, Neurosurgery, and Psychiatry. 81:1324–6.

Almada,A.E., Wu,X., Kriz,A.J., Burge,C.B. and Sharp,P.A. (2013) Promoter directionality is controlled by U1 snRNP and polyadenylation signals. Nature, 499, 360–363.

Andersen, PM., Al-Chalabi, A. (2011). Clinical genetics of amyotrophic lateral sclerosis: what do we really know? Nature reviews. Neurology. 7:603–15.

Baek, D. et al. (2007) Characterization and predictive discovery of evolutionarily conserved mammalian alternative promoters. Genome Res. 17:145–155.

Bentley, DL. (2005). Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors. Curr Opin Cell Biol. 17:251–256.

Brugiolo, M., Herzel, L. and Neugebauer, K.M. (2013). Counting on co-transcriptional splicing. F1000Prime Rep.5.9.

Bultman, S., Gebuhr, T., Yee, D., La Mantia, C., Nicholson, J., Gilliam, A., Randazzo, F., Metzger, D., Chambon, P., Crabtree, G., Magnuson, T. (2000). A Brg1 null mutation in the mouse reveals functional differences among mammalian SWI/SNF complexes. 6:1287-1295.

Byrne, S., Walsh, C., Lynch, C., Bede, P., Elamin, M., Kenna, K., et al. (2011). Rate of familial amyotrophic lateral sclerosis: a systematic review and meta-analysis. Journal of Neurology, Neurosurgery, and Psychiatry. 82 : 623 – 7 .

Cartegni, L., Chew, S. L., & Krainer, AR. (2002). Listening to silence and understanding nonsense: exonic mutations that affect splicing. Nature Reviews Genetics 3, 285–298.

Chan, S., Choi, EA., Shi, Y. (2011). Pre-mRNA 39-end processing complex assembly and function. Wiley Interdiscip Rev RNA 2: 321–335.

Chen, T. and Dent, SYR. (2014). Chromatin modifiers and remodellers: regulators of cellular differentiation. Nature review 15:93-106.

Clapier, C.R., Cairns, B.R. (2009). The biology of chromatin remodeling complexes. Ann. Rev. Biochem., 2009; 78:273,304.

Cooper, GM. (2000). The Cell: A Molecular Approach. 2nd edition. Sunderland (MA): Sinauer Associates;. Available from: http://www.ncbi.nlm.nih.gov/books/NBK9839/

Corti, S., Donadoni, C., Ronchi, D., et al. (2009). Amyotrophic lateral sclerosis linked to a novel SOD1 mutation with muscle mitochondrial dysfunction. J Neurol Sci.276(1-2):170–174.

Courtois, V. et al. (2003) New Otx2 mRNA isoforms expressed in the mouse brain. J. Neurochem. 84, 840–853.

Davuluri, RV., Suzuki, Y., Sugano, S., Plass, C., Huang, TH. (2008). The functional consequences of alternative promoter use in mammalian genomes. Trends Genet 24: 167–177.

Dechassa, M.L., Sabri, A., Pondugula, S., Kassabov, S.R., Chatterjee, N., Kladde, M.P., Bartholomew, B. (2010). SWI/SNF has intrinsic nucleosome disassemblly activity that is dependent on adjacent nucleosomes. Mol. Cell.38:590-602.

Ebbert, R., Birkmann, A., Schuller, H.J. (1999). The product of the SNF2/SWI2 paralogue INO80 of Saccharomyces cerevisiae required for the

efficient expression of various yeast structural genes is a part of a high-molecular-weight protein complex. Mol.Microbiol.,32:741-751.

Elkon, R., Ugalde, A.P.and Agami, R. (2013). Alternative cleavage and polyadenylation: extent, regulation and function . Nature Reviews Genetics 14,496–506.

Ferri, A., Cozzolino, M., Crosio, C., et al. (2006). Familial ALS- superoxide dismutases associate with mitochondria and shift their redox potentials. Proc Natl Acad Sci U S A.;103(37):13860–13865.

Figueroa-Romero, C., Hur, J., Bender, DE., Delaney, CE., Cataldo, MD., Smith, AL., et al. (2012) Identification of Epigenetically Altered Genes in Sporadic Amyotrophic Lateral Sclerosis. PLoS ONE 7(12): doi:10.1371/journal.pone.0052672

Flowers, S., Nagl, N.G., Beck, G.R., Moran, E. (2009). Antagonistic roles for BRM and BRG1 SWI/SNF complexes in differentiation. J. Biol. Chem., 284:10067-10075.

Guo, C., Sun, L., Chen, X. and Zhang, D. (2013). Oxidative stress, mitochondrial damage and neurodegenerative diseases. Neural Regen Res. 8(21): 2003–2014.

Gurney, ME., Pu, H., Chiu, AY., et al. (1994). Motor neuron degeneration in mice that express a

human Cu,Zn superoxide dismutase mutation. Science. 264:1772–1775.

Hargreaves, D.C., Crabtree, R.C. (2011). ATP-dependent chromatin remodeling: genetics, genomics and mechanisms. Cell. Res. 21:396-420.

Hirose, Y, Manley, JL (1998).RNA polymerase II is an essential RNA polyadenylation factor.Nature. 3; 395(6697):93-6.

Ho, L. & Crabtree, G. R. (2010). Chromatin remodelling during development. Nature. 463, 474–484.

Hochheimer,A. and Tjian, R. (2003). Diversified transcription initiation complexes expand promoter selectivity and tissue-specific gene expression. Genes Dev. 17, 1309–1320.

Kadam,S., Emerson, B.M. (2003). Transcriptional specificity of human SWI/SNF BRM chromatin remodeling complexes. Mol Cell.11:377-389.

Kleiman, FE., Manley, JL. (1999). Functional interaction of BRCA1-associated BARD1 with polyadenylation factor CstF-50. Science. 3; 285(5433):1576-9.

Kleiman, FE., Wu-Baer, F., Fonseca, D., Kaneko, S., Baer, R., Manley, JL. (2005). BRCA1/BARD1 inhibition of mRNA 3' processing involves targeted degradation of RNA polymerase II.Genes Dev.15; 19(10):1227-37.

Kornblihtt, A.R., de la Mata, M., Fededa, J.P., Munoz, M.J. and Nogues, G. (2004) Multiple links between transcription and splicing. RNA, 10, 1489–1498.

Kornblihtt, A.R., Schor, I.E., Allo, M., Dujardin, G., Petrillo, E. and Munoz, M.J. (2013) Alternative splicing: a pivotal step between eukaryotic transcription and translation. Nat. Rev. Mol. Cell Biol., 14, 153–165.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, K., et al. (2001). Initial sequencing and analysis of the human genome. Nature. 409:860-920.

Landry, J.R. et al. (2003) Complex controls: the role of alternative promoters in mammalian genomes. Trends Genet. 19, 640–648.

Levine, M. and Tjian, R. (2003) Transcription regulation and animal diversity. Nature 424, 147–151.

Liu, N., Hayes, J.J. (2010). When push comes to shove: SWI/SNF uses a nucleosome to get rid of a nucleosome. Mol. Cell. 38:484-486.

Liu, S. and Cheng, C.,(2013). Alternative RNA splicing and cancer. Wiley Interdiscip Rev RNA. 2013 ; 4(5): 547–566.

Luger, K., Dechassa, M.L. & Tremethick, D.J. (2012) New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? Nature Rev. Mol. Cell Biol. 13, 436–447.

Lutz, CS., Moreira, A. (2011). Alternative mRNA polyadenylation in eukaryotes: An effective regulator of gene expression. Wiley Interdiscip Rev RNA 2: 22–31.

Marfella, C.G., Imbalzano, A.N. (2007). The Chd family of chromatin remodeling complexes. Mutat. Res.618:30-40.

Mehler, MF. (2008). Epigenetic principles and mechanisms underlying nervous system functions in health and disease . Progress in Neurobiology. 6 : 305 – 41.

Moore, MJ., Proudfoot, NJ. (2009). Pre-mRNA processing reaches back to transcription and ahead to translation. Cell 136: 688–700.

Muchardt, C., Yaniv, M. (1999). ATP-dependent chromatin remodelling: SWI / SNF and Co. are on the job. J. Mol. Biol.293:187-198

Muller, M., Schleithoff, ES., Stremmel, W., Melino, G., Krammer, PH., Schilling, T. (2006). One, two, three–p53, p63, p73 and chemosensitivity. Drug Resist Updat 9: 288–306.

Munch, C., O'Brien, J., Bertolotti, A. (2011). Prion-like propagation of mutant superoxide dismutase-1 misfolding in neuronal cells. Proc Natl Acad Sci U S A. 108:3548–3553

Papoulas, O., Beek, SJ., Moseley, SL., McCallum, CM., Sarte, M., Shearn, A., Tamkun, JW. (1998). The Drosophila trithorax group proteins BRM, ASH1 and ASH2 are subunits of distinct protein complexes. Development. 125:3955-3966.

Phelan, ML., Sif, S., Narlikar, GJ., Kingston, RE. (1999). Reconstitution of a core chromatin remodeling complex from SWI/SNF subunits. Mol. Cell. 3:247-253.

Pozner, A. et al. (2000) Transcription-coupled translation control of AML1/RUNX1 is mediated by cap- and internal ribosome entry sitedependent mechanisms. Mol. Cell. Biol. 20, 2297–2307 28

Resch, A.M., Ogurtsov, A.Y., Rogozin, I.B., Shabalina, S.A. and Koonin, E.V. (2009) Evolution of alternative and constitutive regions of mammalian 5'UTRs. BMC Genomics, 10, 162.

Reyes, JC., Barra, J., Muchardt, C., Camus, A., Babinet, C., Yaniv, M. (1998). Altered control of cellular proliferation in the absence of mammalian brahma (SNF2α). EMBO J.17:6979-6991.

Rigo,F. and Martinson,H.G. (2008) Functional coupling of last-intron splicing and 3’-end processing to transcription in vitro: the poly(A) signal couples to splicing before committing to cleavage. Mol. Cell. Biol., 28, 849–862.

Rosen, DR., Siddique, T., Patterson, D., et al. (1993). Mutations in Cu=Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. Nature. 362:59–62.

Shabalina,S.A., Spiridonov,A.N., Spiridonov,N.A. and Koonin,E.V. (2010) Connections between alternative transcription and alternative splicing in mammals. Genome Biol. Evol., 2, 791–799.

Sif, S., Saurin, A.J., Imbalzano, A.N., Kingston, R.E. (2001). Purification and characterization of mSin3A-containing Brg1 and hBRM chromatin remodeling complexes. Genes Dev. 15:603-618.

Strobeck, M.W., Reisman, D.N., Gunawardena, R.W., Betz, B. L., Angus, S.P., Knudsen, K.E., Kowalik, T.F., Weissman, B.E., Knudsen, E.S. (2002). Compensation of BRG-1 function by BRM. Biochemistry. 277:4782-4789.

Strong, MJ. (2010). The evidence for altered RNA metabolism in 53. amyotrophic lateral sclerosis (ALS). J Neurol Sci. 288 : 1 – 12 .

Synofzik, M., Ronchi, D., Keskinm I., et al. (2012). Mutant superoxide dismutase-1 indistinguishable from wild-type causes ALS. Hum Mol Genet . 21:3568–3574.

Tang, L., Nogales, E., Ciferri, C. (2010). Structure and Function of SWI/SNF Chromatin Remodeling Complexes and Mechanistic Implications for Transcription. Prog Biophys Mol Biol.; 102:122,128.

Tomasini, R., Tsuchihara, K., Wilhelm, M., Fujitani, M., Rufini, A., Cheung, CC., Khan, F., Itie-Youten, A., Wakeham, A., Tsao, MS., et al. (2008). TAp73 knockout shows genomic instability with infertility and tumor suppressor functions. Genes Dev 22: 2677–2691.

Trapnell, C., Williams, BA., Pertea, G., Mortazavi, A., Kwan, G., van Baren, MJ., Salzberg, SL., Wold, BJ., Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol. 28: 511–515

Tsukiyama, T., Daniel, C., Tamkun, J., Wu, C. (1995). ISWI, a member of the SWI2/SNF2 ATPase family, encodes the 140 kDa subunit of the nucleosome remodeling factor. Cell. 83:1021-1026.

Wang, W., Xue, Y., Zhou, S., Kuo A., Cairns, B.R., Crabtree G.R. (1996). Diversity and specialization of the mammalian SWI/SNF complexes. Genes Dev. 10:2117-2130.

Winston, F., Allis, C.D. (1999). The bromodomain: a chromatin-targeting module? Nature, 6:601-604

Yamashita,R., Wakaguri,H., Sugano,S., Suzuki,Y. and Nakai,K. (2010) DBTSS provides a tissue specific dynamic view of Transcription Start Sites. Nucleic Acids Res., 38, D98–D104.

Chapter 2:

Oxidative stress controls the choice of alternative last exons via a Brahma-BRCA1-CstF pathway

Aurora Rigamonti[1]*, Gabriele A. Fontana[1#]*, Silvia C. Lenzken[1], Giuseppe Filosa[1], Reinaldo Alvarez[1], Raffaele Calogero[3], Marco E. Bianchi[2], Silvia M.L. Barabino[1]

1 Department of Biotechnology and Biosciences, University of Milan-Bicocca Piazza della Scienza 2, 20126, Milan, Italy

2 Division of Genetics and Cell Biology, San Raffaele Scientific Institute and University, Via Olgettina 60, 20132 Milan, Italy.

3 Department of Biotechnology and Health Sciences, University of Torino, Via Nizza 52, I-10126 Torino, Italy.

#Present address: Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058, Basel

* Both authors contributed equally to the work.

Nucleic Acids Research 2016;

doi: 10.1093/nar/gkw780

2.1 Introduction

Pre-mRNA splicing and 3’ end processing are essential metazoan genes. Recently, RNAseq analysis showed that not only alternative splicing (AS) but also alternative polyadenylation (APA) is more frequent and complex than previously anticipated (Di Giammartino et al.,2011-Wang et al., 2008-Shi et al., 2012-Pickrell et al. 2010). The choice of alternative poly(A) sites generates different 3’UTRs that can affect translation, stability and localization of the mRNA. Alternative pre-mRNA processing changes the length of the 3’ UTR during cell differentiation contributing to the regulation of gene expression (Sandberg et al., 2008-Ji et al., 2009). When coupled to the inclusion of an alternative last exon (ALE), alternative polyadenylation leads to the generation of mRNA variants that differ in their 3’UTR and that may encode proteins with different C-terminal regions. Whereas the molecular details of pre-mRNA 3’ end processing are rather well known, how the choice of APA sites is regulated is only partially understood. The mature 3’ ends of most eukaryotic mRNAs are generated by endonucleolytic cleavage of the primary transcript followed by the addition of a poly(A) tail to the upstream cleavage product (Di Giammartino et al., 2011),Proudfoot et al.,2011).

Maturation of the 3’ end is executed by a large multicomponent complex that is assembled in a cooperative manner on specific cis-acting sequence elements in the pre-mRNA (Danckwardt et al., 2008,Elkon et al., 2013). In mammalian cells the Cleavage and Polyadenylation Specificity Factor (CPSF) recognizes the consensus hexanucleotide AAUAAA, whereas the Cleavage stimulation Factor (CstF), a hexameric complex of subunits of 77, 64 and 50 kDa (Edwards et al., 2008, Bai et al.,2007), binds a more degenerate GU- or U-rich element downstream of the poly(A) site. Both CPSF and CstF interact with RNA polymerase II (RNAPII) at the promoter and appear to remain associated with it during elongation (Dantonel et al., 1997, McCracken et al., 1997). CstF50 and CstF77 interact specifically with the carboxy-terminal domain (CTD) of RNAPII largest subunit. CstF50 was also shown to bind the BARD1/BRCA1 ubiquitin ligase after DNA damage, resulting in the inhibition of 3’ end processing(Kleiman et al., 2001). This observation and the fact that APA is modulated in development, differentiation and neuronal activation (reviewed in Elkron et al., 2013) indicates that 3’ end processing can be regulated in response to physiological and pathological stimuli.

Oxidative stress is a widely occurring phenomenon in biological systems, and is due to an imbalance between the intracellular production or influx of reactive oxygen species (ROS) and the availability of antioxidant compounds, such as glutathione. Oxidative stress has been implicated in the etiology of many neurodegenerative disorders, including Alzheimer's and Parkinson's diseases (Lin et al., 2006). At the cell level, oxidative stress elicits a wide spectrum of responses ranging from proliferation to growth arrest, senescence, or cell death. The particular outcome reflects the balance between a variety of intracellular stress signalling pathways that are activated in response to the oxidative insult and that ultimately modulate gene expression. We recently described that exposure of human neuroblastoma SH-SY5Y cells to different sources of ROS leads to genome-wide alternative splicing changes, modifying the relative proportion of alternatively spliced forms (Lenzken et al., 2011). Here, we show that oxidative stress specifically affects the choice of ALEs increasing the production of transcripts variants terminating at a more proximal poly(A). Oxidative stress induces the transcriptional downregulation of Brahma (BRM), one of the two alternative ATPase subunits of the SWI/SNF complex. We find that in normal condition BRM is enriched on the

proximal ALE. In addition we observe the accumulation of BARD1, a protein that forms a functional heterodimer with BRCA1, which has E3 ubiquitin-ligase activity and interacts with the 50 kDa subunit of CstF inhibiting 3' end processing (Kleiman et al., 1999). Consistent with these observations, we detect an ubiquitinated pool of CstF50 and show that ubiquitination is mediated by BARD1/BRCA1. Taken together, our results suggest that the presence of BRM on the proximal exon leads to the BARD1/BRCA1-mediated ubiquitination of CstF50 and the inhibition of 3' end processing at the proximal poly(A). This in turn allows transcription to proceed to the distal terminal exon.

2.2 Results

BRM is down-regulated in response to oxidative stress.

Oxidative stress has long been implicated in neuronal cell death that is associated with neurodegenerative disorders. For example, exposure of cells to paraquat (N,N'-dimethyl-4,4' bipyridinium dichloride, PQ), a neurotoxic herbicide, increases the production of the superoxide anion (Cocheme et al., 2008), and has been linked to the incidence of Parkinson's disease.

Moreover, the second most frequent cause of hereditary Amyotrophic Lateral Sclerosis (ALS) are mutations in the cytoplasmic superoxide dismutase 1 (SOD1) protein and have been shown to increases intracellular ROS levels (Ciriolo et al., 2000). We recently showed that treatment with PQ or expression of the mutant SOD1(G93A) protein cause extensive changes in both mRNA expression and alternative splicing in human neuroblastoma SH-SY5Y cells (Lenzken et al., 2011). Among the most down-regulated genes was SMARCA2, the gene encoding BRM (Figure 1A). SOD1(G93A) expression almost abrogated BRM expression at the protein level, while expression of the hSWI/SNF alternative ATPase subunit BRG1 was not significantly affected (Fig.1B). Similarly, treatment of SH-SY5Y cells with PQ strongly reduced BRM mRNA and protein levels without affecting BRG1 expression (Fig. 1A and Fig.1B). To test a possible causal link between oxidative stress and BRM expression, we exposed primary adult fibroblasts to different sources of ROS. Treatment with H2O2 or with PQ (Fig. 1C) effectively reduced BRM expression.

We set out to identify regulatory elements in the SMARCA2 promoter that respond to oxidative stress. Inspection of the -3444/+57 region identified features

associated to promoter sequences (Fig. 2A): peaks of evolutionary conservation, a CpG island, a putative DNase hypersensitive region, and peaks of enrichment of acetylation of histone 3 lysine 27 (H3KAc27). A high GC-content and the absence of a detectable TATA-box suggested that the promoter belongs to the class of GC promoters (Fig. 2B). We then cloned a 3.4 kb genomic fragment (from position –3344 to +57) into a promoter-less luciferase reporter. This region conferred sensitivity to PQ and showed reduced luciferase expression when transfected in SH-SY5Y/SOD1(G93A) cells (Fig. 3A). Further 5' deletion analysis pointed to the -146/+57 fragment as critical for the oxidative stress response (Figure 3B). Addition of resveratrol, a natural antioxidant, to SH SY5Y/SOD1(G93A) cells reduced ROS production (Figure 3B) and restored luciferase expression to both the long 3344/+57 and the minimal -146/+57 promoter constructs (Figure 3C). Overall these results indicate that SMARCA2 transcription and hence BRM expression are reduced by oxidative stress.

Selection of the proximal alternative last exon is favored in BRM-depleted cell

Using splicing-sensitive microarrays we had previously detected a large number of AS changes (262 genes,

involving 418 exons) in SH-SY5Y/SOD1(G93A) cells (16). Within this dataset, 89 exons (in 78 genes) appeared as ALEs. We validated by PCR six of these genes: in 5 out of 6 genes, the distal ALE was favoured in the presence of BRM (i.e. in SH-SY5Y/SOD1 cells), whereas the proximal ALE was preferred when BRM was expressed at low level as in SH-SY5Y/SOD1(G93A) cells (Figure 4).

We chose the RPRD1A gene for mechanistic analysis because of the ease of PCR quantification of the ALE choice. As shown in Figure 5A, alternative splicing of the RPRD1A primary transcript generates two mRNA variants with different terminal exons. When exon 8 (e8) is chosen as proximal ALE, a shorter mRNA is produced. In contrast, a longer mRNA isoform is generated when e8 is skipped and cleavage and polyadenylation occur in the distal ALE exon 10 (e10). Exon e8 was preferentially skipped in SOD1 cells, where BRM is expressed, while it was preferentially included in SOD1(G93A) cells in which BRM is hardly detected (Figure 5B & C). An inverse correlation between BRM expression and e8 skipping was also found in SOD1 cells where BRM expression was silenced with a specific shRNA (Figure 5B). Moreover, when BRM was exogenously expressed in the SOD1(G93A)

background, e8 was mostly skipped and the mRNA variant ending in e10 was favoured (Figure 5C, and Figure 11 for expression levels). In contrast, neither silencing of BRG1 in SOD1 cells nor its overexpression in SOD1(G93A) cells affected the ratio of the two mRNA isoforms (Fig.5B & C). The transient transfection of expression constructs for BRM and BRG1 in HEK293T cells further confirmed the BRM-dependent inhibition of e8 selection (Figure 5D and Figure 11 for expression levels). Similar results in the different cell lines were obtained for the SLC6A15 gene (Figure 5E & F).

Overall, these results indicate that a high expression level of BRM favours the skipping of the proximal ALE of RPRD1A.

BRM accumulates on the skipped proximal ALE

A well-known example of selection of alternative poly(A) sites located in different terminal exons is provided by the immunoglobulin mu (Igμ) gene. The regulation of this processing event was shown to depend on the different levels of the 3’ end processing factor CstF during the maturation from B cell to a plasma cell (Takagaki et al., 1996 and 1998). We

therefore tested the expression levels of the core cleavage and polyadenylation factors in SH-SY5Y/SOD1 and in SH-SY5Y/SOD1(G93A) cells, but did not detect significant differences in the levels of CPSF, CstF and CF Im subunits (data not shown).

BRM was previously shown to accumulate on the variant exon v5 of the CD44 gene and to promote its inclusion by modulating the elongation rate of RNAPII, and by interacting with the splicing factor Sam68 (Batsche et al., 2006). We thus explored the possibility that BRM could contribute to the choice of ALE by regulating the elongation rate of the polymerase. First, we characterized the distribution of BRM and BRG1 along the terminal region of the RPRD1A gene from e6 to e10 by chromatin immunoprecipitation (ChIP, Figure 6A). In SOD1(G93A) cells, where it was hardly expressed, BRM was distributed equally on all exons. In contrast, in SOD1 cells BRM accumulated on the proximal ALE (Figure 6B). BRG1, on the other hand, was distributed equally on all exons in both the cellular lines (Figure 6C). Next, we analysed the distribution of the elongating polymerase, using the N20 antibody that recognizes the N-terminal region of RNAPII. An increased density of RNAPII was detected on e7 and e8

in SOD1(G93A) cells and in both cell lines on e10 (Figure 6D).

We also performed ChIP assays with the H14 antibody to RNAPII phosphorylated at Ser5 of the CTD, a modification associated with a paused polymerase (Morris et al., 2005, Boehm et al., 2003) (Figure 6E). We detected a slight accumulation of p-Ser5 RNAPII on e8 in SOD1(G93A) cells. In addition, we observed a high density on e10 in both cell lines, where this exon is used as terminal exon, albeit with different efficiency. This result is reminiscent of the observed pausing of RNAPII on terminal exons in yeast (Carrillo et al, 2010).

Our findings indicate a modest inverse correlation between the accumulation of BRM and the accumulation of RNAPII on included exons. Thus, the contribution of RNAPII pausing to the inclusion of the proximal exon appears small, at variance with what previously observed for the inclusion of internal exons (26).

BRM does not require its ATPase activity to affect ALE choice

The SWI/SNF chromatin-remodeling complex destabilizes histone-DNA interactions and alters nucleosome positions. Recent studies indicate that

nucleosome positioning may influence inclusion of internal exon sequences (Tilgner et al., 2009, Schwartz et al., 2009). We therefore wondered whether BRM might contribute to the choice of ALE by moving nucleosomes around the proximal last exon. We thus identified nucleosome positions and occupancy in SOD1 and SOD1(G93A) cells, using micrococcal nuclease (MNase) digestion in cells synchronized in G0. We used mononucleosomal DNA fragments (150 bp) generated by MNase as a template for qPCR with overlapping primer sets across a ~1 kb region centered on e8 of the human RPRD1A genomic locus. As shown in Figure 7A, the MNase digestion pattern did not significantly differ between the two cell lines, indicating that the level of BRM protein had no influence on nucleosome distribution in the interrogated region of the RPRD1A gene.

To further exclude that the chromatin remodeling activity of BRM-containing SWI/SNF complexes may contribute to e8 skipping in SOD1 cells, we analysed the alternative splicing pattern in SH-SY5Y cells stably expressing a BRM mutant carrying the K755A substitution in the ATPase domain that abrogates the chromatin remodeling activity (Richmond et al., 1996) (Figure 7B, and Figure 11 for expression levels) we also

analysed HEK293T cells transiently transfected with the same construct (Figure 7C, and Figure 11 for expression levels). Similarly to the wild type, expression of the mutant BRM stimulated exon 8 skipping and the selection of the distal ALE.

Taken together these data indicate that the chromatin remodeling function of BRM is not required for the inhibition of the inclusion of the proximal ALE in SOD1 cells. Therefore, we hypothesized that BRM might contribute to e8 skipping by negatively acting on pre-mRNA processing, possibly by recruiting an inhibitory complex.

BARD1 and CstF50 associate with BRM on the skipped proximal ALE

Previously, a BARD1-CstF complex was shown to inhibit 3' processing of nascent transcripts deriving from premature transcription termination at sites of DNA damage (Mirkin et al., 2008). Therefore, we speculated that BARD1 (and its partner BRCA1) might also be involved in the inhibition of transcription termination at the proximal poly(A) site. Indeed, we found that endogenous BRCA1 and BARD1 in HEK293T cells co-immunoprecipitate with SNF5 and BRM (Figure 8A),

indicating that both proteins can associate with the SWI/SNF complex.

Next, we determined that in SOD1 cells, where BRM is present at higher level, BARD1 is recruited on the RPRD1A gene on e8 (Figure 8B). Conversely, in SOD1(G93A) cells BARD1 was enriched on e7. We then showed that BRM and BARD1 were present on the same DNA fragments by performing two successive ChIP assays (ChIP-reChIP) using anti-BRM followed by anti BARD1 antibodies (Figure 8C). Finally, overexpression of BRM and BARD1 proteins (Figure 11 for expression levels) favoured the inclusion of the distal ALE (Figure 8D).

Taken together, these data indicate that BARD1 and BRCA1 are integral to the process of ALE choice.

The BARD1/BRCA1 complex promotes the ubiquitination of CstF50

BARD1 and CstF50 were shown to interact both in vitro and in vivo (Edwards et al.,2008; Kleiman et al., 1999 e 2001). We confirmed this interaction by cotransfecting HEK293T cells with plasmids expressing myc-tagged BARD1 and Flag-tagged CstF50. Immunoprecipitation with either anti-myc or anti-Flag

antibodies followed by Western blotting demonstrated that BARD1 and CstF50 interact physically (Figure 9A).

Moreover, ChIP-reChIP experiments using anti-BARD1 followed by CstF50 antibodies demonstrated that BARD1 and CstF50 co-assemble on e8 of RPRD1A in SOD1 cells, where the distal ALE is preferentially selected (Figure 9B).

Since BRCA1 has E3-ubiquitin ligase activity (Manley et al., 2002), we wondered whether the interaction between BARD1 and CstF50 may lead to CstF50 ubiquitination. HEK293T cells were contransfected with plasmids encoding Flag-CstF50 and histidine-tagged ubiquitin. Indeed, CstF50 immunoprecipitated with an anti Flag-antibody and tested positive to the His-specific antibody (Figure 9C). We further verified if a fraction of CstF is always ubiquitinated in unperturbed cells.

To this end, cells were lysed under denaturing conditions and the ubiquitinated protein fraction was immunoprecipitated with an anti-ubiquitin antibody. Western blot analysis with an anti-CstF50 antibody confirmed the ubiquitination of CstF50 (Figure 9D).

We next investigated whether the BRCA1 E3 ubiquitin ligase activity was responsible for the ubiquitination of CstF50. To this end, BRCA1 was transiently silenced

with a specific siRNA. Three days after transfection the level of the BRCA1 protein was reduced to ~25% (Figure 11). Immunoprecipitation of the ubiquitinated protein fraction with an anti-ubiquitin antibody from BRCA1-silenced cell extracts showed a reduction in the amount of ubiquitinated CstF50 protein (Figure 10A). Moreover, expression of deltaN-BRCA1 protein, lacking the amino-terminal RING domain (Chiba et al.,2002), reduced the amount of ubiquitinated CstF50 (Figure 10B). Together these results confirm the BRCA1/BARD1-mediated ubiquitination of CstF50.

Next, we determined the effect BRCA1 silencing on pre-mRNA splicing. As shown in Figure 10C, siRNA-mediated knockdown of endogenous BRCA1 resulted in the preferred inclusion of the proximal ALE.

Overall these results suggest a model in which BRM positioned on the proximal ALE mediates the recruitment of the BARD1/BRCA1 complex, which in turn ubiquitinates CstF50 impairing 3' end processing at the proximal poly(A) site. This allows transcription to proceed to the distal terminal exon.

2.3 Discussion

Oxidative stress is a common biological event, and has been associated with cancer, neurodegenerative

diseases and aging. We previously reported that oxidative stress affects splice site selection (Lenzken et al., 2011). Here, we show that it affects in particular the choice of ALEs, generally promoting the inclusion of the proximal ALE; it does so by inducing a severe decrease in the expression of the SWI/SNF subunit BRM. Our data also provide evidence for a regulatory mechanism that partially inhibits 3' end processing at the proximal poly(A) site.

Recently, knock-down of BRM and of other SWI/SNF subunits was shown to affect splicing associated polyadenylation in Drosophila S2 cells by an unknown mechanism (Waldholm et al, 2011). We suggest that the involvement of BRM-containing SWI/SNF complexes in the regulation of pre-mRNA processing is similar in Drosophila and mammalian cell, and thus is conserved in evolution, and possibly not restricted to the choice of ALE choice following oxidative stress.

We specifically analyse the ALE choice on RPRD1A gene, which generates two main mature transcripts: one including exons 1 to 7 followed by exon 8 (the proximal ALE) and the polyA, and one where exons 1-7 are followed by exons 9-10 and polyA. Both mature transcripts are present in cells, whether undergoing

oxidative stress or not, but their relative abundance changes.

We find that BRM is associated with RPRD1A e8 in normal conditions, whereas oxidative stress strongly reduces BRM expression and its association with e8. BRM and SNF5 (another core subunit of the SWI/SNF complex) co-immunoprecipitate the BRCA1/BARD1 E3 ubiquitin ligase. Indeed, our ChIP and ChIP-ReChIP experiments reveal that BRM, BARD1 and CstF co-accumulate on e8 in SOD1 cells, in which the choice of the distal ALE is promoted. Likewise, overexpression of BRM and BARD1 proteins promoted the choice of the distal ALE. Further, we show that BRCA1/BARD1 mediates the ubiquitination of CstF50. Mutation or silencing of BRCA1 reduces CstF50 ubiquitination, and promoted the choice of the proximal ALE. We thus propose that the presence of BRM on the proximal ALE promotes the recruitment of the BARD1/BRCA1 ubiquitin ligase, which by ubiquitinating CstF50 causes the inhibition of 3' end processing at the proximal poly(A) site and allows the continuation of transcription through exons 9-10. A fraction of the pool of CstF50 is always monoubiquitinated, which is consistent with the fact that some transcripts incorporating the distal ALE are always present.

Our findings are consistent with previous knowledge. The coupling of transcription and 3’ end processing is well documented (Elkron et al., 2013). RNAPII and other transcription factors interact with components of the 3’ end processing machinery (Shi et al., 2009, Nagaike et al.,2011, Rozenblatt-Rosen et al., 2009, Katahira et al., 2013). Generally, these interactions appear to promote the use of the proximal poly(A) site, since this is transcribed first and is encountered first by the 3’-end-processing machinery. Cleavage and polyadenylation at the proximal poly(A) site is favoured by a reduction of the elongation rate of the transcribing polymerase (Pinto et al., 2011). In this respect, it not surprising that the mechanism we describe restrains the cleavage of the nascent transcript.

In vivo BARD1 can interact with BRCA1 (Wu et al.,1996) and the BARD1/BRCA1 heterodimer has E3 ubiquitin ligase activity. Moreover, BRCA1 interacts with the SWI/SNF chromatin remodeling complex (Bochard et al., 2000). Finally, the interaction between CstF50 and BARD1 in UV-treated cells was shown to inhibit erroneous 3’ processing of nascent, truncated transcripts deriving from premature transcription termination at sites of DNA damage.

BRM has already been implicated in alternative splicing: it was previously shown to promote inclusion of variant internal exons of the CD44 gene by inducing a transient pausing of the transcribing RNAPII. However, in Drosophila, BRM-containing SWI/SNF complexes, while inducing RNAPII pausing, can negatively regulate splicing (Bochar et al., 2000). In this case, RNAPII stalling, and the consequent splicing reduction, were not induced by BRM but were rather dependent on SNF5. Instead, BRM remodeling activity was shown to be necessary for the subsequent chromatin remodeling following the release of the stalled polymerase (Zraly et al., 2012). In line with the report by Zraly and coauthors, we found that BRM negatively affects the inclusion of the proximal ALE. This effect does not require its ATPase activity. Moreover, we did not detect differences in the distribution of nucleosomes on the proximal ALE. These findings suggest that the choice of ALE is not correlated with the nucleosome remodeling activity of BRM. We also observed that RNAPII is not greatly enriched on the skipped proximal ALE. Instead, we observed an inverse correlation between the accumulation of BRM and RNAPII particularly on the distal ALE, where there was a distinct enrichment of RNAPII but BRM was virtually absent. Our data indicate that in the case of the choice of alternative terminal

exons, the modulation of the elongation rate of RNAPII is not the main role of BRM, and rather implicate a different molecular mechanism (Elkon et al., 2013). The choice of alternative poly(A) sites located in different terminal exons occurs in conjunction with splicing, and splicing factors are known to influence 3' processing (reviewed in Millevoi et al., 2010). Thus, the mechanism we describe is certainly not the unique determinant of the choice between proximal and distal terminal exons. Moreover, we cannot exclude that BRM and BRCA1/BARD1 may also affect binding of splicing factors to the terminal exon. Future work will determine how the interplay between chromatin remodeling complexes, splicing and polyadenylation factors eventually determines which terminal exon will be included in the mature transcript.

2.4 Materials and Methods

Cell culture and transfections

Human neuroblastoma SH-SY5Y, SH-SY5Y/SOD1, SH-SY5Y/SOD1(G93A) and HEK293T cells were cultured in D-MEM High Glucose medium (Gibco, Invitrogen), 10% fetal bovine serum (FBS), 2.5 mM L-glutamine, 100

U/ml penicillin, and 100 µg/ml streptomycin (Euroclone) at 37°C with 5% CO2. Paraquat (Sigma-Aldrich) treatment of SH-SY5Y cells was carried out as described (16). Resveratrol (Sigma Aldrich) was dissolved in DMSO, and cells were treated with 10 µM resveratrol for 24 hr. Control cells were treated with vehicle. Adult dermal fibroblats (ATCC, PCS 201-012) were cultured D-MEM High Glucose medium (Gibco, Invitrogen), 10% fetal bovine serum (FBS), 5 mM L-glutamine, 100 U/ml penicillin, and 100 µg/ml streptomycin (Euroclone) at 37°C with 5% CO2. Treatment with 2 mM paraquat was performed for 48 hr, while treatment with 0.2 mM H2O2 (Sigma Aldrich) was performed for 24 hr. Cells were transfected using polyethylenimine (PEI, Sigma Aldrich) according to the manufacturer's instructions. Stably transfected cells were obtained by selection with 1µg/ml puromycin (Sigma Aldrich). For depletion of BRCA1, HEK293T cells were transfected twice (day 0 and 3) in 6-well plates with 75 pmoles of the corresponding Silencer® select siRNAs and Negative control siRNA #1 (Ambion). Transfection was carried out using Lipofectamine® RNAiMax reagent (Invitrogen) following manufacturer's instructions. Cells were expanded and lysed at day 6.

For BRM-K755A overexpression, cells were electroporated with Amaxa® nucleofector® system (Lonza) using Amaxa® Cell Line Nucleofector® Kit V according to the manufacturer's instructions.

Intracellular ROS were measured with 2',7'-dichlorodihydrofluorescein diacetate (H2DCF-DA; Sigma Aldrich, D6883). Cells were exposed to H2DCF-DA (20 μM) during the last 30 min of culture, then collected and washed with PBS. A blank sample (cells not exposed to H2DCF-DA) was also prepared. The H2DCF-DA fluorescence was measured by flow cytometry after addition of propidium iodide (PI) to the samples. Only the cellular population of PI impermeable cells was considered for measuring the fluorescence intensity of H2DCF-DA (Rizzardini et al., 2005).

Plasmids and Antibodies

The cDNAs encoding the N-terminal flag-tagged human BRM and N-terminal flag-tagged BRG1 proteins were a gift from B. Emerson. These cDNAs were subcloned into the SmaI – SalI (New England Biolabs) sites of the pAD5-CMV-Wpre-PGK-Puro expression vector (a kind gift from S. Philipsen). The mutation K755A (Richmond et al.,1996) was introduced using a two-step mutagenic

PCR procedure using Phusion polymerase (Finnzymes). The interfering shRNA oligonucleotides targeting human SMARCA2 and SMARCA4 transcripts were designed using the SiDesign Center (Dharmacon). The shRNA primers were cloned into pSUPuro vector. pSUPuro, and the pSUPuro ß2 T-Cell Receptor Beta used as an unrelated control shRNA were gifts from M.D. Ruepp. BARD1 and BRCA1 constructs were a gift of N. Chiba. The cDNA encoding CstF50 was subcloned from pcDNA3 HA CstF50 (a gift from M.D. Ruepp) into a p3XFLAG-myc-CMV26 removing the myc tag. The Histidine-tagged Ubiquitin was a gift of M.L. Guerrini. All constructs were verified by sequencing (BMR Genomics). All oligonucleotide sequences are listed in Table 1. The commercial antibodies used are listed in the Supplementary Table 1. Non-immune rabbit IgGs (Millipore) were used as a control in the immunoprecipitation assays.

Luciferase Assay

SH-SY5Y cells were transiently co-transfected with the indicated pGL2 luciferase reporter plasmids and with the Renilla-encoding pRL-TK plasmid (Promega Inc.). 24h after transfection, cells were lysed and luciferase activity quantified using the Dual Luciferase Reporter

kit (Promega Inc.) and a Berthold luminometer (Berthold Technologies). For the luciferase experiments, paraquat was added 3 hours after transfection.

bioinformatics analysis

Transcripts characterized by alternative splicing events at their 3' end were detected by R scripting using Bioconductor 2.12 packages GenomicRanges, TxDb.Hsapiens.UCSC.hg19.knownGene and HuExExonProbesetLocation (www.bioconductor.org). Transcripts were extracted from TxDb.Hsapiens.UCSC.hg19.knownGene (80922 transcripts). After removal of transcripts lacking a link to Entrez Gene Identifier (Maglott et al., 2011), 71350 transcripts (265661 exons), associated to 22932 EG, were left for further analysis. Subsequently we selected all genes associated to the presence of alternative splicing even at the 3' end (12839 genes, 58451 transcripts, 26608 exons involved in ALE). Affymetrix Human Exon 1.0 ST Array (HuEx-1_0-st) exon-level probe sets chromosomal locations were extracted from HuExExonProbesetLocation (Sanges et al.,2007). Only exon-level probesets associated to the Affymetrix core annotation were considered (284805 exon-level probesets). These exon-level probesets mapped on

12839 genes (59986 UCSC transcripts, 230112 exons). Out of the 230112 exons 22983 were associated to ALE. Alternatively spliced exon-level probesets were retrieved by Lenzken et al. (16). 406 exon-level probesets mapping on 418 exons were associated to 262 genes (1191 UCSC transcripts, 4844 exons). Within the 418 exons 89 exons (78 UCSC genes) were involved in ALE.

RNA analysis

RNA preparation and RT-PCR reactions were performed as described (16). The sequences of the exon-specific primers are listed in Table 2. Assay conditions were optimized for each gene with respect to primer annealing temperatures, primer concentrations, and MgCl2 concentrations. Quantification was performed with Bioanalyzer 2100 (Agilent Technologies).

Chromatin immunoprecipitation and qPCR analysis

107 SH-SY5Y cells were cross-linked with 1% formaldehyde (Sigma-Aldrich), and quenched with the addition of 125 mM glycine (Sigma Aldrich). Cell pellets

were lysed in Lysis Buffer (50 mM HEPES-KOH pH 7.5, 140 mM NaCl, 1 mM EDTA pH 8, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1 % SDS, protease inhibitors (Roche) for 1 hour on ice, and then sonicated with a Branson 250 sonifier (Branson Inc.) to obtain chromatin fragments of ~400 nt. Aliquots corresponding to 5x106 cells were flash-frozen in liquid nitrogen and then maintained at -80°C. For each immunoprecipitation, 1% of the total genomic DNA was saved as input DNA. The chromatin solution was precleared at 4°C for 1h with sepharose beads. Immunoprecipitation was carried out using the Chromatin IP Assay Kit (Millipore), following manufacturer's instructions. Bound material was eluted with 500 µl of Elution Buffer (1% SDS, 100 mM NaHCO3) for 1 hour at room temperature. Crosslinking was reversed at 65°C overnight, with the addition of 250 mM NaCl and 2 µg/ml Proteinase K (Sigma). DNA was purified by phenol:choloform:isoamyl alcohol (SigmaAldrich) extraction and ethanol precipitation. For ChIP-ReChIP experiments, eluates from the first ChIP assays were diluted in the ChIP Dilution buffer (Millipore) to reduce the SDS concentration to 0.1% (w/v). Then, the second ChIP was performed. Immunoprecipitated DNA was analyzed in triplicate by qPCR using the SYBR Green method with Mesa Green

qPCR master mix (Biosense) in an ABI PRISM 7500 (Applied Biosystem). The primers used for the qPCRs are listed in the Supplemental Table S2. Data were normalized by the Fold Enrichment Method as follows: Relative Enrichment = 2 – (ΔCt antibody- ΔCt IgG).

Micrococcal nuclease assays

107 SH-SY5Y cells were synchronized in G0 by serum starvation as assessed by FACS analysis (FACSCalibur, BD Biosciences). Cells were cross-linked with 1% (v/v) formaldehyde and quenched with the addition of 125 mM glycine. Cell pellets were lysed in Lysis Buffer (50 mM HEPES-KOH pH 7.5, 140 mM NaCl, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1 % SDS, protease inhibitors (Roche). Nuclei were disrupted by five bursts of sonication. Aliquots of 5x106 cells were flash-frozen in liquid nitrogen and then maintained at -80°C. One undigested aliquot was used as control. Digestion was performed with 20 U of Micrococcal Nuclease (Fermentas) in 1 mM final CaCl2 and MgCl2 for 30 minutes at room temperature, then stopped with 20 mM EDTA. Crosslinking was reversed at 65°C overnight, with the addition of 250 mM NaCl and 2 µg/ml Proteinase K followed by phenol:choloform:isoamyl alcohol extraction and ethanol precipitation. qPCR analysis was performed on

4 ng. The primers used for the qPCRs are listed in Supplemental Table S2. Each sample was analyzed in triplicate, and the average Ct was calculated for each primer set. In order to determine the Relative Nucleosome Occupancy associated with each primer set, the following equation was used: Relative Nucleosome Occupancy = 10(Ct Untreated – Ct MNase).

Nuclear extracts preparation and co-immunoprecipitations

Nuclear protein extracts (NEs) were prepared using according to the Lamond protocol (www.Lamondlab.com). For co-immunoprecipitations, 300 µg of NEs were first precleared with protein A sepharose (GE) for 2 h at 4°C and then incubated with the indicated antibodies or with rabbit IgG as a control. The antibodies were coupled to protein A-Sepharose beads following manufacturer's instructions. Immunoprecipitations were carried out for 2 hours at 4°C on a rotating wheel. Beads were then washed three times with a solution composed of 10 mM TrisHCl, 100 mM NaCl, 0.01% Nonidet P-40, and 0.04% bovine serum albumin (BSA). The proteins bound to the beads were eluted by boiling the samples at 95°C for 10

minutes. Aliquots were analyzed by SDSPAGE and immunoblotting. For ubiquitin immunoprecipitation, cells were lysed in 1% SDS lysis buffer (50 mM Tris HCl pH 7.4, 0.5 mM EDTA, 1% SDS) and further boiled for an additional 10 minutes. Lysates were clarified by centrifugation at 16,000 g for 10 minutes. Supernatant was diluted 10 times with a buffer composed of 50 mM Tris HCl pH 7.4, 150 mM NaCl, 1% Triton and complete Protease Inhibitor Cocktail (Roche). Immunoprecipitation was performed with anti-Multi Ubiquitin mAbAgarose (MBL) overnight at 4°C. The precipitates were washed three times and the samples were resolved on 6% SDS-PAGE.

For Flag-tagged CstF50 immunoprecipitation, cells were lysed as for ubiquitin immunoprecipitation. Immunoprecipitation was performed with ANTI-FLAG® M2 Affinity Gel (Sigma) overnight at 4°C. The precipitates were washed three times and eluted in a rotating wheel for 1h at 4°C with a buffer containing 50 mM TrisHCl pH 7.4, 150 mM NaCl, 1% Triton, complete Protease Inhibitor Cocktail (Roche) and 3Xflag-peptide (5 µg/µl). The samples were resolved on 7% SDS-PAGE.

For myc-tagged BARD1 immunoprecipitation, cells were lysed in 50 mM Tris HCl, pH 7.4, with 150 mM NaCl, 1

mM EDTA, and 1% Triton X-100. Lysates were incubated overnight at 4°C with anti-myc antibody (71D10 Cell Signalling). The complexes were coupled to protein A-Sepharose beads following manufacturer's instructions. Immunoprecipitations were carried out for 2 hours at 4°C on a rotating wheel. Beads were then washed three times with a solution containing 10 mM TrisHCl, 100 mM NaCl, 0.05% Nonidet P-40, and 0.04% bovine serum albumin (BSA). The proteins bound to the beads were eluted by boiling the samples at 95°C for 10 minutes. The sample were resolved on 7% SDS-PAGE. Quantification of western blots was performed by using NIH Image J software the Image Studio software (Odissey FC-Licor) or Image Studio v1.1.

Statistical analysis

Statistical analysis of RT-PCR experiments was performed either using a one-way Anova test followed by a post-hoc Tukey-Kramer multiple comparison test, or a two-tailed, paired t-test. Analysis of FACS experiments was carried out using a one-way ANOVA test followed by a posthoc Tukey-Kramer multiple comparison test. Analysis of Western Blot data was performed using a two-tailed, paired t-test.

2.5 Figures and Tables

Figure 1:

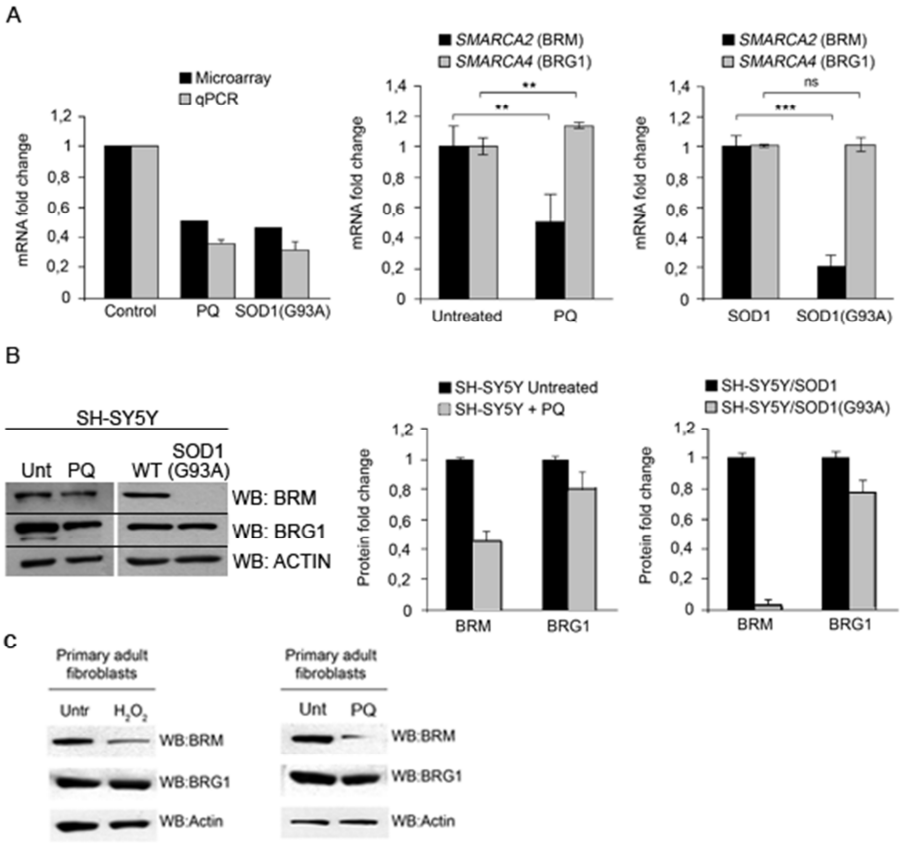


Figure 1: Both PQ and mutant SOD1(G93A) expression reduce the level of the SMARCA2 transcript.

A. qPCR validation of the gene-level microarray data of BRM expression in the indicated cell lines. Assays were performed in triplicate from three biological replicates. Quantification of BRM and BRG1 mRNA levels from six independent biological replicates in SH-SY5Y cells stably expressing wild type SOD1 or the mutant SOD1(G93A) protein and SH-SY5Y cells treated with vehicle (untreated) or paraquat (+ PQ). Quantification of BRM and BRG1 mRNA levels from six independent biological replicates. Variability is expressed as standard deviation. *** = $p < 0.001$.

B. Representative Western blot analysis of BRM and BRG1 in SH-SY5Y cells stably expressing wild type SOD1 or the mutant SOD1(G93A) protein and cells treated with vehicle (untreated) or paraquat (+ PQ) and quantification of BRM and BRG1 protein levels from three independent biological replicates. Error bars represent standard deviation.

C. Representative Western blot analysis of BRM and BRG1 in primary fibroblasts treated with 0.2 mM H2O2 for 24 h and with 2 mM PQ for 48 h.

Figure 2:

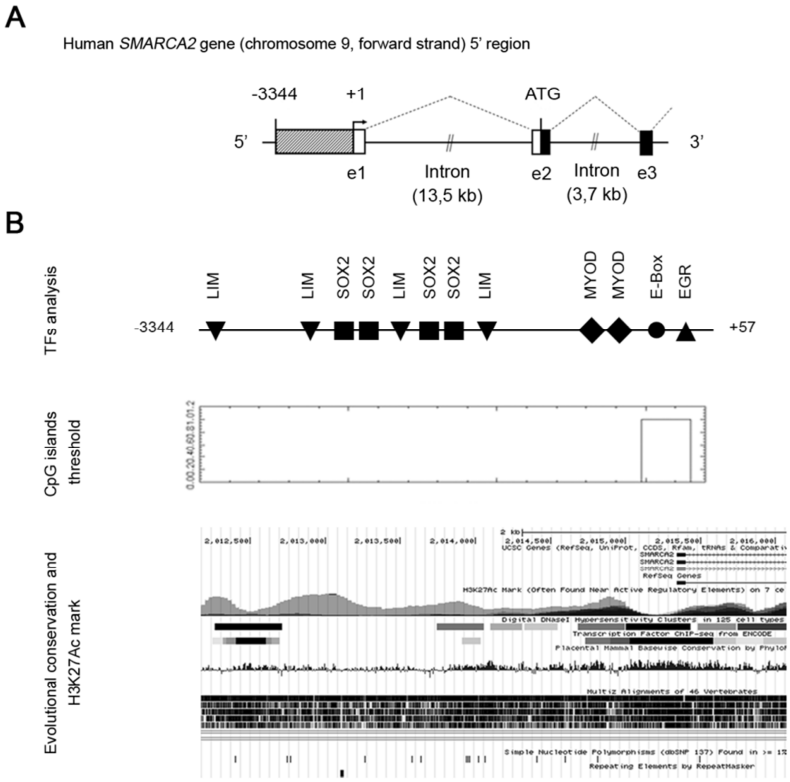

A
Human SMARCA2 gene (chromosome 9, forward strand) 5' region
-3344
+1
ATG
5'
3'
e1
Intron
(13,5 kb)
e2
Intron
(3,7 kb)
e3
B
TFs analysis
LIM
LIM
SOX2
SOX2
LIM
SOX2
SOX2
LIM
MYOD
MYOD
E-Box
EGR
-3344
+57
CpG islands
threshold
Evolutional conservation and
H3K27Ac mark
2 kb
UCSC Genes (RefSeq, UniProt, CCDS, Rfam, tRNAs & Comparati
SMARCA2
SMARCA2
SMARCA2
RefSeq Genes
H3K27Ac Mark (Often Found Near Active Regulatory Elements) on 7 ce
Digital DNaseI Hypersensitivity Clusters in 125 Cell Types
Transcription Factor ChIP-seq from ENCODE
Placental Mammal Basewise Conservation by Phylo
Multiz Alignments of 46 Vertebrates
Simple Nucleotide Polymorphisms (dbSNP 137) Found in >= 1%
Repeating Elements by RepeatMasker

Figure 2: Bioinformatics analysis of *SMARCA2* promoter.

A. Schematic representation of the 5' region f the SMARCA2 gene. White boxes indicates the 5'UTR regions, black boxes indicates the coding regions, while introns are represented by lines. The putative SMARCA2 regulatory region examined in this paper is indicated by the striped box.

B. Results of the bioinformatics study carried out on the -3344/+57 region located at the 5' side of human SMARCA2 gene. The presence of transcription factors binding sites was analysed using TFsearch (upper panel). The presence of CpG islands was evaluated using Emboss CpGPlot (middel panel). The evolutionary conservation, the DNase hypersensitive region, and H3K27Ac occupancy were identified with the UCSC Genome Browser (lower panel).

Figure 3:

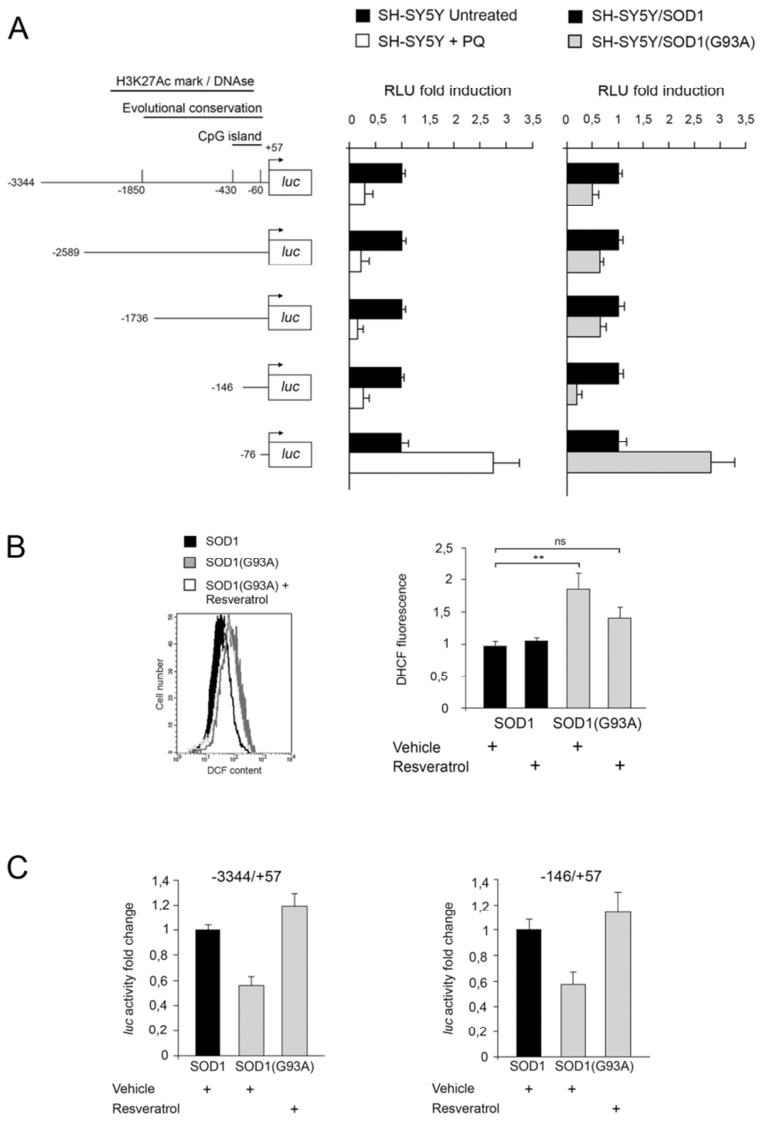

A
SH-SY5Y Untreated
SH-SY5Y + PQ
SH-SY5Y/SOD1
SH-SY5Y/SOD1(G93A)
H3K27Ac mark / DNAse
Evolutional conservation
CpG island
+57
-3344
-1850
-430
-60
-2589
-1736
-146
-76
luc
RLU fold induction
0 0,5 1 1,5 2 2,5 3 3,5
B
SOD1
SOD1(G93A)
SOD1(G93A) + Resveratrol
Cell number
DCF content
DHCF fluorescence
ns
**
SOD1
SOD1(G93A)
Vehicle
Resveratrol
C
-3344/+57
-146/+57
luc activity fold change
SOD1
SOD1(G93A)
Vehicle
Resveratrol

Figure 3: Oxidative stress impairs BRM expression.

A. Promoter activity of the 3.4 kb fragment containing the SMARCA2 regulatory region and its deletion constructs in the indicated cell lines. Cells were transiently cotransfected with the indicated luciferase reporter plasmids and with the Renilla luciferaseencoding pRL-TK plasmid. Renilla luciferase activity was used to normalize the transfection efficiency. Results are expressed as fold induction relative to controls. Error bars represent standard deviations calculated on three independent experiments.

B. Expression of mutant SOD1(G93A) protein induces oxidative stress. Cells were incubated with 20 μM H2DCF-DA for 30 min at 37°C and assayed by FACS as described under "Materials and Methods." Resveratrol was added 18 h prior to H2DHCF-DA incubation. Left panel: FACS profiles. Black filled, SH-SY5Y/SOD1 cells; grey filled, SHSY5Y/SOD1(G93A) cells; empty, resveratrol-treated SH-SY5Y/SOD1(G93A) cells. Right panel: quantification of H2DHCF-DA fluorescence. Statistical analysis was carried out using a one-way ANOVA test followed by a post-hoc Tukey-Kramer multiple comparison test. ** = $p<0.01$. Error bars represent standard deviations calculated on three independent experiments.

C. Resveratrol treatment restores promoter activity of the -3344/+57 fragment and of the 146/+57 minimal promoter in SH-SY5Y(SOD1) cells. Results are expressed as fold induction relative to vehicle-treated cells. Error bars represent standard deviations calculated on three independent experiments.

Figure 4:

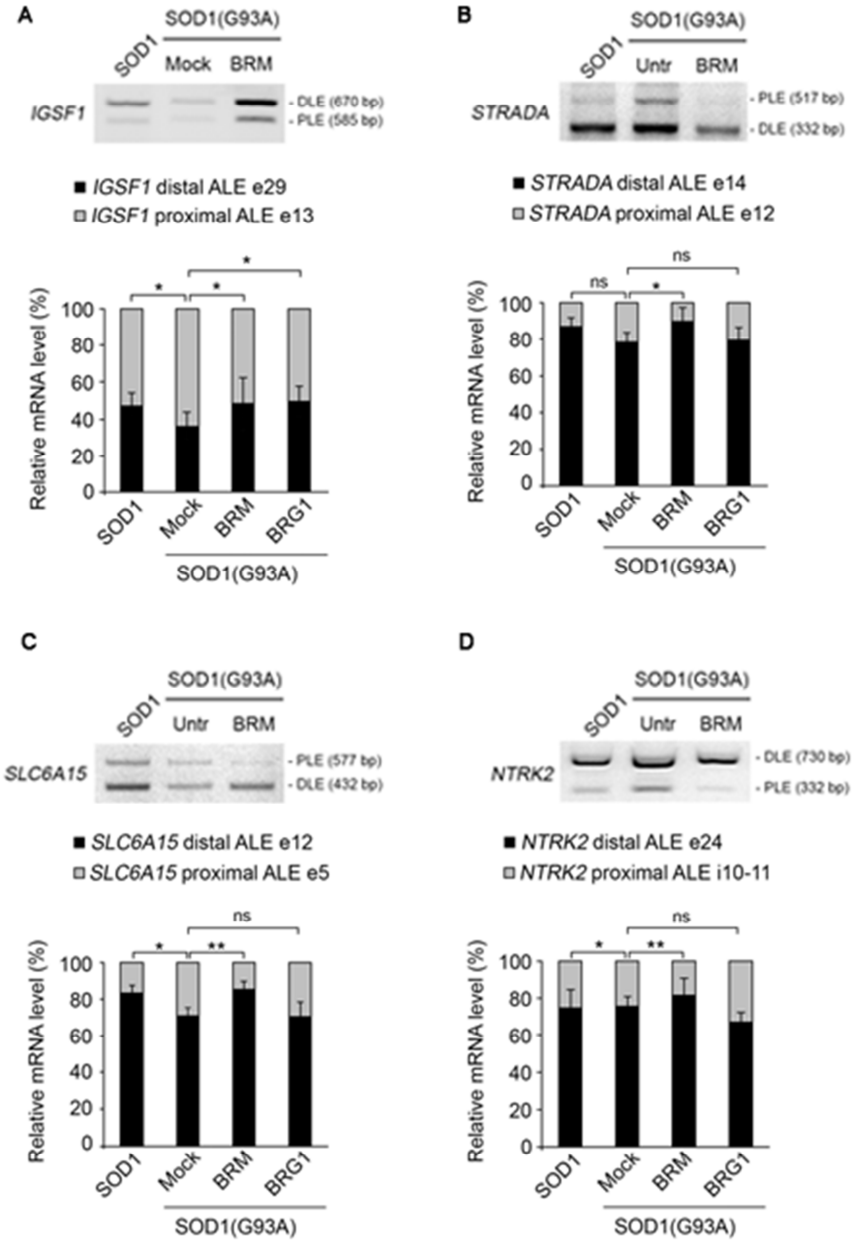

A
SOD1
SOD1(G93A)
Mock
BRM
IGSF1
- DLE (670 bp)
- PLE (585 bp)
IGSF1 distal ALE e29
IGSF1 proximal ALE e13
Relative mRNA level (%)
SOD1
Mock
BRM
BRG1
SOD1(G93A)
B
SOD1
SOD1(G93A)
Untr
BRM
STRADA
- PLE (517 bp)
- DLE (332 bp)
STRADA distal ALE e14
STRADA proximal ALE e12
ns
C
SLC6A15
- PLE (577 bp)
- DLE (432 bp)
SLC6A15 distal ALE e12
SLC6A15 proximal ALE e5
D
NTRK2
- DLE (730 bp)
- PLE (332 bp)
NTRK2 distal ALE e24
NTRK2 proximal ALE i10-11

Figure 4: Alternative splicing pattern of the ALEs of four genes affected by BRM expression.

A-D. *Upper panel*: The splicing pattern in the indicated cell lines was analyzed by RT-PCR and gel electrophoresis. PLE: proximal ALE, DLE: distal ALE.

Lower panel: quantification of the amount of proximal and distal ALE usage for each gene in SH-SY5Y/SOD1 and in SH-SY5Y/SOD1(G93A) cells, either utransfected, or stably transfected with a BRM- or BRG1-expressing plasmid. Data are representative of three independent experiments. Statistical analysis was carried out using a one-way Anova test followed by a post-hoc Tukey-Kramer multiple comparison test. Error bars represent standard deviation. * = $p<0,05$ and **= $p < 0,01$

Figura 5:

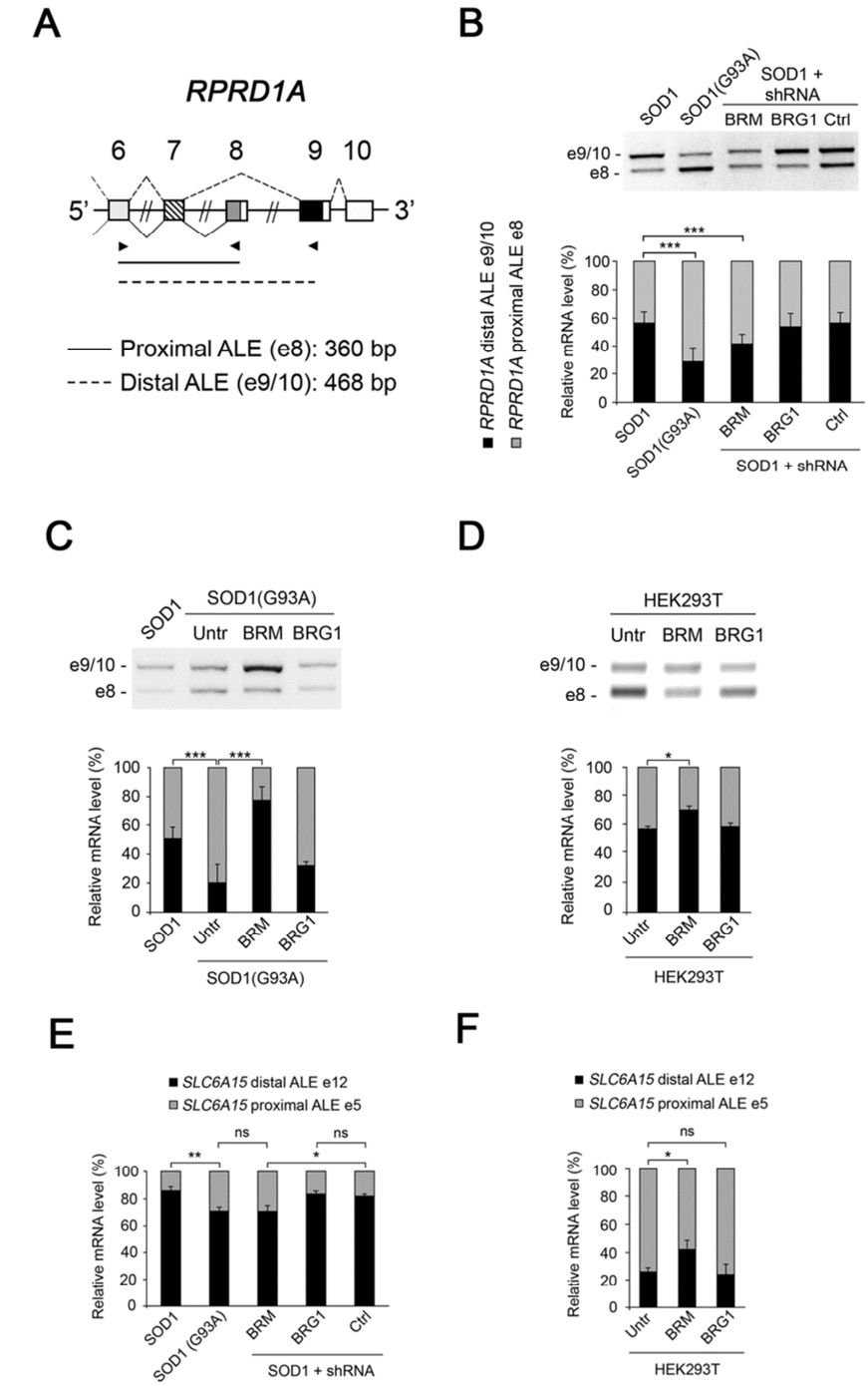

A
RPRD1A
6 7 8 9 10
5'
3'
Proximal ALE (e8): 360 bp
Distal ALE (e9/10): 468 bp
B
SOD1
SOD1(G93A)
SOD1 + shRNA
BRM BRG1 Ctrl
e9/10 -
e8 -
RPRD1A distal ALE e9/10
RPRD1A proximal ALE e8
Relative mRNA level (%)
***
***
SOD1
SOD1(G93A)
BRM
BRG1
Ctrl
SOD1 + shRNA
C
SOD1
SOD1(G93A)
Untr BRM BRG1
e9/10 -
e8 -
Relative mRNA level (%)
***
***
SOD1
Untr
BRM
BRG1
SOD1(G93A)
D
HEK293T
Untr BRM BRG1
e9/10 -
e8 -
Relative mRNA level (%)
*
Untr
BRM
BRG1
HEK293T
E
SLC6A15 distal ALE e12
SLC6A15 proximal ALE e5
Relative mRNA level (%)
**
ns
*
ns
SOD1
SOD1 (G93A)
BRM
BRG1
Ctrl
SOD1 + shRNA
F
SLC6A15 distal ALE e12
SLC6A15 proximal ALE e5
Relative mRNA level (%)
*
ns
Untr
BRM
BRG1
HEK293T

Figure 5: BRM inhibits inclusion of the proximal ALE.

A. Schematic representation of the relevant region of the RPRD1A gene indicating the positions of the PCR primers (arrowheads), the splicing pattern and the size of the expected PCR products.

B. Stable knock-down of BRM, but not of BRG1, favors e8 inclusion in SOD1 cells.

C.Stable overexpression of BRM, but not of BRG1, favors e10 inclusion in SOD1(G93A) cells.

D. Transient overexpression of BRM in HEK293T cells favors e10 inclusion.

B-D. *Upper panels*: The splicing pattern in the indicated cell lines was analyzed by RT-PCR and gel electrophoresis.

Lower panel: Quantification of proximal and distal ALE usage in the indicated cell lines. Data are representative of six independent experiments. Error bars represent standard deviation. * = $p<0.05$ and *** = $p< 0.001$

Figure 6:

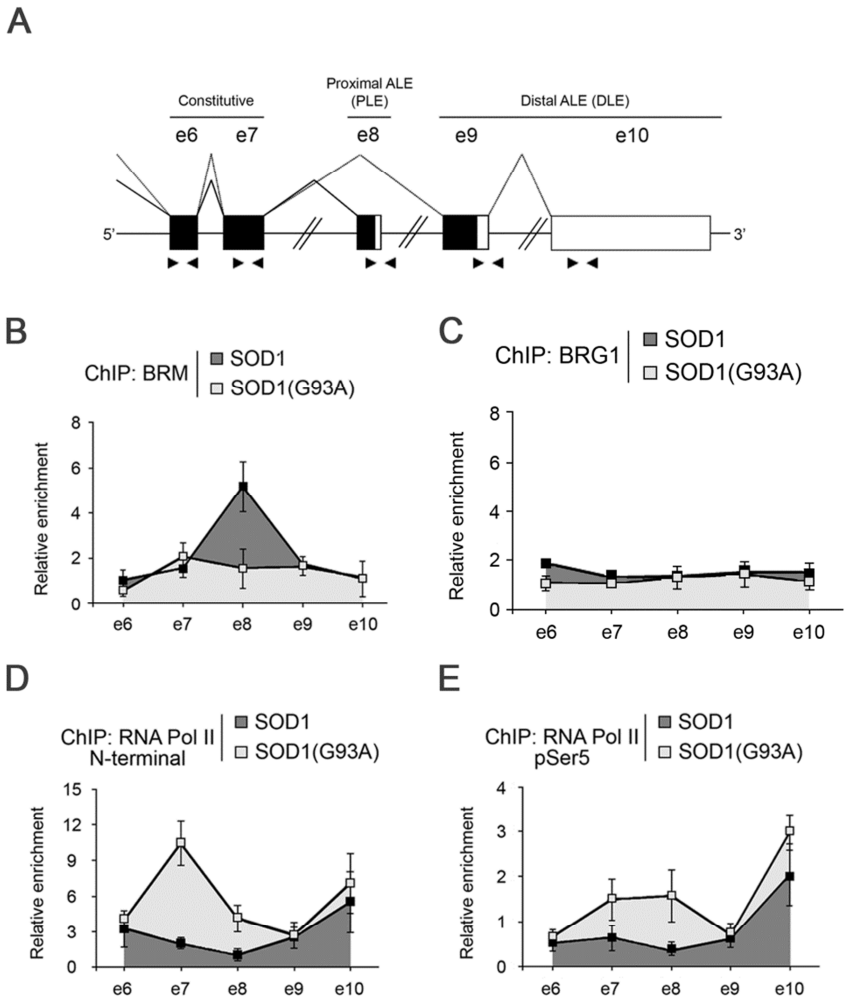

A
Constitutive
Proximal ALE (PLE)
Distal ALE (DLE)
e6
e7
e8
e9
e10
5'
3'
B
ChIP: BRM
SOD1
SOD1(G93A)
Relative enrichment
C
ChIP: BRG1
SOD1
SOD1(G93A)
Relative enrichment
D
ChIP: RNA Pol II N-terminal
SOD1
SOD1(G93A)
Relative enrichment
E
ChIP: RNA Pol II pSer5
SOD1
SOD1(G93A)
Relative enrichment

Figure 6: BRM localized on the excluded proximal ALE of RPRD1A gene.

A. Schematic representation of the relevant region of the RPRD1A gene indicating the splicing pattern and the positions of the qPCR primers used for ChIP analysis (arrowheads).

B-E. BRM, BRG1, and RNAPII on the RPRD1A gene. ChIP assays were performed with the indicated antibodies on bulk chromatin from SH-SY5Y/SOD1 and SHSY5Y/SOD1(G93A) cells. The graphs show the recruitment of the indicated proteins relative to the values obtained with the non-immune sera. Data are representative of three independent experiments. Error bars represent s.e.m.

Figure 7:

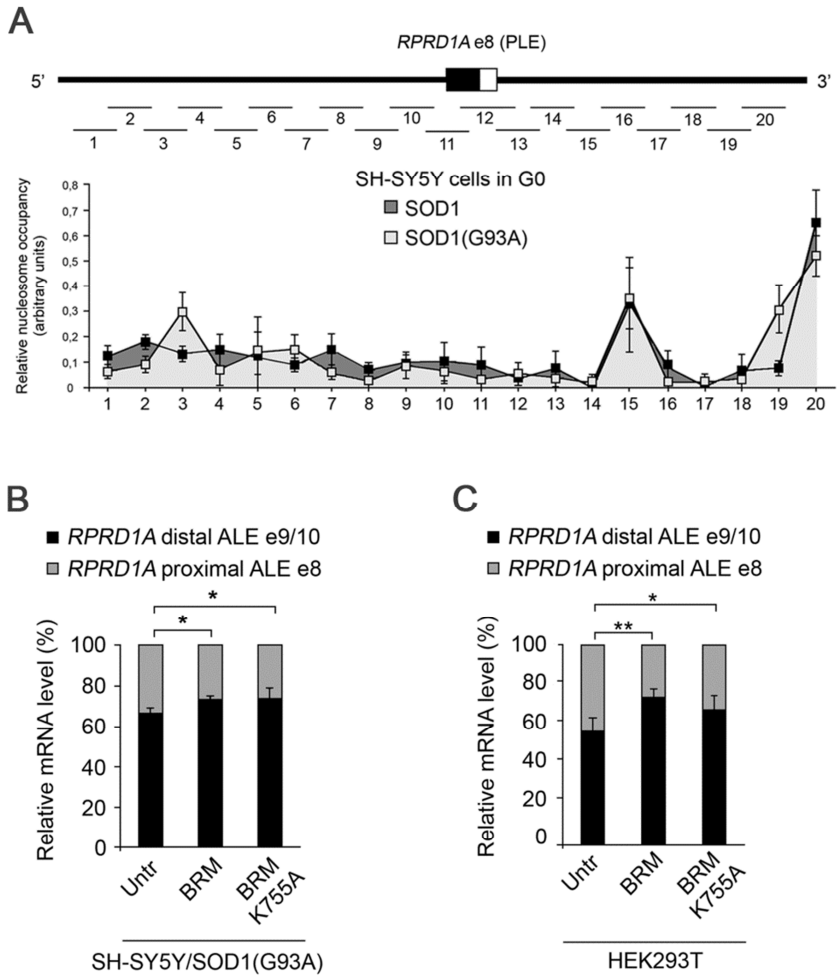

A
RPRD1A e8 (PLE)
5'
3'
SH-SY5Y cells in G0
SOD1
SOD1(G93A)
Relative nucleosome occupancy (arbitrary units)
B
RPRD1A distal ALE e9/10
RPRD1A proximal ALE e8
Relative mRNA level (%)
Untr
BRM
BRM K755A
SH-SY5Y/SOD1(G93A)
C
RPRD1A distal ALE e9/10
RPRD1A proximal ALE e8
Relative mRNA level (%)
Untr
BRM
BRM K755A
HEK293T

Figure 7: BRM ATPase activity is not reqired for ALE choice.

A. Nucleosome occupancy profile of the relevant region of the RPRD1A gene in G0synchronized cells. Lines indicate the positions of the qPCR primers. Nucleosome positioning was analyzed with MNase digestion for the 33,605,100-33,606,100 region of human chromosome 18 containing the RPRD1A gene in SH-SY5Y/SOD1 and in SHSY5Y/SOD1(G93A) cells. Data are representative of three independent biological replicates. Error bars represent s.e.m.

B-C BRM ATPase activity is not required for exon 8 skipping. The splicing pattern was analyzed by RT-PCR and quantified as percentage of proximal and distal ALEs. Data are representative of six independent experiments. Error bars represent standard deviations. * = $p<0.05$ and ** = $p<0.01$.

Figure 8.

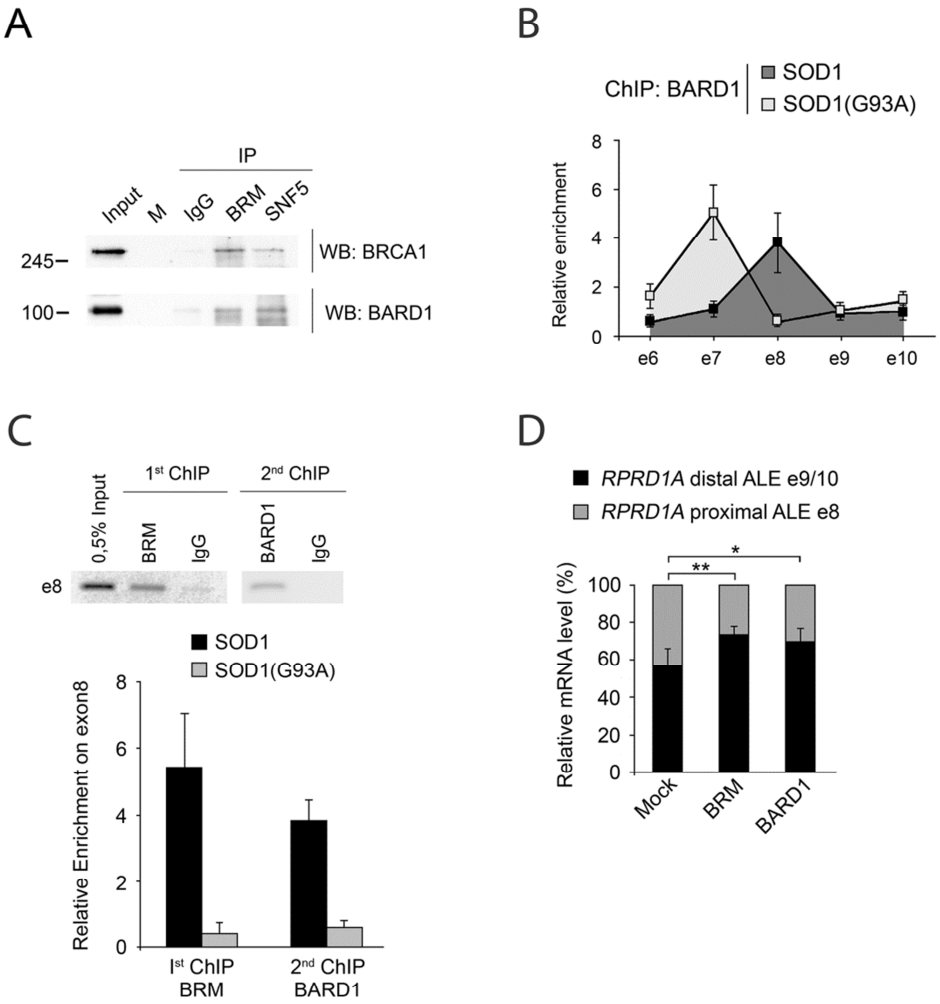

A
IP
Input
M
IgG
BRM
SNF5
245–
100–
WB: BRCA1
WB: BARD1
B
ChIP: BARD1
SOD1
SOD1(G93A)
Relative enrichment
8
6
4
2
0
e6
e7
e8
e9
e10
C
1st ChIP
2nd ChIP
0,5% Input
BRM
IgG
BARD1
IgG
e8
SOD1
SOD1(G93A)
Relative Enrichment on exon8
8
6
4
2
0
Ist ChIP
BRM
2nd ChIP
BARD1
D
RPRD1A distal ALE e9/10
RPRD1A proximal ALE e8
Relative mRNA level (%)
100
80
60
40
20
0
*
**
Mock
BRM
BARD1

Figure 8: BRM and BARD1 accumulated on the excluded proximal ALE.

A. Immunoprecipitations were carried out with the indicated antibodies (IP) from HEK293T nuclear extracts. Filters were sequentially probed with the indicated antibodies (WB).

B. BARD1 is enriched on the RPRD1A gene. ChIP assays were performed with the indicated antibodies on bulk chromatin from SH-SY5Y/SOD1 and SH-SY5Y/SOD1(G93A) cells. The graphs show the recruitment of the indicated proteins relative to the values obtained with the non-immune Ig. Data are representative of three independent experiments. Variability is expressed as s.e.m.

C. ChIP-reChIP was performed for BRM and BARD, and analyzed by PCR (upper panel) andqPCR (lower panel). Data are representative of three independent experiments. Error bars represent standard deviations.

D. Transient overexpression of BRM, and BARD1 in HEK293T cells promotes skipping of the proximal ALE. Error bars represent standard deviations. * = $p<0.05$, ** = $p<0.01$.

Figure 9:

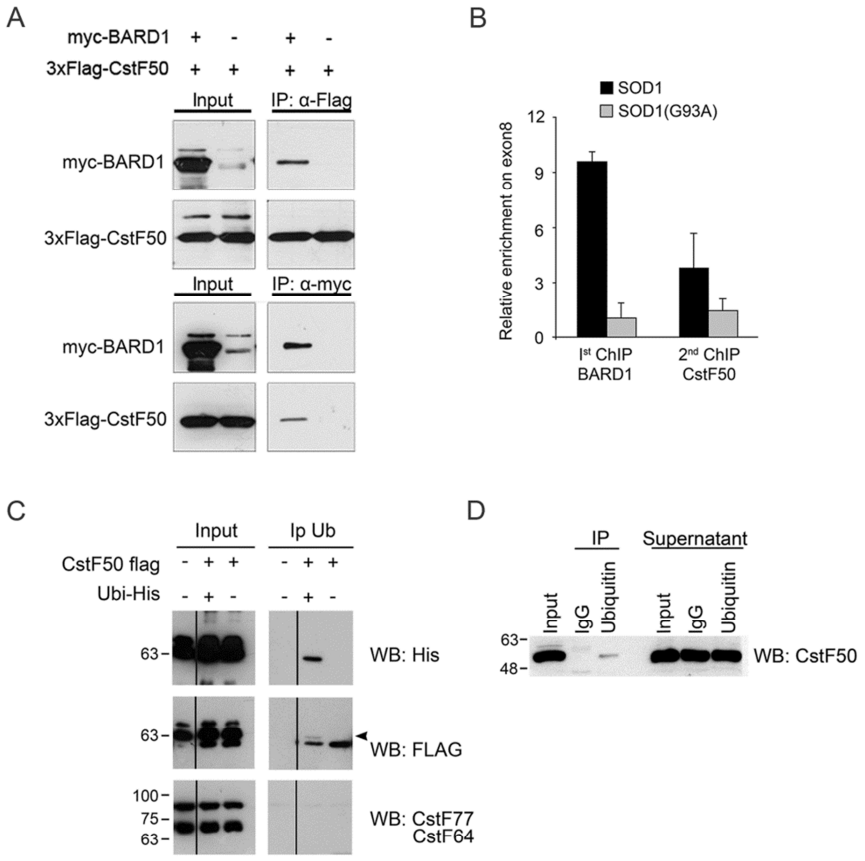


Figure 9: A fraction of CstF50 is ubiquitinated.

A. BARD1 and CstF50 physically interact. Cell lysates prepared from HEK293T cells and transfected with the indicated plasmids were immunoprecipitated with anti-FLAG (upper panel) or anti-Myc (lower panel) coniugated beads. Input (10% of the extract) and coimmunoprecipitated

proteins were analyzed by Western blots with the indicated antisera.

B. Both BARD1 and CstF50 are enriched on the RPRD1A gene. Quantitative PCR analysis of ChIP-reChIP assays for BARD1 and CstF50. Data are representative of three independent experiments. Error bars represent standard deviations.

C. CstF50 is ubiquitinated in HEK cells overexpressing his-tagged ubiquitin. HEK293T cells were cotransfected with plasmids expressing Flag-tagged CsF50 and His-ubiquitin. Total extracts were boiled in 1% of SDS lysis buffer, diluted tenfold, and immunoprecipitated with antiFLAG coniugated beads (IP). Input (10% of the extract) and immunoprecipitated proteins were analyzed by Western blotting with the indicated antisera. The arrowhead indicates the His tagged CstF50 protein. The shift is consistent with a monoubiquitination. (The samples were run on the same gel but the lanes were rearranged for clarity).

D. A fraction of CstF50 is ubiquitinated in unperturbed HEK cells. Whole cell extracts were prepared from HEK293T cells and immunoprecipitated with an anti-ubiquitin conjugated resin. Input and coimmunoprecipitated proteins were analyzed by Western blotting with an anti CstF50 antiserum.

Figure 10:

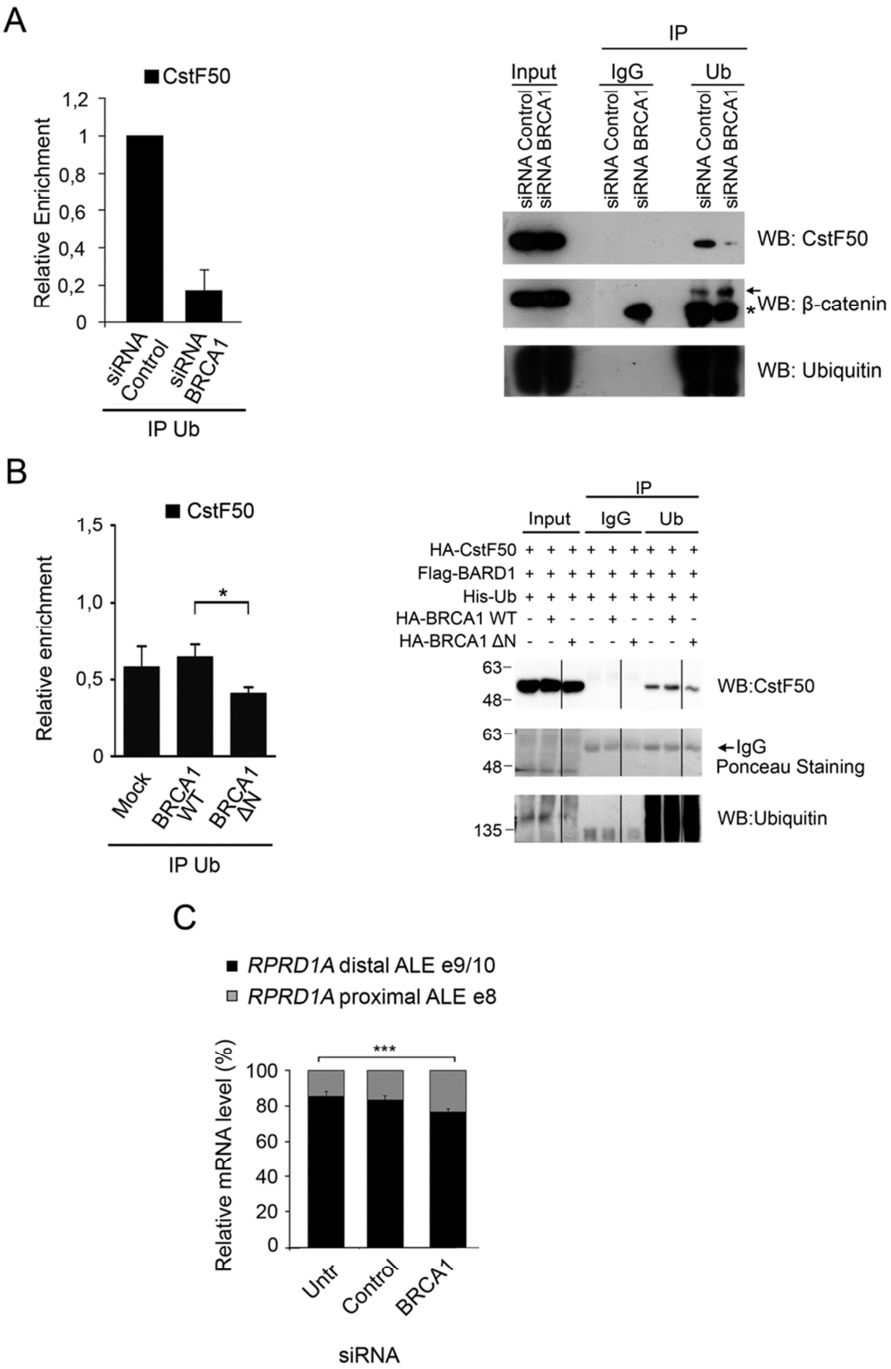

A
CstF50
Relative Enrichment
1,2
1
0,8
0,6
0,4
0,2
0
siRNA Control
siRNA BRCA1
IP Ub
IP
Input
IgG
Ub
siRNA Control
siRNA BRCA1
WB: CstF50
WB: β-catenin
WB: Ubiquitin
B
CstF50
Relative enrichment
1,5
1,0
0,5
0
*
Mock
BRCA1 WT
BRCA1 ΔN
IP Ub
IP
Input
IgG
Ub
HA-CstF50
Flag-BARD1
His-Ub
HA-BRCA1 WT
HA-BRCA1 ΔN
63
48
WB:CstF50
63
48
IgG
Ponceau Staining
135
WB:Ubiquitin
C
RPRD1A distal ALE e9/10
RPRD1A proximal ALE e8
Relative mRNA level (%)
***
100
80
60
40
20
0
Untr
Control
BRCA1
siRNA

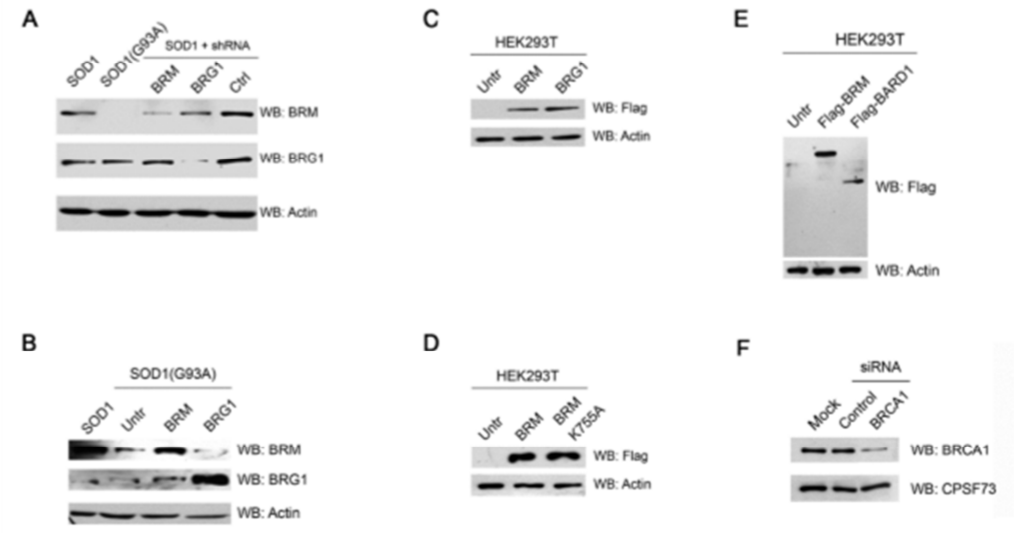
Figure 10: Uquitination of CstF50 is catalyzed by BRCA1.

A. Whole cell extracts from HEK293T cells transfected with control or BRCA1 siRNAs were immunoprecipitated with a control or an antiubiquitin antibody. Western blots were analysed with the indicated antibodies. β-catenin was used as control for BRCA1-independent ubiquitination and is highlighted by an arrow. The asterisk indicates non-specific bands. Upper panel: Quantification of CstF50 ubiquitination upon BRCA1 silencing. Quantification was performed on three independent immunoprecipitation experiments. Immunoprecipitated CstF50 protein was normalized on βcatenin. Error bars represent standard deviations. Lower panel: representative Western blot.

B. HEK293T cells were cotransfected with plasmids expressing CstF50, BARD1 and either wild type BRCA1 or the BRCA1-ΔN mutant. Ubiquitinated proteins were immunoprecipitated with an anti-ubiquitin antibody as described in "Materials and Methods" and analysed by Western blotting with the indicated antisera. Upper panel: Quantification was performed on three independent immunoprecipitation experiments. Immunoprecipitated CstF50 protein was normalized on IgG bands that are indicated in the Ponceau Staining. Error bars represent standard deviations. * = $p<0.05$, and ** = $p<0.01$. Lower panel: representative Western blot. The arrow indicates the bands corresponding to the IgGs.

C. Quantification of the amount of proximal and distal ALE usage in the indicated cell lines. Data are representative of three independent experiments. Error bars represent standard deviation. *** = p< 0.001

Figure 11:



Representative Western blots of:

A. SH-SY5Y/SOD1 cells stably silenced for BRM or BRG1.

B. SH-SY5Y/SOD1(G93A) stably expressing either BRM of BRG1

C. HEK293T cells transiently expressing BRM or BRG1

D. HEK293T cells transiently expressing BRM or BRM/K755A

E. HEK293T cells transiently expressing BRM and BARD1.

F. HEK293T cells untransfected (mock) or transfected with the indicated siRNAs were subjected to Western blot analysis with an anti-BRCA1 and anti-CPSF73 antibodies.

Table1:

Antibody	***Supplier***
rabbit polyclonal anti-BRM ChIP Grade,	Abcam
rabbit polyclonal anti-BRG1 ChIP grade,	Abcam
mouse monoclonal anti-ß-actin AC-40,	Sigma Aldrich
rabbit polyclonal anti-RNAPII N20	Santa Cruz Biotec.
mouse monoclonal anti-RNA Polymerase II H-14	Covance
rabbit polyclonal anti-Bard1	Bethyl Lab.
rabbit polyclonal anti-CstF50	Bethyl Lab.
rabbit polyclonal anti-CstF64 H-300	Santa Cruz Biotec.
rabbit polyclonal anti-CstF64	Bethyl Lab.
rabbit polyclonal anti-CstF77	Bethyl Lab.
rabbit polyclonal anti-SMARCB1/SNF5	Bethyl Lab.
rabbit polyclonal anti-BRCA1	Cell Signaling
rabbit polyclonal anti-Ubiquitin FL-76	Santa Cruz Biotec.
Mouse monoclonal anti-Ubiquitin	Enzo Life Sciences
monoclonal anti-Flag M2,	Sigma Aldrich
mouse monoclonal anti-HA Tag 6E2	Cell Signaling
mouse monoclonal anti-CstF50 (A5)	Santa Cruz Biotec.
Mouse monoclonal anti-β catenin (E-5)	Santa Cruz Biotec.
mouse monoclonal anti-multi Ub (FK29)	MBL
Mouse monoclonal anti-histidine Sigma	Sigma Aldrich
Rabbit polyclonal anti-myc	Cell Signaling

(71D10)	

Table2:

Oligos for the introduction of the K755A mutation in the SMARCA2 cDNA; underlined the codon mutation	
Mut_K755A External Fw (BstXI site)	GAAGCTCTCCAAAGCAGTGGC
Mut_K755A Internal Rev	CTGTATGGTCgcTCCAAGCCCC
Mut_K755A Internal Fw	GGGGCTTGGAgcGACCATACAG
Mut_K755A External Rev (Bsu361 site)	GCTCAAACTTCCCTGAGGC
Oligos for Quantitative Real-Time PCR (qPCR)	
qPCR_ SMARCA2_ Fw	AAACCTGTAGTGAGCGATT
qPCR_SMARCA2_ Rev	TCATCATCCGTCCCACTT
qPCR_ SMARCA4_ Fw	TGCCGTGATCAAGTACAAGG
qPCR_SMARCA4_ Rev	GAAGACCTCGCTGAGCTGAC
qPCR_GAPDH_ Fw	ACGGATTTGGTCGTATTGGG
qPCR_GAPDH_ Rev	TGATTTTGGAGGGATCTCG
Oligos for short hairpin-mediated RNA interference (shRNAi); underlined the shRNA region	
shRNAi_SMARCA2_Fw	gatccccGCAGGAAACCGAAGAGAAAttcaag agaTTTCTCT TCGGTTTCCTGCtttttggaaa
shRNAi_SMARCA2_Rev	agcttttccaaaaaGCAGGAAACCGAAGAGAA AtctcttgaaTTTC TCTTCGGTTTCCTGCggg
shRNAi_SMARCA4_Fw	gatccccGCTCAGAAGAAGAGGAAGAttcaag agaTCTTCCT CTTCTTCTGAGCtttttggaaa
shRNAi_SMARCA4_Rev	agcttttccaaaaaGCTCAGAAGAAGAGGAAG AtctcttgaaTCTT CCTCTTCTTCTGAGCggg
Oligos for Retro-Transcription PCR (RT-PCR) RNA analysis	

IGSF1_ex12_Fw (common)	GGTGCCCTTACTGAGTCCAA
IGSF1_ex13_Rev (PLE)	ATTGGCTCCCATACATCTGC
IGSF1_ex14_Rev (DLE)	TCACCCAGATTTTCAGGACA
NTRK2_ex9_Fw (common)	CTGCAAATCTGGCCGCACCTAAC
NTRK2_intr10-11_Rev (PLE)	AATTTAAGCAGCACCCAGAGTGCC
NTRK2_ex14_Rev (DLE)	GACGCAATCACCACCACAGCATAG
RPRD1A_ex6_Fw (common)	CTGGGATCTCCAAGTGAACC
RPRD1A_ex8_Rev (PLE)	GTAGATGTCTCCCGCAAAGG
RPRD1A_ex9_Rev (DLE)	CGTTAGAAGATACGCCCATGT
SLC6A15_ex4_Fw (common)	GGGATCAGTGTCCTTTGGTG
SLC6A15_ex5_Rev (PLE)	CCACTTTCCCCAATTTCCAT
SLC6A15_ex7_Rev (DLE)	AGCTTGAAAAGGCAATGACA
STRADA_ex11_Fw (common)	GCCATGTCCCCTTTAAGGAT
STRADA_ex12_Rev (PLE)	CCGAAATCCTGCCACTTATG
STRADA_ex13-14_Rev (DLE)	ACGTCGCTTGATCTGCTTG
Oligos for Chromatin Immunoprecipitation (ChIP) and ChIP-ReChIP qPCR analysis	
ChIP_RPRD1A_ex6_Fw	CTCTAGATCTCGTTAGAGCATTACAA
ChIP_RPRD1A_ex6_Rev	TCTTGGACTTCAACAGGTAAAGAA
ChIP_RPRD1A_ex7_Fw	TTTGGGAGTTGGGATGAGAG
ChIP_RPRD1A_ex7_Rev	CATCTATTTCTGCCGCCAAT

ChIP_RPRD1A_ex8_Fw	TGGACATGGGCGTATCTTCT
ChIP_RPRD1A_ex8_Rev	AAACAGTGACAAATGACCATCA
ChIP_RPRD1A_ex9_Fw	ACAGGAGTACAAGCGCAAGC
ChIP_RPRD1A_ex9_Rev	GTGACATTGGGCAATCGAG
ChIP_RPRD1A_ex10_Fw	GGCAATAGCACATGGGAAGA
ChIP_RPRD1A_ex10_Rev	ACTTTGCTTCCCTCCCAGTC
Oligos for Micrococcal Nuclease (MNase) assays qPCR analysis	
MN_RPRD1A_01_Fw	CTGTAAATTTGTAGGAATGTCATGGT
MN_RPRD1A_01_Rev	AGAGACTGCTGAATCAATAACTAAACA
MN_RPRD1A_02_Fw	TTATTGATTCAGCAGTCTCTAATTGTG
MN_RPRD1A_02_Rev	ATGGGTAGCAATGGCAATCT
MN_RPRD1A_03_Fw	GCTTAGATTAGATTGCCATTGCTAC
MN_RPRD1A_03_Rev	CTTGAGAGCTGGGAATTTGAG
MN_RPRD1A_04_Fw	TTAATTATTCCCACCTCAAATTCC
MN_RPRD1A_04_Rev	AGAGAAGGAGTAAATTGCAAAGGA
MN_RPRD1A_05_Fw	TTTGCAATTTACTCCTTCTCTGC
MN_RPRD1A_05_Rev	TATAGTGGCCACCAACTATCATCA
MN_RPRD1A_06_Fw	AATACCCTCAGATGATGATAGTTGGT
MN_RPRD1A_06_Rev	TTTTCCTTTCCATGTAGAAGTGG
MN_RPRD1A_07_Fw	GGTTGCCCACTTCTACATGG
MN_RPRD1A_07_Rev	GTGGCAAGTTAATTATCAAATTCATC
MN_RPRD1A_08_Fw	TTTGATAATTAACTTGCCACTATTTCA
MN_RPRD1A_08_Rev	ATGTTTAAAATTGAGGTTTCATCACA
MN_RPRD1A_09_Fw	CCTTGTGATGAAACCTCAATTTTA
MN_RPRD1A_09_Rev	TTACTTACCCTTAATTCTGTGGACAAT
MN_RPRD1A_10_Fw	TCCACAGAATTAAGGGTAAGTAATGTC

MN_RPRD1A_10_Rev	GAAGCTCAATAGAAGTTAAAGAAACCA
MN_RPRD1A_11_Fw	TGGTTTCTTTAACTTCTATTGAGCTTC
MN_RPRD1A_11_Rev	GGAAGAATATACTAATGGTGACGACAT
MN_RPRD1A_12_Fw	TCGTCACCATTAGTATATTCTTCCTG
MN_RPRD1A_12_Rev	AAAAACATGAGACAATACAGGCTTC
MN_RPRD1A_13_Fw	GCCTGTATTGTCTCATGTTTTTGA
MN_RPRD1A_13_Rev	TGACCATCATCATAAGTGAGAAAGTAT
MN_RPRD1A_14_Fw	TTTCTCACTTATGATGATGGTCATTT
MN_RPRD1A_14_Rev	TGGCATCTGATCTATTTTCTACAGTTA
MN_RPRD1A_15_Fw	ACTGTAGAAAATAGATCAGATGCCACT
MN_RPRD1A_15_Rev	CTATGAAATGTGGCTGGAAGG
MN_RPRD1A_16_Fw	CCACATTTCATAGTTGACAGTCATAG
MN_RPRD1A_16_Rev	GCCTAGAATAGTACCTGACACATGAAT

2.6 References:

Bai, Y., Auperin, T.C., Chou, C.Y., Chang, G.G., Manley, J.L. and Tong, L. (2007) Crystal structure of murine CstF-77: dimeric association and implications for polyadenylation of mRNA precursors. Mol Cell, 25, 863-875.

Batsche, E., Yaniv, M. and Muchardt, C. (2006) The human SWI/SNF subunit BRM is a regulator of alternative splicing. Nat Struct Mol Biol, 13, 22-29.

Bochar, D.A., Wang, L., Beniya, H., Kinev, A., Xue, Y., Lane, W.S., Wang, W., Kashanchi, F. and Shiekhattar, R. (2000) BRCA1 is associated with a human SWI/SNF-related complex: linking chromatin remodeling to breast cancer. Cell, 102, 257-265.

Boehm, A.K., Saunders, A., Werner, J. and Lis, J.T. (2003) Transcription factor and polymerase recruitment, modification, and movement on dhsp70 in vivo in the minutes following heat shock. Mol Cell Biol, 23, 7628-7637.

Carrillo Oesterreich, F., Preibisch, S. and Neugebauer, K.M. (2010) Global analysis of nascent RNA reveals transcriptional pausing in terminal exons. Mol Cell, 40, 571-581.

Chiba, N. and Parvin, J.D. (2002) The BRCA1 and BARD1 association with the RNA polymerase II holoenzyme. Cancer Res, 62, 4222-4228.

Ciriolo, M.R., De Martino, A., Lafavia, E., Rossi, L., Carri, M.T. and Rotilio, G. (2000) Cu,Zn-superoxide dismutase-dependent apoptosis induced by nitric oxide in neuronal cells. J Biol Chem, 275, 5065-5072.

Cocheme, H.M. and Murphy, M.P. (2008) Complex I is the major site of mitochondrial superoxide production by paraquat. J Biol Chem, 283, 1786-1798.

Danckwardt, S., Hentze, M.W. and Kulozik, A.E. (2008) 3' end mRNA processing: molecular mechanisms and implications for health and disease. Embo J, 27, 482-498.

Dantonel, J.C., Murthy, K.G., Manley, J.L. and Tora, L. (1997) Transcription factor TFIID recruits factor CPSF for formation of 3' end of mRNA. Nature, 389, 399-402.

Di Giammartino, D.C., Nishida, K. and Manley, J.L. (2011) Mechanisms and consequences of alternative polyadenylation. Mol Cell, 43, 853-866.

Edwards, R.A., Lee, M.S., Tsutakawa, S.E., Williams, R.S., Nazeer, I., Kleiman, F.E., Tainer, J.A. and Glover, J.N. (2008) The BARD1 C-terminal domain structure and interactions with polyadenylation factor CstF-50. Biochemistry, 47, 11446-11456.

Elkon, R., Ugalde, A.P. and Agami, R. (2013) Alternative cleavage and polyadenylation: extent,

regulation and function. Nature reviews. Genetics, 14, 496-506.

Ji, Z., Lee, J.Y., Pan, Z., Jiang, B. and Tian, B. (2009) Progressive lengthening of 3' untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development. Proc Natl Acad Sci U S A, 106, 7028-7033.

Katahira, J., Okuzaki, D., Inoue, H., Yoneda, Y., Maehara, K. and Ohkawa, Y. (2013) Human TREX component Thoc5 affects alternative polyadenylation site choice by recruiting mammalian cleavage factor I. Nucleic Acids Res.

Kleiman, F.E. and Manley, J.L. (1999) Functional interaction of BRCA1-associated BARD1 with polyadenylation factor CstF-50. Science, 285, 1576-1579.

Kleiman, F.E. and Manley, J.L. (2001) The BARD1-CstF-50 interaction links mRNA 3' end formation to DNA damage and tumor suppression. Cell, 104, 743-753.

Lenzken, S.C., Romeo, V., Zolezzi, F., Cordero, F., Lamorte, G., Bonanno, D., Biancolini, D., Cozzolino, M., Pesaresi, M.G., Maracchioni, A. et al. (2011) Mutant SOD1 and mitochondrial

damage alter expression and splicing of genes controlling neuritogenesis in models of neurodegeneration. Hum Mutat, 32, 168-182.

Lin, M.T. and Beal, M.F. (2006) Mitochondrial dysfunction and oxidative stress in neurodegenerative diseases. Nature, 443, 787-795.

Maglott, D., Ostell, J., Pruitt, K.D. and Tatusova, T. (2011) Entrez Gene: gene-centered information at NCBI. Nucleic Acids Res, 39, D52-57.

Mallery, D.L., Vandenberg, C.J. and Hiom, K. (2002) Activation of the E3 ligase function of the BRCA1/BARD1 complex by polyubiquitin chains. Embo J, 21, 6755-6762.

McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., Patterson, S.D., Wickens, M. and Bentley, D.L. (1997) The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. Nature, 385, 357-361.

Millevoi, S. and Vagner, S. (2010) Molecular mechanisms of eukaryotic pre-mRNA 3' end processing regulation. Nucleic Acids Res, 38, 2757-2774.

Mirkin, N., Fonseca, D., Mohammed, S., Cevher, M.A., Manley, J.L. and Kleiman, F.E. (2008) The 3' processing factor CstF functions in the DNA repair response. Nucleic Acids Res, 36, 1792-1804.

Morris, D.P., Michelotti, G.A. and Schwinn, D.A. (2005) Evidence that phosphorylation of the RNA polymerase II carboxyl-terminal repeats is similar in yeast and humans. J Biol Chem, 280, 31368-31377.

Nagaike, T., Logan, C., Hotta, I., Rozenblatt-Rosen, O., Meyerson, M. and Manley, J.L. (2011) Transcriptional activators enhance polyadenylation of mRNA precursors. Mol Cell, 41, 409-418.

Pickrell, J.K., Marioni, J.C., Pai, A.A., Degner, J.F., Engelhardt, B.E., Nkadori, E., Veyrieras, J.B., Stephens, M., Gilad, Y. and Pritchard, J.K. (2010) Understanding mechanisms underlying human gene expression variation with RNA sequencing. Nature, 464, 768-772.

Pinto, P.A., Henriques, T., Freitas, M.O., Martins, T., Domingues, R.G., Wyrzykowska, P.S., Coelho, P.A., Carmo, A.M., Sunkel, C.E., Proudfoot, N.J. et al. (2011) RNA polymerase II kinetics in polo

polyadenylation signal selection. EMBO J, 30, 2431-2444.

Proudfoot, N.J. (2011) Ending the message: poly(A) signals then and now. Genes Dev, 25, 1770-1782.

Richmond, E. and Peterson, C.L. (1996) Functional analysis of the DNA-stimulated ATPase domain of yeast SWI2/SNF2. Nucleic Acids Res, 24, 3685-3692.

Rizzardini, M., Mangolini, A., Lupi, M., Ubezio, P., Bendotti, C. and Cantoni, L. (2005) Low levels of ALS-linked Cu/Zn superoxide dismutase increase the production of reactive oxygen species and cause mitochondrial damage and death in motor neuron-like cells. Journal of the neurological sciences, 232, 95-103.

Rozenblatt-Rosen, O., Nagaike, T., Francis, J.M., Kaneko, S., Glatt, K.A., Hughes, C.M., LaFramboise, T., Manley, J.L. and Meyerson, M. (2009) The tumor suppressor Cdc73 functionally associates with CPSF and CstF 3' mRNA processing factors. Proc Natl Acad Sci U S A, 106, 755-760.

Sandberg, R., Neilson, J.R., Sarma, A., Sharp, P.A. and Burge, C.B. (2008) Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. Science, 320, 1643-1647.

Sanges, R., Cordero, F. and Calogero, R.A. (2007) oneChannelGUI: a graphical interface to Bioconductor tools, designed for life scientists who are not familiar with R language. Bioinformatics (Oxford, England), 23, 3406-3408.

Schwartz, S., Meshorer, E. and Ast, G. (2009) Chromatin organization marks exon-intron structure. Nat Struct Mol Biol, 16, 990- 995.

Shi, Y. (2012) Alternative polyadenylation: new insights from global analyses. RNA (New York, N.Y.), 18, 2105-2117.

Shi, Y., Di Giammartino, D.C., Taylor, D., Sarkeshik, A., Rice, W.J., Yates, J.R., 3rd, Frank, J. and Manley, J.L. (2009) Molecular architecture of the human pre-mRNA 3' processing complex. Mol Cell, 33, 365-376.

Takagaki, Y. and Manley, J.L. (1998) Levels of polyadenylation factor CstF-64 control IgM heavy chain mRNA accumulation and other events

associated with B cell differentiation. Mol Cell, 2, 761-771.

Takagaki, Y., Seipelt, R.L., Peterson, M.L. and Manley, J.L. (1996) The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B cell differentiation. Cell, 87, 941-952.

Tilgner, H., Nikolaou, C., Althammer, S., Sammeth, M., Beato, M., Valcarcel, J. and Guigo, R. (2009) Nucleosome positioning as a determinant of exon recognition. Nat Struct Mol Biol, 16, 996-1001.

Waldholm, J., Wang, Z., Brodin, D., Tyagi, A., Yu, S., Theopold, U., Farrants, A.K. and Visa, N. (2011) SWI/SNF regulates the alternative processing of a specific subset of premRNAs in Drosophila melanogaster. BMC Mol Biol, 12, 46.

Wang, E.T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P. and Burge, C.B. (2008) Alternative isoform regulation in human tissue transcriptomes. Nature, 456, 470-476.

Wu, L.C., Wang, Z.W., Tsan, J.T., Spillman, M.A., Phung, A., Xu, X.L., Yang, M.C., Hwang, L.Y., Bowcock, A.M. and Baer, R. (1996) Identification

of a RING protein that can interact in vivo with the BRCA1 gene product. Nat Genet, 14, 430-440.

Zraly, C.B. and Dingwall, A.K. (2012) The chromatin remodeling and mRNA splicing functions of the Brahma (SWI/SNF) complex are mediated by the SNR1/SNF5 regulatory subunit. Nucleic Acids Res, 40, 5975-5987.

Chapter 3.

Characterization of SMARCA2 isoforms.

Aurora Rigamonti[1], Giuseppe Mauro[1], Gabriele A. Fontana[1#] and Silvia M.L. Barabino[1]

1 Department of Biotechnology and Biosciences, University of Milan-Bicocca Piazza della Scienza 2, 20126, Milan, Italy

#Present address: Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058, Basel

3.1 Introduction

Recent evidence suggests that 30–50% of the human and about one half of the mouse genes have multiple alternative promoters (APs) that can span up to thousands of kilobases (Beak et al.,2007). Use of APs enables diversification of transcriptional regulation within a single locus and there by plays a significant role in the control of gene expression in various cell lineages, tissue types and developmental stages. Alternative promoters can generate mRNA isoforms that encode distinct proteins, sometimes having opposing biological activities. In most cases the activities of the protein isoforms differ because the promoters are separated by one or two exons that encode an important functional domain. For example, LEF1, which encodes lymphoid enhancer factor proteins that mediate the transcriptional regulation of Wnt/beta-catenin target genes, is transcribed by two alternative promoters P1 and P2. The P1 promoter produces the full-length LEF1 protein (LEF1), which recruits b-catenin to Wnt target genes, whereas the intronic promoter P2 derives the shorter LEF1 (ΔNLEF1), which cannot interact with beta-catenin and instead suppresses Wnt regulation of target genes (Arce et al., 2006).

Here, we report the identification of an alternative promoter, present in the human SMARCA2 gene which controls the transcription of a class of isoforms of the chromatin remodeling factor Brahma (BRM, SNF2α). BRM is one of the two mutually ATPase subunits present in the SWI/SNF-BAF complex, a multi-enzymatic complex which epigenetically controls the transcription of genes involved in differentiation and development (Muchardt et a., 1999). The AP of *SMARCA2* drives the expression of a short isoforms of BRM, in contrast to the full length BRM isoforms (Brm FL). The regulatory region which controls the expression of Brm short isoforms, as well as the one which controls the expression of the Brm FL isoforms (Rigamonti et al., in Submission), respond to mitochondrial stress, an observation which is consistent with the data collected from mRNA analysis. BRM FL and BRM short transcripts are evolutionarlyly conserved, their product proteins differ for functional domains: BRM FL shows an ATPase domain that it's missed in the short isoforms while both isoforms maintain the Bromodomain at their N-terminal region. For this reason BRM FL and BRM short show different protein interactors: both proteins bind histones, but only the BRM FL isoforms it is able to enter in the SWI/SNF complex. This observation leads us to

hypothesize a role of "dominant negative" for these shorter proteins.

3.2 Results

bioinformatics analysis of the transcripts encoded by the human SMARCA2 gene.

The human *SMARCA2* gene is located on the forward strand of chromosome 9 (position 2,015,342-2,193,624).

Ensemble Genome Browser database shows different SMARCA2 transcripts obtained from analysis using different experimental approaches (Fig. 1A). We found that human *SMARCA2* gene can potentially encodes at least two classes of transcripts. The first class is constituted by the transcripts encoding the full length BRM protein (*SMARCA2*-003,-004,-008,-017,-019). The protein prediction of these transcripts (Fig.1B) shows that they encode proteins of nearly 180 KDa. Starting from the N-terminal region these proteins contain a proline and glutamine rich domain (P/Q), an helicase-SANT–associated (*HSA*) *domain*, an ATPase/Helicase Domain, a tyrosine and arginine rich domain (K/R), an

E7 domain and a Bromodomain (Bro). They also contain four nuclear localization signals (NLS). Among these NLS, NLS2, 3 and 4 (Loe-Mie et al., 2010) are experimentally validated. On the other hand, NLS1 is a result of bioinformatics prediction.

The second class is constituted by several shorter transcripts (Fig.1A). These transcripts have different transcription start sites but the protein prediction suggests that these isoforms could encodes proteins of 30-40 KDa containing only a C-terminal Bromodomain (Fig.1B). Most of these proteins contain only NLS3 and 4.

Taken together these observations suggest that human SMARCA2 gene could encode for two class of proteins: the BRM FL protein, which present all the domains already described in literature, and BRM Short protein, which is characterized by only the Bromodomain.

The expression of BRM protein isoforms is impaired by mitochondrial stress.

In the same microarray dataset that was the starting point of my first project we also detected short *SMARCA2* transcript isoforms. The microarray showed that the ratio of long and short isoforms changed upon oxidative stress. To verify this observation we

monitored the expression of these two type of transcripts using two paradigms of mitochondrial stress: an acute stress model, SH-SY5Y human neuroblastoma cells treated by Paraquat (PQ, N,N'-dimethyl-4,4'-bipyridinium dichloride, a chemical that inhibits mitochondrial complex 1)(Maracchioni et al., 2007) and a chronic stress model constituted by the same cell line overexpressing the SOD1 protein carrying the G93A mutation, one of the genetic causes of Amyotrophic Lateral Sclerosis (ALS) (Shi et al., 2010).

In order to monitor the expression of the two classes of transcripts, we set up a retro-transcription PCR (RT-PCR) approach. Using a three primers PCR we were able to discriminate the relative expression of both groups of transcripts. From these experiments we noticed a change in the relative ratio of long and short transcript upon mitochondrial stress. In particular PQ treatment and SOD1(G93A) expression induce a relative downregulation of the expression of the long transcripts encoding BRM FL and a concomitant relative upregulation of the expression of transcripts encoding BRM short. This is true for the acute stress (Fig.2A) as well for the chronic model (Fig.2B).

We next tried to examined the expressions of the different BRM isoforms at the protein level, focusing our attention on the SOD1 (G93A) model (Fig2C). We used a commercial antibody (anti-BRM/SNF2α KR-17) that recognizes a common epitope located in the C-terminal bromodomain. Unfortunately this antibody recognizes several other proteins that are not predicted by bioinformatics analysis. This problem was also reported by Yang and colleagues, (2011). In this article they used two different antibodies that recognized the C-terminal domain of BRM and in both cases they observed multiple proteins. They tried to explain these bands as 1) degraded or proteolyzed products of the 180-kD BRM proteins, 2) proteins that are unrelated to SMARCA2 but immunoreactive to the BRM C-terminal antibodies, i.e. nonspecific bands on the blot, 3) BRM short proteins as they are within the estimated molecular weights, or 4) BRM short proteins translated from some in-frame ATGs. The first possibility is less likely true because they demonstrated the high quality of the protein preparation without obvious degradation. The second or third possibility needs to be determined by additional studies. If short isoforms exist, each mRNA variant needs to be causally linked to one or several specific bands on the western blot. Cloning the full-length cDNA of various SMARCA2 short variants

that encode different ORFs and raising the corresponding antibodies are prerequisites for this line of work, which are tedious and thus beyond this communication. The fourth possibility is rarely discussed.

Taken together these results indicate that mitochondrial stress affects the expression of both the two classes of SMARCA2 transcripts. In particular SOD1(G93A) overexpression induces a strong downregulation of BRM FL mRNA expression and a concomitant relative upregulation of BRM short mRNAs. Unfortunately we could not determine the expression of the protein.

Long and short isoforms localized in the nucleus.

In order to functionally characterize BRM short isoforms we cloned one of the short transcripts, we focused on the *SMARCA-001* isoforms because is one of the validated short isoforms of the UCSC Genome Browser that seems to be the most representative for its exon compositions. We cloned this isoforms with a FLAG-tag to specifically detect it by immunofluorescence. We transiently transfected HeLa cells with this construct or with a plasmid encoding BRM FL FLAG-tag fusion protein and we performed immunofluorescence (IF).

We found that both proteins (short and full length) localize in the nucleus (Fig.3). This observation is according with the presence of NLS (specifically NLS3 and 4) in both (Fig.1B). We also monitored the localization of endogenous BRM and of a FLAG-tagged BRG1, the other alternative ATPase subunit of the SWI/SNF complex. We analized endogenous BRM FL using an antibody that recognized the N-terminal domain present only in the long isoforms.

In conclusion we confirmed the nuclear localization of BRM short similarly to BRM FL and BRG1 subunit of the SWI/SNF complex.

Isolation of the alternative promoter that controls the expression of BRM short variants.

Long and short BRM isoforms have a different N-terminal region. The presence of this different region suggests us that long and short isoforms are not product of alternative splicing. Their production could be better explained with the presence of alternative promoter usage.

To test this hypothesis we examined the genomic region located upstream the transcription start site of the *SMARCA-001* short transcript. The first two exons

of this transcript are only present in the short isoforms and are located in a 34 Kb intron of the large isoform (Fig.4A). Using UCSC Genome Browser we found a peak of evolutionarly conservation in the region upstream the transcription start site of *SMARCA-001* transcript. Starting from this point we analysed the -1230/+198 region (the number are relative to the *SMARCA-001* transcription start site, indicated as +1). This region contains not only peaks of evolutionarly conservation but also five putative CpG island predicted with Emboss CpG Plot (Fig.4B). The UCSC Genome Browser, based on submitted ChIP-Seq data obtained from different cell lines, identified in the -912/+198 region peaks of enrichment of trimethylation of histone 3 lysine 4 (H3K4me3), an histone marker usually associated to regulatory regions located in close proximity to the transcription start sites (Kolasinska-Zwierz et al., 2009). In the same region this browser also identified a DNAse hypersensible region, another feature associated with transcription start sites. We used ChIp Seq data of UCSC Genome Browser to map Transcription Factors (TF) binding site found in this genomic region, the most relevant are shown in Fig. 4C.

In order to experimentally verify if this region contains a putative promoter, we cloned the entire region (from -1230 to 198) upstream the luciferase gene in a promoter-less vector. Following 5'deletions we tested the luciferase (luc) activity of four different constructs (Fig.4C). Initially, we monitored the activities of the different constructs in different cellular models. In addition to SH-SY5Y cells, we used HEK 293 and NIH 3T3 cells as a model of non-neuronal cells. In order to directly compare the relative luciferase units (RLUs) obtained from these different cell lines, we normalized over the background (calculated using the empty luciferase vector) and an unrelative promoter (the ubiquitously active SV40 promoter controlling the luciferase transcription inside the same vector backbone). We found that the constructs containing the -1230/+198 region and the -850/+198 region had a very low activity comparable to the background in all the tested cell lines (Fig.5A). However the regions -513/+198 and -166/+198 increased their luciferase activity. These observations suggest that the region between -1230/-513 is not part of the promoter region or alternatively could contains same elements that act as silencer elements. The results obtained with the other two constructs indicate that the basic regulatory elements that control the transcription of the human

SMARCA2 short isoforms may reside in -513/+198 region. We did not found tissue specific elements that could control *SMARCA2* short isoforms transcription.

In another set of experiments we analysed the response of the four constructs in our model of chronic mitochondrial stress (Fig.5B). From these experiments we noticed that only the luciferase activity of the -850/+198 region changed in response to mitochondrial stress. The activities of -513/+198 and -166/+198 fragments were slightly inhibited in SOD1 (G93A) cells. These observations lead to us to think that also in the -850/+198 regions could be present promoter regulatory elements, in particular a sort of "mitochondrial stress responsive" elements that increase the expression of the BRM short variants.

Thanks to the ENCODE project it is now possible to identify TFs binding a particular genomic region. The comparison between TFs of BRM FL and BRM short promoter (Fig.5C) show only one common TF: C/EBP alfa (CCAAT/enhancer-binding protein alpha). However to better characterize BRM short promoter we were looking for a TF able to increase specifically the activity of this promoter. We focused our attention on the Cyclic AMP-responsive element-binding protein (CREB) binding site. CREB protein is one of the major

transcriptional factors that regulate the expression of genes necessary for the development and function of the nervous system (Lonze and Ginty, 2002). The transcriptional activity of CREB is induced through serine phosphorylation in its conserved kinase inducible domain by the cAMP-dependent protein kinase (PKA) (Sands and Palmer, 2008). PKA is activated by cyclic AMP that is product by the cell with the adenylate cyclase enzyme. For this reason we activated CREB protein with forskolin an activator of adenylate cyclase (Monaghan et al., 2008). To confirm the activation of CREB protein we controlled the upregulation of the expression of the aspartate/glutamate carrier gene (SLC25A12 gene) that is induced via CREB by forskolin (data not shown) (Menga et al., 2015). We performed a luciferase assay using the -1230/+198, -850/+198, -166/+198 regions and the –3344 to +57 region of the BRM FL promoter (Rigamonti A.et al, submetted). Only the minimal region of BRM short promoter show an increase of their activity. This result allows to speculate that upstream this region could be a possible regulatory feature that block and control CREB activity on BRM short promoter.

Taken together these results indicate that the short isoforms of Brm are generated by alternative promoter

usage. Moreover our region of interest contains elements that are able to drive the ectopic expression of the luciferase reporter gene. Our study suggests the presence of putative silencer elements and "mitochondrial stress responsive" elements that could control BRM short promoter activity.

Evolutionarly conservation of the short BRM variants.

SWI/SNF complex contain two closely related alternative ATPases: BRM or Brahma-related gene 1 (BRG1). Although BRM and BRG1 share a high degree of amino acid sequence identity, they are not equally important for development. An ES cell-specific assembly of BAF, called esBAF, is characterized by the presence of BRG1 (Ho et al., 2009) and not BRM. Then, when neural precursors start to appear, the chromatin landscape starts to require the neural progenitor-specific BAF (npBAF) complex (Lessard et al., 2007) that includes either Brg1 or Brm as the core ATPase.

Looking for an *in vivo* model where we could control the expression of BRM FL and BRM short isoforms, we focused our attention on Zebrafish model. Approximately 70% of human genes have at least one zebrafish orthologue (Howe et al., 2013). We verified

the presence of BRM short trascripts in *Danio rerio* genome using bioinformatics analysis (Fig.6A). The zebrafish *SMARCA2* gene is located on Chromosome 5, in the forward strand between positon 46,757,516-46,902,718. It encodes one long isoforms and two short isoforms that have in common with the longest one the first four or five exons. Ensemble Genome Browser database doesn't show 3' short isoforms of this gene. Zebrafish are members of the teleost infraclass, a monophyletic group that is thought to have arisen approximately 340 million years ago from a common ancestor. Compared to other vertebrate species, this ancestor underwent an additional round of whole-genome dupli-cation (WGD) called the teleost-specific genome duplication (TSD)(Meyer at al.,1999).

We tried to performe an alignment using the Basic Local Alignment Searching Toll (BLAST) with the predicted protein of the *SMARCA-001* transcript and the zebrafish database of non-redundant protein sequences (nr). This analysis gave us five putative locus and we concentrated our attention on the hits that were similar to the short protein in their first 33 aminoacids (the only region that differ this short isoform to the FL isoform). In this way we found two hits with a common uncharacterized locus (Fig.6B). The cDNA sequence of

this uncharacterized locus shows with the hBRM short sequence an homology of 66% (by CLUSTAL W Multiple Sequence Alignment analysis). Moreover a nucleotide BLAST obtained using a RefSeq genomic of *Danio rerio* shows that this zebrafish cDNA is contained in the sequence with accession number NC007116.5. This region encoded from a gene of 6 exons that is located in the same Chromosome of SMARCA2 gene (Chromosome 5) but in the reverse strand.

Starting from this cDNA, we designed primers to performed quantitative-PCR to check the effective presence of this short BRM transcripts (Fig.6C). From these experiments we noticed that in zebrafish BRM short transcript displaye the same behavior of BRG1 transcript: BRM short slightly decreases until 24 hour post fertilization (hpf) and then increases its expression. Instead BRM FL transcript seems to decrease their expression until maternal to zygotic transition (MZT) and then increases its level. It's important to note that the BRM short transcripts level is very low compared to BRM and BRG1 level. In fact BRM short amplification appears 4 cycles after BRM FL amplification, this means approximately 16 times less abundance than BRM FL (the data are plot with their –

ΔCt that inversely correlates with the initial template amount).

Taken together these results show that BRM short isoforms are evolutionary conserved in Zebrafish. In this organism this isoform is produced as an independent gene on the same chromosome of the long isoform but in its reverse strand. Real time PCR confirmed the presence of the transcript during different zebrafish developmental stage. The fact that this short isoform has maintained its expression independently from BRM after teleosts gene duplication provides an indication that its expression could be physiologically important.

BRM short binds histones but does not interact with SWI/SNF complexes.

Since the short isoforms lack the enzymatic domain but retain the bromodomain, we speculated that BRM short could act as "dominant negative".

As a first step towards the elucidation of the function of BRM short we tested its interactions with SWI/SNF complex by Co-Immunoprecipitation (Co-Ip). In particular we tested the interaction with Histone 3 (H3) and two core subunits of the SWI/SNF complex, SNF5 and BAF155. We know that BRM FL interacts with the

globular domain of Histone H3 (Lavigne et al., 2009) and interacts with SWI/SNF thanks to SNF5 (Muchardt et al., 1995) and BAF155 (Shilpa et al.,2003) subunit. We used HEK 293 cells transiently transfected with BRM FL FLAG-tagged protein, BRMshort FLAG-tagged protein or eGFP FLAG-tagged protein and we performed a FLAG Co-Ip. Our results show that our BRM short protein immunoprecipitates with H3 but not with SNF5 or BAF155 (Fig.7).

Taken together this data show that BRM short variants could bind histones independently from their interactions with SWI/SNF complex. This result allows us to speculate that BRM short variants could play a role as dominant negative on the long isoform.

3.3 Discussion.

A previous report (Yang et al., 2011) suggested the presence of different murine Brm isoforms, that could potentially be generated by alternative splicing. Our present work further characterizes the complexity of *SMARCA2* gene expression, experimentally validating the presence of an alternative promoter (AP) in this gene.

SMARCA2 gene encodes from Brahma (BRM), one of the two alternative ATPase subunits of the SWI/SNF-

BAF complex. Starting from a bioinformatics analysis of this gene we noticed the presence of short 3' transcripts that display a C-terminal bromodomain in common to the BRM full length (Brm FL) isoforms. However, Brm short variants differ from Brm FL in the N-terminal region: as a matter of facts this isoforms do not display the catalytic ATPase domain (Figure 1).

We demonstrated that the ratio between the two classes of Brm isoforms changes during mitochondrial stress. In particular we noticed a relative downregulation of Brm FL transcripts and a concomitant upregulation of Brm short transcripts (Figure2). We cloned a shorter transcript (*SMARCA-001*) and we confirmed its nuclear localization already hypothesized from our bioinfomatic NLS analysis (Figure 3).

More than half of human genes (at least 53%) have alternative promoters (Kimura et al.,2006). The contribution of alternative promoters to the structural and functional diversity of protein isoforms in eukaryotic cells is considered, including their role in synthesis of identical proteins from different mRNAs, generation of protein isoforms with different and even opposite functions, expression of housekeeping genes, and the variation of the recognition domains of adhesion proteins and receptors. In some cases,

alternative promoters allow one gene to produce mRNAs with different open reading frames and, consequently, proteins with no amino acid sequence homology (Pankratova, 2008). In this work we reported the identification of an alternative promoter in the human *SMARCA2* gene. By bioinformatics analysis (Figure 4) and subsequent experimental validation by luciferase reporter assays (Figure 5), we identified a regulatory region that controls the expression of the short Brm isoforms.

We demonstrated that short transcripts are evolutionarly conserved also in Zebrafish. In particular we found that in this organism the short BRM isoform is produced by an independent gene on the same chromosome of the long isoform but in its reverse strand. This is not unusual because the zebrafish underwent to a teleost- specific genome duplication (TSD) that increased its number of genes. Our hypothesis is that if the short isoform has maintained its expression independently from Brm after TSD maybe its expression could be important in a same physiological context.

Searching for a potential functional role for this short isoforms we tested Brm short protein interactors. Using Co-Immunopreciptation we found that BRM short could

bind histones independently from their interactions with SWI/SNF complex. The consideration that both proteins could interact with histones, but that BRM short variants lack of the ATPase domain, suggest a scenario in which Brm short isoforms act as dominant negative proteins. In particular, the increase in the Brm short abundance in the nucleus may saturate the sites of interaction between Brm FL and the acetylated histones (Fig.8). This intriguing hypothesis, that has yet to be verified, may constitute a regulatory loop that could modulates Brm activity in case of mitochondrial stress.

3.4 Materials and Methods

Bioinformatics analyses

The analysis of the human and zebrafish SMARCA2 transcripts was carried out using the Ensemble Genome Browser database (http://www.esemble.org).

The bioinformatics analysis of the human SMARCA2 putative promoter was performed using UCSC Genome Browser and EmbossCpGplot. The bioinformatics prediction of BRM FL and BRM short secondary

structure and NLS localization was performed with PredictProtein software and Expasy Prosite.

The alignment between human and zebrafish sequence was performed with BLASTp (NCBI) and CLUSTAL W Multiple Sequence Alignment (EMBL).

Plasmid construction:

The -1230/+198 region of the human SMARCA2 putative promoter was amplified by PCR using the Phusion Hot Start High-Fidelity DNA Polymerase (Finnenzyme) according the manufacturer's instructions. Human genomic DNA extract from Hela cells was used as template of the PCR. The primers used for the PCR are listed in Table 1. The resulting 1430bp PCR product was purified and cloned into the pGEM-T-easy vector system (Promega). The insert was verified by oligonucleotide sequencing. The SMARCA2 promoter -1230/+198 region was then subcloned into a SmaI-SacI sites of the pGL2 Basic vector (Promega), a promoterless vector which allows the cloning of putative promoter sequences upstream of the *Firefly* luciferase gene. In order to obtain shorter promoter sequences, subsequent 5' deletions were operated using restriction endonuclease digestions (Figure 4C); all the restriction enzymes and corresponding buffers were from New England Biolabs.

The cDNAs encoding the N-terminal flag-tagged human BRM and N-terminal flag-tagged BRG1 proteins were a gift from B. Emerson (Kadam et al., 2003). This cDNA encodes BRM as an N-terminal FLAg fusion protein, and it was subcloned into the Kpn– Xho I sites of pcDNA3 expression vector (Invitrogen). The cDNA encoding one of the human BRMS isoforms (Ensembl SMARCA2-001) was obtained performing a PCR on the cDNA from SH-SY5Y cells. The PCR was performed using the Phusion High Fidelity polymerase (Finnymes) according to manufacturer's instructions. The primers used for the PCR are listed in Table 1.. The resulting band was subcloned into pGEM®-T Easy vector system, and verified by sequencing. The cDNA encoding BRMS was then cloned into the EcoRI – EcoRV sites of the p3xFLAG-Myc-CMV-26 vector (Sigma), that allows to expressed this isoform as a N-terminal 3xFLAG-tagged protein.

Cell culture, transfections and treatments

Human neuroblastoma SH-SY5Y, SH-SY5Y/SOD1, SH-SY5Y/SOD1(G93A), HEK293T cells and NIH3T3 cells were cultured in D-MEM High Glucose medium (Gibco, Invitrogen), 10% fetal bovine serum (FBS), 2.5 mM L-glutamine, 100 U/ml penicillin, and 100 µg/ml streptomycin (Euroclone) at 37°C with 5% CO2. The

treatment with Paraquat (PQ, N,N'-dimethyl-4,4'-bipyridinium dichloride, Sigma, prepared as a 100 mM stock solution in H_2O) was performed for 18 hours at a final concentration of 750 µM.

Cells were transfected using polyethylenimine (PEI, Sigma Aldrich) according to the manufacturer's instructions. Stably transfected cells were obtained by selection with 1µg/ml puromycin (Sigma Aldrich).

For the luciferase assay experiments, 1,5 x 10^5 SH-SY5Y cells and 10^5 HEK293 and NIH3T3 cells were seeded in 24 multiwells, and the next day plasmids (a total of 0,75 µg) were co-transfected using Polyethylenimine (PEI, Sigma, 100 mM in H20 pH 7.00)

according to the manufacturer's instruction. Transfected cells were maintained for 24 hours before lysis with Passive Lysis Buffer 1x (Promega). For forskolin (sigma) treatment cells were transfected and the next day treated with 10uM forskolin. Transfected cells were maintained for 24 hours before lysis with Passive Lysis Buffer 1x (Promega).

Plasmids and Antibodies

The cDNAs encoding the N-terminal flag-tagged human BRM and N-terminal flag-tagged BRG1 proteins were a gift from B. Emerson.

Luciferase Assay

SH-SY5Y cells were transiently co-transfected with the indicated pGL2 luciferase reporter plasmids and with the Renilla-encoding pRL-TK plasmid (Promega Inc.). 24h after transfection, cells were lysed and luciferase activity quantified using the Dual Luciferase Reporter kit (Promega Inc.). 10 µl of the lysates were subjected to the luciferase assays, which was carried out using the Dual-Luciferase Reporter Assay System (Promega) and a Berthold luminometer (Berthold Inc.). The relative luminescence units (RLUs) were obtained normalizing the *Firefly* luciferase readings to the corresponding *Renilla* luciferase readings. The RLUs were further normalized to the bakground signal (constituted by the RLUs of the empty vectors).

RNA analysis

SH-SY5Y SOD1 cells were seeded on 10 cm plates, and 24 hours after plating the RNAs were extracted using TRIzol Reagent (Invitrogen) and subsequently purified using silica membrane spin columns from the RNeasy Mini Kit (Qiagen). RNA quantity and purity were assessed using a NanoDrop® instrument (Thermo Fisher Scientific Inc.). 2 µg of total RNA were reverese-transcribed using the random hexamers-based High Capacity cDNA Reverse-Transcription Kit (Applied

Biosystems), according to manufacturer's instructions. In order to inhibit RNAse activity, RNAsin Plus reagent (Promega) was added to the reverse-transcription reactions.

Zebrafish RNAs were extracted using TRIzol Reagent (Invitrogen) and subsequently purified using silica membrane spin columns from the GenElute™ Total RNA Miniprep Kit (Sigma). 500 ng of total RNA were reverese-transcribed using the random hexamers-based High Capacity cDNA Reverse-Transcription Kit (AppliedBiosystems), according to manufacturer's instructions.

In order to monitor the relative amount of the mRNAs encoding the human *SMARCA2* isoforms in the SH-SY5Y SOD1 cells, we used a three primers retro-transcription PCR (RT-PCR) approach. cDNA synthesis was performed as described above.

The RT-PCRs were perfomed using the polymerase GoTaq Flexi (Promega), following the manufacturer's instructions. We run the following PCR program: 95°C 30'' seconds, 60°C 30'', 72°C 40'', and the cycling was repeated for 35 cycles (we previously ensure that transcript amplification was within a linear range). The sequences of the exon-specific primers are listed in Table 1. Assay conditions were optimized for each gene

with respect to primer annealing temperatures, primer concentrations, and MgCl2 concentrations. Quantification was performed with Bioanalyzer 2100 (Agilent Technologies).

In order to monitor the presence of mRNAs encoding the human *SMARCA2* short isoforms in zebrafish was performed a Real time PCR (qPCR). The primers used for the qPCRs are listed in the Supplemental Table 1.

Antibodies

The antibodies used for immunoblotting were: rabbit polyclonal anti-hBRM/SNF2α (KR17), mouse monoclonal anti-ß-actin AC-40 (Sigma), rabbit polyclonal anti-BRM ChIP Grade (Abcam), monoclonal anti-Flag M2 (Sigma), rabbit polyclonal anti-SMARCB1/SNF5 (Bethyl), rabbit polyclonal anti SMARCC1/BAF155 Antibody (Bethyl), rabbit polyclonal anti-H3 (AbCam).

Immunofluorescence

Cells were seeded on a glass coverslip at day 1. At day 2, cells were either fixed or transfected (and fixed on day 3). The cells were washed twice in PBS1x, and then fixed with 4% Paraformaldehyde (Sigma, diluted in PBS1x) for 10 minutes at room temperature. After washing, permeabilization was performed for 5 minutes on ice using cold CKS solution (20mM HEPES pH 7.4;

3mM $MgCl_2$; 50mM NaCl; 300mM Saccharose; 0,2% Triton-X100). The coverslips were then washed twice with PBS1x-BSA 0,2% (w/v, Sigma). Blocking was performed with a PBS1x-Tween 0,05% solution supplemented with 10%FBS. The samples were then incubated with the following antobodies: mouse monoclonal anti-FLAG M2 (Sigma, diluted 1:200) and rabbit polyclonal anti-BRM (ab15597, Abcam, 1:100

dilution). The incubation with the primary antibodies was performed for 1 hour at 37°C, in a humified chamber. The coverslips were then washed three times with the PBS1x-BSA 0,2% solution. The incubation with the secondary antibody was performed in the same conditions of the primary antibodies. The

secondary antibodies were: anti-mouse Alexa 488 (Molecular Probes, 1:2500 dilution) or anti-rabbit Alexa 633 (Molecular Probes , 1:2500 dilution). After the secondary antibody incubation, DAPI staining (4,6-diamidino-2-phenylindole, Sigma) was performed for 10 minutes at room temperature. After three washings with the PBS1x-BSA 0,2% solution, the coverslips were fixed using FluorSave Reagent (Calbiochem). The images were collected with a confocal Leica microscope (60x lent).

Total extracts preparation and Co-immunoprecipitations

Total protein extracts were prepared using amodified RIPA buffer (50mM TrisHC pH7.4, 150mM NaCl, 1% triton, 1mM EDTA, complete Protease Inhibitor Cocktail (Roche)) For co-immunoprecipitations, 1 mg of extract was first precleared with protein A sepharose (GE) for 2 h at 4°C and then incubated with the indicated antibodies or with rabbit IgG as a control. The antibodies were coupled to protein A-Sepharose beads following manufacturer's instructions. Immunoprecipitations were carried out for 2 hours at 4°C on a rotating wheel. Beads were then washed three times with a solution composed of 10 mM TrisHCl, 100 mM NaCl, 0.01% Nonidet P-40, and 0.04% bovine serum albumin (BSA). The proteins bound to the beads were eluted by boiling the samples at 95°C for 10 minutes. Aliquots were analyzed by SDSPAGE and immunoblotting. For ubiquitin immunoprecipitation, cells were lysed in 1% SDS lysis buffer (50 mM Tris HCl pH 7.4, 0.5 mM EDTA, 1% SDS) and further boiled for an additional 10 minutes. Lysates were clarified by centrifugation at 16,000 g for 10 minutes. Supernatant was diluted 10 times with a buffer composed of 50 mM Tris HCl pH 7.4, 150 mM NaCl, 1% Triton and complete

Protease Inhibitor Cocktail (Roche). Immunoprecipitation was performed with anti-Multi Ubiquitin mAbAgarose (MBL) overnight at 4°C. The precipitates were washed three times and the samples were resolved on 6% SDS-PAGE.

For Flag-tagged CstF50 immunoprecipitation, cells were lysed as for ubiquitin immunoprecipitation. Immunoprecipitation was performed with ANTI-FLAG® M2 Affinity Gel (Sigma) overnight at 4°C. The precipitates were washed three times and eluted in a rotating wheel for 1h at 4°C with a buffer containing 50 mM TrisHCl pH 7.4, 150 mM NaCl, 1% Triton, complete Protease Inhibitor Cocktail (Roche) and 3Xflag-peptide (5 µg/µl). The samples were resolved on 7% SDS-PAGE.

3.5 Figures and Tables

Figure 1:

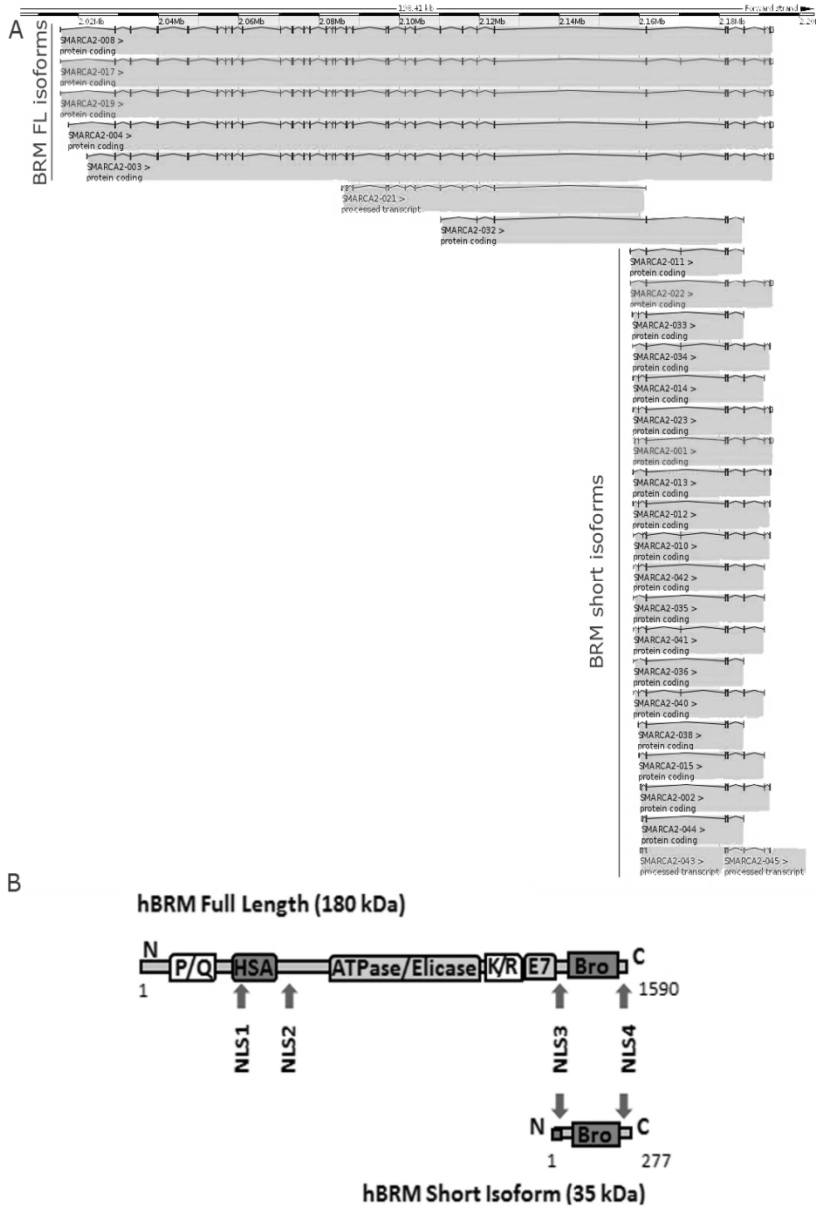


Figura 1: *SMARCA2 gene encodes for two groups of transcripts that generate proteins with different domain composition.*

A. Schematic representation of the human SMARCA2 transcripts as reported by Ensemble Genome Browser. Two classes of transcripts could be identified: the first group encode for BRM FL protein , while the second group for BRM short isoforms.

B. Schematic representation of the domain composition of BRM FL and BRM short isoforms as predicted by various bioinformatics tools (see Material and Methods). P/Q= Proline/Glutamine rich domain, HSA = Helicase Sant domain, ATPase-helic= ATPase and helicase domain, K/R= Lysine/Arginine rich domain, E7 = domain of interaction with Rb proteins, Bro=bromodomain, NLS= nuclear localization signal.

Figura 2:

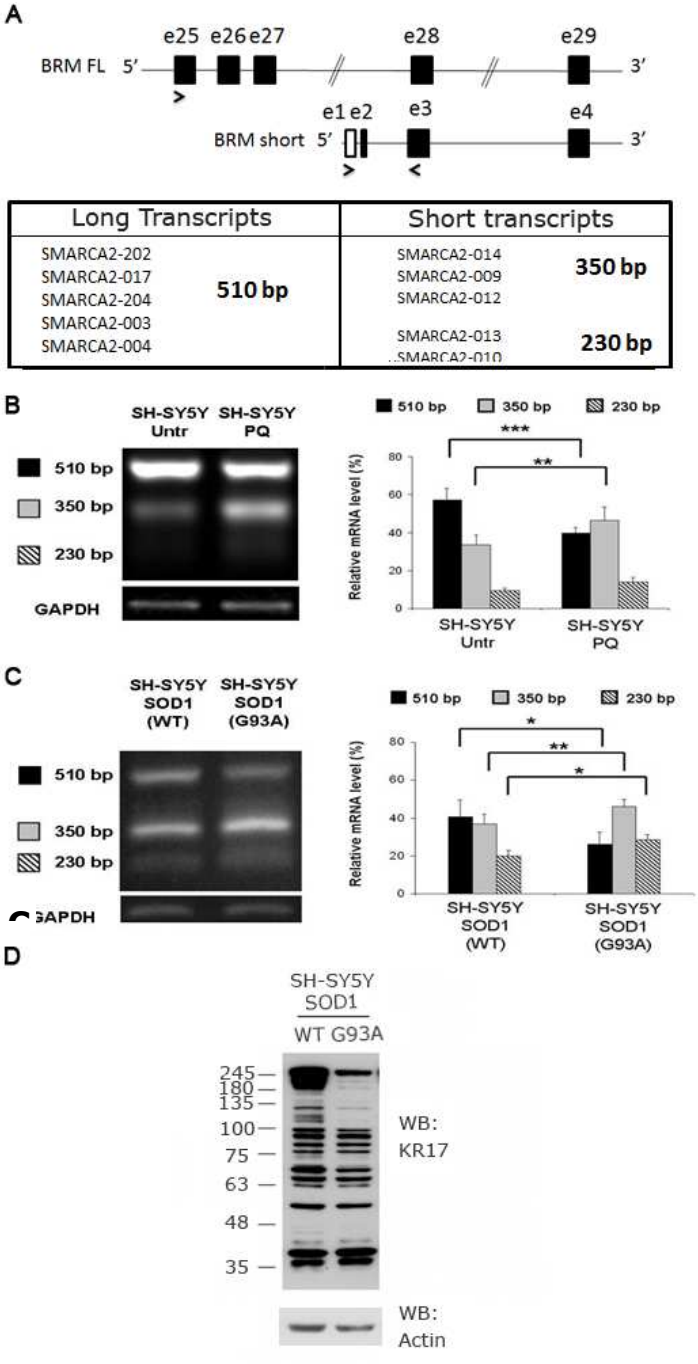

A
e25 e26 e27
e28
e29
BRM FL 5′
3′
e1 e2
e3
e4
BRM short 5′
3′
Long Transcripts
Short transcripts
SMARCA2-202
SMARCA2-017
SMARCA2-204
SMARCA2-003
SMARCA2-004
510 bp
SMARCA2-014
SMARCA2-009
SMARCA2-012
350 bp
SMARCA2-013
230 bp
B
SH-SY5Y Untr
SH-SY5Y PQ
510 bp
350 bp
230 bp
GAPDH
Relative mRNA level (%)
C
SH-SY5Y SOD1 (WT)
SH-SY5Y SOD1 (G93A)
D
SH-SY5Y SOD1
WT G93A
245
180
135
100
75
63
48
35
WB: KR17
WB: Actin

Figure 2: Oxidative stress affects BRM FL and short expression.

A. Upper panel: Schematic representation of BRM FL and BRM short representative isoforms indicating the positions of the primers (arrows) used for the RT-PCR. Lower panel: size of the RT-PCR products: the 510bp band represents the corresponding transcripts encoding the BRM FL isoforms, while the 350bp and the 230bp bands represent the corresponding transcripts encoding the BRM short isoforms.
B. RT-PCR analysis of SH-SY5Y cells untreated or treated with Paraquat (PQ). *Left panel*: representative gel. GAPDH was used as a loading control. *Right panel*: quantification of six independent experiments. The experimental variability is expressed as a standard deviation T-test:**=$p<0.01$, ***=$p<0.001$.
C. RT-PCR analysis of SH-SY5Y cells overexpressing the wildtype (WT) o mutated (G93A) SOD1. *Left panel*: representative gel. GAPDH was used as a loading control. *Right panel*: quantification of six independent experiments. The experimental variability is expressed as a standard deviation T-test:*=$p<0.05$, **=$p<0.01$.
D. Western blot analysis on SH-SY5Y SOD1 total extracts with the KR17 antibody. Actin was used as a loading control.

Figure 3:

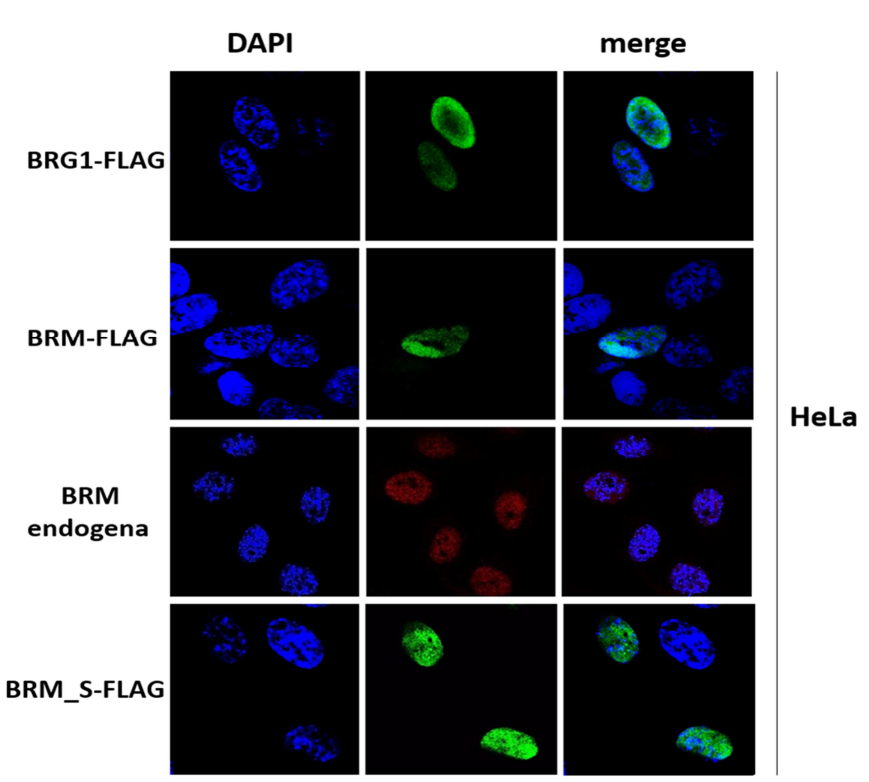


Figure 3: BRM short localize in the nucleus.

Immunofluorescence of transiently transfected Hela cells.

First row: Hela cells transiently transfected with a BRG1-Flag protein.

Second row: Hela cells transiently transfected with a BRM FL-Flag protein.

Third row: Hela cells not transfected

Fourth row: Hela cells transiently transfected with a BRM short-Flag protein.

In green the Alexa 488 signals of the indicated protein,in red the Alexa 633 signals of the endogenous BRM FL proteins. DAPI (blue) indicates the location of nuclei.

Figure 4:

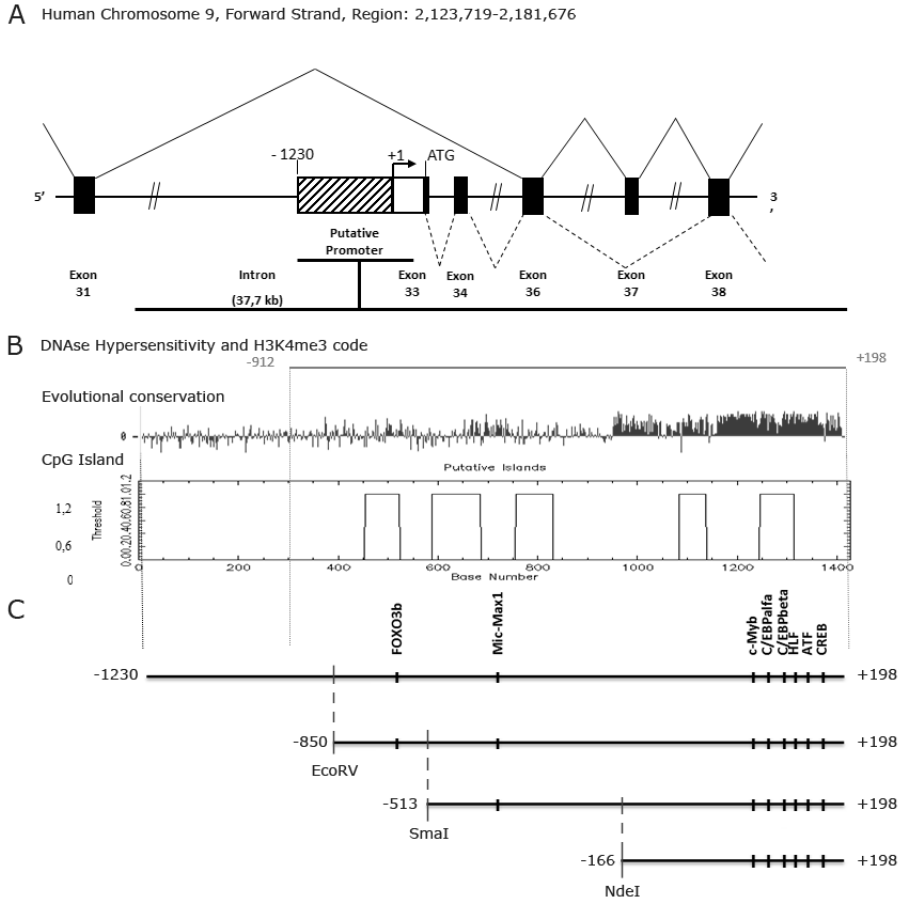

A
Human Chromosome 9, Forward Strand, Region: 2,123,719-2,181,676
- 1230
+1
ATG
5'
3'
Putative Promoter
Exon 31
Intron (37,7 kb)
Exon 33
Exon 34
Exon 36
Exon 37
Exon 38
B
DNAse Hypersensitivity and H3K4me3 code
-912
+198
Evolutional conservation
CpG Island
Putative Islands
Threshold
1,2
0,6
0
0 200 400 600 800 1000 1200 1400
Base Number
C
FOXO3b
Mic-Max1
c-Myb
C/EBPalfa
C/EBPbeta
NF1
ATF
CREB
-1230
+198
-850
EcoRV
-513
SmaI
-166
NdeI

Figure 4: Bioinformatics analysis of *SMARCA2* short promoter.

A. Schematic representation of the genomic region of interest containing the putative human SMARCA2 short promoter.
B. Bioinformatics analysis of the region of interest. +1 denotes the transcriptions site of the SMARCA-001 isoform. The evolutionarly conservation, histone marks, DNAse hypersensibility and transcription factors are evaluated using UCSC Genome Browser. The CpG islands analyses was performed with Emboss CpGplot. The 5' deletion regions show the restriction enzyme used for the cloning.

Figure 5:

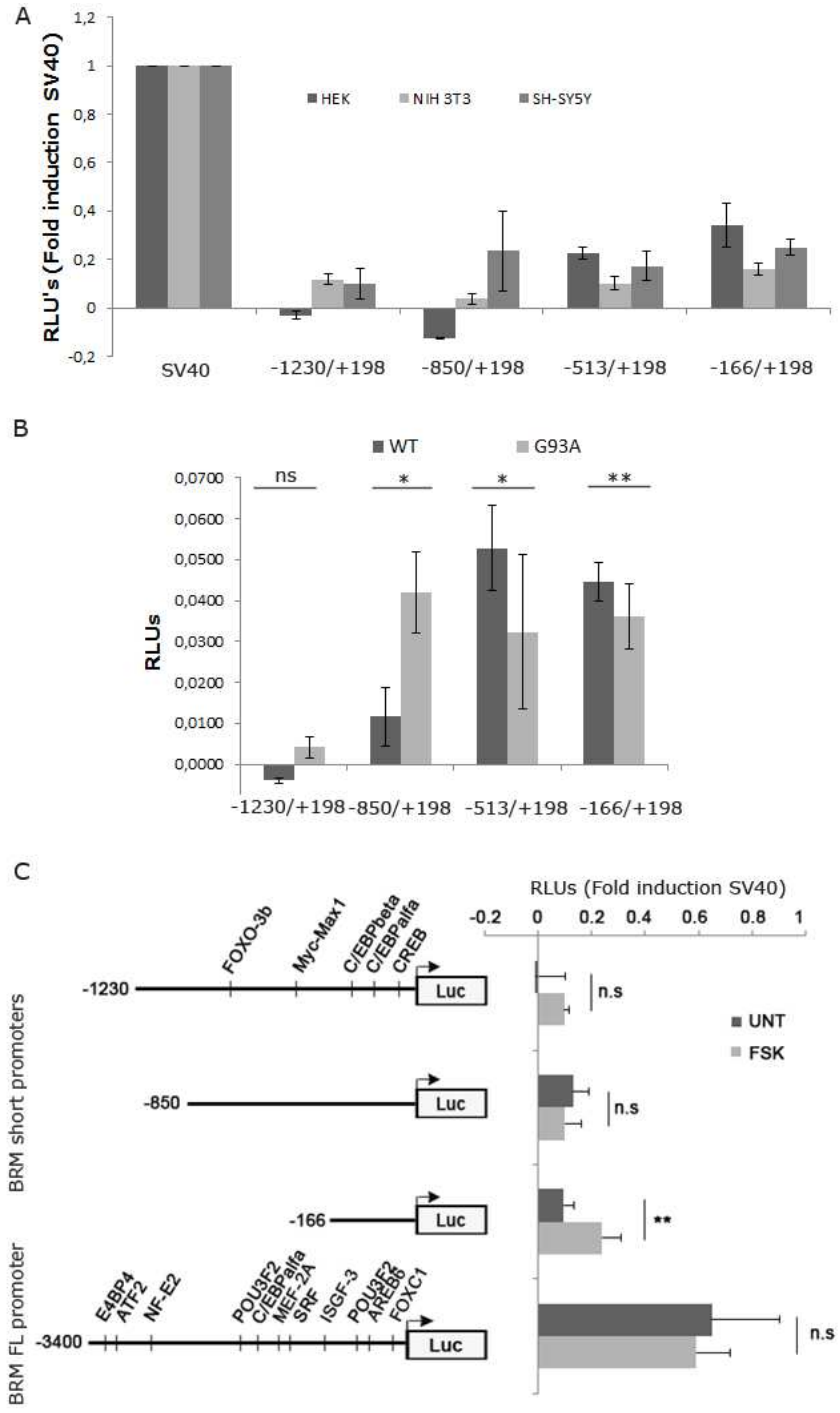

A
RLU's (Fold induction SV40)
1,2
1
0,8
0,6
0,4
0,2
0
-0,2
HEK
NIH 3T3
SH-SY5Y
SV40
-1230/+198
-850/+198
-513/+198
-166/+198
B
WT
G93A
RLUs
0,0700
0,0600
0,0500
0,0400
0,0300
0,0200
0,0100
0,0000
ns
*
*
**
-1230/+198
-850/+198
-513/+198
-166/+198
C
RLUs (Fold induction SV40)
-0.2
0
0.2
0.4
0.6
0.8
1
BRM short promoters
BRM FL promoter
FOXO-3b
Myc-Max1
C/EBPbeta
C/EBPalfa
CREB
-1230
Luc
-850
Luc
-166
Luc
E4BP4
ATF2
NF-E2
POU3F2
C/EBPalfa
MEF-2A
SRF
ISGF-3
POU3F2
AREB6
FOXC1
-3400
Luc
UNT
FSK
n.s
n.s
**
n.s

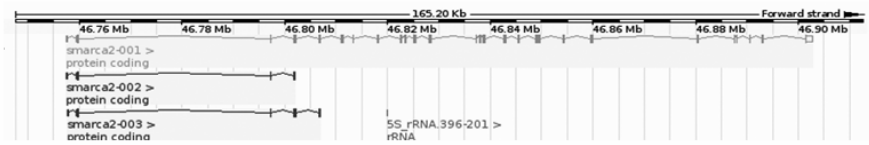
Fig. 5: Brm short promoter regions show different activity in response to oxidative stress and CREB activation.

Results of the luciferase (luc) reporter assay. Luciferase units (RLUs) were normalized to the background signal obtained with the empty vector signal. For each cell line a second normalization was then operating setting the normalized over-background RLUs of the pGL2SV40 construct to 1. The other RLUs constructs are expressed as fractions of this value. (A,C). The graph shows the results of 3 independent experiments in duplicate. The experimental variability is expressed as standard deviation. Where present T-test:*=p<0.05, **=p<0.01.

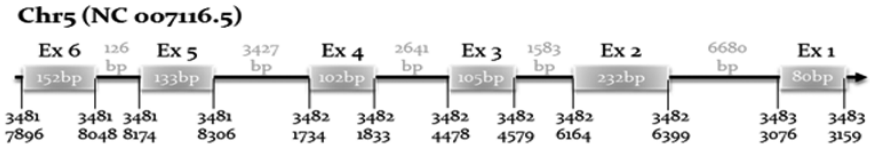
A. Results on HEK, NIH3T3 and SH-SY5Y cells.
B. Results on WT or G93A SOD1 SH-SY5Y cells.
C. Results on forskolin (FSK) treated and untreated SH-SY5Y cells. Left panel: schematic representation of the SMARCA2 promoter with their transcription factors. Right panel: Results of the luciferase (luc) reporter assay.

Figure 6:

A **SMARCA 2:**
Chromosome 5: 46,757,516-46,902,718 forward strand



B



C

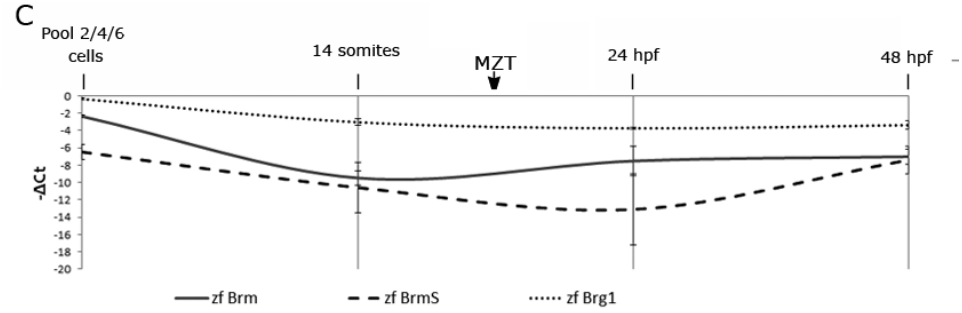


Figure 6: BRM short isoform is conserved in zebrafish.

A. Schematic representation of the zebrafish SMARCA2 transcripts as reported by Ensemble Genome Browser. It present one long isoforms and two short isoforms that have in common with the longest one the first four or five exons. 3' short trascripts are not reported for this gene.

B. Results of the protein Blast performed using hSMARCA2-001 protein in a non redundant protein sequences database of Danio rerio.

C. Results of the Real time PCR performed on zebrafish BRMFL, BRMS and BRG1 during different stages of zebrafish development. Data are plot with their $-\Delta Ct$ that inversely correlates with the initial template amount. MZT=maternal to zygotic transition, hpf = hours post fertilization.

Figure 7:

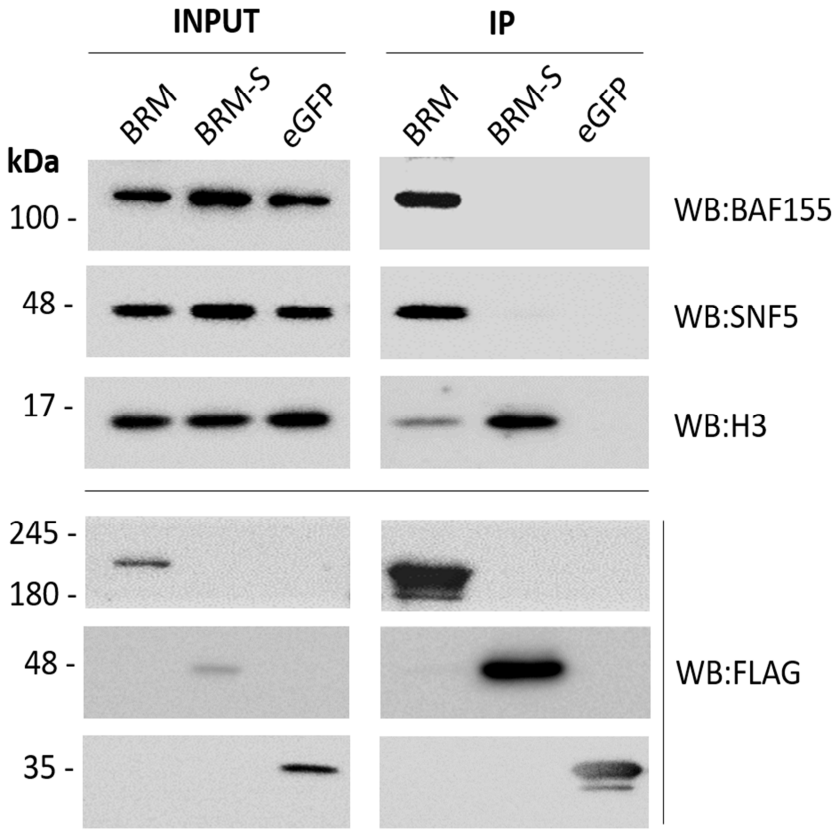


Figure 7: BRM Short isoform interacts with histone 3 but not with the SWI/SNF complex.

Hek cells were transiently transfected with BRM FL FLAG-tagged protein, BRMshort FLAG-tagged protein or eGFP FLAG-tagged protein as a negative control. Proteins were immunoprecipitated with a resin conjugated with an anti-FLAG antibody as described in "Materials and Methods" and analysed by Western blotting with the indicated antisera. Inputs are 1% of the immunoprecipitates.

Figure 8:

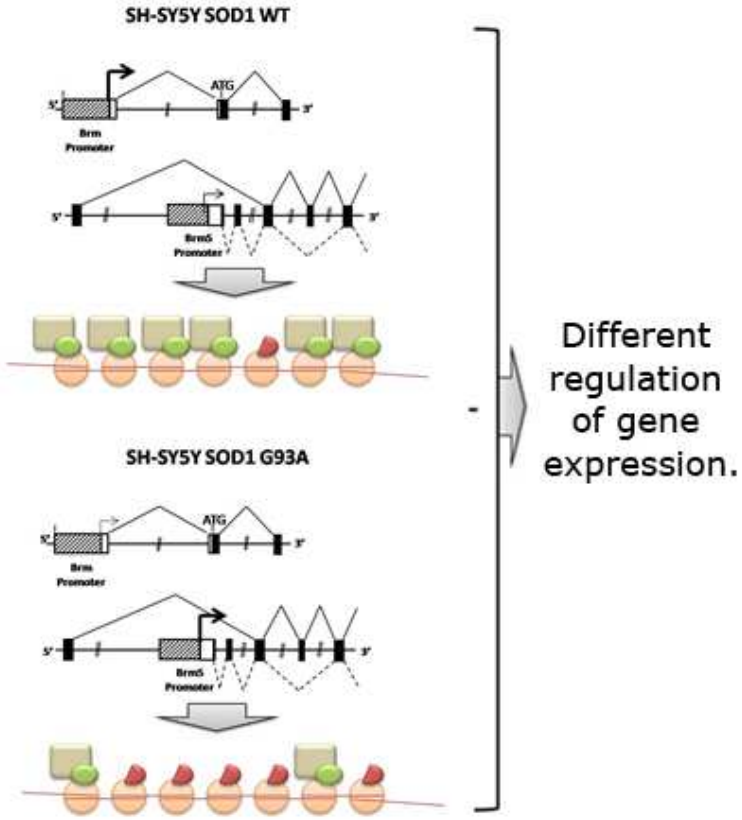


Figure 8: BRM could play a role as dominant negative. Schematic representation of an hypothetic functional BRM short model. Schematic representation of the possible chromatin scenario: pink circles =nucleosomes, red half circles = BRM short proteins, green circles =BRM proteins, grey rectangle =SWI/SNF complexes. Bold narrow indicates an higher promoter activity while normal narrow a lower activity.

Table1:

Oligos for the putative SMARCA2 short promoter	
-1230/+198 FW	TTTCTTGTTTGGGGGATCA
-1230/+198 REV	CCCCAAAACTTGCTACACAA
Oligos for BRM short 001 cloning	
SMARCA-001 FW	AAGAGACTAGCAGCTCGCTGC
SMARCA-001 REV	TCACTCATCATCCGTCCCACTTC
Oligos for Retro-Transcription PCR (RT-PCR) RNA analysis	
BRM FL FW	ATCTTGGAGCATGAGGAGG
BRMS FW	GGGGTTTGCTTCTGTGATTT
BRM common REV	GCGTTCATCTGCTTTGTCAG
Oligos for Quantitative Real-Time PCR (qPCR)	
zf BRM FW	GCAGCCAAGCACATCATAGA
zf BRM REV	TTGAATGGTCTTCCCCAGAC
zf BRMS FW	AGGGCTCTGAGACTGGTCAA
zf BRMS REV	GCCCAACGTGTCCTTGTAGT
zf BRG1 FW	TGCCAAGCAAGATGTAGACG
zf BRG1 REV	GCCCGTTCACTAGCAGACTC

3.6 References

Arce, L., Yokoyama, NN., Waterman, ML. (2006). Diversity of LEF/TCF action in development and disease. Oncogene 25, 7492–7504.

Baek, D., Davis, C., Ewing, B., Gordon, D., Green, P. (2007). Characterization and predictive discovery of evolutionarily conserved mammalian alternative promoters. Genome Res. 17, 145–155.

Batschè, E., Yaniv, M., Muchardt, C. (2006). The human SWI/SNF subunit BRM is a regulator of alternative splicing. Nat. Struct. & Ml. Biol.,13:22-29.

Bourachot, B., Yaniv, M., Muchardt, C. (1999). The activity of mammalian BRM/SNF2alpha is dependent on a high-mobility-group protein I/Y-like DNA binding domain. Mol Cell Biol. 19(6):3931-9.

Ho, L, Jothi, R, Ronan, JL, Cui, K, Zhao, K, Crabtree, GR. (2009). An embryonic stem cell chromatin remodeling complex, esBAF, is an essential component of the core pluripotency transcriptional network. Proc Natl Acad Sci U S A.,106:5187–5191.

Howe, K., Clark, M.D., Torroja, C.F, …, Stempl, D.L. (2013). The zebrafish reference genome sequence and its relationship to the human genome. Nat. Letter.;496:498-503.

Kadam, S, Emerson, BM.(2003). Transcriptional specificity of human SWI/SNF BRG1 and BRM chromatin remodeling complexes. Mol Cell. 11(2):377-89.

Kimura, K, Wakamatsu, A, Suzuki, Y, Ota, T, Nishikawa, T, et al. (2006) Diversification of transcriptional modulation: large-scale identification and characterization of putative alternative promoters of human genes. Genome Res 16: 55–65.

Kolasinska-Zwierz, P., Down, T., Latorre, I., Liu, T., Liu, X.S., Ahringer, J. (2009) Differential chromatin marking of introns and expressed exons by H3K36me3. Nat. Gen., 41:376-381

Lavigne, M.,Eskeland, R.,Azebi, S.,Saint-André, V.,Min Jang, S.,Batsché, E.,Ying, Fan H.,Kingston, R.E.,Imhof, A. and Muchardt C. (2009) Interaction of HP1 and Brg1/BRM with the Globular Domain of Histone H3 Is Required for HP1-Mediated Repression. PLoS Genet. 2009 December; 5(12): e1000769.

Lessard, J, Wu, JI, Ranish, Ja, Wan, M, Winslow, MM, Staahl, BT, Wu, H, Aebersold, R, Graef, Ia, Crabtree, GR. (2007). An essential switch in subunit composition of a chromatin remodeling complex during neural development. Neuron; 55:201–15.

Loe-Mie, Y, Lepagnol-Bestel, AM, Maussion, G, Doron-Faigenboim, A, Imbeaud, S, Delacroix, H, Aggerbeck, L, Pupko, T, Gorwood, P, Simonneau, M, Moalic, JM. (2010) SMARCA2 and other genome-wide supported schizophrenia-associated genes: regulation by REST/NRSF, network organization and primate-specific evolution. Hum Mol Genet. 15;19(14):2841-57

Lonze, B.E., Ginty, D.D. (2002) Function and regulation of CREB family transcription factors in the nervous system. Neuron, 35:605–623

Machida, Y., Murai, K., Miyake, K. and Lijima, S. (2001) Expression of chromatin remodeling factors during neural differentiation. J. Biochem. 129:43-49

Maracchioni, A.,Totaro, A.,Angelini, D.F.,Di Penta, A., Bernardi, G., Carri, M.T., Achsel, T. (2007). Mitochondrial damage modulates alternative splicing in neuronal cells: implications for

neurodegeneration. J Neurochem., 207;100:142-153

Menga, A., Iacobazzi, V., Infantino, V., Avantaggiati, M.L., Palmieri, F. (2015) The mitochondrial aspartate/glutamate carrier isoform 1 gene expression is regulated by CREB in neuronal cells. Int. J. Biochem. Cell Biol., 60:157–166

Meyer, A. & Schartl, M. (1999). Gene and genome duplications in vertebrates: the one-to-four(-to-eight in fish) rule and the evolution of novel gene functions. Curr.Opin.Cell Biol.11:699–704.

Monaghan, T.K., Mackenzie, C.J., Plevin, R., Lutz, E.M. (2008). PACAP-38 induces neuronal differentiation of human SH-SY5Y neuroblastoma cells via cAMP-mediated activation of ERK and p38 MAP kinases. J Neurochem, 104:74–88

Muchardt, C, Sardet, C, Bourachot, B, Onufryk, C, Yaniv, M. (1995) A human protein with homology to Saccharomyces cerevisiae SNF5 interacts with the potential helicase hBRM. Nucleic Acids Res. 11;23(7):1127-32.

Pankratova, E.V. (2008). Alternative promoters in expression of genetic information. Mol Biol. 42: 371.

Sands, W.A., Palmer, T.M. (2008). Regulating gene transcription in response to cyclic AMP elevation. Cell Signal, 20:460–466

Shi, P., Gal, J., Kwinter, D.M., Liu, X., Zhu, H. (2010). Mitochondrial dysfunction in Amyotrophic Lateral Sclerosis. Biochem Biophys Acta. 1802:45-51.

Shilpa, K., Emerson, B.M.,(2003). Transcriptional Specificity of Human SWI/SNF BRG1 and BRM Chromatin Remodeling Complexes. Molecular Cell , 11 , 2 : 377 - 389

Xin, D, Hu, L, Kong, X (2008) Alternative Promoters Influence Alternative Splicing at the Genomic Level. PLoS ONE 3(6): e2377.

Yang, M., Sun, Y., Ma, L., Wnag, C., Wu, J.M., Bi, A., Liao, J.(2011) Complex Alternative Splicing of the Smarca2 Gene suggests the importance of Smarca2-B variants. Journal of Cancer 2:386-400

Chapter 4
Conclusion

4.1 Summary

In the present thesis, I report the results obtained from my research activity carried out during this PhD program.

My work was focused on the human protein Brahma (BRM, SNF2α), encoded by the SMARCA2 gene. BRM is one of the two mutually exclusive alternative ATPases that could be present in the SWI/SNF-BAF (Brahma-Associated Factors) complexes. SWI/SNF complexes are huge multi-enzymatic complexes that remodel the chromatin whitin these complex the main task of BRM is to hydrolize the ATP to obtain the required energy to expose particular sites present in the chromatin or to alter the nucleosome composition. These modifications of the chromatin state have crucial impacts on gene expression (Clapier et al., 2009). It has been proposed that BRM can also regulate the alternative splicing of pre-mRNA (AS), modulating the inclusion of internal cassette exons. More specifically, BRM exerts this activity interacting with Sam68 (an exon inclusion enhancer) and components of the spliceosome. BRM acts as exon inclusion enhancer, and its activity is exerted cotranscriptionally on internal cassette exons (Batsché et al., 2006).

My first project derived from a comparative evaluation and validation of microarray data from two mitochondrial stress models. The first model represented by an acute mitochondrial stress is constituted by human SH-SY5Y neuroblastoma cells treated with Paraquat (PQ), while the second model is a chronic model constituted from the same cell line stably overexpressing the Superoxide Dismutase 1 carrying the most common mutation found in familiar ALS (SOD1 G93A). The merge of this microarrays data showed that oxidative stress affects the choice of specific ALEs increasing the production of transcripts variants terminating at a more proximal ALE. Moreover, Oxidative stress induces the transcriptional downregulation of the SMARCA2 gene product: BRM. I found that in normal condition BRM is enriched on the proximal ALE. In addition I observed the accumulation of BARD1, a protein that forms a functional heterodimer with BRCA1 which has E3 ubiquitin-ligase activity and interacts with the 50 kDa subunit of CstF inhibiting 3' end processing. Consistent with these observations, I detected an ubiquitinated pool of CstF50 and showed that ubiquitination is mediated by BARD1/BRCA1. Taken together, these results suggested that the presence of BRM on the proximal exon leads to the BARD1/BRCA1-mediated ubiquitination of CstF50 and

the inhibition of 3’ end processing at the proximal poly(A). This in turn allows transcription to proceed to the distal terminal exon.

My second project started from the same microarray data used as a starting point of my first project. In these microarray data detected an increase of short *SMARCA2* gene isoforms. In particular, the microarray results showed that the ratio of long and short isoforms changed after oxidative stress. For this raison my second project was focused on the characterization of the *SMARCA2* gene. Starting from a bioinformatics analysis we noticed that the arrangement of long and short isoforms in the locus leads to think that short isoforms are not a product of alternative splicing, but that are transcribed by using an alternative promoter. Using bioinformatics tools I located a possible promoter region and I cloned and tested this region. As a matter of fact this region works in different cell lines and is modulated upon oxidative stress. The isoforms encoded by this promoter lack the N-terminal region that contains the catalytic domain of BRM. BRM-containing SWI/SNF complexes are enriched in differentiated cells, where they regulate the expression of genes involved in differentiation. Moreover, BRM has an important role in co-transcriptional alternative splicing (AS). Concerning the short isoforms, it is known that they are transcribed

but there are no evidences about their biological function. We discovered that short isoforms are evolutionarily conserved. We found one short isoform also in zebrafish and interestingly in this organism this isoform is produced as an independent gene on the same chromosome of the long isoform but from the reverse strand. If this short isoform has maintained its expression independently from BRM after teleosts gene duplication maybe its expression could be important.

bioinformatics analyses reveal that the short protein in comparison with the long one has in common only the C-terminal domain where there is a Bromodomain. The N-terminal domain that contains the catalytic domain of BRM is missing in the short isoforms.

To establish a possible functional role of short isoforms and I tested if the short protein similary to BRM FL is able to interact with histones. By Co-Immunoprecipitation (Co-IP), I verified that both isoforms interact with histone H3. Considering that short isoform lacks the ATPase domain, the Co-IP suggests a possible "dominant negative" role for this protein. If the short isoforms work as negative dominant of BRM, the ratio between long and short expressed proteins could become very important for BRM target genes. In particular this alteration of ratio could have a negative effect on the cells since several

tumor cell lines show a very low level of the long protein. (i.e. human lung tumor cell lines). The idea that Brm short could act as dominant negative is supported by the evidence that other proteins that lack their catalytic domain negative regulate the long isoforms activity. For example, Homer1 is present as two forms: Homer1a and Homer 1b/c. Homer 1a is the short isoforms and is a splice variant of Homer1b/c. This protein lacks the ability of linking mGluR1/5 to synaptic proteins and functions as an endogenous negative modulator of the mGluR1/5 inositol 1,4,5-trisphosphate receptor signaling complex (Tappe et al., 2006).

4.2 Future perspectives

The present thesis shows a novel role for BRM in the regulation of alternative splicing. BRM has already been implicated in alternative splicing: it was previously shown that BRM promotes the inclusion of variant internal exons of the CD44 gene by inducing a transient pausing of the transcribing RNAPII (Batsché et al., 2006). However, in Drosophila, BRM-containing SWI/SNF complexes, while inducing RNAPII pausing, can negatively regulate splicing (Bochar et al., 2000). In this case, RNAPII stalling, and the consequent splicing reduction, were not induced by BRM but were

rather dependent on SNF5. Instead, BRM remodeling activity was shown to be necessary for the subsequent chromatin remodeling following the release of the stalled polymerase (Zraly et al., 2012). In line with Zraly and coauthors, we found that BRM negatively affects the inclusion of the proximal ALE. This effect does not require its ATPase activity. The choice of alternative poly(A) sites located in different terminal exons occurs in conjunction with splicing, and splicing factors are known to influence 3' processing (reviewed in Millevoi et al., 2010). Thus, the mechanism we describe is certainly not the unique determinant of the choice between proximal and distal terminal exons.

We demonstrated that BRM downregulation in human neuroblastoma cells is triggered by a mitochondrial stress that induced impairment in SMARCA2 promoter activity. Oxidative stress is a common biological event, and has been associated with cancer, neurodegenerative diseases and aging. In our laboratory, my collaborators are collecting microarray data regarding the mitochondrial stress induced impairments in the expression of miRNAs. It will be very interesting to verify if the same miRNAs which downregulate BRM in tumor cells also participate to the downregulation of BRM in our cellular models. From my data also emerges that mitochondrial stress impairs the

BRM-dependent AS of the terminal exons of five out of six genes involved in the regulation of axon growth and guidance or, more generally, in the regulation of neuronal functions. Recent reports suggested that axon retraction is one of the first hallmarks of neurodegeneration, and that these proteins could represent a novel therapeutic targets for Amyotrophic Lateral Sclerosis (ALS) clinical trials (Schmidt et al., 2009). In my screening, I found that genes like RPRD1A (or p15RS, a gene involved in the inhibition of the WNT/β-catenin pathway and/or in the regulation of termination of the RNA Pol II-dependent transcription) (Yang et al., 2010, Wu et al., 2010; Blakely et al., 2011; Ni et al., 2011), SLC6A15 (a gene encoding a brainrestricted aminoacid transporter) (Takanaga et al., 2005) IGSF1 (encoding a protein that plays crucial roles in cell-to-cell interactions) (Mazzarella et al., 1998) and STRADA (a regulator of neuronal polarity and synaptic organization) (Kim et al.,
2010) respond in a dose-dependent fashion to BRM expression. The BRM-dependent, mitochondrial stress-induced alterations of the choice of the 3' terminal exons of these genes may generate proteins which display profound changes in their domain compositions (data not shown). In order to establish a possible link between BRM expression, these genes, and

neurodegeneration, it will be of major interest to validate our observation in an in vivo system. A BRM knock-out mouse model is available (Reyes et al., 1998), and it exhibits a mild phenotype. No signs of neurodegeneration were reported in these mice, probably because the animals were sacrified at early adult stages, and neurodegeneration usually onsets in late adult life stages. So far, no reports have specifically connected BRM to the onset of neurodegeneration and/or to other diseases which specifically alters neuronal functionality. However, alterations the expression of SMARCA2 gene has been linked to the emergence of schizophrenia. This role has been established by a genome-wide survey of genes involved in the pathophysiology of this psychiatric disease (Leo-Mie et al., 2010), as well as by direct evaluation of the social behavior of BRM knock out mice (Koga et al., 2010). A direct evidence concerning the roles that BRM may play in the onset of neurodegenerative diseases, and more specifically in the onset of ALS, is still missing. However, the results reported in the present thesis, together with my preliminary observations regarding the BRM-dependent transcriptional regulation of the expression of CXCR4 and HGF (data not shown), two genes linked to ALS onset (Luo et al., 2007; Kodoyama et al., 2007),

suggest that this link may exist. If direct causal links would be established between BRM, the deregulation in the expression of genes involved in axon growth and guidance, and neurodegeneration, then these results would fit in family of "RNA-related" genes whose functions are impaired in ALS pathology. As a matter of fact, a long list of genes which confer major risk of ALS onset are linked to various aspect of RNA biology, ranging from transcription (ELP3, ANG2, STX3), to splicing (FUS, TDP43) and editing (SMN, TLS) (Blitterswijk et al., 2010). Finally, it is notewhorty to observe that the putative role of BRM in the control of the AS of the last exons of genes involved in axon growth and guidance fits in the frame of the crucial roles played by BRM in the regulation of the neuron-specific gene expression. As a matter of fact, it has been demonstrated that BRM is enriched in neurons and plays crucial roles in regulating the expression of genes involved in neurogenesis (Olave et al., 2002), such as Neurogenin1 (Wu et al., 2009). In the present thesis, I also report the identification of a novel alternative promoter (AP), localized in one intron of the human SMARCA2 gene. This promoter controls the production of the transcripts encoding BRM short isoforms. Another research team reported the presence of isoforms of murine SMARCA2 gene, but this paper

focused on the murine BRM isoforms derived only from alternative splicing events (Yang et al., 2011). BRM short isoforms display a C-terminal bromodomain, a domain involved
in the interaction with the acetylated histones (Lavigne et al., 2009). They differ from the BRM FL isoforms in the N terminal region, because they do not display the N-terminal catalytic domains and the regions which allow to BRM FL to enter in the SWI/SNF-BAF complex (Muchardt et al., 1995). These observations and the Co-IP experiments suggest that short variants may interact with acetylate histones, but also that their activity may be exerted outside of the chromatin remodeling complex. Interestingly, I have demonstrated that both BRM isoforms are nuclear, suggesting that a cross-talk between them is possible. From my analysis also emerges that the promoter activity of the regulatory region that control the expression of short transcripts increases in response to SOD1 (G93A) overexpression, while the activity of the promoter from which the full length isoforms originate is donwregulated by the same stressful stimulus. These data are consistent with the results obtained from the analysis of the relative abundance of the mRNAs levels of the two groups of isoforms. Taken together, these observations suggest the intriguing hypothesis that

BRM short isoforms may modulate BRM FL activity during mitochondrial stress, "buffering" the sites of interaction with acetylated histones. The idea is that short isoforms could have a possible role of "dominant negative" for the BRM FL protein.

I think that demonstrating the existence of dominant negative of the FL isoforms could help to better understand the role of BRM in pathology. These complex patterns of transcription initiation suggest that in some situations the functions of the *SMARCA2* gene may be very complex.

In the last year BRM has become a novel anticancer gene, which is frequently inactivated in a variety of tumor types. Unlike many anticancer genes, BRM is not mutated, but rather epigenetically silenced (Glaros et al.,2007). Histone deacetylase complex (HDAC) inhibitors are known to reverse BRM silencing, but they also inactivate it via acetylation of its C-terminus (Bourachot et al., 2007). High-throughput screening has uncovered many compounds that are effective at pharmacologically restoring BRM and thereby inhibit cancer cell growth. In order to identify any potential clinically effective drugs that could be used to restore BRM, it is necessary to understand the endogenous proteins involved in BRM silencing, as well as its inactivation. It could be also interesting to test the

activities of the *SMARCA2* promoter constructs that I have generated in tumor cell lines, to check if BRM downregulation in tumors is not only due to epigenetic silencing but also to a transcriptional impairment. Interestingly, it has been very recently demonstrated that a transcriptional impairment of BRM expression, caused by two sequences insertion in BRM promoter, is associated to lung cancer risk (Liu et al., 2011).

4.3 references:

Batsché E., Yaniv M., Muchardt C. (2006). The human SWI / SNF subunit BRM is a regulator of alternative splicing. Nat. Struct. & Mol. Biol.13:22-29.

Blakely B.D., Bye C.R., Fernando C.V., Horne M.K., Macheda M.L., Stacker S.A., Arenas E., Parish C.L. (2011). Wnt5a regulates midbrain dopaminergic axon growth and guidance. PloS One. 6:e18373

Blitterswijk M.V., Landers J.E. (2010). RNA processing pathways in amyotrophic lateral sclerosis. Neurogenetics. 11:275-290.

Bochar, D.A., Wang, L., Beniya, H., Kinev, A., Xue, Y., Lane, W.S., Wang, W., Kashanchi, F. and Shiekhattar, R. (2000). BRCA1 is associated with a human SWI/SNF-related complex: linking

chromatin remodeling to breast cancer. Cell, 102, 257-265.

Bourachot B, Yaniv M, Muchardt C. (2003). Growth inhibition by the mammalian SWI-SNF subunit BRM is regulated by acetylation. EMBO J; 22(24): 6505–6515.

Clapier C.R., Cairns B.R. (2009). The biology of chromatin remodeling complexes. Ann. Rev. Biochem. 78:273-304.

Glaros S, Cirrincione GM, Muchardt C, Kleer CG, Michael CW, Reisman D. (2007). The reversible epigenetic silencing of BRM: implications for clinical targeted therapy. Oncogene. 26(49): 7058–7066.

Kadoyama K., Funakoshi H., Ohya W., Nakamura T. (2007). Hepatocyte growth factor (HGF) attenuates gliosis and motoneuronal degeneration in the brainstem motor nuclei of a transgenic mouse model of ALS. Neurosci. Res. 59:446-456.

Kim S., Hyunmin K., Fong N., Erickson B., Bentley D.L. (2011). Pre mRNA splicing is a determinant of histone H3K36 methylation. PNAS,; 108:13564-13569.

Koga M., Ishiguro H., Yazaki S., et al. (2009). Involvement of SMARCA2 / BRM in the SWI / SNF chromatin-remodeling complex in schizophrenia. Hum.Mol Gen. 18:2483-2494.

Lavigne M., Eskeland R., Azebi S., Saint-André V., Min Jang S., Batsché E., Ying Fan H., Kingston R.E., Imhof A., Muchardt C. (2009). Interaction of HP1 and Brg1/BRM with the globular domain of histone H3 is required for HP1-mediated repression. PLoS Gen. 5:1-11.

Liu G., Gramling S., Munoz D., Cheng D., Azad A.K., Mirshams M., Chen Z., Xu W., Roberts H., Sheperd F.A:, Tsao M.S., Reisman D. (2011). Two novel BRM insertion promoter sequence variants are associated with loss of BRM expression and lung cancer risk. Oncogene. 30:3295-3304.

Loe-Mie Y., Lepagnol-Bestel A.M., Maussion G., Doron Faigenboim A., Imbeaud S., Delacroix H., Aggerbeck L., Pupko T., Gorwood P., Simonneau M., Moalic J.M. (2010). SMARCA2 and other genome-wide supported schizophrenia-associated genes: regulation by REST/NRSF, network organization and primate specific evolution. Hum Mol Genet. 19:2841-57

Luo Y, Xue H., Pardo A.C., Mattson M.P., Rao M.S., Maragakis N.J. (2007). Impaired SDF1 / CXCR4 signaling in glial progenitors derived from SOD1 G93A mice. Jour. Neurosci. Res. 85:2422-2432.

Mazzarella R., Pengue G., Jones J., Jones C., Schlessinger D. (1998). Cloning and expression of

an Immunoglobulin Superfamily Gene (IGSF1) in Xq25. Mol. Microbiol.162:157-162.

Millevoi, S. and Vagner, S. (2010) Molecular mechanisms of eukaryotic pre-mRNA 3' end processing regulation. Nucleic Acids Res, 38, 2757-2774.

Muchardt C., Sardet C., Bourachot B., Onufryk C., Yaniv M. (1995). A human protein with homology to Saccharomyces cerevisiae SNF5 interacts with the potential helicase hBRM. Nuc. Ac. Res. 23:1127-1132.

Ni Z., Olsen J.B., Guo X., Zhong G., Ruan E. D., Marcin E., Young P., Guo H., Li J., Moffat J., Emili A., Greenblat J.F. (2011). Control of the RNA polymerase II phosphorylation state in promoter regions by CTD interaction domain-containing proteins RPRD1A and RPRD1B. Transcription. 25:237-242.

Olave I., Wang W., Xue Y., Kuo A., Crabtree G.R.(2002). Identification of a polymorphic, neuron-specific chromatin remodeling complex. Gen. Dev. 16:2509-2517.

Reyes J.C., Barra J., Muchardt C., Camus A., Babinet C., Yaniv M. (1998). Altered control of cellular proliferation in the absence of mammalian brahma (SNF2α). EMBO J. 17:6979-6991.

Schmidt E.R., Pasterkamp R.J., van den Berg L.H. (2009) Axon guidance proteins: novel therapeutic targets for ALS ? Progr. Neurobiol. 88:286-301.

Tappe, A. and Kuner, R. (2006) Regulation of motor performance and striatal function by synaptic scaffolding proteins of the Homer1 family. Proc. Natl. Acad. Sci. U. S. A. 103, 774–779

Takanaga H., Mackenzie B., Peng J., Hediger M.A. (2005). Characterization of a branched-chain amino-acid transporter SBAT1 (SLC6A15) that is expressed in human brain. Biochem. Biophys.Res. Comm.337:892-900.

Wu Y., Zhang Y., Zhang H., Ren F., Liu H., Zhai Y., Jia R., Yu J., Chang Z. (2010). p15RS attenuates Wnt/Beta-Catenin signaling by disrupting Beta-Catenin/TCF4 interaction. J. Biol. Chem. 285:34621-34631.

Yang G.Y., Luo Z. Implication of Wnt signaling in neuronal polarization. (2010). Dev Neurobiol., 18:495-507.

Yang M., Sun Y., Ma L., Wang C., Wu J., Bi A., Liao J. (2011). Complex alternative splicing of the Smarca2 gene suggests the importance of smarca2-B variants. Journal Canc. 2:386-400

Zraly, C.B. and Dingwall, A.K. (2012) The chromatin remodeling and mRNA splicing functions of the

Brahma (SWI/SNF) complex are mediated by the SNR1/SNF5 regulatory subunit. Nucleic Acids Res, 40, 5975-5987.