THESIS ON NATURAL AND EXACT SCIENCES B199

Alternative Splicing in Health and Disease

ANNA KAZANTSEVA

TUT PRESS

TALLINN UNIVERSITY OF TECHNOLOGY
Faculty of Science
Department of Gene Technology

This dissertation was accepted for the defense of the degree of Doctor of Philosophy in Gene Technology on November 11, 2015.

Supervisor: **Senior Research Scientist Kaia Palm, PhD**
Department of Gene Technology
Tallinn University of Technology, Estonia
Protobios LLC, Estonia

Reviewed by: **Research Scientist Indrek Koppel, PhD**
Department of Gene Technology
Tallinn University of Technology, Estonia

Opponents: **Professor Eugene Makeyev, PhD**
MRC Centre for Developmental Neurobiology, King's College London, United Kingdom

Senior Research Scientist Ana Rebane, PhD
Institute of Biomedicine and Translational Medicine
Faculty of Medicine
University of Tartu, Tartu, Estonia

Defense of the thesis: January 15, 2016

Declaration:
Hereby I declare that this doctoral thesis, my original investigation and achievement, submitted for the doctoral degree at Tallinn University of Technology has not been submitted for any academic degree.

/Anna Kazantseva/

ISSN 1406-4723
ISBN 978-9949-23-874-3 (publication)
ISBN 978-9949-23-875-0 (PDF)

LOODUS- JA TÄPPISTEADUSED B199

Alternatiivne splaising haiges ja terves koes

ANNA KAZANTSEVA

TTÜ
KIRJASTUS

CONTENTS

ORIGINAL PUBLICATIONS

I Pruunsild, P.∗, **Kazantseva, A**.∗, Aid, T., Palm, K., Timmusk, T. (2007). Dissecting the human BDNF locus: bidirectional transcription, complex splicing, and multiple promoters. Genomics, 90(3), 397-406

II **Kazantseva, A**.∗, Sepp, M.∗, Kazantseva, J., Sadam, H., Pruunsild, P., Timmusk, T., Neuman, T., Palm, K. (2009). N-terminally truncated BAF57 isoforms contribute to the diversity of SWI/SNF complexes in neurons. Journal of Neurochemistry, 109(3), 807-818

III Lend, AK.∗, **Kazantseva, A**.∗, Kivil, A., Valvere, V., Palm, K. (2015). Diagnostic significance of alternative splice variants of REST and DOPEY1 in the peripheral blood of patients with breast cancer. Tumour Biology, 36(4), 2473-80

∗ Equal contribution

INTRODUCTION

Alternative splicing is one of the major mechanisms in the eukaryotic cells, which significantly modulates proteomic diversity and regulates the expression of most genes. This mechanism specifies how different exonic sequences are combined within a single gene to generate multiple mRNAs with distinct functions. It is established by genome-wide sequencing-based analyses that more than 95% of the mammalian multi-exon genes undergo alternative splicing, providing a source of huge protein diversity (Johnson et al., 2003; Pan et al., 2008).

Alternative splicing was first discovered about 4 decades ago and since then it was shown to be a common feature of all metazoans with the highest abundance and complexity found in primates (Chow et al., 1977; Berget et al., 1977; Barbosa-Morais et al., 2012). Highly specialized organs like brain, testis, heart and skeletal muscle have the highest levels of alternatively spliced transcripts (Yeo et al., 2004).

The mechanism of splicing is extremely complex, where multiple interactions between nascent mRNAs, small nuclear RNAs, and splicing factor proteins are involved. Proper cell functioning requires the accurate regulation of splicing by *cis*-acting regulatory elements and *trans*-acting protein factors in response to external stimuli (Luco et al., 2010; Kornblihtt et al., 2013).

Three processes of eukaryotic gene expression - chromatin remodeling, transcription, and splicing are highly interconnected in the cell. Splicing and transcription are mutually coordinated and reciprocal with transcription factors and response elements in promoters regulating splicing, and *vice versa,* the factors involved in maturation and processing of RNA influencing transcription. This functional interplay is essential for the multilayered regulation of gene expression.

Alternative splicing is a determining factor in a wide range of cellular physiological processes and developmental programs, including cell growth and differentiation, tissue morphogenesis, immune system specialization, neuronal development and death (Chen & Manley, 2009).

Disruption of normal regulation of splicing and consequent aberrant expression of proteins has been found to cause severe human diseases, including disabling neurological conditions, autoimmune disorders, and cancer. This explains the multiplicity of study programs on the use of reprogramming of the aberrant splicing for therapeutical purposes.

In my thesis, the major focus is on analysis of the role of alternative splicing in a tissue-specific gene expression, where two genes are chosen for detailed characterization. Another emphasis is on the diagnostic and prognostic significance of alternatively spliced variants in tumors.

ABBREVIATIONS

AON – antisense oligonucleotide
AS – alternative splicing
ASV – alternative splicing variant
BAF – BRM/BRG1-associated factor
BAF57 - BRM/BRG1-associated factor of 57 kDa
BC – breast cancer
BCL2L1 - B-cell lymphoma 2-like 1
BDNF – brain-derived neurotrophic factor
BPS – branch-point sequence
BRCA – breast cancer
BRG – Brahma-related gene
BRM – Brahma (protein)
CaRF – calcium response factor
CCND1 – cell cycle regulator cyclin D1
CDK – cyclin-dependent kinase
CELF – CUG-BP ELAV-like family member
CNS – central nervous system
CREB - cAMP response element-binding protein
CTD – carboxy-terminal domain (of RNA polymerase II)
EGFR – epidermal growth factor receptor
ESE – exonic splicing enhancer
ESS – exonic splicing silencer
EST – expressed sequence tag
FGFR - fibroblast growth factor receptor
FOX – forkhead box
GABA - gamma-aminobutyric acid
GTF – general transcription factor
HMG – high mobility group
hnRNP – heterogeneous nuclear ribonucleoprotein
ISE – intronic splicing enhancer
ISS – intronic splicing silencer
LABC – locally advanced breast cancer
LTP – long-term potentiation
mAb – monoclonal antibody
MAPK – mitogen-activated protein kinase
MeCP - methyl CpG binding protein
MED – mediator complex subunit
Met - methionine
miRNA – microRNA

N-BAF57 – neuronal BRM/BRG1-associated factor of 57 kDa
NMD – nonsense-mediated decay
NOVA – neuro-oncological ventral antigen
nPTB – neural polypyrimidine tract-binding protein
NRSE - neuronal-restrictive silencing element
NRSF - neuron-restrictive silencer factor
nSR100 – neural-specific SR-related protein of 100 kDa
NT – neurotrophin
ORF – open reading frame
PIC – pre-initiation complex
PNS – peripheral nervous system
PRP(F) – pre-mRNA processing factor
PSF – PTB-associated splicing factor
PTB(P) – polypyrimidine tract-binding protein
REST - RE1-silencing transcription factor
RRM – RNA recognition motif
RS domain – arginine/serine repeat-containing domain
SHH – sonic hedgehog
SMA – spinal muscular atrophy
sRNA – small RNA
siRNA – small interfering RNA
SMARCE1 - SWI/SNF related matrix-associated actin-dependent regulator of chromatin, subfamily E, member 1
SMN – survival of motor neuron
snRNP - small nuclear ribonucleic protein
SR protein – serine/arginine-rich protein
SRSF – serine/arginine-rich splicing factor
SSO – splice-switching oligonucleotide
SWI/SNF - SWItch/Sucrose Non-Fermentable
TAF – TATA-box binding protein associated factor
TFIIA(-B,-D,-E,-F,-H) – transcription factor IIA(-B,-D,-E,-F,-H)
TBP – TATA-box binding protein
TNC – tenascin C
TrkB - tropomyosin receptor kinase B
VEGF-A - vascular endothelial growth factor A
VEGFR - vascular endothelial growth factor receptor
WNT– *wingless integrated*

1. REVIEW OF THE LITERATURE

1.1. The basics of alternative splicing (AS)

Alternative splicing (AS) is a process by which a single gene, as a result of combinatorial use of different exons, enables the production of a variety of mRNA transcripts and proteins, thereby enhancing transcriptomic and proteomic diversity. About 24000 genes of human genome are estimated to be translated into millions of protein isoforms (Kobeissy et al., 2014). More than 95% of human genes undergo AS (Modrek & Lee, 2002; Wang et al., 2008). In addition, AS has a regulatory role in defining biological properties of the proteins, such as binding capacity, subcellular localization, protein-protein interaction, enzymatic activity, and stability in tissues (Smith et al., 1989). AS is a crucial regulatory mechanism in stem cell renewal and differentiation, organ morphogenesis, immune system and neural development and maturation (Gamazon & Stranger, 2014). It is estimated that the highest frequency of AS occurs in genes coding for the transcription factors, mostly in the gene family encoding Krüppel-associated box domains of Zn-finger proteins (Scarpato et al., 2015).

The basic splicing reaction is catalyzed by the huge multiprotein conglomerate, called spliceosome, which consecutively removes non-coding intronic sequences by two sequential phosphodiester transfer reactions to form mature mRNA prior to its export and translation (Zhou et al., 2002; Hoskins & Moore, 2012; Matera & Wang, 2014). In higher eukaryotes, the core RNA splicing signals recognized by spliceosome include a canonical 5' splice site with a conserved GU dinucleotide and 3' splice site with conserved AG dinucleotide, ~10 basepair long polypyrimidine tract (PPT) at the 3'end of the intron, and the branch point sequence (BPS) with highly conserved adenosine ~10-50 nt upstream of the 3' splice site. All these splice elements together define intron-exon boundaries (Mount, 1982; Matera & Wang, 2014). However, a vast majority of splice sites diverge significantly from the canonical sequences and are thus called "weak" splice sites. The nature of splice sites determines their affinity for spliceosome and the final splicing outcome (Kornblihtt et al., 2013).

There are several major modes of AS described, including exon skipping, usage of alternative 3' or 5' splice sites, intron retention, and intra-exon splicing. In some cases, a combination of different patterns of splicing may take place in a single pre-mRNA (Keren et al., 2010). In higher eukaryotes, the most prevalent event of AS is exon skipping, which occurs in up to 40% of all cases and contributes the most to proteome diversity and phenotypic

complexity, whereas the less frequent splicing mode is intron retention, accounting less than 5% of all splicing events (Gamazon & Stranger, 2014).

It is estimated that roughly a third of alternative splicing events results in mRNA isoforms with a premature stop-codon in their reading frame, which leads to the nonsense-mediated decay (NMD) and later degradaion of truncated products in the cytoplasm (Lewis et al., 2003). Recent studies suggest that NMD, in addition to controlling the splicing errors, is also involved in the global regulation of gene expression levels (Mendell et al., 2004; Hamid & Makeyev, 2014). As an example, the expression of many splicing factors, such as SR proteins and hnRNPs, as well as other core spliceosomal proteins, is mainly regulated via alternative splicing coupled with NMD, in a process known as regulated unproductive splicing and translation (RUST). This autoregulation method of genes is peculiar to the proteins of splicing or translation machinery and is evolutionarily conserved from plants to humans (Lareau et al., 2007; Lareau & Brenner, 2015).

1.2. The general mechanism of AS regulation

AS is a very complicated process, which is dependent on many factors, such as environmental signals, signaling pathways, chromatin structure, mRNA transcription, pre-mRNA secondary structure, length of exons and introns, sequence features (strong or weak splice sites), combinatorics of various regulatory factors, and interactions stemming from this (Barash et al., 2010).

At the local level, in addition to core splicing signals, there are multiple intronic and exonic *cis*-acting elements as well as *trans*-acting factors, which regulate the proper outcome of AS. *Cis*-regulatory elements are classified into four different groups depending on their role and position on pre-mRNA - ESEs (exonic splicing enhancers), ESSs (exonic splicing silencers), ISEs (intronic splicing enhancers), and ISSs (intronic splicing silencers) (Figure 1). Typically, these elements are located not far from exon-intron boundaries, mainly in a single-stranded conformation. They are highly variable and their activities are dependent on positions relative to the regulated exon (Matera & Wang, 2014). They recruit and directly interact with *trans*-acting factors, which control splicing efficiency by activating or repressing the splice site recognition and spliceosome assembly (Kornblihtt et al., 2013).

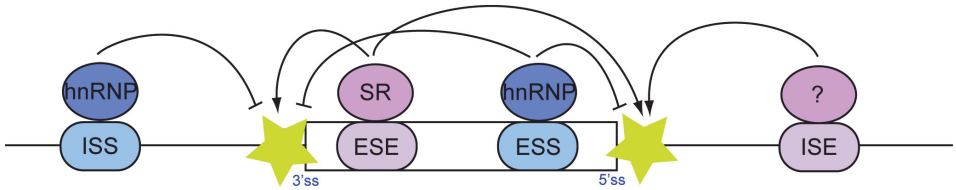

Figure 1. *Cis*-acting regulatory elements (ISSs, ISEs, ESEs and ESSs) and *trans*-acting factors (SRs, hnRNPs, and others) involved in the regulation of AS. General positive splicing factors, including SR proteins, recognize ESEs to promote splicing. General negative splicing factors, including hnRNPs, recognize ISSs and ESSs to inhibit splicing. Regulatory factors that interact with ISEs are unknown. The exon is shown as a box, the introns as lines. The splicing site is marked as a star. ESE - exonic splicing enhancer, ESS - exonic splicing silencer, ISE - intronic splicing enhancer, ISS - intronic splicing silencer. SR - serine/arginine-rich proteins, hnRNPs - heterogeneous nuclear ribonucleoproteins.

Trans-acting factors act as auxiliary RNA-binding splicing regulators, which assist the spliceosome in finding the correct splice sites and regulate the exon usage. There are numerous RNA-binding proteins that function to control and regulate AS. The general splicing factors include SR proteins (serine/arginine-rich proteins) and hnRNPs (heterogeneous nuclear ribonucleoproteins), which are widely expressed and conserved throughout the eukaryotes. These modular factors act in both constitutive and AS regulation (Chen & Manley, 2009).

The SR proteins belong to a well-characterized family of positive regulators. All the members of this family have a common structure with two important functional domains – one or two RNA recognition motifs (RRMs) at the amino-terminus, which provide RNA-binding specificity, and arginine/serine repeat-containing domain (RS domain) at the carboxy-terminus, which is important for protein-protein interactions with other splicing factors. Via these domains, SR proteins are able to bind simultaneously RNA and other regulatory proteins (Black, 2003; Shepard & Hertel, 2009). The serines of the RS domain can be extensively phosphorylated, which influences their protein-protein and protein-RNA interactions and modulates their function (Xiang et al., 2013). In humans, there are twelve canonical SR proteins (SRSF1-SRSF12) and several SR-related polypeptides, that either have different or no RNA-binding domains, or unable to assist in the splicing process (Chen & Manley, 2009; Busch & Hertel, 2012).

In opposition to SR proteins, the hnRNP proteins are the best-studied negative regulators of splicing. However, these proteins do not belong to a single family. Most of the hnRNP proteins also contain several functional domains – RRM domain and several domains that support protein-protein

interactions, including RGG box motif (a cluster of arginine and glycine residues), glycine-rich, acidic, or proline-rich domains (Busch & Hertel, 2012). There are more than 20 kinds of hnRNP proteins (ranging from hnRNP A to hnRNP U) found in human cells. The most abundant and most-studied hnRNPs are the six "core" hnRNP proteins - A1, A2, B1, C1, B2 and C2 as well as PTB (Polypyrimidine Tract-Binding Protein), also known as hnRNP I (Dreyfuss et al., 1993; Black, 2003).

Most often, SR proteins recognize and interact with ESEs through their RRMs, whilst the targets of hnRNPs are usually ESSs and ISSs. The interaction partners for ISEs are not sufficiently described (Figure 1) (Chen & Manley, 2009). Thus, multiple SR protein family members are normally believed to act as splicing activators by recruiting spliceosome and other positive regulatory factors to the splice sites of pre-mRNA, resulting in inclusion of exons (Lam & Hertel, 2002). On the contrary, hnRNP proteins are mostly known as negative regulators of splicing that cause exon skipping. Their proposed mechanism of action is cooperative binding along exons to the transcript to prevent the recruitment of spliceosome complex. PTB proteins can bind specifically to short polypyrimidine tract (CU-rich motifs) near the 3' end of introns and cause downstream exon exclusion (Coutinho-Mansfield et al., 2007). By binding to ESEs, PTB can cause exon skipping by inhibiting formation of the splicing complex (Matera & Wang, 2014).

However, recent findings indicate that action of both types of regulators is modulated by their relative abundance, post-translational modifications (*e.g.* phosphorylation status), location of the binding site relative to the regulated exon, and affinity of the binding sites. Depending on the context, both types of proteins can either facilitate or prevent spliceosome assembly at particular splice sites (Lin & Fu, 2007; Huelga et al., 2012; Pandit et al., 2013).

To conclude, spliceosome recruitment, binding affinity and splicing events are influenced by activatory and inhibitory proteins, which recognize and bind to specific sets of RNA elements, including splice site signals and SR/hnRNP protein binding sites.

1.3. The epigenetic regulation of AS

Traditionally, the regulation of AS has been thought to include mostly RNA-binding splicing factors. However, the most recent studies have demonstrated that the control of splice site choice is far more complex and efficiency and accuracy of AS are dependent on the cellular environment, chromatin structure, histone modifications, DNA topology and transcriptional status (Figure 2) (Luco et al., 2011).

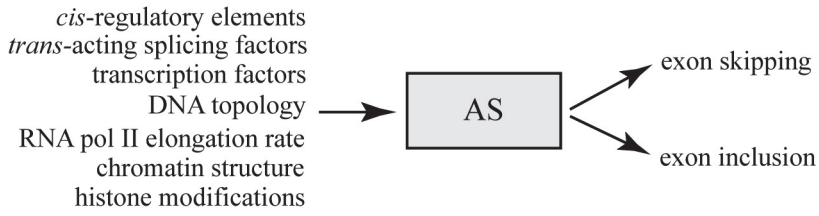

Figure 2. Factors involved in the global regulation of alternative pre-mRNA splicing (AS) leading to exon inclusion or skipping.

1.3.1. Coupling of splicing and transcription

It is estimated that AS of about 80% of all mammalian genes occurs co-transcriptionally (Brugiolo et al., 2013). In other words, the decison of whether alternative exon is skipped or inserted is usually made during ongoing transcription, or simultaneously with transcription, within the same cell nuclear environment (Pandya-Jones & Black, 2009). Moreover, these two events are coupled not only temporally and spatially, but also functionally (Beyer & Osheim, 1988; Ameur et al., 2011; Kornblihtt et al., 2013; Bentley, 2014). This implies that most of the factors that are involved in the control of transcription also impact splicing, and this process is mostly reciprocal. Namely, the spliceosomal components and *trans*-acting splicing factors were shown to facilitate transcription through interactions with transcription factors (Fong & Zhou, 2001; Lin et al., 2008).

Two general molecular mechanisms have recently been proposed that physically and functionally couple transcription with splicing, known as "recruitment coupling" and "kinetic coupling" (Montes et al., 2012). The first model postulates that the recruitment of splicing factors to pre-mRNA works in association with transcription machinery. The CTD (Carboxyl Terminal Domain) of RNA polymerse II plays the major role in this process acting as a target for the recruitment of processing factors to the regions of active transcription (McCracken et al., 1997; Fong et al., 2003). In mammals, CTD is composed of up to 52 copies of heptapeptide sequence (Tyr-Ser-Pro-Thr-Ser-Pro-Ser) (Meininghaus et al., 2000). Post-translational modifications of CTD residues, in particular phosphorylation at two specific serine residues (Ser2 and Ser5), are very important for its interaction with the components of transcription machinery and spliceosome to generate the elongation-ready RNA pol II (Millhouse & Manley, 2005; Montes et al., 2012). Thus, the recruitment of splicing factors to proper splice sites is dependent on the RNA pol II CTD.

Multiple regulators of AS that associate with CTD and participate in coupling have been found in both yeast and mammalian cells. The most important of these are SR proteins, which associate with both RNA pol II CTD and other splicing factors and spliceosome components (Millhouse & Manley, 2005). Amongst the other factors involved in the process of co-transcriptional splicing by direct interactions with CTD are splicing factors U2AF65 (U2 Auxiliary Factor 65 kDa subunit) and PRP19 (Pre-mRNA Processing Factor 19), transcription elongation factor CA150 (Coactivator of 150 kDa), transcriptional regulators MED23 (Mediator complex subunit 23), TFIIH (Transcription Factor IIH), PSF (PTB-associated Splicing Factor), and others (Morris & Greenleaf, 2000; Smith et al., 2004; Endoh et al., 2004; O'Gorman et al., 2005; Lin & Fu, 2007; David et al., 2011; Ji & Fu, 2012).

The second, less direct "kinetic coupling" model suggests that the outcome of co-transcriptional RNA splicing and the ability of splicing factors to recognize splice sites is modulated by RNA pol II elongation rate. The rate, by which RNA pol II transcribes, is highly dynamic. It varies between and within the genes and depends on multiple factors - the chromatin structure (histone marks), DNA topology, the phosphorylation level of Ser2 and Ser5 residues on CTD of RNA Pol II, elongation factors, nucleosome remodelers, and even on number of exons per gene. For example, the low-complexity DNA sequence and low GC-content are associated with higher elongation rate (Eperon et al., 1988; Jonkers & Lis, 2015). Or, specific histone marks, like H3K36me3 (Histone H3 trimethylated at lysine 36) and H4K20me1 (Histone H4 monomethylated at lysine 20), can reduce elongation rate (Jonkers & Lis, 2015). The rate of transcription synthesis dictates the inclusion or skipping of exons. The fast elongation speed of RNA Pol II, which occurs when *e.g.* CTD is hyperphosphorylated or chromatin has a low nucleosome density at the site of transcription, usually leads to alternative exon skipping or inclusion of exons with "strong" splice sites. And *vice versa,* a slow elongation rate, caused *e.g.* by hypophosphorylation of CTD or high nucleosome density, allows the splicing machinery to assemble on "weak" splice sites, resulting in exon inclusions (Mata et al., 2003; Pandya-Jones & Black, 2009; Montes et al., 2012; Fong et al., 2014; Dujardin et al., 2014; Bentley, 2014). Thus, the optimal rate of mRNA synthesis is required to achieve a proper co-transcriptional splicing (Fong et al., 2014).

However, not always transcription and splicing occur together. In some situations, for example after the induction of transcription in macrophages or repression of splicing in thrombocytes, splicing rather occurs after the process of transcription is completed. Alternative introns that are spliced post-transcriptionally might be possibly coordinated by different regulatory

mechanisms. This field of research certainly needs future investigations (Brugiolo et al., 2013).

1.3.2. AS and chromatin environment

The most recent findings reveal that selection of exons is also regulated at the level of chromatin structure.

Nucleosomes, as subunits of chromatin, are composed of two copies of histone proteins H2A, H2B, H3 and H4, around which approximately 147 base pairs of double-helical DNA is wrapped (Luger et al., 1997). In a variety of species, nucleosomes are found to be associated with DNA sequence preferentially around exons, especially alternative exons, and their positioning is highly dynamic. Presumably, such a specialized nucleosome positioning is regulated by transcription machinery as well as chromatin remodeling complexes to help in splice site recognition (Schwartz et al., 2009). In a standard model, increased nucleosomal density at internal cassette exons is associated with higher levels of exon inclusion, whereas weaker nucleosome positioning is generally associated with exon skipping. Thus, nucleosome positioning and density not only influence exon definition but also contribute to the efficiency of splicing (Mata et al., 2010).

Additionally, the outcome of splicing is directly modulated by the post-translational modifications of histones. The histones located in a close proximity to intron-exon boundaries are usually modified, generally by methylation, acetylation, phosphorylation, ribosylation, and ubiquitination, which adds an extra mark for exon definition and spliceosome positioning (Luco et al., 2010; Montes et al., 2012; Zhou et al., 2014). For instance, tri-methylation of histones H3K36, H3K79, H3K4 and di-methylation of histone H3K27 have shown to be epigenetic splicing-associated marks for internal exons (Andersson et al., 2009; Kim et al., 2011). Depending on the histone code, the rate of RNA pol II elongation can be repressed or increased, which, in turn, determines the outcome of alternative exon usage. Thus, the acetylation of histones located next to the cassette exons has been shown to destabilize histone-DNA interactions leading to chromatin relaxation and enhanced elongation, which promotes exon skipping (Almeida & Carmo-Fonseca, 2012).

The chromatin remodeling SWI/SNF (mating-type SWItch/Sucrose NonFermenting) complex has been shown to be involved in the regulation of AS, independently of its remodeling activity. Several studies reported that hBRM (human Brahma), and possibly BRG-1 (Brahma-Related Gene 1), the two mutually exclusive catalytic subunits of SWI/SNF complex, can facilitate the assembly of the spliceosome to exons with “weak” splice sites and promote the inclusion of variant exons during the splicing of a subset of pre-mRNAs by

reducing the elongation rate of RNA pol II (Batsché et al., 2006; Ito et al., 2008). However, later studies in *Drosophila* revealed that the elongation rate of RNA pol II is not always affected and that BRM together with BRG-1 are not the only subunits of SWI/SNF involved in the regulation of splicing (Waldholm et al., 2011).

Several recent studies have implicated that certain small non-coding RNA molecules, namely siRNAs (small interfering RNAs) and sRNAs (endogenous small RNAs) as well as long non-coding RNAs and pre-mRNAs, can specifically guide histone-modifying complexes to distinct and specific sites and regulate their activity to modulate the proper local chromatin structure providing accurate control of AS (Ameyar-Zazoua et al., 2012; Zhou et al., 2014).

In conclusion, the studies performed during the past few years suggest that synergistic or competing activities of positively and negatively acting *cis*-elements and *trans*-acting factors along with bimodel interactions between splicing, transcription and chromatin structural subunits form a highly complex multi-layered splicing regulatory network in the cell. Despite of these findings, further analysis is needed to underastand the global rules that define the regulation of AS in multicellular organisms.

1.4. Tissue- and cell type-specific AS

AS has a central role in generating morphological differences between species and developing of cell type- or tissue-specific molecular repertoire. Indeed, cell type- or tissue-specificity of splicing is a necessary prerequisite for development, differentiation or organs morphogenesis (Barbosa-Morais et al., 2012; Merkin et al., 2012). It is estimated that about 10-30% of all the alternatively spliced genes have tissue-specific transcripts (Xu et al., 2002). Brain, followed by the liver, muscles and testis have the maximum abundance of AS events across the whole human body, mostly with exon skipping mode (Yeo et al., 2004; Shargunov et al., 2014).

Tissue-specific genes can be dividied into two separate classes – regulators, or genes participating in the transcriptional regulation of target genes, and effectors, or genes involved in the modulation of physiological properties characteristic to certain types of cells (Choi et al., 2006). In turn, tissue-specific regulators can be divided into several subgroups - extracellular signaling proteins, transmembrane receptors, signal transducers, and cellular transcription factors (Choi et al., 2006). It is established that tissue-specific AS is very common to the genes, which are involved in multiple signaling pathways, such as EGF/MAPK (Epidermal Growth Factor/Mitogen-Activated Protein Kinase), Hedgehog, WNT (*Wingless iNTegrated*), TGF-b

(Transforming Growth Factor beta), and JAK/STAT (Janus Kinase/Signal Transducer and Activator of Transcription) pathways (Buljan et al., 2012).

It has been recently suggested that tissue-specific alternative exons usually encode structurally disordered regions with undefined structure, which are enriched in evolutionarily conserved binding motifs important for molecular interactions with other proteins as well as in sites of post-translational modifications. This is opposed to constitutive exons, which often tend to encode protein domains with defined tertiary structure (Buljan et al., 2012; Ellis et al., 2012).

Different tissues exhibit splice-specific programs as a result of combinatorial use of tissue-specific splicing factors. Thus, there are multiple regulatory proteins described that modulate AS in a spatio-temporal manner independent of the activity of auxiliary splicing factors (Wang et al., 2008; Merkin et al., 2012).

1.4.1. AS in neural tissue

AS is one of the major processes important for the maturation and proper function of the mammalian nervous system. Many neuron-specific proteins have diversity of isoforms generated through differential inclusion or exclusion of specific sequences into the final mRNA. Bioinformatics studies together with microarray analyses revealed that mammalian brain, especially fetal brain, is a tissue, where most of AS events take place, contributing to its transcriptome and proteome diversity (Blencowe, 2006). This can be explained by the fact, that brain contains the highest number of specialized cell types that are constantly modulated in response to various stimuli (Ameur et al., 2011). Within the brain, cerebellum was found to exhibit the highest degree of AS (Grange et al., 2010). Interestingly, some genes can show splicing differences within distinct brain segments indicating their specialized function in these regions (Grabowski & Black, 2001). About 40% of all the genes expressed in human brain show significant changes in splicing during lifespan, mainly at infancy and at a very old age. Most of these genes are connected to neuronal function and dysfunction and then associated with neurodegenerative disorders (Mazin et al., 2013).

AS plays an important role in the developing and mature nervous systems. Such diverse neuronal developmental events like neural tube patterning, neural progenitor cell fate determination, axon guidance, synaptogenesis, and synapse maintenance are all regulated by AS (Norris & Calarco, 2012). In the mature nervous system, splicing is known to regulate key physiological processes like synaptic strength and plasticity, ion channels functions, neurotransmission, calcium signaling, receptor specificity, neuronal

depolarization, and more (Grabowski & Black, 2001; Li et al., 2007; Kalsotra & Cooper, 2011).

AS in neurons is often modulated by physiological neuronal activity and stress. Increasing evidence suggests that calcium influx as a result of cellular depolarization or upon drug treatment is involved in the regulation of AS of many key neuronal transcripts encoding channel proteins, neurotransmitter receptors, or modulators of synaptic strength (Daoud et al., 1999). Recent comprehensive studies have found that neuronal depolarization can modulate AS of certain genes independently of calcium by changing transcription kinetics of RNA pol II through histone modifications in the proximity of regulated exon (Schor et al., 2009). In either case, changes in the external or internal stimuli regulate splicing of many neuronal genes often yielding in significantly altered functional activity of these proteins and thus contributing to molecular plasticity of the brain (Daoud et al., 1999; Flavell & Greenberg, 2008; Lee et al., 2009). AS of NMDAR1 (Neurotransmitter N-Methyl-D-Aspartate Receptor 1) is one of the best-described examples of how the regulated production of different protein isoforms modulates the receptor function, subcellular distribution and synaptic efficacy, which leads to the changes in cell physiology (Michaelis, 1993; Zukin & Bennett, 1995; Okabe et al., 1999). Another example is the complex expression of neurexins - major proteins of cell adhesion that establish and maintain synapses in the brain. Through the use of AS, thousands of neurexin protein isoforms can be expressed, each with its own function and affinity for their post-synaptic ligands (Chih et al., 2006).

The proper regulation of AS in neural tissue is very cruical for maintaining its normal functioning. Of note, brain has a large variety of differentially expressed splicing factor genes. There are several *trans*-acting neuron-specific or enriched splicing regulators described so far, including RNA-binding proteins and miRNAs, which are expressed in a temporal- or cell type-specific manner and act by enhancing or repressing splicing depending on the context (Li et al., 2007; Grabowski, 2011). The most-characterized tissue-restricted neuronal regulators are NOVA-1, -2 (Neuro-Oncological Ventral Antigen-1, -2), nSR100 (Neural-specific SR-related Protein of 100 kDa), nPTB (neural Polypyrimidine Tract-Binding Protein), FOX-1, -2, -3 (Forkhead Box Protein-1, -2, -3), CELF (CUG-BP Elav-Like Family), and MBNL2 (Muscleblind-Like 2) (Chen & Manley, 2009).

NOVA-1, NOVA-2, nPTB and nSR100 are the major splicing factors expressed in the mammalian brain that regulate more than 700 AS events (Zhang et al., 2010). It has been reported that transcripts with NOVA-regulated exons encode proteins, which are mostly involved in synaptic function and neuronal migration (Ule et al., 2005; Yano et al., 2010). NOVA-1 and NOVA-2 are paralogs, that are expressed in different subregions of the postnatal brain and

control diverse events of brain development, with NOVA-1 predominantely expressed in brainstem and spinal cord neurons, and NOVA-2 in neocortex and hippocampus (Jensen et al., 2000). Interestingly, NOVA binding to pre-mRNA determines the outcome of splicing. Namely, binding of NOVA upstream of alternative exon leads to its skipping, while binding to downstream site activates exon inclusion (Ule et al., 2006).

nPTB, also known as PTBP2 (Polypyrimidine Tract Binding Protein 2), is a neuron-restricted, or brain-enriched counterpart of PTB (Polypyrimidine Tract-Binding Protein), also known as PTBP1. Its expression is activated in neural tissue as a result of PTB suppression by brain-specific miRNAs (*e.g.* miR-124 in mice) (Makeyev et al., 2007). The switch from PTB to nPTB expression induces many events of neuron-specific AS that are important for neuronal maturation and differentiation (Coutinho-Mansfield et al., 2007; Grabowski, 2011; Li et al., 2014). nPTB acts by binding to CU-rich regulatory sequences of alternative exons to repress or activate their splicing, or sometimes to cause intron retention (Keppetipola et al., 2012). Neurons lacking this protein fail to develop correctly and die (Li et al., 2014).

nSR100 is another crucial regulator of up to 30% of all the AS events important for mammalian neuronal differentiation, which acts by increasing the levels of nPTBs and other factors. It also interacts directly with targeted genes, such as those involved in the remodeling of cytoskeleton during neuronal morphogenesis, promoting neuritic outgrowth or axon guidance (Calarco et al., 2009; Quesnel-Vallières et al., 2015). Additionally, nSR100 activates the expression of several target genes important for neurogenesis by modulating the splicing of REST (RE-1 Silencing Transcription factor), which acts as a repressor of neuronal genes in nonneural tissues. This is achieved by production of a truncated REST isoform in neural cells with no repressor activity. Respectively, in nonneural cells, REST itself maintains its repressor activity in part by inhibiting the activity of nSR100 (Raj et al., 2011).

1.4.2. BDNF as an example of tissue-specific gene expression

BDNF (Brain-Derived Neurotrophic Factor) along with NGF (Nerve Growth Factor), NT-3 (Neurotrophin-3) and NT-4/5 (Neurotrophin-4/5) belongs to the mammalian family of neurotrophic factors (Snider, 1994; Skaper, 2012). It was first identified and purified more than 30 years ago as a factor supporting the survival and fiber outgrowth of the embryonic sensory neurons (Barde et al., 1982). It is now well established that members of this family have a very high degree of sequence conservation between mammalian species and their key role is the regulation of differentiation, survival, plasticity, growth and synaptic transmission of certain neuronal populations of CNS (Lessmann et al., 2003). In

brain, BDNF has been shown to play a critical functional role in triggering neuronal activity-dependent long-term potentiation (LTP) in the hippocampus and neocortex, a process that is underlying learning and long-term memory formation (Korte et al., 1995; Soulé et al., 2006). Disruption of BDNF expression is associated with a set of human neurological and psychiatric conditions, including Alzheimer's disease, Parkinson's disease, Huntington's disease, epilepsy, schizophrenia, multiple sclerosis and depression (Howells et al., 2000; Ferrer et al., 2000; Binder et al., 2001; O'Bryant et al., 2009; Castrén & Rantamäki, 2010; Kalinowska-Lyszczarz & Losy, 2012; Nieto et al., 2013). BDNF is very crucial to postnatal survival, knockout mice that lack this neurotrophin die shortly after birth as a result of severe PNS and CNS deficiencies (Jones et al., 1994).

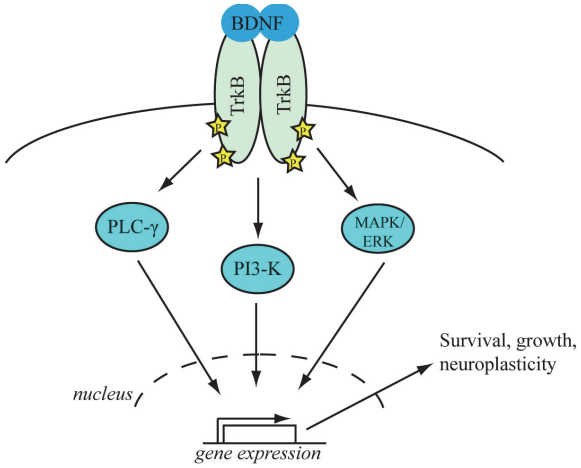

Figure 3. Simplified schematic representation of BDNF-TrkB signaling pathways. BDNF binds to the extracellular domain of TrkB with high affinity, induces its dimerization and autophosphorylation triggering the activation of three main intracellular signaling pathways: MAPK/ERK, PLC-γ and PI3-K, which mediate transcription of genes necessary for neuronal cells plasticity, survival, and growth. MAPK/ERK - mitogen-activated protein kinase/extracellular signal-regulated kinase, PLC-γ - phospholipase C-γ, PI3-K - phosphatidyl inositol-3 kinase. Stars indicate the sites of autophosphorylation.

BDNF is synthesized in the endoplasmic reticulum as a prepro-BDNF molecule of 32 kDa, which is further N-terminally cleaved either extracellularly or intracellularly to form a mature 14 kDa protein (Mowla et al., 2001; Park & Poo, 2013). Following secretion by target tissue (either constitutive or activity-induced), BDNF binds to and activates both its specific cell surface receptors - high-affinity TrkB (Tropomyosin Kinase B) receptor and low affinity $p75^{NTR}$

(p75 Neurotrophin Receptor) (Reichardt, 2006). BDNF binding results in receptor dimerization and autophosphorylation, leading to activation of various signaling cascades, including the MAPK/ERK (Mitogen-Activated Protein Kinase/Extracellular Signal-Regulated Kinase), PI3-K (Phosphoinositide 3-Kinase), and PLC-γ (Phospholipase C-γ) pathways, which further promote different aspects of neuronal survival, neurite outgrowth, neurogenesis, synapse formation, and plasticity (Figure 3) (Heerssen & Segal, 2002; Murray & Holmes, 2011; Park & Poo, 2013).

BDNF is one of the well-characterized examples of neuronal genes with a distinct pattern of tissue-specific AS. The initial studies on *BDNF* genomic organization and regulation have begun more than twenty years ago by T. Timmusk and colleagues, who first characterized in detail its complex structure and expression pattern in rodents (Timmusk et al., 1993). Since then it has become obvious that the structure of this gene is unusual, contains several upstream noncoding exons spliced to a common coding exon, resulting in multiple alternative mRNAs, which are regulated from separate alternative promoters (Timmusk et al., 1993). Later on, several other groups revised the structure of the *Bdnf* in rodents (Bishop et al., 1994; Aoyama et al., 2001; Liu et al., 2006). Since the finding of its complex genomic structure, the great interest to *BDNF* gene has been constantly growing because of its involvement in multiple behavioral, physiological and pathological processes in the brain. However, the most recent studies on the rodent *Bdnf* gene were conducted by neurobiology group supervised by T. Timmusk in 2007. They showed that the organization of both mice and rat *Bdnf* genes is very identical, both genes consist of nine exons - eight 5' untranslated exons transcribed from separate promoters upstream of each exon, and one major protein-coding 3' exon. Additionally, there are two different polyadenylation sites located within the last coding exon. All 5' noncoding exons are independently spliced to the major coding exon to produce at least 22 bipartite alternative transcripts with differential patterns of expression in various brain regions and nonneural tissues (Aid et al., 2007).

The structure and differential expression of the human *BDNF* gene have been less extensively investigated. However, similarly to rodent species, it has been found to have a rather complex organization with seven upstream exons each, which are independently spliced to the single protein coding exon to form diversity of *BDNF* alternative mRNAs, all of which initiated by several individual promoters and expressed differentially within the brain regions (Liu et al., 2005).

Numerous transcription regulatory elements controlling *BDNF* expression that are located in different regulatory regions of the gene have been demonstrated to define the spatiotemporal regulation of its expression. Namely,

binding sites for transcription factors like NRSF (Neuron-Restrictive Silencer Factor), CaRF (Calcium Response Factor), CREB (Calcium/cAMP Response Element binding protein), MeCP2 (Methyl CpG binding Protein 2), USF (Upstream Stimulatory Factor), BHLHB2 (Basic Helix-Loop-Helix B2), NFκB (Nuclear Factor κB), MEF2D (Myocyte Enhancer Factor 2D), NPAS4 (Neuronal PAS domain protein 4), and ARNT2 (Aryl hydrocarbon Receptor Nuclear Translocator 2) have been shown to act in specific promoter contexts as *cis*-acting elements that positively or negatively contribute to regulation of expression of corresponding alternative transcripts in various tissues (Timmusk et al., 1999; Tao et al., 2002; Chen et al., 2003; Martinowich et al., 2003; Zuccato et al., 2003; Lubin et al., 2007; Lubin et al., 2008; Jiang et al., 2008; Pruunsild et al., 2011). The regulation of transcription of the *BDNF* gene occurs predominantly through exons I and IV. The most sensitive to neuronal activity promoter is located upstream of *BDNF* exon IV, containing multiple calcium- and cAMP-responsive elements, to which CaRF and CREB, respectively, bind to initiate transcription (Zheng et al., 2011; Zheng et al., 2012; Park & Poo, 2013). In addition to multiple transcription factors, such epigenetic mechanisms like chromatin remodeling and DNA methylation are also involved in the regulation of isoform-specific *BDNF* transcription. For instance, it was shown that acetylation of histones H3 or H4 at *BDNF* promoter IV is upregulated following fear conditioning leading to the increased mRNAs expression, or DNA methylation in promoter IV following neural stimulation reduces *BDNF* exon IV transcription as a result of interaction with transcriptional repressor MeCP2 (Zhou et al., 2006; Bredy et al., 2007; Lubin et al., 2008).

The expression of *BDNF* mRNA and protein is usually low at birth and increases into adulthood, or in response to neural activity induced by various stimuli, such as GABAergic and glutamatergic neurotransmission (*e.g.* via kainate receptors), KCl treatment-dependent membrane depolarization, or drug administration (*e.g.* acute cocaine or morphine treatment) (Maisonpierre et al., 1990; Zafra et al., 1990; Liu et al., 2006; Aid et al., 2007). It has been shown that BDNF protein and mRNAs are expressed predominantly in the CNS, especially in the brain areas like hippocampus, cerebral cortex, amygdala, and hypothalamus (Hohn et al., 1990; Hofer et al., 1990; Liu et al., 2005; Liu et al., 2006; Aid et al., 2007). Moreover, the relative level of expression of alternative transcripts was shown to be differential not only between various tissues but also in the single cell (*e.g.* neuronal stroma, dendrites or axons). These findings reveal the highly precise functional role of BDNF controlled by compartment-specific selective expression (Leibrock et al., 1989; Hofer et al., 1990; Aid et al., 2007; Chiaruttini et al., 2008; Baj et al., 2011; Garcés et al., 2014). Outside of CNS, the expression of *BDNF* mRNA and protein has been detected at low

levels in thymus, heart, liver, skin, spleen, muscles, pancreas, colon, intestine, lungs, bladder, testis, and placenta.

1.5. Transcription pre-initiation complex and role of AS in generation of its diversity

During every round of transcription, about sixty proteins assemble together at every RNA pol II-regulated promoter. Half of these assembled proteins form transcription pre-initiation complex (PIC). PIC consists of RNA pol II and general transcription factors (GTF), which function together with co-activators (also known as Mediator) and chromatin-remodeling complex to form the huge multi-subunit conglomerate important for the first-step gene expression control in eukaryotes (Conaway & Conaway, 1993; Kornberg, 2007; Murakami et al., 2013).

Previously, the components of general transcription machinery, including core promoter recognition complexes, Mediator complex, and chromatin modifiers were thought to be mostly conserved and invariant in all tissues from yeast to human. However, recent studies support the notion that eukaryotes have evolved an unexpected diversification of GTFs to modulate the complicated process of cell-type or tissue-specific gene expression. Different genes and their promoter elements recruit unique combinations of GTFs. Furthermore, various protein isoforms of PIC components, which are generated by AS, or paralogs, which emerged during evolution, are involved in the formation of highly context-specific transcription complexes. Thus, differentially expressed members of the PIC family are the primary drivers of highly complex patterns of gene expression (D'Alessio et al., 2009).

1.5.1. GTFs and their alternative isoforms

Initiation of eukaryotic transcription requires the assembly of RNA pol II and a set of GTFs, including TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH on the core promoter. Formation of this complex is necessary for promoter recognition, polymerase recruitment, interaction with regulatory proteins, and DNA unwinding. Each GTF has a specialized role during transcription initiation. Thus, TFIIA and TFIIB are necessary for TBP-TATA complex stabilization, TFIIB is involved in RNA pol II/TFIIF recruitment and start-site selection, TFIIE recruits TFIIH and facilitates the formation of the complex and promoter melting, TFIIF facilitates the recruitment of polymerase and other GTF components to the promoter and enhances the efficiency of elongation, and TFIIH has several enzymatic activities important for transcription initiation and

elongation (Thomas & Chiang, 2006; Luse, 2013; Grünberg & Hahn, 2013). TFIID is a major multisubunit component of the general transcription machinery, which is composed of TBP (TATA-Binding Protein) and approximately 14 coactivator subunits called TAFs (TBP-Associated Factors) (Lemon & Tjian, 2000; Thomas & Chiang, 2006; Papai et al., 2011). TFIID has several roles in regulating transcription, including promoter recognition, pre-initiation complex assembly and stabilization, and interaction with multiple transactivators (Albright & Tjian, 2000; Papai et al., 2011).

There are multiple GTF paralogs identified, including gonad-specific TAF1L, TAF4b, TAF7L, and ALF (TFIIA paralog) (Freiman, 2009). However, the GTF isoforms derived by AS are far less studied. AS yielding in GTF subunits is usually modulated by external stimuli. Up to date, alternatively spliced isoforms of TAF1, TAF4 and TAF6 have been described (Albright & Tjian, 2000).

Multiple alternatively spliced isoforms of TAF1, the largest and one of the most essential subunits of TFIID, have been found in eukaryotes. The studies in *Drosophila* reported that all identified isoforms of TAF1 have differential DNA-binding affinity and patterns of expression during development (Katzenberger et al., 2006; Katzenberger et al., 2009). The expression of one of *Drosophila* TAF1 isoforms was found to be enriched in testis and therefore might be involved in the regulation of genes necessary for sexual reproduction (Metcalf & Wassarman, 2007). In mammals, alternative isoforms of TAF1 are highly distinctly expressed across tissues. Neuron-specific isoforms of TAF1 have been described with expression limited to brain only. It was observed that reduced expression of neuron-specific TAF1 causes X-linked dystonia-parkinsonism syndrome (XPD) with neuronal degeneration in striatum (Herzfeld et al., 2007; Sako et al., 2011; Jambaldorj et al., 2012). Thus, neuron-specific isoform of TAF1 seems to be important for survival of striatal neurons.

TAF4 is another member of TFIID complex, whose function is regulated by alternative splicing. TAF4 is primarily known as a co-activator of many transcription factors (Wright et al., 2006). The studies in mice identified five alternatively spliced isoforms of TAF4 with context-specific roles in transcriptional regulation (Brunkhorst et al., 2004). In humans, more than 10 different alternative isoforms of TAF4 have been described with different patterns of expression in various cells and tissues. Of note, one of the alternative isoforms of human TAF4 with structurally altered co-activator TAFH domain, which is important for protein-protein interactions, was shown to be involved in the regulation of mesenchymal stem cell proliferation and differentiation into osteoblasts, adipocytes, and chondrocytes through activation of non-canonical WNT signaling (Kazantseva et al., 2013). In other studies, alternative TAF4

isoforms with modified TAFH domain were shown to be involved in the process of neuronal and glial differentiation from neural progenitor cells, emphasizing the importance of TAF4 involvement in the regulation of neurogenesis (Kazantseva et al., 2015). Unlike canonical TAF4 supporting the pluripotency of the cells, alternative isoforms of TAF4 promote differentiation and control motility.

Alternative isoform of human TAF6 protein, called TAF6δ, is quite well characterized due to its function in apoptosis. TAF6δ isoform has a partial histone-fold motif, which is critical for its interactions with TAF9. In contrast to the canonical TAF6 protein, TAF6δ forms a TAF9 depleted general transcription complex, known as TFIIDπ complex, which is specific to the cells undergoing the programmed cell death. Several studies have confirmed the signal-controlled regulation of TAF6δ in pro-apoptotic cells and its impacts on the regulation of genes triggering apoptosis (Bell et al., 2001; Wilhelm et al., 2010). Although the molecular mechanisms underlying the involvement of TAF6δ in the regulation of pro-apoptotic genes remain to be defined, it has been shown to orchestrate transcription in a promoter-dependent manner (Wilhelm et al., 2010; Kamtchueng et al., 2014).

1.5.2. Mediator complex and its alternative isoforms

Mediator is a large, up to 30 polypeptides containing complex, which plays a central role during the initiation of transcription by RNA pol II in all eukaryotes. According to the nomenclature established in 2004, it consists of MED1 to MED31 proteins, together with CDK8 (Cyclin-Dependent Kinase 8), cyclin C and several paralogs (MED1L, MED12L, MED13L and CDK19) (Bourbon et al., 2004). Mediator complex regulates transcription by transmitting signals from DNA-binding transcription factors to the RNA pol II transcription machinery, resulting in the activation or repression of gene expression in a promoter-specific fashion. In addition to its basic function, recent studies have also established its additional roles in processes of transcription elongation, epigenetic regulation, transcription termination, and mRNA processing. Based on these multiple specific roles in gene regulation, Mediator can be considered as a key coordinator of cellular developmental programs (Poss et al., 2013; Yin & Wang, 2014). The structural organization of the Mediator complex is largely conserved among eukaryotes from yeast to humans. Its numerous subunits can be divided into 4 separate modules designated as head, middle, tail and CDK8 kinase. However, the sequence conservation of the distinct subunits between species is rather low (Malik & Roeder, 2010; Ansari & Morse, 2013; Yin & Wang, 2014).

Recent studies in mammals showed that the subunit composition of the Mediator complex varies and is not restricted to single subunits. There are at least two main distinct Mediator complexes discovered, which differ in composition by the presence or absence of the CDK8 kinase module (Malik & Roeder, 2010). In addition, Mediator complex variations can also arise from the incorporation of subunit paralogs (*e.g.* MED13L instead of MED13, or CDK11 instead of CDK8) or presence/absence of certain subunits (*e.g.* MED1, MED26) (Zhang et al., 2005; Malik & Roeder, 2010; Yin & Wang, 2014). Bioinformatics analysis and expression studies have identified that multiple components of human Mediator complex undergo alternative splicing, potentially yielding in plethora of isoforms (A. Kivil, personal communication; US patent 2010/0087376 A1; Rienzo et al., 2012; Rienzo et al., 2014). For example, it was shown that expression of MED12, MED19 and MED30 alternative transcripts is restricted only to the cells undergoing endothelial progenitor cells (EPC) differentiation (Rienzo et al., 2010; Rienzo et al., 2012; Rienzo et al., 2014). However, up to date the isoforms of Mediator complex are not sufficiently studied and need to be further examined.

1.5.3. Chromatin remodeling complex and its tissue-specific activity

The primary role of chromatin remodeling complexes during transcription initiation is to move, eject or restructure the nucleosomes using ATP hydrolysis, so that the DNA sequence is accessible for the transcription machinery. In addition to remodeling activities, they are also known for their role as positive or negative regulators of gene expression during development (Kwon & Wagner, 2007). There are at least five main classes of chromatin remodeling complexes in eukaryotes: SWI/SNF, ISWI (Imitation SWI), INO80 (INOsitol requiring 80), Mi2/CHD (Mi2/Chromo-Helicase ATPase, DNA binding) and SWR1 (SWI2/SNF2-related 1). These complexes are recruited to DNA by sequence-specific transcription factors or GTFs. The mammalian SWI/SNF family comprises several large multisubunit complexes, among which the BRM/BRG-associated factor complex (BAF complex) is the best studied. The mammalian BAF complex contains up to 15 proteins with one of two mutually exclusive catalytic proteins, BRG1 or BRM, and a set of variable BRG1/BRM-associated factors, including BAF250, BAF200, BAF170, BAF155, BAF60, BAF57, BAF53, BAF47, or BAF45, which dynamic incorporation into the complex is highly genomic context dependent (Kingston & Narlikar, 1999; Vignali et al., 2000; Tang et al., 2010). Structurally different BAF complexes might have specialized functions enabling the cell-type- or developmental-stage-specific regulation of gene expression (Choi et al., 2015).

For instance, embryonic stem cell BAF complexes, esBAFs, contain the homodimers of BAF155, lack BAF170 and have BRG as a catalytic subunit, which is sufficient for the maintenance of self-renewal and pluripotency in embryonic stem cells (Ho et al., 2009). The expression and function of individual BAF proteins are being studied extensively.

There are several tissue-specific subunits of BAF complex described, which are involved in the regulation of various cellular programs. For example, the transition of the mammalian cells from the multipotent neural stem cells (NSCs) into neuronal or glial cell lineages requires the substitution of BAF45a/BAF53a subunits of the BAF complex into the specific homologous neuron-specific isoforms BAF45b, BAF45c and BAF53b. These isoforms are essential for the cell to initiate differentiation from the proliferating neural stem cells into mature post-mitotic neurons (Lessard et al., 2007; Choi et al., 2015). Another example is the tissue-specific role of BAF60 isoforms in the regulation of different cellular processes. For instance, BAF60a is involved in the regulation of hepatic lipid homeostasis and tumor suppression via interaction with PGC1-α (Peroxisome Proliferator-Activated Receptor Gamma, Coactivator 1 alpha) or p53, respectively; while BAF60c plays a role in the regulation of heart development and muscle differentiation, as well as lipogenesis (Lickert et al., 2004; Li et al., 2008; Oh et al., 2008; Meng et al., 2013; Wang et al., 2013). The highly homologous isoforms of the largest subunit BAF250, named BAF250a and BAF250b, are incorporated into separate BAF complexes and regulate the expression of different subset of genes important in early development. Studies in mice have shown that BAF250a is necessary for the pluripotency of embryonic stem cells and formation of early mesoderm by regulating the expression of vital genes, including *Sox2* (*Sex Determining Region Y-Box 2*), *Utf1* (*Undifferentiated Embryonic Cell Transcription Factor 1*) and *Oct4* (*Octamer-binding transcription factor 4*) (Gao et al., 2008). BAF250b, in turn, is found to be assembled into E3 ubiquitin ligase complex that targets histone H2B for ubiquitintion (Li et al., 2010).

Currently, little data is available concerning the alternatively spliced isoforms of BAF complex subunits. Only BRM, an ATPase component of the BAF complex, was shown to have multiple alternative mRNA variants with a complex pattern of expression in mice and humans. Moreover, in different cell lines the expression of these isoforms can be regulated by cyclin D1 and serum concentration, suggesting that in some cases the function of this protein may be very complex (Yang et al., 2011).

1.6. AS and diseases

AS defects may result in numerous human pathologies, including autoimmune and neurological disorders, cardiovascular and metabolic diseases, and cancer. Recent advances in the genome-wide screening revealed thousands of human genes with mutation-driven defective splicing, suggesting that they may be responsible for various diseases (Xiong et al., 2015). As many as 60% of all known human genetic disorders are associated with aberrant splicing, which is more abundant than previously expected (Ward & Cooper, 2010).

Splicing alterations that lead to diseases can be divided into primary and secondary splicing defects. Primary splicing defects emerge as a result of single nucleotide mutations affecting constitutive exon usage, or mutations located within splicing consensus sequences. Consequently, correct splicing cannot occur. Mutations affecting constitutive exon usage are mostly single-point nucleotide substitutions located at the conserved 5' or 3' splice sites, polypyrimidine tract, or branch point sequences (Wang & Cooper, 2007). They constitute approximately 15% of all human pathogenic mutations (Krawczak et al., 2007). This type of mutation significantly decreases the rate of intron-exon site recognition and inhibits spliceosome assembly to the site, which most frequently leads to exon skipping. However, to some extent, it can also lead to intron retention or activation of cryptic splice sites (Nakai & Sakamoto, 1994). Taken together, mutations that cause defective splicing usually result in the formation of aberrant proteins or introduction of premature downstream stop-codons, which occurs more frequently than missense point mutations. A number of diseases associated with this type of mutations have been described, including familial dysautonomia, neurofibromatosis type I, Frasier syndrome, or atypical cystic fibrosis (Ward & Cooper, 2010). As an example, familial dysautonomia is an inherited recessive neurodevelopmental disorder, which affects the development and survival of sensory and autonomic neurons. It is caused by mutation in the canonical 5' splice site of intron 20 in the transcription factor gene *IKBKAP* (*Inhibitor Of Kappa Light Polypeptide Gene Enhancer In B-Cells, Kinase Complex-Associated Protein gene*), which leads to the exon 20 skipping and a truncated protein. It has a dramatic effect on target genes, particularly on genes regulating cell migration in the nervous system (Ibrahim et al., 2007; Douglas & Wood, 2011).

In addition to mutations in the splice site consensus sequences, the aberrant splicing can be caused by mutations located within ESEs or ISEs of pre-mRNA. Such alternations in auxiliary *cis*-acting sequences can lead to the increased or decreased exon inclusion changing the normal ratio of alternative mRNAs. These mutations result in global defects of splicing with severe phenotypic outcomes (Ward & Cooper, 2010). Spinal muscular atrophy (SMA)

is one of the well-known diseases caused by this type of mutation, which affects the control of muscle movements as a result of the degeneration of lower motor neurons. The functioning of motor neurons is under the control of two nearly identical genes - *SMN1* (*Survival of Motor Neuron 1*) and *SMN2* (*Survival of Motor Neuron 2*). Usually, genetic disruption of *SMN1* results in motor neuron degeneration and progressive paralysis. In general, *SMN2* paralog can partially compensate for the loss-of-function of *SMN1* and relieve the symptoms. However, in some cases *SMN2* is mutated at the ESE in exon 7, which promotes its skipping and production of truncated unstable protein, so that it does not rescue the loss of *SMN1* with consequences to disease progression (Lorson et al., 1999). The other examples of such kind of mutation related diseases are frontotemporal dementia and Parkinsonism linked to chromosome 17 (Faustino & Cooper, 2003).

In secondary splicing defects, mutations affect the recognition elements of core splicing machinery, or change intracellular levels, localization and/or activity of the *trans*-acting regulatory proteins. The outcome and severity of the diseases directly depend on the location and type of mutation (Wang & Cooper, 2007; Singh & Cooper, 2012; Pedrotti & Cooper, 2014; Cieply & Carstens, 2015). These types of mutations potentially can affect the splicing of multiple genes. However, they occur relatively rare and usually lead to very dramatic developmental consequences, which are often lethal during embryonic development (Kalsotra & Cooper, 2011). There are still several examples of diseases caused by mutations in the components of the spliceosome as well as other regulators of AS (Ward & Cooper, 2010; Douglas & Wood, 2011; Pedrotti & Cooper, 2014). One of the best-described examples is autosomal dominant retinitis pigmentosa (ADRP), the most common form of progressive degenerative eye disease, which results in loss of vision. This disorder is caused by the mutations in four splice factors PRPF3 (Pre-mRNA Processing Factor 3), PRPF8 (Pre-mRNA Processing Factor 8), PRPF31 (Pre-mRNA Processing Factor 31) and snRNP200 (Small Nuclear Ribonucleoprotein 200kDa), which are important mediators of spliceosome assembly (Waseem et al., 2007). The other described diseases are myelodysplastic syndrome (MDS), acute myeloid leukemia (AML), chronic lymphocytic leukemia (CLL), amyotrophic lateral sclerosis (ALS), autism spectrum disorder (ASD), and others (Singh & Cooper, 2012; Cieply & Carstens, 2015).

1.6.1. AS and cancer

Genome instability and erroneous proteomes as a result of splicing defects are commonly known to cause cancer by promoting tumor cells to grow and survive, ending with metastasis. There are numerous cancer-associated

splice products revealed from melanoma, breast, ovarian, prostate, colon, lung, liver and other cancers, which are emerged as a result of mutations in *cis*-acting elements or changes in expression or activity of *trans*-acting splicing factors (Ward & Cooper, 2010). Current studies indicate that cancer-specific splice variants affect a plethora of cellular events, including metabolism, cell cycle control, immune escape, activation of signal transduction pathways, induction of genes involved in angiogenesis, cell proliferation, adhesion, and invasion, as well as change the ratio of pro-apoptotic/anti-apoptotic proteins (Venables, 2004; David & Manley, 2010; Oltean & Bates, 2014; Tsai et al., 2015).

The aberrant AS involves mutations in tumor suppressor genes and proto-oncogenes. Activation of proto-oncogenes and inactivation of tumor suppressors lead to cancer progression and metastatic dissemination. For instance, several germline mutations that modulate splicing by disrupting the ESEs and ISEs have been described for *BRCA1* (*Breast Cancer 1*) gene, a well-known tumor suppressor. These mutations perturb the structure and function of the protein causing the hereditary breast and ovarian cancer (Easton et al., 1993; Wappenschmidt et al., 2012). The aberrant splicing of *CCND1* (*Cell Cycle regulator Cyclin D1*) gene leads to the production of protein isoform CCND1b, which is frequently upregulated in breast and prostate cancer (David & Manley, 2010). Splicing mutation in the *c-KIT* gene (*Cellular tyrosine protein kinase KIT*) results in constitutive activation of this proto-oncogene leading to gastrointestinal stromal cancer (Chen et al., 2005). There is a continually growing number of findings on mutations in the *cis*-elements of other cancer predisposition genes, *e.g. p53*, *BCL2L1* (*B-Cell Lymphoma 2-Like 1*), *APC* (*Adenomatous Polyposis Coli*), *FHIT* (*Fragile Histidine Triad protein*), *CDKN2A* (*Cyclin-Dependent Kinase Inhibitor 2A*), *MLH1* (*MutL Homolog 1*), *CD44, KLK12* (*Kallikrein-Related Peptidase 12*), and *PTEN* (*Phosphatase and Tensin Homolog*), which increase the risk of various tumors (Petronzelli et al., 2001). Each of the aberrantly spliced genes contribute to different hallmarks of cancer, some of them are even involved in the regulation of more than one hallmarks, as in the case with RAC1 (Ras-Related C3 Botulinum Toxin Substrate 1), cancer-specific protein isoform of which is involved in both increased cell proliferation and invasiveness, or p53 protein isoforms, which are involved in apoptosis and cell proliferation (Surget et al., 2013; Zhou et al., 2013; Mehner et al., 2014).

Numerous recent reports have shown that in addition to mutations in *cis*-elements, changes in the expression of splicing factor genes have a significant role in cancer development and progression. Many SR proteins (*e.g.* SRSF1, SRSF2, and SRSF3) and hnRNP proteins (*e.g.* hnRNPA1, hnRNPA2, and hnRNPH) are frequently expressed at high levels in a wide variety of tumor tissues, which makes them acting as proto-oncogenes and promote cell

transformation (Karni et al., 2007; David & Manley, 2010; Ward & Cooper, 2010; Das & Krainer, 2014). For instance, high expression of splicing factor SRSF1 interferes with the regulation of AS of numerous genes relevant to tumor, such as *BIN1* (*Bridging Integrator 1*), controls the expression of proto-oncogene c-MYC, CCND1b, VEGF (Vascular Endothelial Growth Factor), CASP9 (Caspase 9), or others (Oltean & Bates, 2014; Cieply & Carstens, 2015). Of note, splicing factors themselves can act either as oncogenes or as tumor suppressors, depending on the context.

Recent studies using transcrpitome-wide analysis of AS have found, that cancer-specific splicing predominantly affects the proteins functioning in cell cycles, cell adhesion/migration, and signaling pathways (Tsai et al., 2015). It is known that AS regulates the activities of many signaling effectors. Among the best-studied examples of signaling proteins contributing to cancer growth and progression are members of the receptor tyrosine kinase family, including EGFR (Epidermal Growth Factor Receptor), FGFR (Fibroblast Growth Factor Receptor), INSR (Insulin Receptor), VEGFR (Vascular Endothelial Growth Factor Receptor), and others, as well as non-receptor proteins, such as SRC, RAS, and RAF. Thus, splicing of cytosolic RAF in MAPK pathway results in expression of constitutively active isoforms inducing tumor formation (Siegfried et al., 2013). Various oncogenic isoforms of FGFR protein with altered FGF-binding capacity can cause imbalanced FGFR signaling and promote tumor growth (Haugsten et al., 2010). Aberrant splicing of VEGFR leads to both more angiogenic and tumorigenic isofoms that are frequent in cancer (Zhang et al., 2000).

However, enormous number of natural noisy tissue-specific events, high cell-to-cell variability in splicing, and the heterogeneity of the disease, all these complicate the studies of splicing in cancer.

1.6.2. AS and breast cancer

Breast cancer (BC) is a second leading cause of death in women in Western countries. This heterogeneous disease represents tumors with broad spectra of molecular features, also in respect to prognosis and responses to therapy. The main event leading to death of patients with BC is the development of metastases. Therefore, the early detection of disease, particularly before the clinical diagnosis of metastases, is of great importance. The identification of molecular markers is necessary for early stage disease diagnostics and prognosis of BC. Alternatively spliced cancer-associated mRNA variants can ideally serve as source of such markers, emphasizing once again the role that AS plays in development of cancer.

BC has the highest number of aberrant splicing events among the solid cancers, indicating to the dysregulation of splicing as a general feature of tumor cells (Sveen et al., 2015). There are numerous publications reporting on AS events associated with BC, with exon skipping and intron retention to be the most prevalent (Lixia et al., 2007; Ng et al., 2008; Lapuk et al., 2010; Dutertre et al., 2010; Okumura et al., 2011; Eswaran et al., 2013). Large-scale RT-PCR analysis supported by microarray data of hundreds of cancer-related genes revealed 41 ASVs specific to breast tumor as compared to normal breast tissue (Venables et al., 2008; Venables et al., 2009). For example, the AS of such genes like *TNC* (*Tenascin C*), *BRCA1*, *CD44*, *ESR1* (*Estrogen Receptor 1*), *ESR2* (*Estrogen Receptor 2*), *CALD1* (*Caldesmon 1*), *COL6A3* (*Collagen type VI alpha 3*), *LRRFIP2* (*Leucine Rich Repeat Interacting Protein 2*), *PIK4CB (Phosphatidylinositol-4-kinase*), *PIK3CA* (*Phosphatidylinositol-3-kinase*), *MST1R* (*Macrophage-stimulating Protein Receptor*), *TP53* (*Tumor Protein 53*), *SYK (Spleen Tyrosine Kinase*), *MUC1* (*Mucin 1*), and *TPM1* (*Tropomyosin alpha 1*) can lead to cancer-specific proteins with modulated function resulting in disease progression (Lixia et al., 2007; Ng et al., 2008; Okumura et al., 2011). The more recent global studies using RNA sequencing analysis discovered, in addition to previously reportered BC-specific ASVs, multiple novel cancer-specific alternatively spliced genes, such as *CDK4* (*Cyclin-Dependent Kinase 4*), *LARP1* (*La Ribonucleoprotein Domain Family, Member 1*), *ADD3* (*Adducin 3*), and *PHLPP2* (*PH Domain And Leucine Rich Repeat Protein Phosphatase 2*) (Eswaran et al., 2013).

Most of the BC-specific AS changes are identified for genes with functions tightly linked to tumor progression including cell morphology, migration, proliferation and survival regulating factors. For example TNC, a large extracellular matrix glycoprotein involved in cell migration and proliferation, is one of the most extensively studied proteins, which isoforms are tightly linked to BC. It was demonstrated, that TNC isoforms of higher molecular weight containing two additional domains as a result of AS, are highly induced in BC resulting in tumor cell invasion and metastasis formation (Howeedy et al., 1990; Guttery et al., 2010). Survivin, a member of the inhibitor of apoptosis protein family, is another example of the protein, which gene specific ASVs were found to be correlated with BC aggressiveness and thus served as diagnostic and prognostic markers of disease (Ryan et al., 2005; Khan et al., 2014; Jaiswal et al., 2015).

As spliceosome itself is often altered in cancer, it has recently been discovered that mutations affecting genes of different spliceosomal components are present in 5,6% of diverse subtypes of BCs. Of all the spliceosome components, most frequently mutated gene is *SF3B1* (*Splicing Factor 3B subunit 1*), which affects AS of almost 80 downstream target genes driving

tumor growth, including *TMEM14C* (*Transmembrane Protein 14C*), *RPL31* (*Ribosomal Protein L31*), *UQCC* (*Ubiquinol-Cytochrome C reductase complex assembly factor 1*), and *ABCC5* (*ATP-Binding Cassette sub-family C member 5*) (Maguire et al., 2015). Increasing evidence shows that expression changes of splicing factors can trigger oncogenesis. For instance, SR-related proteins SRp40 and SRp55 have been shown to be highly expressed in BC serving as potent oncogenes (Dutertre et al., 2010). Interestingly, splicing of one-third of BC-specific ASVs is regulated by a single RNA binding splicing factor FOX2, which itself is often misspliced in BC (Venables et al., 2009). All this indicates that misregulation of splicing factors also may have an impact on many splicing events related to cancer.

ASVs can potentially serve as RNA biomarkers for the diagnostics and progression of breast cancer (Venables et al., 2008). However, to most recent years, only a minority of genes has been explored at the level of splicing due to the limitation in available technologies. The new large-scale methods, such as deep RNA-Seq and DNA microarray-based analyses, enable to identify and quantify novel cancer-associated ASVs (Dutertre et al., 2010; Eswaran et al., 2013).

1.6.3. AS as a diagnostic and therapeutic tool

The connection between aberrant splicing and the development of disease is of great interest in terms of contemporary diagnostics and therapeutics. The determination of ASVs from human blood or other bodily fluids and tissues could dramatically increase the pool of potential molecular biomarkers for the early diagnostics and prognosis of certain diseases, predominantly for cancer diagnostics, and have the potential to function as new therapeutic targets (Figure 4) (Sveen et al., 2015).

Any splice variant found specifically in cancer cells may serve as a candidate of biomarker. In cancer, molecular biomarkers can potentially distinguish tumor tissue from normal surrounding tissues, identify certain cancer subtypes, predict clinical outcome, and even predict the susceptibility of cancer cells to chemotherapy. Cancer-related biomarkers can be very helpful in understanding the processes of disease progression and developing targets for therapeutics. However, the heterogeneity of tumors and ASV expression variability across patients immensely complicate these efforts (Brinkman, 2004; Kim & Kim, 2012). Currently, only a few individual splice variants can potentially serve as cancer biomarkers, which need to be validated for diagnostic practice. Among the best-known examples are tumor-specific ASVs of *CD44*, *AR* (*Androgen Receptor*), *WT1* (*Wilms' tumor 1*), *BRCA1*, *MDM2* (*Mouse Double Minute 2 homolog*), *FN1* (*Fibronectin 1*), and *FGFR* genes

(Brinkman, 2004; Gambino et al., 2015).

In diseases other than cancer, the use of ASVs as disease-specific diagnostic markers has been studied less extensively. The potential implication of this diagnostic method has been explored, for example, in patients with Wilson disease, a progressive liver degeneration caused by mutation-induced aberrant AS of *ATP7B* (*ATPase, Cu^{2+} Transporting, Beta Polypeptide*) gene (Wan et al., 2010).

Today, several approaches have been developed for large-scale genome-wide detection of disease-specific AS events, including specialized splice junction specific DNA microarray analyses and deep RNA-sequencing technologies. These methods are constantly developing to allow highly quantitative and qualitative analysis of the abundance of specific ASVs and the assessment of differential usage of alternative exons (Ward & Cooper, 2010). However, despite numerous reports of prognostic biomarkers in the literature, only the small part of them have been approved by USA Food and Drug Administration (FDA) (Goossens et al., 2015).

As for therapeutics, there are several kinds of molecular approaches that can correct the misregulation of splicing or mutations in *cis*-acting and canonical splice site elements (Figure 4).

The correction of aberrantly spliced RNAs by the use of synthetic single-stranded antisense oligonucleotides (AONs), or splice-switching oligonucleotides (SSOs), has shown promise in the area of disease therapy. SSOs are usually short, 15-25 bp oligonucleotides that target the transcript at the site of interest in order to regulate AS. They can act by covering the *cis*-regulatory sequences or 5'/3' splice sites, which prevents *trans*-factors or snRNPs from binding or recognition of these sites. Depending on the sequence blocked, the corresponding exon will be skipped or included allowing the restoration of normal reading-frame of the transcript. Alternatively, SSOs can block aberrant alternative splicing events caused by the activation of cryptic splice sites, restoring normal splicing (Bauman & Kole, 2011; Nancy et al., 2015). The main obstacle of SSOs use as therapeutic agents is their poor cellular delivery. Efficient delivery vectors or synthetically conjugated cell penetrating peptides (CPPs) are used for successful transfection (Nancy et al., 2015). The first demonstration of the use of SSOs in tumors was made by a tumor xenograft study where the aberrant splicing of apoptotic regulator gene *BCL2L1* (*B-Cell Lymphoma 2-Like 1*) was efficiently redirected from anti-apoptotic form into its pro-apoptotic variant by SSOs thus reducing the tumor load in lung metastases (Bauman et al., 2010; Bauman & Kole, 2011). Lately, preclinical studies performed in mice demonstrated the effective treatment of Spinal Muscular Atrophy (SMA) by orally delivered brain-penetrating small molecule splicing modifiers of *Smn2* gene, which induce splicing correction and elevation of SMN

protein levels promoting motor function recovery and survival (Naryshkin et al., 2014).

Another promising therapeutic strategy is the use of small molecule modulators of splicing, which interfere with spliceosome or *trans*-acting splicing regulatory factors. The best example of their action is the modulation of the phosphorylation/dephosphorylation status of SR proteins, which, in turn, regulates their activity and exon recognition. Small molecule kinase inhibitors (*e.g.* indole derivatives) are able to inhibit specifically the phosphorylation of the SR factors (such as SRSF1), thus altering splice site selection and potentially treating the disease. The outcome of this method depends on the specificity and efficiency of the targeted small molecules. Future improvements and additional studies are necessary for the complete optimization of this method and development of new methods for treating aberrant splicing-related diseases (Ward & Cooper, 2010; Bonomi et al., 2013; Nancy et al., 2015).

Finally, aberrant splicing can be modulated by specific monoclonal antibodies (mAbs) targeting unique tumor-associated splicing isoforms, which results in down-regulation or inhibition of their function. The best-studied mAbs currently used in cancer therapy are cetuximab and panitumumab, which recognize tumor-specific alternative isoforms of EGFR, specifically their extramembrane epitope, resulting in the suppression of tumor proliferation and angiogenesis (Bonomi et al., 2013). Another example of promising antibody-based therapy was demonstrated with mAb targeting tumor-specific CD44 isoform, which is expressed on the surface of various human epithelial cancer cells (Masuko et al., 2012). However, this method is rather challenging due to the poor penetration of antibodies into the tumor tissue.

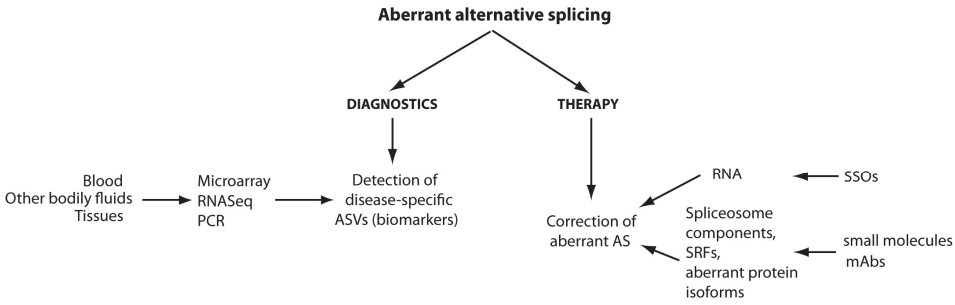

Figure 4. Schematic representation of molecular approaches used for the diagnostics and therapy of diseases caused by aberrant alternative splicing. SSOs - splice-switching oligonucleotides, AS – alternative splicing, ASVs – alternative splice variants, mAbs – monoclonal antibodies, RNASeq – RNA sequencing, SRFs – splicing regulatory factors.

2. AIMS OF THE STUDY

The aims of the current thesis were as follows:

1. To analyze the structure and function of tissue-specific alternative splicing of human *BDNF* and mammalian *BAF57* genes
2. To identify splicing alterations specfic to cancer cells
3. To evaluate the alternatively spliced transcripts as biomarkers of breast cancer

3. MATERIALS AND METHODS

The following methods were used during the studies:

1. Semiquantitative RT-PCR analysis (publications I, II, III)
2. Real-time quantitative RT-PCR analysis (publications II, III)
3. 5'RACE analysis (publication I)
4. CAT assay (publication I)
5. Sense/antisense RNA duplex analysis (publication I)
6. Construction of expression vectors (publications I and II)
7. RNA extraction (publications I, II, III)
8. Cell culture and transient transfection assay (publications I and II)
9. Preparation of nuclear extracts (publication II)
10. Immunocytochemistry (publication II)
11. Panomics QuantiGene® Plex assay on Luminex® platform (unpublished)

The standard QuantiGene® Plex protocol (Panomics, QuantiGene 2.0 reagent system, user manual) was followed. Briefly, whole blood lysates were prepared from 12 µl of whole blood and 84 µl of working lysis mixture. The lysates were incubated with working bead mix for 20 h at 54°C. Sequential hybridization reactions were conducted at 50°C for 1 h in 100 µl of pre-amplifier solution, then at 50°C for 1 h in 100 µl of amplifier solution, and finally at 50°C for 1 h in 100 µl of lable probe solution. Wells were washed with wash buffer after each hybridization. After the final wash, 100 µl of SAPE (streptavidin-R-phycoerythrin) working reagent was added to each well and the plate was left at room temperature for 30 min. The developed fluorescence signal was measured on Luminex® instrument and analyzed using xPonent® software. The results of target genes expression were normalized to the expression level of *TBP* gene.

4. RESULTS AND DISCUSSION

4.1. Tissue-specific alternative splicing and expression of human *BDNF* and mammalian *BAF57* genes (publications I and II)

4.1.1 Organization of the human *BDNF* gene locus, its complex splicing and expression (**publication I**)

Unlike rodent *Bdnf* gene, the structure and expression of which have been previously well described, the human *BDNF* gene was not sufficiently investigated. Thus, by means of the previously published data together with the bioinformatics studies and newly performed experiments using 5'RACE (5' rapid amplification of cDNA ends) and RT-PCR methods, we re-examined the structure of the human *BDNF* gene and analyzed in more details the expression of its multiple alternative mRNAs in various human peripheral tissues and brain regions. Our results indicated that the revised human *BDNF* gene spans more than 70 kb of chromosome 11 genomic sequence and consists of eleven exons, ten 5' exons of which are mostly non-coding (exons I-VIIIh) and one is a protein-coding 3'exon (exon IX) (Figure 5). We introduced a new nomenclature for the human *BDNF* gene with the novel numeration of exons, which is consistent with the rodent exons described by Aid *et. al.* (Aid et al., 2007). Human exons I, II, III, IV, V, VI, VII, VIII and IX are highly homologous to the previously described mice and rat *Bdnf* exons, while exons Vh and VIIIh are newly-discovered and specific only to humans (Aid et al., 2007). It is worth noting that the previous data published by Liu *et al.* have described only seven 5' exons and one protein-coding exon in humans, alternative exons V, Vh, VIII and VIIIh have not been discovered before (Liu et al., 2005). Additionally, we have also found that the final exon IX is much longer in its 5' end than it was previously considered (Figure 5).

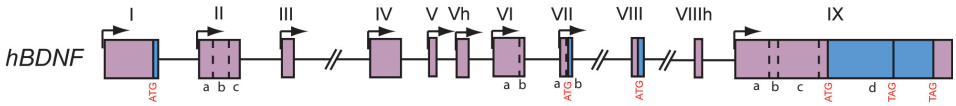

Figure 5. Schematic structure of the human *BDNF* gene. Exons (I-IX) are shown as boxes and introns as lines. Purple colored boxes indicate untranslated regions of the exons and blue colored boxes indicate translated regions of the exons. ATG and TAG mark the positions of putative translation initiation and termination sites, respectively. Arrows indicate the transcription start sites. Vertical dashed lines indicate alternative splice sites. Small letters under the exons (a, b, c, and d) indicate the regions of the exon divided by alternative splice sites.

According to our results we showed that the splicing pattern of human *BDNF* is very complex and diverse, and can be summarized as follows:

- *BDNF* exons I, II, III, IV, V, Vh, VI, and VII are in most cases spliced as 5' exons to the common protein-coding exon IX, specifically to its IXd region; the only exclusion is exon V, which can be also spliced to the internal exons VIII and/or VIIIh;
- *BDNF* 5' exons I, II, III, IV, V, Vh, VI, VII have separate transcription start sites located upstream of each of them;
- *BDNF* final protein-coding exon has its own transcription initiation site located upstream of exon IX, which makes him also transcribed as a separate unit, however, mostly it is used in conjunction with the upstream exons;
- *BDNF* rare exons VIII and VIIIh are exclusively spliced as internal exons, and only to exon V as a 5' exon;
- *BDNF* exons II, VI and VII have alternative splice donor sites;
- *BDNF* exon IX is subject to internal splicing within its non-coding region (IXb and/or IXc regions can be spliced out).

Thus, if we take into account the previously published data indicating that the final exon of *BDNF* employs two separate polyadenylation sites at the protein noncoding region, generating two pools of mRNAs with short (0.35 kb) or long (2.85 kb) 3'UTRs, we can postulate that at least 34 distinct *BDNF* bipartite or tripartite mRNAs can be transcribed from the gene with functional promoters present upstream of nine 5' exons. Before, the only one *BDNF* promoter has been described in human (promoter IV, according to new nomenclature) (Fang et al., 2003). The activities of all identified potential alternative promoters were confirmed in human and mouse cell lines, the intensities of which showed a great variability and differed between the cells. Recent data suggested that the use of multiple alternative promoters and alternative polyadenylation allows gene transcription in different cell types and under different physiological conditions, which increases the regulatory flexibility over the gene (Ayoubi & Van De Ven, 1996; Singer et al., 2008). In case of *BDNF*, we showed that numerous identified alternative transcripts have individual pattern of expression in different adult human tissues, with the highest expression detected throughout the brain, where BDNF exerts its major biological function. Within the brain, the expression of *BDNF* alternative transcripts is highly differential, which might be partially explained by the multiple roles BDNF has in CNS. We observed that human *BDNF* transcripts starting with exons II, III, IV, V, and VII are predominantly brain-specific, whereas transcripts containing exons I, III, Vh, VIab and IXabcd are also detected at high or moderate levels in several peripheral tissues, such as heart,

kidney, lung, muscle, placenta, prostate, testis, and some others. Throughout the adult brain, the expression of individual *BDNF* alternative products was the highest in olfactory bulb and tract, frontal cortex, pituitary stalk (*infundibulum*), mammillary body (*corpus mammilare*), cerebellum, pons, *colliculus,* hippocampus, amygdala, and brain lobes. Conversely, with the exception of exon IXabcd mRNAs, which were expressed at relatively high level in all brain regions, low expression was detected in such brain parts as dentate nucleus, cerebellum white matter, peduncle (*pedunculi*), *substantia nigra*, red nucleus (*nucleus rube*r), callous body (*corpus callosum*), *globus pallidus*, thalamus, and epiphysis, some of which contain only glial cells or axons. The expression of alternative transcripts was also very low in striatum (putamen and *caudate nucleus*), which was previously considered to be a brain region totally lacking BDNF (Lessmann et al., 2003). Of note, the overall expression of transcripts starting with exons III, V and VII was lower if compared to others. The expression results we obtained from our studies in various human tissues and brain regions were mostly similar to those described in rodents by Aid *et al.* and in accord with previous reports described in human by Liu *et. al.* (Liu et al., 2005; Aid et al., 2007).

In silico open reading frame (ORF) analysis of the identified *BDNF* mRNAs predicted that translation of transcripts starting with exons II, III, IV, Vh and VI always initiates from the translation initiation codon positioned in exon IX (IXd part), thus encoding the canonical protein product. However, transcripts starting with exons I and VII encode an in-frame Met in their 5' exons, which potentially leads to the translation of the proteins with several additional amino acids at their N-terminus, 8 amino acids for exon I and 15 amino acids for exon VII, thus extending the signaling sequences of prepro-BDNF. Also, as some transcripts starting with exon V can be spliced with internal exons VIII and/or VIIIh, the N-terminal sequence of the translated proteins might have additional 82 amino acids as a result of an introduced translation initiation codon located in exon VIII. These alternative additional signaling sequences located at the N-terminus of the protein might alter its stability, subcellular sorting, or affect post-translational modifications. However, the most recent studies performed by I. Koppel *et al.* have shown that exon I-specific translation start site leads to more efficient synthesis of the protein as compared to the canonical start site located in exon IX, which is independent of its stability or secretion (Koppel et al., 2015). Finally, after the cleavage of signal peptides from the prepro-BDNF isoforms, all the encoded proteins in most cases share the common primary sequence. Such a diverse pattern of *BDNF* transcripts in most cases does not affect the structure of mature BDNF but is rather necessary for its cell-type- or tissue-specific expression.

Multiple recent studies have demonstrated that *BDNF* promoters contain different regulatory elements, such as calcium-response elements (CaRE), calcium/cAMP response elements (CRE) and upstream stimulatory factor binding element (UBE) located in promoter upstream of exon IV, neuronal-restrictive silencing element (NRSE) located in promoter upstream of exon II, basic helix-loop-helix (bHLH)-PAS transcription factor response element (PasRE) upstream of exon I, and many others, which are recognized by specific transcription factors, also in response to neuronal activation (Tao et al., 1998; Zuccato et al., 2003; Zuccato et al., 2007; Pruunsild et al., 2011). It means, that depending on the nature of external stimuli, corresponding *BDNF* mRNAs can be either expressed or silenced in certain tissues. For instance, the studies in rat have demonstrated that following stress exposure the expression of *Bdnf* exon I and exon IV mRNAs in hippocampus was significantly reduced, whereas fear memory extinction resulted in increased expression of the same mRNAs in prefrontal cortex of mice (Bredy et al., 2007; Fuchikami et al., 2010). The untranslated regions at the 3' end of mRNA transcripts also play a significant role in post-transcriptional regulation of gene expression, where the length of 3'UTR and some important sequences located there influence the efficiency of translation, stability of mRNA and its location (Tanguay & Gallie, 1996; Matoulkova et al., 2012). For instance, it has been shown that the length of *BDNF* 3'UTRs defines its subcellular localization, with the short 3'UTR mRNAs localized primarily at the cell body, whereas the long 3'UTR enables localization to dendrites of the neurons, suggesting their divergent roles at different cellular locations. Moreover, the loss of long 3'UTR of *BDNF* was shown to impair long-term potentiation and leads to altered dendritic spine morphology (An et al., 2008; Lau et al., 2010; Miura et al., 2014).

To our surprise, our RT-PCR analysis revealed that several alternative *BDNF* mRNAs are transcribed in antisense direction in relation to the human *BDNF* gene, which was confirmed by orientation-specific PCR and RNase protection assay. These findings resulted in the discovery of *BDNF* natural antisense transcripts encoded by *antiBDNF* gene located at the same gene locus with *BDNF*. According to our results, this newly discovered *antiBDNF* gene spans about 191 kb and consists of at least 10 exons, with exons 1-4 located downstream of the gene and exons 7-10 upstream of *BDNF* protein-coding exon IX in intronic sequences of the gene. Furthermore, two antisense exons (exons 5 and 6) are complementary to the regions of the major BDNF protein-coding exon IX. Our studies demonstrated that the splicing of newly identified antisense transcripts is highly diverse and results in production of hundreds of alternative transcripts. Unlike *BDNF*, all the *antiBDNF* alternative mRNAs were shown to be transcribed from the single functional promoter located upstream of exon 1 with no ORF. In parallel with our discoveries, Liu and

colleagues had published the similar experimental results ahead of us, thus confirming our findings (Liu et al., 2005). Later experiments have shown that the existence of antisense *BDNF* is unique only to humans. Recent studies provided evidence that about 70% of all the mammalian genome have antisense transcription potential, which might be involved in inhibition of gene transcription and translation, or regulation of alternative splicing (Chen et al., 2004; Ling et al., 2013). Moreover, naturally occurring antisense non-coding mRNAs are rather prevalent in the central nervous system of eukaryotes suggesting their potential role in brain function (Korneev & O'Shea, 2005; Korneev et al., 2013). Antisense transcripts have been reported, for example, for the rat *Kcna2* (*Potassium Voltage-Gated Channel Subfamily A Member 2*) gene, marine mollusk neurotransmitter gene *Srn* (*Sensorin*), mouse gene *Nos1* (*Nitric Oxide Synthase 1*), and many others (Korneev et al., 2008), (Zhao et al., 2013; Kadakkuzha et al., 2014). In case of BDNF, this phenomenon might serve as an additional mechanism of genomic regulation, which needs further to be investigated. The RT-PCR analysis performed with antisense transcripts indicated their prominent expression throughout the adult human tissues analyzed, although at different levels – highest expression was detected in the brain, kidney, spinal cord and testis; whilst the expression in the adrenal gland, bone marrow, liver, pancreas, placenta, small intestine, and trachea was significantly lower. In the brain, *antiBDNF* transcripts were expressed at almost similar levels in all structures under study. The pattern of *BDNF* and *antiBDNF* expression was only partially overlapping, suggesting that these transcripts are regulated by independent promoters, or their splicing is independently regulated.

In sum, our findings confirm that BDNF expression is highly tissue-specific and tightly regulated at the level of transcription in many ways - through alternative promoters usage, AS, antisense transcripts, and alternative polyadenylation sites. To conclude, our detailed analysis of the human *BDNF* gene locus and its expression in adult human peripheral tissues and multiple adult brain regions elucidates its complex organization and highly diverse tissue-specific expression, which might be helpful for the future functional studies.

Since then, the ongoing studies of BDNF have not elaborated its gene structure and expression data, but rather concentrated on the role of this neurotrophin in the brain. It was discovered, for example, that BDNF is not a major survival factor for the majority of CNS neurons, including hippocampal neurons, but is required for the postnatal growth of the striatal neurons and their dendrites (Rauskolb et al., 2010). A recent study revealed that the prodomain region of BDNF is not proteolytically degraded but present at high levels in the adult hippocampus and other brain regions with further role in depolarization-

induced secretion of neurons, indicating that it can act as an independent ligand involved in modulation of neuronal morphology (Dieni et al., 2012; Anastasia et al., 2013; Hempstead, 2015). Recent reports also showed the cooperation of BDNF with WNT signaling to regulate dendritic spine formation, stimulate proliferation and differentiation of neural stem cells, or promote the growth of neurons (Hiester et al., 2013; Chen et al., 2013; Yang et al., 2015). Outside of the nervous system, BDNF together with trkB were shown to share a developmental role in the formation of cardiac endothelium and regulation of cardiac contraction force, which reveals that BDNF has a global role in the maintaining of homeostasis of the whole organism (Anastasia et al., 2014; Fulgenzi et al., 2015).

4.1.2. N-terminally truncated BAF57 isoforms contribute to the diversity of SWI/SNF complexes in neurons (**publication II**)

Analysis of the expression profile of genes encoding PIC complex proteins in humans, including the GTFs, the Mediator complex subunits, and chromatin remodeling complex subunits, resulted in identification of novel ASVs with a tissue- or cell type-specific mode of expression. These studies were performed using RT-PCR, DNA sequencing, and *in silico* analysis. Our widespread experiments revealed that most of the genes encoding PIC proteins have multiple alternatively spliced mRNAs, which are differentially expressed throughout various human tissues and cell lines (data not shown). Out of dozens of genes analyzed, the human *BAF57* gene exhibited a distinct tissue-specific expression of its alternative mRNAs. Therefore, we aimed to study the effects of splicing on the activity of BAF57 in more detail.

BAF57, also known as SMARCE1 (SWI/SNF Related, Matrix Associated, Actin Dependent Regulator Of Chromatin, Subfamily E, Member 1), is a core subunit of the SWI/SNF chromatin remodeling complex, which is present only in higher eukaryotes and is found in all mammalian SWI/SNF complexes. It has been shown to recruit the SWI/SNF complex to gene promoters and directly interact with various transcriptional factors, such as oncogenic transcription factor c-MYC, SRC1 (Steroid Receptor Coactivator 1), MeCP2 (Methyl-CpG Binding Protein 2), TSHZ3 (Teashirt Zinc Finger Homeobox 3), and several hormone receptors (*e.g.* estrogen receptor and androgen receptor) to negatively or positively regulate the expression of multiple target genes (Belandia et al., 2002; Pal et al., 2003; Kiskinis et al., 2006; Harikrishnan et al., 2005; Faralli et al., 2011). Among the genes regulated by BAF57 are those, which are involved in cell cycle progression, implying its role in the transcriptional control of cell proliferation (Hah et al., 2010). Additionally, BAF57 was shown to facilitate REST-mediated repression of neuronal genes by targeting the SWI/SNF complex to NRSE (Neuron-Restrictive Silencer Element) through direct interaction of its coiled-coil domain with transcriptional co-repressor coREST (Battaglioli et al., 2002). It is also demonstrated that the maintenance of BAF57 optimal protein level in the cells is coordinated by SWI/SNF complex subunits BAF155 and BAF170, which is critical for its role as a transcriptional regulator (Chen & Archer, 2005). Of note, BAF57 is expressed at high levels in cancer, such as prostate cancer, ovarian cancer and endometrial cancer (Kagami et al., 2012; Balasubramaniam et al., 2013; Yamaguchi et al., 2015). On the other hand, loss-of-function mutations in *BAF57* gene or alterations of its expression were shown to contribute to severe oncogenic transformations causing breast cancer or spinal meningiomas (Kiskinis et al., 2006; Villaronga et al., 2011; Smith et al., 2013).

Mammalian *BAF57* is a 12-exon gene located on chromosome 17q21 that encodes a 57 kDa protein. BAF57 protein contains several functional domains, including Proline-rich region and a DNA-binding HMG (High Mobility Group) domain located in its N-terminus, highly conserved NHRLI domain together with a kinesin-like coiled-coil domain located in its central part, and a glutamic acid-rich sequence in its C-terminal region (Wang et al., 1998; Chen & Archer, 2005). Previous data and *in silico* analysis of EST database done by us revealed that at least six different BAF57 protein isoforms can be produced from the human *BAF57* gene as a result of combinatorial alternative splicing of exons III, IV and XI. To specify, the canonical form of human BAF57 protein is 411 amino acids long and is translated from mRNA containing exons I to XII with exon XI spliced out; the inclusion of exon XI results in C-terminally truncated protein of 363 amino acids; deletion of exon IV generates protein isoforms missing amino acids 17-51, while deletion of exons III-IV results in N-terminally shorter protein lacking the first 70 amino acids, including part of the HMG box domain (Battaglioli et al., 2002).

We analyzed the expression of *BAF57* gene in various human tissues, cancer cell-lines and adult brain regions and observed the expression of ASVs with deletions of exon IV (*BAF57_v3* and *BAF57_v4*) or exons III-IV (*BAF57_v5* and *BAF57_v6*), both transcribed with or without exon XI. The most intriguing finding was that mRNAs generated by AS of exons III-IV or exon IV (hereafter N-BAF57 isoforms), encoding proteins with truncated N-terminus and either long or short C-terminus, are expressed exclusively in neuronal tissues, including fetal brain, but not in nonneural tissues and cancer cells, underscoring their unique regulatory role during neuronal development. We have demonstrated that expression of all identified *N-BAF57* mRNAs varied within the different adult brain regions. The highest expression was detected in olfactory bulb, cerebral cortex, brain lobes, pituitary stalk (*infundibulum*), mammillary body, cerebellum, caudate nucleus, and thalamus. The moderate expression was in olfactory tract, pons, midbrain tectum, hyppocampus, amygdala, and putamen. However, low or no expression was observed in such brain regions like optic nerve, cerebellar white matter, pedunculus, *corpus callosum*, medulla, cerebellar nuclei, *substantia nigra*, red nucleus (*nucleus ruber*), *globus pallidus*, and epiphysis. Additionally, we have shown that neuron-specific alternative mRNAs were also expressed in adult human neural stem cells as well as in stem cells subjected to neuronal differentiation, which also indicates their potential involvement in a process of neuronal development. Thus, our results indicated that the heterogeneity of SWI/SNF complexes, in addition to the incorporation of common ubiquitously expressed subunits, can also be related to the recruitment of different alternative isoforms to the core complex in a cell-specific manner.

The expression analysis of *BAF57* gene in mouse and rat tissues confirmed the existence of the same alternative mRNAs as described in humans, as well as revealed three additional minor splice variants with deletions of exon II (*BAF57_v7*), II and IV (*BAF57_v8*), and II, III and IV (*BAF57_v9*), which were also present in mouse and rat EST databases. However, the expression of these minor alternative mRNAs was very low. *In silico* analysis has shown that these minor transcripts can potentially encode proteins lacking 19 amino acids (encoded by *BAF57_v7*) or 70 amino acids (encoded by *BAF57_v8* and *BAF57_v9*) at their N-terminus. Similarly to humans, the expression of all identified mouse and rat alternative mRNAs was restricted to neuronal tissues with regional difference within analyzed brain regions.

We observed that in mouse and rat the expression of neuron-specific alternative mRNAs was differential during early and late neurogenesis with high expression during embryonic stage of brain development, after which it gradually decreased (postnatal days 7-21) and increased again in the adult brain. Moreover, the expression of neuron-specific alternative mRNAs in adult brain was observed only in neurons, but not in glial cells. All these findings suggested that various SWI/SNF complexes with mutually exclusive subunit isoforms have distinct functions during development and in adult nervous system.

In vitro experiments in human melanoma cells showed that exogenously expressed canonical BAF57 and N-BAF57 proteins were distributed exclusively in the nucleus of the cells. Moreover, the nuclear localization of all the expressed proteins matched with the distribution of Ser5 phosphorylated form of RNA polymerase II CTD, which is a mark of active RNA polymerase II complex ready for transition from transcription initiation to elongation. This indicates that N-BAF57 isoforms along with canonical BAF57 are recruited to active transcription complexes, where they perform their specialized actions essential for the formation and maintenance of complex nervous system.

Since the splicing of exons III and IV results in disruption of Pro-rich and/or HMG domain of the N-BAF57 isoforms, it has not been sufficiently studied if these deletions can change the functional properties of these proteins. We have confirmed by the fractionation and immunoprecipitation studies that both N-BAF57 isoforms, despite the partially disrupted HMG box domain in one of them, can still be assembled into functional human SWI/SNF complexes containing core subunits BRM or BRG1 along with BAF155 and BAF170, and potentially guide this complex to chromatin as effectively as the endogenous canonical BAF57 protein. Previous studies have already shown that, generally, either Pro-rich region or the HMG domain are not essential for recruiting SWI/SNF complexes to chromatin or their DNA binding and nucleosome disruption activities (Paull et al., 1993; Wang et al., 1998). Considering BAF57

was already known as a regulator of several genes, the similar role of N-BAF57 isoforms was obscure. Since endogenous BAF57 was previously shown to affect the REST-mediated transcription of neuronal genes in nonneural cells by targeting SWI/SNF complex to NRSE-containing promoter sequences, we analyzed if neuronal BAF57 isoforms can also be involved in the modulation of certain genes expression (Battaglioli et al., 2002; Watanabe et al., 2006). In response to exogenous expression of N-BAF57 proteins in human melanoma cells, we observed changes in the expression of such NRSE-containing neuronal genes like *CHRM4* (*Muscarinic Cholinergic Receptor 4*) and *L1CAM* (*L1 Cell Adhesion Molecule*), which suggested that neuron-specific isoforms might potentially be involved in the regulation of transcription of certain neuronal genes in nonneural cells in a promoter-specific manner. However, in order to make broader conclusions concerning the global role of N-BAF57 isoforms in the regulation of genes expression, further studies are needed.

To date, there are no additional studies performed on the functional roles of BAF57 isoforms. However, it has been demonstrated that in some breast cancer cells, mutations in *BAF57* gene generating truncated proteins resulted in abnormal activation of estrogen receptor α (ERα), suggesting that these truncated variants might be involved in estrogen-dependent oncogenic proliferation (Villaronga et al., 2011). Therefore, it would be interesting to study the role of alternative BAF57 isoforms not only in neuron-specific contexts but also during tumorigenesis.

4.2. The studies of alternative splicing in cancer tissues

Identification of AS markers from the peripheral blood of patients with breast cancer (**publication III**)

Breast cancer is a heterogeneous disease characterized by various morphological and molecular features. The outcome and management of the disease relies on the availability of robust early-stage, predictive and prognostic biomarkers. To date, there are several prognostic and diagnostic nucleic-acid based BC biomarker assays available, which are approved by FDA. Among them are GeneSearch™ Breast Lymph Node (BLN) assay that detects the presence of breast tumor cell metastasis in lymph nodes through the detection of *MG* (*Mammaglobin*) and *CK19* (*Cytokeratin 19*) genes expression; Prosigna® Prognostic Breast Cancer Gene Signature assay that analyzes the activity of 58 genes in early-stage hormone-receptor-positive BC; and MammaPrint™ assay that analyzes the expression of 70 genes to estimate a women's recurrence risk for early-stage BC (Goossens et al., 2015). These assays are based on the evaluation of gene expression levels of target genes.

The development of minimally invasive approaches for diagnosing and monitoring of cancer is of great priority in ongoing research efforts. Non-invasive tests for detection of disease biomarkers circulating in peripheral blood are considered as a promising diagnostic tool in oncology. It is known that circulating cell-free DNA and mRNA are present at large concentrations in the blood of cancer patients. Tumor-specific genetic and epigenetic alterations of these can serve as easily accessible and reliable biomarkers (Alix-Panabières et al., 2012; Schwarzenbach, 2013). However, one of the major challenges of the use of this method is its low specificity and sensitivity. In our studies, we have screened dozens of selected candidate molecular markers for BC from the peripheral blood samples of 46 BC patients and 43 healthy controls using Panomics QuantiGene MultiPlex assay on a Luminex® platform, which provides direct quantitative measurement of multiple RNA targets with high accuracy and precision directly from blood. Out of the 32 selected candidate markers, 4-6 showed promising results and will be further tested in the next phase of clinical validation (data not shown).

ASVs represent a novel class of diagnostic markers. However, the early detection of ASVs is rather challenging due to the limitations in methods and their sensibilities. Recently, the new approach to searching the alternatively spliced isoforms as biomarkers from human plasma has been performed at the proteome level, which showed a great potential for further discoveries of new types of BC biomarkers (Zhang et al., 2013). In our work, we focused on the identification of novel potential molecular markers for BC from the peripheral

blood of patients based on the tumor-specific alternatively spliced transcripts. For this purpose, we have chosen by using *in silico* approach potential cancer-specific ASVs of 94 human genes. RT-PCR and consequent qRT-PCR experiments with 29 primary BC and healthy breast tissue samples resulted in the selection of 12 ASVs with the most significant differential expression in cancer relative to healthy individuals. Splice variants of *CASP9* (*Caspase 9*), *PPARG* (*Peroxisome Proliferator-activated Receptor Gamma*) and *CCNH* (*Cyclin H*) genes were detected at low levels in cancer, while the levels of splice variants of *PRC1* (*Protein Regulator of Cytokinesis 1*) and *BUB1* (*Budding Uninhibited by Benzimidazoles 1*) genes were significantly higher in cancer as compared to normal tissue. The expression of *REST-N50*, ASV of *REST* gene, was detected exclusively in cancer tissues. Previously, this splice variant of *REST* with a 50-bp insert located between exons 5 and 6 and leading to a C-terminally truncated protein has been already shown to be associated with different types of cancers, including small cell lung cancer and BC (Coulson et al., 2000; Wagoner et al., 2010).

Further analysis of the expression of 12 selected cancer-specific ASVs was done by using the peripheral blood samples of 26 BC patients diagnosed with early-stage BC (BCI/II), 10 inoperable patients with locally advanced BC (LABC) receiving neoadjuvant treatment and 26 healthy individuals. Our results revealed that overall expression of selected splice variants is low in the PBMC (peripheral blood mononuclear cells) population of healthy individuals and of patients diagnosed with cancer. However, expression of 2 ASVs was statistically significantly higher in blood cells of cancer patients. The expression of *REST-N50* was detected at higher levels in LABC group as compared to BCI/II group. Furthermore, the average expression of *DOPEY1* ASV (referred to as *DOPEY1v2*) with intron retention between exons 32 and 33 was significantly higher in the blood cells of all cancer patients, especially LABC group, as compared to healthy controls. Of note, the expression of *REST-N50* and *DOPEY1v2* ASVs did not show any correlation with cancer stage, grade, or hormone-receptor status. This indicates that these biomarkers can potentially be used for the detection of BC from the peripheral blood of diseased. Although, the additional studies with a larger number of disease and control samples is needed.

In further analysis of the suitability and prognostic significance of *REST-N50* and *DOPEY1v2* as candidate blood markers for monitoring BC treatment efficacy we used qRT-PCR method and peripheral blood samples of 10 LABC patients subjected to neodajuvant therapeutic regimen. Our results showed that before the treatment, the average expression of *REST-N50* was prominent in five out of ten patients with visible decrease during the course of the treatment. The average decrease in the levels of *REST-N50* expression

measured before and after completing the therapy differed significantly and varied from 2- to 10-fold depending on individual. Expression of *DOPEY1v2* did not show any significant changes during the course of the treatment. Thus, our results indicated that *REST-N50,* but not *DOPEY1v2,* could potentially serve as blood biomarkers for monitoring the neoadjuvant treatment efficacy in an inoperable group of patients with locally advanced disease. However, further validation studies with larger cohorts are needed for proper conclusions.

It is very important to find biomarkers that will predict the disease at the early stage with possibly least-invasive method. BC-specific ASVs discovered in the current study may provide novel biomarkers for early diagnosis of breast cancer from the peripheral blood of patients.

CONCLUSIONS

1. The studies on human *BDNF* gene

    - The structure of *BDNF* gene has been revised. In human, *BDNF* gene spans 11 exons with 9 transcription start sites in relevant functional promoters and two alternative polyadenylation signals, which together result in a variety of alternatively spliced mRNAs with a distinct tissue- and brain region-specific expression.
    - Human-specific non-coding natural antisense *BDNF* (*antiBDNF*) RNAs have been discovered. Multiple alternatively spliced *antiBDNF* mRNAs are transcribed from the single promoter in antisense direction to human *BDNF* gene locus.
    - *BDNF* and *antiBDNF* transcripts form RNA duplexes in adult human brain, implying their role in *BDNF* gene regulation.

2. The studies on mammalian *BAF57* gene

    - The structure and expression of mammalian *BAF57* gene has been analyzed. Mammalian *BAF57* contains 12 exons, which are alternatively spliced to encode at least six (nine in mice and rats) different protein isoforms. Alternative mRNAs encoding protein isoforms with modified N-termini are expressed exclusively in neurons of developing and adult brain as compared to nonneural tissues.
    - Neuron-specific isoforms of BAF57 are co-localized with an active form of RNA Pol II in the nucleus of the cells and can be assembled to SWI/SNF chromatin remodeling complexes.
    - Neuron-specific isoforms of BAF57 affect the expression of certain NRSE/RE1-containing genes in nonneural cells.

3. The identification of molecular biomarkers of breast cancer from the peripheral blood of diseased

    - Expression levels of ASVs of *REST* and *DOPEY1* detected in the blood of cancer patients can differentiate between diseased with BC and healthy controls, and discriminate between patient groups.
    - *REST-N50,* ASV of *REST,* has a sound potential in evaluating the efficacy of neoadjuvant therapy.

BIBLIOGRAPHY

Aid, T., Kazantseva, A., Piirsoo, M., Palm, K., Timmusk, T. (2007) Mouse and rat BDNF gene structure and expression revisited. *J Neurosci Res* **85**, 525-35

Albright, S. R., Tjian, R. (2000) TAFs revisited: more data reveal new twists and confirm old ideas. *Gene* **242**, 1-13

Alix-Panabières, C., Schwarzenbach, H., Pantel, K. (2012) Circulating tumor cells and circulating tumor DNA. *Annu Rev Med* **63**, 199-215

Ameur, A., Zaghlool, A., Halvardson, J., Wetterbom, A., Gyllensten, U., Cavelier, L., et al. (2011) Total RNA sequencing reveals nascent transcription and widespread co-transcriptional splicing in the human brain. *Nat Struct Mol Biol* **18**, 1435-40

Ameyar-Zazoua, M., Rachez, C., Souidi, M., Robin, P., Fritsch, L., Young, R., et al. (2012) Argonaute proteins couple chromatin silencing to alternative splicing. *Nat Struct Mol Biol* **19**, 998-1004

An, J. J., Gharami, K., Liao, G.-Y., Woo, N. H., Lau, A. G., Vanevski, F., et al. (2008) Distinct role of long 3' UTR BDNF mRNA in spine morphology and synaptic plasticity in hippocampal neurons. *Cell* **134**, 175-87

Anastasia, A., Deinhardt, K., Chao, M. V., Will, N. E., Irmady, K., Lee, F. S., et al. (2013) Val66Met polymorphism of BDNF alters prodomain structure to induce neuronal growth cone retraction. *Nat Commun* **4**, 2490

Anastasia, A., Deinhardt, K., Wang, S., Martin, L., Nichol, D., Irmady, K., et al. (2014) Trkb signaling in pericytes is required for cardiac microvessel stabilization. *PLoS One* **9**, e87406

Andersson, R., Enroth, S., Rada-Iglesias, A., Wadelius, C., Komorowski, J. (2009) Nucleosomes are well positioned in exons and carry characteristic histone modifications. *Genome Res* **19**, 1732-41

Ansari, S. A., Morse, R. H. (2013) Mechanisms of Mediator complex action in transcriptional activation. *Cell Mol Life Sci* **70**, 2743-56

Aoyama, M., Asai, K., Shishikura, T., Kawamoto, T., Miyachi, T., Yokoi, T., et al. (2001) Human neuroblastomas with unfavorable biologies express high levels of brain-derived neurotrophic factor mRNA and a variety of its variants. *Cancer Lett* **164**, 51-60

Ayoubi, T. A., Van De Ven, W. J. (1996) Regulation of gene expression by alternative promoters. *FASEB J* **10**, 453-60

Baj, G., Leone, E., Chao, M. V., Tongiorgi, E. (2011) Spatial segregation of BDNF transcripts enables BDNF to differentially shape distinct dendritic compartments. *Proc Natl Acad Sci U S A* **108**, 16813-8

Balasubramaniam, S., Comstock, C. E. S., Ertel, A., Jeong, K. W., Stallcup, M. R., Addya, S., et al. (2013) Aberrant BAF57 signaling facilitates prometastatic phenotypes. *Clin Cancer Res* **19**, 2657-67

Barash, Y., Calarco, J. A., Gao, W., Pan, Q., Wang, X., Shai, O., et al. (2010) Deciphering the splicing code. *Nature* **465**, 53-9

Barbosa-Morais, N. L., Irimia, M., Pan, Q., Xiong, H. Y., Gueroussov, S., Lee, L. J., et al. (2012) The evolutionary landscape of alternative splicing in vertebrate species. *Science* **338**, 1587-93

Barde, Y. A., Edgar, D., Thoenen, H. (1982) Purification of a new neurotrophic factor from mammalian brain. *EMBO J* **1**, 549-53

Batsché, E., Yaniv, M., Muchardt, C. (2006) The human SWI/SNF subunit Brm is a regulator of alternative splicing. *Nat Struct Mol Biol* **13**, 22-9

Battaglioli, E., Andrés, M. E., Rose, D. W., Chenoweth, J. G., Rosenfeld, M. G., Anderson, M. E., et al. (2002) REST repression of neuronal genes requires components of the hSWI.SNF complex. *J Biol Chem* **277**, 41038-45

Bauman, J. A., Kole, R. (2011) Modulation of RNA splicing as a potential treatment for cancer. *Bioeng Bugs* **2**, 125-8

Bauman, J. A., Li, S.-D., Yang, A., Huang, L., Kole, R. (2010) Anti-tumor activity of splice-switching oligonucleotides. *Nucleic Acids Res* **38**, 8348-56

Belandia, B., Orford, R. L., Hurst, H. C., Parker, M. G. (2002) Targeting of SWI/SNF chromatin remodelling complexes to estrogen-responsive genes. *EMBO J* **21**, 4094-103

Bell, B., Scheer, E., Tora, L. (2001) Identification of hTAF(II)80 delta links apoptotic signaling pathways to transcription factor TFIID function. *Mol Cell* **8**, 591-600

Bentley, D. L. (2014) Coupling mRNA processing with transcription in time and space. *Nat Rev Genet* **15**, 163-75

Berget, S. M., Moore, C., Sharp, P. A. (1977) Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proc Natl Acad Sci U S A* **74**, 3171-5

Beyer, A. L., Osheim, Y. N. (1988) Splice site selection, rate of splicing, and alternative splicing on nascent transcripts. *Genes Dev* **2**, 754-65

Binder, D. K., Croll, S. D., Gall, C. M., Scharfman, H. E. (2001) BDNF and epilepsy: too much of a good thing?. *Trends Neurosci* **24**, 47-53

Bishop, J. F., Mueller, G. P., Mouradian, M. M. (1994) Alternate 5' exons in the rat brain-derived neurotrophic factor gene: differential patterns of expression across brain regions. *Brain Res Mol Brain Res* **26**, 225-32

Black, D. L. (2003) Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem* **72**, 291-336

Blencowe, B. J. (2006) Alternative splicing: new insights from global analyses. *Cell* **126**, 37-47

Bonomi, S., Gallo, S., Catillo, M., Pignataro, D., Biamonti, G., Ghigna, C. (2013) Oncogenic alternative splicing switches: role in cancer progression and prospects for therapy. *Int J Cell Biol* **2013**, 962038

Bourbon, H.-M., Aguilera, A., Ansari, A. Z., Asturias, F. J., Berk, A. J., Bjorklund, S., et al. (2004) A unified nomenclature for protein subunits of

mediator complexes linking transcriptional regulators to RNA polymerase II. *Mol Cell* **14**, 553-7

Bredy, T. W., Wu, H., Crego, C., Zellhoefer, J., Sun, Y. E., Barad, M. (2007) Histone modifications around individual BDNF gene promoters in prefrontal cortex are associated with extinction of conditioned fear. *Learn Mem* **14**, 268-76

Brinkman, B. M. N. (2004) Splice variants as cancer biomarkers. *Clin Biochem* **37**, 584-94

Brugiolo, M., Herzel, L., Neugebauer, K. M. (2013) Counting on co-transcriptional splicing. *F1000Prime Rep* **5**, 9

Brunkhorst, A., Neuman, T., Hall, A., Arenas, E., Bartfai, T., Hermanson, O., et al. (2004) Novel isoforms of the TFIID subunit TAF4 modulate nuclear receptor-mediated transcriptional activity. *Biochem Biophys Res Commun* **325**, 574-9

Buljan, M., Chalancon, G., Eustermann, S., Wagner, G. P., Fuxreiter, M., Bateman, A., et al. (2012) Tissue-specific splicing of disordered segments that embed binding motifs rewires protein interaction networks. *Mol Cell* **46**, 871-83

Busch, A., Hertel, K. J. (2012) Evolution of SR protein and hnRNP splicing regulatory factors. *Wiley Interdiscip Rev RNA* **3**, 1-12

Calarco, J. A., Superina, S., O'Hanlon, D., Gabut, M., Raj, B., Pan, Q., et al. (2009) Regulation of vertebrate nervous system alternative splicing and development by an SR-related protein. *Cell* **138**, 898-910

Cao, X., Geradts, J., Dewhirst, M. W., Lo, H.-W. (2012) Upregulation of VEGF-A and CD24 gene expression by the tGLI1 transcription factor contributes to the aggressive behavior of breast cancer cells. *Oncogene* **31**, 104-15

Carpenter, R. L., Lo, H.-W. (2012) Hedgehog pathway and GLI1 isoforms in human cancer. *Discov Med* **13**, 105-13

Castrén, E., Rantamäki, T. (2010) The role of BDNF and its receptors in depression and antidepressant drug action: Reactivation of developmental

plasticity. *Dev Neurobiol* **70**, 289-97

Chen, B.-Y., Wang, X., Wang, Z.-Y., Wang, Y.-Z., Chen, L.-W., Luo, Z.-J. (2013) Brain-derived neurotrophic factor stimulates proliferation and differentiation of neural stem cells, possibly by triggering the Wnt/β-catenin signaling pathway. *J Neurosci Res* **91**, 30-41

Chen, J., Archer, T. K. (2005) Regulating SWI/SNF subunit levels via protein-protein interactions and proteasomal degradation: BAF155 and BAF170 limit expression of BAF57. *Mol Cell Biol* **25**, 9016-27

Chen, J., Sun, M., Kent, W. J., Huang, X., Xie, H., Wang, W., et al. (2004) Over 20\% of human transcripts might form sense-antisense pairs. *Nucleic Acids Res* **32**, 4812-20

Chen, L. L., Sabripour, M., Wu, E. F., Prieto, V. G., Fuller, G. N., Frazier, M. L. (2005) A mutation-created novel intra-exonic pre-mRNA splice site causes constitutive activation of KIT in human gastrointestinal stromal tumors. *Oncogene* **24**, 4271-80

Chen, M., Manley, J. L. (2009) Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nat Rev Mol Cell Biol* **10**, 741-54

Chen, W. G., Chang, Q., Lin, Y., Meissner, A., West, A. E., Griffith, E. C., et al. (2003) Derepression of BDNF transcription involves calcium-dependent phosphorylation of MeCP2. *Science* **302**, 885-9

Chiaruttini, C., Sonego, M., Baj, G., Simonato, M., Tongiorgi, E. (2008) BDNF mRNA splice variants display activity-dependent targeting to distinct hippocampal laminae. *Mol Cell Neurosci* **37**, 11-9

Chih, B., Gollan, L., Scheiffele, P. (2006) Alternative splicing controls selective trans-synaptic interactions of the neuroligin-neurexin complex. *Neuron* **51**, 171-8

Choi, K.-Y., Yoo, M., Han, J.-H. (2015) Toward understanding the role of the neuron-specific BAF chromatin remodeling complex in memory formation. *Exp Mol Med* **47**, e155

Choi, S. S., Bush, E. C., Lahn, B. T. (2006) Different classes of tissue-specific genes show different levels of noncoding conservation. *Genomics* **87**, 433-6

Chow, L. T., Gelinas, R. E., Broker, T. R., Roberts, R. J. (1977) An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell* **12**, 1-8

Cieply, B., Carstens, R. P. (2015) Functional roles of alternative splicing factors in human disease. *Wiley Interdiscip Rev RNA* **6**, 311-26

Conaway, R. C., Conaway, J. W. (1993) General initiation factors for RNA polymerase II. *Annu Rev Biochem* **62**, 161-90

Coulson, J. M., Edgson, J. L., Woll, P. J., Quinn, J. P. (2000) A splice variant of the neuron-restrictive silencer factor repressor is expressed in small cell lung cancer: a potential role in derepression of neuroendocrine genes and a useful clinical marker. *Cancer Res* **60**, 1840-4

Coutinho-Mansfield, G. C., Xue, Y., Zhang, Y., Fu, X.-D. (2007) PTB/nPTB switch: a post-transcriptional mechanism for programming neuronal differentiation. *Genes Dev* **21**, 1573-7

D'Alessio, J. A., Wright, K. J., Tjian, R. (2009) Shifting players and paradigms in cell-specific transcription. *Mol Cell* **36**, 924-31

Daoud, R., Da Penha Berzaghi, M., Siedler, F., Hübener, M., Stamm, S. (1999) Activity-dependent regulation of alternative splicing patterns in the rat brain. *Eur J Neurosci* **11**, 788-802

Das, S., Krainer, A. R. (2014) Emerging functions of SRSF1, splicing factor and oncoprotein, in RNA metabolism and cancer. *Mol Cancer Res* **12**, 1195-204

David, C. J., Boyne, A. R., Millhouse, S. R., Manley, J. L. (2011) The RNA polymerase II C-terminal domain promotes splicing activation through recruitment of a U2AF65-Prp19 complex. *Genes Dev* **25**, 972-83

David, C. J., Manley, J. L. (2010) Alternative pre-mRNA splicing regulation in cancer: pathways and programs unhinged. *Genes Dev* **24**, 2343-64

Dieni, S., Matsumoto, T., Dekkers, M., Rauskolb, S., Ionescu, M. S., Deogracias, R., et al. (2012) BDNF and its pro-peptide are stored in presynaptic dense core vesicles in brain neurons. *J Cell Biol* **196**, 775-88

Douglas, A. G. L., Wood, M. J. A. (2011) RNA splicing: disease and therapy. *Brief Funct Genomics* **10**, 151-64

Dreyfuss, G., Matunis, M. J., Piñol-Roma, S., Burd, C. G. (1993) hnRNP proteins and the biogenesis of mRNA. *Annu Rev Biochem* **62**, 289-321

Dujardin, G., Lafaille, C., de la Mata, M., Marasco, L. E., Muñoz, M. J., Le Jossic-Corcos, C., et al. (2014) How slow RNA polymerase II elongation favors alternative exon skipping. *Mol Cell* **54**, 683-90

Dutertre, M., Vagner, S., Auboeuf, D. (2010) Alternative splicing and breast cancer. *RNA Biol* **7**, 403-11

Easton, D. F., Bishop, D. T., Ford, D., Crockford, G. P. (1993) Genetic linkage analysis in familial breast and ovarian cancer: results from 214 families. The Breast Cancer Linkage Consortium. *Am J Hum Genet* **52**, 678-701

Ellis, J. D., Barrios-Rodiles, M., Colak, R., Irimia, M., Kim, T., Calarco, J. A., et al. (2012) Tissue-specific alternative splicing remodels protein-protein interaction networks. *Mol Cell* **46**, 884-92

Endoh, M., Zhu, W., Hasegawa, J., Watanabe, H., Kim, D.-K., Aida, M., et al. (2004) Human Spt6 stimulates transcription elongation by RNA polymerase II in vitro. *Mol Cell Biol* **24**, 3324-36

Eperon, L. P., Graham, I. R., Griffiths, A. D., Eperon, I. C. (1988) Effects of RNA secondary structure on alternative splicing of pre-mRNA: is folding limited to a region behind the transcribing RNA polymerase?. *Cell* **54**, 393-401

Eswaran, J., Horvath, A., Godbole, S., Reddy, S. D., Mudvari, P., Ohshiro, K., et al. (2013) RNA sequencing of cancer reveals novel splicing alterations. *Sci Rep* **3**, 1689

Fang, H., Chartier, J., Sodja, C., Desbois, A., Ribecco-Lutkiewicz, M., Walker, P. R., et al. (2003) Transcriptional activation of the human brain-derived neurotrophic factor gene promoter III by dopamine signaling in NT2/N

neurons. *J Biol Chem* **278**, 26401-9

Faralli, H., Martin, E., Coré, N., Liu, Q.-C., Filippi, P., Dilworth, F. J., et al. (2011) Teashirt-3, a novel regulator of muscle differentiation, associates with BRG1-associated factor 57 (BAF57) to inhibit myogenin gene expression. *J Biol Chem* **286**, 23498-510

Faustino, N. A., Cooper, T. A. (2003) Pre-mRNA splicing and human disease. *Genes Dev* **17**, 419-37

Ferrer, I., Goutan, E., Marín, C., Rey, M. J., Ribalta, T. (2000) Brain-derived neurotrophic factor in Huntington disease. *Brain Res* **866**, 257-61

Flavell, S. W., Greenberg, M. E. (2008) Signaling mechanisms linking neuronal activity to gene expression and plasticity of the nervous system. *Annu Rev Neurosci* **31**, 563-90

Fong, N., Bird, G., Vigneron, M., Bentley, D. L. (2003) A 10 residue motif at the C-terminus of the RNA pol II CTD is required for transcription, splicing and 3' end processing. *EMBO J* **22**, 4274-82

Fong, N., Kim, H., Zhou, Y., Ji, X., Qiu, J., Saldi, T., et al. (2014) Pre-mRNA splicing is facilitated by an optimal RNA polymerase II elongation rate. *Genes Dev* **28**, 2663-76

Fong, Y. W., Zhou, Q. (2001) Stimulatory effect of splicing factors on transcriptional elongation. *Nature* **414**, 929-33

Freiman, R. N. (2009) Specific variants of general transcription factors regulate germ cell development in diverse organisms. *Biochim Biophys Acta* **1789**, 161-6

Fuchikami, M., Yamamoto, S., Morinobu, S., Takei, S., Yamawaki, S. (2010) Epigenetic regulation of BDNF gene in response to stress. *Psychiatry Investig* **7**, 251-6

Fulgenzi, G., Tomassoni-Ardori, F., Babini, L., Becker, J., Barrick, C., Puverel, S., et al. (2015) BDNF modulates heart contraction force and long-term homeostasis through truncated TrkB.T1 receptor activation. *J Cell Biol* ,

Gamazon, E. R., Stranger, B. E. (2014) Genomics of alternative splicing: evolution, development and pathophysiology. *Hum Genet* **133**, 679-87

Gambino, G., Tancredi, M., Falaschi, E., Aretini, P., Caligo, M. A. (2015) Characterization of three alternative transcripts of the BRCA1 gene in patients with breast cancer and a family history of breast and/or ovarian cancer who tested negative for pathogenic mutations. *Int J Mol Med* **35**, 950-6

Gao, X., Tate, P., Hu, P., Tjian, R., Skarnes, W. C., Wang, Z. (2008) ES cell pluripotency and germ-layer formation require the SWI/SNF chromatin remodeling component BAF250a. *Proc Natl Acad Sci U S A* **105**, 6656-61

Garcés, M. F., Sanchez, E., Torres-Sierra, A. L., Ruíz-Parra, A. I., Angel-Müller, E., Alzate, J. P., et al. (2014) Brain-derived neurotrophic factor is expressed in rat and human placenta and its serum levels are similarly regulated throughout pregnancy in both species. *Clin Endocrinol (Oxf)* **81**, 141-51

Goossens, N., Nakagawa, S., Sun, X., Hoshida, Y. (2015) Cancer biomarker discovery and validation. *Transl Cancer Res* **4**, 256-269

Grabowski, P. (2011) Alternative splicing takes shape during neuronal development. *Curr Opin Genet Dev* **21**, 388-94

Grabowski, P. J., Black, D. L. (2001) Alternative RNA splicing in the nervous system. *Prog Neurobiol* **65**, 289-308

Grünberg, S., Hahn, S. (2013) Structural insights into transcription initiation by RNA polymerase II. *Trends Biochem Sci* **38**, 603-11

Guttery, D. S., Shaw, J. A., Lloyd, K., Pringle, J. H., Walker, R. A. (2010) Expression of tenascin-C and its isoforms in the breast. *Cancer Metastasis Rev* **29**, 595-606

Hah, N., Kolkman, A., Ruhl, D. D., Pijnappel, W. W. M. P., Heck, A. J. R., Timmers, H. T. M., et al. (2010) A role for BAF57 in cell cycle-dependent transcriptional regulation by the SWI/SNF chromatin remodeling complex. *Cancer Res* **70**, 4402-11

Hamid, F. M., Makeyev, E. V. (2014) Emerging functions of alternative splicing coupled with nonsense-mediated decay. *Biochem Soc Trans* **42**, 1168-73

Harikrishnan, K. N., Chow, M. Z., Baker, E. K., Pal, S., Bassal, S., Brasacchio, D., et al. (2005) Brahma links the SWI/SNF chromatin-remodeling complex with MeCP2-dependent transcriptional silencing. *Nat Genet* **37**, 254-64

Haugsten, E. M., Wiedlocha, A., Olsnes, S., Wesche, J. (2010) Roles of fibroblast growth factor receptors in carcinogenesis. *Mol Cancer Res* **8**, 1439-52

Heerssen, H. M., Segal, R. A. (2002) Location, location, location: a spatial view of neurotrophin signal transduction. *Trends Neurosci* **25**, 160-5

Hempstead, B. L. (2015) Brain-Derived Neurotrophic Factor: Three Ligands, Many Actions. *Trans Am Clin Climatol Assoc* **126**, 9-19

Herzfeld, T., Nolte, D., Müller, U. (2007) Structural and functional analysis of the human TAF1/DYT3 multiple transcript system. *Mamm Genome* **18**, 787-95

Hiester, B. G., Galati, D. F., Salinas, P. C., Jones, K. R. (2013) Neurotrophin and Wnt signaling cooperatively regulate dendritic spine formation. *Mol Cell Neurosci* **56**, 115-27

Ho, L., Jothi, R., Ronan, J. L., Cui, K., Zhao, K., Crabtree, G. R. (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**, 5187-91

Hofer, M., Pagliusi, S. R., Hohn, A., Leibrock, J., Barde, Y. A. (1990) Regional distribution of brain-derived neurotrophic factor mRNA in the adult mouse brain. *EMBO J* **9**, 2459-64

Hohn, A., Leibrock, J., Bailey, K., Barde, Y. A. (1990) Identification and characterization of a novel member of the nerve growth factor/brain-derived neurotrophic factor family. *Nature* **344**, 339-41

Hoskins, A. A., Moore, M. J. (2012) The spliceosome: a flexible, reversible macromolecular machine. *Trends Biochem Sci* **37**, 179-88

Howeedy, A. A., Virtanen, I., Laitinen, L., Gould, N. S., Koukoulis, G. K., Gould, V. E. (1990) Differential distribution of tenascin in the normal, hyperplastic, and neoplastic breast. *Lab Invest* **63**, 798-806

Howells, D. W., Porritt, M. J., Wong, J. Y., Batchelor, P. E., Kalnins, R., Hughes, A. J., et al. (2000) Reduced BDNF mRNA expression in the Parkinson's disease substantia nigra. *Exp Neurol* **166**, 127-35

Huelga, S. C., Vu, A. Q., Arnold, J. D., Liang, T. Y., Liu, P. P., Yan, B. Y., et al. (2012) Integrative genome-wide analysis reveals cooperative regulation of alternative splicing by hnRNP proteins. *Cell Rep* **1**, 167-78

Ibrahim, E. C., Hims, M. M., Shomron, N., Burge, C. B., Slaugenhaupt, S. A., Reed, R. (2007) Weak definition of IKBKAP exon 20 leads to aberrant splicing in familial dysautonomia. *Hum Mutat* **28**, 41-53

Ito, T., Watanabe, H., Yamamichi, N., Kondo, S., Tando, T., Haraguchi, T., et al. (2008) Brm transactivates the telomerase reverse transcriptase (TERT) gene and modulates the splicing patterns of its transcripts in concert with p54(nrb). *Biochem J* **411**, 201-9

Jaiswal, P. K., Goel, A., Mittal, R. D. (2015) Survivin: A molecular biomarker in cancer. *Indian J Med Res* **141**, 389-97

Jambaldorj, J., Makino, S., Munkhbat, B., Tamiya, G. (2012) Sustained expression of a neuron-specific isoform of the Taf1 gene in development stages and aging in mice. *Biochem Biophys Res Commun* **425**, 273-7

Jensen, K. B., Dredge, B. K., Stefani, G., Zhong, R., Buckanovich, R. J., Okano, H. J., et al. (2000) Nova-1 regulates neuron-specific alternative splicing and is essential for neuronal viability. *Neuron* **25**, 359-71

Ji, X., Fu, X.-D. (2012) The mediator couples transcription and splicing. *Mol Cell* **45**, 433-4

Jiang, X., Tian, F., Du, Y., Copeland, N. G., Jenkins, N. A., Tessarollo, L., et al. (2008) BHLHB2 controls Bdnf promoter 4 activity and neuronal

excitability. *J Neurosci* **28**, 1118-30

Johnson, J. M., Castle, J., Garrett-Engele, P., Kan, Z., Loerch, P. M., Armour, C. D., et al. (2003) Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science* **302**, 2141-4

Jones, K. R., Fariñas, I., Backus, C., Reichardt, L. F. (1994) Targeted disruption of the BDNF gene perturbs brain and sensory neuron development but not motor neuron development. *Cell* **76**, 989-99

Jonkers, I., Lis, J. T. (2015) Getting up to speed with transcription elongation by RNA polymerase II. *Nat Rev Mol Cell Biol* **16**, 167-77

Kadakkuzha, B. M., Liu, X.-A., Narvaez, M., Kaye, A., Akhmedov, K., Puthanveettil, S. V. (2014) Asymmetric localization of natural antisense RNA of neuropeptide sensorin in Aplysia sensory neurons during aging and activity. *Front Genet* **5**, 84

Kalinowska-Lyszczarz, A., Losy, J. (2012) The role of neurotrophins in multiple sclerosis-pathological and clinical implications. *Int J Mol Sci* **13**, 13713-25

Kalsotra, A., Cooper, T. A. (2011) Functional consequences of developmentally regulated alternative splicing. *Nat Rev Genet* **12**, 715-29

Kamtchueng, C., Stébenne, M.-É., Delannoy, A., Wilhelm, E., Léger, H., Benecke, A. G., et al. (2014) Alternative splicing of TAF6: downstream transcriptome impacts and upstream RNA splice control elements. *PLoS One* **9**, e102399

Karni, R., de Stanchina, E., Lowe, S. W., Sinha, R., Mu, D., Krainer, A. R. (2007) The gene encoding the splicing factor SF2/ASF is a proto-oncogene. *Nat Struct Mol Biol* **14**, 185-93

Katzenberger, R. J., Marengo, M. S., Wassarman, D. A. (2006) ATM and ATR pathways signal alternative splicing of Drosophila TAF1 pre-mRNA in response to DNA damage. *Mol Cell Biol* **26**, 9256-67

Katzenberger, R. J., Marengo, M. S., Wassarman, D. A. (2009) Control of alternative splicing by signal-dependent degradation of splicing-

regulatory proteins. *J Biol Chem* **284**, 10737-46

Kazantseva, J., Kivil, A., Tints, K., Kazantseva, A., Neuman, T., Palm, K. (2013) Alternative splicing targeting the hTAF4-TAFH domain of TAF4 represses proliferation and accelerates chondrogenic differentiation of human mesenchymal stem cells. *PLoS One* **8**, e74799

Kazantseva, J., Tints, K., Neuman, T., Palm, K. (2015) TAF4 Controls Differentiation of Human Neural Progenitor Cells Through hTAF4-TAFH Activity. *J Mol Neurosci* **55**, 160-6

Keppetipola, N., Sharma, S., Li, Q., Black, D. L. (2012) Neuronal regulation of pre-mRNA splicing by polypyrimidine tract binding proteins, PTBP1 and PTBP2. *Crit Rev Biochem Mol Biol* **47**, 360-78

Keren, H., Lev-Maor, G., Ast, G. (2010) Alternative splicing and evolution: diversification, exon definition and function. *Nat Rev Genet* **11**, 345-55

Khan, S., Bennit, H. F., Turay, D., Perez, M., Mirshahidi, S., Yuan, Y., et al. (2014) Early diagnostic value of survivin and its alternative splice variants in breast cancer. *BMC Cancer* **14**, 176

Kim, S., Kim, H., Fong, N., Erickson, B., Bentley, D. L. (2011) Pre-mRNA splicing is a determinant of histone H3K36 methylation. *Proc Natl Acad Sci U S A* **108**, 13564-9

Kim, Y.-J., Kim, H.-S. (2012) Alternative splicing and its impact as a cancer diagnostic marker. *Genomics Inform* **10**, 74-80

Kingston, R. E., Narlikar, G. J. (1999) ATP-dependent remodeling and acetylation as regulators of chromatin fluidity. *Genes Dev* **13**, 2339-52

Kiskinis, E., García-Pedrero, J. M., Villaronga, M. A., Parker, M. G., Belandia, B. (2006) Identification of BAF57 mutations in human breast cancer cell lines. *Breast Cancer Res Treat* **98**, 191-8

Kobeissy, F. H., Gulbakan, B., Alawieh, A., Karam, P., Zhang, Z., Guingab-Cagmat, J. D., et al. (2014) Post-genomics nanotechnology is gaining momentum: nanoproteomics and applications in life sciences. *OMICS* **18**, 111-31

Koppel, I., Tuvikene, J., Lekk, I., Timmusk, T. (2015) Efficient use of a translation start codon in BDNF exon I. *J Neurochem* **134**, 1015-25

Kornberg, R. D. (2007) The molecular basis of eucaryotic transcription. *Cell Death Differ* **14**, 1989-97

Kornblihtt, A. R., Schor, I. E., Alló, M., Dujardin, G., Petrillo, E., Muñoz, M. J. (2013) Alternative splicing: a pivotal step between eukaryotic transcription and translation. *Nat Rev Mol Cell Biol* **14**, 153-65

Korneev, S., O'Shea, M. (2005) Natural antisense RNAs in the nervous system. *Rev Neurosci* **16**, 213-22

Korneev, S. A., Kemenes, I., Bettini, N. L., Kemenes, G., Staras, K., Benjamin, P. R., et al. (2013) Axonal trafficking of an antisense RNA transcribed from a pseudogene is regulated by classical conditioning. *Sci Rep* **3**, 1027

Korneev, S. A., Korneeva, E. I., Lagarkova, M. A., Kiselev, S. L., Critchley, G., O'Shea, M. (2008) Novel noncoding antisense RNA transcribed from human anti-NOS2A locus is differentially regulated during neuronal differentiation of embryonic stem cells. *RNA* **14**, 2030-7

Korte, M., Carroll, P., Wolf, E., Brem, G., Thoenen, H., Bonhoeffer, T. (1995) Hippocampal long-term potentiation is impaired in mice lacking brain-derived neurotrophic factor. *Proc Natl Acad Sci U S A* **92**, 8856-60

Krawczak, M., Thomas, N. S. T., Hundrieser, B., Mort, M., Wittig, M., Hampe, J., et al. (2007) Single base-pair substitutions in exon-intron junctions of human genes: nature, distribution, and consequences for mRNA splicing. *Hum Mutat* **28**, 150-8

Kwon, C. S., Wagner, D. (2007) Unwinding chromatin for development and growth: a few genes at a time. *Trends Genet* **23**, 403-12

Lam, B. J., Hertel, K. J. (2002) A general role for splicing enhancers in exon definition. *RNA* **8**, 1233-41

Lapuk, A., Marr, H., Jakkula, L., Pedro, H., Bhattacharya, S., Purdom, E., et al. (2010) Exon-level microarray analyses identify alternative splicing programs in breast cancer. *Mol Cancer Res* **8**, 961-74

Lareau, L. F., Brenner, S. E. (2015) Regulation of splicing factors by alternative splicing and NMD is conserved between kingdoms yet evolutionarily flexible. *Mol Biol Evol* ,

Lareau, L. F., Brooks, A. N., Soergel, D. A. W., Meng, Q., Brenner, S. E. (2007) The coupling of alternative splicing and nonsense-mediated mRNA decay. *Adv Exp Med Biol* **623**, 190-211

Lau, A. G., Irier, H. A., Gu, J., Tian, D., Ku, L., Liu, G., et al. (2010) Distinct 3'UTRs differentially regulate activity-dependent translation of brain-derived neurotrophic factor (BDNF). *Proc Natl Acad Sci U S A* **107**, 15945-50

Lee, J.-A., Tang, Z.-Z., Black, D. L. (2009) An inducible change in Fox-1/A2BP1 splicing modulates the alternative splicing of downstream neuronal target exons. *Genes Dev* **23**, 2284-93

Leibrock, J., Lottspeich, F., Hohn, A., Hofer, M., Hengerer, B., Masiakowski, P., et al. (1989) Molecular cloning and expression of brain-derived neurotrophic factor. *Nature* **341**, 149-52

Lemon, B., Tjian, R. (2000) Orchestrated response: a symphony of transcription factors for gene control. *Genes Dev* **14**, 2551-69

Lessard, J., Wu, J. I., Ranish, J. A., Wan, M., Winslow, M. M., Staahl, B. T., et al. (2007) An essential switch in subunit composition of a chromatin remodeling complex during neural development. *Neuron* **55**, 201-15

Lessmann, V., Gottmann, K., Malcangio, M. (2003) Neurotrophin secretion: current facts and future prospects. *Prog Neurobiol* **69**, 341-74

Lewis, B. P., Green, R. E., Brenner, S. E. (2003) Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proc Natl Acad Sci U S A* **100**, 189-92

Li, Q., Lee, J.-A., Black, D. L. (2007) Neuronal regulation of alternative pre-mRNA splicing. *Nat Rev Neurosci* **8**, 819-31

Li, Q., Zheng, S., Han, A., Lin, C.-H., Stoilov, P., Fu, X.-D., et al. (2014) The splicing regulator PTBP2 controls a program of embryonic splicing

required for neuronal maturation. *Elife* **3**, e01201

Li, S., Liu, C., Li, N., Hao, T., Han, T., Hill, D. E., et al. (2008) Genome-wide coactivation analysis of PGC-1alpha identifies BAF60a as a regulator of hepatic lipid metabolism. *Cell Metab* **8**, 105-17

Li, X. S., Trojer, P., Matsumura, T., Treisman, J. E., Tanese, N. (2010) Mammalian SWI/SNF--a subunit BAF250/ARID1 is an E3 ubiquitin ligase that targets histone H2B. *Mol Cell Biol* **30**, 1673-88

Lickert, H., Takeuchi, J. K., Von Both, I., Walls, J. R., McAuliffe, F., Adamson, S. L., et al. (2004) Baf60c is essential for function of BAF chromatin remodelling complexes in heart development. *Nature* **432**, 107-12

Lin, S., Coutinho-Mansfield, G., Wang, D., Pandit, S., Fu, X.-D. (2008) The splicing factor SC35 has an active role in transcriptional elongation. *Nat Struct Mol Biol* **15**, 819-26

Lin, S., Fu, X.-D. (2007) SR proteins and related factors in alternative splicing. *Adv Exp Med Biol* **623**, 107-22

Ling, M. H. T., Ban, Y., Wen, H., Wang, S. M., Ge, S. X. (2013) Conserved expression of natural antisense transcripts in mammals. *BMC Genomics* **14**, 243

Liu, Q.-R., Lu, L., Zhu, X.-G., Gong, J.-P., Shaham, Y., Uhl, G. R. (2006) Rodent BDNF genes, novel promoters, novel splice variants, and regulation by cocaine. *Brain Res* **1067**, 1-12

Liu, Q.-R., Walther, D., Drgon, T., Polesskaya, O., Lesnick, T. G., Strain, K. J., et al. (2005) Human brain derived neurotrophic factor (BDNF) genes, splicing patterns, and assessments of associations with substance abuse and Parkinson's Disease. *Am J Med Genet B Neuropsychiatr Genet* **134B**, 93-103

Lixia, M., Zhijian, C., Chao, S., Chaojiang, G., Congyi, Z. (2007) Alternative splicing of breast cancer associated gene BRCA1 from breast cancer cell line. *J Biochem Mol Biol* **40**, 15-21

Lo, H.-W., Zhu, H., Cao, X., Aldrich, A., Ali-Osman, F. (2009) A novel splice variant of GLI1 that promotes glioblastoma cell migration and invasion.

Cancer Res **69**, 6790-8

Lorson, C. L., Hahnen, E., Androphy, E. J., Wirth, B. (1999) A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. *Proc Natl Acad Sci U S A* **96**, 6307-11

Lubin, F. D., Ren, Y., Xu, X., Anderson, A. E. (2007) Nuclear factor-kappa B regulates seizure threshold and gene transcription following convulsant stimulation. *J Neurochem* **103**, 1381-95

Lubin, F. D., Roth, T. L., Sweatt, J. D. (2008) Epigenetic regulation of BDNF gene transcription in the consolidation of fear memory. *J Neurosci* **28**, 10576-86

Luco, R. F., Allo, M., Schor, I. E., Kornblihtt, A. R., Misteli, T. (2011) Epigenetics in alternative pre-mRNA splicing. *Cell* **144**, 16-26

Luco, R. F., Pan, Q., Tominaga, K., Blencowe, B. J., Pereira-Smith, O. M., Misteli, T. (2010) Regulation of alternative splicing by histone modifications. *Science* **327**, 996-1000

Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F., Richmond, T. J. (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**, 251-60

Luse, D. S. (2013) The RNA polymerase II preinitiation complex: Through what pathway is the complex assembled?. *Transcription* **5**,

Maguire, S. L., Leonidou, A., Wai, P., Marchiò, C., Ng, C. K., Sapino, A., et al. (2015) SF3B1 mutations constitute a novel therapeutic target in breast cancer. *J Pathol* **235**, 571-80

Maisonpierre, P. C., Belluscio, L., Friedman, B., Alderson, R. F., Wiegand, S. J., Furth, M. E., et al. (1990) NT-3, BDNF, and NGF in the developing rat nervous system: parallel as well as reciprocal patterns of expression. *Neuron* **5**, 501-9

Makeyev, E. V., Zhang, J., Carrasco, M. A., Maniatis, T. (2007) The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing. *Mol Cell* **27**, 435-48

Malik, S., Roeder, R. G. (2010) The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation. *Nat Rev Genet* **11**, 761-72

Martinowich, K., Hattori, D., Wu, H., Fouse, S., He, F., Hu, Y., et al. (2003) DNA methylation-related chromatin remodeling in activity-dependent BDNF gene regulation. *Science* **302**, 890-3

Masuko, K., Okazaki, S., Satoh, M., Tanaka, G., Ikeda, T., Torii, R., et al. (2012) Anti-tumor effect against human cancer xenografts by a fully human monoclonal antibody to a variant 8-epitope of CD44R1 expressed on cancer stem cells. *PLoS One* **7**, e29728

Matera, A. G., Wang, Z. (2014) A day in the life of the spliceosome. *Nat Rev Mol Cell Biol* **15**, 108-21

Matoulkova, E., Michalova, E., Vojtesek, B., Hrstka, R. (2012) The role of the 3' untranslated region in post-transcriptional regulation of protein expression in mammalian cells. *RNA Biol* **9**, 563-76

Mazin, P., Xiong, J., Liu, X., Yan, Z., Zhang, X., Li, M., et al. (2013) Widespread splicing changes in human brain development and aging. *Mol Syst Biol* **9**, 633

McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., et al. (1997) The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature* **385**, 357-61

Mehner, C., Miller, E., Khauv, D., Nassar, A., Oberg, A. L., Bamlet, W. R., et al. (2014) Tumor cell-derived MMP3 orchestrates Rac1b and tissue alterations that promote pancreatic adenocarcinoma. *Mol Cancer Res* **12**, 1430-9

Meininghaus, M., Chapman, R. D., Horndasch, M., Eick, D. (2000) Conditional expression of RNA polymerase II in mammalian cells. Deletion of the carboxyl-terminal domain of the large subunit affects early steps in transcription. *J Biol Chem* **275**, 24375-82

Mendell, J. T., Sharifi, N. A., Meyers, J. L., Martinez-Murillo, F., Dietz, H. C. (2004) Nonsense surveillance regulates expression of diverse classes of

mammalian transcripts and mutes genomic noise. *Nat Genet* **36**, 1073-8

Meng, Z.-X., Li, S., Wang, L., Ko, H. J., Lee, Y., Jung, D. Y., et al. (2013) Baf60c drives glycolytic metabolism in the muscle and improves systemic glucose homeostasis through Deptor-mediated Akt activation. *Nat Med* **19**, 640-5

Merkin, J., Russell, C., Chen, P., Burge, C. B. (2012) Evolutionary dynamics of gene and isoform regulation in Mammalian tissues. *Science* **338**, 1593-9

Metcalf, C. E., Wassarman, D. A. (2007) Nucleolar colocalization of TAF1 and testis-specific TAFs during Drosophila spermatogenesis. *Dev Dyn* **236**, 2836-43

Michaelis, E. K. (1993) Two different families of NMDA receptors in mammalian brain: physiological function and role in neuronal development and degeneration. *Adv Exp Med Biol* **341**, 119-28

Millhouse, S., Manley, J. L. (2005) The C-terminal domain of RNA polymerase II functions as a phosphorylation-dependent splicing activator in a heterologous protein. *Mol Cell Biol* **25**, 533-44

Miura, P., Sanfilippo, P., Shenker, S., Lai, E. C. (2014) Alternative polyadenylation in the nervous system: to what lengths will 3' UTR extensions take us?. *Bioessays* **36**, 766-77

Modrek, B., Lee, C. (2002) A genomic view of alternative splicing. *Nat Genet* **30**, 13-9

Montes, M., Becerra, S., Sánchez-Álvarez, M., Suñé, C. (2012) Functional coupling of transcription and splicing. *Gene* **501**, 104-17

Morris, D. P., Greenleaf, A. L. (2000) The splicing factor, Prp40, binds the phosphorylated carboxyl-terminal domain of RNA polymerase II. *J Biol Chem* **275**, 39935-43

Mount, S. M. (1982) A catalogue of splice junction sequences. *Nucleic Acids Res* **10**, 459-72

Mowla, S. J., Farhadi, H. F., Pareek, S., Atwal, J. K., Morris, S. J., Seidah, N. G., et al. (2001) Biosynthesis and post-translational processing of the

precursor to brain-derived neurotrophic factor. *J Biol Chem* **276**, 12660-6

Murakami, K., Elmlund, H., Kalisman, N., Bushnell, D. A., Adams, C. M., Azubel, M., et al. (2013) Architecture of an RNA polymerase II transcription pre-initiation complex. *Science* **342**, 1238724

Murray, P. S., Holmes, P. V. (2011) An overview of brain-derived neurotrophic factor and implications for excitotoxic vulnerability in the hippocampus. *Int J Pept* **2011**, 654085

Nakai, K., Sakamoto, H. (1994) Construction of a novel database containing aberrant splicing mutations of mammalian genes. *Gene* **141**, 171-7

Nancy, M.-M., Nora, R.-M., Rebeca, M.-C. (2015) Peptidic tools applied to redirect alternative splicing events. *Peptides* **67**, 1-11

Naryshkin, N. A., Weetall, M., Dakka, A., Narasimhan, J., Zhao, X., Feng, Z., et al. (2014) Motor neuron disease. SMN2 splicing modifiers improve motor function and longevity in mice with spinal muscular atrophy. *Science* **345**, 688-93

Ng, W., Loh, A. X. W., Teixeira, A. S., Pereira, S. P., Swallow, D. M. (2008) Genetic regulation of MUC1 alternative splicing in human tissues. *Br J Cancer* **99**, 978-85

Nieto, R., Kukuljan, M., Silva, H. (2013) BDNF and schizophrenia: from neurodevelopment to neuronal plasticity, learning, and memory. *Front Psychiatry* **4**, 45

Norris, A. D., Calarco, J. A. (2012) Emerging Roles of Alternative Pre-mRNA Splicing Regulation in Neuronal Development and Function. *Front Neurosci* **6**, 122

O'Bryant, S. E., Hobson, V., Hall, J. R., Waring, S. C., Chan, W., Massman, P., et al. (2009) Brain-derived neurotrophic factor levels in Alzheimer's disease. *J Alzheimers Dis* **17**, 337-41

O'Gorman, W., Thomas, B., Kwek, K. Y., Furger, A., Akoulitchev, A. (2005) Analysis of U1 small nuclear RNA interaction with cyclin H. *J Biol Chem* **280**, 36920-5

Oh, J., Sohn, D. H., Ko, M., Chung, H., Jeon, S. H., Seong, R. H. (2008) BAF60a interacts with p53 to recruit the SWI/SNF complex. *J Biol Chem* **283**, 11924-34

Okabe, S., Miwa, A., Okado, H. (1999) Alternative splicing of the C-terminal domain regulates cell surface expression of the NMDA receptor NR1 subunit. *J Neurosci* **19**, 7781-92

Okumura, N., Yoshida, H., Kitagishi, Y., Nishimura, Y., Matsuda, S. (2011) Alternative splicings on p53, BRCA1 and PTEN genes involved in breast cancer. *Biochem Biophys Res Commun* **413**, 395-9

Oltean, S., Bates, D. O. (2014) Hallmarks of alternative splicing in cancer. *Oncogene* **33**, 5311-8

Pal, S., Yun, R., Datta, A., Lacomis, L., Erdjument-Bromage, H., Kumar, J., et al. (2003) mSin3A/histone deacetylase 2- and PRMT5-containing Brg1 complex is involved in transcriptional repression of the Myc target gene cad. *Mol Cell Biol* **23**, 7475-87

Pan, Q., Shai, O., Lee, L. J., Frey, B. J., Blencowe, B. J. (2008) Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet* **40**, 1413-5

Pandit, S., Zhou, Y., Shiue, L., Coutinho-Mansfield, G., Li, H., Qiu, J., et al. (2013) Genome-wide analysis reveals SR protein cooperation and competition in regulated splicing. *Mol Cell* **50**, 223-35

Pandya-Jones, A., Black, D. L. (2009) Co-transcriptional splicing of constitutive and alternative exons. *RNA* **15**, 1896-908

Papai, G., Weil, P. A., Schultz, P. (2011) New insights into the function of transcription factor TFIID from recent structural studies. *Curr Opin Genet Dev* **21**, 219-24

Park, H., Poo, M.-m. (2013) Neurotrophin regulation of neural circuit development and function. *Nat Rev Neurosci* **14**, 7-23

Paull, T. T., Haykinson, M. J., Johnson, R. C. (1993) The nonspecific DNA-binding and -bending proteins HMG1 and HMG2 promote the assembly

of complex nucleoprotein structures. *Genes Dev* **7**, 1521-34

Pedrotti, S., Cooper, T. A. (2014) In Brief: (mis)splicing in disease. *J Pathol* **233**, 1-3

Petronzelli, F., Sollima, D., Coppola, G., Martini-Neri, M. E., Neri, G., Genuardi, M. (2001) CDKN2A germline splicing mutation affecting both p16(ink4) and p14(arf) RNA processing in a melanoma/neurofibroma kindred. *Genes Chromosomes Cancer* **31**, 398-401

Poss, Z. C., Ebmeier, C. C., Taatjes, D. J. (2013) The Mediator complex and transcription regulation. *Crit Rev Biochem Mol Biol* **48**, 575-608

Pruunsild, P., Sepp, M., Orav, E., Koppel, I., Timmusk, T. (2011) Identification of cis-elements and transcription factors regulating neuronal activity-dependent transcription of human BDNF gene. *J Neurosci* **31**, 3295-308

Quesnel-Vallières, M., Irimia, M., Cordes, S. P., Blencowe, B. J. (2015) Essential roles for the splicing regulator nSR100/SRRM4 during nervous system development. *Genes Dev* **29**, 746-59

Raj, B., O'Hanlon, D., Vessey, J. P., Pan, Q., Ray, D., Buckley, N. J., et al. (2011) Cross-regulation between an alternative splicing activator and a transcription repressor controls neurogenesis. *Mol Cell* **43**, 843-50

Rauskolb, S., Zagrebelsky, M., Dreznjak, A., Deogracias, R., Matsumoto, T., Wiese, S., et al. (2010) Global deprivation of brain-derived neurotrophic factor in the CNS reveals an area-specific requirement for dendritic growth. *J Neurosci* **30**, 1739-49

Reichardt, L. F. (2006) Neurotrophin-regulated signalling pathways. *Philos Trans R Soc Lond B Biol Sci* **361**, 1545-64

Rienzo, M., Casamassimi, A., Schiano, C., Grimaldi, V., Infante, T., Napoli, C. (2012) Distinct alternative splicing patterns of mediator subunit genes during endothelial progenitor cell differentiation. *Biochimie* **94**, 1828-32

Rienzo, M., Costa, V., Scarpato, M., Schiano, C., Casamassimi, A., Grimaldi, V., et al. (2014) RNA-Seq for the identification of novel Mediator transcripts in endothelial progenitor cells. *Gene* **547**, 98-105

Rienzo, M., Nagel, J., Casamassimi, A., Giovane, A., Dietzel, S., Napoli, C. (2010) Mediator subunits: gene expression pattern, a novel transcript identification and nuclear localization in human endothelial progenitor cells. *Biochim Biophys Acta* **1799**, 487-95

Ryan, B., O'Donovan, N., Browne, B., O'Shea, C., Crown, J., Hill, A. D. K., et al. (2005) Expression of survivin and its splice variants survivin-2B and survivin-DeltaEx3 in breast cancer. *Br J Cancer* **92**, 120-4

Sako, W., Morigaki, R., Kaji, R., Tooyama, I., Okita, S., Kitazato, K., et al. (2011) Identification and localization of a neuron-specific isoform of TAF1 in rat brain: implications for neuropathology of DYT3 dystonia. *Neuroscience* **189**, 100-7

Scarpato, M., Federico, A., Ciccodicola, A., Costa, V. (2015) Novel transcription factor variants through RNA-sequencing: the importance of being "alternative". *Int J Mol Sci* **16**, 1755-71

Schor, I. E., Rascovan, N., Pelisch, F., Alló, M., Kornblihtt, A. R. (2009) Neuronal cell depolarization induces intragenic chromatin modifications affecting NCAM alternative splicing. *Proc Natl Acad Sci U S A* **106**, 4325-30

Schwartz, S., Meshorer, E., Ast, G. (2009) Chromatin organization marks exon-intron structure. *Nat Struct Mol Biol* **16**, 990-5

Schwarzenbach, H. (2013) Circulating nucleic acids as biomarkers in breast cancer. *Breast Cancer Res* **15**, 211

Shargunov, A. V., Krasnov, G. S., Ponomarenko, E. A., Lisitsa, A. V., Shurdov, M. A., Zverev, V. V., et al. (2014) Tissue-specific alternative splicing analysis reveals the diversity of chromosome 18 transcriptome. *J Proteome Res* **13**, 173-82

Shepard, P. J., Hertel, K. J. (2009) The SR protein family. *Genome Biol* **10**, 242

Shimokawa, T., Tostar, U., Lauth, M., Palaniswamy, R., Kasper, M., Toftgård, R., et al. (2008) Novel human glioma-associated oncogene 1 (GLI1) splice variants reveal distinct mechanisms in the terminal transduction of the hedgehog signal. *J Biol Chem* **283**, 14345-54

Siegfried, Z., Bonomi, S., Ghigna, C., Karni, R. (2013) Regulation of the Ras-MAPK and PI3K-mTOR Signalling Pathways by Alternative Splicing in Cancer. *Int J Cell Biol* **2013**, 568931

Singer, G. A. C., Wu, J., Yan, P., Plass, C., Huang, T. H. M., Davuluri, R. V. (2008) Genome-wide analysis of alternative promoters of human genes using a custom promoter tiling array. *BMC Genomics* **9**, 349

Singh, R. K., Cooper, T. A. (2012) Pre-mRNA splicing in disease and therapeutics. *Trends Mol Med* **18**, 472-82

Skaper, S. D. (2012) The neurotrophin family of neurotrophic factors: an overview. *Methods Mol Biol* **846**, 1-12

Smith, C. W., Patton, J. G., Nadal-Ginard, B. (1989) Alternative splicing in the control of gene expression. *Annu Rev Genet* **23**, 527-77

Smith, M. J., Kulkarni, S., Pawson, T. (2004) FF domains of CA150 bind transcription and splicing factors through multiple weak interactions. *Mol Cell Biol* **24**, 9274-85

Smith, M. J., O'Sullivan, J., Bhaskar, S. S., Hadfield, K. D., Poke, G., Caird, J., et al. (2013) Loss-of-function mutations in SMARCE1 cause an inherited disorder of multiple spinal meningiomas. *Nat Genet* **45**, 295-8

Snider, W. D. (1994) Functions of the neurotrophins during nervous system development: what the knockouts are teaching us. *Cell* **77**, 627-38

Soulé, J., Messaoudi, E., Bramham, C. R. (2006) Brain-derived neurotrophic factor and control of synaptic consolidation in the adult brain. *Biochem Soc Trans* **34**, 600-4

Surget, S., Khoury, M. P., Bourdon, J.-C. (2013) Uncovering the role of p53 splice variants in human malignancy: a clinical perspective. *Onco Targets Ther* **7**, 57-68

Sveen, A., Kilpinen, S., Ruusulehto, A., Lothe, R. A., Skotheim, R. I. (2015) Aberrant RNA splicing in cancer; expression changes and driver mutations of splicing factor genes. *Oncogene*

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**, 122-8

Tanguay, R. L., Gallie, D. R. (1996) Translational efficiency is regulated by the length of the 3' untranslated region. *Mol Cell Biol* **16**, 146-56

Tao, X., Finkbeiner, S., Arnold, D. B., Shaywitz, A. J., Greenberg, M. E. (1998) Ca2+ influx regulates BDNF transcription by a CREB family transcription factor-dependent mechanism. *Neuron* **20**, 709-26

Tao, X., West, A. E., Chen, W. G., Corfas, G., Greenberg, M. E. (2002) A calcium-responsive transcription factor, CaRF, that regulates neuronal activity-dependent expression of BDNF. *Neuron* **33**, 383-95

Thomas, M. C., Chiang, C.-M. (2006) The general transcription machinery and general cofactors. *Crit Rev Biochem Mol Biol* **41**, 105-78

Timmusk, T., Palm, K., Lendahl, U., Metsis, M. (1999) Brain-derived neurotrophic factor expression in vivo is under the control of neuron-restrictive silencer element. *J Biol Chem* **274**, 1078-84

Timmusk, T., Palm, K., Metsis, M., Reintam, T., Paalme, V., Saarma, M., et al. (1993) Multiple promoters direct tissue-specific expression of the rat BDNF gene. *Neuron* **10**, 475-89

Tsai, Y. S., Dominguez, D., Gomez, S. M., Wang, Z. (2015) Transcriptome-wide identification and study of cancer-specific splicing events across multiple tumors. *Oncotarget* ,

Ule, J., Stefani, G., Mele, A., Ruggiu, M., Wang, X., Taneri, B., et al. (2006) An RNA map predicting Nova-dependent splicing regulation. *Nature* **444**, 580-6

Ule, J., Ule, A., Spencer, J., Williams, A., Hu, J.-S., Cline, M., et al. (2005) Nova regulates brain-specific splicing to shape the synapse. *Nat Genet* **37**, 844-52

Venables, J. P. (2004) Aberrant and alternative splicing in cancer. *Cancer Res* **64**, 7647-54

Venables, J. P., Klinck, R., Bramard, A., Inkel, L., Dufresne-Martin, G., Koh, C., et al. (2008) Identification of alternative splicing markers for breast cancer. *Cancer Res* **68**, 9525-31

Venables, J. P., Klinck, R., Koh, C., Gervais-Bird, J., Bramard, A., Inkel, L., et al. (2009) Cancer-associated regulation of alternative splicing. *Nat Struct Mol Biol* **16**, 670-6

Vignali, M., Hassan, A. H., Neely, K. E., Workman, J. L. (2000) ATP-dependent chromatin-remodeling complexes. *Mol Cell Biol* **20**, 1899-910

Villaronga, M. A., López-Mateo, I., Markert, L., Espinosa, E., Fresno Vara, J. A., Belandia, B. (2011) Identification and characterization of novel potentially oncogenic mutations in the human BAF57 gene in a breast cancer patient. *Breast Cancer Res Treat* **128**, 891-8

Wagoner, M. P., Gunsalus, K. T. W., Schoenike, B., Richardson, A. L., Friedl, A., Roopra, A. (2010) The transcription factor REST is lost in aggressive breast cancer. *PLoS Genet* **6**, e1000979

Waldholm, J., Wang, Z., Brodin, D., Tyagi, A., Yu, S., Theopold, U., et al. (2011) SWI/SNF regulates the alternative processing of a specific subset of pre-mRNAs in Drosophila melanogaster. *BMC Mol Biol* **12**, 46

Wan, L., Tsai, C.-H., Hsu, C.-M., Huang, C.-C., Yang, C.-C., Liao, C.-C., et al. (2010) Mutation analysis and characterization of alternative splice variants of the Wilson disease gene ATP7B. *Hepatology* **52**, 1662-70

Wang, E. T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., et al. (2008) Alternative isoform regulation in human tissue transcriptomes. *Nature* **456**, 470-6

Wang, G.-S., Cooper, T. A. (2007) Splicing in disease: disruption of the splicing code and the decoding machinery. *Nat Rev Genet* **8**, 749-61

Wang, W., Chi, T., Xue, Y., Zhou, S., Kuo, A., Crabtree, G. R. (1998) Architectural DNA binding by a high-mobility-group/kinesin-like subunit in mammalian SWI/SNF-related complexes. *Proc Natl Acad Sci U S A* **95**, 492-8

Wang, Y., Wong, R. H. F., Tang, T., Hudak, C. S., Yang, D., Duncan, R. E., et al. (2013) Phosphorylation and recruitment of BAF60c in chromatin remodeling for lipogenesis in response to insulin. *Mol Cell* **49**, 283-97

Wappenschmidt, B., Becker, A. A., Hauke, J., Weber, U., Engert, S., Köhler, J., et al. (2012) Analysis of 30 putative BRCA1 splicing mutations in hereditary breast and ovarian cancer families identifies exonic splice site mutations that escape in silico prediction. *PLoS One* **7**, e50800

Ward, A. J., Cooper, T. A. (2010) The pathobiology of splicing. *J Pathol* **220**, 152-63

Waseem, N. H., Vaclavik, V., Webster, A., Jenkins, S. A., Bird, A. C., Bhattacharya, S. S. (2007) Mutations in the gene coding for the pre-mRNA splicing factor, PRPF31, in patients with autosomal dominant retinitis pigmentosa. *Invest Ophthalmol Vis Sci* **48**, 1330-4

Watanabe, H., Mizutani, T., Haraguchi, T., Yamamichi, N., Minoguchi, S., Yamamichi-Nishina, M., et al. (2006) SWI/SNF complex is essential for NRSF-mediated suppression of neuronal genes in human nonsmall cell lung carcinoma cell lines. *Oncogene* **25**, 470-9

Wilhelm, E., Kornete, M., Targat, B., Vigneault-Edwards, J., Frontini, M., Tora, L., et al. (2010) TAF6delta orchestrates an apoptotic transcriptome profile and interacts functionally with p53. *BMC Mol Biol* **11**, 10

Wright, K. J., Marr, 2nd, M. T., Tjian, R. (2006) TAF4 nucleates a core subcomplex of TFIID and mediates activated transcription from a TATA-less promoter. *Proc Natl Acad Sci U S A* **103**, 12347-52

Xiang, S., Gapsys, V., Kim, H.-Y., Bessonov, S., Hsiao, H.-H., Möhlmann, S., et al. (2013) Phosphorylation drives a dynamic switch in serine/arginine-rich proteins. *Structure* **21**, 2162-74

Xiong, H. Y., Alipanahi, B., Lee, L. J., Bretschneider, H., Merico, D., Yuen, R. K. C., et al. (2015) RNA splicing. The human splicing code reveals new insights into the genetic determinants of disease. *Science* **347**, 1254806

Xu, Q., Modrek, B., Lee, C. (2002) Genome-wide detection of tissue-specific alternative splicing in the human transcriptome. *Nucleic Acids Res* **30**, 3754-66

Yamaguchi, T., Kurita, T., Nishio, K., Tsukada, J., Hachisuga, T., Morimoto, Y., et al. (2015) Expression of BAF57 in ovarian cancer cells and drug sensitivity. *Cancer Sci* **106**, 359-66

Yang, J.-W., Ru, J., Ma, W., Gao, Y., Liang, Z., Liu, J., et al. (2015) BDNF promotes the growth of human neurons through crosstalk with the Wnt/β-catenin signaling pathway via GSK-3β. *Neuropeptides* ,

Yang, M., Sun, Y., Ma, L., Wang, C., Wu, J.-M., Bi, A., et al. (2011) Complex alternative splicing of the smarca2 gene suggests the importance of smarca2-B variants. *J Cancer* **2**, 386-400

Yano, M., Hayakawa-Yano, Y., Mele, A., Darnell, R. B. (2010) Nova2 regulates neuronal migration through an RNA switch in disabled-1 signaling. *Neuron* **66**, 848-58

Yeo, G., Holste, D., Kreiman, G., Burge, C. B. (2004) Variation in alternative splicing across human tissues. *Genome Biol* **5**, R74

Yin, J.-w., Wang, G. (2014) The Mediator complex: a master coordinator of transcription and cell lineage development. *Development* **141**, 977-87

Zafra, F., Hengerer, B., Leibrock, J., Thoenen, H., Lindholm, D. (1990) Activity dependent regulation of BDNF and NGF mRNAs in the rat hippocampus is mediated by non-NMDA glutamate receptors. *EMBO J* **9**, 3545-50

Zhang, C., Frias, M. A., Mele, A., Ruggiu, M., Eom, T., Marney, C. B., et al. (2010) Integrative modeling defines the Nova splicing-regulatory network and its combinatorial controls. *Science* **329**, 439-43

Zhang, F., Wang, M., Michael, T., Drabier, R. (2013) Novel alternative splicing isoform biomarkers identification from high-throughput plasma proteomics profiling of breast cancer. *BMC Syst Biol* **7 Suppl 5**, S8

Zhang, H. T., Scott, P. A., Morbidelli, L., Peak, S., Moore, J., Turley, H., et al. (2000) The 121 amino acid isoform of vascular endothelial growth factor is more strongly tumorigenic than other splice variants in vivo. *Br J Cancer* **83**, 63-8

Zhang, X., Krutchinsky, A., Fukuda, A., Chen, W., Yamamura, S., Chait, B. T., et al. (2005) MED1/TRAP220 exists predominantly in a TRAP/ Mediator subpopulation enriched in RNA polymerase II and is required for ER-mediated transcription. *Mol Cell* **19**, 89-100

Zhao, X., Tang, Z., Zhang, H., Atianjoh, F. E., Zhao, J.-Y., Liang, L., et al. (2013) A long noncoding RNA contributes to neuropathic pain by silencing Kcna2 in primary afferent neurons. *Nat Neurosci* **16**, 1024-31

Zheng, F., Zhou, X., Luo, Y., Xiao, H., Wayman, G., Wang, H. (2011) Regulation of brain-derived neurotrophic factor exon IV transcription through calcium responsive elements in cortical neurons. *PLoS One* **6**, e28441

Zheng, F., Zhou, X., Moon, C., Wang, H. (2012) Regulation of brain-derived neurotrophic factor expression in neurons. *Int J Physiol Pathophysiol Pharmacol* **4**, 188-200

Zhou, C., Licciulli, S., Avila, J. L., Cho, M., Troutman, S., Jiang, P., et al. (2013) The Rac1 splice form Rac1b promotes K-ras-induced lung tumorigenesis. *Oncogene* **32**, 903-9

Zhou, H.-L., Luo, G., Wise, J. A., Lou, H. (2014) Regulation of alternative splicing by local histone modifications: potential roles for RNA-guided mechanisms. *Nucleic Acids Res* **42**, 701-13

Zhou, Z., Hong, E. J., Cohen, S., Zhao, W.-N., Ho, H.-Y. H., Schmidt, L., et al. (2006) Brain-specific phosphorylation of MeCP2 regulates activity-dependent Bdnf transcription, dendritic growth, and spine maturation. *Neuron* **52**, 255-69

Zhou, Z., Licklider, L. J., Gygi, S. P., Reed, R. (2002) Comprehensive proteomic analysis of the human spliceosome. *Nature* **419**, 182-5

Zuccato, C., Belyaev, N., Conforti, P., Ooi, L., Tartari, M., Papadimou, E., et al. (2007) Widespread disruption of repressor element-1 silencing

transcription factor/neuron-restrictive silencer factor occupancy at its target genes in Huntington's disease. *J Neurosci* **27**, 6972-83

Zuccato, C., Tartari, M., Crotti, A., Goffredo, D., Valenza, M., Conti, L., et al. (2003) Huntingtin interacts with REST/NRSF to modulate the transcription of NRSE-controlled neuronal genes. *Nat Genet* **35**, 76-83

Zukin, R. S., Bennett, M. V. (1995) Alternatively spliced isoforms of the NMDARI receptor subunit. *Trends Neurosci* **18**, 306-13
de Almeida, S. F., Carmo-Fonseca, M. (2012) Design principles of interconnections between chromatin and pre-mRNA splicing. *Trends Biochem Sci* **37**, 248-53

de la Grange, P., Gratadou, L., Delord, M., Dutertre, M., Auboeuf, D. (2010) Splicing factor and exon profiling across human tissues. *Nucleic Acids Res* **38**, 2825-38

de la Mata, M., Alonso, C. R., Kadener, S., Fededa, J. P., Blaustein, M., Pelisch, F., et al. (2003) A slow RNA polymerase II affects alternative splicing in vivo. *Mol Cell* **12**, 525-32

de la Mata, M., Lafaille, C., Kornblihtt, A. R. (2010) First come, first served revisited: factors affecting the same alternative splicing event have different effects on the relative rates of intron removal. *RNA* **16**, 904-12

ACKNOWLEDGEMENTS

This work has been carried out at the Department of Gene Technology, Tallinn University of Technology, Estonia, Protobios LLC, Estonia, and The Competence Centre for Cancer Research, Estonia.

First and foremost, I would like to thank Tõnis Timmusk and my supervisor, Kaia Palm, for their scientific guidance, for sharing their knowledge and experience, and for their tremendous support, inspiration and belief in me.

I want to thank all my lab-mates and co-authors I had a pleasure to work with, for the pleasant atmosphere and all the excellent results we have achieved together. Thank you Priit Pruunsild, Helle Sadam, Tamara Aid, Alla Piirsoo, Anri Kivil, Ave-Kris Lend, Mari Sepp, Kairit Tints, Anzelika Sarapuu, Ave Laas and many others.

Many thanks to my parents and my sister Jekaterina Kazantseva, who encouraged me to make this work happen and helped me in many ways to complete my studies despite all the obstacles I encountered on my way.

I am also very grateful to my beloved daughters: Victoria and Anastasia, who brightened my way during the last 7 years and taught me to believe in miracles.

And finally, my particular thanks go to my husband Aleksey for making me realize how mysterious, and at the same time interesting life really is, and for giving me unremitting motivation and teaching me to love life.

PUBLICATION I

I Pruunsild, P.∗, **Kazantseva, A**.∗, Aid, T., Palm, K., Timmusk, T. (2007). Dissecting the human BDNF locus: bidirectional transcription, complex splicing, and multiple promoters. Genomics, 90(3), 397-406

Available online at www.sciencedirect.com
ScienceDirect
Genomics 90 (2007) 397–406
GENOMICS
www.elsevier.com/locate/ygeno

Dissecting the human *BDNF* locus: Bidirectional transcription, complex splicing, and multiple promoters☆

Priit Pruunsild [1], Anna Kazantseva [1], Tamara Aid, Kaia Palm, Tõnis Timmusk *

Department of Gene Technology, Tallinn University of Technology, Akadeemia tee 15, Tallinn 19086, Estonia

Received 10 January 2007; accepted 14 May 2007
Available online 12 July 2007

Abstract

Brain-derived neurotrophic factor (BDNF), a member of the nerve growth factor family of neurotrophins, has central roles in the development, physiology, and pathology of the nervous system. We have elucidated the structure of the human *BDNF* gene, identified alternative transcripts, and studied their expression in adult human tissues and brain regions. In addition, the transcription initiation sites for human *BDNF* transcripts were determined and the activities of *BDNF* promoters were analyzed in transient overexpression assays. Our results show that the human *BDNF* gene has 11 exons and nine functional promoters that are used tissue and brain-region specifically. Furthermore, noncoding natural antisense RNAs that display complex splicing and expression patterns are transcribed in the *BDNF* gene locus from the *antiBDNF* gene (approved gene symbol *BDNFOS*). We show that *BDNF* and *antiBDNF* transcripts form dsRNA duplexes in the brain in vivo, suggesting an important role for *antiBDNF* in regulating *BDNF* expression in human.

Keywords: Neurotrophic factor; Brain; neuron; Alternative splicing; BDNF; Natural antisense transcript; RNA duplex; Lin-7c/Mals-3/veli3

Brain-derived neurotrophic factor (BDNF) is a member of the nerve growth factor family of neurotrophins. During development BDNF supports the survival and differentiation of selected neuronal populations of the peripheral and central nervous systems and participates in axonal growth and pathfinding and in the modulation of dendritic growth and morphology [1,2]. BDNF also has a prominent role in later stages of nervous system development and in the adult by regulating synaptic transmission and plasticity and acts as a central modulator of pain [3]. BDNF has been shown to be a modulator of synaptogenesis in vivo [4] and has a functional role in the expression of LTP in the hippocampus [5]. Data accumulated in recent years suggest that neuronal activity regulates transcription of *BDNF,* transport of *BDNF* mRNA and protein into dendrites, and secretion of BDNF protein, which are important for the formation of appropriate synaptic connections and for learning and memory during development and in adults [6]. A single-nucleotide polymorphism in the human *BDNF* gene, resulting in a valine to methionine substitution (Val66Met) in the prodomain, has been shown to lead to reduced activity-induced BDNF secretion and memory impairment [7]. BDNF signaling has been shown to be critical in several neuropsychiatric and neurodegenerative diseases [1], for example, Huntington disease [8]. These results taken together show that BDNF has numerous important roles in brain development, physiology, and pathology.

During development, BDNF protein expression is more abundant in the nervous system compared to other tissues and its levels are dramatically increased in the brain during postnatal development [9]. In the adult nervous system, *BDNF* displays a widespread distribution pattern, with the highest levels of mRNA and protein in the hippocampus, amygdala, cerebral cortex, and hypothalamus [9–12]. *BDNF* mRNA expression is mostly confined to neurons and there are only a few brain areas where *BDNF* mRNA is not detected [10–12]. *BDNF* expression in adult tissues is detectable also outside of the central nervous

☆ Sequence data from this article have been deposited in the GenBank Data Libraries under Accession Nos. EF674517 - EF674521 and EF689009 - EF689021 (human *BDNF* mRNAs), and EF689022 - EF689042 (human *antiBDNF* RNAs).

* Corresponding author. Fax: +372 620 4401.
E-mail address: tonis.timmusk@ttu.ee (T. Timmusk).

[1] These authors contributed equally to this work.

0888-7543
doi:10.1016/j.ygeno.2007.05.004

system. Lower *BDNF* mRNA levels than in the hippocampus have been detected in the thymus, liver, spleen, heart, and lung [9,11,13,14].

The structure and regulation of the *BDNF* gene have been studied in rodents [15–17]. Mouse and rat *BDNF* genes have eight 5′ exons containing separate promoters upstream of each exon and one 3′ exon encoding the mature BDNF protein [17]. Multiple promoters determine tissue-specific expression of the *BDNF* transcripts [15,17]. The human *BDNF* has also been shown to consist of multiple 5′ noncoding exons and one coding exon, which give rise to alternatively spliced transcripts [18–20]. According to the most recent description of the organization of the human *BDNF* gene locus and alternatively spliced transcripts the human *BDNF* has seven noncoding and one coding exon [19]. Provided that only a partial description of the transcripts was given and expression of the alternative transcripts was studied in only a few brain regions and was not investigated in human nonneural tissues, we undertook the current study to analyze thoroughly the structure of the human *BDNF* gene, characterize the expression of alternatively spliced *BDNF* mRNAs in different human tissues and brain regions, and identify and study the activities of alternative human *BDNF* promoters.

Results

Structure and alternative splicing of human BDNF and antiBDNF (approved gene symbol BDNFOS)

To reexamine the human *BDNF* gene structure and identify mRNAs transcribed from the gene, in silico analysis, 5′ rapid amplification of cDNA ends (5′ RACE), and RT-PCR analyses were performed. First, all *BDNF* mRNAs and expressed sequence tags (ESTs) available at the NCBI database (http://www.ncbi.nlm.nih.gov) were analyzed. Primers designed for PCR analyses are presented in Supplementary Table 1. Total RNAs of the adult human frontal cerebral cortex, medulla, and hippocampus were used as templates in the RT-PCR. Second, to identify novel *BDNF* transcripts and to determine the transcription start sites for the human *BDNF* transcripts, 5′ RACE of human adult hippocampal and cerebellar RNA was performed using antisense primers specific for the 3′ exon and for the 5′ exons (Supplementary Table 1).

BDNF gene exon–intron boundaries and genomic locations were determined by BLAT algorithm (http://genome.ucsc.edu/cgi-bin/hgBlat) and by direct comparison of PCR-amplified sequences with genomic DNA sequence from the NCBI database (Accession No. AF411339). The designation of the human *BDNF* exons in this study is consistent with the naming of the mouse and rat *BDNF* exons described in the study by Aid et al. [17]. The exons unique to human *BDNF* are marked with the letter “h” and are named with the same number as the neighboring upstream exon. Our analysis showed that the human *BDNF* gene spans ~70 kb and consists of 11 exons (Fig. 1). *hBDNF* exons named by Liu et al. [19] are designated here as old exons. Comparison of our data with that of Liu et al. [19] shows that exons I–IV correspond to the respective old exons; exons V, Vh, VIII, and VIIIh are novel exons; exons VI and VII correspond to old exons V and VI, respectively; and exon IX variants IXb and IXd correspond to old exons VII and VIII, respectively (Fig. 2A). Nine of the exons, I, II, III, IV, V, Vh, VI, VII, and IX, can be defined as 5′ exons (Fig. 1). Cloning and sequencing of the 5′ RACE products revealed that the transcription start sites relative to the 3′ end of the respective exon are located as follows: −647 and −428 nt for exon I; −433, −423, −422, −416, −407, −400, −226, −224, −204, −200, −78, and −47 nt for exon II; −237 and −191 nt for exon III; −337, −333, −274, and −215 nt for exon IV; −82, −80, and −79 nt for exon V; −225 and −222 nt for exon Vh; −324; −323, −319, −318, and −315 nt for exon VI; and −184 nt for exon VII (Supplementary Fig. 1). We determined that human *BDNF* transcription starts also from the last exon, exon IX, −1102 nt upstream of the translation start site in this exon (Supplementary Fig. 1). No major differences in the transcription start site locations were observed when 5′ RACE products from hippocampal or cerebellar RNA were analyzed. Two exons, VIII and VIIIh, are rarely used and always in combination with exon V as the 5′ exon (Fig. 1). Exons II, III, IV, V, Vh, VI, and VIIIh are untranslated exons and translation of the transcripts containing these exons starts from the ATG positioned in exon IX (Supplementary Fig. 1). Exons I, VII, and VIII contain in-frame ATG codons that could be used as translation start sites leading to the prepro-BDNF proteins with longer N-termini (Fig. 2B).

In addition, we identified alternative splice donor sites in exons II, V, and VI (Fig. 1, Supplementary Fig. 1, and Table 2). Usage of these splice sites leads to the formation of transcripts with different 5′ UTR lengths but does not affect the coding region of *BDNF.* Characterization of the exon–intron boundary sequences showed that not all exon–intron splice junctions adhere to the GU–AG rule characteristic of eukaryotes. We found that exon VII is unique because the splice donor site used contains nucleotides GG instead of the conventional GU sequence (Supplementary Table 2). Exon IX, which encodes the BDNF protein and 3′ UTR, is subjected to internal splicing and/or transcription initiation upstream of exon IX that leads to the generation of alternative transcripts containing variants of exon IX. These exon IX variants comprise different regions or combinations of regions of exon IX that were designated “a”, “b”, “c”, and “d” (Fig. 1). Exon IX is used mostly in conjunction with the upstream exons (I–VIII, VIIIh), and in that case only the most 3′ region of exon IX, IXd, is included in the mature transcripts. On rare occasions when exon VI is the 5′ exon, alternative splicing occurs within exon IX leading to the inclusion of two regions, IXb and IXd, in the mRNAs. When transcription is initiated upstream of exon IX the transcripts are not subjected to internal splicing and contain all the regions of exon IX: IXa, IXb, IXc, and IXd. In extremely rare cases exon IX region “c” is spliced out (Fig. 1 and Supplementary Fig. 1).

Recently it was reported that natural antisense transcripts are transcribed from the human *BDNF* gene locus [19]. We analyzed the exon–intron junctions of amplified *BDNF* cDNAs and also noticed that several mRNAs are transcribed in an antisense direction compared to BDNF mRNAs. This finding was

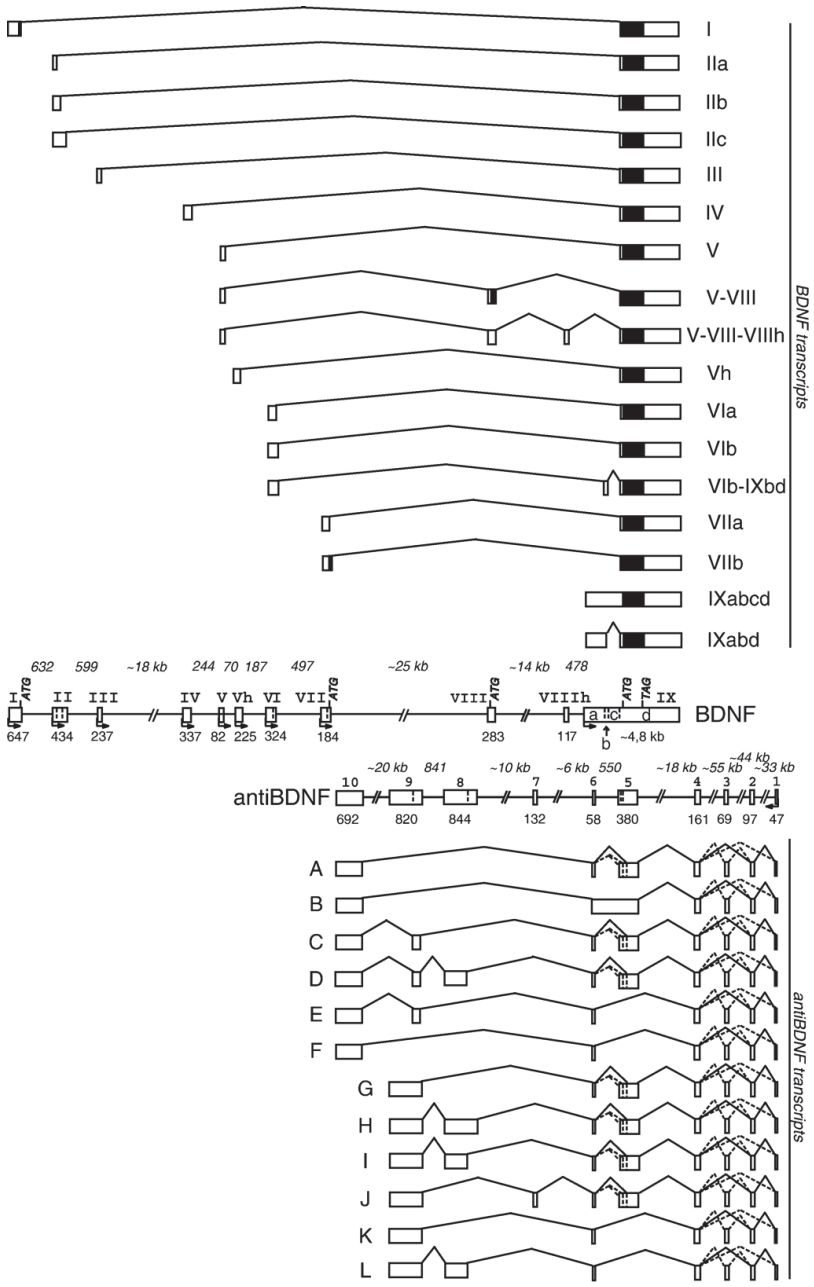

Fig. 1. Structure and alternative transcripts of the human *BDNF* (top) and *antiBDNF* (bottom) genes. The structural organization of the exons and introns was determined by analyzing genomic and mRNA sequence data using bioinformatics, RT-PCR, and 5′ RACE. Exons are shown as boxes and introns as lines. Filled boxes and open boxes indicate the translated regions of the exons and the untranslated regions of the exons, respectively. The numbers below the exons and above the introns indicate their sizes. Exon and intron sizes are in base pairs, if not indicated otherwise. Arrows indicate the transcription start sites. ATG and TAG mark the positions of the translational start and stop codons, respectively. Vertical dashed lines indicate alternative splicing sites for the respective exons. *BDNF* exon IX is divided into regions “a”, “b”, “c”, and “d” as indicated in the box marking the position of exon IX. *BDNF* transcript names relate to the upstream exons used in front of the major 3′ exon IXd. “A”–“L” mark *antiBDNF* transcripts. Solid lines connecting the exons of transcripts represent the major splicing patterns of exons. Dashed lines connecting the exons of transcripts represent the minor splicing patterns of *antiBDNF*. Exon numbers are shown in Roman numerals for the *BDNF* gene and in Arabic numerals for the *antiBDNF* gene.

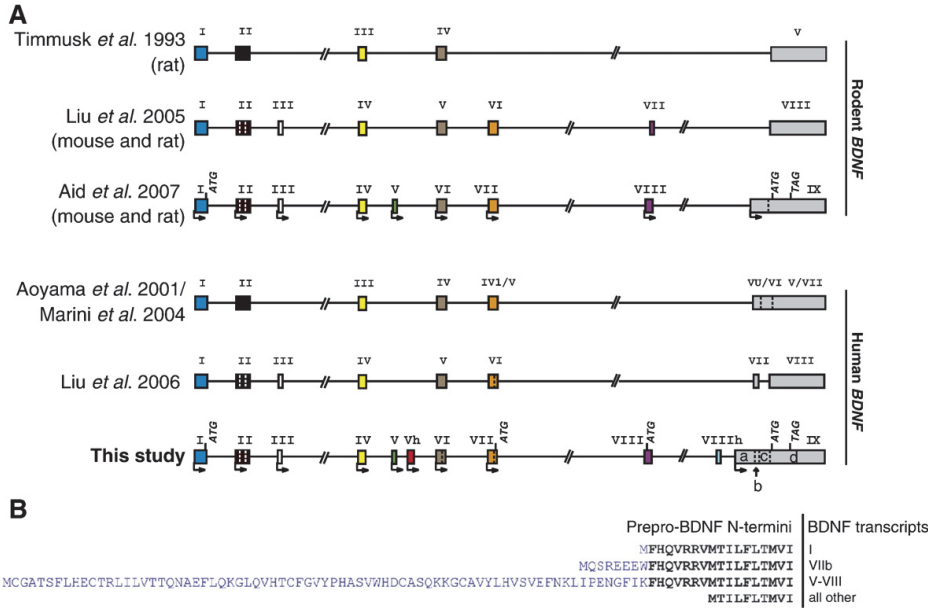

Fig. 2. (A) Comparison of the human and rodent *BDNF* gene structures proposed by different studies. Structures presented are according to Timmusk et al. [15], Liu et al. [16], Aid et al. [17], Aoyama et al. [20], Marini et al. [40], and Liu et al. [19] and the human *BDNF* gene structure determined in this study. Exons are shown as boxes and introns as lines. The identical human exons and the rodent exons homologous to the human exons are shown in the same color in all the structures. Novel exons determined by this study are in green (V), red (Vh), purple (VIII), and light blue (VIIIh). (B) Amino acid sequences of different potential prepro-BDNF N-termini. Amino acids encoded by exon IX are in black and sequences encoded by alternative 5′ exons are in blue. The transcripts encoding the respective N-termini of BDNF are listed adjacent to the N-terminal sequences.

also confirmed by orientation-specific RT-PCR and by RNase protection assay (data not shown). We named the gene and transcripts transcribed from the opposite strand of *BDNF* as *antiBDNF* (for antisense *BDNF;* part of it is described by Liu and colleagues and designated as OSBDNF—Liu et al. [19]). The *antiBDNF* gene spans ~191 kb and consists of at least 10 exons (exons 1–10) and is transcribed from one promoter as shown by our 5′ RACE analyses (Fig. 1). All intron–exon boundaries of the *antiBDNF* gene are consistent with the GU–AG consensus (Supplementary Table 2).

Exons 1–4 of the *antiBDNF* gene are located downstream of the *BDNF* gene (Fig. 1). Exon 5, 345 bp in length, overlaps regions IXc and IXd and exon 6 overlaps region IXa of the *BDNF* coding exon. Exons 7–10 of *antiBDNF* are located in the introns of *BDNF*. In silico and RT-PCR analyses showed that alternative splicing from the *antiBDNF* pre-mRNA produces more than 300 transcripts, but exon 1 of *antiBDNF* is always used as the most 5′ exon for all the transcripts. Of note, our bioinformatics analysis showed that *antiBDNF* exon 1 is in head-to-head orientation with exon 1 of *Lin-7c/Mals-3/veli3,* suggesting that a bidirectional promoter controls the expression of these genes. The majority of the *antiBDNF* alternative transcripts contain exons 5 and 6, which are complementary to the *BDNF* protein-coding exon IX. However, in several *antiBDNF* transcripts exon 5 is skipped out. In addition, exon 5 of *antiBDNF* could be spliced using three alternative splice donor sites (Fig. 1, Supplementary Fig. 2), and the lengths of exons 8 and 9 can vary because of usage of internal alternative splice acceptor sites (Fig. 1). There are no potential open reading frames in any of the identified mRNAs transcribed from the *antiBDNF* gene, suggesting that these transcripts are non-protein-coding, as proposed also by Liu et al. [19].

Expression of alternatively spliced BDNF and antiBDNF mRNAs in adult human tissues

Expression of *BDNF* and *antiBDNF* transcripts was determined by RT-PCR in 22 different adult human tissues (Fig. 3). The results showed that human *BDNF* alternative transcripts are expressed in a tissue-specific manner. The levels of the majority of the human *BDNF* transcripts were highest in the brain. However, several alternative *BDNF* mRNAs showed relatively high expression levels in nonneural tissues. For example, expression levels of transcripts containing exons VI and IXabcd were high in the heart, placenta, and prostate. Transcripts containing exons I, Vh, VI, and IXabcd were highly expressed in the testis. High levels of transcripts containing exon VI were expressed also in the lung. Several *BDNF* mRNAs were expressed at moderate or low levels in the adrenal gland (exon Vh and exon IXabcd transcripts), bone marrow (exons I, VI, and IXabcd transcripts), kidney, muscle, stomach, spinal cord (exons Vh and VI transcripts), liver (exon IXabcd transcripts), small intestine (exon VI transcripts), and trachea (exons Vh, VI, and IXabcd transcripts). Low levels of exon IXabd transcripts

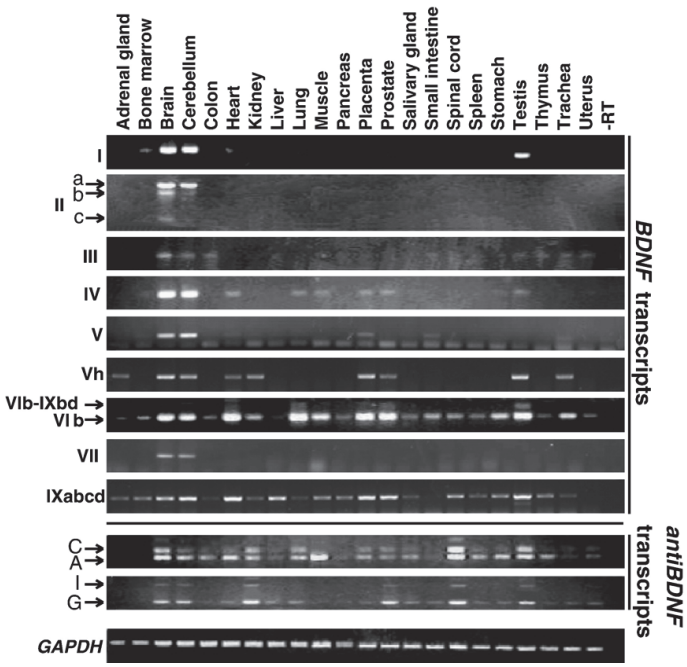

Fig. 3. Semiquantitative analysis of human *BDNF, antiBDNF,* and control *GAPDH* mRNA expression in adult human tissues by RT-PCR. Roman numerals on the left indicate the detected *BDNF* transcripts and the 5′ exon-specific primers used in combination with an antisense primer located in the *BDNF* coding region in exon IXd (Supplementary Table 1). A, C, G, and I refer to the respective *antiBDNF* transcripts shown in Fig. 1.

were observed only in some cell lines (data not shown). *BDNF* mRNAs containing exons II and VII were expressed exclusively in the brain. Altogether, the results indicate that transcripts containing exons II, III, IV, V, and VII are predominantly brain-specific. Transcripts containing exons I and Vh are, in addition to brain, expressed in certain peripheral tissues, and transcripts containing exons VI and IXabcd show a wide pattern of expression.

antiBDNF transcripts were present at different levels in almost all human tissues analyzed (Fig. 3). High expression of human *antiBDNF* transcripts was detected in the brain, kidney, spinal cord, and testis. Moderate levels of *antiBDNF* RNA were seen in the lung, prostate, salivary gland, spleen, stomach, and uterus. Low *antiBDNF* expression levels were detected in the adrenal gland, liver, placenta, small intestine, and trachea. Certain alternative transcripts of *antiBDNF* were expressed in a tissue-specific manner. For example, *antiBDNF* transcripts with exon 10 were present in the colon and muscle, whereas transcripts with exon 9 were not expressed in these tissues. Taken together, *BDNF* and *antiBDNF* expression patterns were distinct, although partially overlapping.

Expression of alternatively spliced BDNF and antiBDNF mRNAs in adult human brain regions

Expression analysis of human *BDNF* and *antiBDNF* transcripts in 30 different adult brain regions was performed by RT-PCR (Fig. 4). Several differences in the expression of alternatively spliced human *BDNF* transcripts were detected. The results showed that all *BDNF* transcripts were expressed at high levels in the corpus mammillare (mammillary body), pons, hippocampus, frontal cortex, colliculi, and olfactory tract. All *BDNF* transcripts except the ones containing exons V and VII were expressed at high levels in the cerebellum and medulla, and all transcripts but those containing exon IXabcd were expressed at high levels in the infundibulum. *BDNF* expression in the dentate nucleus, white matter of the cerebellum, substantia nigra, nucleus ruber (red nucleus), and epiphysis was very low. *BDNF* expression was also very low in the globus pallidus, striatum (caudate nucleus and putamen), and thalamus, with the exception of exon IXabcd transcripts, which were expressed at relatively high levels in these regions. In the amygdala only transcripts containing exons I, IV, and VI were expressed at high levels. In the corpus callosum only exon VI and IXabcd transcripts were detected. Notably, comparison of expression levels of individual transcripts in different brain areas indicated that *BDNF* exon II transcript levels were much higher in the cerebellum than in other brain areas. Exon IXabcd mRNAs were expressed at relatively similar levels in all brain regions, with only infundibulum having very low expression levels. Interestingly, in the brain structures that contain only glial cells and axons and do not contain neuronal cell bodies, such as corpus callosum and optic nerve, exon IXabcd transcripts were predominantly detected. Transcripts containing exons I, Vh, and VI

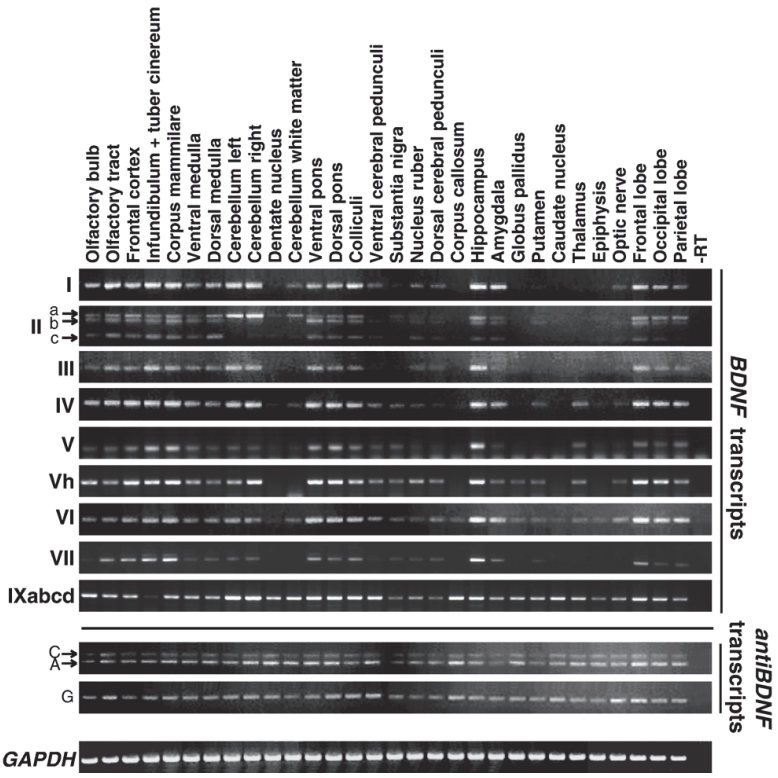

Fig. 4. Semiquantitative analysis of human *BDNF, antiBDNF,* and control *GAPDH* mRNA expression by RT-PCR in different human brain regions. Roman numerals on the left indicate the detected BDNF transcripts and the 5′-exon-specific primers used in combination with an antisense primer located in the *BDNF* coding region in exon IXd (Supplementary Table 1). A, C, and G refer to the respective *antiBDNF* transcripts shown in Fig. 1.

were also present in the optic nerve, although at low levels. *antiBDNF* transcripts were expressed in all studied brain structures at similar levels (Fig. 4).

Promoter activities of the 5′ flanking regions of human BDNF and antiBDNF upstream exons

Since the promoter regions of the human *BDNF* gene have not been analyzed previously and hypothesizing that a functional promoter precedes each of the identified 5′ exon of *BDNF* and that there is a promoter upstream of *antiBDNF* exon 1, the activities of nine potential promoter regions within the *BDNF* gene (namely the upstream sequences of exons I, II, III, IV, V, Vh, VI, VII, and IXabcd) and the region upstream of exon 1 of *antiBDNF* were analyzed for transcription-promoting activity using chloramphenicol acetyltransferase (CAT) assays. The putative promoter regions, each ~0.2–1.3 kb in length and containing a part of the respective 5′ UTR and 5′ flanking genomic sequence (Supplementary Fig. 1), were isolated and cloned into the pBLCAT2 vector in front of the *CAT* gene. The promoter constructs were transfected into human embryonic kidney HEK293T and mouse neuroblastoma N2a cells and the promoter activities were analyzed.

The results showed that all the regions upstream of the 5′ exons of the *BDNF* gene and exon 1 of the *antiBDNF* gene were functional and could activate CAT expression (Fig. 5). Thus it was concluded that the regions upstream of the 5′ exons of the *BDNF* gene and exon 1 of the *antiBDNF* gene act as separate promoters. However, the activities of the promoters varied and differences were detected also between the cell lines used. The activities of promoters upstream of exons II, V, Vh, and VII were somewhat lower compared to other promoters in N2a cells. In HEK293T cells the activities of these promoters could be detected only after longer reaction times. Other promoters showed similar activities in both of the cell lines with *BDNF* promoters upstream of exons III and VI and the *antiBDNF* promoter being the strongest in both cell lines. However, promoters upstream *BDNF* exons I, IV, and IXabcd were slightly (about twofold) more active in HEK293T cells than in N2a cells.

Human BDNF and antiBDNF transcripts form RNA duplexes in adult human brain in vivo

According to our data human *BDNF* and *antiBDNF* are co-expressed in many tissues studied. The complementary region

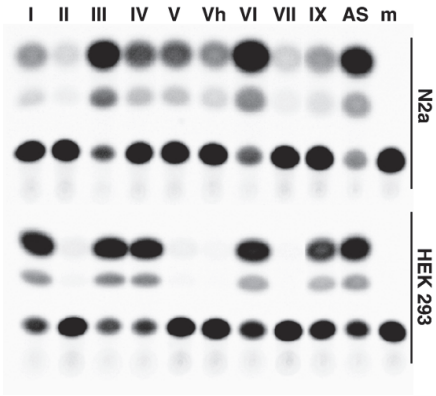

Fig. 5. Analyses of *BDNF* and *antiBDNF* promoter activities in HEK293T and N2a cells. The relative activities of the 5′ flanking regions of *BDNF* exons I, II, III, IV, V, Vh, VI, VII, and IXabcd and *antiBDNF* to promote CAT expression are shown. The promoter regions cloned in front of the *CAT* gene are shown in Supplementary Fig. 1. Note that the activities of *BDNF* promoters II, V, Vh, and VII in HEK239T cells were detectable using longer reaction times. m, mock-transfected cells, negative control.

of the majority of spliced *BDNF* and *antiBDNF* transcripts spans 222 nt or more depending on the splicing donor site used for *antiBDNF* exon 5 (Supplementary Fig. 2 and Supplementary Table 2). Based on this knowledge we hypothesized that if *antiBDNF* has a regulatory role in *BDNF* expression, the complementary RNAs might form RNA–RNA duplexes in vivo. To study this hypothesis we performed a PCR-based assay. Briefly, RNase A/T1-treated RNA from adult human cerebellum was used as a potential double-stranded RNA (dsRNA) template for cDNA synthesis and the existence of the duplexes was analyzed by PCR with primers specific for the complementary region of *BDNF* and *antiBDNF* (Supplementary Table 1). Our results showed that *BDNF*/*antiBDNF* dsRNA duplexes are present in the human brain in vivo (Fig. 6). Control experiments using a primer targeting the region of *antiBDNF* RNA outside of the complementary sequence in combination with a primer specific for the complementary region and experiments using RNA template in which the reverse transcription reaction was omitted showed that the RNA duplex-specific product was not the result of single-strand RNA (ssRNA) or genomic DNA contamination, respectively.

Discussion

Previous studies have revealed that the human *BDNF* gene consists of seven putative 5′ exons and one protein-coding exon [19,20]. However, the expression patterns of different exons have not been thoroughly studied and possible linkage of these exons to separate promoters has not been investigated. Here we show that the human *BDNF* gene, extending over 70 kb, contains 11 exons. The 3′ exon encodes all or most of the protein depending on the 5′ exon used. Independent of the 5′ exon usage, two separate polyadenylation signals in exon IX can be utilized in *BDNF* transcripts. In addition, our data showed that the human *BDNF* gene comprises nine functional promoters.

The structures of the human *BDNF* gene and transcripts determined in this study are in good agreement with the results obtained for the rat and mouse *BDNF* genes [17]. Some differences are present, though. First, human *BDNF* contains two more exons than rodent *BDNF*. Compared to the rat and mouse genes [17] there is an additional exon, exon Vh, linked to a promoter between exons V and VI. Human *BDNF* exon VIIIh, which is not linked to a separate promoter, is also not present in

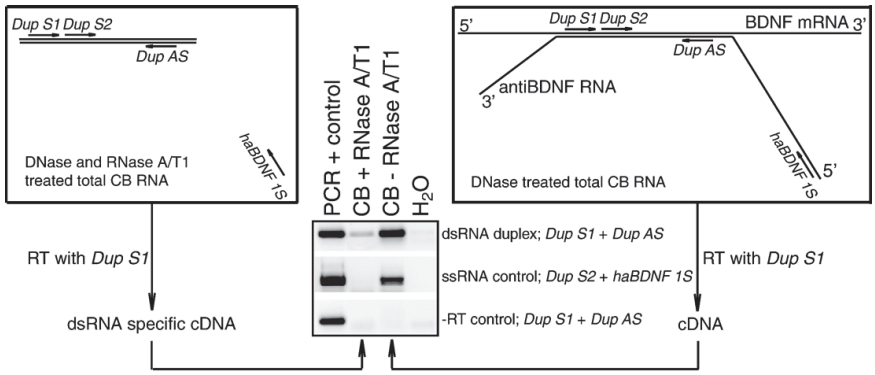

Fig. 6. *BDNF* and *antiBDNF* transcripts form dsRNA duplexes in the human brain in vivo. A schematic representation of the RNA duplex detection assay is shown. Briefly, total human cerebellar RNA was DNase treated. This RNA was divided into two—one half was treated with RNase A/T1 (box on the left) and the other was used as a −RNase control (box on the right). Both RNAs were reverse transcribed (RT) with a *BDNF*/*antiBDNF* complementary region-specific primer, Dup_S1. Subsequently these cDNAs were used as templates in PCR to detect *BDNF*/*antiBDNF* duplex with primers Dup_S1 and Dup_AS. ssRNA contamination control reaction was conducted with primers Dup_S2 and haBDNF_S1. The −RT reaction was used for detection of genomic DNA contamination using primers Dup_S1 and Dup_AS. Lines indicate RNAs, double line marks the complementary region of *BDNF* exon IXd and *antiBDNF* exon 5. Primer positions are indicated with arrows parallel with the lines and primer names are in italic. Human hippocampal cDNA synthesized using an oligo(dT) primer (PCR+control) was used as positive control for all the reactions.

rodent *BDNF*. Furthermore, cryptic splicing donor and acceptor sites are used in human exon IX leading to transcripts containing exons IXbd and IXabd. These transcripts have not been detected in rodents [17]. All this adds more complexity to the regulation of human *BDNF*. Second, in most cases, the usage of alternative promoters in the human *BDNF* gene leads to the expression of transcripts with different 5′ UTRs and with the protein-coding region in the common 3′ exon IXd. However, usage of an alternative upstream in-frame translation start site containing exon I, VII, or VIII could potentially lead to human BDNF prepro-proteins with longer N-termini. Only the translation initiation codon within exon I is characteristic of rodent *BDNF* genes, suggesting that there are more BDNF protein isoforms in human than in rodents. Third, although the transcription initiation sites are generally in good agreement with the respective regions in rodents, we identified more transcription start sites for human exons II and IV than had been reported previously for the rodent *BDNF* respective exons. Fourth, in contrast to the rodent *BDNF* genes we found that exon VIII of the human *BDNF* is not used as a 5′ exon as determined by the 5′ RACE analysis. We show that in human the rarely used exon VIII of *BDNF* is exclusively spliced to exon V. Exon V can also be spliced to exon IXd without including exon VIII. Exon VIII was not detected in any transcript other than the ones starting with exon V, pointing to a possible functional regulation between the usage of a certain promoter and subsequent splicing. Similar promoter-governed splicing regulation has been identified for the human nitric oxide synthase (*NOS1*) [21] and mouse *bcl-X* [22] genes, for example. This kind of splicing regulation is especially interesting provided the notion that exon VIII of *BDNF* contains one of the alternative ATGs that may lead to the synthesis of a prepro-BDNF protein with an alternative N-terminus.

We analyzed the splicing of the human *BDNF* pre-mRNAs and expression of consequent alternative mRNAs in detail. We found that *BDNF* transcripts containing exons II, III, IV, V, and VII are mostly brain-specific, whereas other *BDNF* mRNAs are also expressed at variable levels in nonneural tissues. Similar to the expression pattern of *BDNF* mRNAs in rodents [16,17], the most abundant transcripts in human nonneural tissues were transcripts containing exons VI and IXabcd that were expressed at high levels in several tissues, particularly heart, lung, skeletal muscle, testis, prostate, and placenta. To the best of our knowledge, these are the first data about the expression of alternatively spliced *BDNF* mRNAs in human nonneural tissues. In the human brain, expression of BDNF has been studied at the protein level using many different antibodies, the specificity of which is not always clear [23]. Fewer data are available about *BDNF* mRNA expression. In most studies on human *BDNF*, mRNA expression has been studied in only some regions of brain using postmortem tissue [23–25]. In two studies the expression of human *BDNF* mRNAs with alternative 5′ exons was examined in a few adult brain regions using RT-PCR [19,26]. Our study is the first to examine *BDNF* exon-specific mRNA levels across the whole human brain, thus adding important new data to *BDNF* expression in adult human brain. In the adult human brain, high levels of *BDNF* mRNAs were present in the hippocampus, cerebral cortex, amygdala, and cerebellum, which is similar to the previously reported data on the *BDNF* expression in rodent brain [10,11,15,27,28]. *BDNF* is expressed predominantly in neurons, although some studies have identified *BDNF* expression also in rodent astrocytes [29,30], microglia [31], and oligodendrocytes [32], both in vivo and in vitro. Here we show that some of the alternatively spliced human *BDNF* mRNAs, particularly transcripts containing exons VI and IXabcd, are present in vivo in the corpus callosum and optic nerve containing mostly oligodendroglial cells and axonal projections.

Gene expression in eukaryotes is a highly coordinated process involving regulation at many different levels, among which the regulation of transcription is one of the most important. Several types of *cis*-acting DNA sequence elements, including promoters, contribute to this process. About 18% of human genes have multiple promoters, which regulate and increase their transcriptional and translational potential [33]. Human *BDNF* promoter IV is the only promoter of the human *BDNF* gene that had been characterized so far [18]. In this study we show that *BDNF* gene expression is under the control of at least nine alternative tissue-specific promoters linked to separate 5′ exons. Alternative promoters of the *BDNF* gene could also be involved in developmental stage-specific expression and cell-type-specific expression, giving additional flexibility to the control of *BDNF* expression. Therefore, the data presented in this study show that the expression of the human *BDNF* gene is highly regulated at the level of transcription.

One of the results of our study was the characterization of endogenous noncoding antisense RNAs transcribed from the human *BDNF* gene locus. According to our data the *antiBDNF* gene consists of 10 exons and one functional promoter upstream of exon 1. We show that *antiBDNF* transcripts are expressed in almost all adult human tissues analyzed. High levels of *antiBDNF* mRNAs are present in the brain, kidney, spinal cord, and testis. Expression levels are low in adrenal gland, bone marrow, pancreas, small intestine, uterus, and some other tissues. In the adult brain, all *antiBDNF* transcripts are expressed at similar levels in all brain regions analyzed. We found that hundreds of different noncoding RNAs might be generated from the *antiBDNF* gene as a result of alternative splicing. Alternatively spliced isoform diversity is common to many eukaryotic organisms and it is particularly widely used in the nervous system [34]. Interestingly, *antiBDNF* is not present in rodents [16,17]. *antiBDNF* ESTs are also not available for chimpanzee and rhesus monkey although highly homologous sequences are present in the genomes of these animals (data not shown). All this suggests that *antiBDNF* could have evolved during primate/hominid evolution, as was proposed also by Liu et al. [19].

In this study we have shown that in the human brain *BDNF* and *antiBDNF* transcripts form dsRNA duplexes in vivo. This indicates that *antiBDNF* transcripts could have an important role in the regulation of *BDNF* expression in human. Several studies have shown that natural antisense transcripts (NATs) are involved in the regulation of gene expression in eukaryotes [35,36]. For example, NATs have been suggested to play an

important role in the regulation of several genes encoding transcription factors that are important in eye development and function in mice [37]. Characterization of overlapping transcripts in various species indicates that this form of RNA-mediated gene regulation represents a widespread phenomenon [36,38]. NATs are particularly prevalent in the nervous system where they regulate the expression of several genes [35]. In the case of human *BDNF* and *antiBDNF,* the transcripts could act as *cis*-antisense RNAs and generate siRNAs targeting one of the initial transcripts, as do the natural *cis*-siRNAs described for genes involved in salt tolerance in *Arabidopsis thaliana* [39]. Other possible regulatory functions of *antiBDNF* would be direct inhibition of *BDNF* transcription or translation and/or regulation of *BDNF* pre-mRNA splicing. Our results show that the expression of *antiBDNF* and *BDNF* transcripts in different tissues is not mutually exclusive and that the levels of *BDNF* mRNA do not appear to be specifically reduced in tissues that express high levels of *antiBDNF* transcripts. However, it is possible that the *antiBDNF* transcripts could modulate the levels of *BDNF* provided they are coexpressed in the same cell.

In conclusion, this detailed characterization of the human *BDNF* gene locus opens up insights into the mechanisms governing *BDNF* gene regulation in human.

Materials and methods

RNA isolation, RT-PCR, and cloning and sequencing of RT-PCR products

Total RNAs from 23 human tissues were obtained from Clontech. Total RNAs from postmortem adult human brain regions were isolated using the RNAwiz RNA isolation reagent and treated with DNase (Ambion, USA) according to the supplier's protocol. All experiments with human tissues were approved by the local ethical committee. Five micrograms of total RNA was reverse-transcribed to cDNA with an oligo(dT) primer (Proligo, France) and SuperScript III reverse transcriptase using the SuperScript III First-Strand Synthesis System (Invitrogen, USA). PCR amplification was carried out using HOT FIREPol DNA polymerase (Solis Biodyne, Estonia), according to the manufacturer's instructions. One-fortieth of the first-strand cDNA reaction mix was used in the PCR. The exon-specific PCR primers were designed based on the sequence of the human *BDNF* gene (NCBI Accession No. AF411339), ESTs, and mRNA sequences from GenBank. Sequences for all primers are listed in Table 1 of the supplementary material. The lengths of the PCR products using the primer hBDNF_IXbAS in combination with the following primers were hBDNF_IS, 472 bp; hBDNF_IIS, 610, 527, and 312 bp; hBDNF_IIIS, 347 bp; hBDNF_IVS, 412 bp; hBDNF_VS, 673, 556, and 273 bp; hBDNF_VhS, 340 bp; hBDNF_VIS, 494, 387, and 369 bp; hBDNF_VIIS, 429 and 328 bp; and hBDNF_IXS, 597 and 363 bp. The lengths of the longest PCR products with haBDNF_1S in combination with the following primers were haBDNF_9AS, 947 bp, and haBDNF_10AS, 1483 bp. All products from the RT-PCR were cloned into the pCRII-TOPO vector (Invitrogen) and sequenced.

Analyses of transcription start sites

The transcription start sites for the *BDNF* and *antiBDNF* transcripts were detected with 5′ RACE using the GeneRacer Kit (Invitrogen) for full-length, RNA ligase-mediated rapid amplification of 5′ cDNA ends, according to the manufacturer's instructions. Briefly, 5 μg of total RNA from human hippocampus and cerebellum was dephosphorylated and decapped. The GeneRacer RNA oligo was ligated to the decapped 5′ ends of the full-length mRNAs and reverse transcription of the mRNAs was performed. RACE-ready cDNAs were used as templates for subsequent PCR using the GeneRacer 5′ primer in combination with *BDNF* exon-specific primers (Supplementary Table 1). The PCR products were gel purified and cloned into the pCRII-TOPO (Invitrogen) vector and verified by sequencing.

BDNF promoter–CAT reporter plasmids and CAT assay

PCR was performed to amplify promoter fragments of the human *BDNF* and *antiBDNF* genes with the appropriate primers (Supplementary Table 1). Human genomic DNA was used as a template. The amplified fragments were cloned into the pBL-CAT2 plasmid upstream of the coding region of the *CAT* gene, replacing the thymidine kinase promoter, and verified by sequencing. HEK293T and N2a cells were used for the analysis of promoter activities. The cells were maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum at 37°C in a 5% CO_2 atmosphere. Transfection of cells with the promoter–CAT reporter plasmids was performed with FuGENE 6 (Roche Diagnostics, USA) according to the supplier's instructions. The transfected cells were harvested 40 h after transfection. The samples were incubated with ^{14}C-labeled chloramphenicol and acetyl-CoA and the radioactive products were separated by thin-layer chromatography silica gel (Merck, USA) and visualized by autoradiography as described before [15].

BDNF/antiBDNF RNA duplex analyses

Ten micrograms of human cerebellar RNA, isolated with RNAwiz (Ambion), was treated with DNase (Ambion Turbo DNA-*free*) for 30 min at 37°C according to the manufacturer's instructions, precipitated, and treated with RNase A/T1 (Ambion) for 30 min at 37°C in RNase buffer (300 mM NaCl, 10 mM Tris, pH 7.4, and 5 mM EDTA). The reaction was terminated with 0.4 mg/ml proteinase K (Roche) in 200 mM Tris, pH 7.4, 25 mM EDTA, and 1% SDS for 30 min at 37°C, and the dsRNA was phenol/chloroform extracted. cDNA was synthesized from this dsRNA using the *BDNF/antiBDNF* complementary region-specific primer Dup_S1 (Supplementary Table 1) with Superscript III (Invitrogen) according to the manufacturer's instructions in the presence of 5% DMSO. In subsequent RNA duplex detection PCR 1/20 of the cDNA was used along with primers specific for the *BDNF/antiBDNF* complementary region, Dup_S1 and Dup_AS (Supplementary Table 1; product size 156 bp). Primer haBDNF_1S recognizing *antiBDNF* exon 1 in combination with Dup_S2 (Supplementary Table 1; product size 490 bp) was used for detection of ssRNA contamination. In addition, a −RT reaction was used for detection of genomic DNA contamination using primers Dup_S1 and Dup_AS. All PCRs were performed as follows: 40 cycles of 94°C for 30 s, 57°C for 30 s, and 72°C for 30 s.

Acknowledgments

We thank Enn Jõeste and Liina Kiho from North Estonian Regional Hospital, Tallinn, for help in collecting tissue samples; Epp Väli for technical assistance; Richard Tamme for critically reading the manuscript; and Mart Speek for experimental advice. This work was supported by the following grants to T. T.: Wellcome Trust International Senior Research Fellowship (Grant 067952) and Estonian Ministry of Education and Research (Grant 0222602).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ygeno.2007.05.004.

References

[1] D.K. Binder, H.E. Scharfman, Brain-derived neurotrophic factor, Growth Factors 22 (2004) 123–131.
[2] M. Bibel, Y.A. Barde, Neurotrophins: key regulators of cell fate and cell shape in the vertebrate nervous system, Genes Dev. 14 (2000) 2919–2937.

[3] S. Pezet, S.B. McMahon, Neurotrophins: mediators and modulators of pain, Annu. Rev. Neurosci. 29 (2006) 507–538.
[4] B. Alsina, T. Vu, S. Cohen-Cory, Visualizing synapse formation in arborizing optic axons in vivo: dynamics and modulation by BDNF, Nat. Neurosci. 4 (2001) 1093–1101.
[5] M. Korte, P. Carroll, E. Wolf, G. Brem, H. Thoenen, T. Bonhoeffer, Hippocampal long-term potentiation is impaired in mice lacking brain-derived neurotrophic factor, Proc. Natl. Acad. Sci. USA 92 (1995) 8856–8860.
[6] B. Lu, BDNF and activity-dependent synaptic modulation, Learn. Mem. 10 (2003) 86–98.
[7] M.F. Egan, M. Kojima, J.H. Callicott, T.E. Goldberg, B.S. Kolachana, A. Bertolino, E. Zaitsev, B. Gold, D. Goldman, M. Dean, B. Lu, D.R. Weinberger, The BDNF val66met polymorphism affects activity-dependent secretion of BDNF and human memory and hippocampal function, Cell 112 (2003) 257–269.
[8] C. Zuccato, E. Cattaneo, Role of brain-derived neurotrophic factor in Huntington's disease, Prog. Neurobiol. (2007).
[9] R. Katoh-Semba, I.K. Takeuchi, R. Semba, K. Kato, Distribution of brain-derived neurotrophic factor in rats and its changes with development in the brain, J. Neurochem. 69 (1997) 34–42.
[10] J.M. Conner, J.C. Lauterborn, Q. Yan, C.M. Gall, S. Varon, Distribution of brain-derived neurotrophic factor (BDNF) protein and mRNA in the normal adult rat CNS: evidence for anterograde axonal transport, J. Neurosci. 17 (1997) 2295–2313.
[11] P. Ernfors, C. Wetmore, L. Olson, H. Persson, Identification of cells in rat brain and peripheral tissues expressing mRNA for members of the nerve growth factor family, Neuron 5 (1990) 511–526.
[12] T. Timmusk, H. Persson, M. Metsis, Analysis of transcriptional initiation and translatability of brain-derived neurotrophic factor mRNAs in the rat brain, Neurosci. Lett. 177 (1994) 27–31.
[13] P.C. Maisonpierre, M.M. Le Beau, R. Espinosa III, N.Y. Ip, L. Belluscio, S.M. de la Monte, S. Squinto, M.E. Furth, G.D. Yancopoulos, Human and rat brain-derived neurotrophic factor and neurotrophin-3: gene structures, distributions, and chromosomal localizations, Genomics 10 (1991) 558–568.
[14] P.C. Maisonpierre, L. Belluscio, B. Friedman, R.F. Alderson, S.J. Wiegand, M.E. Furth, R.M. Lindsay, G.D. Yancopoulos, NT-3, BDNF, and NGF in the developing rat nervous system: parallel as well as reciprocal patterns of expression, Neuron 5 (1990) 501–509.
[15] T. Timmusk, K. Palm, M. Metsis, T. Reintam, V. Paalme, M. Saarma, H. Persson, Multiple promoters direct tissue-specific expression of the rat BDNF gene, Neuron 10 (1993) 475–489.
[16] Q.R. Liu, L. Lu, X.G. Zhu, J.P. Gong, Y. Shaham, G.R. Uhl, Rodent BDNF genes, novel promoters, novel splice variants, and regulation by cocaine, Brain Res. 1067 (2006) 1–12.
[17] T. Aid, A. Kazantseva, M. Piirsoo, K. Palm, T. Timmusk, Mouse and rat BDNF gene structure and expression revisited, J. Neurosci. Res. 85 (2007) 525–535.
[18] H. Fang, J. Chartier, C. Sodja, A. Desbois, M. Ribecco-Lutkiewicz, P.R. Walker, M. Sikorska, Transcriptional activation of the human brain-derived neurotrophic factor gene promoter III by dopamine signaling in NT2/N neurons, J. Biol. Chem. 278 (2003) 26401–26409.
[19] Q.R. Liu, D. Walther, T. Drgon, O. Polesskaya, T.G. Lesnick, K.J. Strain, M. de Andrade, J.H. Bower, D.M. Maraganore, G.R. Uhl, Human brain derived neurotrophic factor (BDNF) genes, splicing patterns, and assessments of associations with substance abuse and Parkinson's disease, Am. J. Med. Genet. B Neuropsychiatr. Genet. 134 (2005) 93–103.
[20] M. Aoyama, K. Asai, T. Shishikura, T. Kawamoto, T. Miyachi, T. Yokoi, H. Togari, Y. Wada, T. Kato, A. Nakagawara, Human neuroblastomas with unfavorable biologies express high levels of brain-derived neurotrophic factor mRNA and a variety of its variants, Cancer Lett. 164 (2001) 51–60.
[21] Y. Wang, D.C. Newton, G.B. Robb, C.L. Kau, T.L. Miller, A.H. Cheung, A.V. Hall, S. VanDamme, J.N. Wilcox, P.A. Marsden, RNA diversity has profound effects on the translation of neuronal nitric oxide synthase, Proc. Natl. Acad. Sci. USA 96 (1999) 12150–12155.
[22] A. Pecci, L.R. Viegas, J.L. Baranao, M. Beato, Promoter choice influences alternative splicing and determines the balance of isoforms expressed from the mouse bcl-X gene, J. Biol. Chem. 276 (2001) 21062–21069.
[23] M.G. Murer, Q. Yan, R. Raisman-Vozari, Brain-derived neurotrophic factor in the control human brain, and in Alzheimer's disease and Parkinson's disease, Prog. Neurobiol. 63 (2001) 71–124.
[24] H.S. Phillips, J.M. Hains, M. Armanini, G.R. Laramee, S.A. Johnson, J.W. Winslow, BDNF mRNA is decreased in the hippocampus of individuals with Alzheimer's disease, Neuron 7 (1991) 695–702.
[25] M.J. Webster, C.S. Weickert, M.M. Herman, J.E. Kleinman, BDNF mRNA expression during postnatal development, maturation and aging of the human prefrontal cortex, Brain Res. Dev. Brain. Res. 139 (2002) 139–150.
[26] D. Garzon, G. Yu, M. Fahnestock, A new brain-derived neurotrophic factor transcript and decrease in brain-derived neurotrophic factor transcripts 1, 2 and 3 in Alzheimer's disease parietal cortex, J. Neurochem. 82 (2002) 1058–1064.
[27] S. Ceccatelli, P. Ernfors, M.J. Villar, H. Persson, T. Hokfelt, Expanded distribution of mRNA for nerve growth factor, brain-derived neurotrophic factor, and neurotrophin 3 in the rat brain after colchicine treatment, Proc. Natl. Acad. Sci. USA 88 (1991) 10352–10356.
[28] M. Hofer, S.R. Pagliusi, A. Hohn, J. Leibrock, Y.A. Barde, Regional distribution of brain-derived neurotrophic factor mRNA in the adult mouse brain, EMBO J. 9 (1990) 2459–2464.
[29] D.F. Condorelli, P. Dell'Albani, G. Mudo, T. Timmusk, N. Belluardo, Expression of neurotrophins and their receptors in primary astroglial cultures: induction by cyclic AMP-elevating agents, J. Neurochem. 63 (1994) 509–516.
[30] F. Zafra, D. Lindholm, E. Castren, J. Hartikka, H. Thoenen, Regulation of brain-derived neurotrophic factor and nerve growth factor mRNA in primary cultures of hippocampal neurons and astrocytes, J. Neurosci. 12 (1992) 4793–4799.
[31] S. Elkabes, E.M. DiCicco-Bloom, I.B. Black, Brain microglia/macrophages express neurotrophins that selectively regulate microglial proliferation and function, J. Neurosci. 16 (1996) 2508–2521.
[32] X. Dai, L.D. Lercher, P.M. Clinton, Y. Du, D.L. Livingston, C. Vieira, L. Yang, M.M. Shen, C.F. Dreyfus, The trophic role of oligodendrocytes in the basal forebrain, J. Neurosci. 23 (2003) 5846–5853.
[33] J.R. Landry, D.L. Mager, B.T. Wilhelm, Complex controls: the role of alternative promoters in mammalian genomes, Trends Genet. 19 (2003) 640–648.
[34] D. Lipscombe, Neuronal proteins custom designed by alternative splicing, Curr. Opin. Neurobiol. 15 (2005) 358–363.
[35] S. Korneev, M. O'Shea, Natural antisense RNAs in the nervous system, Rev. Neurosci. 16 (2005) 213–222.
[36] C. Vanhee-Brossollet, C. Vaquero, Do natural antisense transcripts make sense in eukaryotes? Gene 211 (1998) 1–9.
[37] G. Alfano, C. Vitiello, C. Caccioppoli, T. Caramico, A. Carola, M.J. Szego, R.R. McInnes, A. Auricchio, S. Banfi, Natural antisense transcripts associated with genes involved in eye development, Hum. Mol. Genet. 14 (2005) 913–923.
[38] C. Morey, P. Avner, Employment opportunities for non-coding RNAs, FEBS Lett. 567 (2004) 27–34.
[39] O. Borsani, J. Zhu, P.E. Verslues, R. Sunkar, J.K. Zhu, Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in Arabidopsis, Cell 123 (2005) 1279–1291.
[40] A.M. Marini, X. Jiang, X. Wu, F. Tian, D. Zhu, P. Okagaki, R.H. Lipsky, Role of brain-derived neurotrophic factor and NF-kappaB in neuronal plasticity and survival: from genes to phenotype, Restor. Neurol. Neurosci. 22 (2004) 121–130.

PUBLICATION II

II Kazantseva, A.∗, Sepp, M.∗, Kazantseva, J., Sadam, H., Pruunsild, P., Timmusk, T., Neuman, T., Palm, K. (2009). N-terminally truncated BAF57 isoforms contribute to the diversity of SWI/SNF complexes in neurons. Journal of Neurochemistry, 109(3), 807-818

Journal of Neurochemistry

JOURNAL OF NEUROCHEMISTRY | 2009 | 109 | 807–818 doi: 10.1111/j.1471-4159.2009.06005.x

N-terminally truncated BAF57 isoforms contribute to the diversity of SWI/SNF complexes in neurons

Anna Kazantseva,*[,1] Mari Sepp,†[,1] Jekaterina Kazantseva,‡[,1] Helle Sadam,* Priit Pruunsild,† Tõnis Timmusk,† Toomas Neuman†[,§] and Kaia Palm*

*Protobios LLC, Tallinn, Estonia

†The Department of Gene Technology, Tallinn University of Technology, Tallinn, Estonia

‡The Competence Center for Cancer Research, Tallinn, Estonia

§OncoTx, Inc., Pasadena, California, USA

Abstract

The SWItch/Sucrose NonFermentable, a nucleosome remodeling complex (SWI/SNF) chromatin-remodelling complexes act upon the nucleosomal structure and regulate transcription, replication, repair of chromatin and splicing. In this study, we present evidence that human, mouse and rat genes encoding one of the SWI/SNF complex subunits, BAF57, undergo neuron-specific splicing of exons II, III and IV. Alternative splicing yields in at least three isoforms of BAF57 protein that have truncated N-termini (N-BAF57s). The transcripts encoding N-BAF57 isoforms are predominantly expressed in the nervous system. The biochemical fractionation data supported by the results of the co-immunoprecipitation analysis show that N-BAF57 isoforms associate into protein complexes together with Brg1, Brm, BAF155 and BAF170. Transient over-expression of N-BAF57 isoforms in non-neural cells affects the level of expression of certain neuron-restrictive silencer element-containing genes. Together these data suggest that neuronal isoforms of BAF57 contribute to functional SWI/SNF complexes regulating neurogenesis.

Keywords: alternative splicing, BAF57, chromatin regulator, neuron-specific, NRSF/RE1, SWI/SNF.

J. Neurochem. (2009) **109**, 807–818.

Chromatin remodelling is one of the primary routes of regulation of developmental gene expression. The functional state of chromatin can be regulated by two different strategies: chromatin remodelling and histone tail modification (De la Serna *et al.* 2006). In mammals, the SWItch/Sucrose NonFermentable, a nucleosome remodeling complex (SWI/SNF) family of nucleosome-remodelling complexes comprises of a number of 2-MDa multiprotein complexes containing 8–10 subunits (Kingston and Narlikar 1999; Vignali *et al.* 2000; Lessard *et al.* 2007). These complexes include a central catalytic subunit, Brm/SNF2a or Brg1/SNF2b, and several variable Brg1-associated factors (BAF) that contribute to the enzymatic activity of the complex and facilitate the recruitment to specific transcription factors (Nie *et al.* 2000; Sif *et al.* 2001). The role of each BAF has not been completely elucidated, but it is clear that Brg1 or hBrm in combination with BAF170, BAF155, and BAF45 can reconstitute chromatin remodelling, while others such as BAF60 and BAF57 mediate interaction with transcriptional activators or repressors (Phelan *et al.* 1999; Ito *et al.* 2001; Belandia *et al.* 2002; Hsiao *et al.* 2003; Pal *et al.* 2003). Structurally different SWI/SNF complexes might have specialized functions (Wang *et al.* 1996). Indeed, BAF complexes containing subunits of mSin3A complex alter nucleosome structure in a different manner (Underhill *et al.* 2000; Sif *et al.* 2001). Most recently, SWI/SNF chromatin remodelling complexes were identified that exhibit opposing

Received January 14, 2009; revised manuscript received February 12, 2009; accepted February 17, 2009.

Address correspondence and reprint requests to Kaia Palm, Protobios LLC, Akadeemia tee 15A, Tallinn 12618, Estonia.
E-Mail: kaia@protobios.com

[1]These authors contributed equally to this study.

Abbreviations used: CHRM4, cholinergic receptor, muscarinic 4; COX1, cyclooxygenase 1; CTD, C-terminal domain; HMG, high-mobility group; L1CAM, L1 cell adhesion molecule; NRSE, neuron-restrictive silencer element; NRSF/REST, neuron-restrictive silencer factor/repressor element-1 silencing transcription factor; RE1, repressor element-1; SDS, sodium dodecyl sulfate; SWI/SNF, SWItch/Sucrose NonFermentable, a nucleosome remodeling complex; SYP, synaptophysin.

roles in cell-cycle control (Nagl *et al.* 2007). Also, distinct complexes have been purified that contain different BAF60 family members with differential pattern of expression (Wang *et al.* 1996). Recent data also show that BAF45a and BAF53a subunits are necessary for proliferation of neural progenitor cells, whereas BAF45b, BAF45c, and BAF53b are recruited as neural progenitors exit the cell cycle (Lessard *et al.* 2007). These data clearly show that the SWI/SNF complex can be involved in either transcriptional activation or repression. It is suggested that DNA-binding activators as well as repressors require SWI/SNF activity when their binding sites lie within nucleosomes or their binding affinity is low (Sudarsanam and Winston 2000). Human BAF57 is a modular protein containing a Proline-rich region, a high-mobility group (HMG) domain, a conserved region termed the NHRLI (protein domain, named after a group of conserved amino acids in this domain) domain, a putative coiled-coil domain, and a charged C-terminal region (Wang *et al.* 1998; Papoulas *et al.* 2001). Activity and function of mammalian BAF57 proteins have been assayed in different systems. An essential role for BAF57 has been established in CD4 silencing during T-cell lineage commitment (Chi *et al.* 2002). The role of BAF57 in gene silencing is supported by the finding that BAF57 interacts with methyl CpG binding protein 2 (Rett syndrome) and participates in methyl CpG binding protein 2 (Rett syndrome)-dependent transcriptional repression (Harikrishnan *et al.* 2005). Alternatively, BAF57 has been shown to interact directly with ERalpha and the androgen receptor to mediate hormone-dependent transcriptional activation (Belandia *et al.* 2002; Link *et al.* 2005; García-Pedrero *et al.* 2006). Growing evidence shows that SWI/SNF complex has a role in cancer development and BAF57 is essential for the regulation of tumour suppressor genes *CYLD* (Wang *et al.* 2005) and *BRCA1* (Baker *et al.* 2003). Neuronal restricted silencing factor/repressor element-1 silencing transcription factor (NRSF/REST) a key transcriptional repressor for neuron-specific genes in non-neuronal and neuronal cells (Schoenherr *et al.* 1996; Chen *et al.* 1998; Palm *et al.* 1998; Huang *et al.* 1999). BAF57 via its C-terminus interacts with transcriptional co-repressor CoREST and represses neuron-restrictive silencer element (NRSE)-containing neuronal genes (Battaglioli *et al.* 2002). Also, Brg1, another member of SWI/SNF family facilitates NRSF/REST repression by increasing the interaction between REST and chromatin (Ooi *et al.* 2006). In this work, we report that alternative splicing modulates the function of the *BAF57* gene products by in-frame deletions of exons 2, 3 and 4 in the region of BAF57 mRNA that encodes the N-terminus preceding the HMG domain. Our data show that *N-BAF57* mRNAs are exclusively expressed in neurons and not in glial cells. Presented functional studies provide evidence that N-BAF57 proteins are associated with SWI/SNF-like chromatin remodelling complexes containing SWI2/SNF2-like ATPases, Brg1 and Brm. In addition, we show that N-BAF57 proteins may affect the expression of NRSE/repressor element-1 (RE1)-containing genes. Altogether these findings suggest a role of N-BAF57s in the nervous system.

Materials and methods

Construction of BAF57 and N-BAF57 expression vectors

The coding regions of human *BAF57* (*hBAF57*) and its two alternatively spliced isoforms (*hBAF57Δ4* and *hBAF57Δ3-4*) were PCR amplified using human total brain cDNA as a template and cloned into pcDNA3.1/V5-His mammalian expression vector (Invitrogen, Carlsbad, CA, USA). The primers used for generation of expression constructs were as follows: 5′-GCGGTGTCTCAGATTCATTC-3′ and 5′-TTCTTTTTTCTCATCTTCTGGTA-3′. The resulting plasmids *BAF57-V5*, *BAF57Δ4-V5* and *BAF57Δ3-4-V5* contain C-terminal V5 and His tags and do not contain exon XI. The expression constructs were verified by sequencing and *in vitro* translation assay (Promega, Madison, WI, USA).

Cell culture and transient transfection assays

Human melanoma WM266-4 cells and mouse neuroblastoma Neuro2A cells were grown in Dulbecco's modified Eagle's medium (PAA, Linz, Austria) containing 10% foetal bovine serum (PAA), 50 units/mL penicillin (PAA) and 50 μg/mL streptomycin (PAA) at 37°C in 5% CO_2. Primary hippocampal and cortical neurons from E19 rat embryo were plated at near-confluence onto poly-L-lysine coated dishes and grown in Neurobasal™-A medium (GIBCO™, Invitrogen) supplemented with B-27 Supplement (GIBCO™, Invitrogen), glutamine (1 mM), penicillin and streptomycin at 37°C in 5% CO_2. Primary cell cultures of astrocytes were prepared as described earlier (O'Malley *et al.* 1994). Human neural stem cells were grown as in Palm *et al.* (Palm *et al.* 2000).

For transient transfection assays, WM266-4 cells were plated onto 6-well plates 24 h before transfection. The cells were transfected using FuGene HD (Roche, Basel, Switzerland) with DNA to reagent ratio 1 : 5 according to the manufacturer's recommendations. 3 μg of plasmids *BAF57-V5*, *BAF57Δ4-V5* or *BAF57Δ3-4-V5* was used for transfection. pcDNA3.1/V5-His was used as a reference control. Cells were co-transfected with 0.5 μg of a β-galactosidase expression vector. The activities of β-galactosidase were measured with the β-gal assay kit (Promega) to normalize transfection efficiency. All assays were performed in triplicates.

For immunoanalysis, WM266-4 cells were seeded onto glass coverslips in 24-well plates. After 24 h, the cells were transfected with *BAF57-V5*, *BAF57Δ4-V5* or *BAF57Δ3-4-V5* constructs and pcDNA3.1/V5-His vector with DNA to reagent ratio 1 : 3 using FuGene HD. 2 μg of each of the construct was used per transfection.

For biochemical fractionation and western blot analysis, Neuro2A cells were seeded onto 60 mm dishes and transfected with 3 μg of *BAF57-V5*, *BAF57Δ4-V5*, *BAF57Δ3-4-V5* or pcDNA3.1/V5-His constructs using FuGeneHD with DNA to reagent ratio 1 : 3.

RNA extraction, reverse transcription and RT-PCR analysis

Human post-mortem tissues were obtained from the North-Estonian Regional Hospital (Tallinn, Estonia) with the approval of the local

ethical committee. Total RNA from various human brain regions, human non-neural tissues, developing and adult mouse and rat tissues was purified by RNAWiz (Ambion, Austin, TX, USA) as recommended by the manufacturer. Total RNA from different human cell lines and human stem cells was isolated using RNAqueous™ RNA isolation kit (Ambion) according to manufacturer's instructions. RNA from rat primary hippocampal/cortical culture and glial cells was isolated using RNeasy Kit (Qiagen, Valencia, CA, USA) according to manufacturer's instructions.

Two micrograms of RNA from each sample was subjected to reverse transcription using a SuperScript III first-strand cDNA synthesis kit (Invitrogen). Oligo(dT) was used to prime the reverse transcription catalyzed by SuperScript III in the presence of RNase Inhibitor (Bioron, Ludwigshafen, Germany). The resulting cDNAs were used in subsequent PCR reactions. RT-PCR were performed using HotFire® Polymerase (Solis Biodyne, Tartu, Estonia), 35 amplification cycles and an annealing temperature of 58°C. The products were analyzed by 2% agarose gel electrophoresis and visualized with ethidium bromide. PCR products were verified by sequencing. Hypoxanthine-guanine phosphoribosyl transferase 1 (*HPRT1*) or glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) mRNA expression was analyzed as an internal control. PCR primers for *BAF57* gene were designed on the basis of the ESTs available in NCBI database and were as follows: for the human gene *hBAF57*s 5′-GCGGTGTCTCAGATTCATTC-3′ and *hBAF57as* 5′-CTCATGTACGGTTCTCCTTTC-3′, for the mouse and rat genes *mrBAF57s* 5′-CAAAGGGAAGCGAAGCCGGA-3′ and *mrBAF57as* 5′-CTCATGTACGGTTCTCCTTTC-3′. The PCR primers for the human *REST* gene were *hRESTs* 5′-ATACAGCACACCTGAAACACCACA-3′ and *hRESTas* 5′-CCTGTCTTGCATGGCGGGTTACTTCA-3′, and for the human *CoREST* gene were *hCoRESTs* 5′-CGAGGATGAACTGGAAGAGG-3′ and *hCoRESTas* 5′-CCTCTTCGGGCATCTTAATGG-3′. Alternative splicing of *BAF57* exons II, III and IV was further confirmed by RT-PCR using primers that span the appropriate exons/ introns and by sequence analysis. Sequences of the different transcripts were submitted to NCBI database.

Real-time quantitative RT-PCR

Forty-eight hours post-transfection, total RNA from WM266-4 cells was isolated using RNAqueous™ RNA isolation kit (Ambion) and first-strand cDNA was synthesized as described. Real-time PCR was carried out with Platinum® SYBR® Green qPCR SuperMix-UDG kit (Invitrogen) using Roche LightCycler 2.0 according to manufacturer's instructions. PCR cycling conditions consisted of 95°C for 2 min and 40 cycles of 95°C for 15 s, 60°C for 30 s, 72°C for 30 s. The primers for human cholinergic receptor, muscarinic 4 (*CHRM4*) were *hCHRM4s* 5′-CCGTGGCTGATAAGGACACTT-3′ and *hCHRM4as* 5′-GTCTGCTTCGTCACAATCTGG-3′, for L1 cell adhesion molecule (*L1CAM*) *hL1CAMs* 5′-GAACCCATTGACCTCCGGG-3′ and *hL1CAMas* 5′-CAGCGGTACTCGCCATCAT-3′, for cyclooxygenase 1 (*COX1*) *hCOX1s* 5′-CTCCCAGGAGTACAGCTACGA-3′ and *hCOX1as* 5′-CCAGCAATCTGGCGAGAGA-3′, and for synaptophysin (*SYP*) *hSYPs* 5′-AGTTGGGGACTACTCCTCGTC-3′ and *hSYPas* 5′-GGCCCTTTGTTATTCTCTCGGTA-3′. Relative expression was normalized to *GAPDH* mRNA levels. The relative expression level of each of the transcripts was quantified using the comparative $\Delta\Delta$Ct method. For each primer set, qRT-PCR was repeated at least three times per sample.

Immunoanalysis of subcellular localization of BAF57-V5 proteins

Twenty-four hours post-transfection, coverslips with WM266-4 cells were rinsed with phosphate-buffered saline buffer and fixed with 4% paraformaldehyde for 15 min. Subsequently, cells were blocked and permeabilised with 2% bovine serum albumin (Sigma, St Louis, MO, USA) and 0.5% Triton X-100 (Scharlau Chemie S.A, Barcelona, Spain), respectively. Then the cells were incubated with mouse monoclonal anti-V5 antibody (Invitrogen), dilution 1 : 300, and rabbit polyclonal anti-RNA polymerase II C-terminal domain (CTD) repeat YSPTSPS (phospho S5) antibody (Abcam, Cambridge, UK; ab5131), dilution 1 : 300, overnight at 4°C. After washing three times with phosphate-buffered saline with 0.1% Tween20 (BioTop, Oulu, Finland), the cells were incubated with Alexa Fluor® 488-conugated anti-mouse (Invitrogen) and Alexa Fluor® 546-conjugated anti-rabbit (Invitrogen) secondary antibodies at 22°C for 1 h. After a final wash, the cells were mounted with ProLong® Gold anti-fade reagent containing 4′,6-diamidino-2-phenylindole (DAPI) (Invitrogen) and analyzed using a laser scanning microscope (LSM 5 *LIVE*, Zeiss, Jena, Germany).

Preparation of nuclear extracts, biochemical fractionation and western blot analysis

Nuclear extracts were prepared from Neuro2A cells transfected with *BAF57-V5, BAF57Δ4-V5, BAF57Δ3-4-V5* or pcDNA3.1/V5-His constructs. Shortly, cells were resuspended in hypotonic buffer (10 mM HEPES, 1.5 mM MgCl2, 10 mM KCl, 0.5 mM dithiothreitol), homogenized and the nuclei were extracted by 300 mM KCl. The extracted proteins were dialyzed against dialysis buffer (20 mM HEPES, pH 7.9, 20% glycerol, 100 mM KCl, 0.2 mM EDTA, 0.5 mM dithiothreitol) with 100 mM KCl and protein concentration was measured using bicinchoninic acid protein assay kit (Pierce, Rockford, IL, USA). Nuclear extracts were fractionated using phosphocellulose P11 resin (Whatman) pre-cycled in the dialysis buffer. Transcription complexes were fractionated at gradient KCl concentrations, 0.2–1.0 M, using batch technology. For western analysis, fractions were resolved in a 10% sodium dodecyl sulfate (SDS)-polyacrylamide gel (20 μg of each fraction) and transferred to Polyvinylidene Fluoride (PVDF) membrane (Hybond-P, Amersham Pharmacia Biotech, Piscataway, NJ, USA) using semi-dry blotter (TE 77 PWR, GE Healthcare, Freiburg, Germany). For further analysis, a non-block technique was used. For immunodetection, mouse anti-V5 monoclonal antibody (Invitrogen) or rabbit anti-BAF155 polyclonal antibody (Aviva Systems Biology, San Diego, CA, USA) was used at 1 : 10 000 dilution along with anti-mouse-horseradish peroxidase or anti-rabbit- horseradish peroxidase secondary antibody at 1 : 100 000 dilution. The signals were visualized using enhanced chemiluminescence reagents (Enhanced Chemiluminescence Advance Western Blotting Detection Kit, Amersham).

Co-immunoprecipitation

Immunoprecipitation assay was performed as described previously (Baker *et al.* 2003). Briefly, Neuro2A cells were lysed in high-salt buffer [20 mM Tris-HCl pH 7.5, 400 mM NaCl, 1 mM EDTA, 0.5% NP-40 (Sigma)] containing protease inhibitors (Roche). Lysates were sonicated shortly and centrifuged at 4°C for 15 min. Protein concentration of the supernatants (whole-cell extracts) were determined using bicinchoninic acid assay (Pierce). 1000 μg of protein was diluted with an equal volume of dilution buffer (lysis buffer without

NaCl) and the total volume was brought to 500 μL using immunoprecipitation (IP) buffer (lysis buffer with 200 mM NaCl). Immunoprecipitation was carried out by incubating the extracts with 20 μL of 1 : 1 anti-V5 agarose slurry (Sigma) for 12 h at 4°C. Beads were washed six times with IP buffer and bead-bound complexes were eluted by boiling in SDS loading buffer for 10 min. The immunoprecipitates and 200 μg of input were separated on a 7% SDS-polyacrylamide gel electrophoresis and transferred to polyvinylidene difluoride membrane (Immobilon-P, Millipore Corporation, Bedford, MA, USA) by wet-blotting procedure. Rabbit polyclonal antibodies to human Brm (1 : 2500, Abcam), V5 (1 : 3000, Sigma) and BAF155, BAF170, Brg1 (1 : 1000) were used for western blotting analysis.

Results

Mammalian BAF57 genes encode protein isoforms with different N-termini

Mining of Expressed Sequence Tags (EST) databases revealed that the N-terminal region of human BAF57 (hBAF57) protein is affected by mRNA alternative splicing. These splicing events yield in putative alternative protein isoforms BAF57_i3, encoded by transcripts with exon IV deleted (GenBank ESTs with ID: CB154008, BX441359, BX441815, and others), and BAF57_i5, encoded by transcripts with exons III and IV deleted (GenBank EST with ID: BY796605, DC353609, DC313771) (Fig. 1a). BAF57 isoform BAF57v (here BAF57_i2) was identified previously and contains a unique exon XI that introduces an in-frame STOP codon generating a shorter form (363 amino acids) of the protein (Battaglioli *et al.* 2002). Alternative splicing of exons III, IV and XI generates at least six different BAF57 isoforms (Fig. 1b). Isoforms BAF57_i1 and BAF57_i2 correspond to the annotated proteins with the full open reading frame (ORF) of 411 amino acids and the C-terminally truncated BAF57 of 363 amino acids (BAF57v) (Battaglioli *et al.* 2002). Isoforms BAF57_i3 and BAF57_i4 share N-terminus that lacks amino acids 17–51 and correspond to alternatively spliced transcripts with exon IV skipping. Isoforms BAF57_i5 and BAF57_i6 are generated because of alternative splicing of exons III-IV resulting in N-terminally shorter proteins where the first 70 amino acids including part of the HMG box is missing. In these isoforms translation is initiated from an alternative, downstream AUG.

Similar alternative splicing yielding in different N-termini of BAF57 isoforms is conserved in mouse and rat. In mouse, ESTs with IDs: CX237977, BY242131, BY097565 and BY253362, and others encode for protein isoform of BAF57_i3 displaying deletion of exon IV, whereas ESTs with IDs: BY242256, BU847403, BY263867, encode for protein isoform BAF57_i5 with deleted exons III and IV. In rat, ESTs with IDs: CO394818, CV106737, encode for protein isoform BAF57_i3 displaying deletion of exon IV. In addition, three more ESTs with deletions of exons II (mouse ID: CX226500; rat ID: CK364384), II and IV (mouse ID: BY244370) and II, III and IV (mouse ID: BQ442537, CA315143 and BY774202; rat ID: CB737352) are present in mouse and rat EST databases. These data indicate that alternative splicing generates protein isoforms of BAF57 with different N-termini that may have different, tissue specific function.

High expression of N-BAF57 mRNAs in developing and adult brain

In human, RT-PCR analysis of the expression of *BAF57* mRNAs non-neural tissues, cancer cells and adult brain regions (Fig. 2a and b) revealed expression of transcripts with deletion of exons III and IV (*BAF57_v5,_v6*) or exon IV (*BAF57_v3,_v4*). These transcripts of *hBAF57* encode N-terminally truncated protein isoforms (BAF57_i5,_i6 and BAF57_i3,_i4, respectively, Fig. 1). By using transcript-specific primers, analysis of the usage of exon XI in *hBAF57* alternative mRNAs showed that splicing of exons III and IV or exon IV occurs independent of exon XI, since alternative transcripts with and without exon XI were expressed at relatively same levels (data not shown).

RT-PCR analysis of *BAF57* mRNA expression in mouse and rat tissues revealed three murine-specific mRNA splice variants with deletion of exon II only (*BAF57_v7*), with deletion of exons II and IV (*BAF57_v8*) and with deletion of exons II, III and IV (*BAF57_v9*) (Fig. 1a and b). RT-PCR analysis employing transcript-specific primer sets and subsequent sequencing confirmed the expression of *BAF57* alternative transcripts in mouse and rat tissues (data not shown). *BAF57_v7* transcripts are predicted to produce a murine-specific protein lacking 19 amino acids at the N-terminus and initiating from alternative in-frame downstream AUG (BAF57_i7,_i8). *BAF57_v8* and *BAF57_v9* transcripts encode similar protein isoform, namely BAF57_i5 (Fig. 1b). Genomic sequences surrounding exons II, III and IV were compared allowing the measure of the homologies between these regions in the three species. It was confirmed that the sequences of exons II, III and IV are flanked by consensus splice donor/acceptor signals GT/AG in all three species.

RT-PCR analysis data demonstrate that BAF57 splice variants with deletion of exon IV or exons III and IV (*BAF57_v3,_v4* and *BAF57_v5,_v6,* respectively) are expressed exclusively in the brain as compared to non-neuronal tissues (Fig. 2a). However, regional differences in the expression of *BAF57_v3,_4* and *BAF57_v5,_v6* were observed in adult human brain regions. The results showed that human *BAF57_v3,_4* and *BAF57_v5,_v6* mRNAs were highly expressed in the olfactory bulb, frontal cerebral cortex, frontal, parietal and occipital lobes, infundibulum, mammillary body, cerebellum, parietal lobe, caudate nucleus and thalamus (Fig. 2b). However, very low expression of *BAF57_v3,_4* and *BAF57_v5,_v6* mRNAs was found in all regions of white matter (optic nerve, cerebellar white matter, fibre tracts of cerebellar pedunculi, corpus callosum,

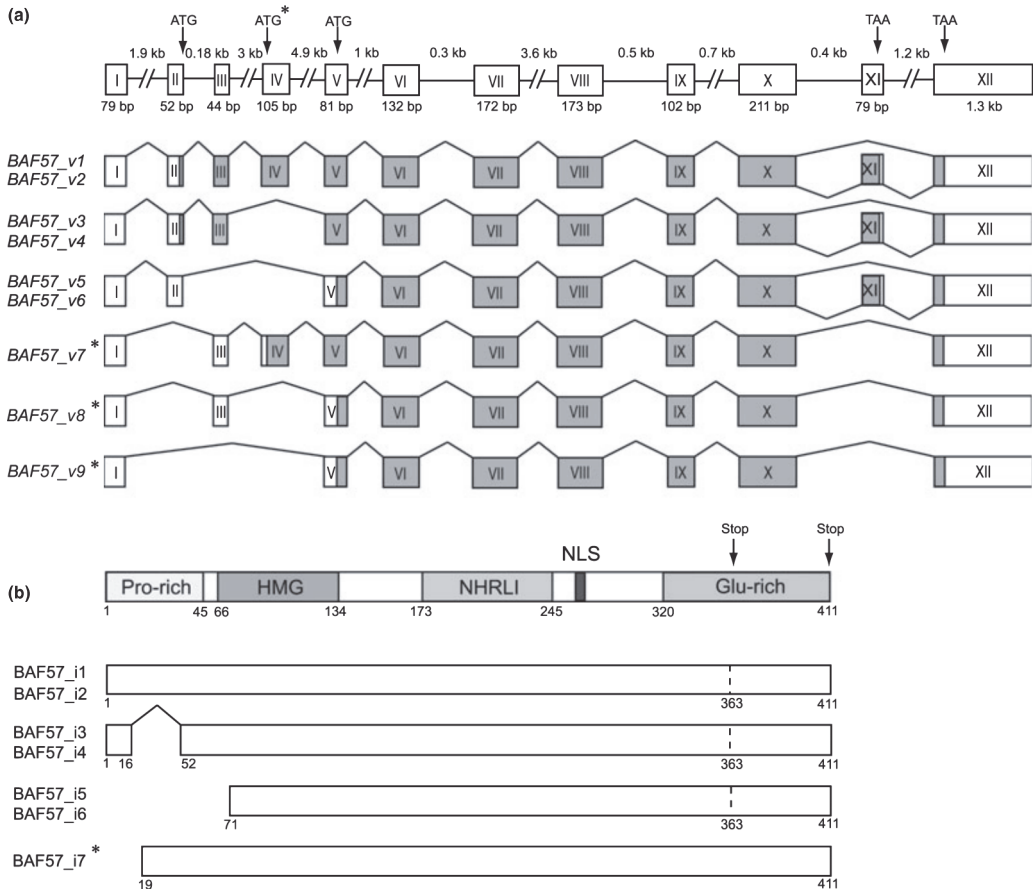

Fig. 1 *BAF57* genomic structure, alternative transcripts and protein isoforms. (a) Schematic structure of the human *BAF57* gene and alternative mRNAs of human and mouse/rat (marked with '*') *Baf57*. Exons are shown as boxes and introns as lines. Filled boxes represent the translated regions of the exons. The numbers below the exons and above the introns indicate their sizes, respectively. Positions of translation start-sites (ATG for human or ATG* for mouse and rat) and terminal codons (TAA) are shown. Two different transcripts with alternative 3′-ends correspond to splicing schemes characteristic to human *BAF57* only (e.g. *BAF57_v1* and *BAF57_v2* have exon XI deleted or included, respectively). (b) Schematic structure of the human and rodent BAF57 proteins. Alternative human and mouse/rat (marked with '*') protein isoforms are shown. Two different protein isoforms with alternative C-terminus correspond to alternative splicing schemes (e.g. BAF57_i1 and BAF57_i2 stop with amino acid 411 or 363, respectively). Pro-rich (proline-rich domain), HMG-box (high mobility group domain), NHRLI (conserved block of amino acids), NLS (nuclear localization signal) and Glu-rich (glutamic acid-rich domain) regions are indicated. Domains are drawn to scale. Arrows indicate the alternative translation termination sites (at position 363 aa and at position 411 aa). Numbers below the protein sequences indicate amino acids.

subcortical white matter, Fig. 2b). Similar neural-specific expression of *BAF57_v3* and *BAF57_v5* splice variants was observed in mouse (Fig. 2c) and rat (Fig. 2d).

In mouse, RT-PCR analysis of different developmental stages of the brain [embryonic (E) day 13 to postnatal (P) day 60] revealed that the expression of *BAF57_v3* and *BAF57_v5* mRNAs was high at the stages of development when neural stem/progenitor cells in the developing brain have ceased to proliferate and started to differentiate into neurons (E13). The expression of *BAF57_v3* increased up to P1, decreased at postnatal days 9–21 and was high in the adult (P60) (Fig. 2c). In rat, the expression of *BAF57_v3* and *BAF57_v5* mRNAs was also high during brain development, both in embryonic stages and in adult

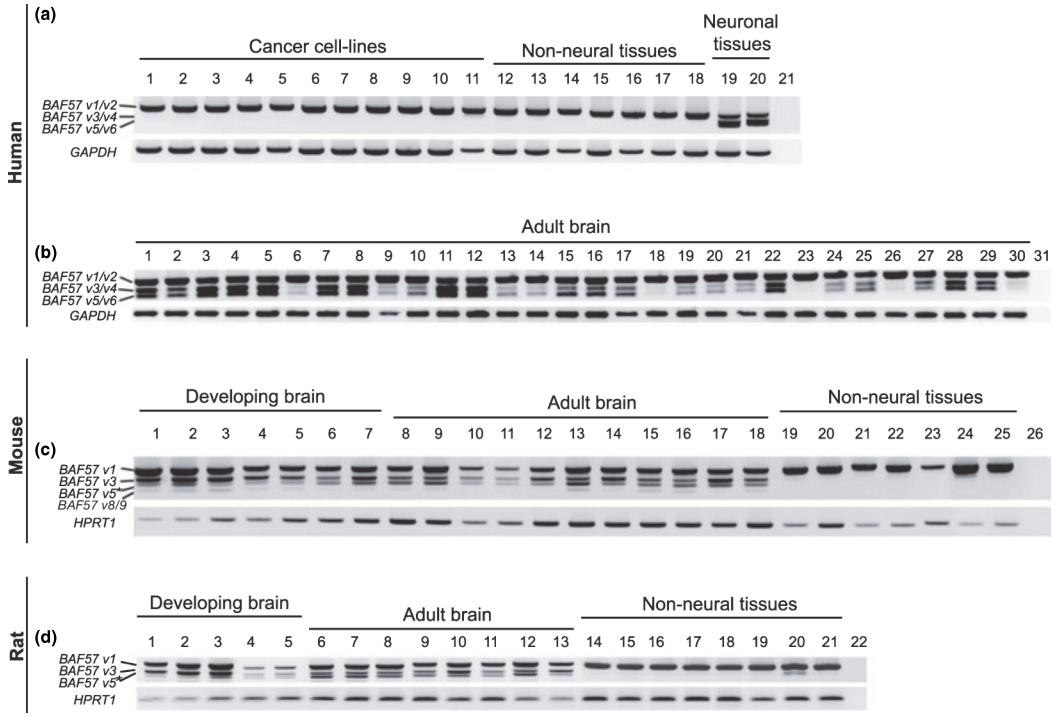

Fig. 2 Semi-quantitative RT-PCR analysis of human, mouse and rat *BAF57* mRNA expression. Tissue-specific expression of *BAF57* mRNA transcripts (*BAF57_v1,_v2*, *BAF57_v3,_v4* and *BAF57_v5,_v6*) in human (Panels a, b), mouse (Panel c) and rat (Panel d) are shown. Expression of *BAF57* mRNA rare transcripts with exon II deleted (*BAF57_v7,_v8*, and *BAF57_9*) are shown in mouse (Panel c) and rat (Panel d). Analysis of *GAPDH* or *HPRT1* mRNA expression levels was performed to normalize the differences in the amount of mRNA across samples. The data are representative of at least three independent experiments. (a) Human cancer cell lines and tissues: 1. Colo-205 (colon cancer cell line), 2. LoVo (colon cancer cell line), 3. Caco-2 (colon cancer cell line), 4. MDA-MB231 (breast cancer cell line), 5. MCF-7 (breast cancer cell line), 6. Hs578T (breast cancer cell line), 7. SkMel28 (melanoma cell line), 8. WM266-4 (melanoma cell line), 9. 1321N1 (astrocytoma cell line), 10. U373 (glioblastoma-astrocytoma cell line), 11. Kelly (neuroblastoma cell line), 12. fetal skin, 13. fetal colon, 14. colon, 15. heart, 16. kidney, 17. liver, 18. testis, 19. fetal brain, 20. cerebellum, 21. –RT. (b) Human brain: 1. Olfactory bulb, 2. olfactory tract, 3. frontal cerebral cortex, 4. frontal lobe, 5. occipital lobe, 6. optic nerve, 7. infundibulum, 8. mammillary body, 9. ventral medulla, 10. dorsal medulla, 11. left cerebellum, 12. right cerebellum, 13. cerebellar nuclei, 14. cerebellar white matter, 15. ventral pons, 16. dorsal pons, 17. tectum, 18. pedunculi, 19. substantia nigra, 20. nucleus ruber, 21. dorsal caudal pedunculi, 22. parietal lobe, 23. corpus callosum, 24. hippocampus, 25. amygdala, 26. globus pallidus, 27. putamen, 28. caudate nucleus, 29. thalamus, 30. epiphysis, 31. –RT. (c) Mouse panel: 1. E13 brain, 2. E15 brain, 3. P1 brain, 4. P9 brain, 5. P14 brain, 6. P21 brain, 7. P60 brain, 8. colliculi, 9. cerebellum, 10. frontal cortex, 11. caudal cortex, 12. hippocampus, 13. medulla, 14. olfactory bulb, 15. midbrain, 16. pons, 17. striatum, 18. thalamus, 19. heart, 20. kidney, 21. liver, 22. lung, 23. muscle, 24. testis, 25. thymus, 26. –RT. (d) Rat panel: 1. E13 brain, 2. E16 brain, 3. E21,5 brain, 4. P7 brain, 5. P21 brain, 6. frontal cortex, 7. colliculi, 8. hypothalamus, 9. medulla, 10. pons, 11. striatum, 12. thalamus, 13. olfactory bulb, 14. heart, 15. kidney, 16. liver, 17. lung, 18. muscle, 19. spleen, 20. testis, 21. thymus, 22. –RT. Abbreviations: *BAF57 v5**, mouse and rat transcripts containing deleted exons III and IV (*BAF57_v5*), or deleted exon II (*BAF57_v7*); *BAF57 v8/9*, mouse transcripts containing deleted exons II and IV (*BAF57_v8*), or deleted exons II, III and IV (*BAF57_v9*); E, embryonic day; P, postnatal day; –RT, negative control containing reverse transcription and PCR reagents with no template.

(Fig. 2d). We were able to recover murine-specific *BAF57* mRNA splice variants *BAF57_v7*, *BAF57_v8* and *BAF57_v9* using exon-spanning primers (Fig. 2, C *lanes 1–9, 13–15*, and *18*) as well as transcript-specific primers (data not shown), although these were barely detected as compared to *BAF57_v3* and *BAF57_v5* mRNAs. Splice variants corresponding to *BAF57_v7*, *BAF57_v8* and *BAF57_v9* mRNAs were also detected in rat using exon-spanning primers (Fig. 2d) The failure to recover rare transcripts with exon II deleted in human neural tissues may

reflect difficulty in amplifying such scarce transcripts rather than their absence.

N-BAF57 transcripts with deletion of exon IV or exons III and IV are expressed exclusively in neurons

In order to assess whether splicing involving exons III and IV of BAF57 are characteristic to all cells of the nervous system or to a certain lineage, the expression of *BAF57* mRNA was analyzed in cultures of embryonic neurons and glia (Fig. 3a). As shown in Fig. 3, cultured neurons expressed *BAF57* splice variants including transcripts with deletion of exon IV or exons III and IV. In contrast, cultured glial cells displayed no expression of *BAF57* splice variants with deletion of exon IV or exons III and IV mRNAs (Fig. 3a).

Furthermore, RT-PCR analysis of *BAF57_v3,_v4* and *BAF57_v5,_v6* splice variants in adult human neural stem cells that were grown as neurospheres or subjected to differentiation to neurons (15% of the differentiated culture was bIIItubulin- and microtubule-associated protein 2-positive, data not shown), revealed that transcripts encoding N-BAF57s are detected in neurospheres as well as upon dibutyryl cAMP or dibutyryl cAMP + all-trans retinoic acid (RA) induced neuronal differentiation (Fig. 3b). In all experiments, RT-PCR analysis of *GAPDH* or *HPRT1* was used to normalize the differences in the amount of mRNA across samples.

From these data, we concluded that splicing mechanisms generating alternative transcripts of *BAF57* with exons III and IV or IV deleted are evolutionarily conserved and may accompany neuronal differentiation.

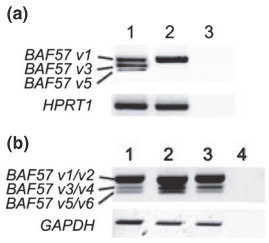

Fig. 3 Semi-quantitative RT-PCR analysis of *BAF57* mRNA expression in rat cultured primary neurons, astroglial cells and human stem cells. (a) 1. Rat cultured primary neurons, 2. Rat cultured primary astrocytes, 3. –RT. (b) 1. Human neural stem cells, 2. human neural stem cells subjected to neuronal differentiation with 1 mM dBcAMP, 3. human neural stem cells subjected to neuronal differentiation with 1 mM dBcAMP and 1 mM all-*trans*-retinoic acid, 4. –RT. Analysis of *GAPDH* or *HPRT1* mRNA expression levels was performed to normalize the differences in the amount of mRNA across samples. The data are representative of at least three independent experiments. Abbreviations: dBcAMP, dibutyryl cAMP; RT, negative control containing reverse transcription and PCR reagents with no template.

N-BAF57 isoforms interact with complexes containing Brm or Brg1 and can function as regulators of expression of NRSE-containing gene

To examine the function of N-BAF57s *in vivo,* a series of vectors expressing V5-tagged BAF57 isoforms was constructed, including *BAF57-V5, BAF57Δ4-V5* and *BAF57Δ3-4-V5* in which the V5 tag was located in the N-terminus of the respective protein. First, we analyzed the cellular localization of V5 tagged-BAF57 isoforms by transient transfection assays in human melanoma WM266-4 cells. Immunocytochemical analysis showed that all BAF57 V5-tagged isoforms were distributed exclusively in the nucleus (Fig. 4). These data supported the suggested nuclear localization signal of BAF57 to span amino acids 255–258 (Chen and Archer 2005). Both, V5-tagged BAF57 (BAF57_i1) and N-BAF57-protein isoforms (BAF57_i3 and BAF57_i5) revealed punctuate distribution throughout the nucleus. Uneven distribution of N-BAF57s in the nucleus was confirmed by line scans through the nuclei (Fig. 4). Furthermore, extensive overlay in the nuclear localization patterns of V5-tagged N-BAF57 proteins and Ser5 phosphorylated form of RNA Pol II-CTD was observed. The phosphorylation status of the CTD detects active RNA Pol II complexes and determines the transition from transcriptional initiation to elongation (Dahmus 1996). Co-distribution of N-BAF57s with Ser5 phosphorylated RNA Pol II was also confirmed by line scans through the nuclei (Fig. 4). These data suggested that N-BAF57 isoforms are nuclear proteins and can be recruited to active transcription complexes.

Next, the ability of N-BAF57 isoforms to be assembled to SWI/SNF complexes was analyzed using a biochemical fractionation approach. Recent studies have demonstrated that the physiological levels of BAF57 are kept constant such that BAF155 and BAF170 serve as molecular chaperones for the correct assembly of BAF57 to the SWI/SNF remodelling complexes (Chen and Archer 2005). Therefore the co-fractionation of N-BAF57 with BAF155 was specifically analyzed. Mouse neuroblastoma Neuro2A cells were transfected with constructs of *BAF57-V5, BAF57Δ4-V5* or *BAF57Δ3-4-V5*, cells were harvested 48 h post-transfection, nuclear extracts were prepared and subjected to chromatographic fractionation at gradient KCl concentrations (see Materials and methods). The elution profiles of *BAF57Δ4-V5* and *BAF57Δ3-4-V5* were similar to that of *BAF57-V5*. Both N-BAF57 isoforms eluted along with endogenous BAF57 and BAF155 proteins (Fig. 5a). These data evidenced that if BAF155 functions as a scaffold for the assembly of the SWI/SNF remodelling complex through increasing the protein stability of BAF57 (Baker *et al.* 2003), then based on our biochemical fractionation data N-BAF57 proteins may similarly be stabilized by BAF155 and assembled into a functional human SWI/SNF complex for transcription.

To address whether N-BAF57 isoforms are assembled into chromatin remodelling SWI/SNF complexes, we used

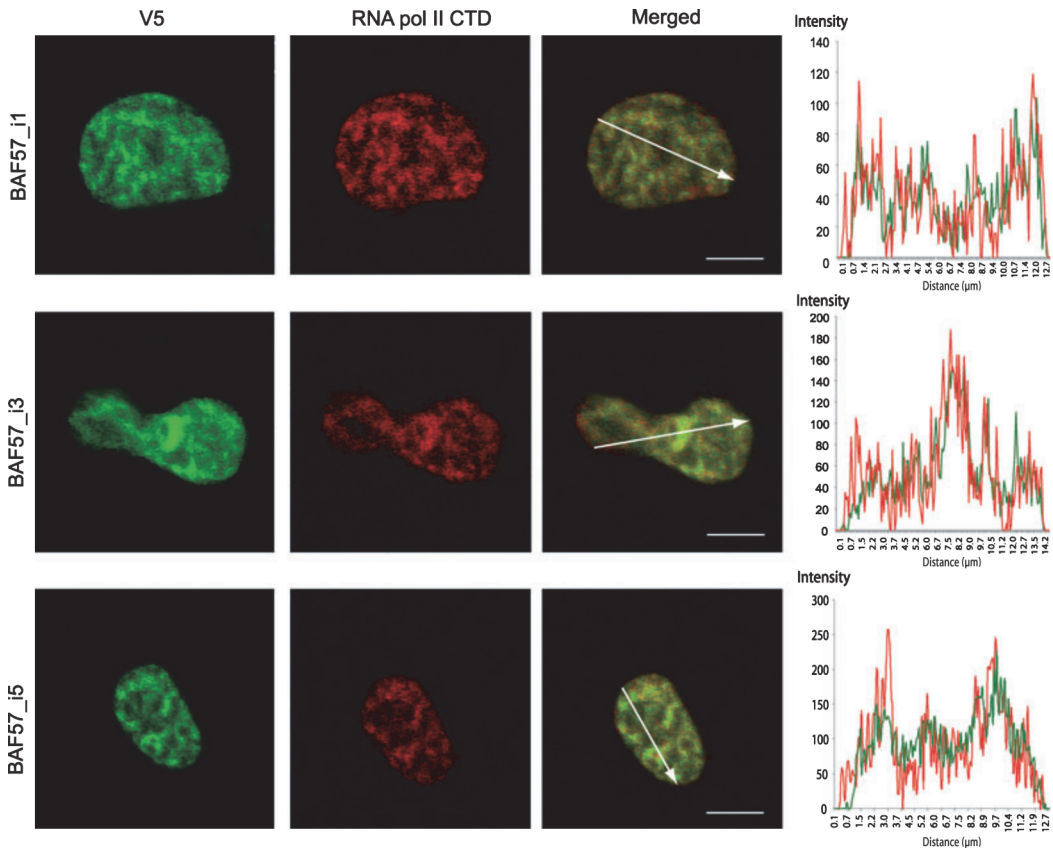

Fig. 4 BAF57 and N-BAF57 isoforms are nuclear proteins and co-localize with Ser5 phosphorylated RNA Pol II. WM266-4 cells were transfected with *BAF57-V5* (BAF57_i1) or *N-BAF57-V5* (BAF57_i3 and BAF57_i5) expression constructs. Twenty-four hours post-transfection, the cells were fixed and labeled with anti-V5 (*green signal*) or anti-RNA polymerase II Ser5 phosphorylated CTD (*red signal*) antibodies and analyzed using confocal microscopy. Merged images of the different images (*yellow signal*) indicate areas of co-localization. The positions and directions of the line scans are indicated by *arrows*. The graphs to the right show the overlay of the V5-tagged protein signal in *green* and the RNA Pol II Ser5 phosphorylated CTD in *red*. The transfected WM266-4 cells incubated with the secondary antibodies alone showed no cross-reactivity (data not shown). The data shown are representative of two independent experiments. Scale bars, 5 μm.

immunoprecipitation analysis. Provided that Neuro2A cells express endogenous *N-Baf57* mRNAs at low levels (data not shown), this cell line was considered as a suitable model system for N-BAF57 complex analysis. Forty-eight hours post-transfection with *BAF57-V5, BAF57Δ4-V5* and *BAF57Δ3-4-V5* constructs, V5-tagged BAF57 protein isoform-containing complexes were immunoprecipitated using anti-V5 antibodies and precipitated complexes were analyzed by western blot method.

As shown in Fig. 5(b), Brm, Brg1, BAF170 and BAF155 proteins were detected in the immunoprecipitates of V5-tagged BAF57 and N-BAF57 proteins. Provided that N-BAF57 isoforms co-immunoprecipitated with Brm, Brg1, BAF170 and BAF155 proteins (Fig. 5b), it was concluded that the first 70 amino acids of BAF57 protein are not essential for the interaction of N-BAF57s with SWI/SNF complex subunits. Recent data reported that CoREST interacts with the C-terminus of BAF57 (Battaglioli *et al.* 2002; Ooi *et al.* 2006). These data evidenced that N-BAF57 isoforms have retained the ability to bind to and be incorporated in SWI/SNF multiprotein complexes and possibly act as regulators of chromatin remodelling.

To understand the molecular mechanisms of N-BAF57 function in SWI/SNF/CoREST complexes in more detail, we examined whether N-BAF57 protein activity affects the REST/NRSF-mediated suppression of neuronal genes by

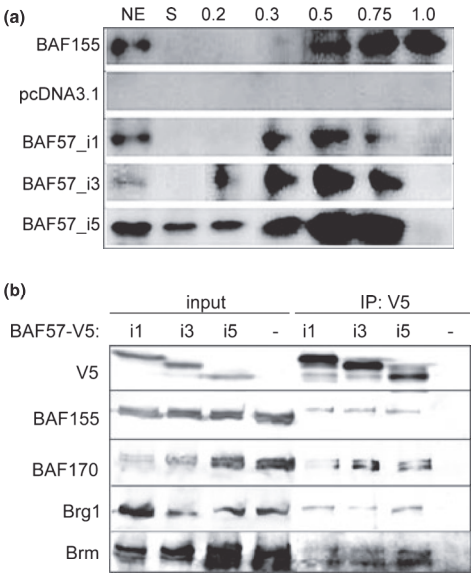

Fig. 5 N-BAF57 proteins are stabilized by BAF155 and assembled into Brg1 and Brm-containing SWI/SNF complexes. (a) N-BAF57 isoforms have similar biochemical fractionation profiles as endogenous BAF57 and BAF155 proteins. Western blot analysis of different nuclear fractions of mouse neuroblastoma Neuro2A cells transfected with V5-tagged BAF57 expression constructs *BAF57-V5* (BAF57_i1), *BAF57Δ4-V5* (BAF57_i3), *BAF57Δ3-4-V5* (BAF57_i5) or plasmid pcDNA3.1/V5-His (pcDNA 3.1). Detection with of endogenous BAF155 was performed using anti-BAF155 antibodies and detection of V5-tagged BAF57 isoforms in transfected cells was performed using anti-V5 antibodies (as shown on the left). Each fraction is marked on the top of the lanes as 0.2–1.0 indicating the molarity of KCl in the process of salt gradient fractionation. Cell extract transfected with pcDNA3.1/V5-His was used as control and analyzed using anti-V5 antibodies. (b) The immunoprecipitated complexes of V5-tagged BAF57 protein isoforms contain chromatin remodulators Brg1, Brm, BAF155, and BAF170. Western blot analysis of co-immunoprecipitated complexes of V5-tagged BAF57 isoforms. Neuro2A cells were transfected with *BAF57-V5* (i1), *BAF57Δ4-V5* (i3), *BAF57Δ3-4-V5* (i5) or pcDNA3.1/V5-His (-) constructs. Whole-cell extracts were prepared from transiently transfected cells and immunoprecipitated with anti-V5 antibodies. The precipitated material was analyzed by western blotting with antibodies as indicated on the *left*, specifically V5, BAF155, BAF170, Brg1 and hBrm. About 5% of the total input lysates used for immunoprecipitation were directly analyzed by western blotting using respective antibodies. Abbreviations: NE, nuclear extract; S, flow-through fraction.

using transient transfection assays. At 48 h post-transfection of constructs *BAF57-V5, BAF57Δ4-V5* or *BAF57Δ3-4-V5* into human melanoma WM266-4, the cells were harvested for total RNA extraction and the level of expression of several NRSE-containing neuronal genes, including *CHRM4, L1CAM, SYP, COX1, BDNF, NaChII* and *SCG10*, were analyzed by real-time RT-PCR method. WM266-4 cells expressed endogenous *REST* and *CoREST* and did not express *N-BAF57* mRNAs (Fig. 6a). We found that *CHRM4* mRNA expression was specifically induced (a 3.5 to 5-fold increase of expression) in cells transfected with *BAF57Δ4-V5* and *BAF57Δ3-4-V5*, whereas a 2-fold increase in the expression of *CHRM4* mRNA was observed in cells transfected with *BAF57-V5* (Fig. 6b). In contrast, exogenous expression of *BAF57Δ4-V5* resulted in a slightly (28.5 ± 5.0%) decreased *L1CAM* mRNA levels (Fig. 6b). No significant change was observed in the expression of *COX1* and *SYP* mRNAs by introducing *BAF57Δ4-V5* or *BAF57Δ3-4-V5* to WM266-4 cells (Fig. 6b). As the expression of *BDNF, NaChII* and *SCG10* mRNAs was not detected pre- or post-treatment (data not shown), it was concluded that exogenous expression of BAF57 protein isoforms was not sufficient to execute transcription from these gene loci in WM266-4 cells. These results suggest that N-BAF57 isoforms could affect the expression of NRSE/RE1-containing genes in non-neural cells, possibly in a promoter-specific manner. Our data are in good agreement with a recently published findings that the BAF57-containing SWI/SNF complex affects the expression of NRSE/RE1-containing genes in non-neural cells (Watanabe *et al.* 2006).

Discussion

The current understanding is that SWI/SNF complexes are critical for maintenance of normal tissue homeostasis and that disruption of these complexes contributes to severe developmental defects (Lessard *et al.* 2007), neoplasia and transformation to a malignant phenotype (Bultman *et al.* 2000; Roberts *et al.* 2000, 2002; Guidi *et al.* 2001). Earlier studies have suggested that BAF57 as a core-subunit of SWI/SNF complexes plays a critical role in the transcriptional repression of neuron- (Battaglioli *et al.* 2002) and T lymphocyte-specific genes (Chi *et al.* 2002), as well as tumour suppressor genes such as *BRCA1* (Baker *et al.* 2003). Here we demonstrate that N-BAF57 isoforms are predominantly expressed in the nervous system suggesting a role of N-BAF57 isoforms in the regulation of neurogenesis.

In this study, we have identified alternatively spliced mRNAs encoding N-terminally truncated BAF57 protein isoforms that are exclusively expressed in neurons. Recent data of Battaglioli *et al.* (2002) identified a human-specific splice variant, namely *BAF57v* that encoded a protein isoform truncated in its C-terminus with ubiquitous pattern of expression (Battaglioli *et al.* 2002). Our data demonstrate that N-BAF57 isoforms generated by the alternative splicing of exons II to IV of the *BAF57* gene yield in N-terminally truncated proteins with either long or short C-terminus. Since exons II, III and IV encode a Proline-rich domain with unknown function, the functional consequence of the

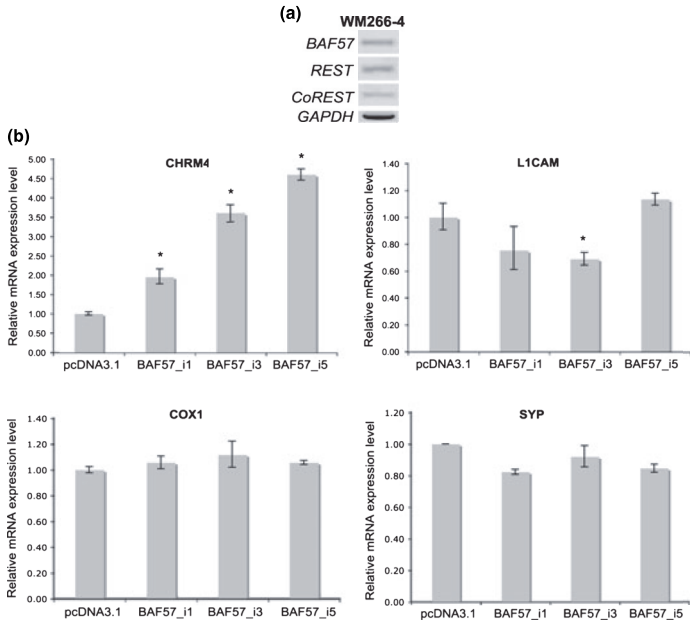

Fig. 6 Over-expression of BAF57 and N-BAF57 affects the expression of NRSE/RE1-containing genes. (a) Expression analysis of *BAF57*, *REST* and *CoREST* mRNA in human melanoma WM266-4 cells. The RT-PCR primers of *hBAF57* recognize both, *BAF57* and *N-BAF57* mRNA isoforms. (b) Over-expression of BAF57 and N-BAF57 affects the expression of NRSE/RE1-containing target genes in WM266-4 cells. WM266-4 cells were transiently transfected with BAF57 expression vectors *BAF57-V5* (BAF57_i1), *BAF57Δ4-V5* (BAF57_i3), *BAF57Δ3-4-V5* (BAF57_i5) or plasmid pcDNA3.1. Twenty-four hours post-transfection, the levels of endogenous *CHRM4*, *L1CAM*, *COX1* and *SYP* mRNAs were measured using qPCR analysis. The *bars* indicate the relative mRNA expression of analyzed genes. The error bars indicate the variation between three independent real-time PCR experiments. Expression levels of all genes were normalized to *GAPDH*. The data shown are the mean ± SD of three independent experiments, each determined with qRT-PCR in triplicates, $^*p < 0.05$ vs. plasmid pcDNA3.1 transfected cells.

alternative splicing yielding in N-terminally truncated BAF57 protein isoforms is not clear. Previous studies revealed that the HMG domain juxtaposing to the Proline-rich region is not essential for the DNA binding activity and/or for the ATP-dependent nucleosome disruption activity of BAF57 (Giese *et al.* 1992; Paull *et al.* 1993). Despite the fact that the Proline-rich region is thought to be important for the assembly of the SWI/SNF complexes (Wang *et al.* 1998) the precise role of this region had not previously been identified. The biochemical fractionation data and results from co-immunoprecipitation analysis presented here demonstrate that N-terminally truncated isoforms of BAF57 lacking the Proline-rich region are readily assembled to the Brg1-Brm containing SWI/SNF complexes. These results suggest that the Proline-rich region is not essential for the assembly of the SWI/SNF complexes. The function of SWI/SNF complexes is associated with the regulation of transcriptional activation and repression (Wang *et al.* 1998). The fact that recruitment of N-BAF57 isoforms to Brg1-Brm containing complexes was observed, further support the concept that N-BAF57s are involved in regulating the transcription.

Our data provide evidence that the heterogeneity of SWI/SNF complexes is related to the expression of different isoforms of BAF57 in a cell-specific manner. Several studies have shown that the heterogeneity of SWI/SNF complexes stems from a subset of BAF proteins that are differentially expressed in different cell types and are encoded by families of highly homologous genes (Wang *et al.* 1996; Olave *et al.* 2002; Lessard *et al.* 2007). Distinct SWI/SNF complexes that have been characterized during the development of the mouse nervous system contain different subunits belonging to the same gene family (Aigner *et al.* 2007; Lessard *et al.* 2007). The results presented here show that the expression of *BAF57* transcripts with the alternative usage of exons II, II-IV, II and IV, is differentially regulated during early and late neurogenesis. In particular, the expression of *BAF57_v3,_v4,_v5,_v6* mRNAs is high in the adult brain and exclusive to neurons (Figs 2 and 3). These findings have

raised the intriguing possibility that SWI/SNF complexes include mutually exclusive isoforms of BAF57 protein thus providing the basis for the combinatorial variety of SWI/SNF complexes with distinct functions during development and in the adult nervous system.

Surprisingly, N-BAF57 isoforms were not detected in mouse P0 brain in the biochemical approach by Lessard and colleagues (Lessard *et al.* 2007). Given that no N-terminus discriminating antibodies are available, it is highly possible that the proteomic analysis of SWI/SNF subunit composition from mouse P0 brain nuclear extracts overlooked the N-BAF57 isoforms as true part of nBAF complexes. Interestingly, *Baf45a/PHF10* that was identified as a subunit of npBAF complex (Lessard *et al.* 2007) has also at least two isoforms, namely isoforms PHF10a and PHF10b. These two isoforms differ by only two amino acids encoded by the splice junction sequences between exons 3 and 4. The precise role of PHF10a and PHF10b proteins is not known and it remains to be identified which of these isoforms is a true subunit of npBAF complex. The results presented here show that *BAF57* is expressed in brain development (Fig. 2), in neurons (Fig. 3a) and upon differentiation of neural stem cells (Fig. 3b) and co-purifies with other SWI/SNF complex subunits (Fig. 5). These data predict that N-BAF57s are assembled to nBAF complexes.

According to the present understanding, the control of higher order chromatin structure is a well-established mechanism for the regulation of a number of gene loci which expression needs tight regulation across tissues (Felsenfeld 1996; Kadonaga 1998). Given that BAF57 isoforms show nuclear localization and were identified as subunits of SWI/SNF complexes along with Brg1, hBrm, BAF155, and BAF170, the role of N-BA57s may therefore be to modify local chromatin environment as part of transcription regulatory complexes. Earlier findings have suggested that the interaction of BAF57 and CoREST recruits the SWI/SNF complexes to promoters that contain binding sites for the transcriptional repressor NRSF/REST (Battaglioli *et al.* 2002; Ooi *et al.* 2006). Potential co-expression of BAF57 and N-BAF57 isoforms may apparently affect the expression of NRSE-containing neuronal genes. Based on our data we suggest that N-BAF57 isoforms affect the expression of certain NRSE-containing neuronal genes non-neural cells. This is in a good agreement with the findings of Ooi *et al.* showing that inhibition of BRG1 nucleosomal remodelling activity in non-neuronal cells with the BRG1 dominant negative (BRG1 DN) mutant results in the derepression of certain NRSF/REST target genes (Battaglioli *et al.* 2002; Ooi *et al.* 2006). Our data predict that the role of N-BAF57 isoforms is to support distinct cellular (neuronal) phenotypes by regulating the expression of certain NRSE-containing genes, e.g. in a promoter-specific manner.

Taken together, N-BAF57 isoforms can influence the activity of BAF complexes in the nervous system through at least three modes of action: (i) the expression of alternatively spliced variants encoding N-BAF57 isoforms in certain cell populations, (ii) affect BAF57_i1 and BAF57_i2 protein turnover by formation of stabilized N-BAF57/BAF155 protein dimers, and (iii) oppose BAF57_i1 and BAF57_i2 activity in the regulation of certain chromatin loci.

Acknowledgements

This work was sponsored in part by OncoTx, Inc (USA). We thank Weidong Wang (National Institutes of Health) for the antibodies against BAF155, BAF170 and Brg1 proteins. We also thank Enn Jõeste (The North Estonian Regional Hospital) for enthusiastic collaboration and Artyom Dugan, Anzelika Sarapuu and Epp Väli for excellent technical assistance and gratefully acknowledge Heiti Paves (Tallinn University of Technology) for support with confocal microscopy studies. TT, PP and MS were supported by a grant from the Estonian Ministry of Education and Research (Grant 0140143).

References

Aigner S., Denli A. M. and Gage F. H. (2007) A novel model for an older remodeler: the BAF swap in neurogenesis. *Neuron* **55**, 171–173.

Baker K. M., Wei G., Schaffner A. E. and Ostrowski M. C. (2003) Ets-2 and components of mammalian SWI/SNF form a repressor complex that negatively regulates the BRCA1 promoter. *J. Biol. Chem.* **278**, 17876–17884.

Battaglioli E., Andrés M. E., Rose D. W., Chenoweth J. G., Rosenfeld M. G., Anderson M. E. and Mandel G. (2002) REST repression of neuronal genes requires components of the hSWI SNF complex. *JBC* **277**, 41038–41045.

Belandia B., Orford R. L., Hurst H. C. and Parker M. G. (2002) Targeting of SWI/SNF chromatin remodelling complexes to estrogen-responsive genes. *EMBO J.* **21**, 4094–4103.

Bultman S., Gebuhr T., Yee D. *et al.* (2000) Brg1 null mutation in the mouse reveals functional differences among mammalian SWI/SNF complexes. *Mol. Cell* **6**, 1287–1295.

Chen J. and Archer T. K. (2005) Regulating SWI/SNF subunit levels via protein-protein interactions and proteasomal degradation: BAF155 and BAF170 limit expression of BAF57. *Mol. Cell. Biol.* **25**, 9016–9027.

Chen Z.-F., Paquette A. J. and Anderson D. J. (1998) NRSF/REST is required in vivo for repression of multiple neuronal target genes during embryogenesis. *Nat. Genet.* **20**, 136–142.

Chi T. H., Wan M., Zhao K., Taniuchi I., Chen L., Littman D. R. and Crabtree G. R. (2002) Reciprocal regulation of CD4/CD8 expression by SWI/SNF-like BAF complexes. *Nature* **418**, 195–199.

Dahmus M. E. (1996) Reversible phosphorylation of the C-terminal domain of RNA polymerase II. *J. Biol. Chem.* **271**, 19009–19012.

De la Serna I. L., Ohkawa Y. and Imbalzano A. N. (2006) Chromatin remodelling in mammalian differentiation: lessons from ATP-dependent remodellers. *Nat. Rev. Gen.* **7**, 461–473.

Felsenfeld G. (1996) Chromatin unfolds. *Cell* **86**, 13–19.

García-Pedrero J. M., Kiskinis E., Parker M. G. and Belandia B. (2006) The SWI/SNF chromatin remodeling subunit BAF57 is a critical regulator of estrogen receptor function in breast cancer cells. *J. Biol. Chem.* **281**, 22656–22664.

Giese K., Cox J. and Grosschedl R. (1992) The HMG domain of lymphoid enhancer factor 1 bends DNA and facilitates assembly of functional nucleoprotein structures. *Cell* **69**, 185–195.

Guidi C. J., Sands A. T., Zambrowicz B. P., Turner T. K., Demers D. A., Webster W., Smith T. W., Imbalzano A. N. and Jones S. N. (2001) Disruption of Ini1 leads to peri-implantation lethality and tumorigenesis in mice. *Mol. Cell. Biol.* **21**, 3598–3603.

Harikrishnan K. N., Chow M. Z., Baker E. K. *et al.* (2005) Brahma links the SWI/SNF chromatin-remodeling complex with MeCP2-dependent transcriptional silencing. *Nat. Genet.* **37**, 254–264.

Hsiao P.-W., Fryer C. J., Trotter K. W., Wang W. and Archer T. K. (2003) BAF60a mediates critical interactions between nuclear receptors and the BRG1 chromatin-remodeling complex for transactivation. *Mol. Cell. Biol.* **23**, 6210–6220.

Huang Y., Myers S. J. and Dingledine R. (1999) Transcriptional repression by REST: recruitment of Sin3A and histone deacetylase to neuronal genes. *Nat. Neurosci.* **2**, 867–872.

Ito T., Yamauchi M., Nishina M., Yamamichi N., Mizutani T., Ui M., Murakami M. and Iba H. (2001) Identification of SWI·SNF complex subunit BAF60a as a determinant of the transactivation potential of Fos/Jun dimers. *J. Biol. Chem.* **276**, 2852–2857.

Kadonaga J. T. (1998) Eukaryotic transcription: an interlaced network of transcription factors and chromatin-modifying machines. *Cell* **92**, 307–313.

Kingston R. E. and Narlikar G. J. (1999) ATP-dependent remodeling and acetylation as regulators of chromatin fluidity. *Genes Dev.* **13**, 2339–2352.

Lessard J., Wu J. I., Ranish J. A., Wan M., Winslow M. M., Staahl B. T., Wu H., Aebersold R., Graef I. A. and Crabtree G. R. (2007) An essential switch in subunit composition of a chromatin remodeling complex during neural development. *Neuron* **55**, 201–215.

Link K. A., Burd C. J., Williams E., Marshall T., Rosson G., Henry E., Weissman B. and Knudsen K. E. (2005) BAF57 governs androgen receptor action and androgen-dependent proliferation through SWI/SNF. *Mol. Cell. Biol.* **25**, 2200–2215.

Nagl N. G., Wang X., Patsialou A., Scoy M. V. and Moran E. (2007) Distinct mammalian SWI/SNF chromatin remodeling complexes with opposing roles in cell-cycle control. *EMBO J.* **26**, 752–763.

Nie Z., Xue Y., Yang D., Zhou S., Deroo B. J., Archer T. K. and Wang W. (2000) A specificity and targeting subunit of a human SWI/SNF family-related chromatin-remodeling complex. *Mol. Cell. Biol.* **20**, 8879–8888.

O'Malley E. K., Sieber B.-A., Morrison R. S., Black I. B. and Dreyfus C. F. (1994) Nigral type I astrocytes release a soluble factor that increases dopaminergic neuron survival through mechanisms distinct from basic fibroblast growth factor. *Brain Res.* **647**, 83–90.

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

Ooi L., Belyaev N. D., Miyake K., Wood I. C. and Buckley N. J. (2006) BRG1 chromatin remodeling activity is required for efficient chromatin binding by repressor element 1-silencing transcription factor (REST) and facilitates REST-mediated repression. *J. Biol. Chem.* **281**, 38974–38980.

Pal S., Yun R., Datta A., Lacomis L., Erdjument-Bromage H., Kumar J., Tempst P. and Sif S. (2003) mSin3A/histone deacetylase 2- and PRMT5-containing Brg1 complex is involved in transcriptional repression of the Myc target gene cad. *Mol. Cell. Biol.* **23**, 7475–7487.

Palm K., Belluardo N., Metsis M. and Timmusk T. (1998) Neuronal expression of zinc finger transcription factor REST/NRSF/XBR gene. *J. Neurosci.* **18**, 1280–1296.

Palm K., Salin-Nordström T., Levesque M. F. and Neuman T. (2000) Fetal and adult human CNS stem cells have similar molecular characteristics and developmental potential. *Brain Res. Mol. Brain Res.* **78**, 192–195.

Papoulas O., Daubresse G., Armstrong J. A., Jin J., Scott M. P. and Tamkun J. W. (2001) The HMG-domain protein BAP111 is important for the function of the BRM chromatin-remodeling complex in vivo. *PNAS* **98**, 5728–5733.

Paull T. T., Haykinson M. J. and Johnson R. C. (1993) The nonspecific DNA-binding and -bending proteins HMG1 and HMG2 promote the assembly of complex nucleoprotein structures. *Genes Dev.* **7**, 1521–1534.

Phelan M. L., Sif S., Narlikar G. J. and Kingston R. E. (1999) Reconstitution of a core chromatin remodeling complex from SWI/SNF subunits. *Mol. Cell* **3**, 247–253.

Roberts C. W. M., Galusha S. A., McMenamin M. E., Fletcher C. D. M. and Orkin S. H. (2000) Haploinsufficiency of Snf5 (integrase interactor 1) predisposes to malignant rhabdoid tumors in mice. *PNAS* **97**, 13796–13800.

Roberts C. W. M., Leroux M. M., Fleming M. D. and Orkin S. H. (2002) Highly penetrant, rapid tumorigenesis through conditional inversion of the tumor suppressor gene Snf5. *Cancer Cell* **2**, 415–425.

Schoenherr C. J., Paquette A. J. and Anderson D. J. (1996) Identification of potential target genes for the neuron-restrictive silencer factor. *PNAS* **93**, 9881–9886.

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

Sudarsanam P. and Winston F. (2000) The SWI/SNF family nucleosome-remodeling complexes and transcriptional control. *Trends Genet.* **16**, 345–351.

Underhill C., Qutob M. S., Yee S.-P. and Torchia J. (2000) A novel nuclear receptor corepressor complex, N-CoR, contains components of the mammalian SWI/SNF complex and the corepressor KAP-1. *J. Biol. Chem.* **275**, 40463–40470.

Vignali M., Hassan A. H., Neely K. E. and Workman J. L. (2000) ATP-dependent chromatin-remodeling complexes. *Mol. Cell. Biol.* **20**, 1899–1910.

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

Wang W., Chi T., Xue Y., Zhou S., Kuo A. and Crabtree G. R. (1998) Architectural DNA binding by a high-mobility-group/kinesin-like subunit in mammalian SWI/SNF-related complexes. *PNAS* **95**, 492–498.

Wang L., Baiocchi R. A., Pal S., Mosialos G., Caligiuri M. and Sif S. (2005) The BRG1- and hBRM-associated factor BAF57 induces apoptosis by stimulating expression of the cylindromatosis tumor suppressor gene. *Mol. Cell. Biol.* **25**, 7953–7965.

Watanabe H., Mizutani T., Haraguchi T., Yamamichi N., Minoguchi S., Yamamichi-Nishina M., Mori N., Kameda T., Sugiyama T. and Iba H. (2006) SWI/SNF complex is essential for NRSF-mediated suppression of neuronal genes in human nonsmall cell lung carcinoma cell lines. *Oncogene* **25**, 470–479.

PUBLICATION III

III Lend, AK.∗, **Kazantseva, A**.∗, Kivil, A., Valvere, V., Palm, K. (2015). Diagnostic significance of alternative splice variants of REST and DOPEY1 in the peripheral blood of patients with breast cancer. Tumour Biology, 36(4), 2473-80

Tumor Biol. (2015) 36:2473–2480
DOI 10.1007/s13277-014-2860-6

RESEARCH ARTICLE

Diagnostic significance of alternative splice variants of *REST* and *DOPEY1* in the peripheral blood of patients with breast cancer

Ave Kris Lend · Anna Kazantseva · Anri Kivil · Vahur Valvere · Kaia Palm

Received: 18 September 2014 / Accepted: 14 November 2014 / Published online: 26 November 2014

Abstract Changes in alternative splicing have been linked to cancer development. We hypothesized that changes occurring in tumor tissue can also be detected in the peripheral blood of cancer patients leading to discovery of blood biomarkers of breast cancer. Alternative splicing profiles of 94 genes were examined in cancerous breast tissue. Discriminating splice variants were analyzed in the peripheral blood of early stage (BCI/II) (stage I–II; n=26), neoadjuvant receiving locally advanced breast cancer patients (LABC) (stage IIb–IIIa, b; n=10) and healthy volunteers (n=26) using qRT-PCR analysis. Changes in marker expression during neoadjuvant therapy were analyzed at 15 timepoints. High expression of *REST-N50*, the alternatively spliced variant of *REST*, was detected in the blood of LABC patients but not in BCI/II and healthy controls (p=0.0032 and p=0.0029, respectively). Expression levels of *DOPEY1v2*, the alternative splice variant of *DOPEY1*, in the blood could differentiate cancer from healthy controls (p=0.024) and discriminate between patient groups (BCI/II vs LABC, p=0.002). Positive response to neoadjuvant therapy of *REST-N50*-positive LABC patients correlated with a decrease in *REST-N50* levels (p<0.0001). Assessment of *REST-N50* and *DOPEY1v2* may prove useful in diagnostic blood tests of breast cancer. *REST-N50* shows a high potential as a blood biomarker for evaluating the effectiveness of therapy in the neoadjuvant setting.

Keywords Alternative splicing · Locally advanced breast cancer · Neoadjuvant therapy · Blood · *REST* · *DOPEY1*

Ave Kris Lend and Anna Kazantseva contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s13277-014-2860-6) contains supplementary material, which is available to authorized users.

A. K. Lend · A. Kazantseva · A. Kivil · K. Palm (✉)
Protobios LLC, Mäealuse 4, 12618 Tallinn, Estonia
e-mail: kaia@protobios.com

A. K. Lend
e-mail: avekris@protobios.com

A. Kazantseva
e-mail: anna@protobios.com

A. Kivil
e-mail: anri@protobios.com

A. K. Lend · V. Valvere · K. Palm
Competence Center for Cancer Research,
Akadeemia tee 15, 12618 Tallinn, Estonia

V. Valvere
e-mail: Vahur.Valvere@regionaalhaigla.ee

A. Kivil · K. Palm
The Department of Gene Technology, Tallinn University of Technology, Akadeemia tee 15, 12618 Tallinn, Estonia

V. Valvere
Clinic of Hematology and Oncology, North Estonia Medical Center, J. Sütiste tee 19, 13419 Tallinn, Estonia

Introduction

Recent developments in the field of cancer research focusing on biomarkers are towards simple, cost-effective, and non-invasive tests. These tests are expected to indicate cancer risk, allow early detection, patient stratification for therapy, and monitor disease progression. For a long time, breast cancer, the second leading cause of death in women, has been classified according to the clinical and pathological criteria for treatment specification. However, molecular characterization has identified at least five different subtypes of breast cancer, allowing more detailed insight into the nature of this disease [1–3]. Given that breast cancer is a highly heterogeneous disease, the use of molecular analysis methods is highly justified, hence pushing the road towards a personalized medicine approach.

Locally advanced breast cancer (LABC) refers to most advanced stage non-metastatic breast tumors covering a spectrum of clinical features. It includes all of stage III disease and a subset of stage IIB tumors. Thus, patients with locally advanced breast cancer are a heterogeneous group with variable outcomes regarding local recurrence rates and survival. Neoadjuvant treatment is commonly used before primary therapy, i.e., surgery. However, the duration of treatment (neoadjuvant and postoperative adjuvant therapy) is long and there are no good markers for real-time monitoring of the effectiveness of the treatment. Given the toxic nature and the cost, monitoring of therapeutic response would be highly feasible. Efforts have been made to understand the mechanisms underlying the tumor responsiveness and to identify parameters that correlate closely with response to therapy. Nevertheless, there are currently no diagnostic biomarkers available to provide this information.

Though cancer tissue-based messenger RNA (mRNA) diagnostics are well ahead (e.g., MammaPrint®, Oncotype DX®, MapQuant Dx™ Genomic Grade for clinical use), there are few studies on blood-based multimarker mRNA tests for cancer [4]. Partly based on the study of Aarøe et al. [5], a commercial blood-based diagnostic test has been launched. BCtec® (DiaGenic ASA, Norway) is intended for detection of breast cancer at an early stage of disease with sensitivity of 68 % and specificity of 74 % [6]. It is clear that gene expression alone cannot fully explain the cellular phenotype. Among other aberrant changes, cancer development is accompanied by flawed transcript processing [7]. Alternative splicing is thought to be responsible for about 10–20 % of all cancer-related gene alterations [8], adding another layer of complexity to cancer formation. Differential patterns of splicing are common to breast cancer [9].

Thus, alternatively spliced mRNA variants (ASVs) are potential biomarker candidates for cancer diagnostics. In the present study, we demonstrate that ASVs identified in the primary tumor can also be found in the peripheral blood of cancer patents. Moreover, our data strongly indicate that detection of ASVs in the blood could be beneficial in efficacy monitoring of the neoadjuvant therapy of breast cancer.

Materials and methods

Study groups

Twenty-nine randomly selected tissue samples of primary breast carcinoma (clinical stages I–IV and tumor grades 1–3, age range 25–83 years), peripheral blood samples of randomly selected 26 patients diagnosed with early stage breast cancer (BCI/II) (stages I–II, age range 42–77 years) and serial peripheral blood samples taken from 10 patients with LABC receiving neoadjuvant treatment (stages IIb–IIIa, b; age range 36–67 years) were obtained from The North Estonia Medical Centre (NEMC). In the group of individuals receiving neoadjuvant therapy, only patients with inoperable breast cancer (either tumor greater than 4 cm or some other indication to be considered as inoperable with breast-preserving surgery) were included to the 2-year study (launched November 2007). The patients in the study group were treated with CEF (Epi 50) regimen: cyclophosphamide (500 mg/m^2 i.v. in days 1 and 8), epirubicin (50 mg/m^2 i.v. in days 1 and 8), and 5-fluorouracil (500 mg/m^2 i.v. in days 1 and 8) and a repeated course of treatment on day 29. The first peripheral blood sample was obtained before the first therapy session and the following samples were taken once a week throughout the four courses of CEF (Epi 50) with a total of 15 samples from different timepoints. Total RNA from breast tissues of four healthy individuals was obtained from Biochain (Biochain Institute Inc., CA, USA), Ambion (Life Technologies, CA, USA), and Stratagene (Stratagene, CA, USA), and normal breast tissue 5 donor pool was obtained from Biochain (Biochain Institute Inc., CA, USA). Peripheral blood of healthy female individuals with no known personal or family history of breast cancer was obtained from 26 volunteers (age range 17–66 years).

In the current study, the principles stated in the Declaration of Helsinki were followed. The National Ethical Committee provided ethical approval for the use of human tissue specimens for RNA extraction and analysis (decision no. 1813, issued 17 September 2009). Also, ethics approval for the neoadjuvant study was obtained from The National Ethical Committee (decision no. 1161, issued 13 September 2007). An informed consent was obtained from each subject.

RNA extraction and semi-quantitative RT-PCR

Tissue samples were taken during operation, immediately snap frozen on dry ice, and subsequently stored at −80 °C. Total RNA from primary breast tissue was isolated using RNAwiz RNA isolation reagent (Invitrogen, CA, USA) according to the manufacturer's protocol. A 10-ml peripheral blood sample was collected into the EDTA containing Vacutainer (BD Biosciences, CA, USA), and processed within 3 h of collection. Total RNA of nucleated cells from the samples of whole blood was extracted using LeukoLOCK™ Total RNA Isolation System (Invitrogen, CA, USA) according to the manufacturer's recommendations. DNase I treatment was performed using TURBO DNase (Invitrogen, CA, USA) according to the manufacturer's instructions. Reverse transcription using the SuperScript III First-Strand Synthesis System for RT-PCR (Invitrogen, CA, USA) was performed according to the standard protocol (with an exception of using 5 U of RNase Inhibitor (Bioron GmbH, Germany) and 100 U SuperScript III per reaction). First strand synthesis with 2 μg

of total RNA was primed using oligo (dT) primers (Microsynth, Switzerland).

List of genes analyzed is provided in Online Resource Table S1. Primers for RT-PCR were designed so that all exons covering the expected alternative splice site(s) located in the same amplicon. Twenty-nine samples of primary cancer tissue along with the normal breast tissue donor pool were used for the initial expression analysis of ASVs. PCR was carried out in a total volume of 20 μl using 1× FIREPol® Master Mix (Solis Biodyne, Estonia), 250 nM of primers, and 1/100 of complementary DNA (cDNA). The following PCR program was used: initial denaturation at 95 °C for 3 min; cycling (35–40 cycles) 95 °C for 30 s, 58 °C for 30 s, and 72 °C for 1 min; final extension at 72 °C for 5 min. PCR products were separated on a 1.2–2 % agarose-TBE gel.

Cloning and sequencing

Prior sequencing PCR products were purified with QIAquick® PCR Purification Kit (QIAGEN, Inc., CA, USA) according to the supplier's protocol and cloned into the pSC-A vector using StrataClone™ PCR Cloning Kit (Agilent Technologies, CA, USA). Cloning reaction was performed in a total volume of 0.6 μl according to the manufacturer's protocol. For sequencing, GATC Biotech (Germany) services were used. Alternatively, spliced transcripts were identified using BLAT (http://genome.ucsc.edu).

Quantitative RT-PCR

The expression of selected cancer-specific ASVs was further analyzed in solid tissue and blood samples by qRT-PCR, using splice junction-specific primers for amplification of the selected ASVs (primers provided in Online Resources Table S2). The reaction of qRT-PCR was carried out in a reaction volume of 6 μl (3 μl 2× SYBR® Mix, 0.1 μl ROX, 0.6 μl 10 mM forward and reverse primers, 0.3 μl BSA and 2 μl cDNA) using Platinum® SYBR® Green qPCR SuperMix-UDG (Invitrogen, CA, USA) with the following program: initial denaturation at 95 °C for 2 min; cycling (50 cycles) 95 °C for 10 s, 60 °C for 10 s, and 72 °C for 10 s. Melting curve analysis was performed at the end of each run. For all amplifications, 7900HT Fast Real-Time PCR System (Applied Biosystems, CA, USA) and MicroAmp Optical 384-well reaction plates (Applied Biosystems, CA, USA) were used. All samples were analyzed in triplicates. ASV expression was normalized to the expression of two reference genes (*GAPDH* and *HPRT* or *ACTB*).

Statistical analysis

The normalized qRT-PCR data were analyzed using the $2^{-\Delta\Delta Ct}$ method [10]. Heat map was created using matrix visualization and analysis platform GENE-E (http://www.broadinstitute.org/cancer/software/GENE-E/index.html). The average expression levels among different study groups were compared using unpaired t test. Values of $p<0.05$ were considered statistically significant. In order to set an optimal threshold value for determining the sensitivity and specificity of the ASV found in the blood of patients, a Receiver operating characteristic (ROC) analysis of the different study groups was performed using MedCalc statistical software (MedCalc software, Belgium). Tumor response to neoadjuvant therapy was evaluated by the treating clinician. Correlation analysis of the expression of ASVs in the blood and effectiveness of the neoadjuvant treatment was performed with coefficient of determination (R^2) and unpaired t test. Values of $p<0.05$ were considered statistically significant.

Results

Assessment of diagnostic significance of tumor-specific alternatively spliced transcripts as candidate biomarkers for breast cancer

Selection of cancer-prone ASVs from 94 human genes was made using in silico approach. The names of the genes and selected ESTs are provided in Online Resource Table S1. RT-PCR analysis performed on 29 primary breast cancer and healthy breast tissue pooled samples revealed that 75 ASVs of 51 genes showed differential expression in cancer versus control tissues. Furthermore, 22 of them were identified as novel ASVs. By including only breast carcinomas ($n=23$) and expanding normal sample size, 24 ASVs were then selected and analyzed by qRT-PCR. The expression profile of 12 ASVs as the most discriminative signature for normal and cancer tissue is shown as a heat map graph in Fig. 1. In majority of the cases, ASVs were expressed at higher levels in cancer than in normal tissue samples. Downregulated expression of ASVs of *CASP9* (*CASP9_DA909628*), *PPARG* (*PPARG_AB107271*), and *CCNH* (*CCNH_AK094534*) was observed in cancer samples as compared to normal tissue. In contrast, the ASVs of *PRC1* (*PRC1_BG827840*) and *BUB1* (*BUB1_BF665520*) showed at least fivefold higher expression in cancer as compared to normal tissue. A human-specific transcript of *REST* (also known as *NRSF*), namely *REST_AF228045* (hereinafter referred to as *REST-N50*) was detected exclusively in the cancer tissues ($n=10$, 43 %). Also, *REST* transcript encoding for the canonical form of the protein, which expression has been demonstrated to be lost in a subset of cancers [11], was not detected in samples that showed extremely high expression of *REST-N50* ($n=3$, 13 %). When correlating clinical parameters with ASV expression data,

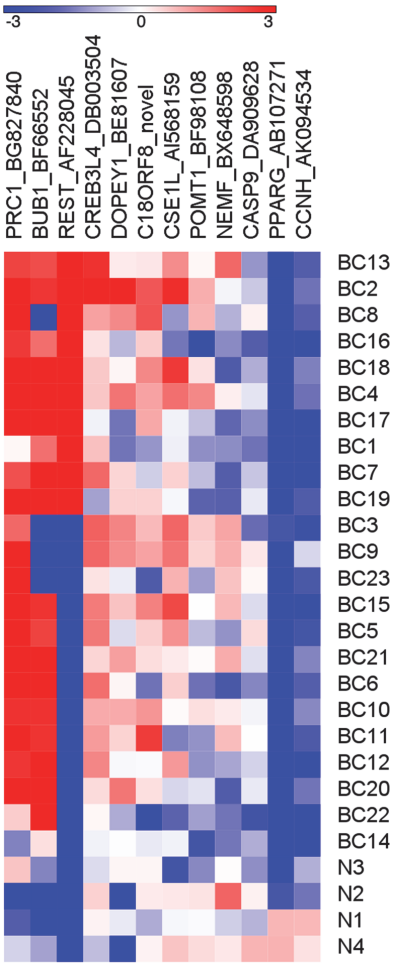

Fig. 1 Expression heat map illustrating the relative expression of the 12 most significantly differentially expressed ASVs in samples of breast cancer (*BC*) and healthy tissue (*N*). The expression data was calculated as relative fold difference compared to average normal tissue expression data. Heat map was then created using the minimum and maximum values across the entire data set. Results are presented in log scale. Positive values (*red*) indicate high expression relative to other samples in the dataset, negative values (*blue*) correspond to lower expression relative to other samples in the dataset, and mean values are represented by *zero*. ASVs that were analyzed are provided with GenBank accession numbers (except for ASV of *C18ORF8*, which is novel) above the heat map

overall, the expression of ASVs showed no correlation with cancer stage, grade, or hormone-receptor (PR, ER, HER2) status (data not shown). However, there was a significant positive correlation between the levels of *PRC1_BG827840* expression and Ki67 proliferative index (p=0.0023).

Tumor-specific ASVs detected in the peripheral blood of breast cancer patients

In disease diagnostics, heading towards the less invasive sample collection (e.g., peripheral blood sample vs tissue biopsy) is highly desired. Based on the tissue findings, we proposed that tumor-specific ASVs would also be detected in the peripheral blood of breast cancer patients. Expression of the 12 selected cancer-specific ASVs (Fig. 1) was tested and found to be low in the peripheral blood leukocytes of healthy individuals (n=5). Hence, expression of these ASVs was examined in the leukocyte fraction of the peripheral blood of the 26 BCI/II patients, 10 inoperable LABC patients, and 26 healthy individuals with no known history of breast cancer in family.

Out of 12 ASVs tested, two showed statistically significant expression in cancer patients. The expression of *REST-N50* was detected at significantly higher levels in the peripheral blood of LABC patients as compared to that of BCI/II group and healthy controls (p=0.0032 and p=0.0029, respectively), but did not discriminate BCI/II patients from healthy controls (p=0.452) (Fig. 2a). Using ROC analysis, the sensitivity of 70 % (95 % CI 34.8–93.3) and specificity of 96.2 % (95 % CI 80.4–99.9) with the positive likelihood ratio (LR) of 18.2 was determined for *REST-N50* in the LABC group (Fig. 2b). The average expression levels of *DOPEY1_ BE816078* (hereinafter referred to as *DOPEY1v2*) were significantly higher in the blood of cancer patients as compared to healthy individuals (p=0.024). The discriminatory effect of *DOPEY1v2* remained when cancer samples were separated by disease severity (BCI/II and LABC) and the expression of *DOPEY1v2* in these study groups was compared to the expression in healthy controls (p=0.026 and p=0.0004, respectively). Also, *DOPEY1v2* showed significantly higher levels of expression in LABC group when compared to BCI/II patients (p=0.002) (Fig. 2c). ROC analysis revealed 60 % sensitivity (95 % CI 26.2–87.8) and 96.2 % specificity (95 % CI 80.4–99.9) for this ASV in LABC patients (Fig. 2d). The correlation between the expression of *DOPEY1v2* or *REST-N50* with other clinical data (hormonal receptor status, tumor Ki67 index) was not detected.

The significance of *REST-N50* as a monitoring blood biomarker of neoadjuvant therapy in breast cancer

Next, we examined the expression of cancer-specific ASVs in blood samples of LABC patients receiving neoadjuvant therapy. During a 2-year study in collaboration with NEMC, 10 patients considered inoperable with breast-preserving surgery were enrolled and treated with a regimen of four courses of CEF (Epi 50) (see "Materials and methods" for the administration schedule and doses). During the study, one planned therapy was postponed due to leukopenia, one was

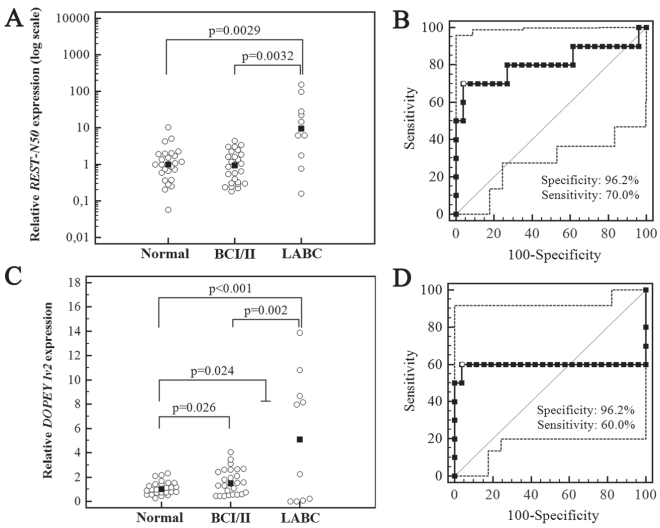

Fig. 2 Expression analysis of *REST-N50* and *DOPEY1v2* in the peripheral blood of breast cancer patients and healthy individuals. Multiple comparison blot of **a** *REST-N50* and **c** *DOPEY1v2* in different study groups: healthy individuals (*normal*), patients with early stage (*BCI/II*), and locally advanced breast tumors (*LABC*). Each *white dot* represents an individual and *black squares* mark median values of each study group. The expression data is given as relative fold-change values compared to the median expression of healthy controls. The relative expression data of *REST-N50* is presented in log rank. Statistically significant expression differences between study groups are shown as *p*-values. ROC curve analysis of **b** *REST-N50* and **d** *DOPEY1v2* in LABC group is shown. *White dot* represents the farthest point from the diagonal line, defining the criteria for marker specificity and sensitivity determination. Specificity and sensitivity for each marker is shown. *Black dotted lines* indicate confidence interval range. Y-axis shows true positive rate (sensitivity), and on X-axis a false positive rate is shown (100-specificity)

discontinued due to the disease progression during neoadjuvant therapy, and one terminated for non-attendance. Out of the remaining seven patients receiving neoadjuvant therapy, six showed a favorable response to the treatment and were subsequently operated.

To determine the suitability and prognostic significance of selected candidate markers for monitoring cancer treatment, the expression dynamics of *REST-N50* and *DOPEY1v2* was analyzed by qRT-PCR in peripheral blood sampled at different timepoints of neoadjuvant therapy. Five out of 10 patients showed a prominent expression of *REST-N50* before the treatment began (Fig. 2a). All these five patients completed the full CEF (Epi 50) therapy regimen with a positive response and were subjected to surgery. Intriguingly, when observing the average levels of *REST-N50* during the course of the treatment, there was a visible trend towards decreased levels of expression already upon completion of two CEF (Epi 50) courses (r=0.80, p=0.009; measured at 2 weeks after the second course) (Fig. 3, 9th blood sample). *REST-N50* levels continued to decline upon completion of the four-course therapy (r=0.85, p<0.0001) (Fig. 3). The average decrease in the levels of *REST-N50* measured before and after completing the neoadjuvant treatment differed significantly (p=0.017) and in

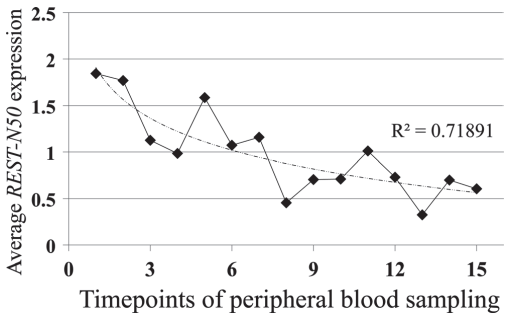

Fig. 3 Expression dynamics of *REST-N50* in the blood of *REST-N50*-positive patients with LABC during CEF (Epi 50) therapy. The expression data was calculated as relative fold difference compared to median normal expression data. The relative expression of *REST-N50* of each individual was unified by dividing each timepoint of *REST-N50* expression value with the average total expression value of individual sample. Group average for each timepoint was then calculated, and the normalized average *REST-N50* expression dynamics after four courses of neoadjuvant therapy is shown. The 1st sample is taken prior to therapy, the 9th sample is taken 2 weeks after the second course of CEF (Epi 50), and the 15th sample is taken a week after finishing the therapy. Each *square* represents one timepoint of peripheral blood sampling. Along the logarithmic trend line, a coefficient of determination (R^2) is shown

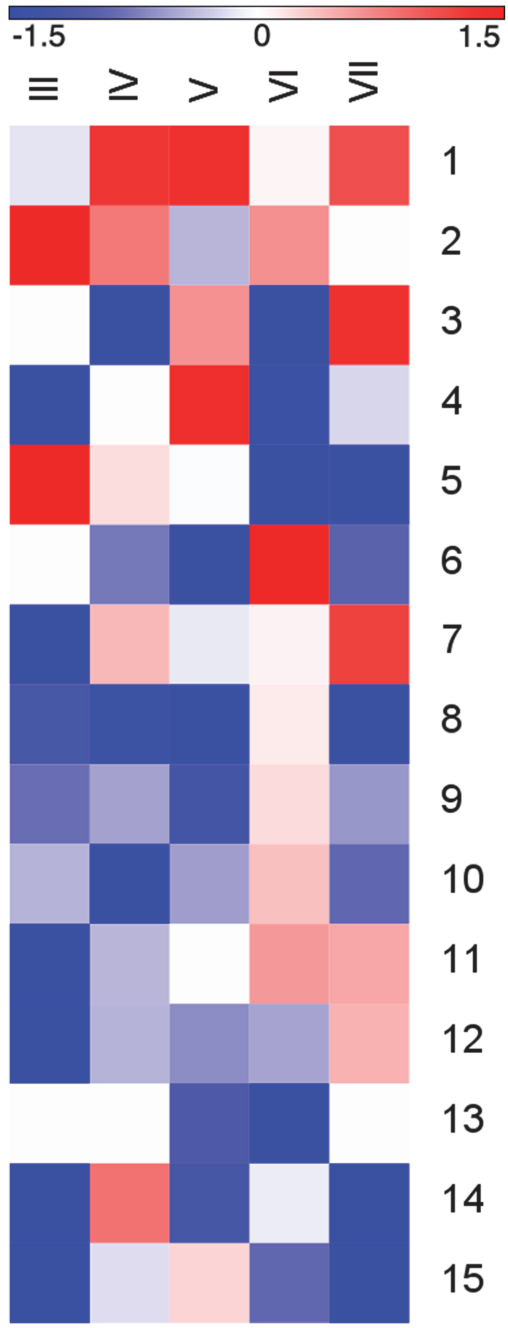

Fig. 4 Heat map presentation of relative expression of *REST-N50* in the peripheral blood of five *REST-N50*-positive LABC patients (indicated in *Roman numerals*) at 15 timepoints (shown on the *right*) spanning four courses of CEF (Epi 50). The expression data was calculated relative to median normal expression data. The heat map was created using the minimum and maximum values across the entire data set. Results are presented in a log scale. Positive values (*red*) indicate high expression relative to other samples in the dataset, negative values (*blue*) correspond to lower expression relative to other samples in the dataset, and mean values are represented by *zero*

different individuals varied from two- to tenfold. The individual expression dynamics of *REST-N50* in five LABC patients is shown in Fig. 4. Expression levels of *DOPEY1v2* showed no significant change during therapy.

Discussion

Herein, we report the use of cancer-specific ASVs for breast cancer detection and assessment of therapeutic response. Altogether, 94 genes were analyzed for cancer-related ASV occurrence. As a result, 12 ASVs were identified which showed differential expression in breast cancer as compared with normal breast tissue. Although some of them (e.g., *REST-N50*) had previously been associated with different types of cancers, none of the ASVs have earlier been annotated as targets for diagnostic analysis of breast pathologies. By examining whether the selected ASVs would specifically be expressed in the peripheral blood of breast cancer patients, we found that alternatively spliced transcripts of *REST* and *DOPEY1* could serve as blood markers in diagnostic tests for advanced stage breast cancer. For the first time, this study illustrates the power of expression analysis of alternatively spliced transcripts, foremost *REST-N50* as a blood biomarker correlating with the efficacy outcome of neoadjuvant therapies.

Alternatively spliced variants *REST-N50* and *DOPEY1* are elevated in peripheral blood of patients with breast cancer

In the current study, the expression of *REST-N50* mRNAs was observed in 43 % of the biopsy samples and in the peripheral blood of 50 % of patients diagnosed with inoperable breast cancer. qRT-PCR profiling revealed that *REST-N50* ASV is not expressed in normal tissues. Moreover, tissues that express *REST-N50* at high levels do not show simultaneous expression of full-length *REST* mRNAs. *REST-N50* has a 50-bp insert incorporated between exons 5 and 6, which, by introducing a stop codon, leads to C-terminally truncated protein. The same splice variant has been detected at high levels in small cell lung cancer [12] and a subset of breast cancers [11]. Neural-specific alternative splice (*REST-N62*) leading to the same truncated version of REST has been found in neuroblastomas [13]. Wagoner et al. [11] defined REST-less breast tumors by

using an expression profile of 24 target genes of REST and argued that tumors with no REST protein were more aggressive and correlative to poor disease-free survival. Interestingly, ASV of *REST* incorporating the 50-bp exon was expressed at high levels in REST-less tumors [11]. To our knowledge, our study is the first to detect *REST-N50* transcript in the blood of cancer patients. The fact that *REST-N50* was well detected in the blood of LABC patients with inoperable tumors might possibly implicate the tumor load in blood circulation of those patients as the expression level of this ASV was low in the blood of patients with stages I–II breast cancer. According to Wagoner et al. [11] it might also correlate with the aggressive tumor phenotype. In good agreement with our data, they established the lymph node involvement and the size of REST-less tumors that was significantly larger than that of tumors expressing *REST* and not expressing *REST-N50*.

In contrast, elevated levels of *DOPEY1v2*, the splice variant of *DOPEY1*, were detected in the blood of BCI/II patients, but also in patients diagnosed with LABC. Our data show that expression of *DOPEY1v2* was higher in the blood of LABC patients in comparison with BCI/II patients, thus correlating with disease severity. *DOPEY1v2* has a retained intron between exons 32 and 33 (exon 12/13 in transcript variant 1). According to GNF Expression Atlas 2 (U133A chip) data, expression of *DOPEY1* is detected in different brain regions and in different types of peripheral blood cells (mast cells, T-lymphocytes, NK cells) [14]. It remains unclear whether *DOPEY1v2* emerged due to the load of cancer cells in the blood or as a result of blood cell imbalances or cellular homeostasis, given that expression profiles of blood cells are affected already at early stages of cancer [4].

It needs to be pointed out that one of the shortcomings of the current study stems from the number of patients and healthy controls analyzed. Because of that, results regarding *REST-N50* lack sufficient statistical power (0.66 for two-sided test, $\alpha=0.05$). For *DOPEY1v2*, the number of patients fulfils the statistical requirements, with the power of 0.8 in case of α 0.05. Nevertheless, detecting *DOPEY1v2* and, most importantly, *REST-N50* in the peripheral blood of patients with breast cancer might prove highly useful as a criterion for stratification and/or disease staging.

Expression of *REST-N50* declines in the blood of patients with breast cancer receiving neoadjuvant therapy

Nowadays breast cancer is increasingly being diagnosed at an earlier stage. However, there remains a group of patients presenting locally advanced disease who are considered inoperable. These patients are subjected to the neoadjuvant treatment prior to surgery. Emphasizing the heterogeneity of cancer, it is practically impossible to predict the outcome of the treatment. Because of this, there is a high need of biomarkers for monitoring of therapeutic response as early identification of non-responders might help to reorganize the therapy.

With reasoning that ASVs found in the peripheral blood of cancer patients could be useful in monitoring the efficacy of neoadjuvant therapy, we specifically examined the expression of *REST-N50* and *DOPEY1v2* in the blood of patients with breast cancer receiving CEF (Epi 50) therapy. Circulating epithelial cancer cells have been shown to respond in a manner identical to the tumor during the first 3–4 cycles of neoadjuvant therapy [15]. Hence, if markers originate from the cancer cells, a decrease in the levels of these markers would be expected. This was indeed what we observed, as patients with declining expression of *REST-N50* showed a positive therapy effect on tumor size. In the current study, patients were treated with four rounds of CEF (Epi 50). Already during the first two courses, a significant decrease in the expression of *REST-N50* was observed. By the end of the therapy the individual differences were up to tenfold between the initial and final levels of *REST-N50*. For the first time, we demonstrated a decline in *REST-N50* levels upon therapy, supporting the use of this ASV as a biomarker to assess the treatment response. However, as *REST-N50* positive neoadjuvant therapy non-responder samples were not available, and the group of LABC patients was limited in number, a larger sample size is needed for definitive conclusions.

The application of tumor-associated transcripts as biomarkers in peripheral blood depends largely on their levels of expression in hematopoietic cells. One of the main obstacles in finding a biomarker suitable for detection in the blood is a fact that ~80 % of genes of the human genome is expressed in human peripheral blood cells [16]. Also, some tumor-associated transcripts such as epithelial-specific genes are induced in lymphocytes in reaction to activation [17]. Hence, this is probably the reason why some of the candidate markers (e.g., ASVs derived from cell proliferation-related molecules such as *PRC1* and *BUB1*) were suited for analysis of cancer tissues with high mitotic activity and inflammatory infiltrates but were discarded as reliable biomarkers of cancer in the blood, which inherently has a relatively high mitotic index. It is also possible, that cancer cells change their gene expression profiles when reaching the bloodstream. However, immanently characteristic markers have still shown to persist [3]

Conclusion

Although it is too early to discuss the use of ASVs described in the current study as clinically valuable surrogate biomarkers, our results raise novel opportunities for molecular identification of breast cancer. In good agreement with other studies, we found that expression of *REST-N50* does not

correlate with known characteristics of the tumor, allowing to hypothesize that *REST-N50* is a hallmark of yet another subtype of breast cancer. In fact, it has been shown that loss of *REST* can contribute to tumor aggression [18]. It has also been proposed that alternatively spliced *REST-N50* could be a target for therapy [11]. Therefore, detection of this ASV from the peripheral blood, as our data show, could well serve as a marker for the assessment of the therapy effectiveness. Taken together, our study underscores the value of detection of alternatively spliced events, in particular these of *REST-N50* and *DOPEY1v2*, in the blood of breast cancer patients as markers for cancer detection as well as for surveillance of therapy.

Acknowledgments This study was supported by Protobios's grants from the Enterprise of Estonia and baseline financing from the Estonian Ministry of Education and Research. We are grateful to Dr. Jõeste and NEMC for the tissue samples. We thank K.J.M. Rand for the work with in silico database, primer design, and cloning; Dr. Pruunsild and Prof. Timmusk for the resources provided; and K. Taal and H. Verev for the technical help in sample handling.

Conflicts of interest None.

References

1. Herschkowitz JI, Simin K, Weigman VJ, Mikaelian I, Usary J, Hu Z, et al. Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. Genome Biol. 2007;8:R76. doi:10.1186/gb-2007-8-5-r76.
2. Hu Z, Fan C, Oh DS, Marron JS, He X, Qaqish BF, et al. The molecular portraits of breast tumors are conserved across microarray platforms. BMC Genomics. 2006;7:96. doi:10.1186/1471-2164-7-96.
3. Perou CM, Sørlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature. 2000;406:747–52. doi:10.1038/35021093.
4. Sharma P, Sahni NS, Tibshirani R, Skaane P, Urdal P, Berghagen H, et al. Early detection of breast cancer based on gene-expression patterns in peripheral blood cells. Breast Cancer Res. 2005;7:R634–44. doi:10.1186/bcr1203.
5. Aarøe J, Lindahl T, Dumeaux V, Saebø S, Tobin D, Hagen N, et al. Gene expression profiling of peripheral blood cells for early detection of breast cancer. Breast Cancer Res. 2010;12:R7. doi:10.1186/bcr2472.
6. Tobin D, Karlsson M, Hagen N, Børresen-Dale A, Sharma P. Use of the blood based, 96-assay set for breast cancer detection. Poster presented at the 21st meeting of the EACR 26–29 June 2010 Oslo, Norway. 2010.
7. Venables JP. Aberrant and alternative splicing in cancer. Cancer Res. 2004;64:7647–54. doi:10.1158/0008-5472.CAN-04-1910.
8. Faustino NA, Cooper TA. Pre-mRNA splicing and human disease. Genes Dev. 2003;17:419–37. doi:10.1101/gad.1048803.
9. Venables JP, Klinck R, Bramard A, Inkel L, Dufresne-Martin G, Koh C, et al. Identification of alternative splicing markers for breast cancer. Cancer Res. 2008;68:9525–31. doi:10.1158/0008-5472.CAN-08-1769.
10. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(−delta delta C(T)) method. Methods. 2001;25:402–8. doi:10.1006/meth.2001.1262.
11. Wagoner MP, Gunsalus KTW, Schoenike B, Richardson AL, Friedl A, Roopra A. The transcription factor REST is lost in aggressive breast cancer. PLoS Genet. 2010;6:e1000979. doi:10.1371/journal.pgen.1000979.
12. Coulson JM, Edgson JL, Woll PJ, Quinn JP. A splice variant of the neuron-restrictive silencer factor repressor is expressed in small cell lung cancer: a potential role in derepression of neuroendocrine genes and a useful clinical marker. Cancer Res. 2000;60:1840–4.
13. Palm K, Metsis M, Timmusk T. Neuron-specific splicing of zinc finger transcription factor REST/NRSF/XBR is frequent in neuroblastomas and conserved in human, mouse and rat. Brain Res Mol Brain Res. 1999;72:30–9.
14. Su AI, Wiltshire T, Batalov S, Lapp H, Ching KA, Block D, et al. A gene atlas of the mouse and human protein-encoding transcriptomes. Proc Natl Acad Sci U S A. 2004;101:6062–7. doi:10.1073/pnas.0400782101.
15. Pachmann K, Dengler R, Lobodasch K, Fröhlich F, Kroll T, Rengsberger M, et al. An increase in cell number at completion of therapy may develop as an indicator of early relapse: quantification of circulating epithelial tumor cells (CETC) for monitoring of adjuvant therapy in breast cancer. J Cancer Res Clin Oncol. 2008;134:59–65. doi:10.1007/s00432-007-0248-3.
16. Liew C, Ma J, Tang H, Zheng R, Dempsey AA. The peripheral blood transcriptome dynamically reflects system wide biology: a potential diagnostic tool. J Lab Clin Med. 2006;147:126–32. doi:10.1016/j.lab.2005.10.005.
17. Kowalewska M, Chechlińska M, Markowicz S, Kober P, Nowak R. The relevance of RT-PCR detection of disseminated tumour cells is hampered by the expression of markers regarded as tumour-specific in activated lymphocytes. Eur J Cancer. 2006;42:2671–4. doi:10.1016/j.ejca.2006.05.036.
18. Gunsalus KTW, Wagoner MP, Meyer K, Potter WB, Schoenike B, Kim S, et al. Induction of the RNA regulator LIN28A is required for the growth and pathogenesis of RESTless breast tumors. Cancer Res. 2012;72:3207–16. doi:10.1158/0008-5472.CAN-11-1639.

ABSTRACT

Alternative splicing is an essential and highly regulated molecular mechanism affecting the gene expression in eukaryotes. During splicing, coding sequences of DNA, or exons, are either retained in the mRNA or targeted for removal in different combinations to generate various mRNAs from a single gene resulting in expression of multiple proteins, often with distinct functions.

By means of modern genomics methods, it is now possible to get a detailed understanding of complex cellular mechanisms of splicing. The interest in splicing and its crucial role in the processes of differentiation, development or immunological events is growing day by day. The role of alternative splicing in human diseases, including cancer, is only starting to be unveiled.

During the studies, I have analyzed thousands of genes for their modes of tissue-specific alternative splicing. These genes included those involved in the regulation of transcription, associated with neuronal features and cancer. The genes with the most marked tissue-specific alternatively spliced mRNAs were studied in more detail and the results of these studies have been published in several scientific papers.

First, I showed that the gene coding for brain-derived neurotrophic factor (BDNF), a protein that promotes the survival and maintenance of nerve cells, is alternatively spliced and its multiple transcripts are differentially expressed in various human tissues. However, high expression of *BDNF* alternative mRNAs was characteristic to specific regions of the adult brain. The complex organization of *BDNF* gene locus with its multiple promoters, polyadenylation signals, and antisense mRNAs indicates that expression of this gene is highly regulated at different levels.

Second, I analyzed the expression of *BAF57* gene, which encodes one of the subunits of a chromatin remodeling complex SWI/SNF. As a result of alternative splicing, *BAF57* generates several protein isoforms, which are exclusively expressed in neurons of developing and adult brain. These neuron-specific isoforms are able to assemble into functional SWI/SNF complexes and might be involved in the regulation of expression of some neuronal genes.

Finally, I analyzed alternative splice variants of hundreds of genes as potential biomarkers of breast cancer. Splice variants of *REST* and *DOPEY1* genes appeared to be the best marker candidates for the detection and monitoring of breast cancer from the peripheral blood of patients.

Taken together, studying of alternative splicing, which plays a central part in generating complex proteomes of eukaryotes, broadens our knowledge and understanding of the nature and operation of genetic codes.

KOKKUVÕTE

Alternatiivne splaising (AS) on väga oluline ja tugevalt reguleeritud molekulaarne mehhanism, mis mõjutab eukarüootsete geenide ekspressiooni väga suurel määral. Valke kodeerivad järjestused genoomis, teisisõnu eksonid, liidetakse splaisingu käigus eri kombinatsioonidena mRNA-de koosseisu. Selle tulemusena loetakse ühelt geenilt mitu erinevat mRNA-d, millest igaüks võib kodeerida sageli unikaalse funktsiooniga valku. Nüüdisaegse genoomika meetodite abil on võimalik splaisingu keerulisi mehhanisme väga täpselt kirjeldada. Huvi splaisingu uurimise vastu organismi arengus, rakkude diferentseerumisel või immuunsüsteemi küpsemisel on järjepidevalt kasvanud. AS-i mehhanismide uuringud haigustes, ka vähis, on aga alles algusjärgus.

Doktoritöö tegemise ajal analüüsisin tuhandete geenide koespetsiifilise splaisingu mustreid. Nende seas oli geene, millelt sünteesitud valgud osalevad transkriptsiooni regulatsioonis, ning geene, mille avaldumine on iseloomulik närvi- või vähikoele. Kõige paremini eristuva koespetsiifilisusega alternatiivseid mRNA-sid uurisin detailsemalt ja avaldasin saadud tulemused teadusartiklites.

Oma töös leidsin esiteks, et närvirakkude kasvu ja ellujäämist soodustava valgu, aju neurotroofse faktori BDNF ekspressioon on AS-i poolt oluliselt reguleeritud ning vastavad mRNA-d koe- ja rakuspetsiifliselt ekspresseeritud. Oma töös näitasin, et täiskasvanud aju teatud piirkondadele on *BDNF*-i alternatiivsete mRNA-de olemasolu väga iseloomulik. *BDNF*-i geeni struktuurne keerukus, sealhulgas erinevate promootorite ja polüadenülatsiooni signaalide olemasolu, samuti kattuvus *antisense* mRNA-ga võimaldas järeldada, et selle geeni ekspressioon hõlmab mitmeid regulatsiooni tasandeid.

Teiseks analüüsisin AS-i mõju *BAF57* geeni ekspressioonile. *BAF57* kodeerib kromatiini remoduleerimise SWI/SNF kompleksi üht alaühikut. *BAF57* AS-i tulemusena tekib mitu valgu isovormi, mis nii areneva kui ka täiskasvanu aju neuronites on erinevalt ekspresseeritud. Näitasin oma uurimistöös, et BAF57 neuraalspetsiifilised isovormid on võimelised moodustama funktsionaalseid SWI/SNF komplekse ning seega oluliselt mõjutama neuraalsete geenide avaldumist.

Kolmandaks analüüsisin paljude geenide alternatiivseid splaiss-variante kui potentsiaalseid rinnavähi tuvastamise markereid. Leidsin, et *REST* ja *DOPEY1* alternatiivselt splaissitud transkriptid on sobivad markerikandidaadid patsientide vereproovidest rinnavähi tuvastamiseks ja monitooringuks.

Kokkuvõtteks: AS-i uurimine on erakordselt oluline meie teadmiste avardamiseks, sest sellel protsessil on keskne roll eukarüootidele iseloomuliku keeruka proteoomi kujunemisel.

CURRICULUM VITAE

Name: Anna Kazantseva
Date and place of birth: 04.04.1981, Tallinn
Citizenship: Estonian
Tel: +372 58172900
E-mail: anna@protobios.com

Education

2005–2015 PhD, Tallinn University of Technology, Department of Gene Technology
2003–2005 MSc, Tallinn University of Technology, Department of Gene Technology
1999–2003 BSc, University of Tartu, Institute of Molecular and Cellular Biology, Chair of Molecular Biology
1988–1999 Tallinn 64. Secondary school

Employment

01.2009–06.2015 The Competence Centre for Cancer Research, Research scientist
01.2006–01.2009 Protobios LLC, Research scientist
06.2005–01.2006 The Competence Centre for Cancer Research, Research scientist
01.2005–06.2005 Tallinn University of Technology, Department of Gene Technology, Chair of Molecular Biology, Engineer
09.2003–12.2004 National Institute of Chemical Physics and Biophysics, Engineer

Publications

Aid, T., **Kazantseva, A**., Piirsoo, M., Palm, K., Timmusk, T. (2007). Mouse and rat BDNF gene structure and expression revisited. Journal of Neurosci Research, 85(3), 525-35.

Pruunsild, P.[1], **Kazantseva, A**. [1], Aid, T., Palm, K., Timmusk, T. (2007). Dissecting the human BDNF locus: bidirectional transcription, complex splicing, and multiple promoters. Genomics, 90(3), 397-406

Kazantseva, A. [1], Sepp, M. [1], Kazantseva, J., Sadam, H., Pruunsild, P., Timmusk, T., Neuman, T., Palm, K. (2009). N-terminally truncated BAF57 isoforms contribute to the diversity of SWI/SNF complexes in neurons. Journal of Neurochemistry, 109(3), 807-818

Kazantseva, J., Kivil, A., Tints, K., **Kazantseva, A**., Neuman, T., Palm, K. (2013). Alternative splicing targeting the hTAF4-TAFH domain of TAF4 represses proliferation and accelerates chondrogenic differentiation of human mesenchymal stem cells. PLoS One, 2013; 8(10)

Lend, AK. [1], **Kazantseva, A**. [1], Kivil, A., Valvere, V., Palm, K. (2015). Diagnostic significance of alternative splice variants of REST and DOPEY1 in the peripheral blood of patients with breast cancer. Tumour Biology, 36(4), 2473-80

Sadam, H., Liivas. U., **Kazantseva, A**., Pruunsild, P., Kazantseva, J., Timmusk, T., Neuman, T., Palm, K. (2015). GLI2 cell-specific activity is controlled at the level of transcription and RNA processing: consequences to melanoma progression. *BBA* (*Biochimica et Biophysica Acta*) Molecular Basis of Disease, 1862(1):46-55

[1] – equal contribution

Intellectual properties

CANCER RELATED ISOFORMS OF COMPONENTS OF TRANSCRIPTION FACTOR COMPLEXES AS BIOMARKERS AND DRUG TARGETS. Owner: Oncotx. Inc; Inventors: **A. Kazantseva**, J. Kazantseva; Patent number US 2010/0087376 A1.

COMPOSITIONS AND METHODS FOR TARGETING CANCER-SPECIFIC TRANSCRIPTION COMPLEXES. Owner: Oncotx. Inc; Inventors: **A. Kazantseva**, A. Liik; Patent number 7973135.

Courses and conferences

June 2005. Tallinn, Estonia. IBRO Summer School in Neurobiology. *Oral presentation "The complex organization of the brain-derived neurotrophic factor gene locus in humans"*.

December 7 – 9, 2005. Berlin, Germany. Ambion® siRNA course.

April 21 – 26, 2006. Taos, New Mexico, USA. Keystone symposia. "Regulation of Eukaryotic Transcription: From Chromatin to mRNA". *Poster presentation "Cell-type specific isoforms of transcription cofactors"*.

January 11–14, 2007. La Jolla, California, USA. Nature symposium. "Biological Complexity: Diseases of Transcription". *Poster presentation "Novel neuronal specific BAF57 isoforms description"*.

August 29 – September 2, 2007. Cold Spring Harbour, New York, USA. Cold Spring Harbour Conference. "Mechanisms of eukaryotic transcription". *Poster presentation "Neuron-specific Isoforms of Human BAF57 Regulate Transcription of RE1/NRSE-containing genes"*.

June 7–10, 2010. Oosterhout, The Netherlands. Luminex xMAP® Technology course. "Fundamental Assay Techniques Nucleic Acid and Protein".

October 20–23, 2010. Wien, Austria. Planet xMAP Europe Conference on Luminex Technology.

ELULOOKIRJELDUS

Nimi: Anna Kazantseva
Sünniaeg ja -koht: 04.04.1981, Tallinn
Kodakondsus: Eesti
Tel: +372 58172900
E-mail: anna@protobios.com

Hariduskäik

2005–2015 PhD, Tallinna Tehnikaülikool, Matemaatika-loodusteaduskond, Geenitehnoloogia instituut
2003–2005 MSc, Tallinna Tehnikaülikool, Matemaatika-loodusteaduskond, Geenitehnoloogia instituut
1999–2003 BSc, Tartu Ülikool, Molekulaar- ja rakubioloogia instituut, Molekulaarbioloogia õppetool
1988–1999 Tallinna 64. keskkool

Teenistuskäik

01.2009–06.2015 Vähiuuringute Tehnoloogia Arenduskeskus (VTAK), teadur
01.2006–01.2009 Protobios OÜ, teadur
06.2005–01.2006 Vähiuuringute Tehnoloogia Arenduskeskus (VTAK), teadur
01.2005–06.2005 Tallinna Tehnikaülikool, Geenitehnoloogia instituut, insener
09.2003–12.2004 Keemilise ja Bioloogilise Füüsika Instituut, insener

Publikatsioonid

Aid, T., **Kazantseva, A**., Piirsoo, M., Palm, K., Timmusk, T. (2007). Mouse and rat BDNF gene structure and expression revisited. Journal of Neurosci Research, 85(3), 525-35.

Pruunsild, P. [1], **Kazantseva, A**. [1], Aid, T., Palm, K., Timmusk, T. (2007). Dissecting the human BDNF locus: bidirectional transcription, complex splicing, and multiple promoters. Genomics, 90(3), 397-406

Kazantseva, A. [1], Sepp, M. [1], Kazantseva, J., Sadam, H., Pruunsild, P., Timmusk, T., Neuman, T., Palm, K. (2009). N-terminally truncated BAF57 isoforms contribute to the diversity of SWI/SNF complexes in neurons. Journal of Neurochemistry, 109(3), 807-818

Kazantseva, J., Kivil, A., Tints, K., **Kazantseva, A**., Neuman, T., Palm, K. (2013). Alternative splicing targeting the hTAF4-TAFH domain of TAF4 represses proliferation and accelerates chondrogenic differentiation of human mesenchymal stem cells. PLoS One, 2013; 8(10)

Lend, AK. [1], **Kazantseva, A**. [1], Kivil, A., Valvere, V., Palm, K. (2015). Diagnostic significance of alternative splice variants of REST and DOPEY1 in the peripheral blood of patients with breast cancer. Tumour Biology, 36(4), 2473-80

Sadam, H., Liivas. U., **Kazantseva, A**., Pruunsild, P., Kazantseva, J., Timmusk, T., Neuman, T., Palm, K. (2015). GLI2 cell-specific activity is controlled at the level of transcription and RNA processing: consequences to melanoma progression. *BBA* (*Biochimica et Biophysica Acta*) Molecular Basis of Disease, 1862(1):46-55

[1] - samaväärsed autorid

Patentsed leiutised

CANCER RELATED ISOFORMS OF COMPONENTS OF TRANSCRIPTION FACTOR COMPLEXES AS BIOMARKERS AND DRUG TARGETS. Omanik: Oncotx. Inc; Autorid: **A. Kazantseva**, J. Kazantseva; Patendinumber US 2010/0087376 A1.

COMPOSITIONS AND METHODS FOR TARGETING CANCER-SPECIFIC TRANSCRIPTION COMPLEXES. Omanik: Oncotx. Inc; Autorid: **A. Kazantseva**, A. Liik; Patendinumber 7973135.

Kursused ja konverentsid

Juuni 2005. Tallinn, Estonia. IBRO suvekool neurobioloogias. *Presentatsioon teemal "The complex organization of the brain-derived neurotrophic factor gene locus in humans".*

Detsember 7 – 9, 2005. Berliin, Saksamaa. Ambioni® siRNA praktiline kursus.

Aprill 21 – 26, 2006. Taos, New Mexico, USA. Keystone sümpoosion. "Regulation of Eukaryotic Transcription: From Chromatin to mRNA". *Poster teemal "Cell-type specific isoforms of transcription cofactors"*.

Jaanuar 11 – 14, 2007. La Jolla, California, USA. Nature sümpoosion. "Biological Complexity: Diseases of Transcription". *Poster teemal "Novel neuronal specific BAF57 isoforms description"*.

August 29 – September 2, 2007. Cold Spring Harbour, New York, USA. Cold Spring Harbour konverents. "Mechanisms of eukaryotic transcription". *Poster teemal "Neuron-specific Isoforms of Human BAF57 Regulate Transcription of RE1/NRSE-containing genes"*.

Juuni 7 – 10, 2010. Oosterhout, Holland. Luminex xMAP® Technology praktiline kursus. "Fundamental Assay Techniques Nucleic Acid and Protein".

Oktoober 20 – 23, 2010. Viin, Austria. Planet xMAP konverents. Luminex Technology.

DISSERTATIONS DEFENDED AT
TALLINN UNIVERSITY OF TECHNOLOGY ON
NATURAL AND EXACT SCIENCES

1. **Olav Kongas**. Nonlinear Dynamics in Modeling Cardiac Arrhytmias. 1998.
2. **Kalju Vanatalu**. Optimization of Processes of Microbial Biosynthesis of Isotopically Labeled Biomolecules and Their Complexes. 1999.
3. **Ahto Buldas**. An Algebraic Approach to the Structure of Graphs. 1999.
4. **Monika Drews**. A Metabolic Study of Insect Cells in Batch and Continuous Culture: Application of Chemostat and Turbidostat to the Production of Recombinant Proteins. 1999.
5. **Eola Valdre**. Endothelial-Specific Regulation of Vessel Formation: Role of Receptor Tyrosine Kinases. 2000.
6. **Kalju Lott**. Doping and Defect Thermodynamic Equilibrium in ZnS. 2000.
7. **Reet Koljak**. Novel Fatty Acid Dioxygenases from the Corals *Plexaura homomalla* and *Gersemia fruticosa.* 2001.
8. **Anne Paju**. Asymmetric oxidation of Prochiral and Racemic Ketones by Using Sharpless Catalyst. 2001.
9. **Marko Vendelin**. Cardiac Mechanoenergetics *in silico*. 2001.
10. **Pearu Peterson**. Multi-Soliton Interactions and the Inverse Problem of Wave Crest. 2001.
11. **Anne Menert**. Microcalorimetry of Anaerobic Digestion. 2001.
12. **Toomas Tiivel**. The Role of the Mitochondrial Outer Membrane in *in vivo* Regulation of Respiration in Normal Heart and Skeletal Muscle Cell. 2002.
13. **Olle Hints**. Ordovician Scolecodonts of Estonia and Neighbouring Areas: Taxonomy, Distribution, Palaeoecology, and Application. 2002.
14. **Jaak Nõlvak**. Chitinozoan Biostratigrapy in the Ordovician of Baltoscandia. 2002.
15. **Liivi Kluge**. On Algebraic Structure of Pre-Operad. 2002.
16. **Jaanus Lass**. Biosignal Interpretation: Study of Cardiac Arrhytmias and Electromagnetic Field Effects on Human Nervous System. 2002.
17. **Janek Peterson**. Synthesis, Structural Characterization and Modification of PAMAM Dendrimers. 2002.
18. **Merike Vaher**. Room Temperature Ionic Liquids as Background Electrolyte Additives in Capillary Electrophoresis. 2002.
19. **Valdek Mikli**. Electron Microscopy and Image Analysis Study of Powdered Hardmetal Materials and Optoelectronic Thin Films. 2003.
20. **Mart Viljus**. The Microstructure and Properties of Fine-Grained Cermets. 2003.
21. **Signe Kask**. Identification and Characterization of Dairy-Related *Lactobacillus.* 2003
22. **Tiiu-Mai Laht**. Influence of Microstructure of the Curd on Enzymatic and Microbiological Processes in Swiss-Type Cheese. 2003.

23. **Anne Kuusksalu**. 2–5A Synthetase in the Marine Sponge *Geodia cydonium*. 2003.
24. **Sergei Bereznev**. Solar Cells Based on Polycristalline Copper-Indium Chalcogenides and Conductive Polymers. 2003.
25. **Kadri Kriis**. Asymmetric Synthesis of C_2-Symmetric Bimorpholines and Their Application as Chiral Ligands in the Transfer Hydrogenation of Aromatic Ketones. 2004.
26. **Jekaterina Reut**. Polypyrrole Coatings on Conducting and Insulating Substracts. 2004.
27. **Sven Nõmm**. Realization and Identification of Discrete-Time Nonlinear Systems. 2004.
28. **Olga Kijatkina**. Deposition of Copper Indium Disulphide Films by Chemical Spray Pyrolysis. 2004.
29. **Gert Tamberg**. On Sampling Operators Defined by Rogosinski, Hann and Blackman Windows. 2004.
30. **Monika Übner**. Interaction of Humic Substances with Metal Cations. 2004.
31. **Kaarel Adamberg**. Growth Characteristics of Non-Starter Lactic Acid Bacteria from Cheese. 2004.
32. **Imre Vallikivi**. Lipase-Catalysed Reactions of Prostaglandins. 2004.
33. **Merike Peld**. Substituted Apatites as Sorbents for Heavy Metals. 2005.
34. **Vitali Syritski**. Study of Synthesis and Redox Switching of Polypyrrole and Poly(3,4-ethylenedioxythiophene) by Using *in-situ* Techniques. 2004.
35. **Lee Põllumaa**. Evaluation of Ecotoxicological Effects Related to Oil Shale Industry. 2004.
36. **Riina Aav**. Synthesis of 9,11-Secosterols Intermediates. 2005.
37. **Andres Braunbrück**. Wave Interaction in Weakly Inhomogeneous Materials. 2005.
38. **Robert Kitt**. Generalised Scale-Invariance in Financial Time Series. 2005.
39. **Juss Pavelson**. Mesoscale Physical Processes and the Related Impact on the Summer Nutrient Fields and Phytoplankton Blooms in the Western Gulf of Finland. 2005.
40. **Olari Ilison**. Solitons and Solitary Waves in Media with Higher Order Dispersive and Nonlinear Effects. 2005.
41. **Maksim Säkki**. Intermittency and Long-Range Structurization of Heart Rate. 2005.
42. **Enli Kiipli**. Modelling Seawater Chemistry of the East Baltic Basin in the Late Ordovician–Early Silurian. 2005.
43. **Igor Golovtsov**. Modification of Conductive Properties and Processability of Polyparaphenylene, Polypyrrole and polyaniline. 2005.
44. **Katrin Laos**. Interaction Between Furcellaran and the Globular Proteins (Bovine Serum Albumin β-Lactoglobulin). 2005.
45. **Arvo Mere**. Structural and Electrical Properties of Spray Deposited Copper Indium Disulphide Films for Solar Cells. 2006.

46. **Sille Ehala**. Development and Application of Various On- and Off-Line Analytical Methods for the Analysis of Bioactive Compounds. 2006.
47. **Maria Kulp**. Capillary Electrophoretic Monitoring of Biochemical Reaction Kinetics. 2006.
48. **Anu Aaspõllu.** Proteinases from *Vipera lebetina* Snake Venom Affecting Hemostasis. 2006.
49. **Lyudmila Chekulayeva**. Photosensitized Inactivation of Tumor Cells by Porphyrins and Chlorins. 2006.
50. **Merle Uudsemaa**. Quantum-Chemical Modeling of Solvated First Row Transition Metal Ions. 2006.
51. **Tagli Pitsi**. Nutrition Situation of Pre-School Children in Estonia from 1995 to 2004. 2006.
52. **Angela Ivask**. Luminescent Recombinant Sensor Bacteria for the Analysis of Bioavailable Heavy Metals. 2006.
53. **Tiina Lõugas**. Study on Physico-Chemical Properties and Some Bioactive Compounds of Sea Buckthorn (*Hippophae rhamnoides* L.). 2006.
54. **Kaja Kasemets**. Effect of Changing Environmental Conditions on the Fermentative Growth of *Saccharomyces cerevisae* S288C: Auxo-accelerostat Study. 2006.
55. **Ildar Nisamedtinov**. Application of ^{13}C and Fluorescence Labeling in Metabolic Studies of *Saccharomyces* spp. 2006.
56. **Alar Leibak**. On Additive Generalisation of Voronoï's Theory of Perfect Forms over Algebraic Number Fields. 2006.
57. **Andri Jagomägi**. Photoluminescence of Chalcopyrite Tellurides. 2006.
58. **Tõnu Martma**. Application of Carbon Isotopes to the Study of the Ordovician and Silurian of the Baltic. 2006.
59. **Marit Kauk**. Chemical Composition of $CuInSe_2$ Monograin Powders for Solar Cell Application. 2006.
60. **Julia Kois**. Electrochemical Deposition of $CuInSe_2$ Thin Films for Photovoltaic Applications. 2006.
61. **Ilona Oja Açik**. Sol-Gel Deposition of Titanium Dioxide Films. 2007.
62. **Tiia Anmann**. Integrated and Organized Cellular Bioenergetic Systems in Heart and Brain. 2007.
63. **Katrin Trummal**. Purification, Characterization and Specificity Studies of Metalloproteinases from *Vipera lebetina* Snake Venom. 2007.
64. **Gennadi Lessin**. Biochemical Definition of Coastal Zone Using Numerical Modeling and Measurement Data. 2007.
65. **Enno Pais**. Inverse problems to determine non-homogeneous degenerate memory kernels in heat flow. 2007.
66. **Maria Borissova**. Capillary Electrophoresis on Alkylimidazolium Salts. 2007.
67. **Karin Valmsen**. Prostaglandin Synthesis in the Coral *Plexaura homomalla*: Control of Prostaglandin Stereochemistry at Carbon 15 by Cyclooxygenases. 2007.

68. **Kristjan Piirimäe**. Long-Term Changes of Nutrient Fluxes in the Drainage Basin of the Gulf of Finland – Application of the PolFlow Model. 2007.
69. **Tatjana Dedova**. Chemical Spray Pyrolysis Deposition of Zinc Sulfide Thin Films and Zinc Oxide Nanostructured Layers. 2007.
70. **Katrin Tomson**. Production of Labelled Recombinant Proteins in Fed-Batch Systems in *Escherichia coli*. 2007.
71. **Cecilia Sarmiento**. Suppressors of RNA Silencing in Plants. 2008.
72. **Vilja Mardla**. Inhibition of Platelet Aggregation with Combination of Antiplatelet Agents. 2008.
73. **Maie Bachmann**. Effect of Modulated Microwave Radiation on Human Resting Electroencephalographic Signal. 2008.
74. **Dan Hüvonen**. Terahertz Spectroscopy of Low-Dimensional Spin Systems. 2008.
75. **Ly Villo**. Stereoselective Chemoenzymatic Synthesis of Deoxy Sugar Esters Involving *Candida antarctica* Lipase B. 2008.
76. **Johan Anton**. Technology of Integrated Photoelasticity for Residual Stress Measurement in Glass Articles of Axisymmetric Shape. 2008.
77. **Olga Volobujeva**. SEM Study of Selenization of Different Thin Metallic Films. 2008.
78. **Artur Jõgi**. Synthesis of 4'-Substituted 2,3'-dideoxynucleoside Analogues. 2008.
79. **Mario Kadastik**. Doubly Charged Higgs Boson Decays and Implications on Neutrino Physics. 2008.
80. **Fernando Pérez-Caballero**. Carbon Aerogels from 5-Methylresorcinol-Formaldehyde Gels. 2008.
81. **Sirje Vaask**. The Comparability, Reproducibility and Validity of Estonian Food Consumption Surveys. 2008.
82. **Anna Menaker**. Electrosynthesized Conducting Polymers, Polypyrrole and Poly(3,4-ethylenedioxythiophene), for Molecular Imprinting. 2009.
83. **Lauri Ilison**. Solitons and Solitary Waves in Hierarchical Korteweg-de Vries Type Systems. 2009.
84. **Kaia Ernits**. Study of In_2S_3 and ZnS Thin Films Deposited by Ultrasonic Spray Pyrolysis and Chemical Deposition. 2009.
85. **Veljo Sinivee**. Portable Spectrometer for Ionizing Radiation "Gammamapper". 2009.
86. **Jüri Virkepu**. On Lagrange Formalism for Lie Theory and Operadic Harmonic Oscillator in Low Dimensions. 2009.
87. **Marko Piirsoo**. Deciphering Molecular Basis of Schwann Cell Development. 2009.
88. **Kati Helmja**. Determination of Phenolic Compounds and Their Antioxidative Capability in Plant Extracts. 2010.
89. **Merike Sõmera**. Sobemoviruses: Genomic Organization, Potential for Recombination and Necessity of P1 in Systemic Infection. 2010.

90. **Kristjan Laes**. Preparation and Impedance Spectroscopy of Hybrid Structures Based on $CuIn_3Se_5$ Photoabsorber. 2010.
91. **Kristin Lippur**. Asymmetric Synthesis of 2,2'-Bimorpholine and its 5,5'-Substituted Derivatives. 2010.
92. **Merike Luman**. Dialysis Dose and Nutrition Assessment by an Optical Method. 2010.
93. **Mihhail Berezovski**. Numerical Simulation of Wave Propagation in Heterogeneous and Microstructured Materials. 2010.
94. **Tamara Aid-Pavlidis**. Structure and Regulation of BDNF Gene. 2010.
95. **Olga Bragina**. The Role of Sonic Hedgehog Pathway in Neuro- and Tumorigenesis. 2010.
96. **Merle Randrüüt**. Wave Propagation in Microstructured Solids: Solitary and Periodic Waves. 2010.
97. **Marju Laars**. Asymmetric Organocatalytic Michael and Aldol Reactions Mediated by Cyclic Amines. 2010.
98. **Maarja Grossberg**. Optical Properties of Multinary Semiconductor Compounds for Photovoltaic Applications. 2010.
99. **Alla Maloverjan**. Vertebrate Homologues of Drosophila Fused Kinase and Their Role in Sonic Hedgehog Signalling Pathway. 2010.
100. **Priit Pruunsild**. Neuronal Activity-Dependent Transcription Factors and Regulation of Human *BDNF* Gene. 2010.
101. **Tatjana Knjazeva**. New Approaches in Capillary Electrophoresis for Separation and Study of Proteins. 2011.
102. **Atanas Katerski**. Chemical Composition of Sprayed Copper Indium Disulfide Films for Nanostructured Solar Cells. 2011.
103. **Kristi Timmo.** Formation of Properties of $CuInSe_2$ and $Cu_2ZnSn(S,Se)_4$ Monograin Powders Synthesized in Molten KI. 2011.
104. **Kert Tamm**. Wave Propagation and Interaction in Mindlin-Type Microstructured Solids: Numerical Simulation. 2011.
105. **Adrian Popp**. Ordovician Proetid Trilobites in Baltoscandia and Germany. 2011.
106. **Ove Pärn**. Sea Ice Deformation Events in the Gulf of Finland and This Impact on Shipping. 2011.
107. **Germo Väli**. Numerical Experiments on Matter Transport in the Baltic Sea. 2011.
108. **Andrus Seiman**. Point-of-Care Analyser Based on Capillary Electrophoresis. 2011.
109. **Olga Katargina**. Tick-Borne Pathogens Circulating in Estonia (Tick-Borne Encephalitis Virus, *Anaplasma phagocytophilum*, *Babesia* Species): Their Prevalence and Genetic Characterization. 2011.
110. **Ingrid Sumeri**. The Study of Probiotic Bacteria in Human Gastrointestinal Tract Simulator. 2011.
111. **Kairit Zovo**. Functional Characterization of Cellular Copper Proteome. 2011.

112. **Natalja Makarytsheva**. Analysis of Organic Species in Sediments and Soil by High Performance Separation Methods. 2011.
113. **Monika Mortimer**. Evaluation of the Biological Effects of Engineered Nanoparticles on Unicellular Pro- and Eukaryotic Organisms. 2011.
114. **Kersti Tepp**. Molecular System Bioenergetics of Cardiac Cells: Quantitative Analysis of Structure-Function Relationship. 2011.
115. **Anna-Liisa Peikolainen**. Organic Aerogels Based on 5-Methylresorcinol. 2011.
116. **Leeli Amon**. Palaeoecological Reconstruction of Late-Glacial Vegetation Dynamics in Eastern Baltic Area: A View Based on Plant Macrofossil Analysis. 2011.
117. **Tanel Peets**. Dispersion Analysis of Wave Motion in Microstructured Solids. 2011.
118. **Liina Kaupmees**. Selenization of Molybdenum as Contact Material in Solar Cells. 2011.
119. **Allan Olspert**. Properties of VPg and Coat Protein of Sobemoviruses. 2011.
120. **Kadri Koppel**. Food Category Appraisal Using Sensory Methods. 2011.
121. **Jelena Gorbatšova**. Development of Methods for CE Analysis of Plant Phenolics and Vitamins. 2011.
122. **Karin Viipsi**. Impact of EDTA and Humic Substances on the Removal of Cd and Zn from Aqueous Solutions by Apatite. 2012.
123. **David Schryer**. Metabolic Flux Analysis of Compartmentalized Systems Using Dynamic Isotopologue Modeling. 2012.
124. **Ardo Illaste**. Analysis of Molecular Movements in Cardiac Myocytes. 2012.
125. **Indrek Reile**. 3-Alkylcyclopentane-1,2-Diones in Asymmetric Oxidation and Alkylation Reactions. 2012.
126. **Tatjana Tamberg**. Some Classes of Finite 2-Groups and Their Endomorphism Semigroups. 2012.
127. **Taavi Liblik**. Variability of Thermohaline Structure in the Gulf of Finland in Summer. 2012.
128. **Priidik Lagemaa**. Operational Forecasting in Estonian Marine Waters. 2012.
129. **Andrei Errapart**. Photoelastic Tomography in Linear and Non-linear Approximation. 2012.
130. **Külliki Krabbi**. Biochemical Diagnosis of Classical Galactosemia and Mucopolysaccharidoses in Estonia. 2012.
131. **Kristel Kaseleht**. Identification of Aroma Compounds in Food using SPME-GC/MS and GC-Olfactometry. 2012.
132. **Kristel Kodar**. Immunoglobulin G Glycosylation Profiling in Patients with Gastric Cancer. 2012.
133. **Kai Rosin**. Solar Radiation and Wind as Agents of the Formation of the Radiation Regime in Water Bodies. 2012.
134. **Ann Tiiman**. Interactions of Alzheimer's Amyloid-Beta Peptides with Zn(II) and Cu(II) Ions. 2012.

135. **Olga Gavrilova**. Application and Elaboration of Accounting Approaches for Sustainable Development. 2012.
136. **Olesja Bondarenko**. Development of Bacterial Biosensors and Human Stem Cell-Based *In Vitro* Assays for the Toxicological Profiling of Synthetic Nanoparticles. 2012.
137. **Katri Muska**. Study of Composition and Thermal Treatments of Quaternary Compounds for Monograin Layer Solar Cells. 2012.
138. **Ranno Nahku**. Validation of Critical Factors for the Quantitative Characterization of Bacterial Physiology in Accelerostat Cultures. 2012.
139. **Petri-Jaan Lahtvee**. Quantitative Omics-level Analysis of Growth Rate Dependent Energy Metabolism in *Lactococcus lactis*. 2012.
140. **Kerti Orumets**. Molecular Mechanisms Controlling Intracellular Glutathione Levels in Baker's Yeast *Saccharomyces cerevisiae* and its Random Mutagenized Glutathione Over-Accumulating Isolate. 2012.
141. **Loreida Timberg**. Spice-Cured Sprats Ripening, Sensory Parameters Development, and Quality Indicators. 2012.
142. **Anna Mihhalevski**. Rye Sourdough Fermentation and Bread Stability. 2012.
143. **Liisa Arike**. Quantitative Proteomics of *Escherichia coli*: From Relative to Absolute Scale. 2012.
144. **Kairi Otto**. Deposition of In_2S_3 Thin Films by Chemical Spray Pyrolysis. 2012.
145. **Mari Sepp**. Functions of the Basic Helix-Loop-Helix Transcription Factor TCF4 in Health and Disease. 2012.
146. **Anna Suhhova**. Detection of the Effect of Weak Stressors on Human Resting Electroencephalographic Signal. 2012.
147. **Aram Kazarjan**. Development and Production of Extruded Food and Feed Products Containing Probiotic Microorganisms. 2012.
148. **Rivo Uiboupin**. Application of Remote Sensing Methods for the Investigation of Spatio-Temporal Variability of Sea Surface Temperature and Chlorophyll Fields in the Gulf of Finland. 2013.
149. **Tiina Kriščiunaite**. A Study of Milk Coagulability. 2013.
150. **Tuuli Levandi**. Comparative Study of Cereal Varieties by Analytical Separation Methods and Chemometrics. 2013.
151. **Natalja Kabanova**. Development of a Microcalorimetric Method for the Study of Fermentation Processes. 2013.
152. **Himani Khanduri**. Magnetic Properties of Functional Oxides. 2013.
153. **Julia Smirnova**. Investigation of Properties and Reaction Mechanisms of Redox-Active Proteins by ESI MS. 2013.
154. **Mervi Sepp**. Estimation of Diffusion Restrictions in Cardiomyocytes Using Kinetic Measurements. 2013.
155. **Kersti Jääger**. Differentiation and Heterogeneity of Mesenchymal Stem Cells. 2013.
156. **Victor Alari**. Multi-Scale Wind Wave Modeling in the Baltic Sea. 2013.

157. **Taavi Päll**. Studies of CD44 Hyaluronan Binding Domain as Novel Angiogenesis Inhibitor. 2013.
158. **Allan Niidu**. Synthesis of Cyclopentane and Tetrahydrofuran Derivatives. 2013.
159. **Julia Geller**. Detection and Genetic Characterization of *Borrelia* Species Circulating in Tick Population in Estonia. 2013.
160. **Irina Stulova**. The Effects of Milk Composition and Treatment on the Growth of Lactic Acid Bacteria. 2013.
161. **Jana Holmar**. Optical Method for Uric Acid Removal Assessment During Dialysis. 2013.
162. **Kerti Ausmees**. Synthesis of Heterobicyclo[3.2.0]heptane Derivatives *via* Multicomponent Cascade Reaction. 2013.
163. **Minna Varikmaa**. Structural and Functional Studies of Mitochondrial Respiration Regulation in Muscle Cells. 2013.
164. **Indrek Koppel**. Transcriptional Mechanisms of BDNF Gene Regulation. 2014.
165. **Kristjan Pilt**. Optical Pulse Wave Signal Analysis for Determination of Early Arterial Ageing in Diabetic Patients. 2014.
166. **Andres Anier**. Estimation of the Complexity of the Electroencephalogram for Brain Monitoring in Intensive Care. 2014.
167. **Toivo Kallaste**. Pyroclastic Sanidine in the Lower Palaeozoic Bentonites – A Tool for Regional Geological Correlations. 2014.
168. **Erki Kärber**. Properties of ZnO-nanorod/In_2S_3/$CuInS_2$ Solar Cell and the Constituent Layers Deposited by Chemical Spray Method. 2014.
169. **Julia Lehner**. Formation of Cu_2ZnSnS_4 and $Cu_2ZnSnSe_4$ by Chalcogenisation of Electrochemically Deposited Precursor Layers. 2014.
170. **Peep Pitk**. Protein- and Lipid-rich Solid Slaughterhouse Waste Anaerobic Co-digestion: Resource Analysis and Process Optimization. 2014.
171. **Kaspar Valgepea**. Absolute Quantitative Multi-omics Characterization of Specific Growth Rate-dependent Metabolism of *Escherichia coli*. 2014.
172. **Artur Noole**. Asymmetric Organocatalytic Synthesis of 3,3'-Disubstituted Oxindoles. 2014.
173. **Robert Tsanev**. Identification and Structure-Functional Characterisation of the Gene Transcriptional Repressor Domain of Human Gli Proteins. 2014.
174. **Dmitri Kartofelev**. Nonlinear Sound Generation Mechanisms in Musical Acoustic. 2014.
175. **Sigrid Hade**. GIS Applications in the Studies of the Palaeozoic Graptolite Argillite and Landscape Change. 2014.
176. **Agne Velthut-Meikas**. Ovarian Follicle as the Environment of Oocyte Maturation: The Role of Granulosa Cells and Follicular Fluid at Pre-Ovulatory Development. 2014.
177. **Kristel Hälvin**. Determination of B-group Vitamins in Food Using an LC-MS Stable Isotope Dilution Assay. 2014.

178. **Mailis Päri**. Characterization of the Oligoadenylate Synthetase Subgroup from Phylum Porifera. 2014.
179. **Jekaterina Kazantseva**. Alternative Splicing of *TAF4*: A Dynamic Switch between Distinct Cell Functions. 2014.
180. **Jaanus Suurväli**. Regulator of G Protein Signalling 16 (RGS16): Functions in Immunity and Genomic Location in an Ancient MHC-Related Evolutionarily Conserved Synteny Group. 2014.
181. **Ene Viiard**. Diversity and Stability of Lactic Acid Bacteria During Rye Sourdough Propagation. 2014.
182. **Kristella Hansen**. Prostaglandin Synthesis in Marine Arthropods and Red Algae. 2014.
183. **Helike Lõhelaid**. Allene Oxide Synthase-lipoxygenase Pathway in Coral Stress Response. 2015.
184. **Normunds Stivrinš**. Postglacial Environmental Conditions, Vegetation Succession and Human Impact in Latvia. 2015.
185. **Mary-Liis Kütt**. Identification and Characterization of Bioactive Peptides with Antimicrobial and Immunoregulating Properties Derived from Bovine Colostrum and Milk. 2015.
186. **Kazbulat Šogenov**. Petrophysical Models of the CO_2 Plume at Prospective Storage Sites in the Baltic Basin. 2015.
187. **Taavi Raadik**. Application of Modulation Spectroscopy Methods in Photovoltaic Materials Research. 2015.
188. **Reio Põder**. Study of Oxygen Vacancy Dynamics in Sc-doped Ceria with NMR Techniques. 2015.
189. **Sven Siir**. Internal Geochemical Stratification of Bentonites (Altered Volcanic Ash Beds) and its Interpretation. 2015.
190. **Kaur Jaanson**. Novel Transgenic Models Based on Bacterial Artificial Chromosomes for Studying BDNF Gene Regulation. 2015.
191. **Niina Karro**. Analysis of ADP Compartmentation in Cardiomyocytes and Its Role in Protection Against Mitochondrial Permeability Transition Pore Opening. 2015.
192. **Piret Laht**. B-plexins Regulate the Maturation of Neurons Through Microtubule Dynamics. 2015.
193. **Sergei Žari**. Organocatalytic Asymmetric Addition to Unsaturated 1,4-Dicarbonyl Compounds. 2015.
194. **Natalja Buhhalko**. Processes Influencing the Spatio-temporal Dynamics of Nutrients and Phytoplankton in Summer in the Gulf of Finland, Baltic Sea. 2015.
195. **Natalia Maticiuc**. Mechanism of Changes in the Properties of Chemically Deposited CdS Thin Films Induced by Thermal Annealing. 2015.
196. **Mario Öeren**. Computational Study of Cyclohexylhemicucurbiturils. 2015.
197. **Mari Kalda**. Mechanoenergetics of a Single Cardiomyocyte. 2015.
198. **Ieva Grudzinska**. Diatom Stratigraphy and Relative Sea Level Changes of the Eastern Baltic Sea Over the Holocene. 2015.